Терри Дж. Уотшем
Кейт Паррамоу

Количественные методы в финансах

Перевод с английского под редакцией
профессора ***М.Р. Ефимовой***

Рекомендовано Министерством общего и профессионального образования Российской Федерации в качестве учебного пособия для студентов высших учебных заведений, обучающихся по экономическим специальностям

Москва
"Финансы"
Издательское объединение "ЮНИТИ"
1999

ББК 65.26в6.я73
У65

Перевод с английского
Ю.В. Воличева, С.М. Таскаева, Э.Б. Аникеенко

Рецензент — д-р экон. наук, проф. ***В.С. Мхитарян***

Главный редактор издательства ***Н.Д. Эриашвили***

Уотшем Т. Дж., Паррамоу К.
У65 Количественные методы в финансах: Учеб. пособие для вузов/Пер. с англ. под ред. М.Р. Ефимовой. — М.: Финансы, ЮНИТИ, 1999. – 527 с.
ISBN 5-238-00036-7 (русск.)
ISBN 0-412-60820-0 (англ.)

Принятие решений в финансовом и банковском менеджменте опирается на точный расчет и количественную оценку последствий принимаемых решений. Учебное пособие предлагает читателю широкий спектр математических и статистических методов, применяемых при исследовании финансовых рисков, управлении рисками, принятии инвестиционных решений.

Для студентов и аспирантов экономических специальностей, а также специалистов банковских и финансовых структур.

ББК 65.26в6.я73

ISBN 5-238-00036-7 (русск.)
ISBN 0-412-60820-0 (англ.)

ПРЕДИСЛОВИЕ

В течение последних 30 лет теория и практика финансов во все большей степени стала опираться на количественные методы математики, статистики и эконометрии. Это привело к более частому использованию количественного анализа при изучении поведения финансовых рынков. Расширенное применение количественного анализа в последнее десятилетие обусловлено также распространением персональных компьютеров в среде академических и финансовых профессионалов.

Параллельно и отчасти по этим же причинам появился новый раздел литературы, посвященный управлению рисками и производным финансовым инструментам, таким, как опционы, фьючерсы и свопы. Это связано с развитием новых сфер в финансовой практике: работой с деривативами, управлением финансовыми рисками, количественным анализом инвестиций.

Новая литература и новые методы задействовали количественные приёмы, ранее применявшиеся только в физике, в то же время развив или адаптировав технику количественного анализа к экономике.

Дальнейшие научные разработки сопровождались значительным увеличением числа людей, работающих на финансовых рынках и в сфере финансовых услуг. Этим специалистам необходимо хорошее понимание количественных методов, лежащих в основе их деятельности. Однако лишь незначительная часть этих людей имеет соответствующую подготовку в указанной области. Таким образом, у студентов, аспирантов и финансовых профессионалов существует возрастающая потребность в развитии навыков применения количественных методов для работы на современных финансовых рынках.

Основная цель написания данной книги — способствовать этому процессу, для чего был собран ряд важных количественных методов в приложении к финансам. Прежде всего книга является учебным пособием, объединяющим математическую и финансовую теории и их применение. Книга написана в стиле, делающем её полезной как для студентов, аспирантов, лиц, занимающихся исследованиями в области финансов, так и для профессионалов финансовых рынков.

Нам известно, что существует большой разброс в математических знаниях как среди студентов, так и среди тех, кто работает на финансовых рынках. Это отражено и в содержании книги, и в стиле подачи информации. В силу данной необходимости мы должны были быть осмотрительны в выборе тем. Поэтому они избраны в соответствии с их полезностью в развитии навыков применения количественных методов, наиболее часто используемых в финансовых исследованиях.

Книга состоит из 11 глав, начинаюшихся с основ и плавно переходящих к более сложным приёмам, применяющимся в анализе управления рисками. Стиль изложения удобен для читателей, желающих познакомиться с методами, которые сейчас используются, но не совсем уверенных в своих навыках количественного анализа. В то же время разнообразие тем делает книгу хорошей отправной точкой для тех, кто уже знаком с количественными методами, но чувствует необходимость расширить свои познания в применении современных методик в финансах.

Глава 1 охватывает математику процентных ставок и рентабельность капитала, в гл. 2 рассказывается об описательной статистике, применяющейся в финансовой практике, гл. 3 и 4 знакомят с дифференциальным и интегральным исчислениями, а также с теорией вероятностей. В гл. 5 развивается тема применения теории вероятностей для проверки гипотез и определения доверительных интервалов, что важно в управлении рисками. Гл. 6 рассматривает регрессионный анализ, а гл. 7 демонстрирует некоторые современные приёмы анализа временных рядов, в частности, охватывает авторегрессионную условную гетероскедастичность (ARCH), общую авторегрессионную условную гетероскедастичность (GARCH) и коинтеграцию Гл. 8 знакомит с рядом численных методов, которые все более активно используются в финансах. В частности, она охватывает численное интегрирование и метод Монте-Карло. Гл. 9 рассказывает о методах оптимизации и их применении в построении портфелей ценных бумаг. Гл. 10 объясняет и рассматривает применение стохастического исчисления, на котором базируются многие известные способы ценообразования опционов, включая модель Блэка—Сколса и ее разновидности. Гл. 11 охватывает многофакторный анализ, особенно метод главных компонент и факторный анализ, которые полезны в управлении рисками.

Написание книги, подобной этой, требует многих жертв, в том числе и от наших жён, Джо и Джоан, которым мы хотим выразить свою благодарность.

Мы также благодарим Сару Хэндерсон и производственную команду Интернешнл Томсон Паблишинг за содействие в данном проекте.

1

ПРОЦЕНТНЫЕ СТАВКИ И РЕНТАБЕЛЬНОСТЬ АКТИВОВ

ВВЕДЕНИЕ

Более всего финансовая наука связана с анализом прибыльности инвестиционной деятельности. Цель инвестиций — увеличение благосостояния инвестора. Это увеличение называется доходом, а при выражении в процентах от стоимости инвестиций — ставкой дохода. Кроме измерения доходности финансовые аналитики имеют дело также с неопределенностью получения дохода, с этой неопределенностью связан анализ риска. В этой главе мы сосредоточимся на измерении доходности, а в следующих главах рассмотрим различные виды и аспекты рисков.

Инвесторы приобретают активы, такие, как акции компаний, облигации или недвижимость, в надежде получить доход либо от продажи их по более высокой цене, либо в виде дивидендов, процентов по купонам или рентных платежей. Кредиторы ссужают деньги в надежде получить доход в виде процентных платежей при полном погашении кредита заемщиком. Таким образом, кредиторы и инвесторы имеют общую цель — получить доход, или процент, как результат инвестиционной или кредиторской деятельности.

Соответственно в этой главе под терминами "процент" будет подразумеваться и доход от инвестиций, "заимодавец" — также и инвестор, а "кредит" будет синонимом и инвестиции.

Эта глава начинается с краткого изложения экономической теории процента, за которой последует математический анализ процентных ставок и дохода с активов. Затем математический анализ будет применен к краткосрочным денежным инструментам, таким, как банковские депозиты и ипотечные ссуды, банковские сертификаты, казначейские векселя, банковские акцепты и краткосрочные коммерческие векселя. В конце главы рассмотрены различные способы оценки доходности, используемые на рынках акций и облигаций.

ЭКОНОМИЧЕСКАЯ ТЕОРИЯ ПРОЦЕНТА

Процентные ставки, вероятно, — наиболее часто используемые финансовые показатели. Многие читатели в своей жизни берут кредит и платят проценты по этому кредиту. Многие читатели помещают деньги в банк или другое финансовое учреждение и получат за это процентные платежи. Во время этих действий они заметят, что существует большое разнообразие процентных ставок по кредитам и вкладам. Эти ставки отличаются не только по величине, но и по методу их вычисления. Одни процентные ставки фиксированы в течение всего периода действия договора, другие же могут изменяться на оговоренных условиях в определенные промежутки времени. Существуют и такие, например по ипотечным ссудам, которые могут меняться по желанию кредитора. Но почему же вообще выплачиваются проценты по кредитам и депозитам?

Чтобы ответить на этот вопрос, необходимо осознать, что деньги приносят выгоду или обеспечивают благосостояние только косвенно, являясь средством обмена. Это означает, что они должны быть обменены на другие товары или услуги, чтобы принести прямую пользу. Следовательно, деньги сами по себе (банкноты, монеты, банковские счета) мало удовлетворяют жизненные потребности. Это происходит путем обмена денег на товары и услуги, такие, как продукты питания, одежда и жилье.

Таким образом, когда кто-либо инвестирует деньги, он отказывается от возможности обратить их в товары и услуги, которые приносят пользу напрямую. Поэтому ему придется довольствоваться более низким уровнем полезности, чем если бы деньги были употреблены для приобретения товаров и услуг вместо инвестирования. Эта потеря потенциальной полезности должна быть компенсирована — в этом состоит важнейшая функция процента.

Далее, кредитор сталкивается со значительной неопределенностью относительно стоимости денег, когда они к нему возвращаются (будущая стоимость этих денег неопределенна). Количественная мера этой неопределенности известна как риск. Кредиторы встречаются с различными видами рисков, и это может уменьшить их выгоду или благосостояние. Одной из функций процента является компенсация этой потери выгоды, существующей из-за рисков.

Первый вид риска, который мы рассмотрим — это *риск потери покупательной способности*, более известный как *инфляционный*. Если экономика характеризуется наличием инфляции в период инвестирования суммы денег или выдачи займа, на эту сумму можно будет приобрести меньшее количество товаров и услуг, когда они будут возвращены, чем до выдачи займа или инвестирования. Поэтому инвесторы будут требовать компенсации за потерю покупательной способности.

Второй вид риска — это то, что деньги могут быть не возвращены вследствие либо неудачи проекта, либо обмана со стороны заемщика, или других непредвиденных обстоятельств. Этот риск называется *риском невыполнения обязательс*ТВ, и снова инвесторы должны получить компенсацию за него.

Если мы объединим риск потери потенциальной выгоды, инфляционный риск и риск невыполнения обязательств, мы получим группу факторов, которые делают обладание деньгами в настоящем более предпочтительным по сравнению с их обладанием в будущем. Люди предпочитают иметь деньги сейчас, чем довольствоваться обещаниями получить их позднее. О деньгах говорят, что они характеризуются **положительным временным предпочтением** (positive time preference).

Проценты компенсируют заимодавцу невозможность удовлетворить эти предпочтения в момент инвестирования средств. Заемщики готовы заплатить за использование средств, потому что это позволяет им иметь дополнительную выгоду раннего потребления в результате получения средств от инвестора.

Из того, что было изложено, должно быть ясно, что существует множество процентных ставок. В любое время на финансовых рынках существует ряд процентных ставок, поэтому полезно разделить факторы, определяющие эти ставки, на две группы: те, которые определяют общий уровень процентных ставок, и те, которые определяют различие процентных ставок.

Факторы, влияющие на уровень процентных ставок:

- политика правительства
- денежная масса
- ожидания относительно будущей инфляции.

Факторы, влияющие на различие процентных ставок:

- время до погашения финансовых обязательств
- риск невыполнения обязательств
- ликвидность финансовых обязательств
- налогообложение
- другие различные факторы, специфические для конкретных финансовых обязательств, например, предоставлено ли обеспечение активами, включены ли права выбора в договор.

Для иллюстрации влияния этих факторов на процентные ставки рассмотрим табл. 1.1, в которой приведены процентные ставки, соответствующие различным финансовым инструментам или рынкам в конце декабря 1995г.

Таблица 1.1. Процентные ставки (%), декабрь 1995

Ипотечная ссуда	7,85
Банковский депозит с недельным сроком извещения о снятии средств	4,5
Банковский овердрафт	22,0
Трехмесячный банковский депозитный сертификат	6,375
Трехмесячный коммерческий вексель	6,45
Трехмесячный казначейский вексель английского правительства	6,32
Шестимесячный межбанковский кредит	6,34
Государственная облигация со сроком погашения 5 лет (Великобритания)	7,0
Государственная облигация со сроком погашения 10 лет (Великобритания)	7,4
Государственная облигация со сроком погашения 10 лет (Германия)	5,88
Корпоративная облигация со сроком погашения 5 лет, обеспеченная активами корпорации (Великобритания)	8,1
Корпоративная облигация со сроком погашения 5 лет, не обеспеченная активами корпорации (Великобритания)	9,2
Конвертируемая облигация со сроком погашения 5 лет (Великобритания)	6,5

Временная стоимость денег

Временна́я стоимость денег имеет отношение к процессу определения текущей стоимости, т.е. сегодняшней стоимости суммы, обещанной в какой-либо момент в будущем, или к расчету будущей стоимости, т.е. стоимости суммы в будущем, полученной или уплаченной сегодня. Процесс определения текущей стоимости денег называется **дисконтированием**, а будущей — **наращением**.

Текущая стоимость находится путем дисконтирования каждого из потоков платежей на процент, который мог бы быть заработан, если бы эти средства были получены сегодня. Наиболее распространенное применение дисконтирования — это оценка облигаций путем дисконтирования будущих купонных платежей, а также оценка акций на основе использования модели дисконтирования дивидендов. Оба эти приложения продемонстрированы ниже в этой главе.

Финансовые активы оцениваются при помощи расчета текущей стоимости ожидаемых потоков платежей от этих активов. Некоторые финансовые инструменты, такие, как фьючерсы и форварды, оцениваются исходя из будущей стоимости денег. **Будущая стоимость находится наращением** всех процентных платежей, которые можно было бы получить на данную сумму до наступления определенного момента в будущем.

Детально проанализируем математические методы наращения и дисконтирования. Начнем с будущей стоимости денег.

Будущая стоимость денег

Рассмотрим ситуацию, когда деньги помещаются на банковский депозит под определенный процент. Стоимость вклада будет расти по мере накопления процентов, следовательно, ожидаемая будущая стоимость вклада будет выше первоначальной на сумму, равную начисленным процентам. Например, будущая стоимость $1000, помещенных под 6% годовых, через год будет равна $1060, т. е. $1000 плюс $60 в качестве начисленных процентов.

В общем случае можно сказать, что размер начисленных процентов и соответственно **будущая стоимость** зависят от **процентной ставки, срока до погашения обязательства**, а также от того, является ли процентная ставка **простой** или **сложной**. Если ставка сложная, то важное значение имеет также и **частота**

начисления процентов. Все эти понятия объяснены и проиллюстрированы ниже. Тип применяемой в расчетах ставки — простой или сложной — зависит от вида конкретного финансового инструмента.

Простые проценты. Если ставка простая, то начисляемые проценты на депозит или по кредиту рассчитываются как произведение процентной ставки, количества лет (или их соответствующих долей) до срока погашения и суммы вклада. Например, если \$1000 разместить на 6 месяцев под 6% годовых, то проценты, начисленные по простой ставке, будут равны:

$$0{,}06 \cdot 0{,}5 \cdot \$1000 = \$30.$$

Будущая стоимость депозита составит:

$$\mathbf{1000\ [1 + (0{,}06 \cdot 0{,}5)] = 1030.}$$

Заметьте, как показаны 6% в данной ситуации: 6% — это 6/100, или 0,06, т. е. проценты выражаются в виде десятичной дроби, например, 6% = 0,06, 10% = 0,1 и т.д.

Таким образом, в общем виде формула для нахождения будущей стоимости по простой процентной ставке выглядит так:

$$FV = P[1 + (r \cdot n)], \qquad (1.1)$$

где FV — будущая стоимость;
P — сумма основного долга;
n — срок вклада в годах;
r — простая процентная ставка в долях единицы.

Обычно, если срок действия финансового инструмента больше одного года, то применяются сложные проценты. О них и пойдет далее речь.

Сложные проценты. Наращение по сложным процентам относится к периодическому добавлению накопленных процентов к основной сумме долга, т.е. накопленные проценты добавляются к основной сумме, увеличивая тем самым ее размер. Проценты в дальнейшем начисляются уже на эту увеличенную сумму. Например, денежная сумма в размере 1000 единиц помещается на банковский депозит на срок 3 года с годовой процентной ставкой 6% и ежегодным начислением процентов. В конце первого года 60 единиц, т.е. (1000 · 0,06), будет добавлено к первоначальному взносу, который возрастет до 1060 единиц. Это можно определить и так: 1000 · 1,06 = 1060. В течение второго года проценты будут начисляться уже на 1060 единиц. В конце второго года 63,6 (1060 · 0,06) будет добавлено к 1060, таким образом в третьем году базой для начисления процентов будет сумма 1123,6, т.е. (1000 · 1,06 · 1,06), и т. д.

Стоимость денег, помещенных на депозит, через 3 года (будущая стоимость первоначальной суммы) рассчитывается так: 1000 · 1,06 · 1,06 · 1,06 = 1191,02. Это пример роста в геометрической прогрессии и записывается следующим образом:

$$1000(1{,}06)^3 = 1191{,}016.$$

Общая формула для расчета будущей стоимости денег при ежегодном начислении процентов выглядит так:

$$FV = P(1 + r)^n, \tag{1.2}$$

где обозначения аналогичны (1.1).

Во многих финансовых операциях начисление происходит чаще, чем один раз в год. Например, процентные платежи могут добавляться к общей сумме депозита или кредита ежеквартально или ежемесячно. Будущая стоимость денег в этом случае будет выше, так как на проценты, начисляемые через более короткие промежутки времени, процентные платежи начисляются раньше.

Для получения формулы наращения, когда проценты начисляются чаще, чем раз в год, необходимо изменить выражение (1.2). Годовая процентная ставка делится на количество периодов начисления в году, а степень n умножается на количество периодов начисления в году:

$$P\left(1+\frac{r}{m}\right)^{mn}, \tag{1.3}$$

где m — количество периодов начисления в году.

Следующий пример, в котором рассмотрено ежеквартальное начисление процентов в течение трех лет, проиллюстрирует использование этой формулы:

$$1000\left(1+\frac{0{,}06}{4}\right)^{4\cdot 3} = 1000(1{,}015)^{12} = 1195{,}618.$$

Следует заметить, что переход от ежегодного к ежеквартальному начислению процентов влечет за собой увеличение будущей стоимости, или приносит дополнительную прибыль. В нашем случае разница составляет 4,602 за 3 года.

Можно прийти к ложному заключению, что с увеличением m (начисляя проценты чаще) происходит бесконечное увеличение значения будущей стоимости

$$P\left(1+\frac{r}{m}\right)^{mn}.$$

Однако это не так, причиной чего является множитель наращения

$$\left(1+\frac{r}{m}\right)^{mn}, \qquad (1.4)$$

который ограничен в росте по мере увеличения m. Табл. 1.2 демонстрирует это для случая, когда $r = 1$ и $n = 1$. Следует отметить, что так как $n = 1$, то $mn = m$ и поэтому n можно в этом примере проигнорировать.

Таблица 1.2

m	2	3	4	5	10	20	100	1000	10000
$(1 + (1/m))^m$	2,25	2,370	2,441	2,488	2,593	2,653	2,705	2,717	2,718

Таким образом, при $m = 2$ мы имеем $(1 + 0,5)^2 = 2,25$. Увеличивая частоту начислений в году до 10, мы получаем множитель наращения 2,593, увеличивая m до 100, получаем множитель 2,704, а при увеличении m до 1000 — 2,716, и т. д. Важным является предел этого увеличения, выражающийся математической константой, что мы далее увидим. Эта константа — иррациональное число, т.е. имеет бесконечное число знаков после запятой, поэтому она не может быть выражена десятичной дробью. Математики назвали ее экспонентой и обозначили е. Конечно, можно записать приближенное значение е, такое, как 2,71828182845904523536287, но даже это число не является абсолютно точным.

Мы можем обобщить эффект увеличения частоты начислений (m), заметив, что

$$\left(1+\frac{r}{m}\right)^{mn}$$

стремится к $e^{r \cdot n}$ по мере увеличения m.

В пределе можно предположить, что начисления становятся настолько частыми, что проценты начисляются непрерывно, тем самым увеличивая основную сумму в экспоненциальной зависимости. Это известно как **непрерывное наращение**. Будущая стоимость денег при непрерывном наращении определяется как

$$Pe^{r \cdot n}, \qquad (1.5)$$

где e — экспоненциальная константа (2,71828...),

n — количество лет или соответствующие доли лет.

Непрерывное наращение — это допущение, существующее только в теории и применяющееся в финансовых моделях, таких, как определение стоимости опционов. Как будет показано далее, приведение процентных ставок с различной частотой начисления к эквивалентным ставкам с непрерывным наращением позволяет сопоставить их между собой.

Приведение дохода с дискретным наращением к эквивалентному доходу с непрерывным наращением. Для преобразования доходов с дискретным наращением к доходам с непрерывным наращением необходимо, чтобы сумма денег, инвестированная под непрерывную процентную ставку, имела такую же будущую стоимость, что и подобная сумма, инвестированная под **эквивалентную** дискретную процентную ставку. Таким образом:

$$Pe^{r_{cc}n} = P\left(1+\frac{r_{dc}}{m}\right)^{mn}, \qquad (1.6)$$

где r_{cc} — непрерывная процентная ставка,

r_{dc} — эквивалентная дискретная процентная ставка,

n — общий срок (в годах),

m — количество периодов начисления в году.

Деля обе части уравнения на P и извлекая корень степени n, получим:

$$e^{r_{cc}} = \left(1+\frac{r_{dc}}{m}\right)^{m}, \qquad (1.7)$$

преобразовав это уравнение, имеем:

$$r_{cc} = m\ln\left(1 + \frac{r_{dc}}{m}\right), \qquad (1.8)$$

где ln — знак натурального логарифма.

Заметим, что ln — функция, обратная экспоненциальной. Например, e = 2,71828... и ln(2,71828...) = 1, e^2 = 7,38905... и ln(7,38905...) = 2. Соответственно $e^{0,1}$ = 1,10517 и ln1,10517 = 0,1.

Следовательно, для получения процентной ставки с непрерывным наращением необходимо дискретную процентную ставку разделить на количество периодов начисления в году, прибавить 1, прологарифмировать полученный результат и умножить на количество периодов начисления в году.

Следующий пример проиллюстрирует использование формулы. Эквивалентная непрерывная процентная ставка, соответствующая ставке 6% годовых с ежеквартальным начислением, равна:

$$r_{cc} = 4\ln\left(1 + \frac{0{,}06}{4}\right) = 4 \cdot 0{,}014889 = 0{,}059554 \approx 5{,}955\%.$$

Следует заметить, что процентная ставка с непрерывным наращением меньше процентной ставки с дискретным наращением, так как проценты начисляются чаще, и таким образом, сами проценты зарабатывают (или приносят) больше процентов.

Приведение доходов с непрерывным наращением к доходам с дискретным наращением. Мы можем получить эквивалентную дискретную процентную ставку, зная непрерывную процентную ставку, следующим образом:

$$r_{dc} = m\left(e^{r_{cc}/m} - 1\right). \qquad (1.9)$$

Значение экспоненты возводится в степень, равную отношению непрерывной процентной ставки к количеству периодов начисления в году при дискретном наращении. Из полученного результата вычитается единица и разность умножается на количество дискретных периодов начисления в году. В результате получаем эквивалентную ставку при дискретном наращении.

В качестве примера рассмотрим непрерывную процентную ставку 12,5% годовых. Какая дискретная процентная ставка с начислением 4 раза в год эквивалентна данной?

$$r_{dc} = m\left(e^{0{,}125/m} - 1\right) = 4\left(e^{0{,}125/4} - 1\right) = 12{,}7\%.$$

Обратим внимание, что эквивалентная дискретная процентная ставка должна быть больше непрерывной эквивалентной ставки, так как при дискретном наращении проценты добавляются реже и, следовательно, на них в свою очередь начисляется меньше процентов.

Текущая стоимость денег

Уже отмечено выше, что сегодня стоимость некоторой суммы денег, обещанной в будущем, меньше, чем ее будущая стоимость. Даже в мире, характеризующемся отсутствием рисков, это условие по-прежнему будет выполняться, потому что денежные средства могли быть инвестированы по процентной ставке, свободной от риска. Поэтому будущая стоимость денег, инвестированных сегодня, будет больше стоимости этой же суммы, обещанной в будущем. В мире, в котором мы живем и который характерен наличием рисков, существует ряд неопределенностей, касающихся стоимости денег, обещанных в будущем, таких, как инфляция и невыполнение договорных обязательств.

Для сравнения стоимостей различных денежных потоков в разные периоды времени в будущем необходимо дисконтировать будущие потоки наличности и привести их к текущей стоимости. **Текущая стоимость** — это сумма, которая при инвестировании под существующую процентную ставку до определенной даты платежа имела бы стоимость, равную по величине сумме платежа, обещанного в этот момент в будущем.

Дискретное дисконтирование. Во многих финансовых операциях, даже краткосрочных, при дисконтировании используются сложные проценты. В этом случае для дисконтирования будущих денежных потоков и приведения их к текущей стоимости необходимо денежный поток разделить на дисконтный множитель (1 + ставка дисконтирования, выраженная в виде десятичной дроби) в степени, равной количеству лет до получения денежных средств. Это может быть представлено в следующем виде:

$$PV = \frac{CF}{(1+r)^n}, \qquad (1.10)$$

где PV — текущая стоимость будущего денежного потока,
CF — величина денежного потока,
r — ставка дисконтирования,
n — количество лет до поступления денежных средств.

На численном примере продемонстрируем использование формулы (1.10). Предположим, что 1000 единиц будет получено через 5 лет, и текущая рыночная ставка дисконтирования составляет 10% годовых. Текущая стоимость в этом случае будет:

$$PV = \frac{1000}{1{,}10^5} = 620{,}92 \text{ ед.}$$

Если дисконтирование происходит чаще, чем один раз в год, тогда r делится на количество периодов дисконтирования в году, а n умножается на количество периодов дисконтирования в году. Используя те же обозначения, что и в тексте о наращении, получим формулу:

$$PV = \frac{CF}{\left(1 + \frac{r}{m}\right)^{mn}}. \tag{1.11}$$

Теперь рассмотрим приведенный выше пример, но изменим частоту дисконтирования до четырех раз в году:

$$PV = \frac{1000}{\left(1 + \frac{0{,}1}{4}\right)^{20}} = 610{,}27 \text{ ед.}$$

Таким образом, текущая стоимость данного денежного потока уменьшается с увеличением частоты дисконтирования при одной и той же процентной ставке.

Непрерывное дисконтирование. Так же, как и в случае с наращением, период между процессами дисконтирования может быть уменьшен до такой степени, что дисконтирование происходит непрерывно.

Формула для непрерывного дисконтирования выглядит так:

$$PV = CFe^{-r \cdot n}, \tag{1.12}$$

т.е. при непрерывном дисконтировании e возводится в отрицательную степень $-r \cdot n$. Например, текущая стоимость 1000 единиц, получаемая через 5 лет и при непрерывном дисконтировании в 10% годовых, равна:

$$1000e^{(-0{,}1 \cdot 5)} = 606{,}53 \text{ ед.}$$

Процентные ставки спот, форвардные процентные ставки и качественные спреды

До сих пор мы рассматривали процессы наращения и дисконтирования. Сейчас же мы обсудим, как определяются реальные процентные ставки, использующиеся в этих процессах. Начнем с двух основных и важных видов процентных ставок — спот и форвардных ставок.

Процентная ставка спот — это процентная ставка, которая используется на финансовых рынках для дисконтирования будущих денежных потоков и приведения их к **текущей стоимости**. Возьмем, например, денежный поток через один год. Одногодичной спот-ставкой будет ставка, по которой этот денежный поток будет дисконтирован и приведен к его текущей рыночной стоимости. Если же мы возьмем денежный поток, скажем, через 5 лет, ставкой, по которой эта сумма будет дисконтирована и приведена к текущей рыночной стоимости, будет пятилетняя спот-ставка, и т.д.

Обозначив одногодичные и пятилетние спот-ставки через r_1 и r_2 и предполагая **непрерывное дисконтирование**, определим текущую стоимость по:

$$PV_1 = CF_1 e^{-r_1},$$

$$PV_5 = CF_5 e^{-r_5 \cdot 5},$$

а спот-ставка за период n рассчитывается так:

$$r_n = \frac{\ln\left(\dfrac{CF_n}{PV_n}\right)}{n}. \qquad (1.13)$$

Однако большинство финансовых рынков на практике используют **дискретное дисконтирование**, поэтому спот-ставкой периода n будет значение r_n в следующем уравнении:

$$PV_n = \frac{CF_n}{(1+r_n)^n}. \tag{1.14}$$

Типичным примером применения спот-ставок является нахождение доходности по облигациям с нулевым купоном. По таким облигациям не выплачиваются периодические процентные платежи и они погашаются по номинальной стоимости в конце срока. Поэтому эти облигации выпускаются с дисконтом к номинальной стоимости, и во время всего периода обращения торговля по ним производится с учетом дисконта, так как в любой момент времени текущая цена облигаций соответствует современной стоимости выкупного платежа.

Процентная ставка, используемая на рынке для дисконтирования выкупной цены облигации, является спот-ставкой соответствующего периода до погашения. Однако эта спот-ставка не фигурирует во время торгов на рынке, определяется лишь текущая цена облигаций с нулевым купоном. Спот-ставку можно рассчитать, зная текущую стоимость облигации, из следующего уравнения:

$$r_n = \left(\frac{CF_n}{PV}\right)^{\frac{1}{n}} - 1, \tag{1.15}$$

где r_n — спот-ставка периода n;
CF_n — стоимость при погашении через n лет (номинальная стоимость);
PV — современная стоимость, или текущая цена облигации с нулевым купоном и со сроком до погашения n лет.

Предположим, что облигация с нулевым купоном имеет срок до погашения 30 лет. Номинальная цена облигации равна 100 единицам, а в настоящее время ее рыночная цена равна £16,97. Тогда 30-летняя спот-ставка при расчете по формуле (1.15) равна:

$$r_n = \left(\frac{100}{16,97}\right)^{\frac{1}{30}} - 1 = 6,09\%.$$

Форвардная процентная ставка — это процентная ставка, зафиксированная сегодня для ссуды, предоставляемой в будущем. Таким образом, процентная ставка для трехмесячного займа, предоставляемого через три месяца — это трехмесячная форвардная процентная ставка на три месяца. Если ссуда одногодичная и предоставляется через 12 месяцев, то процентная ставка для нее будет одногодичная форвардная ставка на 12 месяцев. На денежных рынках эти ставки известны как форвардные форварды и обозначаются в виде 3/6 или 12/24, что означает начало будущих финансовых операций через 3 и 12 месяцев, а окончание соответственно через 6 и 24 месяца. Эти ставки могут также называться как три против шести или 12 против 24.

Форвардные ставки не котируются напрямую на рынке. Однако относительно денежных рынков они могут быть получены по информации о краткосрочных займах и депозитах. Для рынка облигаций, включая казначейские векселя, форвардные процентные ставки можно определить исходя из соотношения между последовательными спот-ставками.

Непосредственный расчет форвардных ставок зависит от того, простые или сложные проценты применяются на данном конкретном рынке. В случае сложных процентов расчет зависит также от того, применяется непрерывное или дискретное наращение. Простые проценты на финансовых рынках применяются к финансовым инструментам со сроком менее одного года. На рынках облигаций используется дискретное наращение, а для опционов облигаций — непрерывное наращение.

Форвардные процентные ставки в случае простых процентов

На финансовых рынках форвардная ставка является ставкой на вклад или ссуду, предоставляемую в некоторый момент в будущем. Форвардная ставка по депозиту может быть создана путем заимствования на короткий срок и помещения этих денег на депозит на длительный срок. Форвардный заем, или кредит, создается в случае займа на длительный срок и помещения этой суммы на депозит на более короткий срок.

Например, если ваш банк предлагает вам процентную ставку на трехмесячный депозит через три месяца, то в этом случае банк предлагает вам форвардный депозит. Банк должен захеджировать форвардное обязательство, данное вам, заняв деньги на три месяца и одновременно ссудив эту же сумму на шесть месяцев. Банковский баланс будет уравновешен в течение первых трех месяцев, так как ссуда соответствует займу. Однако привлеченные средства (депозит) должны быть возвращены через три месяца, тогда как ссуда не будет возвращена в течение шести месяцев. Поэтому банк будет использовать депозит, обещанный вами, для того чтобы уравновесить баланс второго трехмесячного периода.

Безубыточной процентной ставкой, которую банк сможет предложить вам по форвардному депозиту, будет ставка, соответствующая разнице полученных банком процентов по шестимесячной ссуде и процентов, которые он уплатит по трехмесячному займу.

Рассмотрим теперь ситуацию, когда вы попросили банк о трехмесячном кредите через три месяца с процентной ставкой, установленной сегодня. Банк постарается захеджировать свою позицию, взяв шестимесячный заем и ссудив эти средства на три месяца. Через три месяца эта ссуда будет возвращена банку, который использует ее для вашего кредитования. Когда через шесть месяцев вы вернете кредит, банк использует эти средства для погашения своего шестимесячного займа. Безубыточная ставка, по которой банк будет кредитовать вас, соответствует разнице издержек банка по шестимесячному займу и доходов по ссуде за первые три месяца.

Конкретный пример поможет понять нам это. Предположим, что процентная ставка по трехмесячному межбанковскому кредиту составляет 6,0—6,1% годовых, а соответствующая ставка по шестимесячному кредиту — 6,15—6,25%, т.е. банки берут деньги на три месяца под 6,0% и ссужают их под 6,1%, аналогично банки берут шестимесячные кредиты под 6,15% и ссужают их под 6,25%. Необходимо подсчитать процентную ставку для форвардного депозита, предлагаемого банком. Издержки по данному виду хеджа вычисляются исходя из следующего: если банк предоставляет ссуду размером в 100 единиц на шесть месяцев под 6,25%, его доходы составят 3,125 единиц. Одновременно банк занимает деньги на три месяца под 6,0% годовых, и его издержки равны 1,5 единицам. Следовательно, размер процентов, которые могут быть уплачены по депозиту в течение вторых трех месяцев, равен 1,625 единиц, что составляет 6,5% годовых (при умножении на 4).

В действительности приведенный пример имеет несколько неточностей:

а) данная процедура пересчитывает форвардную ставку в годовом исчислении, однако нет никакой уверенности, что деньги можно перезанять под ту же процентную ставку в течение оставшихся периодов года;

б) для форвардного кредита банк получает деньги на длительный срок, а ссужает на короткий период. Процентные платежи, получаемые по ссуде, могли бы быть реинвестированы в течение вторых трех месяцев, и таким образом снижена безубыточная процентная ставка кредита, т.е. безубыточная форвардная ставка. Однако первоначально о реинвестировании ничего не известно. Эта неопределенность может быть уменьшена посредством оценки форвардной ставки в виде чистых доходов по кредиту, выраженных в процентах в годовом исчислении от основной суммы и издержек;

в) в случае форвардного депозита банк занимает деньги на короткий срок, а ссужает на длительный. Издержки по займу в первом периоде должны быть профинансированы во втором, тем самым сокращены платежи по депозиту, т.е. сокращена безубыточная форвардная ставка. Форвардная ставка по форвардному депозиту может быть оценена как чистые издержки, выраженные в процентах в годовом исчислении от основной суммы и доходов.

Для пояснения рассмотрим процесс нахождения процентной ставки для трехмесячного форвардного кредита, начинающегося через три месяца. Допустим, что на денежном рынке для расчетов принято использовать 360 дней в году. Издержки банка по шестимесячному займу составляют 6,25% год, а ставка по трехмесячному реинвестированию — 6,0% в год. Чистые издержки по сумме 100 единиц равны:

$$6{,}25 \cdot \frac{360}{180} - 6{,}0 \cdot \frac{360}{90} = 3{,}125 - 1{,}50 = 1{,}625 \text{ ед.}$$

Основная сумма плюс издержки составит: 100 + 1,50 = 101,50 единиц, тогда форвардная ставка будет равна:

$$\frac{1{,}625}{101{,}5} \cdot \frac{360}{90} = 0{,}064039 \approx 6{,}4\%.$$

Общая формула для нахождения форвардных ставок (RF) на финансовых рынках, учитывающая указанные неопределенности, выглядит следующим образом:

$$RF = \frac{1}{n-m} \cdot \frac{(n \cdot r_n) - (m \cdot r_n)}{1 + \frac{(m \cdot r_m)}{360}}, \tag{1.16}$$

где n — продолжительность в днях более длинного финансового инструмента;
m — продолжительность в днях более короткого финансового инструмента;
r_n — процентная ставка по длинному финансовому инструменту;
r_m — процентная ставка по короткому финансовому инструменту.

Для иллюстрации рассмотрим этот же пример:

$$RF = \frac{1}{90} \cdot \frac{(180 \cdot 0{,}0625) - (90 \cdot 0{,}06)}{1 + \frac{(90 \cdot 0{,}06)}{360}} = 0{,}064039 \approx 6{,}4\%.$$

Соглашения по подсчету дней в году особенно важны при расчете форвардных процентных ставок. Пример, приведенный выше, характерен для большинства финансовых рынков, не торгующих фунтами стерлингов. Если бы расчеты производились для английского межбанковского рынка, необходимо было бы принять 365 дней в году. Соглашения по подсчету дней в году далее будут обсуждены более детально.

Форвардные ставки в случае непрерывного наращения

В случае непрерывного наращения форвардная ставка определяется следующим образом:

$$RF_n = \frac{nr_n - mr_m}{n-m}. \tag{1.17}$$

Применение формулы (1.17) может быть проиллюстрировано на следующем примере. Предположим, что 180-дневная процентная ставка равна 11% годовых, а 90-дневная — 10% годовых.

Форвардная ставка по 90-дневному финансовому инструменту через 90 дней рассчитывается так:

$$RF = \frac{(180 \cdot 11) - (90 \cdot 10)}{180 - 90} = 12\%.$$

В случае дискретного наращения найти форвардную ставку можно с помощью формулы (1.18), где через r_n и r_m обозначены соответствующие ставки-спот, а n и m —сроки, измеряемые в годах:

$$RF_n = \sqrt[n-m]{\frac{(1+r_n)^n}{(1+r_m)^m}} - 1^{n-m}\!\sqrt{\ } . \qquad (1.18)$$

Численный пример поможет понять данную формулу. Предположим, что трехгодичная спот-ставка составляет 12,75% годовых, а двухгодичная спот-ставка — 11,63% годовых. Форвардная ставка равна:

$$RF = \sqrt[3-2]{\frac{(1{,}1275)^3}{(1{,}1163)^2}} - 1 = 15{,}04 .$$

В этом примере (n—m) равно единице, так как разница между двумя периодами составляет один год. Если бы мы использовали ставки, где разница составляла бы, скажем, 6 или 3 месяца, тогда (n—m) было бы равно 0,5 или 0,25 соответственно.

Очевидно, что так как мы живем в мире, полном неопределенностей, форвардные ставки, рассчитанные по этим формулам, могут не быть равны, а точнее, не будут равны тем краткосрочным ставкам, которые реально установятся в будущем. Что действительно выражают форвардные ставки, так это текущие ожидания рынка относительно процентных ставок в будущем.

Для пояснения рассмотрим предлагаемые на рынке облигации с нулевым купоном и со сроком до погашения один и два года. Инвестор, желающий разместить средства на два года, имеет выбор: либо инвестировать средства на один год, а затем вновь вложить деньги в тот же инструмент во втором году, либо же сразу инвестировать средства в двухгодичный инструмент.

Выбор инвестора обусловлен сравнением форвардной ставки по двухгодичному инструменту со своими ожиданиями относительно того, какая доходность по одногодичным инструментам будет через год. Если инвестор ожидает, что доходность через год будет выше текущей двухгодичной спот-ставки, он приобретет годовую облигацию и спустя год пролонгирует данный инструмент. Если же ожидается, что ставки в будущем будут ниже предлагаемых форвардных ставок, инвестор предпочтет двухлетнюю облигацию.

На рынках, где инвесторы могут выбрать финансовые инструменты в соответствии со своими ожиданиями, форвардная ставка отражает текущие ожидания рынка относительно будущей спот-ставки.

Сравнение спот- и форвардных процентных ставок

В табл. 1.3 приведены цены инструментов с нулевым купоном, соответствующие спот-ставки и одногодичные форвардные ставки, рассчитанные исходя как из непрерывного, так и из дискретного наращения (CC и DC соответственно).

Таблица 1.3. Спот- и форвардные процентные ставки

Срок до погашения	Цена инструмента с нулевым купоном	Спот-ставка (CC)	Спот-ставка (DC)	Форвардная ставка (CC)	Форвардная ставка (DC)
1	90,48	10,0	10,52	—	—
2	80,25	11,0	11,63	12,0	12,75
3	69,76	12,0	12,75	14,0	15,04
4	59,45	13,0	13,88	16,0	17,34
5	49,66	14,0	15,03	18,0	19,71

Для того чтобы избежать ошибок из-за округления, значения последней колонки могут быть рассчитаны напрямую, исходя из цен облигаций, а не из подсчитанных спот-ставок. Например:

$$19{,}71 = 100\left(\frac{59{,}45}{49{,}66} - 1\right).$$

Спот-ставки — это ставки дисконтирования, но одновременно они также и годовые ставки доходности от инвестиций в облигации с нулевым купоном по установленным ценам. Например, доходность от инвестирования в облигацию с нулевым купоном и со сроком до погашения в один год равна:

$$(100 - 90{,}48)/90{,}48 = 0{,}\ 1052, \text{ или } 10{,}52\%.$$

Очевидна взаимосвязь между форвардными и спот-ставками. В самом деле, можно видеть, что спот-ставка — это геометрическая средняя между текущей краткосрочной спот-ставкой и всеми краткосрочными форвардными ставками, соответствующими данному сроку до погашения. Например, пятилетняя спот-ставка, равная 15,03% (дискретное наращение), согласно сказанному является средней геометрической из текущей одногодичной спот-ставки и всех четырех одногодичных форвардных ставок, приведенных в табл. 1.3, т.е. ее величина определяется из следующего выражения:

$$\sqrt[5]{(1{,}1052)(1{,}1275)(1{,}1504)(1{,}1734)(1{,}1971)} - 1 = 0{,}1503, \text{ или } 15{,}03\%.$$

Спреды, отражающие различие в качестве

Напомним, что процентная ставка компенсирует потерю выгоды вследствие потери покупательной способности и различных видов риска, связанных с кредитованием, поэтому будет разумным допустить, что чем больше потери выгоды, тем больше должна быть процентная ставка, чтобы компенсировать эту потерю. В самом деле, мы видим, что кредиторы устанавливают тем бóльшие процентные ставки, чем выше их ожидания относительно уровня инфляции, риска невыполнения обязательств и чем ниже их ощущение "качества" инвестиций (например, ликвидности рынка).

Ожидаемая инфляция повлияет на все процентные ставки, но добавочная компенсация за риск и другие количественные отличия специфичны для каждой конкретной инвестиции. Спот-ставка определяется исходя из ставок безрисковых облигаций, примером которых является государственная облигация, деноминированная в валюте данной страны. Считается, что подобные облигации не имеют риска, так как правительство контролирует налогообложение и имеет полномочия по дополнительной денежной эмиссии для того, чтобы выполнить свои обязательства.

Облигации и другие ценные бумаги, выпущенные акционерными компаниями и другими неправительственными организациями, характеризуются наличием риска невыполнения обязательств, поэтому инвесторы требуют по ним процентную ставку больше, чем процентная ставка по аналогичным безрисковым облигациям. Эта добавочная доходность известна как качественный спред. Часть этого качественного спреда возникает из-за риска невыполнения обязательств и называется премией за риск. Однако качество инвестиций определяется не только риском невыполнения обязательств, а также, например, ликвидностью конкретного рынка и тем, облагается ли налогом доход по данному инструменту. Следовательно, качественный спред определяется не только премией за риск. На рис. 1.1 приведены различные кривые доходности.

Рис. 1.1. Кривые спот-ставок

ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ ПРОЦЕНТНЫХ СТАВОК НА ФИНАНСОВЫХ РЫНКАХ

Оценка финансовых инструментов

Выше в этой главе отмечено, что финансовые инструменты оцениваются исходя из суммы текущих стоимостей всех ожидаемых будущих денежных потоков, а фьючерсные и форвардные контракты — по будущей стоимости текущего денежного потока. Здесь мы продемонстрируем использование процентных ставок в процессах дисконтирования и наращения в оценке банковских депозитных сертификатов, казначейских векселей, облигаций с нулевым купоном, купонных облигаций и акций компаний. Помимо этого мы рассмотрим, как оцениваются фьючерсные и форвардные контракты. В первую очередь мы должны познакомиться с некоторыми рыночными соглашениями относительно методов наращения по основным долговым обязательствам, существующим на финансовых рынках.

Соглашения по подсчету дней и простые проценты

Одна из сложностей, которая возникает при вычислении простых процентов, — это соглашения, принятые на данном конкретном рынке по подсчету количества дней. По всем финансовым инструментам наращение происходит ежедневно, однако существуют расхождения между разными рынками в подсчете процентов — одни берут за основу 360 дней в году, а другие — 365 дней или 366 в високосном году. Кроме того, на некоторых рынках принимается в месяце 30 дней, независимо от их реального количества. В Великобритании по денежным инструментам, деноминированным в фунтах стерлингов, используют 365 дней в году при начислении процентов. Во многих других финансовых центрах, а также в Лондоне по финансовым инструментам, деноминированным не в фунтах стерлингов, расчет ведется из 360 дней в году.

Таким образом, в случае 360 дней в году процентные платежи, уплачиваемые по, скажем, 45-дневному депозиту, рассчитываются исходя из общего срока, равного 45/360 лет. Этот метод обозначается как **ACT/360** (actual over 360). Если же процентные платежи рассчитываются исходя из 365 дней в году, тогда для 45-дневного депозита срок в годах выражается через 45/365. Этот метод носит название **ACT/365** (actual over 365).

Для иллюстрации этих методов рассмотрим рынок европейских валют. Допустим, что на этом рынке действует соглашение по использованию 360 дней в году, но проценты начисляются исходя из реального количества дней, т.е. применяется ACT/360. На депозит, размещенный на два 31-дневных месяца, проценты будут начислены исходя из 62 дней, следовательно, доходность будет выражаться так:

$$(62/360) \cdot \text{процентная ставка.}$$

Если процентная ставка по 62-дневному депозиту составляет 6% годовых, то процентные платежи по депозиту размером в 10 000 000 единиц равны:

$$10\,000\,000 \cdot \frac{62}{360} \cdot 0{,}06 = 10\,000\,000 \cdot 0{,}0103333 = 103\,333{,}33.$$

Таким образом, начисленные проценты составляют 103 333 единиц. Сравним этот результат с ситуацией на английском внутреннем межбанковском рынке, где проценты рассчитывают исходя из 365 дней в году. В этом случае проценты составят:

$$10\,000\,000 \cdot \frac{62}{365} \cdot 0{,}06 = 10\,000\,000 \cdot 0{,}010191781 = 101\,917{,}81.$$

Следовательно, 6%-ный депозит на рынке европейских валют — не то же самое, что 6%-ный депозит на некоторых национальных межбанковских рынках, таких, как английский рынок, где проценты начисляются по методу ACT/365. Для того чтобы было возможно сравнить эти годовые ставки, необходимо ставку по евровалюте умножить на 365/360 или же ставку по внутреннему национальному межбанковскому кредиту умножить на 360/365. Таким образом, 6%-ная годовая ставка на европейском рынке соответствует 6,083% [(365/360) · 6] годовых, тогда как на национальном межбанковском рынке Великобритании процентная ставка составляет лишь те самые установленные 6%.

Принимая во внимание существующие различия по подсчету дней как между рынками, так и между финансовыми инструментами, следует различать **номинальную** и **эффективную** процентные ставки. Номинальная ставка — это опубликованная или еще каким-либо образом объявленная процентная ставка. Эффективная же ставка соответствует реальной сумме полученных инвестором платежей, отнесенных к инвестированной сумме.

Ясно, что крайне важно определиться с методами и соглашениями по подсчету дней, принятыми на данном конкретном рынке или же по отношению к конкретным ценным бумагам. Следует отметить, что упомянутые соглашения существуют также и в отношении начисляемых процентов по облигациям, что рассмотрено далее в этой главе.

Если на денежном рынке принято соглашение по применению простых процентов при расчете текущей и будущей стоимости, значит свое отражение должны найти и соглашения по подсчету дней. Вследствие этого выражение (1.14) должно быть изменено таким образом, чтобы его можно было правильно использовать на рынке:

$$PV_n = \frac{CF_n}{\left[1 + \left(\frac{D}{B} \cdot r_n\right)\right]}, \qquad (1.19)$$

где D — количество дней дисконтирования по данному инструменту;

B — 360 или 365 в соответствии с принятым соглашением на данном рынке;

PV — текущая стоимость;

CF — будущий денежный поток, будущая стоимость или выкупная цена;

r_n — годовая простая процентная ставка на период n, используемая для дисконтирования.

Сейчас мы готовы приступить к рассмотрению методов оценки финансовых инструментов, упомянутых выше.

Банковские депозитные сертификаты

Депозитные сертификаты (certificates of deposit — СД) выдаются банком в обмен на средства, размещаемые у них. Подобные депо-ты отличаются от обычных банковских депозитов лишь только , что сертификаты, подтверждающие наличие срочного депо-в банке, могут обращаться на вторичном рынке. *СД* выдаются чо номинальной стоимости, а проценты выплачиваются при чии, когда возвращается основная сумма.

ютрим два СД, оба деноминированы в фунтах стерлин-тин выпущен на европейском рынке, а другой — на Каждый из сертификатов имеет срок действия один ется уже в течение 91 дня, ставка доходности по сертификатам составляет 6%. Стоимость СД размером в £1 000 000 при погашении составит: первоначальная сумма + накопленные за 365 дней проценты. Однако по СД на внутреннем рынке проценты будут начисляться исходя из метода ACT/365, когда как по евродепозитному сертификату будет применяться метод ACT/360. Тогда будущие стоимости каждого из них равны:

$$FV_{CD} = 1\ 000\ 000 \cdot \left[1 + \left(0{,}06 \cdot \frac{365}{365}\right)\right] = 1\ 060\ 000,$$

$$FV_{евроCD} = 1\ 000\ 000 \cdot \left[1 + \left(0{,}06 \cdot \frac{365}{360}\right)\right] = 1\ 060\ 833{,}33.$$

Заметим, что эффективная доходность по евроСД будет $6 \cdot (365/360) = 6{,}0833$.

Депозитный сертификат, обращаясь на вторичном рынке, оценивается исходя из текущей стоимости ожидаемых будущих денежных потоков. Ожидаемые будущие денежные потоки рассчитывают (при отсутствии риска невыполнения обязательств) как было показано выше.

Ставка дисконтирования для расчета текущей стоимости будущих денежных потоков должна соответствовать рыночной ставке аналогичных инструментов при наличии одинакового уровня риска. Предположим, что доходность аналогичных депозитных сертификатов, имеющих такой же уровень рискованности и со сроком 274 (365—91) дня, как на внутреннем, так и на внешнем рынке возросла до 6,1%.

Текущей ценой данных депозитных сертификатов будет стоимость по окончании срока (рассчитанная как было показано выше), дисконтированная по ставке 6,1%. Эффективная доходность по евроСД составит 6,1847% ($6{,}1 \cdot 365/360$).

$$P_{CD} = \frac{1\ 060\ 000}{\left[1 + \left(0{,}061 \cdot \frac{274}{365}\right)\right]} = 1\ 013\ 586,$$

$$P_{евроCD} = \frac{1\ 060\ 833{,}33}{\left[1 + \left(0{,}061 \cdot \frac{274}{365}\right)\right]} = 1\ 013\ 766.$$

Казначейские векселя

Казначейский вексель (treasusy bill — ТВ) — общий термин для обозначения государственных краткосрочных ценных бумаг со сроком обращения обычно 90 дней, но казначейские векселя могут также иметь сроки как от одного месяца, так и до одного года. Обычно рынок казначейских векселей высоколиквиден, а их эмиссия ТВ происходит довольно часто, как правило, еженедельно.

Казначейские векселя выпускаются по дисконтированной стоимости, а погашаются по номиналу, таким образом, дисконт представляет собой доход инвестора. На наличном рынке казначейских векселей их котировка происходит в форме ставок дисконтирования, а не в форме цены. Типичным котировочным предложением может быть 7,0—6,9, что означает желание трейдера покупать ТВ по дисконтной ставке 7% годовых, а продавать — по 6,9% год.

Доходность казначейского векселя находится при известном дисконте следующим образом:

$$\text{Доходность ТВ} = \left(\frac{\text{Дисконт}}{\text{Номинал}}\right) \cdot \left(\frac{365}{n}\right) \qquad (1.20)$$

при допущении ACT/365 метода.

При вычислении доходности ТВ важно обратить внимание на общий срок обращения векселя и соответствующий метод подсчета дней, о чем говорилось выше.

Цена, уплаченная за казначейский вексель, называется **фактурной ценой** (invoice price). Эта дисконтированная цена (INVP) определяется следующим образом:

$$INVP = PAR \cdot \left[1 - r\left(\frac{n}{365}\right)\right], \qquad (1.21)$$

где PAR — номинальная стоимость казначейского векселя;
r — доходность ТВ, выраженная в виде десятичной дроби;
n — число дней до погашения.

Для демонстрации расчета фактурной цены ТВ рассмотрим пример, относящийся к американскому рынку, где принято использовать при подсчете дней метод ACT/360. Период до погашения казначейского векселя равен 44 дням, а текущая доходность казначейского векселя номиналом \$1 000 000 составляет 8,25%. Фактурная цена равна:

$$INVP = 1\ 000\ 000 \cdot \left[1 - 0{,}0825\left(\frac{44}{360}\right)\right] = \$989\ 916{,}67.$$

Облигации с нулевым купоном и купонные облигации

Облигация с нулевым купоном — это облигация, по которой не выплачиваются периодические процентные платежи (известные как купоны) в течение срока ее обращения, отсюда и термин "нулевой купон". Этот вид облигаций выпускается и обращается с дисконтом, так как единственный ожидаемый будущий денежный поток — это номинальная или выкупная цена облигации в конце срока обращения. Доход образуется за счет разницы между текущей стоимостью облигации в момент ее эмиссии или приобретения и стоимостью в момент погашения или продажи. Таким образом, настоящей (текущей) стоимостью или ценой данных облигаций является текущая стоимость выкупного платежа. Формула для расчета текущей стоимости облигации следующая:

$$P_z = \frac{CF_T}{(1 + r_T)^T}, \tag{1.22}$$

где P_z — текущая цена облигации с нулевым купоном,
CF_T — денежный поток через время T (выкупная цена),
r_T — спот-ставка для дисконтирования платежей, причитающихся через время T.

Например, предположим, что облигация со сроком погашения через 5 лет имеет выкупную стоимость, равную 100 единицам, а рыночная ставка для дисконтирования подобных платежей, т.е. пятилетняя спот-ставка, равна 8% годовых. Текущая цена облигации с нулевым купоном определяется так:

$$P_z = \frac{100}{(1{,}08)^5} = 68{,}06 \text{ ед.}$$

Купонные облигации, по которым уже выплачены предпоследние процентные платежи, имеют тот же эффект, что и облигации с нулевым купоном, так как последний купон и выкупная стоимость объединены в единый платеж в будущем периоде. Например, облигация с 10%-ным купоном и выплатой процентов два раза в год погашается по номиналу через четыре месяца. Последний денежный поток по этому инструменту равен 105 единицам. При расчете текущей цены сумма 105 будет дисконтирована по четырехмесячной спот-ставке.

Определение стоимости облигации с нулевым купоном, т.е. дисконтирование будущего платежа по соответствующей ставке, — один из основных элементов вообще в определении стоимости всех видов облигаций. Истинность этого утверждения проиллюстрирована далее.

Оценка стоимости купонных облигаций

По большинству выпускаемых облигаций периодически выплачиваются проценты, называемые купонами. На большинстве рынков государственных и корпоративных облигаций купоны выплачиваются раз в полгода, тогда как на рынках еврооблигаций — один раз в году.

Таким образом, держатель облигации получает право на серию процентных платежей и выплату основной суммы при погашении. Каждый из этих платежей является обособленным денежным потоком в определенный момент в будущем, поэтому отдельный платеж по своей сути не отличается от платежа по облигации с нулевым купоном. Текущая стоимость такого денежного потока находится с помощью дисконтирования по соответствующей процентной ставке. Другими словами, облигацию можно оценить, если рассмотреть ее как портфель облигаций с нулевым купоном. Каждый купон и выкупная стоимость индивидуально дисконтируются по соответствующим спот-ставкам, которые не обязательно должны быть одинаковы для всех платежей. Затем рассчитанные таким образом стоимости суммируются, в результате чего определяется текущая цена облигации.

Математически это можно выразить так:

$$P_B = \sum_{1}^{T} \frac{CF_T}{(1 + r_T)^T}. \quad (1.23)$$

В качестве примера рассмотрим облигацию с пятилетним сроком погашения, по которой выплачивается ежегодный доход в размере 10 единиц, а стоимость ее погашения равна 100 единицам. Спот-ставки на пять лет составят соответственно 10, 11, 12, 13, и 14% в год.

Дисконтированная величина денежных потоков равна:

$$P_B = \frac{10}{1{,}1} + \frac{10}{(1{,}11)^2} + \frac{10}{(1{,}12)^3} + \frac{10}{(1{,}13)^4} + \frac{110}{(1{,}14)^5} = 87{,}59 \text{ ед.}$$

В данном примере подразумевается, что купон был только что выплачен, а выплата следующего состоится только через год. В действительности же облигации должны оцениваться и в периоды между выплатами купонов. В этом случае расчеты также не представляют особых затруднений. Предположим, что по облигации выплачиваются годовые купоны. Если выплата следующего купона будет через три месяца, тогда он должен быть дисконтирован по трехмесячной спот-ставке, а следующий купон, выплата которого состоится через 15 месяцев, дисконтируется по пятнадцатимесячной спот-ставке и т.д.

Для пояснения рассмотрим облигацию со сроком до погашения 4,25 года и выкупной ценой 100 единиц, по которой выплачивается 8% ежегодно. Последний купон был выплачен девять месяцев назад, значит, выплата следующего купона состоится через три месяца, или 0,25 года. Трехмесячная спот-ставка, а также спот-ставки для оставшихся ежегодных платежей равны 9; 10,5; 11,1; 12,2 и 13,5% соответственно. Цена облигации определяется следующим образом:

$$P = \frac{8}{(1{,}09)^{0{,}25}} + \frac{8}{(1{,}105)^{1{,}25}} + \frac{8}{(1{,}111)^{2{,}25}} + \frac{8}{(1{,}122)^{3{,}25}} + \frac{108}{(1{,}135)^{4{,}25}} = 89{,}76 \text{ ед.}$$

Обыкновенные акции

Один из методов оценки обыкновенных акций состоит в расчете текущей стоимости всех ожидаемых будущих дивидендов, т.е. расчете по так называемой модели дисконтирования дивидендов, рассмотренной Уильямсом (1938), Гордоном (1962), а также Фуллером и Хсиа (1984).

Хотя стоимость акций обычно рассматривается как функция ожидаемых будущих доходов, **модель дисконтирования дивидендов** расценивает дивиденды как показатель доходов, и таким образом учитывает будущие доходы. Это допущение приемлемо при условии, что фирма может либо выплатить прибыль в качестве дивидендов, либо реинвестировать ее в производство. В случае реинвестирования и одновременного принятия политики по увеличению дивидендов будущие дивиденды по величине будут больше текущих. Дивиденды будут расти до тех пор, пока часть прибыли направляется на развитие бизнеса. Теоретическая цена обыкновенной акции может быть выражена в следующем виде:

$$P_0 = \frac{D_1}{1+k} + \frac{D_1(1+g_2)}{(1+k)^2} + \frac{D_1(1+g_2)(1+g_3)}{(1+k)^3} + \ldots + \frac{D_1(1+g_2)\cdot\ldots\cdot(1+g_n)}{(1+k)^n}, \tag{1.24}$$

где P_0 — теоретическая цена акции;

D_i — дивиденды, выплачиваемые в конце периода i;

g_i — темп роста дивидендов или доходов за период i;

k — рыночный коэффициент корректировки прибыли в соответствии с риском, или рыночная ставка дисконтирования будущих денежных потоков.

Таким образом, текущая цена — это текущая стоимость всех будущих дивидендов. Будущая величина дивидендов находится наращением суммы дивидендов, выплачиваемых в конце текущего периода, в соответствии с последующими ожидаемыми темпами роста дивидендов. Например, дивиденды, выплачиваемые в конце третьего периода, находятся как ожидаемые дивиденды в конце первого периода, наращенные за один период исходя из темпов роста g_2, а затем наращенные для третьего периода исходя из темпов роста g_3.

Так как прибыль может быть реинвестирована или же выплачена в качестве дивидендов, то $D = E \cdot (1-b)$. Таким образом, уравнение (1.24) может быть скорректировано в соответствии с этим следующим образом:

$$P_0 = \frac{E_1(1-b)}{1+k} + \frac{E_1(1-b)(1+g_2)}{(1+k)^2} + \frac{E_1(1-b)(1+g_2)(1+g_3)}{(1+k)^3} + \ldots$$

$$\ldots + \frac{E_1(1-b)(1+g_2)\cdot\ldots\cdot(1+g_n)}{(1+k)^n}, \qquad (1.25)$$

где E_i — прибыль за период i;
b — часть прибыли, не выплаченная в виде дивидендов.

В приведенном уравнении подразумевается, что процент b постоянный, но он может и меняться от одного периода к другому.

Как находятся темпы роста прибыли и соответственно дивидендов? Этот показатель определяется исходя из объема новых инвестиций и ставки доходности по данному виду инвестирования. При отсутствии внешнего финансирования g является функцией доли нераспределенной прибыли и доходности по акциям r, т.е.:

$$g = br.$$

Рост прибыли и соответственно дивидендов зависит от того, имеются ли в распоряжении у фирмы финансовые средства для осуществления инвестиций, а также от того, существует ли вообще возможность для инвестиций. При внешнем финансировании рост будет усилен посредством доходов, получаемых за счет внешнего финансирования проектов. Тогда $g = r(b + f)$, где f показывает, какую долю от прибыли составляют внешние средства.

Так как акции являются бессрочными, расчет по моделям (1.24) и (1.25) трудно выполним, потому что требуется осуществить прогноз значений переменных вплоть до бесконечности. Ряд подходов предложен для того, чтобы преодолеть данное затруднение, но их рассмотрение лежит за пределами этой книги. Заинтересованные читатели могут обратится к трудам Уотшэма (1993).

Определение стоимости форвардных и фьючерсных контрактов

Форвардные и фьючерсные контракты — это контракты на покупку или продажу определенного количества какого-либо актива на определенную дату в будущем, но по цене, установленной на момент заключения контракта. Выше мы уже отметили, что соглашение, относящееся к процентной ставке на кредит или депозит на некоторый момент в будущем, называется форвардной процентной ставкой. Соглашение на покупку или продажу валюты в определенный момент в будущем, но по согласованному сегодня курсу, называется форвардным валютным контрактом. Аналогичные соглашения существуют также и в отношении облигаций и акций.

Разница между форвардными и фьючерсными контрактами заключается просто в том, что фьючерсами торгуют на специально организованных фьючерсных биржах, тогда как форварды относятся ко вторичному рынку (например, форвард между банком и его клиентом). Разница в ценах по фьючерсам и соответствующим им форвардам на один и тот же вид актива несущественна, поэтому в дальнейшем мы будем употреблять термин "фьючерс" также и по отношению к форварду.

А сейчас представляется необходимым объяснить понятие арбитража. Арбитраж — это процесс одновременной покупки или продажи одного и того же актива, по которому установлены отличающиеся друг от друга цены на разных рынках, что позволяет получать безрисковую прибыль. Например, акция компании "А" котируется на рынках в Лондоне и в Нью-Йорке. В Лондоне цена на акцию равна £2,0, а в Нью-Йорке — $3,0. Текущий обменный курс составляет $1,40 за один фунт стерлингов. Таким образом, акция имеет разную стоимость на двух рынках. Если мы предположим, что издержки по проведению операции отсутствуют, тогда арбитражный доход можно получить так: продать одну акцию в Нью-Йорке за $3,0, затем купить £2,0 за $2,80, на которые приобрести одну акцию в Лондоне и в результате получить доход в размере $0,20. Если бы такая возможность действительно существовала, было бы осуществлено большое количество арбитражных операций, до тех пор пока цены на акции на двух рынках не уравнялись бы (при поправке на обменный курс).

Рассмотрим ценообразование фьючерсов. Вспомним, что ранее мы рассчитывали будущую стоимость денежной суммы как наращение основной суммы по текущей процентной ставке. Значит, цена фьючерса на какой-либо финансовый инструмент — это текущая цена актива, лежащего в основе контракта, наращенная по ставке чистой доходности этого инструмента. Термин "чистая доходность", также известный как "чистые текущие издержки", — это разница между издержками по займу средств для покупки актива и доходом, полученным от этого актива.

Допустим, что возможен арбитражный процесс между фьючерсом и активом, лежащим в основе этого фьючерса (например, валюта, акции, облигации и т.д.), справедливая цена фьючерсного контракта должна отражать ситуацию, когда отсутствуют как прибыли, так и убытки при проведении арбитража. Если фьючерс переоценен, арбитражеры открывают короткую позицию (т.е., продают то, чего у них нет) и, взяв кредит, приобретают основной актив с целью его поставки в будущем согласно фьючерсному контракту. Если же, наоборот, фьючерс недооценен, арбитражеры продают основной актив, одновременно покупая соответствующий фьючерс.

В качестве примера рассмотрим ситуацию, когда цена “спот” на актив равна 100 единицам. Данный актив не приносит никакой прибыли, за исключением чистой доходности в размере 10% в год. Справедливой ценой поставки через один год будет цена покупки данного актива (текущая цена ”спот”) плюс процентный доход за один год. В нашем примере цена ”спот” равна 100, а процентный доход — 10% от 100, (т.е. 10 единиц). Таким образом, справедливая фьючерсная цена равна 110 единицам. Если арбитражным операциям не будет создаваться помех, цена фьючерса будет соответствовать этой цене. Если, например, фьючерсы котируются по цене выше справедливой, скажем 115 единиц, арбитражеры откроют короткую позицию, согласившись поставить актив и получить за него в момент поставки 115 единиц. Одновременно они приобретут актив по цене 100 единиц, воспользовавшись банковским займом и заплатив за это 10 единиц в качестве процентов. Спустя один год приобретенный актив будет поставлен согласно фьючерсному контракту, гарантируя, таким образом, безрисковую прибыль в размере 5 единиц. Этот тип арбитражной операции известен как **арбитраж между наличным и фьючерсным рынками**.

Если фьючерсная цена ниже справедливой, скажем 108 единиц, тогда арбитражеры предпримут **обратный** арбитраж между наличным и фьючерсным рынками. Они приобретут фьючерс, продадут основной актив за 100 единиц и положат эти средства на депозит, получив за это 10 единиц через один год. В момент поставки арбитражер купит согласованный актив, заплатив при этом 108 единиц. Вычитая из 108 единиц полученные проценты, будем иметь чистую стоимость 98 единиц, в то время как текущая рыночная стоимость составляет 100 единиц. Безрисковый доход в данном случае равен 2 единицам.

Изложенные принципы могут быть обобщены для получения формулы ценообразования форвардов и фьючерсов. В случае простых процентов эта формула выглядит так:

$$F = P[1 + (C \cdot (T - t))]. \qquad (1.26)$$

В случае сложных процентов выражение (1.26) примет вид:

$$F = P(1 + C)^{T-t}, \qquad (1.27)$$

где С — величина дохода, выраженная в процентах в виде десятичной дроби;

$T—t$ — время до истечения форвардного или фьючерсного контракта, выраженное в годах.

Попробуем применить формулу (1.26) к оценке трехмесячного форварда или фьючерса.

Предположим, что цена "спот" актива, лежащего в основе контракта, равна 94,6. Процентная ставка по трехмесячному кредиту составляет 10% годовых. Издержки по хранению выплачиваются ежеквартально в размере 0,02% в год, стабильный доход по данному активу равен 5% в год.

Из приведенных данных находим чистый процентный доход, т.е. (10 + 0,02 — 5 = 5,02%), тогда фьючерсная цена должна быть:

$$F = 94.6[1 + (0,0502 \cdot 0,25)] = 94,6(1.01255) = 95,79.$$

Выражения (1.26) и (1.27) могут применяться для расчетов по фьючерсам на валюту, облигации, а также по отношению к биржевым индексам. В каждом из этих случаев P соответствует цене "спот" актива, лежащего в основе контракта, т.е. это либо текущий обменный курс, либо цена облигации или текущий уровень индекса. C — чистый процентный доход — отличается по способу нахождения для каждого из этих типов активов. В случае валютных фьючерсов издержки по займу средств для покупки иностранной валюты составляют внутреннюю (в данной стране) процентную ставку, но следует учесть, что купленная иностранная валюта может быть инвестирована согласно внешней (зарубежной) процентной ставке до того момента, пока не потребуется ее поставка по фьючерсному контракту. Издержки по хранению валюты отсутствуют. Тогда C — это разница между внутренней и внешней процентными ставками.

В случае биржевых индексов покупка акций по индексу дает право владельцу на все дивиденды, уплачиваемые в течение всего времени существования акций. Тогда чистый процентный доход C — это внутренняя процентная ставка минус причитающиеся дивиденды, выраженные в форме процентов от величины индекса.

Аналогичны рассуждения и для облигаций. Так как они приносят доход, выражающийся в виде процентной ставки, то C равно внутренней процентной ставке минус текущая доходность по облигации.

Справедливые ценовые соотношения между фьючерсами или форвардами и ценными бумагами поддерживаются за счет **возможности осуществления арбитража** между наличным и фьючерсным рынками. Эти отношения являются фундаментальными в ценообразовании фьючерсов и форвардов. На финансовых рынках мира обращается большое количество фьючерсных контрактов других видов. Для более детального знакомства заинтересованные читатели могут обратиться к работам Уотшэма (1992).

Доходность ценных бумаг

Рассмотрим, каким образом инвесторы измеряют свои доходы. Доходы могут быть историческими, т.е. реально заработанными, или ожидаемыми — это те доходы, которые рассчитывают получить в будущем. В инвестиционных процессах важную роль играют ожидаемые доходы. Фактически полученные доходы важны постольку, поскольку они являются ориентиром на будущее.

Доход от инвестиций

Доход от инвестиций (holding period return — HPR) — это общий доход от вложений в течение всего периода инвестирования. Он включает в себя всю прибыль, а также доходы или убытки от изменения курса акций во время осуществления инвестирования. Ожидаемый HPR — это то, что рассчитывает заработать инвестор. Исторический HPR — это то, что уже фактически заработано. Метод для расчета каждого из этих HPR один и тот же.

HPR определяется так:

$$\text{HPR} = \frac{INC + V_T}{V_t} - 1, \qquad (1.28)$$

где INC — доход, получаемый в течение периода инвестирования;
V_T — стоимость инвестиций в конце периода;
V_t — стоимость инвестиций в начале периода.

Например, если первоначальные инвестиции составляют \$10, один доллар был получен в качестве дивидендов, а стоимость инвестиции в конце срока — \$11, то HPR определяется так:

$$\text{HPR} = (1 + r) - 1 = \frac{1+11}{10} - 1 = 1{,}2 - 1 = 0{,}2 = 20\%.$$

Доход в 20% относится ко всему сроку капиталовложений, а не к годовым подпериодам. Другими словами, это не является годовым доходом, за исключением случая, когда общий срок инвестирования составляет один год.

Теперь предположим, что инвестирование осуществлялось в течение нескольких лет, например, £1 000 были вложены в инвестиционный фонд, а все дивиденды были реинвестированы. Если сегодняшний день является окончанием периода инвестирования, а текущая стоимость начальных вложений составляет £1 800, то общий HPR будет:

$$1 + r = \frac{1800}{1000} = 1{,}8\,; \quad 1{,}8 - 1 = 0{,}8 \ (80\%).$$

Пусть доход в 80% был получен за пять лет. Этот общий доход складывается из ежегодных доходов следующим образом:

$$1 + \text{HPR} = (1 + r_1)\,(1 + r_2)\,(1 + r_3)\,(1 + r_4)\,(1 + r_5).$$

HPR может быть представлен в виде годовой доходности r путем следующего преобразования:

$$r = \sqrt[T]{(1 + r_1)(1 + r_2)(1 + r_3)\ldots(1 + r_T)} - 1\,, \qquad (1.29)$$

где T — количество годовых периодов.

Подобная форма представляет доход как наращенный годовой доход. Если бы временной период T был сокращен до, скажем, месяцев, тогда результатом был бы месячный наращенный доход.

Возвращаясь к нашему примеру, рассчитаем годовую сложную ставку (ставку при наращении):

$$r = \sqrt[5]{1{,}8} - 1 = 0{,}1247 = 12{,}47\%,$$

что в действительности является средней геометрической ставкой доходности за пятилетний период.

Доходы по банковским депозитам

Примерами банковских депозитов являются **депозиты в евровалюте**, **внутренние межбанковские депозиты** и **депозитные сертификаты**. Так как основная сумма размещается на депозите в начале срока, а выплачивается в конце вместе с **начисленными процентами**, то доход по такого рода инструментам известен как **начисляемый доход**. Помимо этого в данном случае применяются простые проценты.

Для иллюстрации предположим, что по 90-дневному внутреннему межбанковскому депозиту в фунтах стерлингов процентная ставка равна 10% годовых. Доход от вложений на 90 дней составит:

$$\text{Начисляемый доход} = \left(0{,}1 \cdot \frac{90}{365}\right) = 0{,}02466 = 2{,}466\%. \qquad (1.30)$$

Доходы по дисконтным инструментам

Как уже замечено выше, **казначейские векселя**, **краткосрочные коммерческие векселя** и **банковские акцепты** котируются на рынке согласно ставке доходности, а не цены. Котируемая ставка известна как **дисконтная ставка**, которая является дисконтом, выраженным в процентах от **номинальной стоимости**. Например, допустим, что казначейский вексель со сроком до погашения 90 дней имеет цену 97,5, т.е. дисконт равен 2,5 от номинала. Дисконтная ставка вычисляется так:

$$\text{Дисконтная ставка} = \left(\frac{\text{Дисконт}}{\text{Номинальная стоимость}}\right) \cdot 10{,}1389$$

$$\cdot \frac{365}{n} = \frac{2{,}5}{100} \cdot \frac{365}{90} = 10{,}1389, \qquad (1.31)$$

где n -количество дней до погашения.

С помощью уравнения (1.21) было показано, как определить стоимость векселя исходя из известной дисконтной ставки.

Так как на рынках для таких операций принято использовать простые проценты, а не сложные, 2,5/100 умножается на 365/90 вместо того, чтобы [1 + (2,5/100)] возводить в степень 365/90.

Необходимо заметить, что дисконтная ставка не отражает реального дохода, получаемого инвестором, для этого необходимо дисконт отнести к **цене покупки** векселя и умножить на 365/*n* или 360/*n* в зависимости от принятого рыночного соглашения. Таким образом, для нахождения реальной ставки доходности казначейского векселя с целью сравнения ее со ставками по банковским депозитам необходимо вместо дисконтной ставки рассчитать начисляемую доходность. Начисляемая доходность находится следующим образом:

$$\text{Начисляемая доходность} = \left(\frac{\text{Дисконт}}{\text{Цена покупки}}\right) \cdot \left(\frac{365}{n}\right). \qquad (1.32)$$

Численный пример поможет понять, как используется формула. Если дисконт равен 2,5, тогда начисляемая доходность будет:

$$\frac{2{,}5}{97{,}5} \cdot \frac{365}{90} = 10{,}3989\,.$$

Дивидендный доход по обыкновенным акциям и текущая доходность облигаций

Дивидендный доход по обыкновенным акциям определяется как дивиденды, выраженные в процентах от текущей цены акции. Дивиденды могут быть как историческими, т.е., например, выплаченными за последний год, так и ожидаемыми. Расчет выполняется по формуле:

$$\text{Дивидендный доход} = \frac{D}{P}\,, \qquad (1.33)$$

где через D обозначаются исторические или ожидаемые дивиденды. Обычно при расчете дивидендного дохода D включаются также уплаченные налоги для того, чтобы избежать осложнений, связанных с неодинаковым налогообложением разных категорий инвесторов.

Текущая доходность облигаций измеряется как ожидаемый годовой купонный доход, отнесенный к чистой цене облигации, т.е. не включающей накопленных процентов:

$$Y_c = \left(\frac{\text{Купонный доход}}{\text{Чистая цена}}\right), \tag{1.34}$$

где Y_c — текущая доходность.

Отметим, что ни один из этих подходов к оценке доходности не учитывает ни реинвестирования дохода, ни прибылей или убытков по облигациям в момент погашения (облигации, выкупаемые с дисконтом или премией по отношению к номинальной цене), ни прибылей или убытков по акциям при продаже.

Практика использования чистой цены, особенно в Великобритании, имеет тенденцию переоценивать текущую доходность, так как при покупке инвестор выплачивает "грязную" цену (т.е. чистая, или котировочная, цена на рынке плюс причитающиеся накопленные проценты). Проценты начисляются ежедневно согласно купонной ставке и количеству дней со времени последней выплаты.

Ставка общего дохода или общий доход при погашении

Ставка общего дохода широко используется как мера прибыльности и как мера относительной стоимости. Ожидаемые денежные платежи по облигации дисконтируются к текущей цене облигации согласно **ставке внутренней доходности** (internal rate of return — IRR). **Ставка общего дохода** находится путем решения следующего уравнения для r через итерационную процедуру:

$$P_b = \sum_{i=1}^{n} \frac{CF_i}{(1+r)^i}, \tag{1.35}$$

где CF_i — денежные платежи в конце периода i (купоны и выкупная стоимость);

r — ставка общего дохода.

В качестве примера рассмотрим пятилетнюю облигацию с 10%-ным купоном, выплачиваемым ежегодно, текущая рыночная цена — 87,59, а номинальная стоимость составляет £100:

$$\frac{10}{1{,}1358}+\frac{10}{(1{,}1358)^2}+\frac{10}{(1{,}1358)^3}+\frac{10}{(1{,}1358)^4}+\frac{110}{(1{,}1358)^5}=87{,}59\,.$$

Рассчитанное значение ставки общего дохода равно 13,58%. Это значение было получено в результате итерационного процесса проб и ошибок. Для того чтобы найти IRR, мы должны знать как текущую цену, так и распределение денежных потоков во времени. Более конкретно объяснение итерационных процедур дано в гл. 8.

По большинству облигаций, включая подавляющее количество государственных облигаций, купоны выплачиваются раз в полгода. В этом случае CF соответствует полугодовым купонам, а i будет представлена полугодовыми периодами. Следовательно, ставка общего дохода, рассчитанная исходя из выражения (1.35), будет полугодовой доходностью. Для того чтобы полугодовую ставку общего дохода привести к годовой, принято, особенно в США и Великобритании, просто ее удваивать. Таким образом, мы получаем **эквивалентную доходность по облигации**. Фактически же данный метод недооценивает эффективную доходность. Например, допустим, что облигация оценена по номинальной стоимости, по ней выплачивается купон 10% годовых в виде полугодовых платежей размером в 5 единиц. Полугодовая ставка общего дохода будет 5%. Удвоив ее, получаем годовую ставку 10%. Однако если предположить, что первый купонный доход каждого года может быть реинвестирован во втором полугодии под 10% годовых, то эффективная ставка общей доходности составит 10,25%.

В примере выше подразумевалось, что купон был только что уплачен, а выплата следующего состоится только через один купонный период. На практике же может возникнуть необходимость рассчитать ставку общего дохода в период между выплатами купонов. Это можно сделать, решая уравнение

$$P_b=\sum_{i=1}^{n}\frac{CF_i}{(1+r)^V(1+r)^i}\,, \qquad (1.36)$$

где V — часть периода между выплатами купонов, равная количеству дней до выплаты следующего купона.

Единственный случай, когда можно не применять итерационную процедуру для решения уравнений (1.35) и (1.36) — это когда причитается только один платеж, как, например, для облигации с нулевым купоном. Доходность в этом случае рассчитывается следующим образом:

$$Y = \left(\frac{C}{P}\right)^{1/n} - 1, \qquad (1.37)$$

где C — платеж при погашении;
P — текущая цена облигации.

Для иллюстрации рассмотрим облигацию с нулевым купоном, погашаемую через два года согласно номиналу 100 единиц и имеющую цену в настоящем 80 единиц. Ставка общей доходности:

$$Y = \left(\frac{100}{80}\right)^{1/2} - 1 = 11{,}8\%.$$

Данное уравнение аналогично уравнению (1.15), так как ставка общей доходности для облигации с нулевым купоном фактически равна спот-ставке.

Ранее было замечено, что ставка общей доходности часто применяется для сравнения доходности по разным облигациям. Однако в данном случае возникают некоторые серьезные затруднения. В частности, реальный доход инвестора будет равен ставке общей доходности в том случае, если:

а) инвестор владеет облигацией плоть до ее погашения;
б) все выплачиваемые купоны реинвестируются согласно процентной ставке, равной ставке общей доходности, в течение всего периода до погашения.

Очевидно, не все инвесторы имеют желание владеть облигациями вплоть до их погашения и, конечно же, маловероятно, что все купоны будут реинвестированы по одной и той же процентной ставке в течение всего срока обращения облигации.

Кроме того, ставка общей доходности ничего не говорит об относительной стоимости облигации. Для ее расчета необходимо знать рыночную цену облигации, размер и распределение во времени денежных платежей. Т.е. мы должны знать цену облигации до того, как сможем рассчитать ставку общей доходности. Таким образом, она не поможет нам оценить облигацию.

Также необходимо заметить, что ставка общего дохода зависит от размера купонных платежей. Для иллюстрации этого рассмотрим ставку общего дохода по пятилетней облигации с ежегодным купоном 6% годовых. Дисконтная стоимость денежных потоков, если принимать во внимание те же спот-ставки, что и для оценки облигации с 10%-ным купоном, приведенные на с.35, будет равна:

$$P_B = \frac{6}{1{,}1} + \frac{6}{(1{,}11)^2} + \frac{6}{(1{,}12)^3} + \frac{6}{(1{,}13)^4} + \frac{106}{(1{,}14)^5} = 73{,}33 \text{ ед.}$$

Дисконтирование платежей к цене 73,33 дает ставку общего дохода в 13,72% годовых:

$$\frac{6}{1{,}1372} + \frac{6}{(1{,}1372)^2} + \frac{6}{(1{,}1372)^3} + \frac{6}{(1{,}1372)^4} + \frac{106}{(1{,}1372)^5} = 73{,}33\,.$$

Сравним результат со ставкой общего дохода по пятилетней облигации с годовым купоном в 10%, рассмотренной на стр. 35. Текущая стоимость этой облигации при тех же самых спот-ставках равна 87,59, а ставка общего дохода составляет 13,58%, тогда как по облигации с купоном в 6 единиц она равна 13,72%. Распределение платежей во времени одинаково в обоих случаях.

Следовательно, применение ставки общей доходности означает не только, что равноразмерные платежи в различные моменты времени дисконтируются по одной и той же ставке, но и разноразмерные денежные потоки в один и тот же момент времени дисконтируются по разным ставкам. Оба эти утверждения противоречат финансовой теории и практике. Первое потому что не существует одноуровневой структуры процентных ставок. Второе потому, что если бы два денежных потока одинакового "качества" имели разную цену на одном и том же рынке, трейдеры предприняли бы арбитражные операции, покупая дешевые денежные потоки и продавая дорогие, получая таким образом безрисковый арбитражный доход. Следовательно, ставка общей доходности не может быть однозначно применена для сравнительной оценки облигаций.

Реально накопленная доходность

Если учитывать слабость ставки общей доходности как показателя, характеризующего облигации, возникает необходимость в альтернативном показателе, который бы принимал во внимание желание инвесторов относительно продолжительности владения облигациями и их предположения по поводу ставки реинвестирования купонных платежей.

Когда инвестор желает полностью оценить свой доход по активам, находящимся в его распоряжении, до определенной даты в будущем, **реально накопленная доходность** (realized compound yield — RCY) служит подходящим показателем для сравнения альтернативных активов, приносящих доход.

RCY учитывает влияние "процентов на проценты" на стоимость инвестиций. Однако в связи с этим рассчитанная реальная доходность в сильной мере зависит от предположений инвестора относительно ставки реинвестирования. Несмотря на то, что будущая ставка реинвестирования неопределенна, она должна быть точно указана для выполнения расчетов по этому показателю доходности. Следовательно, при условии, что ставка реинвестирования дана в должном соответствии и согласованности с финансовыми инструментами, появляется возможность осуществления значимых сравнений.

Когда желаемый срок владения облигацией меньше периода ее обращения, необходимо определить предполагаемый уровень доходности по облигации в момент ее продажи. Если срок обращения облигации меньше, чем желаемый период размещения средств, необходимо определиться с реинвестированием полной стоимости при погашении.

Реально накопленная доходность — это сложная ставка роста общей стоимости в течение срока владения, приведенная к годовому базису. Вспомнив расчет **годового дохода от инвестиций**, можно увидеть, что это то же самое, что и реально накопленная доходность. В действительности оба показателя могут быть получены в их прежней форме, в этом случае полученный результат точнее называть ожидаемой реально накопленной доходностью или ожидаемым доходом от инвестиций.

Для расчета **реально накопленной доходности** необходимо предпринять следующее:

1. Рассчитать общие купонные платежи и ставку реинвестирования. Последняя может быть найдена с помощью кривой форвардной доходности, представляющей собой рыночные ожидания будущих процентных ставок.
2. Сделать предположения относительно цены облигации в конце периода владения. Эта цена может быть равна, а может

быть и не равна выкупной цене в зависимости от того, будет ли период владения облигацией короче срока ее обращения.

3. Найти конечную стоимость инвестиций исходя из п.п. 1) и 2) и решить следующее уравнение:

$$\mathrm{RCY} = \left(\frac{TV}{P}\right)^{1/n} - 1, \qquad (1.38)$$

где TV — конечная стоимость инвестиций;
P — первоначальная цена покупки облигации;
n — количество периодов выплаты купонов. Если купоны выплачиваются ежегодно, результатом будет годовая RCY, если же купоны выплачиваются раз в полгода, результатом будет полугодовая RCY. В последнем случае результат должен быть удвоен для получения годового эквивалента RCY.

Для иллюстрации расчетов RCY рассмотрим пятилетнюю облигацию с купоном, равным 8% и выплачиваемым раз в год. Первая выплата должна состояться ровно через год, текущая цена составляет 80,46. Также допустим, что форвардные ставки соответствуют значениям, приведенным в табл. 1.3.

Предположим, что периодические купонные платежи реинвестируются по соответствующим форвардным ставкам следующим образом:

8 · 1,1275 · 1,1504 · 1,1734 · 1,1971	=	14,5758
8 · 1,1504 · 1,1734 · 1,1971	=	12,9275
8 · 1,1734 · 1,1971	=	11,2374
8 · 1,1971	=	9,5768
Последний платеж, включая купон	=	108,0
Всего	=	156,3175

Тогда RCY вычисляется следующим образом:

$$\mathrm{RCY} = \left(\frac{156{,}3175}{80{,}45}\right)^{\frac{1}{5}} - 1 = 0{,}142050 \approx 14{,}21\%.$$

Годовая доходность: полезность непрерывного наращения

Приведение простых процентных ставок доходности к годовому базису с помощью как простого умножения, так и наращения по сложным процентам подразумевает **возможность реинве-**

стирования основной суммы и процентных платежей по текущей процентной ставке периодически в течение следующего года! Например, по шестимесячному депозиту с процентной ставкой 6% годовых начисляется 30 единиц на каждую 1 000 единиц от основной суммы. На ту же самую сумму будет начислено 60 единиц за весь год, если она инвестирована под 6% годовых в течение второго полугодия. Тогда приведение к годовому базису 30 единиц и получение в результате 6% правомерно только в том случае, если:

а) проценты несложные,
б) депозит вторично размещается на шесть месяцев под 6% годовых.

Ясно, что это не всегда так. Например, первоначальная сумма вместе с начисленными процентами реинвестируется на вторые шесть месяцев под 6% годовых, тогда общие процентные платежи за год составят 60,90 единиц или 6,09%. Или же возможно, что основная сумма вместе с процентными платежами или без них реинвестируется на период длиннее или короче, чем шесть месяцев, или под процентную ставку, большую или меньшую по сравнению с первоначальной. В этом случае также начисленные проценты за год будут различными. Следовательно, доходы по инструментам с разными сроками обращения не совсем правомерно сравнивать между собой, так как они отличаются количеством периодов реинвестирования.

Поэтому разные доходности следует привести к общей форме для того, чтобы их можно было сравнить. Наилучшим способом это достигается приведением рыночной доходности к **непрерывно начисляемой доходности** с базовым количеством дней в году 365,25.

Для проведения сравнения, во-первых, доходность инструментов, начисляемая исходя из 360 дней в году, должна быть приведена к базе 365,25 дней в году. Тогда реальные проценты, начисляемые за период владения инструментом, рассчитывают следующим образом:

$$\text{Проценты} = R \cdot (n/360) \cdot \text{Основная сумма},$$

где R — объявленная процентная ставка по инструменту, а n — фактическое число дней с момента начала финансовой операции до ее окончания.

В качестве примера предположим, что депозит в \$1 млн. размещен на срок 90 дней под 10% годовых. Ожидаемые процентные платежи будут:

$$(0{,}1)(90/360)\cdot \$1000\,000 = \$25\,000{,}00.$$

Затем вычисляется непрерывно наращенная начисляемая доходность следующим образом:

$$R=\frac{365{,}25}{n}\cdot\ln\left[\frac{\text{Основная сумма + Проценты}}{\text{Основная сумма}}\right]= \tag{1.39}$$

$$=\frac{365{,}25}{90}\cdot\ln\left[\frac{1\,025\,000}{1\,000\,000}\right]=10{,}0211\%.$$

Заметим, что полученный результат больше простой процентной ставки 10% несмотря на то, что это непрерывная процентная ставка. Это объясняется тем, что простые проценты были рассчитаны исходя из 360 дней в году.

Преобразование доходности к этой форме позволяет провести полноценное сравнение между инструментами с различными сроками и характеризующимися различными базами при начислении процентов. Однако в этом случае подразумевается, что текущая процентная ставка неизменна в течение периода наращения.

Годовая процентная ставка (annual percentage rate,* APR*). Альтернативным методом стандартизации эффективной процентной ставки является приведение к годовой процентной ставке. APR определяется как эквивалентная годовая процентная ставка при периодическом наращении только один раз в году. Например, компания кредитных карточек, начисляющая 2% в месяц по непогашенным остаткам, использует номинальную ставку 24%, но производит начисления через месячные интервалы. Эквивалентная APR будет рассчитана как $(1{,}02)^{12} - 1 = 0{,}268$, поэтому объявленная APR компании будет 26,8%.

Непрерывно наращенная доходность активов. Метод непрерывно наращенной доходности активов широко применяется во многих финансовых моделях, включая модель ценообразования опционов. Мы уже рассмотрели, как привести простые или сложные процентные ставки к их эквивалентным непрерывно наращенным процентным ставкам. Но часто бывает необходимо вычислить непрерывно наращенную доходность исходя из цен.

Для этого нужно просто найти натуральный логарифм отношения цен. Например, если цена акции сегодня равна 10, а одну неделю назад она составляла 8, непрерывно наращенная недельная доходность будет:

$$R_{cc} = \ln\frac{P_1}{P_0} = \ln\frac{10}{8} = \ln 1{,}25 = 0{,}22314 \approx 22{,}3\%,$$

где P_1 — текущая цена;

P_0 — цена в конце предыдущего периода.

Полученная доходность соответствует данному периоду, в этом случае одной неделе. Если бы были использованы данные за день или за месяц, доходность была бы непрерывно наращенной дневной или месячной доходностью.

При использовании цен для нахождения доходности актива предполагалось, что никакие процентные платежи, выплаты дивидендов или другие выплаты не производились. Если бы такие выплаты были сделаны в течение периода, ограниченного данными ценами, они должны были быть добавлены к цене P_1.

Природа временной структуры процентных ставок. Многие читатели знакомы с понятием **кривой доходности**, которая представляет собой зависимость ставок полной доходности от срока до погашения. Так как наиболее многочисленными долгосрочными долговыми обязательствами являются купонные облигации, кривые доходности обычно отображают срок до погашения таких облигаций. Иногда считают, что кривые доходности описывают временную структуру процентных ставок, однако **временная структура процентных ставок** в действительности относится к **спот-ставкам** и соответствующим им периодам.

Как мы заметили ранее, спот-ставки на рынке могут быть представлены в виде общей доходности облигаций с нулевым купоном. К сожалению, полная временная структура недоступна в данном случае из-за отсутствия необходимого объема эмиссии таких облигаций. Таким образом, возникает необходимость определения спот-ставок исходя из доходности по другим инструментам, например, начисляемой доходности краткосрочных банковских депозитов, доходности казначейских векселей и цен ожидаемых платежей по безрисковым облигациям. Продемонстрируем это.

Рассмотрим табл. 1.4, содержащую информацию о месячном межбанковском депозите, трехмесячном казначейском векселе и ряде облигаций со сроком до погашения от 0,5 до 3 лет. Шестимесячные и одногодичные облигации — облигации с нулевым купоном, облигации со сроком до погашения 1,5—3 года — купонные облигации.

Используя информацию первых трех колонок табл. 1.4, можно построить кривую доходности.

Первое, что нужно сделать, — это привести начисляемую доходность по банковским депозитам к соответствующей спот-ставке. Необходимо также учесть различие в соглашениях по подсчету дней при применении уравнения (1.30). Затем мы должны скорректировать месячную начисляемую доходность и привести ее к годовой спот-ставке за соответствующий период.

Для того чтобы понять это, рассмотрим депозит с процентной ставкой 7,25% годовых на один месяц. В случае 31 дня в месяце и при методе ACT/365 начисляемая доходность будет:

$$7{,}25 \cdot \frac{31}{365} = 0{,}6157\,.$$

Таблица 1.4

Срок (в годах)	Купон	Цена	Полная доходность	Спот-ставка	Форвардная ставка 0,5 лет
1 месяц, 0,0833	7,25	100	—	7,644	
0,25		98,10	7,98	7,975	
0,5	0	96,18	8,10	8,101	
1,0	0	92,21	8,45	8,448	8,796
1,5	8,5	99,18	9,30	9,347	11,166
2,0	9,0	99,37	9,57	9,624	10,461
2,5	11,0	103,16	9,78	9,853	10,771
3,0	9,5	99,10	10,09	10,187	11,878

Тогда конечная стоимость депозита будет 100,6157 единиц. Для того чтобы получить такую конечную стоимость через один месяц, мы должны вложить 100 единиц сегодня. Эквивалентная годовая спот-ставка может быть рассчитана из уравнения (1.15), которое обозначим (1.40):

$$r_n = \left(\frac{CF_n}{PV}\right)^{\frac{1}{n}} - 1. \qquad (1.40)$$

Таким образом, одномесячная спот-ставка равна:

$$r_{0,0833} = \left(\frac{100,6157}{100}\right)^{\frac{1}{0,08333}} - 1 = 0,07644 = 7,644\%.$$

Заметим, что годовая эквивалентная спот-ставка больше, чем данная одномесячная ставка, вследствие того, что спот-ставка не предполагает ежемесячного наращения.

Так как по следующим двум приведенным в табл. 1.4 ценным бумагам не выплачиваются купонные платежи, их полная доходность является соответствующими спот-ставками для данных периодов. Никаких дополнительных расчетов в этом случае не требуется.

Однако инструменты со сроком до погашения 1,5 — 3 года являются купонными облигациями, и для нахождения соответствующих спот-ставок следует выполнить некоторые расчеты.

Спот-ставка для платежа через 1,5 года может быть рассчитана таким образом. Вспомним, что цена облигации — это текущая стоимость всех ожидаемых будущих денежных потоков, каждый из которых дисконтирован по соответствующей спот-ставке. Другими словами, цена полуторагодовой облигации находится дисконтированием первого полугодового купона по шестимесячной спот-ставке, годового купона — по одногодичной спот-ставке, а выкупной стоимости и последнего купона — по 1,5-годовой спот-ставке:

$$\frac{4,25}{(1,081)^{0,5}} + \frac{4,25}{(1,0845)} + \frac{104,25}{(1+R)^{1,5}} = 99,18.$$

Нашей проблемой является то, что мы не знаем 1,5-годовую спот-ставку. Однако ее можно получить вычитанием текущих стоимостей всех купонов, кроме последнего, из цены облигации, т.е.:

$$99,18 - \left(\frac{4,25}{(1,081)^{0,5}} + \frac{4,25}{(1,0845)}\right) = 91,17.$$

Полученный результат 91,17 — это текущая стоимость купона и выкупной цены 104,25, дисконтированная по 1,5-годовой спот-ставке. Следовательно, так как $91{,}17 = 104{,}25/(1 + R)^{1,5}$, то $(1 + R)^{1,5} = 104{,}25/91{,}17 = 1{,}1434$, откуда $(1{,}1434)^{1/1,5} - 1 = 0{,}0934$. Это и есть 1,5-годовая спот-ставка. Мы можем проверить полученный результат, разделив 104,25 на $(1{,}0934)^{1,5}$, и получим 91,17.

Используя полученную информацию, мы можем продолжить процесс, вычисляя двухгодичную спот-ставку. Зная три спот-ставки, можно дисконтировать три первых из четырех купонов по двухгодичной облигации. Затем вычесть текущие стоимости этих купонов из текущей цены облигации. В результате получим текущую стоимость последнего купона плюс выкупная стоимость, дисконтированная по двухгодичной спот-ставке. Теперь двухгодичная спот-ставка может быть найдена точно так же, как было показано выше.

Эта процедура иногда называется "поиск по линии связи" (bootstrap method). Она применяется до тех пор, пока не будут найдены все необходимые спот-ставки. Однако следует сделать некоторые предостережения. Во-первых, могут быть использованы только купонные облигации без каких-либо дополнительных характеристик, таких, как, например, опционные, в противном случае рыночные цены будут отражать что-то отличное от прямых денежных потоков по облигации. Во-вторых, если доходы от прироста капитала или купонные платежи облагаются налогом по различным ставкам, необходимо принять во внимание преимущества низкодоходных облигаций при налогообложении. Данная проблема может быть решена, если принимать во внимание только недавно выпущенные облигации, но в этом случае может быть учтен не весь спектр сроков до погашения.

Метод поиска по линии связи может оказаться громоздким, когда на рынке обращается большое количество купонных облигаций с одинаковыми сроками до погашения, но с разными другими характеристиками. Поэтому для моделирования временной структуры процентных ставок с разной степенью успеха применяется ряд сложных эконометрических методов. Когда временная структура уже определена, ее элементы являются всего лишь приближенной оценкой реальной временной структуры. Следовательно, получаемые спот-ставки — это всего лишь оценочные значения реальных ставок. Как следствие, цены облигаций, рассчитанные исходя из этих оценочных значений, в свою очередь также только приближенно оценивают реальные цены облигаций.

Ипотечные ссуды и аннуитеты

Ипотечные ссуды. Одна из черт ипотечных ссуд заключается в том, что при определенной процентной ставке периодические выплаты являются фиксированными равноразмерными за весь срок ссуды. Если процентные ставки периодически пересматриваются (плавающие процентные ставки), меняются и периодические выплаты, но по-прежнему они будут равными в каждом из периодов вплоть до следующего изменения процентных ставок. Периодические платежи (обычно ежемесячные) состоят из основной суммы долга и процентов. Сначала, когда основная сумма долга принимает свое максимальное значение, ежемесячные платежи состоят в основном из начисляемых процентов, доля выплачиваемой основной суммы составляет лишь незначительный процент от платежей. Со временем, при уменьшении основной суммы, доля ежемесячных платежей, необходимая для покрытия начисленных процентов, снизится, а доля периодических выплат, отведенная под погашение основной суммы, возрастет. Каждый раз, когда процентная ставка меняется, необходим пересчет периодических платежей, которые покрывают проценты и направляются на выплату основной суммы долга, на оставшийся срок ипотеки.

Для иллюстрации того, как это происходит, рассмотрим прежде всего расчет равных годовых платежей, необходимых для выплаты ипотечной ссуды в £ 100 000 с процентной ставкой 10% годовых (ежегодное наращение), взятой на 20 лет.

Начнем с первого года и первоначального долга £ 100 000. Проценты за первый год могут быть рассчитаны как произведение суммы долга и множителя 1,10, после чего из результата вычитается ежегодный платеж. Если обозначить этот платеж через £ X, тогда долг на начало второго года составит $100\,000 \cdot 1{,}10 - X$.

Применяя этот же способ к полученному результату, найдем значение долга на начало третьего года:

$$((100\,000 \cdot 1{,}10) - X) \cdot 1{,}10 - X = 100\,000 \cdot 1{,}10^2 - 1{,}10X - X.$$

Заметьте, что аргументы в скобках умножены на 1,10.

Повторив эту процедуру 20 раз, мы получим выражение для долга на начало 21-го года, который должен быть равен нулю:

$$100\,000 \cdot 1{,}10^{20} - 1{,}10^{19}X - 1{,}10^{18}X - 1{,}10^{17}X - \ldots$$

$$\ldots - 1{,}10^{2}X - 1{,}10X - X = 0;$$

$$100\,000 \cdot 1{,}10^{20} - X(1{,}10^{19} + 1{,}10^{18} + \ldots + 1{,}10^{2} + 1{,}10 + 1) = 0.$$

Переставив слагаемые в скобках в порядке возрастания, получим $1 + 1{,}10 + 1{,}10^2 + 1{,}10^3 + \ldots + 1{,}10^{18} + 1{,}10^{19}$ — геометрическую прогрессию с первым членом 1 и знаменателем 1,10. Формула для суммы *n* первых членов геометрической прогрессии со знаменателем, большим единицы, выглядит так:

$$\frac{1 \cdot \left(1{,}10^{n} - 1\right)}{1{,}10 - 1}. \qquad (1.41)$$

В общем виде:

$$a \cdot \frac{\left[(1 + r)^{n} - 1\right]}{(1 + r) - 1}, \qquad (1.42)$$

где a — это первый член, а $1 + r$ — знаменатель прогрессии, который больше единицы, так как r всегда положительно.

Расчеты для нашего примера дают такой результат:

$$1 \cdot \frac{\left[1{,}10^{20} - 1\right]}{1{,}10 - 1} = 57{,}275.$$

Таким образом, уравнение для нахождения регулярного платежа по ипотеке X принимает вид:

$$100\,000 \cdot 1{,}10^{20} - 57{,}275X = 0,$$

что в результате дает $X = 11745{,}96$.

Формула для общего случая:

$$X = \frac{P(1 + r)^{n} r}{((1 + r)^{n} - 1)}. \qquad (1.43)$$

Конечно, большинство из нас выплачивает ипотеку ежемесячными погашениями, а не ежегодными. Для нахождения ежемесячных выплат по нашей ипотеке мы можем использовать тот же самый подход (только с 240 шагами) при условии, что годовую процентную ставку преобразуем в эквивалентную месячную ставку. Если обозначить эквивалентную месячную ставку через C, тогда C находится из $C^{12} = 1{,}10$, т.е. $C = \sqrt[12]{1{,}10} \approx 1{,}007974$.

Наше уравнение для X примет вид:

$$100\,000 \cdot C^{240} - \frac{C^{240} - 1}{C - 1} X = 0,$$

или

$$X = \frac{100\,000 \cdot 1{,}007974^{240} \cdot (1{,}007974 - 1)}{1{,}007974^{240} - 1} = 936{,}64$$

(заметим, что этот результат в годовом исчислении меньше примерно на £500 по сравнению с ежегодной выплатой из-за сэкономленных процентных платежей).

Обобщим формулу для ежемесячных выплат по ипотеке:

$$X = \frac{P\left(\sqrt[12]{1+r}\right)^{12n}\left(\sqrt[12]{1+r} - 1\right)}{\left(\sqrt[12]{1+r}\right)^{12n} - 1}, \tag{1.44}$$

где P — размер суммы, взятой в долг, r — годовая процентная ставка, а n — количество лет.

Аннуитеты. В случае с ипотекой клиент получает денежные средства в начале ипотечного периода, а затем выплачивает свой долг посредством серии небольших платежей в течение всего этого периода. В случае с аннуитетом денежные потоки имеют обратное направление. Клиент вносит первоначальную сумму, а в обмен получает серию небольших платежей в течение срока аннуитета, т.е. аннуитет — это ипотека наоборот (или ипотека — это аннуитет наоборот).

Неудивительно, что расчет по аннуитетам очень похож на расчет по ипотекам. Рассмотрим, например, аннуитет, по которому ежегодно выплачивается £1200 в течение 20 лет, а доходность установлена в размере 10% годовых. Какую стоимость имеет данный аннуитет?

Текущая стоимость первого платежа, который состоится через год, равна $1200/1{,}10$, т.е. эта сумма возрастет за один год до величины £1200 по ставке 10% годовых. Аналогично текущая стоимость второго платежа — $1200/1{,}10^2$. Тогда сумма текущих стоимостей всех двадцати платежей равна:

$$P = \frac{1200}{1{,}10} + \frac{1200}{1{,}10^2} + \frac{1200}{1{,}10^3} + \ldots + \frac{1200}{1{,}10^{20}}.$$

Это тоже геометрическая прогрессия, но с первым членом (обозначаемым через a) 1200/1,10 и знаменателем 1/1,10. Знаменатель геометрической прогрессии — это множитель, на который предыдущее значение должно быть умножено для того, чтобы получить текущее значение, например:

$$\frac{1200}{1{,}10^2} \cdot \frac{1}{1{,}10} = \frac{1200}{1{,}10^3}.$$

Формула геометрической прогрессии со знаменателем, меньшим единицы, например, $1/(1 + r)$ или 1/1,10 в нашем примере выглядит так:

$$a\frac{\left(1 - \dfrac{1}{(1+r)^n}\right)}{1 - \dfrac{1}{1+r}}, \qquad (1.45)$$

где $a = 1200/(1 + r)$. Умножим числитель и знаменатель на $(1 + r)$ и получим:

$$1200\left(\frac{1 - \dfrac{1}{(1+r)^n}}{r}\right). \qquad (1.46)$$

Тогда в настоящий момент стоимость всех 20 годовых платежей размером в 1200 с дисконтной ставкой 10% годовых будет равна 10216,28:

$$1200\frac{1 - \dfrac{1}{1{,}10^{20}}}{0{,}10} = 1200\frac{1 - 0{,}1448644}{0{,}10} = 10216{,}28.$$

Так как знаменатель прогрессии меньше единицы, появляется интригующая возможность позволить n — количеству платежей — стремиться к бесконечности ($\to\infty$) или, по крайней мере, принять очень большое значение вследствие того, что при $n\to\infty$ $1/(1 + r)^n\to 0$.

Таким образом, мы подошли к понятию вечного аннуитета, стоимость которого определяется так:

$$\frac{a}{1-\left(\frac{1}{1+r}\right)}.$$

В случае, когда a = 1200/1,10, а знаменатель = 1/1,10, получаем значение 12 000. На практике стоимость аннуитета со сроком, равным жизни клиента, рассчитывается довольно часто, и эта стоимость (в зависимости от возраста клиента) не сильно отличается от значения стоимости вечного аннуитета.

Ежемесячные выплаты. Как и в случае с ипотекой, мы можем рассчитать стоимость аннуитета с 240 ежемесячными выплатами по £100. Для этого снова необходимо рассчитать множитель наращения:

$$C = \sqrt[12]{1{,}10} = 1{,}007974,$$

тогда

$$\frac{100}{1{,}007974}\frac{\left(1-\frac{1}{1{,}007974^{240}}\right)}{\left(1-\frac{1}{1{,}007974}\right)} = 99{,}2063\frac{1-0{,}148644}{1-0{,}99209} = 10676{,}47$$

и получаем значение 10676,47. Это более дорого по сравнению с годовыми выплатами, так как здесь отражен тот факт, что платежи выплачиваются раньше.

УПРАЖНЕНИЯ

1. Какие факторы влияют на:

   а) общий уровень процентных ставок,
   б) временную структуру процентных ставок,
   в) качественную структуру процентных ставок?

2. Рассчитайте процентные платежи, выплачиваемые по одногодичному кредиту размером в 10 000, если процентная ставка составляет 6% годовых и:

   а) проценты простые,
   б) проценты сложные и начисляются раз в полгода,
   в) проценты сложные и начисляются ежеквартально,
   г) проценты сложные и начисляются ежемесячно,

д) проценты сложные и начисляются ежедневно,
е) происходит непрерывное наращение.

3. Какова будущая стоимость инструмента, по которому не выплачиваются периодические доходы, его текущая цена составляет 25 000, а погашение через 210 дней? Текущая 210-дневная процентная ставка равна 7,5% годовых. На рынке для расчетов принято использовать 365 дней в году.

4. Найдите для непрерывной ставки 10% эквивалентную процентную ставку с наращением а) раз в полгода и б) раз в квартал.

5. Исходя из следующих цен облигаций с нулевым купоном определите непрерывную и сложную дискретную спот-ставку, а также трехмесячные форвардные ставки.

Данные по облигациям с нулевым купоном:

Срок до погашения, лет	Цена
0,25	98,4
0,5	96,5
0,75	94,5
1,0	92,4
1,25	90,5
1,5	88,25

6. Используя цены по трехмесячным и шестимесячным облигациям из вопроса 5, рассчитайте простые процентные ставки, которые дают эквивалентные доходности, и исходя из полученных результатов найдите трехмесячную форвардную ставку.
Используйте метод АСТ/365. Каждый квартал равен 91 дню.

7. Найдите эффективную процентную ставку по депозитному сертификату с 5% годовых, который обращается на финансовом рынке, где применяется метод АСТ/360.

8. Найдите текущую стоимость депозитного сертификата размером в DM 5 млн. и с процентной ставкой 4,5%. Сертификат имеет период обращения 270 дней, а его текущий срок до погашения равен 160 дням. 160-дневная спот-ставка составляет 4,3%.

9. Текущая стоимость ценной бумаги равна 102,5, годовая процентная ставка — 6%, стабильная доходность данного актива составляет 4% годовых, издержки по хранению равны трем базисным пунктам (1 базисный пункт равен одной сотой 1%). Рассчитайте цену шестимесячного форвардного контракта, если:

а) проценты простые,
б) проценты сложные.

10. Критически оцените полезность общего дохода при погашении как показателя относительной стоимости облигаций. Проиллюстрируйте ответ с помощью численных примеров.

11. Приведите численный пример для иллюстрации применения реальной наращенной доходности.

12. Используя данные из вопроса 5, примените метод поиска по линии связи для расчета двухгодичной спот-ставки исходя из облигации с купоном 10%, выплачиваемым раз в полгода (т.е. пяти- шестимесячных периодов), текущая стоимость облигации 103,0, а срок до погашения — два года.

13. Рассчитайте ежемесячные платежи по ипотечной ссуде в $150 000, выданной на 20 лет. Текущая процентная ставка равна 4,5%.

14. По двадцатилетнему аннуитету процентная ставка определена в 6% годовых. Какова стоимость аннуитета при ежемесячной выплате 1 500?

15. Какова бы была стоимость аннуитета в вопросе 14, если бы он был вечным?

ОТВЕТЫ К ИЗБРАННЫМ ВОПРОСАМ

2.

Простые проценты		600
Сложные проценты:	2	609
	4	613,6355
	12	616,7781
	365	618,3131
	Непрерывное наращение	618,3655.

3. 26078,77.

4. а) 10,2542.
   б) 10,1260.

5.

Спот-ставки		Форвардные ставки	
CD	CC	CD	CC
6,6644	6,4518		
7,3854	7,1254	8,1113	7,7991
7,8345	7,5427	8,7382	8,3773
8,2251	7,9043	9,4055	8,9891
8,3131	7,9856	8,6660	8,3109
8,4854	8,1545	9,3506	8,9389.

6. 6,4; 7,0; 7,4806.

7. 5,06944.

8. 5071821,8.

9. 103,5404; 103,5351.

12. 8,5382.

13. 941,6839.

14. 212076,5.

15. 308163,2086.

СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ

Anderson, N., Breedon, F., Deacon, M., Derry, A. and Murphy, G. (1996) *Estimating and Interpreting the Yield Curve*. John Wiley, Chichester (особый интерес представляют гл. 1 и 7—9).

Fabbozzi F. *Fixed Income Mathematics*. Probus Publishing.

Fuller, R.J, and Hsia, C.C. (1984) A simplified common stock valuation model. *Financial Analysts Journal*, September/October, pp. 49—55.

Gordon, M.J. (1962) *The Investment Financing and Valuation of the Corporation*. Richard D. Irwin, Homewood, Ill.

Watsham, T. J. (1992) *Options and Futures in International Portfolio Management*. Chapman & Hall, London.

Watsham, T. J. (1993) *International Portfolio Management: A Modern Approach*. Longman, London.

Williams, J. B. (1938) *The Theory of Investment Value*. Harvard University Press, Cambridge, MA.

2

ПРЕДСТАВЛЕНИЕ ДАННЫХ И ОПИСАТЕЛЬНЫЕ СТАТИСТИЧЕСКИЕ ПОКАЗАТЕЛИ

ВВЕДЕНИЕ

Понятие "описательные статистические показатели" — это общий термин, относящийся к группе сводных показателей, каждый из которых одним числом определяет какое-либо качество совокупности данных. Например, при анализе финансовых рынков часто необходимо подсчитать некоторую форму средней из индивидуальных доходностей группы активов или среднюю цену активов в данном периоде. Также нам часто нужно измерить, как рассеяны эти доходности и цены. Далее, нам может понадобиться узнать, распределены ли доходности и цены симметрично вокруг средней величины, имеется ли скошенность в распределении доходов и имеет ли их распределение пик, выраженный больше или меньше, чем предполагалось.

Статистики определяют средние значения как **показатели центра распределения**. Статистические показатели, которые описывают или измеряют, как данные распределены вокруг среднего значения, известны как **показатели вариации**. Величины, определяющие симметричность данных, называются **показателями асимметрии или мерами скошенности**, а выражающие пиковость данных — **показателями эксцесса**.

ТИПЫ ДАННЫХ

Данные могут иметь различный вид, который иногда диктует выбор методов статистического анализа или по меньшей мере влияет на него. Таким образом, до начала изучения различных методов статистического анализа следует рассмотреть различные виды данных.

Непрерывные и дискретные данные

Данные могут быть классифицированы как **непрерывные** и **дискретные**. Непрерывные данные могут принимать любые значения на множестве действительных чисел, т.е. они **измеряются** на непрерывной шкале, а их значения ограничены только степенью точности. Типичным примером является процентная ставка дохода по инвестициям. Она может составлять 10%, 10,1% или вообще 10,0975%. Данные, относящиеся ко времени, расстоянию и скорости, также попадают в эту категорию.

Дискретные данные связаны с процессом счета. Например, число финансовых сделок дискретно, потому что их половины или четверти не имеют рационального смысла. Данные, связанные с ценами активов, могут быть дискретными из-за установленных на индивидуальных рынках правил котировки и минимальных ценовых изменений. Например, в Великобритании государственные облигации котируются в тридцать вторых (1/32) долях фунта. Следовательно, ценовые данные здесь дискретны и изменяются только в тридцать вторых долях или кратных им значениях. Рынки корпоративных облигаций, акций, фьючерсов и опционов служат другими примерами, где правила минимальных ценовых изменений определяют дискретность данных.

Мы увидим далее, что различия между непрерывными и дискретными данными иногда оказывают влияние на методы расчета показателей описательной статистики.

Данные кросс-секций и временных рядов

Кросс-секционные (cross-sectional)[1] данные представляют ситуацию в группе переменных в каждый отдельный момент времени. Например, данные, связанные с ценой каждой из составляющих индекса акций FTSE 100 в конкретный момент времени, являются кросс-секционными. Списки цен акций, процентных ставок или обменных курсов, публикуемые в деловых разделах газет, также представляют собой кросс-секционные данные, потому что относятся к ценам или ставкам нескольких переменных (акций, валют и т. п.) в данный момент времени.

Временные ряды отражают колебания какой-либо переменной на промежутке времени. Например, данные о цене акции, обменном курсе валюты или уровне индекса за каждый день (неделю или месяц) в течение двух лет будут ежедневным (еженедельным или ежемесячным) временным рядом.

Сгруппированные и несгруппированные данные

Только когда обрабатывается небольшое количество исходных данных, они могут быть оставлены в необработанном или **несгруппированном** виде, при этом знакомящиеся с ними будут в состоянии их усвоить. Например, следующие данные, представляющие ежемесячные наблюдения за индексом FTSE 100 на протяжении 12 месяцев еще приемлемы для рассмотрения:

2407,5,	2289,2,	2160,1,	2311,1,	2422,7,	2345,8,
2238,4,	2221,6,	2117,9,	2391,4,	2372,0,	2339,0

---

[1] В переводах на русский язык понятия "cross-sectional data" встречается в качестве одного из вариантов "перекрестные данные". Мы решили оставить в переводе "кросс-секционные данные", поскольку здесь сложно дать точный эквивалент на русском языке (cross можно перевести также как "срез" данных). В действительности различие связано с тем, что в случае так называемых cross-sectional data приводятся сведения о размере признака по совокупности единиц по состоянию на определенный момент. — (*Прим. научн. ред*).

А далее обсудим данные из табл. 2.1, в которой приводятся значения индекса FTSE 100 на конец месяца и месячные доходы с сентября 1989 г. по декабрь 1993 г.

Таблица 2.1

	FTSE 100	Доходы
Сентябрь 1989	2407,5	
	2289,2	—5,0386
	2160,1	—5,8048
	2311,1	6,7569
Январь 1990	2422,7	4,7159
	2345,8	—3,2256
	2238,4	—4,6865
	2221,6	—0,7534
	2117,9	—4,7803
	2371,4	11,3055
	2372	0,0253
	2339	—1,4010
	2166,6	—7,6564
	2030,8	—6,4729
	2028	—0,1380
	2162,7	6,4307
Январь 1991	2143,5	—0,8917
	2165,7	1,0304
	2386,9	9,7252
	2456,5	2,8742
	2508,4	2,0908
	2515,8	0,2946
	2443,6	—2,9118
	2591,7	5,8842
	2679	3,3353
	2645,6	—1,2770
	2549,5	—3,7001
	2414,9	—5,4239
Январь 1992	2493,1	3,1869
	2560,2	2,6558
	2554,3	—0,2307
	2408,6	—5,8733
	2659,8	9,9205
	2697,6	1,4112
	2493,9	—7,8515
	2420,2	—2,9998
	2298,4	—5,1637
	2572,3	11,2587
	2687,8	4,3923

Продолжение таблицы

	2792	3,8035
Январь 1993	2846,5	1,9332
	2851,6	0,1790
	2882,6	1,0812
	2878,4	−1,1458
	2813,1	−2,2948
	2849,2	1,2751
	2888,8	1,3803
	2941,7	1,8146
	3085	4,7564
	3039,3	−1,4924
	3164,4	4,0336
Декабрь 1993	3233,2	2,1509

Так как на бирже индекс рассчитывается с точностью до одного знака после запятой, данные дискретны. Было проведено 52 наблюдения уровней индекса и 51 наблюдение доходов.

Понятно, что когда набор данных велик, как в этом случае, необходимо изложить его в виде **таблицы распределения численностей** для того, чтобы аналитик смог осмыслить его характеристики. Построенная таблица представляет данные в разбивке на группы или интервалы. Пример **сгруппированных дискретных данных**, приведенный в табл. 2.2, описывает уровни индекса FTSE 100.

Таблица 2.2

Уровень FTSE 100	Число наблюдений	Кумулятивная частота
До 2000	0	0
2000,1—2100	2	2
2100,1—2200	6	8
2200,1—2300	4	12
2300,1—2400	6	18
2400,1—2500	9	27
2500,1—2600	7	34
2600,1—2700	5	39
2700,1—2800	1	40
2800,1—2900	7	47
2900,1—3000	1	48
3 000,1—3100	2	50
3100,1—3200	1	51
3200,1—3300	1	52
	52	

Пример **сгруппированных непрерывных данных** показан в табл. 2.3 и связан с уровнем доходов из табл. 2.1.

Таблица 2.3

Месячный доход	Число наблюдений	Кумулятивная частота
До –8%	0	0
Более –8% и до –7%	2	2
Более –7% и до –6%	1	3
Более –6% и до –5%	5	8
Более –5% и до –4%	2	10
Более –4% и до –3%	2	12
Более –3% и до –2%	3	15
Более –2% и до –1%	3	18
Более –1% и до 0%	5	23
Более 0% и до + 1%	3	26
Более + 1% и до + 2%	7	33
Более + 2% и до + 3%	4	37
Более + 3% и до + 4%	3	40
Более + 4% и до + 5%	4	44
Более + 5% и до + 6%	1	45
Более + 6% и до + 7%	2	47
Более + 7% и до + 8%	0	47
Более + 8% и до + 9%	0	47
Более + 9% и до + 10	2	49
Более + 10% и до + 11%	0	49
Более + 11% и до + 12%	2	51
	51	

При группировке данных необходимо уделить особое внимание интервалам. Во-первых, они не должны перекрываться. Во-вторых, они должны быть одинаковых размеров, если не существует специфической потребности в выделении данных в рамках какой-либо "подгруппы", либо если данные в группе настолько малочисленны, что могут быть безопасно слиты с предыдущей или последующей группой без потери информации. В-третьих, интервалы не должны быть столь большими, чтобы скрывать характерные изменения в рамках группы. И наконец, число интервалов должно быть компромиссом между необходимостью передачи деталей и возможностью для аналитика охватить эти детали.

ПРЕДСТАВЛЕНИЕ ДАННЫХ

В этом разделе рассмотрим следующие методы представления данных:

- графики распределения частот;
- графики распределения относительных частот;
- графики распределения накопленных частот или огивы;
- гистограммы.

Распределения частот

В предыдущем тексте необработанные данные преобразовывались в таблицы распределения численности (ряды распределения). Здесь мы будем строить графики распределения частот для визуального отображения "частоты" различных наблюдений.

Первичные данные о доходах были трансформированы в ряд распределения, который имеет интервалы шириной 1%, начиная с —8% и заканчивая + 12%. Эти данные будут использованы для составления графиков **распределения частот**, **распределения относительных частот**, **распределения кумулятивных частот** и **гистограмм**. Данные для построения рядов распределения частот приведены в табл. 2.4.

График распределения частот

Для построения графика распределения частот частота наблюдений откладывается по вертикальной оси, а интервалы (доходы) — по горизонтальной. Затем отмечается каждая из частот и составляется диаграмма, как на рис. 2.1.

Высота каждого столбика представляет частоту наблюдений в рамках каждого интервала.

График распределения относительных частот

Чтобы получить относительные частоты для группы данных, необходимо число наблюдений в этой группе разделить на общее число единиц совокупности. После этого распределение относительных частот можно изобразить графически способом, аналогичным для распределения частот. Получим диаграмму, показанную на рис. 2.2.

Высота каждого столбика представляет относительное распределение отдельных изменений доходов.

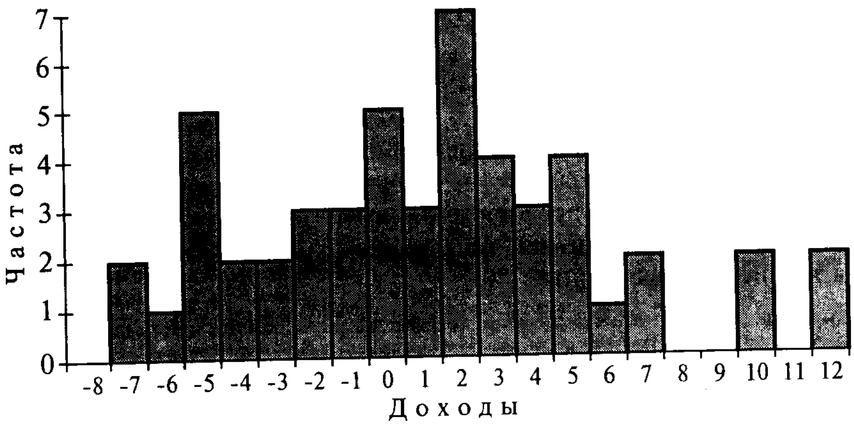

Рис. 2.1. Распределение частот месячных доходов по индексу FTSE 100 с сентября 1989 г. по декабрь 1993 г.

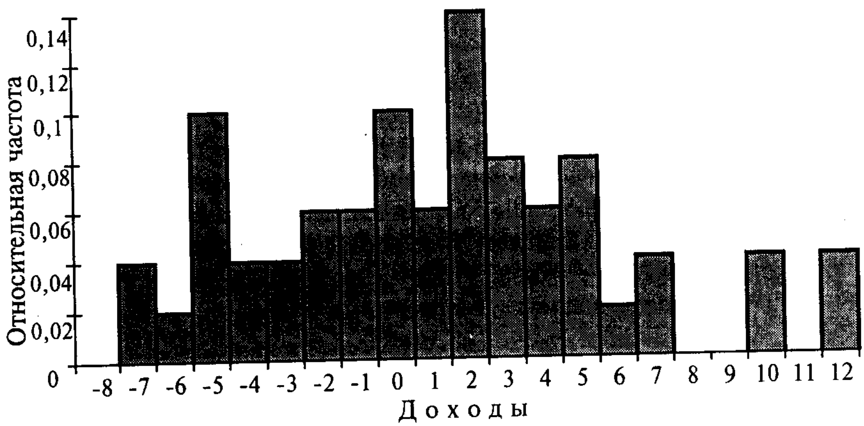

Рис. 2.2. Распределение относительных частот месячных доходов по индексу FTSE 100 с сентября 1989 г. по декабрь 1993 г.

График распределения накопленных частот, или огива

Для построения графиков распределения накопленных частот накопленная частота откладывается по вертикальной оси, а величины интервалов — по горизонтальной.

Огивы строятся нанесением на ось абсцисс данных в возрастающем порядке, и в случае сгруппированных данных нанесением частот напротив верхней границы интервала. Это проиллюстрировано на рис. 2.3.

Гистограмма

Следующим шагом в представлении относительного распределения будет его отображение не высотой, а площадью столбиков. Данное графическое представление известно как **гистограмма**. Для иллюстрации построения гистограммы рассмотрим следующие сгруппированные данные:

Доходы	Частота	Относительная частота
До —0,04	3	0,0015
Более —0,04 и до —0,02	29	0,0144
Более —0,02 и до 0,01	191	0,0947
Более —0,01 и до 0,00	714	0,3538
Более 0,00 и до 0,01	816	0,4044
Более 0,01 и до 0,02	235	0,1165
Более 0,02 и до 0,04	29	0,0144
Более 0,04 и до 0,05	1	0, 0005
Итого	2018	

Гистограмма должна быть построена так, чтобы площадь каждого "столбика" представляла относительную частоту распределения, а общая их площадь составляла 100% распределения. Горизонтальная ось показывает изменения в доходах на 0,01%. Однако два интервала — "более—0,04 и до 0,02" и "более + 0,02 и до + 0,04" — представляют изменения на 0,02%.

Таблица 2.4

Процентные доходы	Частота	Относительная частота	Накопленная частота
—8	0	0	0
—7	2	0,039215	2
—6	1	0,019607	3
—5	5	0,098039	8
—4	2	0,039215	10
—3	2	0,039215	12
—2	3	0,058823	15
—1	3	0,058823	18

Продолжение таблицы

0	5	0,098039	23
1	3	0,058823	26
2	7	0,137254	33
3	4	0,078431	37
4	3	0,058823	40
5	4	0,078431	44
6	1	0,019607	45
7	2	0,039215	47
8	0	0	47
9	0	0	47
10	2	0,039215	49
11	0	0	49
12	2	0,039215	51

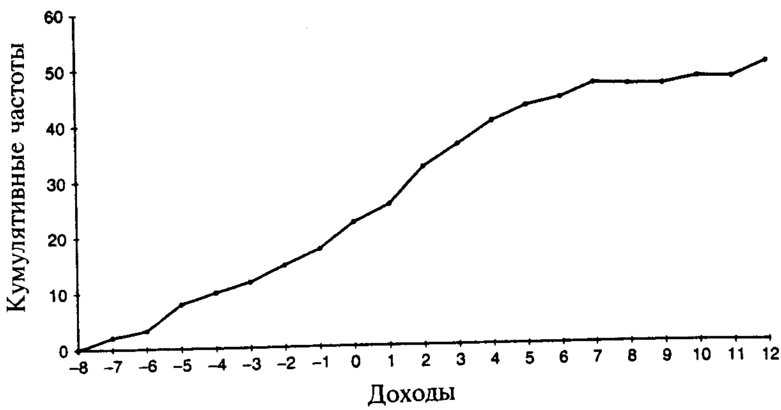

Рис. 2.3. Распределение кумулятивных частот месячных доходов по индексу FTSE 100 с сентября 1989 г. по декабрь 1993 г.

В добавление к этому существует еще и проблема, связанная с тем, что нижняя граница первого интервала не является открытой. Используя наши знания, предположим, что этот интервал ограничен значением —0,05. Для того чтобы убедиться, что площадь столбиков над каждым из интервалов показывает относительную частоту, мы должны разделить относительную частоту на число "единиц" в каждом интервале.

Для иллюстрации рассмотрим 29 наблюдений доходов в группе "более 0,02 и до 0,04", которые представляют интервал размером 0,02%, в то время как остальные интервалы имеют размер 0,01%. Поэтому высота "столбика" над этим интервалом рассчитывается как (0,0144/0,02) = 0,72. Сравним это с высотой "столбика" для 235 наблюдений для предыдущего интервала. Он соответствует 0,01, отсюда высота находится как (0,1165/0,01) = 11,65.

Можно кратко изложить всю эту процедуру, сказав, что при построении гистограммы должен быть определен размер интервала. Затем рассчитывается высота каждого столбика делением относительной частоты на число единиц в конкретном интервале. Потому высота "столбиков", относящихся к интервалам, которые уже, будет больше, а относящихся к тем, которые шире, будет меньше. Это проиллюстрировано на рис. 2.4.

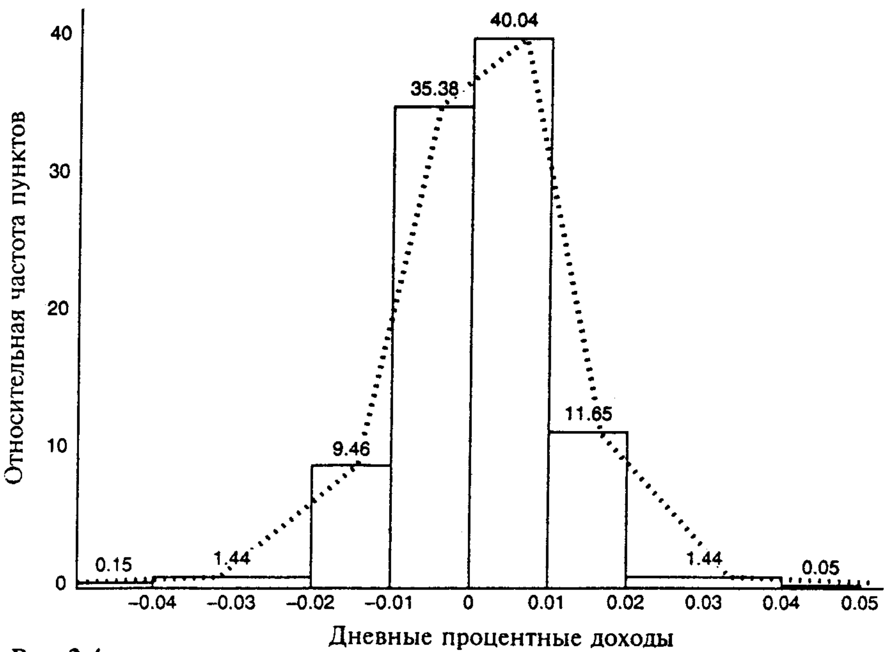

Рис. 2.4

Полигон — это линия, которая соединяет середины вершин каждого из столбиков. Подобная линия построена на рис. 2.4. Если мы предположим, что интервалы (величины изменения доходов) на горизонтальной оси бесконечно малы, то полигон становится плавной кривой. Площадь под кривой составляет 100% случаев наблюдаемых изменений доходов. В гл. 4 мы встретим подобные кривые снова и будем ссылаться на них как на плотности вероятности.

ОПИСАТЕЛЬНЫЕ СТАТИСТИЧЕСКИЕ ПОКАЗАТЕЛИ

Введение

Для представления обобщающих показателей, предположим, доходности активов, цен активов или числа финансовых операций, мы используем **показатели центра распределения, показатели вариации, показатели скошенности (асимметрии)** и **показатели эксцесса**. Средние величины позволяют сделать вывод о центральном или наиболее общем значении, найденном для совокупности данных, меры рассеяния (вариации) показывают, как данные распределены вокруг средней. Показатели скошенности (асимметрии) иллюстрируют степень левосторонней асимметрии, т.е. отрицательной, или правосторонней, т.е. положительной, в распределении частот. Показатели эксцесса определяют уровень островершинности или плосковершинности распределения частот.

Поясним понятие **моментов**. Момент k-го порядка относительно исходной величины A находится как

$$\frac{\sum_{i=1}^{n}(X_i - A)^k}{n}. \tag{2.1}$$

Если $A = 0$ и $k = 1$, то мы получим то, что, как будет показано далее, является средней арифметической. Поэтому средняя арифметическая иногда называется моментом первого порядка относительно нуля. Если же величина A сама является средней арифметической и $k = 2$, мы имеем момент второго порядка относительно средней (центральный момент второго порядка), известный как дисперсия, и характеризующий вариацию признака. При A, равном средней, и $k = 3$ получаем момент третьего порядка относительно средней (центральный момент третьего порядка), который является мерой скошенности, а если $k = 4$, то определяется момент четвертого порядка относительно средней (центральный момент четвертого порядка), измеряющий эксцесс.

Показатели центра распределения

Практически каждый человек знаком с концепцией средней, будь это "средний" размер остатка по кредитной карте, "среднее" число операций по счету или даже "средний" счет для матча в крикет.

Фактически существует несколько показателей "средней", которые особенно интересны в сфере финансов. Это:

- мода;
- медиана;
- средняя арифметическая;
- средняя геометрическая.

Примеры каждого из показателей даны ниже как для несгруппированных, так и для сгруппированных данных.

Мода

Мода — это наиболее часто наблюдаемая величина изучаемой переменной. Для ее иллюстрации рассмотрим следующие данные, которые показывают цену акции, выраженную в пенсах, в течение 15-дневного периода:

10,	12,	9,	8,	10,	15,	14,	12
11,	10,	12,	12,	10,	12,	11.	

Модой, т.е. наиболее часто повторяющимся наблюдением, является величина 12.

Медиана

Медиана — это значение наблюдения, которое находится в середине ранжированного ряда данных, т.е. наблюдение, занимающее срединное положение.

Медиана для несгруппированных данных. Для определения медианы в случае несгруппированных данных мы сначала должны расположить их в возрастающем порядке. Покажем это на примере, использованном при рассмотрении моды

8, 9, 10, 10, 10, 10, 11, 11, 12, 12, 12, 12, 12, 14, 15.

Так как присутствуют 15 наблюдений, медианой является значение восьмого наблюдения, т.е. величина признака, равная 11.

Если бы в примере было четное число наблюдений, то отсутствовало бы "срединное" наблюдение, поэтому была бы рассчитана средняя из двух значений, стоящих посередине. Результатом может являться число, не присутствующее на самом деле в ряду данных.

Медиана для сгруппированных дискретных данных. Когда используются сгруппированные данные, мы не знаем индивидуальных значений наблюдений. Таким образом, мы вынуждены "оценивать" значения описательных статистических показателей. Для этого необходимо сделать некоторые допущения по поводу распределения наблюдений в каждом интервале. Обычно предполагается, что данные равномерно распределены в каждом из интервалов.

Для оценки медианы по сгруппированным дискретным данным мы сначала должны построить распределение накопленных частот. Это достигается простым сложением каждой следующей частоты с текущей общей. Чтобы проиллюстрировать это, рассмотрим снова данные из табл. 2.2.

Для расчета значения медианы необходимо сначала найти позицию срединного наблюдения, оценивая ее как

$$\left(\frac{n+1}{2}\right)\text{-е наблюдение.} \tag{2.2}$$

Используя приведенные выше данные, получаем $(52 + 1)/2 = 26{,}50$. "Двадцать шестое с половиной" наблюдение присутствует в группе, относящейся к уровню индекса 2400,1—2500. Отсюда медиана находится в интервале 2400,1—2500. Так как 26 с половиной меньше накопленной суммы 27, медиана будет очень близка к 2500. Определим ее, интерполируя следующим образом:

$$\left[2400{,}1 + \left(\frac{8{,}5\,*}{9} \times 100\right)\right] \approx 2494{,}5\,.$$

Медиана для сгруппированных непрерывных данных. Для иллюстрации рассмотрим данные из табл. 2.3.

---

* 8,5 получено как разность между номером медианы (26,5) и накопленной частотой интервала, предшествующего медианному (18), т.е. 26,5 — 18,0 = 8,5.

Чтобы оценить медиану для этих данных, применим формулу

$$\text{Медиана} = L + i\left(\frac{\frac{n+1}{2} - F}{f}\right), \tag{2.3}$$

где L — нижняя граница группы, в которой должна находиться медиана;

i — величина интервала, в котором должна находиться медиана;

F — накопленная частота интервала, предшествующего интервалу медианы;

f — частота интервала, в котором должна находиться медиана.

Положение срединного наблюдения (номер медианы) находится как $(n + 1)/2 = 52/2 = 26$. Следовательно, оно расположено в интервале от 0 до + 1%. Эта группа имеет размер интервала (i), равный 1. Накопленная частота до этой группы (F) равняется 23, а частота f интервала, в котором находится медиана, составляет 3.

Таким образом, в результате расчета получим:

$$\text{Медиана} = 0 + 1\left(\frac{26-23}{3}\right) = 0 + [1(1)] = 1{,}0\%.$$

Средняя арифметическая

Средняя арифметическая — наиболее часто используемый показатель центра распределения, именно ее большинство людей рассматривают в качестве средней.

Средняя арифметическая для несгруппированных данных. Средняя арифметическая рассчитывается суммированием всех значений отдельных наблюдений и последующим делением полученной суммы на количество наблюдений. Например, допустим, что мы желаем подсчитать среднюю арифметическую цены какого-либо актива в течение пяти дней. В данный период мы наблюдаем следующие цены

225 225 240 215 230

Найдем среднюю арифметическую, складывая эти пять значений и деля сумму на число наблюдений. В формализованном виде средняя арифметическая выглядит так:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}. \tag{2.4}$$

Здесь показано, что $\overline{X}$ (икс — средняя) — это сумма X_i (отдельных наблюдений X, т.е. цен ценной бумаги в этом примере), поделенная на количество наблюдений. Греческая заглавная буква сигма (Σ) — оператор суммирования, означает, что все X_i необходимо сложить вместе. Так как в нашем примере всего пять X_i, то оператор суммирования будет выглядеть следующим образом

$$\sum_{i=1}^{5},$$

показывая, что первые пять X_i должны быть сложены друг с другом.

Сейчас мы можем подсчитать среднюю арифметическую пяти цен ценной бумаги:

$$\frac{225 + 225 + 240 + 215 + 230}{5} = 227.$$

Таким образом, средняя арифметическая цена равняется 227.

Применим это для данных дохода FTSE в табл. 2.5. Данные, относящиеся к ежемесячным наблюдениям уровня индекса, находятся в левом столбце. Данные в правом столбце отражают ежемесячные непрерывно наращиваемые доходы.

Таблица 2.5

FTSE 100	Непрерывно наращиваемые доходы
2407,5	
2289,2	—5,04
2160,1	—5,80
2311,1	6,76
2422,7	4,72
2345,8	—3,23
2238,4	—4,69
2221,6	—0,75
2117,9	—4,78
2371,4	11,31
2372	0,03
2339	—1,40
2166,6	—7,66
Σ	—10,54
Средняя	—0,88

Начнем с суммирования всех 12 наблюдений дохода, получив —10,54. Следующим шагом будет деление этой суммы на 12. Результат деления равен —0,88. Отсюда средняя арифметическая непрерывно наращиваемых ежемесячных доходов по индексу FTSE в данном периоде составляла —0,88%.

Средняя арифметическая для сгруппированных данных. Часто верхние и нижние пределы крайних интервалов остаются открытыми. Следовательно, необходимо сделать допущение о границах интервалов. Это допущение должно базироваться на характеристиках анализируемых данных. Например, данные из табл. 2.6 относятся к работникам компании финансовых услуг. Нижний интервал обозначен "До 26 лет", а верхний — "Более 55 лет". Если бы было известно, что компания, о которой идет речь, придерживается политики, которая не позволяет содержать в штате людей младше 17 и старше 62 лет, то границы открытых интервалов были бы определены знанием этой политики.

Таблица 2.6

Возраст	Средняя точка (X_i)	f_i	fX_i
До 26 лет	21,5	20	430
26 до 40 лет	33,0	35	1155
40 до 55 лет	47,5	42	1995
Более 55 лет	58,5	15	877,5
		112	4457,5

Средняя арифметическая для сгруппированных данных находится с помощью следующей формулы

$$\overline{X} = \frac{\sum_{i=1}^{m} fX_i}{n}, \tag{2.5}$$

где X_i — центральное значение каждого интервала;

f_i — частота каждого интервала;

m — число интервалов;

n — общее количество наблюдений или единиц совокупности.

Используя приведенные выше данные, рассчитаем среднюю арифметическую возраста работников:

$$\overline{X} = \frac{4457,5}{112} = 39,8 \text{ лет.}$$

Средняя геометрическая

Альтернативным показателем средней арифметической, особенно хорошо подходящим для измерения средних темпов роста, является **средняя геометрическая**. Для иллюстрации этого сначала предположим, что фондовый индекс изменялся со следующими годовыми темпами прироста в течение пяти лет: + 10%, + 20%, + 15%, –30%, + 20%. Средняя арифметическая темпов прироста равна + 35/5 = 7. Однако 100 единиц, инвестированных в первый год, возрастут до следующих значений в каждом году соответственно: 110; 132; 151,80; 106,26; 127,51. Следовательно, фактический прирост за весь пятилетний период составляет лишь 27,5%. Разделив это число на пять, мы получим 5,5% в год; но правильный ли это ответ?

Правду говоря, нет! Что нам действительно нужно, так это единичная мера темпа прироста, которая при повторении ее *n* раз трансформирует начальное значение в конечное. Корректный показатель периодических темпов прироста находится с использованием средней геометрической. Для измерения среднего годового темпа прироста за *n* лет используется следующая формула:

$$\overline{X}_g = \sqrt[n]{X_1 \cdot X_2 \cdot X_3 \ldots \cdot X_n} - 1, \tag{2.6}$$

где X_i имеют вид $(1 + r)$, а величина r, темп прироста, выражена в десятичных дробях, т.е., например, 10% = 0,1.

При использовании приведенных выше данных средняя геометрическая из темпов прироста в течение пяти лет составит 4,98%:

$$\overline{X}_g = \sqrt[5]{(1{,}1) \cdot (1{,}2) \cdot (1{,}15) \cdot (0{,}7) \cdot (1{,}2)} - 1 =$$

$$= 1{,}0498 - 1 = 0{,}0498 \; (4{,}98\%).$$

Таким образом, средняя арифметическая дает преувеличенное значение среднего годового темпа прироста в данном примере. Применение же средней геометрической дает корректную "среднюю" темпа прироста.

Какую среднюю использовать?

Выбор подходящего вида средней зависит, с одной стороны, от природы данных, а с другой — от того, как этот показатель будут использовать.

Средняя арифметическая особо чувствительна к экстремальным (выделяющимся) значениям в одном из направлений, которые называются смещенными данными. Выделяющиеся большие значения увеличивают среднюю выше уровня действительного представляющего точку центра распределения данных. Особо малые значения признаков имеют противоположный эффект. Иногда для того чтобы исключить влияние экстремальных единиц данных, рассчитывается **усеченная средняя**. Для этого просто необходимо удалить 5% наибольших и 5% наименьших наблюдений до расчета средней арифметической.

Экстремальные наблюдения не влияют на медиану и моду, но эти показатели не столь полезны в дальнейшем математическом и статистическом анализе.

Средняя геометрическая лучше других подходит, когда подсчитываются "средние" темпы прироста в течение нескольких временных периодов.

Показатели вариации (меры рассеяния)

Имея представление о точке центра распределения, мы часто хотим знать, как данные рассеяны вокруг нее. Нам предстоит изучить следующие показатели рассеяния (вариации):

- квартильное отклонение;
- дисперсию;
- среднее квадратическое отклонение.

Квартильное отклонение

В то время как серединное наблюдение называется медианой, значения, находящиеся в точках, когда 25% всех значений меньше Q_1 и 75% меньше Q_3 в ряду наблюдений, называются соответственно нижним и верхним квартилями (Q_1 и Q_3). Промежуток между ними (Q_1 — Q_3) известен как межквартильный, а его половина — как квартильное отклонение, которое часто используется как индикатор рассеяния данных вокруг медианы.

Расчет квартилей для несгруппированных данных. Положение (номер) Q_1 и Q_3 определяется следующим образом:

$$Q_1 = \frac{n+1}{4},$$
$$Q_3 = \frac{3(n+1)}{4}. \qquad (2.7)$$

Q_1 и Q_3 — это наблюдения, ближайшие к $(n + 1)/4$ или $3(n + 1)/4$ соответственно.

Используем те же несгруппированные данные, относящиеся к ценам акции, выраженным в пенсах, что и для медианы

8, 9, 10, 10, 10, 10, 11, 11, 12, 12, 12, 12, 12, 14, 15.
Позиция $Q_1 = (n + 1)/4 = 3{,}75$, т.е. 4.
Позиция $Q_3 = 3(n + 1)/4 = 11{,}25$, т.е. 11.

Таким образом, значением Q_1 является 10, а значением Q_3 — 12.

Расчет квартилей для сгруппированных данных. Для оценки Q_1 и Q_3 в случае сгруппированных данных мы применим методологический аналог того, что использовалось при определении медианы для сгруппированных данных. Исходная информация приведена в табл. 2.7.

Чтобы оценить Q_1 и Q_3 для этих данных, применим следующие формулы:

$$Q_1 = L + i\left(\frac{\frac{(n+1)}{4} - F}{f}\right),$$
$$Q_3 = L + i\left(\frac{\frac{3(n+1)}{4} - F}{f}\right), \qquad (2.8)$$

где L — нижняя граница интервала, в котором находится соответствующий квартиль;
i — размер квартильного интервала;
F — накопленная частота до квартильного интервала;
f — частота квартильного интервала.

Номер Q_1 находится как $(n + 1)/4 = 52/4 = 13$. Отсюда находим, что он расположен в интервале между —3 и —2. Величина этого интервала (i) составляет 1. Накопленная частота до данного интервала (F) равна 12, а частота интервала, в котором находится Q_1, равна 3.

Таким образом, Q_1 оценивается так:

$$Q_1 = -3 + 1\left(\frac{13-12}{3}\right) = -3 + 1(0{,}333) = -2{,}667\%.$$

Номер Q_3 составляет $3(n + 1)/4 = 39$, $L = +3$, $F = 37$, $f = 3$. Значение Q_3 рассчитывается так:

$$Q_3 = 3 + 1\left(\frac{39-37}{3}\right) = 3 + 1(0{,}666) = 3{,}666\%.$$

Межквартильный промежуток $Q_3 - Q_1 = 3{,}666 (- 2{,}667) = 6{,}333\%$. Квартильное отклонение равно $(Q_3 - Q_1)/2 = 6{,}333/2 = 3{,}1665$.

Процентили

Так же как мы поделили единицы совокупности на квартили, можно поделить их на процентили. Первый процентиль находится в точке, отделяющей 1% в совокупности данных. Q_1 и Q_3 совпадают с 25-м и 75-м процентилями соответственно, а медиана является 50-м процентилем.

Таблица 2.7

Месячный доход	Число наблюдений	Накопленная частота
До —8%	0	0
Более —8% и до —7%	2	2
Более —7% и до —6%	1	3
Более —6% и до —5%	5	8
Более —5% и до —4%	2	10
Более —4% и до —3%	2	12
Более —3% и до —2%	3	15
Более —2% и до —1%	3	18
Более —1% и до 0%	5	23
Более 0% и до + 1%	3	26
Более + 1% и до + 2%	7	33
Более + 2% и до + 3%	4	37
Более + 3% и до + 4%	3	40
Более + 4% и до + 5%	4	44
Более + 5% и до + 6%	1	45
Более + 6% и до + 7%	2	47
Более + 7% и до + 8%	0	47
Более + 8% и до + 9%	0	47
Более + 9% и до + 10	2	49
Более + 10% и до + 11%	0	49
Более + 11% и до + 12%	2	51
	51	

Чтобы проиллюстрировать оценку процентилей по сгруппированным данным, рассчитаем 90-й процентиль для данных о доходе по FTSE 100, рассмотренных ранее.

Позиция 90-го процентиля равна $0{,}9 \cdot 52 = 46{,}8$. Значение 90-го процентиля находится так:

$$90\text{-й процентиль} = 6 + \left[1 \times \left(\frac{46{,}8 - 45}{2}\right)\right] = 6 + 0{,}9 = 6{,}9\%.$$

Дисперсия и среднее квадратическое отклонение

Если средняя арифметическая выбрана как показатель центра распределения, то соответствующими показателями вариации являются дисперсия и среднее квадратическое отклонение. Дисперсия широко применяется в финансовых расчетах как мера риска и неопределенности и привлекательна, так как имеет свойство аддитивности. Это детально рассмотрено в гл. 4. Среднее квадратическое отклонение имеет сходное применение и используется как мера изменчивости в ценообразовании опционов, которое рассмотрено в гл. 8 и 10. Однако поскольку среднее квадратическое отклонение является квадратным корнем дисперсии, что мы увидим позднее, оно неаддитивно.

Дисперсия для несгруппированных данных. Если отдельные наблюдения сгруппированы вблизи среднего значения, то разности между каждым индивидуальным наблюдением X_i и средней $\overline{X}$ будут небольшими. Если отдельные наблюдения широко рассеяны, то разности между каждым индивидуальным значением X_i и $\overline{X}$ будут велики.

Можно подумать, что с помощью суммирования всех $(X_i - \overline{X})$ мы получим меру рассеяния. К сожалению, это не так, потому что $\Sigma(X_i - \overline{X})$ всегда равна нулю. Для преодоления этого свойства средней необходимо возвести в квадрат каждое отклонение $(X_i - \overline{X})$ и просуммировать получившиеся значения. Деление результата на $n-1$, т.е. на количество наблюдений минус один, дает дисперсию. Это можно показать в формализованном виде:

$$\sigma^2 = \frac{\sum (X_i - \overline{X})^2}{n-1}. \tag{2.9}$$

Если вычисляется дисперсия на основе всей совокупности данных, то делитель в выражении (2.9) равен n. Однако когда данные, которыми мы располагаем, представляют **выборку** из всей совокупности данных, что обычно случается в эмпирических исследованиях в области финансов, следует делить на *n*—1. Происходит это потому, что мы должны делить на то, что статистики называют **степенями свободы**, для предотвращения искажения результата. Степени свободы — это число наблюдений минус число параметров, оцененных по рассматриваемым данным. В случае дисперсии средняя оценивается по тем же данным. Таким образом, когда эта оценка используется при вычислении среднего квадратического отклонения, число степеней свободы равно *n*—1. Наибольший эффект при коррекции *n* на число степеней свободы наблюдается для малых выборок. Понятно, что для большого объема данных эффект коррекции минимален.

Чтобы проиллюстрировать расчет дисперсии для несгруппированных данных, используем табл. 2.8, в которой показаны доходы по индексу FTSE 100 для первых 12 месяцев из приведенных ранее в табл. 2.5. Сначала находятся и располагаются в третьем столбце отклонения от средней ($X_i - \overline{X}$). Затем каждое из них возводится в квадрат и записывается в четвертый столбец. Данные в этом столбце суммируются и получаем 361,9. Это число делится на 11, т.е. на *n*—1, и результат составляет 32,90. Данная величина отражает процентные доли в квадрате, что интуитивно не очень привлекательный результат.

Таблица 2.8

FTSE 100	Доходы X	$X_i - \overline{X}$	$(X_i - \overline{X})^2$
2407,5			
2289,2	–5,04	–4,16	17,31
2160,1	–5,80	–4,93	24,27
2311,1	6,76	7,64	58,3
2422,7	4,72	5,59	31,3
2345,8	–3,23	–2,35	5,51
2238,4	–4,69	–3,81	14,5
2221,6	–0,75	0,13	0,02
2117,9	–4,78	–3,9	15,22
2371,4	11,31	12,18	148,45
2372	0,03	0,90	0,82
2339	–1,40	–0,52	0,27
2166,6	–7,66	–6,78	45,94
Σ	–10,54	0,00	361,90
Средняя	–0,88	Дисперсия	32,90

Дисперсия для сгруппированных данных. При использовании сгруппированных данных формула дисперсии выглядит так:

$$\frac{\sum f_i(X_i - \bar{X})^2}{n-1}, \qquad (2.10)$$

где f_i представляет частоту для каждого интервала, X_i — центральное значение каждого интервала, а $\bar{X}$ —средняя всех X.

Для иллюстрации этого расчета используем сгруппированные данные из табл. 2.9, которые относятся (см. выше приведенный пример расчета средней по сгруппированным данным), к политике набора персонала в финансовой организации.

Таблица 2.9

Возраст (X)	Центральное значение интервала (X_i)	f_i	$X_i - \bar{X}$	$(X_i - \bar{X})^2$	$f_i(X_i - \bar{X})^2$
До 26 лет	21,5	20	−18,30	334,89	6697,8
26 до 40 лет	33,0	35	−6,80	46,24	1618,4
40 до 55 лет	47,5	42	7,70	59,29	2490,18
Более 55 лет	58,5	15	18,70	349,69	5245,35
Всего		112			16051,73

Помня, что $\bar{X}$ = 39,8 лет, определяем дисперсию:

$$\sigma^2 = \frac{16051{,}73}{112-1} = 144{,}61.$$

Величина 144,61 выражена в квадратах лет, что довольно неудобно для рассмотрения.

Среднее квадратическое отклонение

Дисперсия выражается в квадратах единиц, находящихся в основании расчета, что делает ее интерпретацию довольно затруднительной. Эту проблему можно преодолеть, извлекая квадратный корень из дисперсии и получая таким образом **среднее квадратическое отклонение**, обозначаемое как σ.

Среднее квадратическое отклонение для несгруппированных данных. Среднее квадратическое отклонение вычисляется так:

$$\sigma = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}}. \qquad (2.11)$$

Опираясь на расчет дисперсии для несгруппированных данных дохода по индексу FTSE 100 из табл. 2.8, получим квадратный корень из 32,9, который равен 5,74.

Среднее квадратическое отклонение для сгруппированных данных. Формула для нахождения среднего квадратического отклонения для сгруппированных данных выглядит так:

$$\sigma = \sqrt{\frac{\sum f_i (X_i - \bar{X})^2}{n-1}}. \tag{2.12}$$

Снова видим, что это только квадратный корень из дисперсии, но на этот раз для сгруппированных данных. Необходимо также помнить, что в этом примере X_i является центральным значением интервала, а $\bar{X}$ — это средняя арифметическая всех наблюдений. Среднее квадратическое отклонение для сгруппированных данных, относящихся к возрасту нанимаемых работников, составляет 12,02 лет.

Отрицательная полудисперсия и отрицательное полуотклонение

Отрицательная полудисперсия подобна дисперсии, но учету подлежат лишь отрицательные отклонения от средней. Показатель отрицательной полудисперсии иногда предлагается в качестве приемлемой меры риска в финансовой теории, когда доходы распределены несимметрично. Ее величина рассчитывается следующим образом:

$$\bar{\sigma}^2 = \sqrt{\frac{\sum (R_i - \bar{R})^2}{n-1}}, \tag{2.13}$$

где R_i — периодические доходы от портфеля ценных бумаг;
n — общее число отрицательных значений доходов;
$\bar{R}$ — средний доход за период с учетом положительных и отрицательных доходов;
$\bar{\sigma}^2$ — полудисперсия.

Однако заметим, что при условии $R_i < \bar{R}$ в расчет принимаются только отрицательные отклонения от средней. При использовании данных приведенного выше примера с дисперсией для несгруппированных данных отрицательная полудисперсия составляет 20,50.

Квадратный корень отрицательной полудисперсии представляет собой отрицательное полуотклонение, обозначаемое как $\bar{\sigma}$, для тех же данных оно равно 4,53.

Какой показатель вариации использовать?

Выбор показателя вариации диктуется используемым показателем центра распределения. При применении медианы как меры "средней" показателем вариации будет служить квартильное отклонение. Если же используется средняя арифметическая, то будут выбраны дисперсия, среднее квадратическое отклонение или отрицательная полудисперсия.

Относительные показатели вариации

Коэффициент вариации

Среднее квадратическое отклонение выражается в единицах измерения, лежащих в основе расчета. Таким образом, при сравнении степени вариации переменных должны быть учтены различия в величине этих переменных. Для этого нужно рассчитать коэффициент вариации. Он находится как отношение:

$$CV = \frac{\sigma}{\bar{X}}. \qquad (2.14)$$

Примером приложения коэффициента вариации является сравнение вариации изменений уровней индекса FTSE 100 в Великобритании и сводного индекса S&P в США. Первый из них на момент написания книги составлял около 3700, а второй — около 650. Сравнение вариаций изменений уровней с использованием среднего квадратического отклонения ввело бы в заблуждение, поэтому необходимо применить коэффициент вариации. Однако заметьте, что если бы мы измеряли вариацию доходов по каждому из этих двух индексов, то использование среднего квадратического отклонения было бы оправданным. Дело в том, что при подсчете доходности разница в величине основных данных была бы уже учтена.

Коэффициент асимметрии

Важно рассмотреть, есть ли смещения в рассеянии данных. Индикатор этих смещений — скошенность (асимметрия) данных. На рис. 2.5 показаны две типичные формы скошенности и симметричное распределение и отмечено относительное положение средней арифметической, медианы и моды.

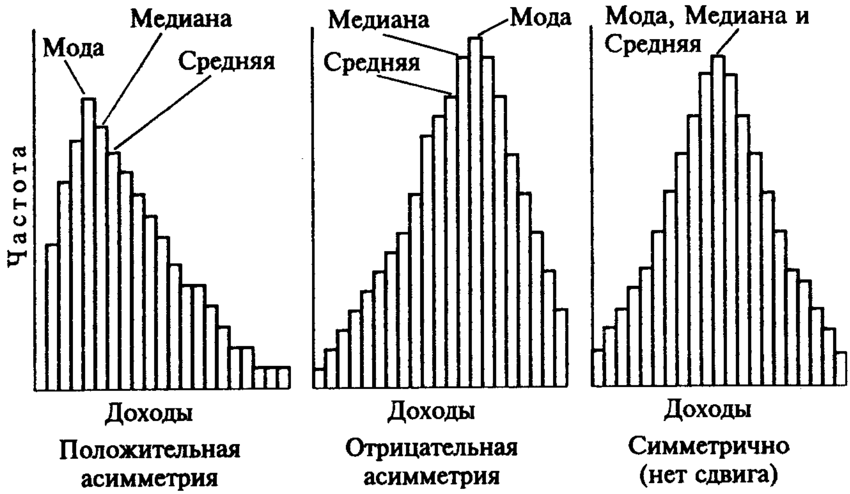

Рис. 2.5. Три примера степени асимметрии распределений

В случае положительной асимметрии распределение имеет длинную правую ветвь. Средняя величина дохода больше медианы, которая в свою очередь больше моды. Значение средней больше медианы и моды, потому что на нее повлияли несколько очень больших значений доходов.

Отрицательная асимметрия проявляется в виде более длинной левой ветви, а величина средней меньше медианы и моды. Большинство наблюдений распределения имеют значения больше средней, но величина средней снижается из-за нескольких очень малых наблюдений.

При симметричном распределении средняя, медиана и мода имеют одно и то же значение.

Наращение доходов по сложной ставке дает увеличение положительной асимметрии. Чтобы понять это, рассмотрим конечную стоимость актива, который приносил положительные доходы в размере 8% годовых в течение двух лет. Если первоначальные инвестиции были равны 100 единицам, то стоимость на конец срока составит

$$100(1{,}08)^2 = 116{,}64 \text{ ед.}$$

Если же наблюдались убытки в размере 8% ежегодно, то через два года стоимость 100 единиц равнялась бы 85,73 единицам:

$$\frac{100}{(1{,}08)^2} = 85{,}73 \text{ ед.}$$

Таким образом, наращение по сложной ставке дает доход 16,64 единиц через два года, в то время как снижение доходности принесет убытки в размере 14,27 единиц за тот же период.

Существует несколько методов для расчета степени асимметрии данных. Коэффициент асимметрии Спирмэна находится следующим образом:

$$\text{Коэффициент асимметрии} = \frac{3(\text{Средняя арифметическая} - \text{Медиана})}{\text{Среднее квадратическое отклонение}}. \tag{2.15}$$

Также возможно определение коэффициента асимметрии с помощью квартилей и процентилей.

Квартильный коэффициент асимметрии равен:

$$\frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_1)} = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}; \tag{2.16}$$

10—90-процентильный коэффициент асимметрии рассчитывается как

$$\frac{(P_{90} - P_{50}) - (P_{50} - P_{10})}{P_{90} - P_{10}} = \frac{P_{90} - 2P_{50} + P_{10}}{P_{90} - P_{10}}. \tag{2.17}$$

Однако показателем асимметрии, который наиболее пригоден для применения в случае сгруппированных данных, является коэффициент асимметрии, основанный на расчете моментов распределения. Он определяется с помощью центрального момента третьего порядка и деления его на куб среднего квадратического отклонения, что можно представить следующей формулой:

$$\frac{\dfrac{\sum (X - \bar{X})^3}{n-1}}{\left(\sqrt{\dfrac{\sum (X - \bar{X})^2}{n-1}}\right)^3}. \tag{2.18}$$

И снова, используя доходы по индексу FTSE 100, обратимся к табл. 2.10, где выражение (2.18) применено к наблюдениям доходов.

Таблица 2.10

FTSE 100	Доходы X	$X_i - \bar{X}$	$(X_i - \bar{X})^3$
2407,5			
2289,2	–5,04	–4,16	–71,99
2160,1	–5,80	–4,93	–119,55
2311,1	6,76	7,64	445,15
2422,7	4,72	5,59	175,10
2345,8	–3,23	–2,35	–12,93
2238,4	–4,69	–3,81	–55,22
2221,6	–0,75	0,13	0,00
2117,9	–4,78	–3,9	–59,40
2371,4	11,31	12,18	1808,77
2372	0,03	0,90	0,74
2339	–1,40	–0,52	–0,14
2166,6	–7,66	–6,78	–311,37
Σ	–10,54	0,00	1799,17
Средняя	–0,88	Центральный момент третьего порядка	163,56
		Среднее квадратическое отклонение в кубе	188,71
		Коэффициент асимметрии	0,87

Отклонения от средней возведены в третью степень и размещены в четвертом столбце. Итог этого столбца поделен на 11, что дает результат, равный 163,56. И, наконец, центральный момент третьего порядка поделен на куб среднего квадратического отклонения для получения коэффициента асимметрии, значение которого равно 0,87.

Если бы распределение доходов по индексу акций было симметричным, то коэффициент асимметрии был бы равен нулю. Так как в результате расчетов мы получили положительное число, то данные доходов по индексу FTSE 100 имеют положительную асимметрию.

Эксцесс

В то время как показатели асимметрии характеризуют симметричность распределения частот, показатели эксцесса описывают пиковость этого распределения. Распределения, имеющие более выраженный пик, чем у нормального распределения (о нем детально рассказано далее), называются островершинными. Те же распределения, у которых степень вытянутости вдоль оси ординат меньше, чем у нормальной кривой, называются плосковершинными, а распределения, которые похожи на нормальное, — средневершинными. Общая форма подобных распределений показана на рис. 2.6.

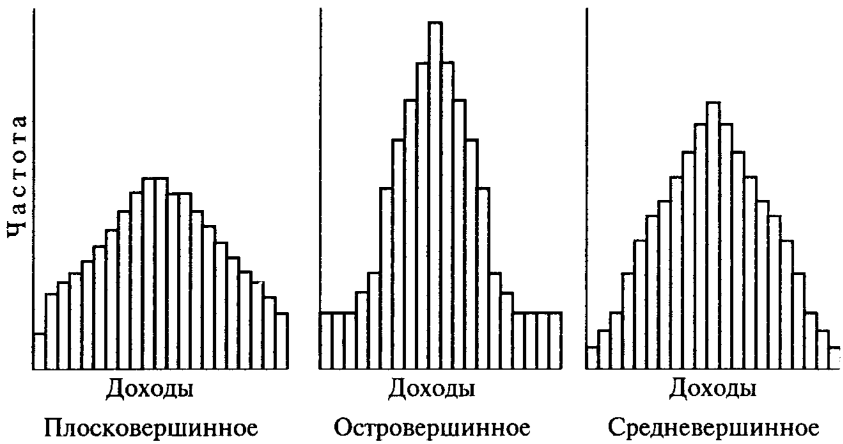

Рис. 2.6. Три вида эксцесса

Островершинные распределения можно увидеть в доходах активов, когда наблюдаются периодические скачки в ценах на эти активы. Рынки с прерывающимся процессом торговли, как, например, рынки ценных бумаг, которые закрываются на ночь и в выходные, имеют большую вероятность для демонстрации скачков цен активов. Причиной служит то, что информация, которая влияет на цены, но публикуется во время закрытия рынков, окажет воздействие на цены, когда рынок откроется вновь, таким образом создавая условия для скачка между предыдущей ценой закрытия и ценой открытия. Этот скачок цен, наиболее заметный для ежедневных или еженедельных данных, приведет к более высокой частоте повторения больших положительных или отрицательных доходов, чем предполагалось бы, если бы торговля на рынках велась непрерывно.

Коэффициенты эксцесса могут быть определены с помощью процентилей и квартилей или центральных моментов распределения.

Процентильный коэффициент эксцесса. Он рассчитывается как отношение:

$$\frac{\frac{1}{2}(Q_3 - Q_1)}{(P_{90} - P_{10})}. \tag{2.19}$$

Коэффициент эксцесса на основе моментов распределения. Коэффициент эксцесса находится делением центрального момента четвертого порядка на среднее квадратическое отклонение, возведенное в четвертую степень. Расчет выглядит следующим образом:

$$\frac{\dfrac{\sum (X - \overline{X})^4}{n}}{\left(\sqrt{\dfrac{\sum (X - \overline{X})^2}{n-1}}\right)^4}. \tag{2.20}$$

Числитель выражения (2.20) $\Sigma(X_i - \overline{X})^4/N$ — центральный момент четвертого порядка, знаменатель — это дисперсия, возведенная в квадрат. Расчет эксцесса по формуле (2.20) на основании данных доходов по индексу акций, использованных ранее, приведен в табл. 2.11.

Таблица 2.11

FTSE 100	Доходы X	$X_i - \overline{X}$	$(X_i - \overline{X})^4$
2407,5			
2289,2	–5,04	–4,16	299,50
2160,1	–5,80	–4,93	588,91
2311,1	6,76	7,64	3398,97
2422,7	4,72	5,59	979,58
2345,8	–3,23	–2,35	30,34
2238,4	–4,69	–3,81	210,26
2221,6	–0,75	0,13	0,00
2117,9	–4,78	–3,9	231,74
2371,4	11,31	12,18	22038,28
2372	0,03	0,90	0,67
2339	–1,40	–0,52	0,07
2166,6	–7,66	–6,78	2110,42
Σ	–10,54	0,00	29888,76
Средняя	–0,88	Момент четвертого порядка	2717,16
		Среднее квадратическое отклонение в четвертой степени	1082,43
		Коэффициент эксцесса	2,51

Центральный момент четвертого порядка равен 2717,16, среднее квадратическое отклонение в четвертой степени составляет 1082,43. Таким образом, расчет коэффициента эксцесса дает значение 2,51.

Если бы данные были нормально распределены (т.е. средневершинны), то коэффициент эксцесса, рассчитанный с помощью моментов, равнялся бы 3,0. Следовательно, пик у рассматриваемых данных дохода по индексу акций выражен меньше, чем у нормального распределения, и их распределение является плосковершинным. Если бы данные были островершинными, т.е. с более выраженным пиком в сравнении с нормальным распределением, коэффициент эксцесса на основе моментов был бы более трех.

ПОКАЗАТЕЛИ СВЯЗИ

Выше объяснялось, что дисперсия случайной переменной показывает, как наблюдаемые значения этой переменной распределены вокруг среднего значения. Здесь мы представим концепцию **ковариации**, которая **показывает, как две случайные переменные ведут себя по отношению одна к другой**. Затем мы перейдем к расчету **коэффициента корреляции**, который является наиболее удобным показателем степени линейной связи между двумя переменными.

Ковариация

В финансовой и других областях встречается много ситуаций, когда важно знать, как две переменные ведут себя по отношению друг к другу. Например, в управлении портфелем для определения риска портфеля активов необходимо знать, как цена ценной бумаги X ведет себя по отношению к ценной бумаге Y для всех пар активов. Другими словами, нам нужно определить ковариацию или корреляцию каждой пары активов.

Если цена ценной бумаги X обычно растет (падает) в то же время, когда растет (падает) цена ценной бумаги Y, ковариация будет положительной. Однако если обычно во время роста цены ценной бумаги X цена ценной бумаги Y падает, то ковариация будет отрицательной. Если же не существует определенной модели связи между движениями цен, т.е. цены двух ценных бумаг ведут себя независимо, ковариация равна нулю.

Формула для вычисления ковариации выглядит следующим образом:

$$\mathrm{cov}_{XY} = \sigma_{XY} = \frac{\sum (X - \bar{X}) \cdot (Y - \bar{Y})}{n-1}. \qquad (2.21)$$

Как видим, величина ковариации согласно расчету зависит от величин наблюдений X и Y. Таким образом, значительная ковариация может быть вызвана в большей степени высокими значениями наблюдений, чем близкой связью между переменными.

Первым шагом в исследовании связи двух переменных является построение точечной диаграммы, отображающей каждую пару наблюдений.

В табл. 2.12 приведены данные, относящиеся к доходам от двух активов, представленных индексами S&P 500 и FTSE 100.[1] Для простоты изложения удобно обозначить их X и Y соответственно. В табл. 2.14 показаны результаты расчета ковариации для этих двух активов.

Таблица 2.12

Доходы по S&P 500, X	Доходы по FTSE 100, Y	$X-\bar{X}$	$Y-\bar{Y}$	$(X-\bar{X})(Y-\bar{Y})$
–0,81	–5,04	–0,043	–4,160	0,179
–2,79	–5,80	–2,026	–4,926	9,980
2,73	6,76	3,495	7,635	26,686
0,79	4,72	1,556	5,594	8,703
–7,22	–3,23	–6,449	–2,347	15,137
1,19	–4,69	1,963	–3,808	–7,475
1,78	–0,75	2,544	0,125	0,319
–1,92	–4,78	–1,154	–3,902	4,503
8,90	11,31	9,664	12,184	117,751
–1,00	0,03	–0,233	0,904	–2,211
–1,20	–1,40	–0,356	–0,522	0,186
–9,73	–7,66	–8,961	–6,778	60,733
Средняя	Средняя		Сумма	236,491
–0,77	–0,88		Ковариация	21,499
			Корреляция	0,793

[1] S&P 500 — индексы Standard and Poor's для 500 корпораций, в том числе 400 промышленных, 20 транспортных, 40 коммунальных и 40 финансовых.
FTSE 100 — фондовый индекс газеты "Financial Times" для 100 компаний.

Выражение (2.21) применено к данным следующим образом. Отклонения каждого из значений X или Y от соответствующих средних располагаются в столбцах $X - \bar{X}$ и $Y - \bar{Y}$. Соответствующие значения из этих столбцов перемножаются и помещаются в крайнем правом столбце. Значения в нем суммируются, в результате чего получаем 236,491. Это число делится на $n-1$ и получается ковариация, равная 21,499. Из рассматриваемого примера мы видим, что активы X и Y имеют положительную ковариацию доходов по ним.

Было бы полезно определить, как образуется знак ковариации. Обратите внимание на то, как плоскость рис. 2.7 разделена на четыре части координатной сеткой с осями, проходящими через $\bar{X}$ и $\bar{Y}$.

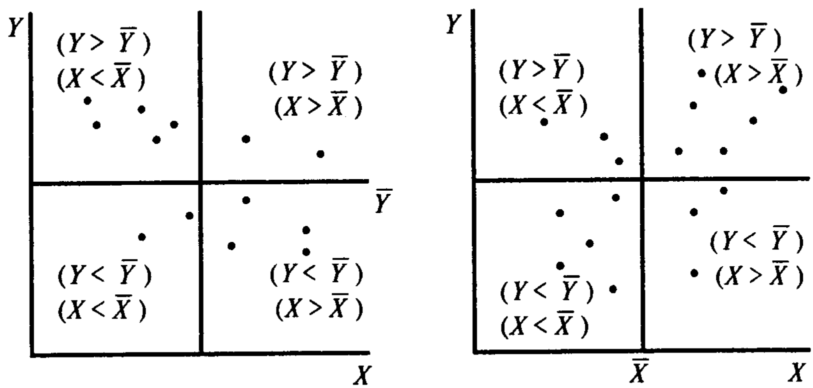

Рис. 2.7

Заметим, что для точек в верхней левой четверти значения Y больше значений $\bar{Y}$, таким образом, разность $Y - \bar{Y}$ положительна. Однако все значения X будут ниже $\bar{X}$. Отсюда отклонения $X - \bar{X}$ отрицательны. Так как отрицательное число при умножении на положительное дает отрицательную величину, то вклад наблюдений из верхней левой четверти будет отрицательным. Аналогично вклад наблюдений из верхней правой и нижней левой частей положителен. Вклад наблюдений из нижней правой части — отрицательный.

Таким образом, данные, которые преобладают в нижней левой и верхней правой четвертях, будут иметь положительную ковариацию, в то время как данные, преобладающие в верхней левой и нижней правой четвертях, имеют отрицательную ковариацию. Если точки данных равномерно рассеяны по всем четвертям, то ковариация равняется нулю.

Дисперсионно-ковариационная матрица

Часто ковариации нескольких переменных изображаются в виде дисперсионно-ковариационной матрицы. В табл. 2.13 показаны ковариации всех возможных пар из группы трех активов.

Таблица 2.13

	(*а*)				(*б*)		
	X	*Y*	*Z*		*X*	*Y*	*Z*
X	***xx***	*xy*	*xz*	*X*	***xx***		
Y	*yx*	***yy***	*yz*	*Y*	*yx*	***yy***	
Z	*zx*	*zy*	***zz***	*Z*	*zx*	*zy*	***zz***

Возможными парами активов в этом примере являются *XX*, *XY*, *XZ*, *YX*, *YY*, *YZ*, *ZX*, *ZY* и *ZZ*. В приведенной выше таблице ковариации записаны строчными буквами. Таким образом, ковариация *X* и *Y* — это *xy*. Заметьте, что это то же самое, что и *yx*. Поэтому все ковариации левее и ниже диагонали повторяются выше и правее диагонали. Значит, эти ковариации введены в матрицу дважды. Важность этой детали проиллюстрируем позднее, когда будем измерять риск портфеля.

Обратите внимание на то, что ковариации на диагонали показаны жирным шрифтом. Это ковариации доходов с ними самими. Ковариация переменной с ней самой равняется дисперсии этой переменной. Для понимания этого используем выражение (2.21) для измерения ковариации актива X с ним самим. Равенство будет выглядеть так:

$$\text{cov}_{XX} = \frac{\sum (X - \bar{X}) \cdot (X - \bar{X})}{n - 1}. \qquad (2.22)$$

Числитель в выражении (2.22) упрощается до $(X - \bar{X})^2$, отсюда оно приобретает следующий вид:

$$\text{cov}_{XX} = \frac{\sum (X - \bar{X})^2}{n - 1} = \sigma_X^2, \qquad (2.23)$$

что совпадает с формулой дисперсии. Таким образом, матрицу ковариаций часто называют **дисперсионно-ковариационной матрицей**.

Коэффициент корреляции

Наиболее широко используемым показателем степени связи между двумя переменными является **коэффициент корреляции**. Это показатель, независимый от единиц измерения, характеризует силу и направление линейной связи между двумя переменными. Следовательно, при его использовании преодолевается недостаток ковариации, так как величина коэффициента корреляции не находится под влиянием значений наблюдений.

Значения коэффициента корреляции находятся в промежутке от —1 (для случая отрицательной связи), проходя через 0, где две переменные независимы друг от друга и заканчиваются в точке + 1 (для случая положительной связи между переменными).

Коэффициент корреляции ρ рассчитывается делением ковариации X и Y на произведение среднего квадратического отклонения X и среднего квадратического отклонения Y. В формализованном виде записывается так:

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}. \tag{2.24}$$

Применив выражение (2.24) к данным из табл. 2.12, получим, что коэффициент корреляции X и Y равен

$$\rho_{XY} = \frac{21{,}499}{4{,}73 \cdot 5{,}74} = 0{,}793.$$

Независимо от того, положительна или отрицательна корреляция, коэффициент корреляции — это только мера статистической связи. В статистике не существует выводов о причинности, т.е. невозможно предположить, что изменение переменной X приведет к изменению переменной Y или наоборот.

Следовательно, коэффициент корреляции может показать, насколько сильна линейная связь между двумя переменными, но не может объяснить их изменений. Для такого объяснения необходимо сначала разработать теорию причинной связи из предварительных рассуждений, построить модель, которая отражает гипотетическую связь, и затем протестировать модель статистически. Для этого используется регрессионный анализ. Эта тема рассматривается в гл. 6.

Корреляционная матрица

Так же как ковариация, коэффициенты корреляции представляются в виде матриц. Рассмотрим для примера табл. 2.14.

Таблица 2.14

	(а)				(б)		
	X	*Y*	*Z*		*X*	*Y*	*Z*
X	1	*xy*	*xz*	*X*	1		
Y	*yx*	1	*yz*	*Y*	*yx*	1	
Z	*zx*	*zv*	1	*Z*	*zx*	*zv*	1

Заметьте, что табл. 2.14 (а) и (б) отличаются от табл. 2.13 (а) и (б) тем, что их диагонали, проходящие из верхнего левого угла до нижнего правого, должны по определению быть равны единице. Чтобы доказать это, вспомним, что числитель в выражении (2.24) представляет собой ковариацию переменной, но согласно выражению (2.23) в этих условиях становится дисперсией. Также обратите внимание на то, что знаменатель — это произведение средних квадратических отклонений двух переменных. В случае корреляции переменной с самой собой знаменатель был бы квадратом среднего квадратического отклонения, т.е. дисперсией. Следовательно, числитель и знаменатель являлись бы одной и той же дисперсией, а результат должен равняться единице.

Приложения ковариации и корреляции

Ковариация и корреляция имеют много приложений в финансовой теории и практике. Рассмотрим здесь, какую роль они играют в измерении риска портфеля активов, обладающих риском. В гл. 6 мы будем исследовать роль корреляции, используемой вместе с регрессионным анализом для хеджирования (страхования) рисков.

Риск портфеля. Рискованность одного актива измеряется дисперсией или средним квадратическим отклонением доходов по этому активу, а риск портфеля — дисперсией или средним квадратическим отклонением доходов портфеля.

Однако чтобы измерить риск портфеля, нам нужно не только знать вариацию доходов отдельных ценных бумаг, но и степень, с которой доходы пар ценных бумаг колеблются вместе. Нам необходимо знать **ковариацию** или же **корреляцию** доходов каждой пары активов в портфеле.

Риск портфеля, измеряемый через дисперсию, рассчитывается как взвешенная сумма ковариаций всех пар активов в портфеле, где каждая ковариация взвешена на произведение весов каждой пары соответствующих активов и дисперсия данного актива рассматривается как ковариация актива с самим собой.

Для демонстрации этого рассмотрим портфель из трех активов A, B и C. Доходы по каждому из них обозначим соответственно a, b и c. Веса, с которыми каждый актив представлен в портфеле, равны W_A, W_B и W_C.

Ковариации доходов по всем возможным парам активов можно отобразить в ковариационной матрице:

	W_A	W_B	W_C
W_A	**cov(*a*,*a*)**	cov(*a*,*b*)	cov(*a*,*c*)
W_B	cov(*b*,*a*)	**cov(*b*,*b*)**	cov(*b*,*c*)
W_C	cov(*c*,*a*)	cov(*c*,*b*)	**cov(*c*,*c*)**

Риск портфеля σ_p^2 можно найти так:

$$\sigma_p^2 = W_A W_A \operatorname{cov} V_{aa} + W_A W_B \operatorname{cov} V_{ab} + \ldots + W_C W_B \operatorname{cov} V_{cb} + W_C W_C \operatorname{cov} V_{cc}. \tag{2.25}$$

Рассмотрев состав ковариационной матрицы, можно заметить, что каждая взвешенная ковариация на самом деле включена в расчет дважды, например, $W_A W_C \operatorname{cov}(a, c)$ — это то же самое, что и $W_C W_A \operatorname{cov}(c, a)$. Вспоминая также, что $\operatorname{cov}(a, a)$ в действительности является дисперсией a, упростим выражение (2.25) и получим:

$$\begin{aligned}\sigma_p^2 &= W_A^2\sigma_a^2 + W_B^2\sigma_b^2 + W_C^2\sigma_b^2 + 2\cdot W_A W_B \operatorname{cov}_{ab} + \\ &+ 2\cdot W_A W_C \operatorname{cov}_{ac} + 2\cdot W_B W_C \operatorname{cov}_{bc}.\end{aligned} \tag{2.26}$$

Выражение (2.26) можно представить в общем виде:

$$\sigma_p^2 = \sum_{i=1}^{N} W_i^2\sigma_i^2 + 2\sum_{i=1}^{N}\sum_{j>i} W_i W_j \operatorname{cov}_{ij}. \tag{2.27}$$

Ясно, что преимущества диверсификации происходят от включения в портфель активов, которые имеют низкие или даже отрицательные ковариации с другими активами портфеля, что снижает сумму ковариаций и, следовательно, общий риск портфеля.

Так как ковариация —показатель связи, неограниченный по величине, то часто в качестве показателя связи используют **коэффициент корреляции**. Преимущество ранжирования пар активов по их коэффициентам корреляции заключается в предоставлении четкой системы включения активов, которые увеличивают преимущества диверсификации, и исключения тех, которые этого не делают.

Вспомним, что коэффициент корреляции рассчитывается как

$$\rho_{ab} = \frac{\sigma_{ab}}{\sigma_a \sigma_b}. \tag{2.28}$$

Следовательно, ковариацию можно выразить как

$$\mathrm{cov}_{ab} = \rho_{ab} \cdot \sigma_a \sigma_b. \tag{2.29}$$

Отсюда дисперсию портфеля двух активов, используя коэффициенты корреляции и дисперсии активов, можно найти следующим способом:

$$\sigma_p^2 = W_A^2 \sigma_a^2 + W_B^2 \sigma_b^2 + 2W_A W_B (\rho_{ab} \cdot \sigma_a \sigma_b), \tag{2.30}$$

где σ_p^2 — дисперсия портфеля;

σ_a^2 и σ_b^2 — дисперсии доходов соответственно по активам a и b;

ρ_{ab} — корреляция доходов по a и b;

σ_a и σ_b — средние квадратические отклонения соответственно по активам a и b;

$(\rho_{ab}\sigma_a\sigma_b)$ — **ковариация** доходов по активам a и b.

Среднее квадратическое отклонение портфеля — это квадратный корень из дисперсии портфеля.

Схожесть с выражением (2.26) понятна, если учесть, что $(\rho_{ab}\sigma_a\sigma_b)$ — это ковариация доходов по активам a и b.

Понижающие риск эффекты диверсификации

Чтобы продемонстрировать понижающие риск эффекты диверсификации, допустим, что ценная бумага A имеет среднее квадратическое отклонение 15%, а ценная бумага *B* имеет среднее квадратическое отклонение 14%. Допустим также, что веса активов в портфеле одинаковы.

Рассмотрим сначала особый случай, когда доходы по активам полностью коррелированны, т.е. коэффициент корреляции равен + 1,0.

Среднее квадратическое отклонение портфеля будет следующим:

$$\sigma_p = \sqrt{(0{,}5)^2(0{,}14)^2 + (0{,}5)^2(0{,}15)^2 + 2(0{,}5 \cdot 0{,}5 \cdot 1 \cdot 0{,}14 \cdot 0{,}15)} =$$

$$= 0{,}145\ (14{,}5\%).$$

Это показывает, что в данном конкретном случае риск портфеля — это просто средний взвешенный риск отдельных активов.

Рассмотрим, что случится, когда ценные бумаги будут иметь корреляцию, равную только 0,6. В этом случае среднее квадратическое отклонение портфеля становится равным 12,97%:

$$\sigma_p = \sqrt{(0{,}5)^2(0{,}14)^2 + (0{,}5)^2(0{,}15)^2 + 2(0{,}5 \cdot 0{,}5 \cdot 0{,}6 \cdot 0{,}14 \cdot 0{,}15)} =$$

$$= 0{,}1297\ (12{,}97\%).$$

Заметьте, что число 12,97% в действительности меньше, чем средние квадратические отклонения отдельных ценных бумаг.

Следующий шаг — исследование другого особого случая, когда два актива полностью отрицательно коррелированны. Среднее квадратическое отклонение портфеля будет равно:

$$\sigma_p = \sqrt{(0{,}5)^2(0{,}14)^2 + (0{,}5)^2(0{,}15)^2 + 2(0{,}5 \cdot 0{,}5 \cdot (-1) \cdot 0{,}14 \cdot 0{,}15)} =$$

$$= 0{,}005\ (0{,}5\%).$$

Риск портфеля практически сведен к нулю, потому что в случае когда один актив растет, другой снижается на подобную величину, и, следовательно, стоимость портфеля не колеблется. Это служит базой для операций хеджирования, в которых отрицательная корреляция достигается продажей позиции по инструменту, который имеет высокую степень положительной корреляции с тем активом, который нуждается в хеджировании. Таким образом, короткая позиция по хеджирующему инструменту создает отрицательную корреляцию между длинной и короткой позициями.

Так как доходы по ценным бумагам в общем не полностью коррелированны, среднее квадратическое отклонение портфеля будет меньше, чем средняя взвешенная средних квадратических отклонений отдельных ценных бумаг. Более того, среднее квадратическое отклонение портфеля падает, когда снижается степень корреляции пар активов.

Следовательно, эффективная диверсификация — это не просто добавление активов к портфелю, но добавление таких активов, доходы которых имеют самые низкие корреляции с активами, присутствующими в портфеле.

Ранее мы видели эффект степени корреляции на диверсификацию. Полезно сейчас рассмотреть, что происходит в портфеле, состоящем из многих активов.

Рассмотрим снова выражение (2.27), которое повторено здесь как (2.31)

$$\sigma_p^2 = \sum_{i=1}^{N} W_i^2 \sigma_i^2 + 2\sum_{i=1}^{N} \sum_{j>i} W_i W_j \operatorname{cov}_{ij}. \tag{2.31}$$

Представим, что имеется очень большое количество активов, доступных для инвестиций, скажем индекс из 100 или 500 акций. Допустим также, что все доходы по активам независимы. Выражение (2.31) сократится до следующего:

$$\sigma_p^2 = \sum_{i=1}^{N} W_i^2 \sigma_i^2 . \tag{2.32}$$

Так как предполагается, что доходы по активам независимы, ковариации равняются нулю. Теперь предположим, что равные суммы инвестированы в каждый из N активов, тогда веса каждого станут равными $1/N$, и дисперсия портфеля примет вид:

$$\sigma_p^2 = \sum_{i=1}^{N} \left(\frac{1}{N}\right)^2 \sigma_i^2 = \frac{1}{N}\left[\sum_{i=1}^{N} \frac{\sigma_i^2}{N}\right]. \tag{2.33}$$

Выражение в прямоугольных скобках является средней дисперсией активов в портфеле. В то время как число активов (N) в портфеле становится больше, $1/N$ уменьшается, и дисперсия портфеля снижается, приближаясь в пределе к нулю.

Однако в действительности не все доходы по активам независимы, особенно когда мы рассматриваем активы, принадлежащие к одному классу, например, акции и облигации. У большинства активов будет присутствовать некоторый уровень ковариации. Отсюда на практике равенство (2.32) превращается в следующее:

$$\sigma_p^2 = \sum_{i=1}^{N}\left(\frac{1}{N}\right)^2 \sigma_i^2 + 2\sum_{i=1}^{N}\sum_{j>1}\left(\frac{1}{N}\right)\left(\frac{1}{N}\right)\text{cov}_{ij}\ . \tag{2.34}$$

Это можно представить так:

$$\sigma_p^2 = \left(\frac{1}{N}\right)\sum_{i=1}^{N}\left[\frac{\sigma_i^2}{N}\right] + \frac{(N-1)}{N}2\sum_{i=1}^{N}\sum_{j>1}\left[\frac{\text{cov}_{ij}}{N(N-1)}\right]. \tag{2.35}$$

Первый член равенства представляет собой среднюю дисперсию, уже встречавшуюся выше в выражении (2.33), а второй — это тоже средняя, т.е. сумма ковариаций, деленная на число ковариаций $N(N-1)$. Выражение (2.35), таким образом, может быть упрощено до

$$\sigma_p^2 = \frac{1}{N}\overline{\sigma}_i^2 + \frac{N-1}{N}\overline{\text{cov}}_{ij}\ . \tag{2.36}$$

Эта формула помогает объяснить, что происходит с риском портфеля, когда в него включено большое количество активов. Когда число активов в портфеле увеличивается, $1/N$ уменьшается, и, таким образом, его произведение на среднюю дисперсию приближается к нулю. Однако $(N-1)/N$ стремится к единице при увеличении N, отсюда второе слагаемое правой части выражения (2.36) приближается к средней ковариации. Следовательно, когда портфель диверсифицирован включением большого числа активов, дисперсия портфеля приближается к средней ковариации отдельных активов.

Значит, общий риск ценной бумаги, находящейся в изоляции, больше, чем у той же ценной бумаги, находящейся в портфеле. Комбинация активов со слабой корреляцией понижает риск портфеля. Таким образом, общий риск состоит из двух частей: а) тот риск, который может быть исключен диверсификацией (**несистематический риск, также известный как случайный или остаточный риск**) и б) тот элемент риска, который не может быть исключен с помощью диверсификации (**систематический риск, также известный как рыночный риск**).

Индексы

Индекс — это единичный описательный статистический показатель, который обобщает относительное изменение одной переменной или группы переменных. Например, публикуемые индексы обобщают в одной величине изменение цен большого количества розничных товаров (**индекс розничных цен**), цен производителей (**индекс цен производителей**), цен акций (**индекс FTSE 100**), изменение стоимости валюты по отношению к набору других валют (**взвешенный индекс торговли**).

Индексы полезны, когда значения, лежащие в основе расчета, очень велики и абсолютные изменения трудно понять, например, при анализе валового национального продукта страны. С другой стороны, индексы применяются для обобщения совокупного изменения в группе составляющих, каждая из которых меняется в разной степени. Например, индекс розничных цен обобщает изменение цен набора (корзины) товаров и услуг, приобретаемых "типичным" домашним хозяйством. Индекс FTSE 100 обобщает изменение цен 100 акций, зарегистрированных на Лондонской фондовой бирже.

Индексы как относительные индикаторы

Индекс цен может быть истолкован как **отношение цен**. Обозначим цену единицы товара сегодня P_1, которая сравнивается со вчерашней ценой P_0. Относительная цена товара, или отношение цен, находится как P_1/P_0.

Если цена выросла, то отношение цен будет больше единицы, а если цена упала, то меньше единицы. Например, если цена в день, который обозначим "0", составляла 12 единиц, а в день, обозначенный как 1, изменилась до 14, то отношение цен равняется 14/12 = 1,1667. Если бы цена снизилась до 11, отношение цен равнялось бы 11/12 = 0,9167.

Ясно, как и в приведенном выше случае, что может быть найдено дальнейшее отклонение цен для характеристики изменения цены между днями 2 и 1, т.е. P_2/P_1. Однако особенностью индексов является их ссылка на **базовый период**. Так, в нашем примере базовым индексом будет отношение P_2/P_0. Таким образом, индекс может быть использован для выражения изменения рассматриваемой переменной(ых) между базовым и текущим периодами.

До сих пор мы рассматривали только отношения цен, но также обоснованно может быть поставлен вопрос о расчете отношения физических объемов или отношения стоимостей, если рассматриваемые данные являются физическими объемами и стоимостями, а не ценами.

Вторая значительная особенность индексов заключается в том, что базовая цена (физический объем или стоимость) приводится к базису 100 или, иногда, 1 000. Для этого надо лишь умножить отношения цен на 100 (или 1 000). Следовательно, если $P_0 = 12$, при приведении к базису 100 мы делим на 12 и умножаем на 100. Если P_1 равно 14, то индекс с базисом 100 составит

$$14/12 \cdot 100 = 116{,}67.$$

Это можно обобщить следующим образом:

$$Z\left(\frac{P_t}{P_0}\right), \tag{2.37}$$

где Z — базовое значение индекса, обычно 100 или 1 000;

P_0 — величина переменной в базовом периоде;

P_t — текущая величина переменной.

Преимущество базового значения индекса, равного 100, в очень доступном отображении процентного изменения индекса со стартовой даты или базового периода. Однако это процентное изменение не отражает изменений между отдельными датами. Например, если цена третьего дня P_3 была равна 15 единицам, индекс цен, отражающий изменение P_3 по сравнению с P_0, составит

$$15/12 \cdot 100 = 125\%.$$

Это показывает 25%-ный прирост со дня старта. Но изменение по сравнению со вторым днем (P_3/P_2) будет только 7,14%:

$$(125 - 116{,}67)/116{,}67 = 7{,}14\%.$$

В приведенных выше примерах мы наблюдали отношение цен только по одной позиции. В то же время на практике индексы составляются для обобщения изменения по совокупности рассматриваемых переменных. Это ставит перед нами две важные проблемы:

1. Каким способом усреднять множество изменений, чтобы прийти к одному индексу?
2. Как относиться к каждой переменной в составе группы? Нужно ли к ним относиться как к одинаково важным? Если нет, то как определить важность каждой из них?

Выбор способа усреднения

Для обобщения изменений большого числа переменных необходимо относительные величины обобщить и привести к средней. Существует два типа нахождения средней, которые применимы к индексам: нахождение средней арифметической и средней геометрической.

Чтобы проиллюстрировать расчет средней арифметической, рассмотрим индекс по совокупности, состоящей из четырех активов, — A, B, С и D. Отношения цен выглядят так:

$$\frac{P_{A_1}}{P_{A_0}}, \frac{P_{B_1}}{P_{B_0}}, \frac{P_{C_1}}{P_{C_0}}, \frac{P_{D_1}}{P_{D_0}}.$$

Предположим, что все составляющие одинаково важны, т.е. имеют равные веса, тогда индекс составит

$$Z\left(\frac{\dfrac{P_{A_1}}{P_{A_0}}+\dfrac{P_{B_1}}{P_{B_0}}+\dfrac{P_{C_1}}{P_{C_0}}+\dfrac{P_{D_1}}{P_{D_0}}}{4}\right). \tag{2.38}$$

Это можно представить в общем виде:

$$Z\left(\frac{\sum\limits_{i=1}^{n}\dfrac{P_{it}}{P_{i0}}}{n}\right). \tag{2.39}$$

Если бы те же отношения цен были включены в индекс, усредненный с помощью геометрической средней, он был бы построен следующим образом:

$$Z\left(\frac{P_{A_1}}{P_{A_0}}\cdot\frac{P_{B_1}}{P_{B_0}}\cdot\frac{P_{C_1}}{P_{C_0}}\cdot\frac{P_{D_1}}{P_{D_0}}\right)^{\frac{1}{4}}. \tag{2.40}$$

В общем виде это выглядит так:

$$Z\prod_{i=1}^{n}\left(\frac{P_{it}}{P_{io}}\right)^{\frac{1}{n}}. \tag{2.41}$$

Средние геометрические индексы имеют присущее им смещение, связанное с тем, что за исключением случаев, когда все составляющие изменяются в одинаковой пропорции, индекс будет недооценивать абсолютную величину роста и переоценивать абсолютную величину падения. Эта переоценка имеет отношение к портфелю акций, составляющих индекс. Более того, если одна из составляющих упадет до нуля, то общий индекс может иметь только нулевое значение.

Смещение в средних геометрических индексах проиллюстрировано в табл. 2.15, где сравниваются изменения средних арифметических и средних геометрических индексов.

Таблица 2.15. Иллюстрация смещения средних геометрических индексов

A	*B*	*C*	*D*	Средняя арифметическая	Средняя геометрическая
100	100	100	100	100	100
103	101	105	106	103,75	103,73
110	106	106	108	107,5	107,49
115	104	108	112	109,75	109,67
120	113	112	123	117	116,91
110	108	107	115	110	109,96
103	105	104	102	103,5	103,49
88	101	101	90	95	94,81
71	95	96	85	86,75	86,13
0	104	108	80	73	0,00

Другой проблемой геометрически усредненных индексов является то, что их поведение не повторяет поведения рассматриваемых составляющих совокупности. Например, если изменения цены акций компании отражены индексом, то доходы, предполагаемые по индексу, не будут совпадать с фактически достигнутыми по портфелю этих акций, построенному с такими же весами, как в индексе. Однако в случае среднего арифметического индекса доходы соответствовали бы подобному рассматриваемому портфелю.

Выбор способа взвешивания

Чтобы индекс отображал общее движение рассматриваемых составляющих, каждое из отношений цен должно быть взвешено для отражения их относительной важности. Существуют четыре известных способа взвешивания составляющих:

- с использованием равных весов;
- по цене или физическому объему базового периода;
- по цене или физическому объему текущего периода;
- по их текущей стоимости, т.е. произведению цена × физический объем.

При расчете индексов с равными весами ко всем элементам группы относятся как к одинаково важным. При этом существуют некоторые недостатки, когда группа, представленная индексом, содержит элементы со значительно разными уровнями важности. Этот подход имеет тенденцию к переоценке важности относительно неважных элементов по сравнению с важными. Например, если мы взвесим с одинаковыми весами изменение цен акций небольших (по рыночной капитализации) компаний, то мы будем переоценивать изменения цен акций небольших компаний, либо недооценивать изменения цен акций больших компаний относительно важности этих компаний для среднего инвестора.

Индексы с равными весами могут быть построены как средняя арифметическая или средняя геометрическая относительных цен способами, описанными в приведенных выше выражениях (2.38) — (2.41).

Если решено использовать систему взвешивания не с "равными весами", то нам необходимо определить принципы нахождения весов. Должны ли веса отражать относительную важность каждой из переменных в базисном периоде? Если да, то это будет **индекс Ласпейреса** или **индекс, в котором используются веса базисного периода**. Альтернативно, должны ли веса отражать текущую важность каждой из переменных? Подобным индексом является **индекс Пааше**. Какой бы вариант ни был выбран, его необходимо использовать постоянно, иначе сравнение значений индекса в разные моменты времени будет бессмысленным.

Индекс Ласпейреса, или индекс, при расчете которого используются веса базисного периода

В индексах Ласпейреса весами могут быть как цена, так и физический объем продукции. Общей чертой является то, что веса определяются в базисном периоде и используются без изменений для расчетов, относящихся к последующим периодам времени.

Индекс цен Ласпейреса измеряет текущую стоимость группы элементов, взвешенную по их количеству в базисном периоде, отнесенную к стоимости того же набора в базисном периоде. $P_n \cdot Q_0$ — это текущая стоимость группы товара, а $P_0 \cdot Q_0$ — это стоимость набора товаров в базисном периоде. Индекс рассчитывается как отношение:

$$P_p = \frac{\sum P_n \cdot Q_0}{\sum P_0 \cdot Q_0} \cdot 100. \quad (2.42)$$

Индекс физического объема Ласпейреса определяется отношением объема текущего выпуска по группе элементов, взвешенных по ценам базисного периода, к стоимости этой группы элементов в базисном периоде. Индекс находится как отношение:

$$I_Q = \frac{\sum P_0 \cdot Q_n}{\sum P_0 \cdot Q_0} \cdot 100. \quad (2.43)$$

Проблема взвешивания по базисным значениям состоит в том, что оно не позволяет учитывать замещение элементов, происходящее из-за изменения вкусов, технологии или относительных изменений цен. Например, в течение последних 15 лет мы стали свидетелями огромного изменения относительных цен говядины и куриного мяса, при котором люди увеличили потребление более дешевого куриного мяса по сравнению с более дорогой говядиной. На данный сдвиг возможно также повлияла перемена отношения к здоровью. Индекс, взвешенный по базисным значениям, не может принять в расчет подобные изменения, по крайней мере не может сделать это быстро.

Индекс Пааше, или индекс, взвешенный по текущим соизмерителям

Индексы Пааше также могут быть индексами цен или физического объема. Индекс цен Пааше измеряет стоимость группы элементов, взвешенную по их ценам в текущем периоде, относительно стоимости той же группы по ценам базисного периода. Текущая стоимость составляет $P_n \cdot Q_n$, а стоимость по ценам базисного периода равна $P_0 \cdot Q_n$. Индекс рассчитывается как отношение

$$I_P = \frac{\sum P_n \cdot Q_n}{\sum P_0 \cdot Q_n} \cdot 100. \tag{2.44}$$

Индекс физического объема Пааше измеряет отношение стоимости объема текущего выпуска, взвешенного по ценам текущего периода, к объему выпуска в базисном периоде, выраженного в ценах текущего периода. Индекс составляется как отношение:

$$I_Q = \frac{\sum P_n \cdot Q_n}{\sum P_n \cdot Q_0} \cdot 100. \tag{2.45}$$

Проблемой взвешивания индекса по текущим значениям весов (соизмерителей) является то, что группа сравниваемых элементов не остается одной и той же в разные периоды времени. Когда вкусы или технология ведут к быстрым изменениям, сравнения могут стать бессмысленными.

Индексы, взвешенные по капитализации

Некоторые индексы ценных бумаг, такие, как большинство индексов фондового рынка, используют взвешивание составляющих по их текущей рыночной капитализации. Она рассчитывается как произведение **текущей цены** на **текущее количество** выпущенных ценных бумаг.

Веса изменяются вместе с изменением относительной стоимости выпуска ценных бумаг. Это может наблюдаться из-за изменения цены ценных бумаг, объема выпуска или того и другого одновременно. Таким образом, во время роста курса акции относительно других важность (весомость) ее в индексе будет увеличиваться, так как ее вес (цена · количество) растет относительно других составляющих. Напротив, если цена акции падает относительно других, весомость этой акции в индексе уменьшится. Добавим также, что если компания выпускает больше акций по отношению к другим участникам фондового рынка, тогда при условии, что цена этих ценных бумаг не упадет, $P \cdot Q$ вырастет, и вес данных ценных бумаг в индексе увеличится.

Следовательно, взвешенные по капитализации индексы имеют смещение в пользу успешно работающих компаний, потому что эти компании имеют растущие цены на свои акции и/или способны выпускать дополнительные акции для инвесторов, отсюда их вес увеличивается. Акции менее успешно работающих компаний, с другой стороны, падают в цене и в зависимости от правил включения составляющих в индекс в конце концов выходят из списка компаний, включаемых в индекс, уступая место более успешным компаниям, которые поначалу не были составляющими индекса.

Примером того, как взвешивание по рыночной капитализации влияет не только на веса в разных периодах, но и на состав индекса, служит поглощение Мидланд Бэнк (Midland Bank) Банком Гонконга и Шанхая (Hong Kong and Shanghai Bank). Как выше уже сказано, индекс FTSE 100 составлен из 100 крупнейших компаний (по рыночной капитализации), зарегистрированных на Лондонской фондовой бирже. По отдельности ни один из банков не был достаточно велик для включения в индекс. Однако объединенная рыночная капитализация двух банков стала достаточно большой для включения нового банка в индекс. При включении укрупненного Банка Гонконга и Шанхая наименьшая из предыдущего списка 100 крупнейших компаний должна быть исключена из списка компаний, составляющих индекс FTSE. Сверх того, цена акций Банка Гонконга и Шанхая выросла относительно остальных составляющих, еще больше увеличив его весомость в индексе.

Расчет индексов акций

Существует большое количество индексов акций, рассчитываемых во всем мире, так как имеется очень много отдельных фондовых рынков, а для многих рынков составляется более одного индекса, каждый из которых отражает различный сектор всего рынка.

Большинство индексов рынков акций являются взвешенными по рыночной капитализации, хотя некоторые представляют собой средние арифметические отношений цен, а другие — средние геометрические отношений цен с равными весами. Проиллюстрируем ниже каждую из этих форм.

Простой агрегат цен. Наиболее простая форма индекса цен — просто сумма цен составляющих. Так как для элементов с высоким уровнем цен выше вероятность больших абсолютных изменений цен, чем для элементов с низким уровнем цен, то колебания цен дорогих элементов будут иметь большее влияние на индекс, чем колебания цен дешевых элементов. Этот недостаток усиливается из-за того, что данный способ не учитывает размер рассматриваемых единиц. Например, простой агрегатный индекс цен акций, взвешенный по цене, будет равно чувствителен как к 20-пенсовому изменению цены акции, оцененной в 10 фунтов стерлингов, так и к такому же изменению цены акции, стоящей 0,25 фунта.

Иногда агрегатные индексы цен отражают среднюю цену ценных бумаг в индексе, тогда изменения индекса представляют среднее ценовое изменение для ценных бумаг в индексе и рассчитываются как средняя арифметическая:

$$\frac{\sum P_i}{n}. \tag{2.46}$$

Промышленный индекс Доу-Джонса (Dow Jones Industrial Average), индекс Фондовой биржи США (MMI) и Никкей 225 (Nikkei 225 Stock Average) имеют агрегатную форму индекса средней цены.

Агрегатные индексы цен дают равные веса всем составляющим, таким образом присваивая больший относительный вес менее важным составляющим. Однако в случае индексов Доу-Джонса и MMI это влияние пренебрежимо мало, потому что в них включено лишь небольшое число акций и все они относительно дороги.

Взвешенная с равными весами средняя геометрическая отношений цен

Некоторые индексы построены как средние геометрические относительных цен, значимость каждого из отношений цен одинакова. Данная техника преодолевает проблему придания равных весов каждому из изменений цен независимо от их относительного размера, но не решает проблему учета относительных размеров составляющих компаний.

Индекс обычных акций Файненшл Таймс (Financial Times Ordinary Index) — это средняя геометрическая отношений цен 30 акций с равными весами.

Взвешенная по рыночной капитализации средняя арифметическая отношений цен

Данный индекс — это **средняя взвешенная цен**, потому что цены отдельных акций взвешены для отражения их относительной важности внутри группы цен акций, которые составляют индекс. Взвешивание произведено по рыночной капитализации. Следовательно, относительная важность каждой акции и, таким образом, относительная важность изменения цены каждой акции при определении значения индекса находятся под влиянием относительного размера каждой включенной в индекс компании.

Чтобы проиллюстрировать эффекты взвешивания по капитализации, построим гипотетический индекс. Допустим, что нам необходимо составить индекс, включающий акции только четырех компаний. Компания A выпустила 1 000 акций, оцененных по 100 ед. каждая, компания B — 10 000 акций по 25 ед. каждая, компания C — 5 000 акций по 50 ед., и компания D — 8 000 акций по той же цене 50 ед.

Рыночная капитализация для каждой компании составит:

A	1 000 · 100	= 100 000
B	10 000 · 25	= 250 000
C	5 000 · 50	= 250 000
D	8 000 · 50	= 400 000

Общее значение рыночной капитализации равно 1 000 000. Это будет базовым значением индекса, хотя оно выражено более простым числом, скажем, 1 000. Достигается это делением на постоянный делитель, составляющий в нашем примере 1 000.

Чтобы показать, как движения цен входящих акций отражаются в изменениях индекса, рассмотрим следующие изменения цен акций только через один день:

A = 105 ед., B = 26 ед., C = 45 ед., D = 55 ед.

Сумма рыночных капитализаций составит 105 000 + 260 000 + 225 000 + 440 000 = 1 030 000 ед. Для нахождения значения индекса это число необходимо разделить на постоянный делитель, т.е. на 1 000. Новый индекс равен 1 030 000/1 000 = 1030, т.е. отражено, что общее значение индекса выросло с начала его вычисления. Это произошло вопреки тому факту, что одна из акций упала в цене. Причина заключается в том, что хотя падение цены акций компании C было таким же большим как рост для компаний A и D, компании A и D имели преобладающую общую рыночную капитализацию. Таким образом, относительная важность падения цены акций C была меньше, чем рост цен акций A и D.

Этот пример может быть обобщен для построения формулы индекса, взвешенного по рыночной капитализации, состоящего из n акций, следующим образом:

$$\text{Индекс}_t = \frac{\sum_{j=1}^{n} m_{j,t} P_{j,t}}{ОД}, \qquad (2.47)$$

где n — количество составляющих индекса;
$m_{j,t}$ — количество акций компании j в период t;
$P_{j,t}$ — цена одной из этих акций в период t;
$ОД$ — общий делитель.

Общий делитель — это базовое значение индекса, что является суммой всех $m_{j,0} \cdot P_{j,0}$, поделенной на начальное значение индекса, т.е. 100 или 1 000 по предпочтению.

Альтернативно, тот же результат может быть достигнут умножением числителя выражения (2.47) на начальное значение индекса и делением этого произведения на базовое значение индекса вместо общего делителя. Это выглядит как

$$\text{Индекс}_t = \frac{Z \sum_{j=1}^{n} m_{j,t} P_{j,t}}{\text{Базовый уровень}}, \qquad (2.48)$$

где Z — начальное значение индекса, т.е. 100 или 1 000 по предпочтению.

Эта формула аналогична используемой при вычислении многих индексов акций, например, индексов FTSE 100 и 250 в Лондоне, индексов S&P 100 и 500 в США и CAC 40 во Франции. Отличие составляет лишь количество компаний в каждом из индексов.

Особенность индексов акций — в периодической смене составляющих. В связи с этим базовое значение заслуживает отдельного обсуждения. Обычно в расчете индекса цен, взвешенного по базовому количеству, базовый объем и базовая цена зафиксированы. Однако в индексах акций это базовое значение необходимо периодически настраивать для учета изменений, происходящих из-за включения и исключения компаний из индекса и изменения количества акций, выпущенных компаниями-составляющими.

Например, выше уже объяснено, что индекс FTSE 100 состоит из обычных акций 100 крупнейших по рыночной капитализации компаний, зарегистрированных на Лондонской фондовой бирже. Так как относительные цены на акции со временем меняются, некоторые компании будут выпадать из классификации, а новые входить в нее. Базовое значение индекса должно быть подстроено для отражения этих изменений, иначе изменения в величине индекса будут отражать не только относительное изменение цен, но и изменения в составляющих индекса и их структуре капитала.

Следовательно, необходим механизм, который позволил бы заменять составляющие, но предоставлял бы при этом преемственность в серии индексов. Это достигается посредством процесса, известного как цепное связывание.

Процесс цепного связывания требует, чтобы базовое значение индекса пересчитывалось для восприятия изменений составляющих и/или структуры их капитала. Хотя могут произойти и комплексные изменения капитала во время включения и исключения составляющих индекса, чистое изменение агрегированной рыночной капитализации составляющих выглядит так:

$$ОД_t = ОД_{t-1}\left[1 + \frac{C_t}{\sum_{i=1}^{n} N_{i(t-1)} \cdot P_{i(t-1)}}\right], \tag{2.49}$$

где $ОД$ — общий делитель;

C_t — изменение капитализации;

N_i — число акций, выпущенных i-й компанией на момент времени t—1;

P_i — цена акций i-й компании в момент времени t—1.

Индексы цен и общая ставка дохода

Необходимо заметить, что большинство публикуемых индексов финансовых рынков являются индексами только цен. Движения такого индекса отражают лишь движения цен рассматриваемых инструментов, они не учитывают дивиденды и другие доходы, получаемые инвесторами, и, таким образом, не рассчитывают общей ставки дохода.

Исключение дивидендов или других выплат ведет к измерению дохода движением индекса, занижая фактические доходы от обладания портфелем, который повторяет этот индекс. Чтобы увидеть, как это происходит, представим, что в нашем приведенном ранее примере гипотетического индекса компания С выплатила дивиденд, равный 5. На следующий день после выплаты дивиденда цена акции падает на 5 (предполагаем отсутствие эффектов налогообложения) до 45. Индекс уменьшится до 975. Хотя в то же время инвестор получил дивиденд, равный 5, и общая стоимость портфеля не изменилась. Следовательно, при сравнении доходов индекса с доходами портфеля влияние дивидендов на индекс должно быть учтено.

Выбор базового периода

Какая бы методология построения индекса не использовалась, необходимо произвести выбор базового периода. Не существует строгих и быстрых правил, позволяющих определить "правильный" базовый период. Это может быть одна дата. Это может быть период с большей или меньшей длиной, а может и средняя на протяжении определенного промежутка времени. Важно то, что он не должен быть "необычны" периодом. Другими словами, он должен быть представительным относительно более ранних периодов. К сожалению, конечно, не существует гарантии, что он будет представительным для последующих периодов.

Далее, нужно, чтобы базовый период не находился так далеко в прошлом, что было бы трудно для современных наблюдателей ссылаться на условия в этом базовом периоде.

Выбор составляющих

Этот выбор зависит от предназначения индекса и опыта тех, кто его составляет. Иногда может потребоваться исследование для подбора наиболее подходящих составляющих. Например, для индекса розничных цен составляющие и веса определяются с помощью обследования расходов домашних хозяйств.

В сфере финансовых рынков выбор составляющих диктуется сектором, который надлежит отобразить с помощью индекса. Так, в индексе, отражающем основные составляющие фондового рынка, будут доминировать более крупные компании. Однако индекс, предназначенный для отображения акций с наименьшими капитализациями, по определению не будет содержать акции крупных компаний. Важно то, чтобы в этих обстоятельствах критерии для включения в индекс были ясно определены и широко известны. Так как компании — динамичные объекты, составляющие индексов акций меняются, некоторые компании покидают индекс, поскольку более не отвечают критериям включения, а другие присоединяются к индексу, потому что впервые стали соответствовать этим критериям.

Примеры индексов рынков ценных бумаг

FTSE 100. Это взвешенный по рыночной капитализации индекс, состоящий из ведущих 100 акций, зарегистрированных на Лондонской фондовой бирже. Базовое значение равно 1 000, а базовая дата — 31 декабря 1984. Он рассчитывается каждую минуту торгового дня. Так как составляющие классифицированы по рыночной капитализации, они меняются при изменении относительной капитализации. Включения и исключения из состава делаются раз в квартал. Этот индекс рассчитывается как для характеристики изменения только цен, так и для оценки изменения общей ставки дохода, но к индексу, характеризующему изменения цен акций, обращаются наиболее часто.

Индекс базируется на средней котировок покупка-продажа, а не на фактических ценах сделок.

FTSE 250. Этот индекс концептуально схож с индексом FTSE 100 за исключением того, что он базируется на акциях 250 крупнейших компаний Лондонской фондовой биржи, занимающих позиции сразу после верхних 100. Существует также индекс FTSE 350, который представляет собой комбинацию индексов FTSE 100 и FTSE 250.

Индексы Standard and Poor's 100 и 500. Это взвешенные по капитализации индексы, содержащие соответственно 100 и 500 акций, наиболее активно участвующих в торгах в США. S&P 500 составляют крупнейшие 456 акций на Нью-Йоркской фондовой бирже, 36 акций, котируемые в системе NASDAQ, и восемь, участвующих в торгах на Американской фондовой бирже.

Композитный индекс Нью-Йоркской фондовой биржи (New York Stock Exchange composite index). Это взвешенный по рыночной капитализации индекс, который включает все акции, котируемые на Нью-Йоркской фондовой бирже (около 1700). Он рассчитывается каждые 15 секунд.

DAX (Deutscher Aktienindex). Это арифметически взвешенный по капитализации индекс 30 германских компаний, зарегистрированных на Франкфуртской фондовой бирже. Он примечателен тем, что игнорирует настройки после выплаты дивидендов и смены собственности. Отсюда рыночная капитализация каждой составляющей (т.е. цена · количество акций) является не тем же самым, что, скажем, в индексе FTSE 100. DAX ведет себя более как индекс общей ставки дохода с реинвестируемыми дивидендами и дополнительными выплатами в форме акций.

Индекс CAC (Compagnie des Agents de Change) 40. Этот взвешенный по рыночной капитализации индекс 40 акций, выбранных из 100 крупнейших компаний, зарегистрированных на Парижской бирже. Он пересчитывается каждые 30 секунд.

УПРАЖНЕНИЯ

1. Используя следующие несгруппированные данные, относящиеся к 20 наблюдениям ежедневных доходов по индексу FTSE 100, подсчитайте:

   а) медиану ежедневных доходов;
   б) моду;
   в) среднюю арифметическую;
   г) среднюю геометрическую.

Несгруппированные данные: 20 наблюдений ежедневных доходов по индексу FTSE 100

–0,43	–0,13	–0,38	–0,50	–0,68	0,84	–0,05	–0,53	–0,04	–0,35
1,07	0,58	–0,75	0,26	–0,04	0,68	–0,51	0,71	–0,12	0,02

2. Предполагая, что приведенные выше доходы наблюдались непрерывно каждый день, подсчитайте ставку наращения на £1 000, ин-

вестированных на рассматриваемые 20 дней. Сравните ее со ставками наращения, полученными с помощью найденных ранее средней арифметической и средней геометрической.

3. Для сгруппированных данных, приведенных ниже, найдите медиану и среднюю арифметическую.

Сгруппированные данные

Доходы	Частота
До –0,05	1
Более –0.05 и до –0.04	2
Более –0.04 и до –0.03	5
Более –0.03 и до –0.02	24
Более –0.02 и до –0.01	191
Более –0.01 и до 0.00	714
Более 0.00 и до 0.01	816
Более 0.01 и до 0.02	235
Более 0.02 и до 0.03	25
Более 0.03 и до 0.04	4
Более 0.04 и до 0.05	1
Всего	2018

4. Для несгруппированных данных из п. 1 рассчитайте дисперсию, среднее квадратическое отклонение и квартильное отклонение.
5. Для сгруппированных данных из п. 3 рассчитайте дисперсию, среднее квадратическое отклонение и квартильное отклонение.
6. Объясните обстоятельства, когда полудисперсия и половина среднего квадратического отклонения могут быть использованы как показатели вариации.
7. Объясните, что имеется в виду под асимметрией. Найдите коэффициент асимметрии с помощью моментов для несгруппированных данных из п. 1 и прокомментируйте свой результат.
8. Объясните, что имеется в виду под эксцессом. Найдите коэффициент эксцесса с помощью моментов для несгруппированных данных из п. 1 и прокомментируйте свой результат.
9. Для следующих данных, относящихся к индексу FTSE 100 и индексу S&P 500, рассчитайте ковариацию и коэффициент корреляции.

FTSE 100	S&P 500
3491,8	481,61

3328,1	467,14
3086,4	445,77
3125,3	450,91
2970,5	456,50
2919,2	444,27
3082,6	458,25
3251,3	475,49
3026,3	462,69
3097,4	472,35
3081,4	453,69
3065,5	459,27
2991,6	470,42
3009,3	487,39
3137,9	500,71
3216,7	514,71

10. а) Для следующих данных постройте взвешенный по рыночной капитализации индекс со значением в базовом периоде 1 000.

Компания	Число акций в выпуске	Цена акции
A	1 000 000	2,50
B	5 000 000	1,75
C	10 000 000	0,80
D	8 000 000	1,60
E	7 500 000	3,00

б) Подсчитайте значение индекса при следующих новых ценах на акции.

A = 2,70, B = 1,30, C = 1,20, D = 1,40, E = 2,70.

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

Bowers, D. (1991) *Statistics for Economics and Business.* Macmillan, London.

Curwin, J. and Slater, R. (1993) *Quantitative Methods for Business Decisions*, 2nd edn. Chapman & Hall, London

Silver, M. (1992) *Business Statistics.* McGraw-Hill, London.

ОТВЕТЫ К ИЗБРАННЫМ ВОПРОСАМ

1. Медиана = —0,085
   Мода = —0,04
   Средняя арифметическая = —0,0175
   Средняя геометрическая = 0, 000189
3. Медиана = 0, 000888
   Средняя = 0, 000525
4. Дисперсия = 0,2905
   Среднее квадратическое отклонение = 0,5390
   Квартильное отклонение = 0,39375
5. Дисперсия = 0, 0000914
   Среднее квадратическое отклонение = 0,00956
7. Асимметрия = 0,560761
8. Эксцесс = 2,0645
9. Ковариация = 1092,14445
   Корреляция = 0,4092
10. 956,17

Приложение : выборочное среднее квадратическое отклонение — почему делитель равен n—1?

Рассмотрим случайную выборку размера n, произведенную из совокупности со средней μ и средним квадратическим отклонением σ, эти параметры генеральной совокупности неизвестны.

Выборочная средняя равна

$$\sum_{i=1}^{n} X_i = \overline{X}.$$

Возьмем это как оценку для μ.

Для оценки дисперсии необходимо рассчитать

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2.$$

Однако мы не знаем μ и, таким образом, должны использовать ее оценку

$$\sum_{i=1}^{n}(X_i - \overline{X})^2.$$

Чтобы увидеть, какой это имеет эффект, рассмотрим

$$\begin{aligned}
\sum_{i=1}^{n}(X_i-\overline{X})^2 &= \sum_{i=1}^{n}(X_i-\mu+\mu-\overline{X})^2 = \sum_{i=1}^{n}\left((X_i-\mu)+(\mu-\overline{X})\right)^2 \\
&= \sum_{i-1}^{n}\left((X_i-\mu)^2+2(X_i-\mu)(\mu-\overline{X})+(\mu-\overline{X})^2\right) \\
&= \sum_{i=1}^{n}(X_i-\mu)^2+2(\mu-\overline{X})\sum_{i=1}^{n}(X_i-\mu)+n(\mu-\overline{X})^2 \\
&= \sum_{i=1}^{n}(X_i-\mu)^2+2(\mu-\overline{X})\left(\sum_{i=1}^{n}X_i-n\mu\right)+n(\mu-\overline{X})^2 \\
&= \sum_{i=1}^{n}(X_i-\mu)^2+2(\mu-\overline{X})(n\overline{X}-n\mu)+n(\mu-\overline{X})^2 \\
&= \sum_{i=1}^{n}(X_i-\mu)^2-2n(\overline{X}-\mu)^2+n(\mu-\overline{X})^2 \\
&= \sum_{i=1}^{n}(X_i-\mu)^2-n(\overline{X}-\mu)^2 .
\end{aligned}$$

По определению ожидаемая величина

$$\sum_{i=1}^{n}(X_i-\mu)^2 \text{ — это } n\sigma^2.$$

По результатам последующего изучения вариации выборочной средней мы знаем, что ожидаемое значение $(\overline{X}-\mu)^2$ равно σ^2/n (см. гл. 5).

Таким образом, ожидаемое значение

$$\sum_{i=1}^{n}(X_i-\overline{X})^2 \text{ — это } n\sigma^2-n\frac{\sigma^2}{n}=(n-1)\sigma^2 .$$

Следовательно, для несмещенной оценки σ^2 нужно рассчитать

$$\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2 .$$

3

ПРИМЕНЕНИЕ ДИФФЕРЕНЦИАЛЬНОГО И ИНТЕГРАЛЬНОГО ИСЧИСЛЕНИЙ В ФИНАНСАХ

ВВЕДЕНИЕ

Исчисления — это раздел математики, который изучает изменения. Он разделен на два основных подраздела: дифференциальное исчисление и интегральное. С помощью дифференциального исчисления рассчитывается как или с какой скоростью изменяется переменная, в частности, на сколько одна переменная изменится в зависимости от бесконечно малого изменения другой переменной. Интегральное исчисление служит для расчетов площадей и объемов, ограниченных графиками функций.

В этой главе мы введем основные принципы исчислений и применим их к решению некоторых конкретных финансовых проблем. Дифференциальное исчисление будет использовано для определения изменения курса облигации в ответ на изменение ее доходности. Интегральное исчисление применяется для нахождения площади под кривой. Это в свою очередь будет использоваться в следующих главах при нахождении вероятности того, что определенная финансовая переменная примет значение в определенных пределах.

ДИФФЕРЕНЦИАЛЬНОЕ ИСЧИСЛЕНИЕ

Дифференциальное исчисление дает возможность определить, насколько одна переменная изменяется в ответ на изменения других, и является ли скорость изменения переменной возрастающей, убывающей либо постоянной. Например, мы можем рассчитать скорость автомобиля, зная время и пройденное расстояние. В области финансов также можно найти чувствительность облигаций к процентным ставкам.

Скорость изменения одной переменной Y в ответ на изменение другой переменной X известна под названием производной Y по X. Вторая производная показывает, является ли скорость изменения Y возрастающей, убывающей или постоянной. Рассмотрим эти понятия подробнее.

Первая производная — скорость изменения

В этом разделе мы рассмотрим скорость изменения переменной Y в зависимости от изменения переменной X. Затем в этой же главе рассмотрим, как изменяется Y, когда только одна из нескольких переменных изменяется, а также когда все эти переменные изменяются одновременно.

Рассмотрим следующие линейные функции: $Y = 12 + 3X$ и $Y = 3X$. Первая функция показывает, что когда $X = 0$, то $Y = 12$, во всех остальных случаях Y будет равен 12 плюс утроенное значение X. Вторая функция показывает, что значение Y равно утроенному значению X. Ниже мы приведем значения Y для возможных значений X для обеих функций:

$Y = 12 + 3X$	X	0	1	2	3	4	5	6	7	8	9	
	Y	12	15	18	21	24	27	30	33	36	39	
$Y = 3X$	X	0	1	2	3	4	5	6	7	8	9	10
	Y	0	3	6	9	12	15	18	21	24	27	30

Графики обеих функций представляют собой прямые линии и изображены на рис. 3.1.

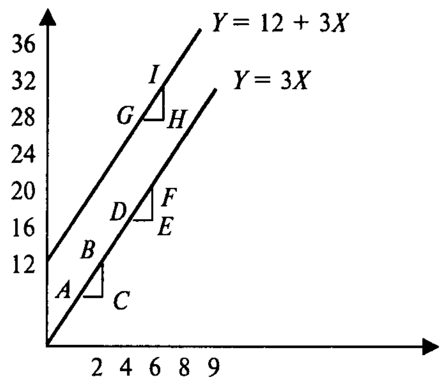

Рис. 3.1

Скорость изменения Y в зависимости от изменения X задана наклоном прямой. Наклон прямой определяется отношением ΔY (изменением Y) к ΔX (изменение X): $\Delta Y/\Delta X$. Если мы рассмотрим прямоугольные треугольники, гипотенузы которых являются участками графиков функций, то увидим, что отношение вертикали к горизонтали всегда равно 3 вне зависимости от участка прямой и функции.

Мы могли бы получить тот же результат, всего лишь взглянув на эти две функции. Мы знаем, что изменение Y будет равно трехкратному изменению X, таким образом, $\Delta Y/\Delta X$ равно 3. Очевидно, что поскольку функция линейна, то наклон графика, или скорость изменения Y в зависимости от изменения X, всегда равен 3, таким образом, $\Delta Y/\Delta X$ постоянно. Кроме того, свободный член первой функции (12) не имеет никакого влияния на наклон прямой, а влияет только на ее расположение. Таким образом, свободный член не влияет на скорость изменения Y.

Теперь обратимся к случаю, когда функция не является линейной. Например, рассмотрим функцию $Y = 4 + 2X^2$. Значения Y при различных значениях X приведены ниже:

X	0	1	2	3	4
Y4	6	12	22	36	

График функции изображен на рис. 3.2.

Важное свойство этой кривой состоит в том, что она становится круче с увеличением значения X. Другими словами, Y возрастает с растущей скоростью по мере того, как X увеличивается, даже если X возрастает с постоянной скоростью. Таким образом, угол наклона графика функции, или скорость изменения, не постоянна.

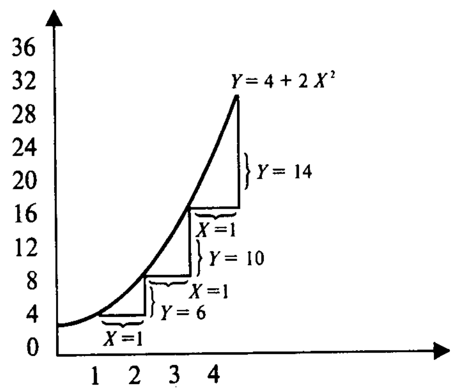

Рис. 3.2

Очевидно, что невозможно построить прямоугольный треугольник с гипотенузой, являющейся частью кривой. Тем не менее, представим себе прямоугольный треугольник, построенный между двумя точками на графике функции, как показано на рис. 3.2. Отношение $\Delta Y/\Delta X$ теперь не дает действительной или моментальной скорости изменения Y в любой из точек на участке кривой. Вместо этого мы получаем среднюю скорость изменения. Эта средняя тем менее будет репрезентировать степень изменения в определенный момент, чем больше расстояние между двумя точками на кривой. И наоборот, чем меньше это расстояние, тем более точно будет измерена скорость изменения Y. Когда расстояние между двумя точками становится меньше и меньше, оно становится бесконечно малым, или близким к нулю. В таком случае математики говорят, что оно приближается к предельному значению, к пределу. Пределом является изменение значения Y в ответ на бесконечно малое изменение X. Чтобы показать это предельное изменение, математики заменили обозначение $\Delta Y/\Delta X$, когда X приближается к пределу, на $\mathrm{d}Y/\mathrm{d}X$.

Теперь нам ясно, что для того чтобы найти моментальное изменение скорости, мы должны рассчитать $\mathrm{d}Y/\mathrm{d}X$. Но как это сделать? Чтобы понять это, рассмотрим функцию $Y = X^2$. Теперь предположим, что X изменяется на незначительную величину δX. Для сохранения равенства Y также возрастает на δY. Таким образом, для функции $Y = X^2$,

$$Y + \delta Y = (X + \delta X)^2. \tag{3.1}$$

Раскрыв правую часть, получим

$$(X + \delta X)^2 = (X + \delta X)\,(X + \delta X) =$$

$$= X{\cdot}X + X{\cdot}\delta X + X{\cdot}\delta X + \delta X{\cdot}\delta X = X^2 + 2(X{\cdot}\delta X) + \delta\, X^2\ .$$

Таким образом,

$$Y + \delta Y = X^2 + 2(X{\cdot}\delta X) + (\delta X)^2. \tag{3.2}$$

Нам нужно найти значение $\mathrm{d}Y$. Так как $Y = X^2$, можно убрать Y из левой части уравнения и X^2 из правой, сохраняя при этом равенство. Получаем:

$$\delta Y = 2(X{\cdot}\delta X) + (\delta X)^2. \tag{3.3}$$

Поскольку предел $\mathrm{d}X$ бесконечно мал, почти равен 0, то умножение $\mathrm{d}X^2$ на $\mathrm{d}X$ дает еще меньшую величину, еще более близкую к нулю. Таким образом:

$$\delta Y = 2(X{\cdot}\delta X). \tag{3.4}$$

Теперь разделим обе части уравнения на δX и получим предел

$$\frac{\mathrm{d}Y}{\mathrm{d}X} = 2X. \tag{3.5}$$

Так, для функции $Y = X^2$, $\mathrm{d}Y/\mathrm{d}X = 2X$. Таким образом, чтобы найти $\mathrm{d}Y/\mathrm{d}X$, мы просто умножим X на число, равное степени, в которую X был возведен, и затем уменьшим значение показателя степени на единицу, т.е. X^2 становится $2{\cdot}X^1$. Это можно представить в общем виде для функции $Y = X^n$ так:

$$\mathbf{d}Y/\mathbf{d}X = nX^{n-1}\ . \tag{3.6}$$

Применяя это к упомянутой выше функции $Y = 4 + X^2$, получаем

$$dY/dX = 2 \cdot 2\, X^1 = 4X. \tag{3.7}$$

Заметьте, что свободные члены исчезают при дифференцировании. Поскольку мы уже видели ранее на рис. 3.1, что свободный член не влияет на наклон графика функции, то можем сказать, что скорость изменения Y не зависит от значения свободного члена.

Теперь, рассмотрев основные принципы дифференцирования, мы можем вывести основные правила дифференцирования.

Нахождение производных от константы

1. Для функции $Y = a$

$$dY/dX = 0. \tag{3.8}$$

В этом случае Y равен константе a, которая по определению не может изменяться.

2. Для функции $Y = bX$

$$dY/dX = b. \tag{3.9}$$

В этом случае Y увеличивается с постоянной скоростью, равной $b \cdot X$, таким образом, прирост Y в зависимости от прироста X постоянен и равен b. График функции $Y = bX$ представляет собой прямую линию.

3. Для функции $Y = a + bX$

$$dY/dX = b. \tag{3.10}$$

В соответствии с правилом 1 мы игнорируем a, и Y изменяется с постоянной скоростью bX, в соответствии с правилом 2. График функции $Y = a + bX$ представляет собой прямую.

Производная степенной функции

4. Для $Y = X^n$

$$dY/dX = nX^{n-1}. \tag{3.11}$$

5. Частный случай правила 4, когда X возведен в отрицательную степень, т.е. $Y = X^{-n}$. В этом случае мы умножаем X на n в соответствии с правилом 4 и уменьшаем значение степени n на 1. Например, для функции $Y = X^{-2}$

$$dY/dX = -2\, X^{-3}\,. \tag{3.12}$$

Этот случай особенно важен для управляющего портфелем ценных бумаг, поскольку используется при определении чувствительности цены облигаций. Это проиллюстрировано далее в этой главе.

6. Второй частный случай правила 4 применяется для X в дробной степени, например $Y = X^{1/n}$; $X^{1/n}$ является корнем n степени из X; $X^{p/n}$ — это корень n степени из X в степени p.

Если $Y = X^{1/n}$

$$dY/dX = \frac{1}{n} \cdot X^{\frac{1}{n}-1}. \tag{3.13}$$

Например, для $Y = X^{1/3}$ (то же самое, что и $\sqrt[3]{X}$)

$$dY/dX = \frac{1}{3} \cdot X^{\left(\frac{1}{3}-1\right)} = \frac{1}{3} \cdot X^{-\frac{2}{3}} = \frac{1}{3\sqrt[3]{X^2}}.$$

Производная суммы двух функций от X

7. Для функции $Y = u + v$, где u и v — функции от X

$$\mathbf{d}Y/\mathbf{d}X = \mathbf{d}u/\mathbf{d}X + \mathbf{d}v/\mathbf{d}X. \tag{3.14}$$

Например, для $Y = 2X^4 + 3X^2$

$$\mathbf{d}Y/\mathbf{d}X = 8X^3 + 6X.$$

8. Для функции $Y = u—v$, где u и v — функции от X

$$\mathbf{d}Y/\mathbf{d}X = \mathbf{d}u/\mathbf{d}X—\mathbf{d}v/\mathbf{d}X. \tag{3.15}$$

Например, если $Y = 3X^3 - 2X$

$$\mathbf{d}Y/\mathbf{d}X = 9X^2 + 2.$$

Производная произведения двух функций от X

9. Если Y — произведение двух функций от X, то dY/dX будет равно произведению одной функции на производную другой плюс произведение другой функции на производную первой. Так, для $Y = u \cdot v$

$$Y/\mathbf{d}X = v(\mathbf{d}u/\mathbf{d}X) + u(\mathbf{d}v/\mathbf{d}X). \tag{3.16}$$

Например, если $Y = (5X + 2)\cdot(4X^2)$, то мы можем заменить правую часть на $u\cdot v$, где $u = 5X + 2$, и $v = 4X^2$. Первый шаг — это нахождение dy/dX и dv/dX:

$$du/dX = 5, \quad dv/dX = 8X.$$

Второй шаг — нахождение $v(du/dX)$:

$$4X^2 \cdot 5 = 20X^2 .$$

Третий — нахождение $u(dv/dX)$:

$$(5X + 2)\cdot 8X = 40X^2 + 16X.$$

Тогда,

$$dY/dX = 20X^2 + 40X^2 + 16X = 60X^2 + 16X.$$

Нахождение производной отношения двух функций от X

10. Для функции $Y = u/v$

$$\mathbf{dY/dX = [v(du/dX) - u(dv/dX)]/v^2.} \tag{3.17}$$

Например, предположим, что $u = 5X + 2$ и $v = 4X^2$

$$du/dX = 5; \; dv/dX = 8X;$$

$$dY/dX = \frac{4X^2\left(\frac{du}{dX}\right) - (5X + 2)\left(\frac{dv}{dX}\right)}{(4X)^2};$$

$$dY/dX = \frac{4X^2(5) - (5X + 2)(8x)}{(4X^2)^2}.$$

(Это выражение может быть дальше упрощено до $\frac{-5x - 4}{4x^3}$).

Производная сложной функции

11. Для функции $Y = f(u)$, где u в свою очередь функция от другой переменной, например $u = y(x)$

$$dY/dX = (dY/du)\cdot(du/dX). \tag{3.18}$$

Это правило цепного дифференцирования.

Для иллюстрации рассмотрим функцию $Y = (2X^3 + 3)^6$. Y здесь будет равен u^6, если $u = 2X^3 + 3$. Отсюда

$$\mathrm{d}u/\mathrm{d}X = 6X^2 \quad \text{и} \quad \mathrm{d}Y/\mathrm{d}u = 6u^5.$$

Соответственно $\mathrm{d}y/\mathrm{d}X = 6u^2 \cdot 6x^2$. Тогда

$$\mathrm{d}Y/\mathrm{d}X = 6(2X^3 + 3)^5 (6X^2).$$

Производная экспоненциальной функции

Экспоненциальная функция является особо важной в исчислении, поскольку кроме нулевой функции это единственная функция, не изменяющаяся при дифференцировании.

12. Например для $Y = \mathrm{e}^X$

$$\mathbf{d}Y/\mathbf{d}X = \mathbf{e}^X. \tag{3.19}$$

13. При $Y = \mathrm{e}^{aX}$

$$\mathbf{d}Y/\mathbf{d}X = a\mathbf{e}^{aX}. \tag{3.20}$$

Например, для $Y = \mathrm{e}^{3X}$

$$\mathrm{d}Y/\mathrm{d}X = 3\mathrm{e}^{3X}.$$

14. Если е возведено в более сложную степень, допустим

$$Y = \mathrm{e}^{X^3 + 2X^2},$$

то мы применим правило цепного дифференцирования (3.18). Например, пусть степень е будет представлена u, тогда $Y = e^u$. Затем находим $\mathrm{d}u/\mathrm{d}X$ и $\mathrm{d}Y/\mathrm{d}X$ как $\mathrm{d}Y/\mathrm{d}u \cdot \mathrm{d}u/\mathrm{d}X$. Согласно приведенному выше правилу (12) $\mathrm{d}Y/\mathrm{d}u = e^A$, и:

$$\mathrm{d}u/\mathrm{d}X = 3X^2 + 4X.$$

Таким образом,

$$\mathrm{d}Y/\mathrm{d}X = e^u \cdot (\mathrm{d}u/\mathrm{d}x). \tag{3.21}$$

Получаем

$$(3X^2 + 4X)\ e^{X^3 + 2X^2}.$$

Производная натурального логарифма

15. Если $Y = \log_e X$, то

$$\mathbf{d}Y/\mathbf{d}X = \mathbf{1}/X. \tag{3.22}$$

Таким образом, производной натурального логарифма является число, обратное X.

16. Если $Y = \log_e$ является натуральным логарифмом более сложной функции от X, скажем $\mathbf{Y = \log(X^3 + 2X^2)}$, то примем u за $(X^3 + 2X^2)$, таким образом, $Y = \log_e(u)$. Теперь мы найдем $\mathrm{d}u/\mathrm{d}X$ и затем $\mathrm{d}Y/\mathrm{d}X$ как $(\mathrm{d}Y/\mathrm{d}u)\cdot(\mathrm{d}u/\mathrm{d}X)$. Тогда $\mathrm{d}Y/\mathrm{d}u$ будет равняться:

$$\mathrm{d}Y/\mathrm{d}u = \frac{1}{u} = \frac{1}{X^3 + 2X^2}. \tag{3.23}$$

Мы знаем, что $\mathrm{d}u/\mathrm{d}X = 3X^2 + 4X$, тогда

$$\frac{\mathrm{d}Y}{\mathrm{d}X} = \left(\frac{\mathrm{d}Y}{\mathrm{d}u}\right) \cdot \left(\frac{\mathrm{d}u}{\mathrm{d}X}\right) = \frac{1}{u}(3X^2 + 4X) = \frac{3X^2 + 4X}{X^3 + 2X^2}.$$

Вторая производная

До сих пор мы рассматривали первую производную функции, которая позволяет найти скорость изменения функции. Чтобы определить, является ли скорость изменения постоянной, следует взять вторую производную функции. Это обозначается как $\mathbf{d^2Y/dX^2}$.

Чтобы взять вторую производную функции, мы всего лишь должны продифференцировать первую производную ее. Так, например $\mathbf{Y = 4 + 2X^2}$. Первая производная будет

$$\mathrm{d}Y/\mathrm{d}X = 4X.$$

Вторая же производная будет производной от $4X$, т.е. 4.

Разложение рядов Тейлора

Зачастую бывает полезно приблизиться к какой-либо функции, используя более простые функции. Разложение рядов Тейлора предоставляет нам методологию для аппроксимации. Например, рассмотрим цену двухгодичной облигации, по которой ежегодно производятся выплаты по купонам. Цена облигации P как функция от совокупной доходности (y) будет следующей:

$$P = f(y) = \frac{CF_1}{(1+y)} + \frac{CF_2}{(1+y)^2}.$$

В графическом виде эта функция известна как кривая цены-доходности и схематически изображена на рис. 3.3.

Функция $f(y)$ достаточно сложная и еще более усложняется по мере того, как увеличивается по облигации количество платежей к получению. Было бы значительно проще, если бы мы имели приближение к функции $f(y)$. Аппроксимации рядов Тейлора являются многочленными выражениями (произведениями чисел и степеней y), которые делают возможным подобные приближения. Так, мы можем построить постоянные, линейные, квадратичные или кубические аппроксимации, а также четвертой степени, пятой, и т.д.

По своей природе аппроксимации обычно имеют силу только на протяжении ограниченных интервалов значений аргумента (y — в данном случае). Мы свободны выбирать любую точку для аппроксимации f, обозначим ее y. Выбрав точку, зафиксируем ее и рассмотрим точки, которые влекут за собой бесконечно малые изменения y, скажем, $y + h$. Здесь h является бесконечно малой величиной и может быть отрицательной. Таким образом, h — это переменная.

Ниже приведен перечень аппроксимаций рядов Тейлора для $f(y)$ в окрестностях точки y. Приближение к константе мы обозначили $P_0(y)$, линейное приближение — $P_1(y)$, квадратичное — $P_2(y)$ и т.д.

$$P_0(y + h) = f(y)$$

$$P_1(y + h) = f(y) + f'(y) \cdot h$$

$$P_2(y + h) = f(y) + f'(y) \cdot h + \frac{f''(y)}{2} \cdot h^2 \tag{3.24}$$

$$P_3(y + h) = f(y) + f'(y) \cdot h + \frac{f''(y)}{2} \cdot h^2 + \frac{f'''(y)}{6} \cdot h^3$$

$$P_4(y + h) = f(y) + f'(y) \cdot h + \frac{f''(y)}{2} \cdot h^2 + \frac{f'''(y)}{6} \cdot h^3 + \frac{f^{iv}(y)}{24} \cdot h^4.$$

Заметьте, что $f'''(y)$ — это третья производная $f(y)$ при $y = y$, $f^{iv}(y)$ — четвертая производная при $y = y$.

Также заметьте, что 2, 6, 24 и т.д. являются факториалами: $2! = (2 \cdot 1)$, $3! = (3 \cdot 2 \cdot 1)$, $4! = (4 \cdot 3 \cdot 2 \cdot 1)$ и т.д.

Нет никакого смысла искать, какая аппроксимация является "лучшей" линейной или "лучшей" квадратической и т.д. Они построены таким образом, что P_0 имеет **то же самое значение** в точке y, что и P_1, и **такое же значение производной**, P_2 имеет **то же самое значение** и **такую же первую и вторую производные**, и т.д.

Как следует интерпретировать эти аппроксимации? Рассмотрим рис. 3.3. Высота $f(y)$ дает нам точку на линии, равную P. Теперь для приближения $P = f(y)$, когда y изменяется на малую величину, мы могли бы добавить линейное приближение. Это будет: $P = f(y + h) = f(y) + hf'(y)$. Будучи линейным приближением, изменение в значении P будет приблизительно равносильно движению вдоль прямой линии, а не кривой.

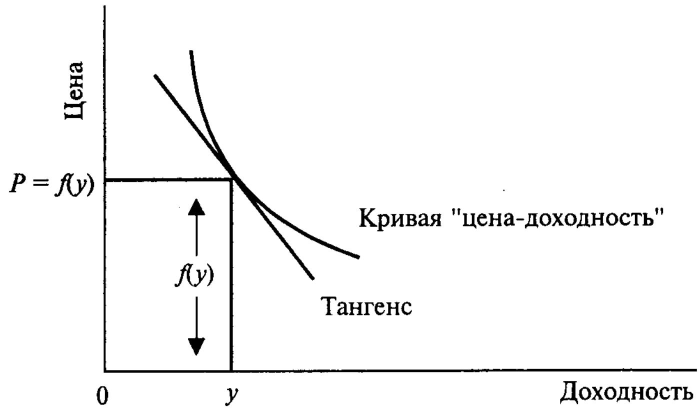

Рис. 3. 3. Кривая "цена-доходность" облигации

Таким образом, для всяких y, кроме наименьших изменений, приближение будет неточным. Мы можем улучшить точность применяя квадратичное приближение, которое повлечет за собой добавление второй производной к линейному приближению. Таким образом, $P = f(y) + hf'(y) + (h^2/2!)\,f''(y)$.

Для некоторых функций, обычно не для анализа облигаций, могут потребоваться кубические и приближения более высокого уровня.

Для иллюстрации применения линейных и квадратичных приближений мы применим эти принципы для нахождения, во-первых, изменения цены облигации и затем волатильности цены облигации.

ПРИМЕНЕНИЕ ДИФФЕРЕНЦИАЛЬНОГО ИСЧИСЛЕНИЯ

Применение рядов Тейлора при оценке изменений цены облигации

Рис. 3.3 представляет собой схематическое изображение соотношения между доходностью и ценой облигации. Кривая, известная как кривая цены-доходности облигации, нелинейна и имеет отрицательный наклон. Моделирование изменения цены в результате изменения доходности облигации может оказаться очень сложным. Тем не менее, исходя из нашего понимания разложения рядов Тейлора, мы должны быть способны приблизиться к функции "цена-доходность" на определенном этапе разложения рядов Тейлора. Можно, например, применить первую производную цены облигации по доходности, вторую, третью и т.д. Фактически мы увидим далее, что применение рядов Тейлора всего лишь первых двух порядков прекрасно позволяет оценить изменение в цене облигации при малом изменении доходности. Более того, если мы разделим разные элементы рядов Тейлора на цену облигации, то получим очень полезный результат, показывающий волатильность цены облигации.

Чтобы показать, как применять ряды Тейлора для нахождения изменений цен облигаций, рассмотрим простой пример однолетней облигации с нулевым купоном, 100% которой соответственно выплачиваются спустя один год. Если доходность на момент погашения составляет 10%, то текущая цена составит 90,91. Таким образом, $P = f(0,10) \mp 90,91$. Какова будет цена при изменении доходности с 0,10 до 0,11? Мы можем сделать линейную аппроксимацию к изменению, прибавляя первую производную, умноженную на изменение доходности, к постоянной и получим:

$$90,91 + h(dP/dy),$$

учитывая, что $P = f(y) = P = 100/(1 + y)$. Получаем:

$$\frac{dP}{dy} = \frac{-100}{(1+y)^2}. \tag{3.25}$$

Таким образом, при $y = 0,10$

$$\frac{dP}{dy} = \frac{-100}{1{,}10^2} = 82{,}6446.$$

Линейная аппроксимация к P при $y = 0{,}11$, т.е. $P = f(0{,}11)$ будет $90{,}9090-(0{,}01 \cdot 82{,}6446) = 90{,}0836$.

В действительности, однако, если доходность мгновенно поднимется до 0,11, цена облигации упадет только до 90,09. Следовательно, линейная аппроксимация недостаточно точна. Мы видим на рис. 3.3, что линейная аппроксимация преувеличивает понижение цены облигации.

Мы можем улучшить приближение, применяя квадратическая аппроксимацию, которая задана следующим образом:

$$P(y + h) = f(y) + \frac{\mathrm{d}P}{\mathrm{d}y}h + \frac{1}{2}\frac{\mathrm{d}^2P}{\mathrm{d}y}h^2. \tag{3.26}$$

Чтобы продолжить, мы должны вспомнить, что $1/y$ равно y^{-1}, $1/y^2 = y^{-2}$, и это может быть обобщено как

$$1/y^n = y^{-n}. \tag{3.27}$$

Соответственно

$$1/(1 + y)^1 = (1 + y)^{-1};\ 1/(1 + y)^2 = (1 + y)^{-2} \text{ и } 1/(1 + y)^T = (1 + y)^{-T}.$$

Мы должны применить это правило при определении второй производной. Вторая производная — это всего лишь производная от первой производной:

$$\frac{\mathrm{d}P}{\mathrm{d}y} = \frac{-100}{(1+y)^2} = -100 \cdot \frac{1}{(1+y)^2},$$

а это то же самое, что и

$$-100\,(1 + y)^{-2}.$$

Дифференцируя, получаем

$$-2\,(-100)\,(1 + y)^{-3}.$$

Таким образом, вторая производная цены данной облигации с нулевым купоном будет:

$$\frac{\mathrm{d}^2P}{\mathrm{d}y^2} = \frac{200}{(1+y)^3}.$$

При $Y = 0{,}10$ соответственно будем иметь:

$$\frac{d^2P}{dy^2} = \frac{200}{(1{,}10)^3} = 150{,}2630.$$

Квадратичное приближение будет равно:

$$90{,}0826 + 1/2\,(150{,}2630 \cdot 0{,}01^2) = 90{,}0826 + (0{,}5 \cdot 0{,}0150) = 90{,}0901.$$

Это равно действительному значению с точностью до четырех знаков после запятой.

Применение исчисления для измерения риска цены облигаций

Первый член рядов Тейлора, деленный на цену облигации, известен как **модифицированная дюрация**, второй член — как **выпуклость**. Члены более высокого порядка обычно считаются незначительными при определении чувствительности цены облигации.

Волатильность облигаций

Теперь мы используем некоторые рассмотренные выше концепции для определения ценовой волатильности облигации. В гл. 1 мы объяснили, что совокупная доходность облигации является внутренней ставкой дохода в год, что равняется стоимости будущих денежных потоков на текущую цену облигации.

Текущая цена купонной облигации рассчитывается следующим образом:

$$P_{CB} = \frac{GF_1}{(1+y)} + \frac{GF_2}{(1+y)^2} + \frac{GF_3}{(1+y)^3} + \ldots + \frac{GF_n}{(1+y)^n}, \qquad (3.28)$$

где y — периодическая совокупная доходность, т.е. внутренняя ставка дохода, отражающая периодичность денежных потоков.

Угол наклона касательной к кривой является первой производной цены по доходности. Первая производная, деленная на цену, дает процентное изменение цены облигации в ответ на изменение доходности на 1%, известное как модифицированная дюрация.

Для того чтобы понять, как берется первая производная dP/dy для каждой облигации, мы сначала рассчитаем ее для текущего значения одного единственного денежного потока

$$P_z = GF_T \cdot \frac{1}{(1+y)^T}, \tag{3.29}$$

где $1/(1 + y)^t$ — дисконтный множитель. Это является текущей стоимостью одной единицы, дисконтированной по соответствующей процентной ставке на T периодов времени. Очевидно, что умножая сегодняшнюю стоимость одной денежной единицы на сумму к получению в будущем, мы получим сегодняшнюю стоимость будущего денежного потока.

Мы можем использовать правило 5 дифференцирования, приведенное выше, но сначала преобразуем функцию (3.29) следующим образом:

$$P = CF_T \cdot (1 + y)^{-T}. \tag{3.30}$$

и dP/dY

$$\frac{dP}{dy} = (-T)CF_T(1 + y)^{-(T+1)}. \tag{3.31}$$

Так как лучше преобразовывать отрицательные показатели степеней в положительные, где это возможно, то получим

$$\frac{(-T)GF_T}{(1+y)^{T+1}}. \tag{3.32}$$

Это может быть распространено и на нахождение первой производной цены облигации по ее доходности на основе применения правила 7, учитывая, что купонная облигация может рассматриваться как портфель с нулевым купоном облигаций. Таким образом, применив правило 5 для нахождения dP/dy для каждого денежного потока, затем в соответствии с правилом 7 складываем первые производные для нахождения dP/dy для облигации в целом. Например, используем уравнение купонной облигации (3.28).

Это может быть выражено так:

$$P_{CB} = CF_1 \cdot (1 + y)^{-1} + CF_2 \cdot (1 + y)^{-2} + CF_3 \cdot (1 + y)^{-3} + \ldots + CF_n \cdot (1 + y)^{-n}.$$

Поскольку мы уже рассчитали dP/dY для беспроцентной облигации, то dP/dY для отдельных денежных потоков будет равняться:

$$\frac{dP_1}{dy} \text{ для } CF_1 = \frac{(-1)CF_1}{(1+y)^2};$$

$$\frac{dP_2}{dy} \text{ для } CF_2 = \frac{(-2)CF_2}{(1+y)^3};$$

$$\frac{dP_3}{dy} \text{ для } CF_3 = \frac{(-3)CF_3}{(1+y)^4}; \tag{3.33}$$

$$\frac{dP_n}{dy} \text{ для } CF_n = \frac{(-n)CF_n}{(1+y)^{n+1}},$$

где dP/dy — сумма $dP_1/dy + dP_2/dy + dP_3/dy + \ldots + dP_n/dy$. Следовательно, сложение отдельных производных в соответствии с правилом 7 дает

$$\frac{dP}{dy} = \frac{(-1)CF_1}{(1+y)^2} + \frac{(-2)CF_2}{(1+y)^3} + \frac{(-3)CF_3}{(1+y)^4} + \ldots + \frac{(-n)CF_n}{(1+y)^{n+1}}. \tag{3.34}$$

Преобразовав уравнение (3.34) и разделив его на P (или умножив на $1/P$), получим

$$\frac{dP}{dy}\frac{1}{P} = \frac{1}{(1+y)}\left[\frac{(-1)CF_1}{(1+y)^1} + \frac{(-2)CF_2}{(1+y)^2} + \frac{(-3)CF_3}{(1+y)^3} + \ldots + \frac{(-n)CF_n}{(1+y)^n}\right]\frac{1}{P}. \tag{3.35}$$

Выражение в скобках называется дюрацией Маколея (Macaulay's duration, 1938). Правая часть уравнения (3.35) известна как **модифицированная дюрация** и используется специалистами рынка облигаций как индикатор процентного риска облигаций. Модифицированная дюрация может быть интерпретирована как приблизительное процентное изменение цены облигации в результате изменения доходности на 1% при бесконечно малых изменениях доходности в следующий период времени.

Заметьте, что в расчетах дюрации совокупная доходность используется как ставка дисконтирования. Это предполагает, что временная структура процентной ставки является прямой, поскольку одна и та же ставка используется при дисконтировании всех денежных потоков независимо от их времени. Отсюда следует предположение, что временная структура может измениться на параллельную. Эти предположения не подтверждаются эмпирическими данными. К этим вопросам мы вернемся в гл. 11.

Численный пример модифицированной дюрации

Для иллюстрации применения дифференциального исчисления при нахождении модифицированной дюрации используем уравнение (3.35). Возьмем двухлетнюю облигацию, по которой выплачивается 5 единиц каждые полгода, с совокупной доходностью 8% годовых. Модель цены и денежных потоков будет выглядеть следующим образом:

$$P = \frac{5}{1{,}04} + \frac{5}{(1{,}04)^2} + \frac{5}{(1{,}04)^3} + \frac{105}{(1{,}04)^4} =$$

$$= 4{,}8077 + 4{,}6228 + 4{,}4450 + 89{,}7544 = 103{,}6299.$$

Сначала мы должны оценить аргументы внутри квадратных скобок в уравнении (3.35)

$$\frac{-1 \cdot 5}{(1{,}04)} + \frac{-2 \cdot 5}{(1{,}04)^2} + \frac{-3 \cdot 5}{(1{,}04)^3} + \frac{-4 \cdot 5}{(1{,}04)^4} =$$

$$= -4{,}8077 + (-9{,}2456) + (-13{,}335) + (-359{,}0175) = -386{,}4058, \quad (3.36)$$

затем умножаем полученную величину на $1/(1 + y)$

$$\frac{1}{1{,}04}(-386{,}4058) = \frac{-386{,}4058}{1{,}04} = -371{,}5440.$$

Наконец, делим этот результат на цену облигации $P = 103{,}6299$:

$$\frac{-371{,}5440}{103{,}6299} = -3.5853.$$

Эта модифицированная дюрация выражена в периодах денежных потоков. Для удобства работы на рынке облигаций дюрация выражается в годах и знак минус опускается. Поскольку по данной облигации выплаты по купону производятся дважды в год, дюрация в годах будет найдена делением 3,5853 на 2, таким образом, получим $3{,}5853/2 = 1{,}7926$.

Модифицированная дюрация может быть интерпретирована как индикатор чувствительности облигации к относительной процентной ставке. В частности, процентное изменение цены облигации может быть найдено как отрицательная модифицированная дюрация, умноженная на процентное изменение совокупной доходности. В данном случае если доходность возрастет на 0,5%, то цена облигации упадет на $0{,}5 \cdot 1{,}7926 = 0{,}8963\%$.

Вторая производная — скорость изменения скорости изменения

Если мы обратимся к рис. 3.2 и 3.3, то увидим, что на обоих графиках тангенс угла наклона касательной в разных точках кривой разный, очевидно, что он непрерывно изменяется по мере того, как мы движемся вверх или вниз по кривой. Это очень важно, поскольку скорость изменения переменной Y, обозначенная как $\mathrm{d}Y/\mathrm{d}X$, может быть найдена только для определенного значения X. Таким образом, также важно знать и то, с какой скоростью $\mathrm{d}Y/\mathrm{d}X$ изменяется по мере изменения X. Это изменение $\mathrm{d}Y/\mathrm{d}X$ называется второй производной Y по X и обозначается $\mathbf{d^2}\boldsymbol{Y}/\mathbf{d}\boldsymbol{X}^2$.

Соответственно если $\mathrm{d}Y/\mathrm{d}X$ обозначает скорость изменения, то $\mathrm{d}^2Y/\mathrm{d}X^2$ показывает, является ли эта скорость возрастающей, убывающей или постоянной.

Для взятия второй производной достаточно всего лишь продифференцировать первую производную. Таким образом, если $Y = X^2$ и $\mathrm{d}Y/\mathrm{d}X = 2X$, то

$$\mathrm{d}^2Y/\mathrm{d}X^2 = 2. \tag{3.37}$$

Аналогично третья производная является производной второй и т.д.

Применение второй производной: выпуклость облигации

Одно из применений второй производной — улучшение измерения чувствительности цены облигации при помощи модифицированной дюрации. Мы можем воспользоваться нашим знанием второй производной функции цены облигации от совокупной доходности для того, чтобы рассчитать то, что называется выпуклостью облигации.

Чтобы показать это на примере, вернемся к анализу одного денежного потока. Взять производную $\mathrm{d}P/\mathrm{d}y$ одного денежного потока можно так:

$$\frac{\mathrm{d}P}{\mathrm{d}y} = (-T)CF_T(1 + y)^{-(T+1)}. \tag{3.38}$$

Для того чтобы взять вторую производную, продифференцируем первую производную:

$$\frac{d^2P}{dy^2} = -(T+1)\cdot(-T)CF_T(1+y)^{-(T+2)}, \tag{3.39}$$

что можно представить в виде

$$\frac{d^2P}{dy^2} = \frac{-(T+1)\cdot(-T)CF_T}{(1+y)^{(T+2)}}. \tag{3.40}$$

Найдем вторую производную каждого денежного потока:

$$\frac{d^2P_1}{dy^2} \text{ для } CF_1 = (-2)\ (-1)CF_1(1+y)^{-3}$$

$$\frac{d^2P_2}{dy^2} \text{ для } CF_2 = (-3)\ (-2)CF_2(1+y)^{-4}$$

$$\frac{d^2P_3}{dy^2} \text{ для } CF_3 = (-4)\ (-3)CF_3(1+y)^{-5}$$

$$\frac{d^2P_n}{dy^2} \text{ для } CF_n = (-(n+1))\ (-n)CF_n(1+y)^{-(n+2)}.$$

Преобразовав эти выражения, получим

$$\frac{d^2P_1}{dy^2} \text{ для } CF_1 = \frac{(-2)(-1)CF_1}{(1+y)^3} = \frac{+2CF_1}{(1+y)^3}$$

$$\frac{d^2P_2}{dy^2} \text{ для } CF_2 = \frac{(-3)(-2)CF_2}{(1+y)^4} = \frac{+6CF_2}{(1+y)^4}$$

$$\frac{d^2P_3}{dy^2} \text{ для } CF_3 = \frac{(-4)(-3)CF_3}{(1+y)^5} = \frac{+12CF_3}{(1+y)^5}$$

$$\frac{d^2P_n}{dY^2} \text{ для } CF_n = \frac{(-(n+1)(-n)CF_n}{(1+y)^{n+2}} = \frac{(n+1)nCF_n}{(1+y)^{n+2}}.$$

Складывая отдельные производные в соответствии с правилом 7, приходим к следующему выражению:

$$\frac{d^2P_2}{dy^2} = \frac{(2)CF_1}{(1+y)^3} + \frac{(6)CF_2}{(1+y)^4} + \frac{(12)CF_3}{(1+y)^5} + \ldots + \frac{n\cdot(n+1)CF_n}{(1+y)^{n+2}}.$$

Численный пример выпуклости

Проиллюстрируем применение второй производной в управлении рисками облигаций путем нахождения выпуклости двухлетней облигации, рассмотренной выше. Выпуклость равняется половине второй производной функции цены облигации, деленной на цену облигации.

Вторая производная цены облигации будет равна:

$$\frac{2\cdot 5}{(1{,}04)^3}+\frac{6\cdot 5}{(1{,}04)^4}+\frac{12\cdot 5}{(1{,}04)^5}+\frac{20\cdot 105}{(1{,}04)^6}=$$

$$8{,}890 + 25{,}644 + 49{,}316 + 1659{,}661 = 1743{,}510.$$

При цене облигации, равной 103,6299, выпуклость будет равна

$$\frac{1}{2}\cdot\frac{1743{,}510}{103{,}6299}=8{,}4122.$$

Как и в случае с модифицированной дюрацией, в практике рынка ценных бумаг выпуклость измеряется в годах в квадрате. Поэтому необходимо разделить полученную величину на число денежных потоков в год, возведенное в квадрат. В данном примере имеются два денежных потока в год, значит, делитель будет равен $2^2 = 4$. Таким образом, выпуклость равна 2,1031.

Также можно использовать исчисление для нахождения приближеных значений модифицированной дюрации и выпуклости купонной облигации, если предположить, что соответствующая учетная ставка равна ставке валовой совокупной доходности.

Купонные денежные потоки по облигации могут быть рассмотрены как аннуитеты плюс выплата номинала облигации. Текущая стоимость аннуитета, выплачивающего C в течение n лет, составляет

$$P = C\left[\frac{1-\left[1/(1+y)^n\right]}{y}\right]. \tag{3.41}$$

Таким образом, цена облигации, выплачивающей C в течение n лет, плюс выплата 100:

$$P = C\left[\frac{1-\left[1/(1+y)^n\right]}{y}\right]+\frac{100}{(1+y)^n}. \tag{3.42}$$

Мы получим значение модифицированной дюрации, беря первую производную и деля ее на P. Вторая производная может быть использована для определения выпуклости. Начнем с первой производной.

Сначала представим уравнение (3.42) в виде, более удобном для дифференцирования:

$$P = C\left(1 - \frac{1}{(1+y)^n}\right)y^{-1} + 100(1+y)^{-n}. \tag{3.43}$$

Первая производная $\mathrm{d}P/\mathrm{d}y$ будет:

$$\frac{\mathrm{d}P}{\mathrm{d}y} = C\left[-\left(1 - \frac{1}{(1+y)^n}\right)y^{-2} + \left(\frac{n}{(1+y)^{n+1}}\right)y^{-1}\right] - \frac{100n}{(1+y)^{n+1}}$$

$$= C\left[\frac{1}{y^2}\left(\frac{1}{(1+y)^n} - 1\right) + \frac{n}{y(1+y)^{n+1}}\right] - \frac{100n}{(1+y)^{n+1}}$$

$$= \frac{C}{y^2}\left(\frac{1}{(1+y)^n} - 1\right) + \frac{n}{(1+y)^{n+1}}\left(\frac{C}{Y} - 100\right). \tag{3.44}$$

Разделив первую производную на грязную цену облигации P, получим модифицированную дюрацию:

$$\text{MOD DUR} = \frac{\left[\frac{C}{y^2}\left(\frac{1}{(1+y)^n} - 1\right) + \frac{n}{(1+y)^{n+1}}\left(\frac{C}{Y} - 100\right)\right]}{P}. \tag{3.45}$$

Аналогично определяем выпуклость. Первая производная уравнения (3.44) будет

$$\frac{\mathrm{d}^2P}{\mathrm{d}y^2} = -\frac{2C}{y^3}\left(\frac{1}{(1+y)^n} - 1\right) + \frac{C}{y^2}\left(-\frac{n}{(1+y)^{n+1}}\right) -$$

$$-\frac{n(n+1)}{(1+y)^{n+2}}\left(\frac{C}{y} - 100\right) + \frac{n}{(1+y)^{n+1}}\left(-\frac{C}{y^2}\right).$$

Разделив сумму аргументов на удвоенную грязную цену облигации P, получим

$$\frac{1}{2}\frac{\left[-\frac{2C}{y^3}\left(\frac{1}{(1+y)^n}-1\right)-\frac{2C}{y^2}\left(\frac{1}{(1+y)^{n+1}}\right)-\frac{n(n+1)}{(1+y)^{n+2}}\left(\frac{C}{y}-100\right)\right]}{P}. \tag{3.46}$$

МАКСИМУМЫ И МИНИМУМЫ

В области финансов существует множество ситуаций, когда бывает нужно знать максимум или минимум функции. Например, мы хотим узнать, какая комбинация активов несет в себе минимальный уровень риска, при этом достигая требуемого уровня доходности. Кроме того, возможны ситуации, когда нам нужно максимизировать вероятность того, что параметры, используемые в модели, являются действительными. Проблемы первого типа подробно рассмотрены в гл. 9, которая раскрывает вопросы оптимизации. Проблемы второго типа разобраны в приложении к гл. 7, где рассмотрены вопросы максимальной вероятности оценки, в частности применительно к анализу временных рядов.

Первая и вторая производные могут быть использованы, когда определяют, является ли данная точка кривой пиком (максимумом), либо дном впадины (минимумом), либо точкой перегиба, где кривая из выпуклой становится вогнутой, или наоборот. Пример пиков и впадин приведен на рис. 3.4.

На рисунке мы ясно видим, что точка A является локальным максимумом. Другими словами, значение функции не возрастает и не убывает, т.е. первая производная равна 0. Точка B находится на дне впадины, она является локальным минимумом, и здесь первая производная также равна нулю. Точка C — точка перегиба, она не является ни максимумом, ни минимумом, но первая производная в этой точке тоже равна нулю. Таким образом, локальный максимум, локальный минимум и точка перегиба имеют одну общую черту — они являются особыми точками, или **экстремумами функции**, где $dy/dx = 0$ в каждом случае.

Если признаком всех экстремумов является равенство нулю первой производной, то как определить, является ли данная точка максимумом, минимумом или точкой перегиба, не пользуясь при этом рисунком графика функции?

Ответ дает вторая производная. Когда $dY/dX = 0$ и вторая производная больше нуля, $d^2Y/dX^2 > 0$, точка является минимумом. Если же $dY/dX = 0$ и вторая производная в этой точке меньше нуля т.е. $d^2Y/dX^2 < 0$, то точка является максимумом. Чтобы понять это, вспомним, что вторая производная показывает, как изменяется первая производная. Теперь посмотрим, что происходит с первой производной в точке *A*. Слева от точки *A* производная выше нуля, а справа ниже, таким образом, в точке *A* значение первой производной меняется с положительного на отрицательное, изменение отрицательное, следовательно, и вторая производная меньше нуля.

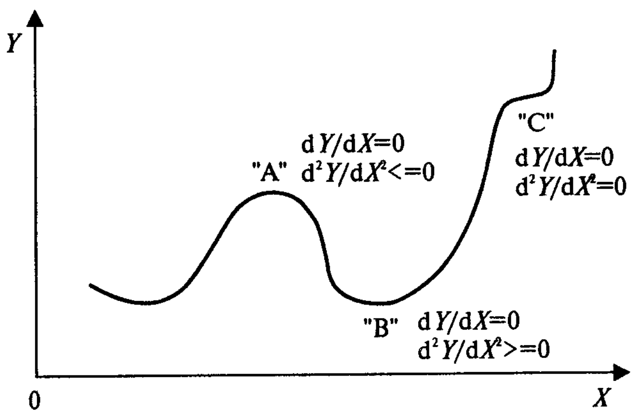

Рис. 3.4. Максимумы и минимумы

Теперь посмотрим, что происходит в точке *B*. Слева от точки первая производная меньше нуля, справа — больше, само изменение положительно, таким образом, и вторая производная больше нуля.

Если же и первая и вторая производные равны нулю, то данная точка является точкой перегиба. Чтобы понять это, снова взглянем на рис. 3.4. В точке *C* первая производная равна нулю. Слева от точки она меньше нуля, а справа — больше, но в точке *C* она равна нулю. Таким образом, точки перегиба — это точки, где обе производные — и первая и вторая — равны нулю, но вторая меняет знак.

Итак, мы можем вывести несколько правил по определению максимумов, минимумов и точек перегиба:

1. Любая точка, где $dY/dX = 0$, является либо локальным минимумом, либо локальным максимумом, либо точкой перегиба.
2. Если в этой точке d^2Y/dX^2 меньше нуля, то это точка локального максимума.
3. Если в этой точке d^2Y/dX^2 больше нуля, то это точка локального минимума
4. Если в этой точке d^2Y/dX^2 равно нулю и меняет знак, то это точка перегиба.

Нахождение минимальных и максимальных значений функции

Теперь мы можем использовать наше понимание минимальных и максимальных точек для нахождения максимальных и минимальных значений функции. Для этого мы должны найти точки, в которых первая производная равна нулю. Таким образом мы определим экстремумы функции. Затем рассчитаем вторую производную. Если эта величина меньше нуля, то точка является точкой локального максимума. Если же найденная величина больше нуля, то в данной точке мы имеем локальный минимум.

Эти позиции проиллюстрируем следующим образом. Найдем максимум функции $Y = 4X^3—2X$.

Первым шагом будет нахождение dY/dX, при которых первая производная равна нулю. Таким образом:

$$\frac{dY}{dX} = 12X^2—2 = 0;$$

$$\therefore 12X^2 = 2;$$

$$\frac{2}{12} = X^2;$$

$$X = \pm\sqrt{\frac{1}{6}} = \pm 0{,}4028.$$

Отсюда одна из точек экстремума $X = 0{,}4028$, а другая — $X = —0{,}4028$.

Следующий шаг — определение знака второй производной данной функции

$$\frac{d^2Y}{dX^2} = 24X.$$

При $X = +\,0{,}4028$ вторая производная больше нуля, следовательно, это точка локального минимума. При $X = -0{,}4028$ вторая производная принимает отрицательное значение и, таким образом, это точка максимума. Для определения значения Y в точке максимума подставим $-0{,}4028$ вместо X в функцию, $Y = 4X^3 - 2X$, т.е.

$$Y = 4\,(0{,}4082)^3 - 2\,(0{,}4082) = 0{,}02721 - 0{,}8164 = -0{,}5443.$$

ДИФФЕРЕНЦИРОВАНИЕ ФУНКЦИЙ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

До настоящего времени мы рассматривали функции одной переменной. В финансах, как и во многих других областях экономики, одна переменная очень часто является функцией нескольких других переменных. Когда мы дифференцируем такую функцию только по одному из ее аргументов, мы вычисляем **частную производную**.

Уравнения, которые содержат частные производные, известны как **уравнения частных производных**, они особенно важны при оценке с помощью производных. В гл. 10 мы увидим, как непрерывно-временные уравнения частных производных используются при определении цены опционов.

Взятие частных производных

Рассмотрим функцию $Y = f(X,Z)$, где X и Z — независимые друг от друга переменные. Можно взять производную такой функции по одной из переменных, допустив, что другая остается неизменной. Такой вид дифференцирования и называется взятием частной производной.

Частная производная обозначается $\partial Y/\partial X$ вместо dY/dX.

В соответствии с правилами дифференцирования функции с более чем одной переменной функция дифференцируется по какой-либо переменной, остальные переменные рассматриваются как константы. Таким образом, для взятия частных производных от

$$Y = X^2 + 2Z^3 + Z^2 \tag{3.47}$$

мы находим $\partial Y/\partial X$, принимая Z за константу, и $\partial Y/\partial Z$, принимая X за константу. Например

$$\frac{\partial Y}{\partial X} = 2X,$$

$$\frac{\partial Y}{\partial Z} = 6Z^2 + 2Z.$$

Теперь рассмотрим функцию трех переменных

$$\begin{aligned} &Y = X^4 + W^3X + XZ - 4Z^3, \\ &\frac{\partial Y}{\partial X} = 4X^3 + W^3 + Z, \\ &\frac{\partial Y}{\partial W} = 3W^2X, \\ &\frac{\partial Y}{\partial Z} = X - 12Z^2. \end{aligned} \tag{3.48}$$

Полученные результаты говорят, что $\partial Y/\partial X$ показывает, что по мере возрастания X Y будет увеличиваться на $(4X^3 + W^3 + Z)$ ∂X. Если возрастает W, то Y возрастает на $3W^2X \cdot \partial W$, и если увеличивается Z, то Y возрастает на $(X - 12Z^2) \cdot \partial Z$.

Вторые производные $\partial^2 Y/\partial X^2$, $\partial^2 Y/\partial W^2$ и $\partial^2 Y/\partial Z^2$ показывает, как ведут себя предельные изменения Y, когда X, W или Z изменяются, но остальные две переменные остаются постоянными.

Если же мы хотим узнать, как изменяется $\partial Y/\partial W$ при изменении X, мы берем производную $\partial Y/\partial W$ по X. Это записывается как $\partial^2 Y/\partial X \partial W$, т.е. показывается, что мы хотим найти, как изменяется $\partial Y/\partial W$ при изменении X. Аналогично, если мы хотим найти влияние изменения Z на $\partial Y/\partial W$, мы должны найти $\partial^2 Y/\partial Z \partial W$. Запись обозначает, что мы берем производную $\partial Y/\partial W$ по Z, т.е. ∂Z.

Можно также использовать обозначение $\partial/\partial Z$, чтобы показать, что мы берем производную по Z

$$\frac{\partial^2 Y}{\partial Z \partial W} = \frac{\partial}{\partial Z}\left(\frac{\partial Y}{\partial W}\right).$$

Полный дифференциал

Полный дифференциал объясняет, как изменяется Y при изменении всех независимых переменных. Рассмотрим еще раз $Y = f(W, X, Z)$. При малых изменениях независимых переменных

$$\Delta Y = \frac{\partial Y}{\partial W}(\Delta W);$$

$$\Delta Y = \frac{\partial Y}{\partial X}(\Delta X);$$

$$\Delta Y = \frac{\partial Y}{\partial Z}(\Delta Z).$$

Заметьте, что δx или ΔX используется для обозначения малого изменения X.

Для обозначения того, как Y изменяется в ответ на одновременное изменение всех аргументов, мы сложим произведения частных производных на малые изменения соответствующих переменных. Полное уравнение выглядит следующим образом:

$$\Delta Y = \frac{\partial Y}{\partial W}(\Delta W) + \frac{\partial Y}{\partial X}(\Delta X) + \frac{\partial Y}{\partial Z}(\Delta Z), \tag{3.49}$$

где ΔY известно как полный дифференциал. Для примера снова обратимся к функции

$$Y = X^4 + W^3X + XZ - 4Z^3;$$

$$\frac{\partial Y}{\partial X} = 4X^3 + W^3 + Z;$$

$$\frac{\partial Y}{\partial W} = 3W^3X;$$

$$\frac{\partial Y}{\partial Z} = X - 12Z^2.$$

Полный дифференциал функции выглядит следующим образом:

$$\Delta Y = + (4X^3 + W^3 + Z)(\Delta X) + (3W^2X)(\Delta W) + (X - 12Z^2)(\Delta Z).$$

Максимумы и минимумы функций нескольких переменных

При рассмотрении этой темы ограничимся разбором случая функции двух переменных. Пусть это будет функция $f(X, Y)$. Теперь мы можем найти две частные производные и четыре вторые частные производные:

$$\frac{\partial f}{\partial X}, \frac{\partial f}{\partial Y} \text{ и} \tag{3.50}$$

$$\frac{\partial^2 f}{\partial X^2}, \frac{\partial^2 f}{\partial Y^2}, \frac{\partial^2 f}{\partial Y \partial X}, \frac{\partial^2 f}{\partial X \partial Y}. \tag{3.51}$$

При этом заметим, что

$$\frac{\partial^2 f}{\partial Y \partial X} = \frac{\partial^2 f}{\partial X \partial Y}. \tag{3.52}$$

В случае с двумя переменными существуют три типа экстремумов — локальный максимум, локальный минимум и седловина, схожие с наивысшими точками горных перевалов — горные пики по сторонам и долины спереди и сзади.

Условия для так называемых "сильных экстремумов" следующие.

Локальный максимум

Критерием для локального максимума является равенство первых производных нулю, т.е.

$$\frac{\partial f}{\partial X} = \frac{\partial f}{\partial Y} = 0. \tag{3.53}$$

Кроме того,

$$\frac{\partial^2 f}{\partial X^2} < 0, \qquad \frac{\partial^2 f}{\partial X^2}\frac{\partial^2 f}{\partial Y^2} > \left(\frac{\partial^2 f}{\partial X \partial Y}\right)^2. \tag{3.54}$$

Локальный минимум

Критерием для локального минимума является равенство первых производных нулю, т.е.

$$\frac{\partial f}{\partial X} = \frac{\partial f}{\partial Y} = 0 \tag{3.55}$$

и также

$$\frac{\partial^2 f}{\partial X^2} > 0, \qquad \frac{\partial^2 f}{\partial X^2}\frac{\partial^2 f}{\partial Y^2} > \left(\frac{\partial^2 f}{\partial X \partial Y}\right)^2. \tag{3.56}$$

Приведенные выше условия относятся к "сильным" локальным экстремумам. "Слабые" локальные экстремумы, которые являются "гребнями" вместо пиков и "долинами" вместо "дна чаши", находятся таким же образом, только знаки $<$ и $>$ заменяются на $\le$ и $\ge$.

Максимумы и минимумы функции на определенном интервале: оператор Лагранжа

Существует множество примеров ситуаций в бизнесе и финансах, где было бы нужно знать максимальное либо минимальное значение функции на определенном интервале. Например, при управлении портфелем желательно знать ожидаемый доход при условии установленного минимального уровня риска. Мы воспользуемся так называемым оператором Лагранжа.

Например, предположим, что хотим найти максимум следующей функции

$$R = 5X + 2X^2 - 4Y \tag{3.57}$$

при ограничении

$$2X + Y = 20. \tag{3.58}$$

Первым шагом будет преобразование функции ограничения таким образом, чтобы в правой части остался 0, т.е.:

$$20 - 2X - Y = 0.$$

Полученное выражение ограничителя затем умножается на неопределенную переменную λ — множитель Лагранжа

$$\lambda(20 - 2X - Y) = 0. \tag{3.59}$$

Затем вычитаем полученную функцию из (3.57) и получаем новую функцию

$$L(X, Y, \lambda) = 5X + 2X^2 - 4Y - \lambda(20 - 2X - Y). \tag{3.60}$$

Новая функция известна под названием оператора Лагранжа или лагранжиана; она дифференцируется по всем переменным, включая λ. Таким образом

$$\frac{\partial L}{\partial X} = 5 + 4X + 2\lambda = 0,$$

$$\text{отсюда } 5 + 4X = -2\lambda,$$

$$\frac{\partial L}{\partial Y} = -4 + \lambda = 0,$$

$$\text{получаем } \lambda = 4,$$

$$\frac{\partial L}{\partial \lambda} = 20 - 2X - Y = 0,$$

$$\text{отсюда } 2X + Y = 20$$

решаем

$$\lambda = 4,$$

$$5 + 4X = -2\lambda,$$

$$2\lambda = 8,$$

$$\therefore 5 + 4X = -8,$$

$$4X = -13,$$

$$\therefore X = -13/4 = -3{,}25,$$

$$2X + Y = 20,$$

$$2(-3.25) + Y = 20,$$

$$-7.5 + Y = 20,$$

$$Y = 27{,}5.$$

ИНТЕГРАЛЬНОЕ ИСЧИСЛЕНИЕ (ИНТЕГРИРОВАНИЕ)

Существует два вида интегрирования, первый является по сути процессом, обратным дифференцированию, и называется вычислением **неопределенного интеграла**. Таким образом, исходя из производной, мы можем судить о виде первоначальной функции. Второй вид интегрирования — это процесс нахождения площади, ограниченной кривыми, известный как нахождение **определенного интеграла**. Мы увидим, что по сути своей оба процесса являются одним и тем же. (Это основная теорема (математического) анализа).

Интеграл обозначается в виде вытянутой буквы *s*, $\int$ и всегда сопровождается обозначением переменной, по которой дифференцирована исходная функция. Таким образом, $\int X^2 \mathrm{d}X$ обозначает “взять интеграл X^2 по X”.

Неопределенный интеграл

Проиллюстрируем неопределенный интеграл, называемый также обратным дифференцированием, на примере функции $Y = aX^n$. Мы знаем, что $\mathrm{d}Y/\mathrm{d}X = naX^{n-1}$. Производя обратный дифференцированию процесс, получим некую исходную функцию. Например,

$$Y = \int naX^{n-1}\mathrm{d}X. \qquad (3.61)$$

Это обозначает взятие интеграла от naX^{n-1} по X. Ответом здесь будет

$$Y = \int naX^{n-1}\mathrm{d}X = aX^n. \qquad (3.62)$$

Для взятия неопределенного интеграла нам нужно выполнить всего лишь несколько простых правил:

1. Для интегрирования $Y = X^n$ мы прибавляем 1 к показателю степени и делим X на новый показатель степени. Например, для функции $4X$ прибавляем 1 к показателю степени, т.е. $4X^{1+1}$ и разделим на $1 + 1 = 2$. Отсюда неопределенный интеграл будет равен:

$$Y = \frac{4X^{1+1}}{1+1} = \frac{4X^2}{2} = 2X^2.$$

2. Вспомните, что все константы исходной функции теряются при дифференцировании. Следовательно, при интегрировании мы должны принять во внимание константы, хотя и не знаем их значения. Так, C, производная постоянная интег-

рирования, обычно прибавляется к результату интегрирования. Именно по причине того, что величина C не определена, этот вид интегрирования называется неопределенным.

Таким образом, правило интегрирования степенной функции можно отобразить в следующем виде: прибавить единицу к показателю степени, разделить на первую экспоненту и прибавить константу, т.е.

$$\int X^n \mathrm{d}X = \frac{X^{n+1}}{n+1} + C. \tag{3.63}$$

Нахождение площади под кривой

Для нахождения площади под кривой мы можем использовать то, что только что узнали о **неопределенных интегралах**. Однако перед этим мы должны ввести новую концепцию —концепцию **первообразной**.

Первообразная — это функция F, первая производная которой равна функции f. Таким образом, если $2X$ — это первая производная от X^2, то X^2 является первообразной от $2X$. Выше мы уже рассмотрели, как взять неопределенный интеграл. Можно воспользоваться этим при нахождении первообразной, поскольку неопределенный интеграл и является первообразной этой функции. Таким образом, первообразная $2X$ будет

$$\frac{2X^{1+1}}{2} = X^2,$$

а первообразная X^2:

$$\frac{X^{2+1}}{2+1} = \frac{X^3}{3}.$$

Довольно скоро понятие первообразной понадобится нам, а пока мы ненадолго оставим эту тему.

Теперь рассмотрим рис. 3.5. Диагональная линия определяется функцией $Y = X$. Площадь квадрата $0abX_1$ равна X^2, где $X = X_1$. Площадь треугольника $0bX_1$ должна быть равна половине площади квадрата, т.е. $X^2/2$. Например, если $X = 2$, площадь под кривой между 0 и $X = 2$ будет $2^2/2 = 4/2 = 2$.

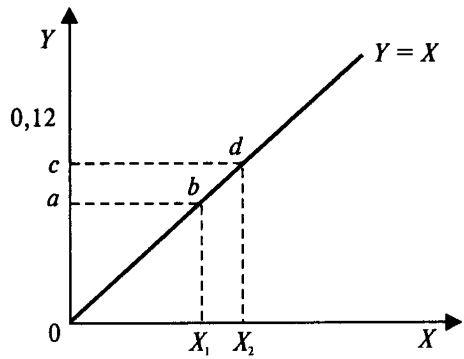

Рис. 3.5

Затем рассмотрим прямоугольник $0cdX_2$. Площадь опять будет X^2, при $X = X_2$, площадь треугольника будет $X^2/2$.

Теперь рассмотрим площадь многоугольника X_1bdX_2. Она равна разности площадей двух треугольников: $0dX_2$—$0bX_1$. Мы можем рассчитать площадь многоугольника, вычислив площадь каждого из треугольников и вычитая меньшую из большей. Однако существует более легкий способ, который также применим к случаям, когда график функции не является прямой. Этот способ использует концепцию первообразной, рассмотренную выше. Как уже выше сказано, график прямой линии описан функцией $Y = X$ и независимо от значения X, площадь двух треугольников выражена функцией $X^2/2$. При этом $X^2/2$ является первообразной от X. Так как $X = X^1$, мы можем показать это следующим образом:

$$\text{первообразная от } X = \frac{X^{1+1}}{1+1} = \frac{X^2}{2}.$$

Таким образом, мы можем найти площадь под кривой между X_1 и X_2 путем нахождения первообразной от X, выразив первообразную в значениях X_1 и X_2 и затем вычитая из большего значения меньшее.

Указание на взятие интеграла с использованием первообразной функции $f(X)$ обычно обозначается так:

$$\int_j^k f(x)\mathrm{d}X = \left[F(X)\right]_j^k, \tag{3.64}$$

где j и k — нижнее и верхнее значения по X интегрируемой функции.

Теперь предположим, что $j = 2$ и $k = 4$. Мы знаем, что первообразная от X — это $X^2/2$. Отсюда для нахождения площади под кривой $Y = X$ между $X = 2$ и $Y = 4$ мы определим следующий определенный интеграл, подставляя значения X в первообразную:

$$\int_2^4 X\mathrm{d}X = \left[\frac{X^2}{2}\right]_2^4.$$

Сначала для $X = 2$ мы получим

$$\frac{2^2}{2} = \frac{4}{2} = 2.$$

Затем для $X = 4$:

$$\frac{4^2}{2} = \frac{16}{2} = 8.$$

Отсюда площадь под кривой будет

$$\mathbf{8—2 = 6.}$$

Мы можем проверить это, обратившись к рис. 3.5. Площадь треугольника $0dX_2$ при $X = 4$ будет $4^2/2 = 8$. Площадь треугольника $0bX$ при $X = 2$ будет $2^2/2 = 2$.Вычитая из площади треугольника $0dX_2$ площадь треугольника $0bX_1$, получим 8—2 = 6. Тот же самый результат получен при помощи первообразной.

Теперь рассмотрим другую линейную функцию, $Y = 4X$, график которой изображен на рис. 3.6.

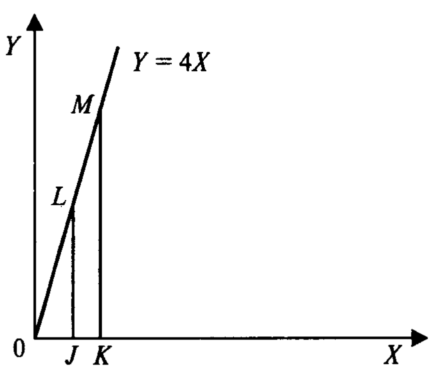

Рис. 3.6

Нам надо найти площадь многоугольника $JLMK$. Для этого надо найти площади треугольников $0LJ$ и $0MK$ и вычесть первую из последней. Этот же результат достигается взятием определенного интеграла

$$\int_j^k 4X\,dX = [F(X)]_j^k ,$$

где F — первообразная $4X$.

Сначала возьмем первообразную от $4X$, т.е.

$$\text{первообразная от } 4X = \frac{4X^{1+1}}{2} = 2X^2 .$$

Чтобы определить площадь под кривой между $J = 2$ и $K = 5$, значения 2 и 5 подставляются в первообразную вместо X и вычисляется разница:

$$\int_2^5 4X\,dX = \left[2X^2\right]_2^5 ,$$

что составляет

$$2(5)^2 - 2(2)^2 = 2 \cdot 25 - 2 \cdot 4 = 50 - 8 = 42.$$

Таким образом, площадь под кривой между $X = 2$ и $X = 5$ равна 42 квадратным единицам.

Это и есть применение основной теоремы математического анализа, которая гласит, что для определения площади под кривой $Y=f(X)$ следует взять первообразную от функции f между точками j и k.

Приведенные выше примеры относятся к ситуации, когда Y — линейная функция от X и верхняя граница площади — прямая линия. Тем не менее аналогичный способ может быть использован для нахождения площади под кривой, где Y является нелинейной функцией от X. Чтобы показать это, мы вновь прибегнем к использованию треугольников — см рис. 3.7.

Здесь мы опять хотим вычислить площадь под кривой между a и b. Можно начать с построения множества прямоугольников с шириной ΔX и высотой, равной расстоянию от оси $0X$ до кривой в верхней границе ΔX. Мы видим, что сумма площадей построенных прямоугольников больше площади под кривой из-за небольших участков, близких по форме к треугольникам, которые выступают над кривой. Заметьте также, что это указывает на интервалы, в течение которых функция возрастает. Схожая ситуация и в случае убывания функции.

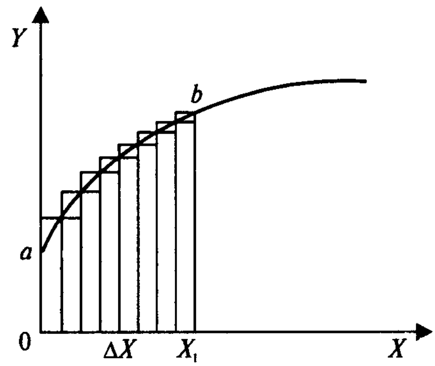

Рис. 3.7

Теперь рассмотрим один из таких прямоугольников поближе, как это показано на рис. 3.8.

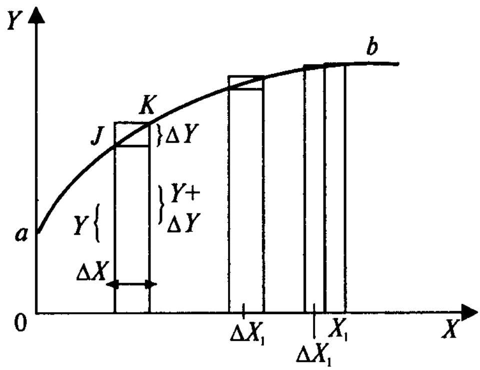

Рис. 3.8

Заметьте, что левый прямоугольник по сути состоит из трех меньших прямоугольников. Первый с шириной ΔX_0 и высотой Y, соответствующей точке J на кривой. Второй — с высотой $Y + \Delta Y$, соответствующей высоте в точке K. Третий, наименьший, наверху, имеет высоту ΔY. Очевидно, что площадь прямоугольника $\Delta X_0(Y + \Delta Y)$ больше площади под кривой между точками J и K, в то время как прямоугольник $\Delta X_0 Y$ — меньше.

Если обозначим всю площадь под кривой $0abX_1$ через A и действительную площадь под кривой над ΔX_0 — как ΔA, то взаимоотношение может быть представлено так:

$$(Y + \Delta Y)(\Delta X_0) > \Delta A > (Y)(\Delta X_0). \tag{3.65}$$

Если мы поделим эту функцию на ΔX_0, то получим

$$Y + \Delta Y > \frac{\Delta A}{\Delta X} > Y. \tag{3.66}$$

Теперь рассмотрим, что происходит по мере того, как ΔX уменьшается. Мы видим на рис. 3.8, что ΔX_q меньше, чем ΔX_p, который в свою очередь меньше ΔX_0, и размер прямоугольника $\Delta Y \cdot \Delta X$ становится меньше и меньше. Соответственно превышение площади прямоугольника над площадью под кривой становится меньше. Посмотрим, что случится, когда ΔX приблизится к нулю. В этом случае

$$\frac{\Delta A}{\Delta X} \to \frac{dA}{dX}, \tag{3.67}$$

и неравенство (3.66) становится таким:

$$Y + \Delta Y > \frac{dA}{dX} > Y, \tag{3.68}$$

тогда

$$\frac{dA}{dX} \to Y. \tag{3.69}$$

Поскольку Y — функция X, то $dA/dX = Y = f(X)$. Таким образом, dA/dX равняется высоте кривой.

Для того чтобы найти площадь под кривой между a и b, мы интегрируем dA/dX для значений X между a и b. Это делается с помощью первообразной функции Y.

Таким образом мы показали, что вне зависимости от того, является ли функция $Y = f(X)$ линейной или нелинейной, площадь под кривой может быть найдена путем взятия определенного интеграла с применением первообразной

$$\int_a^b f(X)dX = F(b) - F(a). \tag{3.70}$$

Используя ранее приведенное обозначение, это можно записать так:

$$\int_a^b f(X)\mathrm{d}X = [F(X)]_a^b, \qquad (3.71)$$

где $F(X)$ — первообразная функции $Y = f(X)$.

Покажем это на примере функции $Y = X^4 + 3X—3$. Нам нужно вычислить площадь под кривой между $X = 3$ и $X = 6$, точнее

$$\int_3^6 \left(X^4 + 3X - 3\right)\mathrm{d}X. \qquad (3.72)$$

Первообразная этой функции

$$\left[\frac{X^5}{5} + \frac{3X^2}{2} - 3X\right]. \qquad (3.73)$$

При определении значения функции между $X = 3$ и $X = 6$ получаем

$$\int_3^6 \left(X^4 + 3X - 3\right)\mathrm{d}X = \left[\frac{X^5}{5} + \frac{3X^2}{2} - 3X\right]_3^6. \qquad (3.74)$$

Теперь подставляем 6 вместо X в правую часть уравнения (3.74):

$$\frac{6^5}{5} + \frac{3(6)^2}{2} - 3(6) = \frac{7776}{5} + \frac{108}{2} - 18 = 1591{,}2. \qquad (3.75)$$

Подставляем 3 вместо X в правую часть уравнения (3.74)

$$\frac{3^5}{5} + \frac{3(3)^2}{2} - 3(3) = \frac{243}{5} + \frac{27}{2} - 9 = 53{,}1. \qquad (3.76)$$

Площадь под кривой будет равна разности между (3.75) и (3,76), т.е. 1591,2—53,1 = 1538,1 кв.ед.

УПРАЖНЕНИЯ

1. Найдите первую производную по Y следующих функций по X:

$Y = X$, $\qquad$ $Y = CX$,

$Y = X^2$, $\qquad$ $Y = X^{-4}$,

$Y = X^{1/4}$, $\qquad$ $Y = X^{1/2}$.

2. Найдите первую производную по Y следующих функций по X:

$$Y = 3X^3 + 4X^5, \qquad Y = 5X + 4X^2,$$
$$Y = 3X^5 - 2X^2, \qquad Y = X^3 - 3X^2.$$

3. Найдите первую производную по Y следующих функций по X:

$$Y = X^2(X + 2)^3, \qquad Y = (3X + 2)2X^3,$$
$$Y = e^{2X}, \qquad Y = X^3e^X,$$
$$Y = X^3e^{3X}, \qquad Y = X\ln(X),$$
$$Y = X^2(X + 2)^3e^{2X}.$$

Примечание: сначала повторите правила дифференцирования степенных и логарифмических функций.

4. Найдите первую производную по Y следующих функций по X:

$$Y = \frac{3X}{X + 2}, \qquad Y = \frac{\ln(X)}{X^2(X + 2)^3}.$$

5. Найдите первую производную по Y следующих функций по X:

$$Y = (X^2 + 2)^3, \qquad Y = (3X^2 + 5)^4,$$
$$Y = e^{X^2(X+2)^3}, \qquad Y = \ln(3X).$$

Проверьте правильность последнего решения, построив графики функций $Y = \ln(X)$ и $Y = \ln(3X)$.

6. Найдите первую производную по Y следующих функций по X:

$$\frac{1}{(1 + X)}, \qquad \frac{1}{(1 + X)^2},$$
$$\frac{1}{(1 + X)^{10}}, \qquad \frac{2X}{(1 + X)^3}.$$

7. Найдите вторую производную по Y следующих функций по X от функций в пп. 1 и 2.

8. Вычислите стоимость трехлетней купонной облигации с совокупной доходностью 7,5% годовых. Используйте разложение рядов Тейлора для определения стоимости облигации, если доходность станет 7,6%.

9. Трехлетняя облигация с ежегодным купоном, равным 10, и с номиналом, равным 100, имеет текущую совокупную доходность 8% годовых. Используя разложение рядов Тейлора, определите стоимость облигации, если доходность станет 8,1%.

10. Используя денежные потоки из п. 9, рассчитайте модифицированную дюрацию и выпуклость облигации.

11. Найдите максимум функции $Y = 2X^2 + 3X$.
12. Найдите полный дифференциал функции:

$$Y = X^3 + 4X^2 + Z^2.$$

13. Найдите максимум функции:

$$R = 3X + X^2 - 2Y$$

при условии, что $2X + Y = 5$.

14. Найдите первообразные следующих функций:

$$X^3, \qquad e^{2X}$$

$$\frac{1}{(1+X)^2}, \qquad \frac{\ln X}{X},$$

$$X^2, \qquad e^{X^3}, \qquad Xe^{X^2}.$$

Для четвертой функции используйте производную от $(\ln X)^2$, для функций в нижней строке используйте производную от e^{X^2}.

СПИСОК РЕКОМЕНДУЕМОЙ И ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ

Hunt, R. (1994) *Calculus of Single Variable.* HarperCollins, New York.

Larson, R. E., Hostetler, R. P., Edwards, B.H. (1994) *Calculus.* D.C. Health, Lexington, MA.

Macaulay, F.R. (1938) *Some Theoretical Problems Suggested by Movements in Interest Rates, Bond Yields and Stock Prices in the US since 1856.* National Bureau of Economic Research, New York.

Watsham T.J. (1993) *International Portfolio Management: A Modern Approach.* Longman, London.

ОТВЕТЫ

1. $1; \quad C; \quad 2X; \quad -4X^{-5}$

$$\frac{1}{4}X^{-\frac{3}{4}} = \frac{1}{4\sqrt[4]{X^3}}$$

$$\frac{1}{2}X^{-\frac{1}{2}} = \frac{1}{2\sqrt{X}}$$

2. $9X^2 + 20X^4$; $5 + 8X$; $15X^2 - 4X$; $3X^2 - 6X$.

3. $2X(X + 2)^3 + 3X^2(X + 2)^2$
$6X^2(3X + 2) + 6X^3$
$2e^{2X}$
$3X^2e^X + X^3e^X$
$\ln(X) + 1$
$2X(X + 2)^3e^{2X} + 3X^2(X + 2)^2e^{2X} + 2X^2(X + 2)^3e^{2X}$

4. $$\frac{3(X+2) - 3X}{(X+2)^2}$$
$$\frac{X(X+2)^3 - \ln(X)\left(2X(X+2)^3 + 3X^2(X+2)^2\right)}{X^4(X+2)^6}$$

5. $6X(X^2 + 2)^2$; $24X(3X^2 + 5)^3$
$(2X(X + 2)^3 + 3X^2(X + 2)^2)\, e^{X^2(X+2)^3}$
$1/X$

6. $\frac{-1}{(1+X)^2}$; $\frac{-2}{(1+X)^3}$; $\frac{-10}{(1+X)^{11}}$
$$\frac{2(1+X)^3 - 6X(1+X)^2}{(1+X)^6}$$

7. На первый вопрос:

0; 2; $20X^{-6}$; $\frac{-3}{16}X^{-\frac{7}{4}}$; $\frac{-1}{4}X^{-\frac{3}{2}}$

На второй вопрос:

$18X + 80X^3$: 8; $30X - 4$; $6X - 6$

8. Стоимость

$$\frac{100}{(1{,}075)^3} = 80{,}496\,.$$

Это значение $\frac{100}{(1+X)^3}$

при $X = 0{,}075$,

если

$$Y = \frac{100}{(1+X)^3}; \frac{dY}{dX} = \frac{-300}{(1+X)^4}; \frac{d^2Y}{dX^2} = \frac{1200}{(1+X)^5}$$

$$\frac{100}{(1+0{,}075+0{,}001)} \approx$$

$$80{,}50 + 0{,}001\left(\frac{-300}{(1{,}075)^4}\right) + \frac{0{,}001^2}{2}\left(\frac{1200}{(1{,}075)^5}\right) =$$

$$80{,}496 — 0{,}225 + 0{,}000 = 80{,}271.$$

9. $$f(X) = \frac{10}{1+X} + \frac{10}{(1+X)^2} + \frac{110}{(1+X)^3}$$

$$f'(X) = \frac{-10}{(1+X)^2} - \frac{20}{(1+X)^3} - \frac{330}{(1+X)^4}$$

$$f''(X) = \frac{20}{(1+X)^3} + \frac{60}{(1+X)^4} + \frac{1320}{(1+X)^5},$$

тогда

$$f(0{,}081) \approx \frac{10}{1+0{,}08} + \frac{10}{(1+0{,}08)^2} + \frac{110}{(1+0{,}08)^3}$$

$$+ 0{,}001\left(\frac{-10}{(1+0{,}08)^2} - \frac{20}{(1+0{,}08)^3} - \frac{330}{(1+0{,}08)^4}\right)$$

$$+ \frac{0{,}001^2}{2}\left(\frac{20}{(1+0{,}08)^3} - \frac{60}{(1+0{,}08)^4} - \frac{1320}{(1+0{,}08)^5}\right)$$

$$= 105{,}154 — 0{,}267 + 0{,}000 = 104{,}89.$$

10. $$MD = \frac{1}{105{,}154}\left(\frac{-10}{1{,}08^2} + \frac{-20}{1{,}08^3} + \frac{-330}{1{,}08^4}\right) = -2{,}54$$

$$CONV = \frac{\frac{1}{105{,}154}\left(\frac{20}{1{,}08^3} + \frac{60}{1{,}08^4} + \frac{1320}{1{,}08^5}\right)}{2} = 4{,}56\,.$$

11. $Y = 2X^2 + 3X$

$$\frac{dY}{dX} = 4X + 3 = 0 \text{ при } X = -3/4$$

$$X = -3/4 \Rightarrow Y = 2(9/16) + 3(-3/4) = 9/8 - 18/8 = -9/8$$

$$\frac{d^2Y}{dX^2} = 4 > 0 \;\therefore \text{минимум.}$$

12. $$\frac{dY}{dX} = 3X^2 + 8X;\ \frac{dY}{dZ} = 2Z$$

$$\therefore \Delta Y = (3X^2 + 8X)\,dX + 2Z\Delta Z.$$

13. $$L(x,\ y,\ \lambda) = 3X + X^2 - 2Y - \lambda(2X + Y - 5)$$

$$\partial L/\partial X = 3 + 2X - 2\lambda$$

$$\partial L/\partial Y = -2 - \lambda,$$

при $\lambda = -2$

$$3 + 2X - 2(-2) = 0 \Rightarrow X = -3{,}5$$

$$2X + Y = 5 \Rightarrow Y = 12.$$

Значение максимума R:

$$3(-3{,}5) + (-3{,}5)^2 - (2 \cdot 12) = 22{,}25.$$

14.

$$\int X^3 dX = \frac{X^4}{4}; \qquad \int e^{2X} dX = \frac{1}{2}e^{2X}; \qquad \int \frac{1}{(1+X)^2} dX = \frac{-1}{1+X};$$

$$\int \frac{\ln X}{X} dX = \frac{1}{2}(\ln X)^2; \quad \int X^2 e^{X^3} dX = \frac{1}{3}e^{X^3}; \quad \int X e^{X^2} dX = \frac{1}{2}e^{X^2}.$$

4

РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ: ПРИМЕНЕНИЕ ПРИ ОЦЕНКЕ РЕНТАБЕЛЬНОСТИ АКТИВОВ

ВВЕДЕНИЕ В ТЕОРИЮ ВЕРОЯТНОСТЕЙ

Вероятность — это мера того, что какое-либо случайное событие произойдет. Вероятность может принимать значения от 0 (невозможное событие) до 1 (достоверное событие). Распределения вероятностей — это математическая модель вероятности наступления случайных событий.

Теория вероятностей играет важную роль в финансах, поскольку практически во всех случаях результаты принятых финансовых решений неопределенны. Например, большинство людей считают будущую цену акций неопределенной, поскольку цены акций изменяются изо дня в день. Можно предполагать, какова будет цена акций в будущем, но мы принимаем, что в действительности цена может отличаться от предполагаемой. С другой стороны, процентные выплаты по банковскому депозиту с фиксированной ставкой также являются неопределенными, поскольку существует вероятность того, что банк обанкротится и мы не сможем получить ожидаемой суммы. Мы вряд ли ожидаем, что банк разорится, иначе мы не стали бы хранить там деньги. Тем не менее, опыт подсказывает нам, что и подобного рода учреждения могут "лопнуть", и, следовательно должна быть некоторая степень, касающаяся неопределенности ожидаемого размера процентов.

Распределения вероятностей близки распределениям частот, с которыми мы столкнулись в гл. 2. Они задают для каждого интервала всех возможных значений вероятность того, что случайная переменная примет значение в данном интервале. Распределения очень важны при принятии решений, поскольку позволяют оценивать вероятность наступления всех видов событий, и поэтому применяются в управлении финансами.

В этой книге мы уделим внимание двум группам распределений вероятностей. Первая группа, которую рассмотрим в этой главе, включает в себя те распределения, которые могут быть использованы для описания поведения рентабельности активов. Использование этих распределений позволяет нам оценить рискованность портфеля финансовых инструментов, таких, как опционы. Эта группа включает в себя нормальное, логнормальное, биномиальное распределения Пуассона и Парето—Леви.

Ко второй группе относятся распределения описательных статистических показателей (известных также как выборочные распределения), которые используются при проверке выдвигаемых гипотез. Эти распределения включают *t*-распределение Стьюдента, $\chi 2$- и *F*-распределения. Их применение будет рассмотрено в гл. 5 и 6.

Перед тем как углубиться в новую тему, необходимо уяснить значение некоторых основных терминов, таких, как "испытание", "событие" и "пространство элементарных событий". Испытание — это любое действие, которое приводит к определенному набору результатов. Конкретные результаты или сочетание их — это события, в то время как множество возможных результатов называется пространством элементарных событий.

Цель этой главы — ознакомление читателя с основами теории вероятностей и некоторыми вероятностными моделями, применимыми при оценке рентабельности активов. Что касается теории, мы рассмотрим три подхода: **классический**, или **априори**, **эмпирический** и **субъективный**. Затем ознакомимся с правилами расчета вероятностей, после чего обсудим математические действия над случайными величинами. Наконец, мы рассмотрим несколько распределений вероятностей и примеры их использования.

Классический, или априори, подход к вероятности

Этот подход применяется, когда возможные неопределенные результаты известны и равновероятны. При помощи простой логики можно определить вероятность каждого исхода.

Рассмотрим подбрасывание монеты. Подразумевается, что должен выпасть либо "орел" либо "решка", причем вероятности каждого результата равны между собой. Количество возможных результатов равно двум, что определяется формой монеты. Вероятность выпадения "орла" должна быть равна 0,5, и вероятность выпадения "решки" также должна быть равна 0,5. Подбрасывание монеты здесь является **испытанием, пространство элементарных событий** — два возможных результата эксперимента, и **событие** — это выпадение "орла" либо "решки".

Таким же образом, в случае с шестигранной игральной костью, где выпадение каждой из сторон равновероятно, испытанием будет метание кости, шесть разных граней представляют собой пространство элементарных событий, и событие — это грань, оказавшаяся наверху. Здесь вероятность выпадения каждой из граней будет 1/6, или 1,66666.

В каждом из описанных выше случаев вероятность результатов была определена формой монеты или кости. Это и лежит в основе *классической* или, априори, теории вероятностей.

При таких обстоятельствах вероятность наступления события определяется так:

$$P(A)=\frac{\text{Число равновероятных результатов, связанных с событием}}{\text{Общее число возможных результатов}}, \quad (4.1)$$

где $P(A)$ — вероятность наступления события A.

Вероятность того, что событие A не наступит, равна P (не A) $= 1-P(A)$.

Эмпирический подход

Однако в финансах, как и во многих других сферах, мы не всегда можем полагаться на точность процесса при определении вероятностей. Например, ученый может быть вынужден повторить испытание множество раз с целью определения вероятности наступления возможных событий. Множество значений доходности активов в финансах практически неограничено, таким образом, у финансового аналитика может возникнуть необходимость отслеживания множества изменений цен активов для того чтобы определить вероятность будущего изменения цены на определенную величину.

В таких случаях вероятность результата Z, $P(Z)$ рассчитывается как отношение числа (количества раз) наступления Z к числу всех событий, т.е. к числу раз проведения испытаний:

$$P(Z)=\frac{\text{Количество наступлений } Z}{\text{Число испытаний}}. \quad (4.2)$$

Читатель может заметить, что это практически анализ относительной частоты наблюдений.

Этот подход анализирует историческую информацию с целью определения вероятности наступления событий в будущем.

Рассмотрим, например, выборку из 100 последовательных ежедневных изменений цены акции, где каждое изменение аналогично одному испытанию, и, таким образом, проведено 100 испытаний. Каждое изменение цены — это событие, а пространство элементарных событий — это множество изменений цены акции на определенную величину, произошедших в действительности. Этот подход известен как **метод относительных частот или эмпирический подход к вероятности**.

Этому подходу мы уделим особое внимание в данной главе, поскольку он позволяет на основании исторических данных выдвигать предположения относительно распределения вероятностей будущей рентабельности активов. Например, предположим, что из 100 изменений цены 5 составляли 0,5 пенса, 15 — 1, 20 — 1,5, 30 — 2,0, 20 — 2,5 и 10 — 3 пенса. Вероятность того, что случайным образом в выбранный день изменение цены составит 1 пенс, будет 15/100, или 0,15. Аналогично вероятность изменения цены на 3 пенса составит 10/100, или 0,1.

Можно записать это как P(1 пенс) = 0,15 и P(3 пенса) = 0,1.

Складывая все вероятности всех возможных изменений цены, получаем:

$$0{,}05 + 0{,}15 + 0{,}2 + 0{,}3 + 0{,}20 + 0{,}1 = 1.$$

Таким образом, одна из черт вероятности состоит в том, что вероятность наступления одного события имеет значение от 0 до 1, а сумма вероятностей всех событий должна равняться единице.

Субъективный подход

Существует также и третий подход к теории вероятностей, известный как **субъективный подход**. Согласно этому подходу вероятность определяется как степень уверенности в наступлении того или иного события.

Субъективная вероятность применяется при решении многих проблем в бизнесе, где вероятность не может быть выведена при помощи логики, либо недостаточно эмпирических данных, на основании которых можно оценить вероятность. Например, субъективная вероятность включается в прогнозирование прибылей компании инвестиционным аналитиком. Она используется также в некоторых методах расчетов ожидаемого дохода от инвестиций.

ОСНОВНЫЕ ПРАВИЛА ТЕОРИИ ВЕРОЯТНОСТЕЙ

Вне зависимости от подхода к теории вероятностей применяется несколько формальных правил. Применимость каждого из правил зависит от того:

1) имеем ли мы дело с отдельным событием, в таком случае результаты соотносятся только с этим событием;
2) имеем ли мы дело с комбинацией нескольких событий, например изменениями индексов FTSE 100 и S&P 500;
3) являются ли совместные события независимыми или взаимоисключающими.

Эти правила — это правила сложения и умножения вероятностей.

Правило сложения применяется в случае, если мы хотим узнать вероятность того, что событие *A* **или** *B* случится, и если мы хотим узнать, являются ли события *A* и *B* взаимоисключающими.

Правило умножения используется для нахождения вероятности одновременного наступления событий *A* **и** *B*. В этом случае нужно также знать, являются ли события *A* и *B* независимыми друг от друга.

Правило сложения применительно к взаимоисключающим событиям

Если два события взаимоисключающие, например, если результат может быть или *A* или *B*, но не оба, то мы применяем **особое правило сложения** и **складываем вероятности** каждого события, для того чтобы определить вероятность наступления хотя бы одного из событий.

Если события *A* и *B* взаимоисключающие, то

$$P(A \textbf{ или } B) = P(A) + P(B). \tag{4.3}$$

Заметьте, что $P(A$ или $B)$ обозначает вероятность наступления события *A* или события *B*. Правильное обозначение этого будет $P(A \vee B)$, однако часто используется и обозначение $P(A \cup B)$, где "$\cup$" обозначает объединение двух наборов.

Это правило будет продемонстрировано далее на примере биномиального представления изменения цен активов.

Правило сложения для взаимонеисключающих событий

Если результаты испытаний не являются взаимоисключающими, то применяется **общее правило сложения вероятностей**, которое можно представить в общем виде:

$$P(A \text{ или } B) = P(A) + P(B) - P(A \text{ и } B). \quad (4.4)$$

Объяснением этому правилу служит то, что некоторые события могли привести к результату *A*, некоторые — к результату *B*, а некоторые — и к *A* и к *B*, поскольку *A* и *B* не исключают друг друга. Таким образом, если мы хотим узнать вероятность наступления *A* или *B*, мы должны вычесть из суммы результаты, которые приводят к *A* и *B* одновременно, поскольку иначе пересечение будет сосчитано дважды — один раз как часть *A* и другой — как часть *B*.

Рассмотрим в качестве примера вероятности роста или падения двух биржевых индексов. События не являются взаимоисключающими, поскольку очевидно, что оба индекса могут опускаться и подниматься одновременно. Предположим, что индекс FTSE 100 может возрасти с вероятностью 0,55 и упасть с вероятностью 0,45. Также предположим, что в тот же интервал времени индекс S&P 500 возрастет с вероятностью 0,35 и может упасть с вероятностью 0,65. Существует также вероятность, равная 0,3, что оба индекса возрастут одновременно. Какова вероятность того, что индекс FTSE 100 или S&P 500 возрастет:

$$P(\text{FTSE } 100\uparrow) = 0{,}55; \qquad P(\text{S\&P } 500\uparrow) = 0{,}35;$$
$$P(\text{FTSE } 100\downarrow) = 0{,}45; \qquad P(\text{S\&P } 500\downarrow) = 0{,}65;$$
$$P(\text{FTSE } 100\uparrow \text{ и S\&P } 500\uparrow) = 0{,}3; \quad (4.4.1)$$
$$P(\text{FTSE } 100\uparrow \text{ или S\&P } 500\uparrow) = 0{,}55 + 0{,}35 - 0{,}3 = 0{,}60.$$

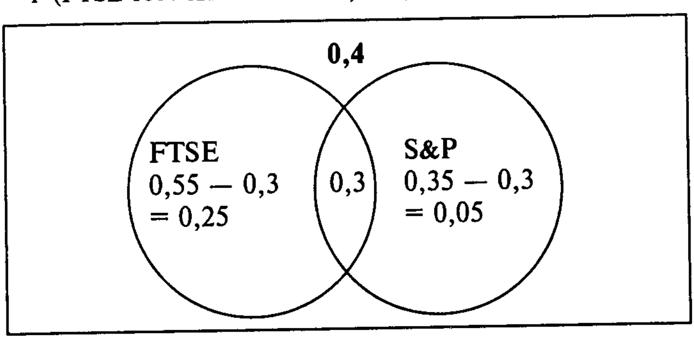

Рис. 4.1

Правило умножения для независимых событий

Два события считаются независимыми в теории вероятностей, если наступление события A никоим образом не сказывается на вероятности наступления события B. Таким образом,

$$P(A|B) = P(A). \tag{4.5}$$

$A|B$ значит "A при условии B".

Если события независимы, то вероятность наступления и события A и события B определяется так:

$$P(A \text{ и } B) = P(A) \cdot P(B). \tag{4.6}$$

Опять-таки правильным обозначением "A и B" будет $A \wedge B$, но часто используется запись $A \cap B$, где знак "$\cap$" обозначает пересечение.

Две переменные считаются независимыми, если обладают ковариацией друг с другом, равной нулю. Например, если два фондовых индекса не влияют друг на друга своими изменениями, то их ковариация равна нулю, и, следовательно, они независимы. Однако следует отметить, что ковариация между основными индексами обычно отличается от нуля.

Тем не менее, представим, что индексы FTSE 100 и S&P 500 не зависят друг от друга. Какова будет тогда вероятность того, что оба индекса возрастут одновременно?

$$P(\text{FTSE } 100\uparrow) = 0{,}55; \qquad P(\text{S\&P } 500\uparrow) = 0{,}35.$$

Отсюда вероятность одновременного возрастания обоих индексов будет

$$P(\text{FTSE } 100\uparrow \text{ и S\&P } 500\uparrow) = 0{,}55 \cdot 0{,}35 = 0{,}1925.$$

Правило умножения применительно к зависимым событиям

Если события не являются независимыми, то вероятность наступления A и B определяется произведением вероятностей наступления события A ($P(A)$) и **условной вероятности** наступления события B при условии наступления A. Условная вероятность обозначается как $P(B|A)$. Вертикальная черта означает "при условии". Так $P(B|A)$ означает вероятность наступления B при условии, что наступает A. Отсюда вероятность наступления A и B при условии, что A и B не являются независимыми событиями, находится следующим образом:

$$P(A \text{ и } B) = P(A) \cdot P(B|A),$$

т.е. (4.7)

$$P(B|A) = P(A \text{ и } B)/P(A).$$

Это правило применимо в случае с двумя фондовыми рынками, если между изменениями их индексов существует положительная или отрицательная ковариация. Каждое изменение состояния рынка — это событие, а условная вероятность — это вероятность роста или падения рынка, обусловленная падением или ростом другого рынка. Мы знаем, что FTSE 100 и S&P 500 не являются независимыми друг от друга, из гл. 2 мы знаем, что коэффициент корреляции между ними равен 0,793.

В нашем примере вероятность роста FTSE 100 в следующий момент времени и роста S&P 500 равна 0,3. Мы также знаем, что вероятность роста FTSE 100 равна 0,55. Отсюда можно вывести вероятность роста S&P 500, обусловленную ростом FTSE 100:

$$P(\text{FTSE } 100\uparrow \text{ и S\&P } 500\uparrow) = P(\text{FTSE } 100\uparrow) \cdot P(\text{S\&P } 500\uparrow | \text{FTSE } 100\uparrow);$$

$$\text{т.е. } 0{,}3 = 0{,}55 \cdot P(\text{S\&P } 500\uparrow | \text{FTSE } 100\uparrow).$$

Таким образом, вероятность $P(\text{S\&P } 500 | \text{FTSE } 100) =$

$$\mathbf{0{,}30/0{,}55 = 6/11 = 0{,}5454...}$$

ДИСКРЕТНЫЕ И НЕПРЕРЫВНЫЕ СЛУЧАЙНЫЕ ПЕРЕМЕННЫЕ

Случайная переменная — это такая переменная, поведение которой неопределенно. А поскольку поведение неопределенно, то мы можем только приписать вероятности возможным значениям таких переменных. Таким образом, случайная переменная определяется ее распределением вероятностей и возможных результатов. В гл. 2 мы классифицировали данные как дискретные и непрерывные; подобным же образом мы можем классифицировать и случайные переменные как дискретные и непрерывные. И поскольку существуют два типа случайных переменных, то также существуют и два типа распределений вероятностей — **непрерыв-**

ные распределения и **дискретные распределения**. В случае дискретных распределений вероятностей мы имеем дело с вероятностями переменными, принимающими только определенные дискретные значения. Например, бросок игральной кости может иметь только определенное количество возможных результатов со значением от 1 до 6. Непрерывные же распределения — это распределения вероятностей переменных, которые могут принимать любое значение в пределах определенного интервала.

Дискретные случайные переменные

Дискретные случайные переменные — это те, которые имеют конечное число возможных результатов. Рассмотрим ситуацию с бросанием шестигранной кости. С каждым из возможных результатов связана определенная вероятность; для нормальной кости каждая из шести вероятностей равна 1/6. Этот процесс можно смоделировать математически в виде дискретной случайной переменной.

В этом случае мы могли бы назвать случайную переменную Z и определить вероятности для Z, принимающей значения от 1 до 6, и вероятности каждого результата. Вероятности вместе со связанными с ними значениями случайной переменной и составляют **функцию частот вероятностей**, определяющую случайную переменную:

Значения	$r =$	1	2	3	4	5	6
Вероятности	$Z = r$	1/6	1/6	1/6	1/6	1/6	1/6

Заметьте, что независимо от количества возможных результатов сумма всех вероятностей должна быть равна единице

$$\sum_{-\infty}^{+\infty} P(X) = 1 . \tag{4.8}$$

Примерами дискретного распределения являются биномиальное и триномиальное распределения. Подбрасывание монеты приводит к биномиальному распределению результатов, поскольку результат может быть либо "орлом", либо "решкой". Цены активов могут падать, расти или оставаться неизменными, что приводит к триномиальному распределению, поскольку могут быть три вида результатов — рост, падение и отсутствие изменений.

Непрерывные случайные переменные

Непрерывные случайные переменные — это такие случайные переменные, которые могут принимать бесконечное количество значений. Например, скорость, время, расстояние, рентабельность активов. Единица измерения может здесь представлять собой бесконечно малую величину. Для примера рассмотрим доход от какой-либо ценной бумаги. Как мы выше уже отметили, это доходность — непрерывная случайная величина. Количество возможных значений доходности может быть бесконечно велико. Например, изменение цены актива со 105 единиц до 109 даст доходность, равную 3,8% или 3,81%, или 3,8095% в зависимости от количества знаков после запятой, допускаемого нами при измерении доходности. В этих обстоятельствах нет никакого смысла в попытках нахождения вероятности значения доходности равной, скажем, 3,81%. Имеет смысл только нахождение вероятности того, что случайная переменная примет значение на каком-то определенном интервале, скажем, между 3,81% и 3,82%.

Очевидно, что найти ожидаемую величину непрерывной случайной переменной путем сложения, как в случае с дискретными переменными, трудно, поскольку пришлось бы искать сумму бесконечного множества вероятностей. Для преодоления этой проблемы мы должны определить непрерывную случайную величину не путем суммирования функции частот вероятностей, которая дает определенные вероятности, а путем интегрирования так называемой **функции плотности вероятностей** (см. гл. 2).

Таким образом, для случайной переменной (X) получаем

$$\int_{-\infty}^{+\infty} f(X)\mathrm{d}X = 1, \tag{4.9}$$

где f — функция плотности вероятностей. Из гл. 3 мы знаем, что выражение вида (4.9) характеризует площадь под кривой. Фактически функция плотности вероятностей является функцией, представляющей вероятность. На рис. 4.2 приведен пример такого графика (заметьте, что площадь под кривой графика функции плотности вероятностей должна равняться единице).

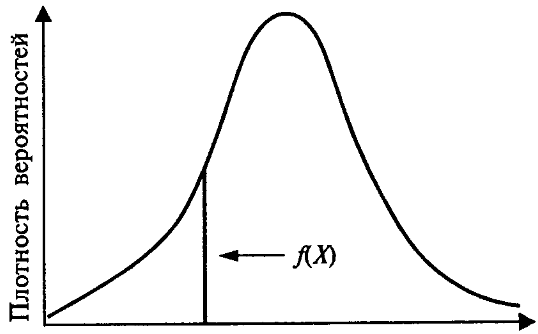

Рис. 4.2

МАТЕМАТИЧЕСКИЕ ДЕЙСТВИЯ НАД СЛУЧАЙНЫМИ ПЕРЕМЕННЫМИ

Умножение случайной величины

При умножении случайной переменной на постоянную величину вероятности остаются неизменными, но возможные результаты умножаются на данное число. Например, если X — дискретная случайная переменная, то $2X$ определяется тем же распределением, что и X, за исключением того, что значения возможных результатов удваиваются (вероятности остаются неизменными).

Таким образом, если X определяется следующим образом:

Значение (r)	0	1	2
Вероятность того, что $X = r$	1/4	1/2	1/4

то $2X$ определяется так:

Значение (r)	0	2	4
Вероятность того, что $X = r$	1/4	1/2	1/4

Сложение двух независимых случайных величин

Теперь рассмотрим сложение случайных переменных Y и X. Предположим, что X определяется так:

Значение (r)	0	1	2
Вероятность того, что $X = r$	1/4	1/2	1/4

Y определяется следующим образом:

Значение (r)	4	5
Вероятность того, что $X = r$	1/2	1/2

Известно, что Y и X независимы друг от друга. Проверяя возможные значения и их вероятности, мы видим, что $Y + X$ задаются следующим образом:

Значение (r)	4	5	6	7
Вероятность того, что $X = r$	1/8	3/8	3/8	1/8

Посмотрим, почему.

Во-первых, возможные значения:

$$0 + 4 = 4,\ \ 0 + 5 = 5,\ \ 1 + 4 = 5,\ \ 1 + 5 = 6,\ \ 2 + 4 = 6,\ \ 2 + 5 = 7.$$

Таким образом, возможные значения: четыре (один способ), пять (два способа), шесть (два способа) и семь (один способ).

Для значения суммы переменных, равной 4, т.е. $0 + 4$, вероятность будет $1/4 \cdot 1/2 = 1/8$.

Значение, равное 5, может быть получено двумя способами: $0 + 5$ с вероятностью $1/4 \cdot 1/2 = 1/8$ и $1 + 4$ с вероятностью $1/2 \cdot 1/2 = 1/4$. Складывая вероятности, получаем вероятность $1/8 + 1/4 = 3/8$.

Таким же образом находим вероятность 6, получаемую двумя способами, которая равна 3/8, и вероятность 7, равная 1/8.

Заметьте, что для случайных переменных $2X$ не то же самое, что и $X + X$! Чтобы убедиться в этом, возьмем приведенную выше переменную X. Возможны ее значения, равные 0, 1, 2. Теперь сложим эту переменную с точно такой же переменной, т.е. $X + X$. Возможны результаты: $0 + 0 = 0$, $1 + 0 = 1$, $0 + 1 = 1$, $1 + 1 = 2$, $2 + 0 = 2$, $0 + 2 = 2$, $1 + 2 = 3$, $2 + 1 = 3$ и $2 + 2 = 4$. Вероятности будут:

0	1	2	3	4
1/16	1/4	3/8	1/4	1/16

МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ И ДИСПЕРСИЯ СЛУЧАЙНОЙ ПЕРЕМЕННОЙ

Математическое ожидание

Математическое ожидание дискретной случайной переменной определяется как

$$E(X) = \sum_r X \cdot \mathrm{prob}(X = r), \tag{4.10}$$

где E — оператор ожидания.

Например, случайная переменная X, определенная выше, имеет математическое ожидание, равное 1:

$$E(X) = 0 \cdot 1/4 + 1 \cdot 1/2 + 2 \cdot 1/4 = 1.$$

Мы видим, что математическое ожидание случайной переменной является вероятностно взвешенным средним всех ожидаемых значений случайной переменной. Однако следует отметить, что математическое ожидание не обязательно должно быть одним из возможных значений дискретной случайной переменной. Рассмотрим пример шестигранной игральной кости. В этом случае математическое ожидание будет равно 3,5:

$$\left(1 \cdot \frac{1}{6}\right) + \left(2 \cdot \frac{1}{6}\right) + \left(3 \cdot \frac{1}{6}\right) + \left(4 \cdot \frac{1}{6}\right) + \left(5 \cdot \frac{1}{6}\right) + \left(6 \cdot \frac{1}{6}\right) = 3{,}5.$$

Это потому, что математическое ожидание является вероятностно взвешенным средним возможных результатов и подобно среднему арифметическому (см. гл. 2) может дать результат, не совпадающий ни с одним из возможных результатов.

Заметьте, что математическое ожидание группы случайных переменных является линейной комбинацией математических ожиданий каждой отдельной случайной переменной:

$$E(X + Y) = E(X) + E(Y). \tag{4.11}$$

Аналогично существует еще одно свойство случайной переменной, которое соответствует значению ожидаемого отклонения в случае большого числа испытаний. Оно называется дисперсией случайной переменной и определяется как

$$\mathrm{var}(X) = E(X - E(X))^2. \tag{4.12}$$

Для иллюстрации этой формулы с помощью числового примера рассмотрим ожидаемую рентабельность какого-либо актива. Предположим, что менеджер по инвестициям предполагает в будущем три возможных варианта развития экономической ситуации: высокий рост, отсутствие роста и спад. Вероятности каждого из изменений оцениваются равными 0,25, 0,5 и 0,25 соответственно. Читатель уже заметил, что это пример субъективной вероятности.

Менеджер по инвестициям ожидает получить 20%-ную доходность с данного актива в случае высокого роста, 10% — в случае отсутствия роста и —4% — в случае спада.

Поскольку математическое ожидание равно сумме произведений каждого из возможных значений на его вероятность, то ожидаемая доходность актива будет составлять:

$$E(r) = (0{,}20 \cdot 0{,}25) + (0{,}10 \cdot 0{,}50) + (0{,}04 \cdot -0{,}25) = 0{,}09 = 9\%.$$

Это может быть представлено в общем виде:

$$E(r) = \sum_{i=1}^{n} r_i P_i \,. \tag{4.13}$$

Поскольку дисперсия случайной переменной равна сумме квадратов отклонений возможных результатов от математического ожидания, умноженных на соответствующие вероятности, то

$$\operatorname{var}(r) = \sum_{i=1}^{n} \left(r_i - \bar{r}\right)^2 P_i \,. \tag{4.14}$$

Отсюда в нашем примере дисперсия ожидаемой доходности будет равна

$$\operatorname{var}(r) = (0{,}20 - 0{,}09)^2\, 0{,}25 + (0{,}10 - 0{,}09)^2\, 0{,}50 + (-0{,}04 - 0{,}09)^2\, 0{,}25 =$$

$$= 0{,}003025 + 0{,}00005 + 0{,}004225 = 0{,}0073.$$

Полученный результат выражается в процентах в квадрате, что само по себе не очень удобно. Поэтому обычно находят квадратный корень из дисперсии, называемый средним квадратическим отклонением. Квадратный корень из 0,0073 равен

0,085, или **8,5%**.

Таким образом, мы видим, что если доходность актива — случайная переменная, то ожидаемый уровень доходности равен математическому ожиданию, и риск, измеряемый вариацией ожидаемых результатов, описывается дисперсией или средним квадратическим отклонением.

Если мы хотим вычислить дисперсию линейной комбинации случайных переменных, то мы должны будем учесть, являются ли эти переменные независимыми или нет, т.е. мы должны посмотреть на их взаимодействие, представленное их ковариацией. В результате получаем

$$\mathrm{var}(aX + bY) = a^2\mathrm{var}(X) + b^2\mathrm{var}(y) + 2ab \cdot \mathrm{cov}(X,Y), \tag{4.15}$$

где $\mathrm{cov}XY$ — это ковариация X и Y, которая рассчитывается следующим образом:

$$\mathrm{cov}(X, Y) = E\,[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y). \tag{4.16}$$

Говорят, что две случайные переменные некоррелированы, если их ковариация равна нулю. В этом случае $\mathrm{var}(X + Y)$ = равно $\mathrm{var}(X) + \mathrm{var}(Y)$.

Теперь рассмотрим математическое ожидание и ожидаемую дисперсию на примере ожидаемой доходности портфеля активов и его дисперсии и среднего квадратического отклонения. Рассмотрим пример портфеля, состоящего из двух активов.

ПРИМЕНЕНИЕ ДИСКРЕТНЫХ СЛУЧАЙНЫХ ПЕРЕМЕННЫХ: РАСЧЕТЫ ДОХОДНОСТИ И СРЕДНЕГО КВАДРАТИЧЕСКОГО ОТКЛОНЕНИЯ ПОРТФЕЛЯ ЦЕННЫХ БУМАГ

Доходность портфеля

Доходность портфеля характеризуется средневзвешенной доходностью его составляющих. Например, доходность портфеля из двух активов рассчитывается следующим образом:

$$E(R_p) = W_aE(r_a) + W_bE(r_b), \tag{4.17}$$

где $E(R_p)$ — доходность портфеля;
W_a — удельный вес актива A;
$E(r_a)$ — доходность актива A;
W_b — удельный вес актива B;
$E(r_b)$ — доходность актива B.

Переменные $E(r_a)$ и $E(r_b)$ — ожидаемые доходности соответствующих активов. Удельные веса равняются части общей стоимости портфеля, инвестированной в данный актив.

Для наглядности рассмотрим это на следующем примере. Предположим, что доходности A и B соответственно равны 10% и 12% и портфель состоит на 60% из актива A и на 40% — из актива B. Ожидаемая доходность портфеля будет равна:

$$0{,}6 \cdot 0{,}1 + 0{,}4 \cdot 0{,}12 = 0{,}108 = 10{,}8\%.$$

Уравнение (4.17) может быть представлено в общем виде:

$$E(R_p) = \sum_{i=1}^{n} W_i E(r_i). \qquad (4.18)$$

Среднее квадратическое отклонение портфеля

Риск портфеля не может быть измерен нахождением средневзвешенной из дисперсий доходности каждого актива. Причина в том, что измеряя риск портфеля, мы должны учесть не только изменения доходности каждого из активов, но и то, в какой степени доходности двух активов изменяются вместе — степень их взаимодействия, или ковариацию. Эта совместная изменчивость измеряется ковариацией, или иначе, корреляцией доходности пар активов.

Из гл. 2 мы знаем, что коэффициент корреляции может принимать значения от + 1 (полная положительная корреляция) до —1 (полная отрицательная корреляция). Положительная корреляция означает, что доходности каждой пары активов в основном изменяются в одном направлении. Это соотношение тем сильнее, чем ближе коэффициент корреляции к + 1. Отрицательная корреляция показывает, что доходности изменяются в противоположных направлениях, при этом соотношение становится сильнее по мере того, как коэффициент приближается к —1.

Важность корреляции при построении портфеля объясняется следующим. Если доходности составляющих портфель активов обладали бы идеальной положительной корреляцией, то мы не получили бы никаких выгод от диверсификации, поскольку рентабельности всех активов изменялись бы в одном и том же направлении, в одно и то же время, в той же степени. И наоборот, если пары активов обладают полной отрицательной корреляцией, т.е. изменяются в противоположных направлениях, в одно и то же время и в одинаковой степени, то изменение доходности всего портфеля будет близко к нулю.

В гл. 2 мы узнали, что среднее квадратическое отклонение портфеля из двух инструментов рассчитывают так:

$$\sigma_p = \sqrt{W_A^2\sigma_a^2 + W_B^2\sigma_b^2 + W_A W_B(\rho_{ab}\sigma_a\sigma_b)}\,, \tag{4.19}$$

где σ_p — среднее квадратическое отклонение портфеля;
W_A и W_B — удельные веса активов A и B;
$\sigma^2{}_a$ и $\sigma^2{}_b$ — дисперсии доходностей активов A и B;
σ_a и σ_b — средние квадратические отклонения доходностей A и B;
ρ_{ab} — корреляция доходностей A и B;
$(\rho_{ab}\sigma_a\sigma_b)$ — **ковариация** доходностей активов A и B;

Заметьте, что коэффициент корреляции равен $\mathrm{cov}AB/(\sigma_a\sigma_b)$. Выраженное при помощи математического ожидания уравнение (4.19) становится следующим:

$$\sigma^2_{(\mathrm{port})} = W_A^2 E\left(\left(r_a - E(r_a)\right)^2\right) + W_B^2 E\left(\left(r_b - E(r_b)\right)^2\right)$$
$$+ 2W_A W_B \rho_{ab}\sqrt{E\left(\left(r_a - E(r_a)\right)^2\right)E\left(\left(r_b - E(r_b)\right)^2\right)}. \tag{4.20}$$

Среднее квадратическое отклонение портфеля равняется квадратному корню из дисперсии.

Математическое ожидание и дисперсия непрерывных случайных величин

Математическое ожидание непрерывной случайной переменной определяется следующим образом:

$$E(X) = \int_{-\infty}^{+\infty} X f(X)\,\mathrm{d}X\,. \tag{4.21}$$

Если X ограничен на каком-либо интервале, скажем, между y и z, то ожидаемое значение будет:

$$E(X) = \int_{y}^{z} Xf(X)\mathrm{d}X\,. \qquad (4.22)$$

Дисперсия и среднее квадратическое отклонение соответственно будут равны:

$$\mathrm{var}(X) = \int_{-\infty}^{+\infty} (X-\mu)^2 f(X)\,; \qquad (4.23)$$

$$\mathrm{SD}(X) = \sqrt{\int_{-\infty}^{+\infty} (X-\mu)^2 f(X)}\,. \qquad (4.24)$$

Поскольку разница между дискретными и непрерывными переменными существенна для построения теоретических моделей, иногда мы можем использовать непрерывные переменные при моделировании дискретных ситуаций, и наоборот. Например, рассмотрим цену некой акции на фондовом рынке в полдень на следующий день. Ясно, что существует только дискретное количество возможных значений (цены акций выражаются в фунтах, пенсах и только иногда в долях пенсов). Тем не менее, мы можем с успехом применять непрерывную случайную переменную при моделировании поведения цены акции.

Наиболее важные характеристики распределений вероятностей в финансах

Оставшаяся часть этой главы посвящена анализу различных распределений вероятностей, применимых при оценке поведения рентабельности активов при условии соответствующих допущений. Начнем с двух непрерывных распределений — нормального и логнормального. Затем рассмотрим два дискретных распределения — биномиальное и Пуассона. Закончим рассмотрение группой других непрерывных распределений, в том числе и распределением Парето—Леви. Объясним наиболее желательные характеристики распределений с точки зрения финансового аналитика.

Вообще финансовые аналитики используют распределения вероятностей в качестве основы для предсказания распределения цены и рентабельности активов. Следовательно, желательно, чтобы распределение обладало такими характеристиками:

- стационарностью;
- стабильностью;
- конечной дисперсией.

Под стационарностью понимают то, что параметры распределения вероятностей не изменяются во времени.

Наибольший интерес в финансовых активах для нас представляет цена. Однако в анализе цен финансовых активов мало пользы от распределения вероятностей, поскольку поведение цен не имеет тенденции быть стационарным. Причина может быть в том, что со временем цены активов изменяются все вместе, например, цены акций со временем растут. Таким образом, математическое ожидание и среднее квадратическое отклонение цен обычно будут выше в данном году, чем в предыдущем. Более того, цены активов могут теоретически возрастать до бесконечности, но не могут упасть ниже нуля, и поскольку цены растут, то будет наблюдаться тенденция правосторонней скошенности распределения цен.

Распределения вероятностей изменений цен также не имеют тенденции быть стационарными, потому что абсолютная величина изменения цены актива также может изменяться по мере того, как цена изменяется. Однако размер изменений в процентах, доходность актива в процентах обычно не зависят от уровня цен активов. Следовательно, доход в процентном выражении может и не меняться во времени только из-за того, что цена актива изменилась.

Стационарные распределения вероятностей доходности активов позволяют производить вероятностную оценку будущих доходов. К тому же историческая информация о доходах может быть основанием для оценки неопределенности, связанной с получением будущих доходов, и отсюда может быть основой для расчета будущих рисков. Более углубленно стационарность обсуждается в гл. 7.

Другое желаемое свойство распределения состоит в том, чтобы линейная комбинация двух распределений одного типа давала в результате распределение такого же типа. Например, линейная комбинация двух нормальных распределений дает также нормальное распределение, разве что только с параметрами, отличными от первых двух. Это свойство распределений называется стабильностью. Для демонстрации важности стабильности распределений рассмотрим распределение вероятностей рентабельности актива в течение двух однодневных периодов. Было бы желательно, чтобы распределение доходности на протяжении двухдневного периода оставалось того же типа.

Конечность дисперсии также важна, поскольку без нее эффективные оценки параметров распределения, сделанные на основе выборки, не будут приближаться к действительным статистическим параметрам генеральной совокупности по мере того, как размер выборки будет увеличиваться. Более того, мы приняли среднее квадратическое отклонение за меру риска. Оно не может быть определено, если дисперсия не является конечной. Присутствие значительного количества выделяющихся значений приводит к тому, что оценки параметров будут изменяться от выборки к выборке.

Сказанное выше позволяет определить те распределения, которые лучше всего описывают распределение вероятностей доходности активов.

Нормальное распределение

Мы отметили в гл. 2, что наиболее широко из распределений частот используется **нормальное распределение**, или **распределение Гаусса**. Отсюда вытекает то обстоятельство, что наиболее широко используемым распределением вероятностей является нормальное распределение. Это распределение непрерывное, но часто применяется при моделировании дискретных случайных переменных.

Кривая нормального распределения имеет форму симметричного колокола, как это изображено на рис. 4.3.

Это распределение полностью определяется **средней арифметической** и **средним квадратическим отклонением**, что делает этот тип распределения весьма привлекательным. Средняя арифметическая указывает на расположение середины колокола, а среднее квадратическое отклонение показывает, насколько "колокол" растянут в стороны.

Если переменная подчиняется закону нормального распределения, то 68,27% всех наблюдений попадут в интервал плюс-минус среднее квадратическое отклонение от средней. Более того, 95,45% всех наблюдений попадают в интервал плюс-минус удвоенное среднее квадратическое отклонение и 99,73% — в пределах плюс-минус утроенное среднее квадратическое отклонение от величины средней.

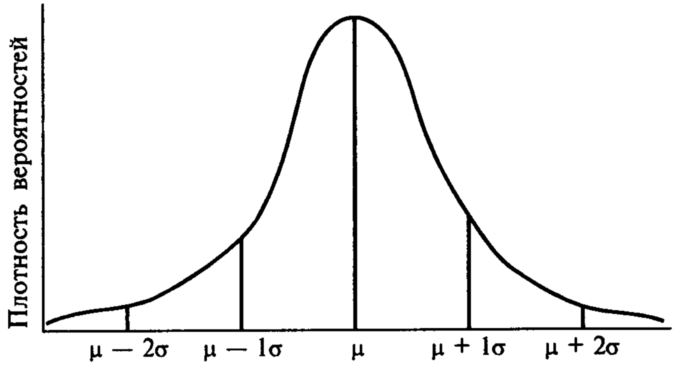

Рис. 4.3

Так как кривая нормального распределения по сути является графиком **функции нормальной плотности частот**, то уравнение, определяющее нормальное распределение, будет:

$$Y = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \tag{4.25}$$

где μ — средняя арифметическая распределения;
σ — среднее квадратическое отклонение;
$\pi = 3{,}1415926$;
$e = 2{,}71828$.

Таким образом, если выражение (4.25) определяется отдельно для всех возможных значений X, то при нанесении полученных точек на график в результате получится нормальная кривая. Мы рассмотрим это более подробно, когда займемся стандартизованными функциями плотности вероятностей.

Центральная предельная теорема

Теоретическое обоснование того, что случайные переменные подчиняются закону нормального распределения, основывается на **центральной предельной теореме**. Теорема утверждает, что математическое ожидание большого числа независимых выборок будет нормально распределено вне зависимости от действительного распределения данных при условии конечной дисперсии.

Для демонстрации центральной предельной теоремы рассмотрим ежеминутные изменения индекса FTSE 100. Допустим, что в течение следующей минуты индекс либо возрастет, либо упадет на один пункт с одинаковой вероятностью. Эта вероятность описывается следующим образом:

Возможный результат	—1	+ 1
Вероятность	0,5	0,5

В следующую минуту индекс может вырасти или упасть на один пункт с равной вероятностью, а потому возможные результаты в конце двухминутного периода будут определяться следующим образом:

Возможный результат	—2	0	+ 2
Вероятность	0,25	0,5	0,25

За третью минуту индекс может измениться таким образом: если он был равен —2, то может возрасти до —1 или упасть до —3. Если он был равен 0, то может возрасти до + 1 или упасть до —1. Если же индекс был равен + 2, то может стать + 3 или + 1. Возможные результаты будут такими:

Возможный результат	—3	—1	+1	+3
Вероятность	1/8	3/8	3/8	1/8

Через четыре минуты результаты будут следующими:

Возможный результат	—4	—2	0	+2	+4
Вероятность	1/16	1/4	3/8	1/4	1/16

Можно увидеть, что уже после четырехминутных интервалов вероятность получения крайних результатов, например, —4 или + 4, становится довольно низкой, а вероятность получения результата, близкого к математическому ожиданию, сравнительно высока. Если бы мы продолжили процесс еще на некоторый промежуток времени, то увидели бы, что функция вероятности принимает колоколообразную форму нормального распределения.

Заметьте, что приведенный выше пример всего лишь описывает тенденцию собранных данных. Центральный предельный процесс описывает поведение **средних** собранных данных.

Стандартизованная функция плотности вероятностей нормальной кривой

Стандартизованная (или нормированная) переменная, обычно обозначаемая *Z*, — это переменная со средней арифметической, равной нулю, и средним квадратическим отклонением, равным единице, вычисляется по следующей формуле:

$$z = \frac{(x - \mu)}{\sigma}. \tag{4.26}$$

Если случайная переменная нормально распределена, то стандартизованная переменная z будет нормально распределена с математическим ожиданием, равным нулю, и средним квадратическим отклонением 1,0. Поскольку нормальное распределение — это распределение вероятностей, то площадь под кривой должна быть равна 1. Кривая изображена на рис. 4.4.

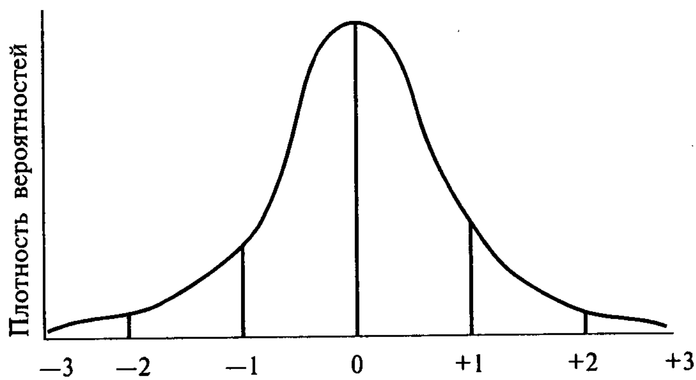

Рис. 4.4

Высота от любой точки кривой до оси абсцисс представляет собой **функцию плотности стандартизованной кривой нормального распределения**. Уравнение **функции плотности** стандартизованного нормального распределения выглядит следующим образом:

$$Y = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}. \tag{4.27}$$

Зависимость между стандартизованной и не стандартизованной функциями плотности нормальной кривой, данной в уравнении (4.25), становится очевидной.

Площадь любого вертикального сегмента под кривой представляет вероятность того, что нормально распределенная случайная переменная примет значение на интервале, ограниченном данным сегментом. Это показано на рис. 4.5.

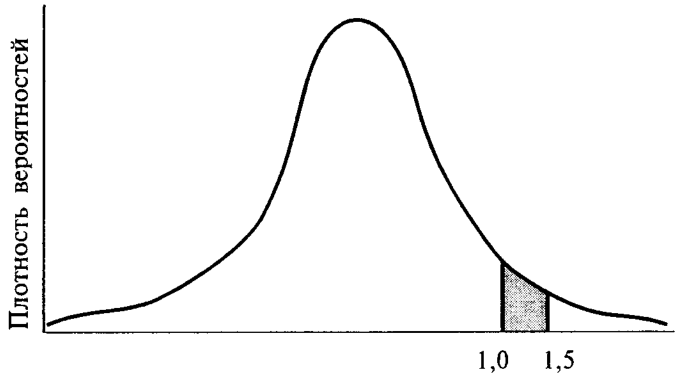

Рис. 4.5

Нахождение площадей под кривой нормального распределения с помощью таблиц

В таблицах приведены площади под стандартной кривой нормального распределения — так называемого стандартизованного нормального распределения. Обычно в таблице при помощи диаграммы указывается, площадь каких сегментов приводится. Но при отсутствии диаграммы, взглянув на данные, приведенные в таблице, можно легко увидеть, приведены ли в таблице площади справа или слева от X для различных значений. Обычно таблицы содержат площади сегментов под кривой справа от X для значений X больше нуля. Прочие площади можно вывести, используя симметричность кривой. Площади (т.е. вероятности) могут быть найдены как разность двух табличных величин для $X =$ a и для $X = b$.

Для определения вероятности того, что нормально распределенная переменная примет значение в пределах какого-либо обусловленного интервала, во-первых, надо найти значение z для верхней и нижней границ интервала.

Во-вторых, используя z-таблицы, найти величину площади под кривой между $z = 0$ и $z =$ нижней границе сегмента.

В-третьих, найти площадь под кривой между $z = 0$ и $z =$ верхней границе интервала. Разность этих двух площадей и будет равна вероятности того, что случайная переменная будет находиться в заданном интервале.

Рассмотрим это на следующем примере.

Нам нужно узнать вероятность того, что некий актив, доходность которого нормально распределена, принесет доход от 4,9% до 5,0%. Средняя величина дохода по этому активу равна 4%, а среднее квадратическое отклонение составляет 1%. Не забудьте, что для непрерывных переменных мы рассматриваем вероятность того, что переменная примет значение в данном интервале, а не в данной точке.

Значения z для 4,9% и 5,0% соответственно будут:

$$z_{4,9} = \frac{4,9 - 4}{1,0} = 0,9; \qquad z_5 = \frac{5 - 4}{1,0} = 1,0\,.$$

Из таблицы стандартизованного нормального распределения находим, что площадь между $z = 0$ и $z = 1$ равна 0,3413, а площадь между $z = 0$ и $z = 0,9$ равна 0,3159. Таким образом, площадь стандартизованного нормального распределения между $z =$ 5% и $z =$ 4,9% равняется 0,3413 — 0,3159 = 0,0254. Это можно интерпретировать так: вероятность того, что доходность актива будет находиться между 4,9% и 5,0%, составляет 2,54%. Этот пример изображен на рис. 4.6.

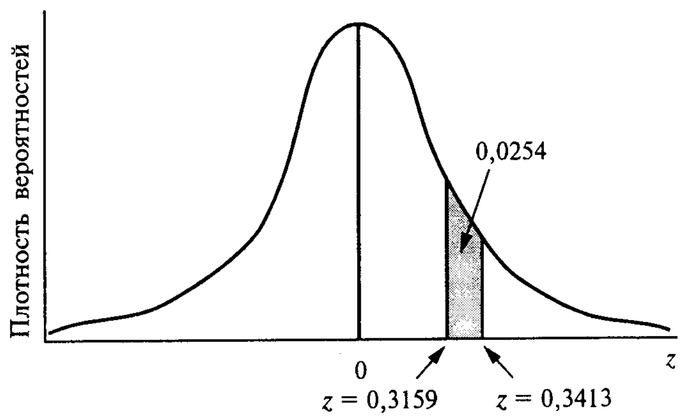

Рис. 4.6

Этот метод можно использовать для определения вероятности того, что случайная величина, например цена актива, примет значение выше или ниже определенной величины. Предположим, например, мы знаем, что ежедневная доходность некой ценной бумаги нормально распределена с математическим ожиданием, равным 0,5%, и средним квадратическим отклонением 0,1%, и мы хотим узнать вероятность того, что ежедневная доходность будет больше 0,525%. Сначала мы найдем значение нормированной величины z:

$$z = \frac{0{,}525 - 0.5}{0{,}1} = 0{,}25\,.$$

Теперь перейдем к таблице и определим площадь под кривой между $z = 0$ и $z = 0{,}25$, которая равна 0,0987. Принимая во внимание, что кривая нормального распределения симметрична и вероятность падения z ниже нуля равна 0,5, определяем вероятность того, что $z < 0{,}25 : 0{,}5 + 0{,}0987 = 0{,}5987$. Соответственно вероятность того, что $z > 0{,}25$ (или доходность выше 0,525%) составит 0,4013, или 40,13% (1 — 0,5987 = 0,4013).

Обычно площадь под кривой определяется интегрированием. К сожалению, для функции нормального распределения не существует первообразной. До изобретения компьютеров это представляло бы собой затруднение, если бы не тот факт, что при стандартизации переменной площадь под кривой остается неизменной, что позволяет использовать табличные данные для стандартизованного распределения. При помощи компьютера можно применить еще несколько удобных способов.

Существуют два основных метода нахождения площади под кривой при помощи компьютера. Первый использует метод численного интегрирования, такого, как правило трапеции и правило Симпсона. Другой метод подразумевает использование многочисленной функции, которая приближается к функции, определяемой площадью под кривой. Оба метода рассмотрены в гл. 8, которая посвящена численным методам.

Логнормальное распределение

Как мы уже отметили выше, в соответствии с центральной предельной теоремой процесс сложения изменений приводит к нормальному распределению. Что же произойдет в результате процесса умножения?

Возьмем, например, относительное изменение цены ценной бумаги за период времени Δt. Пусть $S(t)$ — цена этой ценной бумаги в момент времени t и $S(\ t + \Delta t)$ — цена в момент времени $t + \Delta t$, тогда относительное изменение цены по истечении периода Δt будет равно:

$$\frac{S(t+\Delta t)}{St}. \qquad (4.28)$$

Но если мы рассмотрим изменения цены на протяжении некоторого количества (n) промежутков времени, dt, где $\Sigma dt = \Delta t$, то сможем составить уравнения относительных изменений цены для каждого из таких меньших интервалов. Мы видим, что

$$\frac{S(t+\Delta t)}{S(t)} = \frac{S(t+\Delta t)}{S(t)} \cdot \frac{S(t+2\Delta t)}{S(t+\Delta t)} \cdot \frac{S(t+3\Delta t)}{S(t+2\Delta t)} \cdot \ldots \cdot \frac{S(t+n\Delta t)}{S(t+(n-1)\Delta t)}. \qquad (4.29)$$

Таким образом, относительное изменение цены является результатом процесса умножения. Это может быть выражено суммированием, если использовать натуральные логарифмы относительных изменений цен следующим образом:

$$\ln\frac{S(t+\Delta t)}{S(t)} = \ln\frac{S(t+\Delta t)}{S(t)} + \ln\frac{S(t+2\Delta t)}{S(t+\Delta t)} + \ln\frac{S(t+3\Delta t)}{S(t+2\Delta t)} + \ldots$$

$$\ldots + \ln\frac{S(t+\Delta t)}{S(t+(n-1)\Delta t)}. \qquad (4.30)$$

Предположим, что каждое из отношений цен на протяжении короткого отрезка времени dt было случайной переменной, независимой и идентично распределенной (IID), скажем, X_1, X_2 и т.д., где X_i — идентичная копия случайной переменной X. Тогда отношение $S_{(t+T)}/S_{(t)}$ также будет случайной переменной, скажем Y.

Предположим, что $\ln(X_1)$, $\ln(X_2)$ и т.д. являются IID, тогда мы сможем применить **центральную предельную теорему** для того, чтобы предположить, что $\ln(S_{(t)}/S_{(t-1)})$ приблизительно нормально распределен. Центральная предельная теорема гласит, что если мы рассматриваем большую случайную выборку, то средняя величина ее будет нормально распределена. Таким образом, когда мы разделяем период времени на большое число промежутков (больше 30), с чем мы имеем дело, когда рассматриваем время как непрерывное, то сумма натуральных логарифмов в правой части уравнения будет нормально распределена, и соответственно $\ln(S_t/S_{t-1})$ будет тоже нормально распределено.

Переменная называется **логнормально распределенной**, если **натуральный логарифм** ее **нормально** распределен. Следовательно, если $\ln(S_{t+\Delta t}/S_{(t)})$ нормально распределен, то $S_{t+\Delta t}/S_{(t)}$ должно быть распределено логнормально.

Это очень привлекательная модель распределения отношений цен ценных бумаг, потому что, если цена растет, то отношение цен будет больше единицы, если падает — то отношение цен будет меньше единицы, но оно никогда не принимает отрицательного значения. Теперь рассмотрим график функции логнормального распределения на рис. 4.7. На рисунке видно, что логнормальное распределение вытянуто вправо, но не имеет отрицательных значений. Это совместимо с возможным распределением цен ценных бумаг, поскольку они не могут упасть ниже нуля, и только очень немногие из наблюдений могли быть очень высоки.

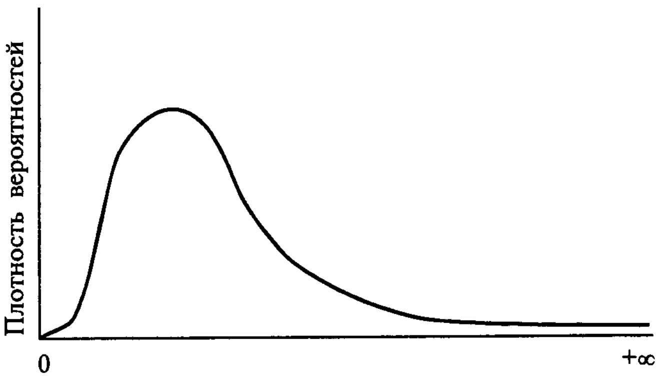

Рис. 4.7

Переменная не может принимать отрицательные значения, и вероятность очень высоких значений приближается к нулю, как это и можно ожидать от переменной, описывающей относительное изменение цены ценной бумаги. Логнормальное распределение очень часто используется для моделирования такого рода случайных переменных, в основе которых лежит процесс умножения.

Мы отметили в гл. 1, что натуральный логарифм относительной цены $\ln(S_t/S_{t-1})$ — это непрерывно начисляемая доходность ценной бумаги S за период времени t—(t—1). Если относительные цены логнормально распределены, непрерывно начисляемая доходность $\ln(S_t/S_{t-1})$ будет также логнормально распределена.

К этому распределению мы вернемся в гл. 10, где используем его для оценки опционов. Следующие результаты доказаны в приложении к этой главе. Если

$$\ln(Y) \sim N(\mu, \sigma^2), \tag{4.31}$$

то

$$E(Y) = e^{\mu + \frac{\sigma^2}{2}}. \tag{4.32}$$

Для удобства это часто записывается так:

$$\exp(\mu + \sigma^2/2); \tag{4.33}$$

$$\mathrm{var}(Y) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1). \tag{4.34}$$

Биномиальное распределение

Одним их наиболее важных **дискретных распределений** в финансах является биномиальное распределение. Для формирования **биномиального распределения** случайная переменная должна отвечать следующим четырем условиям.

1. Переменная может принимать только два значения в данный момент или в результате данного события. Каждое из этих событий или моментов времени называется биномиальной **попыткой**. Два возможных результата называются "**успех**" — в случае благоприятного результата и "**неудача**" — в случае неблагоприятного.
2. Для каждой последовательности попыток вероятность успеха и неудачи постоянна.
3. Все попытки идентичны.
4. Все попытки независимы.

Биномиальная случайная переменная X — это число успехов в результате некоторого количества независимых попыток n. Результат испытания записывается как успех или неудача, и вероятность успеха в каждой из попыток равна p. Мы говорим X~бином (n, p). X, число успехов в n попытках, может принимать значения 0, 1, 2, ..., n. Таким образом, для n попыток может быть n + 1 результат. Один результат будет включать n успехов, другой — n—1 успехов, и т.д. и последний — вообще без успехов.

Это показано на рис. 4.8, где переменная S может возрастать в u раз (u >1) или уменьшаться в d раз ($0 < d < 1$). В данном случае мы имеем две биномиальные попытки и три результата. Результат Su^2 — это результат двух успехов, т.е. $j = 2$. Результат $Sud = Sdu$, т.е. $j = 1$ — это результат одного успеха. И результат Sd^2 — это результат нулевого количества успехов, т.е. $j = 0$.

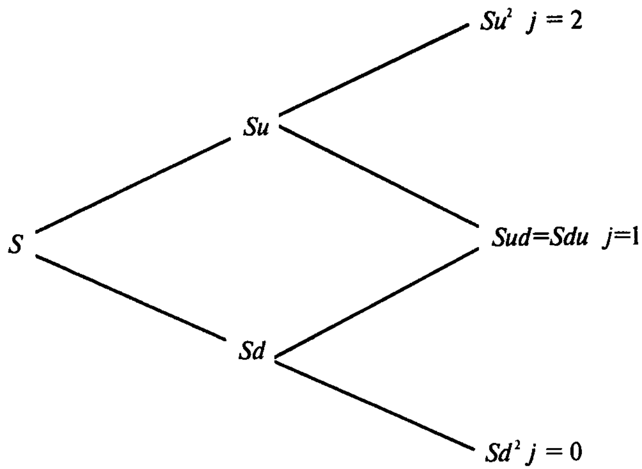

Рис. 4.8

Биномиальное распределение дает вероятности X для каждого из результатов. Вероятность достижения каждого из них зависит от:

1) вероятности успеха, т.е. p;
2) общего числа способов достижения результата.

Чтобы объяснить вероятности, связанные с этим "деревом", рассмотрим вероятность получения двух последовательных успехов. Предположим, что вероятность успеха равна 0,5 и, учитывая, что попытки независимы, вероятность двух успехов $P(Su^2)$ будет равна $0{,}5 \cdot 0{,}5 = 0{,}25$. Таким же образом вероятность успеха, следующего за неудачей, также будет 0,25.

Вероятность успеха, следующего за неудачей, равна вероятности неудачи, следующей за успехом. Отсюда существует два способа достижения одного успеха. Каждый способ имеет вероятность 0,25, и значит, вероятность результата $j = 1$ будет 0,25 + + 0,25 = 0,5.

Для нахождения вероятности в более сложных примерах, например, если вероятности роста и падения не равны или очень много попыток, необходимо знать количество результатов, имеющих определенное число успехов. В рассматриваемом примере мы хотим узнать, сколько результатов соответствуют одному успеху, сколько — двум и сколько — нулю. Можно рассчитать число j — количество успехов для заданного числа n биномиальных попыток, используя формулу:

$$^{n}C_{j} = \frac{n!}{j!(n-j)!}, \tag{4.35}$$

где $^{n}C_{j}$ — количество способов достижения j успехов из n попыток;

n — число биномиальных попыток;

j — число успехов.

Знак "!" после n и j обозначает факториал, т.е. n умножаем на $(n - 1) \cdot (n-2) \cdot (n - 3)$ и т.д. Например, для $n = 4$ найдем $n!$

$$4 \cdot 3 \cdot 2 \cdot 1 = 24,$$

$j!$ рассчитывается подобным же образом.

Для демонстрации использования формулы (4.35) определим число способов получения одного успеха в данном примере:

$$\frac{2!}{1!(2-1)!} = \frac{2 \cdot 1}{1 \cdot 1} = 2.$$

Этот результат можно легко проверить. В рассматриваемом примере *Sud* и *Sdu* приводят к одному успеху каждый.

Теперь, зная число результатов с j успехами, определим вероятность получения j успехов с n попыток, поскольку это даст нам вероятность, связанную с количеством результатов, имеющих j успехов. Эта вероятность находится по формуле

$$p(x = j) = \frac{n!}{j!(n-j)!} p^{j} (1 - p)^{(n-j)}, \tag{4.36}$$

где p — вероятность успеха и $1-p$ — вероятность неудачи.

Для примера вспомним, что мы предположили, что $p = 0,5$ и $(1 - p) = 0,5$. Вероятность конкретного результата (Su^2) после двух попыток отсюда равна:

$$p(Su^2) = \frac{2!}{2!(2-2)!}0,5^2(1-0,5)^{(2-2)} = \frac{2}{2}0,25 = 0,25.$$

Таким образом, существует всего лишь один результат с $j = 2$ числом успехов, и он имеет вероятность 0,25.

Теперь рассчитаем вероятности Sud и Sd^2.

Для Sud

$$p(Sud) = \frac{2!}{1!(2-1)!}0,5^1(1-0,5)^{2-1} = 0,5.$$

Существуют два результата, дающих $j = 1$ успеху, каждый имеет вероятность 0,25. Два таких результата имеют вероятность 0,5.

Для Sd^2

$$p(Sd^2) = \frac{2!}{0!(n-0)!}0,5^0(1-0,5)^{(2-0)} = 0,25.$$

Складывая все полученные вероятности, получаем единицу, как и следовало ожидать.

Математическое ожидание, дисперсия и среднее квадратическое отклонение распределения X составят:

$$E(X) = np;$$
$$\mathrm{Var}(X) = np(1-p);$$
$$\mathrm{SD}(X) = \sqrt{(np(1-p))}.$$

В случае же большого числа наблюдений биномиальной случайной переменной, где значение математического ожидания не приближается ни к нулю, ни к n, распределение вероятностей приближается к нормальному распределению со средней и дисперсией, соответственно равными

$$E(x) = np;$$
$$\sigma^2{}_x = np(1 - p). \tag{4.37}$$

Биномиальное дерево цен активов

Наиболее частое применение биномиального распределения в финансах — это в отношении изменений цен ценных бумаг, где допускается, что в течение следующего малого промежутка времени цена может либо возрасти ("успех"), либо упасть ("неудача") на определенную величину. Биномиальная модель используется как допущение при определении цены опционов. Расчет ожидаемого значения цены актива с применением биномиальной модели состоит из трех шагов: создание биномиального дерева возможных цен активов; определение вероятностей каждого из возможных результатов; умножение каждого из возможных результатов на его вероятность. Сумма произведений и составит ожидаемое значение.

Например, рассмотрим случай, когда надо найти значение цены актива по истечении двух временных интервалов, т.е. после двух биномиальных попыток. Допустим, что в результате каждой попытки существует вероятность роста цены, равная 0,5 (т.е. $p = 0{,}5$), либо падения, равная 0,5 ($1-p = 0{,}5$). Допустим также, что изменение цены в один период времени не зависит от изменения цены в другой период. На рис. 4.9 возможные изменения цены изображены в виде биномиального дерева.

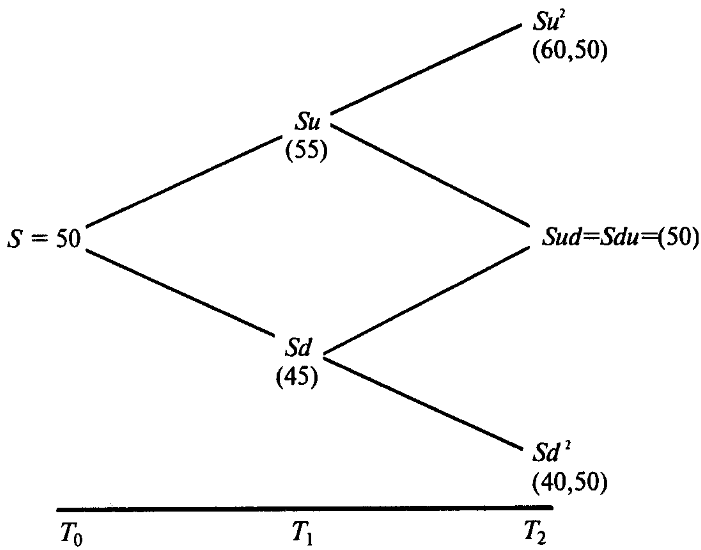

Рис. 4.9

Обозначим настоящий момент времени T_0, конец первого интервала — T_1; цена актива с равной вероятностью либо возрастает до Su, либо падает до Sd. К концу второго промежутка времени T_2, если цена к тому моменту была Su, то она может либо подняться в u раз до Su^2, либо упасть в d раз до Sud. И в альтернативном варианте, если цена в предыдущем периоде упала до Sd, то она могла упасть в d раз до Sd^2, либо возрасти в u раз до Sdu. Тогда после двух временных интервалов, или биномиальных попыток, возможны три результата: Su^2, Sud и Sd^2. Для получения Su^2 нужны два успеха, для получения Sud достаточно одного успеха и для Sd^2 нужны две неудачи.

Чтобы показать процесс на числовом примере, предположим, что $u = 1{,}10$, $d = 0{,}90$ и $S = 50$.

Таким образом, в конце первого периода T_1 цена может либо возрасти до $S \cdot 1{,}1 = 55$, либо упасть до $S \cdot 0{,}9 = 45$, к концу второго периода T_2 — $Su^2 = 60{,}50$, $Sud = 50$ и $Sd^2 = 40{,}50$.

Для определения ожидаемого значения к концу второго интервала следует умножить возможные результаты на их вероятности и сложить полученные произведения. Мы знаем, что существуют три возможных результата — Su^2, Sud и Sd^2, теперь нам надо знать вероятности каждого из этих результатов. Вероятности определяются при помощи уравнения (4.36). При условии, что наш пример подобен предыдущему, неудивительно, что вероятности составят 0,25, 0,50 и 0,25 соответственно.

Отсюда можно найти ожидаемое значение цены актива, изображенной при помощи биномиального дерева, умножая возможные результаты в конце двух временных периодов на их вероятности. Так как вероятность $Su^2 = 60{,}50$ будет 0,25, вероятность Sud =50 будет 0,50 и вероятность $Sd^2 = 40{,}50$ соответствует 0,25, то ожидаемое значение цены актива составит:

$$(60{,}50 \cdot 0{,}25) + (50{,}0 \cdot 0{,}50) + (40{,}50 \cdot 0{,}25) = 50{,}25.$$

Дисперсия будет

$$(60{,}5-50{,}25)^2 \cdot 0{,}25 + (50-50{,}25)^2 \cdot$$
$$\cdot\, 0{,}50 + (40{,}5-50{,}25)^2 \cdot 0{,}25 = 50{,}0625.$$

Может показаться, что такая модель, использующая только два вида изменений — рост и падение, нереалистична. Однако она обладает достаточной гибкостью в выборе значений параметров (u, d и p), чтобы быть мощным инструментом при моделировании поведения цен активов. Следовательно, биномиальное распределение может быть полезно при оценке производных инструментов, в частности при построении численных приближений. Это показано в гл. 8 на примере численных методов.

Распределение Пуассона

Чтобы понять, при каких обстоятельствах применяется распределение Пуассона, предположим, что информация, заставляющая значительно изменяться рыночные цены, поступает на рынок в виде дискретных сообщений независимо, случайным образом и, скажем, со скоростью 10 сообщений в минуту. Возникает вопрос: "Какова вероятность того, что в течение следующей минуты поступит только восемь сообщений?"

Можно попытаться смоделировать этот процесс при помощи биномиального распределения при $n = 60$ и $p = 1/6$, т.е. это можно представить в виде единичного процесса, в ходе которого каждую следующую секунду либо поступает одно сообщение, либо ничего не поступает.

Данное биномиальное распределение (60, 1/6) имеет математическое ожидание, равное 10 ($60 \cdot 1/6$). Существует 61 возможный результат —от 0 (ни одного сообщения не поступило за минуту) до 60 (каждую секунду поступает по сообщению). Каждая из возможностей обладает своей вероятностью. Для $j = 8$ вероятность будет

$$ {}^{60}C_8\left(\frac{1}{6}\right)^8\left(\frac{5}{6}\right)^{52} = \frac{60!}{8!52!}\left(\frac{1}{6}\right)^8\left(\frac{5}{6}\right)^{52} \approx 0{,}1162 . $$

Заметьте, что ${}^{60}C_8$ — это количество способов, которыми могут поступить 8 сообщений за минуту (например, по сообщению в течение первых 8 секунд или по сообщению в течение следующих 8 секунд или любым другим способом).

Эта модель имеет ряд существенных недостатков, в числе которых и то, что в наш век сверхбыстрых коммуникаций может поступить больше одного сообщения в течение одной секунды. Можно улучшить эту модель, введя, скажем, полусекундные интервалы. Соответствующей биномиальной моделью будет бином с параметрами (120, 1/12). В этом случае вероятность получения 8 сообщений будет равна 0,1145:

$$ {}^{120}C_8\left(\frac{1}{12}\right)^8\left(\frac{11}{12}\right)^{112} = \frac{120!}{8!112!}\left(\frac{1}{12}\right)^8\left(\frac{11}{12}\right)^{112} \approx 0{,}1145 . $$

По мере того как мы уменьшаем размер временных интервалов, вычисления все больше усложняются из-за вовлечения в расчеты факториалов и степеней больших чисел. Более того, дополнительная работа не приводит к значительному изменению окончательного ответа, что дает основания подозревать, что мы приближаемся к пределу. В действительности так оно и есть. Можно продемонстрировать это, рассчитав, что бином (240, 1/24) дает вероятность 0,113534, бином (480, 1/48) — 0,113067 и бином (960, 1/96) — вероятность 8 сообщений 0,112834.

Распределение Пуассона является предельным случаем биномиального распределения. Оно применимо в случаях, когда количество попыток (n) приближается к бесконечности, а вероятность успеха (p) — к нулю и математическое ожидание $\lambda = np$ — константа. Формула распределения Пуассона имеет вид:

$$p(X) = \frac{e^{-\lambda}\lambda^{x}}{X!}. \tag{4.38}$$

Распределение Пуассона имеет один параметр λ — математическое ожидание числа появления событий. В нашем примере $\lambda = 10$ (т.е. в среднем 10 сообщений поступает в течение одной минуты). Существует бесконечное множество возможных результатов — 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 и т.д., каждый со своей вероятностью.

Для распределения Пуассона с математическим ожиданием, равным 10, вероятность наступления 8 событий равна:

$$\frac{e^{-10}10^{8}}{8!} = 0{,}1126.$$

Заметьте, что распределение Пуассона довольно симметрично, если имеет достаточно большое математическое ожидание, и в этой ситуации может быть аппроксимировано кривой нормального распределения. Это очень помогает во время работы над группами вероятностей, когда промежуточные вычисления становятся неуправляемыми.

Дисперсия распределения Пуассона равна математическому ожиданию, поэтому можно приблизить распределение Пуассона (λ) к нормальному распределению с параметрами (λ, λ) при условии, что значение λ достаточно велико. Однако надо отметить, что нормальное распределение — это непрерывное распределение, тогда как распределение Пуассона — дискретное. Таким образом, требуется **поправка на непрерывность** при аппроксимации.

Для демонстрации этого мы можем подсчитать, что если X подчиняется распределению Пуассона, вероятность того, что X находится в интервале от 25 до 30 включительно, будет 0,391 (с точностью до трех знаков после запятой).

Записываем $P(25 \leq X \leq 30) = 0{,}391$.

Теперь приближаем X посредством $Y \sim N(30, 30)$.

Для проверки приближения находим $P(24{,}5 < Y < 30{,}5)$. Здесь 0,5 представляет собой поправку на непрерывность, поскольку дискретная величина 25 на распределении Пуассона аппроксимирована интервалом 24,5 — 25,5 на непрерывной кривой нормального распределения

$$P = (24{,}5 < Y < 30{,}5) = P\left(\frac{24{,}5 - 30}{\sqrt{30}} < z < \frac{30{,}5}{\sqrt{30}}\right) = P(1{,}00 < z < 0{,}09) = 0{,}377\,, \tag{4.39}$$

где $z \sim N(0, 1)$.

Приближение будет лучше при больших X, когда оно действительно требуется.

Таким образом, мы видим, что распределение Пуассона применимо в условиях, сходных с условиями для биномиального, за исключением тех случаев, когда число j очень мало, а число попыток n велико. Сейчас мы рассмотрим использование распределения Пуассона на примере больших скачков значения индекса FTSE 100.

Предположим, что нужно смоделировать процесс ежедневного изменения индекса FTSE 100 на более чем 1%. Можно представить, что за определенный период эти изменения случайны, но происходят, скажем, раз в месяц, т.е. 12 раз в год. Как много изменений можно ожидать в течение следующих шести месяцев? Очевидно, мы ожидаем шесть изменений, но мы не очень удивимся, если произойдет четыре, пять, восемь, или любое другое число изменений, близкое к шести.

С первой попытки мы можем построить модель, разбив эти месячные периоды на недельные подпериоды. Затем мы представляем такое изменение происходящим или не происходящим в каждом интервале с вероятностью 12/52, где 52 — число недель в году. Предположив, что данные события должны быть независимы, мы получаем бином с параметрами (26, 12/52). Хотя это и неплохая отправная точка, но это биномиальное распределение является неадекватной моделью практической ситуации, поскольку допускается либо отсутствие изменений, либо одно изменение в течение каждого недельного подпериода. Можно разбить недельные интервалы на дни и использовать биномиальное распределение с параметрами (183, 12/365), но, как мы уже отметили, гораздо эффективнее предположить, что наша случайная переменная подчиняется распределению Пуассона.

Распределение Пуассона описывается следующим образом:

$$P(X = r) = \frac{e^{-\mu}\mu^r}{r!}. \tag{4.40}$$

Продемонстрируем использование распределения Пуассона, вычислив вероятность того, что в течение следующих шести месяцев произойдет больше трех скачков индекса, превышающих 1%.

Анализ информации об индексе FTSE 100 с 3/1/84 по 3/4/92 показывает, что на протяжении 8,25 лет среднее число ежедневных изменений индекса более, чем на 1%, за каждый шестимесячный период, составило 5.

Вероятность наступления одного изменения будет равна:

$$P(X = 1) = \frac{e^{-5}5^1}{1!} = 0{,}0337.$$

Чтобы найти вероятность по крайней мере трех таких изменений в течение следующих шести месяцев, следует найти вероятности $X = 0$, $X = 1$и $X = 2$, сложить их и вычесть из единицы. Получаем:

$$P(X = 0) = \frac{e^{-5}5^0}{0!} = 0{,}0067;$$

$$P(X = 1) = \frac{e^{-5}5^1}{1!} = 0{,}0337;$$

$$P(X = 2) = \frac{e^{-5}5^2}{2!} = 0{,}0842.$$

Таким образом,

$$P(X \geq 3) = 1-[P(X = 0) + P(X = 1) + P(X = 2)] =$$

$$= 1 - 0{,}1246 = 0{,}875.$$

Следовательно, существует вероятность, равная 87,5%, что произойдет по крайней мере три дневных изменения индекса FTSE 100 на 1% и более в течение следующих шести месяцев.

Распределение Парето—Леви

Нормальное распределение не является панацеей для всех проблем при моделировании. В частности, его использование при моделировании относительных цен активов может быть подвергнуто сомнению по крайней мере по двум причинам:

1) предполагается, что относительные цены независимы с течением времени, в то время как они демонстрируют автокорреляцию;
2) вероятность наступления крайних событий в действительности выше, чем это подразумевается нормальным распределением.

Эти факторы указывают на потребность в распределении, которое после логарифмической трансформации будет иметь более высокий "пик" (как следствие автокорреляции) и более "жирные" хвосты (вследствие частоты крайних значений).

Определенно такие симметричные распределения должны иметь коэффициент эксцесса больше нуля (для нормального распределения он равен нулю). Распределения с коэффициентом больше нуля называются **островершинными** (leptokurtic) (см. рис. 4.10), распределения с коэффициентом эксцесса меньше нуля называются **плосковершинными** (platykurtic).

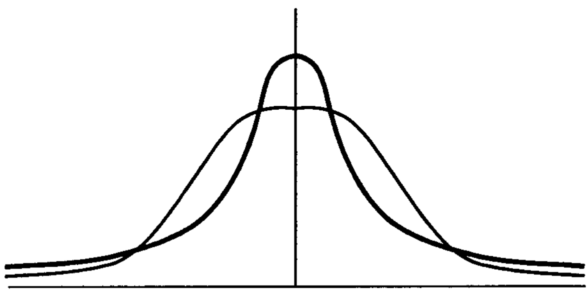

Рис. 4.10

В то время как упомянутые выше распределения определены на основе реальных обстоятельств, в соответствии с которыми они подбираются, было бы полезно иметь теоретическое распределение, наиболее полно соответствующее наблюдаемым явлениям, даже если между ними нет никакой очевидной физической связи. В таких обстоятельствах особенно полезны семейства распределений, которые характеризуются несколькими параметрами.

Такое семейство распределений — это стабильные распределения, называемые так потому, что при сложении распределений (перемножая линейные комбинации характеризующих их функций) этого семейства получается другое распределение, относящееся к этому же семейству. Стабильные распределения в свою очередь состоят из других, лежащих в их основе, распределений. Распределения, построенные на основе распределения Парето (функция плотности вероятности которого $f(X) = \alpha/X^{\alpha+1}$ для $X > 1$), обладают требуемыми характеристиками (симметричность, высокий пик и "жирные" хвосты) при конкретных значениях четырех определяющих параметров. Эти четыре параметра:

α определяет высоты хвостов;

β определяет скошенность;

γ определяет горизонтальный масштаб функции;

δ определяет расположение.

Симметричные распределения получаются при $\beta = 0$. Максимальное возможное значение $\alpha = 2$, в этом случае получается нормальное распределение (для которого $\gamma = \sigma^2/2$). Эмпирические свидетельства указывают на то, что подходящие значения α для моделирования распределения логарифмов доходностей меньше двух, возможны между 1,7 и 1,9, хотя такие распределения страдают недостатком — имеют бесконечно большую дисперсию!

Невозможно записать явную функцию плотности вероятности для таких распределений, и соответствующие вероятности должны рассчитываться с помощью численных значений.

УПРАЖНЕНИЯ

Основные определения

1. Брошена стандартная шестигранная игральная кость. Какова вероятность того, что выпадет:

   а) три;
   б) не три;
   в) меньше трех;
   г) семь.

2. Одновременно выбрасывают две шестигранные кости и полученные числа записывают. Перечислите члены пространства элементарных событий. Результаты, выпавшие на каждой из костей, складываются. Какова вероятность того, что общая сумма равна 6?

Правила сложения

3. Если S — множество натуральных чисел — меньше 13, проклассифицируйте следующие пары подмножеств S как взаимоисключающие и/или исчерпывающие:

   а) {четные}, {нечетные};
   б) {кратные 3}, {кратные 4};
   в) {1, 2, 3}, {≥3};
   г) {1, 2, 3}, {4, 5, 6}.

   Рассчитайте вероятность того, что случайно выбранный член множества S будет членом каждого из 8 вышеперечисленных подмножеств.

Правило умножения, условная вероятность и независимость

4. Случайным образом выбирается число из множества {1, 2, 3, 4, 5, 6, 7, 8, 9}.

   а) Рассчитайте вероятность того, что выбранное число будет больше 5;
   б) рассчитайте вероятность того, что это число будет больше 5 и нечетное;
   в) используя ответы к вопросам а) и б), рассчитайте вероятность того, что выбранное число будет нечетным при условии, что оно больше 5. Проверьте правильность ответа, используя другой метод.

5. Случайным образом выбирается число из множества **S** = {1,2,3,4,5,6,7}.
   Рассмотрите следующие подмножества событий **A** ={1, 2, 3}, **B** = {2, 4, 6}.
   Являются ли **A** и **B** статистически независимыми?

6. Брошены две шестигранные кости — красная и синяя. Предположим, что событие **A** — это выпадение шестерки на синей кости и событие **B** — это то, что сумма выпавших чисел равна 7. Являются ли события **A** и **B** статистически независимыми?

7. В группе из 1000 человек 452 человека имеют текущие счета, 336 — имеют депозитные счета и 302 — и текущий и депозитный счета.

Являются ли атрибуты "обладание депозитным счетом" и "обладание текущим счетом" статистически независимыми?

8. Ценная бумага может подорожать на 1% в течение следующего месяца с вероятностью 0,6. Она также может подешеветь на 1% в течение следующего месяца с вероятностью 0,4. Предполагая, что ежемесячные изменения цены независимы, рассчитайте:

   а) вероятность того, что за три месяца цена станет равной $(1{,}01)^3$ от первоначальной;

   б) вероятность того, что за три месяца цена станет равной $0{,}99 \cdot (1{,}01)^2$ от первоначальной.

   Подтвердите ответ на вопрос б), ссылаясь на правила сложения и умножения вероятностей.

9. Изучение ежедневных изменений цен на двух финансовых рынках выявило следующее:

		Рынок 1	
		Рост цен	Падение цен
Рынок 2	Рост цен	165	26
	Падение цен	32	137

   Рассчитайте вероятности:

   а) роста цен на рынке 1;

   б) роста цен на рынке 1 при условии, что цены на рынке 2 растут;

   в) роста цен на рынке 2 при условии, что цены на рынке 1 растут.

10. Для принятия решений о покупке ценных бумаг была разработана система анализа рынка. Из прошлых данных известно, что 5% рынка представляют собой "плохие" ценные бумаги — неподходящие объекты для инвестирования. Предложенная система определяет 98% "плохих" ценных бумаг как потенциально "плохие", но также определяет 15% пригодных инвестиций как потенциально "плохие". При условии, что ценная бумага была определена как потенциально "плохая", какова вероятность того, что ценная бумага в действительности "плохая"?

    Прокомментируйте пригодность системы для принятия инвестиционных решений.

Случайные переменные

11. Случайная переменная X принимает значение 0 с вероятностью 0,5, значение 1 — с вероятностью 0,3 и 2 — с вероятностью 0,2. Найдите:

    а) $E[X]$; б) $E[3 + X]$; в) $E[3X]$; г) $E[X^2]$; д) $\mathrm{Var}(X)$.

12. Обзор счетов 400 инвесторов на фондовой бирже дал следующую информацию о числе сделок в течение последнего квартала:

X (сделок)	0	1	2	3	4	5	6	7	8	9	10
Число держателей финансовых инструментов	146	97	73	34	23	10	6	3	4	2	2

а) постройте график распределения X;

б) найдите вероятность того, что случайно выбранный инвестор произвел:

ноль сделок;

по крайней мере одну сделку;

больше пяти;

меньше шести;

в) найдите математическое ожидание и дисперсию числа сделок.

13. Ожидаемая рентабельность актива А равна 8% со средним квадратическим отклонением, равным 7%. Ожидаемая рентабельность актива В равна 11% и среднее квадратическое отклонение 10%. Корреляция между этими активами 0,7.

Найдите ожидаемую доходность и среднее квадратическое отклонение портфеля, состоящего на 35% из А и на 65% из В.

Нормальное распределение

14. Если Z подчиняется закону нормального распределения (т.е. $m = 0$, $s = 1$), найдите:

а) $P(Z > 1{,}2)$; б) $P(Z \geq 1{,}34)$; в) $P(Z \leq 1{,}01)$;
г) $P(Z > 0{,}85)$; д) $P(Z \leq 2{,}14)$; е) $P(Z \leq 0{,}07)$;
ж) $P(Z < 1{,}37)$; з) $P(Z \geq -2{,}03)$; и) $P(Z \leq -0{,}17)$;
к) $P(Z \geq -1{,}36)$; л) $P(0{,}34 \leq Z \leq 1{,}29)$; м) $P(-2{,}01 \leq Z \leq 1{,}52)$;
н) $P(-1{,}21 \leq Z \leq -0{,}34)$.

15. Если $X \sim N(5, 36)$ (т.е. X нормально распределен с $m = 5$ и $s = 6$), найдите следующие вероятности:

а) $P(X \geq 14)$; б) $P(X \leq 9{,}5)$; в) $P(X \geq 3{,}5)$; г) $P(2 \leq X \leq 12{,}5)$.

16. Текущая цена акции может быть приблизительно смоделирована при помощи нормального распределения с математическим ожиданием £$15,28 и средним квадратическим отклонением, равным £0,12. Рассчитайте вероятности того, что цена:

а) не ниже £15,5; б) не выше £15,00;
в) между £15,10 и £15,40; г) между £15,05 и £15,10.

17. Цена некой ценной бумаги нормально распределена. В течение последнего года на протяжении 20% рабочих дней цена была ниже 20. В 75% случаев цена была выше 25. Найдите математическое ожидание и среднее квадратическое отклонение цены.

    Критически рассмотрите применение нормального распределения в данной ситуации. Объясните, как логнормальное распределение может быть использовано для преодоления проблемы.

Биномиальное распределение

18. Монета "нечестная" и вероятность выпадения "орла" = 0,8. Монета подброшена пять раз. Рассчитайте вероятность выпадения:

    а) ровно двух "орлов";
    б) ровно четырех "орлов";
    в) по крайней мере двух "орлов".

19. Предположим, что ценная бумага из п. 8 в данный момент стоит £10. Найдите вероятность того, что она будет стоить £10,40 через год.

20. В брокерской конторе для стимулирования прибыльности торговли по отношению к сотрудникам применяется система премий. В соответствии с этой системой сотрудник, не достигающий установленного дневного уровня прибыли на протяжении более трех дней за две недели (10 рабочих дней), теряет свою премию за этот двухнедельный период. Если вероятность того, что сотрудник не выполнит требуемую норму прибыли, равна 0,15, найдите, сколько премий будет потеряно 100 сотрудниками за 50-недельный год?

    Какие предположения вы делали при нахождении ответа? Соответствует ли это действительности?

Распределение Пуассона

21. Поступление информации на торговую площадку в течение напряженного торгового периода подчиняется распределению Пуассона с математическим ожиданием 3,5 сообщения в минуту. Какова вероятность того, что в течение следующей минуты:

    а) не поступит ни одного сообщения;
    б) поступит по крайней мере одно сообщение;
    в) поступят два сообщения;
    г) поступят четыре сообщения.

    Какова вероятность поступления более 20 сообщений в течение 5 минут? Посмотрите, насколько близко к этому результату был бы получен ответ при использовании приближения к нормальному распределению.

ОТВЕТЫ

1. а) 1/6; б) 5/6; в) 2/6 = 1/3; г) 3/6 = 1/2; д) 0.

2. Каждая ячейка таблицы представляет собой результат, число в ячейке — сумма, соответствующая этому результату.

Результат на первой кости; строки — результат на второй кости

	1	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Вероятность (сумма = 6) = $\frac{6}{36} = \frac{1}{6}$.

3. а) взаимоисключающие и исчерпывающие;
   б) ни то ни другое;
   в) исчерпывающие, но не взаимоисключающие;
   г) взаимоисключающие, но не исчерпывающие.

4. а) 4/9; б) 2/9; в) $\frac{2/9}{4/9} = \frac{1}{2}$ (или 2/4).

5. $p(\mathbf{A}) = 3/7$, $p(\mathbf{B}) = 3/7$, $p(\mathbf{A} \wedge \mathbf{B}) = \frac{1}{7}$, $3/7 \cdot 3/7 \neq 1/7$,

следовательно не независимы.

6. $p(\mathbf{A}) = 1/6$, $p(\mathbf{B}) = 6/36 = 1/6$, $p(\mathbf{A} \wedge \mathbf{B}) = 1/36$ $1/6 \cdot 1/6 = 1/36$,

следовательно, независимы.

Или $p(\mathbf{B}) = p(\mathbf{B}|\mathbf{A})$, следовательно, независимы.

7. p (текущий) = 0,452, p (депозитный) = 0,336, p (текущий) · p (депозитный) = 0,151872, но p (текущий $\wedge$ депозитный) = 0,302, , следовательно, не независимы.

8. а) $0{,}6^3 = 0{,}216$, б) $3 \cdot 0{,}4 \cdot 0{,}6^2 = 0{,}432$,
$(0{,}4 \cdot 0{,}6 \cdot 0{,}6) + (0{,}6 \cdot 0{,}4 \cdot 0{,}6) + (0{,}6 \cdot 0{,}6 \cdot 0{,}4)$.

Умножение оправдывается независимостью. Сложение, потому что взаимоисключающи.

9. а) $197/360 \approx 0{,}547$ б) $165/191 \approx 0{,}864$ в) $165/197 \approx 0{,}838$.

10. Для облегчения расчетов рассмотрите выборку из 10 000 ценных бумаг.

		Определены как: Плохие	Определены как: Пригодные	
В действительности	Плохие	490	10	500
В действительности	Пригодные	1 425	8 075	9 500
		1 915	8 085	10 000

p (плохие |определены как плохие) = 490/1 915 = 0,256
упускает значительное число привлекательных инвестиций.

11.

$$E(X) = 0 \cdot 0{,}5 + 1 \cdot 0{,}3 + 2 \cdot 0{,}2 = 0{,}7;$$
$$E(3 + X) = 3 \cdot 0{,}5 + 4 \cdot 0{,}3 + 5 \cdot 0{,}2 = 3{,}7 = 3 + E(X);$$
$$E(3X) = 0 \cdot 0{,}5 + 3 \cdot 0{,}3 + 6 \cdot 0{,}2 = 2{,}1 = 3E(X);$$
$$E(X^2) = 0 \cdot 0{,}5 + 1 \cdot 0{,}3 + 4 \cdot 0{,}2 = 1{,}1\ (\neq (E(X))^2!);$$
$$\mathrm{Var}(X) = E(X^2) - (E(X))^2 = 0{,}61.$$

12.

X	0	1	2	3	4	5	6	7	8	9	10
Частота	146	97	73	34	23	10	6	3	4	2	2
Вероятность	0,365	0,2425	0,1825	0,085	0,0575	0,025	0,015	0,0075	0,01	0,005	0,005
Кумулятивность	0,365	0,6075	0,79	0,875	0,9325	0,9575	0,9725	0,98	0,99	0,995	1

$P(X = 0) = 0{,}365$; $P(X \leq 1) = 1 - 0{,}365 = 0{,}635$;
$P(X > 5) = 1 - 0{,}9575 = 0{,}0425$;
$P(X < 6) = 0{,}9575$;
$E(X) = 1{,}535$; $\mathrm{Var}(X) = 3{,}378775$.

13. Ожидаемый доход = $(0{,}35 \cdot 0{,}08) + (0{,}65 \cdot 0{,}11) = 0{,}0995 \approx 10\%$.

$$\text{Дисперсия} = (0{,}35^2 \cdot 0{,}07^2) + (2 \cdot 0{,}35 \cdot 0{,}65 \cdot 0{,}7 \cdot 0{,}07 \cdot 0{,}10) + (0{,}65^2 \cdot 0{,}10^2) = 0{,}007055.$$

Среднее квадратическое отклонение = 0,084.

14. Все ответы с точностью до двух знаков после запятой

а) 0,12, б) 0,09, в) 0,84, г) 0,20, д) 0,98, е) 0,53, ж) 0,91, з) 0,98, и) 0,43, к) 0,91, л) 0,27, м) 0,91, н) 0,25.

15. а) $P\left(z \geq \frac{14-5}{6} = 1{,}5\right) \approx 0{,}07$,

б) $P\left(z \leq \frac{9{,}5-5}{6} = 0{,}75\right) \approx 0{,}77$,

в) $P\left(z \geq \frac{3{,}5-5}{6} = -0{,}5\right) \approx 0{,}69$,

г) $P\left(\frac{2-5}{6} \leq z \leq \frac{12{,}5-5}{6}\right) = P(-0{,}5 \leq z \leq 1{,}25) \approx 0.59$.

16. а) $P\left(z \geq \frac{15{,}505 - 15{,}28}{0{,}12} = 1{,}875\right) \approx 0{,}03$,

б) $P\left(z \leq \frac{15{,}005 - 15{,}28}{0{,}12} = -2{,}29\right) \approx 0{,}01$,

в) $P\left(\frac{15{,}095 - 15{,}28}{0{,}12} \leq z \leq \frac{15{,}405 - 15{,}28}{0{,}12}\right) = P(-1{,}54 \leq z \leq 1{,}04) \approx 0{,}79$,

г) $P\left(\frac{15{,}045 - 15{,}28}{0{,}12} \leq z \leq \frac{15{,}155 - 15{,}28}{0{,}12}\right) = P(-1{,}96 \leq z \leq -1{,}04) \approx 0{,}12$.

17. $P(z) \leq ((20-\mu)/\sigma) = 0{,}20$, значит $(20-\mu)/\sigma = -0{,}84$,

$P(z) \geq ((75-\mu)/\sigma) = 0{,}25$, значит $(75-\mu)/\sigma = 0{,}67$.

Решаем эту систему уравнений, $\mu \approx 50{,}6$, и $\sigma \approx 36$, если

$$X \sim N(50{,}6,\ 36^2),\ \text{то}\ P(X < 0) = P(z < 50{,}6/36 \approx 1{,}41) \approx 0{,}08.$$

Вместо использования нормального распределения при моделировании поведения цены, которое допускает ее отрицательные значения, можно моделировать натуральные логарифмы рентабельности как нормальное распределение. Это позволяет учесть то, что цена формируется в процессе умножения, а кроме того исключаются отрицательные значения.

18. а) 0,0512, б) 0,4096, в) 0,99328.

19. $10{,}40 = 10 \cdot 1{,}01^n \cdot 0{,}99^{12-n}$
$\log(10{,}40) = \log 10 + n \cdot \log 1{,}01 + (12-n) \cdot \log 0{,}99 \Rightarrow n = 8$.
Вероятность $= {}^{12}C_8(0{,}6)^8(0{,}4)^4 \approx 0{,}213$.

20. Обозначим число дней невыполнения норматива по прибыли X. Предположим, что X подчиняется биномиальному распределению (10, 0,15).

$$P(X > 3) = 1 - P(X = 0) - P(X = 1) - P(X = 2) - P(X = 3) =$$
$$= 1 - {}^{10}C_0\,0{,}15^0\,0{,}85^{10} - {}^{10}C_1 0{,}15\,0{,}85^9 - {}^{10}C_2 0{,}15^2\,0{,}85^8 -$$
$$- {}^{10}C_3 0{,}15^3\,0{,}85^7 \approx 0{,}05.$$

21. X ~ Пуассон (3.5)

$$P(X = 0) = e^{-3{,}5} \approx 0{,}03$$
$$P(X \geq 1) = 1 - P(X = 0) \approx 0{,}97$$
$$P(X = 2) = (3{,}5^2 e^{-3{,}5})/2! \approx 0{,}185$$
$$P(X = 4) = (3{,}5^4 e^{-3{,}5})/4! \approx 0{,}189$$

$Y \sim$ Пуассон$(5 \cdot 3{,}5)$ = Пуассон(17,5)
$P(Y > 20) \approx 0{,}231$

Применяя $Y \approx N(17{,}5;\ 17{,}5)$, получаем

$$P\left(z \geq \left(\frac{20{,}5 - 17{,}5}{\sqrt{17{,}5}}\right) \approx 0{,}717\right) \approx 0{,}236.$$

ПРИЛОЖЕНИЕ: математическое ожидание и дисперсия логнормального распределения

X и Y — случайные переменные. $Y = \ln(X)$ и $Y \sim N(\mu, \sigma^2)$, т.е. Y распределен логнормально. Чему равняется $E(X)$ и дисперсия X?

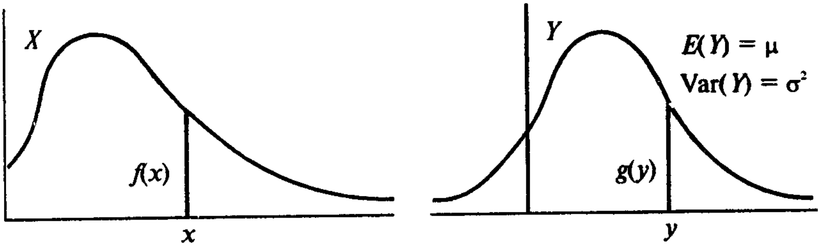

Рис. 4.11

$y = \ln(x)$.

Предварительный результат

$$\int_x^{\infty} f(x)\mathrm{d}x = \int_y^{\infty} g(y)\mathrm{d}y \Rightarrow f(x) = g(y)\frac{\mathrm{d}y}{\mathrm{d}x}.$$

Математическое ожидание

$$E(X) = \int_0^{\infty} xf(x)\mathrm{d}x =$$

$$= \int_{-\infty}^{\infty} \exp(y)g(y)\mathrm{d}y =$$

$$= \int_{-\infty}^{\infty} \exp(y)\left(\frac{1}{\sigma\sqrt{2\pi}}\exp\left(\frac{-(y-\mu)^2}{2\sigma^2}\right)\right)\mathrm{d}y =$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left(y - \frac{(y-\mu)^2}{2\sigma^2}\right)\mathrm{d}y =$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}\left(y^2 + \left(-2\sigma^2 - 2\mu\right)y + \mu^2\right)\right)\mathrm{d}y =$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}\left(\left(y - \left(\sigma^2 + \mu\right)\right)^2 - \frac{1}{2\sigma^2}\left(-2\mu\sigma^2 - \sigma^4\right)\right)\right)\mathrm{d}y$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}\left(y - \left(\sigma^2 + \mu\right)\right)^2\right)\exp\left(\mu + \frac{\sigma^2}{2}\right)\mathrm{d}y$$

$$= \exp\left(\mu + \frac{\sigma^2}{2}\right)\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}\left(y - \left(\sigma^2 + \mu\right)\right)^2\right)\mathrm{d}y =$$

$$= \exp\left(\mu + \frac{\sigma^2}{2}\right).$$

Дисперсия

$\mathrm{VAR}(X) = E(X^2) - (E(X))^2$.

$$E(X^2) = \int_0^{\infty} x^2 f(x)\mathrm{d}x =$$

$$= \int_{-\infty}^{\infty} \exp(2y) g(y)\mathrm{d}y =$$

$$= \int_{-\infty}^{\infty} \exp(2y) \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)\mathrm{d}y =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}\left((y-\mu)^2 - 4\sigma^2 y\right)\right)\mathrm{d}y =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}\left(y^2 - \left(2\mu + 4\sigma^2\right)y + \mu^2\right)\right)\mathrm{d}y =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}\left(y - \left(\mu + 2\sigma^2\right)\right)^2 - \frac{1}{2\sigma^2}\left(-\left(\mu + 2\sigma^2\right)^2 + \mu^2\right)\right)\mathrm{d}y =$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{\left(y - \left(\mu + 2\sigma^2\right)\right)^2}{2\sigma^2}\right) \exp\left(\frac{4\mu\sigma^2 + 4\sigma^4}{2\sigma^2}\right)\mathrm{d}y =$$

$$= \exp(2\mu + 2\sigma^2) \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{\left(y - \left(\mu + 2\sigma^2\right)\right)^2}{2\sigma^2}\right)\mathrm{d}y =$$

$$= \exp(2\mu + 2\sigma^2).$$

$$\therefore VAR(X) = \exp(2\mu + 2\sigma^2) - \exp 2(\mu + \sigma^2/2) =$$

$$= \exp(2\mu + 2\sigma^2) - \exp(2\mu + \sigma^2) = \exp(2\mu + \sigma^2)(\exp(\sigma^2) - 1).$$

5

СТАТИСТИЧЕСКИЕ ВЫВОДЫ: ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ И ПРОВЕРКА ГИПОТЕЗ

ВВЕДЕНИЕ

Данные, с которыми нам предстоит работать, могут представлять собой полный набор данных, который в этом случае называют **генеральной совокупностью** (population), а описательные статистические показатели известны как **статистические показатели генеральной совокупности** (population statistics), или параметры генеральной совокупности (population parameters). Или же данные могут представлять собой выборку из общего объема наблюдений за какой-либо переменной, в этом случае пользуются **описательными статистическими показателями**, известными как показатели (или статистики) **выборочной совокупности** (sample statistics).

Вообще, статистический анализ можно разделить на два вида — **описательный** (descriptive) и **имеющий целью сделать какие-либо выводы** (inferential). Описательная статистика рассмотрена в гл. 2, где мы рассчитывали различные виды описательных статистических показателей для того, чтобы охарактеризовать определенные свойства данных. Однако не было сделано никаких попыток к тому, чтобы, используя полученные результаты, оценить или сделать выводы относительно параметров анализируемой генеральной совокупности. В случае статистики, занимающейся получением выводов, информация, представленная в виде выборочных статистических показателей, используется для оценки параметров генеральной совокупности.

В свою очередь **статистика получения выводов** (inferential statistics) применяется в следующих двух видах анализа: **оценивание** (estimation) и **проверка гипотез** (hypothesis testing). Оценивание применяется в том случае, когда мы заранее (*a priori)* не знаем величину параметров генеральной совокупности. В этом случае мы задаем **доверительные интервалы** (confidence intervals) для оценивания действительных параметров генеральной совокупности с определенной степенью точности нашей оценки.

Если мы располагаем знаниями *a priori* относительно параметров генеральной совокупности, эти сведения могут быть сформулированы в виде гипотезы, которая может быть проверена. Например, мы можем **проверить гипотезу**, что значение параметра генеральной совокупности лежит в определенном интервале. Эти два раздела — оценивание и проверка гипотез — составляют основное содержание данной главы.

На этом этапе полезно ввести обозначения, чтобы можно было различать показатели генеральной совокупности и соответствующие показатели выборки:

μ — генеральная средняя;	$\overline{X}$ — выборочная средняя;
σ — среднее квадратическое отклонение в генеральной совокупности;	s — среднее квадратическое отклонение в выборке.

Если нам достаточно повезло в том смысле, что у нас имеется возможность работать с данными по всей генеральной совокупности, то понятно, что при учете всех возможных данных полученные статистические показатели будут **реальными** показателями генеральной совокупности. Никаких проверок или оценок в этом случае не потребуется. Тем не менее, в финансовых и социальных науках редко можно воспользоваться данными по всей совокупности возможных наблюдений. Приходится иметь дело с выборками, вследствие чего мы не можем точно знать, совпадают ли статистические показатели по выборке с реальными статистическими показателями генеральной совокупности или же значительно отличаются от них. Следовательно, необходимо разработать методику расчета степени достоверности, позволяющей определить параметры генеральной совокупности, исходя из показателей выборки.

ТЕОРИЯ ВЫБОРОЧНОГО НАБЛЮДЕНИЯ

При работе с выборочными данными мы должны использовать **теорию выборочного наблюдения** (sampling theory) для определения распределения вероятностей, относящегося к конкретному выборочному статистическому показателю. Это распределение вероятностей известно как **выборочное распределение** (sampling distribution).

Для того чтобы понять идею, лежащую в основе выборочного распределения, рассмотрим процесс извлечения шаров из барабана для лотерейного розыгрыша. После того как шар был извлечен, а его номер зафиксирован, шар возвращается обратно в барабан и снова может быть выбран. Эта разновидность выборки известна как **случайная повторная выборка** (random sampling with replacement). О ней будет идти речь далее в этой главе, так как она особенно применима в качестве модели выборки из больших или бесконечных совокупностей.

Этот тип выборки используется в статистике, например, следующим образом. Рассмотрим 50 выборок по 20 шаров в каждой с возвращением шаров обратно в барабан. Для каждой из выборок рассчитаем какой-либо статистический показатель, скажем, арифметическую среднюю. Рассчитанные выборочные средние для каждой из выборок будут немного отличаться по значению друг от друга, но все они будут сконцентрированы вокруг реальной средней — одни выше, другие ниже ее. Исходя из предположения о том, что выборки осуществляются случайным образом, выборочные статистические показатели расцениваются как случайные величины. Продолжительный процесс осуществления повторных выборок позволит получить распределение вероятностей выборочного статистического показателя. Это распределение вероятностей называется **выборочным распределением** выборочного статистического показателя. Наши знания относительно выборочного распределения каждого из статистических показателей позволяют сделать выводы относительно величины параметров генеральной совокупности исходя из выборочных показателей.

Мы сосредоточимся на нормальном распределении и t-распределениии Стьюдента для средних величин, χ^2-распределении для дисперсий и F-распределении для коэффициента детерминации. Так как последнее относится к оценке степени пригодности линий регрессии, что будет рассмотрено в следующей главе, мы отложим рассмотрение F-распределения до следующей главы.

Выборочное распределение выборочных показателей

Выборочное распределение выборочной средней арифметической

В гл. 4 мы узнали, что согласно центральной предельной теореме средние аддитивных процессов (арифметические средние) будут нормально распределены независимо от распределения исходных величин при условии, что выборки достаточно велики (объем выборки больше 30). Если первоначальная совокупность нормально распределена, а объем выборки меньше 30, распределение выборочных средних будет следовать t-распределению Стьюдента.

Математическое ожидание средней всех выборочных средних является генеральной средней. **Среднее квадратическое отклонение выборочных средних** известно как **стандартная ошибка** (standard error) и рассчитывается как отношение среднего квадратического отклонения генеральной совокупности к квадратному корню из объема выборки. Обычно эта величина известна как стандартная ошибка средней и определяется по формуле:

$$\mathrm{SE} = \frac{\sigma}{\sqrt{n}}, \tag{5.1}$$

где σ — среднее квадратическое отклонение в генеральной совокупности,

n — объем выборки, по которой рассчитывается средняя.

Таким образом, стандартная ошибка находится через среднее квадратическое отклонение генеральной совокупности, отнесенное к квадратному корню из объема выборки. Однако, так как среднее квадратическое отклонение генеральной совокупности является неизвестным, в качестве его оценки принимается выборочное среднее квадратическое отклонение s. Тогда стандартная ошибка будет определяться так:

$$\mathrm{SE} = \sqrt{\frac{s^2}{n}} = \frac{s}{\sqrt{n}}. \tag{5.2}$$

Таким образом, исходя из центральной предельной теоремы мы можем сказать, что средняя достаточно большой выборки приближенно нормально распределена со средним значением, равным генеральной средней, и средним квадратическим отклонением, равным стандартной ошибке средней, т.е.:

$$\overline{X} \sim N(\mu, s/\sqrt{n}),$$

где знак "~" означает "приближенно распределено".

Объяснение, почему в качестве делителя используется $\sqrt{n}$, дано в приложении 5.1.

Для выборок малого объема, взятых из совокупности с нормальным распределением признака, неопределенность оценки σ с помощью s (которая возрастает при уменьшении объема выборки) допускается при условии применения t-распределения.

Выборочное распределение выборочной дисперсии

Выборочное распределение выборочной дисперсии — это одна из форм гамма-распределения, известная как "хи-квадрат" распределение, обозначаемое через χ^2. Это распределение принимает разную форму для разного числа степеней свободы. Выборочную дисперсию необходимо привести к стандартизованной форме с помощью способа, аналогичного тому, с помощью которого нормальное распределение приводилось к стандартизованной форме в гл. 4. Приведение к стандартизованной форме в нашем случае принимает вид:

$$\chi^2_{n-1} = (n-1)s^2 / \sigma^2 . \tag{5.3}$$

Индекс (n—1) обозначает число степеней свободы, которое в случае χ^2-распределения равно количеству наблюдений минус 1. Для малых выборок форма распределения вероятности смещена вправо, но с увеличением объема выборки распределение становится более симметричным.

Таким образом, если средняя нашей переменной нормально распределена, тогда:

$$(n-1)s^2/\sigma^2 \tag{5.4}$$

будет иметь χ^2- распределение с (n—1) степенями свободы. Несмотря на то, что нам неизвестно σ, полученный результат может быть использован при проверке гипотез относительно величины σ или при определении доверительных интервалов.

ОЦЕНИВАНИЕ И ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ

Располагая выборочными статистическими показателями и знаниями выборочных распределений этих показателей, мы готовы **оценить** параметры генеральной совокупности анализируемых данных. При использовании выборочных показателей в оценивании параметров генеральной совокупности выборочные показатели называются **оценками** (estimators). Желательно, чтобы эти оценки соответствовали характеристике **BLUE**, что означает **B**est (наилучшая) **L**inear (линейная) **U**nbiased (несмещенная) **E**stimator (оценка). Однако может быть необходимым добиться некоторой степени смещенности с целью получения меньшей дисперсии.

Наилучшая (best) означает свойство оценки как имеющей наименьшую дисперсию из всех возможных несмещенных оценок.

Линейная (linear) означает свойство линейной функциональной зависимости оценки от выборочных наблюдений. Например:

$$\hat{\theta} = c_1 X_1 + c_2 X_2 + \ldots + c_n X_n,$$

где c — константа.

Важность данной характеристики может быть не совсем очевидна, но достаточно сказать, что математические свойства линейной оценки гораздо легче анализировать.

Несмещенная (unbiased) означает свойство, состоящее в том, что математическое ожидание оценки (средняя выборочного распределения) равно параметру генеральной совокупности, т.е. в результате осуществления множества выборок для определения оценки одни выборочные показатели будут больше параметра генеральной совокупности, другие меньше, но среднее значение будет равно параметру генеральной совокупности. Напротив, при смещенной оценке среднее значение будет больше или меньше параметра генеральной совокупности.

Оценки, которые одновременно и **несмещенные** и имеют **наименьшую дисперсию**, называются **эффективными оценками**.

Если оценка смещенная и/или неэффективная, то желательно, чтобы она характеризовалась **ассимптотическими свойствами**.

Ассимптотическая несмещенность (asimptotic unbiasedness) означает, что любая существующая смещенность в малых выборках (менее 30) уменьшается с увеличением выборки и стремится к нулю с увеличением объема выборки до бесконечности.

Ассимптотическая эффективность (asimptotic efficiency) — это свойство оценки, когда она одновременно состоятельна и имеет меньшую ассимптотическую дисперсию, чем любая другая состоятельная оценка.

Состоятельность (consistency) — это свойство оценки, согласно которому дисперсия оценки уменьшается до нуля с увеличением объема выборки до бесконечности.

Доверительные интервалы

Вычисление выборочных статистических показателей в качестве оценки параметров генеральной совокупности дает в результате нам то, что мы знаем как **точечную оценку** (point estimates). Однако нам известно, что эта точечная оценка будет сделана с некоторой ошибкой, называемой **оценочной ошибкой** (estimation error). Следовательно, нам нужен механизм, который бы позволил определить степень доверия к этим точечным оценкам. Таким образом, мы подошли к понятию **доверительного интервала** (confidence interval) или **интервальной оценки**.

Проиллюстрируем принцип доверительных интервалов, применив его к средней, также мы адаптируем процесс определения объема выборки, необходимого для получения заданной степени доверия.

Доверительные интервалы для средней (большая выборка)

Напомним, что проблема состоит в том, что мы не знаем среднюю генеральной совокупности, и нам известна только выборочная средняя. Тем не менее, согласно центральной предельной теореме мы знаем, что выборочное распределение средних имеет среднее значение, которое в свою очередь равно генеральной средней, а среднее квадратическое отклонение (стандартная ошибка) равно $\sigma / \sqrt{n}$, где σ — среднее квадратическое отклонение в генеральной совокупности.

Но появляется другая проблема — мы не знаем величины среднего квадратического отклонения в генеральной совокупности, нам известно только выборочное среднее квадратическое отклонение. Однако здесь можно применить другую часть теории выборочного наблюдения, согласно которой наилучшей оценкой σ является:

$$s = \frac{\sum\left(X - \overline{X}\right)^2}{(n-1)}. \qquad (5.5)$$

Другими словами, при условии, что выборочное среднее квадратическое отклонение s найдено при использовании $(n—1)$ в качестве делителя, s является несмещенной оценкой среднего квадратического отклонения генеральной совокупности. Доказательство этого утверждения рассмотрено в приложении 5.1.

Известно, что для нормально распределенной величины 95% наблюдений будет находиться выше или ниже средней не более, чем на 1,96 среднего квадратического отклонения. Так как средние квадратические отклонения выборочных распределений средних называются стандартными ошибками, мы можем сказать, что выборочная средняя в 95% случаев будет находиться внутри **интервала**, равного генеральной средней плюс/минус

1,96 **стандартной ошибки**. Формула **доверительного интервала** выглядит так:

$$\mu \pm 1{,}96\frac{s}{\sqrt{n}}, \tag{5.6}$$

где s — выборочное среднее квадратическое отклонение.

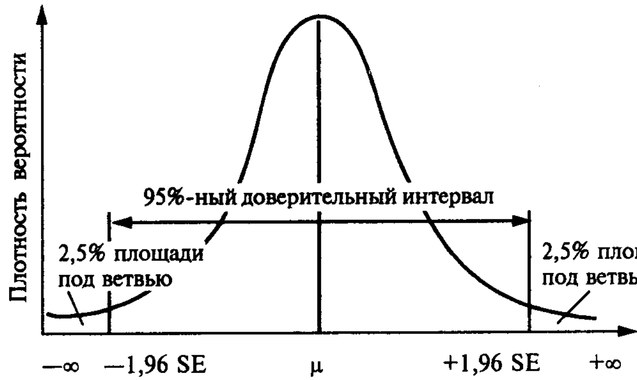

Рис. 5.1

Наглядно это может быть проиллюстрировано с помощью рис. 5.1 (SE — стандартная ошибка). График показывает, что в 95% случаев (2,5% в каждой из граничных областей) выборочная средняя находится в пределах 1,96 стандартной ошибки от генеральной средней. Данный интервал можно представить с 95%-ным уровнем вероятности следующим образом:

$$p\left[\mu - 1{,}96\frac{s}{\sqrt{n}} < \overline{X} < \mu + 1{,}96\frac{s}{\sqrt{n}}\right] = 0{,}95\,. \tag{5.7}$$

Небольшое алгебраическое преобразование в двойном неравенстве даст:

$$p\left[\overline{X} - 1{,}96\frac{s}{\sqrt{n}} < \mu < \overline{X} + 1{,}96\frac{s}{\sqrt{n}}\right] = 0{,}95\,. \tag{5.8}$$

Проиллюстрируем это с помощью конкретного примера. Допустим, что у нас есть данные по 60 месячным наблюдениям доходности по индексу FTSE 100. Выборочная средняя ежемесячной доходности равна 1,125% со средним квадратическим отклонением 2,5%. Каким будет 95%-ный доверительный интервал для этой средней?

Сначала необходимо рассчитать стандартную ошибку:

$$SE = \frac{2{,}5}{\sqrt{60}} = 0{,}3227.$$

Доверительный интервал будет:

$$\mu = 1{,}125 \pm 1{,}96 \cdot 0{,}3227;$$

$$\text{или } 1{,}125 - 0{,}6325 \le \mu \le 1{,}125 + 0{,}6325;$$

$$\text{или } 0{,}4925 \le \mu \le 1{,}7575\%.$$

Доверительный интервал можно проиллюстрировать с помощью рис. 5.2.

Уровень вероятности будет:

$$p\,[\overline{X} - (1{,}96 \cdot 0{,}3227) \le \mu \le \overline{X} + (1{,}96 \cdot 0{,}3227)] = 0{,}95;$$

$$p\,[[1{,}125 - 0{,}6325] \le \mu \le [1{,}125 + 0{,}6325]] = 0{,}95;$$

$$p\,[0{,}4925 \le \mu \le 1{,}7575] = 0{,}95.$$

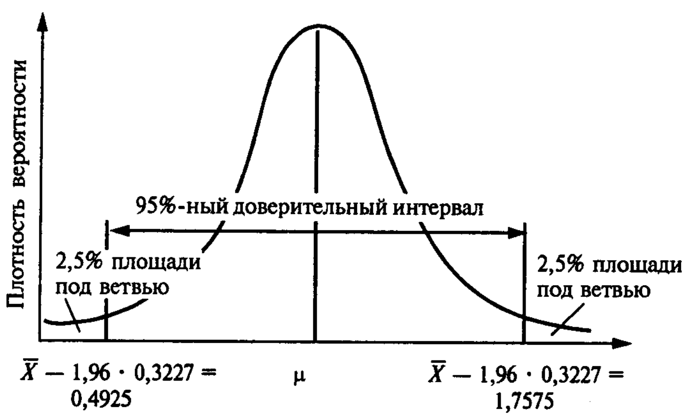

Рис. 5.2

Каким образом аналитик используют эту информацию? Может ли он принять решение относительно того, велик ли диапазон, содержащий в себе значение генеральной средней, с точки зрения практического применения доверительного интервала?

Например, приведенный выше доверительный интервал покрывает отрезок от 0,4925 до 1,7775%, т.е., размах составляет 1,265%, что больше по величине самой средней.

Ясно, что аналитик может не согласиться с таким широким размахом и предпочтет его уменьшить. При условии, что выборочная средняя фиксирована, а значение 1,96 привязано к 95%-ной вероятности, переменной, с помощью которой возможно изменить значение размаха, является стандартная ошибка, которая зависит от выборочного среднего квадратического отклонения и объема выборки. Таким образом, единственный способ уменьшить значение стандартной ошибки — это увеличить объем выборки.

Чтобы проиллюстрировать влияние объема выборки, рассмотрим пример, приведенный выше, но с объемом выборки, увеличенным до 120, выборочное среднее квадратическое отклонение останется неизменным —2,5. Тогда стандартная ошибка будет:

$$SE = \frac{2,5}{\sqrt{120}} = \frac{2,5}{10,95} = 0,2282\,.$$

А доверительный интервал составит:

$$\mu \doteq 1,125 \pm 1,96 \cdot 0,2282,$$

$$\text{или } 1,125 - 0,4473 \le \mu \le 1,125 + 0,4473,$$

$$\text{или } 0,6777 \le \mu \le 1,5723\%,$$

т.е. размах равен уже 0,8946.

Как поступать с малыми выборками?

Центральная предельная теорема может быть использована для доказательства утверждения о том, что выборочная средняя нормально распределена при условии, что объем выборки больше 30. В случае с малыми выборками необходимо допустить, что мы производим выборку из нормально распределенной совокупности для того, чтобы выборочная средняя была нормально распределена. Кроме того, только при выборках малого объема наша оценка генеральной дисперсии не будет надежной. В этом случае t-распределение позволит сделать поправку на эту дополнительную степень изменчивости.

Так же, как и нормальное распределение, t-распределение симметрично, но чуть более пологое. Действительная форма распределения зависит от числа степеней свободы, определяемых (n—1). С увеличением объема выборки t-распределение становится более похожим на нормальное.

Таким образом, двусторонний доверительный интервал для малой выборки будет представлен так:

$$\left[\bar{X} - t_{n-1,\alpha/2}\frac{S_{n-1}}{\sqrt{n}} < \mu < \bar{X} + t_{n-1,\alpha/2}\frac{S_{n-1}}{\sqrt{n}}\right], \tag{5.9}$$

где S_{n-1} — выборочное среднее квадратическое отклонение,
t_{n-1} — значение из таблицы t-распределения для выборки объема n и с n—1 степенями свободы,
α — уровень значимости.

Уровень вероятности, относящийся к этому доверительному интервалу, выглядит так:

$$p\left[\bar{X} - t_{n-1,\alpha/2}\frac{S_{n-1}}{\sqrt{n}} < \mu < \bar{X} + t_{n-1,\alpha/2}\frac{S_{n-1}}{\sqrt{n}}\right] = 1 - \alpha\,. \tag{5.10}$$

Для иллюстрации использования t-критерия приведем вычисление средней квартальной доходности для определенной группы менеджеров, работающих на фондовых рынках. Согласно проведенным 20 наблюдениям (т.е. 20—1 = 19 степеней свободы) выборочная средняя равна 4,5%. Выборочное среднее квадратическое отклонение составляет 5%. Для 95%-ного уровня доверия доверительный интервал будет:

$$\bar{X} - 2{,}093 \cdot 5/\sqrt{19} \le \mu \le \bar{X} + 2{,}093 \cdot 5/\sqrt{19}\,,$$

$$4{,}5 - (2{,}093 \cdot 1{,}47) \le \mu \le 4{,}5 + (2{,}093 \cdot 1{,}47),$$

$$4{,}5 - 2{,}401 \le \mu \le 4{,}5 + 2{,}401,$$

$$2{,}099 \le \mu \le 6{,}901\%.$$

Уровень вероятности имеет вид:

$$p[4{,}5 - 2{,}093 \cdot 1{,}47 \le \mu \le 4{,}5 + 2{,}093 \cdot 1{,}47] = 95\%,$$

$$p[4{,}5 - 2{,}401 \le \mu \le 4{,}5 + 2{,}401] = 95\%,$$

$$p[2{,}099 \le \mu \le 6{,}901] = 95\%.$$

Объем выборки

Мы уже видели, что на величину доверительного интервала влияет объем выборки, поэтому часто бывает полезно определить объем выборки, который бы обеспечил оценку параметра генеральной совокупности с необходимой степенью доверия. Формула для расчета n в этом случае выглядит так:

$$n = \left(\frac{zs}{e}\right)^2, \tag{5.11}$$

где n — необходимый объем выборки;
z — критическое значение из таблицы распределения, соответствующее необходимой степени доверия;
e — половина доверительного интервала, равная разности между μ и границей данного интервала;
s — оценка требуемой величины среднего квадратического отклонения генеральной совокупности.

Доверительный интервал для дисперсии

Выше мы отмечали, что выборочное распределение дисперсии следует, после соответствующего преобразования, χ^2-распределению. Для определения доверительных интервалов для дисперсии нам важно знать не столько само выборочное распределение дисперсии, сколько выборочное распределение этой величины, приведенное к стандартной форме следующим образом:

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}. \tag{5.12}$$

Для того чтобы найти 95%-ный доверительный интервал для точечной оценки дисперсии, мы должны определить значение χ^2, задающее по 2,5% в каждой из граничных площадей под кривой распределения (рис. 5.3). Таким образом, мы должны знать величину χ^2 для 97,5% значений, лежащих справа, и другую величину χ^2 — для 2,5% значений, лежащих справа. Если обозначить степень доверия через 1-α, тогда нам необходимы величины $\chi^2_{1-a/2}$ и $\chi^2_{\alpha/2}$. Если мы работаем с 95%-ным уровнем доверительной вероятности, тогда значение α будет 0,05, а значения хи-квадрат будут соответствовать $\chi^2_{n-1;0,975}$ и $\chi^2_{n-1;0,025}$.

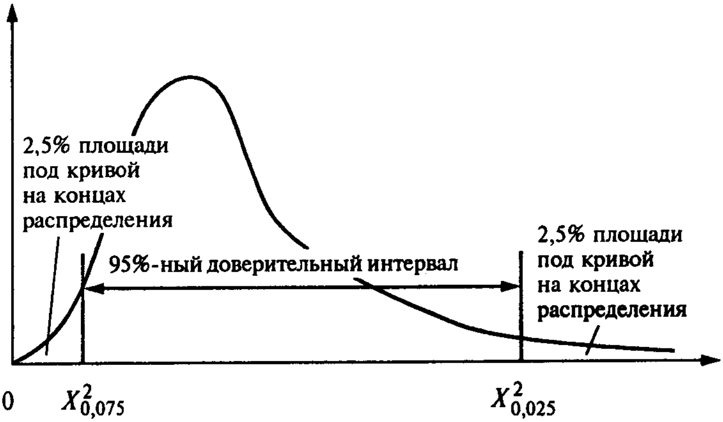

Рис. 5.3

Доверительный интервал определяется так:

$$\left[\frac{(n-1)s^2}{\chi^2_{n-1;1-\alpha/2}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{n-1;\alpha/2}}\right] \tag{5.13}$$

и уровень вероятности будет выглядеть следующим образом:

$$p\left[\frac{(n-1)s^2}{\chi^2_{n-1;1-\alpha/2}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{n-1;\alpha/2}}\right] = 1 - \alpha\,. \tag{5.14}$$

Для иллюстрации расчетов предположим, что согласно выборке объемом в 30 ежемесячных наблюдений дисперсия индекса FTSE 100 составляет 0,0225. Для 29 (n—1 = 29) степеней свободы критические значения χ^2-распределения при 2,5% в каждой из двух критических областей равны 45,72 и 16,05, тогда

$$P\left(\frac{29 \cdot 0{,}0225}{45{,}72} \le \sigma^2 \le \frac{29 \cdot 0{,}0225}{16{,}05}\right) = 0{,}95\,,$$

$$P\,(0{,}0143 \le \sigma^2 \le 0{,}04066) = 0{,}95.$$

Доверительный интервал для среднего квадратического отклонения находится извлечением квадратного корня из границ доверительного интервала для дисперсии, т.е.

$$P\,(0{,}1195 \le \sigma \le 0{,}2016) = 0{,}95.$$

ПРОВЕРКА ГИПОТЕЗ

Существует два подхода к проверке гипотез: **классический**, наиболее ранний, и подход на определении уровня вероятности (***P*-value**), который становится все более популярным с появлением более развитых и продвинутых пакетов прикладных программ. Сначала мы познакомимся с классическим подходом, после чего перейдем к *P*-value методу. Но прежде объясним некоторые термины.

Мы уже узнали, что при известных выборочном распределении различных описательных статистических показателей, объеме выборки и непосредственном значении самих показателей можно построить доверительные интервалы для точечных оценок. Но часто мы располагаем некоторыми предварительными (*a priori*) догадками или предположениями относительно величины параметров генеральной совокупности.

Располагая этими знаниями *a priori*, мы можем проверить гипотезу, что наша догадка действительно верна. **Статистическая гипотеза** — это рассматриваемое предположение о величине параметра распределения генеральной совокупности. Процесс проверки гипотез базируется на формулировании двух гипотез — **нулевой** и **альтернативной**, т.е. формулируются две конкурирующие гипотезы и проверяется, какая из них является верной.

Нулевая гипотеза (null hypothesis), обозначаемая обычно H_0, — это допущение, которое считается верным до тех пор, пока не будет доказано обратное исходя из результатов статистической проверки. **Альтернативная гипотеза** (alternative hypothesis), обычно обозначаемая через H_1, — это гипотеза, которая принимается, если в результате статистической проверки отвергается нулевая гипотеза.

Точная формулировка гипотезы зависит от того, что конкретно мы хотим установить. Например, представим, что просто необходимо знать, равен или нет параметр генеральной совокупности какому-либо значению, скажем, равна ли генеральная средняя μ величине μ_0. В этом случае гипотеза будет сформулирована следующим образом:

$$H_0: \mu = \mu_0, \qquad (5.15)$$
$$H_1: \mu \neq \mu_0.$$

Если же мы хотим знать, превышает или нет параметр генеральной совокупности данное значение, то гипотеза примет вид:

$$H_0: \mu = \mu_0, \qquad (5.16)$$
$$H_1: \mu > \mu_0.$$

В случае, когда необходимо узнать, является ли генеральная средняя меньше по величине μ_0, гипотеза будет выглядеть так:

$$H_0: \mu = \mu_0, \qquad (5.17)$$
$$H_1: \mu < \mu_0.$$

Для того чтобы проверить нашу гипотезу, необходимо выполнить статистическую проверку. **Статистическая проверка** (statistical test) состоит в использовании **стандартизованного статистического критерия** (standardized test statistic), вычисляемого по данным выборки для принятия решения о том, отвергнуть или нет гипотезу, проверяемую относительно величины параметра генеральной совокупности.

Стандартизованный статистический критерий

В гл. 2 мы узнали, что для того чтобы сравнить одну нормально распределенную переменную с другой, необходимо их привести к нормированной форме со средней, равной нулю, и средним квадратическим отклонением, равным единице. Распределение вероятностей такой нормированной переменной известно как нормированная функция кривой нормального распределения. При проверке гипотез мы должны привести статистический критерий к стандартизованной форме, чтобы сделать полноценное сравнение со стандартизованным нормальным распределением или t-распределением в случае проверки гипотезы для средних, или χ^2-распределением для дисперсии.

В случае критерия проверки гипотезы для средних вспомним, что если X нормально распределена (т.е. $X \sim N(\mu,\sigma^2)$), то $\overline{X}$

приближенно распределена как $\overline{X} \sim N(\mu, s^2/n)$. Так как форма нормального распределения зависит от величины средней арифметической и среднего квадратического отклонения, следует привести переменную к стандартизованной форме перед сравнением. В случае проверки гипотезы о величине средней приведение к стандартизованной форме осуществляется следующим образом:

$$\frac{\overline{X} - \mu_0}{s / \sqrt{n}}. \tag{5.18}$$

Эта величина называется **стандартизованным критерием проверки** (*standardized test statistic*). Если $\overline{X}$ равно μ_0, т.е. гипотеза H_0 верна, критерий проверки имеет стандартизованное нормальное распределение (если объем выборки большой) или стандартизованное t-распределение с n—1 степенями свободы (для малой выборки). Оба эти распределения имеют значения средней, равные нулю, и стандартную ошибку (так как они относятся к выборочным распределениям), равную 1. Приведя критерий проверки к стандартизованной форме, можно сравнить его значение напрямую со значениями, относящимися к соответствующим стандартизованным распределениям вероятностей.

Если критерий проверки, полученный в результате стандартизации, расположен в критической области распределения, это доказывает, что средняя $\overline{X}$ не равна μ_0, т.е. наше предположение неверно и H_0 ложна. Смысл сказанного заключается в том, что если критерий проверки находится в критической области, это событие маловероятно согласно нашим допущениям. Однако, как выше доказано, если такое событие произошло, мы должны подвергнуть сомнению принятые допущения.

Проверка гипотезы может быть **односторонней** (one-tailed test) или **двусторонней** (two-tailed test). **Односторонний критерий** используется в тех случаях, когда необходимо знать, является ли параметр генеральной совокупности строго больше (правосторонний критерий) или строго меньше (левосторонний критерий) предполагаемого значения. Гипотезы (5.16) и (5.17) будут проверяться с помощью односторонних критериев. **Двусторонний критерий** приложим к тем случаям, когда нас интересует, отличаются ли реальные значения параметра от предполагаемого значения. Например, гипотезы, представленные выражением (5.15), проверяют с помощью двустороннего критерия.

Критическую область (critical region) составляют те значения выборочных статистических показателей, которые ведут к отказу от нулевой гипотезы. На рис. 5.4 критическая область представлена граничными площадями под кривой.

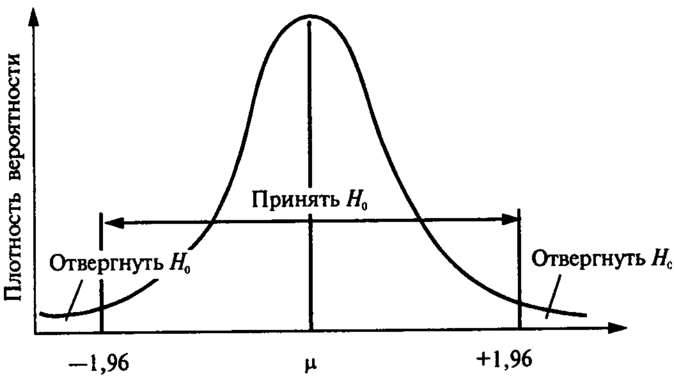

Рис. 5.4

Уровень значимости (significance level) гипотезы — это вероятность того, что критерий проверки находится в критической области при условии верности гипотезы H_0 — обычно равен 5 или 1%. Если проверка при 5%-ном уровне значимости приведет к отказу от гипотезы H_0, то говорят, что величина критерия значима. Если H_0 отвергается при 1%-ном уровне, используется термин "высокий уровень значимости".

Концепция критической области и уровня значимости может быть проиллюстрирована с помощью рис. 5.4, на котором изображен доверительный интервал для средней. Критические области составляют площади под кривой, находящиеся левее −1,96 и правее +1,96. Доверительный интервал — это площадь под кривой, находящаяся между этими двумя точками. Если $\overline{X}$ превосходит в положительном или отрицательном направлении 1,96 стандартной ошибки от μ_0, значит $\overline{X}$ настолько велика (или мала) по сравнению с предполагаемой μ_0, что существует очень малая вероятность (отражаемая площадью в граничных областях), что средняя арифметическая выборки репрезентирует μ_0.

Правило принятия решения (decision rule) для проверки статистической гипотезы — это модель расчета значений выборочных статистических показателей, на основании которых принимается или отвергается нулевая гипотеза.

Прежде чем продолжить объяснение, рассмотрим два очень важных типа ошибок, встречающихся при проверке гипотез.

Ошибки I и II рода

В процессе проверки гипотезы существует вероятность того, что нулевая гипотеза будет отвергнута, когда в действительности она должна быть принята. Это называется **ошибкой I рода**. Вероятность допущения ошибки I рода — это **уровень значимости**. Таким образом, когда мы выбираем 5%-ный уровень значимости для проверки, одновременно мы допускаем, что в 5% случаев мы отвергнем нулевую гипотезу тогда, когда фактически нам следовало бы принять ее.

Второй вид ошибок имеет место при принятии нулевой гипотезы в то время, как в действительности она должна быть отвергнута. Такая ошибка называется **ошибкой II рода**.

Для более четкого объяснения связи между ошибкой I рода и ошибкой II рода рассмотрим аналогию, относящуюся к судебной системе присяжных заседателей и отображенную графически на рис. 5.5.

		Действительность	
		H_0 верна Обвиняемый невиновен	H_0 ложна Обвиняемый виновен
Решение	H_0 принята Обвиняемый освобожден	✓	Ошибка II рода
	H_0 отвергнута Обвиняемый наказан	Ошибка I рода	✓

Рис. 5.5

Обвиняемый может быть либо невиновен, либо виновен, следовательно, присяжные могут принять решение о его виновности или невиновности. По сути присяжные проверяют нулевую гипотезу о невиновности обвиняемого. Если присяжные считают, что обвиняемый виновен, когда он в действительности является таковым, то было принято верное решение и никакой ошибки не произошло. Аналогично если присяжные считают обвиняемого невиновным, когда это в действительности так, то и в этом случае было принято правильное решение. Но если обвиняемый невиновен, а присяжные принимают решение о его виновности, т.е. отвергают нулевую гипотезу, когда фактически она верна — это основная ошибка — **ошибка I рода**. Этой ошибки стараются избегать прежде всего, поэтому уровень значимости проверки устанавливается с целью уменьшения вероятности совершения этой ошибки (обычно 5 или 1%).

Установив уровень значимости, мы не измеряем степень риска, связанного с совершением ошибки II рода (признание виновного человека невиновным).

Проверка гипотезы о величине генеральной средней

Двусторонняя проверка для средней

Если просто нужно проверить, равна ли генеральная средняя выборочной средней, гипотеза формулируется следующим образом:

$$H_0: \ \mu = \mu_0, \qquad H_1: \ \mu \neq \mu_0, \tag{5.19}$$

и затем стандартизованный критерий используется в следующей процедуре для проверки этой гипотезы:

1) определить уровень значимости для проверки, который обычно равен 10, 5 или 1%, т.е. каждая из граничных областей под кривой распределения будет равна соответственно 5, 2,5 и 0,5%;
2) подставить значение μ_0 в нулевую гипотезу;
3) установить соответствующее критическое значение z (или t в случае проверки по малой выборке) исходя из таблиц, отражающих процентные величины для граничных областей в соответствии с выбранным уровнем значимости;
4) применить следующее **правило принятия решения**:

$$\text{Принять } H_0, \text{ если } -z \leq \frac{\bar{X} - \mu_0}{\sqrt{s^2 / n}} \leq z, \tag{5.20}$$

В противном случае отвергнуть H_0.

Для лучшего представления этой процедуры вспомним, что для нормально распределенной переменной 95% всех значений наблюдений будет находиться в интервале не более, чем плюс/минус 1,96 среднего квадратического отклонения от средней. Аналогично если $\overline{X}$ статистически эквивалентно μ_0, с 95%-ной уверенностью стандартизованный критерий проверки будет лежать в интервале не более, чем плюс/минус 1,96 стандартной ошибки от нулевого значения. Число 1,96 является критическим значением для статистической проверки гипотезы с 5%-ным уровнем значимости. Если критерий проверки имеет значение большее, чем + 1,96, или меньшее, чем —1,96, это служит доказательством того, что для 5%-ного уровня значимости $\overline{X}$ — это не то же самое, что μ_0.

Для иллюстрации рассмотрим проверку того, будет ли средняя месячная доходность в 2,4%, полученная управляющим портфелем ценных бумаг, статистически значимо отличаться от среднего уровня в промышленности, составляющего 2,3%. В первую очередь определим критическое значение для критерия проверки в соответствии с заданным уровнем значимости. Для двусторонней проверки это плюс/минус 1,64, 1,96 и 2,58 для случая нормального распределения. Располагая значением $\overline{X}$ в 2,4%, а μ_0 — 2,3%, при 36 наблюдениях со средним квадра тическим отклонением 1,7%, получим значение z, равное 0,3529:

$$\frac{2{,}4 - 2{,}3}{1{,}7 / \sqrt{36}} = 0{,}3529 .$$

Если размер выборки более 30, можно допустить нормальное распределение.

Так как 0,3529 находится между всеми приведенными критическими положительными и отрицательными значениями, можно заключить, что $\overline{X}$ = 2,4% статистически незначимо отличается от μ_0 = 2,3%; таким образом, разница в 0,1% — чистая случайность.

Классическая односторонняя проверка

Правосторонняя проверка

Если необходимо проверить, **больше ли** генеральная средняя μ заданной величины μ_0, следует выполнить проверку с помощью одностороннего критерия, так как нас интересует только одно — больше ли генеральная средняя определенного заданного значения. Формулируем нулевую гипотезу о том, что μ равна μ_0, и альтернативную гипотезу о том, что μ больше, чем μ_0. Нулевую и альтернативную гипотезы запишем так:

$$H_0: \mu = \mu_0, \qquad (5.21)$$
$$H_1: \mu > \mu_0.$$

В данном случае будет осуществляться **правосторонняя проверка** — проверка вероятности того, что μ находится в правой части критической области распределения, так как в этом случае μ будет больше проверяемого значения μ_0, и нулевая гипотеза будет отклонена.

Далее вычисляется стандартизованный критерий проверки таким же образом, что и для двусторонней проверки за исключением того, что уровень значимости будет относиться только к правой части распределения. Правило принятия решения выглядит так:

$$\text{Принять } H_0, \quad \text{если } \frac{\bar{X}-\mu_0}{s/\sqrt{n}} \le z, \qquad (5.22)$$

$$\text{Отвергнуть } H_0, \text{ если } \frac{\bar{X}-\mu_0}{s/\sqrt{n}} \ge z.$$

Для того чтобы понять смысл этого правила, рассмотрим интерпретацию критерия проверки, большего z. Это означает, что $\bar{X}$ больше, чем μ_0, на величину, которая сдвигает ее в критическую область или область отказа. $\bar{X}$ настолько велика, что вероятность для нее быть равной генеральной средней меньше уровня значимости, установленного для проверки гипотезы. Поэтому мы отказываемся от нулевой гипотезы. Если критерий проверки был бы меньше z, это означало бы, что $\bar{X}$ не настолько велика, чтобы попасть в критическую область. Поэтому в данном случае нулевая гипотеза не была бы отвергнута.

В качестве примера предположим, что проверяем гипотезу о том, что средняя месячная доходность по индексу FTSE 100 за данный период составила более 1,2%. По данным 60 наблюдений была рассчитана средняя арифметическая, которая составила 1,25%, и среднее квадратическое отклонение — 2,5%. Тогда критерий проверки составит:

$$\frac{1{,}25-1{,}20}{2{,}5/\sqrt{60}} = 0{,}15492.$$

Поскольку размер выборки большой, можно допустить нормальное распределение выборочной средней.

По таблицам нормального распределения находим, что для 10, 5 и 1% уровней значимости односторонние значения критерия составляют 1,28, 1,64 и 2,33 соответственно. Таким образом, критерий проверки, величина которого составляет 0,1549, не является значимым.

Левосторонняя проверка

Если мы хотим проверить гипотезу о том, что μ меньше μ_0, выполняется левосторонняя проверка вероятности того, что μ находится в левой части распределения. Нулевая и альтернативная гипотезы в этом случае будут:

$$H_0:\ \mu = \mu_0, \qquad (5.23)$$
$$H_1:\ \mu < \mu_0,$$

а критерий для принятия или отказа от гипотезы имеет вид:

$$\text{Принять } H_0, \text{ если } \frac{\bar{X} - \mu_0}{s / \sqrt{n}} \geq z, \qquad (5.24)$$

$$\text{Отвергнуть } H_0, \text{ если } \frac{\bar{X} - \mu_0}{s / \sqrt{n}} < z.$$

Для иллюстрации предположим, что необходимо проверить, что средняя месячная доходность по индексу S&P 50 меньше, чем 1,30%. Известно, что согласно 75 проведенным наблюдениям средняя составила 1,18% со средним квадратическим отклонением 2,2%. Критерий проверки рассчитывается так:

$$\frac{1{,}18 - 1{,}30}{2{,}2 / \sqrt{75}} = -0{,}4724.$$

Опять, выбирая уровни значимости 10, 5 и 1%, получаем критические значения —1,28, —1,64 и —2,33 соответственно. Можно видеть, что критерий проверки не является значимым, следовательно, нет оснований для отказа от нулевой гипотезы.

Проверка гипотезы о величине дисперсии

Стандартизованный критерий проверки для генеральной дисперсии выглядит следующим образом:

$$\frac{(n-1)s^2}{\sigma_0^2}, \tag{5.25}$$

где σ_0^2 — проверяемое значение дисперсии.

Мы уже отметили выше, что стандартизованный критерий проверки следует χ^2-распределению. Правила принятия решения для левосторонней, правосторонней и двусторонней проверок по χ^2-распределению даны ниже.

Для левосторонней проверки

Нулевая и альтернативная гипотезы задаются следующим образом:

$$H_0:\ s^2 = \sigma^2, \tag{5.26}$$
$$H_1:\ s^2 < \sigma^2.$$

Правило принятия решения:

$$\text{Принять } H_0\text{, если } \frac{(n-1)s^2}{\sigma_0^2} \geq \chi^2_{(1-\alpha)}, \tag{5.27}$$

$$\text{Отвергнуть } H_0\text{, если } \frac{(n-1)s^2}{\sigma_0^2} < \chi^2_{(1-\alpha)},$$

где α — выбранный уровень значимости.

Для иллюстрации предположим, что проверяем гипотезу о том, что дисперсия по акции *B* меньше 25. Выборочная дисперсия составила 23, а число наблюдений равно 40 (следовательно, количество степеней свободы будет 40—1 = 39). Так как таблицы χ^2-распределения дают значение вероятностей для левой части распределения, то для левосторонней проверки с уровнем значимости в 5% правая часть площади под кривой распределения будет составлять 95%. Критическое значение χ^2 для данной ситуации с 39 степенями свободы приближенно равно 26,5. Критерий проверки будет:

$$\frac{39 \cdot 23}{25} = 35{,}88.$$

Так как критерий проверки 35,88 больше критического значения 26,5, нулевая гипотеза принимается, т.е. в данном случае нет достаточных доказательств для вывода о том, что дисперсия по акции B меньше 25.

Для правосторонней проверки

$$\text{Принять } H_0, \text{ если } \frac{(n-1)s^2}{\sigma_0^2} \le \chi_\alpha^2, \tag{5.28}$$

$$\text{Отвергнуть } H_0, \text{ если } \frac{(n-1)s^2}{\sigma_0^2} > \chi_\alpha^2.$$

Допустим, что нужно проверить гипотезу, что дисперсия доходности по облигации A больше 8%. Рассчитанная на основе из 35 наблюдений выборочная дисперсия составила 9%. Тогда критерий проверки будет:

$$\frac{34 \cdot 9}{8} = 38{,}25.$$

При 95%-ном уровне значимости правая часть распределения будет соответствовать 5%. Следовательно, критическое значение χ^2 с 34 степенями свободы приближенно равно 48,6. В результате мы принимаем нулевую гипотезу и делаем вывод, что дисперсия доходности по облигации A равна 8%.

Для двусторонней проверки

$$\text{Принять } H_0, \text{ если } \chi^2_{(1-(\alpha/2))} \le \frac{(n-1)s^2}{\sigma_0^2} \le \chi^2_{(\alpha/2)}, \tag{5.29}$$

Отвергнуть нулевую гипотезу в противном случае.

Если бы нужно было проверить утверждение, что дисперсия облигации B равна 7, мы должны были бы найти критические значения χ^2 для 0,975 и 0,025. Для 34 степеней свободы они приближенно равны 19,8 и 51,9 соответственно. Так как критерий проверки, равный

$$\frac{34 \cdot 9}{7} = 43{,}71,$$

находится между этими двумя критическими значениями, принимается нулевая гипотеза, что дисперсия равна 7. Можно заметить, что в приведенных выше примерах многие нулевые гипотезы не могли быть отвергнуты при наличии определенных свидетельств. Фактически же набор проверяемых значений, которые не могут быть отвергнуты, — это соответствующий доверительный интервал.

Проверка гипотезы методом определения уровня вероятности

Величина p — это значение, которое в случае верности нулевой гипотезы представляет собой вероятность получения величины стандартизованного критерия проверки, большего по абсолютному значению, чем рассчитанный критерий проверки.

В случае односторонней проверки p равно площади под кривой слева (левосторонняя проверка) или справа (правосторонняя проверка) от значения критерия проверки. В случае двусторонней проверки оно равно удвоенной площади в части под кривой справа или слева от критерия проверки.

В методе P-value правило принятия решения одинаково независимо от того, выполняется левосторонняя, правосторонняя или двусторонняя проверка. Обозначив степень значимости для проверки через α, получаем следующее правило принятия решения:

$$\text{Принять } H_0, \text{ если } P\text{-value} \geq \alpha, \tag{5.30}$$

$$\text{В противном случае отвергнуть } H_0.$$

При двусторонней проверке величину p, рассчитанную с помощью компьютерных программ, может быть необходимо удвоить для того, чтобы определить вероятность получения критерия проверки, большего по абсолютному значению, чем тот, который был рассчитан.

Расчет величины p

Для того чтобы найти величину p, прежде всего рассчитывают стандартизованный критерий проверки, а затем, зная число степеней свободы, находят вероятности (площади в граничных областях), соответствующие показателям t или z, которые "захватывают в вилку" (охватывают сверху и снизу) рассчитанный критерий проверки. После этого с помощью интерполяции исходя из полученных вероятностей находят величину p.

Проиллюстрируем расчет с использованием таблиц t-распределения (объем выборки, данный в нашем примере, позволяет допустить нормальное распределение, но это допущение не является необходимым). Представим себе, что мы хотим найти такой вид инвестирования, который бы давал твердый доход по крайней мере 13,2%. Нужно проверить, является или нет выборочная средняя доказательством того, что реальный доход больше проверяемого значения средней доходности. Предположим, что средняя месячная доходность, приведенная к годовому базису, по данному индексу облигации составляет 14,4%, а выборочное среднее квадратическое отклонение равно 2,915%. Было проведено 40 наблюдений, известно также, что доходность нормально распределена. Тогда проверяемые гипотезы будут выглядеть так:

$$H_0:\ \mu = 13{,}2,$$
$$H_1:\ \mu > 13{,}2.$$

Критерий проверки находится следующим образом:

$$\frac{(\bar{X}-\mu)}{\frac{s}{\sqrt{n}}} = \frac{14{,}4-13{,}2}{\frac{2{,}915}{\sqrt{40}}} = \frac{1{,}2}{0{,}461} = 2{,}604.$$

Для 39 (40—1) степеней свободы значение $t = 2{,}423$ соответствует 1% в граничной области, когда как значение 2,704 соответствует 0,5% в граничной области. Тогда вероятность, соответствующая значению $t = 2{,}604$, приближенно может быть определена с помощью следующего процесса интерполяции:

1) необходимо интерполировать позицию значения $t = 2{,}604$ в интервале между 0,01 и 0,005, тогда

$$\frac{2{,}604-2{,}423}{2{,}704-2{,}423} = 0{,}644,$$

т.е. это чуть выше середины между 2,704 и 2,423;

2) для определения величины p следует умножить число 0,644, полученное в результате интерполяции, на разницу между двумя уровнями значимости, а затем вычесть полученный результат из величины большего уровня значимости. Тогда величина p равна $\approx 0{,}0068$:

$$0{,}01 - (0{,}644 \cdot (0{,}01 - 0{,}005)) \approx 0{,}01 - 0{,}0032 \approx 0{,}0068.$$

Таким образом, вероятность получения значения критерия проверки большего, чем 2,604, равна 0,0068. Согласно правилу гипотеза H_0 принимается, если p больше α. Так как 0,0068<0,05, нулевая гипотеза, что $\mu = \mu_0$, отвергается.

p-проверка для дисперсии

Вспомним, что для левосторонней проверки величины дисперсии дохода по акции *B* полученное критическое значение было 35,88 (для 39 степеней свободы). Согласно статистическим таблицам для 39 степеней свободы значение 33,93 будет соответствовать 30% в левой граничной области, когда как значение 36,16—40% в левой граничной области.

Для интерполяции вероятности, соответствующей критическому значению 35,88, сначала сделаем следующие вычисления:

$$\frac{35{,}88 - 33{,}93}{36{,}16 - 33{,}93} = \frac{1{,}95}{2{,}23} = 0{,}8744.$$

Затем 0,8744 умножается на разницу вероятностных уровней, содержащую найденное критическое значение, в нашем примере это 10, и прибавляется к наименьшему значению, т.е.:

$$30 + (0{,}8744 \cdot 10) = 38{,}74.$$

Таким образом, величина p равна 0,3874; так как это больше, чем 5%-ный уровень значимости, установленный в начале проверки, мы принимаем нулевую гипотезу, как и раньше.

Проверка степени соответствия

С помощью таблиц χ^2-распределения можно определить, насколько хорошо выборка или генеральная совокупность соответствует данному распределению вероятностей, например, нормальному, биномиальному или распределению Пуассона.

Общие правила, лежащие в основе такой проверки, состоят в следующем. Следует построить предполагаемое распределение в соответствии с определяющими его параметрами, например средней и средним квадратическим отклонением для нормального распределения. Затем это ожидаемое распределение разделяется на равнопропорциональные интервалы, но с единственным ограничением — ни один интервал не должен содержать менее 5 наблюдений. После этого определяется количество фактических наблюдений, попадающих в каждый из равнопропорциональных интервалов. Затем с помощью критерия χ^2 сравнивается сумма разностей в квадрате между количеством ожидаемых и фактических наблюдений, деленная на количество ожидаемых наблюдений.

Для демонстрации этой техники рассмотрим данные по месячной доходности по индексу FTSE 100, приведенные в приложении 5.2. Средняя доходность за месяц составила 0,6905%, а дисперсия — 4,7080. Было сделано 50 наблюдений.

Начнем с определения значения X, которое будет соответствовать 0,1 в граничной области.

Нам известно, что $X = zs + \mu$ (где s — выборочное среднее квадратическое отклонение). Для того чтобы построить равные подходящие интервалы (это удобно, но не является необходимым), следует использовать таблицы нормального распределения для определения значения z, которое относится к 10% в левой граничной области; этим значением является 1,282. Значение z, которое соответствует 20% в левой граничной области, равно 0,842. Полный перечень значений, начиная с 10% и кончая 90%, с интервалами в 10% приведен ниже:

Процент в граничной области слева	Значение z
10	—1,282
20	—0,842
30	—0,524
40	—0,253
50	0
60	0,253
70	0,524
80	0,842
90	1,282

Значение X для первого интервала (с 10% в левой граничной области) рассчитывается так:

$$\overline{X} = -1{,}282 \cdot \sqrt{4{,}708} + 0{,}6905 = -1{,}282 \cdot 2{,}1698 + 0{,}6905 =$$

$$= -2{,}7817 + 0{,}6905 = -2{,}0912.$$

Вычисления производятся аналогичным образом до тех пор, пока не будут найдены значения X для всех интервалов. В нашем примере при $s = 4{,}7080$ и $\mu = 0{,}6905$ они равны:

$$-2{,}0912;\ -1{,}1365;\ -0{,}4465;\ 0{,}1415;\ 0{,}6905;$$

$$1{,}239;\ 1{,}827;\ 2{,}517;\ 3{,}472.$$

Следующий шаг — построение распределения частот по имеющимся данным путем определения числа фактических наблюдений для каждого из интервалов. Это показано в табл. 5.1.

Так как интервалы равнопропорциональны, можно ожидать, что в каждом из них должно содержаться равное число наблюдений. Мы можем определить отклонение количества фактических наблюдений от ожидаемых. Так как одни разности будут положительными, а другие отрицательными, необходимо их возвести в квадрат точно так же, как и для расчета дисперсии. Затем каждый из полученных результатов делится на ожидаемое количество наблюдений в отдельном интервале. После этого расчетные величины для всех интервалов суммируются, а результат должен следовать χ^2-распределению с $k - r - 1$ степенями свободы.

k относится к количеству интервалов, определяемых в ходе анализа, а r равно числу параметров, рассчитанных по данным выборки (средняя и среднее квадратическое отклонение в случае нормального распределения).

Критерий проверки для этой гипотезы выглядит так:

$$\sum\left[\frac{(O_i - E_i)^2}{E_i}\right], \tag{5.31}$$

где O_i — фактическая частота для i -го интервала,
E_i — ожидаемая частота для i -го интервала.

Таким образом, проверяется гипотеза о том, что выборка была произведена из нормально распределенной совокупности:

H_0 — распределение признаков генеральной совокупности подчиняется нормальному закону;

H_1 — распределение признаков генеральной совокупности не является нормальным.

Правило принятия решения выглядит следующим образом:

$$\text{Принять } H_0, \text{ если } \sum\left[\frac{(O_i - E_i)^2}{E_i}\right] \le \chi^2_{k-r-1,(\alpha)}, \tag{5.32}$$

Отвергнуть H_0 в противном случае.

Рассчитанный критерий проверки для этого примера приведен в итоговой строке последней колонки табл. 5.1, он равен 33,6. Однако для 5%-ного уровня значимости (5% в граничной области) значение χ^2 для 7 степеней свободы (10—2—1) составляет 14,06713. Поскольку рассчитанный критерий χ^2 больше, чем критическое значение для 5% уровня значимости, мы отвергаем нулевую гипотезу.

Таблица 5.1

Доход	Частота (O)	Ожидаемая частота (E)	$(O-E)^2$	$(O-E)^2/E$
<—2,0912	14	5	81	16,2
от —2,0912 до —1,1365	3	5	4	0,8
от —1,1365 до —0,4465	2	5	9	1,8
от —0,4465 до 0,1415	4	5	1	0,2
от 0,1415 до 0,6905	2	5	9	1,8
от 0,6905 до 1,239	2	5	9	1,8
от 1,239 до 1,827	4	5	1	0,2
от 1,827 до 2,517	3	5	4	0,8
от 2,517 до 3,472	4	5	1	0,2
3,472 и более	12	5	49	9,8
			Сумма $(O-E)^2/E$	33,6

УПРАЖНЕНИЯ

1. Охарактеризуйте различия между выборочными статистическими показателями и параметрами генеральной совокупности.
2. Что вы понимаете по терминами "выборочное распределение выборочной средней" и "выборочное распределение выборочной дисперсии". Рассчитайте стандартную ошибку средней по отношению к доходу по финансовому индексу со средним значением в 10% и средним квадратическим отклонением 16% на основе 60 наблюдений.
3. Охарактеризуйте значение и важность BLUE характеристик для оценок.

4. Объясните, что значит следующее выражение: "95%-ный доверительный интервал для генеральной средней находится между 0,046 и 0,152".
5. Используя данные из п. 2, рассчитайте 95%-ный доверительный интервал для выборочной средней и выборочного среднего квадратического отклонения.

   Повторите расчеты для 90%-ного доверительного интервала. Сравните длины этих доверительных интервалов. Заметьте, что чем меньше степень доверия, тем больше точность, и наоборот.
6. Выборка из 25 еженедельных наблюдений за доходами по индексу FTSE 100 характеризуется средней 0,005 и средним квадратическим отклонением 0,02. Предположив, что недельные доходы нормально распределены, рассчитайте 95%-ный доверительный интервал для средней доходности.
7. Используя данные п. 6, определите, каким должен быть размер выборки для того, чтобы оценить недельный доход с 95%-ной вероятностью и с максимальной ошибкой не более 0,004.
8. Используя данные п. 6, найдите 95%-ный доверительный интервал для среднего квадратического отклонения недельных доходов.
9. Что вы понимаете под следующим:

   нулевая гипотеза; альтернативная гипотеза; стандартизованный критерий проверки; критическая область; уровень значимости; правило принятия решения; односторонняя проверка; двусторонняя проверка; величина p; ошибка I рода; ошибка II рода.
10. Объясните зависимость между доверительным интервалом и соответствующей проверкой гипотезы.
11. Компания ABC Fund Managers Plc заявила, что месячный доход по ее высокодоходному инвестиционному фонду превысил доход индекса на 0,3%, или на 0,003. В течение одногодичного периода средний доход по индексу составил 0,005, а средний доход фонда — 0,0065, среднее квадратическое отклонение равно 0,019. Выполните одностороннюю статистическую проверку, чтобы выяснить, насколько верно заявление компании.
12. Используя данные п. 11, выполните одностороннюю проверку гипотезы, что доход по портфелю фактически превысил доход по индексу.
13. Управляющий портфелем заботится о том, чтобы не осуществлять инвестиционные вложения в ценные бумаги с дисперсией годовой доходности более чем 0,04. Выборка из 52 наблюдений за доходом по активу "А" показала, что дисперсия равна 0,045. Проверьте гипотезу, что доход по активу "А" характеризуется дисперсией, меньшей или равной 0,04.

14. Объясните, что вы понимаете под термином *P*-value.

15. Том и Дик по очереди готовят чай для старшего женского персонала фирмы биржевых брокеров. Для определения очередности они подбрасывают монету и держат для этого случая специальную монету. В последних 200 случаях Том проиграл 116 раз. Должен ли он настаивать на использовании в дальнейшем другой монеты?

ОТВЕТЫ К ИЗБРАННЫМ ВОПРОСАМ

2. $\text{SE}^* = 0{,}16/\sqrt{(60)} \approx 0{,}02.$

5. 95%-ный доверительный интервал для средней:

$$0{,}10 \pm 1{,}96\frac{0{,}16}{\sqrt{60}} \approx 0{,}10 \pm 0{,}04,$$

т.е. от 0,06 до 0,14, или от 6% до 14%.

95%-ный доверительный интервал для среднего квадратического отклонения:

$$\frac{59 \cdot 0{,}16^2}{\chi^2_{59(0{,}975)}} < \sigma^2 < \frac{59 \cdot 0{,}16^2}{\chi^2_{59(0{,}925)}},$$

$$\frac{59 \cdot 0{,}16^2}{83{,}3} < \sigma^2 < \frac{59 \cdot 0{,}16^2}{40{,}5},$$

$$0{,}0181 < \sigma^2 < 0{,}0373,$$

$$0{,}135 < \sigma < 0{,}193.$$

90%-ный доверительный интервал для средней:

$$0{,}10 \pm 1{,}64\frac{0{,}16}{\sqrt{60}} \approx 0{,}10 \pm 0{,}034,$$

т.е. от 0,066 до 0,134, или от 6,6% до 13,4%.

90%-ный доверительный интервал для среднего квадратического отклонения:

---

* SE — стандартная ошибка.

$$\frac{59 \cdot 0{,}16^2}{\chi^2_{59(0,95)}} < \sigma^2 < \frac{59 \cdot 0{,}16^2}{\chi^2_{59(0,05)}},$$

$$\frac{59 \cdot 0{,}16^2}{79{,}1} < \sigma^2 < \frac{59 \cdot 0{,}16^2}{43{,}2},$$

$$0{,}0191 < \sigma^2 < 0{,}0350,$$

$$0{,}138 < \sigma < 0{,}187.$$

6.

$$0{,}005 \pm t_{24(0,975)} \cdot \frac{0{,}02}{\sqrt{25}} = 0{,}005 \pm 2{,}064 \cdot \frac{0{,}02}{\sqrt{25}},$$

от —0,0033 до + 0,0133.

7. Необходимо, чтобы

$$1{,}96 \cdot \frac{0{,}02}{\sqrt{n}} = 0{,}004,$$

$$\text{тогда } n = \left(\frac{1{,}96 \cdot 0{,}002}{0{,}004}\right)^2 \approx 96.$$

8. 95%-ный доверительный интервал для среднего квадратического отклонения:

$$\frac{24 \cdot 0{,}02^2}{\chi^2_{24(0,975)}} < \sigma^2 < \frac{2480{,}02^2}{\chi^2_{24(0,025)}},$$

$$\frac{24 \cdot 0{,}02^2}{39{,}3641} < \sigma^2 < \frac{24 \cdot 0{,}02^2}{12{,}4012},$$

$$0{,}000244 < \sigma^2 < 0{,}000774,$$

$$0{,}0156 < \sigma < 0{,}0278.$$

10. Доверительный интервал — это набор всех возможных величин параметра, которые не будут отвергнуты как предполагаемые значения при соответствующей проверке.

11.

$$t = \frac{0{,}0065 - 0{,}008}{0{,}019 / \sqrt{12}} = -0{,}2735,$$

$$t_{11(0,05)} = 1{,}796,$$

т.е. гипотеза отвергается.

12.

$$t = \frac{0{,}0065 - 0{,}005}{0{,}019 / \sqrt{12}} = 0{,}2735,$$

$$t_{11(0,05)} = 1{,}796,$$

т.е. отвергается гипотеза, что их доход равен доходу по индексу.

13. Статистический критерий

$$\frac{51 \cdot 0{,}045^2}{0{,}04^2} = 64{,}55,$$

$$\chi^2_{51(0,05)} = 67{,}5,$$

следовательно, нет достаточных доказательств, чтобы отвергнуть гипотезу о том, что дисперсия не превышает 0,04.

14. Это вероятность для гипотезы H_0 получения величины, по абсолютному значению равной критерию проверки или превышающей его.

15. Отвергнуть гипотезу, что монета имеет дефект.

$$\chi^2 = \frac{(-16)^2}{100} + \frac{16^2}{100} = 5{,}12,$$

$$\chi^2_{1(0,05)} = 3{,}84.$$

Замечание. Для $\nu = 1$ (степень свободы) необходимо сделать следующую поправку. Абсолютные рассчи-танные и ожидаемые значения должны быть уменьшены на 0,5. Тогда

$$\chi^2 = \frac{(-15{,}5)^2}{100} + \frac{15{,}5^2}{100} = 4{,}81,$$

что по-прежнему является значимым.

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

Bowers, D. (1991) *Statistics for Economics and Business*. Macmillan, London.

Curwin, J. and Slater, R. (1996) *Quantitative Methods for Business Decisions*, 4th edn. Chapman & Hall, London.

Silver, M. (1992) *Business Statistics*. McGraw-Hill, London.

ПРИЛОЖЕНИЕ 5.1. Стандартная ошибка средней

Рассмотрим случайную переменную X, определяемую следующим образом:

Возможное значение	—1	1
Вероятность	0,5	0,5

(мы с этим уже встречались в гл. 4).

Мы можем промоделировать процесс, обозначенный переменной X, многократным подбрасыванием монеты и фиксированием значений: 1 — при выпадении орла, —1 — при выпадении решки. Проделаем этот эксперимент 16 раз, рассчитывая ожидаемые и фактически полученные значения параметров.

Ожидаемые значения

$$E(X) = 0{,}5 \cdot (-1) + 0{,}5 \cdot 1 = 0$$

$$\mathrm{Var}(X) = 0{,}5 \cdot (-1)^2 + 0{,}5 \cdot 1^2 = 1 = \sigma^2$$

Фактические значения

1, —1, —1, 1, 1, 1, —1, —1, —1, —1, —1, 1, 1, —1, 1, —1

$\bar{X} = -0{,}125$

$s^2 = 1{,}05$ (заметьте, что деление на 15 вместо 16 позволяет в некоторой степени компенсировать использование —0,125 вместо 0 в качестве средней).

А сейчас выясним, что случится, если вместо индивидуальных осуществлять парные наблюдения и для каждого из них находить среднюю. В гл. 3 уже показано, как производить сложение случайных переменных и их умножение на константу. Наш процесс нахождения средней заключается в сложении двух идентичных случайных переменных и умножении полученной суммы на 0,5. Применение методов, описанных в гл. 3, дает следующее распределение:

Возможное значение	—1	0	1
Вероятность	0,25	0,5	0,25

Мы можем использовать те же данные, что и ранее, для сравнения фактических результатов с ожидаемыми.

Ожидаемые значения

$$E((X + X)/2) = 0{,}25 \cdot (-1) + 0{,}5 \cdot 0 + 0{,}25 \cdot 1 = 0$$

Фактические значения

Распределив полученные предыдущие результаты в пары, получим:

(1,—1) (—1,1) (1,1) (—1,—1) (—1,—1) (—1,1) (1,—1) (1,—1)

Это дает следующие средние:

0, 0, 1, —1, —1, 0, 0, 0

$$Var((X+X)/2 = 0{,}25 \cdot (-1)^2 + 0{,}25 \cdot 1^2 = 0{,}5 = \sigma^2/2$$

Исходя из них находим среднюю, равную —0,125, и дисперсию 0,411.
Заметьте, что уменьшение дисперсии достигнуто в результате расчета средних по выборкам размера 2. Мы ожидали 50%-ного уменьшения, но фактическое сокращение составило примерно 60%. Выборка большего размера подтвердит значение 50%.

Можно повторить процедуру для расчета средних по выборкам объемом в 3, 4 и т.д., т.е. нам нужна выборка данных большего объема для того, чтобы получить более надежные результаты. Мы можем ожидать, что средняя меняться не будет, а дисперсия будет сокращаться на величину, соответствующую размеру выборки.

ПРИЛОЖЕНИЕ 5.2. Данные по индексу FTSE 100 для проверки степени соответствия эмпирического и теоретического распределений

FTSE 100	Доход
2407,5	
2289,2	
2160,1	—5,8048
2311,1	6,7569
2422,7	4,7159
2345,8	—3,2256
2238,4	—4,6865
2221,6	—0,7534
2117,9	—4,7803
2371,4	11,3055
2372	0,0253
2339	—1,4010
2166,6	—7,6564
2030,8	—6,4729
2028	—0,1380
2162,7	6,4307
2143,5	—0,8917
2165,7	1,0304
2386,9	9,7252
2456,5	2,8742
2508,4	2,0908
2515,8	0,2946
2443,6	—2,9118
2591,7	5,8842

Продолжение таблицы

2679,6	3,3353
2645,6	−1,2770
2549,5	−3,7001
2414,9	−5,4239
2493,1	3,1869
2560,2	2,6558
2554,3	−0,2307
2408,6	−5,8733
2659,8	9,9205
2697,6	1,4112
2493,9	−7,8515
2420,2	−2,9998
2298,4	−5,1637
2572,3	11,2587
2687,8	4,3923
2792	3,8035
2846,5	1,9332
2851,6	0,1790
2882,6	1,0812
2878,4	−0,1458
2813,1	−2,2948
2849,2	1,2751
2888,8	1,3803
2941,7	1,8146
3085	4,7564
3093,3	−1,4924
3164,4	4,0336
3233,2	2,1509
Средняя	0,6905
Среднее квадратическое отклонение	4,7080

6

РЕГРЕССИОННЫЙ АНАЛИЗ

ВВЕДЕНИЕ

С помощью регрессионного анализа строится и проверяется модель связи между одной зависимой (т.е. эндогенной) и одной или более независимыми (экзогенными) переменными. Зависимая переменная обычно обозначается Y, а независимая (ые), также называемая регрессором, — X.

Направление причинной связи между переменными определяется через *предварительное* обоснование и включается в модель как гипотеза. Регрессионный анализ проверяет статистическую состоятельность модели при данной гипотезе.

Например, предположим, выдвинута гипотеза о том, что уровень фондового индекса FTSE (FTSE 100) линейно зависим от уровня фондового индекса S&P 500, т.е., когда растет S&P 500, растет и FTSE 100, а когда S&P 500 падает, падает и FTSE 100. Можно проверить эту гипотезу, используя простую линейную регрессию с включением только двух переменных.

Альтернативной гипотезой может быть то, что индекс FTSE 100 находится под влиянием не одного фактора, а нескольких. Например, на текущий уровень FTSE 100 могут влиять индекс S&P 500, уровень рынка облигаций Великобритании и обменный курс $/£. Эта гипотеза может быть проверена с помощью множественной регрессии.

Необходимо различать **кросс-секционную регрессию** и **регрессию временных рядов**. Кросс-секционная регрессия проверяет связь между переменными в определенный момент времени. В качестве примера кросс-секционной регрессии можно привести задачу измерения связи между размером компании и доходами по инвестициям в акции. Чтобы измерить эту связь, мы могли бы собрать данные по доходам на акции за один период, скажем за один год, и данные о размерах достаточно большого числа компаний на начало того же периода. Данные о доходах компаний соответствовали бы зависимой переменной, в то время как данные о размерах компаний были бы независимой переменной. Таким образом, регрессионный анализ показывает, как в среднем проявляется связь между переменными в определенный момент времени.

При анализе регрессии во временных рядах данные по каждой из переменных собираются в течение следующих друг за другом периодов времени. Приведенный выше пример фондовых индексов является регрессией временных рядов, потому что данные по FTSE 100 и S&P 500 должны были собираться на протяжении определенного периода времени. Регрессионный анализ позволил бы установить взаимосвязь в среднем в течение того периода времени, по которому имеются данные.

Независимо от того, проводится ли кросс-секционный анализ или анализ временных рядов, основные принципы приложения регрессионного анализа остаются те же.

ПРОСТАЯ ЛИНЕЙНАЯ РЕГРЕССИЯ

Применим регрессионный анализ для простой линейной зависимости между зависимой переменной (Y) и одной независимой переменной (X).

Под линейностью мы имеем в виду, что переменная Y предположительно находится под влиянием переменной X в следующей зависимости:

$$Y = \alpha + \beta X + e, \tag{6.1}$$

где α — постоянная, т.е. если бы даже X была равна нулю, Y имела бы какое-либо положительное или отрицательное значение. Можно ли дать разумное объяснение значению Y даже при X равном нулю, зависит от гипотезы, для которой применяется регрессионный анализ;

β — коэффициент регрессии, отражает наклон линии, вдоль которой рассеяны данные наблюдений. Он может быть истолкован как показатель, характеризующий процентное изменение переменной Y, которое вызвано изменением значения X на единицу. Таким образом, если Y и X — это соответственно индексы FTSE 100 и S&P 500, то β будет указывать, на какое количество пунктов изменится FTSE 100 при изменении индекса S&P 500 на один пункт. Если знак β положителен, то переменные положительно коррелированы. При отрицательном знаке β переменные отрицательно коррелированны;

e — ошибка или значение помехи, также называемая остатком. Она отражает тот факт, что обычно движение Y будет по крайней мере неточно описываться лишь движением X. Присутствуют другие факторы, не включенные в данную модель. Однако если исследуемая гипотеза реалистична, то эти другие переменные должны быть относительно неважными.

Обращаясь снова к взаимосвязи между FTSE 100 и S&P 500, отметим, что индекс FTSE 100 — зависимая переменная Y, так как мы выдвинули гипотезу о том, что движение этого индекса находится под влиянием, т.е. зависит от изменения индекса S&P 500, который представлен переменной X. В данной гипотезе мы предполагаем, что множество других незначительных и несвязанных влияний представлены в модели величиной e.

Если экономические аргументы достаточно сильны, мы можем развить гипотезу о том, что уровень индекса S&P 500 находится под влиянием индекса FTSE 100. При таком допущении величина индекса S&P 500 стала бы переменной *Y*, а индекса FTSE 100 — переменой *X*.

Расположив данные из табл. 6.1 на приведенной ниже точечной диаграмме рассеяния (рис. 6.1), мы действительно видим, что высокие (низкие) значения S&P 500 соответствуют высоким (низким) значениям FTSE 100. Таким образом создается впечатление, что данные по двум индексам растут и падают вместе.

Таблица 6.1

FTSE 100 (Y)	$Y-\bar{Y}$	S&P 500 (X)	$X-\bar{X}$	$(X-\bar{X})^2$	$(X-\bar{X})(Y-\bar{Y})$
2851,6	320,8596	442,52	51,10135	2611,348	16396,358
2882,6	351,8596	442,01	50,59135	2559,484	17801,052
2878,4	347,6596	450,3	58,88135	3467,013	20470,666
2813,4	282,3596	442,46	51,04135	2605,219	14412,015
2849,2	318,4596	453,83	62,41135	3895,176	19875,493
2888,8	358,0596	449,02	57,60135	3317,915	20624,716
2941,7	410,9596	450,15	58,73135	3449,371	24136,211
3085	554,2596	463,15	71,73135	5145,386	39757,788
3039,3	508,5596	461,28	69,86135	4880,608	35528,659
3164,4	633,6596	469,1	77,68135	6034,392	49223,532
3233,2	702,4596	461,89	70,47135	4966,211	49503,275
$\bar{Y}$ =2530,740		$\bar{X}$ =391,419			
				$\Sigma(X-\bar{X})^2=$ 108046,7	$\Sigma(X-\bar{X})(Y-\bar{Y})=$ 644387,4688

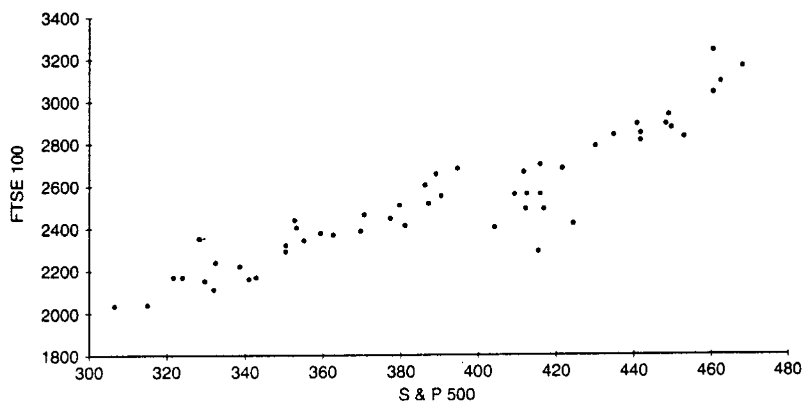

Рис. 6.1. Диаграмма рассеяния уровней FTSE 100 и S&P 500

Однако фактические данные не говорят нам ничего о причинной связи. Наше понимание причинной связи исходит из предварительно выдвинутой гипотезы. Как мы заметили в одном из предыдущих абзацев, указание на причину и следствие, т.е. на то, что является зависимой, а что независимой переменной, определяется выдвинутой гипотезой.

Для иллюстрации этого вернемся снова к гипотезе о том, что уровень FTSE 100 находится под влиянием уровня S&P 500. Фактические данные подтверждают эту идею, но поддержит ли ее ваше понимание экономики финансов? S&P 500 может влиять на FTSE 100 из-за огромного масштаба экономики США и международного оборота капитала. Однако альтернативное предположение заключается в том, что, так как оба рынка открыты для международных инвесторов, они оба могут находиться под влиянием третьего фактора, может быть, ожидания японских или европейских инвесторов.

Ясно, что независимо от регрессионной модели необходимо развивать гипотезу для того, чтобы регрессионный анализ смог обоснованно подтвердить или не подтвердить ее. Регрессионный анализ не в состоянии "доказать" гипотезу, он может лишь подтвердить ее статистически или отвергнуть.

Обращаясь к диаграмме рассеяния (рис. 6.1), отметим, что через точки на графике можно провести несколько прямых линий, удовлетворяющих выражению (6.1), хотя в действительности невозможно построить одну прямую линию, которая пройдет через все точки корреляционного поля. Отсюда очевидно, что нужно выбрать лишь одну линию.

ОПРЕДЕЛЕНИЕ ПАРАМЕТРОВ УРАВНЕНИЯ РЕГРЕССИИ С ПОМОЩЬЮ МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ

Для статистической проверки взаимосвязи между зависимой и независимой переменными необходимо найти значения α, β и e в выражении (6.1). Метод оценки должен быть таким, чтобы это были **наилучшие, линейные, несмещенные оценки** (BLUE — Best, Linear, Unbiased Estimator).

Понятие **наилучшие** относится к требованию для оценок параметров быть наиболее эффективными, т.е., чтобы дисперсии оценок параметров были как можно меньше. Это достигается таким выбором значений α и β, которые минимизируют сумму квадратов значений e_i^2.

Термин **линейные** просто повторяет, что взаимосвязь линейна.

Требование **несмещенные** означает, что ожидаемые значе-ния коэффициентов регрессии должны являться истинными коэффициентами.

Метод, используемый чаще других для нахождения параметров уравнения регрессии и известный как метод наименьших квадратов, дает наилучшие линейные несмещенные оценки. Он называется так потому, что при расчете параметров прямой линии, которая наиболее соответствует фактическим данным, с помощью этого метода стараются найти линию, минимизирующую сумму квадратов значений ошибок или расхождений между величинами Y, которые рассчитаны по уравнению прямой и обозначаются $\hat{Y}$, и фактическими наблюдениями. Это показано на рис. 6.2.

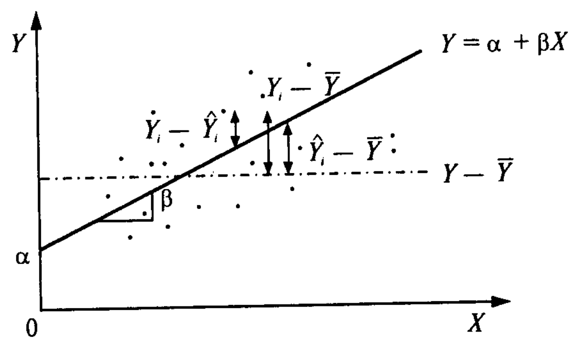

Рис. 6.2

Статистические допущения метода наименьших квадратов

Для обоснованного приложения метода наименьших квадратов (МНК) к данным и проверки взаимосвязи между переменными данные должны соответствовать допущениям, предполагаемым регрессионной моделью. Они представляют собой следующее:

1. Математическая форма взаимосвязи между **истинной зависимой переменной Y и независимой переменной X** выглядит как

$$Y = \alpha + \beta X + e. \qquad (6.2)$$

В данный момент было бы полезным рассмотреть различия между детерминированными и вероятностными моделями. Детерминированные модели — это те, в которых при известном значении X значения Y также точно известны. Например, доход коммивояжера может быть определен через фиксированный оклад "α" и премию "β%" от объема продаж. Как только мы узнаем денежный объем продаж, мы сможем определить Y в точности как $Y = \alpha + \beta X$. Нет необходимости в приложении регрессионного анализа к подобным моделям. Нам просто нужно произвести наблюдение за двумя X и относящимися к ним Y и провести линию через эти точки.

Чаще мы должны относиться к Y как к случайной переменной. При известной величине X мы имеем распределение вероятностей для Y, таким образом, модель вероятностна и выражается следующим образом

$$Y = \alpha + \beta X + e.$$

По фактическим данным мы оцениваем α и β, оценочные значения обозначаются $\hat{\alpha}$ и $\hat{\beta}$ соответственно. Можно затем рассчитать ожидаемое значение Y, т.е. $\hat{Y}$, из выражения

$$\hat{Y} = \hat{\alpha} + \hat{\beta} X. \qquad (6.3)$$

Следовательно, для каждого значения данных X_i существует фактическое значение Y_i, но при использовании выражения (6.3) появляется также оценочное значение $\hat{Y}_i$. Разность между Y_i и $\hat{Y}_i$ — это ошибка оценки e_i. Найденная с помощью МНК линия регрессии представляет собой прямую, которая минимизирует сумму квадратов e_i, т.е. минимизирует $\Sigma e_i^2 = \Sigma(Y_i - \hat{Y}_i)^2$.

2. Значение ошибки e_i **нормально распределено со средней, равной нулю, и постоянной дисперсией** σ^2, что часто записывается как $e_i \sim N(0, \sigma^2)$.
3. Последующие значения ошибок независимы друг от друга, т.е. ковариация в парах значений ошибок равна нулю ($\text{cov}\ e_i e_j = 0$).
4. Независимая переменная является нестохастической.

Первое допущение просто повторяет, что в данной модели мы имеем дело с линейными зависимостями, и что величины зависимой переменной Y определяются только одним значимым фактором — независимой переменной X.

Второе допущение указывает, что хотя существует только один главный фактор (X), определяющий величину Y, присутствует также множество второстепенных факторов, некоторые из них будут оказывать положительное влияние на величину Y, а другие отрицательное. В случае множества отрицательных и положительных влияний значение ошибки будет нормально распределено. Допущение о постоянной дисперсии значения ошибки означает, что как бы ни была велика или мала величина независимой переменной X, разброс значений e постоянен. Говорят, что значение ошибки обладает свойством **гомоскедастичности**. Если разброс значений ошибки непостоянен, то ошибки определяются как **гетероскедастичные**.

Третье допущение о том, что значения e независимы друг от друга, просто означает, что второстепенные факторы или факторы, которые послужили причиной ошибки для одной из величин Y, не приводят автоматически к ошибкам для всех наблюдений Y. Когда значения e независимы, данные являются **неавтокоррелированными**. Если значения e не являются независимыми, говорят, что данные **автокоррелированы** или демонстрируют наличие **автокорреляции**. Иногда автокорреляцию называют "последовательной корреляцией".

Так как Y линейно связана с e, то сама Y — это случайная переменная. Для любых значений X значения Y будут нормально распределены, и, таким образом, статистическое распределение Y_i может быть полностью описано с помощью его средней и дисперсии:

$$E(Y_i) = E(\alpha + \beta_1 X_i + e_i)\,. \qquad (6.4)$$

Так как α и β_1 постоянны, а X_i нестохастична, это выражение преобразуется в

$$E(Y_i) = \alpha + \beta_1 X_i + E(e_i)\,. \qquad (6.5)$$

Однако поскольку математическое ожидание e_i равно нулю, выражение (6.5) превращается в

$$E(Y_i) = \alpha + \beta_1 X_i\,. \qquad (6.6)$$

Так как математическое ожидание значений e_i равно нулю, то дисперсия Y_i, которая также служит дисперсией e_i, является средним значением e_i^2, т.е.

$$\Sigma(e_i - 0)^2/n = \Sigma e_i^2/n = E(e_i^2) = \sigma^2.$$

Отсюда Y_i нормально распределено с параметрами $N(\alpha + \beta_1 X_i, \sigma^2)$. Это показано на рис. 6.3.

Для каждого значения X существует ожидаемое значение Y_i, которое нормально распределено и которому, таким образом, мы можем приписать вероятности, т.е. это вероятностная модель.

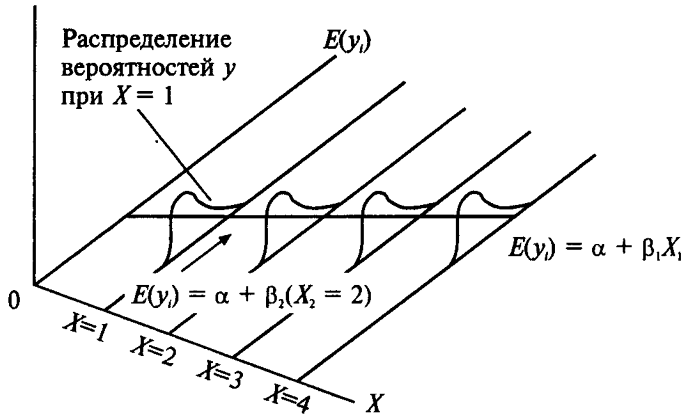

Рис. 6.3

Определение линии регрессии

Величины α и β, минимизирующие суммы квадратов отклонений Y от $\hat{Y}$, находятся следующим образом:

$$\beta = \frac{\mathrm{cov}_{XY}}{\mathrm{var}_X} = \frac{\sum\left[\left(X - \bar{X}\right)\left(Y - \bar{Y}\right)\right]}{\sum\left(X - \bar{X}\right)^2}; \qquad (6.7)$$

$$\alpha = \left(\bar{Y} - \beta\bar{X}\right). \qquad (6.8)$$

Заметьте, что и ковариация и дисперсия обычно имеют делители $n-1$, однако они сокращаются в указанном выше выражении.

Значения ошибок, называемые также остатками, рассчитываются как

$$e_i = \left(Y_i - \hat{Y}_i\right), \tag{6.9}$$

где Y_i — наблюдаемая величина Y, а $\hat{Y}_i$ — это оценка величины Y.

Сумма квадратов (S) составляет

$$SS = \sum\left(Y - \hat{Y}\right)^2 = \sum(Y - \alpha - \beta X)^2, \tag{6.10}$$

где α и β — определяемые параметры.

$$\frac{\partial SS}{\partial \alpha} = \sum\left(-2(Y - \alpha - \beta X)\right) \tag{6.11}$$

и

$$\frac{\partial SS}{\partial \beta} = \sum\left(-2X(Y - \alpha - \beta X)\right)^{*}. \tag{6.12}$$

Сумма квадратов минимальна, когда обе эти частные производные равны нулю, т.е. при

$$\sum(Y - \alpha - \beta X) = 0;$$

$$\sum X(Y - \alpha - \beta X) = 0.$$

В результате получаем:

$$\begin{aligned} \sum Y &= n\alpha + \beta\sum X; \\ \sum XY &= \alpha\sum X + \beta\sum X^2. \end{aligned} \tag{6.13}$$

Сейчас нам необходимо решить систему линейных уравнений. Начнем с умножения первого выражения на ΣX и второго на n. Произведя эти действия, получим:

$$\begin{aligned} \sum X\sum Y &= n\alpha\sum X + \beta\left(\sum X\right)^2; \\ n\sum XY &= n\alpha\sum X + n\beta\sum X^2. \end{aligned} \tag{6.14}$$

---

* $\dfrac{\partial SS}{\partial \alpha}$ и $\dfrac{\partial SS}{\partial \beta}$ — первые производные по α и β.

При вычитании первого выражения из второго получаем:

$$n\sum XY - \sum X\sum Y = n\beta\sum X^2 - \beta\left(\sum X\right)^2 = \beta\left(n\sum X^2 - \left(\sum X\right)^2\right). \tag{6.15}$$

Отсюда

$$\beta = \frac{n\sum XY - \sum X\sum Y}{n\sum X^2 - \left(\sum X\right)^2}. \tag{6.16}$$

Замещение ΣX и ΣY на произведения их средних на n приводит к выражению

$$\beta = \frac{n\sum XY - n\bar{X}n\bar{Y}}{n\sum X^2 - \left(n\bar{X}\right)^2}. \tag{6.17}$$

Поделив числитель и знаменатель этого выражения на n, получим:

$$\beta = \frac{\sum XY - n\overline{XY}}{\sum X^2 - n\bar{X}^2}. \tag{6.18}$$

Это то же самое, что и

$$\beta = \frac{\sum\left[\left(X - \bar{X}\right)\left(Y - \bar{Y}\right)\right]}{\sum\left(X - \bar{X}\right)^2}. \tag{6.19}$$

Для нахождения α разделим первое из уравнений (6.13) на n, отсюда

$$\frac{\sum Y}{n} = \frac{n\alpha + \beta\sum X}{n};$$

$$\bar{Y} = \alpha + \beta\bar{X}; \tag{6.20}$$

$$\alpha = \bar{Y} - \beta\bar{X}.$$

Используя данные из табл. 6.1, относящиеся к индексам FTSE 100 и S&P 500, можно рассчитать β:

$$\beta = \frac{\mathrm{cov}_{XY}}{\mathrm{var}_X} = \frac{\sum\left[\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)\right]}{\sum\left(X_i - \bar{X}\right)^2} = \frac{644387{,}5}{108046{,}7} = 5{,}9640$$

и

$$\alpha = \bar{Y} - \beta\bar{X} = 2530{,}74 - (5{,}9640 \cdot 391{,}4187) = 196{,}3298.$$

Интерпретация уравнения регрессии

По полученным результатам уравнение регрессии выглядит следующим образом:

$$\hat{Y} = +196{,}3298 + 5{,}9640X\ .$$

Это можно истолковать как утверждение о том, что ожидаемое значение зависимой переменной Y (FTSE 100) равно постоянной величине + 196,3298 плюс 5,9640 за каждую единицу независимой переменной X (S&P 500). Постоянная величина представляет собой значение зависимой переменной, когда независимая переменная равна нулю. Графически это выражается как расстояние по вертикали между началом координат и точкой, где линия регрессии пересекает ось ординат на рис. 6.2.

Проверка модели

Мы заметили выше, что вероятностные модели предоставляют лишь оценки коэффициентов регрессии. Важно, таким образом, проверить, насколько представительны данные оценки относительно истинных коэффициентов. Это достигается проверкой статистической значимости коэффициентов регрессии и близости расположения фактических данных к рассчитанной линии регрессии.

Критерии значимости коэффициентов

Как показывает рис. 6.3, в случае вероятностных моделей расчет коэффициентов регрессии с использованием выражений (6.7) и (6.8) дает одну оценку величины Y, т.е. $E(Y_i)$. Оценки коэффициентов регрессии также предположительно нормально распределены. Нам нужно знать статистическую значимость этих коэффициентов. Данная задача решается проверкой того, что коэффициенты регрессии значимо отличаются от нуля.

Обращаясь снова к рис. 6.3, отметим, что поскольку каждая оценка коэффициентов предположительно нормально распределена, то для проверки значимости мы проверяем, попадает ли величина оценки в критическую область распределения, следовательно, являясь случайной, либо же она попадает в основную область распределения.

Таким образом, статистическая значимость коэффициентов измеряется степенью вариации вокруг оценочного значения. Так как ошибки (или остатки) по предположению нормально распределены, то среднее квадратическое отклонение ошибок используется для измерения этой вариации. Эти средние квадратические отклонения известны как **стандартные ошибки коэффициентов**. Для определения степени значимости коэффициентов мы используем ***t*-критерии**. Чтобы их определить, необходимо знать:

- выборочное распределение данных коэффициентов;
- оценки их дисперсий и, таким образом, средних квадратических отклонений.

Затем можно либо проверить гипотезу, относящуюся к коэффициентам, либо определить для них доверительные интервалы.

Выборочное распределение

Выборочное распределение постоянной α записывается так:

$$\hat{\alpha} \sim N\left(\alpha, \frac{\sigma^2 \sum X^2}{n\sum\left(X-\bar{X}\right)^2}\right), \tag{6.21}$$

где σ^2 — дисперсия e_i.

Выборочное распределение β имеет вид

$$\hat{\beta} \sim N\left(\beta, \frac{\sigma^2}{\sum\left(X-\bar{X}\right)^2}\right). \tag{6.22}$$

Оценки дисперсий и средних квадратических отклонений

Выше мы заметили, что средняя или математическое ожидание величины e_i равно нулю, следовательно, $\sum\left(e_i - 0\right)^2$ — это просто $\sum e_i^2$. Для получения s^2 как несмещенной оценки σ^2 необходимо делить эту сумму на (n — 2), так как мы оцениваем два параметра (α и β) по фактическим данным. Отсюда оценка дисперсии e_i, s^2 определяется как

$$s^2 = \frac{\sum \hat{e}_i^2}{n-2}. \tag{6.23}$$

Вспомним, что

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta}_i X_i + \hat{e}_i. \tag{6.24}$$

Тогда

$$\hat{e}_i = \hat{Y}_i - \hat{\alpha} - \hat{\beta}_i X_i \tag{6.25}$$

и, следовательно,

$$Y_i - \hat{Y}_i = \hat{e}_i. \tag{6.26}$$

Используем это при ссылке на дисперсию коэффициентов регрессии.

Дисперсия коэффициента $\hat{\alpha}$ оценивается как

$$\operatorname{var}\hat{\alpha} = \frac{\dfrac{\sum\left(Y_i - \hat{Y}_i\right)^2}{n-2}\sum X_i^2}{n\sum\left(X_i - \overline{X}\right)^2}. \tag{6.27}$$

Дисперсия коэффициента наклона линии регрессии $\hat{\beta}$ оценивается как

$$\operatorname{var}\hat{\beta} = \frac{\left[\dfrac{\sum\left(Y_i - \hat{Y}_i\right)^2}{n-2}\right]}{\sum\left(X_i - \overline{X}\right)^2}. \tag{6.28}$$

Стандартные ошибки

Стандартные ошибки коэффициентов — это просто их средние квадратические отклонения, т.е. квадратные корни дисперсий коэффициентов.

Стандартная ошибка постоянного коэффициента α рассчитывается как

$$\mathrm{SE}_{\hat{\alpha}} = \sqrt{\frac{\dfrac{\sum\left(Y_i - \hat{Y}_i\right)^2}{n-2}\sum X_i^2}{n\sum\left(X_i - \overline{X}\right)^2}}. \tag{6.29}$$

Стандартная ошибка коэффициента наклона β определяется как

$$\mathrm{SE}_{\hat{\beta}} = \sqrt{\frac{\left[\dfrac{\sum (Y_i - \hat{Y}_i)^2}{n-2}\right]}{\sum \left(X_i - \overline{X}\right)^2}} . \qquad (6.30)$$

Мы выяснили в гл. 5, что для данных с нормальным, или Гауссовым, распределением разность между переменной и ее средней, деленная на оценку ее среднего квадратического отклонения, имеет t-распределение. Таким образом, оценки, деленные на стандартные ошибки, как рассчитано ранее, имеют t-распределение с n—2 степенями свободы. Опираясь на эту информацию, можно определить доверительные интервалы вокруг точечных оценок коэффициентов.

Обозначим уровень доверия 1—c, где c — вероятность попадания переменной в критическую область t-распределения. Уровни вероятности составляют

$$P\left(-t_{n-2,c/2} \le \frac{\hat{\alpha} - \alpha}{\mathrm{SE}_{\alpha}} \le t_{n-2,c/2}\right) = 1 - c ,$$

$$P\left(-t_{n-2,c/2} \le \frac{\hat{\beta}_i - \beta_i}{\mathrm{SE}_{\beta_i}} \le t_{n-2,c/2}\right) = 1 - c . \qquad (6.31)$$

Из этих выражений можно определить оценки коэффициентов, попадающие в доверительный интервал

$$P\left(\hat{\alpha} - t_{n-2,c/2}\mathrm{SE}_{\alpha} \le \alpha \le \hat{\alpha} + t_{n-2,c/2}\mathrm{SE}_{\alpha}\right) = 1 - c ,$$

$$P\left(\hat{\beta}_i - t_{n-2,c/2}\mathrm{SE}_{\beta_i} \le \beta_i \le \hat{\beta}_i + t_{n-2,c/2}\mathrm{SE}_{\beta_i}\right) = 1 - c . \qquad (6.32)$$

Следовательно, существует вероятность 1—c того, что действительные значения коэффициентов попадают в установленный промежуток. Если этот промежуток включает нуль, то коэффициенты статистически незначимо отличны от нуля.

На практике все указанные выше и другие вычисления, представленные в этой главе, производятся автоматически с использованием стандартных пакетов прикладных программ.

Проверка гипотез

Выше мы заметили, что уравнение регрессии часто создается, чтобы проверить гипотезы. Для этого выдвигается нулевая гипотеза H_0 о том, что коэффициенты статистически незначимо отличны от нуля. В случае коэффициентов α и β гипотезы будут следующими:

$$H_0:\ \alpha = 0,$$
$$H_1:\ \alpha \neq 0,$$
$$H_0:\ \beta = 0,$$
$$H_1:\ \beta \neq 0.$$

Для проверки этих гипотез нужно определить t-критерий для соответствующего коэффициента регрессии. Эти **t-критерии** рассчитывают делением коэффициентов регрессии на их стандартные ошибки

$$t = \frac{\beta}{SE_\beta}. \qquad (6.33)$$

Эти t-критерии имеют t-распределение при условии действия допущений нормального регрессионного анализа.

Обычно статистическую значимость проверяют при уровне доверительной вероятности 95 или 99%. Это означает, что существует вероятность 95 или 99%, что значения α и β не случайны. Распределением вероятностей t-критерия является t-распределение с n—2 степенями свободы. Степени свободы здесь относятся к числу пар единиц данных, использованных в уравнении регрессии. Коэффициенты регрессии значимы, если t-критерий больше, чем величина, указанная в таблицах t-распределения.

Чтобы проиллюстрировать проверку значимости с использованием таблиц, напомним что величина β составляет 5,9640, а стандартная ошибка для β равна 0,3476. Следовательно, t-критерий равен:

$$t = \frac{5{,}9640}{0{,}3476} = 17{,}1577.$$

Критерий значимости для α равен

$$\frac{\alpha}{SE_\alpha} = \frac{196{,}3298}{136{,}991} = 1{,}4332,$$

и 95%-ные доверительные интервалы выглядят так:

для α: $196{,}3298 \pm 2 \cdot 136{,}991$, т.е. $-77{,}65 \le \alpha \le 470{,}31$;

для β: $5{,}9640 \pm 2 \cdot 0{,}3476$, т.е. $5{,}27 \le \beta \le 6{,}66$.

Односторонняя или двусторонняя проверка?

Нам необходимо решить, будет ли проверка значимости "односторонней" или "двусторонней". Это решение должно быть принято еще до того, как станут известны результаты регрессии. Выбор определяется теоретическим обоснованием модели связи X и Y, проверяемой с помощью регрессии.

Вспомним из гл. 5, что при односторонней проверке проверяется значимость коэффициента регрессии, когда ожидаемая взаимосвязь между X и Y может быть или положительной или отрицательной, но не той и другой одновременно. Двусторонняя проверка проверяет значимость, когда взаимосвязь между X и Y может быть и положительной, и отрицательной.

Чтобы проиллюстрировать процесс принятия решения, рассмотрим взаимосвязь между двумя индексами фондового рынка. Если эта модель базируется на глубокой уверенности в том, что экономика США доминирует в общемировой и определяет ожидания инвесторов, то связь между двумя переменными может быть только положительной, и, следовательно, односторонняя проверка оправдана.

Рассмотрим альтернативное мнение о том, что связь двух рынков зависит от относительного поведения двух экономик. Если не будет высокой корреляции между уровнями деятельности обеих экономик, они могут разойтись, а вместе с ними могут разойтись и характеристики рынков акций. Следовательно, они могут быть связаны положительно или отрицательно. Отсюда для проверки связи между двумя рынками необходима двусторонняя проверка.

Для проведения двусторонней проверки того, что коэффициент регрессии значительно отличается от нуля при уровне доверительной вероятности 95 или 99%, мы должны вспомнить, что число пар наблюдений составляет 52, значит, существует 50 степеней свободы. Обращаясь к таблицам t-распределения, мы сначала двигаемся вниз по столбцу, озаглавленному как число степеней свободы, до 50. Затем мы двигаемся по строке к столбцу, представляющему уровень доверительной вероятности, с которым проводится проверка. Столбец, озаглавленный 0,05, соответствует уровню доверительной вероятности 95%, в то время как столбец, озаглавленный 0,01, представляет уровень доверительной вероятности 99%.

Проиллюстрируем использование одностороннего критерия для уровня доверительной вероятности 99%. Для 50 степеней свободы критическое значение t-критерия составляет 2,4. Если t-критерий для некоторого коэффициента больше, чем 2,4, то этот коэффициент регрессии признается значимым при уровне доверительной вероятности 99%.

В приведенном выше примере t-критерий для α равен 1,4332, а для коэффициента β — 17,15537, что намного больше критического уровня 2,4. Отсюда α не является значимо отличным от нуля, но коэффициент β значим при уровне доверия 99%.

Чтобы провести двустороннюю проверку при уровне доверительной вероятности 95%, мы бы смотрели на столбец, озаглавленный 0,025, что составляет половину от 0,05, и нашли бы, что критическое значение равно 2,0. Для проверки при доверительной вероятности 99% мы бы смотрели на столбец, озаглавленный 0,005, и критическое значение t в этом случае составило бы 2,7.

Степень соответствия: коэффициент детерминации R^2

Регрессионная модель показывает, что вариация Y может быть объяснена вариацией независимой переменной X и значением ошибки e. Мы хотим знать, насколько вариация Y обусловлена изменением X и насколько она является следствием случайных причин. Другими словами, нам нужно знать, насколько хорошо рассчитанное уравнение регрессии соответствует фактическим данным, т.е. насколько мала вариация данных вокруг линии регрессии. Это продемонстрировано ниже на графике. Рис. 6.4(a) показывает, что данные широко рассеяны относительно линии регрессии, следовательно, ошибки велики. На рис. 6.4(b) данные близко концентрируются вокруг линии регрессии. Отсюда делаем вывод, что ошибки малы, и взаимосвязь, характеризуемая линией регрессии, считается в большой степени отражающей истинную взаимосвязь.

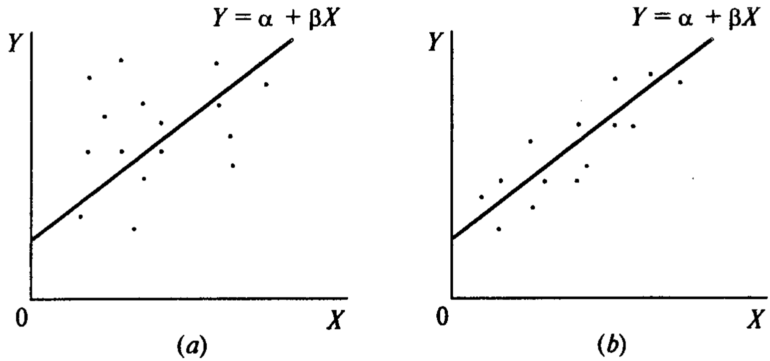

Рис. 6.4

Для оценки степени соответствия линии регрессии нам нужно рассчитать **общую сумму квадратов отклонений, сумму квадратов отклонений, объясняемую регрессией**, и **остаточную сумму квадратов отклонений**, чтобы определить **коэффициент детерминации R^2**.

Общая сумма квадратов отклонений (ОСК) — это сумма возведенных в квадрат разностей между наблюдаемой величиной зависимой переменной Y_i и средней наблюдаемых величин зависимой переменной $\bar{Y}$

$$\text{ОСК} = \sum\left(Y_i - \bar{Y}\right)^2. \tag{6.34}$$

Сумма квадратов, объясняемая регрессией (СКР) — это сумма возведенных в квадрат разностей между прогнозируемыми величинами зависимой переменной $\hat{Y}$ и средней величиной наблюдаемых значений зависимой переменной $\bar{Y}$

$$\text{СКР} = \sum\left(\hat{Y}_i - \bar{Y}\right)^2. \tag{6.35}$$

Остаточная сумма квадратов (СКО) — это сумма возведенных в квадрат разностей между $\hat{Y}_i$ и Y_i

$$\text{СКО} = \sum\left(Y_i - \hat{Y}_i\right)^2. \tag{6.36}$$

Общая сумма квадратов представляет собой сумму двух слагаемых: суммы квадратов отклонений, объясняемых регрессией, и остаточной суммы квадратов. Эти характеристики показаны выше графически на рис. 6.2.

Отношение **суммы квадратов отклонений, объясняемой регрессией, к общей сумме квадратов отклонений** дает пропорцию изменения Y, объясняемого изменением X, и называется **коэффициентом детерминации** (R^2)

$$R^2 = \frac{\text{СКР}}{\text{ОСК}} = \frac{\sum\left(\hat{Y}_i - \overline{Y}\right)^2}{\sum\left(Y_i - \overline{Y}\right)^2}. \qquad (6.37)$$

В случае простой регрессии двух переменных R^2 представляет собой квадрат коэффициента корреляции R.

Коэффициент детерминации будет принимать значения от нуля, когда X не влияет на Y, до единицы, когда изменение Y объясняется изменением X. Значение R^2 для регрессии данных по индексам FTSE 100 и S&P 500 из табл. 6.1 составляет 0,8548. Обычная интерпретация коэффициента детерминации такова: число (значение), скажем $R^2 = 0{,}8548$, умножается на 100 и выражается как процентная доля вариации Y, которая объясняется вариацией X. Таким образом, в этом примере 85,48% изменения Y (индекс FTSE 100) объясняются изменением X (индекс S&P 500). Доверяете вы или нет тому, что рынок акций США имеет сильное определяющее влияние на рынок Великобритании, будет зависеть от того, насколько основательно вы исследовали конкурирующие теории. Вспомните, что данная регрессионная модель — это только математическое выражение той одной гипотезы, которая проверялась.

Коэффициент детерминации R^2 сам по себе является случайной переменной, потому что X и Y — это случайные переменные. Критерий проверки значимости R^2 имеет F-распределение. Это распределение отлично от других распределений тем, что обладает двумя наборами чисел степеней свободы: один (часто обозначаемый ν_1) — в числителе критерия проверки, а другой (обозначаемый ν_2) — в знаменателе. В критерии проверки для R^2 числителю соответствует одна степень свободы и (n—2) степеней свободы соответствуют знаменателю. Расчет критерия проверки для упомянутого выше R^2 выполняется следующим образом:

$$\frac{R^2}{\left(\frac{1-R^2}{n-2}\right)} = \frac{0{,}8548}{\left(\frac{1-0{,}8548}{50}\right)} = 294{,}4.$$

Обратившись к F-таблицам, мы видим, что табличное значение при 5%-ном уровне значимости для $\nu_1 = 1$ и $\nu_2 = 50$ примерно равно 4. Так как значение критерия проверки больше 4, мы отвергаем нулевую гипотезу о том, что $R^2 = 0$.

Однако проверка не является особенно полезной. Она говорит только, что *существует* корреляция между Y и X. При толковании коэффициента детерминации необходима осторожность. Например, как индикатор степени соответствия R^2 часто используется для сравнения уравнений регрессии. Хотя применить R^2 в подобном случае можно, только если зависимые переменные в каждом сравниваемом уравнении идентичны. Также нецелесообразно использовать R^2 для сравнения регрессионных моделей, которые содержат разное число объясняющих переменных. Таким образом, применение R^2 при сравнении степени пригодности простой регрессионной модели и многофакторной модели (имеющей несколько независимых переменных) не оправдано.

ИСПОЛЬЗОВАНИЕ РЕГРЕССИИ ДЛЯ ПРОГНОЗИРОВАНИЯ

Интервал прогнозирования

Регрессионные модели могут быть использованы для прогнозирования. Например, предположим, что мы хотим предсказать уровень индекса FTSE 100 при росте индекса S&P 500 до 550 за данный день. Прогнозное значение индекса составит

$$\hat{Y} = +196{,}3298 + (5{,}963972 \cdot 550) = 3476{,}51 \approx 3477 .$$

Когда мы используем регрессионную модель для прогноза величины Y (в данном случае это уровень индекса FTSE 100), располагая величиной X (уровень индекса S&P 500), мы хотим знать степень доверия к оцениваемому значению. Для этой цели рассчитываются **стандартная ошибка оценки**, а затем **интервал прогнозирования**.

Стандартная ошибка оценки, также известная как стандартная ошибка уравнения регрессии, определяется следующим образом (см. (6.23)):

$$s = \sqrt{\frac{\sum \left(Y_i - \hat{Y}_i\right)^2}{(n-2)}} . \tag{6.38}$$

Фактически это среднее квадратическое отклонение всех e_i.

Интервал прогнозирования рассчитывается так:

$$\hat{Y} \pm t_{99} \cdot s\sqrt{1 + \frac{1}{n} + \frac{(X^* - \overline{X})^2}{\sum\left(X_i - \overline{X}\right)^2}}, \tag{6.39}$$

где нижний индекс 99, относящийся к t, отражает уровень доверительной вероятности, а X^* — это значение X, используемое для прогноза, т.е. 550 для приведенного выше примера.

Стандартная ошибка оценки в отношении регрессии FTSE/S&P равна $s = 114{,}27$. Интервал прогнозирования составляет

$$3476 \pm 2{,}500 \cdot 114{,}27 \cdot \sqrt{1 + 0{,}0192 + \frac{(550 - 391{,}42)^2}{108046{,}7}}$$

$$= 3476 \pm 2{,}500 \cdot 114{,}27 \cdot 1{,}1189 = 3476 \pm 319{,}65.$$

Таким образом, мы можем полагать с вероятностью 99%, что если индекс S&P 500 вырастет до 550, то индекс FTSE 100 увеличится до 3476 ± 320, т.е. будет находиться между 3156 и 3796.

Ложная регрессия

Приведенный выше пример регрессионного анализа основывался на данных, относящихся к фондовым индексам. Однако использование данных, относящихся к уровням цен, может привести к так называемой **ложной регрессии** (spurious regression). Ложная регрессия может возникнуть при анализе данных, когда величина наблюдений каждой из переменных имеет тенденцию к росту (или понижению) с течением времени. Эта тенденция создает уровень корреляции, который переоценивает любую рассматриваемую причинную связь. Данные, относящиеся к переменным, выраженным в виде денежных единиц или финансовых агрегатов, стремятся к росту с течением времени и, таким образом, особенно чувствительны к этой проблеме.

Исследование взаимосвязи между изменениями уровня одной переменной и изменениями уровня другой может привести к отличному от истины результату. По этой причине, когда регрессионный анализ применяется для временных рядов финансовых переменных, часто рассчитываются изменения уровней используемых данных либо в качестве альтернативы изменения могут быть преобразованы в форму процентных доходов (например, $(P_1-P_0)/P_0$ или $\ln P_1/P_0$). По сути мы ищем возможность для преобразования имеющихся данных в стационарные, так как существуют методы для анализа подобных данных (см. гл. 7, где эта тема развивается). Мы проанализировали данные уровней в этом разделе по причине интуитивной, может быть, даже ложной связи между двумя рынками акций. Исправим эту ситуацию в следующем разделе.

МНОЖЕСТВЕННАЯ РЕГРЕССИЯ

Редко поведение зависимой переменной объясняется только с помощью одной независимой переменной. Обычно несколько независимых переменных, используемых в комбинации, предлагают лучшее объяснение. Регрессионная модель, включающая несколько независимых переменных, известна как **множественная регрессия**.

Истинная взаимосвязь между зависимой переменной Y и различными независимыми переменными X_i выражается так:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + e. \tag{6.40}$$

Однако, как и в случае простой линейной регрессии, мы не знаем истинную взаимосвязь и вынуждены делать оценки:

$$\hat{Y} = \hat{\alpha} + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \ldots + \hat{\beta}_n X_n. \tag{6.41}$$

β_i представляют частные производные Y по соответствующим X_i, например,

$$\beta_1 = \frac{\partial \hat{Y}}{\partial X_1}, \quad \beta_2 = \frac{\partial \hat{Y}}{\partial X_2}, \quad \beta_n = \frac{\partial \hat{Y}}{\partial X_n}, \tag{6.42}$$

в предположении, что все остальные X_i постоянны.

Вспомним, что в случае простой регрессии постоянная представляла собой величину зависимой переменной, когда независимая переменная имела нулевое значение. Однако при множественной регрессии толкование постоянной является более сложным. В некоторых моделях постоянный член оценивается *a priori*, в других случаях значимая постоянная может представлять средний эффект, оказываемый на Y любыми независимыми переменными, которые не были включены в модель. Отсюда значимая постоянная может представлять некоторую важную объясняющую переменную, исключенную из проверяемой модели.

Чтобы проиллюстрировать множественную регрессию, рассмотрим гипотезу о том, что доходы по индексу FTSE 100 определяются доходами на рынке государственных облигаций Великобритании (переменная X_1), доходами по индексу S&P 500 (переменная X_2) и доходами по обменному курсу US$/£ ($X_3$).

Существует много пакетов программ, которые способны решить подобную проблему множественной регрессии. Типичный результат будет похож на следующее:

$$Y = -0{,}215 + 0{,}209X_1 + 0{,}934X_2 - 0{,}302X_3$$

$$(-0{,}39) \quad (1{,}02) \quad (6{,}42) \quad (-2{,}54)$$

$$R^2 = 0{,}52 \ CC = 47.$$

Скорректированный $R^2 = 0{,}49$;

$$DW = 2{,}3, \ F = 26{,}0,$$

t-критерий приведен в скобках,
CC означает степени свободы,
DW относится к критерию Дарбина—Уотсона (Durbin—Watson), который связан с автокорреляцией и обсуждается далее в этой главе.

Замечание: данные для этой задачи взяты из приложения 6.2.

Результаты могут быть интерпретированы следующим образом: если X_2 и X_3 постоянны, при изменении X_1 на единицу изменение Y составит 0,209. Аналогично, если X_1 и X_3 постоянны, изменение X_2 на единицу приведет к изменению Y на 0,934 единицы.

Числа в скобках могут представлять стандартные ошибки коэффициентов либо t-критерии.

Допущения относительно использования метода наименьших квадратов для многофакторных моделей те же самые, что и для однофакторной модели. Однако многофакторная модель имеет дополнительное допущение о том, что независимые переменные независимы друг от друга, т.е. cov $(x_j, x_k) = 0$ $(j \neq k)$.

Обратимся к толкованию результата, полученного с помощью компьютера. t-критерии для каждой независимой переменной истолковываются так же, как и раньше. Однако в случае многофакторной регрессии они имеют t-распределение с $n—k—1$ степенями свободы.

Если существует k независимых переменных, то будет $k + 1$ коэффициентов регрессии (включая постоянную), отсюда число степеней свободы составит $n—(k + 1)$ или $n—k—1$.

Поскольку в изложенном выше примере 47 степеней свободы, то при уровне доверительной вероятности 95% значение t-критерия должно быть больше 2,02, а при уровне доверительной вероятности 99% потребовалось бы, чтобы t-критерий был больше для обеспечения значимости коэффициентов регрессий 2,70. Таким образом, в приведенном выше примере постоянный коэффициент и переменная X_1 не будут значимо отличаться от нуля, но две другие переменные значимы при уровне вероятности 95%, а X_2 значима и при уровне доверительной вероятности 99%.

Скорректированный R^2 : $\overline{R}^2$

В многофакторной регрессии добавление дополнительных объясняющих переменных увеличивает коэффициент детерминации. Следовательно, коэффициент детерминации должен быть скорректирован с учетом числа независимых переменных. Скорректированный R^2, или $\overline{R}^2$, рассчитывается так:

$$\overline{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k}, \tag{6.43}$$

где n — число наблюдений;

k — число независимых переменных.

Для иллюстрации вспомним ранее вычисленный R^2 с тремя переменными, равный 0,52, для него

$$\overline{R}^2 = 1 - \left[(1 - 0{,}52)\cdot\frac{50}{47}\right] = 0{,}49223.$$

Скорректированный R^2 уменьшится по величине, если дополнительная переменная незначима. Однако необходимо предостеречь против включения и исключения переменных только лишь из-за их влияния на скорректированный R^2. Рациональной базой для включения и исключения служит теория, стоящая за проверяемой моделью. Отсюда переменная, которая имеет сильное теоретическое основание для включения, должна быть добавлена в модель, даже если скорректированный R^2 от этого не улучшится.

Проверка значимости скорректированного R^2 — это также проверка значимости связи между зависимой переменной Y и любой из независимых переменных X_i. Действительно, если регрессионная модель имеет высокую степень предоставления объяснения формирования взаимосвязи, изменение зависимой переменной происходит из-за изменений независимых переменных, и суммы квадратов отклонений, объясняемые регрессией (СКР) будут относительно больше остаточной суммы квадратов отклонений (СКО). Если же модель имеет низкую степень предоставления объяснения, изменение зависимой переменной происходит из-за изменения значения ошибки, и СКО будет относительно больше СКР.

Критерий проверки вычисляется как

$$\frac{R^2}{\left(1 - R^2\right)} \frac{(n - k)}{k - 1} \sim F_{k-1,\, n-k}. \tag{6.44}$$

Таким образом, этот критерий проверки имеет F-распределение с k-1 степенями свободы в числителе и $n—k$ степенями свободы в знаменателе.

В числовом выражении:

$$F = \frac{0{,}52}{1 - 0{,}52} \cdot \left[\frac{(51 - 3)}{(3 - 1)}\right] = 1{,}0833 \cdot 24 = 26.$$

Критическое значение 1% F-критерия для двух степеней свободы в числителе и 48 — в знаменателе составляет 5,08. Так как правило принятия решения для проверки нулевой гипотезы $R^2 = 0$ состоит в том, чтобы отвергнуть H_0, если F больше критического значения, то мы отвергаем нулевую гипотезу.

Тест Чоу для проверки равнозначности коэффициентов в подпериодах

Иногда бывает необходимо проверить какую-либо гипотезу в различные периоды времени. В подобной ситуации нужно знать, являются ли полученные коэффициенты из отдельных временных периодов действительно значительно разными, либо эти различия случайны. Для этой цели мы можем применить тест Чоу.

Тест Чоу проводится в три этапа:

1. Производим расчеты уравнения регрессии для всего ряда данных и определяем остаточную сумму квадратов отклонений (СКО). Обозначим ее $СКО_1$.
2. Производим расчет регрессионной модели отдельно для различных периодов времени и определим собственные СКО в этих периодах. Допустим, что имеются два подпериода, один из n наблюдений, а другой из m. Таким образом, получим $СКО_2$ и $СКО_3$.
3. Рассчитаем критерий Чоу следующим образом:

$$\frac{\left(СКО_1 - СКО_2 - СКО_3\right) / k}{\left(СКО_2 + СКО_3\right) / \left(n + m - 2k\right)}, \qquad (6.45)$$

где n и m — число наблюдений в каждой соответствующей подгруппе.

Критерий Чоу имеет F-распределение с k степенями свободы в числителе и $m + n—2k$ степенями свободы в знаменателе.

РАССМОТРЕНИЕ ДОПУЩЕНИЙ МНК

Важно определить, выполнялись ли допущения МНК. Особенно важно провести проверку на:

- **гетероскедастичность** — дисперсия остатков не является постоянной;
- **автокорреляцию** — остатки независимы;
- **мультиколлинеарность** — независимые переменные некоррелированны.

Гетероскедастичность

Если остатки имеют постоянную дисперсию, они называются гомоскедастичными, но если они непостоянны, то гетероскедастичными. Гетероскедастичность приводит к тому, что коэффициенты регрессии больше не представляют собой лучшие оценки или не являются оценками с минимальной дисперсией, следовательно, они больше не являются наиболее эффективными коэффициентами.

Воздействие гетероскедастичности на оценку интервала прогнозирования и проверку гипотезы заключается в том, что хотя коэффициенты не смещены, дисперсии и, следовательно, стандартные ошибки этих коэффициентов будут смещены. Если смещение отрицательно, то оценочные стандартные ошибки будут меньше, чем они должны быть, а критерий проверки будет больше, чем в реальности. Таким образом, мы можем сделать вывод, что коэффициент значим, когда он таковым не является. И наоборот, если смещение положительно, то оценочные ошибки будут больше, чем они должны быть, а критерии проверки — меньше. Значит, мы можем принять нулевую гипотезу, в то время как она должна быть отвергнута.

Проверкой на гетероскедастичность служит тест Голдфелда—Кванта. Он требует, чтобы остатки были разделены на две группы из n наблюдений, одна группа с низкими, а другая — с высокими значениями. Обычно срединная одна шестая часть наблюдений удаляется после ранжирования в возрастающем порядке, чтобы улучшить разграничение между двумя группами. Отсюда число остатков в каждой группе составляет $(n—c)/2$, где c представляет одну шестую часть наблюдений.

Критерий Голдфелда—Кванта — это отношение суммы квадратов отклонений (СКО) высоких остатков к СКО низких остатков:

$$\text{ГК} = \frac{\text{СКО}_{\text{В}}}{\text{СКО}_{\text{Н}}}. \tag{6.46}$$

Этот критерий имеет F-распределение с $(n—c)/(2—k)$ степенями свободы.

Чтобы решить проблему гетероскедастичности, нужно исследовать взаимосвязь между значениями ошибки и переменными и трансформировать регрессионную модель так, чтобы она отражала эту взаимосвязь. Это может быть достигнуто посредством регрессии значений ошибок по различным формам функций переменной, которая приводит к гетероскедастичности, например,

$$e_i = \alpha + \beta X_i^H , \tag{6.47}$$

где X_i — независимая переменная (или какая-либо функция независимой переменной), которая предположительно является причиной гетероскедастичности, а H отражает степень взаимосвязи между ошибками и данной переменной, например, X^2 или $X^{1/n}$ и т. д.

Следовательно, дисперсия коэффициентов запишется:

$$E(\sigma_i^2) = \sigma^2 X_i^H. \tag{6.48}$$

Отсюда если $H = 1$, мы трансформируем регрессионную модель к виду

$$\frac{Y_i}{\sqrt{X_i}} = \frac{\alpha}{\sqrt{X_i}} + \beta_i \frac{e_i}{\sqrt{X_i}}. \tag{6.49}$$

Если $H = 2$, т.е. дисперсия увеличивается в пропорции к квадрату рассматриваемой переменной X, трансформация приобретает вид

$$\frac{Y_i}{X_i} = \frac{\alpha}{X_i} + \beta_i \frac{e_i}{X_i}. \tag{6.50}$$

Автокорреляция

Автокорреляция, также известная как сериальная корреляция, имеет место, когда остатки не являются независимыми друг от друга, потому что текущие значения Y находятся под влиянием прошлых значений. Зависимость между остатками описывается с помощью авторегрессионной схемы. Например, допустим, что остаток e_t находится под влиянием остатка из предыдущего периода времени e_{t-1} и какого-либо текущего значения случайной переменной z_t. Остаток e_t будет описываться следующей авторегрессионной функцией:

$$e_t = \rho e_{t-1} + z_t. \tag{6.51}$$

Эта форма авторегрессионной функции называется авторегрессионной функцией первого порядка или АР(1), так как только один предшествующий временной период включен в функцию.

Если бы текущий остаток находился под влиянием, скажем, двух или четырех предшествующих остатков, авторегрессионные функции АР(2) и АР(4) выглядели бы так:

$$\text{АР(2)}: e_t = \rho_{t-1} e_{t-1} + \rho_{t-2} e_{t-2} + z_t; \tag{6.52}$$

$$\text{АР(4)}: e_t = \rho_{t-1} e_{t-1} + \rho_{t-2} e_{t-2} + \rho_{t-3} e_{t-3} + \rho_{t-4} e_{t-4} + z_t. \tag{6.53}$$

Регрессионная модель МНК позволяет получить несмещенную оценку с минимальной дисперсией только тогда, когда остатки независимы друг от друга. Если у остатков существует автокорреляция, то коэффициенты регрессии не смещены, но стандартные ошибки будут недооценены, и проверки коэффициентов регрессии будут ненадежны.

Для проверки на автокорреляцию первого порядка необходимо рассчитать критерий Дарбина—Уотсона. Он определяется так:

$$DW = \frac{\sum (e_t - e_{t-1})^2}{\sum e_t^2}. \tag{6.54}$$

Эмпирическое правило гласит, что если критерий Дарбина—Уотсона равен двум, то не существует положительной автокорреляции, если он равен нулю, то имеет место совершенная положительная автокорреляция, а если он равен четырем, то имеет место совершенная отрицательная автокорреляция. Однако критерий Дарбина—Уотсона имеет выборочное распределение, базирующееся на таблице, представленной в их первоначальной работе (Durbin and Watson, 1950). Это выборочное распределение обладает двумя критическими значениями — d_L и d_U.

Чтобы провести проверку на автокорреляцию данных, проверяем следующую нулевую гипотезу

H_0: нет автокорреляции, если $d_U \le d \le 4-d_U$

H_1: положительная автокорреляция, если $d < d_L$,
отрицательная автокорреляция, если $d > 4-d_L$.

К сожалению, в составе данного распределения существуют некоторые зоны неопределенности, где результаты неокончательны. Это

$$d_L < d < d_U \text{ или } 4-d_U < d < 4-d_L.$$

Автокорреляция может появиться из-за опущения переменных, неверной формы функции в оценочном выражении, например, линейная модель тогда, когда она должна быть нелинейной. Введение переменных с лагами также способно привести к автокорреляции. Как мы обнаружим в гл. 7, данные из временных рядов тоже особенно подвержены автокорреляции.

Чтобы решить проблему автокорреляции, сначала следует рассмотреть возможность исключения переменных или неверной функциональной формы. Затем, если это не приводит к успешному решению проблемы, используют процедуру Оркулта—Кокроуна.

Для применения этой процедуры сначала рассчитаем коэффициент автокорреляции ρ следующим способом:

$$\rho = \frac{\sum (e_t \cdot e_{t-1})}{\sum e_t^2}. \qquad (6.55)$$

Затем переопределим выражение $Y_t = \alpha + \beta X_t$ как

$$Y_t - \rho Y_{t-1} = \alpha + \beta(X_t - \rho X_{t-1}). \qquad (6.56)$$

Это дает результат сдвига автокорреляции первого порядка для данных.

Мультиколлинеарность

Если некоторые или все независимые переменные в множественной регрессии являются высоко коррелированными, регрессионной модели трудно разграничить их отдельные объясняющие воздействия на Y. В результате высококоррелированные независимые переменные действуют в одном направлении и имеют недостаточно независимое колебание, чтобы дать возможность модели изолировать влияние каждой переменной. Не существует точного граничного значения уровня корреляции переменных, при котором возникает проблема мультиколлинеарности, таким образом, необходимо собственное суждение.

Мультиколлинеарность особенно часто имеет место при анализе макроэкономических данных, таких, как доходы и производство, где инфляция, например, может повлиять на оба ряда.

При мультиколлинеарности коэффициенты регрессии нестабильны как в отношении статистической значимости, так и по величине и знаку. Следовательно, они ненадежны. Значения коэффициентов R^2 могут быть высокими, но стандартные ошибки тоже высоки, и отсюда t-критерии малы, отражая недостаток значимости.

В отношении мультиколлинеарности может быть принято несколько мер:

1. Увеличивают объем выборки по принципу, что больше данных означает меньшие дисперсии оценок МНК. Проблема реализации этого варианта решения состоит в трудности нахождения дополнительных данных.
2. Исключают те переменные, которые высококоррелированны с остальными. Проблема здесь заключается в том, что возможно переменные были включены на теоретической основе, и будет неправомочным их исключение только лишь для того, чтобы сделать статистические результаты "лучше".
3. Объединяют данные кросс-секций и временных рядов. При этом методе берут коэффициент из, скажем, кросс-секционной регрессии и заменяют его на коэффициент из эквивалентных данных временного ряда. Покажем это на примере.

Допустим, мы хотим скорректировать следующее уравнение регрессии временного ряда

$$Q_t = \alpha + b_1 P_1 + b_2 Y_t + e_t, \qquad (6.57)$$

где Q представляет объем или стоимость денежных средств пенсионного фонда, инвестированных в британские акции, P отражает уровень индекса FTSE 100, а Y — это доход пенсионного фонда за предыдущий месяц. И уровень, и доход представляются высококоррелированными во времени, потому что оба находятся под влиянием инфляции.

Может быть доступным кросс-секционное исследование по опросу, касающемуся уровня дохода и инвестиций пенсионного фонда. К данным этой кросс-секции применим следующее выражение

$$Q = \lambda_1 + \lambda_2 Y. \qquad (6.58)$$

Затем скорректируем значение Q в первом выражении таким образом:

$$Q - \lambda_1 - \lambda_2 Y = Q^*. \qquad (6.59)$$

Следом за этим проведем регрессию Q^* от P следующим способом:

$$Q^* = \alpha_1 + \alpha_2 P. \qquad (6.60)$$

С помощью данного метода можно получить оценки параметров P и Y из источников, которые не будут высококоррелированными.

ФИКТИВНЫЕ ПЕРЕМЕННЫЕ

Иногда необходимо включение в регрессионную модель одной или более качественных переменных, например, инвесторы мужского и женского пола. Альтернативно может понадобиться сделать качественное различие между наблюдениями одних и тех же данных. Например, если проверяется взаимосвязь между размером компании и месячными доходами по акциям, может быть желательным включение качественной переменной, представляющей месяц январь, по причине хорошо известного "январского эффекта" во временных рядах доходов по ценным бумагам. Данный "январский эффект" — это феномен, заключающийся в том, что средние доходы по акциям, особенно небольших компаний, в среднем выше в январе, чем в другие месяцы. Таким образом, если мы рассматриваем январские наблюдения как качественно отличные от других наблюдений, фиктивная переменная (D) позволит произвести подобное качественное различие.

Фиктивные переменные используются также для отражения действия качественных эффектов изменений политики правительства на анализируемые данные. Например, при исследовании воздействия валюты на уровень фондового индекса следует принять к сведению то, что валюта входит или выходит из валютного блока. Например, исследования индекса FTSE 100, которые включают август/сентябрь 1992 г., будут вынуждены учесть выход фунта стерлингов из механизма обменных курсов Европейской валютной системы (ЕВС).

Фиктивные переменные бывают двух типов — **сдвига** и **наклона**. Фиктивная переменная сдвига — это переменная, которая меняет точку пересечения линии регрессии с осью ординат в случае применения качественной переменной. Фиктивная переменная наклона — это та переменная, которая изменяет наклон линии регрессии в случае использования качественной переменной. Оба типа фиктивных переменных будут иметь значение +1 или —1, когда наблюдения данных совпадают с уместной количественной переменной, но будут иметь нулевое значение при совпадении с наблюдениями, где эта качественная переменная отсутствует. Рис. 6.5 и 6.6 иллюстрируют эти определения.

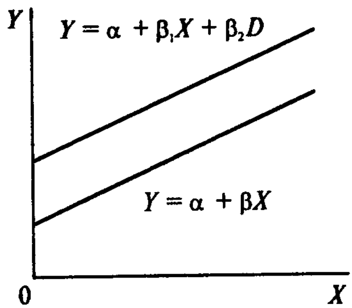

Рис. 6.5. Фиктивная переменная сдвига

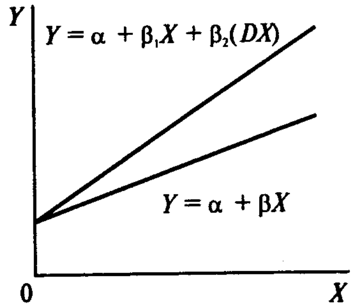

Рис. 6.6. Фиктивная переменная наклона

Чтобы проиллюстрировать фиктивную переменную сдвига, рассмотрим снова "январский эффект". Для проверки взаимосвязи между размером компании и доходами инвесторов необходимы данные, относящиеся к размеру компании за каждый месяц, к доходам по акциям компании за тот же месяц и качественная переменная (январь). Переменная "январь" будет представлена фиктивной переменной со значением единицы для наблюдений января и с нулем для всех других месяцев. Регрессия кросс-секции месячных доходов по размеру компании и отношениям к январю выглядит так:

$$r_i = \alpha + \beta_1 \text{размер}_i + \beta_2 D + e_i, \tag{6.61}$$

где величина D равна единице, когда наблюдения дохода и размера относятся к январю, и нулю, когда эти наблюдения относятся к другим месяцам.

В примере с индексом FTSE 100 в модель добавляется фиктивная переменная со значением 1, когда данные наблюдений совпадают с пребыванием фунта стерлингов в ЕВС, и со значением нуль, когда фунт стерлингов не был в ЕВС.

Для иллюстрации фиктивной переменной наклона рассмотрим эффективность деятельности инвестиционных менеджеров. Если они имеют навыки по выбору времени на рынке, доходы по их портфелям должны расти быстрее или падать меньше, чем в целом по рынку.

Этот подход был использован Хенрикссоном и Мертоном (1981), которые применили фиктивные переменные в линейной регрессии следующей формы

$$r_p - r_f = \alpha + \beta(r_m - r_f) + c[D(r_m - r_f)] + e_p, \tag{6.62}$$

где $r_p - r_f$ — превышение дохода по портфелю;
$r_m - r_f$ — превышение дохода по рынку в целом.

Фиктивная переменная D равна нулю, кода доходы по рынку в целом больше, чем безрисковая ставка r_f, и —1, когда доходы по рынку в целом меньше безрисковой ставки.

Это выражение истолковывается следующим образом. Когда доходы по рынку в целом выше, чем безрисковая ставка, $D = 0$, и, следовательно, выражение (6.62) сокращается:

$$r_p - r_f = \alpha + \beta(r_m - r_f) + e_p, \tag{6.63}$$

а когда доход по портфелю меньше безрисковой ставки, $D = -1$, то выражение (6.62) сокращается до следующего:

$$r_p - r_f = \alpha + (\beta - c)(r_m - r_f) + e_p. \tag{6.64}$$

Отсюда, так как β соответствует коэффициенту регрессии, когда рынок в целом дает доходы выше безрисковой ставки, а (β — c) соответствует портфелю β, когда рынок в целом дает доходы ниже безрисковой ставки, то c — это разность между двумя коэффициентами, отражающая, что инвестиционный менеджер имеет навыки по выбору времени на рынке. Если c положительна, то наклон линии регрессии будет менее крутым, чем при нулевом c.

Положительная α в выражениях (6.63) и (6.64) показывает, что инвестиционный менеджер имеет положительные навыки выбора акций.

Фиктивные переменные могут применяться в более чем одном качественном различии переменных. Например, сезонные колебания данных. Если мы предположим, что в Великобритании из-за окончания налогового года в апреле доходы по ценным бумагам имеют качественные отличия в январе и апреле, то для изучения взаимосвязи между месячными доходами и размером будут нужны две фиктивные переменные. Действительно, когда существует n качественных состояний данных, необходимо иметь n—1 фиктивных переменных. В нашем примере присутствуют три качественных состояния данных дохода и размера — данные, относящиеся к январю, данные, относящиеся к апрелю, и данные, относящиеся к остальным месяцам. Значит, понадобятся две фиктивные переменные.

НЕЛИНЕЙНАЯ РЕГРЕССИЯ

До сих пор обсуждение было сфокусировано на линейной регрессии. Однако может случиться так, что взаимосвязь между зависимой переменной и одной или более независимыми переменными будет нелинейной. Существуют два пути решения этой проблемы:

1) преобразовать данные и применить линейную регрессию;
2) применить методы нелинейной регрессии.

Нелинейная регрессия не обсуждается в этой книге, так что далее остановимся на преобразовании данных.

ПРЕОБРАЗОВАНИЯ ДАННЫХ

Графики на рис. 6.7 показывают разнообразные взаимосвязи между Y и X, не являющиеся линейными. Однако при соответствующем преобразовании Y, α и X взаимосвязь между Y и X может быть трансформирована в линейную для a и b. Таким образом, далее можно использовать МНК.

Рассмотрим три нелинейные формы, отображенные на рис. 6.7. На верхних графиках функциональной формой является $Y = \alpha X^{\beta}$, где $0 < \beta < 1$ или $\beta > 1$. Преобразование в этих случаях заключается во взятии натурального логарифма от Y, α и X. Получающееся уравнение регрессии будет выглядеть так:

$$\ln Y = \ln \alpha + \beta \ln X. \qquad (6.65)$$

Преобразование для нижнего правого графика очень простое, если учесть, что $1/X$ может участвовать в расчете как независимая переменная.

При всех этих преобразованиях необходимо конвертировать результат в нелинейную форму для его правильного истолкования.

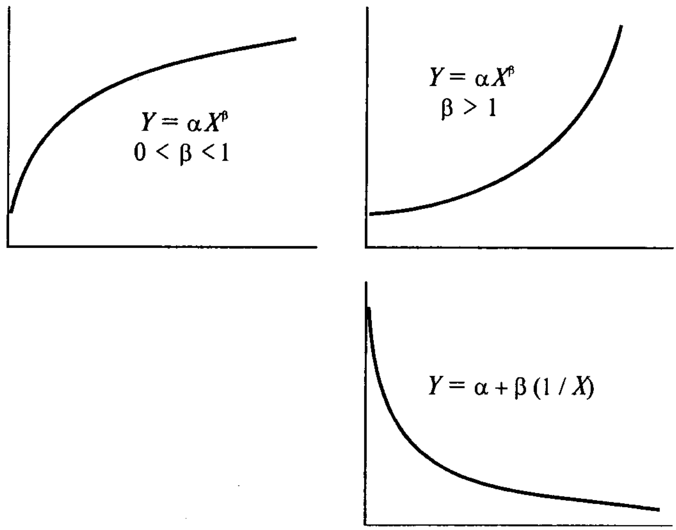

Рис. 6.7

ПРИМЕНЕНИЕ РЕГРЕССИОННОГО АНАЛИЗА В ХЕДЖИРОВАНИИ

Цель хеджирования состоит в устранении риска для портфеля активов. Хеджирование длинной позиции (на покупку) по активу, имеющему риск, достигается занятием короткой позиции (на продажу) в некоторой пропорции от стоимости портфеля по другому, но высококоррелированному рискованному активу. Для иллюстрации этого рассмотрим портфель, имеющий длинную позицию по активу *A*, которую желательно захеджировать путем занятия короткой позиции по фьючерсному контракту на актив *A*.

До установки хеджа необходимо ответить на два вопроса:

1. По какому инструменту занять короткую позицию?
2. Какую пропорцию от стоимости длинной позиции должна представлять короткая позиция, чтобы минимизировать дисперсию всего портфеля?

На первый вопрос легко ответить, рассмотрев коэффициент корреляции изменений цены длинной позиции и потенциальных кандидатур для короткой позиции. Следует выбрать кандидатуру с наивысшей корреляцией с длинной позицией. В данном примере мы допускаем, что это фьючерсный контракт.

Что касается второго вопроса, пропорция короткой позиции называется коэффициентом хеджирования. Для нахождения этих коэффициентов часто используется МНК регрессия.

Чтобы понять это, рассмотрим снова длинную позицию по облигации с доходом по ней R_B. Определено, что высококоррелированным инструментом является фьючерс на облигации. Доход по фьючерсу выражается в виде R_F.

Доходы по захеджированному портфелю, т.е. состоящему из длинной позиции по одной единице облигации и короткой по соответствующей величине h фьючерсного контракта, определяются так:

$$R_p = (R_B - hR_F), \tag{6.66}$$

где h — коэффициент хеджирования.

Дисперсия доходов по портфелю составляет

$$\sigma^2 R_p = \sigma^2 R_B - 2h\ \mathrm{cov}(R_B, R_F) + h^2\sigma^2 R_F. \tag{6.67}$$

Для нахождения коэффициента хеджирования h, который минимизирует дисперсию захеджированного портфеля, следует продифференцировать выражение (6.67) по h и приравнять производную к нулю:

$$\frac{d\left(\sigma^2 R_B - 2h\,\mathrm{cov}\left(R_B, R_F\right) + h^2\sigma^2 R_F\right)}{dh} = -2\,\mathrm{cov}\left(R_B, R_F\right) + 2h\sigma^2 R_F. \tag{6.68}$$

Чтобы приравнять это к нулю, нужно, чтобы

$$2h\sigma^2 R_F = 2\mathrm{cov}(R_B, R_F). \tag{6.69}$$

В таком случае h равно

$$h = \frac{\mathrm{cov}\left(R_B, R_F\right)}{\sigma^2 R_F}, \tag{6.70}$$

h признается аналогичным β — наклону линии регрессии

$$R_B = \alpha + \beta R_F + e. \tag{6.71}$$

Таким образом, коэффициент наклона β дает коэффициент хеджирования. На практике регрессия строится с использованием $\Delta P_B/P_B$ и $\Delta P_F/P_F$, где P представляет цены облигации и фьючерса. Этот пример может быть обобщен для применения регрессионного анализа в хеджировании любого рискованного актива, однако существуют альтернативные методы, которые показаны в гл. 7, относящейся к анализу временных рядов.

УПРАЖНЕНИЯ

Регрессия

1. Что вы понимаете под терминами:

   кросс-секционная регрессия,
   регрессия временных рядов,
   ложная регрессия.

   Следующие упражнения относятся к рассмотрению приведенного ниже результата регрессионного анализа. Числа в скобках — это t-критерии.

$$\mathbf{Y = -0{,}01 + 0{,}04X_1 + 0{,}9X_2 - 0{,}1X_3}$$
$$(-0{,}2) \quad (+\,1{,}75) \quad (+\,3{,}2) \quad (-2{,}6)$$

$R^2 = 0{,}85$; $CC = 51$.
Скорректированный $R^2 = 0{,}80$; $DW = 2{,}4$;
$F = 27$.

2. Дайте общее объяснение уравнения множественной регрессии
   $$Y = -0{,}01 + 0{,}04X_1 + 0{,}9X_2 - 0{,}1X_3,$$
   включая обсуждение значимости каждой переменной.
3. Объясните значение R^2 и скорректированного R^2 и истолкуйте критерии, приведенные выше.
4. Что такое автокорреляция, каковы некоторые из ее причин? Объясните, как она влияет на толкование уравнения регрессии и расскажите о роли критерия Дарбина—Уотсона в определении присутствия автокорреляции.
5. Каково толкование F-критерия для указанных выше данных?
6. Что такое гетероскедастичность, каковы некоторые из ее причин? Как она влияет на толкование уравнения регрессии?
7. Какую проверку вы бы предпочли, чтобы определить, будет или нет гетероскедастичность проблемой при анализе данных? Как бы вы решили эту проблему?
8. Что такое мультиколлинеарность, каковы некоторые из ее причин? Как она влияет на толкование регрессионного уравнения, и как бы вы решили эту проблему, если бы она существовала?

9. Что такое фиктивные переменные, и как они применяются в регрессионном анализе?
10. Объясните, как использовать регрессионный анализ для нахождения коэффициентов хеджирования с минимальным риском, когда для хеджирования рискованных активов применяются финансовые фьючерсы?

Матричная алгебра

11.

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 6 & 0 & 5 \\ 9 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix},\ B = \begin{bmatrix} 1 & 0 \\ 5 & 2 \\ 7 & 1 \end{bmatrix},\ C = \begin{bmatrix} 0 & 8 \\ 5 & 0 \\ 1 & 1 \end{bmatrix},\ D = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}$$

а) Найдите $B + C$. Найдите $A[B + C]$. Найдите AB. Найдите BC. Найдите $AB + BC$.

б) Транспонированная матрица (обозначенная X^{T}) строится перестановкой строк и столбцов.
Найдите $B^{\mathrm{T}}C$ и BC^{T}.
Найдите $B^{\mathrm{T}}[DC]$ и $[B^{\mathrm{T}}C]D$.

в) Найдите D^{-1} и $[B^{\mathrm{T}}C]^{-1}$ Покажите, что матрица $[BC^{\mathrm{T}}]$ не имеет обратной.

г) Докажите, что матрица $[B^{\mathrm{T}}C]^{\mathrm{T}} = C^{\mathrm{T}}B$.
Докажите, что $[[B^{\mathrm{T}}C]D]^{-1} = D^{-1}[B^{\mathrm{T}}C]^{-1}$.

12. Доходы по ценной бумаге Y предположительно зависят от доходов по ценным бумагам X_1 и X_2. Годовые ставки месячных доходов по трем ценным бумагам показаны ниже для шестимесячного периода:

Ценная бумага	Доходы					
X_1	0,054	0,053	0,049	0,049	0,054	0,060
X_2	0,063	0,062	0,061	0,058	0,057	0,057
Y	0,092	0,092	0,091	0,090	0,087	0,086

Регрессионная модель должна быть построена в форме $Y_i = \alpha_0 + \alpha_1 X_{i1} + \alpha_2 X_{i2} + e_i$, где e_i — независимые и нормально распределенные переменные, все имеющие одинаковую дисперсию.

а) Постройте матрицу X так, чтобы модель определялась посредством

$$\underline{Y} = X \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} + \underline{e}$$

Замечание: $\underline{Y}$ и $\underline{e}$ — это векторы Y и e соответственно.

б) Найдите $X^{\mathrm{T}}X$.

в) Опишите как рассчитать $[X^{\mathrm{T}}X]^{-1}$ с использованием разделенной матрицы.

г) Подтвердите, что обратная матрица приблизительно имеет вид

$$\begin{bmatrix} 189{,}11155 & -1084{,}56021 & -2200{,}26425 \\ -1084{,}56021 & 13338{,}36668 & 6291{,}68240 \\ -2200{,}26425 & 6291{,}68240 & 31269{,}66151 \end{bmatrix}.$$

д) Рассчитайте $X^{\mathrm{T}}\underline{Y}$.

е) Используйте $[X^{\mathrm{T}}X]^{-1}$ и $X^{\mathrm{T}}\underline{Y}$ для расчета оценок наименьших квадратов для α_i, $i = 0, 1, 2$.

ж) Рассчитайте дисперсионно-ковариационную матрицу для ценных бумаг X_1 и X_2.

ОТВЕТЫ

11. а)

$$B + C = \begin{bmatrix} 1 & 8 \\ 10 & 2 \\ 8 & 2 \end{bmatrix} \quad A[B + C] = \begin{bmatrix} 40 & 24 \\ 46 & 58 \\ 37 & 78 \\ 35 & 16 \end{bmatrix}$$

$$AB = \begin{bmatrix} 24 & 7 \\ 41 & 5 \\ 26 & 5 \\ 27 & 5 \end{bmatrix} \quad AC = \begin{bmatrix} 16 & 17 \\ 5 & 53 \\ 11 & 73 \\ 8 & 11 \end{bmatrix}$$

$$AB + BC = \begin{bmatrix} 40 & 24 \\ 46 & 58 \\ 37 & 78 \\ 35 & 16 \end{bmatrix}.$$

б)

$$B^{\mathrm{T}}C = \begin{bmatrix} 32 & 15 \\ 11 & 1 \end{bmatrix}, \; BC^{\mathrm{T}} = \begin{bmatrix} 0 & 5 & 1 \\ 16 & 25 & 7 \\ 8 & 35 & 8 \end{bmatrix}$$

$$B^{\mathrm{T}}[CD] = \begin{bmatrix} 64 & 111 \\ 22 & 34 \end{bmatrix}, \; [B^{\mathrm{T}}C]D = \begin{bmatrix} 64 & 111 \\ 22 & 34 \end{bmatrix}$$

в)

$$D^{-1} = \begin{bmatrix} 0{,}5 & -1{,}5 \\ 0{,}0 & 1{,}0 \end{bmatrix}, [B^{\mathrm{T}}C]^{-1} = \begin{bmatrix} -0{,}0075 & 0{,}1128 \\ 0{,}0827 & -0{,}2406 \end{bmatrix}$$

г)

$$[B^{\mathrm{T}}C]^{\mathrm{T}} = \begin{bmatrix} 32 & 11 \\ 15 & 1 \end{bmatrix}$$

$$C^{\mathrm{T}}B = \begin{bmatrix} 0 & 5 & 1 \\ 8 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 5 & 2 \\ 7 & 1 \end{bmatrix} = \begin{bmatrix} 32 & 11 \\ 15 & 1 \end{bmatrix}$$

$$[[B^{\mathrm{T}}C]D]^{-1} = \begin{bmatrix} -0{,}1278 & 0{,}8346 \\ 0{,}0827 & -0{,}2406 \end{bmatrix}$$

$$D^{-1}[B^{\mathrm{T}}C]^{-1} = \begin{bmatrix} 0{,}5 & -1{,}5 \\ 0{,}0 & 1{,}0 \end{bmatrix} \begin{bmatrix} -0{,}0075 & 0{,}1128 \\ 0{,}0827 & -0{,}2406 \end{bmatrix} = \begin{bmatrix} -0{,}1278 & 0{,}8346 \\ 0{,}0827 & -0{,}2406 \end{bmatrix}$$

12. а)

$$X = \begin{bmatrix} 1 & 0{,}054 & 0{,}063 \\ 1 & 0{,}053 & 0{,}062 \\ 1 & 0{,}049 & 0{,}061 \\ 1 & 0{,}049 & 0{,}058 \\ 1 & 0{,}054 & 0{,}057 \\ 1 & 0{,}060 & 0{,}057 \end{bmatrix}$$

б)

$$X^{\mathrm{T}}X = \begin{bmatrix} 6{,}000000 & 0{,}319000 & 0{,}358000 \\ 0{,}319000 & 0{,}017043 & 0{,}019017 \\ 0{,}358000 & 0{,}019017 & 0{,}021396 \end{bmatrix}$$

в)

$$X^{\mathrm{T}}\underline{Y} = \begin{bmatrix} 0{,}538000 \\ 0{,}028571 \\ 0{,}032132 \end{bmatrix}$$

г)

$$\alpha = [X^{\mathrm{T}}X]^{-1}[X^{\mathrm{T}}\underline{Y}] = \begin{bmatrix} 0{,}056 \\ -0{,}239 \\ 0{,}774 \end{bmatrix}$$

д)

$$\overline{X}_1 = 0{,}053167;\ \overline{X}_2 = 0{,}059667$$

$$[\mathrm{cov}(X_1, X_2)] = \frac{1}{6}\begin{bmatrix} 0{,}017043 & 0{,}019017 \\ 0{,}019017 & 0{,}021396 \end{bmatrix} \left[\begin{bmatrix} \overline{X}_1 \\ \overline{X}_2 \end{bmatrix} [\overline{X}_1 \quad \overline{X}_2]\right]$$

$$= \begin{bmatrix} 0{,}002841 & 0{,}003170 \\ 0{,}003170 & 0{,}003566 \end{bmatrix} - \begin{bmatrix} 0{,}002827 & 0{,}003172 \\ 0{,}003172 & 0{,}003560 \end{bmatrix}$$

$$= \begin{bmatrix} 0{,}000014 & -0{,}00002 \\ -0{,}000002 & 0{,}000006 \end{bmatrix}.$$

СПИСОК ИСПОЛЬЗУЕМОЙ И РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

Henriksson, R. D. and Merton, R. C. (1981) On market timing and investment performance II. Statistical procedures for evaluating forecasting skills. *Journal of Business*, October, 513—33.

Mansfield, E. (1991) *Statistics for Business and Economics*, 4th edn. Norton, New York.

Gujarati, D. (1995) *Basic Econometrics*, 3rd edn. McGraw-Hill.

Pindyck, R. (1990) *Econometric Models and Economic Forecasting*, 3rd edn. McGraw-Hill.

ПРИЛОЖЕНИЕ 6.1
Матричная алгебра

Матрицы — это массивы данных, имеющие прямоугольную форму. Данные располагаются в строках и столбцах. Размер матрицы выражается числом строк и столбцов, т.е. матрица с пятью строками и четырьмя столбцами будет матрицей размера 5×4. Например, следующая матрица имеет размер 3×3

$$Y = \begin{bmatrix} 1 & 4 & 6 \\ 3 & 7 & 5 \\ 8 & 6 & 3 \end{bmatrix}.$$

Положение отдельных ячеек в матрице определяется сперва по позиции их строки, а затем по позиции их столбца. Таким образом, тройка слева в средней строке в приведенной выше матрице находится в ячейке 2, 1, в то время как тройка справа в нижней строке находится в ячейке 3, 3.

Векторы — это строки или столбцы данных, фактически это тоже матрицы, только с одной строчкой или одним столбцом. По правилу определения ячеек в матрицах вектор строки из n элементов будет матрицей размера $1 \times n$, а вектор столбца — матрицей размера $n \times 1$.

Часто для идентификации элемента в матрице используются двузначные подстрочные индексы. Например, элемент на пересечении i-ой строки и j-го столбца матрицы Y будет отображен как y_{ij}.

Матрицы могут перемножаться, если их формы удовлетворяют особым условиям. Они могут складываться и вычитаться, только если это матрицы одной формы.

Сложение и вычитание матриц

Матрицы складываются друг с другом посредством сложения каждого элемента из одной матрицы с соответствующим элементом другой матрицы. Вычитание матриц достигается вычитанием каждого элемента второй матрицы из соответствующего элемента первой. В обобщенном виде это записывается так:

$$X + Y = [x_{ij} + y_{ij}];$$

$$X - Y = [x_{ij} - y_{ij}],$$

где X и Y — это матрицы, x_{ij} и y_{ij} — отдельные элементы соответствующих матриц. Для иллюстрации этого рассмотрим следующие две матрицы:

$$X = \begin{bmatrix} 8 & 6 \\ 3 & 2 \end{bmatrix}, \; Y = \begin{bmatrix} 9 & 4 \\ 6 & 1 \end{bmatrix};$$

$X + Y$ составит

$$X + Y = \begin{bmatrix} (8+9) & (6+4) \\ (3+6) & (2+1) \end{bmatrix} = \begin{bmatrix} 17 & 10 \\ 9 & 3 \end{bmatrix}$$

$X - Y$ составит

$$X - Y = \begin{bmatrix} (8-9) & (6-4) \\ (3-6) & (2-1) \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ -3 & 1 \end{bmatrix}.$$

Умножение матриц

Умножение матрицы на единичное число известно как скалярное умножение. Каждая ячейка в матрице просто умножается на это единичное число, или скаляр. Чтобы проиллюстрировать это, умножим приведенную выше матрицу X на скаляр 5

$$5X = \begin{bmatrix} (5 \cdot 8) & (5 \cdot 6) \\ (5 \cdot 3) & (5 \cdot 2) \end{bmatrix} \begin{bmatrix} 40 & 30 \\ 15 & 10 \end{bmatrix}.$$

Для умножения одной матрицы на другую сначала необходимо убедиться, что матрицы подходят для умножения: каждая строка в первой (левой) матрице должна иметь такое же количество ячеек, сколько во второй (правой) матрице. Другими словами, число столбцов в первой (левой) матрице должно совпадать с числом строк во второй (правой) матрице. Для иллюстрации этого рассмотрим матрицы

$$X = \begin{bmatrix} 6 & 2 & 1 \\ 8 & 9 & 4 \end{bmatrix}, \quad Y = \begin{bmatrix} 2 & 8 \\ 3 & 4 \\ 1 & 6 \end{bmatrix}.$$

Чтобы умножить X на Y, необходимо умножить каждый элемент в первой строке X на соответствующий элемент первого столбца Y и расположить их в первой ячейке матрицы Z. Этот процесс повторяется, берутся элементы из первой строки X и умножаются на элементы второго столбца матрицы Y и результат располагается во второй ячейке первой строки матрицы Z. Эта процедура повторяется для нахождения второй строки Z.

Например, значение первой ячейки в матрице Z определяется как $(6 \cdot 2) + (2 \cdot 3) + (1 \cdot 1) = 19$. Вторая ячейка первой строки матрицы Z определяется как $(6 \cdot 8) + (2 \cdot 4) + (1 \cdot 6) = 62$.

Первая ячейка второй строки матрицы Z находится как $(8 \cdot 2) + (9 \cdot 3) + (4 \cdot 1) = 47$. Вторая ячейка во второй строке матрицы Z находится как $(8 \cdot 8) + (9 \cdot 4) + (4 \cdot 6) = 124$.

$X \cdot Y$, таким образом, выглядит как

$$\begin{bmatrix} 19 & 62 \\ 47 & 124 \end{bmatrix}.$$

Заметьте, что размер матрицы Z составляет только $2 \cdot 2$. Это потому, что размер матрицы-произведения равен числу строк в первой матрице и числу столбцов во второй.

Важно понять, что в матричной алгебре XY не является тем же самым, что YX, т.е. матричное умножение некоммутативно. Читателю предлагается доказать в качестве упражнения, что, используя вышеуказанные матрицы X и Y, получим YX, равное

$$\begin{bmatrix} 76 & 76 & 34 \\ 50 & 42 & 19 \\ 54 & 56 & 25 \end{bmatrix}.$$

Более того, видно, что матрица YX даже имеет отличную от матрицы XY форму.

Вычисление обратной матрицы

Только квадратные матрицы и к тому же только некоторые из квадратных матриц имеют обратные матрицы. Обратная матрица — это такая, произведение которой на исходную матрицу равно единичной матрице. В единичной матрице каждая ячейка равна нулю за исключением главной диагонали, содержащей единицы. Например, единичная матрица 3×3 выглядит так:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Единичная матрица обладает следующим свойством: когда она используется в матричном умножении, другая матрица-множитель остается неизменной.

Матрицы нельзя делить. Однако если квадратная матрица имеет обратную, тогда можно вместо этого умножить на обратную для получения желаемых результатов (смотрите ниже для применения). Матрица, обратная матрице X, обозначается X^{-1}. Чтобы найти обратную для матрицы, можно сформировать разделенную матрицу посредством расположения единичной матрицы рядом с инвертируемой (той, для которой находится обратная), например,

$$\left[\begin{array}{ccc||ccc} 6 & 2 & 1 & 1 & 0 & 0 \\ 8 & 9 & 4 & 0 & 1 & 0 \\ 3 & 1 & 0 & 0 & 0 & 1 \end{array}\right].$$

Далее исходная матрица должна быть преобразована в единичную через сложение, вычитание, умножение или деление каждой строки. Когда матрица слева превратится в единичную, матрица, полученная справа, будет представлять собой обратную к исходной матрице. Это показано ниже.

Возьмем записанную выше разделенную матрицу X

$$\left[\begin{array}{ccc||ccc} 6 & 2 & 1 & 1 & 0 & 0 \\ 8 & 9 & 4 & 0 & 1 & 0 \\ 3 & 1 & 0 & 0 & 0 & 1 \end{array}\right].$$

Для формирования новой матрицы Z, такой, что единичная матрица возникнет слева, начнем с преобразования ячейки x_{11} до единицы с помощью деления строки на шесть.

$$\left[\begin{array}{ccc|ccc} 1 & 0{,}33 & 0{,}17 & 0{,}17 & 0 & 0 \\ 8 & 9 & 4 & 0 & 1 & 0 \\ 3 & 1 & 0 & 0 & 0 & 1 \end{array}\right].$$

Чтобы привести 8 в ячейке x_{21} к нулю, отнимем 8 · строка 1 из строки 2.

$$\left[\begin{array}{ccc|ccc} 1 & 0{,}33 & 0{,}17 & 0{,}17 & 0 & 0 \\ 0 & 6{,}33 & 2{,}67 & -1{,}33 & 1 & 0 \\ 3 & 1 & 0 & 0 & 0 & 1 \end{array}\right].$$

Для обращения 3 из ячейки x_{31} в 0 отнимем 3 · строка 1 из строки 3.

$$\left[\begin{array}{ccc|ccc} 1 & 0{,}33 & 0{,}17 & 0{,}17 & 0 & 0 \\ 0 & 6{,}33 & 2{,}67 & -1{,}33 & 1 & 0 \\ 0 & 0 & -0{,}5 & -0{,}5 & 0 & 1 \end{array}\right].$$

Для преобразования 6,33 в ячейке x_{22} к 1 разделим строку 2 на 6,33.

$$\left[\begin{array}{ccc|ccc} 1 & 0{,}33 & 0{,}17 & 0{,}17 & 0 & 0 \\ 0 & 1 & 0{,}42 & -0{,}21 & 0{,}16 & 0 \\ 0 & 0 & -0{,}5 & -0{,}5 & 0 & 1 \end{array}\right].$$

Для преобразования 0,33 в ячейке x_{12} к 0 отнимем 0,33 · строка 2 из строки 1.

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0{,}03 & 0{,}24 & -0{,}05 & 0 \\ 0 & 1 & 0{,}42 & -0{,}21 & 0{,}16 & 0 \\ 0 & 0 & -0{,}5 & -0{,}5 & 0 & 1 \end{array}\right].$$

Чтобы привести —0,5 из ячейки x_{33} к 1, умножим строку 3 на —2.

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0{,}03 & 0{,}24 & -0{,}05 & 0 \\ 0 & 1 & 0{,}42 & -0{,}21 & 0{,}16 & 0 \\ 0 & 0 & 1 & 1 & 0 & -2 \end{array}\right].$$

Для обращения 0,03 в ячейке x_{13} в 0 вычтем 0,03 · строка 3 из строки 1.

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0{,}21 & -0{,}05 & 0{,}05 \\ 0 & 1 & 0{,}42 & -0{,}21 & 0{,}16 & 0 \\ 0 & 0 & 1 & 1 & 0 & -2 \end{array}\right].$$

Для приведения 0,42 из ячейки x_{23} к 0 отнимем 0,42 · строка 3 из строки 2.

$$\left[\begin{array}{ccc||ccc} 1 & 0 & 0 & 0{,}21 & -0{,}05 & 0{,}05 \\ 0 & 1 & 0 & -0{,}63 & 0{,}16 & 0{,}84 \\ 0 & 0 & 1 & 1 & 0 & -2 \end{array}\right].$$

Таким образом, обратная матрица к

$$\begin{bmatrix} 6 & 2 & 1 \\ 8 & 9 & 4 \\ 3 & 1 & 0 \end{bmatrix}$$

выглядит как

$$\begin{bmatrix} 0{,}21 & -0{,}05 & 0{,}05 \\ -0{,}63 & 0{,}16 & 0{,}84 \\ 1 & 0 & -2 \end{bmatrix}$$

(с точностью до двух знаков после запятой).

Решение систем уравнений

Матричная алгебра может быть использована для решения систем уравнений, и именно это свойство позволяет применить ее в вычислении коэффициентов регрессии по методу наименьших квадратов.

Чтобы показать, как матричная алгебра может быть приложена к решению систем уравнений, рассмотрим следующее уравнение:

$$6a + 8b = 25,$$

$$8a - 3b = 20.$$

Наша цель заключается в нахождении значений a и b, которые приводят к решению обоих уравнений.

Это можно записать в матричной форме:

$$\begin{bmatrix} 6 & 8 \\ 8 & -3 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 25 \\ 20 \end{bmatrix}.$$

Мы можем найти требуемые значения a и b следующим способом.

Предварительно умножим обе стороны на матрицу, обратную матрице коэффициентов

$$\begin{bmatrix} 6 & 8 \\ 8 & -3 \end{bmatrix}^{-1} \begin{bmatrix} 6 & 8 \\ 8 & -3 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 8 & -3 \end{bmatrix}^{-1} \begin{bmatrix} 25 \\ 20 \end{bmatrix}.$$

Это дает

$$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}a\\b\end{bmatrix}=\begin{bmatrix}6 & 8\\8 & -3\end{bmatrix}^{-1}\begin{bmatrix}25\\20\end{bmatrix},$$

т.е.

$$\begin{bmatrix}a\\b\end{bmatrix}=\begin{bmatrix}6 & 8\\8 & -3\end{bmatrix}^{-1}\begin{bmatrix}25\\20\end{bmatrix}.$$

Теперь

$$\begin{bmatrix}6 & 8\\8 & -3\end{bmatrix}^{-1}=\begin{bmatrix}0{,}036585 & 0{,}097561\\0{,}097561 & -0{,}07317\end{bmatrix},$$

таким образом,

$$\begin{bmatrix}a\\b\end{bmatrix}=\begin{bmatrix}0{,}036585 & 0{,}097561\\0{,}097561 & -0{,}07317\end{bmatrix}\begin{bmatrix}25\\20\end{bmatrix}=\begin{bmatrix}2{,}866\\0.976\end{bmatrix}.$$

Получаем $a = 2{,}866$ и $b = 0{,}976$.

Приложение к МНК регрессии

Для того чтобы увидеть, как эта методология применяется в регрессионном анализе, возьмем следующее оценочное уравнение регрессии

$$\hat{Y} = \hat{B}_0 + \hat{B}_1 X_1 + \hat{B}_2 X_2 + \ldots + \hat{B}_k X_k + e\,.$$

Для каждого из наблюдений Y и каждого набора наблюдений X переменных может быть составлена система уравнений, как показано ниже:

$$\begin{aligned}
Y_1 &= \hat{B}_0 + \hat{B}_1 X_{11} + \hat{B}_2 X_{21} + \ldots + \hat{B}_k X_{k1} + e_1\\
Y_2 &= \hat{B}_0 + \hat{B}_1 X_{12} + \hat{B}_2 X_{22} + \ldots + \hat{B}_k X_{k2} + e_2\\
&\quad\vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots\\
Y_n &= \hat{B}_0 + \hat{B}_1 X_{1n} + \hat{B}_2 X_{2n} + \ldots + \hat{B}_k X_{kn} + e_n
\end{aligned}$$

Можно затем составить вектор значений Y из n элементов, $n \cdot k$ матрицу значений X и вектор значений e из n элементов, например,

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ . \\ . \\ Y_n \end{bmatrix}; \; X = \begin{bmatrix} 1 & X_{11} & X_{21} & . & . & X_{k1} \\ 1 & X_{12} & X_{22} & . & . & X_{k2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 1 & X_{1n} & X_{2n} & . & . & Xkn \end{bmatrix}; \; e = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ . \\ . \\ e_n \end{bmatrix}.$$

Столбец единиц в матрице X представляет постоянный член в уравнении.

Цель — нахождение вектора B_i.

$$\begin{bmatrix} 1 & X_{11} & X_{21} & . & . & X_{k1} \\ 1 & X_{12} & X_{22} & . & . & X_{k2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 1 & X_{1n} & X_{2n} & . & . & Xkn \end{bmatrix} \cdot \begin{bmatrix} B_0 \\ B_1 \\ . \\ . \\ . \\ B_k \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_n \end{bmatrix}.$$

Если приведенная выше матрица квадратна, т.е. имеет то же число наблюдений, что и неизвестных ($k = n$), тогда она может иметь обратную матрицу. В таком случае мы приходим к

$$\begin{bmatrix} B_0 \\ B_1 \\ . \\ . \\ . \\ B_k \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & . & . & X_{k1} \\ 1 & X_{12} & X_{22} & . & . & X_{k2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 1 & X_{1n} & X_{2n} & . & . & Xkn \end{bmatrix}^{-1} \cdot \begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_n \end{bmatrix}.$$

Проблема эмпирического исследования заключается в том, что обычно существует больше наблюдений (n), чем независимых переменных (k), матрица X не является квадратной, а мы помним, что можно инвертировать только квадратные матрицы. МНК ведет к предварительному умножению матрицы X и вектора Y на транспонированную из X матрицу X^T. Матрица, транспонированная из приведенной выше матрицы X, — это

$$\begin{bmatrix} 1 & 1 & \ldots & 1 \\ X_{11} & X_{12} & \ldots & X_{1n} \\ X_{21} & X_{22} & \ldots & X_{2n} \\ X_{k1} & X_{k2} & \ldots & X_{kn} \end{bmatrix}.$$

Формула для вектора коэффициентов регрессии выглядит так:

$$\hat{B} = [X^{\mathrm{T}}X]^{-1}X^{\mathrm{T}}Y.$$

$[X^{\mathrm{T}}X]^{-1}$ — это обратная к матрице, получающейся от умножения транспонированной из X матрицы на матрицу X. Эта обратная матрица снова умножается на транспонированную из X, а произведение умножается на вектор Y.

Пример с решением

Имеются следующие данные по одной зависимой переменной Y и трем независимым переменным V_1, V_2, V_3:

Y	Постоянная	V_1	V_2	V_3
7	1	13	6	13
9	1	15	8	11
11	1	10	7	12
8	1	12	7	9
8	1	11	8	12
7	1	13	7	16
8	1	11	7	11

После выражения в матричной форме имеем

$$X = \begin{bmatrix} 1 & 13 & 6 & 13 \\ 1 & 15 & 8 & 11 \\ 1 & 10 & 7 & 12 \\ 1 & 12 & 7 & 9 \\ 1 & 11 & 8 & 12 \\ 1 & 13 & 7 & 16 \\ 1 & 11 & 7 & 11 \end{bmatrix}, Y = \begin{bmatrix} 7 \\ 9 \\ 11 \\ 8 \\ 8 \\ 7 \\ 8 \end{bmatrix}.$$

Матрица X^{T}, транспонированная из X, выглядит так:

$$X^{\mathrm{T}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 13 & 15 & 10 & 12 & 11 & 13 & 11 \\ 6 & 8 & 7 & 7 & 8 & 7 & 7 \\ 13 & 11 & 12 & 9 & 12 & 16 & 11 \end{bmatrix}.$$

$X^{\mathrm{T}} \cdot X$ — это

$$\begin{bmatrix} 7 & 85 & 50 & 84 \\ 85 & 1049 & 608 & 1023 \\ 50 & 608 & 360 & 598 \\ 84 & 1023 & 598 & 1036 \end{bmatrix};$$

$[X^{\mathrm{T}}X]^{-1}$ — это

$$\begin{bmatrix} 32{,}54 & -0{,}48 & -2{,}76 & -0{,}57 \\ -0{,}48 & 0{,}06 & -0{,}02 & -0{,}01 \\ -2{,}76 & -0{,}02 & 0{,}38 & 0{,}03 \\ -0{,}57 & -0{,}01 & 0{,}03 & 0{,}04 \end{bmatrix}$$

$[X^{\mathrm{T}}X]^{-1}X^{\mathrm{T}}$ — это

$$\begin{bmatrix} 2{,}31 & -3{,}01 & 1{,}56 & 2{,}33 & -1{,}68 & -2{,}17 & 1{,}66 \\ 0{,}07 & 0{,}16 & -0{,}13 & 0{,}02 & -0{,}09 & 0{,}02 & 0{,}06 \\ -0{,}43 & 0{,}22 & -0{,}002 & -0{,}14 & 0{,}35 & 0{,}04 & -0{,}06 \\ -0{,}002 & -0{,}04 & 0{,}01 & -0{,}12 & 0{,}03 & 0{,}14 & -0{,}03 \end{bmatrix}.$$

$[X^{\mathrm{T}}X]^{-1}X^{\mathrm{T}}Y$ — это

$$\begin{bmatrix} 9{,}53 \\ -0{,}32 \\ 0{,}60 \\ -0{,}14 \end{bmatrix}.$$

Таким образом, $B_0 = 9{,}53$, $B_1 = -0{,}32$, $B_2 = 0{,}60$ и $B_3 = -0{,}14$. Эти показатели представляют собой постоянный коэффициент и коэффициенты регрессии V_1, V_2, V_3.

ПРИЛОЖЕНИЕ 6.2

FTSE 100 Доходы	Государственные облигации Доходы	S&P 500 Доходы	£/$ Доходы
–5,03865	0,97062	–0,81181	3,875111
–5,80479	–0,99772	–2,7947	–2,25009
6,756908	2,726713	2,726272	–1,01653
4,715902	–3,24295	0,786902	2,926117
–3,22561	–2,14766	–7,21813	4,249501
–4,68652	–3,13807	1,194216	–1,01554
–0,75337	–2,70719	1,775336	–2,09321
–4,78026	6,621534	–1,92271	0,641713

Продолжение таблицы

11,30555	3,500561	8,895581	2,585833
0,025298	0,276949	–1,00181	4,583169
–1,401	–0,35164	–1,12439	5,058423
–7,65644	0,330367	–9,72923	0,965673
–6,47293	4,124323	–2,3907	0,877082
–0,13797	3,771859	–2,54693	3,228606
6,430736	0,257748	5,413922	–1,67941
–0,89174	4,775062	1,870699	0,571579
1,030363	1,203434	3,811711	2,38076
9,725182	0,994964	7,689626	–4,05269
2,874212	0,189595	0,223789	–8,74905
2,090753	–0,1096	2,392376	–0,83731
0,294574	0,307949	2,022585	–1,57806
–2,91185	2,740759	–2,64774	–4,98216
5,884167	2,163633	2,405219	3,588102
3,33535	2,862377	2,123906	0,654959
–1,27696	0,204208	–1,58804	3,670671
–3,70006	–0,34128	0,543229	0,085776
–5,42394	1,512134	–2,56769	1,53151
3,186904	2,761412	8,945333	5,182791
2,655846	1,525032	–1,72668	–3,86864
–0,23072	–2,98952	0,607979	–2,50329
–5,87325	4,863675	–2,0131	–1,81065
9,920527	2,543311	2,032491	3,477124
1,411156	–0,58666	1,149646	2,106508
–7,85147	–0,24178	–1,06484	4,714317
–2,99976	–1,25275	2,914392	0,677969
–5,1637	3,832018	–2,14474	3,72397
11,25872	5,746209	0,052862	–13,9774
4,392257	–1,33108	1,539886	–12,5624
3,803517	2,737089	1,881653	0,325946
1,933198	1,197246	1,140232	–1,47519
0,179007	2,755257	1,548583	–4,07807
1,081242	1,040079	–0,11532	–0,93313
–0,14581	–1,80704	1,858152	5,767911
–2,29475	0,560659	–1,7564	2,811625
1,275118	4,08267	2,537261	–1,15386
1,380294	2,698462	–1,06552	–2,77841
1,814645	4,402739	0,251343	–1,26753
4,756401	0,078037	2,847012	1,201216
–1,49244	1,258159	–0,40457	–0,13276
4,033624	2,560811	1,681073	–1,64075
2,150889	4,099294	–1,54892	–0,03377

7

АНАЛИЗ ВРЕМЕННЫХ РЯДОВ

ВВЕДЕНИЕ

Анализ временных рядов включает в себя очень широкий спектр проблем. В этой главе мы ограничимся четырьмя целями. Первая — это объяснить доступным языком значение наиболее важных терминов, используемых в анализе временных рядов (динамических процессов). Вторая — проанализировать процесс построения временных рядов как однофакторный стохастический процесс, т.е. стохастический процесс, составляющие которого являются функциями одной рассматриваемой переменной. Третья и четвертая цели — объяснить два эконометрических метода, используемых для анализа временных рядов. Термин "эконометрические методы" здесь показывает, что процесс моделируется как функция, зависящая от нескольких переменных, не только от рассматриваемой. Два метода, которые в последнее время все чаще используются при анализе финансовой информации, — это коинтеграция и авторегрессионная условная гетероскедастичность (ARCH) и ее обобщенная форма — GARCH. Однако перед тем, как приступить к анализу этих концепций, мы должны определиться с некоторыми понятиями и объяснить некоторые основные формы анализа временных рядов.

ОСНОВЫ

Существует несколько терминов, описывающих статистические характеристики временных рядов, и читателю было бы полезно ознакомиться с этими понятиями, поскольку они часто будут встречаться на протяжении этой главы, а также в другой литературе на эту тему. Некоторые специфические термины уже использованы выше.

В частности, читателю надо запомнить определения **случайного блуждания**, **мартингала**, **стационарности** и **белого шума**.

Случайное блуждание и мартингалы

Случайное блуждание — это часто используемая модель финансовых временных рядов. Случайное блуждание определяет путь случайной переменной, где каждое изменение или "инновация" не зависят от всех предыдущих изменений и каждое подчиняется **идентичному** распределению вероятностей.

Независимость означает, что изменение в какой-либо момент времени не имеет никакого влияния на все последующие изменения. Это может проявляться в виде нулевой корреляции между следующими друг за другом парами наблюдений.

Под **идентичностью** мы понимаем то, что каждое из изменений подчиняется одному и тому же распределению вероятностей, с одними и теми же параметрами распределения, такими, как средняя величина и среднее квадратическое отклонение.

Случайное блуждание — это стохастический процесс, где изменения уровня достигаются прибавлением случайной переменной ε, с постоянной дисперсией и средней, равной нулю. Также этот процесс характеризуется нулевой корреляцией отдельных наблюдений. Это можно выразить в виде формулы:

$$Y_t = Y_{t-1} + \varepsilon. \tag{7.1}$$

Заметьте, что только ε имеет среднюю, равную нулю, и постоянную дисперсию. Математическое ожидание Y_t : $E(Y_t) = \mu$; дисперсия — σ_Y^2.

Иногда случайное блуждание может включать элемент сдвига. Сдвиг по сути означает временну́ю тенденцию (тренд). Таким образом, случайное блуждание со сдвигом — это случайное блуждание с трендом. Например

$$Y_t = Y_{t-1} + \alpha + \varepsilon. \tag{7.2}$$

Случайное блуждание имеет два свойства, которые особенно важны в анализе финансовых временных рядов. Это свойства — **свойство Маркова** и **мартингальное свойство**.

Свойство Маркова состоит в том, что вся информация, необходимая для определения условной вероятности будущего (следующего) значения случайной переменной, содержится в текущем состоянии этой переменной, а не в историческом распределении ее вероятностей. Для случайного блуждания это следует из предположения независимости, поскольку каждое из последующих изменений не зависит от предыдущего уровня. Однако будущее значение зависит от текущего уровня.

Мартингальное свойство означает, что условное ожидание будущего значения случайной переменной равно текущему значению переменной. Это применимо к случайному блужданию, поскольку все изменения при нулевом сдвиге имеют математическое ожидание, равное нулю.

По сути мартингал — это более общий стохастический процесс, чем случайное блуждание, потому что в случае мартингала изменения задаются значением случайной переменной, которая хотя и должна обладать нулевым математическим ожиданием, но не обязательно должна иметь постоянную дисперсию. Кроме того, изменения не должны быть независимыми.

Случайное блуждание с положительным сдвигом является примером **субмартингала**. При отрицательном сдвиге это будет примером **супрамартингала**.

Например, рассмотрим процесс, определенный как

$$Y_t = Y_{t-1} + 0{,}8 + \varepsilon.$$

Это пример субмартингала. В общем виде это выглядит как

$$Y_t = Y_{t-1} + \alpha + \varepsilon; \qquad (7.3)$$

- если $\alpha>0$, то мы имеем дело с субмартингалом,
- если $\alpha<0$, то это супрамартингал;
- если же $\alpha = 0$, то это будет случайное блуждание.

Графики примеров случайного блуждания, субмартингалов и супрамартингалов приведены на рис. 7.1—7.3.

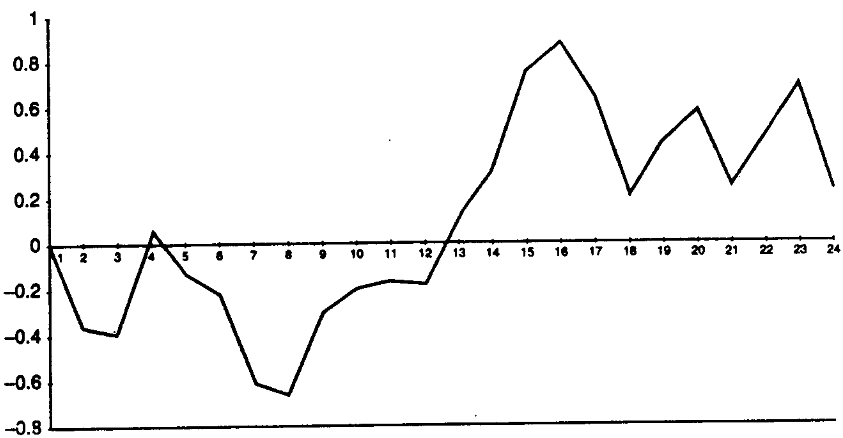

Рис. 7.1. Случайное блуждание

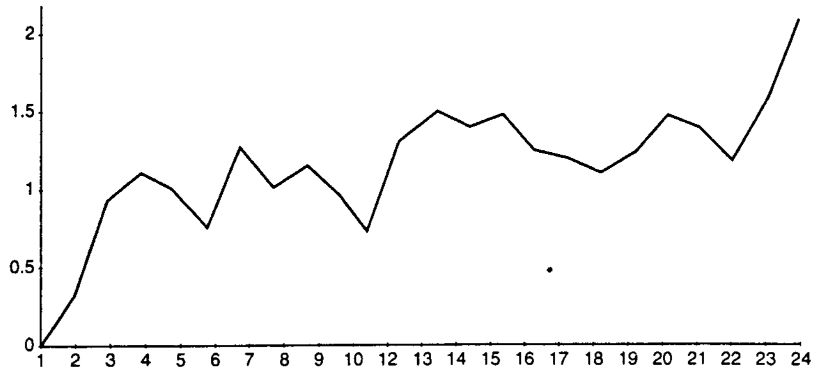

Рис. 7.2. Субмартингал

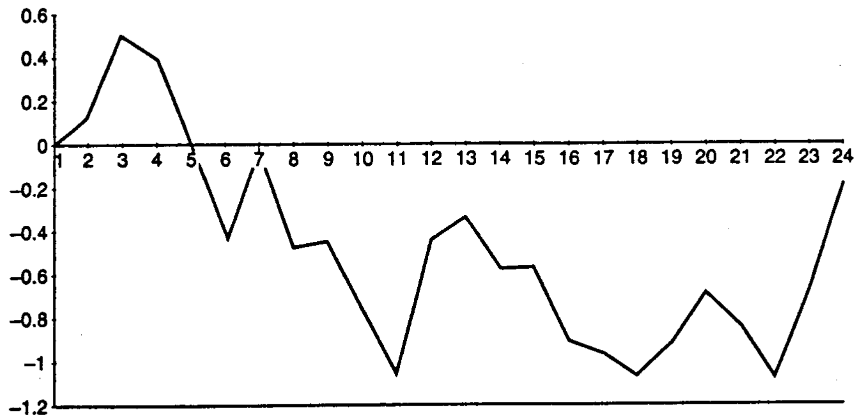

Рис. 7. 3. Супрамартингал

Белый шум

Временные ряды называются "белым шумом", если лежащая в их основе переменная имеет среднюю, равную нулю, постоянную дисперсию и нулевую корреляцию последовательных наблюдений, т.е. нулевую автокорреляцию. Вы помните из гл. 6, что допущения для значения остаточного члена регрессионной модели МНК схожи с этим. Таким образом, остатки рассматриваемые в МНК, можно считать "белым шумом".

Если $E(\varepsilon_t) \sim N(0, \sigma^2)$, то мы имеем дело с гауссовским белым шумом, хотя переменная белого шума необязательно должна подчиняться нормальному распределению.

Пример белого шума изображен на рис. 7.4.

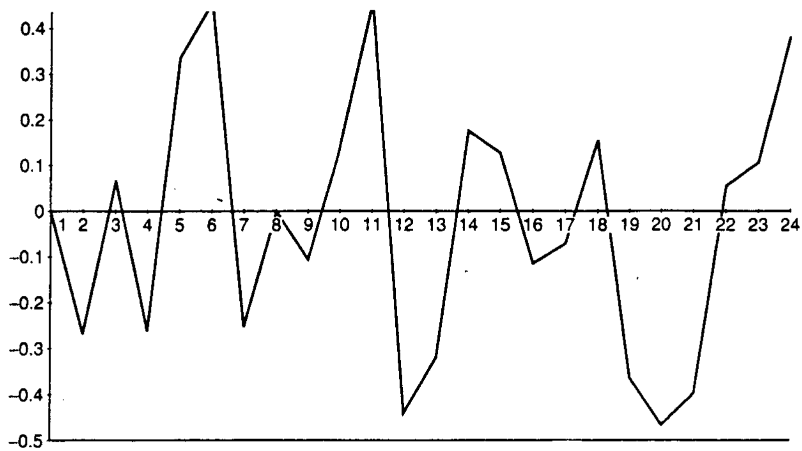

Рис. 7.4. Белый шум

Стационарность

Временные ряды называются стационарными, если они обладают постоянной средней и дисперсией, а ковариация зависит только от временного интервала между двумя отдельными наблюдениями.

Возьмем, например, индекс FTSE 100. Этот индекс возрос с уровня 1000 на момент открытия до 3600 на момент написания этой книги, спустя 12 лет. Представьте, что вы анализируете среднюю и среднее квадратическое отклонение дневного уровня за каждый календарный год. Поскольку индекс возрастает год за годом, то средняя величина и среднее квадратическое отклонение в течение первого года будут ниже, чем в течение второго года и т.д., и очевидно, что в тот год, когда индекс возрос до 3500, затем упал до 3100, средний уровень будет выше, чем в первый год начала исчисления индекса.

В гл. 2 мы отметили, что величина дисперсии и среднего квадратического отклонения может быть функцией от значения индекса. Таким образом, дисперсия индекса, колеблющаяся вокруг 1000, вполне может быть ниже дисперсии индекса, колеблющегося вокруг отметки 3000.

Ковариация также может зависеть от уровня значений анализируемых данных. В этом случае мы имеем ковариацию между последовательными наблюдениями.

По интуиции мы ожидаем, что немногие (если вообще какие-то) временные ряды курсов валют или уровней индексов будут стационарными, поскольку возрастающие и падающие значения являются основной чертой финансовых переменных. Хотя норма прибыли, требуемая инвесторами, обычно зависит от неопределенности, связанной с инвестициями, и не зависит от уровня индекса. Таким образом, норма прибыли может обладать постоянными средней и средним квадратическим отклонением и ковариацией наблюдений, зависящей только от промежутков между наблюдениями.

На рис. 7.5 изображен временной ряд значений FTSE 100 с 1984 по 1992 гг., а на рис. 7.6 показан временной ряд непрерывно наращенной нормы доходности. Очевидно, что ряд значений индекса FTSE 100 не является стационарным, в то время как ряд значений норм доходности вполне может быть. Далее в этой главе мы увидим, как проверять, являются ли временные ряды доходности стационарными в действительности.

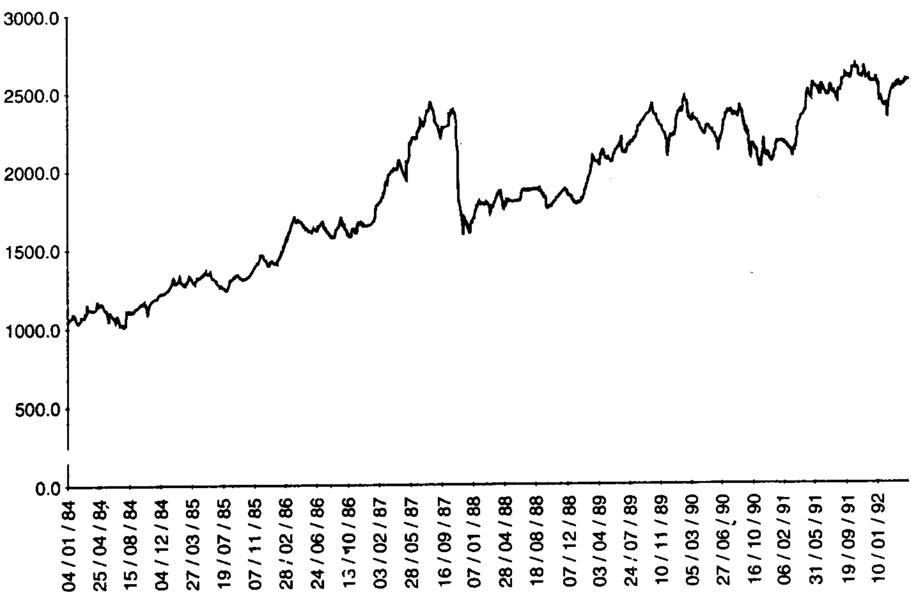

Рис. 7.5. Временной ряд значений индексов FTSE 100 с 1984 по 1992 гг.

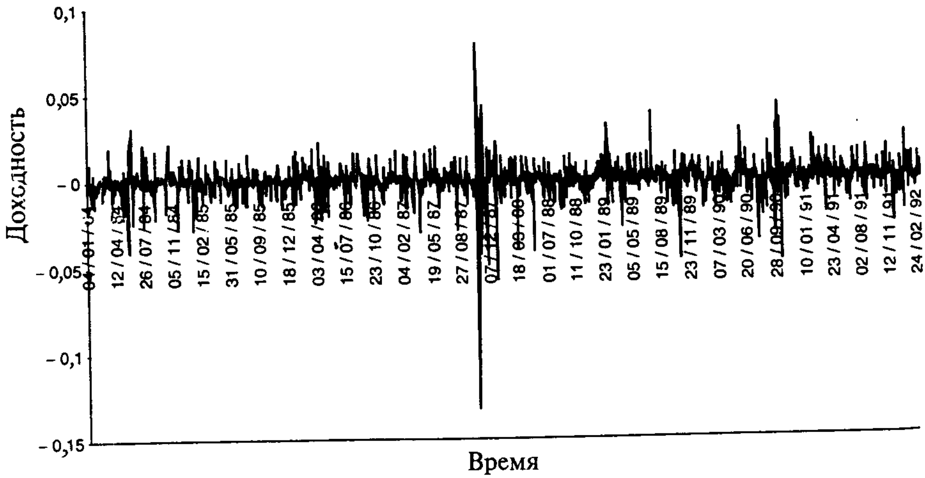

Рис. 7.6. Временной ряд непрерывно наращенной нормы доходности: 1984—1992

Рассматривая приведенное выше определение **белого шума**, мы видим, что он является стационарным рядом. Хотя стационарный ряд необязательно будет белым шумом, поскольку может иметь среднюю, отличную от нуля, или ковариацию.

ОДНОФАКТОРНЫЕ СТОХАСТИЧЕСКИЕ МОДЕЛИ ДИНАМИЧЕСКИХ ПРОЦЕССОВ

В рамках **однофакторных стохастических моделей** процесс построения временных рядов рассматривается как состоящий из компонентов анализируемых временных рядов. Другими словами, если будущие значения рассматриваемой переменной сколько-нибудь предсказуемы, то они являются функцией от прошлых значений этой переменной. В этой ситуации ограничимся анализом составляющих однофакторного стохастического процесса в контексте **авторегрессионных процессов, процессов со скользящей средней** и степени **интегрирования**. Эти три подпроцесса удобно объединить под названием **авторегрессионные интегрированные модели со скользящей средней** (autoregressive integrated moving average processes — ARIMA).

Авторегрессионные процессы

Начнем анализ ARIMA с рассмотрения **авторегрессионного процесса**. Авторегрессионным называется процесс, при котором значение ряда находится в линейной зависимости от предыдущих значений. Например, если текущее наблюдаемое значение является функцией всего лишь одного значения, непосредственно предшествующего наблюдению, т.е. процесс зависит всего лишь от одного значения рассматриваемой переменной, то процесс называется авторегрессионным процессом первого порядка и обозначается AR(1). Это можно обобщить следующим образом: если анализируемый динамический процесс зависит от значений, отстоящих от 1 до n временных лагов назад, то это авторегрессионный процесс порядка n, т.е. AR(n). Например, процесс AR(3) можно отобразить следующим образом:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_3 Y_{t-3}. \qquad (7.4)$$

Здесь текущее значение Y — функция от трех наиболее недавних предыдущих значений. Отсюда авторегрессионная **модель** — это модель, в которой моделируемые значения задаются линейной функцией от предыдущих наблюдений. Читатели здесь увидят сходство с автокорреляцией или внутрирядовой корреляцией, где существовала корреляция между остатками в уравнении регрессии. В действительности, если мы посмотрим на уравнение (7.7), то узнаем многофакторное уравнение регрессии, где прошлые значения Y являются независимыми переменными

$$\hat{Y}_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_3 Y_{t-3} + \varepsilon_t, \qquad (7.5)$$

где ε_t — остаток или ошибка (погрешность).

Далее в этой главе мы объясним, как определять степень автокорреляции временных рядов, используя **коэффициент автокорреляции** и **частный коэффициент автокорреляции**.

Интеграция

Авторегрессионный процесс и процесс скользящей средней, который мы проанализируем в следующем параграфе, предполагают, что анализируемые данные являются стационарными. Интегрирование означает, какого порядка разности должны быть рассчитаны для того, чтобы получить стационарный временной ряд. Здесь нахождение разностей — это всего лишь нахождение изменений значения переменной в последующие периоды, т.е. величины $\Delta Y_t = Y_t - Y_{t-1}$. Ряд значений ΔY — это ряд разностей.

Если во временном ряду должны быть рассчитаны разности первые, чтобы получить стационарный ряд, то первоначальный ряд называется интегрированным рядом первого порядка, или $I(1)$. Если же требуется рассчитать вторые разности для получения стационарного ряда, то это интегрированный ряд второго порядка, или $I(2)$. Если же в ряду вообще не требуется вычислять разности, то он называется интегрированным рядом нулевого порядка, или $I(0)$.

Вспомним приведенное выше обсуждение стационарности. Мы предположили, что уровни значений индексов FTSE 100 вряд ли являются стационарными, в то время как ряд показателей доходности вполне может быть таким.

Если ряд $I(0)$, т.е. стационарен, то его дисперсия будет конечна. Изменения рассматриваемой переменной будут иметь только промежуточное влияние на временной ряд. Коэффициенты автокорреляции будут постепенно убывать таким образом, что их сумма станет конечной. И наоборот, если ряд $I(1)$, то изменения будут иметь постоянный эффект, дисперсия с течением времени будет возрастать до бесконечности .

Вернемся к примеру с индексом FTSE 100. Исходя из наших заключений, можно сказать, что показатели доходности являются первыми разностями ряда значений индекса. Если ряд значений доходности устойчив, то ряд значений индекса будет рядом $I(1)$.

Моделирование при помощи нестационарных рядов может оказаться проблематичным, например, это может привести к **ложной корреляции**. Чтобы понять, что это такое, взгляните на рис. 7.7, на котором изображены значения индексов FTSE 100 и S&P 500. Коэффициент корреляции этих двух рядов равен 0,81. Однако было бы неуместным предполагать причинную связь между значениями индексов. Конечно, оба индекса вместе возрастают и падают на протяжении длительных периодов времени, но является ли рост одного индекса причиной роста другого? В действительности же, в соответствии с экономической теорией, причиной одновременного роста индексов служит долгосрочный экономический рост в обеих странах при условии высокого уровня интегрированности национальных экономик и положительного уровня инфляции. Таким образом, причиной долгосрочного роста индексов является третий фактор — экономическая деятельность.

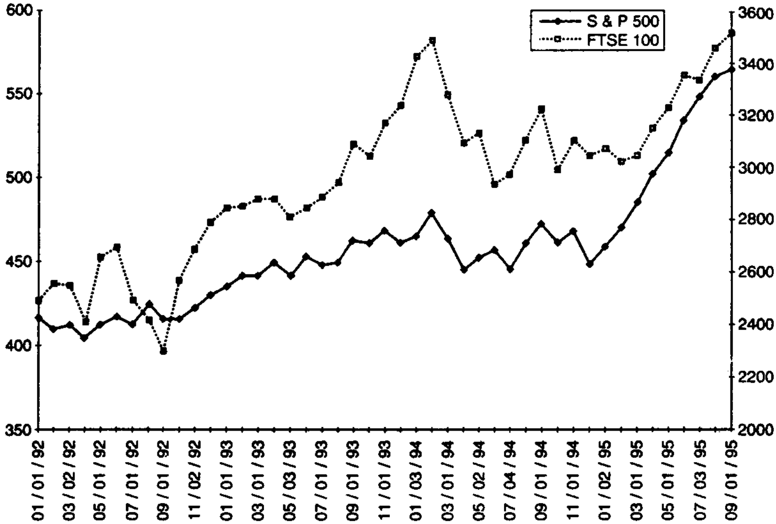

Рис. 7.7. Месячные уровни индексов S&P 500 и FTSE 100

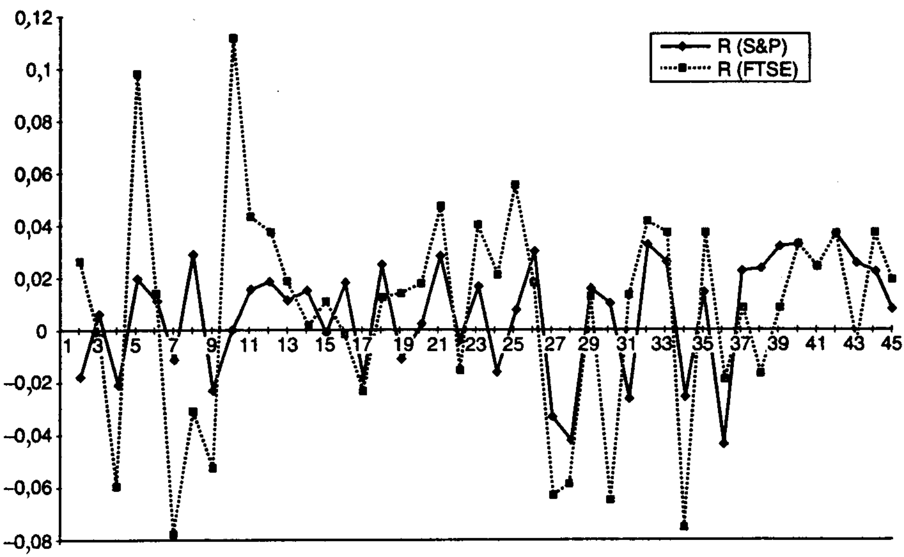

Рис. 7.8. Месячная доходность по S&P 500 и FTSE 100 : с 1.01.92 по 1.09.95

Ложная природа корреляции двух рядов подтверждается при анализе рядов доходности. Если бы корреляция была высокой, то мы могли бы ожидать отрицательную (положительную) рентабельность в одном ряду при отрицательной (положительной) рентабельности в другом. Однако, как это видно на рис. 7.8, это не всегда имеет место. На самом деле корреляция рядов динамики доходностей равна всего лишь 0,3.

Таким образом, построение моделей с переменными $I(1)$ может привести к видимой корреляции, которая может быть принята за причинную связь, в то время как этого на самом деле нет.

Модели скользящей средней

Модель скользящей средней — это модель, где моделируемая величина задается линейной функцией от прошлых ошибок, т.е. разностей между прошлыми смоделированными значениями и прошлыми фактическими наблюдениями.

$$\hat{Y}_t = B_0 + B_1\varepsilon_{t-1} + B_2\varepsilon_{t-2} + B_3\varepsilon_{t-3} + \varepsilon_t, \tag{7.6}$$

где

$$\varepsilon_t = Y_t - \hat{Y}_t. \tag{7.7}$$

Термин "скользящая средняя", используемый в этой главе, не следует путать со схожим термином, относящимся к технике сглаживания данных.

Авторегрессионные модели скользящей средней

Разработаны модели временных рядов, которые сочетают авторегрессионный процесс с моделью скользящей средней. Неудивительно, что эти модели называются **авторегрессиоными моделями скользящей средней** или **ARMA**[1]. Модель ARMA(pq) имеет p временных лагов в авторегрессионном процессе и q интервалов в модели скользящей средней. Например, ARMA (3, 2) выглядит следующим образом:

$$\hat{Y}_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \alpha_3 Y_{t-3} + B_1\varepsilon_{t-1} + B_2\varepsilon_{t-2} + u_t, \tag{7.8}$$

где u_t — остаточный член ошибки в данном уравнении.

[1] Auto-Regressive Moving Average.

Авторегрессионные интегрированные модели скользящей средней (ARIMA)

Если перед применением ARMA необходимо определить разности уровней с целью получения стационарного ряда, то нужно будет знать порядок этих разностей. Таким образом, процесс ARIMA обладает тремя параметрами: p — порядок авторегрессии, d — требуемый порядок предварительно определяемых разностей и q — порядок скользящей средней в модели.

Поскольку ARIMA включает в себя авторегрессионные процессы, модели скользящей средней и интегрирование, то многие динамические процессы можно рассматривать как ARIMA-процессы. Мы уже отметили, что данные могут иметь авторегрессионный компонент (AR). Ряд может обладать определенной степенью интегрирования: $I(1)$, $I(0)$ и даже $I(2)$. В случае $I(1)$ и $I(2)$ нужно единожды или дважды рассчитать разности, чтобы получить стационарный ряд. Наконец, может присутствовать компонент скользящей средней (MA).

Очень важно разбить временной ряд на эти три составляющие для того, чтобы определить структуру моделируемого процесса. Первая стадия — это расчет разностей с целью получения стационарных рядов. Затем можно попытаться смоделировать полученный стационарный ряд с помощью ARMA.

Например, возьмем абсолютно случайный процесс, где Y_t зависит только от среднего уровня ряда и ошибки, т.е.

$$Y_t = \mu + \varepsilon_t. \tag{7.9}$$

В этом процессе нет зависимости от прошлых значений Y, нет разностей Y_t и нет зависимости от прошлых значений ошибки. Мы можем классифицировать этот процесс как ARIMA (0, 0, 0).

Белый шум — это ARIMA(0, 0, 0) с нулевым средним значением.

Процесс ARIMA (1, 0, 0) имеет следующий вид:

$$Y_t = \alpha Y_{t-1} + \varepsilon_t, \tag{7.10}$$

где $-1<\alpha<1$ и ε_t — элемент белого шума.

Процесс зависит от непосредственно предшествующего значения Y, определение разностей уровней не требуется для трансформации этого ряда в стационарный. Это то же самое, что и процесс AR(1).

Если $\alpha = 1$, то процесс не будет стационарным и потребуется вычисление разностей. Это пример процесса ARIMA (0, 1,0), т.е.

$$Y_t = Y_{t-1} + \varepsilon_t, \tag{7.11}$$

потому что для получения стационарного ряда потребуется определение первых разностей Y_t. Это будет определено как процесс случайного блуждания.

Процесс ARIMA (0, 0, 1) имеет следующий вид:

$$Y_t = \varepsilon_t + \theta\varepsilon_{t-1}, \tag{7.12}$$

потому что Y_t зависит только от значений ошибки.

Процесс ARIMA (1, 0, 1) будет:

$$Y_t = \alpha Y_{t-1} + \theta\varepsilon_{t-1} + \varepsilon_t. \tag{7.13}$$

Векторные авторегрессионные процессы и векторные процессы скользящей средней

До сих пор мы рассматривали временные ряды только одной переменной. В финансах, как и во многих других областях, нам может понадобиться рассмотреть соотношения между двумя и больше переменными. Например

$$\begin{aligned} X_t &= \alpha X_{t-1} + \beta Y_{t-1}, \\ Y_t &= \gamma X_{t-1} + \delta Y_{t-1}. \end{aligned} \tag{7.14}$$

Этот процесс можно изобразить в виде матриц с X_t и Y_t, образующими вектор следующим образом:

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix}. \tag{7.15}$$

Таким образом, α, β, γ и δ образуют матрицу 2×2.

Процесс, изображенный выше — это векторный AR(1) процесс или VAR (1) процесс, потому что здесь включается только одна непосредственно предшествующая векторная переменная и нет MA процесса.

Процесс VMA (1) не будет включать предыдущие значения переменной, но только предыдущие значения ошибок:

$$\begin{aligned} X_t &= p\varepsilon_{X_{t-1}} + q\varepsilon_{Y_{t-1}}, \\ Y_t &= r\varepsilon_{X_{t-1}} + s\varepsilon_{Y_{t-1}}. \end{aligned} \tag{7.16}$$

Этот процесс может также быть изображен в матричном виде:

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} p & q \\ r & s \end{bmatrix} \begin{bmatrix} \varepsilon_{X_{t-1}} \\ \varepsilon_{Y_{t-1}} \end{bmatrix}. \tag{7.17}$$

Очевидно, что векторные процессы могут включать как элементы автокорреляции, так и элементы скользящей средней. Примером векторного ARMA, включающего предыдущие значения с лагом в один период и предыдущие значения ошибки с таким же лагом, VARMA (1, 1) будет

$$X_t = \alpha X_{t-1} + \beta Y_{t-1} + p\varepsilon_{X_{t-1}} + q\varepsilon_{Y_{t-1}},$$

$$Y_t = \gamma X_{t-1} + \delta Y_{t-1} + r\varepsilon_{X_{t-1}} + s\varepsilon_{Y_{t-1}}. \tag{7.18}$$

В матричном виде это будет

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} p & q \\ r & s \end{bmatrix} \begin{bmatrix} \varepsilon_{X_{t-1}} \\ \varepsilon_{Y_{t-1}} \end{bmatrix}. \tag{7.19}$$

Мы знаем, что информация во временных рядах может получаться в результате нескольких процессов. Некоторые типы процессов часто встречаются при анализе финансовых временных рядов под общим названием авторегрессионные интегрированные процессы скользящей средней (ARIMA). Они состоят из трех подпроцессов : авторегрессионных процессов (AR), интегрирования (*I*) и процессов скользящей средней (MA). Таким образом, анализ таких динамических процессов может производиться путем разбиения процесса на три упомянутых выше подпроцесса. Очень важно также помнить, что ARMA-процесс предполагает, что временные ряды стационарны с постоянной средней и дисперсией. Таким образом, анализируя ряды, соответствующие ARIMA-процессу, первым делом следует определить степень интегрированности и, если необходимо, рассчитать разностный ряд так, чтобы его среднее значение стало неизменным.

В экономике и финансах многие ряды не обладают устойчивыми средним значением и дисперсией. Поэтому далее в этой главе мы изучим коинтеграцию в рядах динамики и ARCH- и GARCH-методы, разработанные для работы с такими временными рядами.

Коинтеграция была разработана в ответ на растущую потребность в анализе соотношений между группами экономических показателей-переменных, чтобы получать концептуально и эмпирически более значимое измерение этих взаимосвязей в свете нестационарности отдельных временных рядов. Коинтеграция, в частности, адресуется к данным с нестационарными средними значениями.

ARCH и GARCH учитывают неустойчивость дисперсии и были разработаны в связи с потребностью прогнозирования волатильности финансовых временных рядов в преддверии распространения финансовых опционов и вообще растущей неустойчивости финансовых рынков в течение последних двух десятков лет.

ИНСТРУМЕНТЫ АНАЛИЗА ВРЕМЕННЫХ РЯДОВ

Очевидно, что процесс построения временных рядов может принимать различные формы. В нашем обсуждении мы уже ограничили себя тремя элементами и показали, что при анализе временных рядов необходимо обратить внимание на **уровень автокорреляции, интегрированности** и на **компонент скользящей средней**. Далее мы рассмотрим использование **коэффициента автокорреляции** (Auto-correlation coefficient — ACC) и **частного коэффициента автокорреляции** (Partial auto-correlation coefficient — PACC) для идентификации элементов AR и MA в процессе построения временных рядов. Затем мы воспользуемся **расширенным тестом Дики—Фуллера** (angmented Dickey—Fuller) для определения степени интегрирования.

Проверка автокорреляции: коэффициента автокорреляции и функции

Для определения степени автокорреляции временных рядов мы должны определить силу связи между текущими и прошлыми значениями рассматриваемой переменной. Одним из способов измерения этой связи являются коэффициенты автокорреляции (ACC), совокупность которых образует функции автокорреляции (ACF). Коэффициент автокорреляции измеряет связь между текущими и прошлыми наблюдениями временного ряда и рассчитывается следующим образом:

$$\rho_k = \frac{\sum_{t=1}^{n-k}\left(Y_t - \overline{Y}\right)\left(Y_{t+k} - \overline{Y}\right)}{\sum_{t=1}^{n}\left(Y_t - \overline{Y}\right)^2}, \tag{7.20}$$

где k — количество лагов. Таким образом, коэффициент автокорреляции первого порядка будет рассчитан с лагом в один период, коэффициент автокорреляции второго порядка будет учитывать степень связи между значениями, отстоящими на два временных периода, и т.д. Рассчитываются коэффициенты автокорреляции всех порядков и затем проводится статистическая проверка для определения, при каких лагах коэффициенты статистически значимы. Только лаги, являющиеся статистически значимыми, оставляются в модели.

Проверка значимости коэффициентов автокорреляции проводится при помощи критерия стандартной ошибки и Q-критерия Бокса—Пирса. Два критерия предлагаются потому, что существуют два подхода к проверке наличия автокорреляции. При первом подходе подразумевается использование критерия стандартной ошибки, проверяются коэффициенты автокорреляции каждого порядка отдельно, чтобы выявить, какие из них значимы. Второй подход использует Q-критерий Бокса—Пирса для того, чтобы проверить на значимость все множество коэффициентов как группу.

Стандартная ошибка коэффициента корреляции рассчитывается следующим образом:

$$\mathrm{SE}_{r_k} = \frac{1}{\sqrt{n}}. \tag{7.21}$$

Коэффициенты автокорреляции случайных данных обладают выборочным распределением, приближающимся к нормальному с нулевым математическим ожиданием и средним квадратическим отклонением, равным $1/\sqrt{n}$.

Для иллюстрации этого подхода воспользуемся данными об уровнях цен и доходности британских государственных долгосрочных облигаций. Коэффициенты автокорреляции первого порядка рассчитываются на основе выборки из 900 наблюдений. Стандартная ошибка равна $1/\sqrt{900} = 1/30 = 0{,}0333$.

Если r_1 находится в следующем интервале:

$$-1{,}96 \cdot 0{,}0333 \le r_1 \le 1{,}96 \cdot 0{,}0333 = -0{,}065 \le r_1 \le 0{,}065,$$

то можно считать, что данные не показывают наличие автокорреляций первого порядка.

Все рассчитанные коэффициенты автокорреляции для временного ряда цен облигаций были значительно больше 0,065. Это неудивительно при том, что данный временной ряд — это ряд цен облигаций. Однако данные о доходности облигаций показали низкий уровень автокорреляции при том, что только коэффициенты первого, третьего, седьмого и восьмого порядков оказались статистически значимыми. Это показывают данные табл. 7.1.

Таблица 7.1

Лаг	ACC	Лаг	ACC	Лаг	ACC
1	0,095*	2	0,012	3	0,074*
4	–0,009	5	0,022	6	0,031
7	0,080*	8	0,068*	9	0,011

* Значимо при 5%-ном уровне.

Статистический критерий Q рассчитывается так:

$$Q = n\sum_{i=1}^{m} r_i^2 \sim \chi_m^2,$$

где m — максимальный рассматриваемый лаг.

Например, с лагом, равным девяти, получим следующие значения Q-критерия:

$$Q = 900 \cdot 0{,}027 = 24{,}5800,$$
$$\chi_9^2(0{,}005) = 23{,}59.$$

Таким образом, как группа коэффициенты для лагов в девять периодов значимы.

Частный коэффициент и функция автокорреляции

Частный коэффициент автокорреляции (PAC), лежащий в основе частной функции автокорреляции (PAF), измеряет связь между текущим значением переменной X_t и последующими значениями этой переменной X_{t-1}, X_{t-2}, ..., X_{t-k}, когда влияние всех промежуточных временных лагов устранено. Таким образом, частный коэффициент автокорреляции первого порядка будет равен коэффициенту автокорреляции первого порядка, так как нет промежуточных лагов. Но частные коэффициенты второго и следующих порядков будут уже отличаться друг от друга.

Частный коэффициент автокорреляции используется для определения степени автокорреляции внутри временного ряда. Например, ряд, обозначенный AR(m), показывает, что последний статистически значимый частный коэффициент автокорреляции рассчитан с лагом m. Таким образом, в ряде AR(2) текущее значение переменной обладает значимой корреляцией только со значениями, отстоящими на 1 и 2 временных лага назад. В ряде AR(4) значимыми будут частные коэффициенты автокорреляции с лагами от одного до четырех периодов, но коэффициенты с более высокими лагами не будут значимо отличаться от нуля.

В динамическом процессе AR(m) частные коэффициенты автокорреляции значимо отличаются от нуля для временных лагов от 1 до m и затем резко падают до нуля для интервалов $m + 1$ и больше.

Проверка процесса скользящей средней

Зная поведение коэффициента автокорреляции и частного коэффициента автокорреляции, можно попытаться определить, содержит ли ряд элемент скользящей средней. Если ряд скорее MA чем AR, то автокорреляция не будет показывать порядок MA-процесса. Хотя, если значение частных коэффициентов автокорреляции падает по экспоненте, а не опускается резко до нуля, то можно предположить, что ряд содержит процесс скользящей средней, а не AR.

Критерий для ARMA-процессов

Для проверки автокорреляции в рядах, где присутствуют элементы и авторегрессии и скользящей средней, используется критерий Люнга—Бокса (LB) (Ljung—Box, 1978). Критерий LB рассчитывается следующим образом:

$$\text{LB} = n(n+2)\sum_{k=1}^{m}\left(\frac{1}{n-k}\right)r_k^2 \approx \chi^2_{m-p-q}\,, \qquad (7.22)$$

где m — максимальное число временных лагов, рассматриваемых в модели;

p — порядок авторегрессии;

q — порядок процесса скользящей средней.

Проверка степени интеграции и стационарности

Как выше уже говорили, интеграция означает, в какой степени ряд должен быть преобразован с помощью разностей разного порядка, чтобы стать стационарным, что очень важно, так как многие методы анализа временных рядов подразумевают, что анализируемый ряд в действительности является стационарным. Проверка стационарности производится при помощи теста единичного корня (unit root)[1]. Если данные показывают единичный корень, то ряд является $I(1)$.

Ранее подход к проверке стационарности и степени интеграции был назван критерием Дики—Фуллера. С помощью этого критерия проверяется, имеет ли коэффициент α в уравнении (7.23) значение, равное единице или меньше единицы

$$Y_t = \alpha_1 Y_{t-1} + \varepsilon_t. \qquad (7.23)$$

Если α равно единице, то данные имеют единичный корень и степень интегрирования равна 1, т.е. ряд является $I(1)$. Если же α меньше единицы и больше нуля, то ряд стационарен, т.е. $I(0)$. В финансах α обычно бывает не больше 1, поскольку это подразумевает взрывные ряды. Такие ряды маловероятны, поскольку давление экономической среды не позволяет переменной принимать бесконечно большие значения.

Существуют некоторые теоретические проблемы с уравнением (7.23), поскольку возможность нестационарности нарушает допущения регрессии МНК, которая подразумевает постоянную дисперсию остатков. Например, рассмотрим уравнение $Y_t = Y_{t-1} +$

---

[1] Проверка стационарности производится на основе анализа корней характеристического уравнения (единичный корень соответствует границе области стационарности) — *Прим. научн. ред.*

$+ u_t = (Y_{t-1} + u_{t-1}) + u_t = ... = u_t + u_{t-1} + u_{t-2} + ... + u_1$. Поскольку остатки u_t независимы и обладают постоянной дисперсией, то дисперсия Y_t растет до бесконечности по мере того, как t приближается к бесконечности. Тогда требуется уравнение, выражающее изменения Y_t следующим образом:

$$\Delta Y_t = \beta Y_{t-1} + e_t, \tag{7.24}$$

где $\beta = (\alpha_t - 1)$.

Если β равно нулю, то говорят, что ряд Y обладает единичным корнем и является $I(1)$, и ряд ΔY будет стационарным. Если же β меньше нуля, т.е. α меньше единицы, то сам ряд Y является стационарным, $I(0)$.

Уравнение (7.23) предполагает нулевое значение средней и отсутствие тренда. В финансовых временных рядах часто уместно бывает включить положительную среднюю, потому что рисковые активы предполагают положительную норму прибыли. В результате получаем уравнение с положительной средней

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + e_t, \tag{7.25}$$

которое может быть преобразовано так:

$$\Delta Y_t = \alpha_0 + \beta Y_{t-1} + e_t. \tag{7.26}$$

Третья форма уравнения, уместного в финансах, включает тренд следующим образом:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \gamma T + e_t, \tag{7.27}$$

которое можно преобразовать в уравнение:

$$\Delta Y_t = \alpha_0 + \beta Y_{t-1} + \gamma T + e_t. \tag{7.28}$$

Неуместно использовать традиционный t-критерий для проверки значимости β, поскольку, применяя регрессию для оценки β, мы предполагаем, что β меньше нуля ($\alpha<1$). Это видно, когда при $\beta = 0$ слишком большой процент оценок отвергается t-критерием. Таким образом, **нулевая** гипотеза существования единичного корня будет отвергаться слишком часто.

Кроме того, Филлипс (*Phillips*) (1987) показал, что такие единичные корни робастны при разной степени гетероскедастичности, но автокорреляция может создавать проблемы. Проблемы при проверке **стационарности**, когда существует автокорреляция остатков, решаются применением расширенного критерия Дики—Фуллера. При использовании этого метода прошлые значения независимой переменной включаются в уравнение регрессии с лагом, достаточным для того, чтобы избавиться от автокорреляции остатков. Это уравнение может иметь вид:

$$\Delta Y_t = \alpha_0 + \beta_1 Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \gamma_2 \Delta Y_{t-2} + \ldots + \gamma_n \Delta Y_{t-n} + e_t. \qquad (7.29)$$

Точная форма критериев значимости зависит от вида тестируемой модели, т.е. без положительной средней (уравнение (7.24)), со средней (уравнение (7.25)) и со средней и трендом (уравнение (7.28)).

Нулевая гипотеза без средней

При тестировании уравнения (7.24), которое подразумевает отсутствие средней, но учитывает автокорреляцию, как в уравнении (7.29), нулевая гипотеза будет записана так:

H_0: $\beta_1 = 0$.

Если β значительно меньше нуля, то Y стационарен, т.е. $I(0)$.

Нулевая гипотеза будет отвергнута, если статистический критерий $\beta_1/SE(\beta_1)$[1] имеет отрицательное значение, меньшее чем критическое значение из таблиц Дики и Фуллера (1979). Критические значения для уровня значимости 1% и 5% соответственно равны —2,58 и —1,95.

Если нулевая гипотеза принята, то ряд Y — это случайное блуждание без сдвига.

В общем виде этого критерия учитывается размер выборки T. Это достигается путем вычисления модифицированного критического значения по формуле

$$\phi_\infty + \phi_1/T + \phi_2/T^2,$$

где ϕ_∞ равен —2,57 (1%) или —1,94 (5%), ϕ_1 = —1,96 (1%) или —0,398(5%), ϕ_2 = —10,04 (1%) или 0(5%). (Значения ϕ табулированы и имеются у МакКиннона (*McKinnon*) (1991)).

Нулевая гипотеза со средней

Проверка уравнения (7.26), которое включает среднюю, подразумевает использование того же статистического критерия

[1] $SE(\beta_1)$ — стандартная ошибка β_1 (*Прим. научн. ред.*).

$\beta_1/SE(\beta_1)$ и той же самой формулы критических значений. Однако значения ϕ теперь равны:

$\phi_\infty = -3.43$ (1%) или $-2,86$ (5%);
$\phi_1 = -6,00$ (1%) или $-2,74$ (5%);
$\phi_2 = -29,25$ (1%) или $-8,36$ (5%).

Нулевая гипотеза со средней и трендом

Проверка уравнения (7.28), включающего среднюю и тренд, подразумевает использование того же процесса, но при следующих значениях ϕ:

$\phi_\infty = -3,96$ (1%) или $-3,41$ (5%);
$\phi_1 = -8,35$ (1%) или $-4,04$ (5%);
$\phi_2 = -47,44$ (1%) или $-17,83$ (5%).

Пример стационарности доходности обменных курсов валют

В качестве примера мы применим расширенный критерий Дики—Фуллера как к ежедневным уровням, так и к ежедневной рентабельности обменного курса доллара США — фунта стерлингов за период с 1992—1995 гг. Ниже приведены значения статистических критериев для регрессии уровней обменного курса с средней (уравнение (7.26)) и со средней и трендом (уравнение (7.28)), каждое значение для одного, двух и трех временных лагов.

Лаг	Критерий (средняя)	Критическое значение	Критерий (средняя и тренд)	Критическое значение
1	−2,1307	−2,86	−1,8430	−3,41
2	−2,0508	−2,86	−1,8443	−3,41
3	−2,7975	−2,86	−2,0365	−3,41

Критические значения относятся к 95%-ному уровню доверия.

Как мы уже знаем, нулевая гипотеза о том, что обменный курс является рядом $I(1)$, отвергается в пользу альтернативной гипотезы, что он представляет собой ряд $I(0)$, если статистический критерий больше отрицательной критической величины. Во всех приведенных выше случаях это не имеет места, таким образом, нулевая гипотеза принимается, т.е. обменные курсы представляют собой ряд $I(1)$.

Теперь посмотрим на доходность:

Лаг	Критерий (средняя)	Критическое значение	Критерий (средняя и тренд)	Критическое значение
1	−20,2910	−2,86	−20,3090	−3,41
2	−16,5143	−2,86	−16,5463	−3,41
3	−13,8929	−2,86	−13,9255	−3,41

Здесь мы ясно видим, что в отношении доходности обменного курса нулевая гипотеза о том, что ряд является рядом $I(1)$, отвергается в пользу альтернативной гипотезы, т.е. в пользу ряда $I(0)$. Поэтому данные доходности относятся к стационарным рядам.

КОИНТЕГРАЦИЯ

Интуитивное введение

Выше, в гл. 2, мы обсуждали корреляцию как меру линейной зависимости между парами переменных. Теперь, когда мы уже ввели понятие стационарности, ясно, что для того, чтобы коэффициент корреляции являлся статистически значимым показателем связи между двумя временными рядами, необходимо выполнение условия их стационарности. Мы говорим, что два временных ряда должны быть **совместно ковариационно** стационарными. Отдельная переменная является ковариационно стационарными, если и $E(X_t)$ и σ^2 (X_t) — конечные константы для всех значений t, коэффициент корреляции между X_t и X_{t-n} является тем же самым для всех t, и, таким образом, ковариация двух наблюдений X зависит только от времени между наблюдениями. Чтобы две переменные были совместно ковариационно стационарными, индивидуальные ряды должны быть ковариационно устойчивыми, а ковариация X_t и Y_t должна быть неизменной при всех значениях t, т.е. чтобы $\text{cov}(X_t, Y_t)$ не зависела от t.

Проблема использования коэффициента корреляции в финансах заключается в том, что нет особых причин считать финансовые временные ряды ковариационно стационарными. Например, валюты или фондовые биржевые индексы в странах со слабыми экономическими связями вряд ли будут иметь устойчивую взаимосвязь друг с другом. Хотя здесь может присутствовать долгосрочная взаимосвязь, которую требуется определить. Следовательно, нужна иная мера взаимосвязи между переменными, которая может использоваться в свете практических реалий того, что ряды, не будучи совместно ковариационно стационарными на коротком промежутке времени, демонстрируют долгосрочное равновесие. Это понятие соответствует коинтеграции.

Коинтеграция описывает долгосрочную линейную связь между несколькими переменными, которые демонстрируют равновесное отношение друг с другом. Рассмотрим пример с двумя переменными, например, уровень индекса FTSE 100 и курс фьючерсов FTSE 100, которые обозначим соответственно как X и Y. Есть экономические причины полагать, что в долгосрочном плане они будут иметь равновесную связь друг с другом. Чтобы понять это, рассмотрим модель арбитража наличного и фондового рынков, введенную в гл. 1, для объяснения ценообразования финансовых фьючерсов. Что произойдет, если цена фьючерса будет намного выше или ниже теоретического уровня? Если цена фьючерса выше справедливой, то арбитражеры будут продавать фьючерсы и покупать индексы, тем самым опуская цену на фьючерсы до равновесного уровня. И наоборот, если цена фьючерса ниже справедливой, то арбитражеры будут продавать индексы и покупать фьючерсы.

Предположим, что равновесное отношение уровня цены на фьючерс к уровню индекса равно 1,1. Это можно выразить так:

$$Y = 1{,}1X, \tag{7.30}$$

но можно представить и так:

$$Y - 1{,}1\,X = 0. \tag{7.31}$$

Это соотношение верно только для равновесия, на краткосрочных интервалах каждая из этих переменных будет изменяться по-своему и, возможно, будет являться рядом $I(1)$. Но даже если оба ряда относятся к $I(1)$, если существует долгосрочное равновесие между ними, то третья переменная z, заданная как

$$Y - 1{,}1X = z, \tag{7.32}$$

будет стационарной и будет измерять, в какой степени переменные X и Y выведены из равновесия. Переменную z называют ошибкой равновесия, потому что под действием тех сил, что устанавливают равновесие, она устремится к своему среднему значению.

Таким образом, если существует равновесное отношение, то возможно найти такую комбинацию данных двух переменных, т.е. $aY + bX = z$, при которой z будет стационарной. Как выше уже говорили, если z в самом деле стационарна, то ее значение колеблется вокруг постоянной средней, и если z отклоняется от своего среднего значения, то имеет тенденцию возвращаться к нему. Однако на краткосрочных интервалах переменные могут и не изменяться вместе и, таким образом, может не быть краткосрочного равновесия, но все краткосрочные отклонения будут сводиться рыночными силами в долгосрочном плане к нулю. Это показано на рис. 7.9.

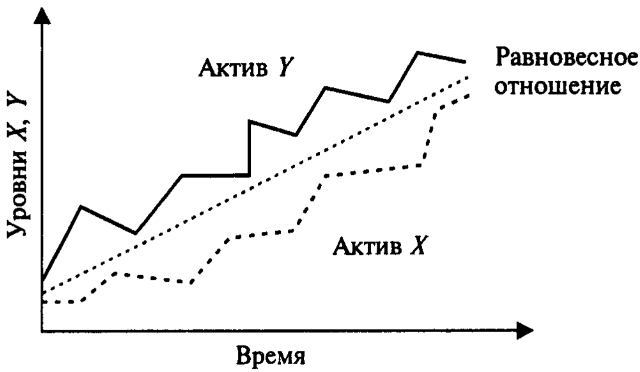

Рис. 7.9

Если существуют такие a и b, что $aY + bX$ является $I(0)$, то $\gamma + (b/\alpha)X$ будет стационарно. Если мы имеем функцию $Y + \beta X = z$, где z стационарна и X и Y являются $I(1)$, то имеется коинтеграция между X и Y. β часто называют коэффициентом коинтеграции.

Таким образом, коинтеграция описывает долгосрочное соотношение двух или более переменных и проистекает из того, что эти переменные демонстрируют общий стохастический тренд во времени.

Одна из задач анализа коинтеграции состоит в анализе преимуществ от диверсификации портфеля в дополнение к корреляционному анализу структуры портфеля с точки зрения его математического ожидания и дисперсии. В частности, анализ коинтеграции позволяет выявить существование долгосрочной зависимости между переменными и скорость, с которой краткосрочные отклонения от долгосрочного равновесия сводятся назад к равновесию (Clare и др., 1995).

Вторая задача анализа коинтеграции была связана с выявлением структурного несовершенства рынков. Смотрите, например, работы Chelley—Steeley и Pentacost (1994) и Cloudhry (1994). Третья функция состояла в обеспечении возможности прогнозирования путем применения так называемой "модели исправления ошибок" (Alexander, 1992).

Коинтеграция между двумя переменными

Коинтеграция двух переменных имеет место, когда порядок интеграции каждого ряда равен b, но некая линейная комбинация которых дает ряд с порядком интегрирации, равным a, где $a < b$. В таком случае говорят, что два ряда $I(b)$ коинтегрированы. Для практических целей в финансах принимают $b = 1$ и $a = 0$. Таким образом, коинтегрированные ряды будут $I(1)$ и их линейная комбинация будет рядом $I(0)$.

Для примера рассмотрим обменные курсы $\pounds/DM$, и $\pounds/FFR$. Предположим, что каждый из курсов является коинтегрированным рядом первого порядка $I(1)$, но если из двух рядов мы можем вывести переменную z, $z = Y_t - \lambda X_t$, которая является рядом $I(0)$, т.е. колеблется вокруг постоянной средней, то говорят, что X и Y коинтегрированы. Параметр λ называется константой коинтеграции.

Коинтеграция характеризует равновесное отношение двух переменных, так как чтобы z было устойчиво, при X и Y, отклоняющихся от их равновесного отношения, они должны будут вернуться к нему, так что z колеблется вокруг определенного (постоянного) среднего значения. Эта тенденция возвращения к равновесию известна как **исправление ошибки**. Модель этого процесса соответственно называется **моделью исправления ошибки** и соответствует интересному свойству стационарности — возвращению стационарных рядов к средней величине.

Чтобы понять идентичность стационарности и возвращения к средней, рассмотрим уравнение (7.33), используемое для проверки на стационарность, пренебрегая эффектами автокорреляции

$$Y_t = \alpha Y_{t-1} + \varepsilon_t. \tag{7.33}$$

Предположим, что постоянное среднее значение равно 0, $\alpha = 0{,}8$ и произошло внезапное отклонение от этой средней, так что значение Y_{t-1} было равно 2. Если пренебречь случайными эффектами, т.е. величинами ε, то Y_t и Y_{t+1} будут равны:

$$Y_t = 0{,}8(2) = 1{,}6;$$

$$Y_{t+1} = 0{,}8(1{,}6) = 1{,}28.$$

С другой стороны, если бы значение Y_{t-1} было равно —3, то Y_t и Y_{t+1} были бы равны соответственно:

$$Y_t = 0{,}8(-3) = -2{,}4;$$

$$Y_{t+1} = 0{,}8(-2{,}4) = -1{,}92.$$

Таким образом, мы видим, что после положительного отклонения от постоянной средней величины следуют отрицательные изменения, и величина этих изменений является функцией размера положительного отклонения. Аналогично отрицательное отклонение вызывает положительные изменения. Таким образом, независимо от изменения и его знака последующие изменения будут приводить значение переменной к ее средней величине.

Можно обобщить это, сказав, что при $\alpha < 1$ ряд будет иметь тенденцию возвращаться к среднему значению. Теперь в случае коинтегрированной регрессии в рядах динамики, ошибки равновесия z возвращаются к их среднему значению и две интегрированные переменные возвращаются к равновесию.

Критерии коинтеграции двух переменных

Первая ступень анализа коинтеграции — это оценка **константы коинтеграции** (при условии, что она существует). В случае с двумя переменными может быть использован МНК. Кроме того, существует другой, более сложный метод, использующий методы оценки **наибольшего правдоподобия** для определения вектора коинтеграции временных рядов. Последний метод сложнее, преимущество его в том, что он имеет общий вид для многофакторных моделей. Здесь мы будем применять МНК. Далее проследим разработку модели исправления ошибок. Затем последует метод, разработанный Йохансеном (*Johansen*, 1988) и Йохансеном и Йезулиусом (*Johansen and Jesulius*, 1990), который использует метод наибольшего правдоподобия.

Признаки коинтеграции двух переменных:

1) их индивидуальные временные ряды должны быть $I(1)$;
2) их линейная комбинация должна быть $I(0)$.

Обратите внимание на то, что не все переменные с рядами $I(1)$ обладают коинтеграцией, а только те, линейная комбинация которых является рядом $I(0)$. Таким образом, проверка коинтеграции рядов динамики происходит в два этапа.

Первый этап

Определить, что переменные являются рядом $I(1)$, можно при помощи расширенного критерия Дики—Фуллера. Исследователь должен произвольно ввести достаточно предшествующих значений X и Y, чтобы превратить остатки в следующих регрессиях в белый шум

$$\Delta X_t = \beta X_{t-1} + \Sigma\theta_i\Delta X_{t-i} + \varepsilon_t;$$
$$\Delta Y_t = \gamma Y_{t-1} + \Sigma\theta_i\Delta Y_{t-i} + \varepsilon_t, \qquad (7.34)$$

где и β и γ аналогичны $(\alpha-1)$ в уравнении (7.3).

Поскольку мы применяем **расширенный критерий Дики—Фуллера**, то надо проверить на значимость β и γ с помощью t-критерия для Y_{t+1} и X_{t-1} соответственно. Если любой из этих параметров не будет значимо отличаться от нуля, то соответствующий ряд (X или Y) будет $I(1)$. Исследователь может ввести в уравнение (7.34) среднее значение или тренд, как это уже обсуждалось выше при рассмотрении интеграции.

Второй этап

Выяснив, что данные ряды являются рядами $I(1)$, мы принимаем МНК в виде, известном как **регрессия коинтеграции**

$$Y_t = \lambda_0 + \lambda X_t + u_t. \qquad (7.35)$$

Эта форма выделяет остатки u_t, так что можно проверить, являются ли они стационарными. Если мы сложим λ_0 и u_t, чтобы получить z_t, то получим $Y_t - \lambda X_t = z_t$ и нужно будет проверить стационарность z. Если z в самом деле стационарна, то λ будет вектором коинтеграции, что уже обсуждалось ранее.

Примечание: обычно регрессия X по Y отличается от регрессии Y по X, но в случае коинтеграции временных рядов "долгосрочная" коинтеграция между переменными равна 1, и обе регрессии по сути одинаковы.

Третий этап

Определим, являются ли остатки рядом $I(0)$. Это достигается использованием следующей регрессии

$$\Delta u_t = \beta u_{t-1} + e_1. \tag{7.36}$$

Это выражение аналогично уравнению (7.24), но скорее с u, чем с Y в качестве независимой переменной. Проверяется нулевая гипотеза, что не существует коинтеграции во временных рядах, т.е. $\beta = 0$. Причина в том, что при β, незначимо отличном от нуля, u_t является рядом $I(1)$, и отсюда нет коинтеграции между рядами Y и X. Мы не можем использовать тот же критерий значимости для U_t, что и для Y_{t-1} в уравнении (7.23), потому что сами остатки уже являются результатом оценки. Поэтому мы должны использовать таблицы МакКиннона (1991).

Этот критерий имеет ряд недостатков. Во-первых, он так же страдает от смещения малых выборок, как и регресии. Во-вторых, он скорее основан на минимизации дисперсии остатков, чем на максимизации стационарности. Однако смещение малой выборки не должно быть проблемой в случае с финансовыми временными рядами, так как размер выборок там обычно достаточно велик.

Модель исправления ошибки

Грейнджер (*Granger*, 1986) и Ингл и Грейнджер (*Engle and Granger*, 1987) показали, что если переменные коинтегрированы, то в них включается **модель исправления ошибки**. Эта модель описывает процесс, в ходе которого переменные $I(1)$ в случае отклонения возвращаются к равновесию. Например, рассмотрим отношения между курсами двух валют. В течение короткого промежутка времени предпочтения инвесторов могут заставить одну валюту возрасти относительно другой. В другой момент времени другая валюта может быть более привлекательной. В обоих случаях валютные курсы отходят от равновесного положения. Однако рыночные силы заставят их вернуться к долгосрочному равновесию. В нашем примере если валюта становится слишком сильной, то политическое и экономическое давление может привести к понижению уровня процентных ставок в соответствующей стране. Если же валюта чрезмерно ослаблена, то процентные ставки в соответствующей стране могут возрасти или экспорт из этой страны возрастет, а импорт сократится до того уровня, пока долгосрочное равновесие снова установится. Модель процесса возврата называется **моделью исправления ошибки**.

Теперь должно быть ясно, почему коинтеграция в рядах динамики подразумевает модель исправления ошибок. Переменная z должна быть стационарной, если переменные $I(1)$ коинтегрированы. Для того чтобы быть стационарной, z должна колебаться вокруг постоянного среднего значения с постоянной дисперсией. Это подразумевает, что в случае отклонений Y и X от равновесного соотношения должны существовать силы (процесс исправления ошибки), приводящие Y и X к равновесию. Отсюда мы видим, что модели исправления ошибки моделируют коинтегрированный процесс и по сути включены в формальный результат, известный как **представительная теорема Грейнджера** (Grander).

Двухстадийный процесс Ингла и Грейнджера

Регрессия полезна, когда анализируются только два ряда, потому что в этом случае может быть не более одного коэффициента коинтеграции. Ингл и Грейнджер (1987) разработали двухстадийный процесс оценки модели исправления ошибки.

Первая стадия — оценка регрессии коинтеграции временных рядов, как описано выше.

Вторая стадия — построение следующей модели исправления ошибки

$$\Delta X_t = \alpha_1 Z_{t-1} + \sum_{i=1}^{n} B_{1i} \Delta Y_{t-i} + \sum_{i=1}^{n} C_{1i} \Delta X_{t-i} + \varepsilon_t;$$

$$\Delta Y_t = \alpha_2 Z_{t-1} + \sum_{i=1}^{n} B_{2i} \Delta Y_{t-i} + \sum_{i=1}^{n} C_{2i} \Delta X_{t-i} + \varepsilon_t, \qquad (7.37)$$

где α_1 и $\alpha_2 \neq 0$.

Таким образом, мы можем использовать МНК для оценки текущих ΔY_t и ΔX_t от прошлых наблюдений Y и X, а также значения коинтегрированной переменной.

В случае с многими переменными может быть больше одного вектора коинтеграции. Следовательно, нужна методология, которая бы определила структуру всех векторов коинтеграции. Такой процесс был разработан Йохансеном (1988) и Йохансеном и Йезулиусом (1990). Он определяет множество временных рядов в качестве векторного авторегрессионного (VAR) процесса. Модель исправления ошибки разрабатывается следующим образом.

Векторное авторегрессионное определение модели исправления ошибки

Для рассмотрения этой разновидности модели исправления ошибки возьмем в качестве примера двухфакторный векторный процесс (AR3) (т.е. такой, в котором значения переменных представляют собой линейную комбинацию последних трех наблюдений). Это может быть записано таким образом

$$X_t = a_1X_{t-1} + a_2X_{t-2} + a_3X_{t-3} + b_1Y_{t-1} + b_2Y_{t-2} + b_3Y_{t-3}, \qquad (7.38)$$

$$Y_t = d_1X_{t-1} + d_2X_{t-2} + d_3X_{t-3} + e_1Y_{t-1} + e_2Y_{t-2} + e_3Y_{t-3}.$$

Это достаточно сложное уравнение может быть выражено в матричной форме

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = A_1\begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + A_2\begin{bmatrix} X_{t-2} \\ Y_{t-2} \end{bmatrix} + A_3\begin{bmatrix} X_{t-3} \\ Y_{t-3} \end{bmatrix}, \qquad (7.39)$$

где A и B — это матрицы $2 \cdot 2$. Например, A_1 — это матрица следующего вида:

$$A_1 = \begin{bmatrix} a_1 & b_1 \\ d_1 & e_1 \end{bmatrix}. \qquad (7.40)$$

Для анализа мы должны преобразовать

$$Y_t = \alpha Y_{t-1}, \qquad (7.41)$$

в первые разности, т.е.

$$\Delta Y_t = (1-\alpha)Y_{t-1}. \qquad (7.42)$$

Далее мы должны преобразовать модель через разности. Это можно сделать следующим образом.

Во-первых, мы вычитаем

$$\begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} \tag{7.43}$$

из обеих частей уравнения, преобразуя левую сторону уравнения в разности, учитывая, что

$$A_1 \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} - \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} = [A_1 - I] \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix}. \tag{7.44}$$

При этом мы получаем:

$$\begin{bmatrix} \Delta X_t \\ \Delta Y_t \end{bmatrix} = [A_1 - I] \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + A_2 \begin{bmatrix} X_{t-2} \\ Y_{t-2} \end{bmatrix} + A_3 \begin{bmatrix} X_{t-3} \\ Y_{t-3} \end{bmatrix}. \tag{7.45}$$

Учитывая, что A_i — это матрицы $2 \cdot 2$, мы можем преобразовать это выражение следующим образом:

$$\begin{bmatrix} \Delta X_t \\ \Delta Y_t \end{bmatrix} = [A_1 - I] \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \left[(A_1 + A_2 - I) - (A_1 - I)\right] \begin{bmatrix} X_{t-2} \\ Y_{t-2} \end{bmatrix} + A_3 \begin{bmatrix} X_{t-3} \\ Y_{t-3} \end{bmatrix}. \tag{7.46}$$

Открывая скобки, получаем:

$$\begin{bmatrix} \Delta X_t \\ \Delta Y_t \end{bmatrix} = [A_1 - I] \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + [A_1 + A_2 - I] \begin{bmatrix} X_{t-2} \\ Y_{t-2} \end{bmatrix} - [A_1 - I] \begin{bmatrix} X_{t-2} \\ Y_{t-2} \end{bmatrix} + A_3 \begin{bmatrix} X_{t-3} \\ Y_{t-3} \end{bmatrix}, \tag{7.47}$$

и группируем члены с общими множителями, например

$$[A_1 - I] \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} - [A_1 - I] \begin{bmatrix} X_{t-2} \\ Y_{t-2} \end{bmatrix}. \tag{7.48}$$

В результате получим выражение

$$[A_1 - I] \left[\begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} - \begin{bmatrix} X_{t-2} \\ Y_{t-2} \end{bmatrix} \right] = [A_1 - I] \begin{bmatrix} \Delta X_{t-1} \\ \Delta Y_{t-1} \end{bmatrix}. \tag{7.49}$$

Схожая процедура применяется по отношению к матрицам A_3. В результате получаем

$$\begin{bmatrix} \Delta X_t \\ \Delta Y_t \end{bmatrix} = [A_1 - I] \begin{bmatrix} \Delta X_{t-1} \\ \Delta Y_{t-1} \end{bmatrix} + [A_1 + A_2 - I] \begin{bmatrix} \Delta X_{t-2} \\ \Delta Y_{t-2} \end{bmatrix} + [A_1 + A_2 + A_3 - I] \begin{bmatrix} X_{t-3} \\ Y_{t-3} \end{bmatrix}. \tag{7.50}$$

Как мы уже говорили выше, полученная матрица $[A_1 + A_2 + A_3 - I]$ — это матрица 2 · 2, обозначим ее П. Матрицу $[A_1 - I]$ обозначим A_1^* и матрицу $[A_1 + A_2 - I] - A_2^*$. В результате имеем

$$\Delta\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = A_1^*\Delta\begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + A_2^*\Delta\begin{bmatrix} X_{t-2} \\ Y_{t-2} \end{bmatrix} + \Pi\begin{bmatrix} X_{t-3} \\ Y_{t-3} \end{bmatrix}. \tag{7.51}$$

Мы видим, что VAR — процесс уровней рядов — может быть записан как VAR — процесс разностей за исключением одного члена

$$\Pi\begin{bmatrix} X_{t-3} \\ Y_{t-3} \end{bmatrix}. \tag{7.52}$$

Ранг матрицы П дает число векторов коинтеграции в рядах динамики (ранг матрицы — это число линейно независимых рядов). Таким образом, ранг матрицы говорит о том, что следует делать. Если П имеет нулевой ранг, то матрица П — нулевая, и мы по сути имеем VAR—процесс в ряде разностей. Это показывает, что коинтеграции в рядах данных нет, и для достижения стационарности требуется нахождение разностей.

Если матрица П — полная, то ряды уже стационарны (матрица П имеет обратную, и, таким образом, выражение может быть решено для уровней, выраженных в разностях. Это будет верным только, если ряды уровней $I(0)$).

Если ранг П лежит между 0 и n ($0 < m < n$, в нашем случае $n = 2$, $m = 1$), то существует m векторов коинтеграции. Эти векторы описывают долгосрочные равновесные соотношения переменных. Модель исправления ошибки **включает в себя** краткосрочные изменения, которые поддерживают долгосрочное равновесие.

Чтобы лучше понять это, разложим матрицу П на матрицу α и γ:

$$\Pi = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}\begin{bmatrix} \gamma_1 & \gamma_2 \end{bmatrix}. \tag{7.53}$$

Так что

$$\Pi\begin{bmatrix} X_{t-3} \\ Y_{t-3} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}\begin{bmatrix} \gamma_1 & \gamma_2 \end{bmatrix}\begin{bmatrix} X_{t-3} \\ Y_{t-3} \end{bmatrix}. \tag{7.54}$$

Если

$$\begin{bmatrix} \gamma_1 & \gamma_2 \end{bmatrix}\begin{bmatrix} X_{t-3} \\ Y_{t-3} \end{bmatrix} = Z, \tag{7.55}$$

стационарно, то существует коинтеграция. α_1 и α_2 интерпретируются как скорость приведения процесса к равновесию.

Таким образом, модель исправления ошибки будет

$$\Delta\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = A_1^*\Delta\begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + A_2^*\Delta\begin{bmatrix} X_{t-2} \\ Y_{t-2} \end{bmatrix} + \begin{bmatrix} \alpha_1 Z \\ \alpha_2 Z \end{bmatrix}, \qquad (7.56)$$

где Z является $I(0)$.

Наиболее точное образование матрицы П получается при методе наибольшего правдоподобия Йохансена (1988) и Йохансена и Йезулеуса (1990), который применяется к коинтеграции нескольких временных рядов.

Коинтеграция нескольких переменных

Теперь мы можем применить анализ коинтеграции к нескольким переменным, например X, Y и W. Существуют четыре возможные линейные комбинации этих переменных, например X и Y, X и W, Y и W, X, Y и W. Однако мы заинтересованы только в независимых комбинациях, так как только они могут быть коинтегрированы. Любая комбинация векторов коинтеграции сама по себе будет вектором коинтеграции. Таким образом, мы можем иметь не более n—1 векторов коинтеграции. Поскольку у нас три переменные, то мы имеем две независимые комбинации.

Теперь рассмотрим упомянутые выше четыре комбинации.

Мы можем доказать, что если $X + Y$ коинтегрированы и Y и W коинтегрированы, то должна существовать коинтеграция в рядах X, Y и W и в рядах X и W.

Так как X и Y коинтегрированы, то существуют такие a, b и c, что $aX + bY + c$ является $I(0)$.

Поскольку коинтегрированы Y и W, то существуют такие p, q и r, что $pY + qW + r$ является $I(0)$.

Сложение дает $aX + (p + b)Y + qW + (c + r)$, что является $I(0)$, отсюда существует коинтеграция X, Y и W.

Умножение на p и b соответственно и вычитание дают ряд $apX - bqW + (cp-br)$, который также является $I(0)$, а значит между X и W существует коинтеграция.

Таким образом, существует не больше двух независимых векторов коинтеграции.

Для определения векторов коинтеграции в случае многих переменных и для построения модели исправления ошибок воспользуемся методом наибольшего правдоподобия Йохансена. Модель исправления ошибок для многих переменных — это всего лишь общий вид модели для двух переменных. Опять мы начнем с построения модели VAR и приведем ее к разностям. Однако на этот раз векторы будут $n \times 1$, а не 2×1, и матрицы будут $n \times n$, а не 2×2.

Можно записать это в матричной форме:

$$\underline{\Delta X_t} = A_1^* \underline{\Delta X_{t-1}} + A_2^* \underline{\Delta X_{t-2}} + \Pi \underline{X_{t-3}}, \tag{7.57}$$

здесь подчеркнутые переменные — это векторы.

В нашем случае мы допускаем три элемента AR, так что конечное уравнение включает лишь два временных лага, точно так же, как и в упомянутом выше примере. Однако матрица П — это матрица $n \times n$.

Число отдельных векторов коинтеграции переменных определяется рангом матрицы П. Если ранг равен m $(0 < m < n)$, то существует m векторов коинтеграции.

В случае существования векторов коинтеграции П может быть разложена на две матрицы — $n \times m$ и $m \times n$. Назовем эти матрицы α и γ, и П будет произведением α и γ, т.е. $\Pi = \alpha\gamma$. Ряды γ таковы, что для каждого ряда γ, $\gamma \times X_{t-3}$ будет $I(0)$. Ряды матрицы γ и формируют векторы коинтеграции. Таким образом, получаем:

$$\gamma X_{t-3} = Z_{t-3}.$$

Z_{t-3} будет $I(0)$-вектором порядка m, если существует m векторов коинтеграции. Опять-таки матрица α представляет скорость приведения к равновесию.

Проверка коинтеграции нескольких переменных

МНК не подходит для определения векторов коинтеграции в условиях многих переменных. Более подходящий тест составляющих вектора коинтеграции — это вероятностное отношение Йохансена, или тест "следа" — ("trace" test) (Йохансен, 1988; Йохансен и Йезулиус, 1990), который привлекает модель исправления ошибок для определения независимых векторов коинтеграции и для проверки их стационарности в пределах матрицы П.

Процедура Йохансена имеет две функции. Первая — определение числа векторов коинтеграции в группе временных рядов, вторая — обеспечение оценок максимального правдоподобия векторов коинтеграции и векторов скорости приведения. Обе модели кратко описаны в приложениях 7.1 и 7.2 соответственно. Однако, помимо этого, многие пакеты прикладных экономических программ содержат процедуры коинтеграции. Мы использовали Microsoft 3.0 для получения результатов в приведенном ниже примере.

Для иллюстрации использования теста коинтеграции в рядах многих переменных и оценки модели исправления ошибок мы выбрали ежедневные курсы фунта стерлингов и гонконгского доллара (HK$), малайского доллара (MD), тайского бхата (TB) и филиппинского песо (*FP*) за 1991—1995 годы включительно.

Первая стадия — это проверка интегрирования рядов обменных курсов, являются ли они рядами *I*(1). Здесь мы применим расширенный критерий Дики—Фуллера, где допускается тренд, как в уравнении (7.28).

*X*1	*X*2	*X*3	*X*4
—1,7976	—1,7447	—1,8849	—1,9562

Мы должны отбросить нулевую гипотезу о нестационарности, если статистический критерий будет иметь большее отрицательное значение, чем критическое. Поскольку критическое значение равно —3,4168, то мы можем заключить, что данные ряды *I*(1).

Вторая стадия — это проверка ранга матрицы П. Так как у нас четыре валюты, то может быть не более трех векторов коинтеграции. Процедура тестирования заключается в следующем. Сначала проверяется нулевая гипотеза о том, что существует один вектор, затем гипотеза о двух векторах и т.д. Мы отвергаем нулевую гипотезу, что *m* — число векторов коинтеграции — меньше чем *n*, если значение статистического критерия больше указанного критического значения. Детали по использованию критерия "следа" по данным четырем валютам приведены в табл. 7.2.

Таблица 7.2

Нулевая гипотеза	Альтернативная гипотеза	Статистический критерий	95%-ное критическое значение
$m = 0$	$m = 1$	62,1827	47,2100
$m \le 1$	$m = 2$	19,5523	29,6800
$m \le 2$	$m = 3$	8,6202	15,4100
$m \le 3$	$m = 4$	2,4095	3,7620

Для определения количества векторов коинтеграции в рядах динамики мы сначала проверяем нулевую гипотезу, что не существует векторов коинтеграции, т.е. $m = 0$, против альтернативной гипотезы, что существует один такой вектор. Мы должны отвергнуть нулевую гипотезу, так как рассчитанное значение критерия равно 62,1827 против критического значения 47,2100, откуда делаем выводы о том, что существует один вектор коинтеграции. Затем проверяем гипотезу, что существует один вектор против альтернативной гипотезы о том, что существуют два вектора коинтеграции. Здесь рассчитанный критерий меньше критического значения, и мы принимаем нулевую гипотезу. То же самое и в случае с альтернативной гипотезой о трех и четырех векторах. Таким образом, мы заключаем, что существует один вектор коинтеграции.

Затем приводим матрицу П в табл. 7.3.

Таблица 7.3

	*X*1	*X*2	*X*3	*X*4
*X*1	–0,046042	–0,076484	0,0003787	0,019572
*X*2	–0,023645	–0,039278	0,0001945	0,058003
*X*3	–0,13644	–0,22666	0,0011223	0,058003
*X*4	–0,085460	–0,14196	0,0007030	0,036329

Эта матрица П может быть разложена на матрицу оценок векторов коинтеграции, заданную вектором 1×4 в табл. 7.4 и на вектор параметров приведения, заданных вектором 4×1 в табл. 7.4.

Произведение стандартизованных переменных в векторе 4×1 и стандартизованных переменных в векторе 1×4 дает матрицу П 4×4 как показано в табл. 7.3.

Таблица 7.4

$$\begin{bmatrix} -0{,}26189 \\ (0{,}04606) \\ -0{,}13449 \\ (0{,}023645) \\ -0{,}77610 \\ (0{,}13644) \\ -0{,}4861 \\ (0{,}085460) \end{bmatrix} \begin{bmatrix} 0{,}17581 & 0{,}29205 & -0{,}0014461 & -0{,}074736 \\ (-1{,}0000) & (-1{,}6612) & (0{,}0082255) & (-0{,}42510) \end{bmatrix}$$

Умножая члены вектора соответствующих прошлых изменений 1 × 4 на нестандартизованные члены вектора 4 × 1, получаем следующее выражение Z:

$$\begin{bmatrix} 0{,}17581 & 0{,}29205 & -0{,}0014461 & -0{,}074736 \end{bmatrix} \begin{bmatrix} HKS_{t-2} \\ MD_{t-2} \\ TB_{t-2} \\ FP_{t-2} \end{bmatrix}. \qquad (7.58)$$

Оценка многофакторной модели исправления ошибок

Чтобы смоделировать

$$\Delta \underline{X}_t, \qquad (7.59)$$

следует оценить матрицы A_1^* и A_2^*. Это достигается путем регрессии

$$\Delta \underline{X}_t - \Pi \underline{X}_{t-3},$$

по

$$\Delta \underline{X}_{t-1} \text{ и } \Delta \underline{X}_{t-2}.$$

Уравнения оценки модели исправления ошибок для двух переменных даны в уравнении (7.39), где регрессия ΔX_t по прошлым значениям Z и прошлым изменениям X и Y. Здесь мы обобщаем этот процесс для четырех переменных и векторного процесса. Таким образом, для оценки модели исправления ошибок для каждой из четырех валют произведем четыре отдельные регрессии. Покажем этот процесс на примере только гонконгского доллара.

Оценочное уравнение будет:

$$\Delta HK\$_t - \alpha Z_{t-2} = B_0 + B_1 \Delta HK\$_{t-1} + B_2 \Delta MD_{t-1} + B_3 \Delta TB_{t-1} + B_4 \Delta FP_{t-1} + C_1 \Delta HK\$_{t-2} + C_2 \Delta MD_{t-2} + C_3 \Delta TB_{t-2} + C_4 \Delta FP_{t-2}$$

Результаты регрессии

$$\Delta HK\$_t - \alpha Z_{t-2} =$$

$$\underset{(28{,}88)}{= 0{,}0806} - \underset{(-3{,}45)}{0{,}3353\,\Delta HK\$_{t-1}} + \underset{(1{,}20)}{0{,}3079\,\Delta MD_{t-1}} + \underset{(3{,}19)}{0{,}0274\,\Delta TB_{t-1}} + \underset{(4{,}03)}{0{,}0877\,\Delta FP_{t-1}}$$

$$\underset{(-2{,}65)}{-0{,}2524\,\Delta HK\$_{t-2}} + \underset{(1{,}36)}{0{,}3489\,\Delta MD_{t-2}} + \underset{(1{,}87)}{0{,}0161\,\Delta TB_{t-2}} + \underset{(2{,}01)}{0{,}0429\,\Delta FP_{t-2}}.$$

Параметры, относящиеся к первому и второму временным лагам малайского доллара и тайского бхата, незначимы при 5%-ном уровне значимости. Следовательно, их можно удалить и оценивать модель исправления ошибок с меньшим числом переменных.

ОБОБЩЕННАЯ АВТОРЕГРЕССИОННАЯ УСЛОВНАЯ ГЕТЕРОСКЕДАСТИЧНОСТЬ (GARCH)

В связи с возрастающей неустойчивостью финансовых рынков и растущим значением опционов в управлении рисками возрос интерес к нестационарности финансовых рисков. Мандлеброт (Mandlebrot, 1963) отметил, что большие изменения цен активов влекут за собой большие изменения в сторону как возрастания, так и убывания, в то время как малые изменения влекут малые изменения. В частности, финансовые переменные имеют спокойные периоды, за которыми следуют периоды сравнительной нестабильности, т.е. нестабильность является непостоянной, а **изменяющейся во времени**. Методы, разработанные в 1980-х (Engle, 1982; Bollerslev, 1986; Nelson, 1991), дают эконометрические инструменты предсказания будущей нестабильности. Ингл ввел понятие авторегрессионной условной гетероскедастичности (Autoregressive conditional heteroscedasticity — ARCH). Боллерслей обощил этот процесс до общей ARCH, или GARCH.

При условии нестабильности, изменчивой во времени, со стороны не предрасположенных к риску экономических посредников разумно требовать изменяющуюся во времени премию за риск в качестве вознаграждения за принятие на себя финансового риска. Модель ARCH математического ожидания, разработанная Инглом и другими (1987) и развитая Нельсоном (1991), задает рамки анализа влияния риска, изменяющегося во времени, на премии за риск, требуемые посредниками.

Вначале рассмотрим условные моменты временных рядов. Затем перейдем к анализу ARCH и GARCH и рассмотрим применение GARCH на примере предсказания изменчивости обменного курса US$. После этого перейдем к применению метода E-GARCH на том же примере, затем рассмотрим модель GARC-M и ее применение к изменяющимся во времени премиям за риск, а в заключение мы перейдем к рассмотрению того, как методы GARCH двух переменых позволяют определять изменяющиеся во времени дисперсии и коинтеграции при управлении рисками.

Условные моменты временных рядов

Для начала определим некоторые основные понятия. Традиционное обозначение математического ожидания случайной переменной Y_t:

$$E(Y_t) = \mu, \tag{7.60}$$

μ в этом случае называется **безусловным математическим ожиданием**. Это всего лишь вероятностно взвешенное среднее ожидаемых значений случайной переменной. Таким же образом безусловная дисперсия определяется как

$$\sigma^2 = E[(Y_t - \mu)^2]. \tag{7.61}$$

Однако мы заинтересованы в **условной средней** m_t и в **условной дисперсии**, которую обозначают h_t. Условная средняя — это математическое ожидание случайной переменной, когда ожидания обусловлены информацией о других случайных переменных. Эта средняя обычно является функцией этих других переменных. Аналогично условная дисперсия — это дисперсия случайной переменной, обусловленная информацией о других случайных переменных.

Условное математическое ожидание формально записывается как

$$m_t = E[Y_t | F_{t-1}]. \quad (7.62)$$

Вертикальная черта | обозначает "при условии".

Из уравнения видно, что m_t — это математическое ожидание Y_t, обусловленное множеством данных F, доступных в предыдущий период времени. Множество F может включать любые источники информации. Естественный источник — это прошлые значения Y_t. Таким образом, мы могли бы, например, найти условную среднюю m_t, рассчитав уравнение регрессии Y_t по прошлым значениям Y_t, с лагом в один временной период.

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \varepsilon. \quad (7.63)$$

Тогда m_t определяется как сумма $\alpha_0 + \alpha_1 Y_{t-1}$ и является, таким образом, изменяющейся во времени оценкой средней величины. Также видно, что это уравнение — AR(1) процесс.

Эта средняя величина — условное математическое ожидание, поскольку обусловлена предыдущими значениями Y. Мы можем увеличить число временных лагов либо по причине априорного знания структуры лагов, либо пока не исчезнет автокорреляция остатков.

Условная дисперсия определяется следующим образом:

$$h_t^2 = E[(Y_t - m_t)^2 | F_{t-1}]. \quad (7.64)$$

Как мы уже видели, разность между Y_t и средней величиной равна ε_t. Отсюда можно вывести условную дисперсию h_t как функцию прошлых остатков уравнения условной средней, возведенных в квадрат. Таким образом, например, мы можем найти значение h_t из уравнения

$$h_t^2 = \alpha_0 + \alpha_1 \varepsilon^2_{t-1}, \quad (7.65)$$

при условии только одного временного лага.

Модели ARCH и GARCH

Итак, мы разобрали все составляющие моделей ARCH и GARCH. Начнем с однофакторной ARCH, затем перейдем к однофакторной GARCH и закончим многофакторными вариантами этих моделей.

Однофакторная модель ARCH

Среднее квадратическое отклонение, как мы определили его в гл. 2, является безусловным средним квадратическим отклонением. Мы полагали, что изменчивость была постоянна на всем протяжении истории ряда данных. Таким образом, мы предполагали, что дисперсия данных была стационарной. В действительности же известно много временных рядов финансовых показателей, обладающих дисперсией, изменяющейся во времени. Из гл. 6 мы знаем, что это условие называется гетероскедастичностью. Процесс ARCH был впервые разработан Инглом (1982) для отражения изменчивости дисперсии во времени. Одно из многочисленных применений этой модели — моделирование волатильности доходности активов, которая, как известно, изменяется во времени.

Определение характеристик ARCH в рядах динамики начнем с моделирования условной средней величины. Для этого определим авторегрессионную модель доходности. Модель AR(p) предполагает, что требуется p временных лагов независимой переменной. Покажем это на примере модели AR(1), предполагая, что текущий уровень доходности зависит только от одного предыдущего значения доходности, например:

$$r_t = \alpha_0 + \alpha_t r_{t-1} + \varepsilon_t. \tag{7.66}$$

Конечно, можно включить любое количество предыдущих значений r_t, пока не исчезнет автокорреляция ε_t. Если же r_t — значение случайной переменной — задается выражением $\alpha_0 + \alpha_1 r_{t-1}$, то оно является условной средней, поскольку обусловлено предыдущим значением (или значениями r_t). Формула может быть представлена в общем виде, чтобы принять во внимание p лагов, следующим образом:

$$r_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i r_{t-i} + \varepsilon_t\,. \tag{7.67}$$

Цель моделирования условной средней состоит в том, чтобы определить ряд квадратов остатков (ε_t^2), на основании которых можно найти условную дисперсию. Вспомните из изложенного в гл. 6, где предполагалось, что остатки в уравнении регрессии, рассчитанной по методу наименьших квадратов, обладают постоянной (равной нулю), средней и средним квадратическим отклонением, равным ε (гомоскедастичным). Таким образом, дисперсия остатков будет равна ε_i^2. При использовании ARCH предполагается, что остатки ε_i обладают непостоянной дисперсией, обозначенной h^2. Таким образом, $\varepsilon_i = \sqrt{h_i^2 \cdot z}$, где z — это единичная нормаль.

Таким образом, на основании временных рядов квадратных остатков уравнения условной средней можно написать следующее уравнение условной дисперсии:

$$h_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2, \tag{7.68}$$

где h_t^2 — условная дисперсия. Здесь мы опять предполагаем наличие только одного временного лага. Таким образом, h_t^2 — **условная дисперсия** — является авторегрессионным процессом квадратов остатков, **авторегрессионной условной гетероскедастичностью**. Этот метод можно представить в общем виде с любым числом (p) лагов при расчете квадратов остатков, включенных в модель:

$$h_t^2 = \beta_0 + \sum_{i=1}^{p} \beta_i \varepsilon_{t-i}^2 . \tag{7.69}$$

Эта форма записи ARCH взята у Ингла (1982) и называется линейной моделью ARCH(p), где p — число лагов при расчете значений квадратов остатков, включенных в модель.

Проверка однофакторной ARCH

Для проверки существования ARCH необходимо возвести в квадрат ошибки из первоначального уравнения условной средней. Этот ряд квадратов регрессируется по константе и прошлым значениям квадратов с лагом p. Критерием является $T \cdot R2$, где T — размер выборки и R^2 — коэффициент множественной регрессии из уравнения регрессии квадратов ошибок. Этот критерий подчиняется χ^2-распределению. Число степеней свободы равно числу временных лагов в регрессии. Если значение критерия больше критического значения из таблиц χ^2, то нулевая гипотеза о том, что ARCH не присутствует, отвергается.

Однофакторная модель GARCH

Одна из проблем, связанных с формулировкой ARCH — это то, что величины $\alpha_i s$ всегда должны быть неотрицательны для того, чтобы условная дисперсия всегда была положительной. Однако при включении большого числа лагов, что требуется для моделирования некоторых процессов, ограничение неотрицательности может быть нарушено. Раньше обычно старались убедиться в том, что число лагов было произвольно ограничено применением *ad hoc* (специальной) линейно убывающей структурой коэффициентов.

Боллерслев (1986) обобщил модель ARCH путем включения предыдущих значений условной дисперсии, чтобы избежать длиннолаговой структуры ARCH(q), разработанной Инглом (1982). Таким образом, обобщенная ARCH, или GARCH(p, q), определяет условную дисперсию как линейную комбинацию p предыдущих квадратов остатков из уравнения условной средней и q лагов предыдущих значений условной дисперсии:

$$h_t^2 = \beta_0 + \sum_{i=1}^{p} \beta_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \alpha_i h_{t-i}^2 , \tag{7.70}$$

где α, β и γ не меньше нуля во избежание вероятности появления отрицательных значений условных дисперсий.

Это и является уравнением GARCH. Оно показывает, что текущее значение условной дисперсии является функцией от константы — некоторого значения квадратов остатков из уравнения условной средней плюс некоторое значение предыдущей условной дисперсии. Например, если условная дисперсия наилучшим образом описывается уравнением GARCH (1, 1), то объясняется это тем, что ряд является AR(1), т.е. значения ε рассчитаны с лагом в один период и условная дисперсия тоже рассчитана с таким же лагом.

Рассмотрим применение GARCH на примере предсказания изменчивости доходности стерлинговых активов в долларах США.

Модель условного математического ожидания здесь будет моделью AR(2) и параметры регрессии следующие:

$$r_{t_{\text{us\$}}} = \alpha_0 + \alpha_1 r_{t-1} + \alpha_2 r_{t-2} + \varepsilon,$$

$$r_{t_{\text{us\$}}} = 0{,}00005 + 0{,}01927\ r_{t-1} - 0{,}0571 r_{t-2},$$

$$(0{,}285) \qquad (0{,}502) \qquad (-1{,}526)$$

В скобках показаны значения t-критерия.

Уравнение условной дисперсии и значения t-критерия выглядят следующим образом:

$$h_t^2 = 0{,}0 + 0{,}04643\varepsilon^2_{t-1} + 0{,}9429h^2_{t-1}.$$
$$(2{,}062) \quad (3{,}572) \quad (57{,}178)$$

Этот результат показывает, что условная дисперсия в момент времени t значимо определяется при помощи одного временного лага квадратов остатков уравнения условной средней и величиной самой условной дисперсии с лагом, равным 1.

Экспоненциальная модель GARCH: E-GARCH

В модели GARCH (p, q) условная дисперсия зависит от размера остатков, а не от их знака. Хотя существует свидетельство, например у Блэка (1976), что волатильность и доходность активов обладают отрицательной корреляцией. Таким образом при росте цен на ценные бумаги при положительной доходности волатильность падает, и наоборот, когда цена активов падает, приводя к снижению доходности, то волатильность растет. В самом деле, периоды высокой волатильности связаны со спадами на фондовых рынках, а периоды низкой волатильности ассоциируются с подъемом на рынках.

Нельсон (1991) разработал E-GARCH для следующей ситуации:

$$\log(h_t^2) = \alpha_0 + \sum_{i=1}^{p} \alpha_i \frac{|\varepsilon_{t-i}|}{h_{t-i}} + \sum_{i=1}^{p} \gamma_i \frac{\varepsilon_{t-i}}{h_{t-i}} + \sum_{i=1}^{q} \beta_i \log(h_{t-i}^2). \qquad (7.71)$$

Заметьте, что ε включаются в уравнение как в виде фактических необработанных данных, так и по модулю, т.е. в форме $|\varepsilon|$. Таким образом, E-GARCH моделирует условную дисперсию как асимметричную функцию значений ε. Это позволяет положительным и отрицательным предыдущим значениям иметь различное влияние на волатильность. Представление в логарифмическом виде позволяет включать отрицательные значения остатков, не получая при этом отрицательную условную дисперсию.

Заметьте, что условные средние квадратические отклонения (h_{t-1}) включаются в качестве деноминаторов в правой части уравнения.

Мы использовали модель E-GARCH применительно к той же информации, что и в случае с GARCH. Параметры регрессии и соответствующие значения t-критерия оказались следующими:

$$\log(h_t^2) = -0{,}01097 + 0{,}118\frac{|\varepsilon_{t-i}|}{h_{t-i}} + 0{,}2496\frac{\varepsilon_{t-i}}{h_{t-i}} + 0{,}9885\log(h_{t-i}^2).$$
$$(-1{,}807) \qquad (5{,}306) \qquad (2{,}070) \qquad (164{,}077)$$

Результаты показывают значимость асимметричной трактовки остатков из уравнения условной средней. Это также подчеркивает значимость GARCH переменной.

Модель GARCH-M

Если рискованность финансовых активов изменяется во времени, то было бы логичным предположить, что требуемая инвесторами доходность будет также изменяться во времени. Поскольку все активы, включая безрисковые, обладают доходностью не ниже безрисковой (обычно за безрисковую доходность принимается доходность краткосрочных государственных облигаций с нулевым купоном, таким, как казначейские векселя), то уместно будет смоделировать премию за риск. Премией за риск считается разность между доходностью рискового актива и актива, не несущего риск.

Модель GARCH-M, разработанная Инглом и др. (1987), выражает условную среднюю как функцию от условной дисперсии, так же как и авторегрессионную функцию прошлых значений рассматриваемой переменной. Общая разновидность GARCH первоначальной модели ARCH выглядит следующим образом:

$$\begin{aligned} y_t &= \beta + \delta h_t + \varepsilon_t; \\ h_t^2 &= \gamma + a\sum_{i=1}^{i}\varepsilon_{t-i}^2 + \sum_{i=1}^{i}h_{t-i}^2. \end{aligned} \tag{7.72}$$

Заметьте, что в уравнении условной средней дисперсия преобразована в условное среднее квадратическое так, что она выражается в тех же самых единицах измерения, что и премия за риск.

Ингл и др. (1987) применили формулировку ARCH приведенной выше модели к премиям за риск одно- и шестимесячных казначейских векселей по отношению к предположительно безрисковым трехмесячным векселям, и к 20-летним корпоративным облигациям против трехмесячных казначейских векселей. В последнем случае в уравнение регрессии математического ожидания включают третью переменную, чтобы отразить спред доходностей трехмесячных и 20-летних облигаций.

Эта же модель была применена Френчем и др. (*French et al*, 1987) к премии за риск американских акций за период 1928–1984 гг. Они использовали модель условной дисперсии GARCH (1, 2).

Проверка модели GARCH

Для проверки адекватности модели GARCH необходимо проверить стандартизованные остатки ε/σ, где σ - условное среднее квадратическое отклонение, рассчитываемое по модели GARCH и ε — остатки в уравнении условного математического ожидания. Если модель GARCH достаточно хорошо определена, то стандартизованные остатки будут независимы и идентично распределены. Этот критерий проводится в два этапа.

Первый этап — расчет критерия Люнга—Бокса (LB) для квадратов значений первичных данных. Это влечет за собой расчет k коэффициентов автокорреляции на основе T наблюдений. Затем коэффициенты автокорреляции γ возводим в квадрат, получая γ^2. Критерий LB рассчитывается следующим образом :

$$\mathrm{LB} = T(T+2)\sum_{k=1}^{m}\left(\frac{1}{T-k}\right)\gamma_k^2 \approx \chi_m^2, \tag{7.73}$$

где m — максимальный временной лаг коэффициентов автокорреляции.

Второй этап — расчет критерия LB по стандартизованным остаткам. Таким образом, каждый остаток делится на соответствующее значение условного среднего квадратического отклонения. Рассчитываются коэффициенты автокорреляции и возводятся в квадрат. Критерий LB рассчитывается следующим образом:

$$\mathrm{LB} = T(T+2)\sum_{k=1}^{m}\left(\frac{1}{T-k}\right)\gamma_k'^2 \approx \chi_{m-p-q}^2, \tag{7.74}$$

где m — максимальный временной лаг автокорреляции, как и до этого, p — число лагов квадратов остатков из уравнения условной средней и q — число лагов значений условной дисперсии.

Если модель GARCH достаточно хорошо определена, то критерий LB стандартизованных остатков будет меньше критического значения χ^2_{m-p-q}.

Возникает вопрос о наиболее уместной модели GARCH и наиболее подходящих параметрах GARCH. Ответ на этот вопрос находят методом проб и ошибок, т.е. сравнивая значения критерия LB по моделям альтернативных типов или структур параметров.

Волатильность GARCH

Как мы уже отметили, волатильность не является постоянной, а изменяется во времени. Следовательно, волатильность GARCH, которая по определению изменяется во времени, это прдходящий измеритель. Конечно, это верно, только если применяется точная модель GARCH. Теория финансов мало говорит о точном определении модели, поэтому это предмет дальнейших эмпирических исследований.

Однако предполагая, что применяется точная модель, для нахождения годовой волатильности нужно определить квадратный корень из условной дисперсии и умножить на квадратный корень из числа наблюдений в год. Эта мера волатильности будет изменяться во времени, т.е. текущая волатильность является функцией от прошлой волатильности.

Для предсказания волатильности при помощи GARCH можно использовать рекурсивную модель следующего вида:

$$h^2_{t+1} = \beta_0 + \beta_1\varepsilon_t^2 + \gamma h_t^2,$$

$$h^2_{t+j} = \beta_0 + (\beta_1 + \gamma_1)h^2_{t+j-1}. \tag{7.75}$$

Во втором уравнении ε^2, величина которого неизвестна, когда выполняется прогноз, заменяется на условную оценку h^2. Таким образом, второе уравнение позволяет предсказывать h^2 в момент времени $t + 1 (j = 1)$, затем h^2 в момент времени $t + 2 (j = 2)$ и т.д. Результат каждого расчета является предсказанием условной дисперсии на отдельный период, на j периодов вперед.

Кроме того, можно получить значения стандартных ошибок коэффициентов и рассчитать изменяющиеся во времени доверительные интервалы для наших прогнозов.

Двухфакторная GARCH

Двухфакторная модель GARCH может применяться для определения условной дисперсии и условной ковариации и коинтеграции двух переменных. Кроме того, можно включить параметр коинтеграции в уравнение условного математического ожидания, который затем делает параметры GARCH пригодными для создания более эффективных коэффициентов хеджирования, поскольку процесс хеджирования эффективен, когда переменные коинтегрированы.

Сначала мы рассмотрим двухфакторную GARCH на примере доходности акции s_t и фьючерса f_t, после чего введем параметр коинтеграции и рассчитаем коэффициент хеджирования.

Для иллюстрации применения этой модели допустим, что в нашем примере модели условных средних имеют следующий вид:

$$s_t = \alpha_0 + \alpha_1 s_{t-1} + \varepsilon_{s_t};$$
$$f_t = \beta_0 + \beta_1 f_{t-1} + \varepsilon_{f_t} \tag{7.76}$$

и уравнения условной дисперсии и ковариации

$$h^2{}_{s_t} = \alpha_0 + \alpha_1 \varepsilon^2{}_{s_{t-1}} + \alpha_2 h^2{}_{s_{t-1}};$$
$$h^2{}_{f_t} = \beta_0 + \beta_1 \varepsilon^2{}_{f_{t-1}} + \beta_2 h^2{}_{f_{t-1}}; \tag{7.77}$$
$$\mathrm{cov}_{st,ft} = \gamma_0 + \gamma_1 \varepsilon_{s_{t-1}} \varepsilon_{f_{t-1}} + \gamma_2 \mathrm{cov}_{s_{t-1}, f_{t-1}}.$$

Условная дисперсия в данном случае будет симметричной матрицей 2 × 2

$$\begin{bmatrix} h^2_{ss_t} & h^2_{sf_t} \\ h^2_{fs_t} & h^2_{ff_t} \end{bmatrix} = \begin{bmatrix} \sqrt{h^2_{s_t}} & 0 \\ 0 & \sqrt{h^2_{f_t}} \end{bmatrix} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \begin{bmatrix} \sqrt{h^2_{s_t}} & 0 \\ 0 & \sqrt{h^2_{f_t}} \end{bmatrix}, \tag{7.78}$$

где диагонали — условные дисперсии, а числа вне диагоналей — условные ковариации.

Когда мы получаем значение волатильности GARCH вышеописанным способом, то GARCH корреляция выглядят следующим образом:

$$\rho_{sf_t} = \mathrm{cor}_{sf_t} = \frac{\mathrm{cov}_{sf_t}}{\sqrt{\sigma_{s_t^2}\sigma_{f_t^2}}}. \qquad (7.79)$$

Таким образом, двухфакторная GARCH может применяться при нахождении изменяющихся во времени коинтеграции и ковариации с их непосредственным использованием при построении портфеля ценных бумаг и определении коэффициентов хеджирования с наименьшей дисперсией.

Можно также использовать и модель исправления ошибки, вытекающую из наличия коинтеграции. Например:

$$s_t = \alpha_0 + \alpha_1(s_{t-1} - \lambda f_{t-1}) + \varepsilon_{s_t};$$

$$f_t = \beta_0 + \beta_1(s_{t-1} - \lambda f_{t-1}) + \varepsilon_{f_t}. \qquad (7.80)$$

Остатки из этих уравнений могут войти в уравнения условной дисперсии, как это описано ранее.

Если двухфакторная GARCH используется для определения коэффициентов хеджирования, то они определяются следующим образом:

$$\hat{b}_t^* = \frac{\hat{\mathrm{cov}}_{sf_t}}{\hat{h}^2_{ff_t}}, \qquad (7.81)$$

где $\hat{\mathrm{cov}}_{sf}$ — условная ковариация между s и f, и $\hat{h}_{ff}$ — дисперсия f. Это аналогично коэффициенту наклона B в уравнении регрессии, рассчитанной по методу наименьших квадратов. Преимущество двухфакторной модели GARCH коэффициента хеджирования в том, что он определяется на основе изменяющихся во времени дисперсий и ковариаций, тогда как коэффициент наклона уравнения регрессии основывается на неизменности дисперсий и ковариаций.

УПРАЖНЕНИЯ

1. Назовите различия между:
   а) случайным блужданием и мартингалом;
   б) марковским свойством и мартингальным свойством;
   в) стационарностью и белым шумом.
2. С точки зрения ARIMA объясните, что означают процессы AR, MA и степень инеграции.

3. Объясните, как используются коэффициент автокорреляции и частный коэффициент автокорреляции при анализе структуры временного ряда.
4. Объясните, как можно проверить степень инеграции данных.
5. Дайте интуитивное объяснение коинтеграции.
6. Объясните, почему в случае коинтегрированных переменных подразумевается включение модели исправления ошибки.
7. Как применяется двухстадийный процесс Ингла—Грейндже-ра для идентификации коинтеграции?
8. Как определить условную дисперсию при:

   а) однофакторном GARCH-процессе;
   б) многофакторном GARCH-процессе.

9. В чем отличие моделей E-GARCH и GARCH-M от обычной GARCH модели?
10. Как определить волатильность рассматриваемой переменной, используя модели GARCH?
11. Что такое двухфакторная GARCH и как она может применяться при определении коэффициентов хеджирования?
12. Что такое оценка максимального правдоподобия? Как она работает?

СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ

Black, F. (1976) Studies in stock price volatility changes. *Proceedings of the 1976 Meetings of the American Statistical Association*, August.

Bollerslev, T. (1986) Generalised autoregressive conditional heteroskedasticity. *Journal of Econometrics*, **31**, 307—27.

Bollerslev, T., Chou, R.V. and Kroner, K.F. (1992) ARCH modeling in finance: a review of the theory and empirical evidence. In Engle, R.F. and Rothchild, M. eds, *ARCH Models in Finance*. Supplement to the *Journal of Econometrics*, **52**.

Clare, A.D., Maras, M. and Thomas S.H. (1995) The integration and efficiency of international bond markets. *Journal of Business Finance and Accounting* **22** (2), 313—22.

Cheeley-Steeley, P.L. and Pentacost, E.J. (1994) Stock market efficiencY, the small firm effect and cointegration. *Applied Financial Economics* **4**, 405—11.

Choudhry, T. (1994) Stochastic trends and stock prices: an international enquirY. *Applied Financial Economics*, **4**, 383—90.

Dickey, D.A. and Fuller, W.A. (1979) Distribution of estimators for autocorrelated time series with a unit root. *Journal of American Statistical Association*, **74**, 427—31/

Engle, R.F. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. *Econometrica*, **50**, 987—1008.

Engle, R.F. and Granger, C.W.J. (1987) Cointegration and error correction: representation, estimation and testing, *Econometrica*, **55**, 251—76.

Engle, R.F., Lillian, D.M. and Robins, R.P. (1987) Estimating time varying risk premia in the term structure: the ARCH-M model. *Econometrica*, **55**, 391—408.

Engle, R.F. and Rothchild, M. (1992) ARCH models in finance. Supplement to the *Journal of Econometrics*, **52**.

French, K.R., Swert, G.W. and Stambaugh, R.F. (1989) Expected stock returns and volatility. *Journal of Financial Economics*, **19**, 3—29.

Granger, C. (1986) Developments in the study of cointegrated variables. *Oxford Bulletin of Economics and Statistics*, **43**, 213—28.

Johansen, S. (1988) Statistical analysis of cointegrating vectors. *Journal of Economic Dynamics and Control*, **12**, 231—54.

Johansen, S. (1991) Estimation and hypothesis testing of cointegrating vectors in Gaussian vector autoregressive models. *Econometrica*, **59**, 1551—81.

Johansen, S. and Jesulius, K. (1990) Maximum likelihood estimation and inference on coinegration — with applications to the demand for money. *Oxford Bulletin of Economics and Statistics*, **52**, 169—210.

Mackinnon J.J. (1990) *Critical Values for Cointegration Tests*.

Mandelbrot B.B. (1963) New methods in statistical economics. *Journal of Political Economy*, **71**, 421—40.

Nelson, D.B. (1991) Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, **2**, 347—70.

Phillips, P.C. (1987) Time series regressions with a unit root. *Econometrica*, **55**, 277—301.

Phillips, P.C. and Perron P.(1988) Testing for a unit root in time series regression. *Biometrica*, **75**.

ПРИЛОЖЕНИЕ 7.1.
Оценка максимального правдоподобия

Существуют две формы оценки максимального правдоподобия — максимальное правдоподобие полной информации и максимальное правдоподобие ограниченной информации. Последнее является методом одиночного уравнения, а первый метод — многофакторный и его мы опишем в этом разделе.

В гл. 6, при изучении регрессионного анализа, мы использовали обычный метод наименьших квадратов для оценки неизвестных параметров выборки. Параметры определяют конкретную модель, которую мы используем для описания наблюдаемых значений. Таким образом, в главе о регрессии мы хотели описать поведение зависимой переменной Y через независимую переменную X_i ($i = 1$, $i = 2$, ..., $i = n$). В результате получили модель

$$\hat{Y} = \alpha_0 + \alpha_1 X_1 + \ldots + \alpha_k X_k, \quad (\text{П.7.1})$$

α_i — параметры нашей модели, и наша задача — выбрать параметры, чтобы $\hat{Y}$ был как можно ближе к Y, представляющим ряд наблюдаемых значений. Это достигалось нахождением таких α_i, при которых

$$\Sigma(Y - \hat{Y})^2, \quad (\text{П.7.2})$$

была минимальной.

Более общий подход к оценке параметров обеспечивается использованием метода **максимального правдоподобия** (maximum likelihood). Согласно этому подходу данные рассматриваются как свидетельство, относящееся к параметрам распределения. Свидетельство выражается как функция неизвестных параметров — **функция правдоподобия**:

$$L(X_1, X_2, X_3, \ldots, X_n; \Phi_1, \Phi_2, \ldots, \Phi_k), \quad (\text{П.7.3})$$

где X_i — наблюдаемые значения и Φ_i — параметры, которые мы хотим оценить. Например, если генеральная совокупность, из которой делаются выборки, нормально распределена, то параметры, которые мы оцениваем — это μ и σ^2.

Функция максимального правдоподобия представляет собой совместную вероятность наблюдений выборки

$$L(X_1, X_2, X_3, \ldots, \Phi_1, \Phi_2, \ldots, \Phi_k) = \\ = P(X_1 \wedge X_2 \wedge X_3 \ldots X_n). \quad (\text{П.7.4})$$

Цель метода максимального правдоподобия состоит в максимизации функции правдоподобия. Это достигается дифференцированием функции максимальной вероятности по каждому из оцениваемых параметров и приравниванием частных производных нулю. Значения параметров, при которых значение функции максимально, и является искомой оценкой.

Обычно для упрощения последующей работы сначала берется логарифм функции правдоподобия.

Оценка параметров регрессии по методу наименьших квадратов — то же самое, что и оценки по методу максимального правдоподобия, если остатки уравнения регрессии нормально распределены. Таким образом, удобно демонстрировать метод максимального правдоподобия на примере оценок МНК.

Из гл. 6 мы знаем, что модель МНК выглядит следующим образом:

$$Y = \alpha + \beta X + \varepsilon, \qquad (\Pi.7.5)$$

где предполагается, что ε нормально распределены $N(0, \sigma^2)$, т.е.

$$Y - (\alpha + \beta X) \sim N(0, \sigma^2). \qquad (\Pi.7.6)$$

Для каждой пары наблюдаемых значений X и Y будет существовать, при условии нормальности функция плотности вероятностей следующего вида:

$$f(X_i, Y_i) = \frac{1}{\sqrt{2\Pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{Y_i - (\alpha + \beta X_i)}{\sigma}\right)^2}. \qquad (\Pi.7.7)$$

При условии n совместных наблюдений X и Y общая вероятность наблюдения всех значений в выборке равна произведению индивидуальных значений функции плотности вероятностей. Таким образом, функция правдоподобия задается следующим образом:

$$L(\alpha, \beta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\Pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{Y_i - (\alpha + \beta X_i)}{\sigma}\right)^2}. \qquad (\Pi.7.8)$$

Поскольку легче дифференцировать сумму, чем произведение, то обычно берется логарифм функции правдоподобия, таким образом:

$$\ln L(\alpha, \beta) = \sum_{i=1}^{n}\left(\ln\left(\frac{1}{\sqrt{2\Pi\sigma^2}}\right) - \frac{1}{2\sigma^2}(Y_i - (\alpha + \beta X_i))^2\right). \qquad (\Pi.7.9)$$

Это полезное преобразование не влияет на конечный результат, потому что $\ln L$ — это возрастающая функция L. Таким образом, те значения α и β, которые максимизируют $\ln L$, также будут максимизировать L. Затем возьмем первую производную функцию (П.7.9) по α и β и приравняем их нулю таким образом:

$$\frac{\partial \ln L(\alpha, \beta)}{\partial \alpha} = \sum_{i=1}^{n} \frac{1}{\sigma^2}(Y_i - (\alpha + \beta X_i)), \qquad (\Pi.7.10)$$

которое становится

$$\frac{\partial \ln L(\alpha, \beta)}{\partial \alpha} = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n} Y_i - \left(n\alpha + \beta \sum_{i=1}^{n} X_i^2\right)\right), \qquad (\Pi.7.11)$$

и

$$\frac{\partial \ln L(\alpha, \beta)}{\partial \beta} = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n} X_i Y_i - \left(\alpha \sum_{i=1}^{n} X_i + \beta \sum_{i=1}^{n} X_i^2\right)\right). \qquad (\Pi.7.12)$$

Приравнивая производные нулю, мы требуем, чтобы:

$$\sum_{i=1}^{n} Y_i = n\alpha + \beta \sum_{i=1}^{n} X_i , \tag{П.7.13}$$

и

$$\sum_{i=1}^{n} X_i Y_i = \alpha \sum_{i=1}^{n} X_i + \beta \sum_{i=1}^{n} X_i^2 . \tag{П.7.14}$$

Эти уравнения являются уравнениями регрессии по методу наименьших квадратов, таким образом мы показали, что решение регрессии с использованием МНК будет решением по методу максимального правдоподобия, когда остатки уравнения МНК нормально распределены.

Йохансен использовал метод максимального правдоподобия для нахождения параметров вектора коинтеграции.

Оценка максимального правдоподобия в моделировании ARCH и GARCH

Вспомните, что мы начинали моделирование ARCH и GARCH с уравнения условного математического ожидания:

$$r_t = \alpha_0 + \sum_{i=1}^{m} \alpha_i r_{t-i} + \varepsilon_t . \tag{П.7.15}$$

Отсюда остатки будут равны:

$$\varepsilon_t = r_t - \left(\alpha_0 + \sum_{i=1}^{m} \alpha_i r_{t-1} \right).$$

Более того, $\varepsilon = h_t^2 z$, где h^2 — это условная дисперсия и $z \sim N(0, 1)$.

Таким образом, $\varepsilon_t \sim N(0, h_t^2)$, где

$$h_t^2 = \beta_0 + \sum_{i=1}^{p} \beta_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \gamma_i h_{t-i}^2 . \tag{П.7.16}$$

(Заметьте, что это запись модели GARCH, для ARCH значения γ равны нулю.)

Итак, мы имеем $m + 1 + p + q + 1$ параметр для оценки ($m + 1$) значений альфа из уравнения условного математического ожидания, ($p + 1$) — бэта и q—гамма из уравнения условной дисперсии.

Плотность вероятностей наших наблюдений равна:

$$f(r_t;\ \alpha_0, \alpha_1, \ldots, \alpha_m;\ \beta_0, \beta_1, \ldots, \beta_p;\ \gamma_1, \gamma_2, \ldots, \gamma_q) =$$

$$= \frac{1}{\sqrt{2\pi h_t}} e^{\frac{1}{2}\left(\frac{r_t - \left(\alpha_0 + \sum_{i=1}^{m} \alpha_i r_{t-i}\right)}{h_t}\right)^2},$$

где

$$h_t^2 = \beta_0 + \sum_{i=1}^{p} \beta_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \gamma_i h_{t-i}^2 .$$

Наша функция правдоподобия является произведением этих функций для различных наблюдений r (при условии достаточного числа наблюдений для расчетов

$$\alpha_0 + \sum_{i=1}^{m} \alpha_i r_{t-i}, \qquad \beta_0 + \sum_{i=1}^{p} \beta_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \gamma_i h_{t-i}^2,$$

для данных значений α и β).

Отсюда

$$\ln L(\alpha_0, \alpha_1, \ldots, \alpha_m;\ \beta_0, \beta_1, \ldots, \beta_p;\ \gamma_1, \gamma_2, \ldots \gamma_q),$$

$$= -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\sum\nolimits_t \ln(h_t) - \frac{1}{2}\sum\nolimits_t \left(\frac{r_t - \left(\alpha_0 + \sum_{i=1}^{m} \alpha_i r_{t-i}\right)}{h_t}\right)^2 .$$

Это выражение зависит не только от α (которые видны), но и от β и γ, которые скрыты, в том смысле, что они являются подсуммами h.

Задача здесь — найти значения α, β и γ, при которых значе-ние $\ln L$ будет максимально. Это достигается числовыми методами поиска.

Оценка максимального правдоподобия при коинтеграции

В интересах доступности изложения в общей модели, приведенной в главе, отсутствует стохастический элемент. Уравнение (7.57), например, должно выглядеть так:

$$\Delta X_t = A_1^* \Delta X_{t-1} + A_2^* \Delta X_{t-2} + \Pi X_{t-3} + \varepsilon_t, \qquad (\text{П}.7.17)$$

ε_t в этом уравнении является вектором.

Предполагается, что компоненты независимы и подчиняются нормальному распределению с неизвестными дисперсиями. Таким образом, функция правдоподобия будет функцией от этих неизвестных дисперсий вместе с элементами матриц A_1^*, A_2^* и П.

Теоретически мы должны раскрыть уравнение (П.7.17) для получения точного выражения компонентов ε_t, выраженных в тех же параметрах, и отсюда построить и максимизировать функцию правдоподобия. На практике это было бы довольно сложно и эффективнее будет использовать для достижения этого матричную алгебру. Для этого был разработан порядок Йохансена.

ПРИЛОЖЕНИЕ 7.2.

Каноническая корреляция и регрессия

В регрессионном анализе мы стараемся смоделировать зависимую переменную как линейную комбинацию совокупности независимых переменных. Наша задача состоит в том, чтобы выбрать "лучшую" линейную комбинацию, которая достигнет наилучшего соответствия моделируемых значений наблюдаемым. Когда же сами независимые переменные являются неизвестными линейными комбинациями других переменных, тогда нам нужно воспользоваться каноническим анализом.

В нашем анализе многофакторной коинтеграции мы показали, что при помощи простых алгебраических действий можно выразить векторный авторегрессионный процесс $\underline{X}_t = A_1\underline{X}_{t-1} + A_2\underline{X}_{t-2} + A_3\underline{X}_{t-3}$ в следующем виде

$$\Delta\underline{X}_t = A_1^*\Delta\underline{X}_{t-1} + A_2^*\Delta\underline{X}_{t-2} + \Pi\underline{X}_{t-3}. \qquad (\text{П.7.18})$$

Мы сказали, что если П имеет полный ранг, мы можем найти решение для $\underline{X}_{t-3}$. Это будет значить, что составляющие $\underline{X}$ являются $I(0)$, поскольку будут выражены в виде разностей. Но это противоречит первоначальному предположению, что они $I(1)$. Это значит, что составляющие X в действительности являются стационарными, что для $\underline{X}$ найдены изменения разности и что правильная модель — это VAR-модель в уровнях.

Мы сказали, что если ранг П равен нулю, то мы имеем VAR в разностях — стандартный подход к моделированию нестационарного процесса.

Отсюда мы заинтересованы в рассмотрении возможности того, что ранг П не является ни полным, ни нулевым. Это говорит о существовании коинтеграции. Проблема заключается в том, что в условиях белого шума мы не будем знать точного ранга П. Таким образом, мы должны построить статистическую процедуру оценки матрицы и ее компонентов (α и γ) и ассоциированный критерий ее ранга.

Одна из таких процедур была разработана Йохансеном (1988). Согласно этому методу уравнение (П.7.18) запишется следующим образом:

$$\Delta \underline{X}_t - \Pi \underline{X}_{t-3} = A_1^* \Delta \underline{X}_{t-1} + A_2^* \Delta \underline{X}_{t-2}. \qquad (\text{П.7.19})$$

В этой форме, при том что члены белого шума для удобства опущены, мы видим, что надо выразить линейную комбинацию $\Delta \underline{X}_t$ и $\underline{X}_{t-3}$ через линейную комбинацию предыдущих разностей.

Если бы мы знали П, то мы могли бы регрессировать $\Delta \underline{X}_t - \Pi \underline{X}_{t-3}$ по $\underline{X}_{t-1}$ и $\underline{X}_{t-2}$ для того, чтобы найти A_1^*, чтобы найти A_2^* (это подразумевает отдельные регрессии для каждого компонента). Мы, таким образом, могли бы проверить несколько предполагаемых матриц П и выбрать те, к которым конечные регрессии подходят лучше всего. Это получается, когда линейные комбинации, представляющие правую и левую стороны уравнения (П.7.19), имеют наивысшую корреляцию, так что анализ называется канонической корреляцией, а регрессии — каноническими.

Как видим, философия этого метода схожа с философией метода максимального правдоподобия. В этом контексте параметры (элементы матрицы П) выбираются таким образом, что максимизируют не функцию правдоподобия, а функцию корреляции.

Конечно, мы не обязаны делать повторные предположения относительно П. Мы можем использовать оценки максимального правдоподобия для регрессий. Подробности алгебраических действий были бы совершенно излишни в этом тексте, но общий план процедуры таков:

1) регрессируйте $\Delta \underline{X}_t$ по предыдущим разностям и запишите остатки R_{0t};
2) регрессируйте $\underline{X}_{t-k}$ по k—1 предыдущим разностям и запишите остатки R_{kt} (в нашем примере $k = 3$);
3) постройте четыре матрицы

$$S_{00} = \frac{1}{n}\sum_{t=1}^{n} R_{0t}R_{0t}^T; \qquad S_{0k} = \frac{1}{n}\sum_{t=1}^{n} R_{0t}R_{kt}^T;$$

$$S_{k0} = \frac{1}{n}\sum_{t=1}^{n} R_{kt}R_{0t}^T; \qquad S_{kk} = \frac{1}{n}\sum_{t=1}^{n} R_{kt}R_{kt}^T;$$

4) оцените П, заданную $S_{0k}\, S_{kk}^{-1}$;

5) собственные векторы матрицы П являются решениями уравнения $|\lambda S_{kk} - S_{ko} S_{00}^{-1} S_{ok}| = 0$, где вертикальная черта значит "определяющий";
6) квадраты канонических коэффициентов корреляции, которые незначимо отличаются от нуля, показывают сниженный ранг P. Таким образом, критерий ранга P основывается на проверке наименьших коэффициентов из ранжированного ряда, или на сумме наименьших;
7) оценками векторов коинтеграции являются соответствующие собственные векторы П;
8) как только известна оценка матрицы П, при помощи МНК можно получить оценки A_i^*.

8

ЧИСЛЕННЫЕ МЕТОДЫ

ВВЕДЕНИЕ

Термин "численные методы" описывает методы решения математических проблем путем многократного повторения математической процедуры либо для поиска решения, либо для агрегирования множества приближенных оценок в одно окончательное решение. Примером первого может служить использование итеративной процедуры для решения уравнений, которые не решаются простыми способами. Пример второго — агрегирование множества небольших площадей под кривой нормального распределения для нахождений общей площади, если она не может быть найдена аналитическим способом интегрирования. Третья форма численных методов известна как метод Монте-Карло. Как следует из названия метода, это процесс нахождения решений посредством имитации случайных процессов, т.е. осуществление многократных расчетов по математической модели с последующим нахождением среднего значения полученных результатов.

РЕШЕНИЕ УРАВНЕНИЙ

Процесс математического моделирования — представление реальных ситуаций математическими выражениями часто ведет к дальнейшим проблемам, которые заключаются в решении уравнения или уравнений этой математической модели. Особенно это относится к нелинейным уравнениям. Нелинейные уравнения — это уравнения, одна из переменных которых возведена в степень, большую или меньшую единицы. Например,

$$2x^2 - 2 = 4$$

является нелинейным уравнением, так как x возведено в степень 2. Уравнение вида

$$2x - 2 = 4$$

линейное.

Все линейные и некоторые нелинейные уравнения могут быть решены довольно просто, поскольку существуют соответствующие формулы. Существуют другие нелинейные уравнения, которые не могут быть решены подобным образом не потому, что они сложные, а потому что не существует соответствующих формул.

Даже если и существует формула, то может быть более удобным или необходимым приближенное решение. В качестве примера рассмотрим значение x в следующем нелинейном уравнении:

$$x^2 - 2 = 0.$$

Решением является квадратный корень из двух. Однако такого рационального числа, которое в квадрате было бы равно двум, не существует. Решение — иррациональное число с точностью до десяти знаков после запятой — 1,4142135624. В действительности же ответ имеет бесконечное количество десятичных цифр.

С помощью метода проб и ошибок мы должны найти решение, которое бы имело допустимую степень точности. Такой поиск посредством проб и ошибок известен как итерация, или итеративный поиск.

В финансовом анализе часто встречается нелинейное уравнение, которое не может быть решено с помощью только одной формулы, — уравнение внутренней ставки доходности (IRR) для серии из пяти или более денежных потоков.

В гл. 1 мы узнали, что IRR — это ставка дисконтирования, которая приводит совокупность будущих денежных потоков к их текущей стоимости. Например, ставка, которая дисконтирует будущие купонные платежи и стоимость облигации при погашении к ее текущей рыночной стоимости, — это IRR. Она называется ставкой общего дохода или полным доходом при погашении.

Для того чтобы рассчитать ставку общего дохода, мы должны решить уравнение (представляющее собой многочлен), полученное исходя из расчета цены облигации. Рассмотрим, например, двухгодичную облигацию, по которой выплачивается годовой купон в размере 10% и оцененную в настоящее время в 100 единиц. Цена этой облигации будет определяться выражением:

$$P = \frac{10}{1+r} + \frac{110}{(1+r)^2}.$$

Для того чтобы определить ставку общего дохода (или IRR), мы должны знать текущую цену или стоимость. Так, если текущая стоимость равна 100 единицам, уравнение принимает вид:

$$100 = \frac{10}{1+r} + \frac{110}{(1+r)^2}.$$

Это уравнение можно привести к многочлену степени два (так как один из аргументов будет возведен в квадрат):

$$100x^2 - 10x - 110 = 0,$$

где $x = 1 + r$. Этот тип уравнения известен также как квадратное уравнение. Для решения этого уравнения мы располагаем формулой:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{8.1}$$

С помощью этой формулы находятся решения квадратных уравнений вида:

$$ax^2 + bx + c = 0. \tag{8.2}$$

Применив формулу к нашему уравнению, где $a = 100$, $b = -10$ и $c = -110$, получим следующие решения:

$$\frac{+10+\sqrt{10^2+4\cdot 100\cdot 110}}{2\cdot 100}=1{,}1,$$

$$\frac{+10-\sqrt{10^2+4\cdot 100\cdot 110}}{2\cdot 100}=-1{,}0.$$

Экономическая теория не признает отрицательных процентных ставок, только положительное решение может иметь экономический смысл. Тогда $(1 + r) = 1{,}1$.

Возможно, хотя и не особенно полезно, решить также кубические уравнения или уравнения четвертой степени таким же образом, что и квадратные. Однако в этом случае процедура решения будет состоять из алгоритма, а не из простой формулы (кубические уравнения в действительности гораздо труднее квадратных и требуют для полного описания нескольких страниц). Несколько столетий математики пытались разработать аналогичные алгоритмы для уравнений пятой степени (их общая форма: $ax^5 + bx^4 + cx^3 + dx^2 + ex + f = 0$) и уравнений более высоких степеней, но только в 1846 г. было доказано, что таких алгоритмов не существует.

Тем не менее, подобные уравнения играют важную роль в финансовом анализе. Например, чтобы найти IRR для облигации со сроком погашения всего 2,5 года и купонами, выплачиваемыми раз в полгода, мы должны решить следующее уравнение:

$$P=\frac{C_1}{(1+r)^{0,5}}+\frac{C_2}{(1+r)}+\frac{C_3}{(1+r)^{1,5}}+\frac{C_4}{(1+r)^2}+\frac{C_5}{(1+r)^{2,5}}. \qquad (8.3)$$

Для того чтобы привести уравнение к многочлену пятой степени, сначала преобразуем искомую IRR в соответствующий полугодовой эквивалент, т.е. $(1 + r)^{0,5}$. Обозначив это через x, получим:

$$P=\frac{C_1}{x}+\frac{C_2}{x^2}+\frac{C_3}{x^3}+\frac{C_4}{x^4}+\frac{C_5}{x^5}. \qquad (8.4)$$

Умножив обе части уравнения на x^5, получим:

$$Px^5 = C_1x^4 + C_2x^3 + C_3x^2 + C_4x + C_5. \qquad (8.5)$$

Заметим: $(C_1/x) \cdot x^5 = C_1x^4$.

Перенесем все члены уравнения в левую часть и получим:

$$Px^5 - C_1x^4 - C_2x^3 - C_3x^2 - C_4x - C_5 = 0. \qquad (8.6)$$

Поясним это на численном примере. Так, для облигации со сроком погашения 2,5 года, по которой выплачиваются полугодовые купоны в размере 5 единиц, текущая стоимость равна 98 единицам. Уравнение для данной облигации будет:

$$98 = \frac{5}{(1+r)^{0,5}} + \frac{5}{(1+r)} + \frac{5}{(1+r)^{1,5}} + \frac{5}{(1+r)^2} + \frac{105}{(1+r)^{2,5}}.$$

Для того чтобы получить уравнение в виде многочлена, необходимо преобразовать $(1 + r)$ в полугодовой наращенный эквивалент, как было показано выше. Получим:

$$98x^5 - 5x^4 - 5x^3 - 5x^2 - 5x - 105 = 0.$$

Для решения этого уравнения, т.е. для нахождения $(1 + r)$, мы должны использовать метод проб и ошибок. Вспомним, что x — это $(1 + r)^{0,5}$, следовательно, мы должны возвести найденное значение x в квадрат, чтобы получить $(1 + r)$.

Возникает вопрос, как использовать итерационные процедуры для нахождения "неуловимого" $(1 + r)$. Существует несколько подходов, но все начинается с догадки относительного искомого значения, а затем осуществляется поиск решения с необходимой степенью точности. Приведем два итерационных метода. Начнем с **метода деления пополам** (bisection), а затем перейдем к **методу Ньютона—Рафсона** (Newton—Raphson).

Метод деления пополам

Вспомним, что мы пытаемся решить уравнение вида $f(x) = 0$. В качестве простого, хотя и не финансового, примера для иллюстрации метода деления пополам рассмотрим проблему нахождения приближенного значения квадратного корня из двух (приближенное решение уравнения $x^2-2 = 0$). Мы знаем, что единица чересчур мала, поскольку $1^2 < 2$ $(1^2-2 = 1-2 = -1 < 0)$, а два — чересчур велико, так как $2^2 > 2$ $(2^2-2 = 2 > 0)$. Таким образом, наше искомое значение лежит между 1 и 2. Ясно, что на данном этапе значение 1,5 должно подвергнуться проверке, и мы находим, что оно чересчур велико, потому что $1{,}5^2 = 2{,}25 > 2$ ($1{,}5^2-2 = 2{,}25-2 = 0{,}25 > 0$). Таким образом, мы узнаем, что искомое решение находится между 1 и 1,5, интервал неопределенности уменьшился в два раза за счет одного вычисления ($1{,}5^2-2$).

Этот процесс может быть продолжен и далее путем деления пополам интервала, содержащего искомое значение.

В этом примере мы далее производим оценку при $x = (1{,}5 + 1)/2 = 1{,}25$ и находим, что $1{,}25^2-2 < 0$, и, следовательно, определяем, что интервал, в котором лежит решение (интервал неопределенности), находится между 1,25 и 1,5. Процедура повторяется до тех пор пока интервал не уменьшается настолько, чтобы удовлетворить в достаточной степени требования к точности.

Процесс деления пополам схематически отображен на рис. 8.1.

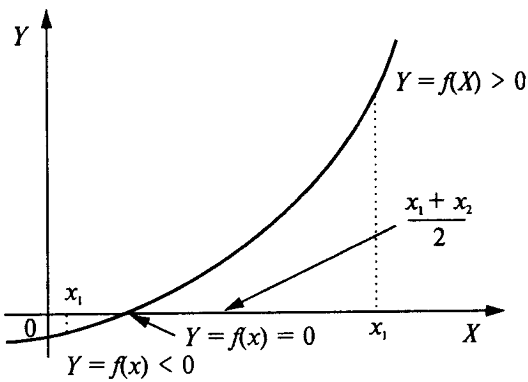

Рис. 8.1.

Формально же процедура может быть определена следующим образом. Прежде всего должны быть найдены такие два значения x (x_1 и x_2), что $f(x_1) < 0$, а $f(x_2) > 0$. Т.е. решение должно находиться между этими двумя значениями. Затем производится оценка функции в средней точке интервала $(x_1 + x_1)/2$ — находится $f[(x_1 + x_1)/2]$. Если $f[(x_1 + x_1)/2] < 0$, тогда x_1 заменяется на $(x_1 + x_1)/2$. Если $f[(x_1 + x_1)/2] > 0$, тогда x_2 заменяется на $(x_1 + x_1)/2$, и все повторяется заново.

Представим расчеты в табличной форме. Первоначально x_1 было выбрано как 1, а x_2 — как 2.

В табл. 8.1 приведены 20 итераций для нахождения решения.

Таблица 8.1

x_1	x_2	$(x_1 + x_1)/2$	$f[(x_1 + x_1)/2]$
1	2	1,5	0,25
1	1,5	1,25	–0,4375
1,25	1,5	1,375	–0,1094
1,375	1,,5	1,4375	0,06641
1,375	1,4375	1,40625	–0,0225
1,4063	1,4375	1,42188	0,02173
1,4063	1,4219	1,41406	–0,0004
1,4141	1,4219	1,41797	0,01064
1,4141	1,418	1,41602	0,0051
1,4141	1,416	1,41504	0,00234
1,4141	1,415	1,41455	0,00095
1,4141	1,4146	1,41431	0,00026
1,4141	1,4143	1,41418	–9E–05
1,4142	1,415	1,41461	0,00112
1,4142	1,4146	1,4144	0,00052
1,4142	1,4144	1,41429	0,00021
1,4142	1,4143	1,41424	6,3E–05
1,4142	1,4142	1,41421	–1E–05
1,4142	1,4142	1,41422	6,4E–06
1,4142	1,4142	1,41422	2,5E–05

Интуитивно мы знали, что 1^2 является чересчур малым значением, а 2^2 — чересчур большим, следовательно, эти числа были взяты как начальные значения для x_1 и x_2 соответственно:

$$f\left(\frac{x_1 + x_2}{2}\right) = f\left(\frac{1+2}{2}\right) = f(1{,}5). \tag{8.7}$$

Затем оценивается значение функции x^2—2 при $x = 1{,}5$. Результат выглядит так:

$$f(1{,}5) = 2{,}25–2– = 0{,}25 > 0.$$

Так как это больше нуля, то $(x_1 + x_2)/2$, 1,5 в данном случае занимает место x_2. Если на следующем этапе значение функции будет меньше нуля, $(x_1 + x_2)/2$ займет место x_1.

Сейчас мы применим ту же самую процедуру для нахождения IRR для рассмотренного ранее случая (облигация со сроком погашения 2,5 года). Для этого решим следующее уравнение:

$$f(x) = 98x^5 - 5x^4 - 5x^3 - 5x^2 - 5x - 105.$$

В этом примере x равно (1 + IRR). Мы догадались, что приближенные значения x_1 и x_2 равны 1,0 и 1,1 соответственно. При $x = 1{,}0$ $f(x) = -27$, т.е. 1,0 чересчур мало. При $x = 1{,}1$ $f(x) = +27{,}3$, т.е. 1,1 чересчур велико. Поэтому следующий шаг — оценка функции в средней точке:

$$f\left(\frac{x_1 + x_2}{2}\right) = f\left(\frac{1{,}0 + 1{,}1}{2}\right) = f(1{,}05),$$

$$f\left(\frac{98 \cdot 1{,}05^5 - 5 \cdot 1{,}05^4 - 5 \cdot 1{,}05^3 - 5 \cdot 1{,}05^2 - 5 \cdot 1{,}05 - 105}{2}\right) = -2{,}5526.$$

Чтобы пояснить процедуру, рассмотрим первую строку табл. 8.2. Функция $f[(x_1 + x_2)/2]$ имеет отрицательное значение —2,5526, тогда $(x_1 + x_2)/2$, т.е. 1,05, заменяет x_1. В следующей строчке (следующая итерация) функция $f[(x_1 + x_2)/2]$ имеет положительное значение 11,6497, тогда $(x_1 + x_2)/2 = 1{,}075$ заменяет x_2 в следующей итерации. Эта процедура повторяется в соответствии с необходимыми заменами, проделываемыми перед каждой итерацией, до тех пор пока не будет достигнута необходимая степень точности, — в нашем случае до тех пор пока $f[(x_1 + x_2)/2]$ не станет равно нулю с точностью до четырех десятичных знаков.

В этом примере была рассчитана полугодовая IRR. Возведение 1,054679 в квадрат дает 1,112348, или годовую IRR, равную 11,2348%.

Этот подход довольно прост, но эффективен. Деление пополам интервала неопределенности на каждой стадии довольно быстро сокращает его до бесконечно малого размера, также на каждой из стадий мы уверены в том, что решение "поймано в ловушку", процесс находится под контролем. Тем не менее, метод имеет один недостаток — необходимо, чтобы на самом первом шаге решение было уже "поймано". Это не всегда легко. Необходим также некоторый мыслительный процесс, создающий трудности для полной автоматизации метода.

Таблица 8.2

x_1	x_2	$(x_1 + x_1)/2$	$f[(x_1 + x_1)/2]$
1	1,1	1,05	–2,5526
1,05	1,1	1,075	11,6497
1,05	1,075	1,0625	4,37346
1,05	1,0625	1,05625	0,86746
1,05	1,0563	1,05313	–0,8532
1,0531	1,0563	1,05469	0,00445
1,0531	1,0547	1,05391	–0,425
1,0539	1,0547	1,0543	–0,2105
1,0543	1,0547	1,05449	–0,103
1,0545	1,0547	1,05459	–0,0493
1,0546	1,0547	1,05464	–0,0224
1,0546	1,0547	1,05466	–0,009
1,0547	1,0547	1,05468	–0,0023
1,0547	1,0547	1,05468	0,00109
1,0547	1,0547	1,05468	–0,0006
1,0547	1,0547	1,05468	0,00025
1,0547	1,0547	1,05468	–0,0002
1,0547	1,0547	1,05468	4,2E–05
1,0547	1,0547	1,05468	–6E–05

Метод Ньютона—Рафсона

Метод Ньютона—Рафсона широко используется в итеративных процедурах и позволяет более быстро находить решения, чем метод деления пополам. Как и метод деления пополам, этот метод начинается с угадывания значения x. Пусть мы предполагаем x_n. Тогда значение функции будет $f(x_n)$, оно также равно высоте вертикальной линии на рис. 8.2. Пусть касательная к кривой функции в точке $(x_n, f(x_n))$ пересекает ось x в точке x_{n+1}.

Вспомним, что в гл. 3, рассматривающей исчисления, мы встречались с понятием наклона кривой в точке, равного соотношению высота/интервал. Высота в нашем примере — это $f(x_n)$, а интервал — $(x_n—x_{n+1})$. Таким образом, зная длины сторон треугольника, можно найти наклон кривой:

$$\frac{f(x_n)}{x_n - x_{n+1}} = f'(x_n) \tag{8.8}$$

Тогда

$$f(x_n) = f'(x_n)(x_n - x_{n+1}), \tag{8.9}$$

или

$$x_{n+1} = x_n - f(x_n)/f'(x_n). \tag{8.10}$$

То есть результат каждой итерации находится через вычитание из результата по предыдущей итерации отношения функции $f(x)$ к ее первой производной $f'(x)$. Однако это не срабатывает только тогда, когда случайным образом одна из x_n является точкой, в которой $f'(x_n)$ равно или близко к нулю (в этом случае, начав процедуру заново с того момента, когда произошел сбой, из новой точки, вы почти всегда решите возникшую проблему).

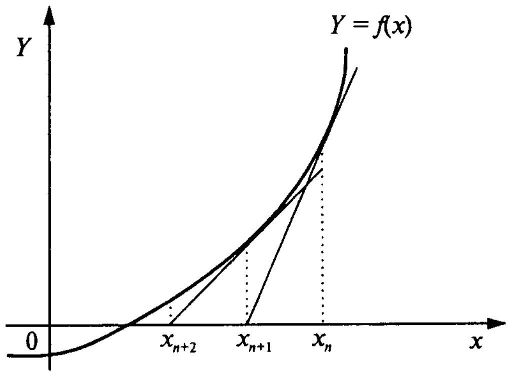

Рис. 8.2

Можно утверждать, что недостаток метода — необходимость предварительного нахождения $f'(x)$. Однако этого можно избежать, если заменить $f'(x)$ на конечное разностное приближение:

$$\frac{f(x+h)-f(x)}{h}, \tag{8.11}$$

где h очень мало и может уменьшаться при выполнении итерационного процесса. Следовательно, теперь в процесс может быть введен компьютер, чтобы рассчитать $f(x_n)$ и $f(x_n + h)$ (для текущих значений x_n и h) и оценить

$$x_n = \frac{f(x_n)}{\left[\dfrac{f(x_n+h)-f(x_n)}{h}\right]}. \tag{8.12}$$

Принимая во внимание сложность и продвинутость современных компьютерных алгебраических пакетов программ (которые должны были бы называться алгебраическо-вычислительными пакетами программ!), непосредственный расчет может производиться автоматически.

Теперь применим процедуру Ньютона—Рафсона к оценке облигации со сроком погашения 2,5 года. Мы уже отметили, что функция $f(x) = 0$ для этой облигации будет являться многочленом пятой степени, а именно:

$$f(x) = 98x^5 - 5x^4 - 5x^3 - 5x^2 - 5x - 105 = 0. \qquad (8.13)$$

В начале итеративного процесса мы предполагаем, что $x = 1{,}055$. Заменив x на 1,055 в выражении (8.13), получим следующий результат:

$$f(x) = 98 \cdot 1{,}055^5 - 5 \cdot 1{,}055^4 - 5 \cdot 1{,}055^3 - 5 \cdot$$
$$\cdot\, 1{,}055^2 - 5 \cdot 1{,}055 - 105 = 0{,}176625.$$

Мы должны скорректировать 1,055 в соответствии с результатом 0,176625 и значением первой производной функции (8.13) 551,2939. Тогда x_1 — значение x во второй итерации равно

$$x_1 = 1{,}055 - \frac{0{,}176625}{551{,}2939} = 1{,}054680.$$

Повторное вычисление по формуле (8.12), но на этот раз для значения 1,054680 в качестве x, дает следующее:

$$f(x) = 98 \cdot 1{,}054680^5 - 5 \cdot 1{,}054680^4 - 5 \cdot 1{,}054680^3 -$$
$$-5 \cdot 1{,}054680^2 - 5 \cdot 1{,}054680 - 105 = 0{,}0001125. \qquad (8.14)$$

Далее мы корректируем 1,054680 в соответствии с результатом 0,0001125 и значением первой производной функции (8.14), равным 550,5916, и получаем x_2 — значение x в третьей итерации:

$$x_2 = 1{,}054680 - \frac{0{,}001125}{550{,}5916} = 1{,}054679.$$

Повторное вычисление по формуле (8.12) при $x = 1{,}054679$ дает

$$f(x) = 98 \cdot 1{,}054679^5 - 5 \cdot 1{,}054679^4 - 5 \cdot 1{,}054679^3 -$$
$$-5 \cdot 1{,}054679^2 - 5 \cdot 1{,}054679 - 105 = 0{,}000017.$$

Следовательно, $x = 1{,}054679$ для $f(x) = 0$ с требуемой точностью до четырех десятичных знаков.

Однако вспомним, что нам нужно найти годовое значение IRR для облигации со сроком погашения 2,5 года, пока же мы только нашли IRR для полугодовых периодов. Поэтому значение x необходимо возвести в квадрат (x^2) для получения искомого 1 + IRR. Значит, $1 + \text{IRR} = 1{,}054679^2 = 1{,}112348$. Это означает, что полный доход при погашении для данной облигации равен 11,2348%. Мы получили точно такой же результат, что и при использовании метода деления пополам, однако он был получен с помощью гораздо меньшего числа итераций.

ЧИСЛЕННЫЕ МЕТОДЫ ИНТЕГРИРОВАНИЯ

Из гл. 3 мы помним, что в процессе интегрирования находится площадь под кривой, а первый этап этого процесса — нахождение первообразной интегрируемой функции. Затем определяется значение первообразной функции в конечных точках интервала для нахождения площади. К сожалению, для множества функций не существует первообразных, хотя это не означает, что не существует и интеграла.

Вспомним из гл. 4, что функция

$$\int_1^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \tag{8.15}$$

это вероятность нормально распределенной переменной принять значение от 1 до 2. Значение этого интеграла представлено в виде заштрихованной области на рис. 8.3.

Поскольку не существует такой функции $f(x)$, чтобы

$$f'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \tag{8.16}$$

методы интегрального исчисления, приведенные в гл. 3, не могут быть использованы, следовательно, необходимо применить численные методы. Покажем два из них. Они заключаются в разбиении площади под кривой на маленькие вертикальные полоски, высота которых равна расстоянию между данной точкой на горизонтальной оси и кривой. Затем рассчитывается площадь каждой из полосок, а их сумма приближенно равна площади под кривой.

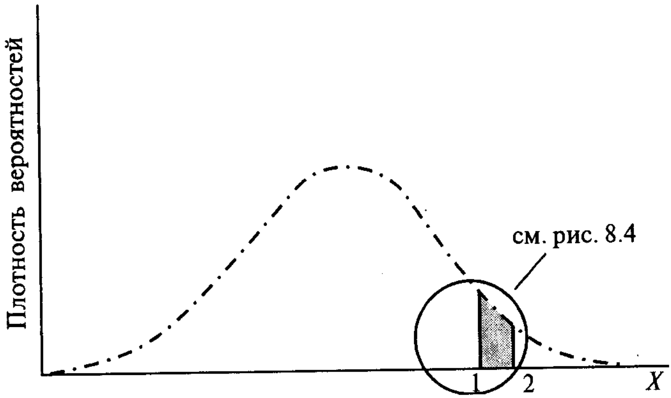

Рис. 8.3

Правило трапеций

Первый метод, который мы продемонстрируем, использует правило трапеций. Согласно этому подходу искомая площадь под кривой разбивается на множество вертикальных узких полосок равной ширины (рис. 8.4). Если каждая линия соединяется сверху с другой линией, полученная форма будет трапецией, которая приближенно соответствует фактической площади под кривой между вертикальными линиями.

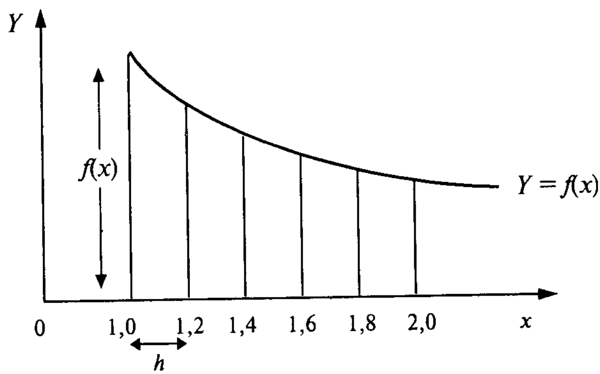

Рис. 8.4

Так как на данном интервале кривая вогнута, то площадь каждой трапеции немного больше фактической площади под кривой между соответствующими линиями. Таким образом, сумма площадей трапеций будет переоценивать искомую площадь. Если бы кривая была выпуклой, площадь трапеций недооценивала бы фактическую площадь под кривой. Этот вид ошибки может быть уменьшен посредством использования большего количества интервалов. Продемонстрируем это позже.

Каждая вертикальная линия выходит из точки x на оси абсцисс, соответствующей аргументу функции $f(x)$. Высота линии определяется значением функции $f(x)$. Тогда площадь трапеции рассчитывается исходя из оценки каждой пары $f(x)$, например, при $f(x = 1{,}0)$ и $f(x = 1{,}2)$ (рис. 8.4) путем нахождения средней этих двух величин (высот) и умножением ее на ширину трапеции (0,2 в нашем примере). В результате площадь под кривой приближенно определяется суммой площадей трапеций.

Для иллюстрации подхода найдем площадь под кривой между точками 1 и 2. Разобьем пространство между ними на пять интервалов.

Если площадь трапеции определяется произведением средней высоты на ширину, тогда общая площадь будет:

$$\left[\frac{f(1{,}0)+f(1{,}2)}{2}\cdot 0{,}2\right]+\left[\frac{f(1{,}2)+f(1{,}4)}{2}\cdot 0{,}2\right]+\left[\frac{f(1{,}4)+f(1{,}6)}{2}\cdot 0{,}2\right]+$$
$$+\left[\frac{f(1{,}6)+f(1{,}8)}{2}\cdot 0{,}2\right]+\left[\frac{f(1{,}8)+f(2{,}0)}{2}\cdot 0{,}2\right].$$

Эту функцию можно представить и в более компактной форме:

$$\frac{0{,}2}{2}[f(1{,}0) + f(2{,}0) + 2(f(1{,}2) + f(1{,}4) + f(1{,}6) + f(1{,}8))].$$

Общий же вид формулы следующий:

$$h/2(\text{“конечные значения”} + 2\cdot\text{”внутренние значения”}), \qquad (8.17)$$

где h — ширина каждой из трапеций, “конечные значения” обозначают **сумму** значений функции в граничных точках, а за “внутренние значения” принимается **сумма** значений функции в остальных (внутренних) точках интервала.

Следовательно, для нахождения интеграла

$$\int_1^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

мы, прежде всего, находим значение функции $f(x)$ для каждого из значений x, разделяющих интервал на пять частей:

x	1	1,2	1,4	1,6	1,8	2
$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$	0,2420	0,1942	0,1497	0,1109	0,0790	0,0540

$h = 0{,}2$,
"граничные значения" $= 0{,}2420 + 0{,}0540 = 0{,}2960$,
"внутренние значения" $= 0{,}1942 + 0{,}1497 + 0{,}1109 + 0{,}0790 = 0{,}5338$.

Тогда

$$\int_1^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \approx \frac{0{,}2}{2}(0{,}2960 + 2 \cdot 0{,}5338) = 0{,}1364.$$

Ранее мы отметили, что точность этого метода может быть улучшена за счет увеличения числа полосок. Поэтому, используя этот же пример, проделаем вычисления снова, но на этот раз для десяти интервалов.

x	1	1,1	1,2	1,3	1,4	1,5
$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$	0,2420	0,2179	0,1942	0,1714	0,1497	0,1295
	1,6	1,7	1,8	1,9	2	
	0,1109	0,0940	0,0790	0,0656	0,0540	

$h = 0{,}1$,
"граничные значения" $= 0{,}2960$ (как и раньше),
"внутренние значения" $= 0{,}2179 + 0{,}1942 + 0{,}1714 + 0{,}1497 +$
$+ 0{,}1295 + 0{,}1109 + 0{,}0940 +$
$+ 0{,}0790 + 0{,}0656 = 1{,}2122$.

Тогда

$$\int_1^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \approx \frac{0{,}1}{2}(0{,}2960 + 2 \cdot 1{,}2122) = 0{,}1360.$$

Правило Симпсона

Правило Симпсона — улучшенный вариант правила трапеций. Площадь под кривой разбивается на **бесконечное** число интервалов одинаковой ширины. Значение функции $f(x)$ определяется для двух граничных и одного среднего значений x для каждой пары интервалов, в результате чего получаем три точки на кривой (рис. 8.5).

Рис. 8.5

Через эти три точки может быть проведена уникальная параболическая кривая, под которой и может быть определена площадь. В большинстве случаев это дает более точное приближение к реальной площади, заключающейся в этих двух интервалах, по сравнению с нахождением и суммированием площадей трапеций.

Не вдаваясь в детали по нахождению соответствующей параболы и площади под ней, так как это довольно утомительно, приведем сразу результат. Площадь оказывается приближенно равна произведению общей ширины двух интервалов и взвешенного среднего высот (значений функций) в соответствии с весами 1, 4 и 1.

Тогда, так же, как и в примере с десятью интервалами (см. правило трапеций), первые два интервала определяются значениями x = 1, 1,1 и 1,2. Приближенно площадь равна:

$$0{,}2 \cdot \frac{f(1{,}0) + 4f(1{,}1) + f(1{,}2)}{6}.$$

Следовательно, приближенная общая площадь:

$$\left[0{,}2\cdot\frac{f(1{,}0)+4f(1{,}1)+f(1{,}2)}{6}\right]+\left[0{,}2\cdot\frac{f(1{,}2)+4f(1{,}3)+f(1{,}4)}{6}\right]+$$

$$+\left[0{,}2\cdot\frac{f(1{,}4)+4f(1{,}5)+f(1{,}6)}{6}\right]+\ldots+\left[0{,}2\cdot\frac{f(1{,}8)+4f(1{,}9)+f(2{,}0)}{6}\right] \tag{8.18}$$

или

$$\frac{0{,}2}{6}\cdot[f(1{,}0)+f(2{,}0)+4(f(1{,}1)+f(1{,}3)+\ldots+f(1{,}9))+$$
$$+2(f(1{,}2)+f(1{,}4)+\ldots+f(1{,}8))].$$

Это также можно записать в виде $h/3$ · ["граничные значения" + 4 · "четные" + 2 · "нечетные"], где "четные" и "нечетные" обозначены подобным образом, так как x = 1,1, 1,3, 1,5, и т.д., т.е. это вторая, четвертая, шестая и т.д. точки, в которых оценивается значение функции, тогда как x = 1,2, 1,4, 1,6 и т.д. являются третьей, пятой, седьмой и т.д. точками.

Таблица функциональных значений из прошлого примера может быть использована для нахождения следующего:

h	= 0,1,
"граничные значения"	= 0,2960,
"четные"	= 0,6784,
"нечетные"	= 0,5338.

Тогда

$$\int_1^2 \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \approx \frac{h}{3}$$

("граничные значения" + 4 · "четные" + 2 · "нечетные") = **0,1359**

$$\frac{0{,}1}{6}[(2\cdot 0{,}2960)+(4\cdot 0{,}6784)+(2\cdot 0{,}5338)]=0{,}1359. \tag{8.19}$$

Точные таблицы стандартного нормального распределения дают результат вероятности между 1 и 2 как 0,1359. Следовательно, правило Симпсона дает более точный результат по сравнению с правилом трапеций при одних и тех же затрачиваемых усилиях на вычисление (это количество расчетов значений функции, которые потенциально дороги с точки зрения затрачиваемого времени).

Нахождение функции в виде многочлена для приближенного описания кумулятивной нормальной кривой

Хотя найти интеграл аналитическим способом для стандартной функции нормальной плотности невозможно, соответствующую площадь под кривой приближенно можно определить численно, используя правила трапеций и Симпсона, которые представляют собой численные методы интегрирования. Альтернативный прием — нахождение или подбор многочлена для описания кумулятивной нормальной кривой.

Описание этого подхода выглядит следующим образом:

1. Рассмотреть площадь под стандартной нормальной кривой, как показано на рис. 8.6, а. Соответствующая кумулятивная функция, или огива, показана на рис. 8.6, б.

   Каждый кумулятивный уровень площади на рис. 8.6, а отображен в виде высоты на рис. 8.6, б, в сущности отображающем кривую кумулятивной площади.
2. Подобрать функцию в виде многочлена, приближенно описывающую верхнюю половину огивы.
3. Для этого следует воспользоваться следующей функцией:

$$N(z) = 1 - \frac{\left(0{,}319382x - 0{,}356564x^2 + 1{,}78148x^3 - 1{,}82126x^4 + 1{,}33027x^5\right) - e^{\frac{1}{2}z^2}}{\sqrt{2\pi}}, \tag{8.20}$$

где $N(z)$ — значение кумулятивной стандартной функции нормальной плотности в точке z, а x определяется следующим отношением:

$$x = \frac{1}{\left(1 + 0{,}231642z\right)}, \tag{8.21}$$

где z — оцениваемая стандартизованная нормальная переменная.

В этом методе подразумевается, что z положительно, так как оценивается только верхняя половина кривой, т.е. рассматривается кумулятивная вероятность, относящаяся к правой половине функции нормальной плотности.

Если z отрицательно, мы должны использовать свойство симметричности функции нормальной плотности. Например, если мы хотим знать кумулятивную частоту для $z = -2$, мы рассчитываем ее для $z = 2$, а полученный результат отнимаем от 1,0.

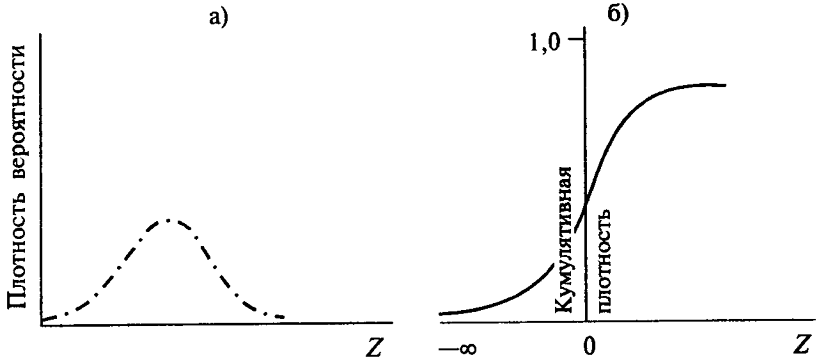

Рис. 8.6

Для иллюстрации метода рассчитаем сначала площадь под кривой для $z = 2$, а затем — для $z = -0{,}8$

$z = 2{,}0$

$$x = \frac{1}{\left(1 + (0{,}231642 \cdot 2)\right)} = 0{,}683394$$

$$\frac{\begin{bmatrix} (0{,}319382 \cdot 0{,}683394) + (1{,}78148 + 0{,}683394^3) \\ -(0{,}356564 \cdot 0{,}683394^2) - (1{,}82126 \cdot 0{,}683394^4) + \\ + (1{,}33027 \cdot 0{,}683394^5) \end{bmatrix}^{e^{-0{,}5 \cdot 2^2}}}{\sqrt{2\pi}} = 0{,}97725$$

$z = -0{,}8$

Проделаем вычисления сначала для $z = +0{,}8$:

$$x = \frac{1}{\left(1 + (0{,}231642 \cdot 0{,}8)\right)} = 0{,}8436585$$

$$\frac{\begin{bmatrix} (0{,}319382 \cdot 0{,}8436585) + (1{,}78148 + 0{,}8436585^3) - \\ -(0{,}356564 \cdot 0{,}8436585^2) - (1{,}82126 \cdot 0{,}8436585^4) + \\ + (1{,}33027 \cdot 0{,}8436585^5) \end{bmatrix}^{e^{-0{,}5 \cdot 0}}}{\sqrt{2\pi}} = 0{,}788328$$

так как мы находим площадь под кривой для $z = -0{,}8$, вычтем значение 0,788328 из 1,0 и получим 0,211672.

Если бы было нужно найти площадь между, например $z = 2$ и $z = 1{,}5$, то расчеты были бы выполнены по формуле (8.20) для $z = 2$ и затем для $z = 1{,}5$, после чего полученное второе значение следовало бы вычесть из первого.

ЧИСЛЕННЫЕ МЕТОДЫ ДЛЯ РЕШЕНИЯ СТОХАСТИЧЕСКИХ ПРОБЛЕМ

Модели поведения переменной могут быть либо **детерминированными**, либо **стохастическими**. Когда в модели предполагается фиксированная (или постоянная) связь между переменными, говорят, что эта модель детерминированная. Например, модель определения заработка трейдера производными финансовыми инструментами, основанная на предопределенной пропорции чистых доходов всего отдела, будет детерминированной, а результат будет получен как только станут известны доходы.

Если модель предполагает наличие некоторого неопределенного элемента, такого, как случайная переменная, говорят, что модель является стохастической. Например, модель прогнозирования будущего уровня курса акций на рынке или фьючерсных цен финансовых активов, содержащая элемент неопределенности, известна как стохастическая. Стохастическая модель приводит не к однозначному результату, а к вероятностному распределению возможных ответов. Модель заработка трейдера была бы стохастической, если использовалась для прогнозирования заработка на основании прошлых уровней доходов в качестве индикатора будущих.

В гл. 4 мы отметили, что переменная, которая предполагает в себе неопределенность, известна как случайная переменная. При построении детерминированных моделей переменным присваивается одно единственное значение, в стохастических же моделях этим переменным присваивается распределение вероятностей.

До этого момента мы имели дело с численными методами, предполагающими детерминированные процессы, т.е. процессы, в которых шанс или случайность не играли никакой роли. В этом разделе мы разберем три широко используемых численных метода в применении к стохастическим проблемам:

- биномиальный процесс,
- триномиальный процесс,
- метод Монте-Карло.

Одно из главных применений этих трех методов заключается в нахождении стоимости опционов, для которых не существует формул или в лучшем случае их трудно использовать. Поэтому продемонстрируем использование этих трех методов в ценообразовании опционов, для чего, немного отступив от основной темы, познакомим читателя с основами ценообразования опционов.

Четвертую группу численных методов — методы конечной разницы — применяют для решения непрерывных во времени уравнений с частными производными. Так как эти уравнения рассматриваются в гл. 10, обсуждение методов конечной разницы отложим до этой главы.

Основы ценообразования опционов

Опцион на покупку (call option) дает покупателю "право, но не обязательство, на покупку актива по предварительно установленной цене до наступления или на момент определенной даты", за это право без обязательства выплачивается опционная премия. Опцион на продажу дает "право, но не обязательство, на продажу актива по предварительно установленной цене до наступления или на момент определенной даты".

Исходя их этого можно заметить, что опцион имеет фиксированный максимальный срок действия, опционное обязательство исполняется согласно фиксированной цене — **цене исполнения** (exercise price, strike price). Покупатель также может отказаться от опциона без угрозы наложения каких-либо штрафов. Текущая стоимость опциона зависит от распределения вероятностей цен по активу, лежащему в основе контракта, в момент истечения срока. В частности, стоимость опциона на покупку зависит от вероятности цены актива быть больше цены исполнения на момент истечения срока, в то время как стоимость опциона на продажу зависит от вероятности цены актива быть меньше цены исполнения.

Определение стоимости на момент исполнения

Никто не станет платить за опцион на покупку актива больше, чем разница между тем, что нужно заплатить для покупки актива на рынке, и тем, что нужно заплатить по опциону, т.е. ценой исполнения. Следовательно, верхний предел стоимости опциона на покупку — это цена самого актива минус цена исполнения.

В случае опциона на продажу актива никто не будет платить за этот опцион больше, чем разница между ценой, по которой актив может быть продан на открытом рынке, и ценой, по которой этот актив будет продан по опциону. Следовательно, верхний предел стоимости опциона на продажу — это цена исполнения минус текущая стоимость актива.

Так как владелец опциона может отказаться от него без последующих обязательств, цена на опцион не может быть отрицательной. Значит, нижний предел цены для опционов обоих типов — это нулевое значение.

Эти ограничивающие условия формально могут быть выражены следующим образом:

$$C = \max\,[0, S - X], \qquad (8.22)$$

$$P = \max\,[0, X - S].$$

Выражение (8.22) определяет стоимость опциона на момент исполнения.

К настоящему времени разработан ряд моделей ценообразования опционов на срок их действия (до момента исполнения). Наиболее широко встречающаяся модель — это модель, разработанная Блэком и Сколсом (1973), которая использует непрерывные временные стохастические исчисления для нахождения стоимости, она будет рассмотрена в гл. 10. Наиболее известная дискретная временная модель — это биномиальная модель, разработанная Коксом и другими (1976), а также Рендельмэном и Бартером (1979). В последнее время также возрастает интерес к триномиальным моделям. В этом разделе мы обсудим биномиальные и триномиальные модели в приложении к ценообразованию опционов.

Биномиальные модели

Эти модели предполагают, что основная случайная переменная (цена ценной бумаги, лежащей в основе опциона) характеризуется биномиальным распределением, которое было обсуждено в гл. 4. Применение биномиальных моделей к ценообразованию опционов подразумевает моделирование цены актива, лежащего в основе опциона как биномиального процесса, будь то цена ценной бумаги, обменный курс или процентная ставка, для определения распределения этой переменной на момент исполнения. Затем с использованием приведенных выше ограничивающих условий, определяется будущая стоимость опциона и дисконтируется к настоящей, определяя, таким образом, текущую цену на опцион.

Биномиальная модель допускает возможность **создания безрискового портфеля посредством хеджирования длинной позиции по активу короткой позицией по ряду справедливо оцененных опционов по покупке этого актива**. Следовательно, только безрисковая процентная ставка должна быть использована для дисконтирования, поскольку предполагается, что если портфель захеджирован наилучшим образом, он будет безрисковым и, следовательно, должен характеризоваться безрисковой процентной ставкой.

Далее продемонстрируем два подхода. Первый, и наиболее популярный — построение биномиальной решетки или дерева, второй подход — использование биномиальной алгебры для получения суммы текущих стоимостей каждой вероятностно взвешенной будущей стоимости опциона. Оба эти подхода предварительно были представлены в гл. 4.

Биномиальная решетка

Стоимость создания безрискового портфеля — это стоимость покупки основного актива минус премии от списанных (проданных) опционов. Следующий пример продемонстрирует, каким образом создается безрисковый портфель.

Предположим, что цена основного актива (S) = 35, цена исполнения опциона (X) = 35, безрисковая процентная ставка (r) = 10%, или 0,1, и $R = 1 + r = 1{,}1$. Срок действия опциона составляет один год. Дополнительно допустим, что в конце года цена актива либо поднимется на 25% с 35 до 43,75, либо упадет на 25% с

35 до 26,25. Графически это может быть проиллюстрировано рис. 8.7, где *u* — это коэффициент движения вверх цены актива (1 + относительный прирост цены), равный в нашем случае 1,25, а *d* — коэффициент движения вниз цены актива (1— относительное падение цены), равный в нашем случае 0,75. Необходимо также условие $d < (1 + r) < u$, поскольку, если *d* и *u* меньше, чем безрисковая ставка, безрисковый актив всегда будет иметь более высокие доходы, чем рисковый актив, что вступает, конечно же, в противоречие с финансовой теорией. Если *d* и *u* больше, чем безрисковая ставка, рисковый актив всегда будет иметь более высокие доходы, чем аналогичный безрисковый актив.

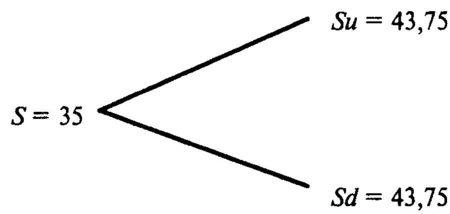

Рис. 8.7

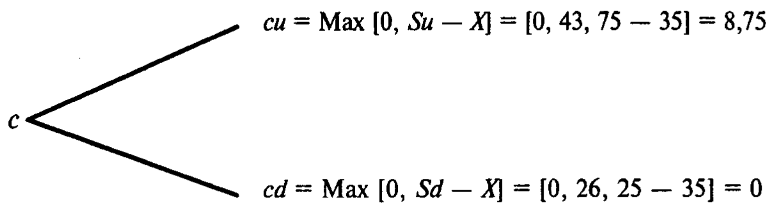

Рис. 8.8

Важно, чтобы процентное движение вверх и вниз было выражено в форме произведения, т.е. $1{,}25 \cdot 35$ или $0{,}75 \cdot 35$. Если бы они были представлены в форме сложения, т.е. плюс или минус 25%, могли бы быть получены отрицательные значения для цен активов.

Исходя из приведенного **одноуровневого биномиального дерева**, цены актива и ограничивающих условий для цены опциона можно построить подобное дерево, относящееся к стоимости опциона.

Из рис. 8.8 видно, что если стоимость актива поднимется до 43,75 в конце одногодичного периода, опцион должен стоить

8,75, обозначим это значение ***cu***. Если стоимость актива упадет до 26,25, опцион не должен стоить ничего, это значение обозначим ***cd***.

Для построения полностью захеджированного портфеля покупается одна единица основного актива с одновременной продажей H опционов. Так как портфель полностью захеджирован, его стоимость будет одинакова независимо от того, поднимется или опустится цена на актив. В обозначениях для биномиального дерева — это $Su—Hcu = 26{,}25 = Sd—Hcd = 26{,}25$ (см. рис. 8.9).

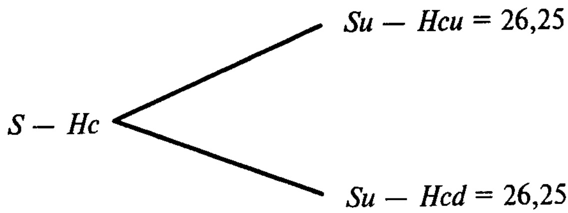

Рис. 8.9

На данном этапе у нас возникают две задачи. **Первая — это найти, сколько опционов необходимо продать, чтобы портфель был безрисковым**, т.е. каково значение H? **Вторая задача заключается в определении справедливой цены, по которой эти опционы должны быть проданы.**

Задача 1: количество опционов, которое следует продать

Определяя количество опционов, которое следует продать для создания безрискового портфеля, заметим, что по приведенному выше биномиальному дереву видно, что размах цен активов определяется выражением $S(u—d)$, или $35(1{,}25—0{,}75) = 17{,}5$. Размах стоимостей опционов, в свою очередь, будет $(cu—cd)$ или $8{,}75—0 = 8{,}75$. Исходя из этой информации мы можем подсчитать количество опционов, которые должны быть проданы, против одной единицы основного актива для того, чтобы портфель стал идеально захеджирован.

Коэффициент хеджирования в этом случае рассчитывается следующим образом:

$$H = \frac{S(u-d)}{(cu-cd)} \tag{8.23}$$

или в численной форме

$$H = \frac{35(1{,}25 - 0{,}75)}{(8{,}75 - 0)} = \frac{17{,}5}{8{,}75} = 2\,.$$

Смысл коэффициента хеджирования может быть объяснен так: $S(u—d)$ — размах возможных цен на актив, а $(cu—cd)$ — размах возможных цен на опцион. Тогда, если размах цен активов в два раза больше размаха цен опционов, возникает необходимость в короткой позиции на два опциона, доход по которым полностью возместит потери от одной длинной позиции по активу.

В итоге, в этом примере безрисковое хеджирование состоит из покупки одной единицы основного актива и продажи двух опционов, цена исполнения которых равна 35. Два возможных результата этой стратегии на момент окончания периода выглядят так:

Цена на актив возрастет до 43,75: $Su—Hcu = 1{,}25(35)—2(8{,}75) = 26{,}25$;

Цена на актив упадет до 26,25: $Sd—Hcd = 0{,}75(35)—2(0) = 26{,}25$.

Задача 2: справедливая цена

Для решения второй задачи, состоящей в определении справедливой цены, по которой следует продать опцион в настоящий момент, заметим, что поскольку стратегия является безрисковой, ставка дохода по опциону также должна быть безрисковой. Текущая стоимость портфеля, состоящего из длинной позиции по S и короткой по двум опционам, а также имеющего безрисковую ставку в течение одного года, должна быть равна текущей стоимости дохода в конце года. Этот доход на конец года составляет $Su—Hcu = 26{,}25$. Значит, текущая стоимость должна быть равна

$$\mathbf{26{,}25/1{,}1 = 23{,}86.}$$

Так как текущая стоимость S равна 35, стоимость двух опционов по короткой позиции должна быть 35—23,86 = 11,14. Следовательно, один опцион должен быть оценен в 11,14/2 = 5,57.

Этот процесс может быть выражен в более обобщенной форме. Если $R = 1 + r$, где r — безрисковая процентная ставка, тогда

$$R(S—Hc) = (Su—Hcu). \qquad (8.24)$$

Это выражение может быть записано следующим образом:

$$RS—RHc = Su—Hcu. \qquad (8.25)$$

Оставив в левой части только *RHc*, получим:

$$-RHc = -RS + Su - Hcu. \quad (8.26)$$

Умножим обе части равенства на —1, тогда

$$RHc = RS - Su + Hcu. \quad (8.27)$$

Разделим обе части равенства на *RH* и поменяем множители местами:

$$c = \frac{S(R-u)+Hcu}{HR}. \quad (8.28)$$

Вспомнив, что

$$H = \frac{S(u-d)}{(cu-cd)}, \quad (8.29)$$

подставим это выражение в (8.28) и после некоторых преобразований получим

$$c = \left[cu\frac{(R-d)}{(u-d)} + cd\frac{(u-R)}{(u-d)}\right] / R, \quad (8.30)$$

что является уравнением опциона на покупку со сроком один год до момента исполнения.

Подставим числовые значения в полученное соотношение ($cu = 8{,}75$, $cd = 0$, $u = 1{,}25$, $d = 0{,}75$, $R = 1{,}1$):

$$c = \left[8{,}75\frac{(1{,}1-0{,}75)}{(1{,}25-0{,}75)} + 0\frac{(1{,}25-1{,}1)}{(1{,}25-0{,}75)}\right] / 1{,}1,$$

$$c = \left[8{,}75\left(\frac{0{,}35}{0{,}5}\right) + 0\right] / 1{,}1 = 5{,}57.$$

Мы можем упростить процедуру, допустив, что $p = (R-d)/(u-d)$ и $(1-p) = (u-r)/(u-d)$, тогда уравнение (8.30) примет вид

$$c = \frac{[pcu + (1-p)cd]}{R}, \quad (8.31)$$

$$c = [(8{,}75 \cdot 0{,}7) + (0{,}3 \cdot 0)]/1{,}1 = 5{,}57.$$

Следовательно, справедливой стоимостью опциона на покупку будет 5,57. Мы можем проверить, действительно ли это справедливая стоимость для данного опциона, так как этот портфель должен приносить доход согласно безрисковой процентной ставке, которая равна 10%.

Установленная стоимость портфеля — это (35—(2 · 5,57)) = 23,86. Эта сумма, инвестированная под безрисковую процентную ставку на один год, увеличится до 23,86 · 1,1 = 26,25, что является точной стоимостью захеджированного портфеля в конце года.

Ценообразование опционов на базе создания безрискового хеджа актива, лежащего в основе опциона, позволяет избежать зависимости цены опциона от ожиданий инвесторов относительно будущей цены этого актива. Все, что необходимо делать в этом случае, — поддерживать эффективность портфеля для того, чтобы он оставался безрисковым.

Многопериодная биномиальная модель

В приведенном примере предполагалось, что время между настоящим моментом и моментом исполнения опциона было разделено лишь на один период, в нашем случае один год. Однако биномиальный подход может быть обобщен таким образом, что срок действия опциона может быть разделен на любое количество временных периодов или биномиальных испытаний. Чем больше количество испытаний за данный период, т.е. чем меньше временной период, соответствующий каждому испытанию, тем более точно определение стоимости опциона.

В самом деле, если временной интервал между испытаниями становится бесконечно малым, т.е. торговля в сущности происходит непрерывно, биномиальная модель становится моделью Блэка—Сколса (см. гл. 10).

Независимо от количества биномиальных испытаний используется один и тот же принцип для нахождения стоимости опциона в каждом узле дерева, начиная от момента исполнения опциона к настоящему временному периоду и, таким образом, к текущей цене опциона. Продемонстрируем эту процедуру в следующем примере, в котором цена актива S равна 35, цена исполнения X составляет 35, годовая безрисковая процентная ставка равна 10%, а годовая волатильность (изменчивость) σ — 20%; одногодичный временной период разделен на четыре квартальных подпериода или биномиальных испытаний.

Перед тем как увеличить количество биномиальных испытаний, годовую безрисковую процентную ставку необходимо скорректировать в соответствии с более короткими временными промежутками между испытаниями. Например, в этой четырехквартальной модели должна быть задействована квартальная сложная процентная ставка, эквивалентная годовой. В этом случае квартальная сложная эквивалентная процентная ставка будет $(1 + r)^{0,25}-1$, где r— годовая процентная ставка, т.е. $(1,1)^{0,25}-1 = 0,024$. Следовательно, $R = 1 + r$, в нашей четырехквартальной модели она составляет 1,024.

В некоторых случаях R может быть представлено как $e^{r \cdot (T-t)/n}$, где r — непрерывно наращенная ставка, эквивалентная безрисковой процентной ставке. В нашем примере 9,53% — непрерывно наращенная ставка, эквивалентная 10% годовых, т.е. $R = e^{0,0953 \cdot 0,25} = 1,024$.

Кроме того, величины потенциального движения в сторону увеличения или уменьшения, т.е. величины u и d, которые относятся к волатильности актива, должны быть определены на основе рыночной информации и скорректированы в соответствии с количеством биномиальных испытаний. Кокс и др. (1979) показали, что u и d соотносятся со средним квадратическим отклонением следующим образом:

$$u = e^{\sigma\sqrt{(T-t)/n}}, \tag{8.32}$$

$$d = e^{-\sigma\sqrt{(T-t)/n}},$$

где $(T-t)$ — срок действия опциона в годах (или соответствующих долях лет), а n — количество биномиальных испытаний. В нашем примере срок действия опциона составляет один год, а количество квартальных биномиальных испытаний равно четырем, следовательно, $(T-t)/n = 0,25$.

Обычно также требуется, чтобы $u = 1/d$, благодаря чему движение цены вверх с последующим движением вниз равнозначно движению вниз с последующим движением вверх. В этом случае видоизмененное биномиальное дерево называется *биномиальной решеткой.*

Следовательно, движение вверх и вниз определяется изменчивостью переменной, средним квадратическим отклонением натурального логарифма отношений цен активов, т.е. средним квадратическим отклонением непрерывно наращенного дохода.

Возвращаясь к нашему примеру, рассчитаем значения u, d, R, p и $1-p$:

$$u = e^{\sigma\sqrt{(T-t)/n}} = e^{0,2\sqrt{0,25}} = 1,10517,$$

$$d = e^{-\sigma\sqrt{(T-t)/n}} = e^{-0,2\sqrt{0,25}} = 0,904837,$$

$$R = (1 + r)^{0,25} = 1,1^{0,25} = 1,024,$$

$$p = \frac{(R-d)}{(u-d)} = 0,59539, \tag{8.33}$$

$$(1-p) = \frac{(u-R)}{(u-d)} = 0,40461.$$

Теперь перейдем к рис. 8.10. Жизненный цикл опциона (или срок его действия) был разделен на четыре квартальных временных периода. Мы можем рассчитать различные стоимости основного актива после четырех периодов, т.е. на момент исполнения опциона. Стоимость опциона зависит от того, какой путь “избрала” цена актива. Например, если цена актива возрастала во всех четырех периодах, соответствующий ей узел будет Su^4 со значением 52,21. Если же, с другой стороны, цена актива возрастала в течение двух и падала также в течение двух периодов, соответствующий узел будет Su^2d^2 со значением 35.

Это дерево также показывает различные потенциально возможные стоимости опциона в конце четвертого периода, т.е. на момент исполнения. Для опциона на покупку стоимость на момент исполнения должна быть max $[0, Su^4 - X]$, если, например, актив возрастал в каждом из периодов. Для расчета текущей стоимости опциона следует начать с этих различных конечных значений и найти стоимость опциона в конце третьего периода с помощью уравнения (8.31).

Затем потенциальные стоимости опциона в конце второго периода находятся исходя из значений, полученных для конца третьего периода, а возможные стоимости опциона рассчитываются на конец первого периода исходя из значений на конец второго периода. И наконец, на основе значений, полученных для первого периода, определяется текущая стоимость опциона. Объясним каждый из этих этапов более подробно, используя биномиальное дерево (рис. 8.10).

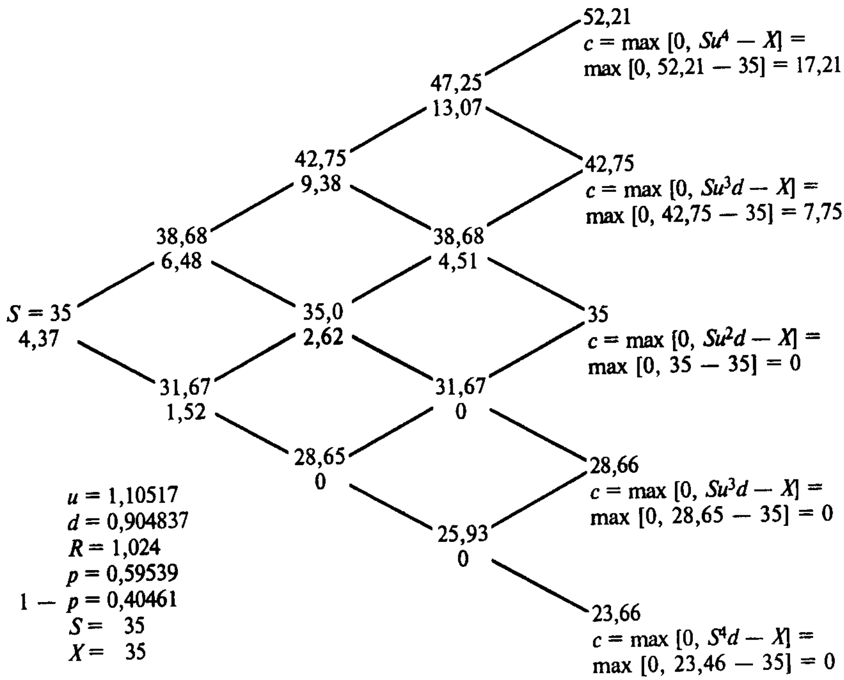

Рис. 8.10. Биномиальная модель опциона на покупку

Этап 1: использование стоимостей на момент исполнения для расчета стоимостей третьего периода

На биномиальном дереве видно, что возможные стоимости опциона равны 17,21; 7,75; 0; 0 и 0. На основе этой информации мы можем рассчитать стоимости опциона на **конец третьего временного периода**.

Прежде всего, вспомним, что u, d и R были изменены с целью учета четырехпериодности этой модели: $u = 1{,}10517$, $d = 0{,}904837$ и $R = 1{,}024$. Следовательно, p будет равно $(1{,}024 - 0{,}904837)/(1{,}10517 - 0{,}904837) = 0{,}59539$, а $1-p$ становится $(1{,}10517 - 1{,}024)/(1{,}10517 - 0{,}904837) = 0{,}40461$.

$$cu^3 = [pcu^4 + (1-p)cu^3d]/R =$$

$$= [0{,}59539 \cdot 17{,}21 + 0{,}40461 \cdot 7{,}75]/1{,}024 =$$
$$= [10{,}25 + 3{,}14]/1{,}024 = 13{,}07.$$

Аналогичным образом мы можем вычислить стоимость опциона для cu^2d:

$$cu^2d = [0{,}59539 \cdot 7{,}75 + 0{,}40461 \cdot 0]/1{,}024 = 4{,}51.$$

Нет необходимости рассчитывать какие-либо значения для других узлов, так как соответствующие величины четвертого периода были равны нулю.

Этап 2: использование стоимостей третьего периода для расчета стоимостей второго периода

Зная, что $cu^3 = 13{,}07$, а $cu^2d = 4{,}51$, мы можем рассчитать значения для соответствующих узлов на **конец второго временного периода** следующим образом:

$$cu^2 = [pcu^3 + (1-p)cu^2d]/1{,}024$$
$$= [0{,}59539 \cdot 13{,}07 + 0{,}40461 \cdot 4{,}51]/1{,}024 = 9{,}38,$$
$$cud = [pcu^2d + (1-p)cud^2]/1{,}024,$$

так как cud^2 равно нулю, то

$$cud = [0{,}59539 \cdot 4{,}51 + 0]/1{,}024 = 2{,}62.$$

Этап 3: использование стоимостей второго периода для расчета стоимостей первого периода

Мы можем найти cu и cd следующим образом:

$$cu = [0{,}59539 \cdot 9{,}38 + 0{,}40461 \cdot 2{,}62]/1{,}024 = 6{,}48,$$
$$cd = [0{,}59539 \cdot 2{,}62 + 0]/1{,}024 = 1{,}52.$$

Этап 4: использование стоимостей первого периода для расчета текущей цены опциона

Наконец, мы можем определить текущую цену опциона следующим образом:

$$c = [0{,}59539 \cdot 6{,}48 + 0{,}40461 \cdot 1{,}52]/1{,}024 = 4{,}37.$$

В гл. 4 мы определили, что ожидаемое будущее значение переменной — это сумма различных возможных будущих значений, помноженных на соответствующие вероятности. Следовательно, мы можем использовать биномиальное вероятностное уравнение из гл. 4 для определения стоимости опциона на покупку путем расчета ожидаемой стоимости актива, превышающего цену исполнения опциона на момент окончания его срока действия , и дисконтирования этой величины к настоящему моменту. Для определения стоимости опциона на продажу необходимо рассчитать ожидаемую стоимость актива ниже цены исполнения опциона. Приведем общую биномиальную формулу для оценки опциона.

Для опциона на покупку:

$$C = \sum_{j=0}^{n}\left(\frac{n!}{(n-j)!\,j!}p^j(1-p)^{n-j}\max[0, u^j d^{n-j}S - X]\right)\Big/(1+r)^n. \tag{8.34}$$

Для опциона на продажу:

$$P = \sum_{j=0}^{n}\left(\frac{n!}{(n-j)!\,j!}p^j(1-p)^{n-j}\max[0, X - u^j d^{n-j}S]\right)\Big/(1+r)^n, \tag{8.35}$$

где p и $1—p$ — описаны ранее, n — общее количество биномиальных испытаний, j — количество произошедших движений вверх, а $n—j$ — количество произошедших движений вниз.

Вспомним, что четырехуровневое биномиальное дерево (рис. 8.10) имеет пять возможных исходов. Su^4 достигается после четырех повышений, стоимость опциона — 17,21 на момент исполнения; Su^3d — после трех повышений и одного понижения цены, стоимость опциона на момент окончания срока действия — 7,75; Su^2d^2 — после двух повышений; Sud^3 — после одного повышения, а Sd^4 — при отсутствии повышений. В каждой из последних трех ситуаций стоимость опциона на момент исполнения нулевая.

Следовательно, для определения стоимости Европейского опциона необходимо выполнить следующие три шага:

1. найти биномиальную вероятность для каждого из результатов на момент окончания срока опциона;
2. умножить дисконтированные стоимости опциона для каждого из этих результатов на соответствующие вероятности;

3. найти сумму полученных выше результатов для получения стоимости опциона.

Используя ту же информацию, что и в четырехпериодном биномиальном процессе, стоимость опциона на покупку можно определить суммой следующих пяти выражений.

Результат после четырех движений вверх (u^4):

$$\frac{\frac{4!}{4!(4-4)!}0{,}59539^4(0{,}40461)^{4-4}[17{,}21]}{(1{,}024)^4}=1{,}967\,.$$

Результат после трех движений вверх и одного движения вниз (u^3d):

$$\frac{\frac{4!}{3!(4-3)!}0{,}59539^3(0{,}40461)^{4-3}[7{,}75]}{(1{,}024)^4}=2{,}408\,.$$

Результат после двух движений вверх и двух движений вниз (u^2d^2):

$$\frac{\frac{4!}{2!(4-2)!}0{,}59539^2(0{,}40461)^{4-2}[0]}{(1{,}024)^4}=0\,.$$

Результат после одного движения вверх и трех движений вниз (ud^3):

$$\frac{\frac{4!}{1!(4-1)!}0{,}59539^1(0{,}40461)^{4-1}[0]}{(1{,}024)^4}=0\,.$$

Результат после четырех движений вниз (d^4):

$$\frac{\frac{4!}{0!(4-0)!}0{,}59539^0(0{,}40461)^{4-0}[0]}{(1{,}024)^4}=0\,.$$

Суммирование этих значений дает точно такой же результат, что и по методу, применявшемуся ранее, т.е. получаем величину 4,37 с точностью до сотых.

В этих примерах предполагалось, что срок действия опциона разбивается лишь на четыре дискретных временных периода. В действительности же время до момента исполнения может быть разделено на множество бесконечно малых временных периодов. Чем больше количество периодов, тем больше степень точности расчетов. Для того чтобы это продемонстрировать, разделим одногодичный срок на 6, 10, 20, 40 и 100 периодов, тогда цены опциона на покупку будут 4,5256, 4,5723, 4,6081, 4,6262 и 4,6371 соответственно. Сравним эти результаты со значением 4,6446, полученным по модели Блэка—Сколса, предусматривающей непрерывный во времени процесс. Таким образом, ошибка для 100-периодной биномиальной модели составила всего 0,0075 единиц опционной премии.

Таким образом, можно отметить интересное обстоятельство, что с уменьшением временного интервала каждого биномиального испытания (т.е. с возрастанием *n*) процесс приближается к непрерывному во времени стохастическому процессу получения дохода по активу, который описывается моделью Блэка—Сколса. Эта модель будет описана в гл. 10, охватывающей финансовую математику непрерывных процессов.

Триномиальный эквивалент биномиальной модели ценообразования опционов

Биномиальная модель ценообразования опционов имеет преимущество умеренной интуитивности и большой гибкости по отношению к опционам, к которым она может быть применена. Однако у нее есть и недостаток — медленное получение результата по сравнению с другими моделями, имеющими однозначное решение.

Представим триномиальный эквивалент биномиальной модели. Эта модель обладает такой же гибкостью, что и биномиальная, но опционная премия рассчитывается быстрее.

В триномиальном процессе цена на основной актив может принять три возможных значения в конце каждого триномиального испытания. Она может возрасти согласно коэффициенту u и составить Su, остаться на том же уровне Sq или упасть до Sd, где $d = 1/u$.

Для того чтобы учесть допущение, что опционы оцениваются в безрисковой среде, а актив следует логнормальному распределению, мы должны скорректировать значения u и d (q равно единице по определению):

$$\begin{aligned} u &= e^{2\sigma\sqrt{\delta t/2}}, \\ d &= e^{-2\sigma\sqrt{\delta t/2}}, \\ q &= 1. \end{aligned} \tag{8.36}$$

Однопериодная триномиальная модель ценообразования имеет вид:

$$C = \frac{p_u cu + p_q cq + p_d cd}{R}. \tag{8.37}$$

Так же как и родственная биномиальная модель, эта модель может применяться для опционов на покупку и на продажу исходя из ограничивающих условий:

$$C = \max\,[S - X{,}0],$$

$$P = \max\,[X - S{,}0].$$

Перед тем как применить эту модель, мы должны рассчитать величины p_u, p_q и p_d. Первый шаг — расчет p, определяемого как

$$p = \frac{e^{r\delta t/2} - e^{-\sigma\sqrt{\delta t/2}}}{e^{\sigma\sqrt{\delta t/2}} - e^{-\sigma\sqrt{\delta t/2}}}. \tag{8.38}$$

Затем рассчитываем $p_u = p^2$, $p_q = 2p(1-p)$ и $p_d = (1-p)^2$.

Для того чтобы проиллюстрировать преимущество экономии времени этой модели, построим двухпериодное дерево и покажем, что точность результата будет такой же, что и при построении четырехпериодного биномиального дерева.

Рис. 8.11 изображает двухпериодное дерево для цены основного актива и стоимости соответствующего опциона на покупку, где $S = 35$, $X = 35$, $r = 0{,}09525$, волатильность $= 0{,}2$ и $(T-t) = 1{,}0$.

Используя выражение (8.37), а также значения параметров, приведенных на рисунке, вместе с конечными значениями для цен активов, стоимости опциона в трех узлах (cu, cq и cd) во время $t + 1$, найдем так же, как и при биномиальном подходе. Например, стоимость опциона на покупку cu:

$$cu = \frac{0{,}354396 \cdot 17{,}2114 + 0{,}481831 \cdot 7{,}7491 + 0{,}163773 \cdot 0}{1{,}0241} = 9{,}3769,$$

а стоимость опциона на покупку cq:

$$cq = \frac{0{,}354396 \cdot 7{,}7491 + 0{,}481831 \cdot 0 + 0{,}163773 \cdot 0}{1{,}0241} = 2{,}6816.$$

Ясно, что значение cd равно нулю.

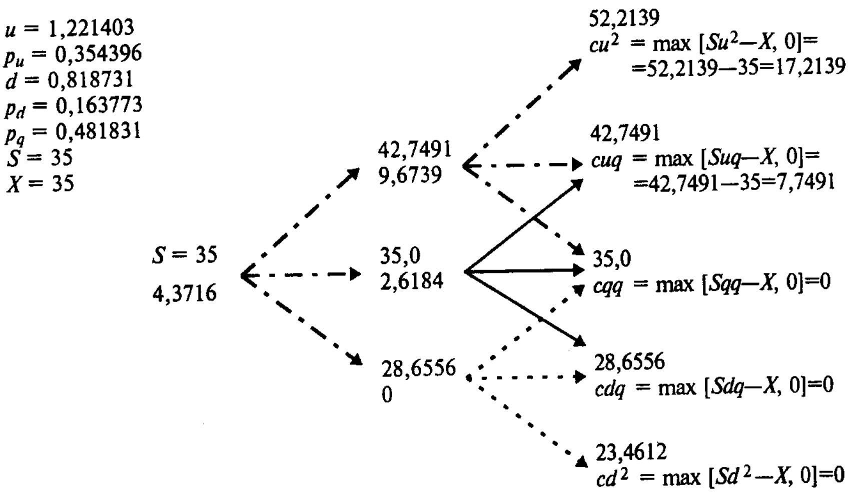

Рис. 8.11. Триномиальная модель ценообразования опциона

И в заключение подстановка в уравнение (8.37) значений стоимостей опциона в узлах *cu*, *cq* и *cd* даст нам стоимость опциона в настоящий временной период. Это значение составит 4,3716. Таким образом, мы приходим к такому же результату только после двух триномиальных испытаний по сравнению с четырьмя испытаниями в биномиальной модели. В общем можно сказать, что экономия усилий заключается в том, что триномиальная модель, представленная выше, может привести к результату (стоимости опциона) с такой же степенью точности, что и биномиальная модель, но в два раза быстрее.

МЕТОД МОНТЕ-КАРЛО

Метод Монте-Карло — это численный метод, позволяющий моделировать будущие значения переменной с помощью имитации ее поведения времени. Хотя множество изящных математических приемов было разработано для стохастических процессов переменных, возможно, что простые задачи могут привести к сложным математическим расчетам или возникшие задачи нецелесообразно решать с помощью аналитических методов. Доступные современным аналитикам возрастающие компьютерные возможности позволяют применять моделирование Монте-Карло к решению многих финансовых задач.

Функция вероятности дискретной случайной переменной (или функция плотности вероятности для непрерывных случайных величин) предоставляет информацию о вероятности для переменной принять определенное значение (или в случае непрерывного процесса — информацию о вероятности нахождения в определенном промежутке). Даже если событие, для которого происходит моделирование, произойдет всего один раз, появляется осознание того, что **если бы** оно было повторено много раз, случайная переменная приняла бы значения, соразмерные с этими вероятностями.

Моделирование процесса заключается просто в имитации поведения основной (входной) случайной переменной и через систему зависимостей получения выходной переменной, которая нас и интересует. Но необходимым элементом здесь является **повторение**. Многократным повторением процесса мы получаем распределение выходной величины или величин, исходя из которого можно построить распределение вероятностей. И именно для получения многократного повторения нужны существующие компьютерные мощности.

Пять этапов метода Монте-Карло

Процесс моделирования Монте-Карло может быть разделен на следующие пять этапов:

1. **Определение стохастической природы** входной переменной. Это позволит выбрать распределение вероятностей, необходимое для осуществления моделирования.
2. **Имитация движения входных переменных** с помощью многократного генерирования случайных чисел, корректируемых с таким расчетом, чтобы иметь такое же распределение вероятностей, как и основная переменная. Это подразумевает преобразование случайных чисел с равномерным распределением, сгенерированных компьютером, в случайные переменные с таким же распределением, что и переменные, предназначенные для моделирования. Скорректированные случайные переменные являются входными переменными.

3. **Осуществление моделирования** — объединение входных переменных вместе в соответствии с логикой системы, описывающей, каким образом связаны входные переменные и как получаются выходные. С помощью многократного генерирования случайных чисел мы получаем будущее значение искомой переменной.
4. **Многократное повторение этого процесса** (возможно несколько тысяч?) позволяет найти среднюю полученных значений. Эта средняя — будущее (ожидаемое) значение моделируемой переменной. Затем для определения настоящей стоимости моделируемой переменной это будущее значение дисконтируется по соответствующей дисконтной ставке.
5. **Применение техники контроля разбросанности** или других методов сокращения дисперсии для повышения точности результатов, полученных в ходе моделирования.

Этап 1: определение распределения вероятностей

Первый этап любого моделирования методом Монте-Карло — определение распределения(ий) вероятностей для входной(ых) переменной(ых). Большинство компьютерных программ, предназначенных для моделирования методом Монте-Карло, содержат встроенную библиотеку распределений вероятностей. Они также имеют возможность построения распределения вероятностей, основанного на суждениях самого исследователя, поскольку современные компьютеры имеют встроенные генераторы случайных чисел (в действительности генераторы псевдослучайных чисел), которые позволяют получать равновероятные числа между 0 и 1. Таким образом получение числа в диапазоне от 0,1 до 0,2 имеет такую же вероятность, что и получение числа между 0,7 и 0,8; или любое число в интервале от 0,3 до 0,5 имеет такую же вероятность, что и число из интервала 0,8—1,0.

Закончив генерирование равновероятных случайных чисел, мы должны перейти к процессу преобразования их в случайные числа, характеризующиеся распределением вероятностей нашего выбора.

Важным является сам механизм преобразования однородно распределенных случайных переменных в случайные переменные, имеющие распределение, схожее с эмпирическим. Для этого нам потребуется **распределение относительных частот** нашей случайной переменной.

В качестве примера возьмем следующее распределение относительных частот доходов по активу. В этом случае высота столбца, где доходность составила данное значение, соответствует проценту от всех наблюдений. Также сумма всех процентов для каждого типа наблюдения должна составлять 100 или 1 в зависимости от представления частностей в процентах или долях единицы (рис. 8.12). Это распределение относительных частот относится к дневным доходам по индексу FTSE 100 с 1984 по 1992 гг. Используя это распределение, мы можем построить кумулятивное распределение относительных частот, изображенное на рис. 8.13.

Заметим, что в случае кумулятивного распределения относительных частот вертикальная ось размечена в интервале до единицы, представляя тем самым эмпирическую вероятность, относящуюся к соответствующему значению, расположенному на горизонтальной оси. На горизонтальной оси же в свою очередь отмечен уровень дохода.

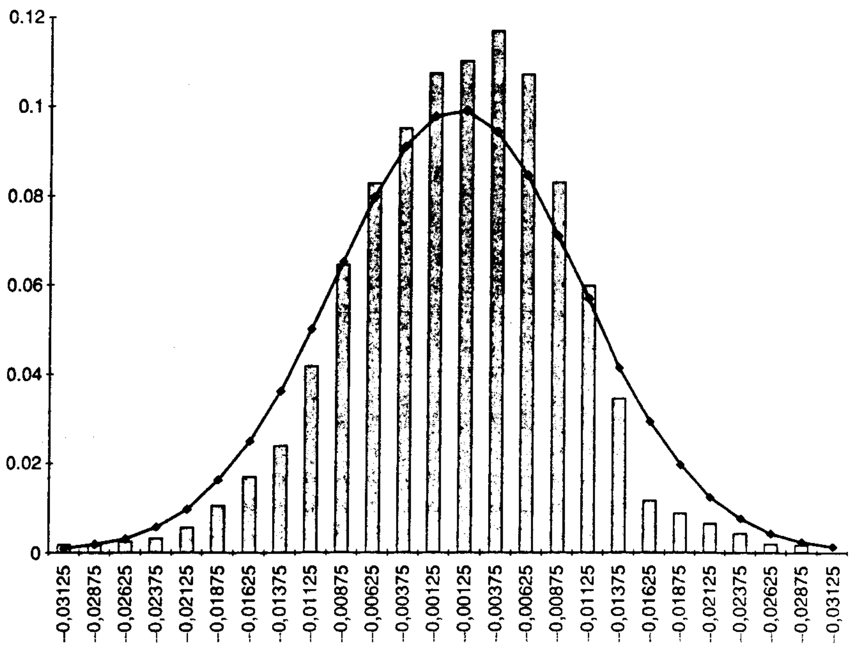

Рис. 8.12

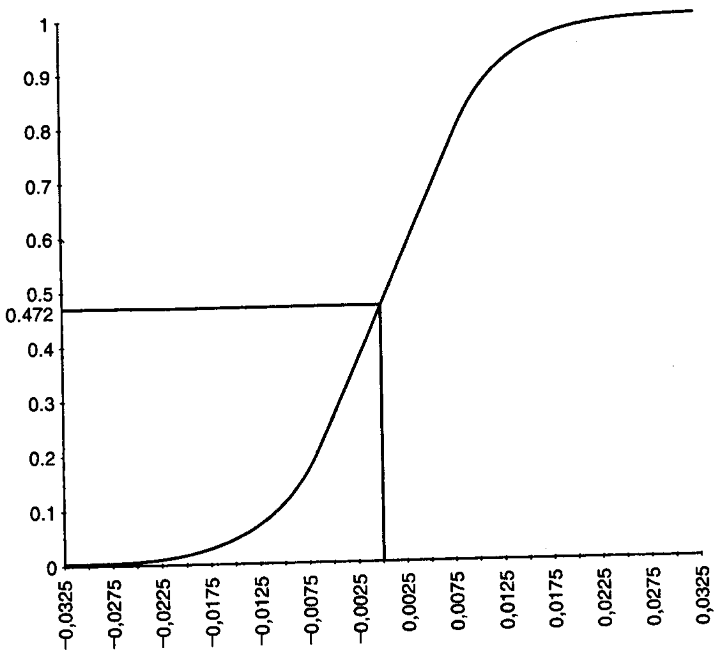

Рис. 8.13

Этап 2: имитация входных переменных

Второй этап заключается в моделировании поведения входных случайных переменных. Для этого необходимо сгенерировать достаточно большое количество равномерно распределенных случайных чисел в интервале от 0 до 1. Затем каждое случайное число откладывается по вертикальной оси кумулятивной плотности, а соответствующее значение случайной переменной находится по горизонтальной оси. Эти значения являются **входными для данной модели**. Так как кумулятивное распределение относительных частот отражает и величину основной переменной (доход в нашем случае), и эмпирическое распределение вероятностей, значения на горизонтальной оси представляют собой случайные наблюдения основной переменной с эмпирическим распределением вероятностей, которое сходно с распределением подлинных данных.

Для иллюстрации этого представим, что сгенерировано случайное число 0,472, располагаемое на вертикальной оси. Двигаясь от него горизонтально по направлению к функции кумулятивной плотности, а от точки пересечения с ней — перпендикулярно к оси абсцисс, получим доход, выбранный случайным образом. Если наши данные относятся к дневным доходам по активу, результатом является случайно сгенерированный однодневный доход, определенный исходя из распределения вероятностей, которое сходно с первоначальным эмпирическим распределением. Конечно, благодаря современным компьютерным программам подобные процедуры, выполняемые "вручную", заменяются компьютеризированными процессами, но в сущности они аналогичны описанным.

Интересно заметить, что числа из генератора случайных чисел, находящиеся в интервале от 0,4 до 0,6, дают результат в диапазоне от —0,25 до 0,25, тогда как случайные числа от 0,7 до 0,9 дают результат от 0,52 до 1,28. Таким образом, крутизна огивы в середине — причина более плотного расположения результатов, тогда как пологость граничных ветвей кривой — причина более разбросанных результатов (в точности то, что и описывает функция вероятности).

Некоторые читатели могут спросить, почему мы прибегаем к такому процессу, когда, используя допущение нормальности, можно получить такой же результат. Ответить можно так. Использование нормального распределения обоснованно только в том случае, если доходы независимы. Однако, если существует корреляция в рядах данных доходов, процедура не приведет к обоснованным результатам. Существует же свидетельство того, что корреляция в рядах валют, процентных ставок и индексов ценных бумаг присутствует.

Этап 3: моделирование основной переменной

Третий этап — объединение входных значений в соответствии с логикой системы. Система описывает, каким образом каждое индивидуальное входное значение участвует в получении результата на выходе, которым в нашем примере является оценка будущего значения индекса FTSE 100. Сейчас мы прервемся на некоторое время, чтобы определить, что мы вообще понимаем под моделированием. В контексте методологии Монте-Карло мы получаем результат всего **моделирования** (simulation) (которое состоит из серии испытаний) в качестве средней выходных значений по каждому отдельному **испытанию, ходу** или **проигрыванию** модели (simulation run).

Для того чтобы это понять, предположим, что мы хотим промоделировать будущую цену актива S_T. Результат или оценку, полученную в ходе моделирования этого актива, обозначим $\hat{S}_T$ — средней множества испытаний, которые дают индивидуальные оценки S_T. Обозначим эти индивидуальные оценки $\hat{S}_j$. Предположим, что мы располагаем данными об эмпирическом распределении непрерывно наращенных доходов по нашему активу. Обозначим индивидуальные наблюдения дохода через r. Тогда оценка $\hat{S}_j$ после одного периода будет:

$$\hat{S}_j = S_0 e^{r_1}. \tag{8.39}$$

Если $\hat{S}_j$ — оценка для n периодов, мы сначала получаем n последовательных наблюдений за доходом, сгенерированных случайным образом, а затем определяем $\hat{S}_j$ как:

$$\hat{S}_j = S_0 e^{r_1} \cdot e^{r_2} + \ldots + e^{r_n}. \tag{8.40}$$

Это то же самое, что и

$$\hat{S}_j = S_0 e^{(r_1 + r_2 + \ldots + r_n)}. \tag{8.41}$$

Этот процесс представляет собой одно **испытание**.

В испытании, которое только что было описано, используется новая случайная переменная для каждого подпериода всего времени, охватываемого исследованием. Следовательно, использование данных дневных доходов, например, случайной траектории, сгенерированной на базе 250 случайных наблюдений (так как один год содержит 250 торговых дней), даст нам смоделированное одногодичное значение стоимости.

Этот процесс генерирования случайных траекторий важен с той точки зрения, что будущее значение моделируемой переменной зависит не только от конечного значения, но и от того, как оно было получено. Примером могут служить опционы, которые основаны как на средней ставке основного актива, так и на минимальном и максимальном ее значениях.

Этап 4: получение распределения будущей стоимости актива

Этап заключается в многократном повторении процесса, описанного в этапе 3, для того чтобы получить распределение будущей стоимости актива. Весь процесс повторного осуществления (скажем, десять тысяч раз) **испытаний**, или **проходов** называется **моделированием**. Средняя будущих значений, полученная в этих испытаниях, является **смоделированным будущим значением** случайной переменной. **Смоделированное текущее значение** находится путем **дисконтирования** смоделированного будущего значения по соответствующей дисконтной ставке.

Тогда конечное смоделированное значение случайной переменной, т.е. средняя значений по каждому испытанию, определяется так:

$$\hat{S}_T = \frac{1}{n}\sum_{j=1}^{n}\hat{S}_j , \tag{8.42}$$

где $\hat{S}_T$ — конечное смоделированное значение случайной переменной, а $\hat{S}_j$ — результат на выходе в каждом испытании.

Для определения дисперсии $\hat{S}_T$ сначала необходимо оценить дисперсию $\hat{S}_j$ следующим образом:

$$\mathrm{D}_{\hat{S}_j} = \frac{1}{n}\sum_{j=1}^{n}\left[\hat{S}_j - \hat{S}_T\right]^2 . \tag{8.43}$$

Дисперсия $\hat{S}_T$ находится делением дисперсии $\hat{S}_j$ на *n*.

Тогда стандартная ошибка определяется как

$$\mathrm{SE}_{\hat{S}_T} = \sqrt{\frac{\mathrm{D}_{\hat{S}_j}}{n}} . \tag{8.44}$$

Для дополнительных пояснений можно обратиться к приложению 5.1 гл. 5.

Этап 5: применение методов по сокращению дисперсии для увеличения точности

Среднее квадратическое отклонение оценки S_T ($\hat{S}$) определяется следующим образом:

$$\sigma_{\hat{S}} = \sqrt{\frac{D(S_T)}{n}}. \qquad (8.45)$$

Из этого следует, что для того чтобы уменьшить среднее квадратическое отклонение в десять раз, необходимо увеличить количество испытаний в 100 раз. Альтернативным же подходом (чтобы сократить такое большое количество испытаний) является применение методов или техник по уменьшению дисперсии. Таких методов было разработано несколько, мы же обсудим два:

- антитетический метод случайной величины,
- метод контроля случайной величины.

Антитетический метод случайной величины

Метод заключается в следующем. Каждый раз, когда генерируется случайная переменная r, рассчитывается дополняющее ее значение $(1-r)$, которое используется для осуществления параллельного испытания. Таким образом, если входная переменная довольно велика по величине, в параллельном испытании входная переменная будет довольно мала. Как правило, это ведет к значениям по каждому испытанию, которые отрицательно коррелируемы.

Для оценки средней будем использовать выражение $S = (S_1 + S_2)/2$, где S_1 и S_2 — результаты параллельных испытаний. Дисперсия S тогда будет $(D(S_1) + D(S_2) + 2\mathrm{cov}(S_1 + S_2))/4$. При отсутствии отрицательной корреляции $D(S) = (D(S_1) + D(S_2))/4$. Но в случае отрицательной корреляции S_1 и S_2 будут иметь отрицательную ковариацию, которая уменьшит дисперсию.

Метод контроля случайной величины

Идея, лежащая в основе метода контроля случайной величины, заключается в нахождении переменной, сходной с моделируемой переменной, значение которой известно. Обозначим эту переменную h.

Следующий шаг — использование такого же метода генерирования случайных чисел для моделирования h, что и для моделирования S_T. Предположим, что полученный результат $\hat{h}$. Таким образом, одно и то же случайное число используется для генерирования выборки S и выборки h. Хотя для получения обеих оценок используется одно и то же случайное число, они не будут одинаковыми вследствие различий в процессах получения S_T и h.

Новая оценка $\hat{S}$, обозначенная S^*, рассчитывается так:

$$S^* = h + (\hat{S} - \hat{h}). \qquad (8.46)$$

Дисперсия S^* равна:

$$D(S^*) = D(\hat{S}) + D(\hat{h}) - 2\operatorname{cov}(\hat{S}, \hat{h}). \qquad (8.47)$$

Среднее квадратическое отклонение тогда составит:

$$D(S^*) = \sqrt{D(\hat{S}) + D(\hat{h}) - 2\operatorname{cov}(\hat{S}, \hat{h})}\ , \qquad (8.48)$$

что будет меньше, чем $D(S^*)$, если

$$p_{\hat{S},\hat{h}} > \sqrt{\frac{D(\hat{h})}{2D(\hat{S})}}. \qquad (8.49)$$

Таким образом, возможности уменьшения дисперсии с помощью метода контроля случайной величины зависят от того, удастся ли найти контрольную случайную переменную, которая была бы высококоррелируема с моделируемой переменной и в то же время имела сходное распределение вероятностей.

Применение метода Монте-Карло к ценообразованию опционов

Проиллюстрируем использование метода Монте-Карло для определения стоимости одногодичного опциона на покупку актива, имеющего распределение дохода, изображенное на рис. 8.12. Текущая цена актива равна 1000 единиц, цена исполнения опциона также составляет 1000 единиц, а безрисковая процентная ставка равна 6% годовых (непрерывно наращенная). Мы используем технику моделирования, поскольку, как видно из рис. 8.12, эмпирическое распределение вероятностей не является нормальным.

На первом этапе необходимо определить распределение. Используя наши данные, мы найдем, что средняя дневного дохода равна 0,000455, а среднее квадратическое отклонение дневного дохода — 0,0100694. Необходимо преобразовать равномерно распределенную случайную переменную в другую случайную переменную с распределением, идентичным эмпирическому распределению актива. Результатом будет серия случайных наблюдений за дневным доходом.

В действительности мы не используем наблюдаемую среднюю дневную доходность *r*, мы осуществляем корректировку. Вспомним, что в биномиальной модели ценообразования опционов опцион был оценен в рамках нейтральности к риску, так как было допущено, что опционная позиция может быть идеально захеджирована. То же самое мы допускаем и в процессе Монте-Карло. Вследствие этого соответствующая непрерывно наращенная ставка дохода будет однодневным эквивалентом безрисковой ставки, относящейся к сроку действия опциона. Предположим, что ставка равна 6% годовых, поэтому следует скорректировать дневную непрерывно наращенную ставку. Для этого мы должны вспомнить, что нормальное распределение со средней μ и средним квадратическим отклонением σ может быть трансформировано в логнормальное распределение со средней $e^{\mu+\sigma^2/2}$. Следовательно, для того чтобы получить годовую ставку 6%, мы должны скорректировать дневной непрерывно наращенный доход *r* так, чтобы $e^{\mu+\sigma^2/2} = e^{0{,}06/250}$. Значит, $r + 0{,}0100694^2/2 = 0{,}06/250$, что дает $r = 0{,}000189$.

В результате текущая цена актива наращивается в соответствии с дневным эквивалентом 6% годовых, а распределение вероятностей сохраняет свою форму. В сущности эмпирическое распределение смещено влево, его форма не изменяется, а средняя становится меньше, что иллюстрирует рис. 8.14.

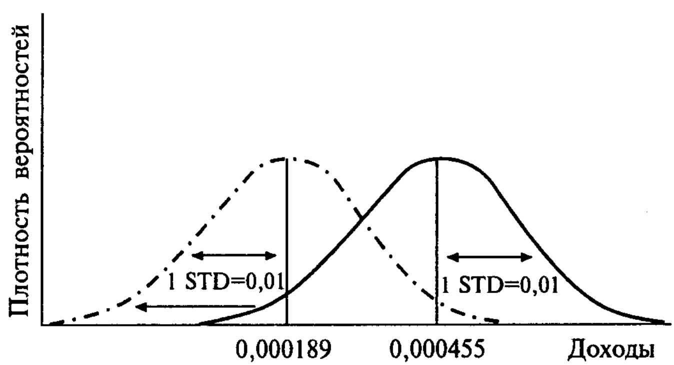

Рис. 8.14. Преобразование распределения дохода

Второй этап заключается в наращении текущей цены актива по случайной дневной ставке дохода для каждого дня торговли в течение срока действия опциона. В нашем примере мы допустили, что один год содержит 250 торговых дней. Так как эмпирическое распределение вероятностей относится к непрерывно наращенному доходу, цена актива наращивается следующим образом:

$$S_T = S_0 e^{r_1} \cdot e^{r_2} \cdot \ldots \cdot e^{r_{250}}.$$

Ранее мы заметили, что это то же самое, что и

$$S_T = S_0 e^{(r_1 + r_2 + \ldots + r_{250})}, \tag{8.50}$$

где r_n — случайное наблюдение однодневной ставки непрерывного наращения дохода, полученное согласно такому же эмпирическому распределению вероятностей, что и для данных основного актива.

Общий эффект этих 250 случайных наблюдений — это одно испытание. Мы должны повторить его много раз, чтобы снизить изменчивость нашей средней и привести в соответствие с требованиями точности. В этом примере мы решили, что для достижения достаточной точности средняя должна иметь ошибку менее 0,50. Годовая волатильность (изменчивость) основной переменной составляет 15,9%. Следовательно, годовое среднее квадратическое отклонение для актива с ценой 1000 единиц равно 159. Поэтому стандартная ошибка полученной средней не будет превышать $159/\sqrt{n}$, где n — количество испытаний (чем выше цена исполнения, тем ниже фактическая стандартная ошибка). Для того чтобы снизить ее до 0,50, n должно быть равно 125000.

Заметим, что правило определения количества испытаний может быть выражено как:

$$n = \left(\frac{S \cdot \sigma}{se}\right)^2,$$

где se — требуемая стандартная ошибка.

Таким образом, процесс был повторен 125000 раз, предоставляя 125000 альтернативных значений S_T.

Затем мы должны применить к опционам те же ограничивающие условия, что и в биномиальной модели, т.е. стоимость опциона на покупку во время T будет $C_T = \max[S_T - X, 0]$. Эти ограничивающие условия применяются к каждому значению S_T, а затем находится средняя всех 125000 значений C_T, которая при дисконтировании дает стоимость опциона.

Результатом наших 125000 испытаний модели является цена опциона 93,52. Биномиальная модель, состоящая из 100 испытаний, или модель Блэка—Сколса, которая будет обсуждена в гл. 10, дают значение 93,03. Как видно из рис. 8.12, наш моделируемый опцион может иметь меньшую цену, чем та, которая была получена из уравнения Блэка—Сколса: эмпирическое распределение немного смещено вправо, а также немного приподнято над значением средней. Так как мы определяем стоимость опциона на покупку, это будет иметь эффект небольшого уменьшения стоимости.

Выбор 125000 испытаний был обусловлен расчетом стандартной ошибки будущей цены актива. Однако требуемая стандартная ошибка цены опциона будет меньше, чем стандартная ошибка цены актива вследствие применения ограничивающих условий к опциону (max[S—X,0]). В результате это приведет к более узкому размаху конечных величин. Среднее квадратическое отклонение 125000 цен опциона было 125, тогда

$$\mathrm{SE}_{S_T} = \frac{125}{\sqrt{125000}} = \frac{125}{353{,}55} = 0{,}35\,.$$

Мы можем сказать с уверенностью 95%, что реальная цена опциона составляет 93,52 плюс/минус 0,7 (т.е. две стандартные ошибки).

УПРАЖНЕНИЯ

1. Облигация имеет срок погашения 2,5 года. По ней выплачиваются полугодовые купоны в 5 единиц. Облигация будет выкуплена по цене 100 единиц. Ее текущая цена составляет 102 единицы.

   Вычислите полный доход при погашении (gross redemption yield — GRY) для этой облигации, используя:

   а) метод деления пополам;

   б) метод Ньютона—Рафсона.

   Сравните и сопоставьте эти два подхода.

2. Случайная величина X распределена как N (6,9). Следовательно, функция плотности вероятности для X:

$$f(x) = \frac{1}{3\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-6}{3}\right)^2}\,.$$

Требуется оценить $p(8 \le X \le 10)$ с помощью:

а) статистических таблиц;
б) правила трапеций для четырех интервалов;
в) правила Симпсона для четырех интервалов;
г) нахождения функции в виде многочлена.

3. Используя однопериодную биномиальную модель и нижеприведенные данные, убедитесь, что справедливая стоимость опциона на покупку равна 7,95.

   Цена актива равна 75 единицам, цена исполнения — 50 единицам, процентная ставка — 10% годовых, срок действия 1 год. Цена актива будет двигаться вверх и вниз на 25% в течение одного года.

4. Убедитесь, что коэффициент хеджирования H в примере выше равен 2.

5. Исходя из данных, приведенных ниже, постройте шестипериодную биномиальную модель и убедитесь, что справедливая цена опциона на покупку составляет 4,52 единицы.

   Данные: цена актива — 35 единиц, цена исполнения — 35 единиц, волатильность — 20%, краткосрочная процентная ставка — 10% годовых, время до исполнения — один год.

6. Используя данные из вопроса 3, разработайте трехступенчатую триномиальную модель и убедитесь, что справедливая цена опциона равна также 4,52 единиц.

7. Кратко опишите пять этапов процесса моделирования методом Монте-Карло в приложении к ценообразованию опционов на актив, имеющий логнормальное распределение.

8. Опишите антитетический метод случайной величины и метод контроля случайной величины как техники по сокращению дисперсии. Объясните преимущества и недостатки каждого из них при использовании метода Монте-Карло к ценообразованию опционов.

ОТВЕТЫ К ИЗБРАННЫМ ВОПРОСАМ

1. GRY определяется выражением x^2-1, где x — решение следующего уравнения:

$$102x^5-5x^4-5x^3-5x^2-5x-105 = 0.$$

а) результат применения метода деления пополам:

$$(f(x) = 102x^5-5x^4-5x^3-5x^2-5x-105 = 0)$$

L	R	(L + R)/2	f[(L + R)/2)
1,01	1,505	1,2575	179,091
1,01	1,2575		58,424
1,01	1,1338	1,071	15,457
1,01	1,0719	1,040	–2,473
1,0409	1,0719	1,0564	6,217
1,0409	1,0564	1,0487	1,805
1,0409	1,0487	1,0448	–0,351
1,0448	1,0487	1,0467	0,723
1,0448	1,0467	1,0458	0,185
1,0448	1,0458	1,0453	–0,083
1,0453	1,0458	1,0455	0,051
1,0453	1,0455	1,0454	–0,016
1,0454	1,0454		

Корень уравнения равен 1,045, тогда GRY = 9,2%.

б) применение процедуры Ньютона—Рафсона дает:

$x_1 = 1,$
$x_2 = 1{,}05,$
$x_3 = 1{,}045479935,$
$x_4 = 1{,}045438601,$
$x_5 = 1{,}045438598,$
$x_6 = 1{,}045438598.$

$GRY = 9{,}3\%.$

2. а) 0,1613.
   б) и в)

x	8	8,5	9	9,5	10
$f(x)$	0,10648267	0,09397063	0,08065691	0,0673329	0,05467002

б) правило трапеций: 0,16126389,
в) правило Симпсона: 0,16128010,
г) использование многочлена: 0,908780897—0,74750848=0,1612848.

СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ

Black, F. and Scholes, M.J. (1973) The pricing of options and corporate liabilities. *Journal of Political Economy*, **81**, 637—59.

Boyle, P.P. (1977) Options: a Monte-Carlo approach. *Journal of Financial Economics*, **4**, 323—38.

Cox, J.C., Ross, S.A. and Rubinstein, M. (1979) Option pricing: a simplified approach. *Journal of Financial Economics*, **7**, 229—63.

Rendelman, R. and Barter, B. (1979) Two state option pricing. *Journal of Finance*, **34**, 1093—110.

Wilmot, P., Dewynne, J. and Howison, S. (1993) *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press, Oxford.

Wilmot, P., Howison, S. and Dewynne, J. (1995) *The Mathematics of Financial Derivatives: A Student Introduction*. Cambridge University Press, Cambridge.

Winston, W.J. (1996) *Simulation Modeling Using @ Risk*. Wadsworth.

9

ОПТИМИЗАЦИЯ

ВВЕДЕНИЕ

В этой главе мы рассматриваем оптимизацию, особое внимание уделяем определению структуры оптимальных портфелей, где оптимальный определяется как имеющий минимальную дисперсию при заданном уровне дохода.

Из гл. 3 мы узнали, что можно использовать математический анализ, чтобы найти оптимальные точки (экстремумы), которые являются либо точками минимума, либо точками максимума функции. Мы также узнали, что, используя множители Лагранжа, можно найти экстремумы при соответствующих ограничениях. Нахождение экстремума любой функции эквивалентно (если функция дифференцируема, а также при небольших трудностях, таких, как стационарные точки перегиба) решению уравнения $f'(x) = 0$. Наша оптимизационная задача состоит в том, чтобы найти портфель с минимальной дисперсией, где дисперсия портфеля — это функция от ковариаций и весов активов. Однако мы также имеем ограничение, касающееся достижения какого-то минимального уровня дохода по портфелю.

В этой главе мы расширим анализ экстремумов до многофакторного примера, а также применим критерии, известные как условия Кюна—Такера, для того чтобы применить множители Лагранжа в том случае, когда ограничения выражены в виде неравенств, а не уравнений.

Таким образом, последовательность изучения оптимизации (и структуры главы) следующая:

- изучение линейного программирования;
- объяснение задачи выбора оптимальных портфелей рискованных активов;
- повторение темы множителей Лагранжа и их приложения к многофакторному примеру;
- объяснение и применение условий Кюна—Такера;
- описание методологии Данцига—Вольфа, позволяющей использовать методы линейного программирования для решения задачи квадратического программирования.

Начнем с определения некоторых терминов, используемых в этой главе.

Определения

Целевая функция. Целевая функция определяет задачу, которая должна быть решена в процессе оптимизации. Например, в этой главе мы занимаемся **минимизацией риска** портфеля активов. Типичной целевой функцией для портфеля рискованных активов будет

$$\text{минимизировать } Z = \sum_{i=1}^{N}\sum_{j=1}^{N} W_i W_j \left[\text{cov}_{ij}\right], \qquad (9.1)$$

где Z — общий риск, W_i — веса активов в портфеле.

Из гл. 2 вспомним, что риск портфеля активов представляет собой функцию дисперсий и ковариаций активов и что дисперсия отдельного актива равна его ковариации с самим собой. Вспомним также, что дисперсия рассчитывается по отклонениям в квадрате. Таким образом, целевая функция нелинейна, это **квадратическая функция**.

Далее увидим, что некоторые портфельные задачи имеют линейные целевые функции.

Ограничения. Целевая функция обычно задается с учетом набора ограничений. Например, какие-то средства необходимо инвестировать в каждый из активов. Может быть ограничение, что все средства должны быть полностью инвестированы. Другое ограничение может заключаться в том, что минимальный риск должен быть достигнут при условии получения какого-то минимального уровня дохода. Эти три ограничения можно отобразить как

$$\begin{aligned} &1)\ W_i > 0; \\ &2)\ \sum_{i=1}^{N} W_i = 1; \\ &3)\ R \geq 0{,}10. \end{aligned} \tag{9.2}$$

Ограничение 1 свидетельствует о том, что все активы должны иметь вес в портфеле, больший нуля. Его формой является **ограничение в виде неравенства**. Ограничение 2 означает, что сумма весов равна единице. Это **ограничение в виде равенства**. Оно просто означает, что все средства должны быть инвестированы в рискованные активы. Ограничение 3 говорит, что по портфелю должна быть получена ставка дохода, которая равна или больше чем 10%. Это еще одно ограничение в виде неравенства.

Могут быть добавлены дальнейшие ограничения, подобные двум указанным ниже. Они также являются ограничениями в виде неравенств и просто утверждают, что вес актива j в портфеле не должен превышать 10%, в то время как вес актива k не должен превышать 15%.

$$\begin{aligned} &4)\ W_j \leq 0{,}10; \\ &5)\ W_k \leq 0{,}15. \end{aligned} \tag{9.3}$$

Эти примеры не являются всеобъемлющими, однако они показывают две формы ограничений — **в виде равенств** и **в виде неравенств**.

Задача математического программирования — это задача, в которой функцию многих переменных (целевую функцию) необходимо оптимизировать при наборе ограничений. Число ограничений, как правило, меньше, чем (обычно значительно меньше) число переменных. Оптимальный вариант устанавливается в условиях максимизации или минимизации.

Задача линейного программирования — это задача, в которой целевая функция и ограничения линейны.

В задаче квадратического программирования целевая функция является квадратической функцией переменных, т.е. значения некоторых из переменных находятся во второй степени, однако ограничения остаются линейными.

ЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ

Ограниченные оптимизационные задачи, в которых целевая функция имеет вид линейной функции переменных и в которых функции ограничений также линейны, известны как задачи линейного программирования.

Задачи линейного программирования для двух переменных могут быть решены с помощью построения графиков. В 1940-х годах Данциг разработал алгоритм, называемый **симплексным алгоритмом**, эффективно преобразующий графический подход в алгебраический метод, который может быть использован для компьютерного приложения и позволяет обрабатывать любое число переменных. Симплексный алгоритм — это **итерационный процесс** нахождения оптимального значения (экстремума) целевой функции.

Приложением линейного программирования (ЛП) к выбору портфеля является построение портфелей в рамках *модели ценообразования активов капитала* (Capital Asset Pricing Model — CAPM). Детализированный анализ этой модели можно найти в книге Уотсхэма (1993). Здесь же дадим краткое его описание.

Модель выражает ожидаемый доход по активу в виде линейной функции от безрисковой ставки дохода, ожидаемого дохода по рыночному портфелю и уровня систематического риска, присущего активу. Ожидаемый доход по активу i определяется как $E(r_i) = r_f + \beta_i(E(r_m) - r_f)$, где $(E(r_m) - r_f)$ — ожидаемая премия за риск по активу i, а β_i — мера систематического риска данного актива.

Когда мы комбинируем активы в портфель, доходы по каждому активу складываются в линейной форме, и риск портфеля, представляемый β портфеля, тоже является линейной комбинацией. В данном случае это средняя взвешенная из β_i отдельных активов.

Для иллюстрации этого предположим, что мы желаем скомбинировать два актива в портфель с пропорциями W_a и W_b ($W_a + W_b = 1$). Допустим, что ожидаемые доходы по этим активам обозначаются $E(r_a)$ и $E(r_b)$ соответственно и что β активов — β_a и β_b.

Тогда ожидаемый доход по портфелю будет равен $W_aE(r_a) + W_bE(r_b)$, т.е. доходы складываются линейно. То же самое отметили в гл. 2, когда рассматривали доходы по портфелю рискованных активов с использованием анализа средних и дисперсий.

Подобно этому, так как в модели CAPM

$$E(r_a) = r_f + \beta_a(E(r_m)-r_f) \tag{9.4}$$

и

$$E(r_b) = r_f + \beta_b(E(r_m)-r_f), \tag{9.5}$$

следует, что доход по портфелю определяется через

$$W_a[r_f + \beta_a(E(r_m)-r_f)] + W_b[r_f + \beta_b(E(r_m)-r_f)] \tag{9.6}$$

$$= r_f + [W_a\beta_a + W_b\beta_b](E(r_m)-r_f) = r_f + \beta(E(r_m)-r_f), \tag{9.7}$$

где $\beta = W_a\beta_a + \beta_b$.

Таким образом, β портфеля активов является средней взвешенной β отдельных активов. Следовательно, если цель процедуры оптимизации заключается в максимизации дохода по портфелю при ограничениях максимального размера β портфеля, перед нами ставится задача, где целевая функция, т.е. доход по портфелю, линейна и ограничения тоже линейны. Следовательно, мы имеем задачу линейного программирования.

Проиллюстрируем приложение линейного программирования к максимизации дохода по портфелю из трех активов при ограничении максимального уровня β портфеля.

Выбор портфеля из трех активов — использование линейного программирования для контроля систематического риска

Рассмотрим задачу построения портфеля с целевой функцией достижения максимального ожидаемого дохода при том ограничении, что β портфеля не должна быть выше 1,1. Допустим, что для выбора у нас есть три актива — *A*, *B* и *C*. Их ожидаемые доходы составляют 0,11, 0,15 и 0,08 соответственно. β для САРМ равны 1, 1,2 и 0,9 соответственно. Доли каждого из активов в портфеле обозначаются как W_a, W_b и W_c. Значения этих весов устанавливаются портфельным менеджером и являются переменными, которые могут корректироваться для достижения цели. Ожидаемые доходы и значения β различных активов зафиксированы с точки зрения портфельного менеджера, потому что они определяются рынком. Однако доходы и величина β портфеля могут формироваться портфельным менеджером посредством подбора доли каждого из активов в портфеле. Цель состоит в том, чтобы найти те комбинации весов, которые максимизируют целевую функцию при ограничениях.

Таким образом, задача заключается в определении оптимальных пропорций (весов) каждого из активов, которые приведут к максимальному ожидаемому доходу при условии данного максимального уровня риска (β). Эта задача может быть сформулирована математически следующим образом.

Целевая функция

Максимизировать (доход)

$$0{,}11W_a + 0{,}15W_b + 0{,}08W_c. \qquad (9.8)$$

Так как доход по каждому активу предопределен, только веса могут быть изменены для достижения целевой функции.

Ограничения

Мы заметили выше, что β портфеля не должна превышать 1,1 — это первое ограничение. В данном примере в каждый актив обязательно должны быть проведены положительные инвестиции — это второе ограничение. И все средства должны быть полностью инвестированы — это третье ограничение. Таким образом, целевая функция и ограничения имеют вид, указанный ниже.

Максимизировать (доход) при

1. $W_a + 1{,}2W_b + 0{,}9W_c \le 1{,}1$ (β портфеля не должна превышать 1,1) (9.9)
2. $0 \le W_a \le 1$

$0 \le W_b \le 1$ (все активы должны иметь неотрицательные веса) (9.10)

$0 \le W_c \le 1$

3. $W_a + W_b + W_c = 1$ (средства должны быть полностью инвестированы) (9.11)

Заметьте, что все ограничения линейны (т.е. нет величин во второй или более высоких степенях) и одновременно присутствуют ограничения в виде равенств и неравенств.

Мы продемонстрируем приложение линейного программирования, показав, как вышеозначенная задача решается сначала **графически**, а затем **симплексным методом**.

Графическое решение

Для иллюстрации графического решения мы используем третье ограничение, чтобы исключить W_c и тем самым получить задачу с двумя переменными. Это возможно, потому что в каждой портфельной задаче, где средства полностью инвестируются в портфель, если мы знаем n—1 весов, то можно определить n-ный вес. Так как все веса должны в сумме дать единицу, то n-ный вес равен

$$W_n = 1 - W_a - W_b - \ldots - W_{n-1}.$$

Заметьте, что этот переход от трех к двум переменным происходит только по причине облегчения графической демонстрации метода. Задача в том виде, в котором она была сформулирована, будет скорее всего передана ЛП — пакету в ее первоначальной форме — преобразования, сокращающее число переменных, обычно не проводятся.

Используя ограничение $W_a + W_b + W_c = 1$ для замены W_c на $1 - W_a - W_b$, переопределим целевую функцию:

максимизировать

$$0{,}11W_a + 0{,}15W_b + 0{,}08(1 - W_a - W_b) = 0{,}03W_a + 0{,}07W_b + 0{,}08 \qquad (9.12)$$

при

$$\begin{aligned} &1W_a + 1{,}2W_b + 0{,}9(1 - W_a - W_b) \le 1{,}1 \\ &0 \le W_a \le 1 \\ &0 \le W_b \le 1 \\ &0 \le (1 - W_a - W_b) \le 1. \end{aligned} \qquad (9.13)$$

Если считать вес актива C как остаток, мы можем максимизировать следующую целевую функцию:

$$\textbf{максимизировать } 0{,}03W_a + 0{,}07W_b. \quad (9.14)$$

Когда мы найдем оптимальные веса для A и B из этой функции, мы сможем вычесть сумму этих весов из целого и найти вес актива C.

Заметьте, что "+0,08" является константой. Как таковую нам не нужно включать ее в функцию для максимизации, потому что, если мы найдем значения W_a и W_b, которые максимизируют $0{,}03W_a + 0{,}07W_b$, то найдем и значения, максимизирующие $0{,}03W_a + 0{,}07W_b + 0{,}08$. Поэтому, как только мы максимизируем $0{,}03W_a + 0{,}07W_b$, то сразу же следует прибавить константу 0,08.

Так как мы заменили выражение для C в целевой функции, необходимо сделать то же самое и с ограничениями. Следовательно, ограничения становятся следующими

$$\begin{aligned} 0{,}1W_a + 0{,}3W_b &\le 0{,}2 \\ 0 \le W_a &\le 1 \\ 0 \le W_b &\le 1 \\ 0 \le W_a + W_b &\le 1. \end{aligned} \quad (9.15)$$

Заметьте, что как и в целевой функции, мы удалили величину β актива C из β активов A и B. Мы также устранили влияние β актива С на максимальное β портфеля.

При нахождении графического решения (рис. 9.1) мы стремимся определить "область возможных решений" между двумя осями, которыми в данном примере являются W_a и W_b.

По второму ограничению максимальная пропорция, которую мы можем инвестировать в A, составляет 100% или единицу. Это показано точкой N. Аналогично точка K отражает 100% портфеля, инвестированные в актив B. Так как и W_a, и W_b должны быть меньше или равны единице, вертикальная и горизонтальная линии проходят через соответствующие оси. Однако эти ограничения перекрываются тем, что требуется, чтобы $W_a + W_b$ было меньше или равно единице. Чтобы понять расположение этого ограничения, допустим, что 100% инвестировано в A и ничего в B, тогда ограничение будет находиться в точке N. Если бы 100% было инвестировано в B и ничего в A, ограничение располагалось бы в точке K. Точки между K и N представляют комбинации A и B, которые дают в сумме целое.

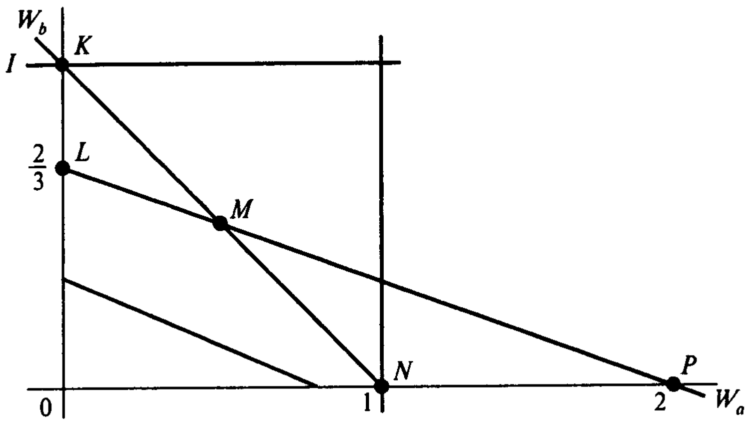

Рис. 9.1

И наконец, существует ограничение $0{,}1W_a + 0{,}3W_b \leq 0{,}2$. Для его иллюстрации мы проводим линию $0{,}1W_a + 0{,}3W_b = 0{,}2$. Она представляет границу между областью точек, удовлетворяющих $0{,}1W_a + 0{,}3W_b < 0{,}2$, и областью, определяемой $0{,}1W_a + 0{,}3W_b > 0{,}2$. Чтобы иметь правильное представление о расположении этой линии, предположим, что все инвестировано только в B и ничего в A. Мы должны были бы инвестировать лишь 2/3 наших денег, чтобы достичь дохода 0,2, потому что актив B приносит доход 0,3 на единицу ($2/3 \cdot 0{,}3 = 0{,}2$). Если мы инвестируем только в актив A, нам необходимо вложить 200% наших средств, потому что по активу A зарабатывается лишь 0,1 ($2 \cdot 0{,}1 = 0{,}2$).

Точки между K и M удовлетворяют ограничению $W_a + W_b = 1$, но не отвечают ограничению $0{,}1W_a + 0{,}3W_b \leq 0{,}2$. Так же точки между M и P соответствуют требованию $0{,}1W_a + 0{,}3W_b = 0{,}2$, но не удовлетворяют ограничению $W_a + W_b = 1$. Таким образом, область возможных решений в этом примере должна быть ограничена линией $0LMN$.

Теперь мы знаем, что оптимальная комбинация актива A и актива B лежит в пределах области $0LMN$. Однако нам нужно знать, какая точка из этой области является оптимальной. Для этого обратимся к целевой функции $0{,}03W_a + 0{,}07W_b$ и построим так называемые линии равной прибыли. Это прямые линии, отображающие целевую функцию, где каждая точка на линии представляет одинаковое значение целевой функции. Рассмотрим возможность приравнивания целевой функции, скажем, к 0,021. Целевая функция $0{,}03W_a + 0{,}07W_b = 0{,}021$ дает нам линию равной прибыли, состоящую из точек, представляющих портфели (веса W_a, W_b и, следовательно, W_c) при цели 0,021. Это отражает доход по портфелю в размере 0,101, потому что мы добавляем 0,08, относящиеся к активу C.

Данная линия проведена на диаграмме и можно увидеть, что некоторые точки на ней попадают в область возможных решений, показывая, что этот доход достижим.

Можно ли добиться большего? Например, можем ли мы достичь целевого значения 0,042 (т.е. доходности портфеля 12,2%, учитывая постоянную 0,08, которая снова прибавлена)?

В обоих случаях линии содержат точки в рамках области возможных решений, так что оба уровня доходов возможны. Можно этот подход далее развить, пока линия равной прибыли, удаляясь от начала координат, не будет касаться лишь точки на границе области возможных решений. Оптимальная комбинация будет представлена той точкой, на которой линия равной прибыли отстоит дальше всего от начала координат, но все еще находится в области возможных решений. В рассматриваемом примере это точка M.

Так как точка M лежит на обеих линиях ограничений, мы можем найти координаты этой точки, решив систему уравнений для двух ограничений

$$\begin{aligned} W_a + W_b &= 1 \\ 0{,}1W_a + 0{,}3W_b &= 0{,}2. \end{aligned} \tag{9.16}$$

Решим эту систему уравнений методом подстановки.

Сначала умножим второе уравнение на 10, тогда

$$\begin{aligned} W_a + W_b &= 1 \\ W_a + 3W_b &= 2. \end{aligned} \tag{9.17}$$

Вычитая первое из второго, получаем

$$2W_b = 1, \tag{9.18}$$

значит,

$$W_b = 0{,}5. \tag{9.19}$$

Подставляя это значение в уравнение (9.16), получаем $W_a = 0{,}5$.

Таким образом, мы получили $W_a = 0{,}5$ и $W_b = 0{,}5$, а целевое значение составляет 0,03×0,5 + 0,07×0,5 = 0,05, и, если мы добавим постоянную 0,08, то получим 0,13, или 13%.

В реальных условиях необязательно строить эти "линии равной прибыли". Тот факт, что они существуют и что они прямые и параллельны друг другу, показывает, что наилучшее решение всегда располагается в вершине (на пересечении двух линий) в области возможных решений. Следовательно, все, что требуется, — это проверить вершины, сужая таким образом задачу до дискретной.

Проверка вершин включает, как было сказано выше, решение систем уравнений. Для двух измерений вершины находятся на пересечении двух линий, и существуют два уравнения с двумя неизвестными. Для трех измерений вершины располагаются там, где пересекаются плоскости, давая три уравнения с тремя неизвестными.

Если бы у нас было 200 переменных, существовало бы 200 уравнений с 200 неизвестными. Задача стала бы чрезвычайно трудной. Следовательно, необходим эффективный алгоритм, организующий проверку как можно меньшего числа вершин. Для этого используется симплексный алгоритм. Компьютерные версии этого алгоритма очень эффективны и могут скрупулезно и эффективно решать задачи, включающие сотни переменных и ограничений.

Симплексный метод

Симплексный метод — это итерационный процесс, который начинается с одного "**предварительного решения**" и в поисках лучшего решения движется по границе **области возможных решений** до тех пор, пока не достигнет оптимального решения. Чтобы увидеть, как собственно работает симплексный метод, применим его для той же портфельной задачи с тремя активами.

Симплексный метод предполагает, что все переменные имеют неотрицательное значение. Это не препятствие, поскольку переменная, которая может принимать отрицательные значения, может быть заменена разностью двух неотрицательных переменных. Например, X может иметь значение —3 при замене ее на $X_1 - X_2$, где $X_1 = 0$ и $X_2 = +3$. Пакеты прикладных программ, включающие симплексный алгоритм, выполняют подобные действия автоматически.

Переопределим нашу целевую функцию и ограничения следующим образом:

максимизировать

$$Z = +\,0{,}03W_a + 0{,}07W_b \quad .$$

при

$$0{,}1W_a + 0{,}3W_b \le 0{,}2 \qquad (9.20)$$

$$W_a + W_b \le 1.$$

Первый шаг — преобразование неравенств в равенства с помощью дополнительных переменных, известных как **свободные переменные**. Так как эта задача содержит два неравенства, то нужно применить две свободные переменные. Поэтому, $0{,}1W_a + 0{,}3W_b \le 0{,}2$ преобразуется в выражение $0{,}1W_a + 0{,}3W_b + s_1 = 0{,}2$, где s_1 — свободная переменная. Она отражает уровень, до которого сумма $0{,}1W_a + 0{,}3W_b$ находится ниже 0,2. Аналогично, $W_a + W_b \le 1$ становится $W_a + W_b + s_2 = 1$.

Таким образом, ограничения в виде неравенств преобразуются в ограничения в виде равенств следующим образом:

1. $0{,}1W_a + 0{,}3W_b \le 0{,}2$ становится $0{,}1W_a + 0{,}3W_b + s_1 = 0{,}2$
2. $W_a + W_b \le 1$ становится $W_a + W_b + s_2 = 1$

Цель состоит в максимизации $0{,}03W_a + 0{,}07W_b$.

Заметьте, что мы опустили ограничения неотрицательности, так как предполагается, что для симплексного метода переменные могут иметь только неотрицательные значения.

Одна из возможных, хотя и не оптимальная точка, определена как $W_a = 0$ и $W_b = 0$, т.е. как начало координат, потому что если ничего не инвестировалось в A и B и мы задаем $s_1 = 0{,}2$ и $s_2 = 1$, удовлетворяются оба ограничения в виде равенств. Назовем это предварительным возможным решением. Функция симплексного метода заключается в том, чтобы искать оптимальное решение посредством повторяемого движения от одного возможного решения к другому, лучшему возможному решению.

Следует отметить, что часто нелегко определить первоначальное возможное решение. Если все ограничения в виде неравенств имеют форму "≤", первоначальным возможным решением служит начало координат. При наличии ограничения “≥”, это не будет верным. К счастью, путем искаженного обращения процедур, что нравится математикам, метод может быть направлен на самого себя, чтобы произвести свое собственное первоначальное возможное решение.

Таким образом, текущий вид задачи состоит в следующем:

максимизировать

$$Z = 0{,}03W_a + 0{,}07W_b$$

при

$$0{,}1W_a + 0{,}3W_b + s_1 = 0{,}2 \tag{9.21}$$

$$W_a + W_b + s_2 = 1.$$

При $W_a = 0$, $W_b = 0$, $s_1 = 0{,}2$ и $s_2 = 1$ решение равно $Z = 0$. Это первоначальное возможное решение. Далее метод продвигается через последовательность улучшающихся возможных решений до тех пор, пока не будет найдено лучшее. Для облегчения поиска лучшего решения представим исходную ситуацию в табличной форме

	W_a	W_b	s_1	s_2	ПС
ограничение 1	**0,1**	**0,3**	**1**	**0**	**0,2**
ограничение 2	**1**	**1**	**0**	**1**	**1**
целевая функция	**–0,03**	**–0,07**	**0**	**0**	**0**
	←	индикаторы		→	

Нижняя строка представляет собой целевую функцию. Мы установили цель в форме $Z - 0{,}03W_a - 0{,}07W_b + 0s_1 + 0s_2 = 0$. Первая и вторая строки показывают два ограничения, связанных с этой задачей. ПС представляет правую сторону неравенств.

Теперь мы должны дать определение **базисных и небазисных** переменных. **Базисные переменные** — это те, в столбцах которых содержатся только нули кроме одной единицы. Значит, в данном предварительном решении s_1 и s_2 являются базовыми переменными. Текущая величина этих переменных может быть найдена, если посмотреть на правый столбец напротив элемента “один” в соответствующих столбцах. Отсюда текущая величина s_1 составляет 0,2, а текущее значение s_2 — 1.

Остальные переменные известны как **небазисные переменные**. В этом процессе небазисные переменные имеют текущее значение ноль.

При каждой итерации (шаге) процесса поиска одна базисная переменная преобразуется в небазисную переменную. В то же время какая-либо небазисная переменная преобразуется в базисную переменную (что называется вхождением в решение).

В каждом шаге процедуры поиска лучшего решения мы должны убедиться в том, что каждое ограничение содержит только одну базисную переменную. Алгоритм также требует, чтобы целевая функция была выражена с помощью небазисных переменных. Наше предварительное возможное решение удовлетворяет этим условиям.

Улучшение решения

Первым шагом в этом процессе служит определение одного элемента в целевой строке, который имеет знак минус. Мы выберем —0,03 (возможен выбор между —0,03 и —0,07). Столбец чисел, где находится —0,03, известен как основной.

Вспомним, что в первоначальной форме целевая функция имеет вид $Z = 0{,}03W_a + 0{,}07W_b$. В таблице Z выражена через текущие небазисные переменные, т.е. в таблице это соотносится с $Z—0{,}03W_a—0{,}07W_b = 0$. Так как в предварительном решении эти небазисные переменные имели нулевые значения и их коэффициенты в таблице отрицательны, мы видим, что Z возрастет, если увеличится W_a или W_b, т.е. если мы сделаем какую-то из них базисной. Имея возможность выбора, мы решили выбрать увеличение W_a.

Второй шаг заключается в следующем: **используя только строки ограничений**, разделим каждый элемент с правой стороны на соответствующий элемент из основного столбца. Это называется проверкой соотношений. Найдем показатель с **минимальной неотрицательной величиной**. Соответствующий элемент из основного столбца определяет основную строку, а элемент, находящийся на пересечении основного столбца и основной строки, называется основой.

Проведя проверку соотношений, мы обнаружили, что единица во втором ограничении является основой. Мы должны преобразовать эту основу в единицу, деля всю основную строку на себя саму. Однако в данном случае в этом нет необходимости, потому что основа уже равна единице.

Следующий шаг — преобразование всех остальных элементов основного столбца в нули с помощью прибавления или вычитания соответствующих кратных количеств основной строки (заметьте: мы используем новые значения основной строки, т.е. после того как основа была приведена к единице). Таким образом, для преобразования 0,1 вверху основного столбца в ноль мы отнимаем 0,1 раза основную строку от каждого соответствующего элемента верхней строки. Чтобы трансформировать —0,03 в ноль, прибавим 0,03 раза основную строку к каждому элементу нижней строки. Получилась такая таблица:

	W_a	W_b	s_1	s_2	ПС
ограничение 1	**0**	**0,2**	**1**	**–0,1**	**0,1**
ограничение 2	**1**	**1**	**0**	**1**	**1**
целевая функция	**0**	**–0,04**	**0**	**0,03**	**0,03**

← индикаторы →

Для толкования этого решения заметьте, что переменные, являющиеся **базисными**, будут иметь значения, которые даны соответствующими элементами с правой стороны. Все значения других переменных будут равны нулю. Следовательно, целевая функция (Z) имеет величину 0,03 (без учета постоянной 0,08), W_a равна единице, s_1 равна 0,1, в то время как W_b и s_2 имеют нулевые значения.

На то, что поиск оптимального решения еще не завершен, указывают индикаторы, а именно — один индикатор со знаком "минус".

Поэтому повторим процесс. Во второй таблице мы видим, что основным является столбец, в котором нижний элемент равен —0,04. Проверяя соотношения, мы находим основу, равную 0,2 в верхней строке ограничений, которая отсюда становится основной строкой. Для преобразования основы в единицу мы делим все элементы основной строки на 0,2. Затем трансформируем все остальные элементы основного столбца в нули, вычитаем заново преобразованную основную строку из нижнего ограничения и прибавляем 0,04 раза новую основную строку к целевой функции. Получившаяся таблица имеет вид:

	W_a	W_b	s_1	s_2	ПС
ограничение 1	**0**	**1**	**5**	**–0,5**	**0,5**
ограничение 2	**1**	**0**	**–5**	**1,5**	**0,5**
целевая функция	**0**	**0**	**0,2**	**0,01**	**0,05**

← индикаторы →

Заметьте, что и W_a и W_b теперь являются базисными и что обе свободные переменные стали небазисными.

Теперь рассмотрим целевую функцию. Она выражается как $Z = 0{,}05-0{,}2s_1-0{,}01s_2$. Однако вспомним постоянную 0,08, следовательно, реально она выражается как

максимизировать

$$Z = 0{,}13-0{,}2s_1-0{,}01s_2$$

при

$$\begin{aligned} W_b + 5s_1-0{,}5s_2 &= 0{,}5 \\ W_a-5s_1 + 1{,}5s_2 &= 0{,}5. \end{aligned} \tag{9.22}$$

Но свободные переменные являются небазисными и поэтому имеют нулевые значения.

Поскольку $Z = 0{,}13-0{,}2s_1-0{,}01s_2$ и мы имеем точку возможного решения, в которой и s_1 и s_2 равны нулю, лучшего не добиться — мы нашли оптимальную точку.

Нулевые значения s_1 и s_2 свидетельствуют, что ограничения, относящиеся к этим расхождениям/избыткам, являются "жесткими". В большей задаче с большим числом ограничений значения других свободных переменных не будут ненулевыми и дадут нам уровень, до которого соответствующие ограничения "ослаблены". Более того, коэффициенты s_1 и s_2 в целевой функции (0,2 и 0,01 соответственно) показывают нам предельную стоимость ослабления соответствующих ограничений (что часто приводит к задействованию дополнительных ресурсов).

Однако вспомним, что первоначальной задачей было создание портфеля из трех активов —*A*, *B* и *C*. Решение, найденное выше, показывает, как и графическое решение, что оптимальный портфель содержит только активы *A* и *B*.

Как можно видеть, симплексный метод выражает алгебраически процесс перехода от вершины к вершине области возможных решений, где движение всегда предпринимается в направлении увеличения величины целевой функции (заметьте, что существуют случаи исключений, в которых величине целевой функции позволяется оставаться постоянной при шаге). Преимущество переключения с геометрии на алгебру состоит, конечно, в том, что алгебра действует при 200 измерениях так же, как и при двух, а графические изображения представить гораздо сложнее!

ПОСТРОЕНИЕ ПОРТФЕЛЕЙ ДЛЯ МИНИМИЗАЦИИ ОБЩЕЙ ДИСПЕРСИИ

CAPM предполагает, что только систематический риск каждого отдельного актива важен при построении портфеля. Однако модель, первоначально разработанная Марковицем (1952) и до сих пор широко применяемая, использует общий риск каждого отдельного актива. Следовательно, при построении портфелей и определении общего риска портфеля должны рассматриваться ковариации в каждой паре потенциальных для портфеля активов.

Из гл. 4 мы узнали, что, когда доходы по рискованному активу являются случайными переменными, доходы по портфелю — это взвешенная по стоимости средняя доходов по отдельным активам, т.е.

$$R_p = \sum_{i=1}^{n} W_i r_i . \tag{9.23}$$

Однако среднее квадратическое отклонение портфеля не равно взвешенной по стоимости средней из средних квадратических отклонений отдельных ценных бумаг, потому что должна быть учтена ковариация в каждой паре активов. Для иллюстрации этого среднее квадратическое отклонение портфеля из двух активов составляет

$$\sigma_p = \sqrt{W_a^2\sigma_a^2 + W_b^2\sigma_b^2 + 2W_aW_b(\rho_{ab}\sigma_a\sigma_b)}\,, \tag{9.24}$$

где σ_p— среднее квадратическое отклонение портфеля;

W_a и W_b— веса активов a и b в портфеле;

σ_a^2 и σ_b^2 — дисперсии доходов по активам A и B;

ρ_{ab} — корреляция доходов по активам A и B;

σ_a и σ_b — средние квадратические отклонения доходов по a и b;

$(\rho_{ab}\sigma_a\sigma_b)$ — **ковариация** доходов по активам A и B.

Выражение (9.24) может быть обобщено:

$$\sigma_p^2 = \sum_{i=1}^{n} W_i\sigma_i^2 + \sum_{i=1}^{n} \sum_{j=1, i\neq j}^{n} W_iW_j\sigma_{ij}\,, \tag{9.25}$$

где σ_{ij} — ковариация в портфеле в парах активов.

Для портфеля активов с *a* по *n* это может быть записано в матричном формате как

$$\sigma_p^2 = [W_a W_b W_c \ldots W_n] \begin{bmatrix} \sigma_a^2 & \sigma_{ab} & \cdots & \sigma_{an} \\ \sigma_{ba} & \sigma_b^2 & \cdots & \sigma_{bn} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ \sigma_{na} & \sigma_{nb} & \cdots & \sigma_n^2 \end{bmatrix} \begin{bmatrix} W_a \\ W_b \\ W_c \\ . \\ . \\ W_n \end{bmatrix} \qquad (9.26)$$

Каждый элемент-дисперсия в дисперсионно-ковариационной матрице умножен дважды на соответствующий ему вес актива, поэтому веса, связанные с дисперсиями, имеют возведенное в квадрат влияние. Поэтому W_i^2. Каждая ковариация умножается один раз на вес каждого актива из пары активов, и существуют две ковариации для каждой возможной пары. Поэтому $2\mathrm{cov}\,W_i W_j$.

Граница эффективности

Как было показано в гл. 2, если коэффициент корреляции в парах активов меньше чем 1,0, то диверсификация может улучшить взаимосвязь между ожидаемым риском портфеля и ожидаемым доходом по портфелю. Это происходит потому, что, если переменная доходности является линейной функцией средней доходности, то фактор риска представляет собой квадратическую функцию дисперсии доходов по ценным бумагам. Степень улучшения портфеля зависит от весов, которые каждый из активов имеет в портфеле, и от корреляции этих активов.

Лучший способ продемонстрировать это — пример с двумя активами. Рассмотрим данные табл. 9.1 — различные средние квадратические отклонения портфеля, составленного из двух рискованных активов, при допущениях, что корреляция (Cor) равна 0,6 или 0,9 и что доли каждого актива в портфеле меняются на 10%. Рис. 9.2 — это диаграмма границ эффективности, относящихся к портфелям, построенным с учетом предположенных Cor = 0,60 и Cor = 0,90. Актив *A* имеет ожидаемый доход 10% со средним квадратическим отклонением 14%, а актив *B* — ожидаемый доход 12% со средним квадратическим отклонением 15%.

Таблица 9.1

Вес W_b	**Вес W_a**	**Доход R_p**	**Среднее квадратическое отклонение, Cor = 0,6**	**Среднее квадратическое отклонение, Cor = 0,9**
0	1,0	10,0	14,0	14,0
0,1	0,9	10,2	13,55	13,965
0,2	0,8	10,4	13,22	13,961
0,3	0,7	10,6	13,01	13,988
0,4	0,6	10,8	12,92	14,045
0,5	0,5	11,0	12,97	14,133
0,6	0,4	11,2	13,15	14,251
0,7	0,3	11,4	13,45	14,397
0,8	0,2	11,6	13,86	14,571
0,9	0,1	11,8	14,38	14,773
1,0	0	12,0	15,0	15,0

Для предположенной степени корреляции среднее квадратическое отклонение рассчитано для некоторых различных портфелей, которые могут быть построены из этих двух активов и нанесены на диаграмму (рис. 9.2).

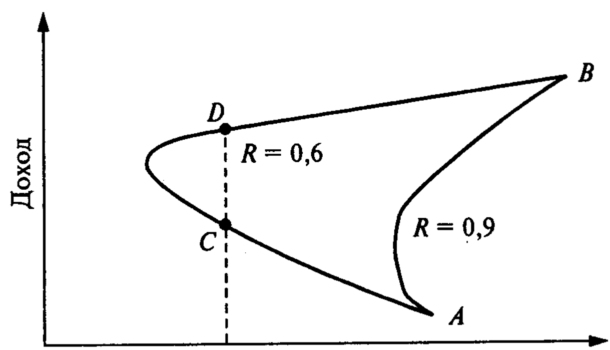

Рис. 9.2. Границы эффективности

Сначала рассмотрим данные в столбце табл. 9.1 для Cor = 0,6 и график на рис. 9.2 для Cor = 0,6, отражающие выгоды от диверсификации для случая, когда активы умеренно коррелированны. Данные и график, обозначенные Cor = 0,9, показывают, что диверсификация имеет благотворное влияние на соотношение риск-доход, даже когда активы высоко, но не полностью коррелированны. Заметьте, что в обоих случаях граница эффективности вогнута. Чем больше степень вогнутости, тем больше выгоды от диверсификации. Учтите, однако, что не все точки на границе эффективны, а эффективна только верхняя часть каждой вогнутой границы (обозначенных *AB* на рис. 9.2).

Верхняя часть каждой из линий *AB* представляет границу эффективности возможных портфелей, так как на границе невозможно достичь большего дохода без несения большего риска. Выше линии находится область недостижимых комбинаций риска и дохода из-за ограниченности характеристик ценных бумаг *A* и *B*. Ниже линии находятся худшие комбинации риска и дохода, которые могут быть улучшены просто перемещением в любую точку на линии *AB*. Это достигается продажей существующих активов и покупкой *A* и/или *B*. Например, портфель *C* располагается на нижней части границы, помеченной Cor = 0,6. Инвестор может повысить свою полезность продажей этого портфеля и покупкой комбинации *A* и *B*, представленной любой из точек на границе эффективности. Например, перемещаясь в точку *D*, инвестор несет тот же уровень риска, но получает более высокий доход, чем в *C*.

Нужно отметить, что не существует единственного наилучшего портфеля. Жирные линии указывают на многие "эффективные портфели". Граница эффективна, потому что невозможно повысить доход без увеличения риска или снизить риск без снижения дохода. Возможная комбинация риска и дохода будет зависеть от целевой функции (функция полезности для инвестора).

Однако давая в реальности обычно менее чем полностью коррелированные доходы по отдельным активам, теория предполагает, что наиболее диверсифицированным и, следовательно, приносящим наилучший доход на единицу риска, будет портфель, который содержит все рискованные активы. Это должны помнить инвестиционные менеджеры, поскольку их портфели обычно ограничены до содержания только денежных средств, облигаций и обычных акций.

Задача оптимизации портфеля

Теперь, понимая взаимосвязь между риском и доходом и влияние ковариации, мы можем определить задачу оптимизации портфеля. **Задача оптимизации портфеля** заключается в том, чтобы определить, какая доля портфеля должна быть отведена для каждой из инвестиций так, чтобы величина ожидаемого дохода и уровень риска **оптимально** соответствовали целям инвесторов. Предположим, что цель инвестора состоит в минимизации риска портфеля, где риск измеряется дисперсией портфеля.

На практике инвестор обычно устанавливает ограничения относительно способа, по которому может быть построен портфель. Например, целевой функцией может быть минимизация риска, но при каком-то минимальном уровне дохода, а также при ограничениях на минимальные и максимальные доли, которые могут быть инвестированы в каждый актив. Как поступать с этими ограничениями — объясним позже.

Сейчас же проиллюстрируем портфельную задачу, рассмотрев оптимизацию при ограничениях для случая портфеля из трех активов.

Риск и доход для портфеля из трех активов *a*, *b* и *c* и их весами, обозначенными как W_i, определяются следующим образом.

Математическое ожидание дохода

$$E(r_p) = W_a r_a + W_b r_b + W_c r_c. \tag{9.27}$$

Дисперсия портфеля составляет

$$\begin{aligned}\sigma_p^2 = {} & W_a^2\,\sigma_a^2 + W_b^2\,\sigma_b^2 + W_c^2\,\sigma_c^2 \\ & + 2W_aW_b(\mathrm{cov}_{ab}) \\ & + 2W_aW_c(\mathrm{cov}_{ac}) \\ & + 2W_bW_c(\mathrm{cov}_{bc}).\end{aligned} \tag{9.28}$$

Как было объяснено ранее, эта дисперсия портфеля может быть выражена с помощью векторов весов и матрицы дисперсий и ковариаций. Выразив дисперсию портфеля через *Z*, получим:

$$Z = [W_a\ W_b\ W_c]\begin{bmatrix}\mathrm{var}_a & \mathrm{cov}_{ab} & \mathrm{cov}_{ac}\\ \mathrm{cov}_{ba} & \mathrm{var}_b & \mathrm{cov}_{bc}\\ \mathrm{cov}_{ca} & \mathrm{cov}_{cb} & \mathrm{var}_c\end{bmatrix}\begin{bmatrix}W_a\\ W_b\\ W_c\end{bmatrix} \tag{9.29}$$

ОПТИМИЗАЦИЯ ПРИ ОГРАНИЧЕНИЯХ

Требования инвестора обычно ограничивают процесс выбора. Например, инвестор может потребовать минимизации риска при ожидаемом доходе не менее или равном данному уровню.

Портфельная задача, таким образом, состоит в минимизации дисперсии портфеля при каком-то минимальном уровне дохода. Как мы видели выше, дисперсия портфеля Z может быть выражена как произведение транспонированного вектора W, т.е. $\underline{W}^{\mathrm{T}}$, дисперсионно-ковариационной матрицы Ω и вектора W, т.е. $\underline{W}$. Следовательно, поставленная задача является задачей квадратического программирования и может быть формально записана как

$$\begin{array}{lll} \text{минимизировать} & Z = \underline{W}^{\mathrm{T}}\Omega\underline{W} & \\ \text{при} & W_a + W_b + W_c = 1 & (9.30) \\ & W_aE(r_a) + W_bE(r_b) + W_cE(r_c) \geq R, & \end{array}$$

где R —это минимальный приемлемый уровень дохода.

З а м е ч а н и е: при матричном отображении выделение выражений жирным шрифтом и их подчеркивание как $\underline{W}$ означает, что эти выражения представляют собой векторы.

Проиллюстрируем процесс оптимизации примером трех активов.

Приложение квадратического программирования к задаче выбора портфеля из трех активов — нахождение оптимального (с минимальной дисперсией) портфеля

Предположим, что мы имеем три актива — A, B и C с ожидаемыми доходами 0,11, 0,15 и 0,08 соответственно. Дисперсионно-ковариационная матрица, которая будет обозначена как Ω, имеет следующий вид:

$$\begin{array}{cc} & \begin{array}{ccc} A & B & C \end{array} \\ \begin{array}{c} \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \end{array} \; \Omega = & \begin{bmatrix} 0{,}00015 & 0{,}00005 & -0{,}00007 \\ 0{,}00005 & 0{,}00025 & -0{,}00003 \\ -0{,}0007 & -0{,}00003 & 0{,}00010 \end{bmatrix} \end{array}$$

Мы хотим найти пропорции W_S для инвестирования в каждый актив, чтобы получить требуемый доход 11% при минимальной дисперсии, т.е. найти $\underline{W}$ (т.е. $[W_aW_bW_c]^{\mathrm{T}}$) для решения следующей задачи

$$\begin{array}{ll} \text{минимизировать} & Z = \underline{W}^{\mathrm{T}}\Omega\underline{W} \\ \text{при} & W_a + W_b + W_c = 1 \\ & 0{,}11W_a + 0{,}15W_b + 0{,}08W_c = 0{,}11. \end{array}$$

В данном примере мы предполагаем, что отрицательные позиции по активам невозможны. Составляющие дисперсии $Z = \underline{W}^{\mathrm{T}}\Omega\underline{W}$ перемножаются, чтобы получить

$$Z = [W_a\ W_b\ W_c]\begin{bmatrix} 0{,}00015 & 0{,}00005 & -0{,}00007 \\ 0{,}00005 & 0{,}00025 & -0{,}00003 \\ -0{,}0007 & -0{,}00003 & 0{,}00010 \end{bmatrix}\begin{bmatrix} W_a \\ W_b \\ W_c \end{bmatrix}$$

$$(= W_a^2\,\sigma_a^2 + W_b^2\,\sigma_b^2 + W_c^2\,\sigma_c^2 + 2W_aW_b\mathrm{cov}_{ab} + 2W_aW_c\mathrm{cov}_{ac} + 2W_bW_c\mathrm{cov}_{bc})$$

$$= 0{,}00015W_a^2 + 0{,}00025W_b^2 + 0{,}00010W_c^2 + 0{,}00010W_aW_b -$$
$$- 0{,}00014W_aW_c - 0{,}00006W_bW_c.$$

Таким образом, наша оптимизационная задача в итоге выглядит как

минимизировать

$$Z = \mathbf{0{,}00015W_a^2 + 0{,}00025W_b^2 + 0{,}00010W_c^2 + 0{,}00010W_aW_b -}$$
$$\mathbf{0{,}00014W_aW_c - 0{,}00006W_bW_c}$$

при

$$\mathbf{W_a + W_b + W_c = 1}$$
$$\mathbf{0{,}11W_a + 0{,}15W_b + 0{,}08W_c = 0{,}11}$$

Оптимизация при ограничениях в виде равенств: использование множителей Лагранжа

Так как наша задача в приведенном выше виде содержит ограничения только в виде равенств, она может быть решена с использованием множителей Лагранжа, одного для каждого ограничения. Применение множителей Лагранжа к оптимизации при одной переменной было описано в гл. 3.

Вспомним, что, когда множители Лагранжа применяются в случае одной переменной, сначала ограничение приравнивается нулю и затем прибавляется к оптимизируемой функции

минимизировать $Z = f(\underline{W})$ (подчеркивание выделенного жирным W указывает, что фактически W — это вектор, например, n активов)

при $g(\underline{W}) = \underline{0}$ (тот факт, что g выражено в виде вектора, показывает, что существует несколько, скажем, m равенств. Заметьте, что $m < n$).

В нашем конкретном случае

$$\underline{g} = \begin{bmatrix} g_1(\underline{W}) \\ g_2(\underline{W}) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad (9.31)$$

где

$$g_1(\underline{W}) = W_a + W_b + W_c - 1$$

$$g_2(\underline{W}) = 0{,}11W_a + 0{,}15W_b + 0{,}08W_c - 0{,}11$$

Построим лагранжиан: $L(\underline{W},\underline{\lambda})$, где $L(\underline{W},\underline{\lambda})$ =

$$L(\underline{W},\underline{\lambda}) = Z - \sum_{i=1}^{m} l_i g_i(\underline{W}), \qquad (9.32)$$

λ_i — это множители Лагранжа.

Следовательно, лагранжиан строится посредством вычитания из первоначальной целевой функции всех отдельных функций ограничений, которые были помножены на соответствующие им множители Лагранжа.

Заметьте, что сейчас мы имеем функцию $n + m$ переменных. Найдем n частных производных по переменным x_i (замечание: мы не определяем m частных производных для множителей Лагранжа, потому что частные производные по λ только лишь возвращают нас к уже имеющимся ограничениям в виде равенств). Затем попробуем найти x_iS и λ_S, которые превращают эти частные производные в ноль и одновременно удовлетворяют ограничениям в виде равенств. Отсюда мы имеем систему $n + m$ уравнений с $n + m$ неизвестными. Так как число уравнений и неизвестных равно, мы можем найти решение через постановку и решение задачи, представленной системой уравнений.

Однако существует или нет решение зависит от того, являются ли ограничения противоречивыми. Если одно ограничение гласит, скажем, $x_{20} < 50$,а другое — $x_{20} > 100$, то ясно, что решения не существует, но в большой и сложной задаче подобные несоответствия заметить бывает очень трудно. В зависимости от сложности функции L мы можем найти решение аналитически или с использованием численного метода при условии, что решение существует.

Применим это к нашей задаче
минимизировать

$$Z = 0{,}00015\,W_a^2 + 0{,}00025\,W_b^2 + 0{,}00010\,W_c^2 + 0{,}00010\,W_aW_b - 0{,}00014\,W_aW_c - 0{,}00006\,W_bW_c$$

при

$$W_a + W_b + W_c = 1$$
$$0{,}11W_a + 0{,}15W_b + 0{,}08W_c = 0{,}11.$$

Отсюда, лагранжиан имеет вид:

$$L(\underline{W},\underline{\lambda}) = 0{,}00015\,W_a^2 + 0{,}00025\,W_b^2 + 0{,}00010\,W_c^2 + 0{,}00010\,W_aW_b -$$
$$0{,}00014\,W_aW_c - 0{,}00006\,W_bW_c - \lambda_1(W_a + W_b + W_c - 1) -$$
$$\lambda_2(0{,}11W_a + 0{,}15W_b + 0{,}08W_c - 0{,}11).$$

Найдем частные производные и приравняем их к нулю (заметьте, что $\underline{W}^{\mathrm{T}}\Omega\underline{W}$ дифференцируется до $2\Omega\underline{W}$ либо это можно сделать почленно. Мы выбираем последнее)

$$\partial L/\partial W_a = 0{,}00030\,W_a + 0{,}00010\,W_b - 0{,}00014\,W_c - \lambda_1 - 0{,}11\lambda_2 = 0$$
$$\partial L/\partial W_b = 0{,}00050\,W_b + 0{,}00010\,W_a - 0{,}00006\,W_c - \lambda_1 - 0{,}15\lambda_2 = 0$$
$$\partial L/\partial W_c = 0{,}00020\,W_c - 0{,}00014\,W_a - 0{,}00014\,W_b - \lambda_1 - 0{,}08\lambda_2 = 0,$$

помня об ограничениях

$$W_a + W_b + W_c = 1$$
$$0{,}11\,W_a + 0{,}15\,W_b + 0{,}08\,W_c = 0{,}11.$$

Таким образом, мы имеем пять уравнений с пятью неизвестными. Решим эту систему уравнений, используя ту же матричную алгебру, что и в гл. 6 при решении систем уравнений для модели множественной регрессии. Однако здесь задача более проста, так как матрица Ω симметрична, она имеет размер 5×5.

Сначала частные производные и ограничения выражаются в формате матрицы

$$\begin{array}{rrrrrl}
0{,}00030\,W_a & +\,0{,}00010\,W_b & -\,0{,}00014\,W_c & -\,\lambda_1 & -\,0{,}11\lambda_2 & = 0\\
0{,}00010\,W_a & +\,0{,}00050\,W_b & -\,0{,}00006\,W_c & -\,\lambda_1 & -\,0{,}15\lambda_2 & = 0\\
-0{,}00014\,W_a & -\,0{,}00014\,W_b & +\,0{,}00020\,W_c & -\,\lambda_1 & -\,0{,}08\lambda_2 & = 0\\
1\,W_a & 1\,W_b & 1\,W_c & 0\lambda_1 & 0\lambda_2 & = 1\\
0{,}11\,W_a & 0{,}15\,W_b & 0{,}08\,W_c & 0\lambda_1 & 0\lambda_2 & = 0{,}11.
\end{array}$$

Затем мы располагаем левосторонние элементы в матрицу 5×5, за которой следует вектор переменных 5×1, включающий λ. Справа мы располагаем единичную матрицу и вектор 5×1, отражающий правую сторону дифференцированного выражения:

$$\begin{bmatrix} 0{,}00030 & 0{,}00010 & -0{,}00014 & -1 & -0{,}11 \\ 0{,}00010 & 0{,}00050 & -0{,}00006 & -1 & -0{,}15 \\ -0{,}00014 & -0{,}00006 & 0{,}00020 & -1 & -0{,}08 \\ 1 & 1 & 1 & 0 & 0 \\ 0{,}11 & 0{,}15 & 0{,}08 & 0 & 0 \end{bmatrix} \begin{bmatrix} W_a \\ W_b \\ W_c \\ \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0{,}11 \end{bmatrix}$$

Начнем с умножения первых строк обеих матриц на 3333,33 для приведения элемента в левом верхнем углу к 1.

В следующем шаге отнимем 0,00010 раз новую первую строку от вторых строк с обеих сторон. Затем добавим 0,00014 раз новую первую строку к третьим строкам с обеих сторон, потом отнимем 1 раз новую первую строку от четвертых строк и в завершение отнимем 0,11 раз новую первую строку от пятых строк с обеих сторон. В результате первый столбец в левой матрице содержит 1 в качестве верхнего элемента и 0 для всех остальных, отсюда

$$\begin{bmatrix} 1 & 0{,}3333 & -0{,}46666 & -3333{,}33 & -366{,}66 \\ 0 & 0{,}000466 & -0{,}000013 & -0{,}66666 & -0{,}11333 \\ 0 & -0{,}000013 & 0{,}00020 & -1{,}4666 & -0{,}131333 \\ 0 & 0{,}9999 & 1{,}46666 & 3333{,}33 & 366{,}666 \\ 0 & 0{,}11333 & 0{,}13133 & 366{,}666 & 40{,}3333 \end{bmatrix} \begin{bmatrix} W_a \\ W_b \\ W_c \\ \lambda_1 \\ \lambda_2 \end{bmatrix} =$$

$$= \begin{bmatrix} 3333{,}3 & 1 & 0 & 0 & 0 \\ -0{,}33333 & 1 & 0 & 0 & 0 \\ -0{,}46666 & 0 & 1 & 0 & 0 \\ -3333{,}33 & 0 & 0 & 1 & 0 \\ 366{,}666 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0{,}11 \end{bmatrix}$$

Продолжим похожим способом, сначала умножая вторые строки так, чтобы второй элемент во втором столбце левой матрицы (0,000466) стал единицей, и затем вычитая или прибавляя произведения второй строки, чтобы обратить остальные элементы второго столбца в нули. Повторение этого процесса для третьего, четвертого и пятого столбцов левой матрицы приводит к преобразованию ее в единичную матрицу, в то время как проведение действий над правой матрицей производит кратную обратную. В конце мы приходим к

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} W_a \\ W_b \\ W_c \\ \lambda_1 \\ \lambda_2 \end{bmatrix} =$$

$$\begin{bmatrix} 1991{,}87 & -853{,}659 & -1138{,}21 & 0{,}79350 & -4{,}22764 \\ -853{,}659 & 365{,}854 & 487{,}805 & -1{,}48293 & 16{,}0976 \\ -1138{,}21 & 487{,}805 & 650{,}407 & 1{,}68943 & -11{,}8699 \\ -0{,}79350 & 1{,}48293 & -1{,}68943 & 0{,}00155 & 0{,}01542 \\ 4{,}22764 & -16{,}0976 & 11{,}8699 & -0{,}01542 & 0{,}15837 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0{,}11 \end{bmatrix} = \begin{bmatrix} 0{,}328 \\ 0{,}288 \\ 0{,}384 \\ -0{,}000157 \\ 0{,}002 \end{bmatrix}$$

получая результат $W_a = 0{,}328$, $W_b = 0{,}288$ и $W_c = 0{,}384$ (при $\lambda_1 = -0{,}000157$ и $\lambda_2 = 0{,}002$).

Конечно, вычисления, описанные выше, очень утомительны, и практически невозможно провести их вручную без ошибок (даже используя калькулятор). Существует много эффективных компьютерных методов, чтобы выполнить эту задачу, и мы могли бы использовать электронные таблицы. Однако можно применить и пакет линейного программирования, если целевая функция просто выражена в какой-либо искусственной форме, и ограничения выглядят так:

$$\begin{aligned} 0{,}00030W_a &+ 0{,}00010W_b & -0{,}00014W_c & -\lambda_1 - 0{,}11\lambda_2 &= 0 \\ 0{,}00010W_a &+ 0{,}00050W_b & -0{,}00006W_c & -\lambda_1 - 0{,}15\lambda_2 &= 0 \\ -0{,}00014W_a &- 0{,}00014W_b & +0{,}00020W_c & -\lambda_1 - 0{,}08\lambda_2 &= 0 \\ 1W_a &+ 1W_b & +1W_c & \;0\lambda_1 \quad 0\lambda_2 &= 1 \\ 0{,}11W_a &+ 0{,}15W_b & +0{,}08W_c & \;0\lambda_1 \quad 0\lambda_2 &= 0{,}11. \end{aligned}$$

В действительности пакет линейного программирования просто используется для нахождения возможного решения системы уравнений. Так как уравнений столько же, сколько неизвестных, обычно будет существовать единственное возможное решение, которые и должно быть оптимальным.

Квадратическое программирование с неравенствами

При практическом применении задачи выбора портфелей часто могут включать некоторое количество ограничений в виде неравенств, обычно устанавливающих пределы для инвестирования в конкретных направлениях. Представим, например, что в рассматриваемой задаче мы налагаем дополнительное ограничение $W_c \leq 0{,}25$.

Проблема здесь заключается в том, что мы не знаем заранее, будет ли наше дополнительное ограничение задействовано. Конечно, мы можем разрешить эту дилемму, просто решая две задачи, в одной из которых мы не имеем ограничений для x_3, и, если ответ дает значение $W_c > 0{,}25$, тогда то ограничение, на котором мы настаиваем, составляет $W_c = 0{,}25$. Однако с двумя подобными ограничениями будут существовать четыре возможности для проверки — ни одно из ограничений не является жестким; первое — жесткое, а второе — нет; второе жесткое, а первое — нет; оба — жесткие. При десяти подобных ограничениях существует $2^{10} = 1024$ возможности. Так как избежать этого комбинаторного взрыва нельзя, было бы полезным иметь возможность каким-то образом автоматизировать поиск. Условия Юна—Такера предоставляют такой механизм.

Заметьте, что мы теперь не способны гарантировать, что окажемся на границе эффективности, поэтому мы также должны ослабить наши ограничения в виде равенств по поводу полного инвестирования и/или требуемого дохода — мы можем прийти к более рискованному портфелю, но нам не нужно будет использовать все свои деньги, и/или мы можем прийти к большему доходу!

УСЛОВИЯ КЮНА—ТАКЕРА (Kuhn—Tucker)

Первый шаг в этом анализе — манипулирование неравенствами, так чтобы задача была выражена в **стандартной форме**

минимизировать $Z = f(\underline{W})$

при $g_i(\underline{W}) = 0$ для i в некоторой совокупности E (т.е. совокупность ограничений в виде равенств)

$g_i(\underline{W}) \geq 0$ для i в некоторой совокупности I (ограничения в виде неравенств)

Следовательно, для нашей задачи (заметьте, что теперь не существует ограничений в виде равенств) минимизировать $Z = 0{,}00015W_a^2 + 0{,}00025W_b^2 + 0{,}00010W_c^2 + 0{,}00010W_aW_b - 0{,}00014W_aW_c - 0{,}00006W_bW_c$

при

$$1 - W_a - W_b - W_c \geq 0$$
$$0{,}11W_a + 0{,}15W_b + 0{,}08W_c - 0{,}11 \geq 0$$
$$0{,}25 - W_c \geq 0$$

Заметьте, что необходимо представить неравенство $W_a + W_b + W_c \leq 1$ в **стандартной форме**, т.е. $1 - W_a - W_b - W_c \geq 0$.

Снова строится лагранжиан и рассматриваются аналитические методы для нахождения его неограниченного минимума. Однако появляется сложность — не все ограничения в виде неравенств должны действовать. Анализ справляется с этим посредством приведения соответствующих множителей Лагранжа к нулю (существуют обстоятельства исключений, когда множители Лагранжа для действующих выражений могут также быть равны нулю).

Это выражается формально в виде поиска точек Кюна—Такера, т.е. точек, удовлетворяющих условиям Кюна—Такера

$$\frac{\delta L(\underline{W}, \lambda)}{\delta W_i} = 0 \text{ для каждого значения } i \tag{9.33}$$

и

$$g_i(\underline{W}) = 0 \text{ для } i \text{ в } E$$
$$g_i(\underline{W}) \geq 0 \text{ для } i \text{ в } I$$
$$\lambda_i \geq 0 \text{ для всех } i$$
$$\lambda g_i(\underline{W}) = 0 \text{ для всех } i.$$

Последнее условие называется *условием комплиментарности* и говорит о том, что недействующие ограничения имеют нулевой множитель.

И снова для нашей задачи (принимая во внимание третий множитель Лагранжа, относящийся к $0{,}25 - W_c \geq 0$)

$$0{,}00030W_a + 0{,}00010W_b - 0{,}00014W_c - \lambda_1 - 0{,}11\lambda_2 + \lambda_3 = 0$$
$$0{,}00010W_a + 0{,}00050W_b - 0{,}00006W_c - \lambda_1 - 0{,}15\lambda_2 + \lambda_3 = 0$$
$$0{,}00020W_c - 0{,}00014W_a + 0{,}00006W_c - \lambda_1 - 0{,}08\lambda_2 + \lambda_3 = 0$$
$$1 - W_a - W_b - W_c \geq 0$$
$$0{,}11W_a + 0{,}15W_b + 0{,}08W_c - 0{,}11 \geq 0$$
$$0{,}25 - W_c \geq 0$$
$$\lambda_1 \geq 0$$
$$\lambda_2 \geq 0$$
$$\lambda_3 \geq 0$$
$$\lambda_i g_i(\underline{W}) = 0 \text{ для всех } i.$$

Анализ Кюна—Такера предоставляет базу для адаптации линейного программирования к решению этой задачи.

МЕТОД ДАНЦИГА—ВОЛЬФА (Dantzig—Wolfe)

Установив условия Кюна—Такера, будем искать точку Кюна—Такера, т.е. точку, удовлетворяющую каждому из этих условий. Заметьте, что у нас есть девять линейных ограничений. Следовательно, за исключением комплиментарности и отсутствия целевой функции мы имеем случай линейного программирования. Наша задача заключается в нахождении точки возможного решения для определения оптимальной точки.

Нахождение первой точки возможного решения — это фактически первый шаг симплексного алгоритма. Оно включает в себя создание искусственного решения, в котором все "реальные" переменные имеют нулевые значения, а искусственные переменные $a_1,..., a_9$ добавлены по одной к каждому ограничению при значениях, равных соответствующим правым сторонам. В этом случае цель состоит в минимизации суммы a_is. Когда решение найдено при этой сумме, равной нулю, мы будем иметь возможное решение нашей задачи.

Таким образом, все, что требуется, — это взять начальную часть симплексного алгоритма и добавить дополнительные ограничения. В нашем примере присутствуют три подобных условия:

по меньшей мере один из λ_1 и $(W_a + W_b + W_c—1)$ должен быть нулем;
по меньшей мере один из λ_2 и $(0{,}11W_a + 0{,}15W_b + 0{,}08W_c—0{,}11)$ должен быть нулем;
по меньшей мере один из λ_3 и $(0{,}25—W_c)$ должен быть нулем.

Эти условия легко проверить. Следовательно, использование до сих пор чисто теоретических условий Кюна—Такера дает возможность преобразовать нашу конкретную задачу квадратического программирования в модифицированную задачу линейного программирования.

Результат, полученный для нашей задачи, составляет W_a = 0,266, W_b = 0,405 и W_c = 0,25. Это решение с минимальной дисперсией, но заметьте, что не все наши деньги использованы: $W_a + W_b + W_c = 0{,}921$. Это не должно быть удивительным, так как наше дополнительное ограничение не дает нам использовать наименее рискованный актив, как этого мы хотим. Данное решение показывает, что возможно достичь нашего целевого дохода при использовании меньшего, чем мы можем себе позволить, объема более рискованных активов.

Если мы настаиваем на инвестировании всех наших денег, тогда можем вернуться к ЛП, меняя $W_a + W_b + W_c \leq 1$ на $W_a + W_b + W_c = 1$. Это приводит к результату $W_a = 0{,}281$, $W_b = 0{,}469$ и $W_c = 0{,}25$. Доход сейчас составляет 0,121, что на 1,1% больше, чем нам было нужно.

Краткое изложение результатов дано в табл. 9.2.

Таблица 9.2

	Эффективный портфель с минимальной дисперсией	Портфель с минимальной дисперсией при дополнительном ограничении	Лучшие инвестиции при дополнительном ограничении (т.е. со всеми инвестированными средствами)
Доля, инвестированная в ценную бумагу №1	0,328	0,266	0,281
Доля, инвестированная в ценную бумагу №2	0,288	0,405	0,469
Доля, инвестированная в ценную бумагу №3	0,384	0,25	0,25
Общая сумма инвестированных долей	1	0,921	1
Доход	0,11	0,11	0,121
Риск (дисперсия)	0,0000368	0,0000533	0,0000694
(среднее квадратическое отклонение)	0,00607	0,00730	0,00833100

КРАТКИЙ ОБЗОР МЕТОДОВ ВОСХОЖДЕНИЯ НА ХОЛМЫ

Все большее число компьютерных пакетов, электронных таблиц и т. п. предоставляют встроенные функции оптимизации. Существует много доступных мощных алгоритмов, и нет какого-либо одного метода, который бы предусматривал все, что можно встретить. Однако мы попытаемся дать понятие о них, показывая, как задача Марковица может быть решена более общими методами.

Метод Данцига—Вольфа ограничен задачей управления портфелем Марковица. Поскольку существует много различных обстоятельств, при которых возникают задачи квадратического программирования, то, нужны более общие подходы. Действительно, портфельные задачи Марковица сейчас чаще решаются именно подобными пакетами.

Подход математического программирования, представленный ранее, предполагает движение от одной точки к другой по границе области допустимых решений, всегда увеличивая значение целевой функции (для задачи максимизации). Таким образом, фокус находится на ограничениях, а цель предоставляет направленность для итерации. Альтернативный метод предполагает движение от точки к точке в границах области возможных решений, фокус находится на цели и перемещается при каждой итерации насколько возможно далеко в рамках ограничений.

Смысл этого подхода состоит в использовании характеристик целевой функции в текущей точке для определения направления движения. Затем в этом направлении происходит перемещение в рамках области возможных решений в новую точку, и процесс повторяется. Иногда для описания этого подхода используется термин "восхождение на холмы". Выбор направления может быть сделан посредством изучения или оценки градиента функции (отсюда термин "метод градиентов"), или другими способами (методы прямого поиска).

Один из таких подходов широко используется при решении задач квадратического программирования, они известны как **методы активной группы**.

Методы активной группы для задач квадратического программирования

Методы активной группы — это методы поиска, в которых при каждой итерации подгруппа ограничений в виде неравенств вместе со всеми ограничениями в виде равенств являются **активными** (действующими). В этом случае итерация состоит в попытке движения из текущей точки, которая удовлетворяет всем активным ограничениям как равенствам, в наилучшую точку, которая удовлетворяет всем активным ограничениям как равенствам.

Нахождение наилучшей точки, т.е. решение задачи квадратического программирования при ограничениях в виде равенств, является сравнительно простым. Оно может быть, например, выполнено с помощью равенств для сокращения переменных или множителей Лагранжа, описанных ранее. Но движение в эту точку может быть невозможным, так как оно способно привести к нарушению одного или более неактивных ограничений. Соответственно перемещение делается как можно дальше в направлении от текущей точки к наилучшей точке (для текущей активной группы).

Если шаг не достигает наилучшей точки, то это означает что какое-то другое ограничение стало активным. Затем это ограничение добавляется к активной группе, и процесс готов начать следующую итерацию.

Если наилучшая точка (для текущей активной группы) достигнута, т.е. если она возможна, то для активных ограничений рассчитываются множители Лагранжа. Если все они неотрицательны, то решение найдено. Если хотя бы один отрицателен, то это означает, что соответствующему неравенству может быть позволено стать неактивным. Следовательно, удаляется из активной группы, подготавливая путь для новой итерации.

УПРАЖНЕНИЯ

1. Что вы понимаете под терминами “целевая функция”, “ограничения”, “математическое программирование”, “линейное программирование” и “квадратическое программирование”?
2. Представьте графическое решение для следующей задачи линейного программирования.

Максимизировать (доход): $0{,}10W_a + 0{,}18W_b + 0{,}07W_c$

при:

(а) $0{,}8W_a + 1{,}3W_b + 1{,}1W_c \leq 1$

(б) $0 \leq W_a \leq 1$

$0 \leq W_b \leq 1$

$0 \leq W_c \leq 1$

(в) $W_a + W_b + W_c = 1$

3. Как вы понимаете термины "базисная" и "небазисная" в отношении линейного программирования?
4. Используйте симплексный метод для решения задачи из вопроса 2.
5. Рассмотрите три актива A, B и C с ожидаемыми доходами 8%, 18% и 7%. Дисперсионно-ковариационная матрица выглядит как

$$\Omega = \begin{bmatrix} 0{,}00013 & 0{,}00004 & -0{,}00008 \\ 0{,}00004 & 0{,}00022 & -0{,}00002 \\ -0{,}00008 & -0{,}00002 & 0{,}00009 \end{bmatrix}$$

Поставьте оптимизационную задачу, которая минимизирует дисперсию портфеля при условии достижения минимального дохода 12%. Замечание: вам не нужно решать эту задачу, определите только целевую функцию и ограничения.

6. Определите лагранжиан, применимый к указанной выше задаче.
7. Что вы понимаете под "условиями Кюна—Такера"? Определите условия Кюна—Такера, соответствующие задачам из вопросов 5 и 6.

ОТВЕТЫ К ИЗБРАННЫМ УПРАЖНЕНИЯМ

2. Используйте $W_a + W_b + W_c = 1$, чтобы выразить W_c как $1 - W_a - W_b$. Подстановка дает:

максимизировать $0{,}10W_a + 0{,}18W_b + 0{,}07(1 - W_a - W_b) = 0{,}03W_a + 0{,}11W_b + 0{,}07$

при $0{,}8W_a + 1{,}3W_b + 1{,}1(1 - W_a - W_b) \le 1$,

что упрощается до $W_b \le 1{,}5W_a$

$0 \le W_a \le 1$

$0 \le W_b \le 1$

$0 \le W_a + W_b \le 1$

Решение: $W_a = 0{,}4$; $W_b = 0{,}6$; $W_c = 0$; доход = 0,148.

4. Использование преобразованной задачи с двумя переменными как в вопросе 2.

(Заметьте, что вторая строка показывает $-3W_a + 2W_b \le 0$, что то же самое, что и $W_b \le 1{,}5W_a$).

Доход	W_a	W_b	S_1	S_2	S_3	S_4	ПС
1	—0,03	—0,11	0	0	0	0	0,07
0	—3	2	1	0	0	0	0
0	1	0	0	1	0	0	1
0	0	1	0	0	1	0	1
0	1	1	0	0	0	1	1
1	—0,195	0	0,055	0	0	0	0,07
0	—1,5	1	0,5	0	0	0	0
0	1	0	0	1	0	0	1
0	1,5	0	—0,5	0	1	0	1
0	2,5	0	—0,5	0	0	1	1
1	0	0	0,016	0	0	0,078	0,148
0	0	1	0,2	0	0	0,6	0,6
0	0	0	0,2	1	0	—0,4	0,6
0	0	0	—0,2	0	1	—0,6	0,4
0	1	0	—0,2	0	0	0,4	0,4

5. Пусть W_a, W_b и W_c будут весами каждого из активов в портфеле.

Минимизировать $0{,}00013W_a^2 + 0{,}00022W_b^2 + 0{,}00009W_c^2 + 0{,}00008W_aW_b — 0{,}00016W_aW_c — 0{,}00004W_bW_c$

при

$$0{,}08W_a + 0{,}18W_b + 0{,}07W_c \geq 0{,}12$$

$$0 \leq W_a \leq 1$$

$$0 \leq W_b \leq 1$$

$$0 \leq W_c \leq 1$$

$$W_a + W_b + W_c \leq 1.$$

6. Задача значительно упрощается при условии, что $W_a + W_b + W_c = 1$, т.е. что все деньги инвестируются. Более того, если позволяется создание коротких позиций (продажа при отсутствии), то неравенства $0 \leq ... \leq 1$ не нужны. Лагранжиан для данной упрощенной задачи выглядит как

$$L(W_a, W_b, W_c; 1, \mu) = 0{,}00013W_a^2 + 0{,}00022W_b^2 + 0{,}00009W_c^2$$
$$+ 0{,}00008W_aW_b — 0{,}00016W_aW_c — 0{,}00004W_bW_c$$
$$—\lambda(0{,}12 — 0{,}08W_a — 0{,}18W_b — 0{,}07W_c) — \mu(W_a + W_b + W_c — 1).$$

7. Игнорируя неравенства $0 \le \ldots \le 1$,

$$
\begin{aligned}
0{,}00026W_a + 0{,}00008W_b - 0{,}00016W_c + 0{,}08\lambda - \mu &= 0\\
0{,}00044W_b + 0{,}00008W_a - 0{,}00004W_c + 0{,}18\lambda - \mu &= 0\\
0{,}00018W_c - 0{,}00016W_a - 0{,}00004W_b + 0{,}07\lambda - \mu &= 0\\
0{,}08W_a + 0{,}18W_b + 0{,}07W_c &\ge 0{,}12\\
W_a + W_b + W_c &\le 1\\
\lambda &\ge 0\\
\mu &\ge 0\\
\lambda(0{,}12 - 0{,}08W_a - 0{,}18W_b - 0{,}07W_c) &= 0\\
\mu(W_a + W_b + W_c - 1) &= 0.
\end{aligned}
$$

СПИСОК ИСПОЛЬЗУЕМОЙ И РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

Fletcher R. (1987) *Practical Methods of Optimization*, 2nd edn. John Wiley, New York.

Wilkes F. M. (1994) *Mathematics for Business Finance and Economics.* Routledge, London.

Markowitz H. (1952) Portfolio selection. *Journal of Finance*, pp. 77—99.

10

МАТЕМАТИКА НЕПРЕРЫВНЫХ ПРОЦЕССОВ В ФИНАНСАХ: ЦЕНЫ АКТИВОВ КАК СТОХАСТИЧЕСКИЙ ПРОЦЕСС

ВВЕДЕНИЕ

Стохастический процесс — это процесс, описывающий изменения в одной или нескольких переменных, где эти изменения характеризуются неопределенностью. В особенности эти процессы применимы к анализу будущих изменений в ценах активов, так как эти изменения действительно неопределенны.

В финансах нас особенно интересуют две большие группы стохастических процессов. Процессы дискретного времени/дискретной переменной позволяют дискретным переменным изменяться в дискретные промежутки времени. Мы уже встречались с этими стохастическими процессами в форме биномиальных и триномиальных моделей в гл. 8.

В этой главе мы коснемся стохастических процессов непрерывного времени/непрерывной переменной, позволяющих непрерывным переменным — ценам активов и доходам по активам — изменяться непрерывно во времени. Математика таких процессов известна как математика непрерывных процессов, которые ведут к использованию стохастических исчислений.

Начнем главу с объяснения стохастических непрерывных процессов цен активов и доходов. Затем объясним применение стохастических исчислений к ценообразованию производных финансовых инструментов. И в заключение представим концепцию нейтральности к риску, после чего мы перейдем к детальному объяснению ценообразования производных финансовых инструментов в рамках нейтральности к риску в приложении к модели, разработанной Блэком и Сколсом (1973).

СТОХАСТИЧЕСКИЙ ПРОЦЕСС СТОИМОСТИ АКТИВОВ

Цены активов не могут быть отрицательными, но могут быть бесконечно положительными, поэтому и отношение цен, т.е. P_1/P_2 не может быть отрицательным, но может быть бесконечно положительным. Этот эмпирический факт является базой для выдвижения предположения, что отношение цен ценных бумаг распределено логнормально и что непрерывно наращенный доход по ценным бумагам, т.е. $\ln P_1/P_2$, распределен нормально. Если наблюдения проводятся за индивидуальной случайной переменной, то, поскольку они происходят в отдельном промежутке времени, случайная переменная (доход по ценной бумаге в нашем случае) может следовать случайному блужданию (random walk), что является частным случаем мартингального процесса. Непрерывное во времени случайное блуждание называется диффузионным процессом. Характеристики мартингальных процессов и частный случай случайных блужданий были изучены в гл. 7.

Существует группа случайных переменных, где смещение непосредственно зависит от предыдущего состояния данной переменной. Такие стохастические процессы предполагают, что только текущее состояние стохастической (случайной) переменной является важным в прогнозировании будущих значений этой переменной. Такие стохастические процессы известны как **марковские процессы**. Однако важным свойством всех интересующих нас процессов является то, что стохастическое изменение или инновация переменной независима и тождественно распределена (independent and identically distributed, IID)

Согласно части финансовой теории, известной как **гипотеза эффективных рынков** (efficient market hypothesis, EMH), цены активов отображают всю историческую информацию, касающуюся этого актива, и немедленную реакцию на поступающую новую информацию по этому активу. Эта ответная реакция проявляется в виде изменения цены. Если действительно рынки немедленно реагируют на новую информацию и каждая часть новой информации независима от предыдущей, изменения в ценах активов будут следовать марковскому процессу.

Хотя история движения случайной переменной на протяжении некоторого времени не требуется для прогнозирования будущих движений, статистические характеристики прошлых движений могут быть полезны в прогнозировании распределения вероятностей будущих движений. Например, средняя и волатильность прошлых движений могут быть полезны в прогнозировании будущих движений в вероятностном смысле.

Существует целое семейство марковских процессов; некоторые, такие, как биномиальные и триномиальные модели, мы уже встречали. В этой главе мы рассмотрим три процесса непрерывного времени/непрерывной переменной: основной процесс Винера, обобщенный процесс Винера и процесс Ито.

Процесс Винера, известный также как броуновское движение

Разновидность марковского процесса, которая используется как отправная точка для определения стохастических процессов цен активов, это **основной процесс Винера** (basic Wiener process), или **броуновское движение** (Brownian motion). Стохастический процесс переменной описывается как подверженный многочисленным небольшим импульсам. На данном этапе представляется подходящим начать описание процесса движения цен активов. Часто считается, что цены активов изменяются случайным образом на протяжении периода в результате совокупного эффекта многих независимых случайных импульсов, являющихся следствием получения новой информации.

Основной процесс Винера

Для того чтобы понять, каким образом процесс Винера соотносится с движением цен активов, мы начнем с объяснения основного процесса Винера.

Пусть S будет любой случайной переменной, а t — периодом времени. За малый промежуток времени Δt случайная переменная S изменится на ΔS. Если S следует процессу Винера, т.е. броуновскому движению, изменение S за малый промежуток времени будет соотноситься с Δt следующим образом:

$$\Delta S = \varepsilon\sqrt{\Delta t}\,, \tag{10.1}$$

где ε — получено на основе случайной выборки из нормально распределенной переменной со средней, равной нулю, и средним квадратическим отклонением, равным единице.

В пределе это можно записать так:

$$dS = \varepsilon\sqrt{dt}\,. \tag{10.2}$$

Так как ε — нормально распределенная переменная, то и ΔS должна быть нормально распределена со средней, равной нулю, дисперсией Δt и средним квадратическим отклонением $\sqrt{\Delta t}$.

Следовательно, по сути мы имеем переменную S, которая изменяется случайным образом на величину ΔS, которая зависит от другой случайной переменной $\varepsilon\sqrt{\Delta t}$ (эффект случайного получения новой информации на рынке), имеющей среднюю, равную нулю, дисперсию Δt и среднее квадратическое отклонение $\sqrt{\Delta t}$.

Независимость переменных — важное свойство процесса Винера. Это означает, что разные значения ΔS независимы и, так как ε нормально распределена, значения ΔS характеризуются как независимые и тождественно распределенные IID ($\Delta S \sim N(0, \sqrt{\Delta t})$). Если дисперсия отдельной ΔS равна Δt, дисперсия за более длинный промежуток времени будет $\Sigma\Delta t$.

Для того чтобы это понять, рассмотрим независимые наблюдения ΔS за два очень малых, но следующих друг за другом временных периода t_1 и t_2. Так как наблюдения S независимы, дисперсия за общий промежуток времени $T = t_1 + t_2$ будет равна сумме дисперсий за каждый короткий временной промежуток, т.е. $D_T = D_{t_1} + D_{t_2}$.

Таким образом, если случайная переменная следует процессу Винера, изменение за данный промежуток времени ($T = \sum\Delta t$), $\Delta S(T)$ будет иметь среднюю (математическое ожидание, равное нулю), но дисперсию, равную T, и среднее квадратическое отклонение, равное $\sqrt{T}$. То есть изменение случайной переменной будет иметь математическое ожидание, равное нулю, и среднее квадратическое отклонение, равное квадратному корню из заданного будущего временного периода.

Рис. 10.1 отображает процесс Винера. Заметьте, что степень изменчивости (variability) уровня увеличивается с увеличением интервала Δt, тогда как степень изменчивости инновации постоянна, что далее иллюстрирует рис. 10.2, где представлены три графика одного и того же процесса.

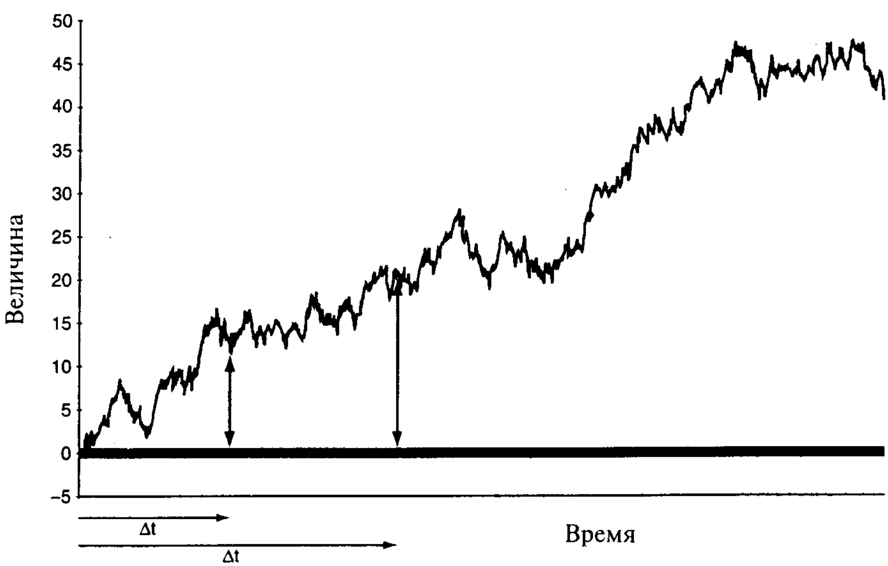

Рис. 10.1. Броуновское движение

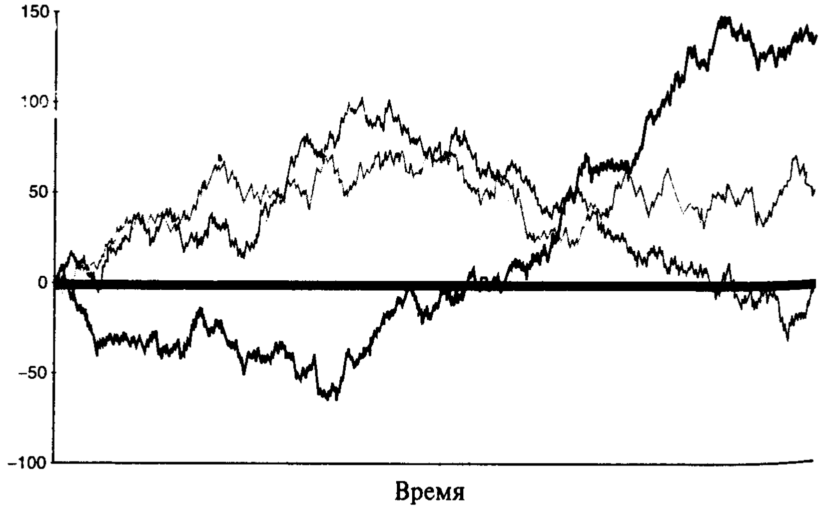

Рис. 10.2. Несколько броуновских движений

Можем ли мы применить процесс Винера для описания стохастических процессов цен активов? К сожалению, не в его настоящей форме. Применение процесса Винера невозможно по следующим трем причинам:

1. Активы характеризуются различными степенями волатильности. В процессе же, описанном выше, волатильность была одна.
2. Рисковые активы имеют положительное ожидаемое среднее значение дохода. В процессе, описанном выше, средняя значений ΔS предполагалась равной нулю, следовательно, в среднем будущая цена не будет отличаться от настоящей.
3. В процессе Винера предполагается, что абсолютные изменения в цене ΔS независимы от величины S. Однако в реальности это не совсем так. Абсолютное изменение цены более дорогого актива будет ожидаться в среднем большим, чем абсолютное изменение более дешевого актива. Мы скорее будем ожидать, что пропорциональные изменения в цене актива $\Delta S/S$ будут независимы от S; пропорциональные или процентные изменения могут быть одинаковыми независимо от цены актива. Более того, как показатель важности данного

изменения цены актива, более подходящим будет измерение относительного дохода $\Delta S/S$, чем просто ΔS.

Проблема 1: разные активы имеют разные степени случайности или волатильности

Во-первых, мы знаем, что разные ценные бумаги характеризуются разной степенью волатильности, поэтому влияние ε будет разным на ценные бумаги, что и отражено в их волатильности.

Эта проблема легко решаема путем умножения ε на σ, что является годовым средним квадратическим отклонением ΔS. Следовательно, ε будет иметь среднее нулевое значение, а среднее квадратическое отклонение будет равно $\sigma \cdot 1 = \sigma$.

Таким образом, уравнение (10.2) принимает вид:

$$\Delta S = \sigma\varepsilon\sqrt{\Delta t}\,. \tag{10.3}$$

А в пределе оно будет выглядеть так:

$$\mathrm{d}S = \sigma\varepsilon\sqrt{\mathrm{d}t}\,. \tag{10.4}$$

Математическое ожидание ΔS равно нулю, среднее квадратическое отклонение — $\sigma\sqrt{\Delta t}$, а дисперсия — $\sigma^2\Delta t$.

Проблема 2: рисковые активы имеют положительный ожидаемый доход

Вторая проблема возникает вследствие того, что математическое ожидание случайной переменной ε равно нулю и, следовательно, ожидаемое изменение значения S — ΔS также равно нулю. Однако ожидаемый доход от рисковых активов, находящихся во владении инвесторов, в среднем должен быть положительным, чтобы вознаградить инвесторов за несение риска. Поэтому мы должны адаптировать основной процесс Винера и привести его к обобщенному процессу Винера, чтобы учесть положительный ожидаемый доход.

Обобщённый процесс Винера (generalized Wiener process) — это основной процесс Винера с добавлением тенденции.

По отношению к стохастическим процессам термин *тенденция* (drift) используется для обозначения положительного или отрицательного тренда во временных рядах стохастической переменной. Когда речь идет о финансовых активах, следует принять **положительную тенденцию** (positive drift), поскольку, как уже отмечалось, рисковые активы должны иметь положительный доход, чтобы компенсировать несение риска инвесторам. Таким образом, тенденция аналогична ожидаемому доходу.

Параметр тенденции α представляет собой изменение S за малый промежуток времени dt. Если бы мы рассматривали только тенденцию, ds, получаемая из ожидаемого дохода α за малую единицу времени, была бы αdt. Однако объединив это с процессом Винера, мы получаем стохастический процесс случайной переменной, характеризующийся как **скоростью тенденции**, так и основными свойствами процесса Винера. Эта случайная переменная имеет ожидаемое изменение по следующим двум причинам: 1) ожидаемый доход за малый интервал времени составляет αdt; 2) случайное изменение $\sigma\varepsilon\sqrt{\Delta t}$, которое было обсуждено в основном процессе Винера. Таким образом, малое изменение в цене актива за малый временной интервал может быть смоделировано следующим **стохастическим дифференциальным уравнением** (stohastic differential equation):

$$\mathrm{d}S = \alpha \mathrm{d}t + \sigma\varepsilon\sqrt{\mathrm{d}t}\,. \tag{10.5}$$

Данное уравнение также можно записать и в дискретной форме:

$$\Delta S = \alpha\Delta t + \sigma\varepsilon\sqrt{\Delta t}\,. \tag{10.6}$$

Так как ΔS распределено нормальным образом, то средняя (математическое ожидание) равно $\alpha\Delta t$, среднее квадратическое отклонение составляет $\sigma\sqrt{\Delta t}$, а дисперсия равна $\sigma^2\Delta t$. Следовательно, σ^2 становится дисперсией за единицу времени; параметр σ известен как волатильность.

Поэтому обобщенный процесс Винера включает основной процесс Винера наравне с элементом тенденции. Элемент тенденции является детерминированным, т.е. неслучайным, а основной процесс Винера представляет собой стохастический элемент. На рис. 10.3 изображен один из таких процессов, где ясно заметно влияние положительной тенденции. Обобщенный процесс Винера может рассматриваться как непрерывный во времени эквивалент субмартингального процесса.

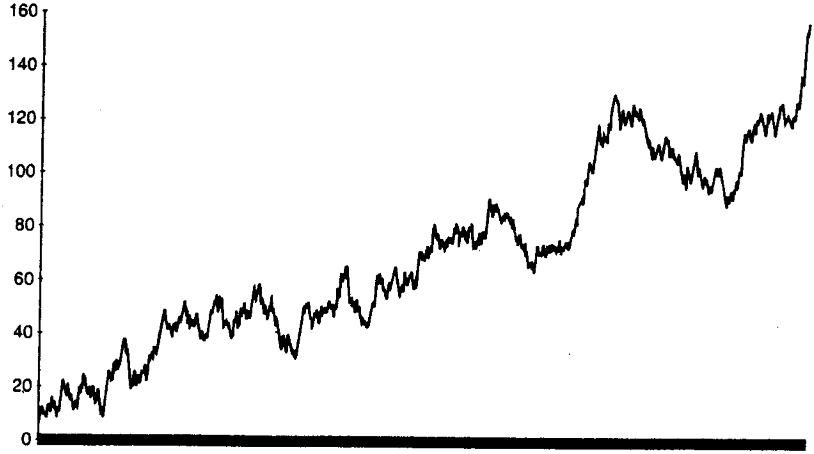

Рис. 10.3. Броуновское движение (с тенденцией)

Проблема 3: величина изменения цены актива должна быть независима от величины цены актива

Для того чтобы понять эту проблему, вспомним, что α — это **абсолютный** доход за единицу времени, являющийся постоянным, но не зависимым от цены актива. Однако инвесторы нуждаются в **процентной ставке** доходности, зависящей от принимаемого риска, и ,следовательно, не зависящей от уровня цены актива. Поэтому если инвесторам нужна 8%-ная доходность по данному активу, они требуют эти 8% независимо от того, стоит актив 1 или 10 фунтов стерлингов.

Следовательно, желаемый стохастический процесс цен активов должен включать абсолютный доход, который является функцией от цены актива, но ставка дохода не должна зависеть от цены актива.

Таким образом, для того чтобы удовлетворить наши требования к модели цен активов, необходимо обобщенный процесс Винера, разработанный до этого момента, заменить на более общий тип стохастического процесса, известный как **процесс Ито** (Ito process).

Процесс Ито — это обобщенный процесс Винера, в котором параметры α (ожидаемый доход) и σ^2 (дисперсия) являются функциями от основных переменных. В общем виде процесс

Ито выглядит как $dx = \alpha(x, t)dt + \sigma(x, t)\varepsilon\sqrt{dt}$. В нашем случае основные переменные — это цена актива S и время t, тогда процесс Ито записывается так: $dS = \delta(S, t)dt + \sigma(S, t)\varepsilon\sqrt{dt}$. Следовательно, если меняются основные переменные, меняется и абсолютная скорость тенденции. Например, с увеличением S увеличивается и α, а с увеличением t увеличиваются α и σ.

Для преобразования нашего процесса Винера в процесс Ито обозначим ожидаемую ставку доходности, выраженную в десятичной форме, как μ, тогда μS будет абсолютным доходом. Для малого промежутка времени Δt ожидаемый абсолютный доход будет $\Delta S = \mu S\Delta t$, что в пределе выглядит как $dS = \mu S dt$. Разделив обе части равенства на S, получим ставку доходности $dS/S = \mu dt$.

Таким образом, мы имеем абсолютное изменение цены актива ΔS, которое является функцией от цены актива ($\mu S\Delta t$), и ставку доходности dS/S, которая не зависит от цены актива (μdt).

Хотя мы и можем установить функциональную зависимость абсолютного ожидаемого дохода от цены актива, степень неопределенности, касающаяся ожидаемого дохода, в течение малого периода времени будет независима от цены актива. Это значит, что инвестор испытывает одинаковую неопределенность относительно будущих доходов независимо от того, равна цена 1 или 10 фунтам стерлингов.

Следовательно, dS/S является функцией от μdt и $\sigma\varepsilon\sqrt{dt}$. Объединив это, получим

$$\frac{dS}{S} = \mu dt + \sigma\varepsilon\sqrt{dt} \qquad (10.7)$$

$$\text{или } dS = \mu S dt + \sigma S\varepsilon\sqrt{dt},$$

что и является процессом Ито. Если процесс Винера часто называется броуновским движением, то рассматриваемый нами процесс Ито иногда называется *геометрическим броуновским движением.*

Рис. 10.4 показывает пример процесса Ито (геометрического броуновского движения). Следует отметить рост, подобный экспоненциальному, что является следствием пропорциональной зависимости ставки роста от уровня S в процессе.

Мы уже заметили, что хотя степень неопределенности будет независима от цены актива, абсолютный доход будет тем больше, чем выше цена актива. Поэтому фактическая разбросанность ожидаемых цен активов будет зависеть от величины текущей цены актива. В результате среднее квадратическое отклонение абсолютного изменения будет тем больше, чем выше цена актива.

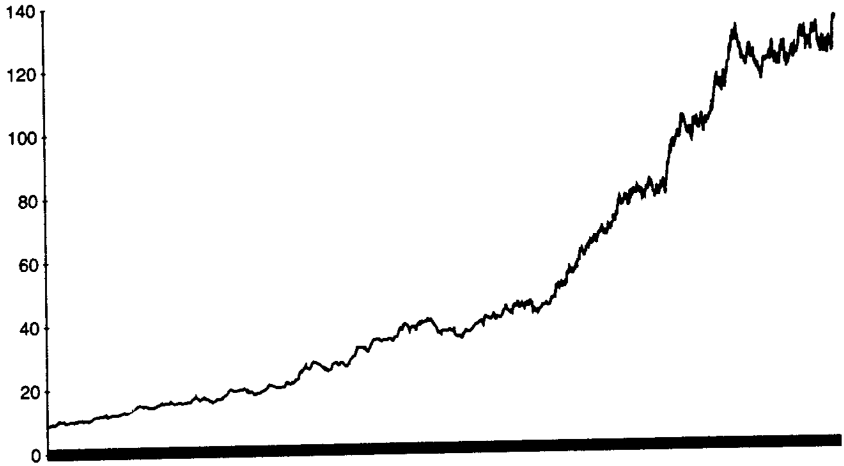

Рис. 10.4. Процесс Ито

Дисперсия фактических изменений цены актива, зависящих от величины этой цены, определяется как

$$\sigma^2 S^2 \Delta t, \tag{10.8}$$

а получаемая мгновенная дисперсия равна

$$\sigma^2 S^2. \tag{10.9}$$

Этот результат получается потому, что если мы установили пропорциональную зависимость среднего квадратического отклонения от цены актива, дисперсия будет пропорциональна квадрату этой цены.

Выражая изменчивость в качестве среднего квадратического отклонения, мы имеем ожидаемый (детерминированный) элемент дохода, представленный как функция от цены актива и времени, также мы имеем стохастический элемент изменчивости цены как функцию от цены актива и времени. Такая модель движения цены актива может быть представлена в форме:

$$dS = \mu S dt + \sigma S \varepsilon \sqrt{dt}, \tag{10.10}$$

которая является процессом Ито, так как цены актива и изменчивость этих цен являются функцией от основных переменных — цены актива S и времени t.

Первый элемент — средняя $\mu S\mathrm{d}t$ — фиксированный, т.е. не стохастический. Второй элемент — стохастический компонент $\sigma S\varepsilon\sqrt{\mathrm{d}t}$ — является причиной стохастической природы всей функции. Кроме того, так как ε выбирается из стандартного нормального распределения, $\mathrm{d}S$ также нормально распределена со средней $\mu S\mathrm{d}t$ и средним квадратическим отклонением $\sigma S\varepsilon\sqrt{\mathrm{d}t}$; также и $\mathrm{d}S/S$ нормально распределена со средней $\mu\mathrm{d}t$ и средним квадратическим отклонением $\sigma\sqrt{\mathrm{d}t}$.

ПРИМЕНЕНИЕ ЛЕММЫ ИТО К ЦЕНООБРАЗОВАНИЮ ПРОИЗВОДНЫХ ФИНАНСОВЫХ ИНСТРУМЕНТОВ

Мы объяснили, каким образом цены активов следуют процессу Ито:

$$\mathrm{d}S = \mu S\mathrm{d}t + \sigma S\varepsilon\sqrt{\mathrm{d}t}\,. \tag{10.11}$$

Мы можем использовать наши знания процесса Ито для определения стоимости производных финансовых инструментов таких, как форварды, фьючерсы и опционы. Ито (1951) доказал, что любая переменная, являющаяся функцией другой переменной, которая следует процессу Ито, сама будет следовать процессу Ито в форме:

$$\mathrm{d}W = \left(\frac{\mathrm{d}W}{\mathrm{d}X}\alpha + \frac{\mathrm{d}W}{\mathrm{d}t} + \frac{1}{2}\frac{\mathrm{d}^2W}{\mathrm{d}X^2}\sigma^2\right)\mathrm{d}t + \frac{\mathrm{d}W}{\mathrm{d}X}\sigma\varepsilon\sqrt{\mathrm{d}t}\,, \tag{10.12}$$

где W — функция (производная) от X, а X следует процессу Ито:

$$\mathrm{d}X = \alpha\mathrm{d}t + \sigma\varepsilon\sqrt{\mathrm{d}t}\,. \tag{10.13}$$

Следовательно, W также следует процессу Ито. Параметры α, σ^2 и σ в уравнении (10.12) — это ожидаемый доход, мгновенная дисперсия и волатильность соответственно, характеризующие переменную x, которая следует процессу Ито. Заметьте, что составным элементом является основной процесс Винера $\varepsilon\sqrt{\delta t}$.

Цена производного финансового инструмента W является функцией от цены ценной бумаги, лежащей в основе контракта, и времени. Переменная W имеет скорость тенденции:

$$\frac{\delta W}{\delta X}\alpha + \frac{\delta W}{\delta t} + \frac{1}{2}\frac{\delta^2 W}{\delta X^2}\sigma^2 \qquad (10.14)$$

и норму волатильности:

$$\left(\frac{dW}{dX}\right)\sigma. \qquad (10.15)$$

Применяя лемму Ито к производным финансовым инструментам, вспомним, что S — это цена актива, лежащего в основе контракта, а ожидаемый доход, дисперсия и среднее квадратическое отклонение основного актива обозначены через μS, $\sigma^2 S^2$ и σS соответственно. Тогда производная переменная W, являющаяся функцией от S и t, будет следовать процессу Ито:

$$dW = \left(\frac{\delta W}{\delta S}\mu S + \frac{\delta W}{\delta t} + \frac{1}{2}\frac{\delta^2 W}{\delta S^2}\sigma^2 S^2\right)dt + \frac{\delta W}{\delta S}\sigma S\varepsilon\sqrt{\delta t}. \qquad (10.16)$$

Ниже мы увидим, каким образом полученный результат (10.16) может быть использован при определении стоимости производных финансовых инструментов.

Ценообразование производных финансовых инструментов в безрисковой среде

Нейтральность к риску — это искусственное (абстрактное) состояние среды, где инвесторы, как предполагается, безразличны к риску. Вследствие этого они не требуют премии за риск, но и не платят другим за несение риска. В результате в безрисковом мире по всем рисковым активам выплачивается безрисковый доход. Подобный мир не относится к тем, которые обычно распознаются как миры, описывающие наши финансовые рынки. Однако данное интеллектуальное порождение довольно полезно в преодолении трех взаимозависимых проблем, связанных с ценообразованием производных финансовых инструментов.

Прежде всего, нам необходимо знать ожидания инвесторов относительно дохода по данному активу. Это даст нам величину скорость тенденции μ.

Во-вторых, нам необходимо знать ожидания инвесторов относительно риска — диапазона распределения вероятностей будущих доходов по активу. Это даст нам среднее квадратическое отклонение.

В-третьих, нам необходимо знать требуемый доход для инвесторов. Это даст нам ставку дисконтирования, которая будет использована в определении стоимости производных финансовых инструментов. Требуемый доход будет связан с ожиданиями инвесторов относительно риска и со степенью неприятия риска ими. Пока же мы не представляем себе, какими будут эти отношения к риску.

Все три переменные — ожидаемый доход, ожидаемый риск и степень неприятия риска — будут различны для инвесторов. Но для двух наших проблем существует искусное решение. Вспомним, что процесс Ито для основной переменной и процесс Ито для производной содержат стохастический элемент $\sigma S\varepsilon\sqrt{dt}$, но в разных пропорциях. Следовательно, возможно занять длинную позицию по основной переменной и короткую по производной переменной таким образом, что эти два стохастических процесса исключат друг друга. На рынке активов это возможно осуществить, например, путем покупки основного актива и продажи некоторого количества производных финансовых инструментов, скажем, опционов на покупку. Вспомним, что это метод, с помощью которого была разработана однопериодная биномиальная модель в гл. 8.

Вспомним также, что стохастический элемент или процесс $\sigma S\varepsilon\sqrt{dt}$ отвечает за стохастическую природу и, следовательно, за рискованность как основной переменной, так и производного финансового инструмента. В результате портфель ценных бумаг, состоящий из длинной позиции по основной переменной и короткой по производной переменной таким образом, что два стохастических процесса исключают друг друга, будет безрисковым. Поэтому на эффективных финансовых рынках этот портфель (или комбинация) должен иметь безрисковую процентную ставку, что означает, что только безрисковая процентная ставка может быть использована для дисконтирования будущей стоимости и приведения ее к текущей.

Для того чтобы это стало ясным, вспомним, что процесс Ито для основного актива выглядит так:

$$dS = \mu S dt + \sigma S\varepsilon\sqrt{dt}\,, \qquad (10.17)$$

а процесс Ито для W — производной от S — так:

$$\delta W = \left(\frac{\delta W}{\delta S}\mu S + \frac{\delta W}{\delta t} + \frac{1}{2}\frac{\delta^2 W}{\delta S^2}\sigma^2 S^2\right)dt + \frac{\delta W}{\delta S}\sigma S\varepsilon\sqrt{dt}\,. \qquad (10.18)$$

Представив актив S и соответствующий производный финансовый инструмент W с помощью соответствующих процессов Ито, построим портфель П, состоящий из одной короткой позиции по производному финансовому инструменту и длинной позиции по $\delta W/\delta S$ единиц основного актива. Тогда **изменение** в стоимости портфеля $\delta\Pi$ будет равно:

$$\delta\Pi = -\left[\left(\frac{\delta W}{\delta S}\mu S + \frac{\delta W}{\delta t} + \frac{1}{2}\frac{\delta^2 W}{\delta S^2}\sigma^2 S^2\right)\mathrm{d}t + \frac{\delta W}{\delta S}\sigma S\varepsilon\sqrt{\mathrm{d}t}\right] +$$
$$+\frac{\delta W}{\delta S}(\mu S\mathrm{d}t + \sigma S\varepsilon\sqrt{\mathrm{d}t}) = \qquad (10.19)$$
$$= \left(-\frac{\delta W}{\delta t} - \frac{1}{2}\frac{\delta^2 W}{\delta S^2}\sigma^2 S^2\right)\delta t\,.$$

Чтобы показать, как был получен конечный результат, раскроем скобки в уравнении (10.19) и сократим слагаемые:

$$\delta\Pi = -\left[\frac{\delta W}{\delta S}\mu S\mathrm{d}t + \frac{\delta W}{\delta t}\mathrm{d}t + \frac{1}{2}\frac{\delta^2 W}{\delta S^2}\sigma^2 S^2\mathrm{d}t + \frac{\delta W}{\delta S}\sigma S\varepsilon\sqrt{\mathrm{d}t}\right] +$$
$$+\frac{\delta W}{\delta S}\mu S\mathrm{d}t + \frac{\delta W}{\delta S}\sigma S\varepsilon\sqrt{\mathrm{d}t} = \qquad (10.20)$$
$$= -\frac{\delta W}{\delta t}\mathrm{d}t - \frac{1}{2}\frac{\delta^2 W}{\delta S^2}\sigma^2 S^2\mathrm{d}t =$$
$$= \left(-\frac{\delta W}{\delta t} - \frac{1}{2}\frac{\delta^2 W}{\delta S^2}\sigma^2 S^2\right)\mathrm{d}t\,.$$

Вспомним, что наш портфель П состоит из одной короткой позиции по производному финансовому инструменту W и длинной позиции по $\delta W/\delta S$ единиц основного актива S. Так как портфель является безрисковым в каждом коротком промежутке времени, он имеет мгновенную безрисковую процентную ставку. Поэтому любое изменение П за малый временной промежуток должно быть равно произведению безрисковой процентной ставки и стоимости безрискового портфеля за тот же малый временной интервал, тогда

$$\delta\Pi = \left(-\frac{\delta W}{\delta t} - \frac{1}{2}\frac{\delta^2 W}{\delta S^2}\sigma^2 S^2\right)\delta t = r\left(-W + \frac{\delta W}{\delta t}S\right)\delta t\,. \qquad (10.21)$$

Раскрыв скобки и переставив слагаемые, получим

$$\frac{\delta W}{\delta t} + rS\frac{\delta W}{\delta S} + \frac{1}{2}\sigma^2 S^2 \frac{\delta^2 W}{\delta S^2} = rW \,. \tag{10.22}$$

Этот результат особенно полезен, так как он означает, что ожидаемый доход или скорость тенденции μ может быть заменена безрисковой процентной ставкой r. Вследствие этого отпадает необходимость в том, чтобы знать ожидаемую ставку дохода по основному активу в процессе определения цены производного финансового инструмента, следовательно, первая проблема решена. Помимо этого, так как портфель является безрисковым, по крайней мере в течение мгновенного временного интервала dt, только безрисковая процентная ставка должна быть использована для дисконтирования будущей стоимости производного финансового инструмента к настоящей или текущей цене.

Однако наша вторая проблема еще не решена, нам попрежнему необходимо знать среднее квадратическое отклонение или волатильность основной переменной! В сущности мы должны знать распределение вероятностей основного актива.

Уравнение с частными производными Блэка—Сколса и ценообразование опционов

Уравнение (10.22) — это уравнение в частных производных, выведенные Блэком и Сколсом (1973) для определения стоимости любого производного финансового инструмента. Для разных типов производных финансовых инструментов уравнение имеет разные решения, которые зависят от ограничивающих условий для каждого из этих типов оцениваемых производных финансовых инструментов. Для Европейских опционов ограничивающие условия те же, что и приведенные в гл. 8, а именно:

$$C = \max\,[0, S-X],$$

$$P = \max\,[0, X-S].$$

Ограничивающие условия играют особенно важную роль в решении уравнений с частными производными, поскольку решение таких уравнений аналогично процессу интегрирования, рассмотренному в гл. 3. Уравнения с частными производными описывают скорость изменения переменной, но не описывают уровень значения этой переменной. Для этого нам нужна другая информация, которая и предоставляется ограничивающими условиями.

ДОПУЩЕНИЯ — ПРОЦЕСС ИТО И ЛОГНОРМАЛЬНОСТЬ

Среди различных допущений, сделанных Блэком и Сколсом, есть два особенно важных. Первое — цены ценных бумаг следуют процессу Ито. Второе — цены ценных бумаг распределены логнормально и, как следствие, непрерывно наращенный доход распределен нормальным образом. Эти два допущения, как будет сейчас показано, неразрывно связаны друг с другом.

Процесс Ито

Предполагается, что цена ценной бумаги S следует процессу

$$\mathrm{d}S = \mu S \mathrm{d}t + \sigma S \varepsilon \sqrt{\mathrm{d}t}\,. \tag{10.23}$$

Натуральный логарифм этой переменной будет нормально распределен:

$$\ln S_t \sim N\left[\ln S_t + \left(\mu - \frac{\sigma^2}{2}\right)(T - t), \sigma\sqrt{T - t}\right], \tag{10.24}$$

где t — текущее время, T — будущая точка во времени, а $(T—t)$ — период времени, за который анализируется ценная бумага.

Например, пусть $W = \ln S$, так как W — функция S, тогда если S следует процессу Ито, W также будет следовать процессу Ито:

$$\begin{gathered} W = \ln S, \\ \frac{\delta \ln S}{\delta S} = \frac{1}{S}, \\ \frac{\delta^2 \ln S}{\delta S^2} = -\frac{1}{S^2}, \\ \frac{\delta \ln S}{\delta t} = 0\,. \end{gathered} \tag{10.25}$$

Вспомнив гл. 3, можно увидеть, что это действительно так.

Вспомним также, что процесс Ито для переменной W был дан в виде

$$\delta W = \left(\frac{\delta W}{\delta S}\mu S + \frac{\delta W}{\delta t} + \frac{1}{2}\frac{\delta^2 W}{\delta S^2}\sigma^2 S^2\right)\mathrm{d}t + \frac{\delta W}{\delta S}\sigma S\varepsilon\sqrt{\mathrm{d}t}, \quad (10.26)$$

тогда, подставив значения выражений (10.25) в уравнение (10.26), получим

$$\mathrm{d}\ln S = \left(\frac{1}{S}\mu S + 0 + \left(-\frac{1}{2}\frac{1}{S^2}\sigma^2 S^2\right)\mathrm{d}t + \frac{1}{S}\sigma S\varepsilon\sqrt{\mathrm{d}t}\right). \quad (10.27)$$

Далее, сделав преобразование и вспомнив, что ε по допущению равно единице, получим:

$$\mathrm{d}\ln S = \left(\mu - \frac{\sigma^2}{2}\right)\mathrm{d}t + \sigma\sqrt{\mathrm{d}t}. \quad (10.28)$$

Таким образом, $\ln S$ следует процессу Винера.

Логнормальное распределение

В гл. 4 мы показали, что переменная распределена логнормально, если ее натуральный логарифм распределен нормально. Кроме того, логнормальное распределение является привлекательным с точки зрения применения его по отношению к ценам активов, так как диапазон этого распределения находится в интервале от 0 до $+\infty$. Этот диапазон в точности схож с теоретическим диапазоном цен активов потому, что они не могут быть отрицательными, но могут принимать очень высокие положительные значения.

Также в гл. 4 мы показали, что если текущая цена актива S_t в действительности логнормально распределена, цена в будущий момент времени T, $\ln S_t$ будет нормально распределена:

$$\ln S_t \sim N\left[\ln S_t + \left(\mu - \frac{\sigma^2}{2}\right)(T - t), \sigma\sqrt{T - t}\right]. \quad (10.29)$$

Известно, что $\ln S_T—\ln S_t = \ln S_T/S_t$. Следовательно, разница между двумя логарифмами переменной является непрерывно наращенной ставкой доходности за период $T—t$, которая нормально распределена со следующими средней и средним квадратическим отклонением:

$$\text{Средняя} = \left(\mu - \frac{\sigma^2}{2}\right)(T - t), \tag{10.30}$$

$$\sqrt{D} = \sigma\sqrt{T - t}\ .$$

Если мы перепишем уравнение (10.30) для очень малых временных интервалов, являющихся основой математики непрерывных процессов, то получим следующее:

$$\text{Средняя} = \left(\mu - \frac{\sigma^2}{2}\right)\delta t, \tag{10.31}$$

$$\sqrt{D} = \sigma\sqrt{\delta t}\ ,$$

что является в точности тем, что мы получили исходя из предположения, что цены активов следуют процессу Ито. Следовательно, предположения, что процесс Ито лежит в основе движения цен активов и цены активов логнормально распределены, являются совместимыми.

Блэк и Сколс признали, что уравнение с частными производными, которое они вывели, аналогично уравнению, описывающему тепловые диффузионные процессы в твердых телах. Решение так называемого "теплового уравнения" было уже описано Черчилем (Churchill) (1963). Поэтому с заданными ограничивающими условиями и предположениями, что r и σ постоянны, они имели возможность вывести точное и однозначное решение для стоимости Европейского опциона на актив, в течение срока действия которого не выплачивается денежных средств таких, как дивиденды.

Точное решение для Европейского опциона на покупку выглядит так:

$$Xc = SN(d_1) - Xe^{-r(T-t)}N(d_2), \tag{10.32}$$

где

$$d_1 = \frac{\ln(S / X) + (r + \sigma^2 / 2)(T - t)}{\sigma\sqrt{(T - t)}}, \tag{10.33}$$

$$d_2 = d_1 - \sigma\sqrt{(T - t)}\ , \tag{10.34}$$

c — премия опциона на покупку,
S — текущая цена актива,
X — цена исполнения,
T—t — время до момента исполнения, выраженное в десятичных долях года,

σ	— среднее квадратическое отклонение цены актива в виде десятичной дроби,
ln	— натуральный логарифм,
$N(\cdot)$	— кумулятивное стандартное нормальное распределение вероятностей,
d_1 и d_2	— стандартизованные нормальные переменные,
r	— безрисковая процентная ставка, представленная десятичной дробью.

Рассмотрим пример, предположив, что текущая цена актива равна 35,0 единиц, цена исполнения равна 35,0 единиц, безрисковая процентная ставка 10%, волатильность составляет 20%, а время до момента исполнения один год. Т.е. $S = 35$, $X = 35$, $(T-t) = 1{,}0$, $r = 0{,}1$ и $\sigma = 0{,}2$.

Сначала, мы должны рассчитать d_1, d_2 и текущую величину цены исполнения $Xe^{-r(T-t)}$:

$$d_1 = \frac{\ln(35/35) + (0{,}1 + 0{,}2^2/2)\cdot 1{,}0}{0{,}2\sqrt{1{,}0}} = 0{,}60\ ,$$

$$d_2 = d_1 - 0{,}2\sqrt{1{,}0} = 0{,}40$$

и

$$Xe^{-r(T-t)} = 35e^{-(0{,}1\ \cdot\ 1{,}0)} = 31{,}66934.$$

Тогда уравнение для опциона на покупку примет вид:

$$c = 35N(0{,}6) - 31{,}6693N(0{,}4).$$

На следующем шаге по таблицам находятся значения **кумулятивного стандартизованного нормального распределения вероятностей** в точках 0,6 и 0,4. Альтернативным методом для расчета значений кумулятивного стандартного нормального распределения вероятностей является нахождение функции в виде многочлена, что было объяснено в гл. 8.

Зная, что d_1 — это стандартизованная нормально распределенная случайная переменная и что $N(d_1)$ — кумулятивное стандартизованное нормальное распределение, $N(d_1)$ отображает величину площади под стандартизованной нормальной кривой от $z = -\infty$ до $z = d_1$.

Эти значения равны 0,7257 и 0,6554. Подставим их в уравнение и получим:

$$c = 35(0{,}7257) - 31{,}6693(0{,}6554) = 4{,}6434.$$

Таким образом, искомая цена опциона на покупку равна 4,6434. Читатель может сравнить этот результат с теми, что были получены по биномиальной и триномиальной моделям, рассмотренным в гл. 8.

УПРАЖНЕНИЯ

1. Объясните, как вы понимаете термин "марковский процесс".
2. Объясните, что такое основной процесс Винера. Почему если переменная следует основному процессу Винера, средняя изменений будет 0, а среднее квадратическое отклонение $\sqrt{T}$?
3. Что вы понимаете под термином «обобщенный процесс Винера»?
4. Что такое процесс Ито? Почему он более приемлем для описания стохастических процессов цен активов, чем основной процесс Винера или обобщенный процесс Винера?
5. Объясните, используя процесс Ито, каким образом возможно оценить производные финансовые инструменты в нейтральной к риску среде.
6. Объясните, каким образом два предположения модели Блэка—Сколса ценообразования опционов: 1) цены ценных бумаг следуют процессу Ито; 2) цены ценных бумаг логнормально распределены, совместимы друг с другом.
7. Используя модель Блэка—Сколса для Европейского опциона, по которому не выплачиваются дивиденды, найдите стоимость трехмесячного опциона **на покупку**, где цена актива и цена исполнения равны 35 единицам, краткосрочная процентная ставка 10% годовых, а волатильность 20%.
8. Сравните ваш ответ на вопрос 7 с ответом на вопрос 3 из гл. 8. Объясните различия, если они имеются.
9. Объясните, каким образом методы конечной разницы могут быть применены для решения уравнений с частными производными. Объясните разницу между применением определенного метода конечной разницы и неопределенного метода конечной разницы.
10. Объясните, как можно преодолеть возникающие проблемы для достижения сближения и устойчивости в определенном методе конечной разницы.
11. Используя данные из вопроса 7, определите стоимость опциона на покупку с помощью шестипериодной решетки. Сравните ваш результат с ответом на вопрос 3 из гл. 8.

ОТВЕТЫ К ИЗБРАННЫМ ВОПРОСАМ

7. 1,85.

11.

	X	Δt	r	ΔS	волатильность		
	35	0,04	0,1	2,5	0,2		
			0,0066667				
50							15
47,5						12,645228	
12,5							
45					10,289854	10,145228	10
42,5				7,9338793	7,7898538	7,6452282	7,5
40			5,603448	5,443196	5,2898538	5,1452282	5
37,5		3,4588128	3,2569941	3,0500355	2,8418287	2,6452282	2,5
35	7,7827782	1,5772272	1,3558335	1,1109992	0,8284516	0,4792531	0
32,5		0,4649837	0,3272028	0,1958625	0,0801407	0	0
30			0,0370406	0,0115722	0	0	0
27,5				0	0	0	0
25					0	0	0
22,5						0	0
20							0

СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ

Black, F. and Scholes, M. (1973) The pricing of options and corporate liabilities. *Journal of Political Economy*, **81**, (May—June), pp. 637—59.

Churchill, R.V. (1963) *Fourier Series and Boundary Value Problems*, 2nd edn. McGraw—Hill, New York.

Hull, J.C. (1993) *Options, Futures and other Derivatives*, 2nd edn. Prentice—Hall, Englewood Cliffs, NJ.

Ito, K. (1951) On stochastic differential equations. *Memoirs, American Mathematical Society*, **4**, pp.1—51.

Watsham, T.J. (1992) *Options and Futures in International Portfolio Management*. Chapman & Hall, London.

ПРИЛОЖЕНИЕ 10.1: Методы конечной разницы в приложении к уравнению в частных производных Блэка—Сколса

В гл. 10 мы уже показали, что объединение допущения о нейтральности риска с допущением, что цены активов следуют процессу Ито, позволяет нам вывести уравнение частных производных, пригодное для применения по отношению к любым производным финансовым инструментам. Это уравнение Блэка—Сколса еще раз приведено ниже:

$$\frac{\delta W}{\delta t}+rS\frac{\delta W}{\delta S}+\frac{1}{2}\sigma^2 S^2\frac{\delta^2 W}{\delta S^2}=rW\,. \qquad (\text{П.10.1})$$

Подобные уравнения называются уравнениями с частными производными, так как они содержат частные производные одной или нескольких переменных. В этом случае мы имеем уравнение, включающее первую производную цены опциона по отношению ко времени $\delta W/\delta t$, первую производную цены опциона по отношению к цене основного актива $\delta W/\delta S$ и производную $\delta W/\delta S$ также по отношению к цене данного актива.

Однозначное решение для таких дифференциальных уравнений можно получить только при наличии строгих ограничений (одно из которых было рассмотрено ранее в этой главе). В случае уравнения Блэка—Сколса мы имеем Европейский опцион, по которому не выплачиваются дивиденды, а его стоимость определяется ограничивающими условиями, приведенными ранее. Другие типы опционов, например, так называемые "экзотические опционы", имеют более сложные ограничивающие условия, а однозначное решение часто невозможно получить, в этих случаях необходимо прибегнуть к числовым методам. К счастью, существует множество таких подходов, которые легко понимаются и применяются. Метод конечной разницы — один из таких подходов. В действительности существует два подхода — определенный метод конечной разницы и неопределенный метод конечной разницы.

В этом методе частные производные в дифференциальных уравнениях заменяются на **конечные разностные приближения** (finite difference approximations). В определенном методе конечной разницы эти приближения выглядят следующим образом:

$$\begin{aligned}
&\frac{\delta W}{\delta t}\approx\frac{W(t+\Delta t,S)-W(t,S)}{\Delta t}\\
&\frac{\delta W}{\Delta S}\approx\frac{W(t+\Delta t,S+\Delta S)-W(t+\Delta t,S-\Delta S)}{2\Delta S}\\
&\frac{\delta^2 W}{\delta S^2}\approx\frac{W(t+\Delta t,S+\Delta S)-2W(t+\Delta t,S)+W(t+\Delta t,S-\Delta S)}{\Delta S^2}
\end{aligned} \qquad (\text{П.10.2})$$

Проделав замену в дифференциальном уравнении Блэка—Сколса в соответствии с этим, получим следующее:

$$\frac{W(t+\Delta t,S)-W(t,S)}{\Delta t}+rS\frac{W(t+\Delta t,S+\Delta S)-W(t+\Delta t,S-\Delta S)}{2\Delta S}+$$
$$+\frac{1}{2}\sigma^2 S^2\frac{W(t+\Delta t,S+\Delta S)-2W(t+\Delta t,S)+W(t+\Delta t,S-\Delta S)}{\Delta S^2}=rW(t,S) \quad (\text{П.10.3})$$

После некоторых алгебраических манипуляций уравнение примет вид:

$$W(t,S)=\frac{1}{1+r\Delta t}\left[\frac{S}{2\Delta S}\Delta t\left(\frac{S}{\Delta S}\sigma^2+r\right)W(t+\Delta t,S+\Delta S)+\right.$$
$$\left.\left(1-\left(\frac{S}{\Delta S}\right)^2\sigma^2\Delta t\right)W(t+\Delta t,S)+\frac{S}{2\Delta S}\Delta t\left(\frac{S}{\Delta S}\sigma^2-r\right)W(t+\Delta t,S-\Delta S)\right].$$

(П.10.4)

Мы проиллюстрируем эту технику, применив ее к тому же самому опциону, что и для модели Блэка—Сколса.

Процесс начинается с построения решетки, в которой по вертикали помещены цены актива, возрастающие через равные интервалы, а по горизонтали — равномерно распределенные временные отрезки. Разница между каждыми равномерно изменяющимися ценами составляет ΔS, а временными точками — Δt. Ограничивающие условия для опциона дадут нам возможные значения стоимости опциона на момент исполнения (в момент времени T), они составят значение самого правого столбца решетки. Эти значения будут аналогичны стоимостям на момент исполнения, расположенным у правой границы биномиального или триномиального дерева.

В нашем примере таблица с ценой актива S, изменяющейся от 10 до 60 с шагом ΔS = 2,5, и временным шагом Δt, равным одной десятой года 0,1, представлена на рис. 10.5.

Так как оцениваемый опцион имеет цену исполнения 35, то правый граничный столбец решетки соответствует ограничивающим условиям для опциона, т.е. $c = \max[S_T—35,0]$.

Рассмотрим рис. 10.6, представляющий собой уменьшенную копию рис. 10.5.

Самый правый столбец решетки (время T) связан с ограничивающими условиями для цены опциона, относящимися в данном случае к опциону с ценой исполнения 35.

Рассмотрим три верхних значения столбца. Ячейка со значением стоимости опциона 22,50 относится к точке $W(t + \Delta t, S)$. Значение 25,00 в самой верхней ячейке соответствует $W(t + \Delta t, S + \Delta S)$, а значение 20,00 — $W(t + \Delta t, S—\Delta S)$.

	X	Δt	r	ΔS	волатильность						
	35	0,10	0,1	2,5	0,2						
60											25
57,5										22,84653	22,5
55									20,68964	20,34653	20
52,5			0,016					18,52934	18,18964	17,84653	17,5
50							16,36569	16,02934	15,68964	15,34653	15
47,5						14,1987	13,86569	13,52934	13,18964	12,84653	12,5
45					12,05241	11,73283	11,36569	11,02934	10,68961	10,34653	10
42,5				9,975935	9,658943	9,26024	8,927471	8,529345	8,189638	7,846535	7,5
40			8,015471	7,687618	7,285476	6,939675	6,511311	6,155919	5,689638	5,346535	5
37,5		6,210491	5,877035	5,49624	5,141786	4,728753	4,352501	3,876839	3,485565	2,846535	2,5
35	4,600594	4,281332	3,938053	3,600486	3,232397	2,871997	2,459606	2,070626	1,546642	1,143564	0
32,5		2,647781	2,355079	2,051135	1,74779	1,427594	1,113621	0,763501	0,456294	0	0
30			1,187463	0,963614	0,742513	0,532283	0,329068	0,157218	0	0	0
27,5				0,337589	0,220145	0,120385	0,046231	0	0	0	0
25					0,036596	0,011443	0	0	0	0	0
22,5						0	0	0	0	0	0
20							0	0	0	0	0
17,5								0	0	0	0
15									0	0	0
12,5										0	0
10											0

Рис. 10.5

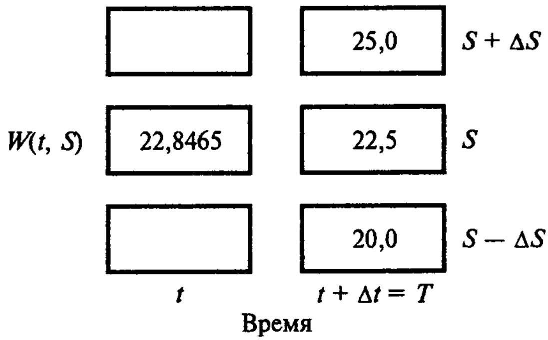

Рис. 10.6

Затем эти три верхних конечных значения подставляются в уравнение (П.10.4) наравне с величинами S, ΔS, Δt, волатильностью и безрисковой процентной ставкой. Значение S — это значение, находящееся в самом левом столбце решетки и соответствующее среднему значению конечных стоимостей опциона во время T, т.е. 57,5 для нашей тройки.

Непосредственный расчет показан ниже. $\Delta t = 0{,}1$, $r = 0{,}10$, $\Delta S = 2{,}50$ и $\sigma = 0{,}20$. Если $S = 57{,}5$, то $S + \Delta S = 60{,}0$ и $S{-}\Delta S = 55{,}0$. Цена исполнения опциона равна 35, а соответствующие значения цен опциона на момент времени T — 25, 22,50 и 20. Подставив эти значения в уравнение (П.10.4), получим:

$$W(t,S) = \frac{1}{1+(0{,}1\cdot 0{,}1)}\left[\frac{\left(\dfrac{57{,}5}{2{,}5}\right)0{,}1}{2}\left(\frac{57{,}5}{2{,}5}0{,}2^2+0{,}1\right)25+\right.$$

$$\left.+\left(1-\left(\frac{57{,}5}{2{,}5}\right)^2 0{,}2^2 0{,}1\right)22{,}5+\frac{\dfrac{57{,}5}{2{,}5}0{,}1}{2}\left(\frac{57{,}5}{2{,}5}0{,}2^2-0{,}1\right)20\right] = 22{,}8465$$

Следовательно, стоимость опциона на покупку в точке $W(S,t)$ в решетке равна 22,8465. Этот процесс повторяется после смещения на одну ячейку вниз, т.е. используются три соответствующих значения для момента времени $t + \Delta t$. В нашем примере эти значения — 22,5, 20 и 17,5. Далее, подставив эти значения в выражение (П.10.4), получим стоимость опциона 20,3465. Т.е. мы просчитываем значения, двигаясь от ограничивающих условий к левому краю решетки до тех пор, пока не достигнем текущего временного периода и цены опциона 4,6006. Аналогия между этим методом и биномиальной и триномиальной моделями, обсужденными в гл. 8, должна быть очевидна.

Заметим, что может быть необходимо использовать следующие ограничивающие условия:

$W(t,0) = 0$ для опциона на покупку, или X для опциона на продажу (X — цена исполнения),
Для больших S $W(t, S) = S—X$ для опциона на покупку, или 0 для опциона на продажу.

Для Американского опциона на покупку на каждом этапе вычислений рассчитанное значение $W(t,S)$ сравнивается с $S—X$ ($X—S$ в случае опциона на продажу) и заменяется на него, если $S—X$ ($X—S$) больше.

Это сходно с ранним примером и иллюстрирует гибкость численного подхода в возможности справиться с опционами, отличными от самых простых и элементарных.

Сближение и устойчивость

Определенный метод конечной разницы позволяет найти цены опционов для любых комбинаций цен активов и временных интервалов до момента исполнения. Однако численные методы порождают проблемы, связанные с многочисленными повторяющимися округлениями результатов, которые могут привести к серьезным отклонениям при расчете соседних значений и/или отрицательным результатам, т.е. **неустойчивости**.

Сравним значение 4,6006 со значением 4,57 для десятиступенчатой биномиальной модели и со значением 4,64 для модели Блэка—Сколса. Как следствие, возникает необходимость в увеличении точности наших приближенных оценок. Как и в биномиальной модели, мы можем рассмотреть сокращение временных интервалов и интервалов между значениями S в сетке (т.е. вертикальных интервалов в вышеприведенной схеме, известных как размер сетки), чтобы уменьшить ошибки приближенных вычислений и **приблизиться** к реальному результату.

К счастью, можно доказать, что две цели — достижение сближения и избежание неустойчивости — могут быть достигнуты при условии, что если при проделанных сокращениях для временных интервалов и размера сетки значение отношения $\Delta S/S$ останется между 0 и 0,5. Это, однако, порождает свои проблемы, так как видно, что уменьшение размера сетки в два раза должно быть компенсировано сокращением временного шага в четыре раза, а время вычислений увеличится соответственно в восемь раз.

Неопределенный метод конечной разницы

На практике требуется большое количество точных результатов, поэтому вычислительная эффективность становится особенно важной. Альтернативный подход — неопределенный метод конечной разницы — использует следующие более явные и более точные конечные разностные приближения для частных производных:

$$\frac{\delta W}{\delta t} \approx \frac{W(t+\Delta t,S)-W(t,S)}{\Delta t},$$
$$\frac{\delta W}{\delta S} \approx \frac{W(t,S+\Delta S)-W(t,S-\Delta S)}{2\Delta S}, \qquad (\text{П.10.5})$$
$$\frac{\delta^2 W}{\delta S^2} \approx \frac{W(t,S+\Delta S)-2W(t,S)+W(t,S-\Delta S)}{\Delta S^2}.$$

Если в определенном методе конечной разницы **точная** формула изменятся с учетом $W(t + \Delta t, S + \Delta S)$, $W(t + \Delta t, S)$ и $W(t + \Delta t, S—\Delta S)$, то в **неопределённой** версии мы должны решить систему уравнений, чтобы найти $W(t, S)$. Это необходимо потому, что структура разностей имеет вид, показанный на рис. 10.7

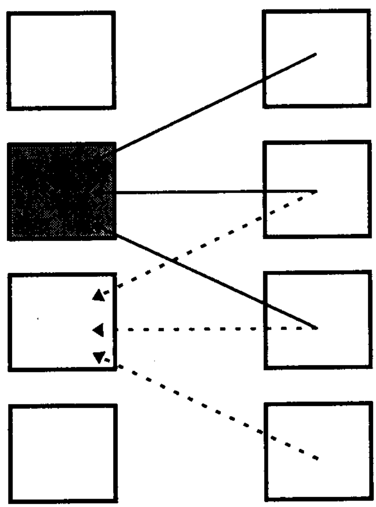

Рис. 10.7

Значение в заштрихованной ячейке будет участвовать в определении трех индивидуальных приближенных оценок, как показывают три отдельные стрелки. Это означает, что обратный итерационный процесс, начинающийся с известных значений в момент времени $t + \Delta t$, будет включать в себя решение системы уравнений. Количество решаемых уравнений является функцией от размера сетки.

Заметим, что, как и раньше, может быть необходимым использовать верхнее и нижнее ограничивающие условия:

$W(t,0) = 0$ для опциона на покупку, или X для опциона на продажу (X — цена исполнения).

Для больших S $W(t,S) = S-X$ для опциона на покупку, или 0 для опциона на продажу.

ПРИЛОЖЕНИЕ 10.2: Получение уравнения Блэка—Скоулса с использованием ожиданий

Мы можем рассчитать стоимость опциона как ожидание — вероятностно взвешенную сумму. Мы осуществим вывод уравнения для опциона на покупку и приведем результат для опциона на продажу. Предпримем элементарный подход, который будет содержать неизбежно скучную алгебру.

Вспомним, что стоимость Европейского опциона на покупку на момент исполнения равна $\max(0, S_T - X)$, где X — это цена исполнения.

$$E\,(\text{стоимость в момент времени } T) = \int_X^{\infty} (S_T - X) f(S_T)\,dS_T =$$

$$= \int_X^{\infty} (S - X) f(S)\,dS \quad \text{(это известно как стохастический интеграл)}$$

$$= \int_X^{\infty} S f(S)\,dS - \int_X^{\infty} X f(S)\,dS,$$

где f — функция плотности вероятности для S_T.

Найдем отдельно каждый интеграл. Начнем со второго, так как он легче:

$$\int_X^{\infty} X f(S)\,dS = \int_X^{\infty} f(S)\,dS =$$

$$= X\left(1 - N\left(\frac{\ln X - \left(\ln S(t) + \left(\mu - \frac{\sigma^2}{2}\right)(T - t)\right)}{\sigma\sqrt{(T - t)}}\right)\right),$$

так как S логнормально распределена, N возвращает левостороннюю вероятность стандартного нормального распределения (см. рис. 10.8):

$$= X\left(1 - N\left(\frac{-\ln X\dfrac{S}{X} - \left(\mu - \dfrac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{(T-t)}}\right)\right) = XN\left(\frac{\ln X\dfrac{S}{X} + \left(\mu - \dfrac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{(T-t)}}\right)$$

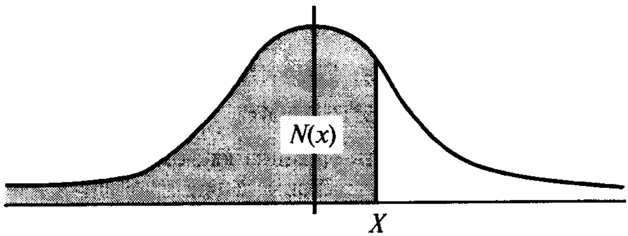

Рис. 10.8

Первый интеграл труднее, так как необходима подстановка $W = \ln S$, что, как следствие, обяжет нас отыскать функцию плотности вероятности для W (см. рис. 10.9)

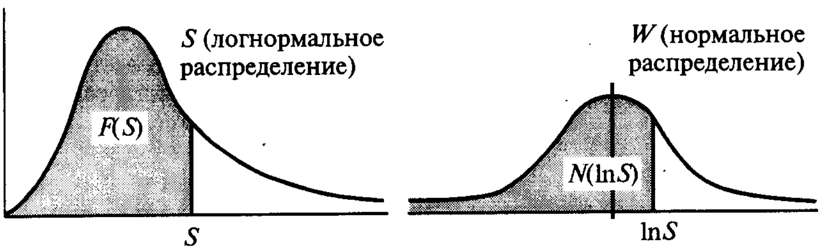

Рис. 10.9

$F(S) = N(\ln S)$, тогда

$$f(S) = n(\ln S)\cdot\frac{1}{S} = \frac{n(w)}{e^w},$$

где n — функция плотности вероятности для нормального распределения.

$$\therefore \int_X^\infty Sf(S)\mathrm{d}S = \int_{\ln X}^\infty e^w \frac{n(w)}{e^w} e^w \mathrm{d}w =$$

$$= \int_{\ln X}^{\infty} e^{w} \frac{1}{b\sqrt{2\pi}} \exp\left(\frac{-(w-a)^2}{2b^2}\right) dw = ,$$

где $\quad a = \ln S + \left(\mu - \frac{\sigma^2}{2}\right)$ и $b = \sigma\sqrt{T-t}$

$$= \int_{\ln X}^{\infty} \frac{1}{b\sqrt{2\pi}} \exp\left(\frac{-(w^2 - 2aw + a^2 - 2wb^2)}{2b^2}\right) dw =$$

$$= \int_{\ln X}^{\infty} \frac{1}{b\sqrt{2\pi}} \exp\left(\frac{-[(w-(a+b^2))^2 - 2ab^2 - b^4]}{2b^2}\right) dw =$$

$$= \int_{\ln X}^{\infty} \frac{1}{b\sqrt{2\pi}} \exp\left(\frac{-(w-(a+b^2))^2}{2b^2}\right) \exp\left(a + \frac{b^2}{2}\right) dw =$$

$$= \exp\left(a + \frac{b^2}{2}\right) \int_{\ln X}^{\infty} \frac{1}{b\sqrt{2\pi}} \exp\left(\frac{-(w-(a+b^2))^2}{2b^2}\right) dw =$$

$$= \exp\left(a + \frac{b^2}{2}\right)\left(1 - N\left(\frac{\ln X - (a+b^2)}{2b^2}\right)\right) =$$

$$= \exp\left(\ln S + \left(\mu - \frac{\sigma^2}{2}\right)(T-t) + \frac{\sigma^2(T-t)}{2}\right)\left(1 - N\left(\frac{\ln X - \ln S - \left(\mu - \frac{\sigma^2}{2}\right)(T-t) + \sigma^2(T-t)}{\sigma\sqrt{(T-t)}}\right)\right)$$

$$= e^{\mu T} S\left(1 - N\left(\frac{-\ln\frac{S}{X} - \left(\mu + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{(T-t)}}\right)\right) = e^{\mu T} S N\left(\frac{\ln\frac{S}{X} + \left(\mu + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{(T-t)}}\right).$$

Таким образом, ожидаемая стоимость опциона в момент времени T — это $e^{\mu(T-t)}S(t)N(d_1)-XN(d_2)$, где

$$d_1 = \frac{\ln\frac{S}{X} + (\mu + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{(T-t)}} \text{ и } d_2 = \frac{\ln\frac{S}{X} + (\mu - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{(T-t)}}.$$

И в заключение мы должны осуществить дисконтирование и привести стоимость к настоящему времени t. Но здесь существует проблема — что использовать в качестве ставки дисконтирования?

Если норма отклонения μ основного актива больше безрисковой ставки r, то доход по ценной бумаге будет более высоким, но и при более высоком риске. В этом случае мы сможем рассчитать ожидаемую будущую стоимость опциона, но не сможем дисконтировать ее, чтобы получить текущую стоимость, поскольку неизвестно, какой ставкой дисконтирования следует воспользоваться.

Тот факт, что дифференциальное уравнение Блэка—Сколса для цены производного финансового инструмента не содержит μ, предполагает, что и решение (стоимость производного финансового инструмента) будет независимо от μ. Таким образом, мы можем обойти нашу трудность, переместившись в абстрактный безрисковый мир, где арбитражные операции обеспечат безрисковый доход для всех ценных бумаг. Это не повлияет на цену опциона на ценную бумагу.

Следовательно, мы можем определить стоимость опциона с помощью расчета его ожидаемой стоимости, допустив, что ценная бумага имеет норму отклонения r, а не μ, и дальнейшего дисконтирования по безрисковой ставке r.

Тогда текущая стоимость опциона на покупку ценной бумаги определяется следующим образом:

$$e^{-rT}(e^{rT}S(t)N(d_1)-XN(d_2)) = S(t)N(d_1)-Xe^{-rT}N(d_2).$$

Это и есть уравнение Блэка—Сколса.

11

МНОГОМЕРНЫЙ АНАЛИЗ: АНАЛИЗ ГЛАВНЫХ КОМПОНЕНТ И ФАКТОРНЫЙ АНАЛИЗ

ВВЕДЕНИЕ

Анализ главных компонент и факторный анализ — это методы анализа структуры данных в рамках многофакторности. Вместе с множественной регрессией (см. гл. 6) и многофакторной корреляцией в рядах динамики (см. гл. 8) эти методы наиболее часто используются в многофакторном анализе. Они отличаются от множественной регрессии тем, что целью регрессии является определение связи между экзогенной переменной и множеством эндогенных переменных. В случае анализа главных компонент и факторного анализа исследуется взаимоотношение только между эндогенными переменными. В отличие от корреляции в рядах динамики отношения между эндогенными переменными не обязательно должны быть устойчивыми.

В финансах и общественных науках, как и во многих других сферах, часто необходимо определить характеристики многофакторных структур. К двум таким характеристикам относятся:

- изменчивость многофакторной структуры;
- корреляция или коллинеарность переменных.

Обычно некоторые переменные будут иметь сильное влияние на общее изменение структуры, в то время как другие будут влиять слабо или же незначительно.

Одна из трудностей состоит в том, чтобы определить, какие переменные включать в модель и соответственно измерять. Например, если две переменные обладают идеальной корреляцией, то можно обойтись одной из них — вторая не несёт в себе никакой дополнительной информации. Это аналогично проблеме мультиколлинеарности во множественной регрессии.

В общем бывает неясно, какие переменные следует включить и какие исключить из рассмотрения, и возникает потребность в механизме отбора переменных или их комбинирования, таким образом, чтобы они включали всю доступную информацию наиболее эффективным образом.

Анализ главных компонент (principal components analysis) применяется при анализе изменчивости многофакторных структур. Факторный анализ (factor analysis) используется при анализе корреляции между переменными в многофакторной структуре. Оба метода основываются на анализе дисперсионно-ковариационной матрицы, поскольку она содержит всю информацию о том, в какой мере исследуемые переменные изменяются вместе, т.е. в какой мере дублируют или дополняют друг друга. В этой главе мы будем обозначать дисперсионно/ковариационную матрицу буквой C.

Хотя в методе главных компонент и факторном анализе используется дисперсионно-ковариационная матрица, они отличаются от анализа дисперсии — математического ожидания, рассмотренных в гл. 4 и 9, тем, что анализ дисперсии — математического ожидания измеряет общую изменчивость группы переменных без определения особого вклада подгруппы переменных в эту изменчивость. Метод главных компонент определяет и ранжирует подгруппы по их вкладу в совокупную изменчивость. Каждая из этих подгрупп — это "главная компонента" и определяется степенью ковариации между компонентами подгруппы. Вклад каждой из главных компонент в совокупную изменчивость ранжируется согласно совокупной дисперсии подгруппы.

При использовании анализа главных компонент, общая изменчивость данных находится как сумма собственных значений (которая будет равна сумме элементов на главной диагонали *C*, известной как ее след). Затем компоненты (линейные комбинации переменных) выбираются в порядке убывания собственных значений, пока главные компоненты не будут отвечать за достаточно большую долю изменчивости. Таким образом, размерность системы признаков снижается и определяются наиболее важные компоненты (направления).

В этой главе потребуется вспомнить математические действия над матрицами, описанные в приложении к гл. 6, применительно к определению дисперсии портфеля.

Из гл. 2 мы знаем, что дисперсия портфеля равна сумме взвешенных ковариаций каждой пары активов, где дисперсия считается ковариацией актива с самим собой. Представим портфель, состоящий из двух активов *A* и *B*. Дисперсия доходности актива *A* равна 0,00015, дисперсия доходности актива *B* равна 0,00025, и ковариация между *A* и *B* равна 0,00005. Дисперсионно-ковариационная матрица *C* будет иметь вид

$$C = \begin{bmatrix} 0{,}00015 & 0{,}00005 \\ 0{,}00005 & 0{,}00025 \end{bmatrix}. \qquad (11.0)$$

Если мы предполагаем, что доли ценных бумаг в портфеле равны, то дисперсия портфеля будет определяться умножением матрицы *C* на горизонтальный вектор весов $1 \times n$ и затем доумножением полученной матрицы на вертикальный вектор весов $n \times 1$. Таким образом будем иметь

$$\begin{bmatrix} 0{,}5 & 0{,}5 \end{bmatrix} \begin{bmatrix} 0{,}00015 & 0{,}00005 \\ 0{,}00005 & 0{,}00025 \end{bmatrix} \begin{bmatrix} 0{,}5 \\ 0{,}5 \end{bmatrix} = 0{,}000125. \qquad (11.1)$$

Следовательно, общая дисперсия портфеля равна 0,000125.

АНАЛИЗ ГЛАВНЫХ КОМПОНЕНТ (АГК)

Метод главных компонент применяется в двух целях. Первая — это уменьшение размерности данных с многих до нескольких переменных. Это достигается путем определения групп первоначальных переменных таким образом, чтобы члены группы обладали корреляцией между собой, но группа в целом была бы линейно независима от других переменных или групп переменных. Линейно независимые группы переменных называются *главными компонентами.*

Вторая цель, обусловленная первой, — это интерпретация данных. Это становится возможным благодаря тому, что метод главных компонент определяет линейные комбинации переменных и выстраивает их в порядке убывания их влияния на совокупную дисперсию первоначальных данных. Таким образом, первая главная компонента будет являться линейной комбинацией переменных, обладающей наиболее высокой дисперсией, вторая компонента — это линейная комбинация со второй по величине дисперсией и т.д. Это делает возможным объяснить бóльшую часть дисперсии наименьшим возможным количеством компонент.

Существует множество ситуаций в финансовой практике, когда желательно определить наиболее важные переменные или линейно независимые комбинации переменных, которые делают наибольший вклад в уровень риска. В нашем примере портфеля из двух активов будут только две компоненты. Однако в большом портфеле из *n* переменных будет *n* главных компонент. Некоторые из компонент будут обладать высокой корреляцией с другими, так что подгруппа в целом будет влиять на степень риска независимо от влияния других переменных или групп переменных. Метод главных компонент позволяет нам определить эти независимые линейные комбинации переменных и их влияние на совокупную дисперсию. Мы, таким образом, получаем более полное представление о том, что влияет на уровень риска и, следовательно, получаем возможность лучше управлять рисками.

Анализ главных компонент представляет собой скорее средство, чем цель. Например, определение главных компонент может служить для построения уравнения регрессии, так что зависимая переменная регрессируется не по первичным независимым переменным, а по главным компонентам. Далее в этой главе мы увидим, как определение главных компонент в изменениях процентных ставок позволяет нам лучше измерить процентный риск портфелей облигаций.

Сначала мы рассмотрим метод главных компонент на примере портфеля из двух инструментов, используя гипотетические данные из уравнения (11.1), приведенного выше. Затем, чтобы лучше понять этот материал, мы рассмотрим пример портфеля из четырех активов с использованием реальных данных, и в конце мы рассмотрим, как эта методика может использоваться для определения риска облигации.

Гипотетический пример с двумя активами

Использование АГК позволяет нам извлекать из дисперсионно-ковариационной матрицы число линейных комбинаций дисперсий и ковариаций активов, которое объясняет ковариационность активов, причем, каждая комбинация не зависит от других комбинаций. Это возможно благодаря тому, что симметричная структура дисперсионно-ковариационной матрицы позволяет это сделать при помощи процесса диагонализации. *Диагонализация* — это процесс, при помощи которого мы определяем линейные комбинации переменных, дисперсий и ковариаций, в данном случае независимых от других линейных комбинаций. Процесс включает три стадии:

1. нахождение собственных векторов и соответствующих собственных значений;
2. построение трех матриц Q, D и Q^{-1};
3. определение линейных комбинаций из собственных векторов, ранжирование комбинаций в порядке убывания собственных значений.

Первая стадия: нахождение собственных векторов и собственных значений

Первая стадия — это нахождение **собственных векторов** и соответствующих **собственных значений** дисперсионно-ковариационной матрицы C. Мы должны найти собственные векторы, потому что они дают нам линейно независимые комбинации переменных — главные компоненты, которые влияют на совокупную дисперсию. Мы должны найти собственные значения, потому что они показывают, за какую долю совокупного риска отвечает каждая главная компонента.

Математически собственные векторы — это векторы $\underline{X}_i$, каждый из которых обладает соответствующим скалярным значением λ_i — собственным значением, таким, что когда дисперионно-ковариационная матрица C умножается на вектор $\underline{X}_i$, то это равно умножению вектора на скалярную величину λ, т.е. $C\underline{X}_i = \lambda\underline{X}_i$. Симметричность матрицы C означает, что существует n таких векторов (при условии, что C — это не единичная матрица, т.е. обладает обратной матрицей) и что они ортогональны (перпендекулярны друг другу).

Пример приведен ниже. Матрица 2 × 2 в левой части — это дисперсионно-ковариационная матрица, использованная ранее. Если эта матрица может быть помножена на вектор, так, что произведение будет равно произведению собственного вектора на скалярную величину, например:

$$\begin{bmatrix} 0{,}00015 & 0{,}00005 \\ 0{,}00005 & 0{,}00025 \end{bmatrix} \begin{bmatrix} 1 \\ m \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ m \end{bmatrix}, \tag{11.2}$$

где

$$\begin{bmatrix} 1 \\ m \end{bmatrix} \tag{11.3}$$

является собственным вектором, то λ — это соответствующее собственное значение.

Раскрыв скобки, мы получаем:

$$0{,}00015 + 0{,}00005m = \lambda;$$

$$0{,}00005 + 0{,}00025m = \lambda m. \tag{11.4}$$

Отсюда, умножая оба уравнения на m и сокращая λ, мы получаем:

$$0{,}00015m + 0{,}00005m^2 = 0{,}00005 + 0{,}00025m$$

$$0{,}00005m^2 - 0{,}00010m - 0{,}00005 = 0$$

$$5m^2 - 10m - 5 = 0$$

$$m^2 - 2m - 1 = 0$$

$$m = \frac{+2 \pm \sqrt{4+4}}{2} = \frac{2 \pm \sqrt{8}}{2} = 1 \pm \sqrt{2}.$$

Таким образом, собственные векторы равны:

$$\begin{bmatrix} 1 \\ 1+\sqrt{2} \end{bmatrix} \text{ и } \begin{bmatrix} 1 \\ 1-\sqrt{2} \end{bmatrix}. \tag{11.5}$$

Теперь мы должны нормализировать векторы так, чтобы длина их стала равна единице. Это значит, что каждая компонента собственного вектора умножается на квадратный корень суммы квадратов каждой компоненты. Полученные нормализированные векторы являются собственными векторами

$$\begin{vmatrix} \dfrac{1}{\sqrt{1^2+\left(1+\sqrt{2}\right)^2}} \\ \dfrac{1+\sqrt{2}}{\sqrt{1^2+\left(1+\sqrt{2}\right)^2}} \end{vmatrix} = \begin{bmatrix} 0{,}383 \\ 0{,}924 \end{bmatrix} \quad \begin{vmatrix} \dfrac{1}{\sqrt{1^2+\left(1-\sqrt{2}\right)^2}} \\ \dfrac{1-\sqrt{2}}{\sqrt{1^2+\left(1-\sqrt{2}\right)^2}} \end{vmatrix} = \begin{bmatrix} 0{,}924 \\ -0{,}383 \end{bmatrix}.$$

Запомним, что существует столько же собственных векторов, сколько переменных в дисперсионно-ковариационной матрице. Таким образом, в матрице 2 × 2 будут два собственных вектора, а в матрице *n* × *n* — *n* собственных векторов.

Мы можем доказать, что полученные векторы являются собственными векторами, потому что собственные векторы — это векторы, которые при умножении на матрицу *C* равны произведению вектора на скалярную величину. Например

$$\begin{bmatrix} 0{,}00015 & 0{,}00005 \\ 0{,}00005 & 0{,}00025 \end{bmatrix} \begin{bmatrix} 0{,}383 \\ 0{,}924 \end{bmatrix} = \begin{bmatrix} 0{,}000104 \\ 0{,}000250 \end{bmatrix} = 0{,}000271 \begin{bmatrix} 0{,}383 \\ 0{,}924 \end{bmatrix}.$$

Здесь мы умножим матрицу *C* на вектор, чтобы получить новый вектор. Такой же вектор мы получим при умножении 0,000271 на первоначальный вектор. Таким образом, мы видим, что вектор в левой части — это собственный вектор, и скаляр 0,000271 — это собственное значение.

Точно такая же процедура применяется и к вектору в правой части, и в результате видим, что этот вектор также является собственным вектором

$$\begin{bmatrix} 0{,}00015 & 0{,}00005 \\ 0{,}00005 & 0{,}00025 \end{bmatrix} \begin{bmatrix} 0{,}924 \\ -0{,}383 \end{bmatrix} = \begin{bmatrix} 0{,}000119 \\ 0{,}000050 \end{bmatrix} = 0{,}000129 \begin{bmatrix} 0{,}924 \\ -0{,}383 \end{bmatrix},$$

здесь собственное значение равно 0,000129.

Вторая стадия: построение матриц Q, D и Q^{-1}

Следующий шаг — это построение трех матриц Q, D и Q^{-1}. Матрицу Q строим из собственных векторов, записывая их как колонки матрицы в порядке убывания соответствующих им собственных значений. Таким образом, матрица Q для нашего примера будет

$$Q = \begin{bmatrix} 0{,}383 & 0{,}924 \\ 0{,}924 & -0{,}383 \end{bmatrix}. \qquad (11.6)$$

Матрица D — это диагональная матрица (все значения, кроме значений на главной диагонали, равны нулю), значения на главной диагонали равны собственным значениям, в порядке убывания их значений. Таким образом, матрица D в нашем примере имеет следующий вид:

$$D = \begin{bmatrix} 0{,}000271 & 0 \\ 0 & 0{,}000129 \end{bmatrix}. \qquad (11.7)$$

Так, мы получим $CQ = QD$. Опять-таки при условии, что C и Q неединичные матрицы, мы сможем записать $C = QDQ^{-1}$. Однако, если длина собственных векторов равна единице, т.е. сумма квадратов компонент собственных векторов равна единице, то матрицей, обратной Q, будет сама матрица Q, тогда можно записать, что $C = QDQ$.

$$\begin{array}{ccccccc} C & & Q & & D & & Q^{-1} \\ \begin{bmatrix} 0{,}00015 & 0{,}00005 \\ 0{,}00005 & 0{,}00029 \end{bmatrix} & = & \begin{bmatrix} 0{,}383 & 0{,}924 \\ 0{,}924 & -0{,}383 \end{bmatrix} & & \begin{bmatrix} 0{,}000271 & 0{,}0 \\ 0{,}0 & 0{,}000129 \end{bmatrix} & & \begin{bmatrix} 0{,}383 & 0{,}924 \\ 0{,}924 & -0{,}383 \end{bmatrix}. \end{array} \qquad (11.8)$$

Третья стадия: определение линейных комбинаций переменных и ранжирование комбинаций согласно их собственным значениям

Теперь рассмотрим дисперсионно-ковариационную матрицу C. Мы предположим, что она относится к двум активам X и Y. Дисперсия этого портфеля из двух активов может быть записана как

$$\begin{bmatrix} X & Y \end{bmatrix} \begin{bmatrix} 0{,}00015 & 0{,}00005 \\ 0{,}00005 & 0{,}00025 \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}. \qquad (11.9)$$

Но поскольку $C = QDQ^{-1}$, она может быть записана как

$$\begin{bmatrix} X & Y \end{bmatrix} \begin{bmatrix} 0{,}383 & 0{,}924 \\ 0{,}924 & -0{,}383 \end{bmatrix} \begin{bmatrix} 0{,}000271 & 0{,}0 \\ 0{,}0 & 0{,}000129 \end{bmatrix} \begin{bmatrix} 0{,}383 & 0{,}924 \\ 0{,}924 & -0{,}383 \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}, \qquad (11.10)$$

что приводит к следующему

$$[0{,}383X + 0{,}924Y \quad 0{,}924Y - 0{,}383X]\begin{bmatrix} 0{,}000271 & 0 \\ 0 & 0{,}000129 \end{bmatrix}\begin{bmatrix} 0{,}383X + 0{,}924Y \\ 0{,}924Y - 0{,}383X \end{bmatrix}.$$

На основе вышеприведенного примера диагонализации мы можем определить линейные комбинации переменных X и Y, которые независимо друг от друга влияют на дисперсию всего портфеля. Эти комбинации определяются собственными векторами. Таким образом, в нашем примере $0{,}383X + 0{,}924Y$ представляет собой одну линейно независимую комбинацию, и $0{,}924Y - 0{,}383X$ — другую.

Мы должны иметь возможность ранжировать собственные векторы согласно их важности для общей дисперсии всего портфеля. Ранжирование производится согласно собственным значениям. Собственный вектор с наибольшим собственным значением дает наибольший вклад в совокупную дисперсию, собственный вектор, оказывающий второе по важности влияние, обладает вторым по величине собственным значением, и т.д.

Доля совокупной дисперсии, которая приходится на счет каждой комбинации переменных, равна доле соответствующего собственного значения в сумме всех собственных значений. Сумма собственных значений задается суммой элементов главной диагонали в матрице D. В нашем примере сумма собственных значений равна 0,000400. Таким образом, собственный вектор с собственным значением, равным 0,000271, отвечает за 0,000271/0,000400, или, 67,75% совокупной дисперсии.

Пример анализа доходности FTSE 100, государственных облигаций, S&P 500 и обменного курса валют

Рассмотрим портфель, содержащий в равных долях четыре актива — индексы FTSE 100 и S&P 500, британские государственные облигации и обменный курс фунт/доллар. Используя ежемесячные значения доходности за период с сентября 1989 по декабрь 1993, получаем следующую дисперсионно-ковариационную матрицу:

	Облигации	*FTSE*	*S&P*	*Доллар/Фунт*
Облигации	5,5884	1,6749	0,2233	—1,0528
FTSE	1,6749	22,3660	10,0886	—5,5765
S&P	0,2233	10,0886	10,7730	—0,5839
Фунт/Доллар	—1,0528	—5,5765	—0,5839	16,2161

Четыре собственных значения и собственный вектор будут составлять:

собственные значения: $\lambda_1 = 30{,}25{,}72$; $\lambda_2 = 14{,}8258$; $\lambda_3 = 5{,}61280$; $\lambda_4 = 4{,}2477$; собственные векторы:

$$\begin{bmatrix} 0{,}074904 \\ 0{,}824166 \\ 0{,}438093 \\ -0{,}351135 \end{bmatrix} \begin{bmatrix} -0{,}058582 \\ 0{,}199571 \\ 0{,}362687 \\ 0{,}908404 \end{bmatrix} \begin{bmatrix} 0{,}921233 \\ -0{,}153459 \\ -0{,}322417 \\ 00{,}154422, \end{bmatrix} \begin{bmatrix} 0{,}377235 \\ -0{,}507399 \\ -0{,}756690 \\ -0{,}166314 \end{bmatrix}.$$

Первому собственному вектору соответствует собственное значение, равное 30,2572, и он отвечает за 55,07% совокупной дисперсии. Второй собственный вектор с собственным значением, равным 14,8258, отвечает за 26,98% совокупной дисперсии. Третий и четвертый собственные векторы имеют собственные значения 5,6128 и 4,2477 соответственно и отвечают за 10,22% и 7,73% совокупной дисперсии.

Пример стандартизованных переменных

Одна из проблем, связанных с методом главных компонент, состоит в том, что различные данные обладают разными порядками величины. Читатель, наверное, помнит, что это проблема ковариаций вообще. Заключается проблема в том, что величина ковариации является функцией величины данных, так же как и разность X и $\bar{X}$. Читатель, возможно, помнит, что предлагаемое решение состояло в определении коэффициента корреляции, который, по сути, представляет собой стандартизованную ковариацию.

Подобная проблема может появиться при использовании метода главных компонент, базирующегося на ковариационной матрице, составленной из нестандартизованных данных. По этой причине, в случае если данные по разным переменным обладают величинами разного порядка или используются различные единицы измерения для разных переменных, обычно данные стандартизируют перед применением метода главных компонент.

В следующем примере исходные данные о доходности были стандартизованы, поскольку дисперсии и ковариации разных переменных существенно различаются между собой. Стандартизация производилась следующим образом:

$$z_i = \frac{(x_i - \bar{x})}{\sigma_x}.$$

	Облигации	*FTSE*	*S&P*	*Доллар/Фунт*	
Облигации	1,0	0,1498	0,0288	–0,1106	
FTSE	0,1498	1,0	0,6499	–0,2928	(11.10.5)
S&P	0,0288	0,6499	1,0	–0,0442	
Фунт/Доллар	–0,1106	–0,2928	–0,0442	1,0	

Четыре собственных вектора и соответствующие собственные значения будут следующими:

собственные значения: $\lambda_1 = 1{,}76345$; $\lambda_2 = 1{,}05032$; $\lambda_3 = 0{,}88860$; $\lambda_4 = 0{,}29764$;
собственные векторы:

$$\begin{bmatrix} 0{,}205868 \\ 0{,}688288 \\ 0{,}612746 \\ -0{,}329272 \end{bmatrix}\begin{bmatrix} -0{,}698552 \\ 0{,}102372 \\ 0{,}424376 \\ 0{,}566966 \end{bmatrix}\begin{bmatrix} 0{,}679614 \\ -0{,}010147 \\ 0{,}166887 \\ 0{,}714262 \end{bmatrix}\begin{bmatrix} -0{,}088133 \\ 0{,}718106 \\ -0{,}645443 \\ 0{,}244866 \end{bmatrix}.$$

Поскольку переменные были стандартизованы перед анализом, сумма собственных значений будет равна четырем, поскольку мы имеем четыре переменные. Принимая это во внимание, соответствующий вклад каждого из собственных векторов в совокупную дисперсию составит:

$$\lambda_1 = 44{,}09;\ \lambda_2 = 26{,}26;\ \lambda_3 = 22{,}22;\ \lambda_4 = 7{,}44.$$

Следует отметить, что процесс стандатризации не является непоследовательным, так как величина собственных векторов и собственных значений, найденных на основе стандартизованных данных, будет отличаться от таковых, найденных на основе нестандартизованных данных. В вышеприведенном случае, например, при использовании исходных данных, первая главная компонента отвечала за 55% совокупной дисперсии и третья компонента — только за 10%.

Однако при использовании стандартизованных данных первая компонента отвечает всего лишь за 44% совокупной дисперсии, а третья главная компонента — за 22%, при том что вторая и четвертая компоненты остались практически без изменений.

Интерпретация главных компонент

Интерпретация главных компонент в определенной мере является субъективной. Например, рассмотрим собственные векторы, относящиеся к вышеприведенным стандартизованным переменным. Так, первый собственный вектор, отражающий первую главную компоненту, может использоваться для измерения подверженности портфеля стохастической изменчивости, соответствующей этой главной компоненте. Это дает управляющему портфелем возможность изменять удельные веса составляющих портфель активов так, чтобы снизить связанный с ними элемент риска.

Мы продемонстрируем это в другом контексте — на примере управления рисками на рынках облигаций.

ПРИМЕНЕНИЕ НА РЫНКАХ ОБЛИГАЦИЙ

Анализ главных компонент применяется при разработке моделей риска на рынках облигаций. Из гл. 3 мы помним, что дюрация и выпуклость широко используются участниками рынка облигаций для обобщения чувствительности отдельных облигаций и портфелей облигаций к изменениям процентных ставок. Однако как дюрация, так и выпуклость облигации основываются на допущении, что кривая доходности является твердой и может переноситься только параллельно. Случайный эмпиризм показывает, что эти допущения в действительности не имеют места. На практике отдельные спот-курсы не обладают идеальной корреляцией, таким образом, временная структура и отсюда кривая доходности не переносятся параллельно.

Анализ главных компонент использовался Каном (Kahn) (1989), Каном и Гульраджани (Kahn and Gulrajani) (1993) и

Карки и Рейесом (Karki and Reyes) (1994) при определении основных тем в общих изменениях большого числа спот-курсов, что составляет временную структуру процентных ставок. Затем информация о совместных изменениях может комбинироваться с информацией о подверженности облигаций изменениям отдельных спот курсов в пределах временной структуры для определения меры риска, связанного с каждой формой совместных изменений.

Изменения цены облигаций происходят по трем причинам:

1. приближается срок погашения облигации;
2. изменяется временная структура облигации;
3. изменяется оценка рынком характеристик облигации, таких как возможность неплатежа, включение в расчет стоимости опциона и т.д.

Приближение срока погашения облигации известно, и таким образом, его влияние на изменение стоимости облигации известно и не является неопределенным. Однако изменение стоимости облигации по причине изменений временной структуры и оценки рынком характеристик облигации являются неопределенными и, следовательно, являются причинами риска.

Мы знаем, что временная структура является рядом спот-курсов по отношению к спектру погашений на рынке облигаций, поэтому облигация будет подвержена изменениям во временной структуре в зависимости от того, как ожидаемые денежные потоки распределены по временной структуре, т.е. как каждый денежный поток связан с каждым спот-курсом. Например, рассмотрим две бескупонные облигации. Срок погашения одной 10 лет, а другой — 3 года. Если временная структура параллельно переносится вверх или вниз, то обе облигации изменятся в одинаковой степени, и величина изменений показывается продолжительностью и выпуклостью. Однако если долгосрочные процентные ставки возрастают, а краткосрочные — падают, то цена десятилетней облигации упадет, а трехлетней — возрастет. Теперь рассмотрим облигацию с ненулевым купоном. Она по сути является портфелем облигаций с нулевым купоном. Каждый из купонов и погасительный платеж дисконтируются по соответствующей спот-ставке. Стоимость облигации будет равна сумме приведенных стоимостей, дисперсия облигации — это взвешенная сумма ковариаций каждого из дисконтированных денежных потоков.

Таким образом, первая задача при применении метода главных компонент для анализа рискованности облигации — это определение погасительного профиля временной структуры процентных ставок. Мы будем ссылаться на каждую из точек этого погасительного профиля как к вершине. Обычно бывают месячные интервалы на протяжении первых трех месяцев, затем квартальные на протяжении следующих девяти месяцев и далее — полугодовые, на интервале, достаточном, чтобы охватить все облигации выпуска.

Анализ главных компонент относится к линейным комбинациям переменных, и из гл. 1 мы знаем, что цены облигаций представляют собой линейную комбинацию текущих стоимостей. Каждая текущая стоимость — это результат применения соответствующей спот-ставки для приведения денежного потока. Мы можем использовать метод главных компонент для определения линейных комбинаций изменений текущих стоимостей в результате изменения процентных спот-ставок. На основе линейных комбинаций изменений текущих стоимостей очень просто соотнести их с комбинациями изменений спот-ставок, которые явились причиной изменения текущей стоимости.

Поскольку доходность облигации — это результат как нестохастического процесса (приближения срока погашения облигации), так и изменения процентных ставок, необходимо удалить нестохастический элемент. Чтобы сделать это, мы работаем не с полной доходностью, а с **избыточной доходностью** и определяем риск как дисперсию избыточной доходности. Избыточная (дополнительная) доходность определяется как разность между доходностью облигации с нулевым купоном и безрисковой доходностью. **Безрисковая доходность** — это доходность данной облигации с нулевым купоном только за счет приближения срока погашения.

Отсюда следует, что мы должны начать с определения доходности облигации и безрисковой доходности. Чтобы понять концепцию доходности облигации, рассмотрим одномесячный период владения облигацией. Мы обозначим этот период как Δt. Текущий момент времени — t. Остаточный срок погашения (срок, остающийся до погашения к концу периода владения облигацией) — wi. Таким образом, срок жизни облигации в момент времени t равен $\Delta t + w_t$. Теперь рассмотрим инвестирование \$1 в момент времени t в облигацию со сроком погашения $t + \Delta t + w_t$ и продажу ее в момент времени $t + \Delta t$. Изменение стоимости инвестиции за время обладания облигацией и будет доходностью за этот период. Доходность будет определяться как приближением срока погашения, так и изменением временной структуры и восприятия рынком характеристик облигации.

Теперь рассмотрим инвестирование \$1 в дисконтную государственную облигацию со сроком погашения $t + \Delta t$. Доходность этой инвестиции будет абсолютно определенной, и, следовательно, безрисковой. Вычитая безрисковую доходность из доходности облигации со сроком погашения $t + \Delta t + w_t$, получаем избыточную доходность денежного потока с погашением в вершине $t + \Delta t + w_t$. При вычислении безрисковой доходности из доходности облигации со сроком погашения $t + \Delta t + w_t$ любое изменение стоимости облигации из-за приближения срока погашения будет удалено. Таким образом, полученная добавочная доходность является чисто стохастическим элементом.

Если мы соберем временные ряды доходностей за период владения для каждой вершины временной структуры, то мы сможем построить дисперсионно-ковариационную матрицу избыточных доходностей.

Из полученной дисперсионно-ковариационной матрицы мы находим собственные векторы и связанные с ними собственные значения. На рынке государственных облигаций три главные компоненты отвечают за 99% риска временной структуры. В вышеупомянутых исследованиях Кана, Кана и Гульраджани и Карки и Рейеса первая компонента может быть интерпретирована как изменение общего уровня временной структуры аналогично параллельному смещению, вторая компонента — как изменение угла наклона кривой временной структуры, третья компонента — как изменение изгиба кривой временной структуры.

Чтобы объяснить, как эта модель может использоваться при определении риска купонной облигации (с ненулевым купоном), рассмотрим сначала, как мы можем использовать математические действия над матрицами для нахождения совокупной дисперсии облигации с пятью ежегодными платежами — CF_1, CF_2, CF_3, CF_4 и CF_5 соответственно. Совокупная дисперсия V будет

$$[CF_1 \quad CF_2 \quad CF_3 \quad CF_4 \quad CF_5]\begin{bmatrix} \mathrm{var}_1 & \mathrm{cov}_{12} & \mathrm{cov}_{13} & \mathrm{cov}_{14} & \mathrm{cov}_{15} \\ \mathrm{cov}_{21} & \mathrm{var}_2 & \mathrm{cov}_{23} & \mathrm{cov}_{24} & \mathrm{cov}_{25} \\ \mathrm{cov}_{31} & \mathrm{cov}_{32} & \mathrm{var}_3 & \mathrm{cov}_{34} & \mathrm{cov}_{35} \\ \mathrm{cov}_{41} & \mathrm{cov}_{42} & \mathrm{cov}_{43} & \mathrm{var}_4 & \mathrm{cov}_{45} \\ \mathrm{cov}_{51} & \mathrm{cov}_{52} & \mathrm{cov}_{53} & \mathrm{cov}_{54} & \mathrm{var}_5 \end{bmatrix}\begin{bmatrix} CF_1 \\ CF_2 \\ CF_3 \\ CF_4 \\ CF_5 \end{bmatrix} = V. \tag{11.11}$$

Теперь, используя метод главных компонент, определим темы совместных изменений, которые имеют влияние на дисперсию, преобразуем дисперсионно-ковариационную матрицу в группу из трех матриц Q, D и Q^{-1} и определим совокупную дисперсию, умножая вектор весов $1 \times n$ на матрицы Q, D и Q^{-1} $n \times n$ и затем полученное произведение — на вектор весов $n \times 1$:

$$[CF_1 \quad CF_2 \quad CF_3 \quad CF_4 \quad CF_5]QDQ^{-1}\begin{bmatrix} CF_1 \\ CF_2 \\ CF_3 \\ CF_4 \\ CF_5 \end{bmatrix} = V. \tag{11.12}$$

Как мы уже говорили, Q — это матрица собственных векторов, ранжированных слева направо в порядке убывания соответствующих собственных значений. Таким образом, в уравнении (11.13) с $EV_{1(1)}$ по $EV_{1(n)}$ — это элементы первого собственного вектора с 1 *пол.*

Таким же образом, умножая вектор размера денежных потоков на матрицу Q, мы получим вектор разложения отдельных денежных потоков по облигации по влиянию главных компонент:

$$[CF_1, CF_2, CF_3, CF_4, CF_5, 0\ 0 \ldots 0_n]\begin{bmatrix} EV_{1(1)} & \cdots & EV_{n(1)} \\ EV_{1(2)} & \cdots & EV_{n(2)} \\ . & \cdots & . \\ . & \cdots & . \\ EV_{1(n)} & \cdots & EV_{n(n)} \end{bmatrix}. \tag{11.13}$$

Заметьте, что вектор денежных потоков включает помимо действительных денежных потоков по рассматриваемой облигации ряд нулей. Эти нули относятся к тем вершинам временной структуры, влиянию которых у данной облигации нет подверженных денежных потоков. Это необходимо для того, чтобы было возможно производить математические действия над матрицами.

Таким образом, произведение вектора денежных потоков и первого собственного вектора даст нам единичную меру (статистическое резюме) подверженности облигации влиянию первой компоненты. Произведение вектора денежных потоков и второго собственного вектора даст нам подверженность облигации влиянию второй компоненты и т.д. Полученный вектор подверженностей будет:

$$[\exp_1, \exp_2, \exp_3, \exp_4, \exp_5, 0\ 0\ 0\ \ldots\ 0]. \qquad (11.14)$$

D — это диагональная матрица, где элементы диагонали являются собственными значениями (λ), записанными в порядке убывания. Умножая транспонированный вектор подверженностей на матрицу D и затем умножая это произведение на вектор подверженностей, мы получаем совокупную дисперсию, выраженную через главные компоненты следующим образом:

$$[\exp_1,\exp_2,\exp_3,\exp_4,\exp_5,\ 0\ 0\ 0\ \ldots\ 0_n]\begin{bmatrix} \lambda & 0 & \ldots & 0_n^1 \\ 0_1^2 & \lambda_2 & \ldots & 0_n^2 \\ . & . & .. & . \\ . & . & \ldots & . \\ 0_1^n & . & \ldots & \lambda_n \end{bmatrix}\begin{bmatrix} \exp_1 \\ \exp_2 \\ \exp_3 \\ \exp_4 \\ \exp_5 \\ 0 \\ . \\ 0_n \end{bmatrix} = v. \qquad (11.15)$$

Дисперсия может быть выражена в виде суммы произведений собственных значений и соответствующих подверженностей:

$$V = (\exp_1{}^2 \cdot \lambda_1) + (\exp_2{}^2 \cdot \lambda_2) + \ldots + (\exp_n{}^2 \cdot \lambda_n). \qquad (11.16)$$

Таким образом, рискованность облигации является функцией от собственных значений каждой главной компоненты и подверженности денежных потоков влиянию каждой из главных компонент.

ФАКТОРНЫЙ АНАЛИЗ

Факторный анализ (ФА) представляет собой иной способ толкования структуры дисперсионно-ковариационной матрицы. Чтобы уяснить использование ФА, мы должны начать с более близкого рассмотрения понятия дисперсии. Вернемся к модели определения риска портфеля Марковитца. Эта модель разделяет совокупную дисперсию на систематическую и несистематическую. *Систематическая дисперсия* (риск) — это такой риск, от которого нельзя избавиться при помощи диверсификации, в то время как от несистематического риска можно избавиться. По сути систематический риск — это общий риск для всех активов в портфеле, в то время как несистематический риск уникален для каждого отдельного инструмента.

Различие между систематическим и несистематическим рисками и лежит в основе ФА. При помощи метода главных компонент мы объясняли совокупную дисперсию. При помощи же факторного анализа мы собираемся определить размер систематического риска (общности — в терминах ФА) внутри ковариационной структуры.

В анализе главных компонент мы извлекали линейные комбинации рассматриваемых переменных так, что на каждой стадии получаемая компонента объясняет наибольшую возможную долю остающейся изменчивости. В ФА мы, по сути, разбиваем совокупную дисперсию данных на две части — на ту, которая разделяется всеми переменными (общность), и ту, которая специфична для каждой отдельной переменной. Факторный анализ использует оценки общности для построения объясняющих факторов.

Для демонстрации использования ФА предположим, что мы применяем метод главных компонент, с помощью которого выявили m значимых компонент.

Затем мы пытаемся выразить каждую из первоначальных переменных как линейную комбинацию того числа факторов. Выбираются те комбинации, которые лучше всего работают — производят модель, наиболее полно объясняющую систематическую дисперсию.

Рассмотрим стандартизованную дисперсионно-ковариационную матрицу четырех активов, которые мы рассматривали в качестве примера применения метода главных компонент:

	Облигации	*FTSE*	*S&P*	*Фунт/Доллар*
Облигации	1,0	0,1498	0,0288	—0,1106
FTSE	0,1498	1,0	0,6499	—0,2928
S&P	0,0288	0,6499	1,0	—0,0442
Фунт/Доллар	—0,1106	—0,2928	—0,0442	1,0

Мы помним, что матрица основывается на 51 значении месячных доходностей с сентября 1989 по декабрь 1993 гг. Также мы знаем, что вклад первых трех компонент в совокупную дисперсию составлял 93%. Таким образом, мы предположили, что существуют три значимых фактора для целей факторного анализа.

Предполагается, что каждая переменная является комбинацией этих трех факторов:

$$\begin{aligned}
\text{GILT} &= t_1 = \lambda_{11} f_1 + \lambda_{12} f_2 + \lambda_{13} f_3 + e_1 \\
\text{FTSE} &= t_2 = \lambda_{21} f_1 + \lambda_{22} f_2 + \lambda_{23} f_3 + e_2 \\
S\&P &= t_3 = \lambda_{31} f_1 + \lambda_{32} f_2 + \lambda_{33} f_3 + e_3 \\
\text{Фунт/Доллар} &= t_4 = \lambda_{41} f_1 + \lambda_{42} f_2 + \lambda_{43} f_3 + e_4.
\end{aligned} \tag{11.17}$$

В результате получаем вектор переменных 4 × 1, матрицу значений λ 4 × 3, вектор факторов 4 × 1 и вектор значений ошибки 4 × 1.

Теперь рассмотрим n наблюдений каждой из четырех переменных. Так, мы получаем матрицу 4 × n значений t

$$T = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ t_{31} & t_{32} & \cdots & t_{3n} \\ t_{41} & t_{42} & \cdots & t_{4n} \end{bmatrix}. \tag{11.18}$$

Похожим образом мы получаем матрицу 3 × n значений F

$$F = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ f_{31} & f_{32} & \cdots & f_{3n} \\ f_{41} & f_{42} & \cdots & f_{4n} \end{bmatrix} \tag{11.19}$$

и матрицу 4 × n значений e

$$E = \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1n} \\ e_{21} & e_{22} & \cdots & e_{2n} \\ e_{31} & e_{32} & \cdots & e_{3n} \\ e_{41} & e_{42} & \cdots & e_{4n} \end{bmatrix}. \tag{11.20}$$

Наша модель становится:

$$T = \lambda F + E \tag{11.21}$$

где T — 4 × n, λ — 3 × n, F —3 × n и E — 4 × n.

Теперь сделаем дополнительные предположения:

1. значения f_i независимы;
2. значения e_i независимы;
3. между f_i и e_i нет корреляции.

Теперь $C = TT^T$, что равно

$$(\lambda F + E)(\lambda F + E)^T. \tag{11.22}$$

Это равно

$$\lambda F(\lambda F)^T + \lambda FE^T + E(\lambda F)^T + EE^T =$$

$$= \lambda FF^T\lambda^T + \lambda FE^T + EF^T\lambda^T + EE^T. \tag{11.23}$$

Поскольку предполагается, что между значениями f и e нет корреляции, матрица FE^T является нулевой, и EF^T — также нулевая. Таким образом, уравнение (11.23) сокращается до

$$\lambda FF^T\lambda^T + [0] + [0] + EE^T. \tag{11.24}$$

Так как мы используем стандартизованные переменные, и поскольку значения f независимы, FF^T становится единичной матрицей, так как корреляция F с самим собой равна единице, а ковариации предполагаются равными нулю.

Матрица EE^T является диагональной матрицей с дисперсией ошибки на диагонали и с нулями в остальных ячейках. Таким образом, ковариация группы активов равна произведению факторных матриц λ и λ^T, прибавленных к матрице EE^T. Если у нас e обозначает несистематический риск (US), то мы увидим, что

$$\mathrm{cov} = \lambda\lambda^T + \begin{bmatrix} US & 0 & \dots & 0 \\ 0 & US & \dots & 0 \\ . & . & \dots & . \\ 0 & 0 & \dots & US \end{bmatrix}, \tag{11.25}$$

где λ — это матрица 4×3, и λ^T — матрица 3×4.

В результате получаем матрицу 4×4, в которой общности располагаются на главной диагонали. US в матрице в правой части уравнения представляет несистематический риск в ковариационной матрице. Задача факторного анализа — это решение уравнения (11.25) таким образом, чтобы $\lambda\lambda^T$ отвечало за максимально возможную долю совокупной дисперсии, в то время как диагональная матрица — за минимально возможную.

Для достижения этой цели используется итерационный процесс. Мы продемонстрируем этот процесс, показав результаты работы компьютера.

Матрица λ будет

$$\begin{array}{lc} & \textit{Фактор } 1 \quad \textit{Фактор } 2 \quad \textit{Фактор } 3 \\ \begin{array}{c} \textit{Gilt} \\ \textit{FTSE} \\ \textit{S\&P} \\ \textit{Фунт / Доллар} \end{array} & \lambda = \begin{array}{rrr} 0{,}27338 & -0{,}71591 & 0{,}64064 \\ 0{,}91401 & 0{,}10492 & -0{,}00956 \\ 0{,}81370 & 0{,}43492 & 0{,}15732 \\ -0{,}43726 & 0{,}58105 & 0{,}67333 \end{array}. \end{array}$$

Матрица λ^{T}

$$\begin{bmatrix} 0{,}27338 & 0{,}91401 & 0{,}81370 & -0{,}43726 \\ -0{,}71591 & 0{,}10492 & 0{,}43492 & 0{,}58105 \\ 0{,}64064 & -0{,}00956 & 0{,}15732 & 0{,}67330 \end{bmatrix}.$$

Матрица EE^T, известная как матрица общностей, имеет вид:

$$EE^T + \begin{bmatrix} 0.997688 & 0 & 0 & 0 \\ 0 & 0.846516 & 0 & 0 \\ 0 & 0 & 0.876006 & 0 \\ 0 & 0 & 0 & 0.982154 \end{bmatrix}.$$

При условии такого первоначального уточнения каждый фактор будет отвечать за следующую дисперсию:

фактор 1 = 1,1763450

фактор 2 = 1,050317

фактор 3 = 0,888597,

как это задается в $\lambda^{\mathrm{T}}\lambda$. Сумма дисперсией, соответствующих этим факторам, равна 3,702354, тогда как совокупная дисперсия равна четырем, поскольку мы используем стандартизованные переменные, т.е. три фактора объясняют 92,56% совокупной дисперсии рассматриваемого портфеля.

В то время как факторный анализ пытается определить факторы, которые объясняют наблюдения, сами факторы не могут быть непосредственно наблюдаемыми (или не находились в фокусе анализа). Таким образом, следует более сосредоточиться на понимании, чем на применении. Тем не менее, идентификация нужных факторов полезна при определении направления последующей работы.

Чтобы еще более усложнить ФА, можно показать, что для данной совокупности факторов любое ортогональное преобразование (т.е. вращение) этих векторов будет иметь тот же самый эффект. Следовательно, мы вольны выбирать, какие из доходностей представляют собой наиболее значимый результат. Процедура, известная под названием *варимаксной* (varimax procedure), применяется для того, чтобы выбрать факторы так, что нагрузка одних факторов велика, а других — мала. Это связывает переменные с меньшим числом более различающихся факторов.

В качестве примера рассмотрим применение варимаксной процедуры в нашем случае.

Во-первых, мы строим "матрицу ортогонального преобразования":

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} \quad 1 \quad & \quad 2 \quad & \quad 3 \quad \end{array} \\ \begin{bmatrix} 0{,}91024 & -0{,}35424 & 0{,}21441 \\ 0{,}39490 & 0{,}58692 & -0{,}70681 \\ 0{,}12454 & 0{,}72804 & 0{,}67413 \end{bmatrix} \end{array}.$$

Мы умножаем эту матрицу 3×3 на первоначальную матрицу λ 4×3. В результате процесса вращения получаем новую матрицу λ

$$\begin{array}{l} \\ \textit{Gilt} \\ \textit{FTSE} \\ \textit{S\&P} \\ \textit{Фунт/Доллар} \end{array} \quad \lambda = \begin{array}{c} \begin{array}{ccc} \textit{Фактор } 1 & \textit{Фактор } 2 & \textit{Фактор } 3 \end{array} \\ \begin{bmatrix} 0{,}04591 & -0{,}05062 & -0{,}99650 \\ 0{,}87221 & -0{,}26917 & -0{,}11537 \\ 0{,}93200 & 0{,}08155 & 0{,}02689 \\ -0{,}08470 & 0{,}98612 & 0{,}05055 \end{bmatrix} \end{array}.$$

Эти факторы имеют следующую интерпретацию. Фактор 1 практически не включает государственных облигаций, имеет малое отрицательное число элементов валютных курсов, но включает большую комбинацию обоих фондовых индексов. Наибольший элемент фактора 2 — это обменные курсы, при том, что этот фактор практически не имеет государственных облигаций и индекса S&P 500 и только очень малую долю FTSE 100. Фактор 3 имеет большой отрицательный элемент государственных облигаций, а все остальные элементы очень малы.

Новая матрица λ^T будет:

$$\lambda^T = \begin{bmatrix} 0.04591 & 0.87221 & 0.93200 & -0.08470 \\ -0.05062 & -0.2917 & 0.08155 & 0.98612 \\ 0.99650 & 0.11537 & -0.02689 & -0.05055 \end{bmatrix}.$$

Матрица EE^T в результате этого процесса не изменяется.

Дисперсия, объясняемая каждым фактором, теперь составляет:

фактор 1 = 1,638670

фактор 2 = 1,054089

фактор 3 = 1,009606

(опять задается $\lambda^T\lambda$). Сумма этих чисел равна 3,702365, таким образом, эти три фактора по-прежнему объясняют 92,56% совокупной дисперсии.

ТЕОРИЯ АРБИТРАЖНОГО ЦЕНООБРАЗОВАНИЯ

Обычно вышерассмотренный процесс применяется при проверке теории арбитражного ценообразования (ТАЦ) (arbitrage pricing theory).

Теория предполагает, что на рынке, где возможен арбитраж, все активы со схожими характеристиками будут иметь схожие цены, поскольку действия арбитражеров будут устранять все различия.

ТАЦ предполагает, что поскольку инвесторы владеют широко диверсифицированным портфелем ценных бумаг, только систематический риск оценивается на рынке. Однако систематический риск является функцией от нескольких факторов. Эти факторы не определены теоретически, но должны быть определены путем эмпирического исследования. Теория арбитражного ценообразования утверждает, что существует множество рассматриваемых факторов риска, которые вместе отвечают за наблюдаемую изменчивость доходности актива. Из этого предположения вытекает, что ожидаемое значение доходности (изменения цены) является линейной комбинацией значений факторов, которое измеряет подверженность актива влиянию каждого фактора риска.

Эти факторы риска ненаблюдаемы — мы можем знать только историческую информацию о доходности — но можем использовать факторный анализ для оценки влияния каждого фактора и отсюда премии за риск, связанный с каждым фактором риска.

Например, рассмотрим рынок, где существуют три фактора риска: F_1, F_2 и F_3. Трехфакторная модель доходности актива соотносит ожидаемую доходность конкретного инструмента с ограниченным числом экономических факторов и некоторыми идиосинкратическими элементами

$$R_{it} = rf + b_{i1}F_{1i} + b_{i2}F_{2t} + b_{i3}F_{3i} + e_{it}, \tag{11.26}$$

где rf — безрисковая доходность;
R_{it} — доходность актива i в момент времени t;
b_{it} — факторные нагрузки;
e_i — идиосинкратические элементы влияющие на ценную бумагу i;
F_{it} — премия за риск фактора i.

Существуют два важных статистических допущения в отношении к этой модели. Первое допущение — значения e не зависят друг от друга. Другое предположение — значения F также независимы.

Есть два источника риска, связанного с владением определенной ценной бумагой. Первый источник — это макроэкономические факторы, которые оказывают влияние на все ценные бумаги. Эти факторы составляют систематический риск. Второй источник — это характерный риск, уникальный для каждой ценной бумаги, и диверсификация может снизить уровень этого риска. Таким образом, на эффективном рынке будут выплачиваться премии только за риск, связанный с систематическими (макроэкономическими) факторами.

Поскольку эти факторы риска оказывают влияние на все активы, то доходность для каждого из этих факторов риска является рыночной ценой несения этого риска. Следовательно, доходность для каждого фактора одинакова для всех активов. Другими словами, вознаграждение за принятие на себя определенного типа риска устанавливается рынком. Однако степень, в которой актив подвержен данному типу риска зависит от характеристик самого актива. Другими словами, эта подверженность влиянию фактора или факторная нагрузка значения b_i будет уникальной для каждого актива.

Систематический риск отдельного актива определяется загруженноситью фактора и ковариациями доходностей факторов. Систематический риск портфеля ценных бумаг определяется на основе ковариаций доходностей факторов, загруженности факторов и удельного веса каждого актива в портфеле.

Для того чтобы теорию арбитражного ценообразования можно было успешно использовать, необходимо составить управляемый, т.е. относительно короткий список факторов, по которому возможно измерить ожидаемые премии за риск, связанные с каждым фактором. Для этого нужно иметь возможность оценить неожиданные изменения фактора. Следовательно, некоторые применения, по сути, моделируют неожиданные значения фактора, моделируя ожидаемое значение и вычитая его из действительного. Считается, что временные ряды некоторых факторов имеют такие шумы, что изменение данных само по себе является адекватной мерой неожиданного изменения.

Также необходимо иметь возможность измерить чувствительность каждой ценной бумаги к каждому фактору. К тому же эти чувствительности должны быть достаточно стабильными.

Ролл и Росс (Roll and Ross) (1980) использовали ФА и обнаружили, что всего лишь три, или, возможно, четыре фактора объясняли процесс формирования доходности американских ценных бумаг. Однако, Дримес, Френд и Гультекин (Dhrymes, Friend and Gultekin) (1984) отметили, что число факторов может зависеть от количества ценных бумаг в каждом портфеле.

Результаты проверки британского рынка с применением факторного анализа были неубедительными (незаконченными?). Бинсток и Чэн (Beenstock and Chan) (1986) выявили 20 факторов, которые объясняют доходность ценных бумаг. Более того, Диакоджианнис (Diacogiannis) (1986) обнаружил, что факторы были нестабильными во времени для одного и того же портфеля. Оба эти исследования показали, что число факторов зависит от того, как много ценных бумаг включено в каждый портфель.

Дальнейшее исследование британского рынка Абейсейкеры (Abeysekera) и Махаджана (Mahajan) (1987) привело к выводу, что они не смогли определить единый состав факторов для семи портфелей, рассматриваемых в исследовании.

УПРАЖНЕНИЯ

1. Объясните, как вы понимаете понятия "факторный анализ" и "анализ главных компонент". Определите разницу между двумя процессами.

2. Что такое собственные векторы и собственные значения? Каково их значение в факторном анализе и анализе главных компонент.
3. Используя методы, продемонстрированные в этой главе, и следующую дисперсионно-ковариационную матрицу 2 × 2, постройте собственные векторы и найдите собственные значения.

$$\begin{bmatrix} 0{,}00020 & 0{,}00006 \\ 0{,}00006 & 0{,}00030 \end{bmatrix}.$$

4. Нормализуйте собственные векторы из п. 3.
5. Постройте матрицы QDQ^{-1} относящиеся к матрице из п. 3. Проверьте, является ли $C = QDQ^{-1}$.
6. Найдите подверженности, выраженные как линейные комбинации переменных X и Y для дисперсионно-ковариационной матрицы из п. 3.
7. Объясните, как использовать стандартизованные переменные в анализе главных компонент.
8. Используя только обозначения, объясните, как использовать метод главных компонент при анализе составляющих дисперсии четырехлетней купонной облигации с ежегодными выплатами по купонам.
9. В контексте факторного анализа объясните, что вы понимаете под термином "общности".
10. Объясните, как факторный анализ применяется в теории арбитражного ценообразования.

ОТВЕТЫ

3. Первый собственный вектор равен $\begin{bmatrix} 1 \\ 2{,}135 \end{bmatrix}$.

Первое собственное значение равно 0,00033.

Второй собственный вектор равен $\begin{bmatrix} 1 \\ -0{,}468 \end{bmatrix}$.

Второе собственное значение равно 0,00017.

4. Нормализованные собственные векторы равны:

$$\begin{bmatrix} 0{,}424 \\ 0{,}906 \end{bmatrix}, \qquad \begin{bmatrix} 0{,}906 \\ -0{,}424 \end{bmatrix}.$$

5. QDQ^{-1} будет:

$$\begin{bmatrix} 0{,}424 & 0{,}906 \\ 0{,}906 & -0{,}424 \end{bmatrix} \begin{bmatrix} 0{,}00033 & 0 \\ 0 & 0{,}00017 \end{bmatrix} \begin{bmatrix} 0{,}424 & 0{,}906 \\ 0{,}906 & -0{,}424 \end{bmatrix}.$$

6. $0{,}424X + 0.906Y$; $0.906X - 0.424X$.

СПИСОК ИСПОЛЬЗУЕМОЙ И РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

Abeyskera, S. And Mahajan, A. (1987) A mest of the APT in pricing UK stocks. *Journal of Accounting and Finance*, **17**, 377—91.

Beenskock, M. and Chan, K. (1986) Testing the arbitrage pricing theory in the UK. *Oxford Bulletin of Economics and Statistics*, **50**, 27—39.

Dhrymes, P., Friend, I. And Gultekin, B. (1984) A critical reexamination of the empirical evidence on the arbitrage pricing theory. *Journal of Finance*, June, 323—46.

Diacogiannis, G. (1986) Arbitrage pricing model: a critical examination of its empirical applicability for the London Stock Exchange. *Journal of Business Finance and Accountancy*, Winter.

Kahn, R. H. (1989)Risk and return in the US bond market: a multi-factor approach. In F. J. Fabozzi (Ed.) *Advances and innovations in the Bond and Mortgage Markets*. Probus publishing, Chicago.

Kahn, R. H., (1990) Estimating the US treasury term structure of interest rates. In F. J. Fabozzi (ed.) *The Handbook of US Treasury and Governmenet Agency securities: Instruments, Strategies and Analysis*. Probus publishing, Chicago.

Kahn, R. H and Gulrajani, D. (1993) Risk and return in the Canadian bond market. *Journal of Portfolio Management*, Spring, 86—93.

Karki, J. and Reyes, C. (1994) Model relationship. *Risk*, December.

Roll, R. and Ross, S. A. (1980) An empirical investigation of the arbitrage pricing theory. *Journal of Finance*, December.

Стандартное нормальное распределение

Пример

Pr $(0 \le z \le 1{,}96) = 0{,}4750$

Pr $(z \ge 1{,}96) = 0{,}5 - 0{,}4750 = 0{,}025$

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0711	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1106	0,1113	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1811	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3310	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3S99	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4117	0,4162	0,1177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,1265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4118	0,4429	0,1141
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,1591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,1678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,1756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,47R8	0,1793	0,4798	0,1803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,1854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4915	0,4946	0,4948	0,4919	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4971
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4981
3,0	0,4987	0,4987	0,1987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990

Процентные пределы для распределения Стьюдента

Степени свободы	Область верхнего хвоста 0,10	0,05	0,025	0,01	0,005
	Доверительный интервал 90%	95%	97,5%	99%	99,5%
1	3,078	6,314	12,706	31,821	63,657
2	1,886	2,920	4,303	6,965	9,925
3	1,638	2,353	3,182	4,541	5,841
4	1,533	2,132	2,776	3,747	4,604
5	1,476	2,015	2,571	3,365	4,032
6	1,440	1,943	2,447	3,143	3,707
7	1,415	1,895	2,365	2,998	3,499
8	1,397	1,860	2,306	2,896	3,355
9	1,383	1,833	2,262	2,821	3,250
10	1,372	1,812	2,228	2,764	3,169
11	1,363	1,796	2,201	2,718	3,106
12	1,356	1,782	2,179	2,681	3,055
13	1,350	1,771	2,160	2,650	3,012
14	1,345	1,761	2,145	2,624	2,977
15	1,341	1,753	2,131	2,602	2,947
16	1,337	1,746	2,120	2,583	2,921
17	1,333	1,740	2,110	2,567	2,898
18	1,330	1,734	2,101	2,552	2,878
19	1,328	1,729	2,093	2,539	2,861
20	1,325	1,725	2,086	2,528	2,845
21	1,323	1,721	2,080	2,518	2,831
22	1,321	1,717	2,074	2,508	2,819
23	1,319	1,714	2,069	2,500	2,807
24	1,318	1,711	2,064	2,492	2,797
25	1,316	1,708	2,060	2,485	2,787
26	1,315	1,706	2,056	2,479	2,779
27	1,314	1,703	2,052	2,473	2,771
28	1,313	1,701	2,048	2,467	2,763
29	1,311	1,699	2,045	2,462	2,756
30	1,310	1,697	2,042	2,457	2,750
40	1,303	1,684	2,021	2,423	2,704
60	1,296	1,671	2,000	2,390	2,660
120	1,289	1,658	1,980	2,358	2,617
∞	1,282	1,645	1,960	2,326	2,576

χ^2 – критические значения

Вероятность	0,250	0,100	0,050	0,025	0,010	0,005	0,001
ν = 1	1,32	2,71	3,84	5,02	6,63	7,88	10,8
2	2,77	4,61	5,99	7,38	9,21	10,6	13,8
3	4,11	6,25	7,81	9,35	11,3	12,8	16,3
4	5,39	7,78	9,49	11,1	13,3	14,9	185
5	6,63	9,24	11,1	12,8	15,1	16,7	20,5
6	7,84	10,6	12,6	14,4	16,8	18,5	22,5
7	9,04	12,0	14,1	16,0	18,5	20,3	24,3
8	10,2	13,4	15,5	17,5	20,3	22,0	26,1
9	11,4	14,7	16,9	19,0	21,7	23,6	279
10	12,5	16,0	18,3	20,5	23,2	25,2	296
11	13,7	17,3	19,7	21,9	24,7	26,8	313
12	14,8	18,5	21,0	23,3	26,2	28,3	32,9
13	16,0	19,8	22,4	24,7	27,7	29,8	34,5
14	17,1	21,1	23,7	26,1	29,1	31,3	36,1
15	18,2	22,3	25,0	27,5	30,6	32,8	37,7
16	19,4	23,5	26,3	28,8	32,0	34,3	39,3
17	20,5	24,8	27,6	30,2	33,4	35,7	40,8
18	21,6	26,0	28,9	31,5	34,8	37,2	423
19	22,7	27,2	30,1	32,9	36,2	38,6	43,8
20	23,8	28,4	31,4	34,2	37,6	40,0	45,3
21	24,9	29,6	32,7	35,5	38,9	41,4	46,8
22	26,0	30,8	33,9	36,8	40,3	42,8	48,3
23	27,1	32,0	35,2	38,1	41,6	44,2	49,7
24	28,2	33,2	36,4	39,4	43,0	45,6	51 2
25	29,3	34,4	37,7	40,6	44,3	46,9	52,6
26	30,4	35,6	38,9	41,9	45,6	48,3	54,1
27	31,5	36,7	40,1	43,2	47,0	49,6	55,5
28	32,6	37,9	41,3	44,5	48,3	51,0	56,9
29	33,7	39,1	42,6	45,7	49,6	52,3	58,3
30	34,8	40,3	43,8	47,0	50,9	53,7	59,7
40	45,6	51,8	55,8	59,3	63,7	66,8	73,4
50	56,3	63,2	67,5	71,4	76,2	79,5	86,7
60	67,0	74,4	79,1	83,3	88,4	92,0	99,6
70	77,6	85,5	90,5	95,0	100	104	112
80	88,1	96,6	102	107	112	116	125
90	98,6	108	113	118	124	128	137
100	109	118	124	130	136	140	149

Где ν — число степеней свободы.

ОГЛАВЛЕНИЕ