

რეზიზან მუქვაბიძე

ეკონომიკური
პრაქტიკაში

რუიზან მექვაბიძე

ეკონომეტიკა პრაქტიკაში

2009

“ეკონომეტრიკა პრაქტიკაში“ წარმოდგენს გზამკვლევეს ელემენტარულ ეკონომეტრიკაში. წიგნი მოიცავს ერთი განტოლებით წარმოდგენილ რეგრესიულ ანალიზს გასაგებ ფორმატში. ტექსტში მოცემულია ეკონომეტრიკის სწავლისათვის აუცილებელი სტატისტიკური კონცეფციები რეგრესიული ანალიზის გაგების თვალსაზრისით. წიგნი გამიზნულია არა მარტო ეკონომიკური პროფილის სტუდენტებისა და ბიზნეს ანალიტიკოსებისათვის, არამედ პრაქტიკოსებისთვისაც, რომლებიც დაინტერესებულნი არიან მოდელირებისა და პროგნოზირების საკითხებით.

წიგნში წარმოდგენილი მასალა ტრადიციულია, თუმცა მასში გათვალისწინებულია სპეციფიური მახასიათებლებიც, რომელთა შესწავლა სასარგებლო იქნება არა მარტო საბაკალავრო, არამედ სამაგისტრო კურსის დონეზეც რაოდენობრივ მეთოდებში. ავტორის მიერ, წიგნში გათვალისწინებულია ლისაბონის ეკონომიკისა და ბიზნეს ადმინისტრირების ინსტიტუტის პროფესორების: მაქსიმიანო პინჰეიროს, ხოსე მატოს პოსოს, კარლოს სილვა რიბეიროს, ხოსე სანტოს სილვოს და ხოსე ანდრადე დე სილვას (TEMPUS, Contract NCD_JEP-21058-2000-2002), მოსკოვის ეკონომიკური სკოლის პროფესორების: ვიქტორ პოლტეროვიჩის, პაველ კატიშევის და ანატოლი პერესეცკის (HESP, project B1639 #H2981639/03) კონსულტაციები და რჩევები და პროფესორების: ჯიმ ლეიტცელის (ჩიკაგოს უნივერსიტეტი), კარინე ტოროსიანის (თსუ-სტან არსებული საერთაშორისო ეკონომიკური სკოლა), რობერტ ჩაიძე (თსუ-სთან არსებული საერთაშორისო ეკონომიკური სკოლა), რენდელ ფილერის (ნიუ იორკის უნივერსიტეტი), მარკ ვაინშტაინი (სამხრეთ კალიფორნიის უნივერსიტეტი), მიშელ ბინსტოკის (იერუსალიმის უნივერსიტეტი) მიერ ჩატარებული მასტერ კლასების მოსმენითა და მონაწილეობით მიღებული გამოცდილება (The program of ISET: Excellent Teaching for Decision Making in Economics, OSI/HESP/Budapest, 2009-2012)

დიზაინერი: ფიქრია მექვაბიძე

Ruizan Mekvabidze

ECONOMETRICS IN PRACTICE

2009

330.1(075..8)
45678910-9998

„ **Econometrics in Practice** “ is a practical guide to the understanding of elementary econometrics. It covers the topic of single-equation linear regression analysis in an easily understandable format. All statistical concepts necessary for econometric study are covered in the text, but they are covered only to the extent needed for understanding of regression analysis. The book is aimed not only for economic profile students and business analysts but also for regression users looking for at experienced practitioners interested in the application of modeling and forecasting.

The material covered by this book is traditional but there are some of specific features and the book can be used not only in undergraduate courses but in magisterial level courses in quantitative methods .

The author used the consultations and advices of the professionals and professors Jose Maximiano Pinheiro, , Carlos Silva Ribeiro, Jose Matos Pasos, Jose Santos Silva, Jose Andrade e Silva of Lisbon Institute of Economics and Business Administration (TEMPUS project , Contract NCD_JEP-21058-2000-2002) and professors Victor Polterovich, Pavel Katyshev, Anatoly Peresetsky of Moscow Economic School (HESP project B1639, #H2981639/02). Also, the experience by attending and participating of the master classes provided by professors : Jim Leitzel (University of Chicago), Karine Torosyan (International school of economics at TSU), Robert Chaidze (International school of economics at TSU), Randal Filer (University of New York), Mark Weinstein (University of Southern California), Michael Beenstock (Hebrew University of Jerusalem) (The program of ISET: Excellent Teaching for Decision Making in Economics, OSI/HESP/Budapest, 2009-2012)

Designer: Pikria Mekvabidze

შინაარსი

რა უნდა ვიცოდეთ 1

What we should know 4

წინასიტყვაობა 6

შესავალი 9

თავი 1. რეგრესიული ანალიზის მიმოხილვა 20

1.1. წრფივი რეგრესია 20

1.2. რეგრესიის სავარაუდო განტოლება 22

1.3. რეგრესიული ანალიზის უმარტივესი მაგალითები 23
რეზიუმე 25

დამოუკიდებელი სამუშაო 26

თავი 2. უმცირეს კვადრატთა მეთოდი 27

2.1. რეგრესიის წრფის შერჩევა 27

2.2. უმცირეს კვადრატთა მეთოდი 28

2.3 მულტივარიაციული რეგრესიის მოდელი 33

2.4. რეგრესიის განტოლების თვისობრივი შეფასება 36

2.5. დეტერმინაციის კოეფიციენტი 37

2.6. ლაგური ცვლადები 40

2.7. ფიქტიური ცვლადები 40

რეზიუმე 42

დამოუკიდებელი სამუშაო 42

თავი 3. კლასიკური მოდელი 48

3.1. გაუს-მარკოვის თეორემა 50

3.2. სტანდარტული ეკონომეტრიკული შენიშვნები 51

რეზიუმე 51

დამოუკიდებელი სამუშაო 52

თავი 4. ჰიპოთეზების ტესტირება 53

4.1. რას არის ჰიპოთეზების ტესტირება 53

4.2. რეგრესიის კოეფიციენტების მიმართ ჰიპოთეზების შემოწმება 54

4.3. ნდობის ინტერვალი 58

4.4. ცალმხრივი და ორმხრივი t ტესტის გამოყენება 58

რეზიუმე 59

დამოუკიდებელი სამუშაო 60

თავი 5. მოდელის სპეციფიკაცია 61

5.1. სტატისტიკური შეცდომების წყაროები 63

5.2. გამოტოვებული ცვლადები 64

5.3. განტოლებისათვის მიუღებელი ცვლადები 70

5.4. სპეციფიკაციის ოთხი მნიშვნელოვანი კრიტერიუმი 71

5.5. ნახტომისებური რეგრესიული ანალიზის პროცედურა 71

რეზიუმე 74

დამოუკიდებელი სამუშაო 74

თავი 6. მოდელის ფუნქციონალური ფორმის შერჩევა 76

6.1. განტოლების მუდმივი წევრი და მისი ინტერპრეტაცია 76

6.2. ალტერნატიული ფუნქციონალური ფორმა 77

6.2.1. წრფივი ფორმა 77

6.2.2. $\log -\log$ ფორმა 78

6.2.3. semi-log ფორმა	78
6.2.4. პოლინომიური ფორმა	79
6.2.5. ინვერსიული ფორმა	79
6.3. პრობლემები არაკორექტული ფუნქციონალური ფორმის შერჩევისას	80
რეზიუმე	82
დამოუკიდებელი სამუშაო	83
თავი 7. ჰეტეროსკედასტიურობა	84
7.1. ჰეტეროსკედასტიურობის ეფექტი ჰიპოთეზების პროგნოზირების ტესტებზე	85
7.2. ჰეტეროსკედასტიურობის ტესტირება	87
7.3. ჰიპოთეზა ჰეტეროსკედასტიურობის შესახებ	90
რეზიუმე	92
დამოუკიდებელი სამუშაო	93
თავი 8. დროითი მწკრივები	94
8.1. ტენდენციის არსი	96
8.2. სეზონური ეფექტები და მათი ანალიზი	98
8.3. დროითი მწკრივების სტატისტიკური მოდულების მახასიათებლები	98
8.4. სტატისტიკური მოდელი ავტორეგრესიის პროცესისასთვის	99
8.5. დარბინ – უატსონის (Durbin-Watson) ტესტი AR(1) პროცესისათვის	101
8.6. მაგალითები AR(1) დროითი მწკრივისა და ავტოკორელაციის ფუნქციებისა	106
8.7. მცოცავი საშუალოს პროცესი	103
8.8. ბოქს-ჯენკინსის მეთოდი	105
8.9. დროითი მწკრივების მოდულები	106
რეზიუმე	107
დამოუკიდებელი სამუშაო	107
თავი 9. ერთდროულ განტოლებათა სისტემები	108
9.1. იდენტიფიკაციის პრობლემა	113
რეზიუმე	116
დამოუკიდებელი სამუშაო	116
თავი 10. პროგნოზირება	118
რეზიუმე	120
დამოუკიდებელი სამუშაო	120
თავი 11 პროექტი ეკონომეტრიკაში	121
11.1. მოდელი	121
11.2. სტატისტიკური მონაცემები	122
11.3. მოდელის შეფასება და ანალიზი	122
11.4. პროექტის გაფორმება	123
თავი 12. ეკონომეტრიკის პროგრამული უზრუნველყოფა	123
12.1. პროგრამული უზრუნველყოფის მოკლე მიმოხილვა	123
12.2. STATA Windows	127
12.3. Small STATA 5.0 –ის ძირითადი ბრძანებები და განმარტებები	-131
12.4. მოდელის შეფასება	141
12.5. გრაფიკები	144
ლიტერატურა	150

რა უნდა ვიცოდეთ

ეკონომეტრიკას სამი ფუნქცია აქვს: ეკონომიკური სინამდვილის აღწერა, ეკონომიკური თეორიის შესახებ ჰიპოთეზების ტესტირება, ეკონომიკური აქტივობის პროგნოზირება.

წიგნის მიზანია ეკონომეტრიკის მეთოდების ინტეგრირება ემპირიულ ეკონომიკურ კვლევაში, რადგან ეკონომიკა, ეკონომიკური მონაცემები და ეკონომეტრიკა განსაზღვრავენ გამოყენებითი ეკონომეტრიკის მიმართულებას, რის გამოც უნდა ვიცოდეთ შემდეგი:

1. რეგრესიული ანალიზი განსაზღვრავს დამოკიდებულ ცვლადს, როგორც დამოუკიდებელი ცვლადების ფუნქციას. რეგრესიული ანალიზი შეიძლება გამოვიყენოთ განტოლებისათვის, რომელიც წრფივია ამ განტოლების კოეფიციენტების მიმართ და ამასთან, რეგრესიის კოეფიციენტები წარმოდგენილია ყველაზე მარტივი ფორმით. სტოქსტური წევრი აუცილებლად უნდა დაემატოს რეგრესიის ყველა განტოლებას, რათა გავითვალისწინოთ დამოკიდებული ცვლადის შემთხვევითი ცვლილებები.

განტოლების შემთხვევითი კომპონენტი მოიცავს:

ა/ ცვლადებს, რომლებიც არ მონაწილეობენ რეგრესიის განტოლებაში;

ბ/ მონაცემების შეგროვებისას დაშვებულ შეცდომებს;

გ/ თეორიული განტოლების შეცდომას, რომელსაც სხვა ფუნქციონალური ფორმა აქვს ვიდრე რეგრესიის განტოლებას.

სავარაუდო რეგრესიის განტოლება აპროქსიმაციაა იმ ჭეშმარიტი განტოლებისა, რომელიც მიღებულია დამოკიდებული ცვლადის Y -ის ფაქტობრივი მნიშვნელობების მიხედვით მოცემული შერჩევიდან. ჩვენ არაფერი ვიცით ჭეშმარიტი განტოლების შესახებ და ეკონომეტრიკული ანალიზი მთელ ყურადღებას მიმართავს სავარაუდო რეგრესიის განტოლებისაკენ და აფასებს ამ განტოლების კოეფიციენტებს.

2. რეგრესიის განტოლების კოეფიციენტების შესაფასებლად უმცირეს კვადრატთა მეთოდი ყველაზე ხშირად გამოყენებული მეთოდია. ამ მეთოდით მიღებული პარამეტრების მნიშვნელობები აკმაყოფილებენ სხვაობების კვადრატების მინიმუმის პირობას ანუ $\min \sum e_i^2$ მიისწრაფის ნულისაკენ. რეგრესიის განტოლების შესაფასებლად ერთ-ერთი მარტივი მაჩვენებელია დეტერმინაციის კოეფიციენტი. არსებითად, ამ კოეფიციენტის მნიშვნელობა ყოველთვის იზრდება, როდესაც რეგრესიის განტოლებაში დამატებულია ცვლადი.

ყოველთვის გვახსოვდეს, რომ შესაფასებელი განტოლებისათვის პროგნოზის ხარისხი დამოკიდებულია რეგრესიის განტოლების კოეფიციენტებზე, რომელთა სიზუსტე უნდა განსაზღვრავდეს ეკონომიკური თეორიის მოთხოვნას რეგრესიის განტოლებაზე.

რეგრესიული ანალიზის ტიპური ეტაპებია:

- მოდელის სპეციფიკაცია: დამოუკიდებელი ცვლადების გამოყოფა და ფუნქციონალური ფორმის აღწერა;
- განტოლების კოეფიციენტების ნიშნების დადგენა;

- მონაცემების შეგროვება;
- განტოლების შეფასება;
- შედეგების ანალიზი.

ფიქტიური ცვლადების შემოტანა რეგრესიის განტოლებაში დაკავშირებულია დამოუკიდებელი ცვლადის შინაარსთან. კერძოდ, თუ დამოუკიდებელი ცვლადი თვისობრივი ხასიათის მატარებელია, მაშინ მისი წარმოდგენა დასაშვებია ფიქტიური ბინარული ცვლადით, რომელიც მხოლოდ ორ მნიშვნელობას იღებს: 0-ს ან 1-ს.

3. რეგრესიის კლასიკური მოდელის ძირითადი თვისებებია:

ა/ რეგრესიის მოდელი წრფივია კოეფიციენტების მიმართ და შეიცავს შემთხვევით წევრს;

ბ/ შემთხვევითი წევრის საშუალო მნიშვნელობა ნულის ტოლია;

გ/ დამოუკიდებელი ცვლადები კორელირებულნი არ არიან შემთხვევით წევრთან და შემთხვევით წევრზე დაკვირვებების შედეგები ასევე, ერთმანეთთან კორელირებულები არ არიან;

ე/ შემთხვევითი წევრის ვარიაცია (დისპერსია) მუდმივია;

ვ/ არც ერთი დამოუკიდებელი ცვლადი არ არის დანარჩენი დამოუკიდებელი ცვლადების წრფივი ფუნქცია;

ზ/ შემთხვევითი წევრის განაწილება ნორმალური განაწილების კანონს ემორჩილება;

უმცირეს კვადრატთა მეთოდის გამოყენებით მიღებულ შეფასებას ორი მნიშვნელოვანი თვისება გააჩნია: გადაუადგილებადობა და ვარიაციის მინიმუმი, რადგან უმცირეს კვადრატთა მეთოდის გამოყენებით მიღებული რეგრესიის კოეფიციენტების - \hat{S}_k შეფასება საუკეთესო შეფასებაა წრფივ შეფასებათა გადაუადგილებად კლასში. მას აქვს საინტერესო თვისებები და უპირველესად ის, რომ \hat{S}_k -ს ვარიაცია დამოკიდებულია დაკვირვებათა რიცხოვნებაზე, ხოლო მისი მნიშვნელობის გაზრდა ან შემცირება დამოკიდებულია შერჩევაში დაკვირვებათა რიცხოვნების გაზრდასა ან შემცირებაზე;

4. ჰიპოთეზის ტესტირება საშუალებას გვაძლევს გენერალური ერთობლიობიდან შერჩევის გზით შევამოწმოთ მოსაზრება ეკონომიკური თეორიის ბაზაზე აგებული რეგრესიის განტოლების კოეფიციენტების მიმართ მათი ჭეშმარიტების შესახებ.

ტესტირება მოიცავს ოთხ ძირითად ეტაპს:

ა/ ნულოვანი და ალტერნატიული ჰიპოთეზების ჩამოყალიბება;

ბ/ მნიშვნელოვნების დონის შერჩევა და კრიტიკული მნიშვნელობის t –სტატისტიკის გამოთვლა;

გ/ რეგრესიის განტოლების დადგენა;

დ/ გადაწყვეტილების მიღება t – ტესტის გამოთვლილი და კრიტიკული მნიშვნელობების შედარებით, რომლის საფუძველზეც ვიდებთ ან უარვყოფთ ნულ ჰიპოთეზას.

F -ტესტი გამოიყენება ნულოვანი ჰიპოთეზის შესამოწმებლად იმ შემთხვევაში, როდესაც რეგრესიის განტოლება შეიცავს ერთზე მეტ კოეფიციენტს.

ჰიპოთეზების შემოწმებისას ადგილი აქვს ორი ტიპის შეცდომას:

ა/ პირველი ტიპის შეცდომა: უარვყოფთ ნულოვან ჰიპოთეზას, როდესაც ის ჭეშმარიტია;

ბ/ მეორე ტიპის შეცდომა: არ უარვყოფთ ნულოვან ჰიპოთეზას, როდესაც ის მცდარია.

განტოლებიდან გამოტოვებული ცვლადი დარჩენილი კოეფიციენტების შეფასებაში წანაცვლების გამომწვევი მიზეზია, რადგან გამოტოვებული ცვლადი განტოლებაში ჩართულ ცვლადებთან კორელირებულია.

ცვლადის შეყვანა რეგრესის განტოლებაში, რომელიც ამ განტოლებისათვის არსებითი არ არის, გამოიწვევს იმ ცვლადების კოეფიციენტების წანაცვლებას, რომლებიც განტოლებაშია შეყვანილი და ასევე გამოიწვევს მათი ვარიაციის მნიშვნელობების გაზრდას, t -ტესტის მნიშვნელობისა და კორექტირებული დეტერმინაციის კოეფიციენტის დაწევას.

განტოლებაში ცვლადის შესაყვანად სასარგებლოა გამოვიყენოთ შემდეგი:

ა/ თეორია;

ბ/ t -ტესტი;

გ/ წანაცვლება;

დ/ \bar{R}^2 .

თეორია და არა სტატისტიკური მაჩვენებლების მორგება წარმოადგენს მნიშვნელოვან მომენტს სამიხედელ განტოლებაში შესაყვანი ცვლადის დასადგენად.

5. განტოლების ფუნქციონალური ფორმის შერჩევის საფუძველს ეკონომიკური თეორია წარმოადგენს. წრფივი ფორმა მისაღებია იმ შემთხვევაში, თუ სპეციფიური ჰიპოთეზა არ არის შემოთავაზებული. წინააღმდეგ შემთხვევაში, ფუნქციონალური ფორმა, რომელიც ცვლადების მიმართ არაწრფივია, შეიცავს ორმაგ ლოგარითმულ $\sim \log\text{-}\log$) ფორმას ან ნახევრად ლოგარითმულ $\sim \text{semi-log}$) ფორმას, ან პილინომიალურ ფორმას ან ინვერსიულ ფორმას.

არაწრფივი ფუნქციონალური ფორმის გამოყენება ხასიათდება პრობლემებით. კერძოდ, კორექტირებული დეტერმინაციის კოეფიციენტი შედარების თვალსაზრისით მიუღებელია, თუ დამოკიდებულმა ცვლადმა ტრანსფორმაცია განიცადა ან ნაშთების მნიშვნელობები საკმარისად დიდია.

6. დროითი მწკრივებისა და პროგნოზირების ანალიზისას მნიშვნელოვანია შემდეგი:

ლაგ-მოდელი განიხილავს დამოკიდებული ცვლადის Y -ის მიმდინარე მნიშვნელობას, როგორც X დამოუკიდებელი ცვლადის მიმდინარე და გასული პერიოდის მნიშვნელობათა ფუნქციას. მცირე მოცულობის შერჩევებში უმცირეს კვადრატთა მეთოდით შეფასებული ლაგ-მოდელი კოიკის განაწილებით წანაცვლებადია და მაჩვენებლების ტესტირებისას ჰიპოთეზა არასაიმედოა.

პროგნოზირება დამოკიდებულია ცვლადის მოსალოდნელი მნიშვნელობის იმ დაკვირვებისათვის, რომელიც შერჩევის მონაცემების მიმდევრობის ნაწილს არ წარმოადგენს.

პროგნოზი პირობითია, თუ რეგრესიის განტოლებაში შემავალი ყველა დამოუკიდებელი ცვლადის მნიშვნელობები პროგნოზირებადია, ხოლო თუ რეგრესიის განტოლებაში შემავალი ყველა დამოუკიდებელი ცვლადის მნიშვნელობები განსაზღვრულია, მაშინ პროგნოზი უპირობოა. პროგნოზის შედეგი განსაკუთრებით საინტერესოა მაშინ, როდესაც მისი მნიშვნელობა ნდობის ინტერვალითაა შეფასებული.

ARIMA (autoregressive integrated moving average) – ავტორეგრესიის მოდელი ინტეგრირებული მცოცავი საშუალოთი იყენებს რა დამოკიდებული ცვლადის

მიმდინარე და გასული პერიოდის მნიშვნელობებს, მისაღები სიზუსტით იძლევა ამ ცვლადის მოკლევადინ პროგნოზს.

ARIMA(p,d,q) აერთიანებს p-რიგის ავტორეგრესიის პროცესს და q-რიგის მცოცავი საშუალოს პროცესს.

What we should know

The objective of this book is to integrate at an elementary level econometrics methods in empirical economic research. The intersection of economics, econometrics and economic data defines the field of applied econometrics. In the analysis of economic problems, there often exists much prior information about the unknown economic parameters.

1. Econometrics has three major uses: describing economic reality; testing hypotheses about economic theory; forecasting future economic activity. While regression analysis specifies that a dependent variable is a function of one or more independent variables, regression analysis alone can not prove or even imply causality.

Linear regression can only be applied to equation that are linear in the coefficients. A stochastic error term must be added to all regression equations. The components of this error term include:

a/ left-out variables;

b/ measurement errors in the data;

c/ purely random and unpredictable events.

An estimated regression equation is an approximation of the true equation that is obtained by using data from a sample of actual Y and X.

2. Ordinary Least Squares –OLS is the most frequently used method of obtaining estimates of the regression coefficients from a set of data. The chooses parameters minimize the condition for residuals $\sum e_i^2 \rightarrow 0$ for a particular sample.

The coefficient of determination is a simplest measure of the estimated equation. It always increases when a variable is added to an equation.

The quality of fit of an estimated equation is only one of the measures of the overall quality of that regression.

The regression analyses has the next steps:

a/ Review and develop the theoretical model

b/ Specify the model

c/ Hypotheses

d/ Collect the data

e/ Estimate and evaluate the equation

f/ Analyses of the results

3. The classical regression model includes the next properties:

a/ The regression model is linear with an error terms;

b/ Error term has a mean of zero;

c/ An error term is uncorrelated with the independent variables and other observations of the error terms are not correlated;

d/ An error term has a constant variance ;

e/ The independent variables must not be perfect linear functions of each other;

f/ An error term is normally distributed;

Estimator has two most important properties: unbiasedness and minimum variance;

The sampling distribution of the OLS estimator of $\hat{\beta}_k$ is excellent, it has desirable properties. The variance or the degree of dispersion of the sampling distribution of $\hat{\beta}_k$, decreases as the number of observations increases.

4. Hypothesis testing makes inferences about the validity of specific economic theories from a sample of the population for which the theories are supposed to be true. The four basic steps of Hypothesis testing are:

- a/ set up the null and alternative Hypotheses;
- b/ Choose a level of significance and critical t-value;
- c/ Run the regression and obtain an estimated t-value;
- d/ Apply the decision rule by comparing the calculated t-value with the critical t-value in order to reject or accept the null hypothesis.

The two kinds of error we can make in hypothesis testing are:

Type I: We reject a null hypothesis that is true;

Type II: We do not reject a null hypothesis that is false.

The **F-test** is a method of testing a null hypothesis when the regression equation includes more than one coefficient

The omission of a variable from an equation will cause bias in the estimates of the remaining coefficients because the omitted variable is correlated with included variables.

Including a variable in an equation in which it is actually irrelevant does cause bias, but it will usually increase the variances of the included variables' estimated coefficients and lowering their t-values and lowering \bar{R}^2 .

Four useful criteria for the inclusion of a variable in an equation are:

- a/ Theory;
- b/ t-test;
- c/ Bias;
- d/ \bar{R}^2

Theory, not statistical fit, should be the most important criterion for the inclusion of a variable in a regression equation

5. The choice of a functional form should be based on the economic theory. A form that is linear in the variables should be used unless a specific hypothesis suggests otherwise. Functional forms that are nonlinear in the variables include the double-log (log-log) form, the semi-log form, the polynomial form, the inverse form. The use of nonlinear functional forms has a number of potential problems. In particular, the \bar{R}^2 are difficult to compare if Y has been transformed and the residuals are potentially large

6. Distributed lag model explains the current value Y as a function of current and past values of X. In small samples, OLS estimates of a Koyck distributed lag model are biased and have unreliable hypotheses testing properties.

Forecasting is the estimation of the expected value of a dependent variable for observation that are not part of the sample data set. A forecast in which all the values of the independent variables are known with certainty is called an unconditional forecast.

A forecast in which one or more of the independent variables have to be forecasted it is a conditional forecast.

Forecasts are often useful if they are accompanied a confidence interval.

ARIMA is a refined curve fitting technique that uses current and past values of the dependent variable to produce accurate short-term forecasts of that variable.

The **ARIMA(p,d,q)** approach combines an autoregressive process of order p with a moving average process of order q .

წინასიტყვაობა

რაოდენობრივი მეთოდების მიმართ მნიშვნელოვნად გაიზარდა ინტერესი, რაც ძირითადად გამოწვეულია ფინანსური ბაზრების განვითარებით. პროფესიონალების მიერ ამ დარგში წარმატებით გამოიყენება სტატისტიკური ტექნიკა პორტფოლიოების მართვისას ფასიანი ქაღალდების რეგულირების მიზნით. დღეს პროფესიონალების ყურადღება მიმართულია ფინანსური დროითი მწკრივების მოდელირებისკენ, რომელიც მოიცავს შემოსავლების პროგნოზირებას, ფინანსური ბაზრების მიკროსტრუქტურას, კაპიტალაქტივების ფასწარმოქმნის მოდელს, ფასწარმოქმნის საარბიტრაჟო თეორიას, საპროცენტო განაკვეთების სტრუქტურას, ფინანსური ბაზრების არაწრფივ მოდელს, ქაოსის თეორიას და სხვა.

სტუდენტების ერთი ნაწილისათვის ეკონომეტრიკა რატომღაც განყენებული და ზოგჯერ გაუგებარია. რა არის ამის მიზეზი? ერთის მხრივ, ეკონომისტები თეორიულად ასაბუთებენ, რომ მაგალითად, ფასების წონასწორობა დაფუძნებულია კარგად გააზრებულ ზღვრულ შემოსავლებსა და ღირებულებებზე, ხოლო მეორეს მხრივ, ფირმები მუშაობენ ისე, თითქოს ასეთი კონცეფცია არასოდეს არ გაუგონიათ. ეკონომეტრიკა საშუალებას გვაძლევს შევისწავლოთ ესა თუ ის ეკონომიკური მოვლენა ან პროცესი მათ შესახებ მონაცემების მიხედვით და რაოდენობრივად შევაფასოთ ფირმების, მომხმარებლების, მთავრობების ქმედებების შედეგები.

90-იან წლებიდან გამოჩნდა დღეისათვის საკმაოდ ცნობილი მონოგრაფიები და სახელმძღვანელოები ეკონომეტრიკაში, რომელთა გამოცემა ბიზნესად გადაიქცა. ამ მიმართულებით, დღეს მიმოქცევაშია საკმარისი რაოდენობა წიგნებისა და მათ მიმართ სხვადასხვა მოსაზრება არსებობს. მათი ძირითადი მიმართულება ეკონომეტრიკის თეორიაა. 1994 წ. ნობელის პრემიის ლაურეატმა კლაივ გენჯერმა გამოსცა 4 სახელმძღვანელო, რომელთაგან სამი მათგანი შემდეგი ავტორებისაა: გრინი, გოლდბერგერი და დევიდსონ-მაკკინონი. გრენჯერის აზრით "მასალა თანმიმდევრობითაა დალაგებული, მაგრამ რეალური პრობლემებიდან, რომლებთანაც პრაქტიკოსს უხდება შეხება, ძალიან შორსაა. მართალია, სტუდენტები ეუფლებიან სტანდარტული ეკონომეტრიკის ტექნიკურ ასპექტს, მაგრამ მათ ნაკლები იციან რეალური მონაცემების მოდელირების შესახებ სპეციფიკაციის პირობებში".

ელის ნაკამურა 2Nakamura A./ – Ernst R. Berndt. The Practice of Econometrics: Classic and Contemporary. Addison – Wesley, 1991, 702 p. სახელმძღვანელოს შესახებ აღნიშნავს: "არ არის წიგნი ნაკლის გარეშე. პრაქტიკოსი ეკონომეტრისტი სისტემატიურად სრულყოფენ თავიანთ გაგებას. მე კმაყოფილი ვარ ამ წიგნის კითხვით. მე მან ზევრი რამ მასწავლა".

საინტერესოა ეკონომეტრიკაში თანამედროვე სახელმძღვანელოების შეფასება სტანისლავ ანატოლევის (,2008) მიერ მათში განხილული თემების მიხედვით.

თემა GO D GR JD H R JM CT B
M

რეგრესიული გეომეტრია	-	+	-	-	-	+	-	-	-
			+	+					
სიმულაციაზე დაფუძნებული შეფასება	-	-	-	-	-	-	-	+	-
						+			
მოდელის შერჩევა	-	-	-	-	-	-	+	+	-
	+		+	+					
არაპარამეტრული მეთოდები	-	-	-	+	-	-	+	+	-
			+						
ავტორეგრესია	+	+	+	+	+	+	-	-	-
									+
ემპირიული ილუსტრაცია	+	-	+	+	+	+	-	+	+

აღნიშვნები

ავტორები

GO	Arthur S. Goldberger
DM	Davidson&James G. Maccinnon
Gr	Davidson&James G. Maccinnon
JD	John Johnston & John Dinardo
H	Fumio Hayashi
R	Paul A. Rund
MJM	Ron C. Mittelhammer
CT	Colin Cameron & Pravin K. Trivedi
B	Ernst R. Berndt

ეკონომეტრიკის პრაქტიკული გამოყენების თვალსაზრისით სამი ძირითადი მიმართულება მაინც შეიძლება დავინახოთ:

1. ეკონომიკური სინამდვილის აღწერა;
2. ეკონომიკური თეორიის ჰიპოთეზების ტესტირება;
3. ეკონომიკური აქტივობის პროგნოზი .

სწორედ ეკონომეტრიკის პრაქტიკული დანიშნულებიდან გამომდინარე, ეკონომეტრიკის მიზანია გამოუმუშავოს სტუდენტს ეკონომიკური მოდელების აგების ანალიტიკური აზროვნება და მოამზადოს სტუდენტი პრაქტიკული ეკონომიკური ამოცანების ამოსახსნელად, რისთვისაც მასალა დალაგებულია შემდეგი მიმდევრობით:

1. ეკონომიკური მოდელის ფორმულირება;
2. მონაცემების გენერირების პროცესის ჩამოყალიბება;
3. მონაცემების თანმიმდევრობითი იდენტიფიცირება;
4. ეკონომეტრიკული მოდელის ჩამოყალიბება;
5. ეკონომეტრიკული მოდელის შეფასების პროცედურის შერჩევა;

6. მიღებული შედეგების ანალიზი ეკონომიკურ მოდელთან მიახლოებაში;

7. წინააღმდეგობრივი შედეგების მიღების შემთხვევაში ახალი მოდელის ძიება.

მოტივაცია ასეთია: მარტივი მოდელის განხილვიდან, რომლის მიხედვითაც შეიძლება წამოიჭრას რიგი კითხვებისა, სტუდენტმა შეძლოს მათზე გააზრებული პასუხის გაცემა და შედარებით რთულ მოდელზე გადასვლა.

ამ მიზნის მიღწევა შეიძლება, თუ განვიხილავთ ამოხსნილი მაგალითების გარჩევას, რომლებშიც გამოყენებულია ეკონომეტრიკის პროგრამული უზრუნველყოფა (STATA , EcoView, SPSS ან სხვა), რაც მნიშვნელოვან მომენტს წარმოადგენს მასალის ათვისებისათვის, რადგან სტუდენტს გამოუმავდება ეკონომეტრიკის პროგრამული უზრუნველყოფის გამოყენების პრაქტიკული უნარ-ჩვევები. ევროპის ქვეყნების უნივერსიტეტებისათვის STATA -ს გარდა, შედარებით პოპულარულია პროგრამული უზრუნველყოფები: SAS, SHASAM და SPSS პროგრამები, თუმცა შეიძლება გამოვიყენოთ ნებისმიერი პროგრამა ხელმისაწვდომი მოცემულ მომენტში.

ფაქტობრივად, "ეკონომეტრიკა პრაქტიკაში"-ს ჩარჩოებში ერთ მთავარ მიზანში გაერთიანებულია სამი მიზანი: ეკონომეტრიკული მოდელის აგება ეკონომიკური მოდელის ბაზაზე, მოდელის შეფასება მისი დახვეწის მიზნით და ეკონომეტრიკის პროგრამული უზრუნველყოფის გამოყენება მოდელის ამოახსნელად, მის შესაფასებლად და შემდეგ მის დასახვეწად.

წიგნში თითქმის ტრადიციული მიმდევრობით დალაგებულ მასალას აქვს შემდეგი მიზნობრივი დანიშნულება:

1. შესასწავლი მასალის სიმარტივე;
2. ეკონომეტრიკის შესწავლა მაგალითების განხილვით და სავარჯიშოები შესასწავლი მასალის მიხედვით;
3. სპეციფიკაციის შესასწავლი მასალის მნიშვნელოვნების ხედავ და მისი წარმოდგენა რეგრესიული ანალიზის მომხმარებლებისათვის როგორც ერთ-ერთი მნიშვნელოვანი პრობლემა;
4. სავარჯიშოების აწყობა რეგრესიულ ანალიზის გამოყენებაზე ისეთი მიმდევრობით, რომ მოხდეს ეკონომეტრიკული ანალიზის იმიტაცია უკუკავშირის გამოიყენებით სხვა და სხვა ამონახსნებში;
5. პროგრამული უზრუნველყოფის STATA - ს გამოყენება განხილული მასალის გაგების გასაადვილებლად.

წიგნის დამუშავებისას გამოყენებულია ის მასალაც, რომელიც ავტორი გაეცნო TEMPUS -ის პროგრამის ჩარჩოებში (Contract NCD_JEP-21058-2000-2002) ლისაბონის ეკონომიკისა და ბიზნეს ადმინისტრირების ინსტიტუტში და მოსკოვის ეკონომიკური სკოლაში ყოფნისას 1998წ. და 2004წ . გარდა ამისა, წიგნში გათვალისწინებულია ამ ინსტიტუტისა და ეკონომიკური სკოლის პროფესორ-მასწავლებლებთან შეხვედრის შედეგები ეკონომეტრიკის სასწავლო პროგრამის აგებასთან დაკავშირებით, რის გამოც კიდევ ერთხელ მადლობას ვუხდით პროფესორებს: ანტონიო მენდოსას, კარლოს ალბერტო დო სილვა რიბეირუს, ჯოზე მონტეირო ბარათას, ანდრადე სილვას, ლოაო ფერეირა დო ამარალს, პ. კატიშევს და ა. პერესეცკის სასარგებლო შეხვედრებისა და კონსულტაციებისათვის.

შესავალი

რა არის ეკონომეტრიკა?

ეკონომიკური სინამდვილე მრავალფეროვანი, ცვალებადი და წინააღმდეგობრივია, ამიტომ ეკონომიკური მეცნიერება არც აცხადებს პრეტენზიას რეალური მოვლენებისა და პროცესების ურთიერთკავშირის ადექვატურ ასახვაზე. ეკონომიკური მეცნიერება ჭეშმარიტებას აღწევს ერთგვარი მიახლოებით. მას აინტერესებს ფუნქციონალური კავშირები და მოვლენის განვითარების ტენდენცია. ამჟამად ეკონომიკური მეცნიერება ეძებს ახალ მიმართულებებს, რომელთა გამოყენებითაც შედარებით სრულად აისახება ცვლილებები საზოგადოებაში, ტექნოლოგიებში, წარმოების სტრუქტურასა და მოხმარებაში. ცხადია ისიც, რომ ცალკეული ქვეყნების ეკონომიკის განვითარებაზე სულ უფრო მეტად მოქმედებენ გლობალური პრობლემები და პროცესები და ბუნებრივია, ეკონომიკა არაეკონომიკური ფაქტორების მნიშვნელოვან გავლენას განიცდის. რის გამოც, ტრადიციული ეკონომეტრიკის გამოყენება ითვალისწინებს მონაცემების შეგროვებას, მოდელის აგებას და ამოხსნას, მოდელის შესაფასებელი მეთოდის შერჩევას, მოდელის დახვეწას, ფუნქციის პარამეტრების შეფასებას, ჰიპოთეზების შემოწმებას და პროგნოზს ეკონომეტრიკული პოლიტიკის შესახებ. მაგრამ ცნობილია, რომ ეკონომიკაში თითქმის არ არსებობს ეკონომიკური მოდელი, რომელიც ნებისმიერ პირობებში სამართლიანი იქნება. საუკეთესო შემთხვევასთან გვაქვს სამე, როდესაც მოდელი ლოკალურ პირობებშია სამართლიანი და მოდელი შეესაბამება ამოცანაში დასმულ მიზანს.

უცნაურია, მაგრამ პასუხი ამ კითხვაზე ” რა არის ეკონომეტრიკა ” საკმაოდ მარტივი იყო. კერძოდ, ეკონომეტრიკა არის ის რასაც აკეთებს ეკონომეტრისტი ანუ ეკონომეტრიკა სტატისტიკური მეთოდების გამოყენებაა ეკონომიკური ფენომენის ანალიზისათვის. ეს მოსაზრება ჯერ კიდევ 1953 წელს გამოთქვა ტინტნერმა / Tintner G, 1953/.

მართლაც, თუ გავითვალისწინებთ იმას, რომ ეკონომეტრისტები მრავალფეროვან სამუშაოს ასრულებენ, ამ შეკითხვაზე პასუხი შეიძლება ასეთიც იყოს / Paul A. Samuelson, T.C.Koopmans and J.R. Stone, 1954/:

”ეკონომეტრიკა მეცნიერებაა, რომელიც შეისწავლის ეკონომიკური კანონებისა და მათი ურთიერთქმედების ემპირიულ ანალიზს, აყალიბებს ეკონომიკურ მოდელებს, საფუძვლად იყენებს რა ეკონომიკურ თეორიას იძლევა მოდელის პარამეტრების შეფასებას, პროგნოზს და რეკომენდაციებს ეკონომიკური პოლიტიკის ნორმატივების შესახებ” /, 1998/

”ეკონომეტრიკა ეკონომიკური მოვლენებისა და პროცესების რაოდენობრივი ანალიზია” / Studemund A.H.,1996/

ეკონომეტრიკა შუალედური ხიდია, რომელიც აერთებს აბსტრაქტულ ეკონომიკურ სამყაროსა და კაცობრიობის შემოქმედების რეალურ სამყაროებს.

ეკონომეტრისტები მრავალფეროვან სამუშაოს ასრულებენ მიზნის მისაღწევად. ბუნებრივია, იბადება კითხვა იმის შესახებაც, თუ ვინ არიან ეკონომეტრისტები. უპირველეს ყოვლისა, ეკონომეტრისტები არიან ეკონომისტები, რომლებსაც შეუძლიათ ეკონომიკური თეორიის გამოყენება ეკონომიკური პრობლემის ემპირიული ანალიზის გასაუმჯობესებლად. იმავე დროს,

ეკონომეტრისტიკები მათემატიკოსები არიან, რომლებიც ეკონომიკური თეორიის ისეთ ფორმულირებას აყალიბებენ, რომ იგი სტატისტიკური ტესტირებისათვის ხელმისაწვდომი გახდეს. იმავე დროს, ეკონომეტრისტიკები აღმრიცხველები არიან, რომლებიც ეკონომიკურ ცვლადებზე დაკვირვებას აწარმოებენ და მათ მნიშვნელობებს აფიქსირებენ. იმავე დროს, ეკონომეტრისტიკები არიან გამოყენებითი სტატისტიკის სპეციალისტები: რომლებიც კომპიუტერებისა და პროგრამული უზრუნველყოფის გამოყენებით ეკონომიკურ კავშირებს აფასებენ. იმავე დროს, ეკონომეტრისტიკები თეორიული სტატისტიკის სპეციალისტები არიან, რომლებიც იყენებენ თავიანთ პროფესიონალიზმს სტატისტიკური ტექნიკის განვითარებისათვის. ეკონომეტრისტიკის ხელოვნებაა სავარაუდო მდგომარეობებიდან შეარჩიოს საკმაოდ სპეციფიური, საკმაოდ რეალისტური, მოცემულ განხილულ ეკონომიკურ სიტუაციას მორგებული გადაწყვეტილება.

ეკონომეტრისტი იყენებს რეგრესიულ ანალიზს იმ ეკონომიკური დამოკიდებულებების შესაფასებლად, რომლებსაც უპირველეს ყოვლისა, თეორიული საფუძველი გააჩნია.

რეგრესიული ანალიზი სტატისტიკური ტექნიკის ის საშუალებაა, რომლის გამოყენებითაც ის ცდილობს ახსნას ერთი ცვლადის ცვლილება, რომელიც დამოკიდებულ ცვლადს წარმოადგენს როგორც ფუნქცია სხვა ცვლადებისა, რომლებსაც დამოუკიდებელი ცვლადები ეწოდებათ და დამოკიდებულ და დამოუკიდებელ ცვლადებს შორის დამოკიდებულება ჩაწეროს განტოლებით.

მაგალითად, განტოლება

$$C = f(P, P_s, P_d),$$

სადაც: C - დამოკიდებული ცვლადია, ხოლო P, P_s, P_d - დამოუკიდებელი ცვლადებია, რეგრესიული ანალიზის შესწავლის ობიექტია.

ეკონომისტსა და ეკონომეტრისტს შორის განმასხვავებელი ნიშანი მოგვიანებით ჩამოყალიბდა. ეს განსხვავება კარგად ჩანს შემდეგი მაგალითის მიხედვით: ეკონომისტისათვის მოხმარება შემოსავლის ფუნქციაა და ამ მაჩვენებლებს შორის დამოკიდებულება მათემატიკურად ჩაიწერება შემდეგი სახით:

$$C = f(I),$$

სადაც, C - მოხმარებაა; I - შემოსავალია.

ეკონომეტრისტი განაცხადებს, რომ ეს დამოკიდებულება აუცილებლად შეიცავს მაწონასწორებელ შემთხვევით წევრს და განტოლებას ჩაწერს შემდეგი სახით:

$$C = f(I) + e$$

სადაც, e - მაწონასწორებელი წევრი სტოქასტურ ხასიათს ატარებს. მეორეს მხრივ, როდესაც სამომხმარებლო ფუნქცია ჩაიწერება სახით:

$$c = S_1 + S_2 y + e$$

სადაც β პარამეტრები ახასიათებენ სამომხმარებლო ფუნქციას, ეკონომისტი ინტერესდება შეისწავლოს ამ უცნობი პარამეტრების მნიშვნელობები, მაგრამ შემთხვევითი წევრის არსებობა ართულებს ამ საკითხის მოგვარებას. სწორედ ამ პარამეტრების შეფასებისკენაა მიმართული ეკონომეტრისტიკის ერთ-ერთი ძირითადი საქმიანობა.

ეკონომიკურ მაჩვენებლებს შორის დამოკიდებულების შეფასების აუცილებლობა შეიძლება დავინახოთ შემდეგი მაგალითებიდან:

1. ანალიტიკოსები როგორც სამთავრობო, ისე კერძო სექტორიდან დაინტერესებულები არიან სხვადასხვა პროდუქტისა და მომსახურების მოთხოვნა-მიწოდების შეფასებით;
2. კერძო ფირმები და კომპანიები დაინტერესებულები არიან გაყიდვებისა და მოგების შეფასებით მათი პროდუქციის რეკლამირების ეფექტების მიხედვით;
3. მთავრობას სურს შეაფასოს მისი ფისკალური პოლიტიკის გავლენა ისეთ მაჩვენებლებზე, როგორცაა: დასაქმება, დაუსაქმებლობა, შემოსავალი, იმპორტი და ექსპორტი, ინფლაცია, საბიუჯეტო დეფიციტი და სხვა;
4. მუნიციპალიტეტები დაინტერესებულები არიან კომპანიების განთავსებით რეგიონში. კონკრეტულ შემთხვევაში მათი ინტერესი შეიძლება დაკავშირებული იყოს შინამეურნეობების მოთხოვნებთან, დასაქმების დონესთან, ელექტროენერჯის მოხმარებასთან, უკუგებასთან, წყალმომარაგებასთან და სხვა;
5. ფინანსური ბაზრების ანალიტიკოსები ეძებენ აქციების ფასების დამოკიდებულებებს კომპანიების ეკონომიკურ მდგომარეობასთან მიმართებაში.

ეკონომეტრიკის კავშირი სხვა დისციპლინებთან

ეკონომეტრიკა და ეკონომიკური ანალიზი

ჩვენ 21-ე საუკუნეში ვცხოვრობთ და როგორც ბევრი, მეცნიერების სხვა დარგი, ეკონომიკაც საინფორმაციო ტექნოლოგიებისა და პროგრამული უზრუნველყოფების აქტიური მომხმარებელი გახდა. ეკონომიკური თეორია, რომელიც ეკონომიკური ინფორმაციის ძირითად ბაზას წარმოადგენს, მნიშვნელოვანი ეკონომიკური ცვლადების იდენტიფიცირების საშუალებას იძლევა და განსაზღვრავს საბაზრო ეკონომიკის პირობებში ეკონომიკურ ცვლადებს შორის ურთიერთდამოკიდებულების ბუნებას. თუ ინფორმაცია საკმარისი აღმოჩნდება, მაშინ ეკონომიკურ ცვლადებს შორის დამოკიდებულების შესწავლისას სტუდენტს უნდა შეეძლოს უპასუხოთ შესასწავლი ამოცანის ირგვლივ დასმულ კითხვებს და შესაძლებელია, მაგალითად ისეთ კითხვებსაც, როგორცაა:

1. რა გავლენა ექნება ბიუჯეტზე ინფლაციისა და საპროცენტო განაკვეთის დონის ცვლილებას?
2. არის თუ არა კავშირი დოუ-ჯონსის ინდექსსა და საპროცენტო განაკვეთების დონეებს შორის?
3. როგორ აისახება სავაჭრო დეფიციტი დასაქმებაზე?
4. რა დამოკიდებულებაა M1 ფულის ოდენობასა და ეკონომიკურ საქმიანობის დონეს შორის?
5. როგორ აისახება ეკონომიკური უზრუნველყოფის დონე კრიმინალურ სიტუაციებზე ქალაქებში?
6. როგორ აისახება საგადასახადო პოლიტიკა ინვესტიციების დონეზე და სხვა.

საიდან და როგორ მოდის ამ ტიპის ინფორმაცია?

უნდა შევნიშნოთ, რომ ეკონომიკური თეორია ასეთი ინფორმაციის ერთ-ერთი ძირითადი წყაროა. იგი არა მარტო გვეხმარება მოვახდინოთ მნიშვნელოვანი ეკონომიკური ცვლადების იდენტიფიცირება, არამედ წარმოადგენს იმ ეკონომიკური მონაცემების წაროს, რომლებსაც ეკონომიკურ მაჩვენებლებზე დაკვირვების შედეგად ვიღებთ, როგორცაა: ფასები, მოხმარება, შემოსავლები, მიწოდების დონეები და სხვა. მაგრამ რეალურად, ეკონომეტრიკის მიზანი უნდა მოიცავდეს პასუხს კითხვაზე: როგორ შევისწავლოთ რეალური ეკონომიკური გარემო ეკონომიკური მონაცემების მნიშვნელობების მიმდევრობის მიხედვით?

თუ ამ კითხვას ეკონომისტს დავუსვამთ აუცილებლად მივიღებთ პასუხს რომელიც იწყება ან მთავრდება ასე: "ეს ყველაფერი დამოკიდებულია"

ასეთი პასუხი მოსალოდნელიცაა, რადგან ბევრი რამ მართლაც დამოკიდებულია სხვა და სხვა ეკონომიკურ ცვლადებზე და იმ უცნობ კავშირებზე, რომლებიც მათ შორისაა. ამ მოსაზრების ჩამოყალიბების მიზნით შეგვიძლია გამოვიყენოთ ამ დამოკიდებულებების წარმოდგენა მათემატიკური სახით. მაგალითად, თუ თქვენ ფიქრობთ, რომ არსებობს დამოკიდებულება შემოსავალსა და მოხმარებას შორის შეგვიძლია ჩავწეროთ ეს დამოკიდებულება შემდეგი სახით:

$$C = f(I) \quad (1)$$

სადაც:

C – მოხმარებაა, I – შემოსავალია და მოხმარება შეგვიძლია განვიხილოთ როგორც შემოსავლის ფუნქცია.

მიღებული ჩანაწერი ეკონომიკური მოდელია და იგი ასახავს ეკონომიკურ ცვლადებს შორის დამოკიდებულებას. ამ ტიპის ეკონომიკური მოდელები ეკონომიკური ანალიზის საფუძველია, მაგრამ ამ სახით მათი წარმოდგენა ტესტირებისა ან პროგნოზისათვის საკმარის საფუძველს არ იძლევა. ამიტომ მოდელი უნდა შეიცავდეს ე. წ. პარამეტრებს. უცნობ ეკონომიკურ პარამეტრებს, რომელთა შეფასებაც შესაძლებელია იმ ეკონომიკური მონაცემების მიხედვით, რომლებიც თან ახლავს ეკონომიკურ პოლიტიკას. ეკონომიკურ მოდელში ამ პარამეტრების შემოტანა და შემდეგ მათი შეფასება. ეკონომიკური მოდელის აგების შემდეგ ეკონომეტრიკული ანალიზის ერთ-ერთი მნიშვნელოვანი ეტაპია.

ეკონომიკური თეორია არ იძლევა ერთეული ეკონომიკური ობიექტის დახასიათების სპეციფიკას, რადგან იგი ახასიათებს საერთოდ ეკონომიკურ სისტემას და თუ ამოცანა ინდივიდუალურ ეკონომიკურ ობიექტს ითვალისწინებს, აუცილებელია მოდელში გათვალისწინებული იყოს ის შემთხვევითი კომპონენტი, რომელიც ამ ობიექტს თან სდევს და რომელსაც შემთხვევით შეცდომას ვუწოდებთ. მოდელი, რომელიც ამ შემთხვევით სიდიდეს გაითვალისწინებს ჩაიწერება შემდეგი სახით:

$$c = f(i) + e, \quad (2)$$

სადაც,

e - შემთხვევითი სიდიდეა, რომელიც თავის თავში აერთიანებს იმ მრავალ ფაქტორს, რაც განხილულ მოდელზე მოქმედებს. მოდელი (2) ეკონომიკური მოდელი აღარ არის. იგი სტატისტიკური მოდელია, რომლის სპეციფიკაციის მიზნით აუცილებელია ვიცოდეთ ეკონომიკურ ცვლადებს შორის დამოკიდებულების სახე. მაგალითად, თუ აგრეგირებული მოხმარება - $f(i)$ გამოსახულია როგორც წრფივი ალგებრული ფუნქცია აგრეგირებული შემოსავლისა, მაშინ მოხმარების სისტემატური ნაწილი შეიძლება წარმოვადგინოთ ასე:

$$f(i) = S_1 + S_2 i$$

რომლის შესაბამისი სტატისტიკური მოდელია

$$c = S_1 + S_2 i + e$$

და თქვენი მიზანია განვსაზღვროთ ის ფორმა, რომელიც თავსებადია ეკონომიკურ თეორიასა და მონაცემებს შორის.

ყველა ეკონომიკური ფუნქცია, როგორცაა მოთხოვნის ფუნქცია, მიწოდების ფუნქცია, საწარმოო ფუნქცია და სხვა ასევე შეიცავს სისტემატურ კომპონენტსა და შემთხვევითი შეცდომის კომპონენტს, რომელიც დაკვირვებას არ ექვემდებარება.

ეკონომიკური თეორია აღწერს სისტემატურ კომპონენტს. შემთხვევითი კომპონენტის შესახებ კი ვიცით, რომ ის არსებობს, მაგრამ არაფერი ვიცით მის შესახებ რის გამოც, მას ეკონომიკური მოდელი სტატისტიკურ მოდელში გადაყავს, რაც სტატისტიკური ლოგიკური დასკვნის გაკეთების საფუძველს იძლევა. ეს შედეგი უცნობი პარამეტრების შეფასებისა და მათ შესახებ ჰიპოთეზების ტესტირების აუცილებლობას საფუძველსაც გვაძლევს.

ეკონომეტრისტი მიზანია განსაზღვროს ცვლადებს შორის და ასევე, ცვლადების მოქმედებასა და შედეგს შორის დამოკიდებულებები, რისთვისაც გამოიყენება შერჩევა T დაკვირვებათა რიცხოვნებით, ე.ი. ეკონომეტრისტს სურს მონაცემების მიხედვით შეაფასოს β -ს ეფექტი c -ს, გამოსავალზე ე.ი. ეკონომეტრისტი კომბინირებას უკეთებს ეკონომიკურ თეორიას, სტატისტიკას, პროგრამულ უზრუნველყოფას, გამოთვლით ტექნიკას და სწორედ ეს კომბინაციაა ეკონომეტრიკის ქმედების შედეგი.

ეკონომიკური მონაცემების უმეტესობა არ არის კონტროლირებადი ექსპერიმენტის შედეგი, რაც ბუნებრივიცაა და ეკონომეტრისტი ასტრონომს ემსგავსება და როდესაც საქმე გვაქვს არაკონტროლირებად ექსპერიმენტთან, ეკონომისტი იღებს თავის თავზე დამკვირვებლის როლს და ეკონომიკური თეორია მიუთითებს იმ აქტიურ ცვლადებზე, რომელთა განხილვა აუცილებელია.

ამგვარად, გვაქვს:

1. ეკონომიკური მოდელი, რომელიც გვებმარება აქტიური ცვლადებისა და პარამეტრების იდენტიფიკაციაში, საფუძველს გვაძლევს ეკონომიკური დასკვნების გასაკეთებლად;

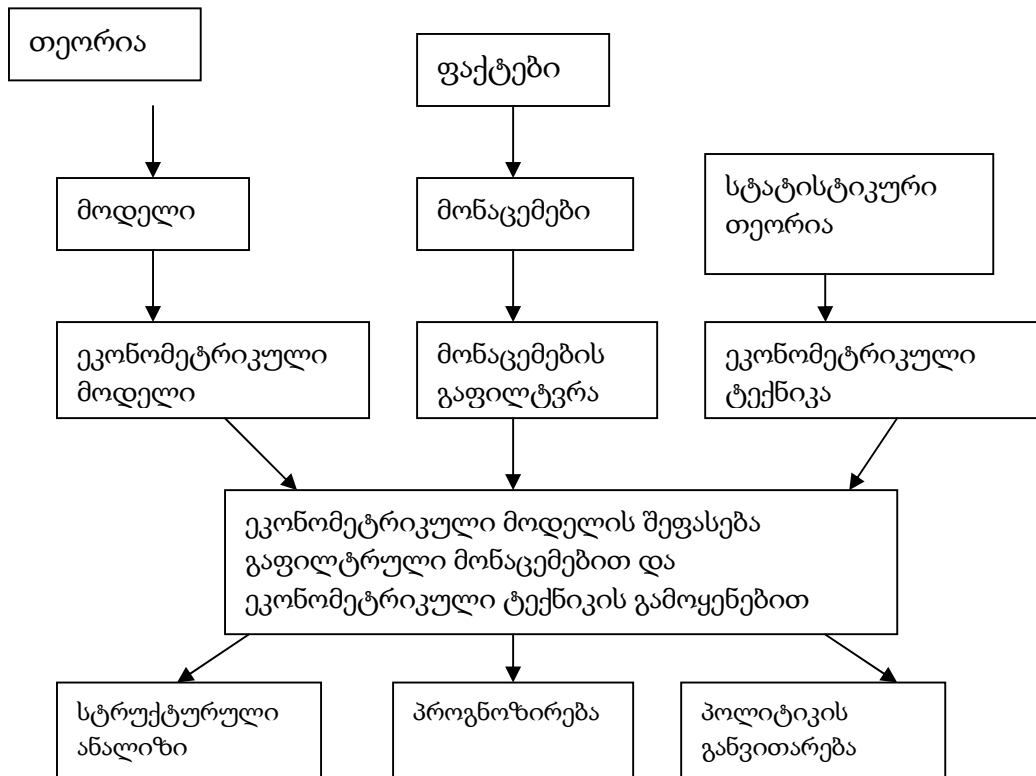
2. სტატისტიკური მოდელი, რომელიც განსაზღვრავს სპეციფიკაციის პროცესს, რომლის გამოყენებითაც ხდება შერჩევის მონაცემების გენერირება და რომელიც უცნობი პარამეტრების იდენტიფიცირებას ახდენს.

3. ეკონომიკურ ცვლადებზე დაკვირვების მნიშვნელობები.

ამ შედეგების საფუძველზე ეკონომეტრისტმა უნდა განსაზღვროს, თუ როგორ გამოიყენოს ეს ინფორმაცია, რომ შეისწავლოს უცნობი ეკონომიკური პარამეტრები და გააუმჯობესოს გადაწყვეტილების მიღება ეკონომიკური ცვლადების შერჩევის მიზნით.

ეკონომიკური და სტატისტიკური მოდელების განხილვისას ყურადღება გადაგვაქვს ეკონომიკურ ცვლადებს შორის კავშირის ძიებაზე და გვსურს მივიღოთ ინფორმაცია უცნობი და დაკვირვებას დაუქვემდებარებელი პარამეტრების შესახებ.

სქემატურად ეკონომეტრიკული მეთოდის ბუნება შეიძლება ასე წარმოვადგინოთ.



ნახაზი 1. ეკონომეტრიკული მეთოდის გამოყენების ბუნება

ნახაზის მიხედვით ეკონომეტრიკის სამი პრინციპული მიზანი აშკარაა: სტრუქტურული ანალიზი, პროგნოზირება, პოლიტიკის განვითარება.

სტრუქტურული ანალიზი ეკონომიკური დამოკიდებულებების რაოდენობრივი წარმოდგენაა ანუ ეკონომიკური დამოკიდებულებების რაოდენობრიობის გასაზომის საშუალებაა ეკონომეტრიკული მოდელის გამოყენებით. სტრუქტურული ანალიზი საშუალებას იძლევა პასუხი გაეცეს კითხვას, თუ რა უნდა განვიხილოთ ეკონომეტრიკული კვლევის მიზნის ქვეშ, რომ გასაგები გახდეს ეკონომიკური დამოკიდებულებების რეალური არსი, რაც შემდეგში შეიძლება საფუძვლად დაედოს ეკონომიკური თეორიის განვითარებას.

პროგნოზირება შეფასებული ეკონომეტრიკული მოდელით გათვლის-წინებული ცვლადების რაოდენობრივი შეფასების საშუალებაა მათზე დაკვირვების რეალიზაციის გარეთ და შეიძლება ქმედების საფუძველი გახდეს. მაგალითად, საქონლის შესყიდვებისა და მისი რეალიზაციის შემთხვევაში ფირმამ შეიძლება გაითვალისწინოს დამატებითი სამუშაო ადგილები, თუ მომდევნო პერიოდში პროგნოზირებული იქნება გაყიდვების გაზრდა.

პოლიტიკის განვითარება შეფასებული ეკონომეტრიკული მოდელის მიხედვით გამოიყენება ალტერნატიული პოლიტიკის არსებობის შემთხვევაში რომელიმე მიმართულების შესაფასებლად. გადაწყვეტილების მისაღებად

სასარგებლოა პირობითი პროგნოზირების გაკეთება ცვლადის მნიშვნელობის შესახებ მომავალში თითოეული ალტერნატივის მიხედვით.

მათემატიკური ეკონომიკა და ეკონომეტრიკა

მათემატიკური ეკონომიკა ეკონომიკური პროცესების მათემატიკური მოდელების აგებას, ამოხსნასა და ანალიზს ემსახურება. ზოგიერთ შემთხვევაში ეს მოდელები შეიძლება განვიხილოთ როგორც მათემატიკური თეორიის შემადგენელი ნაწილი ეკონომიკური მეცნიერების საზღვარზე. მათემატიკური ეკონომიკა გამიჯნულია ეკონომეტრიკისაგან, რადგან ეკონომეტრიკის ძირითადი მიზანია ეკონომიკური მოვლენებისა და პროცესების ანალიზი და შეფასება ემპირიული მონაცემების საფუძველზე, ხოლო მათემატიკური ეკონომიკა იკვლევს თეორიულ მოდელებს, რომლებიც დაფუძნებულია განსაზღვრულ პირობებზე, როგორცაა: ამოზნექილობა, წრფივობა, მონოტონურობა და სხვა. მისი ამოცანაა მოდელის ამონახსნის არსებობის შესწავლა. იგი ეკონომიკური მეცნიერების დამოუკიდებელი ნაწილია და მისი ერთ-ერთი ინსტრუმენტია. მათემატიკური ეკონომიკის მოდელებს შორის შეიძლება გამოვყოთ ორი, დიდი არსებითი კლასი. ეკონომიკურ სისტემებში წონასწორობის მოდელები და ეკონომიკური ზრდის მოდელები. მათი გამოყენება ეკონომიკური მოვლენის დინამიკაში აღწერის შესაძლებლობას იძლევა.

მათემატიკური ეკონომიკისაგან გასხვავებით ეკონომეტრიკა ეკონომიკაში იკვლევს რაოდენობრივ კანონზომიერებებსა და ურთიერთ დამოკიდებულებებს მათემატიკური სტატისტიკის გამოყენებით. კვლევის ძირითად მეთოდად ეკონომეტრიკა იყენებს კორელაციულ-რეგრესიულ ანალიზს.

ეკონომეტრიკაში პირველი შრომები XIX საუკუნის ბოლოს და XX საუკუნის დასაწყისში გამოჩნდა. კერძოდ, 1897წ. გამოჩნდა პირველი შრომა, რომელიც ეკუთვნოდა ეკონომიკურ თეორიაში მათემატიკური სკოლის ფუძემდებელს ვ. პარეტოს. ეს შრომა მიემდგვნა სხვადასხვა ქვეყნებში მოსახლეობის შემოსავლების სტატისტიკურ შესწავლას. შემოთავაზებული იყო ე.წ. პარეტოს მრუდი

$$y = A(x - a)^{-r}$$

სადაც:

x - შემოსავალია;

y - იმ პირების რიცხოვნებაა, რომელთა შემოსავალი x -ზე მეტი იყო;

a - მინიმალური შემოსავალია;

A, r - პარამეტრებია, რომლებიც კვლევის შედეგად მიიღება.

XX ს. დასაწყისში გამოვიდა ინგლისელი სტატისტიკოსის ჰუკერის რამდენიმე ნაშრომი, რომლებშიც მან გამოიყენა კორელაციურ-რეგრესიული ანალიზის მეთოდები, რომლებიც დაამუშავა პირსონმა და მისმა სკოლამ ეკონომიკური მაჩვენებლების ურთიერთდამოკიდებულებების შესასწავლად. კერძოდ, მარცვლეულის სასაქონლო ბირჟაზე გაკოტრებათა რიცხოვნების გავლენა მარცვლეულის ფასზე. ჰუკერის შრომები შეიცავდა ლაგური ცვლადების იდეას და ეკონომიკური მაჩვენებლების ნაზრდების კორელაციურ ანალიზს და არა თვითონ მაჩვენებლების კორელაციურ ანალიზს. მათემატიკური სტატისტიკისა და მისი გამოყენებითი ელემენტების განვითარებასთან ერთად საგრძნობლად გაიზარდა

ეკონომეტრიკაში ახალი შრომების ნაკადი. მათ მიეკუთვნება რ. ფიშერის შრომები დისპერსიულ ანალიზში და ეკონომიკური მაჩვენებლების შეფასების მცდელობა როგორც კობისა და დუგლასის კლასიკური შრომა 1928 წ.

მოდელები, ეკონომიკური და ეკონომეტრიკული მოდელები

განსაზღვრების მიხედვით მოდელი ნებისმიერი სისტემის ან პროცესის რეპრეზენტაციაა, რომელიც გათვლილია ამოხსნის, პროგნოზირების და კონტროლის მიზნით. მოდელირება მოდელის აგების ხელოვნებაა და მეცნიერების განუყოფელი ნაწილია, ვინაიდან ნებისმიერი, ბუნებაში არსებული პროცესი ან სისტემა მეტისმეტად კომპლექსურია და მათი წარმოდგენა მოდელების სახით მათ შესწავლას აადვილებს. მოდელირება იწყება იმით, რომ პროცესი ან სისტემა განიხილება როგორც შავი ყუთი თავისი შესავლით და გამოსავლით და შემდეგ მუშავდება ვარიანტები “ რა არის ამ ყუთის შიგნით“.

მოდელის სხვადასხვა ტიპებს შორის ალგებრული მოდელები ეკონომეტრიკისათვის განსაკუთრებული მოდელებია, რომლებიც შეიძლება განტოლებით ან განტოლებათა სისტემის სახით იყოს წარმოდგენილი. ეროვნული შემოსავლების ალგებრული მაკროეკონომიკური მოდელი წარმოდგენილია შემდეგი სახით:

$$C = C(Y) \\ Y = C + Z$$

სადაც ცვლადების აღნიშვნებია:

C – მოხმარება;

Y – ეროვნული შემოსავალი;

Z (ეგზოგენური ცვლადი) - დანახარჯები.

ამ მარტივი მოდელის Z - ეგზოგენური ცვლადში გაერთიანებულია ადგილობრივი ინვესტიციები, სამთავრობო დანახარჯები და წმინდა ექსპორტი. მოცემული განტოლებათა სისტემა განსაზღვრავს წონასწორობის მნიშვნელობას მოხმარებასა და ეროვნულ შემოსავალს შორის. ამასთან, პირველი განტოლება სამომხმარებლო ფუნქციაა, ხოლო მეორე – წონასწორობის პირობა, რომლის მიხედვითაც დგინდება რომ საერთო შემოსავალი სამომხმარებლო და ეგზოგენური ცვლადით გამოწვეული დანახარჯების ჯამია.

მათემატიკური მოდელი წარმოადგენს მათემატიკურ თანაფარდობათა სისტემას, რომელიც რეალური ობიექტის არსებით თვისებებს ასახავს და საკვლევი ამოცანის პარამეტრებს შორის ურთიერთკავშირს ადგენს. მათემატიკური მოდელების აგების პროცესს მათემატიკური მოდელირება ეწოდება. ეკონომიკურ-მათემატიკური მოდელების გამოყენებისას ხშირად ვიზუალურად ამოცანის უფრო მნიშვნელოვანი დეტალებით და არ ვითვალისწინებთ ზოგიერთ მნიშვნელოვან დეტალს, რომელიც გამოწვეულია ეკონომიკური ამოცანების სირთულით. ამიტომ, ეკონომიკური ამოცანის მოდელი ისე უნდა აიგოს, რომ მაქსიმალურად გავითვალისწინოთ მოვლენების არსებითი მხარეები: მაგრამ მოდელი ისე არ უნდა გავართულოთ, რომ შემდეგ მისი ამოხსნა ვერ შევძლოთ.

შემთხვევითი ფაქტორების გათვალისწინებით მათემატიკური მოდელები პირობითად ორ კლასად შეიძლება დაყვით: ალბათური და დეტერმინირებული. ალბათური მოდელები რეალური ეკონომიკური პროცესის შემთხვევით ხასიათს ითვალისწინებენ. ალბათურ მოდელებს მიეკუთვნება რეგრესიის, პროგნოზირების,

იმიტაციური, გადაწყვეტილებათა მიღების მოდელები და სხვა. დეტერმინირებული მოდელებია: მათემატიკური პროგრამირების, ქსელური და სხვა.

ალბათური მოდელები დეტერმინირებულ მოდელებთან შედარებით საშუალებას გვაძლევს გავითვალისწინოთ უფრო მეტი ფაქტორი, რაც მოდელის გამარტივების ერთ-ერთი წინა პირობაა, თუმცა მოდელირების შედეგები ძნელად ექვემდებარებიან ანალიზს.

ეკონომიკური პროცესების ანალიზი გვიჩვენებს, რომ მათი კვლევისას აუცილებელია ერთდროულად გამოვიყენოთ როგორც ალბათური, ისე დეტერმინირებული მოდელები: რომლებიც ფაქტობრივად ერთმანეთს ავსებენ. მაგალითად, ალბათური მოდელი შეიძლება გამოვიყენოთ ეკონომიკური ანალიზის პირველ ეტაპზე იმ ძირითადი ფაქტორების გამოვლენის მიზნით, რომლებიც მოდელშია შესაყვანი და ამ ფაქტორებს შორის ურთიერთკავშირის გამოსავლენად. ამ მიზანს ძირითადად ემსახურებიან რეგრესიის მოდელები. შემდეგში, ამოცანა შეიძლება ამოვხსნათ დეტერმინირებული მოდელის გამოყენებით.

ნებისმიერ შემთხვევაში, მათემატიკური მოდელის აგებისას აუცილებელია ამოცანის ეკონომიკურ არსს კარგად ვიცნობდეთ.

პროცესის მათემატიკური მოდელის აგება შეიძლება შესრულდეს რამდენიმე ეტაპად:

1. განისაზღვრება ამოცანის შინაარსი და მოდელირების მიზანი. წინასწარ შეფასდება მოსალოდნელი შედეგები და ყალიბდება ის პირობები, რომელთა შესრულებითაც მიზანს ვაღწევთ, გამოიყოფა მნიშვნელოვანი ფაქტორები და დგინდება ის კანონზომიერებები, რომლებსაც ეს ფაქტორები ემორჩილებიან. ამ ეტაპზე ვითვალისწინებთ სტატისტიკური მონაცემების მოპოვებასაც;

2. შეირჩევა მოდელის მიზნის ფუნქცია. ექსპერიმენტით, თუ კი ეს შესაძლებელია, შემოწმდება მოდელის შეზღუდვების პირობები და განისაზღვრება მოდელის სტრუქტურა;

3. იწყება მოდელის ამოხსნა. ე.ი. მიზნის ფუნქციის ექსტრემუმის ძიება. ალბათური მოდელებისათვის ამ ეტაპზე ტარდება სტატისტიკური ანალიზი. ე.ი. იმ უცნობი ფაქტორების სტატისტიკური შეფასება, რომლებიც მოდელში შედიან. კერძოდ, მიღებული შეფასებებისა და მათი სიზუსტეების კვლევა;

4. შეფასდება რეალური პროცესის და მოდელის ამოხსნით მიღებული შედეგები. დადგინდება მოდელის ადექვატურობის ხარისხი გამოყენებული ინფორმაციის სიზუსტის გათვალისწინებით. აქ ორ შემთხვევას შეიძლება ჰქონდეს ადგილი:

ა/ მოდელი დამაკმაყოფილებელ შედეგებს იძლევა და იგი შეიძლება მივიღოთ;

ბ/ შედეგები არადაკმაყოფილებელია და გადავდივართ მათემატიკური მოდელის აგების მეორე ციკლზე მოვლენის (ობიექტის) შესახებ საწყისი ინფორმაცია ზუსტდება. კიდევ ერთხელ შემოწმდება ამოცანის არსი. კვლავ აიგება მათემატიკური მოდელი და ამოიხსნება. შედეგები კვლავ შედარდება რეალურს და ა.შ.

5. დამუშავდება სათანადო რეკომენდაციები მოდელის გამოყენების შესახებ, სადაც მიეთითება მოდელის გამოყენების შესაძლებლობები, რადგან არც ერთი მოდელი უნივერსალური არ არის.

მოდელების ამოსახსნელად მათემატიკური მეთოდების ფართო სპექტრი გამოიყენება. ეს მეთოდები ამა თუ იმ გაგებით უნივერსალურებად ითვლებიან, რადგან სხვა და სხვა ტიპის ამოცანების ამოსახსნელად გამოიყენება, მაგრამ უნდა

შევიშნოთ, რომ ერთი და იგივე ამოცანა შეიძლება სხვა და სხვა მეთოდებით ამოვხსნათ. კონკრეტული, განსაზღვრული მათემატიკური მეთოდის შერჩევა განისაზღვრება საკვლევი ამოცანის სპეციფიკით.

ალბათური მეთოდების ფართო სპექტრიდან პრაქტიკაში ფართოდაა გავრცელებული მათემატიკური სტატისტიკის მეთოდები კერძოდ, რეგრესიული ანალიზი.

რეგრესიული ანალიზის მეთოდები გამოიყენება ეკონომიკურ სისტემაში საკვლევ ფაქტორებს შორის ურთიერთკავშირის დასადგენად: ამოცანის არსის დასახვეწად და მართვისა და დაგეგმვის ამოცანების ამოსახსნელად და გარდა ამისა, დიდი მნიშვნელობა ენიჭება პროგნოზირების მათემატიკური მეთოდების გამოყენებას ეკონომიკური მაჩვენებლების დასადგენად, რომელთა შორის მნიშვნელოვანია მრავლფაქტორიანი მაპროგნოზირებელი მეთოდები.

მაპროგნოზირებელი მოდელების აგებისას ყურადღებას იმსახურებს მოვლენის დეტერმინირებული (ტრენდის სახით) და შემთხვევითი მდგენელების გამოყოფა.

შემთხვევითი მდგენელების აღწერას ემსახურება შემთხვევითი ფუნქციების თეორიის მეთოდები. ტრენდისა და შემთხვევითი მდგენელების გამოყოფა საშუალებას იძლევა შედარებით სრულად ავლწეროთ ეკონომიკური მოვლენების (პროცესების) ცვლილებების კანონზომიერებები.

გამოთვლითი ტექნიკისა და საინფორმაციო ტექნოლოგიების განვითარებასთან ერთად მათემატიკური მეთოდებისა და მოდელების გამოყენების მნიშვნელობა საგრძნობლად გაიზარდა. ამჟამად, ეკონომეტრიკული მოდელები და მეთოდები წარმოადგენენ მძლავრ იარაღს ეკონომიკაში არა მარტო ეკონომიკური მოვლენებისა და პროცესების უკეთ შეცნობისათვის, არამედ პრაქტიკული გადაწყვეტილებების მისაღებადაც.

მათემატიკური მოდელები ფართოდ გამოიყენება ბიზნესსა, ეკონომიკასა, საზოგადოებრივ მეცნიერებებსა, პოლიტიკური პროცესების კვლევაში და სხვა. შეიძლება გამოვყოთ მოდელების სამი ძირითადი კლასი, რომელიც ანალიზის ან პროგნოზისათვის გამოიყენება:

დროითი მწკრივების მოდელები

მათ მიეკუთვნება:

ტრენდის მოდელი

$$Y(t) = T(t) + v_t$$

სადაც, $T(t)$ -პარამეტრული სახით მოცემული დროითი ტრენდია.

წრფივი ტრენდის შემთხვევაში

$$T(t) = a + bt$$

სადაც, v_t -შემთხვევითი, ალბათური კომპონენტი.

სეზონურობის მოდელი:

$$Y(t) = S(t) + v_t$$

სადაც, S_t -პერიოდული, სეზონური კომპონენტი; v_t -შემთხვევითი, ალბათური კომპონენტი.

ტრენდისა და სეზონურობის მოდელები

$$Y(t) = T(t) + S(t) + v_t \quad (\text{ადიტიური მოდელი})$$

$$Y(t) = T(t)S(t) + v_t \quad (\text{მულტიპლიკატიური მოდელი})$$

სადაც, $T(t)$ - პარამეტრული სახით მოცემული დროითი ტრენდია;

$S(t)$ - პერიოდული, სეზონური კომპონენტი;

v_t - შემთხვევითი, ალბათური კომპონენტი.

რეგრესიის მოდელები ერთი განტოლებით

ამ მოდელებში დამოკიდებული ცვლადების წარმოდგენა შეიძლება შემდეგი ფუნქციის სახით :

$$f(x, S) = f(x_1, x_2, \dots, x_n, S_1, S_2, \dots, S_k)$$

სადაც, x_1, x_2, \dots, x_n - დამოუკიდებელი ცვლადებია;

S_1, S_2, \dots, S_n - პარამეტრებია;

$f(x, S)$ ფუნქციის მიხედვით რეგრესიის მოდელები ორ ჯგუფად იყოფიან: წრფივი და არაწრფივი. რეგრესიის მოდელების გამოყენების სფერო (წრფივი მოდელებისა და კი) უფრო ფართოა, ვიდრე დროითი მწკრივებისა.

ერთდროული განტოლებების სისტემით აღწერილი მოდელები

ამ მოდელებში შემავალი სისტემა შეიძლება შედგებოდეს იგივეობებისა და რეგრესიის განტოლებებისაგან, რომელთაგან თითოეული დამოუკიდებელი ცვლადის გარდა შეიძლება შეიცავდეს ცვლადებს სისტემის სხვა განტოლებებიდან, ე.ი. გვაქვს დამოუკიდებელი ცვლადების ერთობლიობა, რომლებიც დაკავშირებული არიან განტოლებათა სისტემით.

ასეთი მოდელის მაგალითია მოთხოვნა-მიწოდების მოდელი:

$$Q_t^s = a_1 + a_2 p_t + a_3 p_{t+1} + v_t \quad (\text{მიწოდება})$$

$$Q_t^p = b_1 + b_2 p_t + b_3 Y_t + U_t \quad (\text{მოთხოვნა})$$

$$Q_t^s = Q_t^p \quad (\text{წონასწორობის პირობა})$$

სადაც, Q_t^p - მოთხოვნა საქონელზე;

Q_t^s - საქონლის მიწოდება;

p_t - საქონელზე ფასი მოცემულ t მომენტში;

Y_t - შემოსავალი t მომენტში.

$Q_t^s = Q_t^p$ - საქონელზე მოთხოვნა - მიწოდების წონასწორობის პირობა მოიცემა მოდელის განტოლებიდან.

თავი 1. რეგრესიული ანალიზის მიმოხილვა

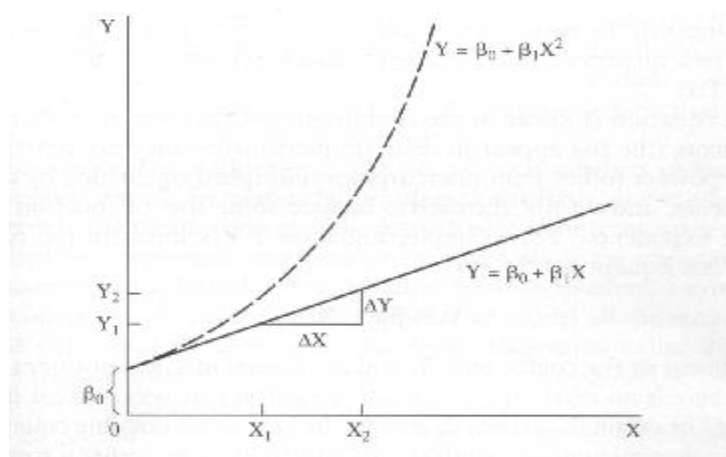
1.1. წრფივი რეგრესია

განვიხილოთ უმარტივესი წრფივი მოდელი:

$$Y = S_0 + S_1 X \quad (1)$$

სადაც Y – დამოკიდებული ცვლადია და დამოუკიდებელი ცვლადის X -ის წრფივი ფუნქციაა.

S_0, S_1 - კოეფიციენტებია ან პარამეტრები, რომლებიც განსაზღვრავენ შესაბამისი წრფის წერტილების კოორდინატებს. S_0 - მუდმივი სიდიდეა და გამოსახავს Y -ის მნიშვნელობას: როდესაც $X = 0$; S_1 - წრფის დახრის კოეფიციენტია X ღერძის მიმართ. წრფის გრაფიკული წარმოდგენა მოცემულია ნახაზზე.



ნახ 1.1. წრფის გრაფიკული წარმოდგენა

ამავე ნახაზზე წარმოდგენილია განტოლება $Y = S_0 + S_1 X^2$, რომელიც არწრფივ დამოკიდებულებას გამოსახავს: როდესაც $S_1 > 0$. წრფის დახრის კოეფიციენტი წრფივი დამოკიდებულებებისათვის მუდმივია და ადგილი აქვს ტოლობას:

$$S_1 = \frac{Y_2 - Y_1}{X_2 - X_1} = \frac{\Delta Y}{\Delta X}$$

განხილული მოდელი (1) ეკონომიკური მოდელია, მაგრამ რაოდენობრივი ეკონომიკურ ანალიზისათვის ვიყენებთ ეკონომეტრიკულ მოდელს, რომლისთვისაც ეკონომიკური მოდელი საბაზო მოდელია, როდესაც განვიხილავთ ეკონომიკურ

მოდელს. ბუნებრივია, ჩვენ ვერ ვითვალისწინებთ იმ დამატებით ფაქტორებს, რომლებიც შესასწავლ ეკონომიკურ მოვლენაზე მოქმედებენ, რაც გამოწვეულია არა ჩვენი "დაუდევრობით", არამედ იმით, რომ ფიზიკურად ამ ფაქტორების კონტროლი არ შეგვიძლია. ეკონომომეტრისტი უშვებს რა გაუთვალისწინებელი (აუხსნელი) ცვლილების არსებობას, შემოაქვს სტოქასტიური (შემთხვევითი) წევრი, რომელიც პასუხისმგებელია იმ შეცდომაზე, რომელიც არსებობს ეკონომიკურ მოდელსა და ფაქტობრივად რეალურ მოდელს შორის. სტოქასტიური წევრი ის წევრია, რომელიც ემატება რგრესიის განტოლებას. ეს წევრი ხშირად გამოიყენება როგორც მოდელის ცდომილება ან შემთხვევითი ცდომილება ან მაწონასწორებელი წევრი ან შემთხვევითი კომპონენტი ან შემთხვევითი წევრი. გამოვიყენოთ ამ წევრის შინაარსისათვის ტერმინი შემთხვევითი ცდომილება და ავღნიშნოთ იგი ε -ით.

თუ ამ შემთხვევით ცდომილებას დავუმატებთ განტოლება (1)-ს, შედეგში მივიყვებთ ტიპიურ რეგრესიის განტოლებას:

$$Y = S_0 + S_1X + v \quad (2)$$

რომელიც შეიცავს ორ კომპონენტს: დეტერმინირებულს და სტოქასტურს ან შემთხვევითს. მათგან $S_0 + S_1X$ გამოსახულება რეგრესიის განტოლების დეტერმინირებული კომპონენტია, რომლითაც განსაზღვრულია Y - ის დამოკიდებულება X -ზე. შემთხვევითი კომპონენტი აუცილებლად უნდა იყოს წარმოდგენილი რეგრესიის განტოლებაში თუნდაც იმიტომ, რომ არსებობს სულ ცოტა სამი წყარო მაინც, რომელთა გავლენითაც მოსალოდნელია Y - ის ცვლილება:

1. Y -ზე მოქმედი მრავალი "უმნიშვნელო" ფაქტორი არ არის გათვალისწინებული რეგრესიის განტოლებაში, რადგან მათ შესახებ მონაცემები არ გაგვაჩნია;

2. ფაქტობრივად შეუძლებელია გამოვრიცხოთ განტოლების ცვლადების გაზომვის შეცდომები;

3. თეორიულ განტოლებას სხვა და სხვა ფუნქციონალური ფორმა აქვს მაშინ, როდესაც რეგრესიის განტოლებისათვის ჩვენ ერთ-ერთ მათგანს ვირჩევთ. მართლაც, თუ დაკვირვებათა რიცხვი n -ია ($i = 1, 2, \dots, n$), მაშინ რეალურად ჩვენ საქმე გვაქვს n განტოლებასთან, რომელთაგან თითოეული მათგანი შეესაბამება თითოეულ დაკვირვებას:

$$\begin{aligned} Y_1 &= S_0 + S_1X_1 + v_1 \\ Y_2 &= S_0 + S_1X_2 + v_2 \\ &\dots\dots\dots \\ Y_n &= S_0 + S_1X_n + v_n \end{aligned}$$

რეგრესიის მოდელი უნდა პასუხობდეს ითოეულ დაკვირვებას. ამასთან, კოეფიციენტების მნიშვნელობები დაკვირვებიდან დაკვირვებამდე უცვლელნი რჩებიან მაშინ, როდესაც X , Y და ε -ის მნიშვნელობები ცვლილებას განიცდიან. მეორეს მხრივ, თუ ჩვენ შემოვიტანთ შემდეგ აღნიშვნებს:

X_{1i} - პირველი დამოუკიდებელი ცვლადის i -ური დაკვირვებაა,

X_{2i} - მეორე დამოუკიდებელი ცვლადის i -ური დაკვირვებაა,

X_{3i} - მესამე დამოუკიდებელი ცვლადის i -ური დაკვირვებაა, მაშინ შესაბამისი რეგრესიის მოდელი უნდა ჩაიწეროს შემდეგი სახით:

$$Y_i = S_0 + S_1X_{1i} + S_2X_{2i} + S_3X_{3i} + V_i$$

სადაც:

S_1 - ასახავს X_1 დამოუკიდებელი ცვლადის გავლენას დამოკიდებულ Y ცვლადზე, როდესაც X_{2i}, X_{3i} -ს მუდმივი მნიშვნელობები აქვთ;

S_2 - ასახავს X_2 დამოუკიდებელი ცვლადის გავლენას დამოკიდებულ Y ცვლადზე, როდესაც X_{1i}, X_{3i} -ს მუდმივი მნიშვნელობები აქვთ.

ეს მულტივარიაციული რეგრესიის კოეფიციენტები იცავენ Y -ს სხვა დამოუკიდებელი ცვლადების გავლენისაგან, როდესაც რომელიმე მათგანი ცვლილებას განიცდის. ეს მომენტი აუცილებელია გავითვალისწინოთ, როდესაც ჩვენ კონტროლირებად ექსპერიმენტს ვატარებთ. სინამდვილეში ეკონომიკური ფაქტორები მოქმედებენ ერთდროულად და ხშირად, ერთმანეთის საწინააღმდეგო მიმართულებებითაც.

ზოგადად, მულტივარიაციული რეგრესიის მოდელი K დამოუკიდებელი ცვლადით ჩაიწერება შემდეგი სახით:

$$Y_i = S_0 + S_1X_{1i} + S_2X_{2i} + \dots + S_kX_{ki} + V_i, (i=1,2, \dots n)$$

1.2. რეგრესიის სავარაუდო განტოლება

თეორიული რეგრესიის განტოლების რაოდენობრივად შეფასებული ვერსია წარმოადგენს სავარაუდო რეგრესიის განტოლებას, თუ თეორიული რეგრესიის განტოლება ჩაწერილია ფორმით:

$$Y_i = S_0 + S_1X_i + V_i$$

ხოლო რეგრესიის სავარაუდო განტოლებას ექნება შემდეგი სახე:

$$\hat{Y}_i = \hat{S}_0 + \hat{S}_1X_i$$

სადაც,

\hat{Y}_i - სავარაუდო მნიშვნელობაა, რომელიც Y_i -ზე დაკვირვებების საფუძველზე გამოთვლის შედეგადაა მიღებულია.

დამოკიდებული ცვლადისათვის მის ფაქტობრივ მნიშვნელობასა და სავარაუდო მნიშვნელობებს შორის სხვაობა e_i განისაზღვრება როგორც გადახრა ფაქტობრივ მნიშვნელობიდან და ჩაიწერება:

$$e_i = Y_i - \hat{Y}_i$$

უნდა შევნიშნოთ, რომ შემთხვევით წევრს თეორიული შინაარსი აქვს და მასზე დაკვირვება შეუძლებელია მაშინ, როდესაც გადახრა რეალური სიდიდეა და მისი გამოთვლა შესაძლებელია თითოეული დაკვირვებისათვის. მართლაც, გადახრებზე შესაძლებელია არა მარტო დაკვირვება, არამედ შეგვიძლია გამოვყოთ \hat{S}_0 -ისა და \hat{S}_1 -ს ის მნიშვნელობები, რომლებისთვისაც გადახრების მნიშვნელობები დაბალია. ეს მომენტი საინტერესოა იმით, რომ რაც მცირეა გადახრის მნიშვნელობა, მით უფრო ახლოსაა \hat{Y} ისა და Y - მნიშვნელობები ერთმანეთთან?

მულტივარიაციული რეგრესიის სავარაუდო განტოლება ჩაიწერება შემდეგი სახით:

$$\hat{Y}_i = \hat{S}_0 + \hat{S}_1 X_{1i} + \hat{S}_2 X_{2i} + \dots + \hat{S}_{ki} X_{ki}$$

1.3. რეგრესიული ანალიზის უმარტივესი მაგალითები

1.1. ვთქვათ, თქვენ დასასვენებელ პარკში აგიყვანეს სამუშაოდ. თქვენი მოვალეობაა პარკში შემომსვლელთა (მამაკაცების) წონის გამოცნობა მათი სიმაღლის მიხედვით. იმ შემთხვევაში, თუ თქვენ მათ წონას გამოიცნობთ 10 კგ-ის სიზუსტით მიიღებთ 50 თეთრს, წინააღმდეგ შემთხვევაში - გადაიხდით 60 თეთრს. შემომსვლელის ზუსტი წონა ფიქსირდება ისე, რომ თქვენ ვერ ხედავთ.

პირველი სამუშაო დღე არა სასიამოვნო იყო თქვენთვის. თქვენ ორი ლარი წააგეთ.

მეორე დღეს გადაწყვიტეთ შეაგროვოთ მონაცემები, რომ დაადგინოთ დამოკიდებულება ადამიანის წონასა და მის სიმაღლეს შორის, გამოიყენოთ რეგრესიული ანალიზი. თქვენთვის თეორიულად ეს დამოკიდებულებაა:

$$Y_i = f(X_i) = S_0 + S_1 X_i + V_i$$

სადაც: Y_i - პარკში შემომსვლელის წონა;

X_i - პარკში შემომსვლელის სიმაღლეა;

V_i - შემთხვევითი წევრია თითოეული i - ური შემომსვლელისათვის.

პარკში 20 შემომსვლელის მიხედვით მათ წონასა და სიმაღლეს შორის მონაცემები მოცემულია ცხრილით:

N	X_i	Y_i	\hat{Y}_i	e_i	დანაკარგი, ლარი
1	5.0	140.0	135.3	4.7	+50
2	9.0	157.0	150.8	-3.8	+50
3	13	205.0	186.3	18.7	-60
4	12.0	198.0	179.9	18.1	-60
5	10.0	162.0	167.2	-5.2	+50
6	11.0	174.0	173.6	0.4	+50
7	8.0	150.0	54.4	-4.4	+50
8	9.0	165.0	160.8	4.2	+50
9	10.0	170.0	167.2	2.8	+50
10	12.0	180.0	179.9	0.1	+50
11	11.0	170.0	173.6	-3.6	+50
12	9.0	162.0	160.8	1.3	+50
13	10.0	165.0	167.2	-2.2	+50
14	12.0	180.0	179.9	0.1	+50
15	8.0	160.0	154.4	5.6	+50
16	9.0	155.0	160.8	-5.8	+50
17	10.0	165.0	167.2	-2.2	+50
18	15.0	190.0	199.1	-9.1	+50
19	13.0	185.0	186.3	-1.3	+50
20	11.0	155.0	173.6	-18.6	-60
ჯამი					6.70 ლარი

ცხრილის მონაცემების გამოყენებით მიღებული რეგრესიის განტოლებაა:

$$\hat{Y}_i = 103.40 + 6.38X_i$$

როგორ იმუშავებს მიღებული რეგრესიის განტოლება?

ამ კითხვაზე პასუხის გასაცემად გამოითვლება სხვაობები $e_i = Y_i - \hat{Y}_i$ ამ განტოლების მიხედვით. ცხრილის ბოლო სვეტში მოცემულია ის მოგება 6.70 ლარი, რასაც თქვენ მიიღებთ ნაცვლად 2.0 ლარის დანაკარგისა.

ამ მაგალითის მიზანი იყო თეორიული რეგრესიის განტოლების წარმოდგენა, თუმცა შეიძლება ჭეშმარიტი განტოლება მიუწვდომელი აღმოჩნდეს ისევე, როგორც სტოქასტური წევრი V . მთავარი, რაც აუცილებელი იყო ამ მაგალითით გვეჩვენებინა, ეს არის რეგრესიის განტოლების ჩასაწერად გამოთვლითი პროცედურის გამოყენება.

1.2. ვთქვათ, გინდათ შეიძინოთ სახლი და მონახეთ შესაფერისი სახლი თქვენთვის სასურველ რაიონში. სახლის მეკატრონემ მოგთხოვათ 230000 ლარი სახლისათვის, რომლის ფართობი 1600კვ.მ-ია, რამაც თქვენში ეჭვი გამოიწვია, რომ ეს ფასი არ შეესაბამებოდა ამ სახლის რეალურ ფასს. როგორ მოიქცევით? გადაიხდით მოთხოვნილ თანხას, თუ შეეცდებით დააგინოთ სახლის რეალური ფასი?

მოდით, მივმართოთ რეგრესიული ანალიზის გამოყენებას რეალური ფასის დასადგენად და თქვენთვის სასურველი სახლის შესაძენად.

თეორიულად სახლის ფასი სახლის ფართობის მიხედვით შეიძლება გამოვსახოთ რეგრესიის განტოლებით:

$$P_i = f(S_i) = S_0 + S_1S_i + V_i$$

სადაც: P_i - სახლის ფასია (ლარი);

S_i - სახლის ფართობია (კვ.მ);

V_i - შემთხვევითი წევრია.

ინტერაქტიული გამოკითხვის შედეგად მიღებულია 43 მონაცემი თქვენთვის სასურველ რაიონში სახლების ფასებისა და მათი ფართობების შესახებ მოცემულია ცხრილით:

N	P	S	N	P	S	N	P	S
1	107	736	16	300	1643	31	215	1216
2	133	720	17	310	1675	32	228	1447
3	141	768	18	315	1714	33	242	1974
4	165	929	19	350	2150	34	250	1600
5	170	1080	20	365	2206	35	250	1168
6	173	942	21	503	3269	36	255	1478
7	182	1000	22	135	936	37	255	1756
8	200	1472	23	147	728	38	265	1542
9	220	1200	24	165	1014	39	265	1633
10	226	1302	25	175	1661	40	275	1500
11	260	2109	26	190	1248	41	285	1735
12	275	1528	27	191	1834	42	365	1900
13	280	1081	28	195	989	43	397	2468
14	289	1753	29	205	1232			

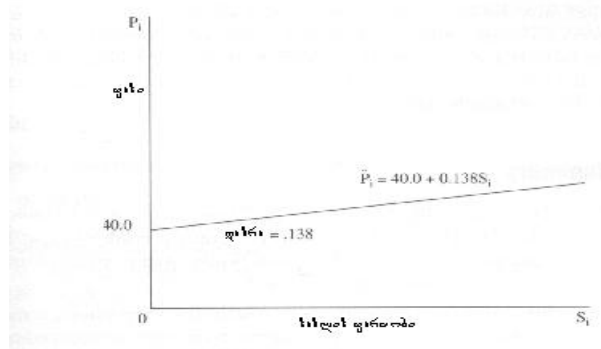
15	295	1528	30	210	1017			
----	-----	------	----	-----	------	--	--	--

ცხრილის მონაცემების მიხედვით მიღებულია რეგრესიის განტოლება:

$$\hat{P}_i = 40.0 + 0.138S_i$$

რას გამოსახავენ კოეფიციენტები რეგრესიის განტოლებაში?

კოეფიციენტი $S_1 = 0.138$ მიუთითებს შემდეგზე: თუ სახლის ფართობის მნიშვნელობა გაიზრდება 1 კვ. მ-ით, მაშინ სახლის გასასყიდი ფასი გაიზრდება 138 ლარით. კოეფიციენტი $S_0 = 40.0$ მუდმივი სიდიდეა და მიუთითებს შემდეგზე: სახლის ფასი 40 ლარის ტოლია, თუ მისი ფართობი ნულის ტოლია. მისი არსი უფრო ღრმად განხილული იქნება შემდეგში. გრაფიკულად ეს დამოკიდებულება წარმოდგენილია ნახ.1.1-ით.



ნახ. 1.1. რეგრესიის მოდელი: $\hat{P}_i = 40.0 + 0.138S_i$

როგორ გამოვიყენებთ მიღებულ რეგრესიის განტოლებას? მიღებულ განტოლებაში შევიტანოთ შესაბამისი მონაცემი შესასყიდი სახლის ფართობის შესახებ. მივიღებთ:

$$\hat{P}_i = 40.0 + 0.138 * 1600 = 40.0 + 220.8 = 260.8 \text{ ათასი ლარი}$$

სახლის მეპატრონემ მოგთხოვათ 230000 დოლარი და როგორც ჩანს, თქვენთვის მისაღები უნდა იყოს სახლის ეს ფასი.

რეზიუმე

1. რეგრესიული ანალიზი განსაზღვრავს, რომ დამოკიდებული ცვლადი ერთი ან მეტი დამოუკიდებელი ცვლადების ფუნქციაა;
2. რეგრესიული ანალიზი შეიძლება გამოვიყენოთ განტოლებისათვის, რომელიც წრფივია ამ განტოლების კოეფიციენტების მიმართ და ამასთან, რეგრესიის კოეფიციენტები წარმოდგენილია ყველაზე მარტივი ფორმით;
3. სტოქსატური წევრი აუცილებლად უნდა დაემატოს რეგრესიის განტოლებას, რომ გავითვალისწინოთ დამოკიდებული ცვლადის შემთხვევითი ცვლილებები. განტოლების შემთხვევითი კომპონენტი მოიცავს:

- ცვლადებს, რომლებიც არ მონაწილეობენ რეგრესიის განტოლებაში;
- მონაცემების შეგროვებისას დაშვებულ შეცდომებს;
- თეორიულ განტოლების შეცდომას, რომელსაც სხვა ფუნქციონალური ფორმა აქვს ვიდრე რეგრესიის განტოლებას;
- ნებისმიერი შემთხვევით და გაუთვალისწინებელ სიდიდეს.

4. სავარაუდო რეგრესიის განტოლება აპროქსიმაციაა იმ ჭეშმარიტი განტოლებისა, რომელიც მიღებულია Y -ის ფაქტობრივი მნიშვნელობების მიხედვით მოცემული შერჩევიდან. ჩვენ არაფერი ვიცით ჭეშმარიტი განტოლების შესახებ და ეკონომეტრიკული ანალიზი მთელ ყურადღებას მიმართავს სავარაუდო რეგრესიის განტოლებისაკენ და იძლევა ამ განტოლების კოეფიციენტების შეფასებას.

დამოუკიდებელი სამუშაო

1.1. განსაზღვრეთ შემდეგი ტერმინები:

- ეკონომიკური მოდელი;
- სტატისტიკური მოდელი;
- ეკონომეტრიკული მოდელი;
- რეგრესიის სავარაუდო განტოლება.

1.2. იქნება თუ არა ქვემოთ ჩამოთვლილ დამოკიდებულ და დამოუკიდებელ ცვლადებს შორის კავშირი პოზიტიური, ნეგატიური ან ცალმხრივი:

- აგრეგატული წმინდა ინვესტიცია და მიმდინარე წელს ერთობლივი ეროვნული პროდუქტი;
- პროფესორის ასაკი და მის თავზე თმის რაოდენობა;
- პლანტაცია ჰექტარებში და მარცვლეულის ფასი სეზონის დასაწყისში;
- წმინდა ინვესტიცია და საპროცენტო განაკვეთი იმავე წელს.

1.3. შესწავლილია დამოკიდებულება სატრანსპორტო თვითმფრინავების ექსპლუატაციის ღირებულებასა და ექსპლუატაციის ხანგრძლივობას შორის. ცხრილში მოცემულია სათანადო ინფორმაცია:

ექსპლუატაციის ხანგრძლივობა, წელი	6 თვიანი ღირებულება, \$	ექსპლუატაციის ხანგრძლივობა, წელი	6 თვიანი ღირებულება, \$	ექსპლუატაციის ხანგრძლივობა, წელი	6 თვიანი ღირებულება, \$
4.5	619	5.0	890	6.0	764
4.5	1049	5.0	1522	6.0	1373
4.5	1033	5.5	987	1.0	978
4.0	495	5.0	1194	1.0	466
4.0	723	0.5	163	1.0	549
4.0	681	0.5	182		

- შეიძლება თუ არა ვივარაუდოთ მაჩვენებლებს შორის წრფივი დამოკიდებულება?
- შეიძლება თუ არა რეგრესიის მოდელის აგება?

1.4. აშშ-ს ფედერალური კომისიის ერთ-ერთ მოსმენაზე ენერგეტიკის შესახებ წარმოდგენილი იყო სტატისტიკური მონაცემები ცხრილის სახით:

X	Y	X	Y
13.3	3.5	26.3	6.0
16.9	5.1	30.1	9.5
19.9	4.8	42.6	8.1
23.2	6.7		

სადაც:

X - მოპოვებული გაზის შემადგენლობაში სითხის პროცენტული შემცველობა;

Y - ერთეული კუბ.მ გაზის ღირებულება (ცენტი).
ცხრილის მონაცემების მიხედვით შეარჩიეთ წრფივი მოდელი

1.5. გაზეთში დაბეჭდილი იყო შეტყობინება, რომ საბითუმო ბაზარზე სამარკო პორტვინის თითოეულ ბოთლზე ფასების ცვლილება შემდეგ ხასიათს ატარებდა:

წელი	მანეთი	წელი	მანეთი	წელი	მანეთი	წელი	მანეთი
1890	50.00	1934	15.00	1941	10.00	1952	4.99
1900	35.00	1935	13.00	1944	5.90	1965	5.98
1920	25.00	1940	6.98	1948	8.98	1985	4.98
1931	11.98			1950	6.98		

ააგეთ მონაცემების მიხედვით გრაფიკი და გამოიკვლიეთ აქვს თუ არა აზრი რეგრესიის მოდელის აგებას.

1.6. საცხოვრებელი ბინის გადასახადსა და პირად შემოსავლებს შორის დამოკიდებულება შეიძლება ჩაიწეროს განტოლებით:

$$\hat{y} = -27.6 + 0.178x$$

საცხოვრებელი ბინის გადასახადი დროსთან დამოკიდებულებაში შეიძლება ჩაიწეროს განტოლებით:

$$\hat{y} = 48.9 + 4.8t$$

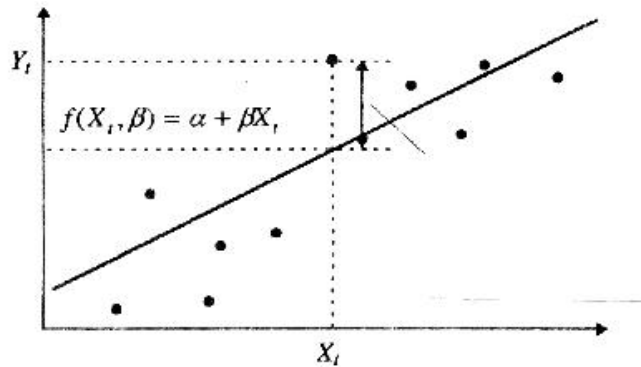
როგორია რეგრესიის ამ განტოლებების ეკონომიკური შინაარსი

თავი 2. უმცირეს კვადრატთა მეთოდი

2.1. რეგრესიის წრფის შერჩევა

ვთქვათ, $(X_n, Y_n), i = \overline{1, n}$ ცვლადებისა და მათი შესაბამისი წყვილები მნიშვნელობები მოცემულია სიბრტყეზე ისე, როგორც ეს ნახ. 2.1-ზეა წარმოდგენილი:

შევარჩიოთ ფუნქცია $f(X, S)$ ფუნქციითა პარამეტრული ოჯახიდან ისეთი, რომელიც Y -ის X -ზე დამოკიდებულების "საუკვეცსო" ვარიანტს შეესაბამება. ეს ნიშნავს: შევარჩიოთ S პარამეტრის "საუკეთესო" მნიშვნელობა. $f(x, S)$ ფუნქციის გადახრის საზომად (x, y) დაკვირვების წყვილების მნიშვნელობიდან შეიძლება შემოვიტანოთ:



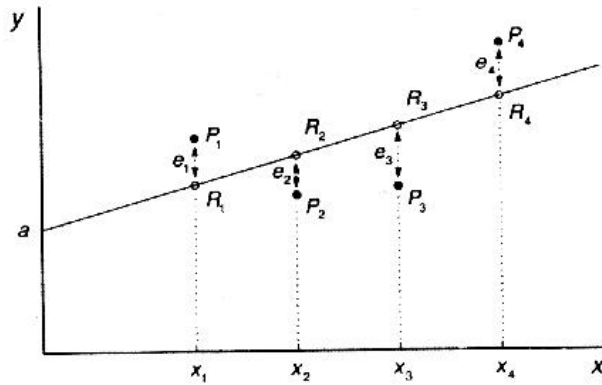
ნახ. 2.1. რეგრესიის წრფის შერჩევა

1. სხვაობების კვადრატების ჯამები: $S = \sum_{i=1}^n (Y_i - f(X_i, S))^2$;
2. სხვაობების მოდულების ჯამები: $S = \sum_{i=1}^n (|Y_i - f(X_i, S)|)$.

2.1. უმცირეს კვადრატთა მეთოდი

აუცილებელია შევნიშნოთ, რომ რეგრესიის მოდელის S_0, S_1, \dots პარამეტრების ჭეშმარიტი მნიშვნელობების დადგენა თითქმის შეუძლებელია და ბუნებრივია, პასუხის მოძებნა კითხვაზე, თუ როგორ განვსაზღვროთ რეგრესიის წრფის მდებარეობა კვლევის საგანს წარმოადგენს. სამიუბელი პარამეტრებისათვის შეგვიძლია მივიღოთ შეფასებები კარგი ან ცუდი, შეიძლება ზუსტიც (მხოლოდ შემთხვევითი დამთხვევის შედეგი). ეს მოსაზრებები მაშინაც სამართლიანია, როდესაც პარამეტრების შეფასების ყველაზე სრულყოფილი მეთოდებია გამოყენებული. უფრო მეტიც, ეს თითქმის შეუძლებელია, თუ ცვლადი დამოკიდებულია არა ერთ, არამედ ორ და მეტ დამოუკიდებელ ცვლადზე.

არსებობს თუ არა ამ პარამეტრების შეფასების საკმარისად ზუსტი მეთოდები? პირველი ნაბიჯი ამ მიმართულებით თითოეული დაკვირვებისათვის ჭეშმარიტ და შეფასებულ მნიშვნელობებს შორის სხვაობის განსაზღვრაა? ბუნებრივია, რეგრესიის წრფე ისე უნდა აიგოს, რომ ამ სხვაობების მნიშვნელობები მინიმალურს შეესაბამებოდნენ. აუცილებელია, შერჩევის ისეთი კრიტერიუმი გამოვიყენოთ, რომელიც სხვაობის ყველა მნიშვნელობას ერთდროულად გაითვალისწინებს. ამ მიზნის მისაღწევად არსებობს კრიტერიუმების ფართო სპექტრი, რომელთა გამოყენება შეიძლება. აღნიშნული პრობლემის გადაწყვეტის ერთ-ერთი გზა სხვაობების კვადრატების ჯამის S -ის მინიმიზაციაა მდგომარეობს. ნახ. 2.2-ზე წარმოდგენილია რეგრესიის წრფის აგებისას სხვაობების მდებარეობის შესაძლო ვარიანტები.



ნახ. 2.2. სხვაობების მნიშვნელობები რეგრესიის წრფის აგებისას

ნახ. 2.2. - ის მიხედვით ჭეშმარიტია ტოლობა:

$$S = e_1^2 + e_2^2 + e_3^2 + e_4^2$$

S – ის სიდიდე დამოკიდებულია რეგრესიის განტოლების პარამეტრების შერჩევაზე, რადგან ისინი განსაზღვრავენ რეგრესიის წრფის მდებარეობას. რაც უფრო ნაკლებია S – ის მნიშვნელობა, მით უფრო უკეთესადაა წარმოდგენილი X, Y ცვლადებს შორის დამოკიდებულება. თუ $S = 0$ ანუ როდესაც სხვაობების კვადრატების ჯამი ნულის ტოლია, მაშინ მიღებული გვაქვს აბსოლუტურად ზუსტი შესაბამისობა. ეს კი ნიშნავს, რომ ამ შემთხვევაში რეგრესიის წრფე გადის ყველა წერტილზე. თუმცა საერთოდ, შემთხვევითი წევრის არსებობის პირობებში ეს თითქმის შეუძლებელია. განვიხილოთ ჩვეულებრივი უმცირეს კვადრატთა მეთოდი (უკმ) (OLS – ordinary least squares).

”ჭეშმარიტი” წრფიდან გადახრების კვადრატების ჯამი ჩაიწერება შემდეგი სახით:

$$S = \sum_{i=1}^n e^2 = \sum_{i=1}^n (Y_i - s_0 - s_1 X_i)^2 \quad (2.1)$$

მოვძებნოთ \hat{S}_0 და \hat{S}_1 – ის ისეთი მნიშვნელობები, რომ (2.1) განტოლებაში \hat{S}_0 და \hat{S}_1 – ის ჩანაცვლებით შესრულდეს პირობა: $S \rightarrow 0$. ? (შენიშვნა: X_i და Y_i მნიშვნელობები მოცემულია)

\hat{S}_0 და \hat{S}_1 –ის მოძებნის პროცედურა შემდეგია:

1. მოვძებნოთ (2.1) ფუნქციის კერძო დიფერენციალი \hat{S}_0 და \hat{S}_1 პარამეტრებით;

$$\frac{\partial S}{\partial \hat{S}_0} = -2 \sum_{i=1}^n (Y_i - \hat{S}_0 - \hat{S}_1 X_i)$$

$$\frac{\partial S}{\partial \hat{S}_1} = -2 \sum_{i=1}^n (Y_i - \hat{S}_0 - \hat{S}_1 X_i) X_i$$

2. ამოვხსნათ განტოლებათა სისტემა \hat{S}_0 და \hat{S}_1 პარამეტრების მიმართ:

$$\begin{cases} \sum_1^n (Y_i - \hat{s}_0 - \hat{s}_1 X_i) = 0 \\ \sum_1^n X_i (Y_i - \hat{s}_0 - \hat{s}_1 X_i) = 0 \end{cases} \Rightarrow \begin{cases} \hat{s}_0 n + \hat{s}_1 \sum_1^n X_i = \sum_1^n Y_i \\ \hat{s}_0 \sum_1^n X_i + \hat{s}_1 \sum_1^n X_i^2 = \sum_1^n X_i Y_i \end{cases} \quad (2.2)$$

საიდანაც, $\hat{s}_1 = \frac{\sum_1^n X_i Y_i - \left[\left(\sum_1^n X_i \right) \left(\sum_1^n Y_i \right) \right] / n}{\sum_1^n X_i^2 - \left(\sum_1^n X_i \right)^2 / n} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$

$$\bar{X} = (X_1 + X_2 + \dots + X_n) / n = \sum_1^n X_i / n$$

რადგან,

$$\bar{Y} = (Y_1 + Y_2 + \dots + Y_n) / n = \sum_1^n Y_i / n$$

გვექნება:

$$\sum (X_i - \bar{X})(Y_i - \bar{Y}) = \sum X_i Y_i - \bar{X} \sum Y_i - \bar{Y} \sum X_i + n \bar{X} \bar{Y} = \sum X_i Y_i - n \bar{X} \bar{Y} = \sum X_i Y_i - \left(\sum X_i \right) \left(\sum Y_i \right) / n$$

შემოვიტანოთ აღნიშვნები:

$$S_{xx} = \sum X_i^2 - n \bar{X}^2$$

$$S_{yy} = \sum Y_i^2 - n \bar{Y}^2$$

$$S_{xy} = \sum X_i Y_i - n \bar{X} \bar{Y}$$

\hat{s}_1 და \hat{s}_0 პარამეტრების გამოსათვლელად მივიღებთ გამოსახულებებს:

$$\begin{cases} \hat{s}_1 = S_{xy} / S_{xx} \\ \hat{s}_0 = \bar{Y} - \hat{s}_1 \bar{X} \end{cases} \quad (2.3)$$

მიყვებული (\hat{s}_0, \hat{s}_1) პარამეტრებით ჩაიწერება შესაფასებელი რეგრესიის განტოლება:

$$\hat{Y}_i = \bar{Y} - \hat{s}_1 (X_i - \bar{X}) \quad (2.4)$$

მაგალითი 1.

ვთქვათ X და Y ცვლადებს შორის დამოკიდებულება მოცემულია დაკვირვების შედეგების წვილებით:

$$X = 1,5; Y = 4; X = 3, Y = 6; X = 4, Y = 7;$$

ჩავეწეროთ რეგრესიის განტოლება:

$$\hat{Y} = \hat{s}_0 + \hat{s}_1 x.$$

გამოთვალოთ რეგრესიის განტოლების პარამეტრები:

$$\begin{aligned} S &= e_1^2 + e_2^2 + e_3^2 = (4 - \hat{s}_0 - 1.5 \hat{s}_1)^2 + (6 - \hat{s}_0 - 3 \hat{s}_1)^2 + (7 - \hat{s}_0 - 4 \hat{s}_1)^2 - \hat{s}_0 \hat{s}_1 = \\ &= 101 + 3 \hat{s}_0^2 + 27.25 \hat{s}_1^2 - 34 \hat{s}_0 - 104 \hat{s}_1 + 17 \hat{s}_0 \hat{s}_1 \end{aligned}$$

შევარჩიოთ \hat{S}_0 და \hat{S}_1 პარამეტრები ისე რომ S - ის მნიშვნელობა მინიმალური იყოს. ამოვხსნათ ნორმალურ განტოლებათა სისტემა:

$$\begin{cases} \frac{\partial S}{\partial \hat{S}_0} = 0 \\ \frac{\partial S}{\partial \hat{S}_1} = 0 \end{cases} \quad \begin{cases} \sum Y_i = \hat{S}_0 n + \hat{S}_1 \sum X_i \\ \sum X_i Y_i = \hat{S}_0 \sum X_i + \hat{S}_1 \sum X_i^2 \end{cases}$$

$$\begin{cases} 17 = 3\hat{S}_0 + 8.5\hat{S}_1 \\ 52 = 8.5\hat{S}_0 + 27.25\hat{S}_1 \end{cases} \quad \begin{cases} \hat{S}_0 = 2.238 \\ \hat{S}_1 = 1.21 \end{cases}$$

ე. ი. რეგრესიის განტოლება შემდეგი სახისაა:

$$\hat{Y} = 2.238 + 1.21x$$

რეგრესიული ანალიზის მიზანია დამოკიდებული ცვლადის ცვლილების შესწავლა ნებისმიერ მოცემულ შერჩევაში. ცალკეულ შემთხვევაში, y -ის მნიშვნელობა შეიძლება დაბალი იყოს, ცალკეულში - მაღალი. რატომ შეიძლება ასე ხდებოდეს? რეგრესიულ ანალიზში ვცდილობთ ავხსნათ Y -ის ცვლილების მიზეზი დამოუკიდებელი ცვლადის X - ის მიხედვით. რეგრესიის განტოლების განსაზღვრის შემდეგ Y -ის შესაბამისი e_i -ს თითოეული მნიშვნელობა თითოეულ დაკვირვებაში შეიძლება წარმოვადგინოთ ორი მდგენელის სახით. კერძოდ,

$$y_i = \hat{y}_i + e_i$$

სადაც:

$\hat{y}_i - i$ -ურ დაკვირვებაში y -ის გამოთვლილი მნიშვნელობაა, რომელიც შეიძლება მიეღო y -ს, თუ შემთხვევით ფაქტორს ადგილი არ ექნებოდა;

$y_i - i$ -ურ დაკვირვებაში x_i -ურის შესაბამისი პროგნოზირებული მნიშვნელობაა;

e_i -სხვაობაა, რომელიც შეესაბამება i -ურ დაკვირვებაში y -ის ფაქტობრივ და პროგნოზირებულ მნიშვნელობებს შორის განსხვავებას. მისი ახსნა რეგრესიის განტოლებით არ ხერხდება.

რადგან $\text{cov}(\hat{y}_i, e)$ ნულის ტოლი უნდა იყოს, მივიღებთ:

$$\text{var}(y) = \text{var}(\hat{y}) + \text{var}(e)$$

მაგალითი 2.

გამოვიყენოთ ცხრილი 1.1 -ის მონაცემები უმცირეს კვადრატთა მეთოდის გამოყენების ილუსტრაციისათვის.

გამოთვლების გასამარტივებლად შევადგინოთ ცხრილი 2.1.

ცხრილი 2.1

N	სიმა ღლე, x_i	წონა Y_i	$y_i - \bar{y}$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})(x_i - \bar{x})$	\hat{y}_i	$e_i = y_i - \hat{y}_i$
1	5.0	140.0	-29.40	-5.35	28.62	157.29	135.3	4.7
2	9.0	157.0	-12.40	-1.35	1.82	16.74	160.8	-3.8
3	13.0	205.0	35.60	2.65	7.02	94.34	186.3	18.7
4	12.01	198.0	28.60	1.65	2.72	47.19	179.9	18.
5	10.0	162.0	-7.40	-0.35	0.12	2.59	167.2	-5.2
6	11.0	174.0	4.60	0.65	0.42	2.99	173.6	0.4
7	8.0	150.0	-19.40	-2.35	5.52	45.59	154.4	-4.4
8	9.0	165.0	-4.40	-1.35	1.82	5.94	160.8	4.2
9	10.0	170.0	0.60	-0.35	0.12	-0.21	167.2	2.8
10	12.0	180.0	10.60	1.65	2.72	17.49	179.9	0.1
11	11.0	170.0	0.60	0.65	0.42	0.39	173.6	-3.6
12	9.0	162.0	-7.40	-1.35	1.82	9.99	160.8	1.2
13	10.0	165.0	-4.40	-0.35	0.12	1.54	167.2	2.8
14	12.0	180.0	10.60	1.65	2.72	17.49	179.9	0.1
15	8.0	160.0	-9.40	-2.35	5.52	22.09	154.4	5.6
16	9.0	155.0	-14.49	-1.35	1.82	19.44	160.8	-5.8
17	10.0	165.0	-4.40	-0.35	0.12	1.54	167.2	-2.2
18	15.0	190.0	20.60	4.65	21.62	95.79	199.1	-9.1
19	13.0	185.0	15.60	2.65	7.02	41.34	186.3	-1.3
20	11.0	155.0	-14.40	0.65	0.42	-9.36	173.6	-18.6
	207.0	3388.0	0.0	0.0	92.50	590.20	3388. 3	-0.3

\hat{Y} - საშუალო მნიშვნელობაა 169.4

თუ გამოვიყენებთ რეგრესიის პარამეტრების შესაფასებლად მიღებულ (2.3) გამოსახულებებს და ცხრილის მონაცემებს, მივიღებთ:

$$\hat{S}_1 = \frac{\sum(X_i - \bar{X}) * (Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} = \frac{590.20}{92.50} = 6.38$$

$$\hat{s}_0 = \bar{Y} - \hat{s}_1 \bar{X} = 169.4 - (6.38 * 10.35) = 103.4$$

თუ ამ შედეგს შევადარებთ ცხრილი 2.1. -ში მიღებულ მნიშვნელობებს დავინახავთ, რომ შედეგები ერთმანეთს ემთხვევა. რასაკვირველია, ამ შრომატევად სამუშაოს ავცდებით, თუ გამოვიყენებთ კომპიუტერს და ეკონომეტრიკის რომელიმე პროგრამულ უზრუნველყოფას. ჩვენ შემთხვევაში ამავე სამუშაოს ჩასატარებლად გამოყენებულია STATA . გავანალიზოთ მიღებული ცხრილი, რომლის მიხედვითაც შეგვიძლია დავადგინოთ უმცირეს კვადრატთა მეთოდის ზოგიერთი თვისება, კერძოდ:

1. \hat{Y} -ის მნიშვნელობების ჯამი (სვეტი 9: ცხრილი 2.1) ტოლია \bar{Y} -ის მნიშვნელობების ჯამისა (სვეტი 2);
2. რეგრესიის წრფიდან გადახრების მნიშვნელობების ჯამი (სვეტი 9) ნულის ტოლია;
3. რეგრესიის წრფე გადის \bar{Y} და \bar{X} -ის ცვლადების საშუალო მნიშვნელობებზე.

2.3. მულტივარიაციული რეგრესიის მოდელი

მულტივარიაციული რეგრესიის მოდელი K დამოუკიდებელი ცვლადით შეიძლება წარმოვადგინოთ განტოლებით:

$$Y_i = S_0 + S_1 X_{1i} + S_2 X_{2i} + \dots + S_k X_{ki} + V_i$$

სადაც,

X_{ki} - დამოუკიდებელი ცვლადის მნიშვნელობას i -ურ დაკვირვებაში.

მნიშვნელოვანი განსხვავება მულტივარიაციულ რეგრესიის მოდელსა და ერთი დამოუკიდებელი ცვლადით რეგრესიის მოდელს შორის დაკავშირებულია დახრის კოეფიციენტების ინტერპრეტაციასთან, რომლებიც გამოხატავენ დამოკიდებული ცვლადის ცვლილებას განტოლების ერთი, რომელიმე დამოუკიდებელი ცვლადის ცვლილებისას, როდესაც დანარჩენი დამოუკიდებელი ცვლადები მუდმივ მნიშვნელობას ინარჩუნებენ. ე.ი. S_1 კოეფიციენტი გამოხატავს X_1 -ის ცვლილების გავლენას Y -ზე, როდესაც X_2, X_3, \dots, X_k დამოუკიდებელი ცვლადები მუდმივებია. კოეფიციენტი S_0 გამოხატავს Y -ის მნიშვნელობას, როდესაც ყველა, როგორც X მნიშვნელობები, ისე შემთხვევითი წევრის მნიშვნელობა ნულის ტოლია.

განვიხილოთ მულტივარიაციული რეგრესიის მაგალითი. გვაქვს მონაცემები წლების მიხედვით აშშ-ი ერთ სულ მოსახლეზე მოხმარებული ხორცის, წლიური შემოსავლისა და ხორცზე ფასების ცვლილების შესახებ. მონაცემები მოცემულია ცხრილით 2.2.

ცხრილი 2.2

წელი	Y	X1	X2
1960	85.1	6.036	20.4
1961	87.8	6.113	20.2
1962	88.9	6.271	21.3
1963	94.5	6.378	19.9
1964	99.9	6.727	18.0
1965	99.5	7.027	19.9

1966	104.2	7.280	22.2
1967	106.5	7.513	22.3
1968	109.7	7.728	23.4
1969	110.8	7.891	26.2
1970	113.7	8.134	27.1
1971	113.0	8.322	29.0
1972	116.0	8.562	33.5
1973	108.7	9.042	42.8
1974	115.4	8.867	35.6
1975	118.9	8.944	32.2
1976	127.4	9.175	33.7
1977	123.5	9.381	34.4
1978	117.9	9.735	48.5
1979	105.4	9.829	66.1
1980	103.2	9.722	62.4
1981	104.2	9.769	58.6
1983	103.7	9.725	56.7
1983	105.7	9.930	55.5
1984	105.5	10.414	57.3
1985	106.5	10.625	53.7
1986	107.3	10.905	52.6
1987	103.3	10.970	61.1

სადაც:

Y - საქონლის ხორცის მოხმარება ერთ სულზე (კგ);

X1- წლიური შემოსავალი ერთ სულზე (ათასი დოლარი);

X2- ერთი კგ ხორცის საშუალო წლიური ფასია, (ცენტი).

საძიებელი რეგრესიის განტოლება ჩავწეროთ შემდეგი სახით:

$$\hat{Y} = \hat{\alpha}_0 + \hat{\alpha}_1 X_1 + \hat{\alpha}_2 X_2$$

განტოლების კოეფიციენტების მოსაძებნად გამოყენებულია პროგრამული უზრუნველყოფა { Small STATA 5.0 და მიღებულია შედეგები მოცემულია ქვემოთ:

reg Y X1 X2

Source	SS	df	MS	Number of obs =	28
-----+-----				F(2, 25) =	5.20
Model	793.833872	2	396.916936	b > F	= 0.0130
Residual	1909.19617	25	76.3678468	ared	= 0.2937
-----+-----				Adj R-squared =	0.2372
Total	2703.03004	27	100.112224	MSE	= 8.7389

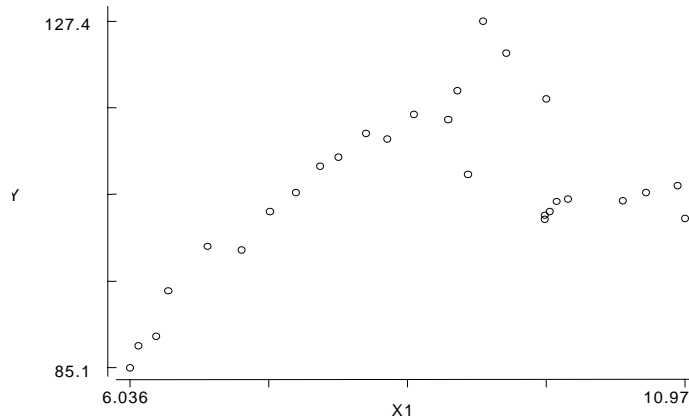
Y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
-----+-----					
X1	3.662684	1.140879	3.210	0.004	1.313 6.012368
X2	-.0504742	.0503924	-1.002	0.326	-.1542593 .0533109
_cons	77.34607	9.717361	7.960	0.000	57.33279 97.35935

როგორც მიღებული შედეგებიდან ჩანს რეგრესიის განტოლებაა:

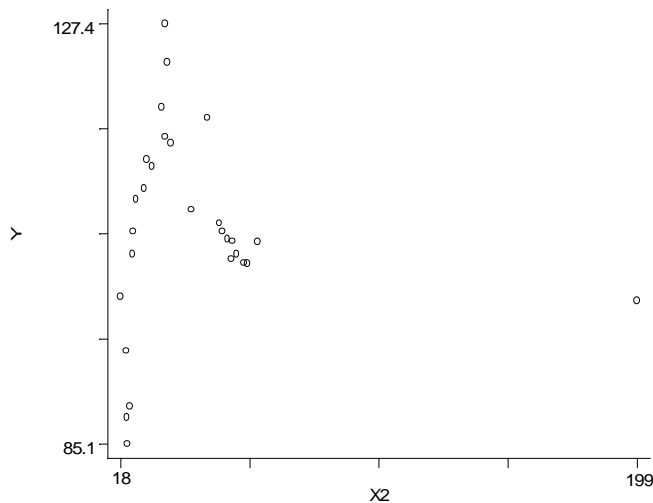
$$Y = 77.3461 + 3.6627X_1 - 0.5047X_2$$

კოეფიციენტების სტანდარტული შეცდომები და მათი საიმედოობის ინტერვალები მოცემულია STATA -ს ცხრილში. რეგრესიის განტოლების უკეთ გაანალიზებისათვის წარმოვადგინოთ გრაფიკული სახე (Y, X1) და (Y, X2) მაჩვენებლებისათვის. ამ პროცედურისათვის კვლავ გამოვიყენოთ პროგრამული უზრუნველყოფა (STATA 5.0).

გრაფიკების სახე წარმოდგენილია ნახ.2.3 და ნახ. 2.4



ნახ.2.3 ერთ სულ მოსახლეზე ხორცის მოხმარებასა და შემოსავალს შორის დამოკიდებულება



ნახ.2.4 ერთ სულ მოსახლეზე ხორცის მოხმარების ცვლილება ფასების გავლენის შედეგად

როგორც ნახაზებიდან ჩანს და რეგრესიის განტოლების კოეფიციენტიც ამაზე მიუთითებს, შემოსავლის ზრდასთან ერთად ხორცის მოხმარება იზრდება. რის გამოც რეგრესიის განტოლებაში კოეფიციენტი შემოსავალთან დადებით ნიშანს ატარებს. ხოლო რაც შეეხება ხორცის მოხმარებაზე ფასების გავლენას, რეგრესიის

განტოლებაში ფასთან კოეფიციენტი უარყოფითი ნიშნის მატარებელია და გრაფიკიც ამასვე მიუთითებს. ფასების ზრდა უარყოფითად მოქმედებს, თუმცა მისი გავლენა ერთ სულ მოსახლეზე ხორცის მოხმარებაზე შედარებით ნაკლებ საგრძნობია, ვიდრე შემოსავლის გავლენა.

ამ მაგალითის განხილვისას მხედველობაშია მისაღები მაჩვენებლებს შორის დამოკიდებულებების ახსნის შესაძლებლობა პროგრამული უზრუნველყოფის გამოყენებით, როდესაც რეგრესიის განტოლების აგების არსი გარკვეულია თუნდაც ერთი დამოუკიდებელი ცვლადის შემთხვევაში. Small Stata 5.0 –ის გამოყენებით ჩვენს მიერ შესასრულებელი იყო შემდეგი პროცედურა:

1. მონაცემების შეყვანა Stata-ს მოქმედების არეში, რაც დაკავშირებულია ბრძანებების შესრულებასთან შემდეგი მიმდევრობით: start, მენიუდან პროგრამის შერჩევა, editor –ის დახმარებით მონაცემების შეტანა კლავიატურიდან (თუ წინასწარ ფაილი არ გვაქვს შექმნილი), მაჩვენებლებისათვის სახელების მინიჭებით (ჩვენს მაგალითში მინიჭებული სახელებია: დამოკიდებული ცვლადისათვის - Y, ხოლო დამოუკიდებელი ცვლადებისათვის - X1 და X2;

2. Stata -ს ბრძანებათა სტრიქონში ვწერთ ბრძანებას რეგრესიის განტოლების მისაღებად. ბრძანების სახეა:

```
regress Y X1 X2 ან  
reg Y X1 X2
```

დისპლეიზე აისახება ამოცანის ამოხსნის შედეგი: რეგრესიის განტოლება, განტოლების კოეფიციენტების სტანდარტული შეცდომები, დაკვირვებების რიცხოვნება, თავისუფლების ხარისხი, ნდობის ინტერვალები და სხვა.

3. დამოუკიდებელ ცვლადებსა და დამოკიდებულ ცვლადს შორის დამოკიდებულების გრაფიკული წარმოდგენის მიზნით Stata -ს ბრძანებათა სტრიქონში ჩაიწერება ბრძანება:

```
graph Y X1
```

(დისპლეიზე წარმოდგენილი შედეგი შეესაბამება ნახ.2.3) და მისი შესრულების შემდეგ - ბრძანებას:

```
graph Y X2
```

(დისპლეიზე წარმოდგენილი შედეგი შეესაბამება ნახ.2.4.).

თუ შედეგების გამობეჭდვა დაგვჭირდება Stata -ს მენიუდან მივცემთ ბრძანებას

```
Print
```

ხოლო თუ შედეგების დახსომება დაგვჭირდება, იმავე მენიუდან შევარჩევთ ბრძანებას

```
Save as
```

და მივუთითებთ ადგილს დასახსომებლად.

2.4. რეგრესიის განტოლების თვისობრივი შეფასება

თუ ჩავთვლით, რომ რეგრესიული ანალიზის საფუძველი უმცირეს კვადრატთა მეთოდია, მაშინ ეკონომეტრისტის საფიქრალია, თუ რამდენად კარგად ერგება უმცირეს კვადრატთა მეთოდი შესასწავლი ამოცანის არსს. ხშირად, მოსაზრება იმის შესახებ, რომ ამოცანის ამოხსნის შედეგი კომპიუტერის დახმარებითაა მიღებული, დავის საგანს აღარ წარმოადგენს, რაც ყოველთვის გამართლებული არ არის. მას შემდეგ, როდესაც პროგრამული უზრუნველყოფის გამოყენების შედეგად

მიღებულია შედეგი, მისი ანალიზის თვალსაზრისით ისეთ კითხვებს მაინც უნდა ვუპასუხოთ, როგორცაა:

1. ეხმაურება თუ არა მიღებული განტოლება თეორიულ მიდგომას;
2. რამდენად მოიცავს რეგრესიულმა ანალიზმა ჩვენს ხელთ არსებული მონაცემები;
3. საკმარისია თუ არა არსებული ინფორმაციის მოცულობა შესასწავლი პრობლემის დასახასიათებლად და რამდენად ზუსტად ასახავს მას;
4. არის თუ არა უმცირეს კვადრატთა მეთოდი საუკეთესო (შეფასების თვალსაზრისით) მიღებული განტოლებისათვის;
5. არის თუ არა ყველა მნიშვნელოვანი ცვლადი, რომლებსაც გავლენა აქვთ დამოკიდებულ ცვლადზე განტოლებაში შეყვანილი;
6. აქვს თუ არა საუკეთესო ლოგიკური ფუნქციონალური ფორმა მიღებულ შედეგს.

2.5. დეტერმინაციის კოეფიციენტი

განვიხილოთ დამოკიდებული ცვლადის Y - ის ვარიაცია თავის საშუალო მნიშვნელობის მიმართ:

$$\text{var}(Y) = \sum (Y_i - \bar{Y})^2$$

და წარმოვადგინოთ ის ორი მდგენელის სახით, რომელთაგან ერთი რეგრესიის განტოლების ვარიაციაა, ხოლო მეორე - დაკავშირებულია რეგრესიის განტოლების შემთხვევით წევრთან. რადგან

$$\hat{Y}_i = \hat{S}_0 + \hat{S}_1 X_i \quad ,$$

$$\text{ამიტომ, } \text{var}(Y) = \sum (Y_i - \bar{Y})^2 = \sum (Y_i - \hat{Y}_i)^2 + \sum (\hat{Y}_i - \bar{Y})^2 + 2 \sum (Y_i - \hat{Y}_i)(\hat{Y}_i - \bar{Y})$$

ამ გამოსახულების მესამე შესაკრები ნულის ტოლია, რის გამოც:

$$\text{var}(Y) = \sum (Y_i - \bar{Y})^2 = \sum (Y_i - \hat{Y}_i)^2 + \sum (\hat{Y}_i - \bar{Y})^2$$

შემოვიტანოთ აღნიშვნები:

კვადრატების ჯამი (total sum of squares)

$$\text{TSS} = \sum (Y_i - \bar{Y})^2 ;$$

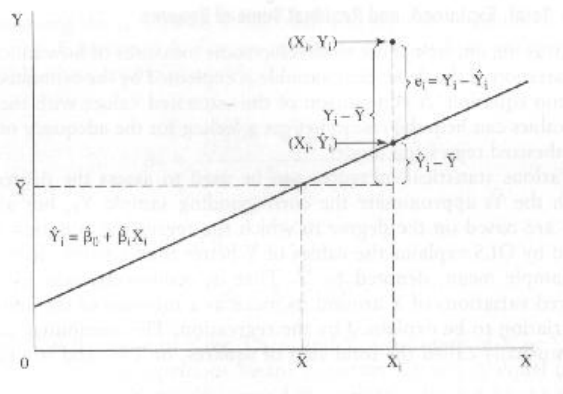
შცდომების (გადახრების) კვადრატების ჯამი (error sum of squares)

$$\text{ESS} = \sum (Y_i - \hat{Y}_i)^2 ;$$

რეგრესიის კვადრატების ჯამი (regression sum of squares)

$$\text{RSS} = \sum (\hat{Y}_i - \bar{Y})^2$$

Y -ის ვარიაციის დეკომპოზიცია ნაჩვენებია ნახ. 2.5-ზე.



ნახ.2.5 Y -ის ვარიაციის დეკომპოზიცია

განსაზღვრება: Y -ის დისპერსიის ნაწილს, რომელიც რეგრესიის განტოლებით აიხსნება დეტერმინაციის კოეფიციენტი ეწოდება და ჩაიწერება შედეგი სახით:

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum e_i^2}{\sum (Y_i - \bar{Y})^2}$$

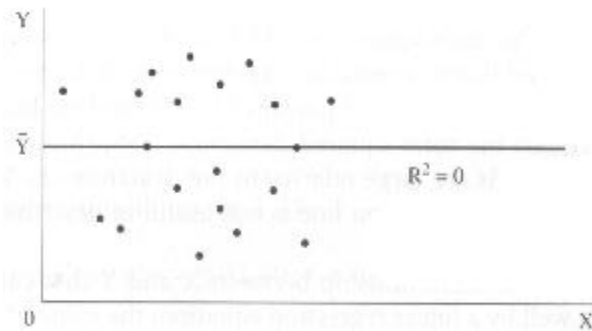
მისი მაქსიმალური მნიშვნელობა ერთის ტოლია $R^2 = 1$, რაც იმ შემთხვევას შეესაბამება, როდესაც რეგრესიის წრფე ზუსტად ასახავს დაკვირვების ყველა შედეგს, ხოლო იმ შემთხვევაში, თუ Y და X ცვლადებს შორის დამოკიდებულება არ არსებობს, მაშინ R^2 ნულის ტოლია. ყველა სხვა შემთხვევაში სასურველია, რომ R^2 -ის მნიშვნელობა მისწრაფვოდეს 1-კენ. ბუნებრივია, ჩვენ დაინტერესებულნი ვიქნებით, რომ რეგრესიის განტოლების პარამეტრების ისე შეირჩეს, რომ R^2 -ის მნიშვნელობა მაქსიმალური იყოს.

ხომ არ ეწინააღმდეგება ეს მოთხოვნა კრიტერიუმს, რომლის შესაბამისად რეგრესიის განტოლების პარამეტრების შერჩევა სასურველია ისე, რომ ადგილი ჰქონდეს სხვაობების კვადრატების ჯამის მინიმიზაციას. ბუნებრივია: არა! მთავარი პრობლემა დეტერმინაციის კოეფიციენტთან დაკავშირებით მდგომარეობს შემდეგში: თუ კონკრეტულ განტოლებაში შემოვიტანთ სხვა დამოუკიდებელ ცვლადს, ამან შეიძლება არ გამოიწვიოს მისი მნიშვნელობის დაწევა. ეს კი ნიშნავს შემდეგს: რეგრესიის ორი განტოლების შედარებისას რეგრესიის იმ განტოლებას, რომელიც დამოუკიდებელი ცვლადების დიდ რაოდენობას მოიცავს ექნება დეტერმინაციის კოეფიციენტის უკეთესი მნიშვნელობა. მართლაც, თუ ამ კოეფიციენტის მნიშვნელობას გავანალიზებთ, შეგვიძლია ასეთი დასკვნა გავაკეთოთ: ვიდრე დამოკიდებული ცვლადი არ იცვლება, TSS -ის მნიშვნელობა იგივე რჩება. ასევე, ვიდრე უმცირეს კვადრატთა მეთოდი იმის გარანტიას იძლევა, რომ ცვლადის დამატება არ გაზრდის გადახრების კვადრატების ჯამს, RSS -ის მნიშვნელობა შეიძლება შემცირდეს ან იგივე დარჩეს. თუ RSS -ის მნიშვნელობა შემცირდება, მაშინ ასევე შემცირდება შეფარდება RSS/TSS და $(1 - RSS/TSS)$ მნიშვნელობა გაიზრდება. ამგვარად, განტოლებაში ცვლადის დამატება იძლევა იმის გარანტიას, რომ დეტერმინაციის კოეფიციენტის მნიშვნელობა გაიზრდება. არსებითად, დეტერმინაციის კოეფიციენტი გვცხმარება გადაწყვეტილების მიღებაში შემოვიტანოთ დამატებითი ცვლადი, თუ არა. ამ მიზნით შეიძლება შემოვიტანოთ

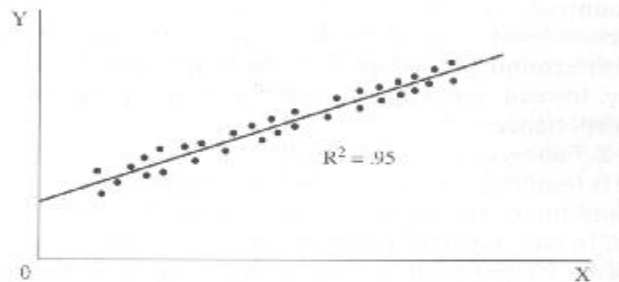
კორექტირებული დეტერმინაციის კოეფიციენტი ე.ი დეტერმინაციის კოეფიციენტი კორექტირებული თავისუფლების ხარისხით.

$$\bar{R}^2 = 1 - \frac{RSS / (n - K - 1)}{TSS / (n - 1)} = 1 - \frac{\sum e_i^2 / (n - K - 1)}{\sum (Y_i - \bar{Y})^2 / (n - 1)}$$

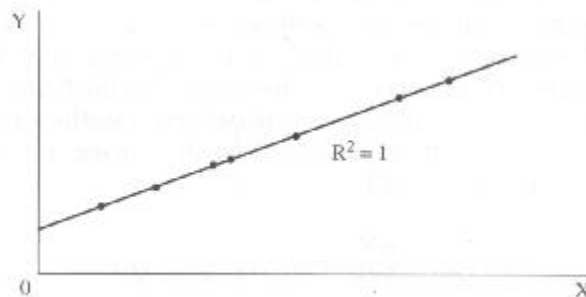
ნახაზებით 2.6 – 2.8 წარმოდგენილია რეგრესიის წრფისა და დეტერმინაციის კოეფიციენტის სხვა და სხვა მნიშვნელობების ურთიერთშესაბამისობა.



ნახ. 2.6. X და Y ცვლადები ერთმანეთთან კავშირში არ იმყოფებიან ამ შემთხვევაში დეტერმინაციის კოეფიციენტის მნიშვნელობა ნულის ტოლია.



ნახ.2.7. X და Y ცვლადების მონაცემების შესაბამისი მნიშვნელობების განლაგება კარგად ეთანადება რეგრესიის წრფეს ამ შემთხვევაში დეტერმინაციის კოეფიციენტის მნიშვნელობა 0.95 -ის ტოლია.



ნახ.2.8. X და Y ცვლადების მონაცემების შესაბამისი მნიშვნელობები განლაგებულია რეგრესიის წრფეზე ამ შემთხვევაში დეტერმინაციის კოეფიციენტი ერთის ტოლია.

2.6. ლაგური ცვლადები

რეგრესიის განტოლებებში შეიძლება გამოვყოთ ორი ძირითადი ტიპი. განტოლებების პირველი ტიპი შეიცავს ცვლადებს როგორც დამოკიდებულ, ისე დამოუკიდებლებს, რომლებიც ერთსა და იმავე პერიოდში განიხილება. მაგალითად:

$$Y_t = S_0 + S_1 X_{1t} + S_2 X_{2t} + V_t$$

სადაც t მიუთითებს დროის კონკრეტულ მომენტზე, რომელშიც განიხილება განტოლებაში შემავალი ცვლადები.

მაგრამ ეკონომიკასა და ბიზნესში დამოკიდებული და დამოუკიდებელი ცვლადების ერთდროული თანხვედრა დროში ყოველთვის არ გვხვდება და ხშირია შემთხვევები, როდესაც ეს ცვლადები ძვრას განიცდიან დროში. დროის ხანგრძლივობის სიგრძეს ლაგი ეწოდება. მრავალი ეკონომეტრიკული განტოლება შეიცავს ერთ ან მეტ ლაგურ დამოუკიდებელ ცვლადს, როგორცაა $X_{1(t-1)}$, სადაც ინდექსი $t-1$ აღნიშნავს, რომ X_1 ცვლადზე დაკვირვება წინ უსწრებდა t პერიოდს: როგორც ეს ქვემოთ მოყვანილ განტოლებაშია:

$$Y_t = S_0 + S_1 X_{1(t-1)} + S_2 X_{2t} + V_t \quad (2.3.)$$

ე.ი X_1 დროის ერთი პერიოდის მიხედვით იყო ლაგური, მაშინ როდესაც დამოკიდებულება Y -სა და X_1 -ს შორის დროის მიხედვით არ შეცვლილა.

მაგალითი. განვიხილოთ პროცესი, რომელიც უკავშირდება სასოფლო-სამეურნეო პროდუქტის მიწოდებას: მოსავლის აღება (მარცვლეული, ბოსტნეული, ბაღჩეული და სხვა) ხორცის დამზადება, კვერცხის დამზადება და სხვა. ასევე, სასოფლო-სამეურნეო ბაზარზე ფასების ცვლილება ძნელად თუ განიხილება ერთსა და იმავე დროში, რადგან მათ ლაგური ხასიათი აქვთ და შეიძლება შემდეგი სახით წარმოვადგინოთ:

$$C_t = S_0 + S_1 PC_{t-1} + S_2 PF_t + V_t \quad (2.4.)$$

სადაც: C_t - პროდუქტზე მოთხოვნილი რაოდენობაა t წელს;

PC_{t-1} - პროდუქტზე ფასი $t-1$ წელს ;

PF_t - პროდუქტის წარმოებაზე შრომის ანაზღაურება t წელს.

ამ განტოლების მიხედვით ლაგი არსებობს პროდუქტის რაოდენობასა და ამ პროდუქტზე ფასს შორის.

რეგრესიის განტოლებაში ლაგური ცვლადების კოეფიციენტების მნიშვნელობა არ არის იგივე, როგორც განტოლებაში, რომელსაც ლაგი არ სდევს. კოეფიციენტი ლაგურ ცვლადთან ზომავს ამ წლის დამოკიდებული Y ცვლადის ცვლილებას გასული წლის X -ის მნიშვნელობის ერთი ერთეულით გაზრდისას (სხვა დამოუკიდებელი ცვლადები მუდმივ დონეზე განიხილებიან). (2.4) განტოლებაში S_1 კოეფიციენტი ზომავს პროდუქტის იმ რაოდენობას: რომელსაც აიღებდა ფერმერი ამ წელიწადს იმ პირობით, რომ გასულ წელს პროდუქტზე ფასი ერთი ერთეულით რომ გაზრდილიყო, ხოლო შრომის ანაზღაურება ამ წელს მუდმივი დარჩენილიყო.

2.7. ფიქტიური ცვლადები

ზოგიერთი ცვლადი, მაგალითად სქესი, თვისობრივი ხარისხის მატარებელია და მათი გამოსახვა რაოდენობრივად არ ხერხდება, რის გამოც ერთ-ერთი გზა

მათი რაოდენობრივი სახით გამოსახვის თვალსაზრისით არის მათი წარმოდგენა ფიქტიური (ბინარული) სახით. ამ შემთხვევაში ფიქტიური ცვლადს შეუძლია მიიღოს მხოლოდ ორი მნიშვნელობა: 0 ან 1.

ვთქვათ, Y წარმოადგენს i -ური სკოლის მასწავლებლის ხელფასს და ხელფასის დონე დამოკიდებულია მასწავლებლის გამოცდილებასა და მის ხარისხზე. ვთქვათ, ყველა მასწავლებელს ბაკალავრის ხარისხი აქვს და ზოგიერთს - მაგისტრის ხარისხი. რეგრესიის განტოლება, რომელიც აკავშირებს ხელფასის დონეს მასწავლებლის ხარისხსა და გამოცდილებასთან ჩაიწერება შემდეგი სახით:

$$Y_i = S_0 + S_1 X_{1i} + S_2 X_{2i} + V_i$$

სადაც: $X_{1i} = \begin{cases} 1 \\ 0 \end{cases}$

X_{2i} - მასწავლებლის გამოცდილება (წელი);

1- თუ მასწავლებელს მაგისტრის ხარისხი აქვს;

0 - თუ მასწავლებელს ბაკალავრის ხარისხი აქვს.

რეგრესიის განტოლების კოეფიციენტების ინტერპრეტაცია შეიძლება ასე წარმოვადგინოთ:

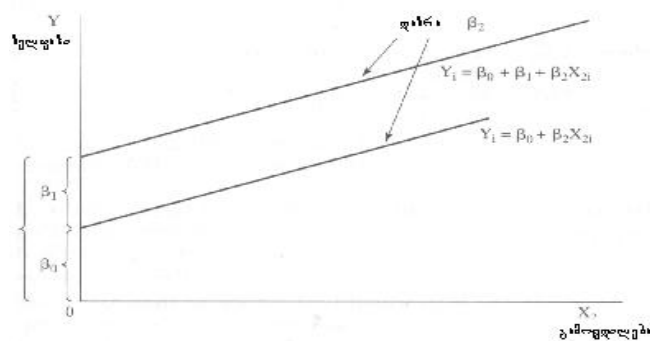
1. თუ მასწავლებელს ბაკალავრის ხარისხი აქვს, მაშინ X_1 -ის მნიშვნელობა ნულის ტოლია და რეგრესიის განტოლებას ეწება სახე:

$$E(Y_i | X_1) = S_0 + S_2 X_{2i}$$

2. თუ მასწავლებელს მაგისტრის ხარისხი აქვს, მაშინ X_1 -ის მნიშვნელობა ერთის ტოლია და რეგრესიის განტოლებას ეწება სახე:

$$E(Y | X_i) = S_0 + S_1 + S_2 X_{2i}$$

შევეცადოთ და შევადაროთ ეს ორი განტოლება: S_1 წარმოადგენს ხელფასზე საშუალო დანამატს, როდესაც მასწავლებელს მაგისტრის ხარისხი აქვს. ეს ინტერპრეტაცია საკმაოდ მნიშვნელოვანია, რადგან მკვლევარს საშუალებას აძლევს დაადგინოს ამ კოეფიციენტის ნიშანი. გრაფიკულად ამ განტოლებების წარმოდგენის შემთხვევაში S_0 და S_1 მნიშვნელობები მონაკვეთებია Y ღერძზე. გრაფიკი მოცემულია ნახ. 2.9-ით.



ნახ. 2.9. ფიქტიური ცვლადის გრაფიკული ინტერპრეტაცია

რეზიუმე

1. რეგრესიის განტოლების კოეფიციენტების შესაფასებლად უმცირეს კვადრატთა მეთოდი ყველაზე ხშირად გამოყენებული მეთოდია. ამ მეთოდით მიღებული პარამეტრების მნიშვნელობები აკმაყოფილებენ სხვაობების კვადრატების მინიმიზაციის პირობას ე.ი. $\min \sum e_i^2$ მიისწრაფის ნულისაკენ;

2. დეტერმინაციის კოეფიციენტი ერთ-ერთი მარტივი მაჩვენებელია რეგრესიის განტოლების შესაფასებლად. არსებითად, დეტერმინაციის კოეფიციენტი გვეხმარება გადაწყვეტილების მიღებაში შემოვიტანოთ დამატებითი ცვლადი თუ არა რეგრესიის განტოლების გაუმჯობესების მიზნით;

3. ყოველთვის უნდა გვახსოვდეს, რომ შესაფასებელი განტოლებისათვის პროგნოზის ხარისხი დამოკიდებულია რეგრესიის განტოლების კოეფიციენტებზე, რომელთა სიზუსტე უნდა განსაზღვრავდეს ეკონომიკური თეორიის მოთხოვნას რეგრესიის განტოლებაზე.

4. რეგრესიული ანალიზის ტიპური ეტაპებია:

- მოდელის სპეციფიკაცია, დამოუკიდებელი ცვლადების გამოყოფა და ფუნქციონალური ფორმის ჩაწერა;
- განტოლების კოეფიციენტების ნიშნების დადგენა;
- მონაცემების შეგროვება;
- განტოლების შეფასება;
- შედეგების ანალიზი.

5. ფიქტიური ცვლადების შემოყვანა რეგრესიის განტოლებაში დაკავშირებულია დამოუკიდებელი ცვლადის შინაარსთან. კერძოდ, თუ დამოუკიდებელი ცვლადი თვისობრივი ხასიათის მატარებელია, მაშინ მისი წარმოდგენა დასაშვებია ფიქტიური (ბინარული) ცვლადით, რომელიც მხოლოდ ორ მნიშვნელობას იღებს: 0-ს ან 1-ს.

დამოუკიდებელი სამუშაო

დევიდ კატმა (David Katz, "Faculty salaries, Promotions, and Productivity at a university", American Economic Review, 1973, pp. 469-477 (ავტ. თარგმანი)) შეისწავლა უნივერსიტეტის მასწავლებლების ხელფასების ცვლილება მასწავლებლების "პროდუქტიულობის" მიხედვით და მიიღო რეგრესიის განტოლება:

$$\hat{S}_i = 11.155 + 230B_i + 18A_i + 102E_i + 489D_i + 189Y_i + \dots$$

სადაც:

S_i - i -ური პროფესორის ხელფასია 1969-70 წ.წ. (დოლარი);

B_i - i -ური პროფესორის მიერ გამოქვეყნებული წიგნების, (ცალი);

A_i - i -ური პროფესორის მიერ გამოქვეყნებული სტატიების (ცალი);

E_i - i -ური პროფესორის მიერ გამოქვეყნებული საუკეთესო სტატიების (ცალი);

D_i - i -ური პროფესორის ხელმძღვანელობით დაცული დისერტაციების (ცალი);

Y_i - უნივერსიტეტში მუშაობის გამოცდილება, სამუშაო სტაჟი (წელი)

ა/ შეესაბამება თუ არა რეგრესიის განტოლების კოეფიციენტების ნიშნები თქვენს შეხედულებას ?

ბ/ რამდენად ლოგიკურია რეგრესიის მოცემულ განტოლებაში ცვლადების კოეფიციენტების ნიშნები?

გ/ ვთქვათ, პროფესორს აქვს საკმარისი დრო წიგნის დასაწერად ან ორი საუკეთესო სტატიის დასაწერად ან გაუწიოს ხელმძღვანელობა სამ დისერტაციას. როგორი იქნება თქვენი რეკომენდაცია და რატომ?

დ/ როგორი შეფასება შეიძლება მიეცეს მოცემულ რეგრესიის განტოლებას?
ხომ არ არის განტოლება შინაარსს მოკლებული? რატომ?

სტუდენტების აკადემიური მოსწრების გამოკვლევის მიზნით დევიდ რომერმა (David Romer, "Do Students Go to class? Should They?" Journal of Economic Perspectives, 1993, pp. 167-174) მიიღო განტოლება:

$$G_i = f(ATT_i, PS_i)$$

სადაც:

G_i - i -ური სტუდენტის შეფასება რომერის კლასში ($A=4, B=3, \dots$);

PS_i - ლექციების წილი, რომლებსაც სტუდენტი ესწრება;

ATT_i - პრობლემების წილი, რომლებიც სტუდენტს ექმნება სემესტრის განმავლობაში;

ა/ რა ნიშნებია მოსალოდნელი განტოლებაში დამოუკიდებელი ცვლადების კოეფიციენტებთან? ახსენით თქვენი მოსაზრება.

ბ/ რომერის მიერ შეფასებული განტოლებაა :

$$\hat{G}_i = 1.07 + 1.74ATT_i + 0.60PS_i$$

$$n = 195; R^2 = 0.33$$

შეესაბამება თუ არა თქვენს მოლოდინს მიღებული შედეგები?

გ/ ვთქვათ, 25 საათი ეთმობა ლექციებს სემესტრში, ხოლო სტუდენტს ესაჭიროება 50 საათი თავის პრობლემების მოსაგვარებლად სემესტრში. თუ სტუდენტს აქვს მხოლოდ ერთი საათი რომელიმე ლექციაზე დასასწრებლად და სურს მიიღოს მაღალი შეფასება, აიღებს თუ არა ის დამატებით საათს?

დ/ შეიძლება თუ არა გააკეთოთ დასკვნა, რომ დამოუკიდებელ ცვლადთან კოეფიციენტის დიდი მნიშვნელობა დიდ გავლენას ახდენს დამოუკიდებელ ცვლადზე?

როგორი იქნებოდა თქვენი პასუხი იმ შემთხვევაში, თუ ლექციების წილი 50 საათი დარჩებოდა სემესტრში და 10 საათი - სტუდენტის პრობლემების მოსაგვარებლად;

ე/ რას ნიშნავს $R^2 = 0.33$?

რა შეიცვლებოდა, თუ თქვენ დაამატებდით ერთ ან რამდენიმე ცვლადს განტოლებაში? ახსენით თქვენი მოსაზრება.

2.3. Charkts A. Lave -ის (Charkts A. Lave, "Speeding, Coordination..." , American economic review, December, 1985, pp. 1159-1164) მიერ შესწავლილი იყო საბედისწერო საგზაო შემთხვევები აშშ-ს მიხედვით 1981 და 1982 წლებისათვის. მან მიიღო რეგრესიის განტოლებები:

$$\hat{F}_i = \hat{S}_0 + 0.176V_i + 0.0136C_i - 7.75H_i$$

1981წ.:

$$\bar{R}^2 = 0.624; n = 41$$

$$\hat{F}_i = \hat{S}_0 + 0.181V_i + 0.0071C_i - 5.29H_i$$

1982წ.

$$\bar{R}^2 = 0.532; n = 44$$

სადაც :

F_i - საბედისწერო შემთხვევები i -ური შტატის მაგისტრალზე;

\hat{S}_0 - მონაკვეთის სიდიდე F_i -ის ღერძზე;

V_i - მანქანის სიჩქარე i -ურ შტატში;

C_i - განაცხადები საძიებელ მძღოლებზე i -ურ შტატში;

H_i - საავადმყოფოს დაშორება შემთხვევის ადგილიდან.

ა/ გაიაზრეთ თითოეული ცვლადის შინაარსი და შეაფასეთ თითოეული კოეფიციენტის ნიშანი.

ბ/ არის თუ არა განსხვავება რეგრესიის განტოლებებს შორის ?

გ/ რომელი განტოლებისთვისაა კორექტირებული დეტერმინაციის კოეფიციენტი ?

2.4. ეკონომიკის ცოდნის შეფასების ერთ-ერთი ცნობილი ტესტი GRE (Graduate Record Examination), რომლითაც შეიძლება შეფასდეს გამოსაცდელის ანალიტიკური შესაძლებლობები ეკონომიკაში, გამოიყენება სტუდენტებისათვის, რომლებსაც სადოქტორო ხარისხის დაცვა სურთ. შემდეგში მიიჩნის, რომ ეს ტესტი, როგორც SAT (Scholastic Aptitude Test) ტესტი მიმართული იყო ქალებისა და ეთნიკური ჯგუფების წინააღმდეგ. იმის შესასწავლად, თუ რამდენად ასახავდა ეს მოსაზრება სინამდვილეს Mary Hirschfeld-მა, Robert Moore-მა და Eleanor Brown-მა შეაფასეს შემდეგი განტოლება:

$$GRE = 172.4 + 39.7G_i + 78.9GPA_i + 0.203ATM_i + 0.110SATV_i$$

$$10.9; 10.4; 0.071; 0.058$$

$$n = 149$$

$$\bar{R}^2 = 0.46$$

სადაც: GRE – ეკონომიკაში სტუდენტის შეფასება ამ ტესტით;

G – ფიქტიური ცვლადია და 1-ის ტოლია, თუ სტუდენტი მამრობითი სქესისაა. დანარჩენ შემთხვევაში - 0-ის ტოლია;

GPA – სტუდენტის შეფასება ეკონომიკაში ([A=4; B=3...]);

SATM - სტუდენტის შეფასება მათემატიკური ტესტით;

SATV – მათემატიკაში სტუდენტის შეფასება ზეპირი გამოცდის სახით.

ა/ ახსენით G -ის კოეფიციენტის მნიშვნელობა განტოლებაში;

ბ/ მიღებული შედეგი ამტკიცებს თუ არა, რომ ტესტი GRE მიმართულია ქალების წინააღმდეგ, რატომ?

გ/ თუ დააპირებთ, რომ ამ განტოლებას დაამატოთ ერთი ცვლადი მაინც, რომელი ცვლადი იქნებოდა და ახსენით მიზეზი;

დ/ ვთქვათ, ავტორებმა სქესის შესაბამისი ცვლადის მნიშვნელობა შეცვალეს: მდებარეობით სქესს შეუსაბამეს 1, ხოლო დანარჩენ შემთხვევას - 0. რა შეიცვლებოდა განტოლებაში?

2.5. თქვენმა ხელმძღვანელმა გადაწყვიტა ახალი პროდუქციის გამოშვება და დაგეგმათ ააგოთ მოდელი ერთობლივი შემოსავლების შესახებ ბოლო ხუთი წლის განმავლობაში გამოშვებული ფილმების მიხედვით. თქვენს მიერ მიღებული რეგრესიის განტოლება:

$$\hat{G}_i = 781 + 15.4T_i - 992F_i + 1770J_i + 302.7S_i - 3160B_i + \dots;$$

5.9;674;800;1006;2886

$$\bar{R}^2 = 0.485; \quad n=254$$

სადაც: G_i - i -ური ფილმიდან მიღებული ერთობლივი შემოსავალი (ათასი \$);

T_i - კინოთეატრების რიცხოვნება, სადაც i -ურ ფილმს აჩვენებდნენ;

F_i - ფიქტიური ცვლადი, რომელიც 1-ის ტოლია თუ ფილმში მონაწილე ვარსკვლავი ქალია, ხოლო სხვა დანარჩენ შემთხვევაში 0-ია;

J_i - ფიქტიური ცვლადი, რომელიც 1-ის ტოლი, თუ ფილმი ივნისში ან ივლისში იყო ნაჩვენები. სხვა დანარჩენ შემთხვევაში 0-ის ტოლია;

S_i - ფიქტიური ცვლადი, რომელიც 1-ის ტოლია თუ ფილმში მონაწილეობს სუპერვარსკვლავი (მაგალითად, ტომ კრუზი ან სხვა). სხვა დანარჩენ შემთხვევაში 0-ის ტოლია;

B_i - ფიქტიური ცვლადი, რომელიც 1-ის ტოლია, თუ ფილმის სპონსორი სუპერვარსკვლავია, სხვა დანარჩენ შემთხვევაში 0-ის ტოლია?

ა/ რა ნიშანია მოსალოდნელი განტოლების თითოეული კოეფიციენტისათვის?

ბ/ თქვენს ხელმძღვანელს სურს უჩვენოს ფილმი 200 ეკრანზე, რაც მას დაუჯდება \$1200000. როგორი იქნება თქვენი შეფასება?

გ/ ფილმის ჩვენება გათვალისწინებულია სექტემბერში და ეღირება \$1000000; რამდენად ღირებულია ასეთი გადაწყვეტილება?

დ/ ვთქვათ, თქვენი განტოლება მისაღებია. გაქვთ თუ არა რაიმე მოსაზრება ამის საწინააღმდეგოდ?

2.6. გვაქვს მონაცემები თვეების მიხედვით ინდივიდის დანახარჯების შესახებ კომერციულ სექტორში მისი შემოსავლების მიხედვით:

წელი. თვე	ინდივიდის დანახარჯი, \$	შემოსავალი, \$
მარტი 92	404,438	440,110
ივნისი 92	409,782	446,32
სექტ. 92	415,398	450,0
დეკემბ. 92	425,133	465,879
მარტი 93	429,456	459,820
ივნისი 93	434,730	467,860
სექტ. 93	440,118	470,05
დეკემბ.93	446,960	477,760
მარტი 94	453,501	483,279
ივნისი 94	458,644	491,35
სექტ. 94	465,746	499,029
დეკემბ. 94	473,481	510,189

მარტი 95	478,208	518,439
ივნისი 95	485,102	520,1

შეისწავლეთ დამოკიდებულება მოცემულ მაჩვენებლებს შორის. ააგეთ რეგრესიის მოდელი. შეაფასეთ მოდელის კოეფიციენტები და გამოთვალეთ დეტერმინაციის კოეფიციენტი.

2.7. ქვემოთ მოყვანილ ცხრილში მოცემულია აშშ-ში სხვადასხვა მაჩვენებლების მიხედვით ხელფასის ცვალებადობა (ათასი დოლარი). ცხრილის მონაცემებით შეისწავლეთ ასაკისა და განათლების მიხედვით ქალებისა და მამაკაცების ხელფასებს შორის დამოკიდებულება. ააგეთ სქესის მიხედვით რეგრესიის განტოლებები. შეაფასეთ რეგრესიის განტოლებების პარამეტრები და გააკეთეთ სათანადო დასკვნები.

N	ხელფასი	სქესი 0-მამრ 1-მდედრ.	განათლება, წელი	ასაკი
1	8,99977	0	10	43
2	5,49973	0	12	38
3	3,79999	1	12	22
4	10,5002	1	12	47
5	14,9992	0	12	58
6	8,99977	1	16	49
7	9,56968	1	12	23
8	14,9992	0	14	42
9	11,0000	0	8	56
10	4,99981	1	12	32
11	24,9756	0	17	41
12	20,3992	0	17	26
13	25,0006	0	14	24
14	13,9796	0	14	34
15	3,50012	0	12	53
16	4,9998	0	14	20
17	10,0001	0	16	29
18	14,9992	0	16	48
19	5,82990	0	13	22
20	9,10023	0	13	35
21	11,2503	1	17	55
22	12,9993	0	12	43
23	7,99966	1	14	35
24	4,28020	1	12	56
25	7,875	0	16	39
26	6,9441	1	12	25
27	24,9756	1	17	28
28	10,0001	0	14	42
29	13,7494	0	14	41

30	5,62488	1	13	21
31	14,6699	0	14	36
32	24,9756	0	16	40
33	11,9999	0	12	32
34	14,9992	0	18	36
35	5,7747	0	16	25
36	12,6505	1	17	36
37	11,710	1	16	64
38	12,9993	0	12	26
39	12,4996	1	15	27
40	22,4996	0	16	44
41	11,2503	1	16	25
42	15,3804	0	16	55
43	7,69984	1	13	27
44	11,8401	0	12	44
45	4,99981	0	14	20
46	6,49998	0	12	46
47	8,99977	0	8	47
48	3,35013	0	12	18
49	4,50010	1	12	25
50	10,5265	1	14	32
51	3,50012	0	9	62
52	14,9994	0	16	34
53	8,50028	1	12	43
54	9,16965	1	12	62
55	9,60035	0	9	48
56	6,87505	1	8	36
57	13,9502	1	18	38
58	10,0001	0	14	42
59	3,75016	1	12	59
60	7,99966	1	12	26
61	11,6662	1	12	61
62	8,88531	1	8	43
63	4,00002	0	12	24
64	5,49973	1	12	39
65	3,35013	0	16	25

თავი 3. კლასიკური მოდელი

რეგრესიის კლასიკური მოდელი ითვალისწინებს შემდეგს:

1. რეგრესიის მოდელი წრფივია კოეფიციენტების მიმართ და შეიცავს შემთხვევით წევრს.

რეგრესიის მოდელი რომ წრფივი არ იყოს კოეფიციენტების მიმართ, მაშინ უმცირეს კვადრატთა მეთოდის გამოყენება მიუღებელი იქნებოდა.

მაგალითად, ექსპონენციალური ფუნქცია

$$Y_i = e^{S_0} X_1^{S_1} e^{v_i} ,$$

სადაც e ნატურალური ლოგარითმის ფუძეა, შეიძლება გარდავქმნათ და თუ განვიხილავთ განტოლების ორივე მხარის ლოგარითმს, მივიღებთ:

$$\ln(Y_i) = S_0 + S_1 \ln(X_i) + v_i \quad (3.1)$$

თუ შემოვიტანთ აღნიშვნებს:

$$\ln(Y_i) = Y_i^* \quad \text{და} \quad \ln(X_i) = X_i^*$$

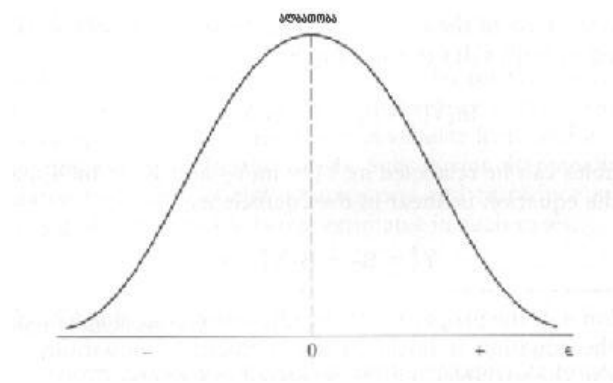
მაშინ განტოლება (3.1) შეიძლება ჩავწეროთ შემდეგი სახით:

$$Y_i^* = S_0 + S_1 X_i^* + v_i \quad (3.2)$$

უმცირეს კვადრატთა მეთოდის გამოყენება (3.2) განტოლების მიმართ ისევ მისაღები იქნება, რადგან განტოლება (3.2) წრფივია კოეფიციენტების მიმართ.

2. შემთხვევითი წევრის საშუალო მნიშვნელობა ნულის ტოლია.

ეკონომეტრისტები რეგრესიის განტოლებას უმატებენ ალბათურ (შემთხვევით) წევრს, რომლის მნიშვნელობა თითოეულ დაკვირვებაში სავსებით შემთხვევითაა განსაზღვრული, რაც კარგად ჩანს გრაფიკიდან.



ნახ. 3.1. შემთხვევითი წევრის განაწილება (საშუალო მნიშვნელობა ნულის ტოლია)

განვიხილოთ რეგრესიის ტიპიური განტოლება:

$$Y_i = S_0 + S_1 X_i + v_i$$

ვთქვათ, შემთხვევითი წევრის მნიშვნელობა 3-ის ტოლია ნაცვლად ნულისა, ხოლო მათემატიკური მოლოდინი $E(v_i - 3) = 0$. თუ განტოლებას დავუმატებთ და გამოვაკლებთ 3-ს, მივიღებთ:

$$Y_i = (S_0 - 3) + S_1 X_i + (V_i - 3)$$

ეს განტოლება შეიძლება ჩაიწეროს აღნიშვნების გათვალისწინებით შემდეგი ფორმით:

$$Y_i = S_0^* + S_1 X_i + V_i^*$$

3. დამოუკიდებელი ცვლადები კორელირებულები არ არიან შემთხვევით წევრთან

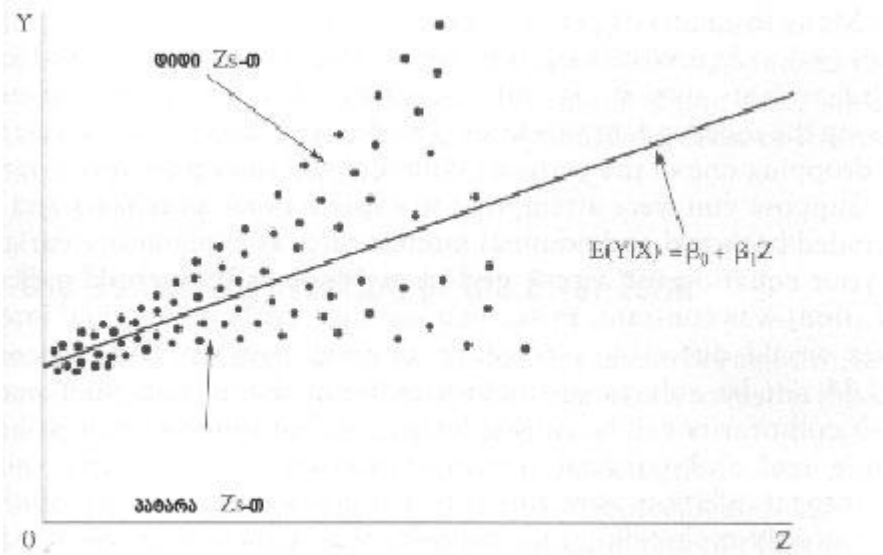
დამოუკიდებელ ცვლადებზე დაკვირვების შედეგები განსაზღვრულია დამოკიდებელი ცვლადსა და შემთხვევითი წევრისაგან დამოუკიდებლად, თუ X და შემთხვევითი წევრი დადებითად კორელირებულნი არიან. მაშინ შეფასებულ კოეფიციენტს უფრო დიდი მნიშვნელობა ექნება, რადგან უმცირეს კვადრატთა მეთოდი შეცდომით მიაწერდა Y -ში გამოწვეულ ვარიაციას V -ს, ნაცვლად X -სა.

4. შემთხვევით წევრზე დაკვირვებების შედეგები ერთმანეთთან კორელირებულები არ არიან

ეს რომ ასე არ იყოს, მაშინ უმცირეს კვადრატთა მეთოდით დამოუკიდებელი ცვლადების კოეფიციენტების ზუსტი შეფასების მიღება შედარებით გაძნელებოდა. ეს მოსაზრება განსაკუთრებით მნიშვნელოვანია დროითი მწკრივების შესწავლისას.

5. შემთხვევითი წევრის ვარიაცია (დისპერსია) მუდმივია

ეს შემთხვევა კარგად არის გამოსახული ნახ.3..-ზე



ნახ.3.2. შემთხვევითი წევრის ვარიაციის ცვლილება

6. არც ერთი დამოუკიდებელი ცვლადი არ არის დანარჩენი დამოუკიდებელი ცვლადების წრფივი ფუნქცია

ორ ცვლადს შორის კოლინეარობა განიხილება როგორც ერთი ცვლადის არსებობა ან როგორც ცვლადი, რომელიც მეორისგან განსხვავებულია მუდმივი

სიდიდით რის გამოც უმცირეს კვადრატთა მეთოდის გამოყენებისას შეფასების პროცედურა ერთ ცვლადს მეორისაგან ვერ განასხვავებს.

7. შემთხვევითი წევრი ნორმალური განაწილების კანონს ემორჩილება

ეს მომენტი მოთხოვნილი არ არის უმცირეს კვადრატთა მეთოდის გამოყენებისას მისი მთავარი გამოყენება ჰიპოთეზების ტესტირებისას ხდება.

3.1. გაუს-მარკოვის თეორემა

გაუს-მარკოვის თეორემა უმცირეს კვადრატთა მეთოდის ორ მნიშვნელოვან თვისებას ამტკიცებს. ეს თეორემა დამტკიცებულია ეკონომეტრიკის თითქმის ყველა სახელმძღვანელოში და მისი დამტკიცება სტანდარტულ კითხვას წარმოადგენს ეკონომეტრიკის გამოცდაზე. რეგრესიული ანალიზის გამოყენებისას უფრო მნიშვნელოვანია ვიცოდეთ, თუ რაში მდგომარეობს ამ თეორემის არსი და რის დამტკიცებას გულისხმობს ეს თეორემა.

გაუს-მარკოვის თეორემა აყალიბებს შემდეგს: უმცირეს კვადრატთა მეთოდით შეფასებულ პარამეტრებს აქვთ მინიმალური დისპერსია "საუკეთესო შეფასება" ყველა წრფივ გადაუადგილებად შეფასებათა კლასში. როდესაც ვამბობთ "საუკეთესო", აუცილებელია შევნიშნოთ, რომ თითოეულს აქვს უმცირესი შესაძლო დისპერსია და მათემატიკური მოლოდინი.

$$E(\hat{s}_k) = s_k \quad (k=1,2,\dots,K).$$

ყველა, ზემოთ აღნიშნული თვისებების საფუძველზე შეიძლება ჩამოყალიბდეს რეგრესიის განტოლების კოეფიციენტების შეფასება უმცირეს კვადრატთა მეთოდით, რომელიც შემდეგი თვისებების მატარებელია:

1. $E(\hat{s}_k) = s_k \quad (k=1,2,\dots,K)$ რაც ნიშნავს, რომ უმცირეს კვადრატთა მეთოდის გამოყენებით შეფასებული რეგრესიის განტოლების კოეფიციენტები კონცენტრირებულნი არიან ერთობლიობის ჭეშარიტი პარამეტრების ირგვლივ;

2. მათ აქვთ ვარიაციის (დისპერსიის) მინიმალური მნიშვნელობა;

3. თუ შერჩევის მოცულობა გაიზრდება, ვარიაციის სიდიდე შემცირდება და თითოეული კოეფიციენტის შეფასება უკეთ მიუახლოვდება კოეფიციენტის ჭეშარიტ მნიშვნელობას;

4. ისინი ნორმალური განაწილების კანონს ემორჩილებიან ე.ი. სტატისტიკური ტესტი, როელსაც საფუძვლად უდევს ნორმალური განაწილების კანონი შეიძლება გამოვიყენოთ \hat{s}_k შეფასებისას.

3.2. სტანდარტული ეკონომეტრიკული შენიშვნები

ერთობლიობის ორი პარამეტრი საშუალო სიდიდე და ვარიაცია აღწერს \hat{s}_k კოეფიციენტების განაწილებას. $E(\hat{s}_k)$ -ს მნიშვნელობა ტოლია კოეფიციენტის მნიშვნელობისა, თუ შეფასება გადაუადგილებადია. კვადრატული ფესვის მნიშვნე-

ლობა ვარიაციიდან სტანდარტული გადახრის მნიშვნელობის ტოლია და ვარიაცია $\hat{\sigma}^2 (\hat{S}_k)$ ერთობლიობის პარამეტრია, რომელზედაც დაკვირვება პრაქტიკულად შეუძლებელია და მის ნაცვლად ვიყენებთ $\hat{\sigma}^2 (\hat{S}_k)$ ან $s^2 (\hat{S}_k)$ სიდიდეებს. უნდა შევნიშნოთ, რომ ჭეშმარიტი S_k -ს ვარიაცია ნულის ტოლია მაშინ, როდესაც შეფასებული კოეფიციენტის ვარიაციაზე დაკვირვება შესაძლებელია. ჭეშმარიტი ვარიაცია დაკვირვებას არ ექვემდებარება. ცხრილში 3.1. მოცემულია გენერალური ერთობლიობისა და შერჩევის პარამეტრებისათვის გამოყენებული აღნიშვნები.

ცხრილი 3.1

დაკვირვებას დაუქვემდებარებელი ერთობლიობის პარამეტრების ჭეშმარიტი მნიშვნელობები		პარამეტრების შეფასება შერჩევაზე დაკვირვების შედეგების მიხედვით	
პარამეტრის დასახელება	სიმბოლო	პარამეტრის დასახელება	სიმბოლო
რეგრესიის კოეფიციენტი	S_k	შეფასებული რეგრესიის კოეფიციენტი	\hat{S}_k
შესაფასებელი რეგრესიის კოეფიციენტის მოსალოდნელი მნიშვნელობა	$E(S_k)$	შესაფასებელი შემთხვევითი წევრის ვარიაცია	s^2 ან $\hat{\sigma}^2$
შემთხვევითი წევრის ვარიაცია	$var(e)$ ან σ^2	რეგრესიის განტოლების სტანდარტული შეცდომა	s
შემთხვევითი წევრის სტანდარტული გადახრა	σ	შეფასებული ვარიაცია	$s^2 (\hat{S}_k)$ ან $var(\hat{S}_k)$
შეფასებული კოეფიციენტის ვარიაცია	$\hat{\sigma}^2 \sim \hat{S}_k \cdot \sigma^2$ ან $var(\hat{S}_k)$	შეფასებული კოეფიციენტის სტანდარტული შეცდომა	$\hat{\sigma} (\hat{S}_k)$
შემთხვევითი წევრი	V_i	სხვაობები	e_i

რეზიუმე

1. რეგრესიის კლასიკური მოდელის ძირითადი მახასიათებლებია:

ა/ რეგრესიის მოდელი წრფივია კოეფიციენტების მიმართ და შეიცავს შემთხვევით წევრს;

ბ/ შემთხვევითი წევრის საშუალო მნიშვნელობა ნულის ტოლია;

გ/ დამოუკიდებელი ცვლადები კორელირებულნი არ არიან შემთხვევით წევრთან;

დ/ შემთხვევითი წევრზე დაკვირვებების შედეგები ერთმანეთთან კორელირებულები არ არიან;

ე/ შემთხვევითი წევრის ვარიაცია (დისპერსია) მუდმივია;

ვ/ არც ერთი დამოუკიდებელი ცვლადი არ არის დანარჩენი დამოუკიდებელი ცვლადების წრფივი ფუნქცია;

ზ/ შემთხვევითი წევრის განაწილება ნორმალური განაწილების კანონს ემორჩილება.

2. უმცირეს კვადრატთა მეთოდის გამოყენებით მიღებულ შეფასებას ორი მნიშვნელოვანი თვისება გააჩნია: გადაუადგილებადობა და ვარიაციის მინიმუმი.

3. რადგან უმცირეს კვადრატთა მეთოდის გამოყენებით მიღებული \hat{S}_k -ს შეფასება საუკეთესო შეფასებაა წრფივ შეფასებათა გადაუადგილებად კლასში, მას აქვს საინტერესო თვისებები, უპირველესად: \hat{S}_k -ს ვარიაცია დამოკიდებულია დაკვირვებათა რიცხოვნებაზე. მისი მნიშვნელობის გაზრდა ან შემცირება დამოკიდებულია შერჩევაში დაკვირვებათა რიცხოვნების გაზრდასა ან შემცირებაზე.

დამოუკიდებელი სამუშაო

3.1. გაუს-მარკოვის თეორემის მიხედვით უმცირეს კვადრატთა მეთოდით მიღებული შეფასება საუკეთესოა გადაუადგილებად წრფივ შეფასებათა კლასში მინიმალური ვარიაციის თვალსაზრისით.

ა/ თუ მოგიწევთ არჩევანის გაკეთება განტოლების კოეფიციენტების შესაფასებლად რომელს აირჩევდით: გადაადგილებადს არამინიმალური დისპერსიით, თუ გადაუადგილებადს მინიმალური დისპერსიით? ახსენით მიზეზი.

ბ/ არის თუ არა ისეთი სიტუაცია, რომ შეცვალოთ თქვენი გადაწყვეტილება ა/ პუნქტში?

გ/ გაქვთ თუ არა მოსაზრება კოეფიციენტების შესაფასებელი გზის მოსაძებნად?

3.2. 1993 წელს ედუარდ სოდერსმა (Eduard M. Souders. "Stock Prices and Wall Street Weather", American Economic Review, December 1993, pp.1337-1346) გამოაქვეყნა სტატია, რომელშიც აღწერა ფასიანი ქაღალდების ბაზრის ცვლილების დამოკიდებულება უოლ სტრიტის "ამინდზე". მან გამოიყენა ყოველდღიური მონაცემები 1962-1989 წლების მიხედვით და შეაფასა განტოლება (კოეფიციენტების სტანდარტული შეცდომები მოცემულია ფრჩხილებში).

$$DJ_t = \hat{S}_0 + 0.10R_{t-1} + 0.0010J_t - 0.017M_t + 0.0005C_t$$

$$(0.1) \quad (0.0006) \quad (0.004) \quad (0.0002)$$

$$n=6.911 \quad \bar{R}^2 = 0.02$$

სადაც:

DJ_t - დოუ ჯონსის ინდექსის ცვლილება t დღეს (პროცენტი);

R_t - ყოველდღიური მოგება ან წაგება კაპიტალის ნაზრდიდან;

J_t - ფიქტიური ცვლადი, რომელიც 1-ის ტოლია, თუ i -რი დღე იანვრის თვეში იყო, სხვა დანარჩენ შემთხვევაში ნულის ტოლია;

M_t - ფიქტიური ცვლადი, რომლის მნიშვნელობა 1-ის ტოლია, თუ i -რი დღე ორშაბათი იყო: სხვა დანარჩენ შემთხვევაში ნულის ტოლია;

CC_t - ცვლადი, რომელიც 1-ის ტოლია, თუ ღრუბლებმა 20 ან მასზე ნაკლები პროცენტი დაფარა 1-ის ტოლია და სხვა დანარჩენ შემთხვევაში ნულის ტოლია.

ა/ გამოქვეყნებულ რეგრესიის განტოლებაში მუდმივი წევრის მნიშვნელობა არ არის შეტანილი, რატომ? (ახსენით პასუხი);

ბ/ თუ განტოლებას ყურადღებით გავანალიზებთ, R ლაგური ცვლადია. ეთანხმებით თუ არა ამ აზრს?

გ/ არის თუ არა C ცვლადი ფიქტიური? რა დანიშნულება აქვს ამ ცვლადს განტოლებაში?

3.3. ვ. ბრაუნმა და ტ. ფაინგენმა (W. Brown and T. Finegan, "Labor force Participation and Unemployment", Employment Policy and Labor Markets (Berkeley), University of California Press, 1965) შეაფასეს 78 ქალაქი სამუშაო ძალისა და დაუსაქმებლობის მიხედვით და ჩაწერეს რეგრესიის განტოლება (ფრჩხილებში მოცემულია თითოეული კოეფიციენტისათვის საშუალო სტანდარტული გადახრა):

$$\hat{L}_i = 94.2 - 0.24U_i + 0.20E_i - 0.69I_i + 0.065S_i + 0.002C_i - 0.80D_i$$

(0.08) (0.06) (0.16) (0.18) (0.03) (0.53)

$n = 78; \bar{R}^2 = 0.51$

სადაც:

L_i - დასაქმებული სამუშაო ძალა i -ურ ქალაქში (მამაკაცების ასაკი 25-54 წ), პროცენტი;

U_i - უმუშევრობის დონე i -ურ ქალაქში, პროცენტი;

E_i - ხელფასიდან საშუალო წლიური შემოსავალი i -ურ ქალაქში, დოლარი;

I_i - სხვა სახის საშუალო წლიური შემოსავალი i -ურ ქალაქში, დოლარი

S_i - საშუალო სრული განათლება i -ურ ქალაქში, წელი;

C_i - სამუშაო ძალა i -ურ ქალაქში შავკანიანების სახით, პროცენტი;

D_i - ფიქტიური ცვლადი 1-ის ტოლია, თუ ქალაქი სამხრეთს ეკუთვნის, სხვა დანარჩენ შემთხვევებში 0-ის ტოლია.

ა/ გააკეთეთ C და D ცვლადებთან კოეფიციენტების ინტერპრეტაცია.

ბ/ რამდენადაა მოსალოდნელი კოლინეარულობა ამ განტოლებაში?

გ/ გაუკეთეთ კომენტარი გამოთქმას: " მე ვიცი, რომ მიღებული შედეგი არ არის საუკეთესო შეფასება წრფივ გადაუადგილებად შეფასებათა კლასში, რადგან საშუალო მნიშვნელობა 94.2 პროცენტი ძალიან მაღალია".

თავი 4. ჰიპოთეზების ტესტირება

4.1. რა ნიშნავს ჰიპოთეზის ტესტირება?

ვიდრე განტოლებას შევაფასებთ, აუცილებელია ჩამოვაყალიბოთ ჰიპოთეზა, რომლის დროსაც უნდა გავიაზროთ რა არის ჰემმარიტი და რა არ არის ჰემმარიტი. ასეთი მიდგომა ალტერნატიული და ნულოვანი ჰიპოთეზების სახით ყალიბდება.

ნულოვანი ჰიპოთეზა ტიპური მტკიცებაა რეგრესიის კოეფიციენტების დიაპაზონის არსებობის მოსალოდნელობის შესახებ, თუ თეორია კორექტული არ არის.

ალტერნატიული ჰიპოთეზა გამოიყენება იმ შემთხვევაში, როდესაც საჭიროა განისაზღვროს რეგრესიის კოეფიციენტების ის დიაპაზონი, რომელსაც ადგილი ექნებოდა, როდესაც თეორია კორექტულია.

ნულოვან ჰიპოთეზას ავლნიშნავთ H_0 და თუ მაგალითად, მოსალოდნელია რეგრესიის კოეფიციენტის უარყოფითი ნიშანი, ნულოვანი ჰიპოთეზა იქნება:

$$H_0 : \beta \geq 0$$

რაც ნიშნავს შემდეგს: ეს არის კოეფიციენტის ის მნიშვნელობა, რომელსაც არ ველით.

ალტერნატიული ჰიპოთეზა აღინიშნება: H_A და თუ მაგალითად, მოსალოდნელია რეგრესიის კოეფიციენტის უარყოფითი ნიშანი, ალტერნატიული ჰიპოთეზა იქნება:

$$H_A : \beta < 0$$

რაც ნიშნავს: კოეფიციენტის მოსალოდნელი მნიშვნელობა ჭეშმარიტი იქნება.

ნულოვანი და ალტერნატიული ჰიპოთეზების გამოყენება შეიძლება რეგრესიის კოეფიციენტის ნულოვანი ან ნულისაგან განსხვავებული მნიშვნელობების დასადგენად, მაშინ ნულოვანი და ალტერნატიული ჰიპოთეზები ჩაიწერება შემდეგი სახით :

$$H_0 : \beta = 0$$

$$H_A : \beta \neq 0$$

4.2. რეგრესიის კოეფიციენტების მიმართ ჰიპოთეზის შემოწმება

თეორიიდან გამომდინარე, ჰიპოთეზის ფორმულირება ყოველთვის წინ უსწრებს ექსპერიმენტს, რომლის ტესტირებაც ჩვენი ინტერესების არეს მიეკუთვნება. მაგალითად, შეიძლება ჩავთვალოთ, რომ ეკონომიკაში საერთო ინფლაციის ტემპი- I დამოკიდებულია ხელფასების ზრდით გამოწვეულ ინფლაციის ტემპზე- w . ეს კავშირი შეიძლება აისახოს დამოკიდებულებით:

$$I = r + sw + v$$

სადაც:

I – ინფლაცია;

r, s – პარამეტრებია;

v – შემთხვევითი წევრია.

გვინდა შევამოწმოთ ჰიპოთეზა: საერთო ინფლაცია ხელფასების ზრდით გამოწვეული ინფლაციის ტოლია (შემთხვევითი წევრით შემოტანილ ეფექტს არ ვითვალისწინებთ). ამ პირობებში ნულოვანი ჰიპოთეზა, რომლის შემოწმებასაც ვაპირებთ, მდგომარეობს შემდეგში:

$$H_0 : \beta = 1$$

ხოლო ალტერნატიული ჰიპოთეზა მდგომარეობს შემდეგში:

$$H_1 : \beta \neq 1$$

განვიხილოთ კიდევ მაგალითი: კვების პროდუქტებზე მოთხოვნის ფუნქცია ჩაიწერება განტოლებით:

$$y = r + sx + v$$

სადაც:

y – მოთხოვნის სიდიდე კვების პროდუქტებზე;

x – შემოსავალია.

ვთქვათ, ვითვალისწინებთ, რომ კვების პროდუქტებზე მოთხოვნა შემოსავალზეა დამოკიდებული. ნულოვან ჰიპოთეზად შეიძლება მივიღოთ მტკიცება იმის შესახებ, რომ y არ არის დამოკიდებული x – ის მნიშვნელობებზე,

$$\text{ე. ი. : } H_0 : \rho = 0$$

ალტერნატიული ჰიპოთეზა კი მდგომარეობს შემდეგში:

$$H_A : \rho \neq 0$$

ე. ი. y განიცდის x –ის გავლენას.

ასევე შეიძლება მივუდგეთ F კოეფიციენტსაც.

შესაძლებელია, გამოვიყენოთ ზოგადი მიდგომა და ნულოვანი ჰიპოთეზა ჩამოვაცალიბოთ ისე, რომ შევამოწმოთ S –ს ტოლობა რაიმე ფიქსირებულ მნიშვნელობასთან, ხოლო ალტერნატიული ჰიპოთეზის მიხედვით - ნულოვანის უარყოფა, ე. ი.

$$H_0 : S = S_0$$

$$H_A : S \neq S_0$$

თუ ნულოვანი ჰიპოთეზა ჭეშმარიტია, მაშინ S –ს შეფასებებს, რომლებიც ანალიზის დროს მიიღება, ექნებათ განაწილება მათემატიკური ლოდინით S_0 - და დისპერსია $\frac{\sigma^2}{n \text{ var}(x)}$. შემდეგ დაუშვებთ, რომ შემთხვევით წევრის - V განაწილება ნორმალურია. აქედან გამომდინარეობს, რომ S –ს განაწილებაც

ნორმალურია და $t_s = \frac{\bar{t}_v}{\sqrt{n \text{ var}(x)}}$. თუ გავითვალისწინებთ ნორმალური განაწილების სტრუქტურას, S – პარამეტრის შეფასებები მოთავსებულია $\pm 2t$ - საზღვრებში, თუ ჭეშმარიტია ჰიპოთეზა $H_0 : S = S_0$.

როგორ შეიძლება განვსაზღვროთ ნულოვანი ჰიპოთეზის ან ალტერნატიული ჰიპოთეზის ჭეშმარიტება?

ცალსახად, ამ კითხვაზე პასუხი არ არსებობს. ეკონომიკაში კრიტიკულ დონედ მიღებულია 5% -იანი ან 1%-ანი მნიშვნელოვნებების დონეები. თუ 5%-ანი კრიტიკული დონე შერჩეულია, მაშინ 1%-ან დონეზე გადასვლა ხდება იმ შემთხვევაში, თუ ჭეშმარიტი ნულოვანი ჰიპოთეზის დროს S –ს ექსტრემალური მნიშვნელობის მიღების ალბათობა 5%-ზე ნაკლებია, ამ შემთხვევაში ნულოვანი ჰიპოთეზა უარყოფილია 5%-ანი მნიშვნელოვნების დონით. ამ შემთხვევას ადგილი აქვს მაშინ, თუ S პარამეტრი S_0 –ს ჩამორჩება მნიშვნელოვანად. ალბათობა იმისა, რომ S –ს მნიშვნელობა აღემატება საშუალო სიდიდეს არა უმეტეს 1.96 t თი :

$$Z > 1.96 \text{ ან } Z < -1.96$$

$$Z = \frac{S - S_0}{t_s}$$

ნულოვანი ჰიპოთეზა უარყოფილი არ იქნება, თუ

$$-1.96 < Z < 1.96$$

ანუ

$$-1.96 < \frac{S - S_0}{t_s} < 1.96$$

$$\begin{aligned} \text{ე.ი. } & -1.96t'_s < S - S_0 < 1.96t'_s \\ & S_0 - 1.96t'_s < S < S_0 + 1.96t'_s \end{aligned}$$

S -თვის მიღებულია მნიშვნელობათა სიმრავლე, რომლის მიხედვითაც ნულოვანი ჰიპოთეზა $H_0 : S = S_0$ უარყოფილი არ იქნება. ამ სიმრავლეს S პარამეტრისათვის 5%-ანი მნიშვნელობის დონით ჰიპოთეზის მისაღები არე ეწოდება.

რატომ არის კვლევის პროცესში ორი დონე 1%-ანი და 5%-ანი გამოყენებული მიზეზი შემდეგია: მოიძებნოს პირველი და მეორე რიგის შეცდომებს შორის ბალანსი, რადგან:

1. I რიგის შეცდომასთან საქმე გვაქვს მაშინ, როდესაც უარყოფილია ჭეშმარიტი ნულოვანი ჰიპოთეზა;

2. II რიგის შეცდომასთან საქმე გვაქვს მაშინ, როდესაც მცდარი ჰიპოთეზა უარყოფილი არ არის.

ცხადია, I რიგის შეცდომის დაშვების რისკი მით უფრო მეტია, რაც უფრო დაბალია კრიტიკული ალბათობა.

თუ ვიყენებთ 5%-ან მნიშვნელოვნების დონეს, მაშინ ჭეშმარიტი ჰიპოთეზის უარყოფის შემთხვევები 5%-ია, ხოლო თუ მნიშვნელოვნების დონე 1%-ია, მაშინ I რიგის შეცდომის შემთხვევა 1%-ს შეესაბამება. ე. ი. 1%-ანი მნიშვნელოვნების დონე უფრო საიმედოა.

თუ ნულოვანი ჰიპოთეზა მცდარია, მაშინ რაც უფრო მაღალია მნიშვნელოვნების დონე, მით უფრო ფართოა ჰიპოთეზის მიღების არე და მით უფრო მეტია ჰიპოთეზის მიღების ალბათობაც, ე. ი. II რიგის შეცდომის დაშვების რისკი მაღალია.

განხილულ სიტუაციებში ვთვლიდით, რომ S პარამეტრის სტანდარტული გადახრა ცნობილია. პრაქტიკულად, ეს დაშვება ნაკლებად რეალურია, ამიტომ ვიყენებთ t სტატისტიკას, რომელიც გამოითვლება ფორმულით:

$$t = \frac{S - S_0}{t'_s} \quad (4.1)$$

სადაც:

t'_s - S -ს სტანდარტული შეცდომაა

t -ს კრიტიკული მნიშვნელობა აღინიშნება t_c -ით და პირობა იმისა, რომ რეგრესიის შეფასება არ მიგვიყვანს ნულოვანი ჰიპოთეზის უარყოფასთან $S = S_0$ ჩაიწერება შემდეგი სახით:

$$-t_c < \frac{S - S_0}{t'_s} < t_c \quad (4.2.)$$

მაგალითი: აშშ-ში კვებაზე დანახარჯების ფუნქცია 25 წლის მონაცემების მიხედვით (1959-1983წ.წ.) წარმოდგენილია რეგრესიის განტოლებით (მნიშვნელობები გაზომილია მლრდ დოლარით მუდმივი ფასებით) (Studemund A. H.):

$$\hat{y}_i = 55.3 + 0.093x_i$$

$$S_0 = 55.3; t'_{s_0} = 2.4; s = 0.093; t'_s = 0.003$$

ვთქვათ, ჰიპოთეზა: $H_0 : S_0 = 0$ (კვებაზე დანახარჯები დამოკიდებულია შემოსავალზე). მისი შესაბამისი სტატისტიკა:

$$t = \frac{s - s_0}{t'_s} = \frac{b - 0}{t'_s} = \frac{0.093}{0.003} = 31.0$$

შერჩევა მოიცავს 25 დაკვირვებას და ვაფასებთ ორ პარამეტრს. ე.ი. თავისუფლების ხარისხია 23. 5%- იანი მნიშვნელოვნების დონისათვის t -ს კრიტიკული მნიშვნელობაა 2.069.

მაგრამ პირობა (4.2) არ სრულდება, ე.ი. ნულოვანი ჰიპოთეზა უარყოფილია, ე.ი. შემოსავალს კვების დანახარჯებზე გავლენა აქვს, რადგან ჰიპოთეზის შესამოწმებლად ვიყენებთ 5% - იანი მნიშვნელოვნების დონეს, ბუნებრივია, არსებობს პირველი რიგის შეცდომის დაწვების 5%- იანი რისკი. ამ რისკის შემცირების საშუალებაა 1%-იანი მნიშვნელოვნების დონის გამოყენება. მისი შესაბამისი t -ს კრიტიკული მნიშვნელობა თავისუფლების ხარისხით 23 ტოლია 2.807. მაგრამ ამ შემთხვევაშიც, ნულოვანი ჰიპოთეზა უარყოფილი იქნება.

მაგალითი: შევაფასოთ რეგრესიის მოდელი, რომელიც წარმოადგენს დამოკიდებულებას საერთო ინფლაციასა და ხელფასის გაზრდით გამოწვეულ ინფლაციას შორის (შერჩევაში დაკვირვების რაოდენობაა 20).

რეგრესიის განტოლებაა:

$$I = 1.21 + 0.82w$$

$$s_0 = -1.21; t'_{s_0} = 0.05; s = 0.82; t'_s = 0.10$$

სადაც:

I -საერთო ინფლაცია, % ;

w -ხელფასით გამოწვეული ინფლაცია, \$

გამოვიკვლიოთ ჰიპოთეზა გრძელვადიან პერიოდში. ხელფასის გაზრდით გამოწვეული ინფლაცია საერთო ინფლაციის ტოლია, ე.ი. $H_0 : S = 1.0$

მისი შესაბამისი სტატისტიკა:

$$t = \frac{s - s_0}{t'_s} = \frac{0.82 - 1.00}{0.10} = -1.8$$

თავისუფლების ხარისხია 18. 5%- იანი მნიშვნელოვნების დონისათვის $t_c = 2.101$, ნულოვანი ჰიპოთეზა მიღებული იქნება.

4.3. ნდობის ინტერვალი

აღნიშნული იყო, რომ ჰიპოთეზა წინ უსწრებს ემპირიულ კვლევას, თუმცა ეს სრულად არ ეთანადება სინამდვილეს. ხშირად, ჰიპოთეზა და ექსპერიმენტი ერთდროულად მოქმედებენ. ამ შემთხვევაში წამოიჭრება შემდეგი საკითხი: რამდენად განსხვავდება ჰიპოთეტიკური შედეგი ექსპერიმენტის შედეგისაგან. ამ კითხვაზე პასუხის შემდეგ ან ვუარყოფთ ან მივიღებთ ჰიპოთეზას. (4.2) დამოკიდებულების მიხედვით რეგრესიის კოეფიციენტი \hat{S} და ჰიპოთეტიკური S -ს მნიშვნელობა არათავსებადი აღმოჩნდება, თუ შესრულდება პირობა: S -ს მნიშვნელობა აკმაყოფილებს ორმაგ უტოლობას:

$$\hat{S} - t(\hat{S}) * t_c < S < \hat{S} + t(\hat{S}) * t_c \quad (4.3)$$

ნებისმიერი, S -ს ჰიპოთეტიკური მნიშვნელობა, რომელიც აკმაყოფილებს (4.3) პირობას თავსებადი იქნება მის შეფასებასთან \hat{S} -სთან. ამ მნიშვნელობების სიმრავლე განსაზღვრული როგორც ინტერვალი, უტოლობის ზედა და ქვედა საზღვრებს შორის, წარმოადგენს S -ს მნიშვნელობისათვის ნდობის ინტერვალს. შევნიშნოთ, რომ ნდობის ინტერვალის შუაში ძვეს თვითონ \hat{S} -მნიშვნელობა და მისგან ინტერვალის საზღვრები თანაბარი მანძილებითაა დაშორებული. გარდა ამისა, t_c -ს მნიშვნელობა დამოკიდებულია მნიშვნელოვნების დონის შერჩევაზე და საზღვრებიც ამ შერჩევაზე იქნება დამოკიდებული. თუ ავიღებთ 5%-იან მნიშვნელოვნების დონეს, მაშინ მისი შესაბამისი ნდობის ინტერვალი შეესაბამება 95%-იან ინტერვალს. თუ შერჩეული იქნება 1%-იანი მნიშვნელოვნების დონე. მისი შესაბამისი ინტერვალი 99% იქნება და ა.შ.

რადგან t_c -ის მნიშვნელობა 1%-იანი მნიშვნელოვნების დონისათვის მეტი იქნება, ვიდრე 5%-იანი მნიშვნელოვნების დონისათვის (თავისუფლების ნებისმიერი ხარისხის მოცემული მნიშვნელობისთვის, ამიტომ 99% ინტერვალი უფრო ფართო იქნება, ვიდრე 95% ინტერვალი. ნდობის ინტერვალი შესაფასებელი კოეფიციენტისათვის ჩაიწერება გამოსახულებით:

$$\hat{S} \pm t_c \cdot \hat{S} \cdot \sqrt{\frac{1}{n}}$$

მაგალითი. გამოვიყენოთ რეგრესიის განტოლება:

$$\hat{Y}_i = 102.192 - 9075C_i + 0.354P_i + 1.288I_i$$

(20.53) (0.0727) (0.543) - სტანდარტული გადახრების მნიშვნელობები
 $t = -4.42 \quad 4.88 \quad 2.37$, t - სტატისტიკის მნიშვნელობები
 $n = 33 \quad \bar{R}^2 = 0.579$

მონაცემების მიხედვით P მაჩვენებლისათვის $\hat{S}_p = 0.3547$; $t(\hat{S}_p) = 0.0727$; თავისუფლების ხარისხი $DF = 29$; $t_c = 1.699$ (ცხრილის მიხედვით 10% -იანი ნდობის ინტერვალი იქნება:

$$\hat{S}_p = 0.3547 \pm 1.699 \cdot 0.0727 = 0.3547 \pm 0.1235$$

ე.ი. კოეფიციენტის ჰემარიტი მნიშვნელობა მოთავსებულია 0.2312-სა და 0.4782-ს შორის.

ნდობის ინტერვალის გამოყენება ეკონომეტრიკაში მეტად მნიშვნელოვანია პროგნოზირებისას.

4.4. ცალმხრივი და ორმხრივი t ტესტის გამოყენება

t ტესტთან მუშაობა მოიცავს შემდეგ ეტაპებს:

1. ნულოვანი და ალტერნატიული ჰიპოთეზების ჩამოყალიბება;
2. მნიშვნელოვნების დონეებისა და შესაბამისი t -ს კრიტიკული მნიშვნელობის შერჩევა;
3. ეკონომეტრიკის პროგრამული უზრუნველყოფის გამოყენებით რეგრესიის განტოლებასა და t -ს შეფასებული მნიშვნელობების ჩაწერა;
4. გამოთვლების შედეგად მიღებული t -ს მნიშვნელობის შედარება t -ს კრიტიკულ მნიშვნელობასთან და ჰიპოთეზის მიღება ან უარყოფა.

ცალმხრივი t - ტესტის გამოყენება იმითაა განსაზღვრული, რომ რეგრესიის კოეფიციენტს აქვს წინასწარ, თეორიით განსაზღვრული ნიშანი. თუ ველოდებით, რომ კოეფიციენტის ნიშანი დადებითია და მივიღეთ უარყოფითი ნიშანი კოეფიციენტისათვის, ბუნებრივია, ეს ნიშნავს შემდეგს: უარყოფით, რომ ჭეშმარიტია კოეფიციენტის უარყოფითი ნიშანი. მეორეს მხრივ, თუ ჩვენ ველოდით, რომ კოეფიციენტს ექნება დადებითი ნიშანი და მივიღეთ, რომ მას უნდა ჰქონდეს უარყოფითი ნიშანი, ეს შემთხვევა ბადებს დამატებით მოსაზრებებს. მაგრამ თუ შეფასებული კოეფიციენტი დადებითი ნიშნის მატარებელია და მისი მნიშვნელობა ნულთან ახლოსაა, მაშინ ცალმხრივი t ტესტი გამოიყენება იმისათვის, რომ განვსაზღვროთ რამდენად ახლოსაა ეს კოეფიციენტი ნულთან.

ორმხრივი t - ტესტი გამოიყენება მხოლოდ იმ შემთხვევაში, როდესაც შეფასებული კოეფიციენტის მნიშვნელობა მნიშვნელოვნად განსხვავდება თეორიულიდან. აქ განიხილება ორი შემთხვევა:

1. შეფასებული კოეფიციენტი მნიშვნელოვნად განსხვავდება ნულისაგან. ამ შემთხვევაში მოწმდება ჰიპოთეზა:

$$H_0 : S_k = 0$$

$$H_A : S_k \neq 0$$

2. შეფასებული კოეფიციენტი მნიშვნელოვნად განსხვავდება არანულოვანი მნიშვნელობისაგან.

\hat{S} კოეფიციენტის ორმხრივი ტესტირების აუცილებლობა წამოიჭრება იმ შემთხვევაში, როდესაც იბადება კითხვა: როგორი იქნება შესაფასებელი კოეფიციენტის ნიშანი, დადებითი თუ უარყოფითი. მაგალითად, განხილული ამოცანისათვის \hat{S} -ს კოეფიციენტის ნიშნის დასადგენად შეიძლება შემოწმდეს ჰიპოთეზა ისევე, როგორც პუნქტი 1-ითაა გათვალისწინებული.

ორმხრივი ტესტი გულისხმობს ჰიპოთეზის უარყოფის განხილვას ორ სხვა და სხვა არეში - დადებითსა და უარყოფითში, რომელიც შეესაბამება მოსალოდნელ ზონებს.

რეზიუმე

1. ჰიპოთეზის ტესტირება საშუალებას გვაძლევს გენერალური ერთობლიობიდან შერჩევის გზით შევამოწმოთ ეკონომიკური თეორიის ბაზაზე აგებული რეგრესიის განტოლების კოეფიციენტების მიმართ მოსაზრება მათი ჭეშმარიტების შესახებ. ტესტირება მოიცავს ოთხ ძირითად ეტაპს:

ა/ ნულოვანი და ალტერნატიული ჰიპოთეზების ჩამოყალიბება;

ბ/ მნიშვნელოვნების დონეებისა და შესაბამისი t -ს კრიტიკული მნიშვნელობის შერჩევა;

გ/ რეგრესიის განტოლებასა და t -ს შეფასებული მნიშვნელობების ჩაწერა;

დ/ t -ს მნიშვნელობის შედარება t -ს კრიტიკულ მნიშვნელობასთან და ჰიპოთეზის მიღება ან უარყოფა.

2. ჰიპოთეზების შემოწმებისას ადგილი აქვს ორი ტიპის შეცდომას:

ა/ პირველი ტიპის შეცდომა უარყოფს ნულოვან ჰიპოთეზას, როდესაც ის ჭეშმარიტია;

ბ/ მეორე ტიპის შეცდომა: არ უარყოფს ნულოვან ჰიპოთეზას, როდესაც ის მცდარია.

3. t -ტესტის გამოყენებით ვაფასებთ რეგრესიის განტოლების კოეფიციენტებს, რისთვისაც გამოთვლილ t -ს მნიშვნელობას შევადარებთ თავისუფლების ხარისხისა და მნიშვნელოვნების დონის მიხედვით ცხრილიდან აღებულ t -ის მნიშვნელობას და ვიღებთ შესაბამის გადაწყვეტილებას: უარყოფთ ნულოვან ჰიპოთეზას, თუ $|t| > t_c$ და ალტერნატიულ ჰიპოთეზას, თუ t -ს ნიშანი იგივეა, რაც გათვალისწინებულია ალტერნატიული ჰიპოთეზით.

დამოუკიდებელი სამუშაო

4.1. ჩამოაყალიბეთ ნებისმიერი ამოცანა და ააგეთ მისი შესაბამისი რეგრესიის მოდელი. შეამოწმეთ ნულოვანი და ალტერნატიული ჰიპოთეზა ამ მოდელის კოეფიციენტებისათვის .

4.2. რა უფრო მნიშვნელოვანია: პირველი რიგის შეცდომის შენარჩუნება შედარებით დაბალ დონეზე, თუ მეორე რიგის შეცდომის შენარჩუნება დაბალ დონეზე?

4.3. რა უფრო მნიშვნელოვანია მეორე რიგის შეცდომის შენარჩუნება შედარებით დაბალ დონეზე, თუ პირველი რიგის შეცდომის შენარჩუნება დაბალ დონეზე?

4.4. განიხილეთ ჰიპოთეზა: ჟურნალებში სტატიების დასათაურება მოკლე სტატიებისათვის უფრო გრძელია, ვიდრე გრძელი სტატიებისათვის? შეამოწმეთ ეს ჰიპოთეზა. აიღეთ ნებისმიერი ჟურნალის რამდენიმე ნომერი და შეადგინეთ სტატიების ჩამონათვალი.

შემოიტანეთ აღნიშვნები:

X - სტატიების ფურცლების რაოდენობა;

Y - სტატიის სათაურში სიტყვების რაოდენობა .

თუ სიტყვა იწერება დეფისით, დააფიქსირეთ როგორც ერთი სიტყვა.

ააგეთ შესაბამისი გრაფიკი, შეარჩიეთ წრფე

$$Y = S_0 + S_1 X + v$$

შეაფასეთ მოდელის პარამეტრები უმცირეს კვადრატთა მეთოდის გამოყენებით; შეამოწმეთ მოდელი ადექვატურობაზე; შეამოწმეთ ჰიპოთეზა

$$H_0 : S_1 = 0$$

$$H_1 : S_1 < 0$$

ალტერნატივის საწინააღმდეგოდ t კრიტერიუმის გამოყენებით. გააკეთეთ შესაბამისი დასკვნები.

4.5. 36 სტუდენტს შემოწმება ჩაუტარეს მათი გონებრივი მონაცემების შემოწმების მიზნით. მონაცემები მოცემულია ცხრილით.

თქვენი აზრით, მიზანშეწონილია თუ არა სასწავლო კურსებზე მათ მისაღებად ტესტის ჩატარება? თუ მიზანშეწონილია, როგორ განსაზღვრავდით გამსვლელ ქულას.

გამოიყენეთ პირველი და მეორე რიგის შეცდომებს შორის კომპრომისი გამსვლელი ბალის დასადგენად.

#	ტესტი რების ბალი	სწავლე ბის შედეგი	#	ტესტი რების ბალი	სწავლების შედეგი
1	30	ვერ გავიდა	20	51	ვერ გავიდა
2	29	გავიდა	21	45	ვერ გავიდა
3	33	ვერ გავიდა	22	22	ვერ გავიდა
4	62	გავიდა	23	30	გავიდა
5	59	ვერ გავიდა	24	40	გავიდა
6	63	გავიდა	25	26	ვერ გავიდა
7	80	გავიდა	26	9	ვერ გავიდა
8	32	ვერ გავიდა	27	36	გავიდა
9	60	გავიდა	28	61	გავიდა
10	76	გავიდა	29	79	გავიდა
11	13	ვერ გავიდა	30	57	ვერ გავიდა
12	41	ვერ გავიდა	31	46	გავიდა
13	26		32	70	ვერ გავიდა
14	51	გავიდა	33	43	გავიდა
15	45	ვერ გავიდა	34	31	გავიდა
16	22	გავიდა	35	68	გავიდა
17	30	გავიდა	36	62	გავიდა
18	40	ვერ გავიდა	37	41	ვერ გავიდა
19	26	გავიდა	38	35	ვერ გავიდა

თავი 5. მოდელის სპეციფიკაცია

ეკონომიკური მოდელის აგება მოიცავს მოდელში შემავალი თანაფარდობების სპეციფიკაციასაც. თითოეულ მათგანში შემავალი ცვლადების შერჩევას, მათემატიკური ფუნქციის განსაზღვრას. თუ უკვე განსაზღვრულია ის ცვლადები, რომლებიც რეგრესიის განტოლებაში შევლენ, მაშინ ჩვენი ამოცანა შემოისაზღვრება მხოლოდ რეგრესიის განტოლების კოეფიციენტების შეფასებით, მათთვის ნდობის ინტერვალის დადგენით და სხვა შეფასებებით. პრაქტიკაში თითქმის არასდროს არ ვართ დარწმუნებულები, რომ რეგრესიის განტოლება ზუსტადაა წარმოდგენილი. ეკონომიკური თეორია თუმცა მიმართულებას აძლევს, მაგრამ შესაძლებელია, რეგრესიის განტოლებაში შევიყვანოთ ცვლადები, რომლებიც შეიძლებოდა არ შეგვეყვანა და პირიქით, არ შეგვეყვანა ის ცვლადები, რომელთა ჩართვა განტოლებაში აუცილებელი იქნებოდა.

რეგრესიის კოეფიციენტების შეფასებაც მნიშვნელოვნადაა დამოკიდებული მოდელის ზუსტ სპეციფიკაციაზე, ამიტომ ცვლადების სპეციფიკაციასთან დაკავშირებით შეიძლება აღვნიშნოთ შემდეგი:

1. თუ რეგრესიის მოდელში გამოტოვებულია ცვლადი, რომელიც მოდელში უნდა შედიოდეს, მაშინ რეგრესიის კოეფიციენტების შეფასება გადაადგილებადია. კოეფიციენტის სტანდარტული შეცდომები და შესაბამისი ტესტები საერთოდ არაკორექტულია;

2. თუ რეგრესიის მოდელში შესულია ცვლადი, რომელიც მოდელისათვის გაუმართლებელია, მაშინ რეგრესიის კოეფიციენტების შეფასება გადაუადგილებადია, თუმცა არაეფექტურია. სტანდარტული შეცდომები საერთოდ კორექტიულია, მაგრამ შეფასების არაეფექტურობის გამო თავისი მნიშვნელობები შედარებით დიდია;

3. თუ ცვლადის გაზომვა შეუძლებელია (ეს შეეხება ცვლადებს, რომლებითაც უნდა გამოისახოს სოციალურ-ეკონომიკური მდგომარეობა ან განათლების დონე ან მათი გაზომვა შეიძლება, მაგრამ შედეგის მიღება დიდ დროსა და ენერჯის ხარჯვასთან არის დაკავშირებული და პრაქტიკულად, ასეთი ცვლადების მოდელში შეყვანაზე უარს ვამბობთ, მაშინ მიზანშეწონილია, ასეთი ცვლადების ჩანაცვლება სხვა ცვლადებით რათა მათი გავლენა მოდელზე მთლიანად არ გამოირიცხოს. მაგალითად, სოციალურ-ეკონომიკური მდგომარეობის მაჩვენებლად შეიძლება გამოვიყენოთ შემოსავლის მაჩვენებელი, ხოლო განათლების დონის მაჩვენებლად - მასწავლებლებისა და თანამშრომლების საერთო რიცხოვნება სტუდენტების რიცხოვნებასთან შედარებით ან ერთ სტუდენტზე გაწეული დანახარჯი. ჩამმნაცვლებელი ცვლადის მოძებნას ორი მიზეზი აქვს: ერთი ის, რომ მნიშვნელოვანი ცვლადის იგნორირებას არ ვაკეთებთ; მეორეც - ჩამმნაცვლებელი ცვლადის მიხედვით შეიძლება ირიბი ინფორმაცია მაინც მივიღოთ ჩანაცვლებული ცვლადის შესახებ.

4. ხანდახან, ცვლადის ჩანაცვლებისას ადგილი აქვს შეცდომას. ეს იმ შემთხვევაშია მოსალოდნელი, როდესაც ვთვლით, რომ მაგალითად, ცვლადი X დამოკიდებულია P ცვლადზე მაშინ, როდესაც X დამოკიდებულია q ცვლადზე. თუ P და q ცვლადებს შორის კორელაცია დაბალია შედეგი შეიძლება უარესი მივიღოთ, ხოლო იმ შემთხვევაში, თუ კორელაცია ამ ცვლადებს შორის მაღალი იქნება R^2 სასურველ დონესთან ახლოსაა და ჩვენ ვერც ვიგრძნობთ, რომ მიღებული თანაფარდობა ჭეშმარიტი არ იქნება.

ვიდრე რეგრესიის განტოლების შეფასებას დავიწყებთ, უნდა დავწმუნდეთ, რომ განტოლება სავსებით განსაზღვრულია. ეკონომეტიკული განტოლების სპეციფიკაცია სამ ეტაპს მოიცავს:

- დამოუკიდებელი ცვლადების შერჩევა;
- ფუნქციონალური ფორმის კორექტირება;
- სტოქასტური (შემთხვევითი) წევრის კორექტირება.

დამოუკიდებელი ცვლადების შერჩევისას ჩვენი გადასაწყვეტია, თუ რომელი ცვლადი უნდა შევიყვანოთ რეგრესიის განტოლებაში, რაც ერთდროულად, ეკონომეტრიკის სუსტ და ძლიერ მხარესაც წარმოადგენს. მისი სიძლიერე იმაში გამოიხატება, რომ ცვლადების შერჩევის შემდეგ შეგვიძლია განტოლების ჩაწერა, ხოლო სუსტი მხარე - ცვლადების შერჩევისას შეიძლება მოგვიწიოს მრავალი სხვა და სხვა სპეციფიკაციის შემოწმება, ვიდრე იმ ცვლადების ერთობლიობას შევარჩევთ, რომელიც საკვლევი ამოცანისათვისაა მისაღები.

პირველი განხილვისას განტოლებაში შეგვყავს ის ცვლადები, რომელთა არსებობა განტოლებაში ეკონომიკური თეორითაა განსაზღვრული. ამ შემთხვევაში

სადავო თითქოს არაფერია, მაგრამ თუ თეორია ვერ ადასტურებს ან ნაკლებ მოსალოდნელობას იჩენს ცვლადისადმი, მაშინ აუცილებლად გვექმნება პრობლემა განტოლებაში შესაყვანი ცვლადების მიმართ. გარდა ამისა, აუცილებლად წამოიჭრება შერჩეულ მოდელებში საკითხი პარამეტრების შეფასების შესახებ, კერძოდ:

1. შესაფასებელ მოდელში გამორჩენილია (გამოტოვებულია) დამოუკიდებელი ცვლადები, რომლებიც ჭეშმარიტ მოდელში არიან;
2. შესაფასებელ მოდელში შეყვანილი დამოუკიდებელი ცვლადები ჭეშმარიტი მოდელისათვის არსებით (მნიშვნელოვან) ცვლადებს არ წარმოადგენენ.

5.1. სტატისტიკური შეცდომების წყაროები

როგრც ცნობილია, ემპირიული სამუშაო მოიცავს გაზომვის, შერჩევის და მოდელის სპეციფიკაციის შეცდომებს. ეს შეცდომები განვიხილოთ კონკრეტული მაგალითის მიხედვით.

ვთქვათ, სამომხმარებლო ფუნქციაში გამოსათვლელია რეგრესიის განტოლების პარამეტრები. მონაცემები, რომლებიც ახასიათებენ მოხმარებასა და შემოსავალს უშუალო გაზომვებით არ მიიღება. ძირითადად, მათი მიღება გაანგარიშებების გზით ხდება იმ მონაცემების საფუძველზე, რომლებიც ასახულია საგადასახადო საბუთებში გარკვეული პროცედურის გამოყენებით. ბუნებრივია, გაანგარიშების მეთოდის და საწყისი მონაცემების შესაძლებელია შეცდომებს მოიცავდენ, თუ სავარაუდოდ, სამომხმარებლო ფუნქციაში

$$C = a + bY_d$$

გამოსათვლელია რეგრესიის განტოლების პარამეტრები, მაშინ ეს გამოსახულება უნდა შემოწმდეს $C = a + bY_d$ და $(Y_d + z)$ ტიპის მონაცემებზე, სადაც x და z უცნობი სიდიდეებია, რომლებიც წლიდან წლამდე გარკვეულ ცვლილებებს განიცდიან. ბუნებრივია, როდესაც საქმე გვაქვს ასეთი ტიპის ინფორმაციასთან, ადვილი გასაგებია, რომ სტატისტიკური მეთოდები a და b პარამეტრების არასრულყოფილ შეფასებას იძლევიან.

ნებისმიერი ეკონომისტისათვის ცხადია, რომ სამომხმარებლო ფუნქცია არა მარტო Y_d -ს გავლენას განიცდის, არამედ მასზე მოქმედებს უამრავი ფაქტორი და თუნდაც ისეთი მაჩვენებლები, როგორიცაა: მომხმარებელთა გემოვნება, საპროცენტო განაკვეთები, ფასთა სხვაობები რეგიონების მიხედვით, შემოსავლების განაწილების ხასიათი, მოსახლეობის ასაკობრივი სტრუქტურა, მომხმარებელთა შემოსავლების ცვლილებები და სხვა, რომელთა გავლენა მოდელში გათვალისწინებული არ იქნება, მაგრამ იმ შემთხვევაში, თუ მოდელში ჩავრთავთ ყველა მაჩვენებელს, მაშინ დამოკიდებულების ჭეშმარიტი სახე საკმაოდ რთული იქნება, ვიდრე იმ შემთხვევაში, თუ მოდელში შევიყვანთ მხოლოდ Y_d მაჩვენებელს.

უნდა შევნიშნოთ, რომ, რადგან არსებობს შესაძლო ალტერნატივების სიმრავლე, მაღალი საიმედოობით შეიძლება ვთქვათ, რომ ჰიპოთეზები, რომელთა შემოწმებასაც ვატარებთ, აუცილებლად შეიცავენ სპეციფიკაციის გარკვეულ შეცდომებს.

5.2. გამოტოვებული ცვლადები

ვთქვათ, რეგრესიის მოდელის აგებისას არსებითი, დამოუკიდებელი ცვლადის შეყვანა დაგვავიწყდა ან არსებითი ცვლადის შესახებ მონაცემები არ გავაჩნია. ორივე შემთხვევაში საქმე გვაქვს გამოტოვებული ცვლადის შემთხვევასთან. შედეგში, შეფასებული რეგრესიის განტოლება საეჭვო ხდება. ეს მდგომარეობა აღინიშნება როგორც სპეციფიკაციის გადაადგილება. განტოლება, რომელიც შეიცავს ერთზე მეტ დამოუკიდებელ ცვლადს, S_k კოეფიციენტი წარმოადგენს დამოკიდებული Y ცვლადის ცვლილებას, რომელიც გამოწვეულია დამოუკიდებელი ცვლადის ერთი ერთეულით გაზრდისას, როდესაც განტოლების დანარჩენი დამოუკიდებელი ცვლადები მუდმივ დონეზე არიან დაფიქსირებულნი.

ვთქვათ, რეგრესიის განტოლების ჭეშმარიტი მოდელია

$$Y_i = S_0 + S_1 X_{1i} + S_2 X_{2i} + V_i \quad (5.1)$$

სადაც,

V - შემხვევითი წევრია და ვთქვათ, წინასწარი განზრახვის გარეშე გამოგვრჩა მნიშვნელოვანი (არსებითი) წევრი - დამოუკიდებელი ცვლადი ან ამ ცვლადის შესახებ ვერ მივიღეთ ინფორმაცია მონაცემების სახით და ნაცვლად ჭეშმარიტი განტოლებისა, გვაქვს განტოლება

$$Y_i = S_0 + S_1 X_{1i} + V_i^*$$

სადაც ,

$$V_i^* = f(S_2 X_{2i} + V_i)$$

შემთხვევითი წევრი V_i^* არ არის დამოკიდებული X_{1i} -ზე, ვიდრე X_{1i} და X_{2i} ცვლადები კორელირებულნი არ არიან და დამოუკიდებელი ცვლადიც შემთხვევით წევრზე არ არის დამოკიდებული, ვიდრე გამოტოვებული ცვლადი კორელირებული არ არის განტოლებაში შემავალ ყველა დამოუკიდებელ ცვლადთან.

ეს კი იმის მაჩვენებელია, რომ უმცირეს კვადრატთა მეთოდით შეფასებული რეგრესიის კოეფიციენტები არ აკმაყოფილებენ გადაუადგილებადობის ან ვარიაციის მინიმუმის პირობას ან ორივეს ერთად.

ზოგადად, რეგრესიის მოდელისათვის ორი დამოუკიდებელი ცვლადით S_1 კოეფიციენტის მოსალოდნელი მნიშვნელობა, როდესაც არსებითი ცვლადი - X_2 გამოტოვებულია, ტოლია

$$E(\hat{S}_1) = S_1 + S_2 * f(r_{12}) \quad (5.2),$$

რომლის მიხედვითაც კოეფიციენტის მოსალოდნელი მნიშვნელობა ტოლია მის ჭეშმარიტ მნიშვნელობას დამატებული გამოტოვებული ცვლადის კოეფიციენტის ჭეშმარიტი მნიშვნელობა გამრავლებული განტოლებაში შეყვანილი ცვლადისა და გამოტოვებული ცვლადის კორელაციის კოეფიციენტზე:

$$\text{ფუნქცია } f(r_{12}) = r_{12} \sqrt{\frac{\sum X_2^2}{\sum X_1^2}}$$

$$\text{სადაც: } \begin{aligned} x_1 &= (X_{1i} - \bar{X}_1) \\ x_2 &= (X_{2i} - \bar{X}_2) \end{aligned}$$

ჩვენს მიერ ჩატარებული ანალიზის ილუსტრაციისათვის განვიხილოთ მაგალითი.

ვთქვათ, ქვეყანაში ინდივიდის შემოსავალი განისაზღვრება სწავლის ხანგრძლივობით - S , ინტელექტის ინდექსით - IQ და იკბლიანობით - u . ძირითად შემოსავალს, რომელიც შეადგენს 10000 ერთეულს, ემატება 1500 ერთეული სწავლის თითოეული წლის მიხედვით 10 წლის ზევით და 250 ერთეული თითოეული IQ ბალზე 85 ბალის ზემოთ. გარდა ამისა, არსებობს იღბლიანობის ფაქტორი - u .

20 ინდივიდისათვის მონაცემები მოცემულია ცხრილით 5.1. თუმცა S და IQ მნიშვნელობები შემთხვევით ხასიათს ატარებენ, ისინი დადებითად კორელირებული არიან, რაც მრავალ ქვეყანაშიც შეიმჩნევა.

ჩავწეროთ რეგრესიის განტოლება მიღებული მონაცემების მიხედვით:

$$Y=f(S,IQ,u)$$

და შედეგი შევაფასოთ STATA -ს პროგრამული უზრუნველყოფის გამოყენებით.

ცხრილი 5.1

#	S	IQ	Y	u
1	10	95	13880	1380
2	10	100	15310	1560
3	10	100	10470	-3280
4	11	105	17280	780
5	11	85	12480	980
6	11	115	18660	-340
7	11	95	14720	720
8	12	100	19390	2460
9	12	100	15510	-1240
10	12	110	19590	340
11	13	90	15770	20
12	13	120	22790	-460
13	14	110	20910	-1340
14	14	95	16720	-1780
15	15	105	23200	700
16	16	110	24690	-560
17	16	100	23130	380
18	17	125	34940	4440
19	18	105	27780	780
20	19	105	30380	1880

reg Y S IQ u

Source	SS	df	MS	Number of obs =	20
-----+-----				F(3, 16) =	.
Model	735079400	3	245026467	Prob > F =	0.0000
Residual	1.1921e-07	16	7.4506e-09	R-squared =	1.0000
-----+-----				Adj R-squared =	1.0000
Total	735079400	19	38688389.5	Root MSE =	8.6e-05

Y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
S	1500	7.87e-06	.	0.000	1500	1500
IQ	250	2.19e-06	.	0.000	250	250
u	1	1.24e-08	.	0.000	1	1
_cons	-26250	.000216		0.000	-26250	-26250

მივიღეთ რეგრესიის განტოლება და მისი შეფასება:

$$Y = -26250 + 1500S + 250IQ + u$$

კოეფიციენტების სტანდარტული გადახრების მნიშვნელობებია:

7.87e-06 2.19e-06 1.24e-08

თითოეული კოეფიციენტისათვის t- სტატისტიკის მნიშვნელობებია:

t 0.000 0.000 0.000

95%-იანი ნდობის ინტერვალებია:

S : 1500 1500

IQ: 250 250

u: 1 1

s : -26250 -26250

დეტერმინაციის კოეფიციენტი $R^2 = 1;000$

კორექტირებული დეტერმინაციის კოეფიციენტი $\bar{R}^2 = 1.00$

დაკვირვებების რიცხოვნება: 20

თავისუფლების ხარისხი:3

F(3; 16) = 0.000

იმ შემთხვევაში, როდესაც ამოცანის კვლევისას შესაძლებელია არ იყოს გათვალისწინებული ინტელექტის IQ -ს გავლენა შემოსავალზე , მაშინ შეფასდება რეგრესიის განტოლება $Y = f(S)$.

კვლავ გამოვიყენოთ Stata -ს პროგრამა რეგრესიის განტოლების შესაფასებლად, მივიღებთ:

. reg Y S

Source	SS	df	MS	Number of obs = 20
-----+-----				F(1, 18) = 64.22
Model	574153814	1	574153814	Prob > F = 0.0000
Residual	160925586	18	8940310.33	R-squared = 0.7811
-----+-----				Adj R-squared = 0.7689
Total	735079400	19	38688389.5	Root MSE = 2990.0

Y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
S	1984.768	247.6692	8.014	0.000	1464.435	2505.102
_cons	-6418.182	3349.034	-1.916	0.071	-13454.24	617.8765

მივიღეთ რეგრესიის განტოლების შეფასება:

$$Y = -6418.182 + 1984.768S$$

კოეფიციენტების სტანდარტული გადახრების მნიშვნელობებია:
3349.034 247.6692

თითოული კოეფიციენტისათვის t სტატისტიკის მნიშვნელობებია:
-1.916 8.014

95%-იანი ნდობის ინტერვალებია:

S : 1500 1500

IQ: 250 250

u: 1 1

s : -26250 -26250

დეტერმინაციის კოეფიციენტი $R^2 = 0;7811$

კორექტირებული დეტერმინაციის კოეფიციენტი $\bar{R}^2 = 0;7689$

დაკვირვებების რიცხოვნება: 20

თავისუფლების ხარისხი: 1

$F(1; 18) = 64.22$

თუ განტოლებას განვიხილავთ დამოუკიდებელი ცვლადებით S და IQ და შევაფასებდით რეგრესიის განტოლებას $Y=f(S; IQ)$, კვლავ გამოვიყენებთ Stata-ს რეგრესიის განტოლების შესაფასებლად, მივიღებთ:

reg Y S IQ

Source	SS	df	MS	Number of obs = 20		
Model	686450081	2	343225041	F(2, 17) =	119.99	
Residual	48629319.0	17	2860548.18	Prob > F =	0.0000	
				R-squared =	0.9338	
				Adj R-squared =	0.9261	

Y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
S	1640.033	150.5115	10.896	0.000	1322.481	1957.584
IQ	267.9744	42.76961	6.266	0.000	177.7384	358.2104
_cons	-29585.79	4154.652	-7.121	0.000	-38351.34	-20820.24

შეფასებული განტოლება:

$$Y = -29585.79 + 1640.033 S + 267.9744IQ$$

კოეფიციენტების სტანდარტული გადახრების მნიშვნელობებია:
4154.652 150.5115 42.76961

თითოული კოეფიციენტისათვის t სტატისტიკის მნიშვნელობებია:
-7.121 10.896 6.266

95%-იანი ნდობის ინტერვალებია:

S : 1322.481 1957.584

IQ: 177.7384 358.2104

s : -38351.34 -20820.24

დეტერმინაციის კოეფიციენტი $R^2 = 0,9338$

კორექტირებული დეტერმინაციის კოეფიციენტი $\bar{R}^2 = 0,9261$

დაკვირვებების რიცხოვნება: 20

თავისუფლების ხარისხი: 2

$F(2; 17) = 119.99$

თუ კიდევ განვიხილავთ მცდარ სპეციფიკაციას და ამჯერად გამოტოვებული ცვლადის როლში თუ გამოვა ცვლადი – S, მაშინ შესაფასებელი რეგრესიის განტოლებას ექნება სახე: $Y=f(IQ)$ და Stata -ს გამოყენების შემდეგ მივიღებთ შედეგებს:
. reg Y IQ

Source	SS	df	MS	Number of obs =	20
-----+-----				F(1, 18) =	16.08
Model	346812986	1	346812986	Prob > F =	0.0008
Residual	388266414	18	21570356.3	R-squared =	0.4718
-----+-----				Adj R-squared =	0.4425
Total	735079400	19	38688389.5	Root MSE =	4644.4

Y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
-----+-----						
IQ	438.338	109.3176	4.010	0.001	208.6702	668.0057
_cons	-25487.98	11361.93	-2.243	0.038	-49358.51	-1617.444
-----+-----						

შეფასებული განტოლებაა:

$$Y = -25487.98 + 438.338 IQ$$

კოეფიციენტების სტანდარტული გადახრების მნიშვნელობებია:
11361.93 109.3176

თითოული კოეფიციენტისათვის t სტატისტიკის მნიშვნელობებია:
-2.243 4.010

95%-იანი ნდობის ინტერვალებია:

IQ: 208.6702 668.0057

s : -49358.51 -1617.444

დეტერმინაციის კოეფიციენტი $R^2 = 0,4718$

კორექტირებული დეტერმინაციის კოეფიციენტი $\bar{R}^2 = 0,4425$

დაკვირვებების რიცხოვნება: 20

თავისუფლების ხარისხი: 2

$F(2; 17) = 16.08$

ცხადია, რომ მიღებული რეგრესიის განტოლება ისევე, როგორც ჭეშმარიტი სპეციფიკაციით მიღებული რეგრესიის განტოლება $Y=f(S; IQ; u)^{\wedge}$ დამოკიდებულია შერჩევაში შემთხვევითი წევრის ფაქტობრივ მნიშვნელობებზე. ბუნებრივია, მცდარი და ჭეშმარიტი სპეციფიკაციის შემთხვევებში დამოუკიდებელ ცვლადებთან კოეფიციენტების რხევა თავის საშუალო მნიშვნელობის ირგვლივ, განსხვავებული იქნება. მაგალითად, S ცვლადთან კოეფიციენტის გაბნევა, რომელიც ჭეშმარიტი

სპეციფიკაციის შემთხვევაში მისი საშუალო მნიშვნელობის 1500 -ის ირგვლივ ხდება. მცდარი სპეციფიკაციის შემთხვევაში ამავე ცვლადთან კოეფიციენტის გაზრდა მისი საშუალო მნიშვნელობის - 1854-ის ირგვლივს. საერთო ანალიზის ჩასატარებლად შევადგინოთ ცხრილი 5.2.

ცხრილი 5.2

დამოუკიდებელი ცვლადები	მუდმივი წევრი	S ცვლადთან კოეფიციენტი S_1	IQ ცვლადთან კოეფიციენტი S_2	დეტერმინაციის კოეფიციენტი R^2
S, IQ (სტანდარტული გადახრები)	-35864 (5808)	1776 (156)	312 (44)	0.89
S (სტანდარტული გადახრები)	3309 (3483)	1213 (258)	-----	0.55
IQ (სტანდარტული გადახრები)	13875 (11070)	-----	54 (109)	0.01

ცხრილი 5.2.-ში მოცემულია რეგრესიის განტოლებების შეფასების შედეგები მრავლობითი რეგრესიისათვის, რომელიც ჭეშმარიტ სპეციფიკაციას შეესაბამება და ორიც - რეგრესიისათვის თითო დამოუკიდებელი ცვლადით, რომლებიც მცდარ სპეციფიკაციას შეესაბამებიან. იმ შემთხვევაში, როდესაც გამოტოვებულია IQ ცვლადი, მაშინ S დამოუკიდებელ ცვლადთან კოეფიციენტის გადაადგილების სიდიდე ტოლი იქნება

$$S_2 \text{ cov}(S; IQ) / \text{var}(S) = 250 * (-1,80) = -450$$

იმ შემთხვევაში, თუ S ცვლადი არ არის შეყვანილი განტოლებაში, მაშინ IQ ცვლადთან კოეფიციენტის გადაადგილების სიდიდე ტოლი იქნება

$$S_1 \text{ cov}(S; IQ) / \text{var}(IQ) = 1500 * (-0.145) = -217$$

განვიხილოთ ზოგადი შემთხვევა. ვთქვათ, მოდელის ჭეშმარიტი სახეა

$$y = \Gamma + S_1 x_1 + S_2 x_2 + \dots + S_n x_n + V$$

და X_1 - ის შესახებ მონაცემები არ გვაქვს, თუმცა შეგვიძლია x_1 ცვლადი P ცვლადით ჩავანაცვლოთ:

$$x_1 = K + }P$$

სადაც K და $\}$ მუდმივი სიდიდეებია, მაგრამ უცნობები. თუ ავაგებთ რეგრესიის მოდელს

$$\hat{Y} = (a + b_1 K) + b_2 x_2 + \dots + b_k x_k + b_1 \} P,$$

მაშინ პარამეტრების შეფასებები და სტანდარტული შეცდომები ისეთივე იქნება როგორც x_1 ცვლადის არსებობის პირობებში. ერთადერთი ნაკლი იმაში მდგომარეობს, რომ x_1 თან კოეფიციენტის შეფასება არ გვექნება და $(a + b_1 K)$ პარამეტრის მიხედვით r – პარამეტრის შეფასებას მივიღებთ, ხოლო $b_1 \}$ – ის შეფასების მიხედვით S_1 – ის შეფასებას.

კორექტირება გამოტოვებული ცვლადების შემთხვევაში

თეორიაში სპეციფიკაციის პრობლემა შედარებით ადვილად გამოიყურება, ვიდრე ეს პრაქტიკულად კეთდება. თითქოს მარტივია, შევიყვანოთ განტოლებაში გამოტოვებული ცვლადი, მაგრამ სამწუხაროდ, სათქმელად ადვილია, ვიდრე შესასრულებლად, რადგან გამოტოვებული ცვლადით გამოწვეული გადაადგილება აღმოსაჩენად ძნელია. ჯერ ერთი, მისი მნიშვნელობა შეიძლება მცირეა და მაშინვე ვერ შევამჩნიოთ. ეს განსაკუთრებით იმ შემთხვევას ეხება, როდესაც მიზეზი არ გვაქვს ვიფიქროთ გამოტოვებული ცვლადის შესახებ. ამაზე დაფიქრება გვიწევს იმ შემთხვევებში, როდესაც შეფასებული კოეფიციენტის მნიშვნელობა მოსალოდნელისაგან მნიშვნელოვნად განსხვავდება. საერთოდ, მოდელის აგებისას მის საფუძველს ჩვენ თვითონ ვქმნით და ჩვენვე ვადგენთ:

- რომელი ცვლადები უნდა შევიყვანოთ მოდელში?
- როგორი ნიშნებია მოსალოდნელი ცვლადების კოეფიციენტებთან?
- გვაქვს თუ არა კოეფიციენტების მნიშვნელობებისათვის მოსალოდნელი ინტერვალი?
- გვაქვს თუ არა (ცნობილია თუ არა) მოდელში შესაყვანი არსებითი ცვლადები და სხვა.

თუ განტოლების შეფასებისას მოულოდნელი შედეგი მივიყვანთ იმ მოსაზრებამდე, რომ განტოლებაში გამოტოვებულია მნიშვნელოვანი ცვლადი და გადავწყვეტთ ცვლადის შეყვანას განტოლებაში. ამ შემთხვევაში, მიზანშეწონილია კოეფიციენტების მოსალოდნელი გადაადგილების მიმართ ანალიზის ჩატარება, თუ მოსალოდნელი გადაადგილების ნიშანი იგივეა, რაც ჩვენთვის მოულოდნელი შედეგისა.

5.3. განტოლებისათვის მიუღებელი ცვლადები

რა მოხდება, თუ რეგრესიის განტოლებაში შევიყვანთ ცვლადს, რომელიც მას არ ეკუთვნის ანუ "განტოლებაში მას არაფერი ესაქმება", თუ ჭეშმარიტი რეგრესიის მოდელის სპეციფიკაციაა:

$$Y_i = S_0 + S_1 X_{1i} + V_i,$$

მაგრამ გაუთვალისწინებელი მიზეზების გამო განტოლებაში შეყვანილია სხვა ცვლადიც, განტოლებისათვის მიუღებელი ცვლადი X_{2i} .

$$Y_i = S_0 + S_1 X_{1i} + S_2 X_{2i} + V_i^{**}$$

მაშინ, შემთხვეითი წევრი წარმოდგენილი იქნება შემდეგი სახით:

$$V_i^{**} = f(V_i - S_2 X_{2i})$$

თუ მიუღებელი ცვლადის ჭეშმარიტი კოეფიციენტი ნულის ტოლია, ამ შემთხვევაში ეს ნიშნავს, რომ კოეფიციენტი გადაუადგილებადია, როდესაც $S_2 = 0$;

მიუღებელი ცვლადის ჩართვა განტოლებაში გამოიწვევს შეფასებული კოეფიციენტების ვარიაციების გაზრდას, ხოლო გაზრდილ ვარიაციებს ექნებათ ტენდენცია დაბლა დასწიონ t - სტატისტიკის შეფასებების მნიშვნელობები. ასევე, მიუღებელ ცვლადებს გავლენა ექნებათ კორექტირებულ დეტერმინაციის კოეფიციენტზე, მაგრამ არა R^2 -ზე და როგორც წესი, დასწევნ მის მნიშვნელობას. უმცირეს კვადრატთა მეთოდით მიღებული S_1 -ის ვარიაცია იქნება:

$$\text{var}(\hat{S}_1) = \frac{t^2}{(1 - r_{12}^2) * \sum (X_1 - \bar{X}_1)^2}$$

მაგრამ, როდესაც $r_{12} = 0$, მაშინ

$$\text{var}(\hat{S}_1) = \frac{t^2}{\sum (X_1 - \bar{X}_1)^2}$$

ცხრილი 5.3. -ში მოცემულია გამორჩენილი და მიუღებელი ცვლადების გავლენა რეგრესიის კოეფიციენტების შეფასებზე.

ცხრილი 5.3.

კოეფიციენტების შეფასების ეფექტი	გამოტოვებული ცვლადი	მიუღებელი ცვლადი
გადაადგილება	დიახ	არა
ვარიაცია	აწევა	დაწევა

5.4. სპეციფიკაციის ოთხი მნიშვნელოვანი კრიტერიუმი

სპეციფიკაციის კრიტერიუმები იმდენად მნიშვნელოვანია, რომ შეიძლება რჩევის სახით მიიღოს მკლევარმა ვიდრე ის მუშაობას დაიწყებდეს რეგრესიის განტოლებში ცვლადის შეყვანას ან გამოყვანაზე.

1. თეორია: ეხმაურება თუ არა ცვლადი თეორიას;
 2. t -ტესტი: არის თუ არა ცვლადების შეფასებული კოეფიციენტების მნიშვნელობები მოსალოდნელ მნიშვნელობებთან შესაბამისობაში;
 3. უმჯობესდება თუ არა განტოლება დამატებითი ცვლადის შეყვანით ან მისი გამოყვანით განტოლებიდან;
 4. გადაადგილება/დობა: აქვს თუ არა მნიშვნელოვანი ცვლილება შეფასებული განტოლების კოეფიციენტებს დამატებითი ცვლადის შეყვანის შემდეგ განტოლებაში.
- შესაძლებელია, ოთხივე პუნქტი ერთდროულად არ სრულდებოდეს, რაც ხშირად ხდება, მაგრამ მათი გათვალისწინება ნამდვილად აუცილებელია რეგრესიის მოდელის დახვეწის მიზნით.

5.5. ნახტომისებური რეგრესიული ანალიზის პროცედურა

ნახტომისებური რეგრესიული ანალიზის პროცედურა ითვალისწინებს კომპიუტერული პროგრამის გამოყენებას დამოუკიდებელი ცვლადის შესარჩევად კონკრეტული განტოლების შეფასებისას.

ეკონომეტრისტიკის უმრავლესობა ირჩევს სპეციფიკაციის თანმიმდევრობით ძიებას, იღებს საწყის განტოლებას და შემდეგ, თანმიმდევრობით გამოყავს განტოლებიდან ცვლადი ან უმატებს ახალ ცვლადს. ეს პროცედურა შეესაბამება ძიებას "დიდი განტოლებიდან პატარასაკენ" მანამ, სანამ არ მიიღება სამიხედო განტოლება მოცემული ამოცანისათვის. ასეთი მიდგომა ეკონომეტრისტს საშუალებას აძლევს შეაფასოს განტოლებების საკმაოდ დიდი რაოდენობა. ასეთი მიდგომა ნაწილობრივ ამახინჯებს ანალიზის შედეგების სტატისტიკურ საიმედოებას ორი მიზეზის გამო:

1. შედეგების სტატისტიკური მნიშვნელოვნება ზედმეტად შეფასებულია, რადგან წინა რეგრესიის განტოლებების შეფასებები იგნორირებულია;
2. მოსალოდნელობათა მიმდევრობა რეგრესიის სხვა და სხვა განტოლებების შეფასებების შედეგებს შორის, რასაც ეკონომეტრისტი იყენებდა.

სამწუხაროდ, არ არსებობს კვლევის უნივერსალური მეთოდი, რომელიც შეიძლება შეთავაზებული უნდა იყოს, მაგრამ ცხადია ის, რომ ეკონომეტრისტის ძირითადი დასაყრდენი ამოცანის შინაარსისა, ეკონომიკურ თეორიისა, ეკონომეტრისტის ინტუიციისა და განხილული მეთოდების ერთობლივი ნაზავია.

განვიხილოთ მაგალითი. ვთქვათ, ვიკლევთ სტუდენტთა 25 კაციან ჯგუფს ეკონომეტრიკის კლასში და მიღებულია ცვლადების შესახებ შემდეგი მონაცემები (Studmund A.H.):

GPA_i - i -ური სტუდენტის კოლეჯში მიღებული საშუალო შეფასება ოთხსკალიანი სისტემით;

$HGPA_i$ - i -ური სტუდენტის უმაღლესი სკოლის შეფასება ოთხსკალიანი სისტემით;

$VSAT_i$ - i -ური სტუდენტის მიერ დაგროვილი ბალური შეფასება მათემატიკაში ზეპირი ტესტის ჩაბარებისას ($SAT_i - MSAT_i + VSAT_i$);

$MSAT_i$ - i -ური სტუდენტის მიერ დაგროვილი ბალური შეფასება SAT ტესტის ჩაბარებისას მათემატიკაში;

$GREK_i$ - ფიქტიური ცვლადი: რომელიც 1-ის ტოლია, თუ i -ური სტუდენტი რომე-ლიმე ორგანიზაციის წევრია. სხვა, დანარჩენ შემთხვევებში 0-ის ტოლია;

HRS_i - i -ური სტუდენტის მიერ კოლეჯში სწავლის პერიოდში დახარჯული საათების რიცხოვნება;

$PRIV_i$ - ფიქტიური ცვლადი: რომელიც 1-ის ტოლია, თუ i -ურმა სტუდენტმა დაამთავრა კერძო სკოლა. სხვა, დანარჩენ შემთხვევებში 0-ის ტოლია;

$JOCK_i$ - ფიქტიური ცვლადი, რომელიც 1-ის ტოლია: თუ i -ური სტუდენტი ატლეტთა გუნდის წევრია. სხვა, დანარჩენ შემთხვევებში 0-ის ტოლია;

$InEX_i$ - i -ური სტუდენტის მიერ კოლეჯში დაგროვილი ყველა კურსის რიცხოვნების ნატურალური ლოგარითმი;

მივიღოთ, რომ GPA_i დამოკიდებული ცვლადია. ავაგოთ რეგრესიის მოდელი ისე, რომ მოვიცვათ შემდეგ კითხვებზე პასუხები :

- რომელ დამოუკიდებელ ცვლადს შეარჩევდით?
- როგორია მოსალოდნელი ნიშანი დამოუკიდებელი ცვლადის თითოეულ კოეფიციენტთან?
- რამდენად მისაღებია თეორია თითოეული ცვლადის შერჩევისას?
- რომელი ცვლადებია მნიშვნელოვანი?
- რომელი ცვლადები შეიძლება მივიჩნიოთ ნაკლებ მნიშვნელოვანი ცვლადის როლში;
- შეგძლიათ თუ არა დაასახელოთ ცვლადი: რომელიც შეიძლება ჩავრთოთ ანლიზში.

ჩაწეროთ განტოლების მოსალოდნელი სპეციფიკაცია. ვთქვათ, ასეთი სახით:

$$GPA_i = f(?, ?, ?, ?, ?)$$

ჩვენ გვჯერა, რომ სტუდენტის შეფასება ისეთი მაჩვენებლების ფუნქციაა, როგორცაა : შესაძლებლობა, მონდომება, გამოცდილება სასაწვლი კურსების შერჩევისას. აქედან გამომდინარე, მოდელის სპეციფიკაცია იქნება:

$$GPA_i = f(HGPA_i, HRS_i, \ln EX_i, \dots)$$

შენიშვნა: ყველასათვის ცნობილია, რომ ტესტის ჩაბარება ერთ-ერთი მნიშვნელოვანი მაჩვენებელია. რამდენად მნიშვნელოვანია კერძო თუ სახელმწიფო სკოლის დამთავრება ან რომელიმე ორგანიზაციაში გაწევრიანება და სხვა მოითხოვენ ამოცანის შინაარსიდან გამომდინარე ანალიზს. გარდა ამისა, ჩვენ არ შეგვიძლია ავტომატურად ვივარაუდოთ, რომ მოცემული ცვლადები ყველა უნდა შევიყვანოთ მოდელში და ვირჩევთ იმ ცვლადებს, რომლებიც მეტად ახასიათებენ ამოცანის შინაარსს. მაგალითად, ჩვენ ვიცით, კვირაში რამდენ საათს უნდა ხარჯავდეს სტუდენტი სწავლაზე, ვიცით, რომ კერძო სკოლა შეიძლება იქ იყვეს, სადაც სახელმწიფო სკოლა არ არის და ვგრძნობთ, რომ ეს ცვლადი ამოცანისათვის არსებითი არ არის და ა.შ.

თუ ჩვენ შევაფასებთ ქვემოთ ჩაწერილი მოდელის

$$\hat{GPA}_i = -0.26 + 0.49HGPA_i + 0.06HRS_i + 0.42 \ln EX_i \quad (5.4)$$

სპეციფიკაციას სასწავლო ჯგუფში 25 სტუდენტის გათვალისწინებით, მივიღებთ შეფასებებს (ფრჩხილებში მოცემულია სტანდარტული გადახრები)

	(0;21)	(0;02)	(0;14)
t	2;33	3;00	3;00
n = 25	A=12;3	$\bar{R}^2=0;585$	

ვიდრე ამ სპეციფიკაციას თეორიულ საფუძველს მოვუძებნით, შეიძლება ასეთი კითხვა დავსვათ: ჩვენი მოლოდინი შეფასებული კოეფიციენტების ნიშნების მიმართ დაემთხვა და შეიძლება თუ არა ჩავთვალოთ, რომ განტოლების შეფასება მისაღებია?

მოვსინჯოთ სხვა ვარიანტები იმ ცვლადების ჩართვით, რომლებიც ამოცანის შინაარსთან ახლოს არიან, კერძოდ:

$$\hat{GPA}_i = -0.92 + 0.47HGPA_i + 0.05HRS_i + 0.44 \ln EX_i + 0.00060SAT_i \quad (5.5.)$$

	(0.22)	(0.02)	(0.14)	(0.00064)
t :	2.12	2.50	3.12	0.93

$$n=25 \quad F=9.4 \quad \bar{R}^2=0.583$$

გამოვიყენოთ სპეციფიკაციის კრიტერიუმები (5.4) და (5.5) განტოლებების შესადარებლად.

1. **t-ტესტი:** SAT ცვლადის ნიშანი დადებითია, როგორც მოსალოდნელი იყო, მაგრამ მისი კოეფიციენტის მნიშვნელობა მცირედ განსხვავდება ნულისაგან;

2. \bar{R}^2 : როგორც მოსალოდნელი იყო მისი მნიშვნელობა SAT ცვლადის დამატებით განტოლებაში მცირედ შეიცვალა;

3. **გადაადგილება (წანაცვლება):** განტოლებაში SAT ცვლადის დამატებით დახრის კოეფიციენტებიდან არც ერთი მნიშვნელოვნად არ შეცვლილა. შეიცვალა მხოლოდ t-ს მნიშვნელობები, რაც იმითაა გამოწვეული, რომ გაიზარდა სტანდარტული გადახრების მნიშვნელობები SAT ცვლადის შემოტანასთან დაკავშირებით.

ამგვარად, შეიძლება ჩავთვალოთ, რომ SAT ცვლადი არსებითი არ არის განხილული ამოცანისათვის და მისი ჩართვა მოდელში აუცილებელი არ არის.

არ გამოვრიცხავთ, რომ სხვადასხვა მკვლევარმა განხილული ამოცანისათვის სხვა და სხვა შედეგი მიიღოს და ამოხსნას სხვა და სხვა ფინალი ჰქონდეს.

რეზიუმე

1. შეფასებული განტოლების კოეფიციენტების წანაცვლების გამომწვევი მიზეზი განტოლებიდან გამოტოვებული ცვლადია, რადგან ის კორელაციაშია განტოლებაში შეყვანილ ცვლადებთან;

2. ცვლადის შეყვანა რეგრესის განტოლებაში, რომელიც ამ განტოლებისათვის არსებითი არ არის, არ გამოიწვევს იმ ცვლადების კოეფიციენტების წანაცვლებას, რომლებიც განტოლებაშია შეყვანილი, მაგრამ ყოველთვის გამოიწვევს მათი ვარიაციის მნიშვნელობების გაზრდას და t-ს მნიშვნელობისა და კორექტირებული დეტერმინაციის კოეფიციენტის დაწევას;

3. განტოლებაში ცვლადის შესაყვანად სასარგებლოა გამოვიყენოთ კრიტერიუმები:

- თეორია;
- t-ტესტი;
- გადაადგილება (წანაცვლება)

4. თეორია და არა სტატისტიკური მაჩვენებლების მორგება წარმოადგენს მნიშვნელოვან მომენტს საძიებელ განტოლებაში შესაყვანი ცვლადის დასადგენად.

დამოუკიდებელი სამუშაო

5.1. თითოეული სიტუაციისათვის განსაზღვრეთ ნიშანი მოსალოდნელი გადაადგილების (წანაცვლების) შემთხვევაში, რომელიც გამოწვეულია გამოტოვებული ცვლადის შეყვანით განტოლებაში:

- მუშების მიერ გამოქვეყნებული ხელფასის შესახებ განტოლებაში ჩართულია ცვლადი, რომელიც შეესაბამება მუშის გამოცდილებას. რა გავლენა ექნება გამოტოვებულ ცვლადს მუშის გამოცდილების კოეფიციენტზე, რომელიც შეესაბამება ასაკს;

- გვაქვს საწარმოო ფუნქცია თვითმფრინავების გამოშვების შესახებ. რა გავლენა ექნება დასაქმების კოეფიციენტზე გამოტოვებულ ცვლადს, რომელიც შეესაბამება საწარმოო კაპიტალს;
- გვაქვს განტოლება ღია კონცერტებზე დასწრების შესახებ. რა გავლენა ექნება გამოტოვებულ ცვლადს გამოსასვლელი დღის შესაბამისი ფიქტიური ცვლადის კოეფიციენტზე, რომელიც შეესაბამება მოსალოდნელი ნალექების ალბათობას კონცერტების დროს;
- გვაქვს განტოლება ატმის ზეთზე მოთხოვნის შესახებ, რომელიც შეიცავს შემოსავლის შესაბამისს ცვლადს. რა გავლენა ექნება შემოსავლის კოეფიციენტზე გამოტოვებულ ცვლადს, რომელიც შეესაბამება ფასს ატმის ზეთზე.

5.2. ვთქვათ, გაზეთების კომპანიას სურს გახსნას მაღაზია (მას უკვე აქვს 30 მაღაზია გახსნილი) და დაგეგვალათ შეარჩიოთ შესაბამისი ადგილი ამ ახალი მაღაზიისათვის. თქვენ ჩაწერეთ ამ ამოცანისათვის რეგრესიის განტოლება და შეაფასეთ იგი (ფრჩხილებში მოცემულია სტანდარტული გადახრები).

$$\hat{Y}_i = 30 + 0.1X_{1i} + 0.01X_{2i} + 10.0X_{3i} + 3.0X_{4i}$$

(0.02) (0.01) (1.0) (1.0)

სადაც:

Y_i - i -ურ მაღაზიაში გაზეთის საშუალო დღიური გაყიდვა;

X_{1i} - i -ურ მაღაზიასთან ყოველ საათში გავლილი მანქანების რიცხოვნება;

X_{2i} - i -ურ მაღაზიიდან საშუალო შემოსავალი;

X_{3i} - i -ურ მაღაზიაში დახლების რაოდენობა;

X_{4i} - i -ური მაღაზიის ახლოს პრესტიჟული მაღაზიები;

- ჩამოაყალიბეთ ჰიპოთეზა მოსალოდნელი ნიშნების შესახებ. გამოთვალეთ t -სტატისტიკა და თითოეული კოეფიციენტისათვის, შეამოწმეთ ტესტი 1%-იანი მნიშვნელოვნების დონისათვის;
- შეიმჩნევა თუ არა პრობლემა შეფასებულ განტოლებაში;
- რა წინადადება გექნებოდათ, თუ მეორედ შეაფასებდით ჰიპოთეტურ განტოლებას.

5.3. ფორმულირებულია მაკროეკონომიკური მოდელი

$$C_t = r + sY_t + v_t$$

$$Y_t = C_t + I_t + G_t$$

სადაც:

C - მოხმარება;

I - ინვესტიციებია;

Y - შემოსავალია;

v - შემთხვევითი წევრია;

G_t - სახელმწიფო სექტორის მიმდინარე ხარჯებია.

ცვლადები: ინვესტიციები და სახელმწიფო სექტორის მიმდინარე ხარჯები შეიძლება განვიხილოთ, როგორც ეგზოგენური ცვლადები, რომელთა შესახებ გვაქვს მონაცემები 25 წლის განმავლობაში და რომელთა მიხედვითაც C -ს მნიშვნელობა

საშუალოდ Y -ის, 70% I -ის 20% და G_t -ის 10% შეადგენდა. დისპერსიებს შორის შემდეგი დამოკიდებულება:

$$t_I^2 > t_G^2$$

უმცირეს კვადრატთა მეთოდის გამოყენებით მოხმარების ფუნქციის შეფასებისას რა მიმართულებით იქნება r, S -ს შეფასებათა გადაადგილება?

5.4. რეგრესიის განტოლების შეფასებისას აღმოჩნდა, რომ განტოლებიდან გამორჩენილია არსებითი ცვლადი, რომელიც დაამატეს განტოლებას. განსაზღვრეთ ქვემოთ მოცემულ შემთხვევებში გადაადგილების ნიშანი:

- შესწავლილია ერთი ჰექტარიდან მარცვლეულის მოსავლის აღება წლების მიხედვით წლის განმავლობაში წვიმიანი დღეების გათვალისწინებით და მიღებულია რეგრესიის მოდელი. მოდელში გამორჩენილია წლების მიხედვით საშუალო ტემპერატურა;
- შესწავლილია ვაშლის წლიური მოხმარება ქვეყანაში და მიღებულია რეგრესიის მოდელი. განტოლებაში არ შედის ფასი ფორთოხალზე;
- რეგრესიის განტოლება გამოსახავს პირველ სემესტრში სტუდენტის შეფასებას სტუდენტის მიერ სწავლებაზე დახარჯული საათების მიხედვით. არ არის გათვალისწინებული გამოცდის წინა ღამეს სტუდენტის მიერ მიღებული საათები.

5.5. ჩატარებულია ბრაზილიური ყავის წარმოების კვლევა საბაზრო მოთხოვნასთან დამოკიდებულებაში (ფასი და ბაზრის სხვადასხვა მახასიათებლები), რომელსაც გავლენა უნდა ჰქონოდა ყავის მოსავალზე.

- რა ნიშანია მოსალოდნელი ფასის გამომსახველი ცვლადის კოეფიციენტთან, რატომ?
- რომელი ცვლადები იქნებოდა მნიშვნელოვანი განტოლებაში შესაყვანად?
- როგორია განტოლების სპეციფიკაცია?

თავი 6. მოდელის ფუნქციონალური ფორმის შერჩევა

მას შემდეგ, როდესაც რეგრესიის განტოლებაში შემავალ ცვლადებს მიჩენილი აქვთ მათთვის განკუთვნილი ადგილი (შერჩეულია და დაფიქსირებულია დამოკიდებული და დამოუკიდებელი ცვლადები), შემდეგი ეტაპი განტოლების ფუნქციონალური ფორმის შერჩევაა. ჩვეულებრივ, თეორიული მიდგომა კარნახობს რეგრესიის ფორმირებისავის აუცილებელი მოდელის ტიპს ვიყენებთ რა ეკონომიკის პრინციპებს, მათემატიკის გამოყენებით ჩავწერთ სამიხედულ დამოკიდებულებას.

6.1. განტოლების მუდმივი წევრი და მისი ინტერპრეტაცია

რეგრესიის განტოლებაში S_0 მუდმივი წევრია ან Y ღერძზე მოკვეთილი მონაკვეთის სიდიდეა ან Y -ის ის მნიშვნელობა, როდესაც ყველა დამოუკიდებელი ცვლადისა და შემთხვევითი წევრის მნიშვნელობა ნულის ტოლია. ამავე დროს, S_0 -ს მნიშვნელოვანი თეორიული შინაარსიც გააჩნია. განვიხილოთ შემდეგი განტოლება:

$$C_i = S_0 + S_1 Q_i + V_i$$

სადაც:

C – გამოშვებული პროდუქციის ღირებულება;

Q – გამოშვებული პროდუქციის რაოდენობა.

განტოლების წევრი - $S_1 Q_i$ გამოხატავს ღირებულებას, რომელიც ასოცირებულია პროდუქციის Q_i რაოდენობის გამოშვებასთან. S_0 - ფიქსირებული ღირებულებაა, რომელიც განსაზღვრულია როგორც ღირებულება, როდესაც გამოშვებული პროდუქციის მნიშვნელობა ნულის ტოლია. კლასიკური მოდელის მეორე დებულების თანახმად მუდმივი წევრი შთანთქმავს ნებისმიერ არანულოვან მნიშვნელობას, რომელიც დაკავშირებულია შემთხვევით წევრთან და რომელსაც შეიძლება ადგილი ჰქონდეს მოცემულ კერძო შერჩევაში. მუდმივი წევრის მიმართ არ შეიძლება დაუდევრობა გამოვიჩინოთ, ვინაიდან ამ წევრის მიხედვით ვმსჯელობთ იმის შესახებ, თუ სად იღებს რეგრესიის განტოლება თავის საწყის მნიშვნელობას.

მეორეს მხრივ, შესაძლებელია, ცვლადებს შორის ჭეშმარიტი კავშირი არა-წრფივი იყოს და გადიოდეს კოორდინატთა სათავეზე. ეს არაწრფივი დამოკიდებულება აპროქსიმირებულია წრფივი რეგრესიით და მუდმივ წევრზე მნიშვნელოვანი დატვირთვა მოდის. ამიტომ მისი გათვალისწინება კიდევ უფრო მნიშვნელოვანი იქნება.

ზოგჯერ, საქმე გვაქვს შემთხვევასთან, როდესაც რეგრესიის განტოლება საშუალებას გვაძლევს მუდმივი წევრი ნულის ტოლად მივიღოთ, მაგრამ უნდა გვახსოვდეს ძირითადი წესი: არასოდეს არ უგულვებელვყოთ მუდმივი წევრი, თუ ეს თეორიას ეწინააღმდეგება!

6.2. ალტერნატიული ფუნქციონალური ფორმა

განტოლებისათვის ფუნქციონალური ფორმის შერჩევა სწორედ რომ განტოლების სპეციფიკაციიდან გამომდინარეობს. ფუნქციონალური ფორმის შერჩევა ძირითადად მომდინარეობს ეკონომიკური ან ბიზნესის თეორიიდან. დამოკიდებულ და დამოუკიდებელ ცვლადებს შორის დამოკიდებულების ლოგიკური ფორმის შესარჩევად ერთმანეთს დარდება სხვა და სხვა ფუნქციონალური ფორმები, რომელთაგან შეირჩევა ის ფუნქციონალური ფორმა, რომელიც თავისი თვისებებით ახლოსაა საკვლევ ამოცანასთან.

6.2.1. წრფივი ფორმა

წრფივი რეგრესიის მოდელი, რომელსაც ჩვენ ვიყენებთ, ძირითადად დაფუძნებულია იმ მოსაზრებაზე, რომ დამოუკიდებელ და დამოკიდებულ ცვლადებს შორის დამოკიდებულების დახრის კოეფიციენტი მუდმივია და წარმოდგენილია შემდეგი სახით:

$$\Delta Y / \Delta X_k = s_k, \quad k=1,2,\dots,K$$

დახრა მუდმივია, მაგრამ Y -ის ელასტიურობა X_k -ს მუდმივი არ არის:

$$E_{Y,X_k} = \frac{\Delta Y / Y}{\Delta X_k / X_k} = \frac{\Delta Y}{\Delta X_k} \frac{X_k}{Y} = s_k \frac{X_k}{Y}$$

თუ ჰიპოთეზა Y და X ცვლადებს შორის დამოკიდებულების შესახებ ისეთია, რომ დახრის კოეფიციენტის არსებობა მოსალოდნელია, მაშინ შეიძლება გამოვიყენოთ წრფივი ფუნქცია. თეორია ძირითადად გვკარნახობს ცვლადებს შორის დამოკიდებულების მხოლოდ ნიშანს და არა მის ფუნქციონალურ ფორმას, ხოლო გამოცდილება საშუალებას გვაძლევს გამოვიყენოთ წრფივი ფუნქციონალური ფორმა.

6.2.2. log –log ფორმა

log –log ფორმა ყველაზე ზოგადი ფუნქციონალური ფორმაა. ამ ფორმის მიხედვით, Y ცვლადის ნატურალური ლოგარითმი დამოკიდებული ცვლადია, ხოლო X ცვლადების ნატურალური ლოგარითმი დამოუკიდებელი ცვლადებია.

$$\ln Y = S_0 + S_1 \ln X_1 + S_2 \ln X_2 + v$$

log –log განტოლებაში რეგრესიის კოეფიციენტების ინტერპრეტაცია ელასტიურობაა. მაგალითად S_k -თვის გვაქვს:

$$E_{Y, X_k} = \frac{\Delta Y / Y}{\Delta X_k / X_k} = \frac{\Delta(\ln Y)}{\Delta(\ln X_k)} = S_k$$

S_k -ს ინტერპრეტაცია **log –log** განტოლებაში შემდეგია: თუ ცვლადის მნიშვნელობა გაიზრდება 1%-ით და დანარჩენი - X ცვლადები მუდმივ დონეზე განიხილება, მაშინ Y ცვლადის მნიშვნელობა შეიცვლება S_k %-ით. **log –log** მოდელი გამოიყენება მხოლოდ იმ შემთხვევებში, როდესაც ცვლადები მხოლოდ დადებით მნიშვნელობებს იღებენ. ფიქტიური ცვლადები, რომლებსაც შეუძლიათ მიღონ ნულოვანი მნიშვნელობა, არ შეიძლება გამოიყენოთ **log –log** განტოლებაში!

6.2.3. ნახევრად ლოგარითმული - semi-log ფორმა

semi-log ფორმა log-log ფორმის ერთ-ერთი ვარიანტი, რომელშიც ზოგიერთი ცვლადი (როგორც დამოკიდებული, ისე დამოუკიდებელი) გამოსახულია ნატურალური ლოგარითმით. მაგალითად,

$$Y_i = S_0 + S_1 \ln X_{1i} + S_2 X_{2i} + v_i$$

ამ შემთხვევაში დახრის ორი კოეფიციენტის ეკონომიკური მნიშვნელობა ერთმანეთისაგან განსხვავდება. X_2 ცვლადი Y ცვლადთან წრფივ დამოკიდებულებაშია, ხოლო X_1 ცვლადი არაწრფივ დამოკიდებულებაშია Y ცვლადთან, კერძოდ:

$$\Delta Y / \Delta X_1 = S_1 / X_1$$

$$\text{ან } S_1 = \Delta Y / (\Delta X_1 / X_1)$$

ე.ი. თუ X_1 გაიზრდება 1%-ით, მაშინ Y შეიცვლება $S_1 / 100$ ერთეულით და Y -ის ელასტიურობა იქნება:

$$E_{Y, X_1} = \frac{\Delta Y}{\Delta X_1} \frac{X_1}{Y} = \frac{S_1}{Y}$$

semi-log ფუნქცია საკმაოდ ხშირად გამოიყენება ბიზნესსა და ეკონომიკაში. მაგალითად, სამომხმარებლო ფუნქციის ტენდენცია შემოსავლის მიხედვით სწორედ

ასეთ ხასიათს ატარებს. ამ ტიპის ფუნქციასთან გვაქვს საქმე როდესაც ხელფასების მოდელს ვაგებთ. ასეთ მოდელში Y ცვლადი შეესაბამება i -ური დასაქმებულის ხელფასს. დამოუკიდებელი ცვლადი შეიძლება i -ური დასაქმებული ინდივიდის გამოცდილების შესაბამის ცვლადს წარმოადგენდეს.

უნდა შევნიშნოთ, რომ სავალდებულო არ არის **semi-log** ფუნქციას ლოგარითმი განტოლების მარჯვენა მხარეს ჰქონდეს. სავსებით შესაძლებელია, თვითონ Y -ის ლოგარითმი წარმოადგენდეს X ცვლადების ფუნქციას შემდეგი სახით:

$$\ln Y = S_0 + S_1 X_1 + S_2 X_2 + V$$

ასეთ გამოსახულებაში კოეფიციენტებს საინტერესო ინტერპრეტაცია აქვთ. თუ X_1 გაიზრდება ერთი ერთეულით, მაშინ Y შეიცვლება $S_1 * 100\%$ -ით, როდესაც ცვლადი მუდმივ დონეზეა დაფიქსირებული.

6.2.4. პოლინომიალური ფორმა

პოლინომიალური ფუნქციონალური ფორმა გამოსახავს Y -ს როგორც დამოუკიდებელი ცვლადების ფუნქციას, რომელთაგან ზოგი შედარებით ძლიერ აღმავალია. მაგალითად, ამ შემთხვევას ადგილი აქვს მეორე რიგის პოლინომიალურ (კვადრატულ) განტოლებაში:

$$Y_i = S_0 + S_1 X_{1i} + S_2 (X_{1i})^2 + S_3 (X_{2i}) + V_i$$

ამასთან,
$$\frac{\Delta Y}{\Delta X_1} = S_1 + 2S_2 X_1 \text{ და } \frac{\Delta Y}{\Delta X_2} = S_3$$

თუ განვიხილავთ ინდივიდის წლიურ გამომუშავებას ასაკის და განათლების მიხედვით ან მხოლოდ ერთ-ერთი მათგანის მიხედვით, შეიძლება დაისვას კითხვა, თუ რა გავლენა აქვს ასაკს გამომუშავებაზე. ასაკის მომატების მიხედვით მუშა მეტს გამოიმუშავებს და თერიული მოდელი შეიძლება წარმოვადგინოთ შემდეგი სახით:

$$Y = S_0 + S_1 X_i + S_2 X_i^2 + \dots + V_i$$

სადაც,

Y – გამომუშავება;

X – ასაკია.

მაგრამ, როდესაც მუშის ასაკი რაღაც ზღვარს ასცდება, მაშინ თანაფარდობა მუშის ასაკსა და გამომუშავებას შორის მკვეთრად იცვლება. ასეთი დამოკიდებულება შეიძლება პოლინომიალური ფუნქციის გამოყენებით შევისწავლოთ.

პოლინომიალური რეგრესიის გამოყენებისას, რეგრესიის კოეფიციენტების ინტერპრეტაცია რთულდება. მაგალითად, მესამე რიგის პოლინომიალური ფორმის გამოყენებისას ფუნქციის მნიშვნელობები გარკვეულ მონაკვეთში შეიძლება დადებითი იყოს, შემდეგ მიიღოს უარყოფითი მნიშვნელობები, შემდეგ კვლავ დადებითი და ა.შ.

6.2.5. ინვერსიული ფორმა

ინვერსიული ფორმის გამოყენებისას Y გამოისახება როგორც ერთი ან რამდენიმე დამოუკიდებელი შებრუნებული ცვლადის ფუნქცია. განტოლება შეიძლება ჩაიწეროს შემდეგი სახით:

$$Y = S_0 + S_1(1/X_{1i}) + S_2X_{2i} + v_i$$

ამ შემთხვევაში ინვერსიულ ფორმას იძლევა დამოუკიდებელი ცვლადი X_1 . ინვერსიული ფუნქციონალური ფორმა გამოიყენება იმ შემთხვევაში, როდესაც მოსალოდნელია, რომ რომელიმე დამოუკიდებელი ცვლადის გავლენა დამოკიდებულ ცვლადზე ნულს უახლოვდება. უნდა შევნიშნოთ, რომ როგორც კი X_1 -ის მნიშვნელობა საკმაოდ დიდი გახდება, Y -ის მნიშვნელობა შემცირდება (X_1 -ს არ შეუძლია მიიღოს ნულის ტოლი მნიშვნელობა, რადგან შედეგი ამ შემთხვევაში განსაზღვრული არ იქნება).

დახრის კოეფიციენტების მნიშვნელობებია:

$$\frac{\Delta Y}{\Delta X_i} = \frac{-S_1}{X_1^2} \text{ და } \frac{\Delta Y}{\Delta X_i} = S_2$$

1. თუ S_1 დადებითია, მაშინ შეფარდების მნიშვნელობა უარყოფითია. დამოკიდებულება Y -სა და X_1 -ს შორის მუდმივ დონეზე დაფიქსირებული X_2 -ის მნიშვნელობის შემთხვევისათვის უახლოვდება ($S_0 + S_2X_2$)-ს (შემთხვევითი წევრის მნიშვნელობა იგნორირებულია);

2. როდესაც S_1 უარყოფითია, მაშინ ადგილი აქვს კვეთას X_1 ღერძთან და ფუნქცია ასიმპტოტურად უახლოვდება ჰორიზონტალურ წრფეს, რომელიც შეესაბამებოდა შემთხვევას $S_1 > 0$.

ინვერსიული ფუნქციონალური ფორმის გამოყენების ერთ-ერთი მაგალითია ფილიპსის მრუდი (Nancy Wulwick, "Phillips' Apprximate Regression", History and Methodology of Econometrics", Oxford, 1989, pp.170-188), რომლითაც მან დაახასიათა დაუსაქმებლობასა და ხელფასს შორის დამოკიდებულება განტოლებით:

$$W_t = S_0 + S_1(1/U_t) + v_t$$

სადაც:

W – ხელფასის პროცენტული ცვლილება;

U – დაუსაქმებლობა.

განტოლება შეაფასა უმცირეს კვადრატთა მეთოდის გამოყენებით და მიიღო განტოლება. ფრჩხილებში მოცემულია სტანდარტული გადახრის მნიშვნელობა.

$$\hat{W}_t = 0.00679 + 0.16842(1/U_t)$$

$$(0.0590)$$

$$t=3.20$$

$$\bar{R}^2 = 0.397$$

თუმცა მიღებული შედეგების მიხედვით ცხადად არ ჩანს, რომ ინვერსიული ფუნქციონალური ფორმა საუკეთესოა განხილული შემთხვევისათვის.

6.3. პრობლემები არაკორექტული ფუნქციონალური ფორმის შერჩევისას

რეგრესიული მოდელისათვის ფუნქციონალური ფორმის შერჩევის საუკეთესო გზაა ისეთი სპეციფიკაციის შერჩევა, რომელიც მაქსიმალურად ეთანადება თეორიას. სირთულეებთან გვაქვს საქმე და ეს სირთულე წამოიჭრება მაშინ, როდესაც მოდელი ლოგიკურად არაწრფივია ცვლადების მიხედვით, მაგრამ ამის აღმოჩენა ძნელია. ამ შემთხვევაში წრფივი ფორმა არაკორექტულია. ყველა შემთხვევაში, როდესაც თეორია ზუსტად არ გვკარნახობს ფორმის სპეციფიკას, ბუნებრივია, პრობლემებიც გვექმნება.

ცხრილში მოცემულია ალტერნატიული ფუნქციონალური ფორმები

ფუნქციონალური ფორმა	განტოლება	დახრა ($\Delta Y / \Delta X$)	ელასტიურობა ($\Delta Y / \Delta X$) (X/Y)
წრფივი	$Y_i = S_0 + S_1 X_i + V_i$	S_1	$S_1 (X_i / Y_i)$
Log-log	$\ln Y_i = S_0 + S_1 \ln X_i + V_i$	$S_1 (Y_i / X_i)$	S_1
Semi-log (lnX)	$Y_i = S_0 + S_1 \ln X_i + V_i$	$S_1 (1 / X_i)$	$S_1 (1 / Y_i)$
Semi-log (lnY)	$\ln Y_i = S_0 + S_1 X_i + V_i$	$S_1 Y_i$	$S_1 X_i$
პოლინომიალური	$Y_i = S_0 + S_1 X_i + S_2 X_i^2 + V_i$	$S_1 + 2S_2 X_i$	$S_1 (X_i / Y_i) + 2S_2 (X_i^2 / Y_i)$
ინვერსიული	$Y_i = S_0 + S_1 (1 / X_i) + V_i$	$-S_1 (1 / X_i^2)$	$-S_1 (1 / X_i Y_i)$

უნდა აღვნიშნოთ, რომ კორექტირებული დეტერმინაციის კოეფიციენტი შესადარებლად ძნელია გამოვიყენოთ, თუ დამოკიდებული ცვლადი ტრანსფორმირებულია.

მაგალითად, ვთქვათ გვინდა შევადაროთ წრფივი განტოლება:

$$Y = S_0 + S_1 X_1 + S_2 X_2 + V$$

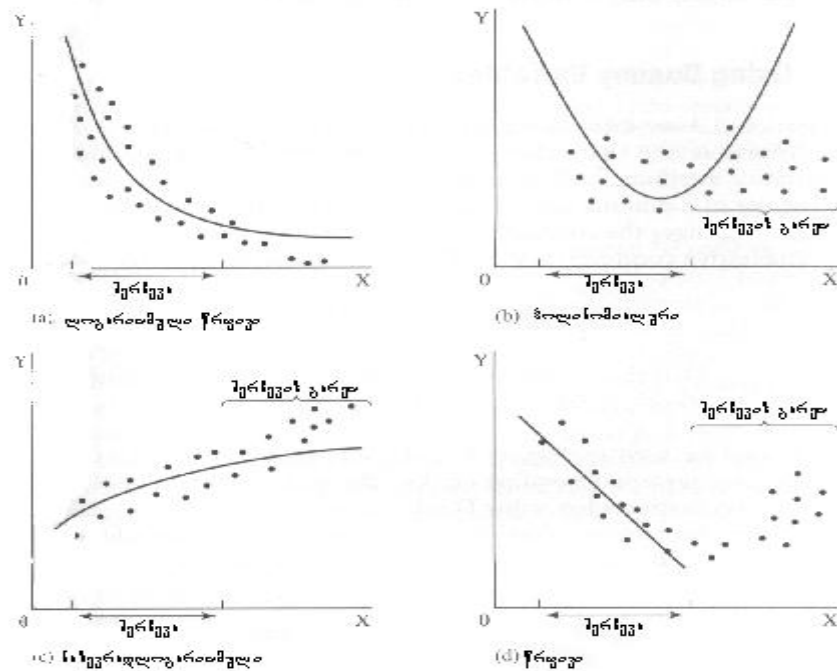
იმავე განტოლების ნახევრად ლოგარითმულ ვერსიას:

$$Y = S_0 + S_1 X_1 + S_2 X_2 + V$$

მათ შორის განსხვავება მხოლოდ დამოკიდებული ცვლადის ფუნქციონალურ ფორმაშია. მიზეზი იმისა, რომ კორექტირებული დეტერმინაციის კოეფიციენტი - \bar{R}^2 ნაკლებად გამოდგება ამ განტოლებების შესადარებლად ის არის, რომ დამოკიდებული ცვლადისათვის მისი საშუალოდან გადახრების კვადრატების ჯამი TSS ამ ორი განტოლებისათვის სხვა და სხვაა და გარდა ამისა, თვით დამოკიდებული

ცვლადის წარმოდგენაც სხვა და სხვაა. თუ TSS -ის მნიშვნელობები განსხვავებულია, \bar{R}^2 -ების მნიშვნელობები შესაძარებლად არ გამოდგება!

თუ კორექტული ფუნქციონალური ფორმა მაინც გამოყენებულია, მაშინ დიდია ალბათობა იმისა, რომ გენერალური ერთობლიობის ჭეშმარიტი მაჩვენებლების მნიშვნელობებიდან ცდომილებები საგრძნობი იქნება. ნახაზზე წარმოდგენილია არაკორექტული ფუნქციონალური ფორმები



არაკორექტული ფუნქციონალური ფორმები

რეზიუმე

1. ფუნქციონალური ფორმის შერჩევის საფუძველს ეკონომიკური თეორია წარმოადგენს. წრფივი ფორმა მისაღებია იმ შემთხვევაში, თუ სპეციფიური ჰიპოთეზა არ არის შემოთავაზებული;
2. ფუნქციონალური ფორმა, რომელიც ცვლადების მიმართ არაწრფივია, შეიცავს ორმაგ ლოგარითმულ Log-Log ფორმას ან ნახევრად ლოგარითმულ Semi-Log ფორმას ან პოლინომიულ ფორმას ან ინვერსიულ ფორმას;
3. არაწრფივი ფუნქციონალური ფორმის გამოყენება ხასიათდება პრობლემებით. კერძოდ, კორექტირებული დეტერმინაციის კოეფიციენტი \bar{R}^2 შედარების თვალსაზრისით მიუღებელია, თუ დამოკიდებულმა ცვლადმა ტრანსფორმაცია განიცადა.

დამოუკიდებელი სამუშაო

6.1. თითოეული (X, Y) წყვილისათვის (Y -დამოკიდებული ცვლადია, X -დამოუკიდებელი ცვლადი). როგორი იქნება ფუნქციონალური ფორმა? ახსენით თქვენნი მოსაზრება.

- Y - ფებსაცმელების შექმნა;
- X - ერთჯერადი შემოსავალი;
- Y – ფულზე მოთხოვნა;
- X - ბანკის საპროცენტო განაკვეთია;

6.2. განვიხილოთ semi-log განტოლების შეფასება (ფრჩხილებში მოცემულია საშუალო კვადრატული გადახრა)

$$\ln \hat{SAL}_i = 8.10 + 0.100ED_i + 0.110EXP_i$$

(0.025) (0.050)

$\bar{R}^2 = 0.48 \quad n = 28$

სადაც:

$\ln SAL$ – i -ური მუშის ხელფასი;

ED – i -ური მუშის განათლება, წელი;

EXP - i -ური მუშის სამუშაო გამოცდილება, წელი.

- როგორია თქვენი ჰიპოთეზა ნიშნების მიმართ?
- გამოთვალეთ t -ს კრტიკული მნიშვნელობა.
- შეამოწმეთ თქვენი ჰიპოთეზები.
- როგორია განტოლების მუდმივი წევრის ეკონომიკური შინაარსი?
- როგორ ფიქრობთ, რატომ არის გამოყენებული დამოკიდებული ცვლადის ლოგარითმული ფორმა (როგორია ფუნქციის დახრა განათლებასა და გამოცდილებასთან დაკავშირებით?)
- დავუშვათ, რომ ამ ამოცანისათვის გამოიყენეთ წრფივი განტოლება და მიღებულია კორექტრებული დეტერმინაციის კოეფიციენტის მნიშვნელობა 0.46-ია. როგორი იქნება თქვენი დასკვნა ამ შედეგთან დაკავშირებით?

6.3. ქვეყანა დაყოფილია ოთხ რეგიონად: ჩრდილოეთი, სამხრეთი, აღმოსავლეთი და დასავლეთი და შესწავლილია არაკვალიფიციური მუშახელის მოძრაობა სამუშაოს საშოვნელად. მოცემულია განტოლების შეფასება:

$$\hat{Y}_i = 4.78 - 0.038E_i - 0.041S_i - 0.048W_i$$

(0.019) (0.010) (0.012)

$\bar{R}^2 = 0.49 \quad n = 7338$

სადაც:

Y – არაკვალიფიციური მუშის საათობრივი ანაზღაურება;

E – ფიქტიური ცვლადია, რომელიც 1-ის ტოლია, თუ i -ური მუშა ქვეყნის ჩრდილოეთში ცხოვრობს. სხვა დანარჩენ შემთხვევაში 0-ის ტოლია;

S - ფიქტიური ცვლადია, რომელიც 1-ის ტოლია, i -ური მუშა ქვეყნის სამხრეთში ცხოვრობს. სხვა დანარჩენ შემთხვევაში 0-ის ტოლია;

W - ფიქტიური ცვლადია, რომელიც 1-ის i -ური მუშა ქვეყნის აღმოსავლეთში ცხოვრობს. სხვა დანარჩენ შემთხვევაში 0-ის ტოლია;

- რა პირობებს არ მოიცავს განტოლება ? რა პირობებია გამოტოვებული?
- თუ დაამატებთ გამოტოვებულ ცვლადს ისე, რომ დანარჩენ ცვლადებს დატოვებთ განტოლებაში, რა შეიცვლებოდა?

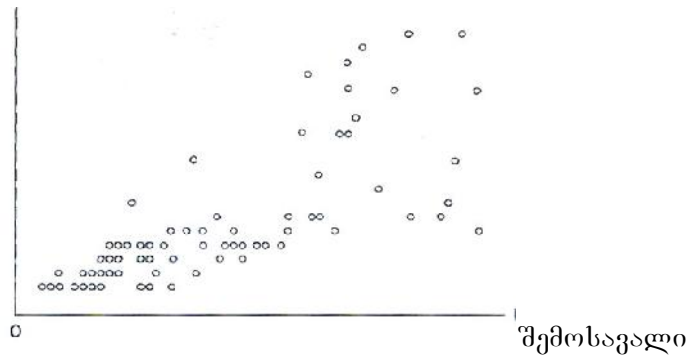
- თუ დაამატებთ განტოლებაში ფიქტიურ ცვლადს ისე, რომ განტოლებიდან გამოიყვანოთ E -ცვლადს: რა იქნებოდა ახალი ცვლადის კოეფიციენტის ნიშანი განტოლების შეფასების შემდეგ?
- არის თუ არა შემდეგი მოსაზრება კორექტული: რეგიონალური ცვლადების კოეფიციენტები საკმაოდ მცირეა საშუალო ხელფასთან შედარებით, ამიტომ მათ არ შეუძლიათ მნიშვნელოვანი გავლენა მოახდინონ რეგიონის მიხედვით ხელფასის ცვლილებაზე.
- თუ დააპირებთ მოდელში ერთი ცვლადის დამატებას, რომელი იქნებოდა ის ცვლადი?

თავი 7. ჰეტეროსკედასტიურობა

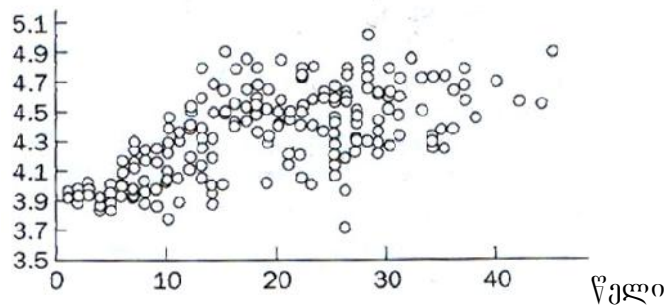
რეგრესიის კოეფიციენტების შეფასებისას უმცირეს კვადრატთა ან მაქსიმალური დასაჯერისობის მეთოდის გამოყენებით ვვარაუდობთ, რომ რეგრესიის მოდელში შემავალი ნაშთები განაწილებულია ნულოვანი საშუალო სიდიდით და დისპერსიით σ^2 რაც ნიშნავს, რომ $\text{var}(e_t) = \sigma^2$ ნებისმიერი t -თვის. ეს ვარაუდი ვარიაციის მუდმივობის შესახებ დაკავშირებულია ე.წ. ჰომოსკედასტიურობასთან რაც ნიშნავს გაბნევის ერთგვაროვნებას. სწორედ σ^2 წარმოადგენს e_t -ნაშთების შესაფასებელ მაჩვენებელს მისი საშუალო მნიშვნელობის მიმართ. ეს ექვივალენტურია იმისა, რომ ეს არის დამოკიდებულ ცვლადზე დაკვირვებების მნიშვნელობების ცვლილებების დისპერსიის ზომა რეგრესიის წრფის მიმართ.

მრავალ შემთხვევაში ეს დაშვება შეიძლება მცდარი აღმოჩნდეს. მაგალითად, ვთქვათ გამოვიკვლიეთ შინამეურნეობები შემთხვევითი შერჩევის მიხედვით და მივიღეთ ინფორმაცია თითოეული შინამეურნეობის მაჩვენებლების შესახებ: მოხმარება, დანახარჯები, შემოსავალი წლის განმავლობაში. როგორც წესი, შინამეურნეობები დაბალი შემოსავლით არ აღინიშნებიან ხარჯვის მოქნილობით, მაშინ როდესაც შინამეურნეობები მაღალი შემოსავლით აღინიშნებიან ხარჯვის მოქნილობით. ეს ბუნებრივიცაა, რადგან დაბალშემოსავლიანი შინამეურნეობების დანახარჯები ძირითადად დაკავშირებულია აუცილებელ საჭიროებებთან როგორცაა: საკვები, ტანსაცმელი, ტრანსპორტი და სხვა, რის გამოც ეს დანახარჯები ერთმანეთისაგან მნიშვნელოვნად არ განსხვავდებიან. მეორეს მხრივ, მდიდარი ოჯახების მიერ გაწეული დანახარჯები ბევრად მოქნილია. შედეგად ვიღებთ, რომ ფაქტობრივად, მოხმარების მაჩვენებლები შეიძლება საკმარისად განსხვავდებოდნენ ერთმანეთისაგან. უფრო ზუსტად, მაღალი შემოსავლის შინამეურნეობებს ექნებათ მაღალი დისპერსია და დაბალი შემოსავლის შინამეურნეობებს ექნებათ დაბალი დისპერსია. აქედან გამომდინარეობს, რომ დიაგრამაზე დატანილი წერტილები დაბალშემოსავლიანი შინამეურნეობებისათვის შედარებით ახლო დალაგდებიან რეგრესიის წრფის მიმართ და უფრო გაფანტულები აღმოჩნდებიან მაღალშემოსავლიანი შინამეურნეობების რეგრესიის წრფის მიმართ. ეს სიტუაცია განიხილება როგორც ჰეტეროსკედასტიურობა. აღნიშნული სიტუაცია წარმოდგენილია ნახაზებით 7.1. და 7.2..

მოხმარება



ნახ.7.1. ჰეტეროსკედასტიურობა შემოსავლისა და მოხმარების მიხედვით ხელფასები



ნახ.7.2. ჰეტეროსკედასტიურობა ხელფასებისა და პროფესიონალიზმის მიხედვით

განვიხილოთ უმცირეს კვადრატა მეთოდის გამოყენება რეგრესიის განტოლების კოეფიციენტების შეფასებისას ჰეტეროსკედასტიურობის არსებობის პირობებში.

ვთქვათ, მოდელია:

$$Y_t = S_1 + S_2 X_{t2} + \dots + S_k X_{tk} + e_t$$

სადაც: $\text{var}(e_t) = \sigma_t^2$, $t = 1, 2, \dots, n$. ცვლილება მხოლოდ იმაში მდგომარეობს, რომ t -ს სხვადასხვა მნიშვნელობების დროს ვარიაციების შეცდომები განსხვავებულია და მასთან ერთად, უცნობია. რადგან უმცირეს კვადრატა მეთოდის გამოყენებისას აუცილებელი მოთხოვნაა რომ e_t -ს საშუალო მნიშვნელობა ნულია და კორელირებული არ არის X_t -თან, ამიტომ ჰეტეროსკედასტიურობის არსებობის პირობებში მტკიცება იმისა, რომ უმცირეს კვადრატა მეთოდის გამოყენება რეგრესიის განტოლების კოეფიციენტების შეფასებისას ეფექტურია, ჭეშმარიტი არ არის.

ვთქვათ იგნორირება გავუკეთეთ ჰეტეროსკედასტიურობას, როდესაც მას ადგილი აქვს და უმცირეს კვადრატა მეთოდი გამოვიყენეთ რეგრესიის კოეფიციენტების შესაფასებლად. როგორი იქნება მათი თვისებები? იქნება სტატისტიკური ტესტები ძალაში? არის თუ არა პროცედურა, რომელიც გვიკარნახებს ჰეტეროსკედასტიურობის გათვალისწინების აუცილებლობას?

ჰეტეროსკედასტიურობის არსებობის შემთხვევაში ორ კითხვაზე პასუხის გაცემა მაინც მოგვიწევს, კერძოდ:

1. რა გავლენა აქვს ჰეტეროსკედასტიურობას ჰიპოთეზების ტესტირებაზე?

2. რა გავლენა აქვს ჰეტროსკედასტიურობას ეკონომიკური ცვლადების პროგნოზირებაზე?

7.1. ჰეტროსკედასტიურობის ეფექტი ჰიპოთეზების პროგნოზირების ტესტებზე

უმცირეს კვადრატთა მეთოდის გამოყენებისას მიღებული რეგრესიის კოეფიციენტების ვარიაციისა და კოვარიაციის შეფასებები წანაცვლებული და წინააღმდეგობრივია, როდესაც ჰეტროსკედასტიურობა სახეზეა და იგნორირებულია. რის გამოც, ტესტები ჰიპოთეზების შემოწმებაზე დიდხანს ვერ იქნებიან ძალაში.

უმცირეს კვადრატთა მეთოდის გამოყენებისას მიღებული რეგრესიის კოეფიციენტების შეფასებები წაუნაცვლებადია და პროგნოზირებაც დაფუძნებული ამ შეფასებებზე, ასევე წაუნაცვლებადია. მაგრამ რადგან შეფასებები არაეფექტურია ჰეტროსკედასტიურობის არსებობის გამო, ასევე არაეფექტურია პროგნოზირების შედეგიც.

ამის დემონსტრირება შეიძლება შემდეგი მოსაზრებით: განვიხილოთ მოდელი

$$y_t = Sx_t + e_t, \text{var}(e_t) = \sigma^2 = \sigma^2 z_t^2$$

სადაც,

z_t -ს მნიშვნელობა ცნობილია. მუდმივი წევრი არ მონაწილეობს, რადგან ცვლადები წარმოდგენილია როგორც შესაბამისი მნიშვნელობის წარმოებული. გავამარტივოთ და გამოვიყენოთ ჰეტროსკედასტიურობის ფორმა, კერძოდ ის რომ შეცდომის სტანდარტული გადახრა პროპორციულია ცნობილი z_t ცვლადისა. ეს შედეგი გამოყენებულია ზოგადი შემთხვევისათვის.

უმცირეს კვადრატთა მეთოდის პროცედურის მიხედვით მიიღება რომ $z_t=1$ და ვიღებთ შეფასებებს:

$$\hat{S} = S_{xy} / S_{xx}; \sigma^2 = (\sum \hat{e}_t^2) / (n-1); \text{var}(\hat{S}) = \sigma^2 / S_{xx}$$

სადაც,

$$S_{xy} = \sum x_t y_t \quad \text{და} \quad S_{xx} = \sum x_t^2$$

თუ ჩავსვამთ y_t -ს მნიშვნელობას მოდელიდან, მივიღებთ:

$$\hat{S} = \frac{\sum x(Sx_t + e_t)}{S_{xx}} = S + \frac{\sum x_t e_t}{S_{xx}}$$

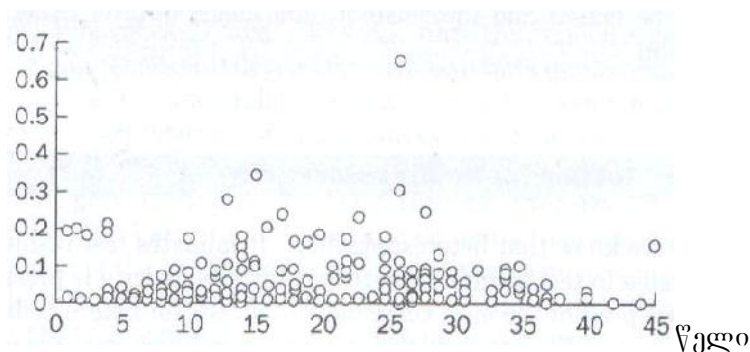
რადგან x_t მოცემულია და $E(e_t) = 0$, მივიღებთ $E(\hat{S}) = S$ და აქედან გამომდინარეობს, რომ უმცირეს კვადრატთა მეთოდით შეფასება ჯერ კიდევ წაუნაცვლებადია. ასევე, $E(x_t e_t) = 0$ და ბუნებრივია, დიდ რიცხვთა კანონის თანახმად \hat{S} -ს ალბათობის ზღვარი, როდესაც n მიისწრაფის ∞ -კენ ტოლია S -სი. ამ შემთხვევაში არ გამოგვიყენებია e_t -ს ვარიაცია. ეს კი მიუთითებს იმაზე, რომ ადგილი აქვს ჰეტროსკედასტიურობას. მართალია შედეგი მიღებულია რეგრესიის მარტივი მოდელიდან გამომდინარე ჰეტროსკედასტიურობის კერძო ფორმის შემთხვევაში, მაგრამ ეს შედეგი სამართლიანია მრავლობითი რეგრესიის შემთხვევისთვისაც.

როგორც წესი, მნიშვნელოვანია:

ა/ შევფასოთ ნაშთების ვარიაცია;

ბ/ ავაგოთ გრაფიკი ნაშთების კვადრატებით და იმ ცვლადით, რომელზედაც ეჭვი გვაქვს, რომ ჰეტროსკედასტიურობის მიზეზია. თუ მოდელში რამდენიმე დამოუკიდებელი ცვლადი შედის, იმდენ გრაფიკს ავაგებთ, იმის შესაბამისად, რამდენი ცვლადიც გვაქვს. შეიძლება გრაფიკის აგება \hat{Y}_i -ს მიმართაც.

გრაფიკის მიხედვით შესაძლო ჰეტროსკედასტიურობის შესაფასებლად ვაგებთ წერტილებს სიბრტყეზე ისე როგორც ნაჩვენებია ნახაზზე 7.3. (გვ. 384, Ramu Ramanatan.). როგორც გრაფიკიდან ჩანს, ნაშთების კვადრატების მიხედვით შეინიშნება ჯერ ზრდის ტენდენცია და შემდეგ შემცირების ტენდენცია. ეს კი ცხადად მიუთითებს მოდელში ჰეტროსკედასტიურობის არსებობაზე.



ნახ.7.3. ნაშთების კვადრატებსა და წლების მიხედვით შექმნილ პროფესიონალიზმს შორის დამოკიდებულება

7.2. ჰეტროსკედასტიურობის ტესტირება

ეკონომეტრიკა დაკავშირებულია რაოდენობრივი ან სტატისტიკური მეთოდების გამოყენებით იმ ამოცანების შესწავლასთან, რომლებიც ეკონომიკური პრინციპების ანალიზსა და გააზრებას ემსახურებიან, რადგან იგი აერთიანებს ეკონომიკურ თეორიას სტატისტიკასთან, რათა გაანალიზოს ეკონომიკური დამოკიდებულებები ტესტირების გამოყენებით. ეკონომიკური მოდელი ეძლევა სტუდენტს, რათა მან შეიცნოს ეკონომეტრიკის თეორიის განვითარების პრინციპები რის გამოც, ძირითადი აქცენტი გადატანილია ეკონომეტრიკული მოდელების ეკონომიკურ ინტერპრეტაციასა და გამოყენებაზე. ლოგორც აღვნიშნეთ, ჰეტროსკედასტიურობას მნიშვნელოვანი როლი ენიჭება ეკონომეტრიკული მოდელების აგებისას და პრობლემებს გვიქმნის, როდესაც რეგრესიის განტოლებაში, სხვადასხვა დაკვირვებებში ცვლადების მნიშვნელობები საგრძნობლად განსხვავდებიან ერთმანეთისაგან. ჰეტროსკედასტიურობამ ასევე, შეიძლება თავი იჩინოს დროითი მწკრივების ანალიზისას იმ შემთხვევებში, როდესაც შემთხვევითი წევრის დისპერსია დროში ზრდას განიცდის.

რეგრესიის მოდელებისათვის მიღებული შედეგები ცხადად ან მიახლოებით მაინც მიუთითებს ჰეტროსკედასტიურობის პირობებზე, ზოგიერთი შედეგი დამოკიდებულია მოსალოდნელობასა და იმ ფაქტზე, რომ ეს შედეგები ნორმალურადაა განაწილებული. ხშირად გვეჩვენება, რომ პრაქტიკაში ჰეტროსკედასტიურობა და ნორმალური განაწილება ორივე დარღვეულია.

თუმცა რთული არ არის შევავასოთ რეგრესიის მოდელი შეწონილი უმცირეს კვადრატთა მეთოდით, როდესაც პირობა დარღვეულია ან შევავასოთ მოდელი მაქსიმალური დასაჯერისობის მეთოდით როდესაც სკედასტიური ფუნქციის პარამეტრები ცნობილი არ არის, თუმცა მათი ფორმა ცნობილია.

ჰეტეროსკედასტიურობის პრობლემა შეიძლება წინასწარ განვჭვრიტოთ მონაცემთა ხასიათის მიხედვით, რამაც შეიძლება თავიდან აგვაცილოს ფორმალური შემოწმების პროცედურა. თანამედროვე პირობებში შემოთავაზებულია ტესტების მრავალფეროვანი სპექტრი და რასაკვირველია, მათთვის სათანადო კრიტერიუმები. განვიხილოთ ზოგიერთი მათგანი, რომლებშიც გათვალისწინებულია სხვადასხვა მოსაზრებები შემთხვევითი წევრის დისპერსიისა და საკვლევი ფუნქციის - y -ის შესახებ, მაგრამ სასარგებლო და მიზანშეწონილია ჰეტეროსკედასტიურობის ტესტების გამოყენებისას, თუ გუპაუხებით კითხვებს:

1. ცხადია თუ არა სპეციფიკაციის შეცდომები ნებისმიერ შემთხვევაში?
2. ნამდვილად აისახება თუ არა ჰეტეროსკედასტიურობა?
3. უეჭველად მიუთითებს ჰეტეროსკედასტიურობაზე ნაშთების გრაფიკული გამოსახვა?
4. რა საშუალებებია ამ პრობლემის თავიდან ასაცილებლად?

1. ჰეტეროსკედასტიურობის აღმოჩენა;
2. ტესტირება ტესტების გამოყენებით:
 - ა/ გოლდფელდ-კვანდტის;
 - ბ/ ბრეუშ-პაგანის;
 - გ/ბრეუშ-გოდფრის;
 - დ/უაიტის;
 - ე/სპირმენის;
 - ვ/გლეიზერის

ბუნებრივია, ჰეტეროსკედასტიურობაზე მინიშნების (აღმოჩენის) შემთხვევაში გადავდივართ ტესტების გამოყენებაზე. ტესტების ფართო სპექტრიდან, რომელსაც პროგრამა ითვალისწინებს, ამჯერად, ყურადღებას ორ ტესტზე - გოლდფელდ-კვანდტისა და ბრეუშ - პაგანის ტესტებზე შევანერებთ, რომელთა გამოყენება, შედარებით მარტივია გამოთვლითი პროცესების გამო და სტუდენტები მეტ ყურადღებას უთმობენ მათ პრაქტიკული გამოყენების თვალსაზრისით.

ა/ გოლდფელდ-კვანდტის ტესტის გამოყენება

ეს ტესტი ყველაზე პოპულარულია და მისი გამოყენებისას გათვალისწინებულია, რომ საშუალო კვადრატული გადახრა - T მოცემულ დაკვირვებაში X -ის მნიშვნელობების პროპორციულია. ასევე, გათვალისწინებულია, რომ შემთხვევითი წევრის განაწილება ნორმალურია და ავტოკორელაციას ადგილი არ აქვს. n - დაკვირვებათა ყველა მნიშვნელობა მოწესრიგებული სახით ჩაიწერება X -ის მნიშვნელობების მიხედვით, ხოლო შემდეგ, შეფასდება რეგრესიები პირველი n' და ბოლო n' დაკვირვებებისათვის. შუალედური დაკვირვებები - $(n - 2n')$ იგნორირდება. იმ შემთხვევაში, თუ მოსაზრება ჰეტეროსკედასტიურობის ბუნების მიმართ ჭეშმარიტია, მაშინ დისპერსია ბოლო n' დაკვირვებაში მეტი იქნება პირველ n' დაკვირვებასთან შედარებით. ეს შედეგი აისახება ორივე, კერძო რეგრესიის სხვაობების ჯამის კვადრატებში. თუ მათ პირველ n' დაკვირვებაში აღვნიშნავთ შესაბამისად -

RS_1, RS_2 -ით და გამოვთვლით შეფარდებას RS_2 / RS_1 , მას ექნება F განაწილება $(n' - k - 1)$ თავისუფლების ხარისხით, სადაც k – რეგრესიის განტოლების ცვლადების რიცხვია.

კრიტერიუმის სიმძლავრე დამოკიდებულია n' -ის შერჩევაზე n -დან. გოლდფელდისა და კვანდტის მტკიცებით n' უნდა შეესაბამებოდეს 11-ს, თუ $n = 30$ და 22-ს, თუ $n = 60$ (თუმცა როგორც პრაქტიკა გვიჩვენებს ეს მონაცემები უფრო მიახლოებითია და დასაზუსტებლად მათთვის სპეციალური გამოსახულებები შეირჩევა). თუ მოდელში ერთზე მეტი ცვლადია, მაშინ დაკვირვების შედეგების მოწესრიგება იმ ცვლადის მიხედვით ხდება, რომელიც \hat{T}_i -სა და n' -თან არის დაკავშირებული და მეტია, ვიდრე $k + 1$, სადაც k – ცვლადების რიცხვია.

აღნიშნული ტესტი შეიძლება გამოვიყენოთ ჰეტეროსკედასტიურობის შესამოწმებლად იმ შემთხვევაში, როდესაც \hat{T}_i უკუპროპორციულია x_i .

მაგალითი:

მოცემულია ოჯახების დღიური შემოსავლები და ხარჯვის შესახებ მონაცემები ლარებში:

N	დანახარჯი, Y	შემოსავალი, X
1	1,80	5,00
2	2,00	5,00
3	2,00	5,00
4	2,00	5,00
5	2,10	5,00
6	3,00	10,00
7	3,20	10,00
8	3,50	10,00
9	3,50	10,00
10	3,60	10,00
11	4,20	15,00
12	4,20	15,00
13	4,50	15,00
14	4,80	15,00
15	5,00	15,00
16	4,80	20,00
17	5,00	20,00
18	5,70	20,00
19	6,00	20,00
20	6,20	20,00

ამ მონაცემების მიხედვით გამოვეთ ოჯახების ორი ჯგუფი: დაბალი და მაღალი შემოსავლებით. დაბალშემოსავლიან ოჯახებს მივაკუთვნოთ ოჯახები, რომელთა შემოსავალი 5 – 10 ლარია და მაღალშემოსავლიან ოჯახებს ის ოჯახები, რომელთა შემოსავალი 10 ლარზე მეტია. მიღებული ჯგუფების მიხედვით ავაგოთ რეგრესიის განტოლებები ოჯახების მიერ დღიური დანახარჯების შესასწავლად.

1. დაბალშემოსავლიანი ოჯახებისათვის რეგრესიის განტოლება და შეფასებებია:

$$Y_t = 0.600 + 0.276X_t$$

$$t_a = 3.1; t_b = 11.3; R^2 = 0.94; e_1 = 0.300$$

2. მაღალშემოსავლიანი ოჯახებისათვის რეგრესიის განტოლება და შეფასებებია:

$$Y_t = 1.54 + 0.20 = X_t$$

$$t_a = 1.14; t_b = 3.1; R^2 = 55; e_2 = 2.02_4$$

გამოთვლილი სხვაობების კვადრატების ჯამების მიხედვით $\hat{Y}_i - \bar{Y}$:

$$e = \sum (Y_i - \hat{Y})^2 \text{ - მოკლებნოთ შეფარდება: } e_2 / e_1 = 6.7$$

F - განაწილების ცხრილის მიხედვით კრიტიკული მნიშვნელობა 5%-იანი მნიშვნელოვნების დონისათვის შეესაბამება 3.44. ამ შედეგის მიხედვით შეგვიძლია დავასკვნათ, რომ ნულოვანი ჰიპოთეზა ჰეტეროსკედასტიურობის ალტერნატიული ჰიპოთეზის სასარგებლოდ, მისაღებია.

ბ/ ბრეუშ - პაგანის ტესტის გამოყენება

განვიხილოთ სამომხმარებლო მოდელი:

$$Y_t = r + sX_t + v_t$$

$$t_i^2 = f(x + uZ_i)$$

ჰეტეროსკედასტიურობის შესამოწმებლად ჯერ გამოთვლება \hat{V}_t -ის მნიშვნელობები უმცირეს კვადრატთა მეთოდით პირველი განტოლებიდან და შემდეგ გამოვიყენებთ მას შეფასების მიზნით: :

$$t^2 = (\sum \hat{V}_i^2) / N$$

შემდეგ განვიხილავთ რეგრესიის განტოლებას:

$$\hat{V}_i^2 / t^2 = x + uZ_i + \epsilon_i$$

და მოკლებნით სხვაობების კვადრატების ჯამს:

$$e = \sum (\hat{Y}_i - \bar{Y})^2$$

რომელიც ექვემდებარება t^2 - განაწილებას. შემდეგ კი გამოითვლება ბრეუშ - პაგანის სტატისტიკა:

$$BP = e / 2t^4$$

ნულოვანი ჰიპოთეზა უარყოფილი იქნება და შეიძლება დავასკვნათ, რომ ჰეტეროსკედასტიურობას ადგილი აქვს, თუ ბრეუშ-პაგანის სტატისტიკის მნიშვნელობა მეტია კრიტიკულ მნიშვნელობაზე

$BP > t_c^2$ მოცემული მნიშვნელოვნების დონისათვის.

7.3. ჰიპოთეზა ჰეტეროსკედასტიურობის შესახებ

ვთქვათ, ϵ_i - შემთხვევითი წევრის სტანდარტული გადახრაა i - ურ დაკვირვებაში. თუ ϵ_i - ცნობილი იქნებოდა თითოეული დაკვირვებისათვის, მაშინ შესაძლებელი იქნებოდა ჰეტეროსკედასტიურობის აცილება, თუ თითოეული დაკვირვების შედეგს გაყოფდით მის შესაბამის ϵ_i - ს მნიშვნელობაზე. შედეგად, შემთხვევითი წევრი i - ურ დაკვირვებაში შეესაბამება ϵ_i / ϵ_i და მისთვის თეორიული დისპერსია ჩაიწერება როგორც:

$$E\{\epsilon_i / \epsilon_i\}^2 = \frac{1}{\epsilon_i^2} (\epsilon_i^2) = 1$$

თითოეულ დაკვირვებაში იქნება შემთხვევითი წევრი, რომელიც მიღებულია გენერალური ერთობლიობიდან დისპერსიით 1 და მოდელი ჰეტეროსკედასტიური იქნება. თვით მოდელს კი ექნება შემდეგი სახე:

$$\frac{y_i}{\epsilon_i} = \frac{a}{\epsilon_i} + \beta \frac{x_i}{\epsilon_i} + \frac{\epsilon_i}{\epsilon_i}$$

ვთქვათ, რაიმე Z პროპორციულია ϵ_i და $Z_i = \beta \epsilon_i$, სადც, β მუდმივია, მაშინ:

$$\frac{Y_i}{Z_i} = \frac{a}{Z_i} + \beta \frac{x_i}{Z_i} + \frac{\epsilon_i}{Z_i}$$

$$E(\epsilon_i / Z_i)^2 = E\{\epsilon_i / \beta \epsilon_i\}^2 = \frac{1}{\beta^2} \epsilon_i^2 / \epsilon_i^2 = \frac{1}{\beta^2}$$

მიღებული სიდიდე მუდმივია ყველა დაკვირვებისათვის და პრობლემაც მოწესრიგებულია.

მაგალითი.

ვთქვათ, მიზანშეწონილია მივიღოთ, რომ ϵ_i - პროპორციულია x - ის მნიშვნელობისა ისე, როგორც გოლდფელდ-კვანდტის კრიტერიუმი ითვალისწინებს. თუ ამ პირობის შემდეგ თითოეული დაკვირვების შედეგს გაყოფთ x - ის შესაბამის მნიშვნელობაზე, მაშინ, $y = r + \beta x + \epsilon$ იქნება:

$$\frac{y}{x} = r \frac{1}{x} + \beta \frac{x}{x} + \frac{\epsilon}{x}$$

და უკეთეს შემთხვევაში, შემთხვევითი წევრს $\frac{\epsilon}{x}$ მუდმივი დისპერსია ექნება.

შემდეგ შეფასდება დამოკიდებულება $\frac{y}{x}$ და განტოლებაში ჩაერთვება მუდმივი წევრი. $\frac{1}{x}$ - ის კოეფიციენტი r - ს ეფექტური შეფასება იქნება, ხოლო მუდმივი წევრი β - ს ეფექტური შეფასება.

თუ განვიხილავთ განათლებაზე გაწეულ დანახარჯებს, სახელმწიფოს მიერ გაწეული დანახარჯები განათლებაზე დამოკიდებული ცვლადია, როგორც მთლიანი ეროვნული პროდუქტის წილი, ხოლო დამოუკიდებელი ცვლადი - მთლიანი ეროვნული პროდუქტის შებრუნებული სიდიდე. თუ საქმე გვაქვს რამდენიმე ცვლადთან, თითოეული მათგანი შეიძლება გამოვიყენოთ განტოლების მასშტაბირებისათვის. ამ მაგალითში ალტერნატიული ცვლადია სახით წარმოვიდგება ქვეყანაში მოსახლეობის რიცხოვნება - n , $\ln n$ განტოლება:

$$\frac{y}{n} = r \frac{1}{n} + s \frac{x}{n} + \frac{v}{n}$$

შეესაბამება დანახარჯებს განათლებაზე და ვიმედოვნებთ, რომ შემთხვევით

წევრს $\frac{v_i}{n_i}$ ყველა დაკვირვებაში მუდმივი დისპერსია აქვს. ამ შემთხვევაში კი

შეფასებას ექვემდებარება ერთეული მოქალაქის განათლებაზე სახელმწიფოს გასაწეობა და ერთ სულ მოსახლეზე ერთობლივ ეროვნულ პროდუქტს შორის დამოკიდებულების რეგრესიის მოდელი.

მაგალითად, ერთობლივ ეროვნულ პროდუქტის შემოსავლიდან განათლებაზე სახელმწიფოს მიერ გაწეული დანახარჯების რეგრესიული დამოკიდებულება შეიძლება აიგოს ორ ვარიანტში:

a/ ერთობლივ ეროვნულ პროდუქტზე გაყოფით;

b/ ქვეყნის რიცხოვნებაზე გაყოფით. მოსახლეობის რიცხოვნების მონაცემები ერთ სულ მოსახლეზე ერთობლივ ეროვნულ პროდუქტისა და განათლებაზე სახელმწიფოს მიერ გაწეული დანახარჯების შესახებ მონაცემების მიხედვით /6/ მივიღებთ:

$$\frac{Z}{L} = -0.066 \frac{1}{L} + 0.053; \quad v_a = 0.084; v_b = 0.004; \\ R^2 = 0.15; F = 0.48.$$

$$\frac{Z}{n} = -0.022 \frac{1}{n} + 0.062 \frac{L}{n}; \quad v_a = 0.057; v_b = 0.003; \\ R^2 = 0.83; F = 160.9.$$

პირველ შემთხვევაში $RS_1/RS_2 = 1.37$. ეს მნიშვნელობა არც თუ ისე დიდია, რომ ის მიუთითებდეს ჰეტეროსკედასტურობაზე. მეორე შემთხვევაში კი $RS_2/RS_1 = 4.60$, რაც მიუთითებს, რომ ნულოვანი ჰიპოთეზა ჰეტეროსკედასტურობის შესახებ უარყოფილი უნდა იყვეს 5%-იანი მნიშვნელოვნების დონით. F – კრიტერიუმის მნიშვნელობაა - 2.98.

რეზიუმე

რეგრესიული მოდელებისათვის მიღებული შედეგები ცხადად ან მიახლოებით მაინც მიუთითებს იმაზე, რომ ჰეტეროსკედასტიკურობასთან გვაქვს საქმე. ზოგიერთი შედეგი დამოკიდებულია მოსალოდნელობაზე რომ ეს შედეგები ნორმალურადაა განაწილებული. ხშირად გვეჩვენება, რომ პრაქტიკაში ჰეტეროსკედასტიურობა და ნორმალური განაწილება ორივე დარღვეულია. თუმცა რთული არ არის შევაფასოთ რეგრესიის მოდელი შეწონილი უმცირეს კვადრატთა მეთოდით, როდესაც პირობა დარღვეულია ან შევაფასოთ მოდელი მაქსიმალური დასაჯერისობის მეთოდით როდესაც სკედასტიური ფუნქციის პარამეტრები ცნობილი არ არის, თუმცა მათი ფორმა ცნობილია. როდესაც პრაქტიკულად ჰეტეროსკედასტიურობის გათვალისწინება და მისი ტესტირება გვიხდება, უნდა გვქონდეს პასუხები კითხვებზე:

1. ნებისმიერ შემთხვევაში ცხადია თუ არა სპეციფიკაციის შეცდომები?
2. კვლევის შედეგებზე ნამდვილად სარძნობლად აისახება ჰეტეროსკედასტიურობა?
3. ნაშთების გრაფიკული გამოსახვა უეჭველად მიუთითებს ჰეტეროსკედასტიურობაზე?
4. რა საშუალებებს ვფლობთ პრობლემის ასაცილებლად?

ტესტებიდან, რომლებიც ფოკუსირებულნი არიან ჰეტეროსკედასტიურობის გამოვლენაზე, ყურადღება შენერებულია ორ ცნობილ ტესტზე: გოლფელდ-კვანდტისა და ბრეუმ-პაგანის ტესტებზე და მათ გამოყენებაზე პრაქტიკული ამოცანების ამოხსნისას გამოთვლით პროცესების სიმარტივის გამო.

დამოუკიდებელი სამუშაო

1. მიეცით ახსნა –განმარტება ჰეტეროსკედასტიურობის შემთხვევას თქვენეული გაგებით და ახსენით თქვენი მოსაზრება მაგალითზე დაყრდნობით, რამდენადაც ეს შესაძლებელია თქვენთვის.
2. ქვემოთ მოყვანილი ცხრილის მიხედვით მოცემულია ინფორმაცია ბინის მეპატრონეების შემოსავლისა და მოხმარების შესახებ *

შემოსავალი, \$	საშ. შემოსავალი, \$	მოხმარების საშუალო, \$
0 -999	556	2760
1000 – 1999	1622	1930
2000 – 2999	2664	2740
3000 – 3999	3587	3515
4000 - 4999	4535	4350
5000 - 5999	5538	5320
6000 – 7499	6585	6250
7500 – 9999	8582	7460
10000 - ზემოთ	14033	11500

1. ააგეთ რეგრესიის მოდელი და ეცადეთ ახსნათ საშუალო მოხმარება როგორც საშუალო შემოსავლის ფუნქცია.
2. ჰეტეროსკედასტიურობის ასახსნელად განტოლებაში მიღებული ნაშთებისათვის გამოიყენეთ ტესტი.
3. ცხრილში მოცემულია კების პროდუქტებზე დანახარჯების - y, ოჯახის შემოსავლებისა -x და ოჯახში სულთა რაოდენობის - n დაკვირვების შედეგები.

N	y, \$	x, \$	n
1	15.998	62.476	1
2	16.652	82.304	5
3	21.741	74.679	3
4	7.431	39.151	3
5	10.481	64.724	5
6	13.548	36.786	3
7	23.256	83.052	4
8	17.976	86.935	1
9	14.161	88.233	2
10	8.825	38.695	2
11	14.184	73.831	7

* Alberto Ando and Franco Modigliani. “The Permanent Income and Life Life Cycle. Hypotheses of Saving Behavior: Comparisons and Tests” in I. Friend and R. Jones, Consumption and Saving, Vol.II, 1960, p.154

12	19.604	77.122	3
13	13.728	45.519	2
14	21.141	82.251	2
15	17.446	59.862	3
16	9.629	26.563	3
17	14.005	61.818	2
18	9.160	29.682	1
19	18.831	50.825	5
20	7.641	71.062	4
21	13.882	41.990	4
22	9.670	37.324	3
23	21.604	86.352	5
24	10.866	45.506	2
25	28.980	69.929	6
26	10.882	61.041	2
27	18.561	82.49	1
28	11.629	44.208	2
29	18.067	49.467	5
30	14.539	25.905	5
31	19.192	79.178	5
32	25.918	75.811	3
33	28.833	82.718	6
34	15.869	48.311	4
35	14.910	42.494	5
36	9.550	40.573	4
37	23.066	44.872	6
38	14.751	27.167	7

განვიხილოთ სტატისტიკური მოდელი:

$$y_t = S_1 + S_2 x_t + S_3 n_t + e_t$$

სადაც e_t დამოუკიდებელი, შემთხვევითი შეცდომაა ნულოვანი საშუალო მნიშვნელობით და ვარიაციით:

$$\text{Var}(e_t) = r_1 + r_2 x_t + r_3 n_t$$

შეამოწმეთ ჰეტეროსკედასტიურობა და გამოიყენეთ:

ა/ გოლდფელდ-კვანდტის ტესტი და ცხრილის მონაცემები X-ის მიმართ;

ბ/ ბრეუშ-პაგანის ტესტი $\text{Var}(e_t)$ -ს განტოლებით როგორც ალტერნატიული ჰიპოთეზა.

ახსენით მიღებული შედეგები.

თავი 8. დროითი მწკრივები

დროითი მწკრივები დინამიკაში წარმოდგენილი პროცესია გარკვეულ ცვლადებზე დაკვირვებების რეალიზაციით დროში. მაგალითად, ასეთი პროცესებია სეისმურ ცვლადებზე, ძაბვის ცვლილებებზე, ჰაერის ტემპერატურაზე დაკვირვებების რეალიზაციები. ასევე, დანახარჯები რაიმე საქონლის (პროდუქტის) წარმოებაზე, გასავალი საცხოვრებელ ფართზე და სხვა.

დროითი მწკრივების ანალიზისათვის საკმარისი არ არის ცვლადზე (ცვლადებზე) მიმდინარე დაკვირვებების რაელიზაციები, რადგან მათზე მოქმედებენ გასული (წინა) პერიოდების შედეგებიც. მაგალითად, დანახარჯების მოცულობა პროდუქტის წარმოებაზე განისაზღვრება არა მხოლოდ მიმდინარე შემოსავლისა და ამ პროდუქტზე მიმდინარე ფასის მიხედვით, არამედ შემოსავლითა და ფასებით გასული (წინა) პერიოდების შედეგების გათვალისწინებით. ასევე, საცხოვრებელ ფართზე დანახარჯის შესწავლისას გასათვალისწინებელია მიმდინარე შემოსავალი და ფასები, რადგან მათი მნიშვნელობები დამოკიდებულია გასულ (წინა) პერიოდებზე.

ეკონომიკური ცვლადების შესწავლისას დინამიკაში, მათ შესახებ დაკვირვებების მონაცემების მიხედვით, ერთ-ერთი მიზანია ეკონომიკური ცვლადების მნიშვნელობების პროგნოზირება, რომლისთვისაც შეიძლება გამოვიყენოთ ეკონომეტრიკული მოდელის აგება. ამ შემთხვევაში აუცილებელია შემდეგი:

1. ეკონომიკური მოდელის ფორმულირება, რომელიც მოიცავს დამოკიდებულ ცვლადზე მოქმედ დამოუკიდებელი ეკონომიკური ცვლადების დადგენას;
2. სტატისტიკური მოდელის აგება, რომლის მიხედვითაც ხდება მონაცემების გენერირება;
3. შერჩევის საფუძველზე პროცედურის განხილვა სტატისტიკური მოდელის უცნობი პარამეტრების შესაფასებლად, რომლებიც შემდეგში გამოიყენება პროგნოზირებისათვის.

მაგრამ ამჯერად, გამოყენებული იქნება სხვადასხვა ტექნიკური მიდგომა მოკლევადიანი პროგნოზირებისათვის ნაცვლად იმისა, რომ ავაგოთ ეკონომიკური და სტატისტიკური მოდელები. ეს არის დროითი მწკრივების ანალიზის მეთოდი, რომელიც საშუალებას გვაძლევს გავაანალიზოთ ცვლადების მიმდინარე მნიშვნელობები და ამ ცვლადების მნიშვნელობები გასულ (წინა) პერიოდში.

დროითი მწკრივების შესწავლისას გასათვალისწინებელია, რომ, ისინი არ ერგებიან ეკონომიკური მოდელის ტრადიციულ ჩარჩოს და ეკონომიკურ ცვლადებს შორის ურთიერთდამოკიდებულების შესასწავლად აუცილებელია სპეციფიკური მიდგომა.

ზოგადად, ეკონომიკური დროითი მწკრივების შეწავლისას მიღებულია, რომ დროითი მწკრივი მოიცავს კომპონენტებს:

1. ტრენდი: ეს არის ცვლადის მოძრაობა სპეციფიკური მიმართულებით, რომელიც გრძელდება ხანგრძლივი დროის განმავლობაში;
2. ციკლი: ეს არის ცვლადის კვაზიპერიოდული მოძრაობა აღმავლობით ან დადმავლობით;
3. სეზონურობა: ეს არის ცვლადის რეგულარული (კვირეული, თვიური, კვარტალური და სხვა) ცვლილება სეზონის მიხედვით;
4. შემთხვევითი: ეს არის ნებისმიერი დროითი მწკრივის დაკვირვების რეალიზაციის თანმდევი რეალიზაცია სტაბილური ალბათობით.

დროითი მწკრივების მიხედვით ძირითადი შესასწავლი ამოცანაა პროცესის ცვლილების განსაზღვრა და ამ ცვლილების მიმართულებისა და სისწრაფის დადგენა. დროითი მწკრივების კვლევისას შეიძლება გამოვიყენოთ სამი ძირითადი პრობლემა;

1. დროში მაჩვენებლების ცვლილების აღწერა და ძირითადი ტენდენციის (ტრენდის) დადგენა;
2. მწკრივის ცვლილების დონის ახსნა;
3. მწკრივის ანალიზის შედეგების გამოყენება სტატისტიკური პროგნოზირების მიზნით.

დროითი მწკრივების შესწავლისას აუცილებელია მოვითხოვით მწკრივის მონაცემების შესადარისობის დაცვა, ხოლო რაც შეეხება მწკრივის სიგრძეს, ეს საკითხი შედარებით რთულია, რადგან:

1. ერთის მხრივ, სურვილი გვაქვს გავზარდოთ მწკრივის სიგრძე მაღალი სიზუსტის მიღწევის მიზნით;
2. მეორეს მხრივ, დროითი მწკრივების დამუშავებისას სასურველი არ არის ძველი მონაცემების გამოყენება. აღნიშნული მოთხოვნების გათვალისწინებით მიზანშეწონილია: მწკრივის დაყოფა, „ბიჯის“ შემცირება (მაგალითად, კვარტალური მონაცემებიდან თვიურ მონაცემებზე გადასვლა, თვიური მონაცემებიდან კვირეულ მონაცემებზე გადასვლა და ა.შ), თუ მწკრივის მონაცემები ამის საშუალებას იძლევიან.

შენიშვნა: დროითი მწკრივების გრაფიკული წარმოდგენისას სიფრთხილვა საჭირო საერთო ტენდენციისა და რხევების შესახებ დასკვნისაგან, რადგან დასკვნა შეიძლება მცდარი აღმოჩნდეს საკოორდინატო ღერძების მასშტაბირების გამო.

8.1. ტენდენციის არსი

დროითი მწკრივების ტენდენციის წარმოდგენა, ძირითადად, მრუდით წარმოდგენას უკავშირდება, რომელიც შეესაბამება რომელიმე მათემატიკურ ფუნქციას და მას **ტრენდი** ეწოდება და რომელიც იმ ძირითადად კანონზომიერებას ახასიათებს, რომელიც დროში მანვენებლის ცვლილებას შეესაბამება და შემთხვევითი ზემოქმედებისაგან თავისუფალია. ტრენდიდან გადახრა შეესაბამება შემთხვევით მდგენელს, ე.ი. ტრენდი განისაზღვრება მუდმივ მოქმედი ფაქტორების გათვალისწინებით, ხოლო ტრენდიდან გადახრა – შემთხვევითი ფაქტორების გათვალისწინებით. ამიტომ ცხადია, რომ მწკრივიდან დეტერმინირებული შემთხვევითი სიდიდის გამოყოფა პირობითია.

ზოგადი ტენდენციის განვითარების განსაზღვრის მიზნით გამოიყენება მოსწორების (გაგლუვების) სხვადასხვა მეთოდები როგორცაა **მცოცავი საშუალოს** მეთოდი ან სხვა ანალიზური მეთოდი.

მცოცავი საშუალოს მეთოდი ერთერთი მარტივი მეთოდია და საკმაოდ გაერცვლებულია. მწკრივის მოსწორებისათვის გამოიყენება სამი წევრი და პროცედურა შემდეგში მდგომარეობს:

1. მოცემული **დროითი** მწკრივის - y_1, y_2, \dots, y_n მიხედვით განისაზღვრება მიმდევრობითი სამი წევრისაგან მცოცავი საშუალოს მნიშვნელობა :

$$\bar{Y}_i = \frac{y_{i-1} + y_i + y_{i+1}}{3}, i = 1, 2, \dots, n-1$$

2. მოცემული **დროითი** მწკრივის მნიშვნელობას განსაზღვრავენ ხუთი მიმდევრობითი წევრის მიხედვით:

$$\bar{Y}_i = \frac{y_{i-2} + y_{i-1} + y_i + y_{i+1} + y_{i+2}}{5}, i = 3, 4, \dots, n-2$$

ოთხი მიმდევრობითი წევრის მიხედვით მცოცავი საშუალოს განსაზღვრა ორი ოპერაციისაგან შედგება:

1. საშუალო არითმეტიკულის განსაზღვრა ოთხი მიმდევრობითი წევრის მიხედვით:

$$\bar{Y}_i' = \frac{y_i + y_{i+1} + y_{i+2} + y_{i+3}}{4}, i = 1, 2, \dots, n-3$$

$$\bar{Y}_i'' = \frac{y_{i+1} + y_{i+2} + y_{i+3} + y_{i+4}}{4}, i = 1, 2, \dots, n-4$$

2. მიღებული საშუალოების მნიშვნელობების მიხედვით განისაზღვრება კიდევ ერთი საშუალო სიდიდე:

$$\bar{Y}_i = \frac{\bar{Y}_i' + \bar{Y}_i''}{2}$$

ოთხი წევრის მიხედვით მცოცავი საშუალოს განსაზღვრა დაკავშირებულია შემდეგ გარემოებასთან: მწკრივში წევრთა ლუწი რაოდენობის შემთხვევაში ყველა წევრი მონაწილეობს გაანგარიშებაში, რომლებისთვისაც არსებობს მოცემული მომენტის შესაბამისი ფაქტობრივი დონე.

მცოცავი საშუალოს მეთოდი საკმარისად მარტივი მეთოდია. მის გარდა, ფილტრის თვალსაზრისით შეიძლება გამოვიყენოთ შეწონილი საშუალო სიდიდეები, ექსპონენციალური საშუალო სიდიდეები, მიმდევრობით სხვაობებზე დაფუძნებული და სხვა.

დროითი მწკრივების ანალიზური მოსწორების ძირითადი ტენდენციის დახასიათება ხდება დროის ფუნქციით და მისი შესაბამისი მრუდით. ეს ფუნქცია მიიღება დროითი მწკრივის ანალიზური მოსწორების გზით უმცირეს კვადრატთა მეთოდის გამოყენებით და მწკრივის დონე ჩაიწერება ორი მდგენელით:

$$y_t = f(t) + v_t$$

სადაც,

$f(t)$ - ტრენდის განტოლებაა;

v_t - შემთხვევითი მდგენელი ნულოვანი საშუალო სიდიდით და ვარიაციით - t^2 .

დროითი მწკრივის ძირითადი ტენდენცია გამოსახავს იმ კანონზომიერებას, რომლითაც მწკრივი ხასიათდება წინა პერიოდების მონაცემების მიხედვით და რომელიც მოცემული პერიოდისათვისაც შენარჩუნებულია. თუ გავითვალისწინებთ პროცესის ინერციულობის ხასიათს, უახლოესი მომავლის ანალიზისათვის მიმართავენ ზოგადი ტენდენციის ექსტრაპოლირებას. ახლო მომავლისათვის ყველაზე საიმედო დონე, თუ პროცესის მიმდინარეობაში რაიმე რადიკალურ ცვლილებას ადგილი არ აქვს, შეესაბამება იმ დონეს, რომელიც გამომდინარეობს დადგენილი დონიდან.

საინტერესოა, რომ დროითი მწკრივები პროგნოზირების საშუალებას იძლევიან მაშინაც, როდესაც სტრუქტურული მოდელი უცნობია. პროგნოზირებისათვის დროითი მწკრივების გამოყენება ეფექტურია შემდეგ შემთხვევებში:

1. მოკლევადიანი პროგნოზისათვის;
2. ეკონომეტრიკული მოდელი პროგნოზისათვის ითხოვს დიდ დროსა და ენერგიას;
3. გვაქვს მნიშვნელოვნად ბევრი მონაცემი პროგნოზირებადი ცვლადის შესახებ.

რის გამოც განხილვას ექვემდებარება:

1. ალტერნატიული გზების ძიება დროითი მწკრივების სტატისტიკური მოდელის სპეციფირებისათვის;
2. სტატისტიკური მოდელის პარამეტრების შეფასების პროცედურა;
3. დროითი მწკრივების გამოყენება ეკონომიკური ცვლადის პროგნოზირებისათვის

და შესასწავლია:

1. შერჩევის საფუძველზე დროითი მწკრივიდან მიღებული სტატისტიკური მოდელის მახასიათებლები;
2. ავტორეგრესიისა (AR) და მცოცავი საშუალოს (MA) მოდელების სტატისტიკური მახასიათებლები;
3. მცოცავი საშუალოთი კომბინირებული ავტორეგრესიის მოდელი (ARMA);
4. ბოქს-ჯენკინსის (Box-Jenkins) მოდელის აგება

8.2. სეზონური ეფექტები და მათი ანალიზი

დროითი მწკრივების ანალიზის მოდელი მოიცავს ტრენდს, სეზონურ რხევებს და შემთხვევით მდგენელს :

$$Y_t = f(t) + S(t) + v_t$$

სადაც: $f(t)$ – ტრენდია, $S(t)$ – სეზონური რხევებია და v_t - შემთხვევითი მდგენელი.

აღნიშნული კომპონენტების გარდა, მოდელი შეიძლება შეიცავდეს ციკლის კომპონენტებსაც და მათი გათვალისწინებით მოდელი ჩაიწერება შემდეგი სახით:

$$Y_t = f(t) + S(t) + C(t) + v_t$$

ციკლის კომპონენტი – $C(t)$ ასახავს პროცესის აღმავლობისა და დაცემის პერიოდების ხანგრძლივობას.

უმცირეს კვადრატთა მეთოდის გამოყენებით შეიძლება შევავასოთ ტრენდი და აღვნიშნოთ იგი $\hat{f}(t)$ -თი და თითოეული i -ური სეზონისათვის $1 \leq i \leq p$ განვიხილოთ სხვაობა:

$$f_i(t) - \hat{f}_i(t), f_{i+p} - \hat{f}_{i+p}, \dots$$

რომელიც მოიცავს მთელ რიცხვით პერიოდებს $h = (m + 1) p$. თითოეული ამ გადახრებიდან შეიძლება განვიხილოთ, როგორც სეზონური ცვლილების გავლენის შედეგი.

კომპონენტის - $S(t)$ –ს შეფასების ყველაზე მარტივი ფორმაა საშუალო სიდიდე:

$$\hat{S}(t) = \frac{1}{m+1} \sum (f_{i+kp}(t) - \hat{f}_{i+kp}(t)), i = (1, 2, \dots, p)$$

სასურველია, რომ სეზონური ეფექტების ჯამი ნულის ტოლო იყოს, რისთვისაც მიმართავენ სეზონური ეფექტების კორექტირებულ შეფასებას:

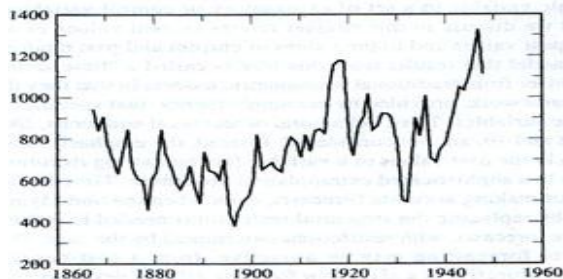
$$S^*_i(t) = \hat{S}_i(t) - \frac{1}{p} \sum \hat{S}_i(t)$$

დროითი მწკრივიდან დეტერმინირებული კომპონენტების: ტრენდის, სეზონურობისა და ციკლური კომპონენტების გამოყოფის შემდეგ დროითი მწკრივი სტაციონალურ პროცესს უნდა წარმოადგენდეს და კვლევის შემდეგი საფეხური სწორდ მისი კვლევაა. ამ ეტაპზე კვლევის მიზანი ითვალისწინებს:

- მწკრივის ასახვას ტრენდის შესაბამისი ამა თუ იმ მოდელით;
- მწკრივის ვარიაციის შეფასების დაზუსტებას;
- სხვაობების სტაციონალურობის შემოწმებას.

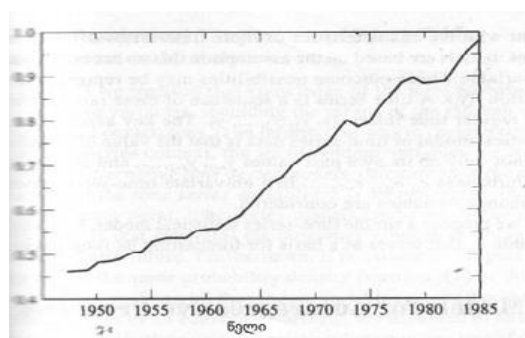
8.3. დროითი მწკრივების სტატისტიკური მოდელების მახასიათებლები

დროითი მწკრივების ანალიზისა და იმ პრობლემების ანალიზისათვის რაც დროითი მწკრივების გამოყენებით ეკონომიკური ცვლადების პროგნოზირების პრობლემებს ეხება, განვიხილოთ გრაფიკები 8.1., 8.2. (William E. Griffiths, R. Carter Hill, George G. Judge, 1997)



ნახ.8.1. მარცვლეულზე ფასების ცვლილება წლების მიხედვით

როგორც ნახაზიდან ჩანს, სხვადასხვა პერიოდების მიხედვით მარცვლეულზე ფასის ცვლილება ხასიათდება პიკებით, ციკლებით დროით მწკრივში. ამ მონაცემებზე დაყრდნობით, ჩვენი მიზანი იქნება მარცვლეულზე ფასის პროგნოზირება ერთი ან ორი ან სამი წლის შემდეგ.



ნახ.8.2. პერსონალური სამომხმარებლო ($x \cdot 10^4$) დანახარჯები წლების მიხედვით

როგორც ნახაზიდან ჩანს, პერსონალური სამომხმარებლო დანახარჯები წლების მიხედვით ხასიათდება აღმავალი ტრენდით. შეიძლება თუ არა ამ ტრენდის მიხედვით გავაკეთოთ პროგნოზი მომავალში პერსონალური სამომხმარებლო დანახარჯების შესახებ?

8.4. სტატისტიკური მოდელი ავტორეგრესიის პროცესისათვის

$x(t)$ შემთხვევითი პროცესი პირველი რიგის ავტორეგრესიის AR(1) პროცესია, თუ მისთვის სრულდება შემდეგი პირობა:

$$X(t) = rX(t-1) + v_t$$

სადაც, r - მუდმივია.

ამ განსაზღვრების საფუძველზე $x(t)$ - ს მნიშვნელობა ნებისმიერ $t > t_0$ მომენტში განსაზღვრულია v_t - თი, თუ t_0 მომენტი ცნობილია. ამ მომენტში ცნობილია $x(t_0)$ - ს მნიშვნელობა.

AR(1) პროცესი ავტორეგრესიის მარტივი შემთხვევაა. იმ შემთხვევაში, თუ ავტორეგრესიაზე მიგვითითებს შემდეგი პირობის შესრულება $Cov(e_t, e_s) \neq 0$ როდესაც $t \neq s$, ეს ნიშნავს, რომ t პერიოდისათვის მოდელის შეცდომა კორელირებულია s პერიოდის შეცდომასთან.

პირველი რიგის ავტორეგრესიისათვის შემდეგი მოთხოვნებია:

$$Y(t) = r + sX_t + e_t \quad (8.1)$$

$$e_t = \dots e_{t-1} + v_t; \quad -1 < \dots < 1 \quad (8.2)$$

როგორც ჩანს, e_t შეცდომა დაკავშირებულია წინა პერიოდის e_{t-1} შეცდომასთან, შეცდომის გამომსახველ ახალ წევრთან v_t და ახალ პარამეტრთან ... -სთან. რადგან ... - კოეფიციენტია ერთი პერიოდით წაძრული შეცდომის გამომსახველი წევრისა, მას პირველი რიგის ავტორეგრესიის კოეფიციენტი ეწოდება, ხოლო პროცესს, რომელიც აღწერილია განტოლებით (8.2) – პირველი რიგის ავტორეგრესიული პროცესი ეწოდება, ან როგორც ის ზოგადად ცნობილია, **AR(1) პროცესი**

$$Y(t) = r + sX_{t-1} + v_t$$

$y(t)$ -ს მნიშვნელობა დროის ნებისმიერ $t > t_0$ მომენტში განსაზღვრულია v_t -თი, თუ t მომენტში ცნობილია $y(t_0)$ -ის მნიშვნელობა. ავტორეგრესიის სტაციონალური პროცესის რიცხვითი მახასიათებლებია: მათემატიკური ლოდინი, კოვარიაცია და კორელაციის კოეფიციენტი, ვარიაცია. **AR(1)** მოდელისათვის გვექნება:

$$EY(t)=0; \text{cov}(y(t), y(t-k)) = E(Y(t)Y(t-k)) = b_k; \quad r_k = \text{corr}(y(t), y(t-k))$$

მეორე რიგის **AR(2)** ავტორეგრესიისათვის სამართლიანია ტოლობა:

$$Y(t) = r_1 Y(t-1) + r_2 Y(t-2) + v_t$$

ამ პროცესისათვის სამართლიანია $EY(t)=0; r_1 + r_2 < 1; r_1 - r_2 < -1; r_2 > -1$ და ნებისმიერი t -თვის ამ პროცესის მახასიათებლებია:

$$E(Y(t)Y(t-k)) = b_k;$$

$$b_1 = r_1 + r_2 b_1; b_2 = r_1 b_1 + r_2 b_2;$$

$$r_1 = r_1 + r_2 r_1; r_2 = r_1 r_1 + r_2 r_2; r_k = r_1 r_{k-1} + r_2 r_{k-2}$$

$$r_1 = \frac{r_1 - r_1 r_2}{1 - r_1^2}; r_2 = \frac{r_2 - r_1^2}{1 - r_1^2}$$

AR(P) პროცესის კოეფიციენტების შესავასებლად ამოიხსნება განტოლებათა სისტემა:

$$r_1 = r_1 + r_2 r_1 + \dots + r_p r_{p-1}$$

$$r_2 = r_1 r_1 + r_2 + \dots + r_p r_{p-2}$$

.....

$$r_p = r_1 r_{p-1} + r_2 r_{p-2} + \dots + r_p$$

$$k = 1, 2, \dots, p$$

ამ სისტემის ამოხსნა გვაძლევს r_1, r_2, \dots, r_p პარამეტრების მნიშვნელობებს და მათ შეფასებას დროითი მწკრივის რეალიზაციის მიხედვით.

მცოცავი საშუალოს პროცესისათვის - **MA(q)** ადგილი აქვს ტოლობას:

$$y(t) = r_1 v_{t-1} + r_2 v_{t-2} + \dots + r_{1q} v_{t-q} + v_t$$

ავტორეგრესიისა და მცოცავი საშუალოს კომბინირებული პროცესის-ARMA(p,q) შემთხვევაში ადგილი აქვს ტოლობას:

$$y(t) = \sum r_i y(t-i) + \sum S_j v_{t-j} + v_t$$

8.5. დარბინ – უატსონის (Durbin-Watson) ტესტი AR(1) პროცესისათვის

მიუხედავად იმისა, რომ გრაფიკული წარმოდგენა კარგი საშუალებაა ავტორეგრესიის იდენტიფიცირებისათვის, მისი ტესტირება არსებითია საბოლოო დასკვნის მისაღებად. პირველი რიგის ავტორეგრესიის ტესტირებისათვის მიზანშეწონილია **დარბინ – უატსონის ტესტი (DW)**, რომელიც დაფუძნებულია ლაგრანჟის მამრავლის მეთოდოლოგიაზე. **დურბინ – უატსონის ტესტი AR(1) პროცესისათვის** აღწერილია შემდეგი რეგრესიის მოდელით:

$$Y_t = S_1 + S_2 X_{t2} + \dots + S_k X_{tk} + e_t$$

$$e_t = \dots e_{t-1} + v_t; -1 < \dots < 1$$

შეფასების პროცედურა შემდეგია:

1. შეფასდება მოდელი უმცირეს კვადრატთა მეთოდით და გამოითვლება სხვაობები \hat{e}_t , როგორც

$$Y_t - \hat{S}_1 - \hat{S}_2 X_{t2} - \dots - \hat{S}_k X_{tk}$$

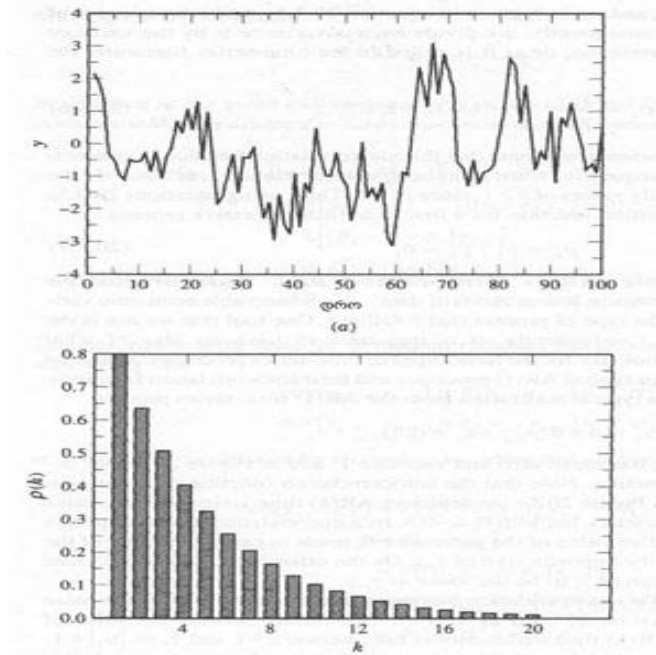
2. გამოითვლება **დარბინ – უატსონის** სტატისტიკა:

$$d = \frac{\sum_{t=2}^{t=n} (\hat{e}_t - \hat{e}_{t-1})^2}{\sum_{t=1}^{t=n} \hat{e}_t^2}$$

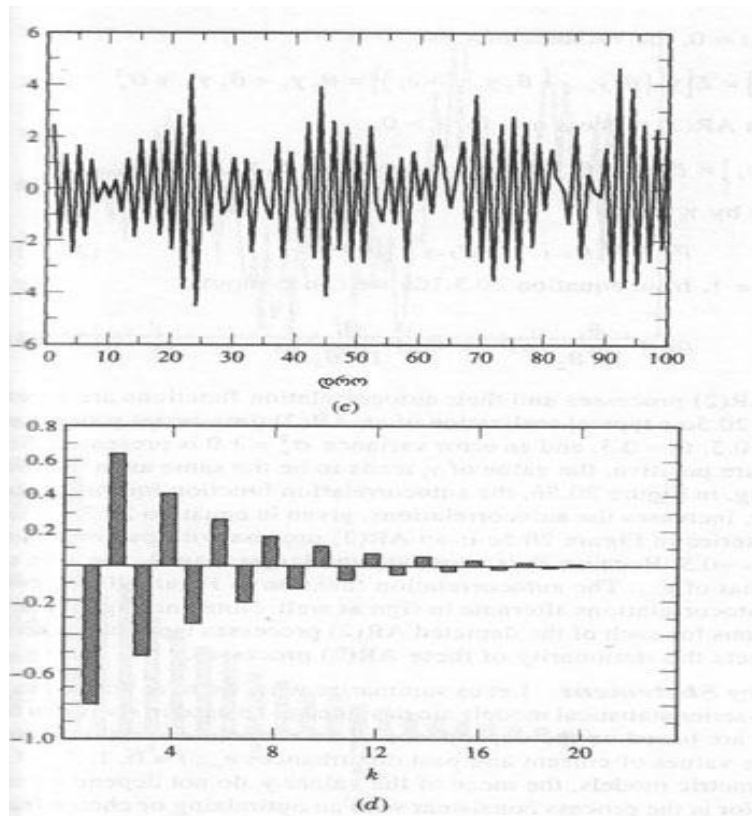
რომლის მნიშვნელობა იცვლება 0-დან 4-მდე

3. რადგან $d \approx 2(1 - \dots)$ და $-1 < \dots < 1$, მაშინ იმ შემთხვევაში, როდესაც $\dots = 0$, $d = 4$. DW სტატისტიკა $d \approx 2$, ეს კი ნიშნავს, რომ ადგილი არ აქვს პირველი რიგის ავტორეგრესიას.

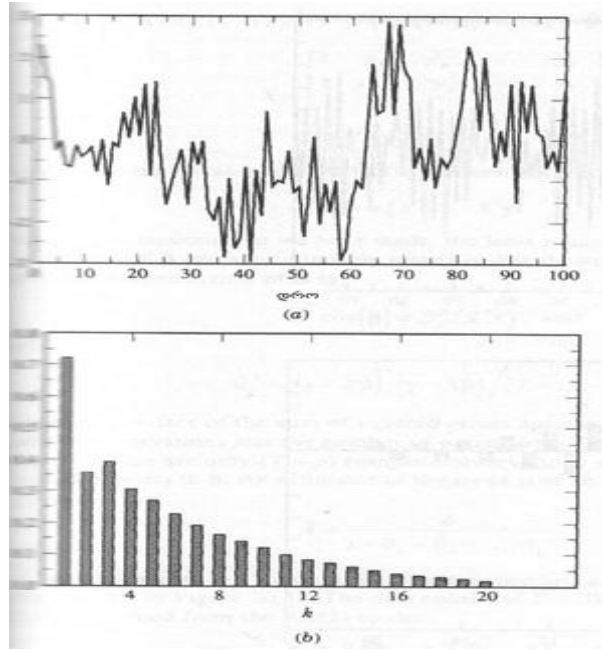
8.6. AR(1) დროითი მწკრივისა და ავტორეგრესიის ფუნქციების მაგალითები



ნახ.8.3. AR(1) დროითი მწკრივი და ავტოკორელაციის ფუნქცია: **a/**რეალიზაცია: $y_t = 0.8y_{t-1} + e_t$; **b/** ავტოკორელაციის ფუნქცია $\rho(k) = 0.8^k$



ნახ.8.4. **c/** რეალიზაცია: $y_t = -0.8y_{t-1} + e_t$ **d/** ავტოკორელაციის ფუნქცია $\rho(k) = (-0.8)^k$



ნახ.8.5. AR(2) დროითი მწკრივი და ავტოკორელაციის ფუნქცია: a/რეალიზაცია: $y_t = 0.5y_{t-1} + 0.3y_{t-2} + e_t$; b/ ავტოკორელაციის ფუნქცია $y_t = 0.5y_{t-1} + 0.3y_{t-2} + e_t$ -თვის

8.7. მცოცხავი საშუალოს პროცესი

ზოგადად, მცოცხავი საშუალოს პროცესი MA(q) განიხილება როგორც ეკონომიკურ ცვლადზე დროითი მწკრივის დაკვირვებების შეწონილი საშუალო. MA(1) პროცესი ჩაიწერება შემდეგი სახით:

$$y_t = \mu + e_t + r_1 e_{t-1}$$

$$E(y_t) = \mu$$

ეს ნიშნავს, რომ

$$\begin{aligned} \text{var}(y_t) &= \sigma_0 = E(y_t - \mu)^2 = \sigma_e^2 (1 + r_1^2) \\ \text{cov}(y_t, y_{t-1}) &= \sigma_1 = E\{(y_t - \mu)(y_{t-1} - \mu)\} = \\ &= E\{(e_t + r_1 e_{t-1})(e_{t-1} + r_1 e_{t-2})\} = r_1 \sigma_e^2 \end{aligned}$$

ანალოგიურად, კოვარიაცია y_t -სა და y_{t-2} -ს შორის იქნება :

$$\begin{aligned} \text{cov}(y_t, y_{t-2}) &= \sigma_2 = E\{(y_t - \mu)(y_{t-2} - \mu)\} = \\ &= E\{(e_t + r_1 e_{t-1})(e_{t-2} + r_1 e_{t-3})\} = 0 \end{aligned}$$

და ყველა ლაგისათვის $k > 1$ კოვარიაციები ნულის ტოლია.

რის გამოც, ავტოკორელაციის ფუნქცია MA(1) პროცესისათვის ჩაიწერება ასე:

$$\rho_k = \frac{\sigma_k}{\sigma_0} = \frac{r_1}{1 + r_1^2}, k = 0$$

$$\rho_k = \frac{\sigma_k}{\sigma_0} = 0, k > 1$$

MA(1) პროცესის სტატისტიკური მოდელი მოცემულია გამოსახულებით:

$$y_t = \mu + e_t + r_1 e_{t-1} + r_2 e_{t-2}$$

ეს ნიშნავს, რომ

$$E(y_t) = \mu$$

$$\text{var}(y_t) = \sigma_e^2 (1 + r_1^2 + r_2^2)$$

$$\begin{aligned} \text{cov}(y_t, y_{t-1}) = \sigma_{11} &= E\{(e_t + r_1 e_{t-1} + r_2 e_{t-2})(e_{t-1} + r_1 e_{t-2} + r_2 e_{t-3})\} = \\ &= r_1 \sigma_e^2 + r_1 r_2 \sigma_e^2 = \sigma_e^2 (r_1 + r_1 r_2) \end{aligned}$$

$$\text{cov}(y_t, y_{t-2}) = \sigma_{12} = E\{(e_t + r_1 e_{t-1} + r_2 e_{t-2})(e_{t-2} + r_1 e_{t-3} + r_2 e_{t-4})\} = r_2 \sigma_e^2$$

$$\text{cov}(y_t, y_{t-3}) = \sigma_{13} = E\{(e_t + r_1 e_{t-1} + r_2 e_{t-2})(e_{t-3} + r_1 e_{t-4} + r_2 e_{t-5})\} = 0$$

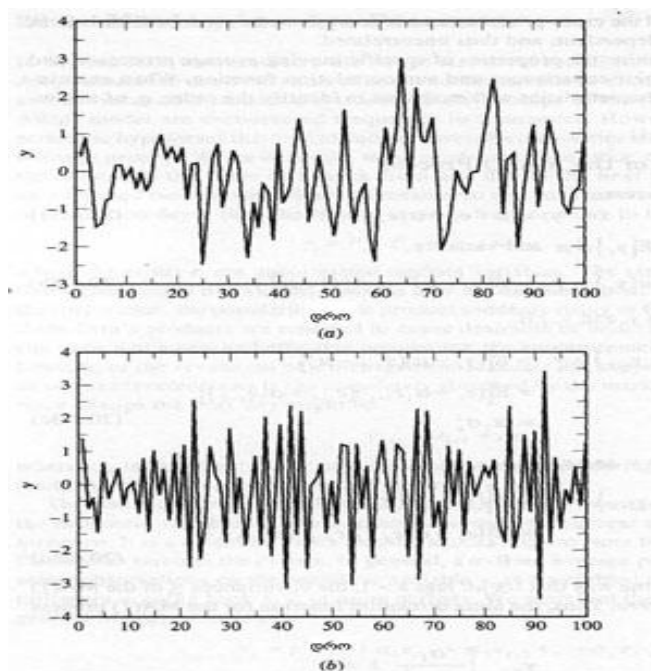
MA(2) პროცესის შესაბამისი ავტოკორელაციის ფუნქციაა:

$$\dots_1 = \frac{r_1(1+r_2)}{1+r_1^2+r_2^2}$$

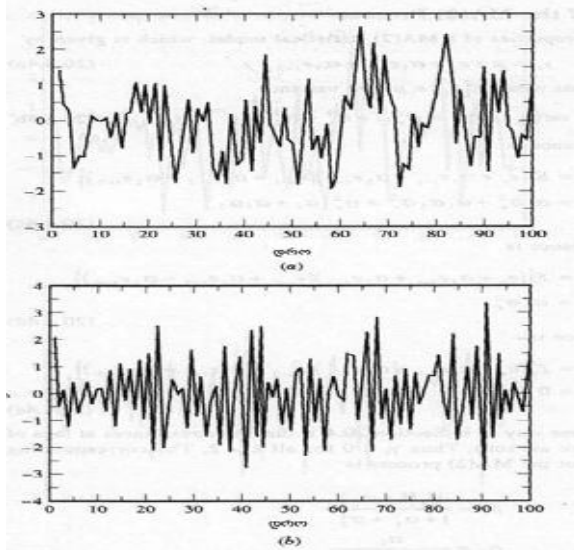
$$\dots_2 = \frac{r_1}{1+r_1^2+r_2^2}$$

$$\dots_k = 0, k > 2$$

მაგალითი:



ნახ.8.6. MA(1) დროითი მწკრივი. **ა/** $y_t = e_t + 0.8e_{t-1}$; **ბ/** $y_t = e_t - 0.8e_{t-1}$



ნახ.8.7. MA(2) დროითი მწკრივი. ა/ $y_t = e_t + 0.5e_{t-1} + 0.3e_{t-2}$;
 ბ/ $y_t = e_t - 0.5e_{t-1} + 0.3e_{t-2}$

ავტორეგრესიის ფუნქცია ნებისმიერი MA(q) პროცესისათვის ჩაიწერება შემდეგი სახით:

$$r_k = \frac{\sum_{i=0}^{q-k} r_i r_{i+k}}{\sum_{i=0}^q r_i^2}, k = 0, 1, 2, \dots, q$$

$$r_k = 0, k > q$$

8.8. ბოქს-ჯენკინსის მეთოდი

დროითი მწკრივების ანალიზისას ეკონომეტრისტიკისათვის მნიშვნელოვანია მოდელის სპეციფიკაცია, რაც აუცილებელია მოდელში შესაყვანი ცვლადების, განტოლებების და ადეკვატური ლაგური სტრუქტურების თვალსაზრისით. ამ მხრივ საინტერესოა ბოქს-ჯენკინსის მეთოდი (Box Jenkins, 1972), რომლის მიხედვითაც ერთცვლადიან სტანდარტულ მოდელში საძიებელი თანაფარდობა დაკავშირებულია დამოკიდებული ცვლადის მიმდინარე და ლაგურ მნიშვნელობასთან. თავდაპირველად, დროითი მწკრივისათვის გამოითვლება პირველი სხვაობა ან უფრო მაღალი რიგის სხვაობა, რათა დროითი მწკრივი “გადავაკეთოთ” სტაციონალურ მწკრივად რაც ფაქტობრივად ნიშნავს: დროითი მწკრივიდან გამოვეყოთ ტრენდი, ციკლისა და სეზონურობის კომპონენტები და სხვა, რომლებიც დაკვირვების შედეგების გადანაწილებას დროის ფაქტორთან აკავშირებენ. შემდეგ შეფასდება რეგრესიის განტოლება:

$$\bar{y}_t = S_0 + S_1 y_{t-1} + \dots + S_k y_{t-k} + X_0 V_t + X_1 V_{t-1} + \dots + X_q V_{t-q}$$

სადაც: $\bar{y}_t - y$ -ის მნიშვნელობების სხვაობებია (t-s) პერიოდებში;

V_t, V_{t-1} - შემთხვევითი სიდიდეებია ნულოვანი საშუალოთი და მუდმივი დისპერსიით.

რადგან \bar{y}_t ნაწილობრივ განსაზღვრულია როგორც ავტორეგრესიული (AR) პროცესი ე.ი. პროცესი, რომელიც დამოკიდებულია მის წინა მნიშვნელობაზე, ხოლო მეორეს მხრივ განსაზღვრულია როგორც შემთხვევითი წევრის მიმდინარე და წინა პერიოდის მცოცავი საშუალოს პროცესი (MA), ამიტომ პროცესი მთლიანობაში წარმოადგენს ARIMA(p,d,q) პროცესს, სადაც: p – AR რიგია, q – ნაწილის რიგი, d – სხვაობების რიგი, რომელიც მწკრივის სტაციონალურობისთვისაა გათვალისწინებული.

ჩატარებულია კოლოსალური სამუშაო დროითი მწკრივების ანალიზის მეთოდების გამარტივების მიზნით და იგი შეივსო მრავალგანზომილებიანი რეგრესიული ანალიზით.

8.9. დროითი მწკრივების მოდელები

დროითი მწკრივების მოდელების აგებისას თუ გამოყენებულია დამოუკიდებელი ცვლადების მიმდინარე და ლაგური მნიშვნელობები, მაშინ მოდელის კოეფიციენტების მოდერნიზაცია გაკვირვებას არ გამოიწვევს. ამიტომ, ლაგური დამოკიდებულების სტრუქტურის შესაფასებლად დამუშავებულია მეთოდები, რომლებიც საშუალებას იძლევიან შევხედოდეთ რეგრესიის განტოლებაში შემაჯავლი დამოუკიდებელი ცვლადების რიცხოვნება, რომ თავი ავარიდოთ მულტიკოლინეარობის პრობლემას, რომელსაც შეიძლება ადგილი ჰქონდეს, ან მაქსიმალურად შემცირდეს მისი ეფექტი.

კოიკის ლაგური მოდელი (Koyck, 1954).

მოდელის ასაგებად კოიკის მიერ შემოტანილია წინადადება რომ დამოუკიდებელ ლაგურ ცვლადებთან კოეფიციენტების მნიშვნელობები კლებადი გეომეტრიული პროგრესიის მოთხოვნას ექვემდებარებიან. ერთი დამოუკიდებელი ცვლადის შემთხვევაში მოდელი ჩაიწერება შემდეგი სახით:

$$Y_t = r + s_0(X_t + \beta X_{t-1} + \beta^2 X_{t-2} + \beta^3 X_{t-3} + \dots) + v_t$$

ამ განტოლებაში შესაფასებელია სამი პარამეტრი: r, s_0, β . ეს განტოლება არაწრფივია კოეფიციენტების მიმართ და მისი შეფასება უმცირეს კვადრატთა მეთოდით არ შეიძლება. ამიტომ აუცილებელია განტოლების ტრანსფორმაცია ისე, რომ მიღებული განტოლება წრფივი იყოს კოეფიციენტების მიმართ. ის შეიძლება წარმოვადგინოთ ასე:

$$\beta Y_{t-1} = \beta r + s_0(\beta X_{t-1} + \beta^2 X_{t-2} + \beta^3 X_{t-3} + \dots) + \beta v_{t-1}$$

ამ ორი განტოლებიდან შეიძლება მივიღოთ წარმოებული განტოლება:

$$Y_t = r_0 + s_0 X_t + \beta Y_{t-1} + u_t$$

სადაც: $u_t = v_t - \beta v_{t-1}$; $r_0 = r - \beta r$

ამგვარად, კოიკმა ტრანსფორმაციის გზით ლაგური განტოლება შეცვალა ლაგური ფორმით დამოკიდებულ ცვლადიანი განტოლებით, რომელიც ხშირად ავტორეგრესიული განტოლების სახელითაა ცნობილი.

რეზიუმე

ეკონომიკური ცვლადების შესწავლისას დინამიკაში, მათ შესახებ დაკვირვებების მონაცემების მიხედვით, ერთერთი მნიშვნელოვანი მიზანია ეკონომიკური ცვლადების მნიშვნელობების პროგნოზირება, რომლისთვისაც შეიძლება გამოვიყენოთ ეკონომეტრიკული მოდელის აგება. ამ შემთხვევაში აუცილებელია შემდეგი:

- ა/ეკონომიკური მოდელის ფორმულირება, რომელიც მოიცავს დამოკიდებულ ცვლადზე მოქმედ დამოუკიდებელ ეკონომიკური ცვლადების დადგენას;
- ბ/სტატისტიკური მოდელის აგება, რომლის მიხედვითაც ხდება მონაცემების გენერირება;
- გ/შერჩევის საფუძველზე პროცედურის განხილვა სტატისტიკური მოდელის უცნობი პარამეტრების შესაფასებლად, რომლებიც შემდეგში გამოიყენება პროგნოზირებისათვის.

დროითი მწკრივების მოდელის აგებისას ლაგური დამოკიდებულების სტრუქტურის შესაფასებლად დამუშავებულია მეთოდები, რომლებიც საშუალებას იძლევიან შევზღუდოთ რეგრესიის განტოლებაში შემავალი დამოუკიდებელი ცვლადების რიცხოვნება, რომ თავი ავარიდოთ მულტიკოლინეარობის პრობლემას, რომელსაც შეიძლება ადგილი ჰქონდეს, ან მაქსიმალურად შემცირდეს მისი ეფექტი.

ARIMA (autoregressive integrated moving average) – მოდელი იყენებს დამოკიდებული ცვლადის მიმდინარე და გასული პერიოდის მნიშვნელობებს და იძლევა ამ ცვლადის მოკლევადიან პროგნოზს.

ARIMA(p,d,q) აერთიანებს p რიგის ავტორეგრესიის პროცესს და q რიგის მცოცავი საშუალოს პროცესს.

დამოუკიდებელი სამუშაო

ცხრილში მოცემულია 100 დაკვირვების შედეგი სამ დროით მწკრივზე . თითოეული მათგანისათვის შეასრულეთ შემდეგი:

- ა/ დაიტანეთ სიბრტყეზე წერტილები დროის მიხედვით;
- ბ/ გამოიყენეთ პირველი T=90 დაკვირვება ბოქს-ჯენკინსის დროითი მწკრივის მოდელის განსაზღვრებისა, შეფასებისა და შემოწმებისათვის

N	y1	y2	y3	N	y1	y2	y3
1	3.6877	12.6893	37.9928	51	0.5894	10.4447	34.3592
2	3.1987	12.4573	37.7335	52	1.6616	11.7054	35.0588
3	2.6961	11.9171	37.0559	53	3.4887	13.5458	37.2654
4	1.1523	10.5487	35.3283	54	1.2313	10.7827	35.6734
5	1.4623	11.2104	35.0744	55	1.7742	11.3608	35.4077
6	1.3849	11.3701	35.0995	56	1.4506	11.3688	35.2749
7	1.5414	11.4815	35.2242	57	2.1189	11.9896	35.8253
8	2.3574	12.2858	36.0946	58	0.3378	10.1575	34.2849
9	2.2112	11.9740	36.2442	59	0.4433	10.4354	33.7464
10	2.2126	11.8472	36.1966	60	2.1517	12.4593	35.4849
11	2.4576	12.1193	36.4435	61	2.8130	12.8156	36.7626
12	1.8156	11.4361	35.8912	62	2.421	11.9147	36.5389
13	2.5679	12.2498	36.4121	63	3.5527	13.0802	37.5858
14	1.5885	11.2649	35.7033	64	5.0817	14.4951	39.5556
15	2.4723	12.1715	36.2342	65	3.4327	12.3644	38.4688
16	2.6411	12.3759	36.7201	66	2.7956	11.7166	37.2451
17	3.3363	12.8777	37.4793	67	4.9430	14.2789	39.1588
18	2.6057	11.9994	37.0034	68	2.6851	11.7816	37.6757

19	3.3782	12.7638	37.5156	69	4.1909	13.2627	8.3708
20	2.4452	11.8384	36.8617	70	3.4817	12.7244	38.2010
21	3.4366	12.8412	37.5179	71	2.7207	11.8002	37.1867
22	1.3704	10.7879	35.8084	72	0.6415	9.9807	34.8296
23	3.3068	12.8799	36.9984	73	1.2354	11.0642	34.6639
24	1.9206	11.5569	36.3061	74	1.0539	11.1724	34.6843
25	0.4035	9.9091	34.2913	75	2.3040	12.3418	35.8645
26	1.2626	11.2808	34.5952	76	1.4617	11.3247	35.4728
27	2.2011	12.3597	35.8351	77	2.0539	11.8237	35.7643
28	2.7095	12.5506	36.6832	78	2.1073	11.9357	36.0308
29	1.2696	10.8502	35.4323	79	2.3208	12.0295	35.2659
30	2.6394	12.3654	36.2845	80	3.0989	12.7602	37.1229
31	2.4200	12.1866	36.5574	81	3.7457	13.2389	38.0541
32	2.7197	12.2667	36.7821	82	4.7940	14.0336	39.3423
33	1.1103	10.6483	35.2828	83	3.9746	12.9182	38.9096
34	0.5919	10.3426	34.1825	84	1.8554	10.7397	36.5018
35	1.5634	11.7017	34.9584	85	3.5450	12.9365	37.4225
36	0.0653	10.4384	34.103	86	2.3365	11.8423	36.8162
37	2.2665	12.3579	35.5743	87	0.8858	10.2773	34.9299
38	1.0188	11.0168	35.0112	88	2.0603	11.9201	35.5730
39	0.8937	10.7445	34.4404	89	1.8923	11.8274	35.8211
40	2.9799	13.0850	36.4784	90	3.6376	13.3810	37.5062
41	0.7530	10.5330	35.0066	91	0.8915	10.3749	35.3932
42	2.5037	12.2653	35.9588	92	3.4940	13.1076	37.0059
43	1.8408	11.7273	35.9249	93	2.8430	12.5382	37.2913
44	3.9846	13.6537	37.8339	94	3.4943	12.8111	37.7155
45	2.2941	11.7292	36.9212	95	2.4369	11.7663	36.8955
46	1.7179	11.0343	35.7425	96	2.0078	11.3921	36.0867
47	1.8324	11.5612	35.6448	97	1.3989	11.0526	35.3186
48	2.7033	12.5208	36.5525	98	1.9590	11.7947	35.6535
49	1.7075	11.3582	35.8710	99	1.7216	11.5777	35.6140
50	1.679	11.2720	35.4236	100	3.2881	13.0790	37.0945

თავი 9. ერთდროულ განტოლებათა სისტემები

ეკონომიკური პროცესებისა და მოვლენების სტატისტიკური მოდელირებისას ხშირად აუცილებელი ხდება რეგრესიის განტოლებათა სისტემის აგება, რომელშიც ერთი და იგივე ცვლადი რეგრესიის სხვადასხვა განტოლებებში ერთდროულად მონაწილეობენ. ასეთი განტოლებათა სისტემა ცნობილია ერთდროულ განტოლებათა სისტემის სახელით. სისტემის განტოლებებში შეიძლება მონაწილეობდნენ არა მარტო ის ცვლადები, რომლებიც მიმდინარე პერიოდს განეკუთვნებიან, არამედ ის ცვლადებიც, რომლებიც გასულ პერიოდს მიეკუთვნებიან. ეს ცვლადები ლაგური ცვლადებია.

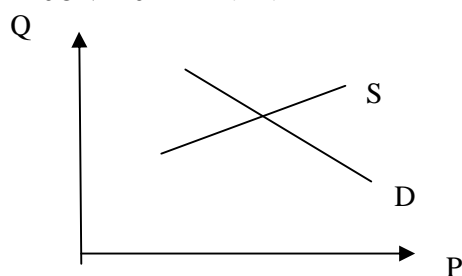
ერთდროულ განტოლებათა სისტემის საუკეთესო მაგალითს წარმოადგენს მოთხოვნა-მიწოდების მოდელი. ცნობილია, რომ მოთხოვნა - D რაიმე პროდუქტზე დამოკიდებულია ამ პროდუქტის P ფასზე. ამავე ცვლადზე, თუმცა საპირისპირო ნიშნით, დამოკიდებულია ამ პროდუქტის მიწოდება. საბაზრო მექანიზმის გავლენით ფორმირდება ისეთი ფასი, რომ მოთხოვნა და

მიწოდება წონასწორდება. ამ სიტუაციის ფორმალიზება მარტივი სახით შეიძლება წარმოვადგინოთ შემდეგი წრფივი მოდელებით:

მოდელი 1:

$$\begin{aligned} D_t &= r_1 p_t + v_t \\ S_t &= r_2 p_t + u_t \\ D_t &= S_t = Q_t \end{aligned} \quad (9.1.)$$

გრაფიკულად ეს სისტემა წარმოადგენს ორი მრუდის – მოთხოვნისა და მიწოდების გადაკვეთას. ამ მოდელში მოთხოვნის ერთი წრფეა და ასევე, მიწოდების ერთი წრფე და დაკვირვების სხვადასხვა მნიშვნელობებისათვის სხვაობები განპირობებულია შეცდომებით v_t, \hat{u}_t (ნახ. 9.1)



ნახ. 9.1. მოთხოვნა-მიწოდების წრფივი მოდელი

ბუნებრივია, არაფრის თქმა შეიძლება ჭეშმარიტი D და S წრფეების შესახებ, რადგან წრფის თითოეული წერტილი შეიძლება რეალიზებული იყოს როგორც ორი წრფის გადაკვეთა, რომლებსაც შეიძლება ჰქონდეთ ნებისმიერი დახრა. ეს ადვილი დასანახია, თუ (9.1) სისტემას დაყვანილი სახით ჩავწერთ:

$$\begin{aligned} q_t &= \frac{r_1 \hat{u}_t - r_2 v_t}{r_1 - r_2} \\ p_t &= \frac{\hat{u}_t - v_t}{r_1 - r_2} \end{aligned} \quad (9.2)$$

მოდელი 2: მოთხოვნა ფასის პროპორციულია პროპორციულობის კოეფიციენტით $r_1 < 0$,

$$D_t = r_1 p_t + b_1 + v_t$$

მიწოდება ფასის პროპორციულია პროპორციულობის კოეფიციენტით $r_2 > 0$,

$$S_t = r_2 p_t + b_2 + \hat{u}_t$$

და

$$S_t \equiv D_t$$

სადაც, v_t, \hat{u}_t - მოდელის შეცდომებია ნულოვანი მათემატიკური მოლოდინით.

პირველი ორი განტოლება, თუ მათ ცალ-ცალკე განვიხილავთ, საესებით ჩვეულებრივი განტოლებებია და თითოეული მათგანისათვის

შეიძლება განვსაზღვროთ რეგრესიის კოეფიციენტები. თუმცა ამ შემთხვევაში ღიად დარჩება საკითხი მოთხოვნა-მიწოდების იკვივობის შესახებ, ე.ი. შეიძლება მესამე პირობა არ შესრულდეს, რომელშიც მოთხოვნა დამოკიდებული ცვლადის სახითაა წარმოდგენილი. სწორედ ამ თვალსაზრისით, განტოლებების ცალ-ცალკე განხილვა აზრს კარგავს.

ეკონომიკური მოდელი, როგორც ერთდროულ განტოლებათა სისტემა შეიძლება წარმოვადგინოთ სტრუქტურული ან დაყვანილი ფორმით. სტრუქტურული ფორმით წარმოდგენისას მის განტოლებებს ისეთი სახე აქვთ, რომელიც ასახავს ცვლადებს შორის უშუალო კავშირს. დაყვანილი ფორმა მიიღება ენდოგენური ცვლადების მიმართ მოდელის ამოხსნის შემდეგ, ე.ი. მისი გამოსახვით ეგზოგენური ცვლადებისა და მოდელის პარამეტრებით.

მაგალითად, მოთხოვნა-მიწოდების მოდელში D_t, S_t, P_t ენდოგენური ცვლადებია, ხოლო r_1, r_2, b_1, b_2 - პარამეტრებია. ეგზოგენური ცვლადები ამ მოდელში არ გვაქვს. დაყვანილ ფორმაში P_t, S_t, D_t ცვლადები აუცილებლად მოდელის პარამეტრებით გამოითვლებიან და თუ მესამე პირობას გამოვიყენებთ, მივიღებთ:

$$r_1 p_t + b_1 + v_t = r_2 p_t + b_2 + v_t \Rightarrow p_t = \frac{b_2 - b_1}{r_1 - r_2} + v_t$$

$$D_t = r_1 \frac{b_2 - b_1}{r_1 - r_2} + b_1 + v_{2t}; S_t = r_2 \frac{b_2 - b_1}{r_1 - r_2} + b_2 + v_{3t}$$

სადაც, v_t, v_{2t}, v_{3t} - გარდაქმნილი გადახრებია. გამოსახულება:

$$\frac{b_2 - b_1}{r_1 - r_2}$$

ჩვენ შეგვიძლია შევაფასოთ

$$r_1 \bar{p} + b_1 = D_t$$

$$r_2 \bar{p} + b_2 = S_t$$

მაგრამ ამ თანაფარდობებიდან შეუძლებელია საწყისი მოდელის პარამეტრების r_1, r_2, b_1, b_2 გამოთვლა, თუმცა მათი შეფასება უმცირეს კვადრატთა მეთოდით შესაძლებელია. მაგრამ ამ მეთოდით საწყისი მოდელის იდენტიფიცირება ყოველთვის არ შეიძლება. იმისათვის, რომ ეს პროცედურა ჩავატაროთ, აუცილებელია შემოვიტანოთ დამატებითი პირობები. კერძოდ, ზოგიერთი კოეფიციენტის ნულთან ტოლობა ან ერთმანეთთან მათი დამოკიდებულების იდეა. სწორედ, მოდელის აგების ეტაპზე უნდა ვეცადოთ მოდელის ისეთი ფორმა შევარჩიოთ, რომლის იდენტიფიცირება შეიძლება. ასეთია, მაგალითად:

$$y_1 = f(x); y_2 = f(y_1, x), \dots, y_k = f(y_1, y_2, \dots, y_{k-1}, x)$$

სადაც: x - დამოუკიდებელი ცვლადების ვექტორია;

y_i - დამოუკიდებელი ცვლადების ვექტორია.

მოდელი 3.

მოთხოვნა-მიწოდების მოდელი ლაგური ცვლადებით:

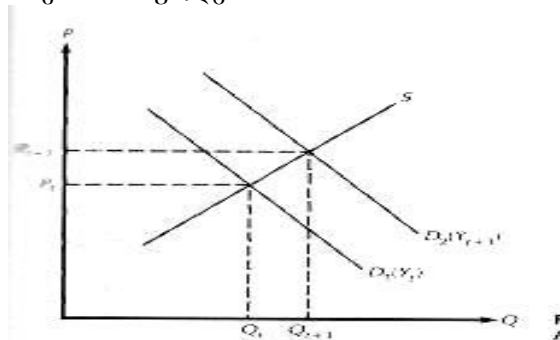
$$\begin{aligned} S_t &= r_1 + r_2 P_t + r_3 P_{t-1} + v_t \\ D_t &= s_1 + s_2 P_t + s_3 Y_t + \hat{\epsilon}_t \\ S_t &\equiv D_t \equiv Q_t \end{aligned} \quad (9.2)$$

სადაც:

- S_t - მიწოდების სიდიდე განსაზღვრული t - პერიოდისათვის;
- D_t - მოთხოვნის სიდიდე განსაზღვრული t - პერიოდისათვის;
- P_t - საქონელზე ფასია განსაზღვრული t - პერიოდისათვის;
- P_{t-1} - საქონელზე ფასია განსაზღვრული წინა $t-1$ პერიოდისათვის;
- Y_t - მომხმარებლის შემოსავალია განსაზღვრული t - პერიოდისათვის.

როგორც აღვნიშნეთ, მიწოდების განტოლება, მოთხოვნის განტოლება და იგივეობის პირობა განსაზღვრავენ საბაზრო ფასს და მოთხოვნა-მიწოდების რაოდენობრივ მახვენებლებს, როდესაც ბაზარი გაწონასწორებულია, ამიტომ D_t, S_t, P_t ცვლადებს ენდოგენურ ცვლადებს უწოდებენ, რადგან ისინი დეტერმინირებულნი არიან განტოლებათა სისტემის შიგნით. (9.2) სისტემის

წარმოდგენა გრაფიკული სახით ნათლად იძლევა P_t და Q_t ცვლადების ენდოგენურობის დანახვის საშუალებას.



ნახ. 9.2. P_t და Q_t ცვლადების ენდოგენურობა

განხილული მოდელი შეიცავს კიდევ ორ ცვლადს, რომელთა სიდიდეები დეტერმინირებულნი არ არიან სისტემის შიგნით. მათ პრედეტერმინირებული ცვლადები ეწოდებათ. ეს ცვლადებია: P_{t-1} ; Y_t . მათ შორის მნიშვნელოვანი განსხვავებაა, კერძოდ, P_{t-1} ცვლადი დეტერმინირებულია სისტემის შიგნით და ის ლაგური ენდოგენური ცვლადია, ხოლო Y_t ცვლადი დეტერმინირებულია მოდელის გარეთ და იგი ენდოგენური ცვლადია. მოთხოვნის მრუდი D_t და მიწოდების მრუდი S_t განსაზღვრული t - პერიოდისათვის და P_{t-1} და Y_t დეტერმინირებული ცვლადებისათვის, შემდგომ $t-1$ პერიოდში ზევით

გადაადგილდება, ე.ი. ადგილი აქვს Y_{t+1} შემოსავლის გაზრდას და სწორედ ამ დროს შეინიშნება P_{t+1} მაღალი ფასი და Q_{t+1} -ის გაზრდილი მნიშვნელობა.

ვინაიდან P_t და Q_t ენდოგენური ცვლადებია, ამიტომ უმცირეს კვადრატთა მეთოდის გამოყენება მოთხოვნის ან მიწოდების შესაფასებლად მცდარ შეფასებას მოგვცემს. თუ საწყისი მოდელიდან გამოვრიცხავთ ლაგურ P_{t-1} ცვლადს, მივიღებთ:

$$\begin{aligned} Q_t &= r_1 + r_2 P_t + v_t \\ Q_t &= s_1 + s_2 P_t + s_3 Y_t + \hat{\epsilon}_t \\ q_t &= Q_t - \bar{Q} \end{aligned}$$

შემოვიტანოთ აღნიშვნები:

$$\begin{aligned} y_t &= Y_t - \bar{Y} \\ p_t &= P_t - \bar{P} \end{aligned}$$

მივიღებთ:

$$\begin{aligned} q_t &= r_2 p_t + v_t \\ q_t &= s_2 p_t + s_3 y_t + \hat{\epsilon}_t \end{aligned}$$

მიღებული მოდელი სტრუქტურული მოდელია. იგი შეიცავს ენდოგენურ ცვლადებს განტოლების მარცხენა მხარეში და პრედექტერმინირებულ ცვლადებს მარჯვენა მხარეში. თუ სისტემას ამოვხსნით თითოეული ენდოგენური ცვლადის მიმართ და განვიხილავთ მათ როგორც პრედექტერმინირებული ცვლადების ფუნქციას, მივიღებთ:

$$\begin{aligned} q_t &= \frac{r_2 s_3 y_t}{r_2 - s_2} + \frac{r_2 \hat{\epsilon}_t - s_2 v_t}{r_2 - s_2} \\ p_t &= \frac{s_3 y_t}{v_2 - s_2} + \frac{\hat{\epsilon}_t - v_t}{r_2 - s_2} \end{aligned}$$

გავითვალისწინოთ, რომ y_t არ არის კორელირებული v_t და $\hat{\epsilon}_t$ -სთან და გამოვიყენოთ უმცირეს კვადრატთა მეთოდი პარამეტრების შესაფასებლად, მივიღებთ:

$$\hat{r}_2 = \frac{\sum_t q_t p_t}{\sum_t p_t^2} = r_2 + \frac{\sum_t v_t p_t}{\sum_t p_t^2}$$

რადგან მიღებული ჯამის მეორე შესაკრების მრიცხველში სიდიდეები კორელირებულები არიან, ხოლო მრიცხველი და მნიშვნელი დამოკიდებული სიდიდეებია, ბუნებრივია, ამ შესაკრების მათემატიკური ლოდინი ნულისგან განსხვავებულია, ე.ი.

$$E\left(\frac{\sum_t p_t v_t}{\sum_t p_t^2}\right) \neq 0$$

და \hat{r}_2 პარამეტრის შეფასება გადაადგილებადია.

ერთდროულ განტოლებათა სისტემის შესაფასებლად არსებობს მთელი რიგი მეთოდებისა, რომელთა დაყოფა შეიძლება ორ ჯგუფად. პირველ ჯგუფს მიეკუთვნებიან ის მეთოდები, რომლებიც გამოიყენება თითოეული განტოლებისათვის, ხოლო მეორე ჯგუფი მოიცავს იმ მეთოდებს, რომლებიც გათვალისწინებულია განტოლებების შეფასებისათვის. მაგალითად, TSP ეკონომეტრიკის პაკეტი მოიცავს თითოეული ჯგუფიდან ერთ მეთოდს. სისტემის თითოეული განტოლების შესაფასებლად ამ პაკეტში ორბიჯიანი უმცირეს კვადრატთა მეთოდია (Two – Stage Least Squares), ხოლო მეთოდების მეორე ჯგუფიდან ამ პაკეტში სამბიჯიანი უმცირეს კვადრატთა მეთოდია (Three Stage Least Squares). ორბიჯიანი უმცირეს კვადრატთა მეთოდი გამოიყენება იმ შემთხვევაში, როდესაც შესაფასებელ მოდელში ლაგური ცვლადებია

მაგრამ ლაგური მნიშვნელობები წინასწარ არ არიან განსაზღვრულები, ამიტომ პირველ ეტაპზე განისაზღვრება მათი ის მნიშვნელობები, რომლებიც არ გვაქვს, რაც უმცირეს კვადრატთა მეთოდის გამოყენებით შესაძლებელია. ამ შემთხვევაში აიგება რეგრესიის მოდელი საწყისი ინფორმაციის გამოყენებით. და მას შემდეგ, როდესაც საწყისი ინფორმაცია შეივსება დამატებითი ინფორმაციით და სახეზე გვექნება სრული მონაცემები, შეგვიძლია საძიებელი პარამეტრების შეფასებას მივხედოთ.

განტოლებათა სისტემის შესაფასებლად გამოიყენება სამბიჯიანი უმცირეს კვადრატთა მეთოდიც. მისი პირველი ორი ეტაპი ორბიჯიანი უმცირეს კვადრატთა მეთოდის ანალოგიურია, ე.ი. შეფასდება ლაგური ცვლადების შემცველი განტოლებების პარამეტრები და ამავე ეტაპზე ხდება რეგრესიის განტოლების კოეფიციენტების გამოთვლა, ხოლო შემდეგ, სისტემის ყველა განტოლება ერთმანეთთან დაკავშირდება. კავშირის ზომად მიიღება მოდელის შეცდომების კოვარიაციის მატრიცა. შემდეგ ეტაპზე, ამ კავშირის გასაძლიერებლად, გამოითვლება $V_t; \hat{t}$ შეცდომების კოვარიაციის მატრიცა. ამ კავშირის გასაძლიერებლად, შემდეგ ეტაპზე რეგრესიის კოეფიციენტების გამოთვლისას გასათვალისწინებელი იქნება შეცდომების კოვარიაციის მატრიცა. მსგავსი თანმიდევრობით მიიღწევა განტოლებათა ურთიერთ დაკავშირება მთლიანობაში.

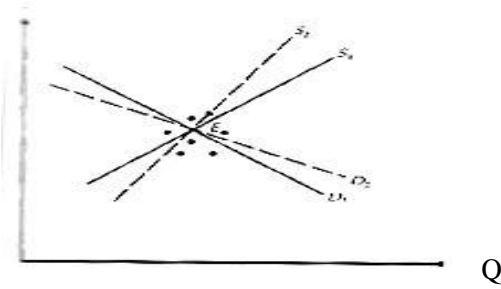
9.1 იდენტიფიკაციის პრობლემა

იდენტიფიკაციის პრობლემა დაკავშირებულია მოდელის პარამეტრების შეფასებასთან. სტრუქტურული ფორმით წარმოდგენილი მოდელის იდენტიფიკაციის საკითხი გადაწერილი იქნება, თუ სტრუქტურული მოდელის პარამეტრები დაყვანილი ფორმის განტოლებით შეიძლება შეფასდეს. იდენტიფიკაციის პრობლემა მოხსნილად ჩაითვლება, თუ პარამეტრების მნიშვნელობები ცალსახადაა განსაზღვრული.

იდენტიფიკაციის პრობლემა შეიძლება განვიხილოთ მოთხოვნა-მიწოდების მაგალითზე. თავდაპირველად განვიხილოთ დროითი მწკრივის მოთხოვნა-მიწოდების მოდელი, რომელიც არ შეიცავს პრედექტერმინირებულ ცვლადებს:

$$\text{მიწოდება: } Q_t = r_1 + r_2 P_t + v_t$$

$$\text{მოთხოვნა: } Q_t = S_1 + S_2 P_t + u_t$$



ნახ. 9.3. მოთხოვნა-მიწოდების მოდელი

მივიღოთ, რომ დროის ნებისმიერ პერიოდში ბაზარი წონასწორობის პირობებშია ისე, რომ მოთხოვნის სიდიდივე მიწოდების სიდიდის ტოლია. ამ შემთხვევაში იდენტიფიკაციის პრობლემის ახსნა იგივეობის პირობაშია. დროის თითოეული პერიოდისათვის მიიღწევა განსაზღვრული P ფასი და გაყიდული საქონლის რაოდენობა – Q . რდესაც ვცდილობთ ცალ-ცალკე შევაფასოთ მოთხოვნა-მიწოდების განტოლებები და ვიყენებთ რა ბაზრის მონაცემებს, აუცილებლად საქმე გვექნება აზრს მოკლებულ შედეგთან, რადგან არ შესრულდა აუცილებელი პირობა, კერძოდ: მოთხოვნა-მიწოდების იგივეობის მოთხოვნის პირობა.

მოდელი, რომელსაც ვიხილავთ ერთერთი იმ მოდელთაგანია, რომელშიც მოთხოვნა-მიწოდების მრუდების მდებარეობა განსაზღვრულია. განსაზღვრულია ისიც, რომ სტრუქტურული მოდელის პარამეტრების მნიშვნელობების დადგენა განტოლების დაყვანილი ფორმიდან არ შეიძლება (განტოლების დაყვანილი ფორმა მარტივი განტოლებაა, რომელიც აღწერს მრუდების გადაკვეთის წერტილს მოდელის შეცდომების გათვალისწინებით). განტოლებების დაყვანილი ფორმაა:

$$p_t = \frac{u_t - v_t}{r_2 - s_2}; q_t = \frac{v_2 u_t - s_2 v_t}{r_2 - s_2}$$

როგორც ნახ.9.3-დან ჩანს მოთხოვნა-მიწოდების ნებისმიერი წყვილი E წერტილში იკვეთება ანუ ჩვენ გვაქვს მოთხოვნა – მიწოდების წრფეების უსასრული რაოდენობა, რომლებსაც ექნებათ ერთი და იგივე დაყვანილი ფორმა. ე.ი. მოთხოვნა-მიწოდების ჯერ კიდევ არ არიან იდენტიფიცირებულნი.

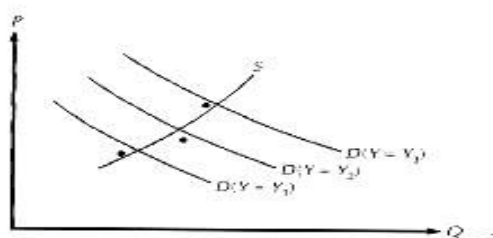
განვიხილოთ მოთხოვნა-მიწოდების კიდევ ერთი მოდელი:

$$\text{მიწოდება: } Q_t = r_1 + r_2 P_t + v_t$$

$$\text{მოთხოვნა: } Q_t = s_1 + s_2 P_t + s_3 Y_t + u_t$$

$$s_3 \neq 0$$

ეს დამოკიდებულება წარმოდგენილია ნახაზზე 9.4.



ნახ. 9.4. მიწოდების მრუდი იდენტიფიცირებულია

მოსალოდნელია, რომ მოთხოვნა-მიწოდების ნებისმიერი წყვილის შესაბამისი მრუდების გადაკვეთა შეესაბამება ჰეშმარიტ მრუდებს ან სხვაგვარად რომ ვთქვათ, საქმე გვაქვს სტრუქტურული ტიპის მოდელებთან, რომლებიც შეიცავენ ერთი და იგივე დაყვანილ ფორმას. ცხადია, მოდელის სისტემის განტოლებების იდენტიფიკაცია ითხოვს დამატებით ინფორმაციას.

ჩვენ არ შეგვიძლია წარმოვადგინოთ მოთხოვნის ერთი მრუდი და მიწოდების ერთი მრუდი დროის მთელი პერიოდის განმავლობაში, რადგან მომხმარებლის შემოსავალი განსაზღვრავს მოთხოვნას, შემოსავალი კი იცვლება დროის მთელი პერიოდის განმავლობაში. ჩვენ შეგვიძლია ჩავთვალოთ ან მივიღოთ ფაქტად ის გარემოება, რომ მოთხოვნის მრუდი იცვლება სისტემატიურად, მთელი t –პერიოდის განმავლობაში, როგორც ნახაზიდან ჩანს: მიწოდების მრუდი იდენტიფიცირებულია, რადგან მოდელის პარამეტრები მიიღება მოდელის დაყვანილი ფორმიდან. განტოლების იდენტიფიკაციის მიზეზი ისიც არის, რომ ეგზოგენური y ცვლადი გამორიცხულია მიწოდების განტოლებიდან. მოთხოვნის განტოლება იდენტიფიცირებული არ არის, რადგან წინა (გასული პერიოდის) მონაცემებით შეუძლებელია მოთხოვნის ცალსახად განსაზღვრა. დასაკვირველია, შესაძლებელია ორივე განტოლების (მოთხოვნა, მიწოდების) იდენტიფიკაცია. ქვემოთ მოყვანილი მოდელი სწორედ ამ ბუნების მატარებელია:

$$\text{მიწოდება: } Q_t = r_1 + r_2 P_t + r_3 T_t + v_t$$

$$\text{მოთხოვნა: } Q_t = s_1 + s_2 P_t + s_3 Y_t + \hat{v}_t$$

სადაც, T რეგიონის ტემპერატურის მაჩვენებელია.

თუ Y და T -ს მნიშვნელობები იცვლებიან დროის მთელი პერიოდის განმავლობაში და ერთმანეთთან კორელირებულები არ არიან, მაშინ მოთხოვნაც და მიწოდებაც შეიძლება დაფიქსირდნენ განსაზღვრულ დონეზე. თუ განვიხილავთ იმ შემთხვევას, რომ მოთხოვნა არა მხოლოდ შემოსავლის ფუნქციაა, არამედ სხვა ცვლადიც მოქმედებს მასზე (მაგ. ჯანმრთელობა) და ორივე ცვლადი მაღალი კორელაციის მატარებელია, მაშინ მოთხოვნის მრუდი დაფიქსირებული იქნება, როგორც ამ ორი ცვლადის ცვლილების შედეგი და მაშინ მიწოდების განტოლება ზეიდენტიფიცირებულია, რადგან მოთხოვნა-მიწოდების მოდელი შეიცავს ორ ეგზოგენურ ცვლადს, რომლებიც გამორიცხულია მიწოდების განტოლებიდან.

განხილული მოსაზრებების მიხედვით, შეიძლება ჩამოყალიბდეს იდენტიფიკაციის პირობები:

ა/ განტოლება იდენტიფიცირებულია, თუ განტოლებიდან გამორიცხული პრედეტერმინირებული ცვლადების რიცხოვნება აღემატება ან ტოლია ერთით ნაკლებ ენდოგენურ ცვლადების რიცხოვნების მაჩვენებლისა;

ბ/ განტოლება არ არის იდენტიფიცირებული, თუ ერთდროულ განტოლებათა დაყვანილი ფორმიდან შეუძლებელია განტოლებათა სტრუქტურული ფორმის ყველა პარამეტრის შეფასება;

გ/ განტოლება ზუსტად იდენტიფიცირებულია, თუ პარამეტრის მნიშვნელობა არსებობს;

დ/ განტოლება ზეიდენტიფიცირებულია, თუ ზოგიერთი პარამეტრისათვის ერთზე მეტი მნიშვნელობის მიღებაა შესაძლებელი.

რეზიუმე

ერთდროულ განტოლებათა სისტემაში შეიძლება მონაწილეობდნენ არა მარტო ის ცვლადები, რომლებიც მიმდინარე პერიოდს განეკუთვნებიან, არამედ ის ცვლადებიც, რომლებიც გასულ პერიოდს მიეკუთვნებიან. ეს ცვლადები ლაგური ცვლადებია. ერთდროულ განტოლებათა სისტემის საუკეთესო მაგალითია მოთხოვნა-მიწოდების მოდელი. ცნობილია, რომ მოთხოვნა - D რაიმე პროდუქტზე დამოკიდებულია ამ პროდუქტის P ფასზე. ამავე ცვლადზე, თუმცა საპირისპირო ნიშნით, დამოკიდებულია ამ პროდუქტის მიწოდება. საბაზრო მექანიზმის გავლენით ფორმირდება ისეთი ფასი, რომ მოთხოვნა და მიწოდება წონასწორდებიან.

ეკონომიკური მოდელი, როგორც ერთდროულ განტოლებათა სისტემა შეიძლება წარმოვადგინოთ სტრუქტურული ან დაყვანილი ფორმით. სტრუქტურული ფორმით წარმოდგენისას მის განტოლებებს ისეთი სახე აქვთ, რომელიც ასახავს ცვლადებს შორის უშუალო კავშირს. დაყვანილი ფორმა მიიღება ენდოგენური ცვლადების მიმართ მოდელის ამოხსნის შემდეგ, ე.ი. მისი გამოსახვით ეგზოგენური ცვლადებისა და მოდელის პარამეტრებით.

ერთდროულ განტოლებათა სისტემის შესაფასებლად არსებობს მთელი რიგი მეთოდებისა, რომელთა დაყოფა შეიძლება ორ ჯგუფად. პირველ ჯგუფს მიეკუთვნებიან ის მეთოდები, რომლებიც გამოიყენება თითოეული განტოლებისათვის, ხოლო მეორე ჯგუფი მოიცავს იმ მეთოდებს, რომლებიც გათვალისწინებულია განტოლებების ერთდროული შეფასებისათვის. მაგალითად, TSP ეკონომეტრიკის პაკეტი მოიცავს თითოეული ჯგუფიდან ერთ მეთოდს. სისტემის თითოეული განტოლების შესაფასებლად ამ პაკეტში ორბიჯიანი უმცირეს კვადრატთა მეთოდია (Two – Stage Least Squares), ხოლო მეთოდების მეორე ჯგუფიდან ამ პაკეტში სამბიჯიანი უმცირეს კვადრატთა მეთოდია (Three Stage Least Squares). ორბიჯიანი უმცირეს კვადრატთა მეთოდი გამოიყენება იმ შემთხვევაში გამოიყენება, როდესაც შესაფასებელ მოდელში ლაგური ცვლადებია.

ერთდროულ განტოლებათა სისტემის ანალიზისას აუცილებელია იდენტიფიკაციის პრობლემასთან დაკავშირებული სიტუაციების გათვალისწინება.

დამოუკიდებელი სამუშაო

1. განვიხილოთ მოდელი:

$$C_t = r_1 + r_2 Y_t + v_t$$

$$I_t = s_1 + s_2 Y_t + s_2 G_{t-1} + u_t$$

$$Y_t = C_t + I_t + G_t$$

ა/ ააგეთ ამ სისტემის დაყვანილი ფორმა და დაყვანილი ფორმიდან განსაზღვრეთ C პირველი ორი პერიოდისათვის;

ბ/ არის თუ არა მოხმარების ფუნქცია იდენტიფიცირებული; არის თუ არა ეს მოდელი ზეიდენტიფიცირებული?

გ/ ინვესტიციის განტოლება იდენტიფიცირებულია? ინვესტიციის განტოლება ზეიდენტიფიცირებულია?

დ/ რა მოხდება მოხმარებას მაჩვენებელს, თუ იგი შეფასდება უმცირეს კვადრატთა მეთოდით?

2. განვიხილოთ მოთხოვნა-მიწოდების მოდელი:

$$\begin{aligned} Q_t^S &= r_1 + r_2 P_t + v_t \\ Q_t^D &= s_1 + s_2 P_t + s_3 Y_t + s_4 P_{t-1} + u_t \\ Q_t^S &= Q_t^D \end{aligned}$$

სადაც, $E(v_i v_j) = 0, i \neq j, E(u_i u_j) = 0, i \neq j$

ა/ არის მიწოდების განტოლება იდენტიფიცირებადი? რა მოხდება თუ მიწოდების განტოლება შეფასდება უმცირეს კვადრატთა მეთოდით?

ბ/ არის მოთხოვნის განტოლება იდენტიფიცირებადი? რა მოხდება თუ მიწოდების განტოლება შეფასდება უმცირეს კვადრატთა მეთოდით?

გ/ თუ ჩათვლით, რომ მიწოდების განტოლების პარამეტრების შესაფასებლად უნდა გამოიყენოთ ორ ბიჯიანი უმცირეს კვადრატთა მეთოდი (2 SLS), როგორ მოიქცევით, როგორ გააკეთებთ ამ პროცედურას?

დ/ გამოიყენებთ თუ არა უმცირეს კვადრატთა მეთოდს მოთხოვნის განტოლების შეფასებლად?

ე/ შეიძლება თუ არა უმცირეს კვადრატთა მეთოდის გამოყენება მოთხოვნის განტოლებისათვის?

ვ/ იქნება თუა არა თქვენი შედეგები განსხვავებული, თუ იცით, რომ V ავტოკორელირებული იყო?

3. განვიხილოთ მოდელი სისტემით, რომელიც ორ განტოლებას მოიცავს:

$$Y_1 = r_1 Y_2 + r_2 Z_1 + u_1$$

$$Y_2 = b_1 + b_2 Y_1 + b_3 Z_1 + b_4 Z_2 + u_2$$

ა/ გამოიყენეთ უმცირეს კვადრატთა მეთოდი პირველი განტოლების შესაფასებლად;

ბ/ გამოიყენეთ ორ ბიჯიანი უმცირეს კვადრატთა მეთოდი (2 SLS) პირველი განტოლების შესაფასებლად.

4. განვიხილოთ მოდელი სისტემით, რომელიც სამ განტოლებას მოიცავს:

$$Y_1 = r_1 + r_2 Y_2 + r_4 X_1 + r_5 X_2 + u_1$$

$$Y_2 = b_1 + b_3 Y_3 + b_5 X_2 + u_2$$

$$Y_3 = x_1 + x_2 Y_2 + u_3$$

ამ განტოლებებიდან რომელია იდენტიფიცირებადი? ზუსტად იდენტიფიცირებადი? ზეიდენტიფიცირებადი?

5. განვიხილოთ რეკურსიული მოდელი ორი განტოლებით:

$$Y_1 = r_1 + r_3 X + u_1$$

$$Y_2 = b_1 + b_2 Y_1 + u_2$$

ა/ ახსენით, რატომ არის მიზანშეწონილი უმცირეს კვადრატთა მეთოდის გამოყენება (ვივარაუდოთ, რომ u_1, u_2 არაკორელირებულია).

6. განვიხილოთ მოთხოვნა-მიწოდების მოდელი:

$$\text{მოთხოვნა: } P_t = r_1 + r_2 Q_t + u_{1t}$$

$$\text{მიწოდება: } Q_t = b_1 + b_2 P_{t-1} + b_3 W + u_{2t}$$

ა/ როგორია თქვენი მოსაზრება თითოეული განტოლების იდენტიფიკაციის შესახებ?

ბ/ რა პირობაა აუცილებელი მოდელის რეკურსიულობისათვის?

გ/ თუ მოდელი რეკურსიულია, როგორ შეფასდება მოთხოვნის განტოლება?

დ/ თუ მოდელი რეკურსიული არ არის, როგორ შეფასდება მოთხოვნის განტოლება?

თავი 10. პროგნოზირება

პროგნოზირება სისტემის მომავალი მდგომარეობის ანალიზია გასული და მიმდინარე პერიოდებში მიღებული ინფორმაციის საფუძველზე. ცვლადების ხასიათის მიხედვით განასხვავებენ პირობით და უპირობო პროგნოზირებას. პროგნოზირება უპირობოა, თუ დამოუკიდებელი ცვლადების მნიშვნელობები ზუსტადაა განსაზღვრული. პრაქტიკაში უფრო ხშირად გვხვდება სიტუაცია, როდესაც დამოუკიდებელი ცვლადების შესახებ ინფორმაცია შეცდომებს შეიცავს. მაგალითად, დროითი მწკრივების პროგნოზირებისას. ამ შემთხვევაში საქმე გვაქვს უპირობო პროგნოზირებასთან.

უპირობო პროგნოზირება

თუ მოცემულია დამოუკიდებელ ცვლადების მიმდევრობა X_1, X_2, \dots, X_{T+1} რომლისთვისაც რეგრესიის მოდელი აგებულია:

$$Y_t = a + bX_t + v_t, t = 1, 2, \dots, T$$

და მოდელის შეცდომა - v_t ნორმალურადაა განაწილებული ნულოვანი მათემატიკური მოლოდინითა და დისპერსიით σ^2 -ით, შეიძლება დამოკიდებული ცვლადის მნიშვნელობის პროგნოზირება $T+1$ პერიოდისათვის (ჩავთვალოთ, რომ პარამეტრები: a, b გამოთვლილია) და შესაბამისი მოდელი შემდეგი სახით წარმოვადგინოთ:

$$\hat{Y}_{T+1} = \hat{a} + \hat{b}X_{T+1}$$

მაშინ პროგნოზის შეცდომაა:

$$\hat{e}_{T+1} = \hat{Y}_{T+1} - Y_{T+1}$$

რომელსაც ორი საინტერესო თვისებები გააჩნია:

$$E(\hat{e}_{T+1}) = 0; E(\hat{e}_{T+1})^2 = \sigma^2$$

რის გამოც, შეფარდება $\hat{e}_{T+1} = \frac{\hat{Y}_{T+1} - Y_{T+1}}{\sigma}$ განაწილების ნორმალურ კანონს ემორჩილება ნულოვანი საშუალო მნიშვნელობით და 1-ის ტოლი სტანდარტული გადახრით.

რადგან σ^2 -ის მნიშვნელობა თითქმის ყოველთვის უცნობია, ის შეიძლება შევაფასოთ გამოსახულებით:

$$S^2 = \frac{1}{T-2} \sum (y_t - \hat{y}_t)^2$$

მაგალითი. მოცემულია სტუდენტის სწავლის დონის შეფასებასა და მისი ოჯახის შემოსავალს შორის დამოკიდებულება. შედეგები მოყვანილია ცხრილში.

ცხრილი 10.1.

სტუდენტის საშ. შეფასება, ბალი Y	მშობლების წლიური შემოსავალი, ათასი \$ X
2.5	12.00
2.5	6.00
3.5	18.00
3.0	12.00
2.0	9.00
3.5	15.00
3.0	15.00
4.0	21.00

ცხრილის მონაცემების მიხედვით წრფივი რეგრესიის განტოლებაა:

$$Y_i = 1.375 + 0.12X_i$$

$$\bar{X} = 13.5; s^2 = 0.109; N = 8$$

Y ცვლადის პროგნოზის ანალიზი X დამოუკიდებელი ცვლადის ცვლილებით გამოწვეული, მოცემულია ცხრილით.

ცხრილი 10.2.

X_{N+1}	Y_{N+1}	s^2	$\hat{Y}_{N+1} - 2.44s^2$	$\hat{Y}_{N+1} + 2.44s^2$
27.5	4.67	0.54	3.44	5.91
24.0	4.15	0.20	0.07	5.24
20.5	3.83	0.16	2.89	4.80
17.0	3.32	0.13	2.43	4.20
13.5	2.99	0.12	2.14	3.85
10.0	2.57	0.13	1.71	3.44
6.5	2.16	0.16	1.18	3.12

პირობითი პროგნოზირება

პირობითი პროგნოზირების შემთხვევაში დამოუკიდებელი ცვლადების მნიშვნელობები შეცდომების მატარებლები არ არიან, მაგრამ ასეთი სიტუაცია ნაკლებადაა მოსალოდნელი, რადგან დამოუკიდებელი ცვლადების სტოქასტური ბუნება გავლენას ახდენს დამოუკიდებელი ცვლადის პროგნოზზე, რის გამოც პროგნოზის შეცდომის გამოსათვლელი ფორმულის დადგენა შედარებით რთულია. გარდა ამისა, პროგნოზის შეცდომის დისპერსია უფრო მაღალია. დამოუკიდებელი ცვლადის პროგნოზირებას შემოაქვს პროგნოზის დამატებითი შეცდომა რაც კიდევ უფრო ართულებს საკითხს.

რეზიუმე

პროგნოზირება სისტემის მომავალი მდგომარეობის ანალიზია გასული და მიმდინარე პერიოდებში მიღებული ინფორმაციის საფუძველზე. ცვლადების ხასიათის მიხედვით განასხვავებენ პირობით და უპირობო პროგნოზირებას. პროგნოზირება უპირობოა, თუ დამოუკიდებელი ცვლადების მნიშვნელობები

ზუსტადაა განსაზღვრული. პრაქტიკაში უფრო ხშირად გვხვდება სიტუაცია, როდესაც დამოუკიდებელი ცვლადების შესახებ ინფორმაცია შეცდომებს შეიცავს. მაგალითად, დროითი მწკრივების პროგნოზირებისას. ამ შემთხვევაში საქმე გვაქვს უპირობო პროგნოზირებასთან.

დამოუკიდებელი სამუშაო

1. მოცემულია ცხრილი ორი ფირმისათვის საერთო ღირებულება და პროდუქციის გამოსავალი.

N	C1	Q1	C1	Q1
1	232	3.65	353	6.99
2	293	4.32	240	4.43
3	564	9.02	271	4.78
4	549	8.79	323	5.81
5	313	5.11	332	5.39
6	217	2.38	256	2.68
7	420	7.19	140	1.35
8	204	2.08	221	3.17
9	733	9.87	418	7.02
10	169	1.30	292	4.33
11	550	9.09	582	9.34
12	221	2.75	193	1.57
13	278	4.93	282	3.68
14	374	6.53	163	1.04
15	497	8.28	326	4.89
16	234	2.72	763	9.93
17	326	5.36	464	7.87
18	252	3.74	350	5.84
19	183	2.09	272	4.33
20	207	3.10	208	1.93
21	280	4.70	192	1.41
23	459	8.04	262	5.49
24	214	2.32	219	3.30
25	373	6.78	208	2.34
26	393	7.42	289	4.13
27	237	3.23	340	6.82
28	445	7.44	132	1.48
29	482	7.99	434	7.49

დაშვებულია ჰიპოთეზა, რომ ამ ფირმებისათვის ღირებულების ფუნქცია კუბურ ხასიათს ატარებს და შეიძლება ჩაიწეროს გამოსახულებებით:

$$\text{ფირმა 1. } C_{1t} = S_1 + S_2 Q_{1t} + S_3 Q_{1t}^2 + S_4 Q_{1t}^3 + e_{1t}$$

$$\text{ფირმა 2. } C_{2t} = X_1 + X_2 Q_{2t} + X_3 Q_{2t}^2 + X_4 Q_{2t}^3 + e_{2t}$$

სადაც,

$$E(e_{1t}) = E(e_{2t}) = 0$$

$$E(e_{1t}^2) = \tau_1^2; \dots; E(e_{2t}^2) = \tau_2^2$$

ამასთან, e_{1t} და e_{2t} ერთმანეთისაგან დამოუკიდებლები არიან.

ა/ შეაფასეთ თითოეული ფუნქცია უმცირეს კვადრატთა მეთოდით. ახსენით შედეგები

ბ/ ჩაატარეთ ჰიპოთეზის ტესტირება:

$$H_0 : \tau_1^2 = \tau_2^2$$

$$H_1 : \tau_1^2 \neq \tau_2^2$$

გამოიყენეთ 10%-იანი მნიშვნელოვნების დონე.

თავი 11. პროექტი ეკონომეტრიკაში

ეკონომეტრიკაში პროექტის დამუშავება ითვალისწინებს: ორიგინალური ეკონომეტრიკული მოდელის აგებას, პროექტით გათვალისწინებული მონაცემების შეგროვებას, მოდელის შესაფასებლად ეკონომეტრიკის ტექნიკის (პროგრამული უზრუნველყოფის) გამოყენებას და მიღებული შედეგების ინტერპრეტირებას. პროექტის დამუშავება აუცილებელია სტუდენტებისათვის მათი ლოგიკური აზროვნებისა და კვლევითი უნარ-ჩვევების გამომუშავების მიზნით. ეკონომეტრიკული პროექტის ძირითადი კვანძებია: **მოდელი, მონაცემები, მოდელის შეფასება, გაფორმება**

11.1. მოდელი

მონაცემების მოპოვება და მოდელის ფორმულირება პროექტის დამუშავების საწყისი მომენტებია. მოდელის ფორმულირებისას პირველი ნაბიჯი სამუშაო თემის შერჩევა და მის მიხედვით მოდელის მიზნისა და დანიშნულების განხილვა. კერძოდ, აუცილებელია დაზუსტდეს შესასწავლი თემის მიზნები, რომელი ჰიპოთეზები უნდა იყვეს ტესტირებული, რომელი ცვლადები უნდა იყვეს განსაზღვრული. ყურადღება უნდა გამახვილდეს ადეკვატური მონაცემების მოძიების შესაძლებლობაზე, რაც ნიშნავს მოდელში შესაყვანი ცვლადების შერჩევას მათ შორის კავშირების გათვალისწინებით. ამ კონტექსტში მოდელის ყველაზე გავრცელებული ტიპია ერთდროულ განტოლებათა მოდელი.

სტუდენტის მიერ შერჩეული მოდელი და მისი კვლევა შეიძლება იყოს ნებისმიერი მიმართულებით. მაგალითად: მარკეტები, ეკონომიკური, ინფლაცია, პოლიტიკური განვითარება, შობადობა და სიკვდილიანობა, დასაქმება, არჩევნები, ხელფასები, მოთხოვნა-მიწოდება, გაათლება, დემოგრაფია, კრედიტები, ქვეყნის შიდა პროდუქტის ზრდა, ბანკების მომგებიანობა, ასაკის მიხედვით წამლების გამოყენება, ქალების წილი შრომის ბზარზე, ლუდის მოხმარება, ჰაერის დაბინძურება, ბავშვთა სიკვდილიანობის სოციალური დეტერმინანტები, ფულის მიმოქცევა, არჩევნები, ოქროს ფასი, რელიგის

მონაწილეობა სამოქალაქო ცხოვრებაში, დაუსაქმებლობა და დანაშაული, ახალგაზრდობის დასაქმება, ტურიზმის განვითარება, ეთნიკური უმცირესობები და უმადლესი განათლება და სხვა.

11.2. სტატისტიკური მონაცემები

მონაცემები, როგორც ეკონომეტრიკის სწავლების არსებითი ინგრედიენტი, მნიშვნელოვანია და ხშირად ეკონომეტრიკული პროექტის საკმარისად კრიტიკული ნაწილია. მონაცემები უნდა იყვეს მოძიებული მოდელში შემავალი ყველა, როგორც ენდოგენური ისე ეგზოგენური, ცვლადებისათვის. მონაცემების მოძიების საინტერესო წყაროს წარმოადგენს ეროვნული და საერთაშორისო წელიწადეულები, სტატისტიკური აბსტრაქტები, გამოკითხვის შედეგები და სხვა.

მონაცემები აუცილებელია შევისწავლოთ და შეძლება გაიფილტროს პროექტის მიზნებიდან გამომდინარე. ქვემოთ, ცხრილი 1-ში მოცემულია პოპულარული სტატისტიკური აბსტრაქტების ჩამონათვალი.

№	დასახელება	№	დასახელება
1	მოსახლეობა	14	მეცნიერება
2	იმიგრაცია	15	ტრანპორტი
3	განათლება	16	სოფლის მეურნეობა
4	დასაქმება	17	სატყეო მეურნეობები
5	სოციალური დაზღვევა	18	მშენებლობა
6	სამუშაო ძალა	19	წარმოება
7	შემოსავალი, დანახარჯები	20	თევზჭერა
8	ფასები	21	დისტრიბუცია
9	ბანკები, ფინანსები, დაზღვევა	22	წიაღისეულის მოპოვება
10	ბიზნესის წარმოება	23	ნავიგაცია
11	არჩევნები	24	შინამეურნეობები
12	კომუნიკაციები	25	ბიუჯეტი
13	ენერჯია	26	ჰაერისა და წყლების დაბინძურება

ასევე შეიძლება გამოყენებული იყვეს წელიწადეულები როგორც ეროვნულ, ისე საერთაშორისო დონეებზე: მოსახლეობა, კომუნიკაცია, მეცნიერება და ტექნოლოგია, ხელფასები და ფასები, სახელმწიფო ფინანსები, საგარეო ვაჭრობა, ენერჯია, ჯანმრთელობა, სამრეწველო პროდუქცია...

11.3. მოდელის შეფასება და ანალიზი

მოდელის აგებისას და მოდელის შესაფასებლად ვიყენებთ ეკონომეტრიკის ტექნოლოგიას, რომელთა შორის ყველაზე პოპულარულია უმცირეს კვადრატთა მთოდი (უკმ - OLS) და 2 საფეხურიანი კვადრატთა მეთოდი (2სკმ – 2SLS). მოდელის აგება ყოველგვარი სირთულეების გარეშე შეიძლება პროგრამული უზრუნველყოფის გამოყენებით.

11.4. პროექტის გაფორმება

1. **თავფურცელი, რომელიც მოიცავს:** სათაურს, ყველა ავტორის სახელსა და გვარს;

2. 1 გვერდიანი აბსტრაქტი თემის მოკლე აღწერით; ცვლადებისა და მოდელის აღწერით; მონაცემების ტიპების გამოყენების შესახებ მსჯელობა გამოყენებული წყაროს მითითებით; შეფასებული განტოლების დამაჯერებლობა და მოდელის ანალიზი; პროგნოზირება და რაც მეტად მნიშვნელოვანია - დასკვნები;
3. შესავალი, რომელიც მოიცავს: კვლევის მიზანს, მოდელის ბუნებას, ცვლადები, რომლებსაც მოდელი მოიცავს და რატომ? რის გაკეთებაა დაგეგმილი მოდელის შესასწავლად და მისი პროგნოზირებისათვის ტესტირების ჩავლით;
4. ლიტერატურის მიმოხილვა, რომელიც შერჩეულ კვლევის თემასთანაა დაკავშირებული მოკლე ანოტაციებით იმ წიგნებისა და სტატიების შესახებ, და რაც კვლევის საფუძველს იძლევა;
5. მოდელის აღწერა მოიცავს: იმ ცვლადების შესახებ მსჯელობას, რომლებიც მოდელში იქნება გათვალისწინებული; დისკუსიას მოდელში შემავალი ცვლადების კოეფიციენტების შესახებ, მათი ნიშნებისა და მათი გავლენის შესახებ მოდელზე.
6. მონაცემების აღწერა მოიცავს ყველა მონაცემის შესახებ ინფორმაციას მათი ბუნებისა და წყაროების შესახებ;
7. დასკვნა შეიცავს იმ ძირითად შედეგების ჩამონათვალს, რისთვისაც მოდელი დამუშავდა, შედეგების შედარებას საკვლევი მიმართულებაში არსებულ მდგომარეობასთან, მომავლის პერსპექტივას.
8. მოდელის შეფასება მოიცავს პარამეტრების სტრუქტურულ ფორმას, შესაბამის სტატისტიკებს და კოეფიციენტების მნიშვნელოვნების დონეების ანალიზს, ასევე: დეტერმინაციის კოეფიციენტისა, t-ტესტისა და დურბინ-უატსონის (Durbin-Watson) სტატისტიკებს თითოეული განტოლებისათვის და მსჯელობას მათი შედეგების მიხედვით; ჰეტეროსკედასტიურობის ანალიზს; შესაძლო პრობლემების შესახებ მსჯელობას და თუ შესაძლებელია, მოდელის აგებას მონაცემთა სხვადასხვა მიმდევრობების მიხედვით და შემდეგ, კვლავ მიღებული შედეგების შეფასებას.

თავი 12. ეკონომეტრიკის პროგრამული უზრუნველყოფა

12.1. პროგრამული უზრუნველყოფების მოკლე მიმოხილვა

ეკონომეტრიკის თავდაპირველი პროგრამული უზრუნველყოფა, რომელიც ეკონომეტრიკის პაკეტების სახითაა წარმოდგენილი, დაიწერა პერსონალური კომპიუტერების გამოჩენამდე და ორიენტირებული იყო დიდი ელექტრონული გამომთვლელი მანქანებისათვის (ეგმ). ფაქტობრივად, ისინი წარმოადგენდნენ FORTRAN -ზე დაწერილ პროცედურების ნაკრებს და პერსონალურ კომპიუტერებზე მათი გადატანისათვის პროცედურების ნაკრებს მიემატა ბრძანებების ენაც რის გამოც, პარამეტრებით ისინი ვერ უტოლდებიან პერსონალური კომპიუტერებისთვისაა დაწერილი ისეთ ეკონომეტრიკულ პაკეტებს, როგორცაა: გამოყენების ეფექტურობა, დიზაინი, ინტერფეისი და სხვა. პერსონალური კომპიუტერებისათვის განკუთვნილი ეკონომეტრიკული პაკეტები ძირითადად დაწერილია "C" ენაზე, უმცა პაკეტები, რომლებიც დაწერილია FORTRAN -ზე (TSP, STATA, SPSS, SYSTAT, SAS) იქმნებოდა და ვითარდებოდა

წლების განმავლობაში. უკანასკნელ წლებში ეს პაკეტები დამუშავდა სხვადასხვა პლატფორმებისთვის როგორცაა: DOS, WINDOWS, UNIX, MACINTOSH.

SAS (WINDOWS): ეკონომეტრიკული (სტატისტიკური) პაკეტიდან მოცულობით ყველაზე დიდია და მოიცავს გამოყენებითი სტატისტიკისა და მონაცემების ანალიზის თითქმის ყველა სფეროს. დამუშავებულია თითქმის ყველა პლატფორმისათვის. პაკეტის სრულყოფილი და სწრაფი ათვისება ფერხდება დოკუმენტაციის დიდი მოცულობის გამო. ლაგური ცვლადების გამოყენებისას პაკეტი ვერ მუშაობს, არ გააჩნია მოდელის კოეფიციენტების შეცდომის სტანდარტული შეფასებების საშუალებები. თანამედროვე ეკონომეტრიკის მეთოდების გამოყენებისას ვერ უთანაბრდება სპეციალიზირებულ ეკონომეტრიკულ პაკეტებს. SAS -ის ინსტიტუტი თავის მომხმარებლებისათვის უშვებს სპეციალურ ჟურნალს.

GAUSS (DOS): დისკზე იკავებს 4,5 მბაიტს. იგი წარმოადგენს შესანიშნავ პროგრამას მატრიცებთან მუშაობის თვალსაზრისით. არსებობს ამ ენაზე დაწერილი ეკონომეტრიკის მეთოდების დიდი რაოდენობა. ზოგიერთ ეკონომეტრისტს მიაჩნია, რომ ამ ენაზე პროგრამების შედგენა მიზანშეწონილია არა სტანდარტული მეთოდების გამოყენებისას. World Wide Web ~WWW)-ის გამოყენებით შეიძლება მივიღოთ ამ ენაზე დაწერილი პროგრამების დიდი რაოდენობა, მაგრამ მათი გამოყენებისას დაზღვეული არ ვართ პროგრამებში დაშვებული შეცდომებისაგან.

STATA 3.0 (1995, DOS), (1998 , WINDOWS): STATA 7.0, 9.0 (2000, WINDOWS) იკავებს 2-მბაიტს ვინჩესტერზე, ორიენტირებულია ეკონომისტებისათვის, ადვილია ასათვისებლად, მოიცავს მოდელის კოეფიციენტების სტანდარტული შეცდომების შეფასებებს, რომლებიც თავსებადია ჰეტეროსკედასტიურობასა და სხვაობების ავტოკორელაციასთან, არ აქვს არაწრფივი რეგრესიის ამოცანების ამოხსნის საშუალებები.

TSP (4.3 1995, DOS) დისკზე იკავებს 2,2 მბაიტს, ფართოდაა წარმოდგენილი არაწრფივი მოდელების ამოხსნის საშუალებები, აქვს ანალიზური დიფერენცირების საშუალება, სრულდება მოდელის კოეფიციენტების სტანდარტული შეცდომების შეფასებები, თავსებადია ჰეტეროსკედასტიურობისა და სხვაობების ავტოკორელაციასთან, აქვს ერთდროულ განტოლებათა სისტემების განყოფილება და რაც არა ნაკლებ მნიშვნელოვანია, აქვს მოხერხებული დოკუმენტაცია.

RATS (DOS): მისი შედარება შეიძლება TSP -თან, მაგრამ ჩამორჩება მას მონაცემთა ბაზებთან მუშაობით და ბრძანებების სრულყოფით. ამ პროგრამის ავტორები ეკონომეტრისტები არიან და მას წარმატებით იყენებენ კვლევით სამუშაოებში.

SPSS v.6.1.2 (1995, WINDOWS): დისკზე იკავებს 20 მბაიტს, აცვს მოხერხებული ინტერფეისი და სტატისტიკური მეთოდების დიდი რაოდენობა, არ აქვს მოდელების კოეფიციენტების სტანდარტული შეცდომების შეფასებების საშუალებები, თავსებადია ჰეტეროსკედასტიურობასა და სხვაობების ავტოკორელაციასთან.

STSTAT v.5.03 (1994, WINDOWS): დისკზე იკავებს 5.2 მბაიტს და ერთ-ერთი პირველი პროგრამაა **WINDOWS** -თვის.

Microsoft v.3.11 (1992, DOS): დისკზე იკავებს 1 მბაიტს, ავტორი სტატისტიკის პროფესორია, მოიცავს მთელ რიგ თავსებად ეკონომეტრიკურ მეთოდებს, მისი მართვა მენიუს გამოყენებით შეიძლება, ადვილია ასათვისებლად.

MicroTSP v.7.03 (1994, DOS): დისკზე იკავებს 1.5 მბაიტს, მოიცავს ეკონომეტრიკის მრავალ თანამედროვე მეთოდს, მოიცავს არაწრფივ რეგრესიას, იმართება მენიუს გამოყენებით.

Econometric View v.2.0 (1995, WINDOWS): დისკზე იკავებს 4.7 მბაიტს, აქვს საკმაოდ მოხერხებული **Help**, რომელიც პრაქტიკულად, ეკონომეტრიკული მეთოდების ცნობარს წარმოადგენს.

- **1.2 (1995, DOS):** დისკზე იკავებს 2.0 მბაიტს. არსებობს მისი სხვა და სხვა ვერსიები: რუსულენოვანი, ინგლისურენოვანი და ფრანგულენოვანი, აქვს კარგი ინტერფეისი, გრაფიკული რედაქტორი და დროითი სკალების ფართო დიაპაზონი, მოხერხებულია სამუშაოდ ტრენდებთან, სეზონურ მოდელებთან, ერთდროულ განტოლებათა სისტემებთან, არ გააჩნია მაქსიმალური დასაჯერისობის მეთოდი, ჰეტეროსკედასტიურობასა და ავტოკორელაციასთან თავსებადია მოდელების კოეფიციენტების სტანდარტული შეცდომების შესაფასებლად, შესაძლებელია: ერთდროულ განტოლებათა სისტემებისათვის სტანდარტული შეცდომების შეფასებები, რეგრესიის კოეფიციენტებისთვის შეცდომების კორექცია სხვაობების ავტოკორელაციის შემთხვევაში.

OLIMP (WINDOWS): პაკეტი გათვალისწინებულია ფინანსური, ეკონომიკური, საინჟინრო და კვლევის მონაცემების ანალიზისა და პროგნოზირებისათვის. საბაზო ვერსია ორიენტირებულია ეკონომიკური დროითი მწკრივების ანალიზზე. პროფესიონალურ ვერსიაში გარდა მითითებულისა, წარმოდგენილია ფაქტორული, კლასტერული, სტრუქტურული ანალიზი და სხვა. მისი გავრცელება ხდება საინფორმაციო ტექნოლოგიების დეპარტამენტის -ს მიერ.

მეზაზორი WINDOWS 1.0 (1995, WINDOWS): დისკზე იკავებს 2.4 მბაიტს, ერთდროულად მუშაობს სხვადასხვა მოდელებთან და სხვადასხვა მონაცემებთან, რაც მუშაობის ეფექტიურობას ამაღლებს. პაკეტი გათვალისწინებულია ძირითადად, დროითი მწკრივების ანალიზისათვის და მოიცავს გაგლუვების მეთოდებს, წრფივ და არაწრფივ ტრენდებს, ავტორეგრესიის მოდელებს, ბოქს-ჯენკინსის მოდელებს და სხვა. ამ პაკეტის გამავრცელებელია - (მოსკოვი).

STATGRAPHICS 3.0; 5.0. (DOS): პაკეტის დამუშავება დაიწყო 80-იანი წლებიდან ამერიკული კორპორაციის { - } ამჟამად პაკეტის რეალიზაციას რუსეთის ბაზარზე ეწევა. პაკეტით წარმოდგენილია სტატისტიკური მეთოდების ფართო სპექტრი: ვარიაციული ანალიზი, დროითი მწკრივების ანალიზი, პროგნოზირება, მცოცავი საშუალოს ანალიზი, რეგრესიული ანალიზი, არაპარამეტრული მეთოდები, მრავალცვლადიანი მეთოდები, ორ- და სამგანზომილებიანი შესანიშნავი გრაფიკა. კორპორაცია კარდინალურ მუშაობას

ეწევა პაკეტის შინაარსის დახვეწისა და გაუმჯობესების მიზნით. პაკეტის გამოყენებას აფერხებს კლავიშების კომბინაციების ათვისებასთან დაკავშირებული სიძნელეები.

FORECAST EXPERT (WINDOWS 3.1 ან მეტი) გამოიყენება ვიწრო სპეციალიზაციით: დროითი მწკრივების ანალიზისათვის, სეზონური და არა სეზონური ARIMA მოდელისათვის და მათ ბაზაზე პროგნოზირებისათვის. პაკეტის საინტერესო მახასიათებელია მოდელის ავტომატური ტესტირება, რომელიც მომხმარებლისაგან არ ითხოვს მათემატიკური სტატისტიკის ცოდნას.

პაკეტი დამუშავებულია მოსკოვის ფირმის - “ - ”-მიერ და წარმოადგენს პროგრამების: Project Expert, Project Questionnaire & Risk, Invest Expert –ის გაგრძელებას.

- (WINDOWS)% პაკეტი გათვალისწინებულია რეგრესიული და ფაქტორული ამოცანების ფართო კლასისათვის და მოიცავს დროით მწკრივებთან დაკავშირებულ საკმაოდ ეფექტურ პროცედურებს. რეალიზაციაშია პაკეტის შემდეგი ვერსიებიც.

პაკეტი დამუშავებულია კარელიის მეცნიერებათა აკადემიის მიერ ფირმა ” ” -თან ერთად და ვრცელდება ამ ფირმისა და მისი დილერების მიერ.

გორის უნივერსიტეტის ეკონომეტრიკის ცენტრი სასწავლო პროცესისათვის იყენებს ეკონომეტრიკის პროგრამულ უზრუნველყოფებს:

1. Small STATA 5.0,
2. STATA 9.0
3. Econometric View
4. SPSS

12.2. ეკონომეტრიკის პროგრამული უზრუნველყოფა

- S T A T A -

Stata სტატისტიკური პაკეტია პროცესების მართვის, ანალიზის და მონაცემების გრაფიკული წარმოსახვისათვის. **Stata** -ს შესაძლებლობები ერგება შემდეგ პლატფორმებს:

Stata for Windows

Stata for windows 3.1

Stata for DOS

Stata for Power Macintosh

Stata for 680x0 Macintosh

Stata for Convex

Stata for DEC Alpha

Stata for HP-9000

Stata for Linux
Stata for RS/6000
Stata for SPARC

Windows და **Macintosh** –თვის გამოიყენება Small Stata და Intercooled Stata. DOS-ის და Unix -თვის არსებობს მხოლოდ Intercooled Stata. მართალია, ორივეს ერთი და იგივე ფუნქცია გააჩნია, მაგრამ Intercooled Stata შედარებით დიდ მონაცემთა ბაზებთან მუშაობს და უფრო სწრაფიცაა.

ნებისმიერ შემთხვევაში, განსხვავებები ტექნიკური და შიდაპროგრამულია. მომხმარებლის თვალსაზრისით Small Stata და Intercooled Stata ერთნაირად მუშაობს. ჩვენს მიერ განხილული ვერსია არის Small Stata . ეს ვერსია იმით განსხვავდება სხვა ვერსიებისგან, რომ მას გააჩნია მონაცემთა მოცულობაში შეზღუდვები, გვაძლევს მცირე მაკროსების შექმნის საშუალებას, აქვს შედარებით მოკლე ბრძანებების ხაზი. ნებისმიერ პლატფორმაზე Stata ბრძანებებით იმართება.

Stata ძალიან სწრაფია. მისი სიჩქარე ნაწილობრივ, ჭკვიანურად დამუშავებული პროგრამის შედეგია, ნაწილობრივ კი იმისა, რომ Stata მონაცემებს ინახავს მეხსიერებაში. Stata -ს მონაცემები შეიძლება იყოს დისკზეც ასლის სახით. მონაცემები იტვირთება მეხსიერებაში, სადაც იგი მუშაობას იწყებს, აკეთებს ანალიზს, შეაქვს ცვლილებები, შემდეგ კი ეს ყველაფერი კვლავ ინახება დისკზე. მონაცემების მეხსიერებაში შენახვა აჩქარებს Stata –ს მუშაობას, ხოლო ასლზე მუშაობა მეხსიერებაში მას ანიჭებს უსაფრთხოებას. ერთადერთი გზა იმისა, თუ როგორ მივაყენოთ ზიანი დისკზე არსებულ მონაცემთა პერმანენტულ ასლს ასეთია: შენახვისას გადავაწერთ ახალ მონაცემებს ძველს. მონაცემების მეხსიერებაში ინფორმაციის არსებობა შეზღუდულია მეხსიერების მოცულობიდან გამომდინარე. Stata მონაცემებს მეხსიერებაში შეკუმშულ ფორმატში ათავსებს და შეუძლია ჩაეტიოს მოცემული მეხსიერების მინაკვეთში. მიუხედავად ამისა, თუ თქვენ დიდი მოცულობის მონაცემთა ბაზებთან მუშაობთ შეგიძლიათ მეხსიერების მთლიანობის შეზღუდვას მიმართოთ. შესაძლებელია, მეხსიერების მაქსიმალურად შეკუმშვა ან Stata –ს მიერ მეხსიერებისთვის განკუთვნილი ერთი მეგაბაიტის გაზრდა.

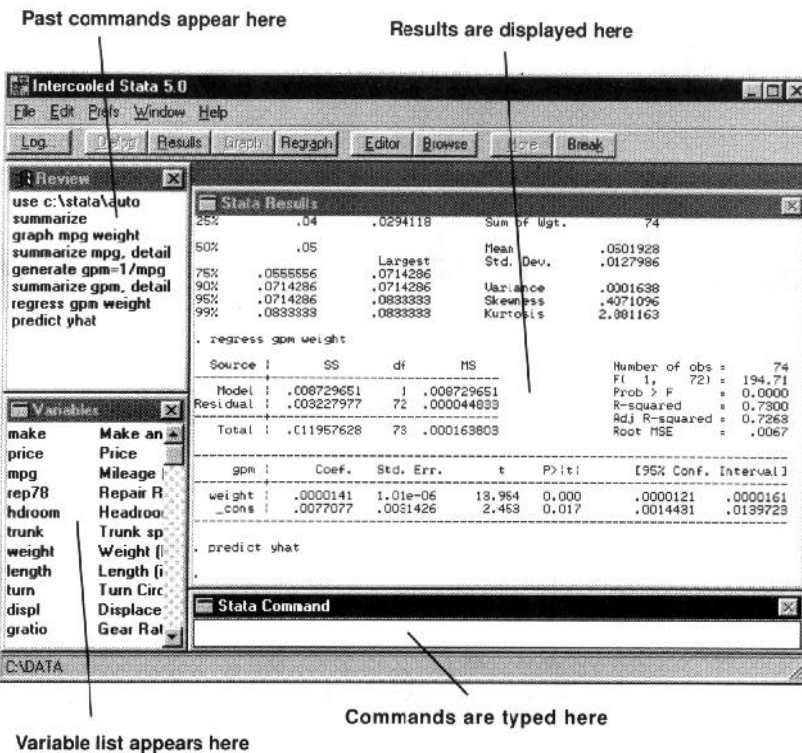
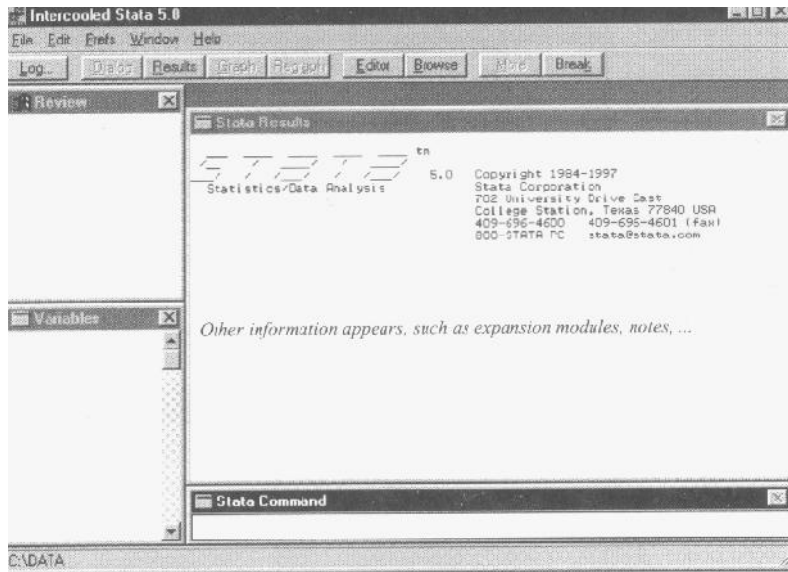
STATA Windows -თვის

მოკლე აღწერა

1. პროგრამის გაშვება

1. Start;
2. Program;
3. Stata – Small Stata 5.0

ეკრანზე გამოქნდება ფანჯარა სამუშაო მდგომარეობაში:



როგორც ვხედავთ, მარცხნივ ორი მოზრდილი ფანჯარაა (Review Variables)^ რომელთაგან ერთი Review განკუთვნილია იმ ბრძანებების დასახსომებლად, რომლებიც ამოცანის ამოხსნისას Stata- - ს ბრძანებათა (STATA command) ფანჯარაში ჩაიწერება. გარდა ამისა, ამავე ფანჯარაში აისახება იმ ფილის სახელი და მისამართი, რომელიც სამუშაო პროცესშია ჩართული, მუდამ ინდიცირდება და ამ ფანჯარაში რჩება მანამდე, ვიდრე მას ჩვენვე არ წავშლით.

მეორე ფანჯარა – Variables ცვლადების წარმოსადგენი ფანჯარაა და მასში აისახება ყველა ის ცვლადი, რომლებიც მოცემულ ამოცანაში მონაწილეობენ. ამავე

ფანჯარაში აისახება ის ახალი ცვლადებიც, რომლებსაც მოცემული ამოცანის მიხედვით შევქმნით.

კომპიუტერის ეკრანზე, ზემოთ განთავსებულია პროგრამის მენიუ, რომლითაც ბრძანებების შესრულება შესაძლებელია “Click”-ით ანუ თავის ღილაკზე დაწკაპუნებით.

2. პროგრამის დახურვა

File – ის მენიუდან შევარჩიოთ ოფცია – Exit . მასზე დაწკაპუნების შემდეგ გამოჩნდება ფანჯარა შეკითხვით ”ვინარჩუნებთ თუ არა ცვლილებებს”. ბუნებრივია, ვითვალისწინებთ ჩვენთვის სასურველ ვარიანტს თავუნთა ღილაკზე დაწკაპუნებით. გარდა ამისა, პროგრამის დახურვა შეიძლება, თუ STATA -ს ბრძანებათა ფანჯარაში ჩაწერთ პროგრამის დახურვის ბრძანებას “Exit” და შევიტანთ მას მეხსიერებაში. შემდეგ, გამოტანილ ფანჯარაში დასმული კითხვის მიხედვით კვლავ ვირჩევთ შესაბამის ვერსიას ჩვენი მიზნის შესაბამისად.

3. დახმარება (Help)

Stata –ს კარგად განვითარებული Help აქვს. მისი გამოძახება შეიძლება ბრძანების სტრიქონში შესაბამისი ბრძანების – Help ჩაწერით ან მენიუდან მისი გამოძახებით.

4. მონაცემები და მონაცემთა ბაზები

•რიცხვები

რიცხვი შეიძლება შეიცავდეს ნიშანს, მთელ ნაწილს, ათობით ნიშანს, წილად ნაწილს, e ან E -ს. ჩაწერილი რიცხვი არ უნდა შეიცავდეს მძიმეს. მაგალითად, რიცხვი 1.024 აუცილებელია ჩაიწეროს როგორც 1024 ან 1024. ან 1024.0 . რიცხვების ჩასაწერად მისაღებია შემდეგი ვარიანტები:

5
-5
5.2
5.2 e+2
5.2 e-2

გამოტოვებული რიცხვი შეგვიძლია დავამატოთ მონაცემთა ბაზას, თუ ეს რიცხვი ჩაწერილი იქნება ზემოთ მოყვანილი რიცხვების ჩაწერის წესების დაცვით.

STATA იგნორირებას უკეთებს გამოტოვებულ რიცხვით მნიშვნელობას. მის ნაცვლად აღიქმება ნულოვანი მნიშვნელობა. მონაცემთა ბაზა შეიძლება შეიცავდეს შენიშვნას. როდესაც მონაცემის აღწერას ვაკეთებთ (ბრძანება: describe), ყურადღება უნდა მივაქციოთ შენიშვნებს მონაცემების ან ცვლადების მიმართ. მაგალითად, ბრძანებათა ფანჯარაში ჩაწერილი ბრძანების – describe - ის მეხსიერებაში შეტანის შემდეგ მივიღებთ:

```
Contain sdata  
Obs:          50  
Vars:         4  
Size: 1,200(99.5% of memory free) *data has notes
```

1. state	str8	%9s	*
2. medage	float	%9.0g	median age
3. mrgrate	long	%12.0g	marriages per 100 000
4. dvcrate	long	%12.0g	indicated variables has notes

5. კლავიატურის გამოყენება

F -კლავიშების დანიშნულება:

F - კლავიშა	დანიშნულება
F 1	help (დახმარება)
F 2	review (გახედვა)
F 3	describe (აღწერა)
F 7	save (შენახვა)
F 8	use (გამოყენება)

6. ოპერატორები

STATA –ს ოთხი სხვადასხვა კლასის ოპერატორები აქვს: არითმეტიკული, სტრიქონის, შედარების და ლოგიკური.

არითმეტიკული ოპერატორები:

1. მიმატება: "+" (addition);
2. გამოკლება: "-" (subtraction);
3. გამრავლება: "*" ("multiplication);
4. გაყოფა: "/" (division)

მაგალითი: გამოსახულება: $-(x+y)/(x*y)$ აღნიშნავს გამოსახულებას: $-\frac{x+y}{x \times y}$

სტრიქონის ოპერატორები:

“+” ნიშანი სტრიქონის ოპერატორია და იგი აერთიანებს ორ მიმდევრობით სტრიქონს.

შედარების ოპერატორები:

შედარების ოპერატორებია:

> (greater than), < (less than), >=greater than or equal), <=(less than or equal), =(equal), and !=(not equal).

ამ ტიპის გამოსახულებები "ჭეშმარიტია" ან "მცდარი". მაგალითად, გამოსახულება $3 > 2$ ჭეშმარიტია ისე, როგორც "zebra" > "cat".

შედარების ოპერატორების გამოყენებით შესაძლებელია გამოვყოთ მონაცემთა ბაზიდან ის მონაცემები, რომლებიც აკმაყოფილებენ მოთხოვნას შედარების ოპერატორის პირობის მიხედვით და შევქმნათ ახალი მონაცემთა ბაზა.

მაგალითად, ვთქვათ მონაცემთა ბაზებიდან ასაკისა და შემოსავლის მიხედვით გვინდა გამოვყოთ შემოსავლის მიხედვით მონაცემები, რომლებიც ასაკის მიხედვით დააკმაყოფილებენ პირობას: ნაკლებია ან ტოლი 25 ($B=25$); ამ შემთხვევაში ჩანაწერს ექნება შემდეგი სახე:

.list if ages <=25

თუ გვინდა გამოვყოთ მონაცემები ასაკის მიხედვით იმ ინდივიდებისათვის, რომლებიც აკმაყოფილებენ პირობას " $=25$ ", საჭიროა ჩავწეროთ:

.list if ages =25

შევნიშნოთ, რომ პროგრამა შეცდომით აღიქვამს გამოსახულებას, რომელიც ჩაწერილია ერთი ტოლობით:

ლოგიკური ოპერატორები:

ლოგიკური ოპერატორებია: **&(and); | (or); ~ (not)**

მაგალითად, ჩანაწერის მიხედვით

list if income > 50000 | income >30000 & age <25

შეიქმნება მონაცემების მიმდევრობა იმ შემოსავლებისა, რომლებიც აკმაყოფილებენ პირობას: მეტი 50000-ზე ან მეტი 30000-ზე და ამ შემოსავლების მქონე პირების ასაკი ნაკლებია 25-ზე.

7. ფუნქციები

STATA -ს მიერ აღიქმება შემდეგი ფუნქციები:

მათემატიკური ფუნქციები:

აბსოლუტური - $\text{abs}(x)$;

არკკოსინუსი რადიანებში გამოსახული არგუმენტით – $\text{acos}(x)$;

არკსინუსი რადიანებში გამოსახული არგუმენტით – $\text{asin}(x)$ *

არკტანგენსი რადიანებში გამოსახული არგუმენტით - $\text{atan}(x)$;

კომბინატორული ფუნქცია – $\text{comb}(n,k)$;

კოსინუსი რადიანებში გამოსახული არგუმენტით – $\text{cos}(x)$;

ექსპონენტა – $\text{exp}(x)$;

ლოგარითმი – $\ln(x)$;

ლოგარითმი ნებისმიერი ფუძით – $\log(x)$;

ათობითი ლოგარითმი – $\lg(x)$;

სინუსი, რადიანებში გამოსახული არგუმენტით – $\text{sin}(x)$;

მოდული – $\text{mod}(x,y)$;

კვადრატული ფესვი – $\text{sqrt}(x)$;

ტანგენსი რადიანებში გამოსახული არგუმენტით – $\tan(x)$.;

ფუნქციების მონაცემები:

STATA შეიცავს სხვა და სხვა ფუნქციების მონაცემების წარმოდგენის სხვა და სხვა სახეს და მათთვის გამოიყენება აღნიშვნები:

s – მონაცემები სტრიქონის სახითაა გამოსახული;

e, m, d, y – ფუნქციების სახითაა გამოსახული;

როგორც $\text{int}(e)$, $\text{int}(9m)$, $\text{int}(9d)$, $\text{int}(y)$

სტრიქონით წარმოდგენილი ფუნქციები

სპეციალური ფუნქციები:

8. მოდელები

მარტივი მოდელები;

მრავალგანტოლებიანი მოდელები;

დისპერსიული მოდელები.

12.3. Small STATA 5.0-ის ძირითადი ბრძანებები და განმარტებები

მონაცემების ფაილის ბრძანებები

describe - მეხსიერებასა ან დისკზე არსებული მონაცემების შინაარსის აღწერა;

edit - მონაცემების შეტანა კლავიატურიდან ან მისი რედაქტირება;

generate - ცვლადების შინაარსის შეცვლა ან ახალი მნიშვნელობების შექმნა;

list - ცვლადების მნიშვნელობების ჩაწერა;

obs - დაკვირვებების რიცხოვნების გაზრდა;

sort - მონაცემების დახარისხება;

display - ინფორმაციის ეკრანზე გამოტანა

მონაცემების მეხსიერებაში შეტანისა და შენახვის ბრძანებები

infile - stata -ში მონაცემების წაკითხვა ფაილიდან;

insheet - stata -ში მონაცემების წაკითხვა XL -ის ფაილიდან;

infile - არაფორმატირებული მონაცემების წაკითხვა;

infx - მონაცემების წაკითხვა ფიქსირებული ფორმატით;

inpt - კლავიატურიდან მონაცემების შეტანა;

outfile - ASCII ფორმატი მონაცემების აწერა;

save - მონაცემების შენახვა.

მონაცემების კომბინირების ბრძანებები

append - ცხრილში მოცემული მონაცემების შეკრება;
merge - მონაცემების შერწყმა.

მონაცემთა ჯგუფის ფორმირების ბრძანებები

collapse - მედიანების, საშუალო მნიშვნელობებისა და სხვა მაჩვენებლების ცხრილში ჩაწერა;
compress - მეხსიერებაში მონაცემების შემჭიდროვება;
expand - დაკვირვებების დუბლირება;
xpose - დაკვირვებების შიგნით ცვლადებისა და მათი შესაბამისი მონაცემების გადაადგილება;
obs - დაკვირვებების რიცხოვნების გაზრდა;

ჭდეებისა და შენიშვნების ბრძანებები

label - ჭდეების მანიპულაცია;
notes - მონაცემების შენიშვნების განთავსება.

ცვლადების ადგილების შეცვლა და სახელების გადარქმევა

rename - ცვლადისათვის სახელის გადარქმევა;
order - მონაცემთა ბაზაში ცვლადის ადგილის შეცვლა.

მონაცემების თვისებების შესწავლა (გამოკვლევა)

cf - მონაცემთა ორი ბაზის შედარება;
codebook - კოდირებული წიგნის შექმნა;
compare - ორი ცვლადის შედარება;
count - დაკვირვებების რიცხოვნების დადგენა;
gsort - ზრდადობისა ან კლებადობის მიხედვით მონაცემების დალაგება;
summarize - კრებსითი მაჩვენებლების გამოყვანა;
table - კრებსითი მაჩვენებლების ცხრილები;
tabdisp - ეკრანზე ცხრილის გამოტანა;
ipolate - წრფივი ინტერპოლაცია;
range - რიცხვითი რანგები, წარმოებულები და ინტეგრალები;
sample - შემთხვევითი შერჩევა;
ძირითადი დამხმარე საშუალებები

about - ინფორმაცია stata -ს ვერსიის შესახებ;
help - დახმარების მიღება;
look up - დოკუმენტაციის გადახედვა;
do - ფაილიდან ბრძანების შესრულება;
exit - **stata**- დან გამოსვლა
obs - მონაცემთა ბაზაში დაკვირვებათა რიცხოვნების გაზრდა;
#review - წინა ბრძანების გადახედვა.

მონაცემთა ტიპები და მეხსიერება

- compress** – მონაცემების შემჭიდროვება მეხსიერებაში;
- data types** – მონაცემების ტიპების სწრაფი მიმოხილვა;
- memory** – მეხსიერების მოცულობის განხილვა;
- recast** - ცვლადის ტიპის შეცვლა.

მაღალი დონის დამხმარე საშუალებები

- msg** – გზავნილების დაბრუნება;
- error messages** – გზავნილების შეცდომა და კოდების დაბრუნება;
- ado** – **ado-file** –ის მანიპულაცია;
- assert** – ჭეშმარიტებისა და სიყალბის შემოწმება;
- cd** - დირექტორიის (კატალოგის) შეცვლა;
- dir** – ეკრანზე ფაილის სახელის გამოტანა;
- discard** – ჩატვირთული პროგრამის ავტომატური ანულირება;
- disp_res** – შენახული შედეგების ეკრანზე გამოჩენა;
- for** – **stata-ს** ბრძანების გამეორება;
- set** – სისტემური პარამეტრების სწრაფი მიმოხილვა;
- type** – ეკრანზე ფაილების შინაარსის გამოტანა.

გრაფიკების ბრძანებები

- graph** – გრაფიკის აგების ბრძანება;
- graph axis labels** – ღერძების ჭდე;
- graph axis scale** – მასშტაბის დადგენა;
- graph bar** – დიაგრამების აგება;
- graph by** – სხვადასხვა ტიპის გრაფიკების გამოსახვა;
- graph connect** – წერტილების შეერთება;
- graph histogram** - ჰისტოგრამით წარმოდგენა;
- graph lines** - აღნიშვნებისა და წრფეების დამატება;
- graph log scales** – ლოგარითმული მასშტაბის შექმნა;
- graph printing** – გრაფიკის გამობეჭდვა ;
- graph saving** – გრაფიკის შენახვა გამობეჭდვისა ან რედაქტირებისათვის;
- graph shading** – გრაფიკის არეების შევსება;
- graph symbol** – გრაფიკის გამოსახვის სპეციალური სიმბოლოები;
- graph textsize** – გრაფიკისადმი ტექსტის კონტროლი;
- graph title** – გრაფიკის სპეციალური დასათაურება;
- gph** – დაბალი რანგის გრაფიკი;
- hist** – ცვლადის ჰისტოგრამა.

სტატისტიკები

- anova** – ვარიაციული და კოვარიაციული ანალიზი;

oneway –	ცალმხრივი ვარიაციული ანალიზი;
bitest -	ბინომიალური ალბათური ტესტი;
correlate –	ცვლადების კორელაცია;
logistic –	ლოგისტიკური რეგრესია;
regress –	წრფივი რეგრესია;
summarize –	კრებსითი მაჩვენებლების შექმნა;
table –	კრებსითი მაჩვენებლების ცხრილი;
tabulate -	სიხშირეების ცალმხრივი და ორმხრივი ცხრილები;
ttest –	საშუალო მნიშვნელობების შესადარებელი ტესტი.

არითმეტიკული, გეომეტრიული და ჰარმონიული საშუალო სიდიდეები:

pcorr –	კერძო კორელაციის კოეფიციენტი;
simul –	მონტე კარლოს მოდელირება.

წრფივი რეგრესია და მაქსიმალური დასაჯერისობის მეთოდი:

areg –	წრფივი რეგრესია დიდი რაოდენობის ფიქტიური ცვლადებით;
cnsmreg –	წრფივი რეგრესიის აგება;
corc –	კოკრან-ორკუტის რეგრესია;
fit –	რეგრესიის მრუდი და დიაგნოსტიკა;
glm –	განზოგადებული წრფივი მოდელები;
hlu -	ჰილდრეტ-ლუს რეგრესია;
mvreg –	მულტივარიაციული რეგრესია;
nl –	არაწრფივი უმცირეს კვადრატთა მეთოდი;
orthpoly –	ორთოგონალური მრავალწევრები;
poisson –	პუასონის რეგრესია;
rreg –	რობასტული რეგრესია;
regress -	წრფივი რეგრესია;
smooth –	არაწრფივი რობასტული მოსწორება;
sw –	მაქსიმალური დასაჯერისობის მეთოდის შეფასება;
vwls –	შეწონილი უმცირეს კვადრატთა მეთოდი

დროითი მწკრივები:

xt –	დროითი მწკრივების ანალიზი;
xtdes -	დროითი მწკრივების აღწერა;
xtgls -	დროითი მწკრივების მოდელის შეფასება განზოგადებული უმცირეს კვადრატთა მეთოდით;
xtreg -	წრფივი მოდელების შემთხვევითი ეფექტები.

ფაქტორული ანალიზი:

- alpha** – კრონზახის ალფა;
- canon** – კანონიკური კორელაცია;
- factor** – ფაქტორული ანალიზი;
- impute** – გამოტოვებული მნიშვნელობების დადგენა.

მატრიცების ბრძანებები:

- matrix** – მატრიცების ბრძანებების შესავალი;
- matrix define** – მატრიცების განსაზღვრა, ოპერატორები, ფუნქციები;
- matrix rowname** – სტრიქონებისა და სვეტების დასათაურება;
- matrix substitute** – მატრიცების შექმნა.

მონაცემები და მასთან დაკავშირებული ოპერაციები

მონაცემების შეტანა:

ამოცანის ამოხახსნელად მონაცემის შეტანის რამდენიმე გზა არსებობს და მათი შერჩევის შესაბამისად სხვადასხვა ბრძანება გამოიყენება:

- მონაცემის შეტანა კლავიატურიდან : გამოიყენება ბრძანება `edit ან input` ;
- მონაცემის შეტანა ფაილიდან: გამოიყენება ბრძანება `insheet`;
- არაფორმატირებული მონაცემების შეტანა ფაილიდან: გამოიყენება ბრძანება `infile (free format)`.
- ფორმატირებული მონაცემების შეტანა: გამოიყენება ბრძანება `infile f y infx (fixed format)` .

მონაცემების შეტანა კლავიატურიდან:

გამოვიყენოთ STATA-ს მონაცემთა ბაზიდან ერთერთი მათგანი - `auto.dta` მონაცემები სხვადასხვა მარკის მანქანების ზრგიერთი პარამეტრების შესახებ მათი მარკების მიხედვით:

Make	Price	MP G	Weight	Gear ratio
VW Rabbit	4967	25	1930	3.78
Olds 98	8814	21	4060	2.41
Chev. Monza	3667	--	2750	2.73
AMC Concors	4099	22	2930	3.58
Datsum 510	5079	24	2280	3.54
-----	5189	20	3280	2.93
Datsum 810	8129	21	2750	3.55

სადაც:

Make – მანქანის მარკის აღმნიშვნელი სტრიქონცვლდი;

Price –მანქანის ფასი;

MPG –ერთი გალონი საწვავის ხარჯვისას განვლილი მანძილი, მილი;

Weight –მანქანის წონა;

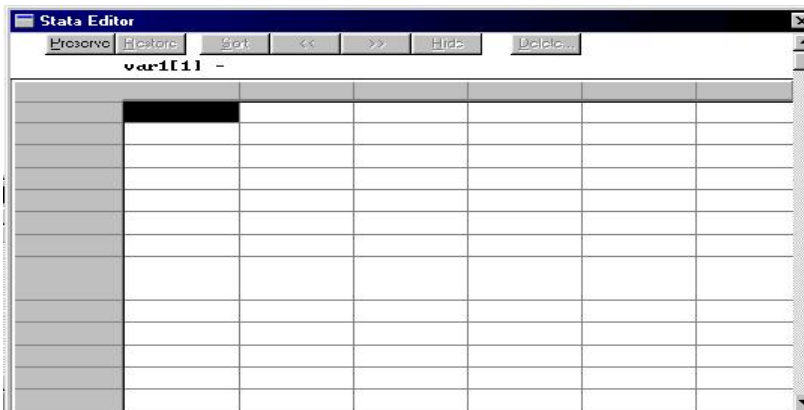
Gear Ratio -შეფასების კოეფიციენტი.

შენიშნეთ რომ არ ვიცით **MPG** მესამე მარკის მანქანისათვის და არც მეექვსე მანქანის მარკა. Stata-ს Editor- ის საშუალებით შევიტანოთ მონაცემები მის მესხიერებაში, რისთვისაც აუცილებელია:

1. გამოვიძახოთ Editor -ი თავუნიას მეშვეობით, რისთვისაც მენიუში დავუწკაპუნებთ Editor -ის ღილაკს ან შევიტანთ ბრძანებას "Edit" STATA -ს ბრძანებათა სტრიქონში;

2. ჩავწეროთ ცხრილის მონაცემები.

შეიძლება შევიტანოთ თითოეული ცვლადის მნიშვნელობა ერთმანეთის მიმდევრობით (ცვლადი ცვლადის შემდეგ) ან დაკვირვება დაკვირვების შემდეგ. სვეტები ასახავენ ცვლადების მნიშვნელობებს, ხოლო სტრიქონები – დაკვირვებების მნიშვნელობებს თითოეული ცვლადის მიხედვით.



Stata -ს მიერ შემოთავაზებული ფურცელი მონაცემების შესატანად

3. დაკვირვება დაკვირვების შემდეგ მონაცემების შეტანისას ვაწკაპუნებთ "Tab"-ის ღილაკზე თითოეული მნიშვნელობის შეტანისას.

მონაცემების შეტანა იწყება პირველი სვეტის პირველი უჯრიდან. შევიტანეთ VW Rabbit და დავაწკაპუნეთ "Tab"-ზე შემდეგ უჯრაში გადასასვლელად. შემდეგ შევიტანოთ ფასი 4697 და ისევ დავაჭირეთ "Tab"-ს. ეს პროცედურა გრძელდება მანამ, ვიდრე არ იქნება შეტანილი პირველი დაკვირვების ყველა მონაცემი. შემდეგ გადავართ პირველი სვეტის მეორე უჯრაში და შევიტანთ მეორე დაკვირვების ყველა მონაცემს და ა.შ.

4. მას შემდეგ, რაც პირველი დაკვირვების შედეგები შეტანილია, Stata აფიქსირებს შესატანი ცვლადების რაოდენობას და მეორე დაკვირვების მნიშვნელობების შეტანის შემდეგ "Tab"-ზე დაჭერისას Stata დაგაბრუნებთ მომდევნო სტრიქონის პირველ უჯრაში.

5. "ცვლადი ცვლადის შემდეგ" მონაცემების შეტანისას ვაწკაპუნებთ Enter -ის ღილაკზე თითოეული მონაცემის შეყვანის შემდეგ.

	var1	var2	var3	var4	var5
1	UV Rabbit	4697	25	1930	3.78

1. სტრიქონ მონაცემის შეყვანისას ბრჭყალები არ არის აუცილებელი, თუმცა Stata -ს სხვა ბრძანების შესრულებისას ბრჭყალები აუცილებელია;
2. ნიშანი “.” აღნიშნავს გამოტოვებულ რიცხვით ცვლადს;
3. გამოტოვებული რიცხვითი ცვლადის შეტანისას საჭიროა დვაწვეთ Tab -ის ან Enter -ის ღილაკს;
4. გამოსახულება var3[4] = აღნიშნავს მესამე ცვლადის მეოთხე დაკვირვების შედეგს.

ცვლადებისათვის სახელის გადარქმევა

როდესაც Stata -ს რედაქტორი ქმნის ახალ ცვლადს. მას ავტომატურად ანიჭებს სახელებს: var1, var2, var3,.....

	var1	var2	var3	var4	var5
1	UV Rabbit	4697	25	1930	3.78
2	Olds 98	8814	21	4060	2.41
3	Chev. Monza	3667	.	2750	2.73
4	AMC Concord	4099	22	2930	3.58
5	Datsun 510	5079	24	2280	3.54
6		5189	20	3280	2.93
7	Datsun 810	8129	21	2750	3.55

ცვლადისათვის სახელის გადასარქმევად საჭიროა:

1. ორჯერ დავაწკაპუნოთ შესაბამისი ცვლადის სვეტში და გამოტანილი ცვლადის ინფორმაციის სადიალოგო ფანჯარაში შევიტანოთ ცვლადის სახელი.

პირველ ცვლადს ვარქმევთ ” make”, მეორეს - ”price”, მესამეს –” mpg”, მეხუთეს –” gratio” და ა.შ.

ცვლადისათვის სახელის გადარქმევასთან დაკავშირებული წესები:

1. მნიშვნელოვანია სახელის ჩაწერის ფორმა. სასურველია სახელში ყველა სიმბოლო პატარა იყოს. მაგალითად, make, price, mpg და ა. შ.;
2. ცვლადის სახელი უნდა შეიცავდეს 1-დან 8 სიმბოლომდე;
3. სიმბოლოები შეიძლება იყოს როგორც ასოები (A-Z), ასევე ციფრები (9-0) და (_);
4. სხვა სიმბოლოები არ არის დაშვებული;
5. ცვლადის სახელში პირველი სიმბოლო აუცილებლად უნდა შეიცავდეს ასოს.

მონაცემების კოპირება და ჩაწება

1. სასურველი მონაცემის შერჩევა:
 - დააწკაპუნეთ ეთხელ ცვლადის სახელზე მთლიანი სვეტის მოსანიშნად;
 - დააწკაპუნეთ ერთხელ დაკვირვების ნომერზე მთლიანი სტრიქონის მოსანიშნად;
 - დააწკაპუნეთ და მონიშნეთ მონაცემების სასურველი არე.
2. შერჩეული მონაცემების კოპირება ბუფერში:
Edit –ის მენიუდან შეირჩევა – Copy;
3. მონაცემების ჩაწება ბუფერიდან:
-შეარჩიეთ ჩასასმელად სასურველი ადგილი

შეასრულეთ ბრძანებები შემდეგი თანამიმდევრობით: Edit – Paste.

	var1	var2	var3	var4	var5
1	VW Rabbit	4697	25	1930	3.78
2	Olds 98	8814	21	4060	2.41
3	Chev. Monza	3667	.	2750	2.73
4	AMC Concord	4099	22	2930	3.58
5	Datsun 510	5079	24	2280	3.54
6		5189	20	3280	2.93
7	Datsun 810	8129	21	2750	3.55

	var1	var2	var3	var4	var5
1	UV Rabbit	4697	25	1930	3.78
2	Olds 98	8814	21	4060	2.41
3	Chev. Monza	3667	.	2750	2.73
4	AMC Concord	4099	22	2930	3.58
5	Datsun 510	5079	24	2280	3.54
6		5189	20	3280	2.93
7	Datsun 810	8129	21	2750	3.55
8	UV Rabbit	4697	25	1930	3.78

გამოსვლა ედიტორიდან :

1. თუ თქვენ იყენებთ Windows, დააწკაპუნეთ ედიტორის ფანჯრის დახურვის ღილაკზე –” X”;
2. ედიტორში გაკეთებულ ცვლილებებს Stata არ შეინახავს მანამ, სანამ თქვენ არ შეასრულებთ სათანადო ბრძანებას.

შეინახეთ ცვლილებები შემდეგი გზით: File -ის მენიუდან შეარჩიეთ ბრძანება - Save as. განხილულ შემთხვევაში ჩვენს მიერ შექმნილ ფაილს ვარქმევთ ”afewcars”, ხოლო გაფართოებას- & dta, Stata თავად მიანიჭებს მას;

4. შენიშვნა: თქვენ არ შეგიძლიათ შეინახოთ ცვლილებები მანამ, ვიდრე არ დახურავთ ედიტორს.

შრიფტის შერჩევა:

შრიფტის შესარჩევად ედიტორის მენიუდან შეარჩიეთ ბრძანება – ”Font”, შემდეგ –შრიფტის სასურველი ტიპი და ზომა;

ჩატარებული ცვლილებების შესანახად საჭიროა შემდეგი ბრძანებების თანმიმდევრობით შესრულება: Preferences – Windowing Preferences.

მონაცემების შეტანა ფაილიდან:

ფაილიდან მონაცემების შესატანი სამი სხვა და სხვა ბრძანება გამოიყენება: insheet, infile, infix.

ბრძანება – insheet გამოიყენება იმ შემთხვევაში, თუ ტექსტური ფაილი შექმნილია. ეს ბრძანება გამოსაყენებლად ადვილია. მას შეუძლია წაიკითხოს ფაილიდან ცვლადების სახელები ან შექმნას ცვლადების სახელები თქვენი სურვილისამებრ. ამ ბრძანების გამოყენებისას Stata -ს ბრძანებათა ფანჯარაში ჩაიწერება:

insheet using ფაილის სახელი, რომელსაც იყენებთ

ბრძანება –” infile” გამოიყენება იმ შემთხვევაში, თუ ყველა სტრიქონი წარმოდგენილია ერთი მთლიანი სიტყვით ან ყველა სტრიქონი ჩასმულია ბრჭყალებში; ვთქვათ, გვაქვს მონაცემები სამი რიცხვითი ცვლადისათვის, რომლებიც ტექსტურ ფაილში გამოყოფილია სივრცით. თუ ისინი შეესაბამებიან, მაგალითად, a, b, c –ს, მაშინ მათი წაკითხვისათვის Stata -ს ბრძანებათა ფანჯარაში ჩაიწერება ბრძანება:

infile a b c using ფაილის სახელი, რომელსაც იყენებთ

სხვა დანარჩენ შემთხვევაში შეიძლება გამოვიყენოთ ბრძანება – ”infix” ფიქსირებული ფორმატით.

ახალი ცვლადების შექმნა

მოცემული ცვლადების მიხედვით შეიძლება შევქმნათ ახალი ცვლადი ახალი ფუნქციით. ახალი ცვლადის შესაქმნელად წარმატებით შეგვიძლია გამოვიყენოთ არითმეტიკული, ლოგიკური და სტრიქონის ოპერატორი. ასევე, მათემატიკური ფუნქციები, სტატისტიკური ფუნქციები, მონაცემების ფუნქციები, სპეციალური ფუნქციები.

generate ახალი ცვლადის სახელი = გამოსახულება

განხილული მაგალითის მიხედვით შეიძლება შევქმნათ ახალი ცვლადი მოცემული ცვლადების მიხედვით:

generate image = mpg/price

და მიღებული ცვლადი ან ასეთი სახით დავიტოვოთ ან მისი გამოსახულება ძველ ცვლადს მივანიჭოთ ბრძანებით: ” replace”

replace mpg = mpg/price

მონაცემთა ფაილის აღწერა:

ფაილის აღწერა თავისი შინაარსით დაკავშირებულია შემდეგი პარამეტრების დადგენასთან: დაკვირვებათა რიცხოვნება, ცვლადების რიცხოვნება, მოცულობა (ზომა), რიცხვების ჩაწერის ტიპი (ფორმატი).

მეხსიერებაში არსებული ფაილის აღსაწერად გამოიყენება ბრძანება –” description”, რომელიც Stata-ს ბრძანებათა ფანჯარაში ჩიწერება, მაგრამ თუ ფაილი დისკზეა შენახული, მაშინ Stata –ს ბრძანებათა ფანჯარაში ჩაიწერება ბრძანება:

description using ფაილის სახელი

12.4. მოდელის შეფასება

წრფივი რეგრესია

მოდელის შეფასების პრობლემა მოდელის ჩაწერის შემდეგ მნიშვნელოვანი მომენტია. თუ მოდელში ცვლადის ისეთი ფორმაა, რომელიც საწყის ჩამონათვალში არ გვაქვს, მაშინ ბრძანებების – generate და replace –ის გამოყენებით შევქმნით შესაფასებელი მოდელისათვის მისაღებ ცვლადებს და შემდეგ მივმართავთ ბრძანებას:

regress ცვლადების ჩამონათვალი, რომლებიც მოდელში შედიან

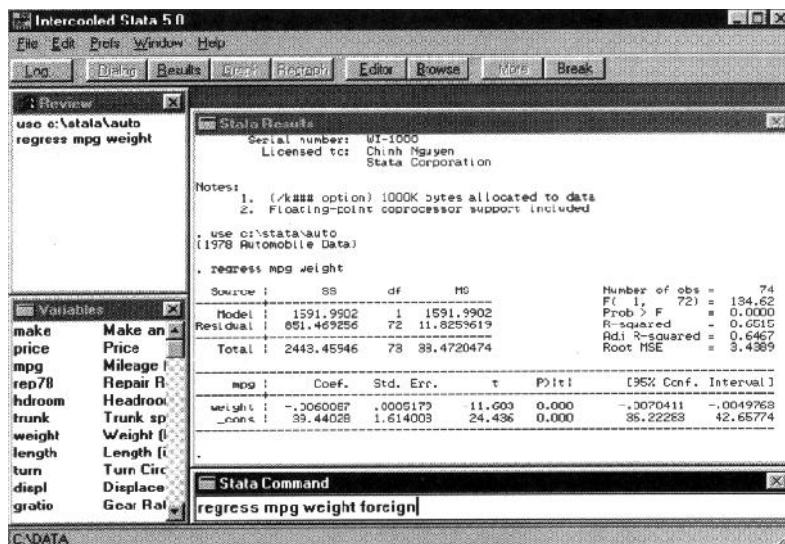
მაგალითად, თუ შესაფასებელია მოდელი:

$$mpg = S_0 + S_1 weight + S_2 foreign + V$$

stata –ს ბრძანებათა ფანჯარაში ჩაიწერება სათანადო ბრძანება:

regress mpg weight foreign

რომლის შესრულების შედეგად stata result -ის ფანჯარაში აისახება შესაბამისი შედეგები:



მაგრამ, თუ შესაფასებელია მოდელი:

$$mpg = S_0 + S_1 weight + S_2 weight^2 + S_3 foreign + V$$

stata –ს ბრძანებათა ფანჯარაში ჩაიწერება სათანადო ბრძანება:

generate wstg + weight^2

რომლის შესრულების შედეგადც მივიღებთ ახალ ცვლადს – wstg = weight^2 და შემდეგ stata -ს ბრძანებათა ფანჯარაში ჩაიწერება მომდევნო ბრძანება:

regress mpg weight wstg foreign

მიღებული შედეგი STATA result -ის ფანჯარაში აისახება ცხრილის სახით, რომელშიც მოცემულია ყველა შესაფასებელი პარამეტრი და მოდელის განტოლება მონაცემების მიხედვით გამოთვლილი პარამეტრების მნიშვნელობებით:

mpg=56.53884-0.0165729 weight+1.59e-06 weight^2-2.2035 foreign

ასევე, STATA result -ის ცხრილში მოცემულია თითოეული გამოთვლილი კოეფიციენტისათვის მისი სტანდარტული კვადრატული გადახრა, 95 პროცენტის ნდობის ინტერვალი, სხვაობების მოდელი თავისუფლების ხარისხის გათვალისწინებით.

STATA –ს გამოყენებით შეგვიძლია სხვა და სხვა ტიპის მოდელის შეფასება როგორცაა: არაწრფივი მოდელები, ლოგისტიკური მოდელები და სხვა. ცალკეული ცვლადისათვის მისი რიცხვითი მონაცემების მიხედვით Stata იძლევა საშუალო მნიშვნელობას, სტანდარტულ გადახრას, მინიმალურ მნიშვნელობას, მაქსიმალურ მნიშვნელობას და დაკვირვებების რიცხოვნებას. მათი მიღება შესაძლებელია, თუ Stata –ს ბრძანებათა ფანჯარაში ჩავწერთ ბრძანებას:

summarize ცვლადის სახელი, რომლისთვისაც გამოთვლებს ვანგარიშობთ

მაგალითად, თუ ბრძანებათა ფანჯარაში ჩაწერილია ბრძანება:

summirize price

მაშინ შედეგი აისახება შემდეგი სახით:

```

_/_/_/_/_/ 5.0 Copyright 1984-1997
  Statistics/Data Analysis      Stata Corporation
                                702 University Drive East
                                College Station, Texas 77840 USA
                                800-STATA-PC http://www.stata.com
                                409-696-4600 stata@stata.com
                                409-696-4601 (fax)
```

Notes:

1. You are running Small Stata.
2. Floating-point coprocessor support included

```
. use "C:\Stata\auto.dta", clear
(1978 Automobile Data)
```

. summarize price

Variable	Obs	Mean	Std. Dev.	Min	Max
price	74	6165.257	2949.496	3291	15906

მაგრამ, თუ ბრძანებათა ფანჯარაში ჩაწერილია ბრძანება:

summarize

მაშინ განხილულ ამოცანაში მონაწილე თითოეული ცვლადისათვის გამოითვლება ზემოთ ჩამოთვლილი ყველა მაჩვენებლის მნიშვნელობა და ეკრანზე აისახება შედეგი ცხრილის სახით:

```
. use "C:\Stata\auto.dta", clear
(Automobile Data)
```

```
. summarize
```

Variable	Obs	Mean	Std. Dev.	Min	Max
make	0				
price	74	6165.257	2949.496	3291	15906
mpg	74	21.2973	5.785503	12	41
rep78	69	3.405797	.9899323	1	5
hdroom	74	2.993243	.8459948	1.5	5
trunk	74	13.75676	4.277404	5	23
weight	74	3019.459	777.1936	1760	4840
length	74	187.9324	22.26634	142	233
turn	74	39.64865	4.399354	31	51
displ	74	197.2973	91.83722	79	425
gratio	74	3.014865	.4562871	2.19	3.89
foreign	74	.2972973	.4601885	0	1

შენიშვნა:

აღნიშნული ბრძანება შეიძლება გამოვიყენოთ ცვლადის არა ყველა მნიშვნელობის გათვალისწინებით, არამედ მისი განსაზღვრული ნაწილისთვისაც, რომელიც გარკვეული პირობით დაფიქსირდება ბრძანებათა ფანჯარაში.

მაგალითად, ბრძანებები:

```
summarize price if mpg<21.3
და
summarize price if mpg=21.3
```

პასუხობენ კითხვებს:

1. როგორია მანქანების საშუალო ფასი, თუ $mpg < 21.3$;
2. როგორია მანქანების საშუალო ფასი, თუ $mpg = 21.3$

ანალიზის შესაბამისი შედეგები მოცემულია Stata -ს შედეგების ცხრილში.

12.5. გრაფიკები

მონაცემების გრაფიკული გამოსახვისათვის STATA -ს ფართო სპექტრი გააჩნია. გრაფიკი შეიძლება x -ღერძის ან y -ღერძის მიმართ ავადგოთ, წარმოვადგინოთ ჰისტოგრამის ან კომბინირებული სახით და სხვა. ნებისმიერ შემთხვევაში STATA -ს ბრძანების სტრიქონში ჩაიწერება ბრძანება – "graph", რომელსაც მოსდევს გრაფიკის ტიპის ამსახველი მენიუ. იგი მოიცავს ორ ოფციას: მონაცემების წარმოდგენას და მოცემების შეერთების ფორმას.

გამოყენებული სიმბოლოებია:

- 0-წერტილები;
- 1-დიდი წრეები;
- 2-კვადრატები;
- 3-სამკუთხედები;
- 4-პატარა წრეები;
- 5-მრავალწახნაგები.
- ”-”გრაფიკის წერტილების შეერთება არ ხდება;
- 1-წერტილები წრფით ერთდება;
- n -გამუალებულ წერტილებს წრფით აერთებს;
- j -წერტილებს გამუალების გარეშე აერთებს.

მონაცემების გრაფიკული სახით წარმოდგენისას ვითვალისწინებთ ორ გარემოებას:

1. ცვლადის წარმოდგენას ინდივიდუალურად (საქმე გვაქვს ჰისტოგრამასთან);
2. ცვლადის წარმოდგენას სხვა ცვლადთან კორელაციაში.

პირველ შემთხვევაში Stata -ს ბრძანებათა ფანჯარაში ჩაიწერება ბრძანება:

graph ცვლადის სახელი

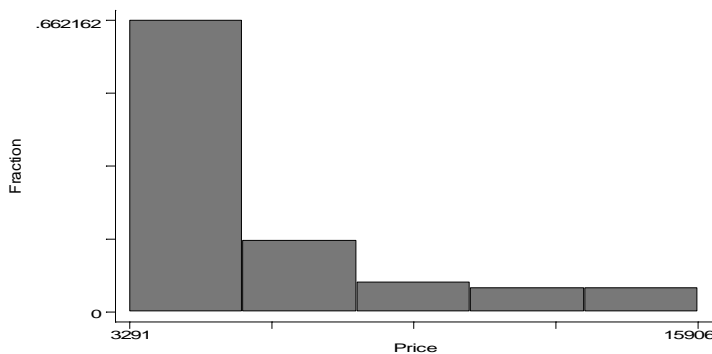
მეორე შემთხვევაში:

graph ცვლადის სახელი ცვლადის სახელი

მაგალითი: Stata –ს ბრძანებათა ფანჯარაში ჩაწერილია ბრძანება:

graph price

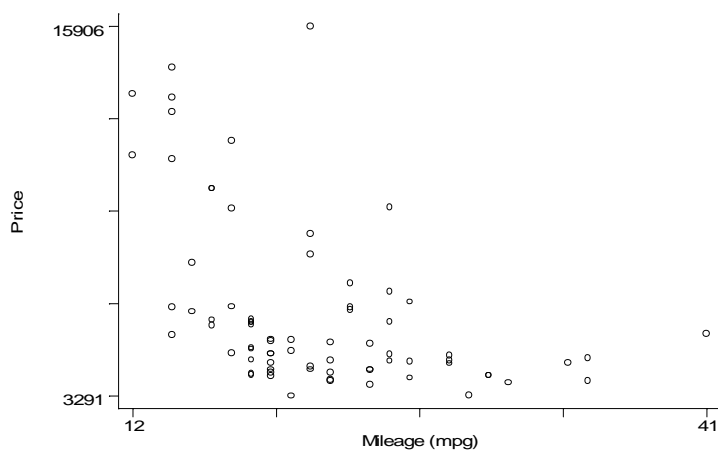
ეკრანზე მივიღებთ ამ ცვლადის ჰისტოგრამას:



თუ Stata -ს ბრძანებათა ფანჯარაში ჩაწერილია ბრძანება:

graph price mpg

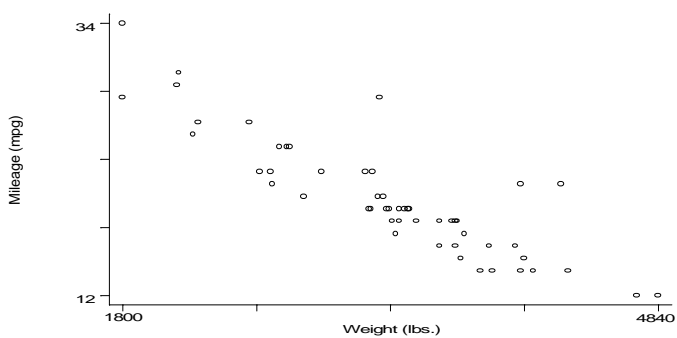
ეკრანზე აისახება ცვლადების – price, mpg მნიშვნელობების შესაბამისი წერტილების განლაგება სიბრტყეზე. ამასთან, Y ღერძზე გადაზომილია ცვლადის price – მნიშვნელობა, ხოლო X ღერძზე –ცვლადის – mpg -ს მნიშვნელობა (შესაძლებელია პირიქითა დამოკიდებულების გამოსახვაც).



მაგრამ თუ stata –ს ბრძანებათა ფანჯარაში ჩაწერილია ბრძანება:

graph mpg weight, by (foreign) total

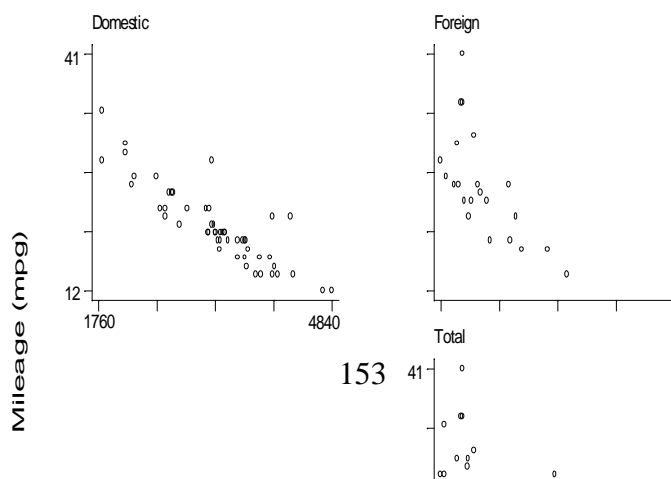
მაშინ გრაფიკს ექნება შემდეგი სახე:



მაგრამ თუ ბრძანებას – graph-ს წინ უძღვის by foreign და ბრძანებას აქვს სახე:

by foreign: graph mpg weight

მივიღებთ ცალ-ცალკე გრაფიკებს ეკრანზე:



გარდა განხილულისა, შეგვიძლია სიბრტყეზე დატანილი წერტილები შევაერთოთ და მივიღოთ სათანადო მრუდი, რისთვისაც ბრძანებაში უნდა მიეთითოს სათანადო პარამეტრები წერტილების შეერთების ტიპისა და წერტილების ფორმის შესახებ.

მაგალითად, ბრძანება:

graph mpg weight if foreign =0, connect (.l) symbol (oi)

12.6. ინფორმაციის გამობეჭდვა

ინფორმაციის გამობეჭდვისთვის საჭიროა შევარჩიოთ ბრძანება Print Log მენიუდან- File.

12.7. დამოუკიდებელი სამუშაო

გვაქვს მონაცემები მოსახლეობის განაწილების შესახებ ასკობრივ ჯგუფებში სიღარიბის ზღვრის მიმართ მდგომარეობის მიხედვით :

გავანალიზოთ ცხრილის მონაცემები და თითოეული ასაკობრივი ჯგუფისათვის გამოვთვალოთ:

1. საშუალო სიდიდე, საშუალო კვადრატული გადახრა, მინიმალური და მაქსიმალური მნიშვნელობები;
2. ავაგოთ გრაფიკები თითოეული ასაკობრივი ჯგუფისათვის.

საქართველოში ღარიბი მოსახლეობის პროცენტული განაწილება მოცემულია ცხრილით.

წელი	6 წლამდე ასაკი	6-15 წლამდე ასაკი	16 წლამდე საპენსიო ასაკი	საპენსიო ასაკი
1997	6.9	19.0	58.1	16.1
1998	6.8	17.5	58.6	17.1
1999	7.4	16.2	50.7	16.7
2000	7.5	19.5	56.5	16.5

1. პროგრამის გაშვება: Start-Programs-Stata-Small stata;

2. დაუწყებულთ ღილაკს – editor. ეკრანზე გამოჩნდება მზა ფურცელი მონაცემების შესატანად;

3. შევიტანოთ მონაცემები კლავიატურიდან მიმდევრობით;
variable by variable პრინციპით;

Stata -ში მონაცემები ჩაიწერება შემდეგი შესაბამისობით:

var1, var2, var3, var4, var5

4. გადავარქვათ ცვლადებს სახელები და მივანიჭოთ მათ სახელები მათი შინაარსის მიხედვით:

var1-წელი – year

var2 –6 წლამდე ასაკის –age6

var3 —6-15 წლამდე ასაკის age 6-15;

var4 –16 წლამდე საპენსიო ასაკის age16

var5 —საპენსიო ასაკი pensage

5. გამოვიდეთ (დავხუროთ) editor და შევინახოთ ჩვენს მიერ შეტანილი მონაცემები სახელის მინიჭებით (ამჯერად, ჩვენი ფაილის მისამართი და სახელი იქნება:

C:\joke.dta)

6. შეგვიძლია ამოცანის ამოხსნას შევუდგეთ:

გამოვთვალოთ თითოეული ცვლადისათვის მოთხოვნილი მაჩვენებლები:

- stata -ს ბრძანებათა ფანჯარაში მიმდევრობით ჩავწეროთ ბრძანებები და შევიტანოთ იგი მებსიერებაში enter -ის ღილაკზე დაჭერით.

summarize age 6

summarize age 15

summarize age 16

summarize age pension

ასეთივე მიმდევრობით შესრულდება ბრძანებები და შედეგი გამოჩნდება ეკრანზე:

800-STATA-PC <http://www.stata.com>
409-696-4600 stata@stata.com
409-696-4601 (fax)

Notes:

1. You are running Small Stata.
2. Floating-point coprocessor support included
. save "C:\joke.dta"
file C:\joke.dta saved

. summarize

Variable	Obs	Mean	Std. Dev.	Min	Max
age6	4	7.15	.3511884	6.8	7.5
age15	11	31.45455	19.56908	16.1	58.6
age16	4	55.975	3.628934	50.7	58.6
pensage	4	16.6	.4163333	16.1	17.1

ასევე, შეგვიძლია გამოვბეჭდოთ შედეგი ცალკეული მაჩვენებლების მიხედვით, რისთვისაც file -ის მენიუდან შევარჩიოთ ოფცია Log print. მიღებული შედეგები გამოიბეჭდება ცხრილის სახით, რომელიც აადვილებს სილატაკის ცვლილებების ანალიზს 1997-2000 წლების მიხედვით.

2. ავაგოთ გრაფიკები ცალკეული ასაკობრივი ჯგუფების მიხედვით, რისთვისაც Stata -ს ბრძანებათა ფანჯარაში მიმდევრობით ჩავწეროთ ბრძანებები და შევიტანოთ იგი მუხსიერებაში Enter -ის ღილაკზე დაჭერით:

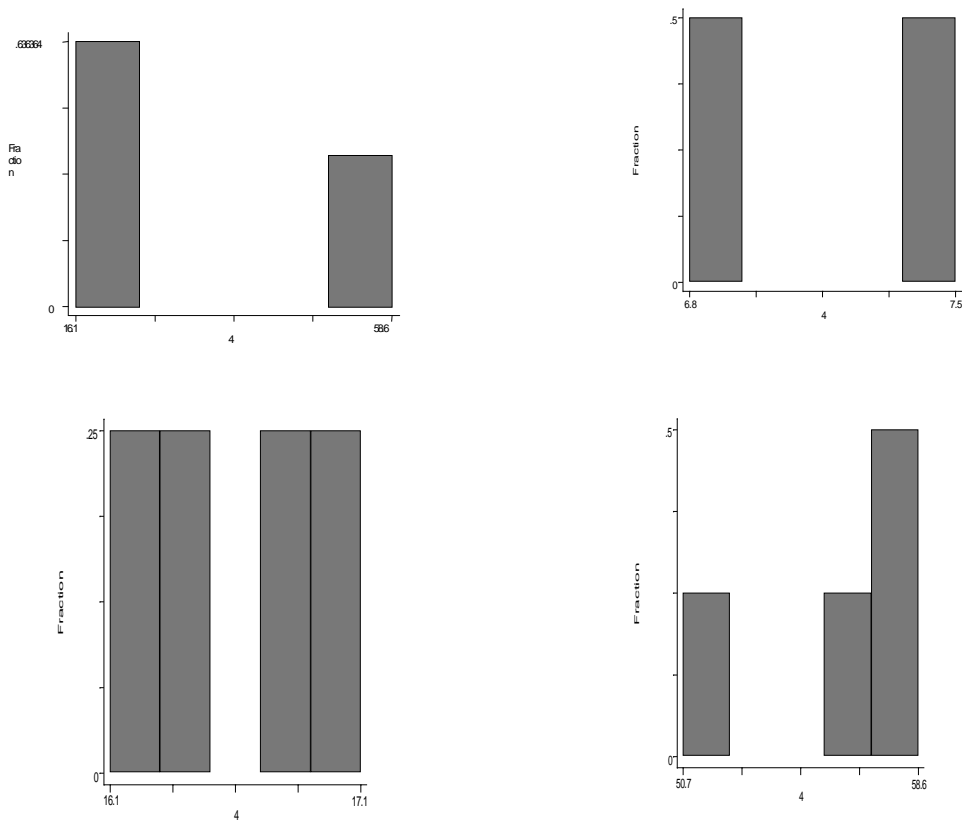
graph age 6

graph age 15

graph age 16

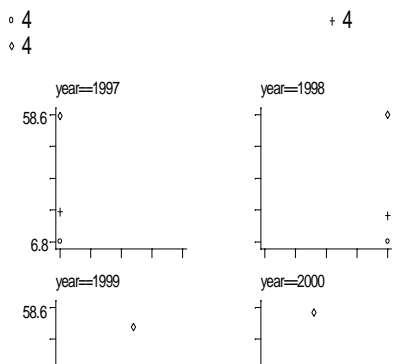
graph pensage

ვემთ განლაგებულია გრაფიკები შესრულებული ბრძანებების შესაბამისად:



3. ავაგოთ გრაფიკები ასაკობრივი ჯგუფების ერთმანეთთან შედარების მიზნით და შევაფასოთ, რისთვისაც Stata -ს ბრძანებათა ფანჯარაში ჩავწეროთ ბრძანება და შევიტანოთ იგი მუხსიერებაში enter -ის ღილაკზე დაჭერით:

graph age 6 age 15 age 16 pensage, by (year) total



ლიტერატურა

- ... , ... 5, 2008
- Tintner G. The Definition of Econometrics. *Econometrica* 21, 1953, p. 31
- Paul A. Samuelson, T.C.Koopmans and J.R. Stone. "Report of the Evaluative Committee for Econometrica", *Econometrica*, 1954,p.141
- 1998

Robert S. Phindych, Daniel L. Rubinfeld. *Econometric Models and Economic Forecasts*. Irvin McGraw Hill, 1998.

William E. Griffiths, R. Carter Hill, George Judge. *Learning and Practicing Econometrics*. John Wiley & Sons, INC, 1997

Adrian O. *The History of Econometrics*. Edward Elgar, 1994

Russell Davidson & James G. MacKinnon. *Estimation and Inference in Econometrics*. Oxford University Press, 1993

John Johnston & John Dinardo. *Econometrics Methods*. McGraw-Hill, 1997

Fumio Hayashi. *Econometrics*. Princeton University Press, 2000

Arthur S. Goldberger. *A Course in Econometrics*. Harvard University Press, 1991

Brenton E. Albert A. *The International Studies in Economics and Econometrics*. Kluwer, 1990

Schmiedt U. Pesaran A. *Handbook of Applied Econometrics*. Blackwell Publishers. 1997

Ernest E. *The Practice of Econometrics*. Addison-Wesley Longman Ed. 1996

Paul A. Ruud. *An Introduction to Classical Econometrics Theory*. Oxford University Press, 2000

Ron C. Mittelhammer, George G. Judge & Douglas J. Miller. *Econometric Foundations*. Cambridge University Press, 2000

Ernst R. Berndt. *The Practice of Econometrics: Classic and Contemporary*. Addison-Wesley, 1991

Darnell Adrian O. *Directory of Econometrics and Forecasting*. Edward Elgar. 1996

Stewart D. *Econometrics*. Prentice Hall Europe. 1997

Gujarati K. *Basic Econometrics*. McGraw Hill HPD. 1996

Theil H. *Principals of Econometrics*. Jacaranda Willey, 1992

Green W. *Econometric Analysis*. McMillan Publishing Company, 1993

White H. Heteroscedasticity Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity. *Econometrica*. V.48/4, pp. 817-838, 1980

Durbin J. Watson G.S. Testing for serial Correlations in Least Squares Regression. *Biometrika*. V.38, pp.159-177

Terence c. Mills. The Econometric Modelling of Financial Time-Series. Cambridge University Press, 1997

Christian Gouriéroux. ARCH Models and Financial Applications. Springer. Verlag. New York, 1997

Holler Manfec J. International Studies in Economics and Econometrics. Kluver, 1987

Gudge G.G. Improved Methods of Inference in Econometrics. North-Holand, 1986

Karlin S. Mathematical Methods and Theory of Games , Programming and Economics. V.1 Addison-Wesley Publishing Co. 1979

Handbook of Econometrics, Vol.3. Amsterdam: North-holand ,1986

Durbin J. Testing for serial correlation in least-squares regression when some of the regression are lagged depended variables ,1970. Econometrica 38(3)

Harvey Andrew C. The Econometric Analysis of Time-series. Deddington, Oxford:Philip Allan, 1981

1997 ,

2001 ,

. , 1980

. , 1981

. ,1973

Kennedy P. A Guide to Econometrics. Mit press, 1992

Karlin S. Mathematical Methods and Theory of Games, Programming and Economics. V.1. Addison-Wesley Publishing Co. 1979

Guidge G.G. Improved Methods of Inference of Econometrics. North-Holand, 1986

Ramu Ramanatan. Introductory Econometrics with Applications. University of Caifornia-San Diego. The Dryden Press. pp.380-384. 1998.

Cagan P.D. The monetary dynamics of Hyperinflation . /Friedman Milton . Studies in the Quantity Theory of money. University of Chicago press .pp 245-256./,1956

Cohrane D., Orcutt G. Application of least squares regression to relationships containing autocorrelated error terms . Journal American Statistical Association 44,pp 32-56

Gleiser H. A New Test for Heteroscedasticity. Journal of the American Statistical Association 64, pp 312-320

Kendrick John W., Grossman Elliot S. Productivity in the United States . trends and Cycles. Baltimore, John Hopkins. 1980

Moser Claus. Calton Graham . Methods in Social Investigation .London. Heinemann . 1970

... , 1997

... , 1979

... , 1974

... , 1976

... , 1975

... 1-2, ...
, 1986

...
, 1986

1984

... , 1968

Steven N. Durlauf University of Wisconsin, Madison; Bruce E. Hansen. Econometric Theory and Practice. University of Wisconsin, Madison. ISBN-13: 9780521807234 | ISBN-0: 0521807239,2008

Wiley and Sons (WIE), 2003. ISBN10: 0471390720

Dimitrios Asteriou, Stephen Hall. Applied Econometrics: A Modern Approach Using Eviews and Microfit. Publisher: Palgrave Macmillan, 2007. ISBN10: 0230506402

Christopher Dougherty . Introduction to Econometrics. Publisher: Oxford University Press, 2006. ISBN10: 0199280967

Philip Hans Franses, Richard Paap . Quantitative Models in Marketing Research. Publisher: Cambridge University Press, 2001. ISBN10: 0521801664

Christian Gourieroux, Joann Jasiak . Financial Econometrics: Problems, Models and Methods (Princeton Series in Finance). Publisher: Princeton University Press

O. Ashenfelter .Statistics and Econometrics: Methods and Applications +D3. Publisher: John, 2001. ISBN10: 0691088721

Christian Gourieroux, Joann Jasiak .The Econometrics of Individual Risk: Credit, Insurance, and Marketing. Publisher: Princeton University Press, 2007. ISBN10: 0691120668

Johnnie E. V. Johnson, Alistair Bruce. Decisions, Risk and Reward. Publisher: Routledge, 2007. ISBN10: 0415426286

Jeffrey Wooldridge . Econometric Analysis of Cross Section and Panel Data. Publisher: MIT Press, 2002. ISBN10: 0262232197

Philip Hans Franses . A Concise Introduction to Econometrics: An Intuitive Guide. Publisher: Cambridge University Press, 2002. ISBN10: 0521520908

Anthony Garratt, Kevin Lee, M.Hashem Pesaran, Yongcheol Shin . Global and National Macroeconometric Modelling: A Long-Run Structural Approach. Publisher: Oxford University Press, 2006. ISBN10: 0199296855

Clive W. J. Granger. Essays in Econometrics, Vol.1. Publisher: Cambridge University Press, 2001. ISBN10: 0521774969

Clive W. J. Granger . Essays in Econometrics, Vol.2. Publisher: Cambridge University Press, 2001. ISBN10: 0521796490

William H. Greene . Econometric Analysis. Publisher: Pearson Education, 2008. ISBN10: 0135137403

Damodar Gujarati . Essentials of Econometrics. Publisher: McGraw Hill Higher Education, 2005. ISBN10: 0071244484

Dennis Halcoussis . Understanding Econometrics. Publisher: South-Western, Div of Thomson Learning, 2004. ISBN10: 0030348064

Christiaan Heij, Paul de Boer, Philip Hans Franses, Teun Kloek, Herman K. van Dijk . Econometric Methods with Applications in Business and Economics. Publisher: Oxford University Press, 2004. ISBN10: 0199268010

Jack Johnston . Econometric Methods –ISE. Publisher: McGraw-Hill UK, 1997. ISBN10: 007115342X

Gary Koop . Bayesian Econometrics. Publisher: Wiley, 2003. ISBN10: 0470845678

Ron C. Mittelhammer, George G. Judge, Douglas J. Miller . Econometric Foundations Pack with CD-ROM. Publisher: Cambridge University Press, 2000. ISBN10: 0521623944

Christian Gourieroux; Alain Monfort . Statistics and Econometric Models: General Concepts, Estimation, Prediction and Algorithms v. 1. Publisher: Cambridge University Press, 1995. ISBN10: 0521405513

Cambridge University Press, 2007. ISBN10: 0521855713

Kerry Patterson. An Introduction to Applied Econometrics. Publisher: Palgrave Macmillan, 2000. ISBN10: 0333802462

Robert Pindyck. Econometric Models & Economic Forecasts w/disk – ISE. Publisher: McGraw-Hill UK, 2000. ISBN10: 0071188312

Paul A. Ruud . An Introduction to Classical Econometric Theory. Publisher: Oxford University Press, 2000. ISBN10: 0195111648

Hamid R. Seddighi, Kevin A. Lawler, Anastasios V. Katos. Econometrics. Publisher: Routledge, 2000. ISBN10: 0415156459

James H. Stock, Mark W. Watson. Introduction to Econometrics. Publisher: Pearson Education, 2006. ISBN10: 0321442539

Leighton Thomas . Modern Econometrics. Publisher: Pearson Education, 1997. ISBN10: 0201876949

Jeffrey Wooldridge . Introductory Econometrics: A Modern Approach. Publisher: Thomson Learning, 2005. ISBN10: 0324323484

Gary Koop, Dale J. Poirier, Justin L. Tobias. Bayesian Econometric Methods. Publisher:

Badi H. Baltagi . Econometric Analysis of Panel Data. Publisher: John Wiley and Sons Ltd, 2005. ISBN10: 0470014563

Alok Bhargava. Econometrics, Statistics and Computational Approaches in Food and Health Sciences. Publisher: World Scientific Publishing, 2007. ISBN10: 9812568417

Michael Clements . Forecasting Economic Time Series. Publisher: Cambridge University Press, 1998. ISBN10: 0521634806

Michael Evans . Practical Business Forecasting. Publisher: Blackwell Publishing, 2002. ISBN10: 0631220666

L. G. Godfrey . Misspecification Tests in Econometrics. Publisher: Cambridge University Press, 1991. ISBN10: 0521424593

Christian Gourieroux, Joann Jasiak . Financial Econometrics: Problems, Models and Methods (Princeton Series in Finance). Publisher: Princeton University Press, 2001. ISBN10: 0691088721

Cambridge University Press, 2002. ISBN10: 0521520916

Christian Gourieroux, Joann Jasiak . The Econometrics of Individual Risk: Credit, Insurance,

Philip Hans Franses . Time Series Models for Business and Economic Forecasting. Publisher: and Marketing. Publisher: Princeton University Press, 2007. ISBN10: 0691120668

Richard Harris, Robert Sollis . Applied Time Series Modelling and Forecasting. Publisher: Wiley, 2003. ISBN10: 0470844434

R. Carter Hill, William E. Griffiths, George G. Judge . Using Excel For Undergraduate Econometrics. Publisher: Wiley, 2000. ISBN10: 0471412376

Terence C. Mills, Raphael N. Markellos . The Econometric Modelling of Financial Time Series. Publisher: Cambridge University Press, 2008. ISBN10: 0521883814

Chandan Mukherjee, Joseph Chai, Sukhan Jackson, Ho . Econometrics and Data Analysis for Developing Countries. Publisher: Routledge, 1997. ISBN10: 0415094003

Peijie Wang . Financial Econometrics. Publisher: Routledge, 2002. ISBN10: 0415224551

ედიბერიძე ა., ნაცვლიშვილი ზ. ალბათობის თეორიისა და მათემატიკური სტატისტიკის ელემენტები. თბილისის სახ. უნივერსიტეტის გამომცემლობა, 1978

მექვაბიძე რ. ეკონომეტრიკა. გამომცემლობა "უნივერსალი", 2004. 330.1 (075.6), 99928-896-2-4

მექვაბიძე რ. ეკონომეტრიკის სალექციო კურსი. გორის სახელმწიფო უნივერსიტეტის გამომცემლობა. ISBN 99926-896-9-1, 2001, პროექტი: HESP 13/012001

მექვამბიძე ფ. ეკონომეტრიკის პროგრამული უზრუნველყოფის გამოყენება პრაქტიკული მეცადინეობისათვის. გორის სახ. უნივერსიტეტის გამომცემლობა, ISBN 99926-896-9-1, 2001, პროექტი: HESP 13/01

ეკონომეტრიკული პაკეტები

STATA . Getting Started with Stata for Windows. Stata Press. College Station. Texas, 1998, 2009

Stata Statistical Software. Statistic Data Management. Graphics. V.1-5.Stata Press. Colege tation. Texas, 1998

EvIEWS. User's Guide. V 1.0. for Windows and Macintosh. QMS, Irvine, California.

Time Series Processor. V.4.3. User's Guide. Browyn H.Holl, 1996

შენიშვნები

შენიშვნები



