საქართველოს ტექნიკური უნივერსიტეტი

თ. შერვაშიძე, ზ. ქვათაძე, ა. კირთაძე, გ. ფანცულაია

მათემატიკური სტატისტიკა ბიზნესსა და ეკონომიკაში

დამტკიცებულია სახელმძღვანელოდ
სტუ-ს სარედაქციო-საგამომცემლო
საბჭოს მიერ. . .2014, ოქმი №

თბილისი
2014

რეცენზენტები:

ISBN
http://www.gtu.ge/publishinghouse/

Verba voland
scripta manent

სარჩევი

წინასიტყვაობა

წიგნი განკუთვნილია უმაღლესი სასწავლებლების ბაკალავრიატისა და მაგისტრატურის სწავლების სტუდენტებისათვის. მოყვანილი მასალის მარტივი ენით გადმოცემა საშუალებას იძლევა ნებისმიერი დაინტერესებული მკითხველი გაეცნოს ალბათობის თეორიისა და მათემატიკური სტატისტიკის ძირითად საკითხებს.

წიგნს საფუძვლად დაედო ლექციათა კურსები, რომლებიც ავტორთა მიერ სხვადასხვა დროს იკითხებოდა თავდაცვის ეროვნულ აკადემიაში (2003-2005), სტუ-ს სატრანსპორტო და მექანიკა-მანქანათმშენებლობის ფაკულტეტის მენეჯმენტის სპეციალობაზე (1998-2005), ეკონომიკური ინფორმატიკის სპეციალობაზე (2000-2006), ივ. ჯავახიშვილის სახელობის თბილისის სახელმწიფო უნივერსიტეტის სოხუმის ფილიალის ეკონომიკურ ფაკულტეტზე (2000-2006), ილიას სახელმწიფო უნივერსიტეტში (2010-2011).

წიგნში განხილულია სხვადასხვა ეკონომიკურ ამოცანებში და მარკეტინგული კვლევის მარტივი საკითხების გადაჭრაში ალბათური მოდელებისა და სტატისტიკური შეფასების თეორიის ელემენტების გამოყენების შესაძლებლობები.

ავტორები შეეცადნენ რთული მათემატიკური დამტკიცებების გარეშე მოეყვანათ საჭირო ფაქტები. მათ აქცენტი გააკეთეს იმაზე, რომ წიგნით სარგებლობა შეეძლოთ არა მხოლოდ სტუდენტებს და იმ პირებს, რომელთაც პრაქტიკული ამოცანების გადაჭრისას სჭირდებათ ალბათობის თეორიისა და მათემატიკური სტატისტიკის მეთოდების გამოყენება, არამედ საშუალო სკოლის დამამთავრებელი კლასების მოსწავლეებსაც.

შესავალი

ალბათობის თეორია, როგორც მეცნიერება, აღმოცენდა XVII საუკუნის შუა პერიოდში. თავდაპირველად იგი დაკავშირებული იყო აზარტული თამაშების ანალიზთან (რომელიც იმ დროს ძალიან გავრცელებული იყო მთელ ევროპაში) და სადაზღვევო საქმესთან, რომელმაც ფართო გამოყენება ჰპოვა საზღვაო ვაჭრობის განვითარების საქმეში. ისტორიამ შემოგვინახა ერთ-ერთი აზარტული მოთამაშის დიდგვაროვან შევალიე დე მერეს სახელი. იგი თამაშის დროს აკვირდებოდა სხვადასხვა კანონზომიერებებს და როდესაც მათ ახსნას ვერ ახერხებდა, მიმართავდა თავის მეგობარს, ცნობილ მეცნიერს ბლეზ პასკალს. ცნობილია, მაგალითად, მისი ერთ-ერთი კითხვის შინაარსი: რა უფრო მოსალოდნელია, ის რომ 4 კამათლის ერთდროულად გაგორებისას ერთზე მაინც მოვა ციფრი 6, თუ ის, რომ ორი კამათლის ერთდროულად 24-ჯერ გაგორებისას ერთხელ მაინც მოვა ექვსიანების წყვილი.

თამაშებთან დაკავშირებულმა საკითხებმა, რომლებიც ვერ იხსნებოდა მაშინ არსებული მათემატიკური მოდელების საშუალებით, მეცნიერები მიიყვანა ახალი იდეებისა და ცნებების შემოტანამდე. ეს ცნებები პირველად გვხვდება ცნობილი მათემატიკოსების: პ. ფერმას, ბ. პასკალის, ი. ბერნულის და სხვათა ნაშრომებში.

ალბათობის თეორიის პირველ სახელმძღვანელოს წარმოადგენდა ქ. ჰიუგენსის (1629-1695) ნაშრომი „აზარტულ თამაშებში დაანგარიშების შესახებ“, რომელიც 1657 წელს გამოვიდა. მასში პირველად არის მოცემული მათემატიკური ლოდინის ცნება. შემდგომში ეს ტერმინი ი. ბერნულიმ (1654-1705) შეიტანა თავის ცნობილ წიგნში „ვარაუდთა ხელოვნება“, რომელშიც მან დაამტკიცა ალბათობის თეორიის პირველი ზღვარითი თეორემა.

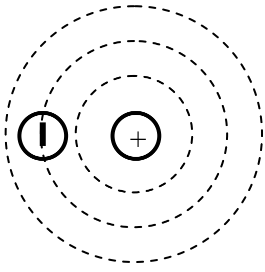
ალბათობის თეორიის განვითარებაში განსაკუთრებული როლი ითამაშეს ლაპლასის, გაუსის და პუასონის, ხოლო შემდგომში რუსი მეცნიერების პ. ჩებიშევის, ა. მარკოვის და ა. ლიაპუნოვის შრომებმა. XX საუკუნის 30-იან წლებში

კი ა. კოლმოგოროვმა მოგვცა ალბათობის თეორიის მკაცრი აქსიომატიკური დაფუძნება.

ალბათობის თეორიის და მათემატიკური სტატისტიკის საგანი

ბუნებაში არსებობს ორი ტიპის მოვლენები: არაშემთხვევითი ანუ დეტერმინისტული და შემთხვევითი ანუ სტოქასტური. მათი ერთმანეთისაგან გასარჩევად განვიხილოთ მაგალითები. დავუშვათ, სხეული მოძრაობს სინათლის სიჩქარეზე ნაკლები მუდმივი V სიჩქარით. ნებისმიერი Δt დროის გასვლის შემდეგ ცალსახად შეგვიძლია განვსაზღვროთ სხეულის მიერ განვლილი მანძილი $S = V \cdot \Delta t$ ფორმულის გამოყენებით. მოვლენებს, რომელთა „ყოფაქცევის“ პროგნოზირება ცალსახადაა შესაძლებელი, ეწოდებათ დეტერმინისტული. მათ შესასწავლად საჭიროა დეტერმინისტული მოდელის შედგენა (მაგალითად, მოყვანილი მოძრაობის მათემატიკური აღწერა მისი დეტერმინისტული მოდელია).

ახლა განვიხილოთ მაგალითი კვანტური მექანიკიდან.



ნილს ბორის პლანეტარული მოდელის ფარგლებში განვიხილოთ ატომგულის ირგვლივ ელექტრონულ შრეებზე მოძრავი ელექტრონი. დავუშვათ, საწყის მომენტში ის მეორე ელექტრონულ შრეზეა. ისმის კითხვა: რა მდებარეობა ექნება მას Δt დროის გასვლის შემდეგ? ამ კითხვაზე პასუხს იძლევა შრედინგერის განტოლება, რომლის

ამონახსნიც გარკვეულ წარმოდგენას გვაძლევს ელექტრონის მოძრაობაზე Δt დროის შემდეგ. თუმცა განსხვავება წინა განხილულ მაგალითთან შედარებით ისაა, რომ შრედინგერის განტოლების ამონახსნი არაცალსახაა. ის შეიძლება იყოს, მაგალითად, ასეთი: 10% შანსით (მოსალოდნელი ალბათობით) ელექტრონი აღმოჩნდება პირველ შრეზე, 40% შანსით მესამეზე, 30% შანსით დარჩება იმავე შრეზე, ხოლო 20% შანსით საერთოდ „გაქრება“. მოვლენებს, რომელთა ევოლუციის პროგნოზირება ცალსახად შეუძლებელია, ეწოდებათ შემთხვევითი ანუ სტოქასტური მოვლენები.

ალბათობის თეორია შეისწავლის შემთხვევითი მოვლენების (ექსპერიმენტების) მათემატიკურ მოდელებს და მათ ფარგლებში აკეთებს დასკვნებს ამ მოვლენებთან დაკავშირებული სხვადასხვა გამონათქვამების შესახებ.

მათემატიკური სტატისტიკა არსებითად იყენებს ალბათობის თეორიის მეთოდებს და ცნებებს, მაგრამ მიუხედავად ამისა იგი გამოყოფილია, როგორც დამოუკიდებელი მათემატიკური დარგი. ეს განპირობებულია იმით, რომ მათემატიკური სტატისტიკის ამოცანები გარკვეული აზრით ალბათობის თეორიის ამოცანების შებრუნებულია. კერძოდ, ალბათობის თეორიაში მოცემულად ითვლება შემთხვევითი მოვლენის მოდელი და მის საფუძველზე კეთდება დასკვნა ამ მოვლენის რეალური მიმდინარეობის შესახებ. მათემატიკურ სტატისტიკაში კი აკვირდებიან რაიმე მოვლენის მოხდენებს (ანუ კონკრეტულ რეალიზაციებს), მათ კანონზომიერებებს და მათგან „ხელოვნურად“ ქმნიან გარკვეულ ალბათურ მოდელს, რომლის ფარგლებშიც აკეთებენ დასკვნებს ამ მოვლენის ირგვლივ სხვადასხვა გამონათქვამებისა და პროგნოზების შესახებ.

განვიხილოთ მარტივი მაგალითი ალბათობის თეორიიდან. დავუშვათ, ვატარებთ შემდეგ ცდას: ვაგორებთ ერთ ცალ ჩვეულებრივ კამათელს. მაშინ ჩვენთვის ალბათური მოდელი უკვე ცნობილია. კერძოდ, სანამ გავაგორებდეთ კამათელს ვიცით, რომ მოსალოდნელი შედეგი შეიძლება იყოს ნებისმიერი ნატურალური რიცხვი 1-დან 6-ის

ჩათვლით და თანაც ტოლშესაძლებლად. თუ მოსალოდნელი შედეგების სიმრავლეს აღვნიშნავთ Ω სიმბოლოთი გვექნება $\Omega = \{1,2,3,4,5,6\}$. ახლა განვიხილოთ გამონათქვამი ამ ცდის შესახებ (ანუ ამ მოვლენის შესახებ): „მოვიდა 4-ზე მეტი რიცხვი“. ეს გამონათქვამი აღვნიშნოთ A ასოთი. ცხადია, A განხორციელდება იმ დროს, როცა მოვა რიცხვი 5 ან 6, ამიტომ A გავაიგივოთ სიმრავლესთან $\{5;\ 6\}$.

ამით ალბათური მოდელი უკვე ჩამოყალიბებულია და ჩვენ მარტივად შეგვიძლია დავასკვნათ, რომ A-ს მოხდენის ალბათობა, ანუ შანსი არის

$$P(A) = \frac{2}{6} = \frac{1}{3}.$$

ამით მოდელის ფარგლებში ცდის (კამათლის გაგორების) შესახებ A გამონათქვამის (4-ზე მეტი რიცხვის მოსვლის) მოხდენის შანსი ანუ „ალბათობა“ გამოთვლილია.

ახლა განვიხილოთ ასევე მარტივი მაგალითი სტატისტიკიდან და დავინახავთ, რომ გარკვეულწილად „შებრუნებულ“ ამოცანას ვიხილავთ. დავუშვათ, ვაკვირდებით რაიმე B მოვლენას და გვაქვს დაკვირვებათა ჩანაწერი (სხვადასხვა დროის მომენტებში)

$$0,\ 0,\ 0,\ 1,\ 0,\ 1,\ 1,\ 0,\ 0,\ 0,\ 1,\ 0,\ 1,\ 1,\ 0,\ 1,$$

სადაც 1 აღნიშნავს იმ დაკვირვებებს, როდესაც B მოხდა, ხოლო 0 – როდესაც B არ მოხდა. გვაინტერესებს შემდგომში B-ს მოხდენის შანსი. ვინაიდან 16 დაკვირვებაში B მოხდა 7-ჯერ, მარტივად შეგვიძლია დავასკვნათ, რომ მისი მოხდენის „შანსი“, ანუ ალბათობა, ფასდება B-ს ფარდობითი სიხშირით ამ ცდებში, რომელიც ტოლია $\frac{7}{16}$-ის. რადგან $\frac{7}{16} < \frac{1}{2}$, ამიტომ უფრო მოსალოდნელია ვიფიქროთ, რომ B არ მოხდება, ვიდრე ის, რომ მოხდება. ე.ი. ჩვენ ჩავატარეთ დაკვირვებები. დაკვირვებების საშუალებით „შევაფასეთ“ B-ს მოხდენის ალბათობა და ამით შევქმენით მოდელი. შემდეგ ამ მოდელის ფარგლებში გავაკეთეთ დასკვნა B-ს მომავალი მოხდენის შესახებ.

თავის მხრივ, მათემატიკური სტატისტიკის წინაშე დასმული ამოცანა შეიძლება ორ ეტაპად დაიყოს. პირველი ეტაპი გულისხმობს შესასწავლი მოვლენის შესახებ მონაცემების (დაკვირვების შედეგების) მოპოვებას, მათ სისტემატიზაციას, კლასიფიკაციას, შესაბამის ცხრილურ და გრაფიკულ წარმოდგენებს და მათთვის სხვადასხვა რიცხვითი მახასიათებლების დადგენას. სტატისტიკის ამ ნაწილს „დესკრიფციულ“, ანუ აღწერით სტატისტიკას უწოდებენ. მეორე ეტაპზე, ალბათობის თეორიის გამოყენებით ხდება ამ მონაცემების ანალიზი და სხვადასხვა დასკვნების მიღება, გარკვეული მახასიათებლების შეფასება და მათი მომავალი „განვითარების“ შესახებ პროგნოზი. ამასთანავე, ხშირ შემთხვევაში, შესაბამისი სტატისტიკური მასალის რაოდენობრივი შეზღუდულობის გამო, საჭირო ხდება ამ დასკვნების საიმედობისა და სიზუსტის შეფასება.

წარმოდგენილი სახელმძღვანელო აგებულია ამ პრინციპის გათვალისწინებით.

I თავი ეძღვნება აღწერითი სტატისტიკას;

II თავში განხილულია ალბათობის თეორიის ძირითადი ცნებები და თეორემები;

III თავში გადმოცემულია სტატისტიკური დასკვნების თეორიის ნაწილი.

უფრო ღრმად ამ საკითხების გაცნობა დოქტორანტებსა და მეცნიერ მუშაკებს შეუძლიათ [1]-ში, რომელიც ავტორთა აზრით ქართველი მეცნიერების მიერ შექმნილი ერთ-ერთი ყველაზე სრულყოფილი სახელმძღვანელოა ამ მიმართულებით.

საქართველოში ალბათობის თეორიისა და მათემატიკური სტატისტიკის სწავლების ტრადიცია დაკავშირებულია ა. რაზმაძესთან (1889-1929), რომელიც კითხულობდა ლექციებს ახლად დაარსებულ თბილისის უნივერსიტეტში (იხ. [2]). შემდგომში ალბათობის თეორიისა და მათემატიკური სტატისტიკის, როგორც მათემატიკის დარგის, განვითარებას საფუძველი ჩაუყარა გ. მანიამ (1918-1985), რომელმაც შექმნა პირველი ქართული სახელმძღვანელოები ამ დარგში ([3]-[7]). მისი დიდი ზრუნვისა და ძალისხმევის

შედეგად 50-იანი წლების მეორე ნახევრიდან საქართველოში შეიქმნა მეცნიერთა კოლექტივები, რომლებიც იკვლევდნენ ალბათობის თეორიისა და მათემატიკური სტატისტიკის პრობლემებსა და მეცნიერების სხვადასხვა დარგის ამოცანებს ალბათურ-სტატისტიკური მეთოდების გამოყენებით. მათი დიდი ნაწილი თავიანთ მოწაფეებთან ერთად მოღვაწეობს სამეცნიერო თუ პედაგოგიურ ასპარეზზე, სტატისტიკურ პრაქტიკაში უშუალო მონაწილეობის ჩათვლით.

I თავი

აღწერითი სტატისტიკა

§ 1. მახასიათებლები და მათი ტიპები

პრაქტიკულ საქმიანობაში ხშირად საჭირო ხდება ობიექტთა გარკვეული ჯგუფის რაიმე კონკრეტულ მახასიათებელზე დასკვნების გაკეთება (მაგალითად, რომელიმე ნაკრძალში მგლების სიცოცხლის ხანგრძლივობაზე; რომელიმე ქვეყნის ამომრჩეველთა პოლიტიკურ შეხედულებებზე; რომელიმე ქარხნის გამოშვებული ნათურების მუშაობის ხანგრძლივობაზე და ა.შ.). შევნიშნოთ, რომ თითოეულ ობიექტს შეიძლება გააჩნდეს მრავალი სხვადასხვა მახასიათებელი (მაგალითად, მგლის მახასიათებელი შეიძლება იყოს წონაც და სიგრძეც, ამომრჩევლის მახასიათებელი შეიძლება იყოს ოჯახური მდგომარეობაც, განათლებაც და ა.შ.).

მახასიათებელი შეიძლება იყოს **თვისებრივი** ან **რაოდენობრივი**.

მახასიათებელს ეწოდება **თვისებრივი** თუ იგი ობიექტის რაიმე თვისებას ან მდგომარეობას გამოხატავს (მაგალითად, ადამიანის თვალის ფერი, პროფესია, განათლება და ა.შ.). თვისებრივი მახასიათებელი იზომება დონეებით. მის ყოველ დონეს **ატრიბუტს** უწოდებენ. მაგალითად, განათლების ატრიბუტებია: არასრული საშუალო, საშუალო, პროფესიული, უმაღლესი და ა.შ.

მახასიათებელს ეწოდება **რაოდენობრივი,** თუ მისი ცალკეული მნიშვნელობა გაზომვის ან თვლის შედეგად მიიღება და რიცხვით გამოისახება (მაგალითად, ადამიანის სიმაღლე ან წონა გაიზომება; პირუტყვის შობადობის მაჩვენებელი დაითვლება და ა.შ.).

მახასიათებლის კონკრეტულ მნიშვნელობას **მონაცემი** ეწოდება. ცხადია, რაოდენობრივი მახასიათებლის მონაცემები რიცხვებია. ხშირად იხმარება ტერმინები: **რაოდენობრივი მონაცემები** ან **თვისებრივი მონაცემები.**

რაოდენობრივი მახასიათებელი შეიძლება იყოს **დისკრეტული** ან **უწყვეტი. რაოდენობრივ მახასიათებელს ეწოდება დისკრეტული** თუ ის ღებულობს მხოლოდ განცალკევებულ რიცხვით მნიშვნელობებს (მაგალითად, პროფესიონალი მოკრივის მიერ რინგზე მოპოვებული გამარჯვებების რაოდენობა; სამორინეში მოთამაშის წარმატებების რაოდენობა და ა.შ.). რაოდენობრივ მახასიათებელს ეწოდება **უწყვეტი**, თუ ის ღებულობს ნებისმიერ რიცხვით მნიშვნელობას რაიმე შუალედიდან (მაგალითად, 10 სმ რადიუსიან მრგვალ სამიზნეზე ტყვიის სროლისას შედეგს წარმოადგენს სამიზნის ცენტრიდან დაშორება და ის შეიძლება იყოს ნებისმიერი რიცხვი [0, 10] ინტერვალიდან). აქვე საჭიროა შევნიშნოთ, რომ ზოგიერთი მახასიათებელი შეიძლება თავისი ბუნებით იყოს უწყვეტი ტიპის, მაგრამ მისი გაზომვა ხდებოდეს როგორც დისკრეტული სიდიდის. მაგალითისათვის, განვიხილოთ ადამიანის სიმაღლე. ყოველი ადამიანი თავის სიმაღლეს ზომავს სანტიმეტრამდე სიზუსტით, თუმცა იგი რეალურად შეიძლება იყოს ამ მონაცემზე რამდენიმე მილიმეტრით მაღალი ან დაბალი.

§ 2. მონაცემები, მათი გაზომვის დონეები და მიღების მეთოდები

დისკრეტული რიცხვითი მონაცემები შეიძლება გაიზომოს ოთხი სხვადასხვა სკალით: **ნომინალური, რიგობრივი, ინტერვალური და ფარდობითი სკალებით**, ხოლო უწყვეტი მონაცემები კი – **ინტერვალური და ფარდობითი** სკალებით.

დავახასიათოდ თითოეული მათგანი.

1. **ნომინალური** ანუ **სახელდებითი** სკალით იზომება ის მონაცემები, რომლებიც შეესაბამებიან მახასიათებლებს, რომელთაც შეიძლება მიიღონ რამდენიმე მნიშვნელობა. ასეთ დროს ხდება მონაცემების დაყოფა კლასებად და თითოეული ამ მნიშვნელობის შესაბამისი მონაცემი ხვდება ერთ კლასში (მაგალითად, ერთიდაიგივე ტიპის საქონელს თანამედროვე საერთაშორისო სავაჭრო ორგანიზაციების გადაწყვეტილებით მინიჭებული აქვთ ერთნაირი კოდი).

სხვადასხვა კლასის ორი ელემენტი ვერ შედარდება ერთმანეთს. ზოგიერთ შემთხვევაში ჩვენ თვითონაც შეგვიძლია სკალის შემოღება. მაგალითად, თუ გვინდა სკოლის მოსწავლეებზე რაიმე მახასიათებლის, დავუშვათ სწავლის მიმართულების მიხედვით, დაკვირვება შეგვიძლია სპორტული სკოლების მოსწავლეები მოვაქციოთ პირველ კლასში, ჰუმანიტარული მიმართულების სკოლების – მეორეში, ხოლო ტექნიკური მიმართულების – მესამე კლასში. შეიძლება ამ მახასიათებლის უფრო დაწვრილებითი გაზომვაც მოვახდინოთ. მაგალითად, საფეხბურთო სკოლის მოსწავლეებს მივანიჭოთ 1.1 ინდექსი, საკალათბურთო სკოლის 1.2 და ა.შ. ცხადია სხვადასხვა ინდექსის მქონე მოსწავლეები სხვადასხვა მიმართულების სკოლებში სწავლობენ და ისინი ერთმანეთს ვერ შედარდებიან. ნომინალური სკალით ყველაზე დაბალი კატეგორიის მონაცემები იზომება.

2. **რიგობრივი სკალით** იზომება მონაცემები, რომელთა შორისაც მახასიათებლის ხარისხის მიხედვით შეიძლება დაწესდეს გარკვეული რიგი. მაგალითად, IV, III, II და I კატეგორიის ინჟინრები (ამ კატეგორიის ხარისხი მატულობს ნომრის კლების მიხედვით), ან განვიხილოთ სამხედრო წოდებების თანამიმდევრობა; ლეიტენანტი, კაპიტანი, მაიორი, ვიცე-პოლკოვნიკი, პოლკოვნიკი, გენერალი. ასეთ შემთხვევაში ჩვენ შეგვიძლია მონაცემები შევადაროთ ერთმანეთს, მაგრამ რამდენით ან რამდენჯერ მეტია ერთი მეორეზე ვერ დგინდება. მონაცემებს რომლებიც რიგობრივი სკალით იზომებიან **რიგობრივ მონაცემებს** უწოდებენ.

3. **ინტერვალური სკალით** იზომება მონაცემები, რომელთაც გააჩნიათ რიგობრივი მონაცემების თვისებები და შეგვიძლია დავადგინოთ ერთი მეორეზე რამდენით მეტია, მაგრამ რამდენჯერ მეტია (ფარდობით დამოკიდებულებას) ვერ ვარკვევთ. ამ დროს ობიექტს მიეწერება ერთეულთა იმ რაოდენობის ტოლი რიცხვი, რომელიც ამ თვისებების ხარისხის ტოლფასია. მაგალითად, სხეულის ტემპერატურა შეიძლება გაიზომოს ცელსიუსის სკალით. ამ დროს ვიტყვით, რომ 18°C-ზე გამთბარი სხეულის ტემპერატურა 9°C-ზე გამთბარზე 9°C-ით მეტია. უცნობია

რაოდენობრივად ორჯერ განსხვავდება თუ არა გათბობის შედეგად მიმდინარე პროცესები მეორე სხეულში, პირველთან შედარებით. თანაც, ასეთ შემთხვევაში 0°C არ ნიშნავს, რომ სხეულს ტემპერატურა არ გააჩნია. ინტერვალურ სკალაში რიცხვი 0 არ ნიშნავს, რომ დაკვირვებადი მახასიათებელი ობიექტს არ გააჩნია.

მეორე მაგალითის სახით განვიხილოთ ადამიანის ასაკი. მართალია, 17-დან 18 წლამდე მიღწევისათვის პიროვნებისათვის ერთი ასტრონომიული წლის „გავლაა“ საჭირო, ისევე როგორც 22-დან 23 წლამდე მიღწევისათვის, მაგრამ პირველ შემთხვევაში, როდესაც 18 წლის ხდება, პიროვნება იძენს ისეთ თვისებებს, როგორებიცააა: არჩევნებში მონაწილეობის უფლება, სრულწლოვანის სტატუსი და აქედან გამომდინარე მისი როლი საზოგადოებრივ ცხოვრებაში მნიშვნელოვნად იცვლება. 22-დან 23 წლამდე გასული ერთი წელი კი ხასიათდება მხოლოდ ორგანიზმში მიმდინარე ფიზიოლოგიური ცვლილებებით ან პიროვნების მიერ მიღწეული წარმატებებით მოღვაწეობის სხვადასხვა სფეროში. ასე რომ, პიროვნება 17 და 18 წლის ასაკში საზოგადოებაში თავისი მდგომარეობით უფრო მეტად განსხვავდება ერთმანეთისაგან, ვიდრე 22 და 23 წლის ასაკში.

4. **ფარდობითი სკალით** იზომება მონაცემები, რომელთაც აქვთ ინტერვალური მონაცემების ყველა თვისება და, გარდა ამისა, შეგვიძლია შევადაროთ რამდენჯერ მეტად აქვს ეს თვისება ერთ ობიექტს, ვიდრე მეორეს. ასეთ სკალაში 0 არსებობს, როგორც გასაზომი თვისების არ ქონის მაჩვენებელი. მაგალითად, განვიხილოთ გარკვეული ქალაქების სიმაღლეები ზღვის დონიდან. ამ დროს თუ რომელიმე გეოგრაფიული ადგილის სიმაღლეა 0, ეს ნიშნავს, რომ ეს ადგილი არ არის აცილებული ზღვის დონეს, ხოლო თუ *A* ქალაქის სიმაღლე ზღვის დონიდან არის 120, ხოლო *B* ქალაქის – 60, მაშინ ცხადია, რომ *A* ქალაქი ორჯერ მაღლაა ზღვის დონიდან, ვიდრე *B* (შევნიშნოთ, რომ ზღვის დონედ მსოფლიოში მიჩნეულია ბალტიის ზღვის დონე).

მონაცემების მოპოვება შეიძლება შემდეგი მეთოდებით:

1) მონაცემების შეგროვება გამოკითხვით;
2) მონაცემების შეგროვება ექსპერიმენტით;
3) მონაცემები შეგროვება დაკვირვების საშუალებით;
4) ადრე არსებული მონაცემების გამოყენება.

პირველი სამი მეთოდით მიღებულ მონაცემებს **პირველადი** მონაცემები ეწოდებათ, ხოლო მეოთხე მეთოდით მიღებულ მონაცემებს **მეორადი მონაცემები**.

გამოკითხვის ჩატარების დროს უნდა დავიცვათ შემდეგი პირობები:

ა) წინასწარ განვსაზღვროთ გამოკითხვის მიზნები და ამოცანები;

ბ) განვსაზღვროთ იმ ადამიანთა კატეგორია, რომელთა გამოკითხვაც გვაინტერესებს;

გ) შევქმნათ კითხვარი და მოვახდინოთ მისი წინასწარი ტესტირება;

დ) განვსაზღვროთ გამოსაკითხ პირთა რაოდენობა და მათი შერჩევის მეთოდი;

ე) შევარჩიოთ გამოსაკითხი პირები და ჩავატაროთ გამოკითხვა.

კითხვები არსებობს ორგვარი: **დახურული ბოლოთი** (რესპონდენტმა პასუხი უნდა შეარჩიოს რამდენიმე ჩამოთვლილი ვარიანტიდან), **ღია ბოლოთი** (რესპონდენტს პასუხის გაცემისას აქვს სრული თავისუფლება) და **ნახევრად ღია ბოლოთი**.

კითხვებს, რომლებიც შეეხება რესპონდენტის პირად მონაცემებს (ოჯახური მდგომარეობა, შემოსავალი და ა.შ.) **დემოგრაფიული** კითხვები ეწოდებათ.

გამოკითხვა შეიძლება ჩატარდეს წერილობით, ტელეფონით ან პირადი ინტერვიუთი.

პირადი ინტერვიუ შეიძლება იყოს **სტრუქტურიზებული** (როდესაც ყველა კითხვა წინასწარაა ცნობილი) და **არასტრუქტურიზებული** (როდესაც იწყება რამოდენიმე ზოგადი კითხვით და შემდეგი კითხვები დამოკიდებულია რესპონდენტის პასუხებზე).

ექსპერიმენტი არის პროცესი, რომელიც გვაძლევს მონაცემებს და მკვლევარს შეუძლია აკონტროლს დასაკვირვებელი ობიექტები. ექსპერიმენტის განმსაზღვრელია

ექსპერიმენტის გეგმა, რაც არის იმ ქმედებების ერთობლიობა, რომლებიც უნდა განხორციელდეს დასაკვირვებელი ობიექტის საინტერესო პარამეტრის განსაზღვრისათვის.

მაგალითი 2.1. თუ გვსურს ქარხნის მიერ დამზადებული აგურის პარტიის სიმტკიცის განსაზღვრა, ექსპერიმენტით უნდა გავზომოთ ერთ-ერთი აგურის სიმტკიცე. ამისათვის ხდება: ა) წნეხის ქვეშ მისი მოთავსება; ბ) წნევის მატება თანდათანობით; გ) ხელსაწყოების საშუალებით აგურის მდგომარეობის დაფიქსირება; დ) დამსხვრევის შემდეგ ნარჩენის შეფასება.

ექსპერიმენტის ჩატარებისას ხშირად ხდება ერთი ან რამდენიმე პირობის შეცვლა ისე, რომ განისაზღვროს მათი ზეგავლენა გამოსაკვლევ პარამეტრზე.

მაგალითი 2.1.-ის შემთხვევაში აგურის სიმტკიცეზე ტენიანობისა და სითბური ეფექტების ზეგავლენის დასადგენად აგური შეიძლება იყოს დასველებული სხვადასხვა ხარისხით ან გახურებული სხვადასხვა ტემპერატურამდე.

დაკვირვებით მონაცემების მიღება ხდება მაშინ, როცა გვაქვს დრო და შესაძლებლობა ობიექტზე დაკვირვებების ჩასატარებლად და მონაცემები გროვდება დაკვირვების პროცესში.

მაგალითი 2.2. ექსპერიმენტში საჭიროა განისაზღვროს ნალექიანობის დონე გაზაფხულზე საქართველოს რომელიმე რაიონში. სხვადასხვა დღეების განმავლობაში ვზომავთ ნალექიანობის დონეს ამ რაიონში და ვიღებთ მონაცემებს. მათი გარკვეული რაოდენობის შეგროვების შემდეგ ვაკეთებთ დასკვნებს ამ მონაცემებზე დაყრდნობით.

ამავე მაგალითში განსახილველი საკითხის გადაწყვეტისათვის შეგვიძლია გამოვიყენოთ ჰიდრომეტეოსამსახურის მიერ **ადრეულ წლებში მიღებული მონაცემები** ნალექიანობის შესახებ. როგორც აღვნიშნეთ ასე მიღებულ მონაცემებს მეორადი მონაცემები ეწოდებათ.

§ 3. პოპულაცია, შერჩევა, შერჩევის მეთოდები

ობიექტთა ერთობლიობის რაიმე მახასიათებლის გამოკვლევისას ხშირად ვერ ხერხდება ყოველი ელემენტისათვის მახასიათებლის მონაცემის მიღება შემდეგი მიზეზების გამო:

1. ობიექტთა ერთობლიობა დიდია და ყოველ მათგანზე მონაცემის მოპოვება მოითხოვს დიდ დროს, რის გამოც გამოკვლევის შედეგებმა შეიძლება აქტუალობა დაკარგოს. მაგალითად, თუ ვიკვლევთ ამომრჩეველთა პოლიტიკურ შეხედულებებს, კვლევა უნდა მოხდეს არჩევნების წინა შეზღუდული დროის შუალედში;
2. მახასიათებლის მონაცემის მიღების შემდეგ გამოსაკვლევი ობიექტი შეიძლება გამოუსადეგარი გახდეს და მთელ ერთობლიობას თუ გავანადგურებთ დასკვნა გაუმართლებელი იქნება ეკონომიკური კუთხით (იხ. მაგალითი 2.1);
3. მახასიათებელი გარკვეული აზრით წარმოსახვითია და გამოკვლევის პერიოდში შეიძლება რეალურად ვერც შევამოწმოთ. მაგალითად, რომელიმე ნაკრძალში მგლების სიცოცხლის ხანგრძლივობის დასადგენად ვერ დაველოდებით როდის მოკვდება მგელი. ასეთ დროს, უბრალოდ, თუ იპოვნიან მკვდარ ცხოველს, ექსპერტები ადგენენ ასაკს. ცოცხალი გამოსაკვლევი ობიექტის ასაკს კი ადგენენ დაკვირვებით და, მისი მდგომარეობის მიხედვით, დარჩენილი სიცოცხლის დროსაც მიახლოებით წინასწარმეტყველებენ. ასე იღებენ სიცოცხლის ხანგრძლივობის მიახლოებით მნიშვნელობას. ცხადია, ამას ყოველ მგელზე ვერ ჩავატარებთ.

ასეთ შემთხვევებში ხდება განსახილველი ობიექტების ერთობლიობიდან რაღაც ნაწილის შერჩევა. ამ ნაწილისათვის შეისწავლება ჩვენთვის საინტერესო მახასიათებელი. ამის შემდეგ, მიღებული დასკვნა „ვრცელდება“ ობიექტთა მთელ ერთობლიობაზე.

განსაზღვრება 3.1. ობიექტთა ერთობლიობის რაიმე კონკრეტული მახასიათებლის ყველა შესაძლო მნიშვნე-

ლობათა სიმრავლეს **პოპულაცია** ეწოდება. თვითონ ამ ობიექტებს კი პოპულაციის **ელემენტარული ერთეულები,** ან **პოპულაციის ერთეულები**, ეწოდებათ.

პოპულაციის ერთეულებმა შეიძლება ბევრი სხვადასხვა პოპულაცია მოგვცეს. მაგალითად, თუ ვიხილავთ *NBA*-ს (ნაციონალური საკალათბურთო ასოციაცია) კალათბურთელთა სიმაღლეების პოპულაციას, ელემენტარული ერთეულები იქნებიან *NBA*-ს კალათბურთელები და მათ შეიძლება მოგვცენ წონების პოპულაცია, ერთ მატჩში საშუალო შედეგიანობების პოპულაცია, პირადი დანაზოგების პოპულაცია და ა.შ.

განსაზღვრება 3.2. მეთოდს, რომლის მიხედვითაც ხდება პოპულაციიდან ნაწილის გამოყოფა (ამოკრეფა, შერჩევა) **სტატისტიკური შერჩევის მეთოდი**, ან **შერჩევის მეთოდი,** ეწოდება. იმ ნაწილს, რომელიც პოპულაციიდან ამოვკრიფეთ ან შევარჩიეთ **შენარჩევი** ანუ **შერჩევა** ეწოდება, ხოლო შერჩევაში ელემენტთა რაოდენობას კი – **შერჩევის მოცულობა**.

ხშირ შემთხვევაში დასაკვირვებელი ობიექტების (ანუ პოპულაციის ელემენტარული ერთეულების) ერთობლიობას **გენერალურ ერთობლიობას** უწოდებენ, ხოლო პოპულაციის მაგივრად იხილავენ დასაკვირვებელი მახასიათებლის მნიშვნელობათა სიმრავლეს.

შემდეგში შერჩევას ორი აზრით გამოვიყენებთ. ერთ შემთხვევაში მასში შერჩევით მეთოდს ვიგულისხმებთ, მეორეში კი – შენარჩევს. ცხადია, რომ ტექსტის მიხედვით გარკვეული იქნება რომელ შემთხვევაში რას ვგულისხმობთ.

როდესაც შერჩევის პროცესი არასწორადაა წარმართული, მიღებული მონაცემები წარმოადგენს რეალური სურათის დამახინჯებას, რაც საბოლოოდ (სტატისტიკური დამუშავების შემდეგ) ხდება არასწორი დასკვნების მიღების მიზეზი. ასეთ დროს ამბობენ, რომ ადგილი აქვს შერჩევის **ჩანაცვლებას.**

დავუბრუნდეთ მაგალითი 2.2-ს. თუ ნალექების ზომაზე დაკვირვებებს ჩავატარებთ მხოლოდ მაისში, მივიღებთ ჩანაცვლებულ შერჩევას. გაზაფხულის პერიოდში ნალექიანობის საშუალო დონის განსაზღვრისას გათვა-

ლისწინებული არ იქნება მარტის და აპრილის მონაცემები, რომლებიც მაისის მონაცემებისაგან განსხვავებულია.

შერჩევის ჩანაცვლების გამომწვევი მიზეზებიდან მოვიყვანოთ რამდენიმე:

1) დაკვირვებულ მნიშვნელობათა შერჩევითი მეთოდით გამოწვეული ჩანაცვლება;
2) შერჩევითი მეთოდის არასწორად გამოყენებით გამოწვეული ჩანაცვლება (იხ. მაგალითი 2.2);
3) გასაზომი ხელსაწყოების უზუსტობა (რომელიც გამოსწორდება, თუ ხელსაწყოების საერთაშორისო დონეზე სტანდარტიზირებას მოვახდენთ და გაზომვათა ერთგვაროვნების პირობებს დავიცავთ);
4) გამოკითხვის დროს უპასუხოდ დარჩენილი კითხვების დიდი რაოდენობა;
5) სუბიექტური აღქმით გამოწვეული ჩანაცვლება.

შერჩევის მიღების ხერხები დავყოთ ორ ჯგუფად:

ა) **არაშემთხვევითი ხასიათის**;

ბ) **შემთხვევითი ხასიათის.**

არაშემთხვევითი ხასიათისაა:

1) **უწესრიგო შერჩევა –** რომელიც წარმოებს უწესრიგოდ. მაგალითად, თუ ქათმების ფერმიდან გამოსაკვლევად საჭიროა 20 ქათამი, ვიჭერთ პირველივე 20 ცალ ქათამს.

2) **მიზანმიმართული შერჩევა** – როდესაც ირჩევენ არაერთგვაროვანი პოპულაციის ისეთ ერთეულებს, რომლებიც იძლევიან საშუალო წარმოდგენას პოპულაციის შესახებ. მაგალითად, თუ შესამოწმებელია ძროხების ფერმის წველადობის პოპულაცია, შერჩევაში იღებენ საშუალო სიდიდის ძროხებს.

3) **ხელმისაწვდომი შერჩევა –** რომელიც სწრაფად ან ადვილად მოსაპოვებელია. მაგალითად, თუ გვინდა შევამოწმოთ მატარებლის ეშელონით შემოსული ხორბლის მარცვლის ზომების პოპულაცია ვიღებთ ვაგონიდან 20-30 სმ სიმაღლის ფენიდან შერჩევას,

შემთხვევითი ხასიათის შერჩევა ექვემდებარება ალბათობის თეორიის მეთოდებით შესწავლას.

განვიხილოთ მისი სახეები.

1. **მარტივი შემთხვევითი შერჩევა**. მარტივი შემთხვევითი შერჩევა ისეთი შერჩევაა, როდესაც პოპულაციის ყოველ ელემენტს შერჩევაში მოხვედრის ერთნაირი შანსი აქვს.

მაგალითად, $N = 400$ რაოდენობის მოკრივის სიმაღლეების პოპულაციიდან $n = 10$ მოცულობის შერჩევის მისაღებად მოვიქცეთ შემდეგნაირად: მუყაოს ფურცლებზე დავწეროთ თითოეულის, მაგალითად, გვარი. ავურიოთ ეს დასტა და შემდეგ იქედან შემთხვევითად ამოვიღოთ 10 ცალი. ასეთ შემთხვევით შერჩევას ლატარიის პრინციპით შერჩევას უწოდებენ.

არსებობს შემთხვევითი რიცხვების საშუალებით შერჩევა. ასეთ დროს ვსარგებლობთ შემთხვევითი რიცხვების ცხრილით (იხ. დანართი, ცხრ. 1), რომელიც შედგება 30 სტრიქონისა და 50 სვეტისაგან. ამ ცხრილში ვიღებთ ნებისმიერ „წერტილს“. მაგალითად, მეოთხე სტრიქონში და მეხუთე სვეტში 9-ს. ამის შემდეგ, მივყვებით რომელიმე მიმართულებით, ვთქვათ, ზემოდან ქვემოთ. გადავთვლით მიყოლებით 10 ცალ ისეთ ციფრთა სამეულს, რომელთაგან თითოეული ადგენს 400-ზე ნაკლებ რიცხვს. მივიღებთ:

093, 254, 140, 306, 097, 173, 139, 221, 217, 248.

იგივე მაგალითში, ანბანის მიხედვით დალაგებული გვარების მიმდევრობიდან ამ შემთხვევითი რიცხვების შესაბამისი ნომრების გვარებს ვიღებთ.

შემთხვევითი რიცხვების არჩევა კომპიუტერული პროგ-რამებითაც არის შესაძლებელი.

2. **სისტემატური შემთხვევითი შერჩევა.** სისტემატური შემთხვევითი შერჩევა გამოიყენება, როდესაც პოპულაციის ერთეულები გადანომრილია. ამ დროს წინასწარ დასახელე-ბული k ბიჯისათვის ვირჩევთ რაიმე m ნომერს $(m \le k)$. შემდეგ ამ ნომრიდან დაწყებული k ბიჯის დამატებით, შერჩევის მოცულობის გათვალისწინებით, ვიღებთ ყველა შესაბამისი ნომრის ერთეულებს.

მაგალითად, გვაქვს $N = 400$ მოცულობის პოპულაციის ერთეულების რაოდენობა და გვინდა გავაკეთოთ $n = 12$ მოცულობის შერჩევა. წინასწარ განსაზღვროთ k ბიჯი, რომელიც არის N/k რიცხის მთელი ნაწილი. ე. ი.

$k = \left[\frac{400}{12} \right] = 33$. შემდეგ შემთხვევით ავიღოთ რაიმე რიცხვი, რომელიც არ აღემატება 33-ს. მაგალითად, $m = 18$ და ამოვარჩიოთ გადანომრილი ერთეულებიდან შემდეგი 12 ცალი ნომრის მქონე ერთეული $(m + kt, \quad t = 0,1,\ldots,11)$:

18, 51, 84, 117, 150, 183, 216, 249, 282, 315, 348, 381.

3. განშრევებული ანუ სტრატიფიცირებული შემთხვევითი შერჩევა. განშრევებული ანუ სტრატიფიცირებული შემთხვევითი შერჩევა წარმოებს მაშინ, როცა პოპულაცია არაერთგვაროვანია და შედგება რამდენიმე ჯგუფისაგან. ამ ჯგუფებს **შრეებს (სტატებს)** უწოდებენ. სასურველია ისე მოხერხდეს, რომ ამ შრეების პროპორციები შერჩევაში თითქმის ისეთივე იყოს, როგორიც პოპულაციაში. თითოეული შრის შერჩევის ელემენტების მისაღებად გამოიყენება მარტივი შემთხვევითი შერჩევა.

ადამიანების პოპულაციის კვლევისას განშრევება შეიძლება მოხდეს პროფესიით, ასაკით, სქესით და ა.შ. მაგალითად, თუ სოციოლოგიური კვლევა წარმოებს ამომრჩეველთა პოლიტიკურ შეხედულებებზე გამოკითხვა რეალურ შედეგს ვერ ასახავს, თუ ჩატარდება მხოლოდ განათლების სისტემის მუშაკებში, ან მხოლოდ სამშენებლო კომპანიების მუშაკებში, ან მხოლოდ სოფლის მეურნეობის მუშაკებში და ა.შ. ამ დროს სასურველია შერჩევის ელემენტები ყველა პროფესიული ფენიდან ავარჩიოთ და თითოეული ფენიდან ასაკისა და საცხოვრებელი ადგილების მიხედვით მარტივი შემთხვევითი შერჩევა მოვაწყოთ. თანაც საჭიროა შეძლებისდაგვარად დავიცვათ თანაფარდობა სხვადასხვა ფენის მოსახლეობისაგან მიღებული შერჩევის მოცულობებისათვის. მაგალითად, თუ სამედიცინო სფეროს მუშაკთა რაოდენობა ორჯერ მეტია პედაგოგთა რაოდენობაზე, მიღებულ შერჩევაში ელემენტარული ერთეულები მედიცინის სფეროს მუშაკებიდან დაახლოებით ორჯერ მეტი უნდა იყოს, ვიდრე პედაგოგებიდან.

ხშირად განშრევებულ შერჩევას **რაიონირებულ შერჩევასაც** უწოდებენ (ტერმინი მოდის გეოგრაფიული რაიონის

შინაარსიდან). მაგალითად, ვთქვათ, პოპულაცია შეადგენს საქართველოს მოსახლეობის მატერიალურ დანაზოგებს ბანკებში. შერჩევისათვის საქართველის ტერიტორია დავყოთ რაიონების მიხედვით და თითოეული რაიონიდან მარტივი შემთხვევითი შერჩევით მივიღოთ პოპულაციის ერთეულები. საბოლოოდ მიღებულ შერჩევაში დაახლოებით დავიცვათ რიცხობრივი თანაფარდობა რაიონების მოსახლეობის რაოდენობებს შორის.

4. **კლასტერული შერჩევა**. კლასტერული შერჩევის გამოყენების დროს ხდება პოპულაციის ერთეულების დაყოფა თანაუკვეთ ჯგუფებად (**კლასტერებად**). შემდეგ ზემოთ ხსენებული მეთოდებით თითოეული კლასტერიდან ხდება შერჩევა (ზოგჯერ, კლასტერულ შერჩევას **ბუდობრივ შერჩევას** უწოდებენ, ხოლო კლასტერს **ბუდეს**).

კლასტერული შერჩევა ძირითადად გეოგრაფიული საზღვრების მიხედვით რაიონირებული შერჩევაა. ასეთი სახის შერჩევა უფრო ეკონომიურია. მაგალითად, მთელი ქალაქის მოსახლეობიდან შემთხვევით არჩეული 50 ოჯახის გამოკითხვა უფრო შრომატევადია, ვიდრე ერთ-ერთი კლასტერის – საბურთალოს რაიონის 50 მეზობელი ოჯახის გამოკითხვა. თუმცა უნდა შევნიშნოთ, რომ განშრევებული შერჩევისაგან განსხვავებით კლასტერული შერჩევა ვერ გამორიცხავს ჩანაცვლებას, ვინაიდან ერთმანეთის მეზობლად მცხოვრები 50 ოჯახიდან უმრავლესობა თითქმის ერთნაირ საზოგადოებრივ ფენას ეკუთვნის (როგორც წესი, სამეზობლო ოჯახები ხშირად მსგავსი სოციალური ფენიდანაა).

ჩვენ საკმაოდ დაწვრილებით განვიხილეთ შერჩევის ძირითადი ტიპები და მათი მიღების მეთოდები (უფრო დაწვრილებით, მკითხველი ამ მეთოდებს შეიძლება გაეცნოს [10]-ში). შერჩევის შედეგად მიღებულ მონაცემებს **ნედლი მონაცემები** ეწოდებათ. აღწერითი სტატისტიკის პირველ ეტაპზე ხდება ნედლი მონაცემების გრაფიკული წარმოდგენები.

§ 4. ვარიაციული მწკრივი. წერტილოვანი და მესერული დიაგრამები. პოლიგონი

განვიხილოთ n მოცულობის შერჩევა: $x_1, x_2, \ldots, x_n$.

განსაზღვრება 4.1. არაკლებადობის მიხედვით დალაგებულ შერჩევის მონაცემთა $x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)}$ მიმდევრობას **ვარიაციული მწკრივი** ეწოდება.

მაგალითი 4.1. ვთქვათ, ვაკვირდებით კერძო საკუთრებაში არსებული მსხვილფეხა რქოსანი პირუტყვის რაოდენობას. შევარჩიეთ სოფლის 16 მცხოვრები. გვაქვს მონაცემები: 0, 2, 3, 4, 2, 4, 5, 0, 2, 3, 3, 7, 4, 3, 2, 2. პოპულაციას წარმოადგენს კერძო საკუთრებაში არსებული მსხვილფეხა რქოსანი პირუტყვის რაოდენობები. შერჩევის მოცულობაა $n = 16$. ვარიაციული მწკრივი იქნება

0, 0, 2, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 5, 7.

ვარიაციული მწკრივიდან ჩანს თუ რამდენჯერ მეორდება თითოეული ელემენტი.

განსაზღვრება 4.2. ზრდადობის მიხედვით დალაგებული შერჩევის განსხვავებული მონაცემების მიმდევრობას

$$x_{(1)} < x_{(2)} < \cdots < x_{(k)} \quad (k \le n)$$

ზრდადი ვარიაციული მწკრივი ეწოდება. ყოველი $x_{(i)}$ მონაცემის **სიხშირე** ეწოდება იმ რიცხვს, რამდენჯერად ეს მონაცემი გვხვდება შერჩევაში და აღინიშნება ν_i სიმბოლოთი. ν_i სიხშირისა და შერჩევის მოცულობის შეფარდებას ეწოდება $x_{(i)}$ მონაცემის **ფარდობითი სიხშირე** და აღინიშნება W_i სიმბოლოთი. ე. ი. $W_i = \dfrac{\nu_i}{n}$.

განსაზღვრება 4.3. თუ ზრდადი ვარიაციული მწკრივის ელემენტებს მივუწერთ შესაბამის სიხშირეებს (ფარდობით სიხშირეებს) მივიღებთ შერჩევის მონაცემების **სიხშირეთა (ფარდობით სიხშირეთა) განაწილების ცხრილს** (ცხრ. 4.1).

$x_{(i)}$	$x_{(1)}$	$x_{(2)}$	…	$x_{(k)}$
v_i	v_1	v_2	…	v_k
W_i	W_1	W_2	…	W_k

ცხრ. 4.1 ვარიაციული მწკრივის სიხშირეთა და ფარდობით სიხშირეთა განაწილების ცხრილი

4.1. მაგალითში ზრდადი ვარიაციული მწკრივი იქნება 0, 2, 3, 4, 5, 7, ხოლო სიხშირეთა და ფარდობით სიხშირეთა განაწილებების გაერთიანებულ ცხრილს ექნება სახე

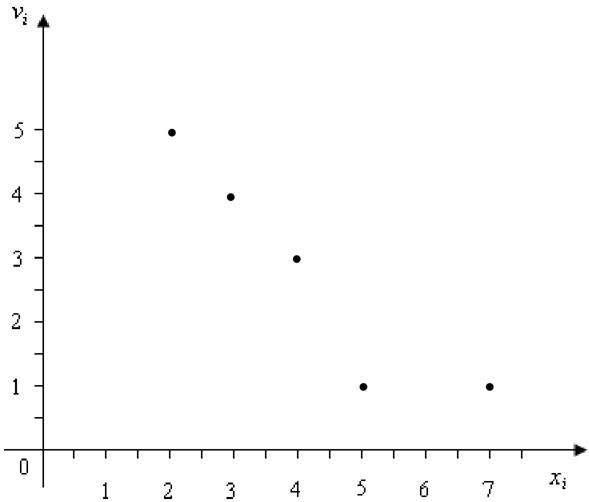
$x_{(i)}$	0	2	3	4	5	7
v_i	2	5	4	3	1	1
W_i	2/16	5/16	4/16	3/16	1/16	1/16

ცხრ. 4.2. სიხშირეთა და ფარდობით სიხშირეთა განაწილების ცხრილი 4.1 მაგალითისათვის

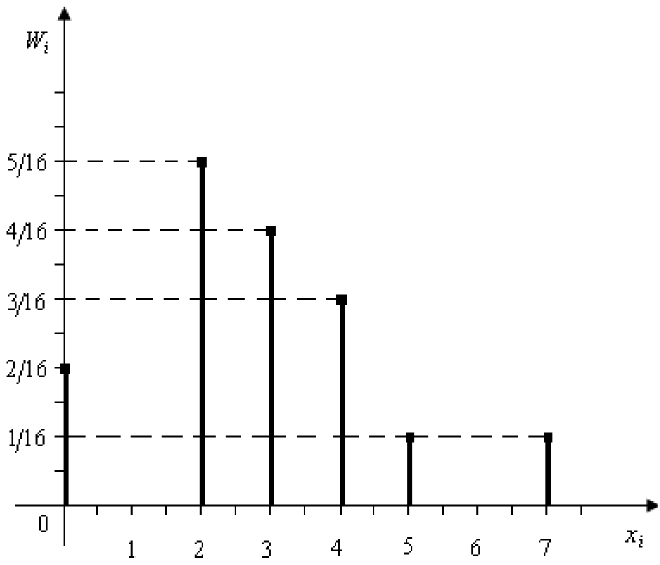
მონაცემების ცხრილური ჩანაწერის გარდა მიმართავენ მათ გრაფიკულ წარმოდგენებს.

განსაზღვრება 4.4 თუ OXY მართკუთხა კოორდინატთა სისტემაში განვიხილავთ ისეთ წერტილებს, რომელთაგან თითოეულის აბსცისაა ზრდადი ვარიაციული მწკრივის ელემენტი, ხოლო ორდინატი მისი შესაბამისი სიხშირე (ფარდობითი სიხშირე), მივიღებთ **სიხშირეთა (ფარდობით სიხშირეთა) წერტილოვან დიაგრამას** (ნახ. 4.1). თუ OX ღერძზე გადავზომავთ ზრდადი ვარიაციული მწკრივის ელემენტებს და თითოეულ მათგანს შევუსაბამებთ მათ შესაბამის სიხშირის (ფარდობითი სიხშირის) სიგრძის მონაკვეთებს. მივიღებთ **სიხშირეთა (ფარდობით სიხშირეთა) განაწილების მესერულ დიაგრამას** (იხ. ნახ. 4.2).

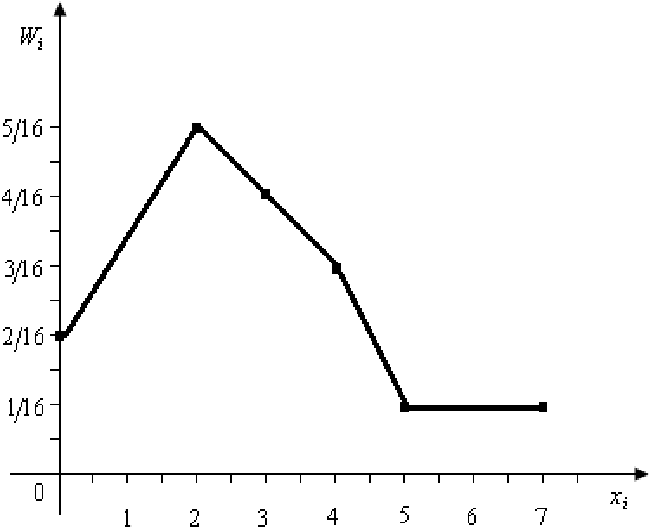
განსაზღვრება 4.5. თუ წერტილოვანი დიაგრამის ბოლო წერტილებს ან მესერული დიაგრამის ბოლო წერტილებს თანამიმდევრულად შევაერთებთ მივიღებთ ტეხილს, რომელსაც **პოლიგონი** ეწოდება. პოლიგონი ხაზოვან დიაგრამას წარმოადგენს (იხ. ნახ. 4.3).



ნახ. 4.1. სიხშირეთა განაწილების წერტილოვანი დიაგრამა 4.1 მაგალითისათვის



ნახ. 4.2. ფარდობით სიხშირეთა განაწილების მესერული დიაგრამა 4.1 მაგალითისათვის



ნახ. 4.3. ფარდობით სიხშირეთა პოლიგონი 4.1 მაგალითისათვის

შენიშვნა. როდესაც პოპულაცია უწყვეტი ტიპისაა იხილავენ **ინტერვალურ ვარიაციულ მწკრივს**. ამისათვის იქცევიან ასე: შერჩევის მინიმალურ და მაქსიმალურ მნიშვნელობებს შორის ინტერვალს ყოფენ ტოლი სიგრძის ინტერვალებად. თითოეულ მათგანს შეუსაბამებენ მასში მოხვედრილ შერჩევის ელემენტთა რაოდენობას და ღებულობენ **ინტერვალურ ვარიაციულ მწკრივს**. მას შეიძლება ქონდეს, მაგალითად, შემდეგი სახე:

Δ_i	$[a_1\ a_2)$	$[a_2\ a_3)$	…	$[a_k\ a_{k+1}]$
v_i	v_1	v_2	…	v_k

სადაც $a_1 = \min\limits_{1 \le i \le n}\{x_i\}$, $a_{k+1} = \max\limits_{1 \le i \le n}\{x_i\}$.

ანალოგიურ საკითხებს დავუბრუნდებით მოგვიანებით (§ 7).

ამოცანები

1. შერჩევის შემდეგი მონაცემებით შეადგინეთ სიხშირეთა და ფარდობით სიხშირეთა ცხრილი.

2, 7, 2, 7, 1, 5, 7, 10, 17, 8, 7, 10, 5, 2, 2, 5.

2. მოცემულია შერჩევა 3, 4, 5, 7, 3, 8, 8, 5, 3, 8, 14. დაწერეთ ვარიაციული მწკრივი და ააგეთ ფარდობით სიხშირეთა პოლიგონი.

3. კლასის ოცმა მოსწავლემ გაიარა გამოცდა მათემატიკაში ათბალიანი შეფასებით. შეფასება 6 დაიმსახურა ორმა მათგანმა, 7 – ცხრამ, 8 ოთხმა, 10 – ორმა, ხოლო დანარჩენმა ვერ მიიღეს დადებითი შეფასება. დაუფიქსირეთ მათ 0 ქულა და ააგეთ მიღებული ქულების სიხშირეთა პოლიგონი.

4. ჩატარდა დაკვირვება 15 დღის განმავლობაში ერთ-ერთ ბენზინგასამართ სადგურზე გაყიდული ბენზინის რაოდენობებზე: 3, 7, 5, 3, 3, 5, 4, 3, 2, 2, 5, 4, 3, 4, 5.

მიღებული მონაცემებით (ისინი აღებულია ტონამდე სიზუსტით დამრგვალებით) ააგეთ სიხშირეთა სვეტოვანი დიაგრამა.

5. ერთ-ერთი არასპორტული პროფილის სკოლის 15 წლის ვაჟმა მოსწავლეებმა მონაწილეობა მიიღეს მძლეოსნობის სექციაში 100 მ დისტანციაზე სირბილში. 30 მონაწილის ნაჩვენები დროები (წამების სიზუსტით) მოცემულია სიხშირეთა ცხრილით

დრო	13	15	16	17	18	19 და მეტი
სიხშირე	2	4	7	8	3	6

ააგეთ ამ ცხრილის შესაბამისი წერტილოვანი დიაგრამა.

6. დააკვირდით რომელიმე ქვეყნის საფეხბურთო ჩემპიონატის სტატისტიკურ ცხრილს და შეადგინეთ გამარჯვებათა რაოდენობების სიხშირეების (გუნდების რაოდენობის მიხედვით) ცხრილი. ააგეთ პოლიგონი და სვეტოვანი დიაგრამა.

7. გააგორეთ კამათელი 100-ჯერ. შეადგინეთ მოსული რიცხვების მიხედვით სიხშირეთა და ფარდობით სიხში-

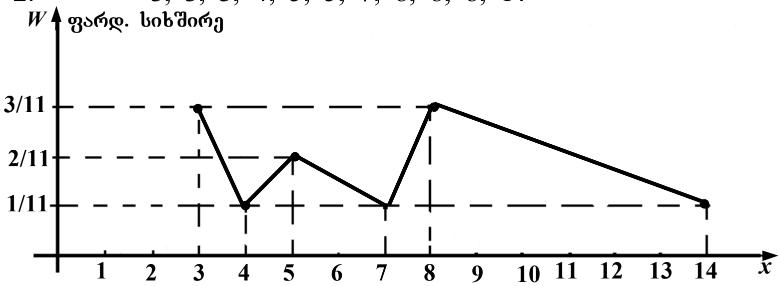
რეთა ცხრილები. ააგეთ შესაბამისი პოლიგინები. რა დასკვნის გაკეთება შეგიძლიათ?

პასუხები.

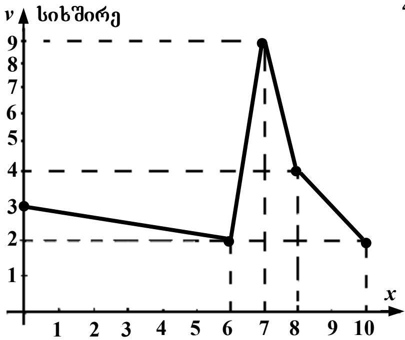
1.

x_i	1	2	5	7	8	10	17
v_i	1	4	3	4	1	2	1
W_i	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

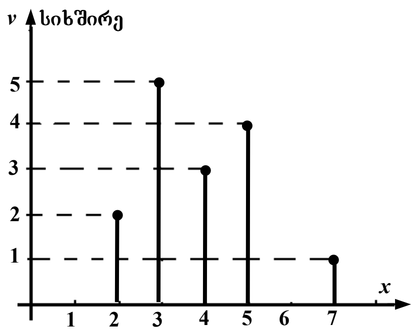
2. 3, 3, 3, 4, 5, 5, 7, 8, 8, 8, 14



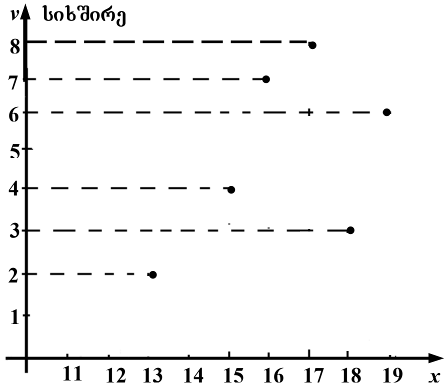
3.



4.



5.



§ 5. უწყვეტი ტიპის მონაცემები. სიხშირეთა ინტერვალური განაწილება. ჰისტოგრამა

როდესაც მონაცემები უწყვეტი ტიპისაა, სიხშირეთა და ფარდობით სიხშირეთა განაწილებები იმ სახით, როგორც წინა პარაგრაფებში განვიხილეთ არაფერს გვაძლევს, ვინაიდან ფაქტობრივად ყველა მონაცემი განსხვავებულია.

ასეთ შემთხვევაში ხდება მონაცემების დაჯგუფება ინტერვალებად. ინტერვალები უნდა იყოს თანაუკვეთი და მონაცემები დაიყოფა თანაუკვეთ კლასებად. ერთ კლასად ჩაითვლებიან ერთი და იგივე ინტერვალში მოხვედრილი მონაცემები.

ხშირად, სასურველია ყველა ინტერვალს ჰქონდეს ტოლი სიგრძე. ასეთ დროს ჯერ ირჩევენ ინტერვალთა რაოდენობას. შერჩევის უმცირესი ელემენტისათვის ნაკლებობით, ხოლო უდიდესისათვის მეტობით შეარჩევენ უახლოეს მთელ a და b რიცხვებს შესაბამისად. კლასის სიგანე იქნება $\Delta = \dfrac{b-a}{k}$ რიცხვი, სადაც k არის

ინტერვალთა რაოდენობა. შემდეგ $[a, b]$ ინტერვალს ყოფენ ტოლი სიგრძის $[a_0, a_1)$; $[a_1\ a_2)$; $[a_2, a_3);\ldots;[a_{k-1},\ a_k]$ (სადაც $a_0 = a$; $a_k = b$) ინტერვალებად. თუ მიღებული ინტერვალებიდან ყოველ i-ურ ინტერვალს მივუწერთ რიცხვს, რომელიც ინტერვალში მოხვედრილი მონაცემების v_i სიხშირის ტოლია, მივიღებთ **მონაცემთა სიხშირეების ინტერვალური განაწილების ცხრილს**. თუ ინტერვალებს შესაბამის W_i ფარდობით სიხშირეებს მივუწერთ მივიღებთ **მონაცემების ფარდობით სიხშირეთა ინტერვალური განაწილების ცხრილს** (იხ. ცხრ 5.1).

ინტერვალი	$[a_0\ \ a_1)$	$[a_1\ \ a_2)$	. . .	$[a_{k-1}\ \ a_k]$
v_i	v_1	v_2	. . .	v_k
W_i	W_1	W_2	. . .	W_k

ცხრ. 5.1. სიხშირეთა და ფარდობით სიხშირეთა ინტერვალური განაწილებების ცხრილი

განსაზღვრება 5.1. ინტერვალის ერთეულზე სიხშირის $\frac{v_i}{\Delta}$ (ფარდობით სიხშირის $\frac{W_i}{\Delta}$) მნიშვნელობას ამ ინტერვალზე **სიხშირის** (შესაბამისად, **ფარდობითი სიხშირის**) **სიმკვრივე** ეწოდება. ცხადია, როდესაც ხდება ტოლ ინტერვალებად დაყოფა სიხშირეთა და ფარდობით სიხშირეთა სიმკვრივეები ისევე განსხვავდებიან ერთმანეთისაგან, როგორც შესაბამისი სიხშირეები და ფარდობითი სიხშირეები.

მაგალითი 5.1. კურსანტმა საცეცხლე მომზადებისას 10სმ რადიუსიან მრგვალ სტენდზე ჩაატარა 80 გასროლიანი სერია. შედეგები (სამიზნის ცენტრიდან დაშორება სანტიმეტრებში) ცხრილშია მოყვანილი. შევადგინოთ სიხშირეთა და ფარდობით სიხშირეთა ინტერვალური განაწილების ცხრილები.

3	4,7	4,1	0,4	1,6	3,2	5,1	7,8	9,6	2,5
7,4	3,1	3,9	4,7	4,8	2,4	5	8,7	4,8	3,4

1,2	0,3	0,8	1,6	3,7	3,2	2,6	2,8	8,3	9,7
5,2	4,8	2,9	2,5	1,4	7,6	3,5	1,1	0,3	0,2
5,3	7	1,8	5,2	9,5	4,8	3,1	2,9	2	5,3
3,4	2,6	2	3,2	5,4	6,7	3,6	3,2	1	0,1
6,7	1,3	0,2	8,4	2,6	1,3	7,3	0,1	2,7	6,9
4,6	3,3	2,8	6,5	1,8	2,9	1,4	4,6	4,8	2,9

ამოხსნა. ამ მაგალითში პოპულაცია წარმოსახვითია. იგი წარმოადგენს კურსანტის ყველა შესაძლო გასროლათა შედეგებს. შერჩევის მოცულობაა $n=80$. უმცირესი მონაცემია 0,1, უდიდესი – 9,7. ე.ი. $a=0$; $b=10$. ინტერვალების რაოდენობად შევარჩიოთ $k=10$ და თითო ინტერვალის სიგრძე იქნება $\Delta=\frac{10-0}{10}=1$. მივიღებთ ინტერვალებს $[0\,,1)$; $[1\,,2);\ldots;[9,\,10]$. სიხშირეთა და ფარდობით სიხშირეთა ინტერვალური განაწილების ცხრილს ექნება სახე (ცხრ. 5.2).

ინტერ-ვალი	[0, 1]	[1, 2)	[2, 3)	[3, 4)	[4, 5)	[5,6)	[6,7)	[7,8)	[8,9)	[9, 10]
v_i	7	11	15	14	10	7	4	5	4	3
W_i	7/80	11/80	15/80	14/80	10/80	7/80	4/80	5/80	4/80	3/80

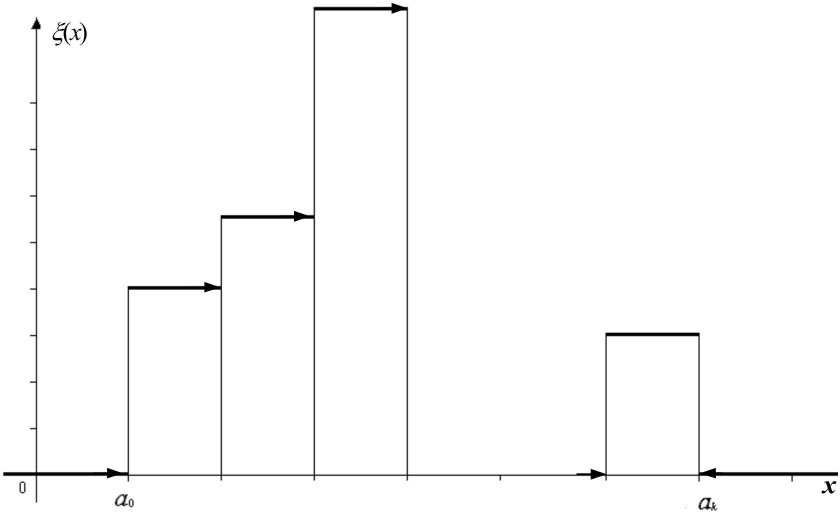
ცხრ. 5.2. სიხშირეთა და ფარდობით სიხშირეთა განაწილების ცხრილი 5.1 მაგალითისათვის დაყოფის ინტერვალთა k =10 რაოდენობის შემთხვევაში

უწყვეტი ტიპის მონაცემების გრაფიკული წარმოდგენისათვის წერტილოვანი და მესერული დიაგრამების ნაცვლად აგებენ **სიხშირეთა და ფარდობით სიხშირეთა ჰისტოგრამებს**. **სიხშირეთა** $\xi(x)$ **(ფარდობით სიხშირეთა** $\eta(x)$**) ჰისტოგრამის** ასაგებად აბსცისათა ღერძზე აიღება დაყოფის ინტერვალები და თითოეულ მათგანზე, როგორც ფუძეზე, დაიდგმება იმ სიმაღლის მართკუთხედი, რომლის ფართობი ტოლია მოცემული ინტერვალის სიხშირის (ფარდობითი სიხშირის). ჰისტოგრამა წარმოადგენს ნამდვილი ცვლადის ფუნქციას, რომელიც ინტერვალის ყოველ წერტილში იღებს მასზე დადგმული მართკუთხედის

სიმაღლის ტოლ მნიშვნელობას, ხოლო ინტერვალების გარეთ კი ნულის ტოლია (იხ. ცხრ. 5.3 და ნახ. 5.1).

ინტერ-ვალი	$x < a_0$	$x \in [a_0\ a_1)$	$x \in [a_1\ a_2)$	...	$x \in [a_{k-1} a_k]$	$x > a_k$
$\xi(x)$	0	$\frac{v_1}{\Delta}$	$\frac{v_2}{\Delta}$	...	$\frac{v_k}{\Delta}$	0
$\eta(x)$	0	$\frac{W_1}{\Delta}$	$\frac{W_2}{\Delta}$	...	$\frac{W_k}{\Delta}$	0

ცხრ. 5.3. სიხშირეთა და ფარდობით სიხშირეთა ჰისტოგრამების ცხრილი



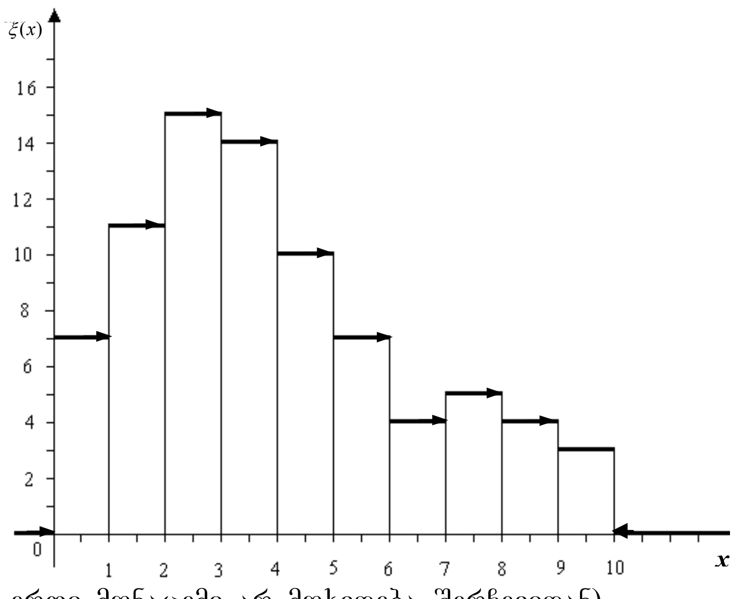
ნახ. 5.1. სიხშირეთა $\xi(x)$ ჰისტოგრამა 5.4 ცხრილის მიხედვით

ცხადია, სიხშირეთა (ფარდობით სიხშირეთა) ჰისტოგრამასა და *OX* ღერძს შორის მოთავსებული არის ფართობი n-ის (შესაბამისად, 1-ის) ტოლია.

ავაგოთ სიხშირეთა ჰისტოგრამა (ნახ. 5.2) დაყოფის ინტერვალთა $k=10$ რაოდენობის შემთხვევაში მიღებული განაწილების ცხრილისათვის (ცხ. 5.2). თუ დავაკვირდებით შევამჩნევთ, რომ მონაცემთა დიდი ნაწილი თავს იყრის

[1, 4) ინტერვალში, ხოლო მათგან ყველაზე უფრო ბევრი მონაცემი [2, 3) ინტერვალში ხვდება.

ბუნებრივია იმისათვის, რომ ჰისტოგრამამ სრულად წარმოაჩინოს შერჩევის სურათი, დიდი მნიშვნელობა აქვს ინტერვალთა რაოდენობის, k-ს სწორად შერჩევას. თუ k ძალიან დიდი იქნება შერჩევის სიმრავლე შეიძლება იმდენად ბევრ ინტერვალად დაიყოს, რომ მივიღოთ მონაცემთა საწყისი სიმრავლის მსგავსი სიმრავლე, ხოლო, პირიქით, ინტერვალების სიმცირემ კი შეიძლება გამოიწვიოს ინფორმაციის დაკარგვა. გასათვალისწინებელია ის ფაქტიც, რომ ინტერვალების რაოდენობის არჩევისას უნდა შევეცადოთ თავიდან ავიცილოთ ცარიელი ინტერვალები (რომლებშიც



არც ერთი მონაცემი არ მოხვდება შერჩევიდან).

ნახ. 5.2. სიხშირეთა ჰისტოგრამა 5.1 მაგალითისათვის ინტერვალთა k =10 რაოდენობად დაყოფის შემთხვევაში

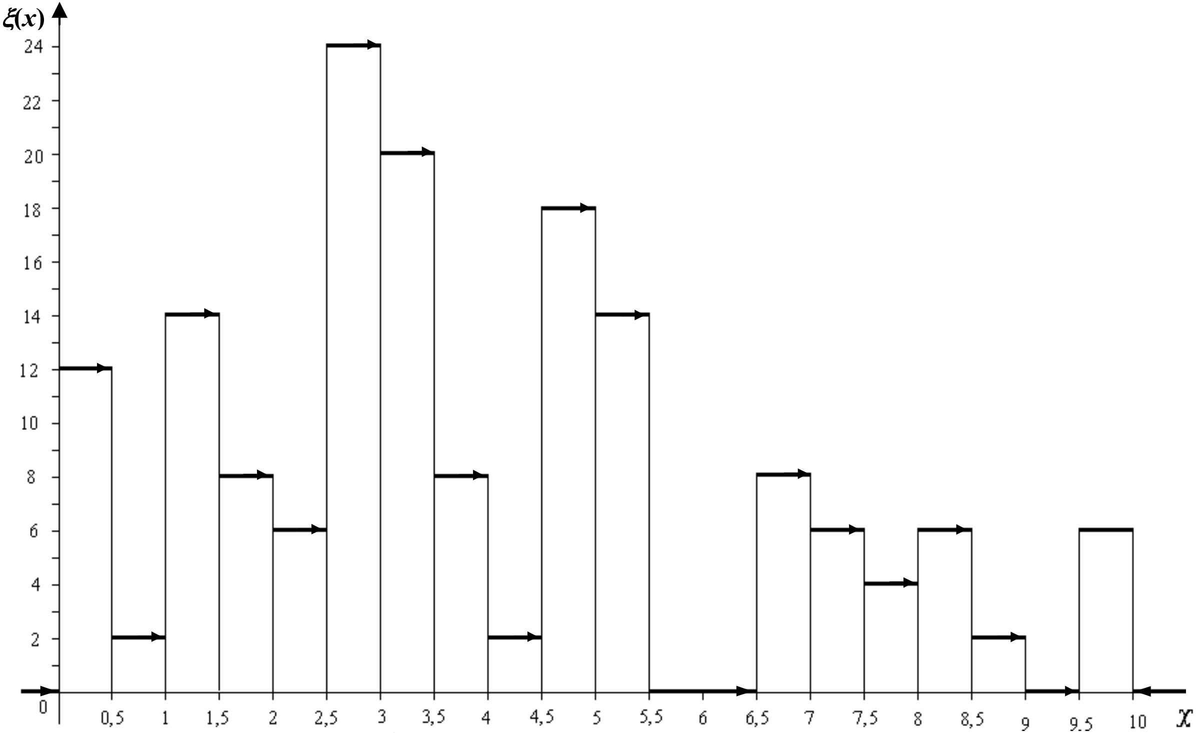
ზემოთ თქმულის საილუსტრაციოდ განვიხილოთ 5.1 მაგალითში $k = 20$ რაოდენობის ინტერვალებად დაყოფა. შევადგინოთ სიხშირეთა განაწილების ცხრილი (იხ. ცხრ. 5.4) და ავაგოთ ჰისტოგრამა (ნახ. 5.3). ცხადია $\Delta = \dfrac{10-0}{20} = 0,5$.

Δ_i	v_i
[0 0.5)	6
[0.5 1)	1
[1 1.5)	7
[1.5 2)	4
[2 2.5)	3
[2.5 3)	12
[3 3.5)	10
[3.5 4)	4
[4 4.5)	1
[4.5 5)	9

Δ_i	v_i
[5 5.5)	7
[5.5 6)	0
[6 6.5)	0
[6.5 7)	4
[7 7.5)	3
[7.5 8)	2
[8 8.5)	3
[8.5 9)	1
[9 9.5)	0
[9.5 10]	3

ცხრ. 5.4. სიხშირეთა განაწილების ცხრილი 5.1 მაგალითისათვის დაყოფის ინტერვალთა $k=20$ რაოდენობის შემთხვევაში

ავაგოთ შესაბამისი ჰისტოგრამა



ნახ. 5.3. სიხშირეთა ჰისტოგრამა 5.1 მაგალითისათვის დაყოფის ინტერვალთა $k = 20$ რაოდენობის შემთხვევაში

ერთი შეხედვითაც ჩანს, რომ გაჩნდა უამრავი გადახრა ძირითადი გრაფიკული წარმოდგენიდან, რომლებიც არავითარ ინფორმაციას არ გვაძლევენ.

კლასების k რაოდენობის არჩევა რეკომენდებულია ე.წ. **ლოგარითმის წესით (სტეიჯსის წესი** – Sturgis' rule)

$$k \approx 1 + 3.322 \log_{10} n$$

ან **რაისის წესით** (Rice rule)

$$k \approx 2\sqrt[3]{n}\ .$$

პირველის გათვალისწინებით შერჩევის n მოცულობიდან გამომდინარე რეკომედირებულია k-ს შემდეგი მნიშვნელობები:

n	$n \le 50$	$n \in [50\ 100]$	$n \in [100\ 250]$	$n \ge 250$
k	$k \in [4\ 7]$	$k \in [6\ 10]$	$k \in [7\ 12]$	$k \in [10\ 20]$

ამოცანები

1. ერთ-ერთ პოლიკლინიკაში, ერთი თვის განმავლობაში ყოველდღიურად აკვირდებოდნენ იმ ბავშვთა რაოდენობებს, რომლებიც გამოკვლევებისას სარგებლობდნენ სახელმწიფო დაზღვევით. მიღებული მონაცემებით შეადგინეთ ყოველდღიურად დაზღვევით მოსარგებლე ბავშვების რაოდენობების სიხშირეთა ინტერვალური განაწილების ცხრილი და ააგეთ სიხშირეთა ჰისტოგრამა.

24	11	17	28	15	2	12	24	10	26
17	13	3	9	20	6	14	14	16	15
15	25	14	15	8	19	8	10	23	9

(**მითითება**: ინტერვალების რაოდენობა დაადგინეთ $k \approx 2\sqrt[3]{n}$ ფორმულით).

2. მოცემულია $n = 50$ მოცულობის შერჩევა

8,6	3,9	5,31	2,3	8,27	6	13,8	11,2	1,6	2,03
1,17	0,4	6,82	9,64	13,05	0,8	5,2	3,72	4,88	3,9
5,01	6,43	0,12	4,5	7	2	8,17	6,49	0	7,58
12,68	2,45	2,14	7,29	1,35	13,14	8,05	1,2	9,35	5,78
1,3	2,83	6,21	1,7	10,8	3,5	9,3	10,9	0,4	1,92

შეადგინეთ ფარდობით სიხშირეთა ინტერვალური განაწილების ცხრილი და ააგეთ ფარდობით სიხშირეთა ჰისტოგრამა (**მითითება**: ინტერვალების რაოდენობა დაადგინეთ $k \approx 1+3{,}322\log_{10} n$ ფორმულით).

3. მოცემულია სიხშირეთა ინტერვალური განაწილების ცხრილი

[−7 −4)	[−4 −1)	[−1 2)	[2 5)	[5 8]
4	6	9	12	8

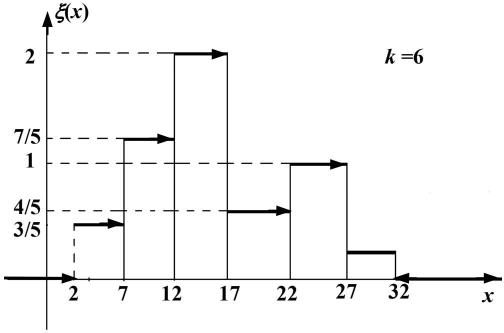
ააგეთ სიხშირეთა ჰისტოგრამა.

4. ჩაატარეთ სამიზნეზე სროლათა სერია. ყოველი ცდის შედეგი იყოს სამიზნის ცენტრიდან დაშორების მანძილი. ააგეთ სიხშირეთა და ფარდობით სიხშირეთა ჰისტოგრამები.

პასუხები

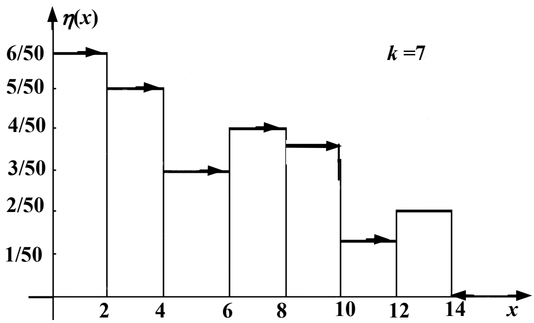
1.

[2 7)	3
[7 12)	7
[12 17)	10
[17 22)	4
[22 27)	5
[27 32]	1

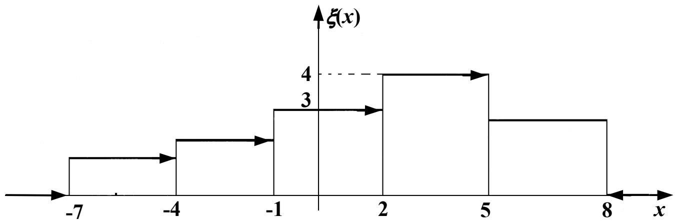


2.

[0 2)	12/50
[2 4)	10/50
[4 6)	6/50
[6 8)	8/50
[8 10)	7/50
[8 12)	3/50
[12 14]	4/50



3.



§ 6. ფოთლებიანი ღეროების მსგავსი დიაგრამა

ხშირ შემთხვევაში საჭირო ინფორმაციის სრულად წარმოჩინებისათვის სიხშირეთა განაწილების გრაფიკულად წარმოდგენასთან ერთად გამოიყენება ე.წ. **ფოთლებიანი ღეროების მსგავსი დიაგრამა.** მისი აგებისათვის საჭიროა მონაცემებიდან წინასწარ გამოიყოს იმდენი ნიშანი, რომ დარჩეს ერთნაირი თანრიგის მაჩვენებელი ციფრები. მიღებული მარცხენა ნაწილები უნდა დალაგდეს ზრდადობის მიხედვით გამეორებების გარეშე (რაც შეადგენს ვერტიკალურ ღეროს). თითოეულ ამ მონაცემს მარჯვნიდან უნდა მიეწეროს გამოყოფილი ნიშნები თანამიმდევრობის მიხედვით (მივიღებთ ჰორიზონტალურ ღეროებს). თუ მონაცემები ათწილადებია შეიძლება ათწილადი ნიშნები გამოვყოთ.

განვიხილოთ კონკრეტული მაგალითი.

მაგალითი 6.1. ავაგოთ ფოთლებიანი ღეროების მსგავსი დიაგრამა 5.1 მაგალითისათვის. ამ დროს ვერტიკალურ ღეროს წარმოადგენენ მთელი რიცხვები 0, 1, 2, …, 9, ხოლო ათწილადი ნაწილები იქნება ფოთლები. ჩამოვწეროთ ვერტიკალური ღეროს ელემენტები სვეტში და მარჯვნივ მივუწეროთ ათწილადი ნაწილები მათი რიგის მიხედვით (იხ. ცხრ. 6.1).

0	4, 3, 8, 3, 2, 2, 1
1	6, 2, 6, 4, 1, 8, 0, 3, 3, 8, 4
2	5, 4, 6, 8, 9, 5, 9, 0, 6, 0, 6, 7, 8, 9, 9
3	0, 2, 1, 9, 4, 7, 2, 5, 1, 4, 2, 6, 2, 3
4	7, 1, 7, 8, 8, 8, 8, 6, 6, 8
5	1, 0, 2, 3, 2, 3, 4
6	7, 7, 9, 5
7	8, 4, 6, 0, 3
8	7, 3, 1, 4
9	6, 7, 5

ცხრ. 6.1. ფოთლებიანი ღეროების მსგავსი დიაგრამა 5.1 მაგალითისათვის

ადვილად მისახვედრია, რომ თუ დიაგრამას მარცხნივ 90°-ით მოვაბრუნებთ მას ჰისტოგრამის სახე ექნება (შეადარეთ ნახ. 5.2-ს). ჰისტოგრამისგან განსხვავებით ეს დიაგრამა არ კარგავს არც ერთ წვრილმან ინფორმაციას. მაგალითად, 5.2 ნახაზზე ჰისტოგრამაზე ჩანს მხოლოდ ის, რომ [1, 2) შუალედში გვხვდება 11 ცალი მონაცემი. კონკრეტულად რომელი მონაცემებია და რამდენჯერ გვხვდება თითოეული ან რა თანმიმდევრობით ეს ფოთლებიანი ღეროების მსგავსი დიაგრამაზე იმ ჰორიზონტალურ ღეროზე შეგვიძლია მოვიძიოთ, რომელიც ვერტიკალური ღეროს (რიცხვი ერთის) მარჯვნივაა.

ფოთლებიანი ღეროების მსგავსი დიაგრამა შეიძლება დისკრეტული მონაცემებისთვისაც ავაგოთ.

მაგალითი 6.2. განვიხილოთ ავტოდრომზე ახლად შექმნილი სატვირთო ავტომობილების გამოცდის დროს გარბენილი მანძილები (კმ-ში):
20150, 24003, 20406, 22127, 24117, 20328, 23215, 20182, 20150, 22054, 22160

ამ მონაცემებიდან გამოვყოთ მძიმით (ასეულების მაჩვენებლის ჩათვლით) მარჯვნიდან სამ-სამი ციფრი. ე.ი. ავტომობილის გარბენი გავზომოთ 1000 კილომეტრის სიზუსტით. შემდეგი ცხრილი გვიჩვენებს ფოთლებიანი ღეროების დიაგრამის შედგენის პროცესს.

გარბენის ნომერი	მონაცე-მები	ვარია-ციული მწკრივი	ვერტიკა-ლური ღეროს აგება		ფოთლებიანი ღეროების დიაგრამა	
					ვერტი-კალური ღერო	ფოთლები
1	20150	20150	20	150	20	150,406,328,182, 150
2	24003	20150	20	406	22	127, 054, 160
3	20406	20182	20	328	23	215
4	22127	20328	20	182	24	003,117
5	24117	20406	20	150		
6	20328	22054	22	127		
7	23215	22127	22	054		
8	20182	22160	22	160		
9	20150	23215	23	215		
10	22054	24003	24	003		
11	22160	24117	24	117		

ცხრ. 6.2. ფოთლებიანი ღეროების მსგავსი დიაგრამის აგება 6.2 მაგალითისათვის

ფოთლები შეიძლება დალაგდეს არაკლებადობის მიხედვით.

იმ შემთხვევაში, როდესაც რომელიმე ღეროს ბევრი ფოთლები აქვს, ხდება ამ ფოთლების დაყოფა რამდენიმე ინტერვალად. შესაბამისად, ეს ღერო ცხრილში იგივე რაოდენობით იქნება წარმოდგენილი თითოეულის გასწვრივ თავისი ფოთლებით.

მაგალითად, თუ გვაქვს ნედლი მონაცემები ღეროს მაჩვენებელი 2-სთვის

25, 21, 22, 21, 20, 22, 21, 21, 25, 27, 25, 25, 21, 22, 23, 24

შეიძლება ფოთლები (0,1); (2,3); (4,5); (6,7); დავალაგოთ განცალკევებით $2^{\bullet}$-ის გასწვრივ 0 და 1-ები, შემდეგ a-ს გასწვრივ 2 და 3-ები, b-ს გასწვრივ 4 და 5-ები და ა.შ. ლათინური ანბანის მიხედვით, ხოლო ბოლო ფოთლები დავალაგოთ 2^{+}-ის გასწვრივ გვექნება (ფოთლები დალაგებული ზრდადი თანმიმდევრობით)

ღერო	ფოთლები
$2^{\bullet}$	0, 1, 1, 1, 1, 1
a	2, 2, 2, 3
b	4, 5, 5, 5, 5
2^{+}	7

შევნიშნოთ, რომ ვერტიკალურ სვეტში ღეროები შეიძლება დავალაგოთ ზემოდან ქვემოთ კლების მიხედვითაც.

ამოცანები

1. მოცემულია $n = 30$ მოცულობის შერჩევა. ააგეთ ფოთლებიანი ღეროების მსგავსი დიაგრამა. მონაცემები დაალაგეთ არაკლებადობის მიხედვით

7,1	6,9	8,15	8,0	10,25	9,7	8,5	7	7	6,84
8,3	8,35	8,52	9,2	9,25	7,8	8,9	8,5	7,4	9,2
10,1	10,15	6,8	8,54	8,6	9,2	9,3	8,7	9,3	8,73

2. მენეჯერს აქვს ერთი თვის განმავლობაში სუპერმარკეტში ყოველდღიურად გაყიდული ყავის (წონა გრამებში) მონაცემები. ააგეთ ფოთლებიანი ღეროების მსგავსი დიაგრამა.

3150	3100	3200	3150	2950	2900	3000	3150
3100	2900	2850	2800	2900	3000	3250	3050
3200	3050	3100	3000	3050	2900	2950	2800
2850	2800	2950	3000	3100	3250		

3. ცხრილის სახით მოცემულია ოთხი წლის განმავლობაში მიკროსაფინანსო ორგანიზაციის მიერ ყოველთვიურად სესხად გაცემული თანხების რაოდენობები (ლარებში). ააგეთ ფოთლებიანი ღეროების მსგავსი დიაგრამა (ფოთლები დაალაგეთ არაკლებადობის მიხედვით).

22700	23300	24750	23200	22350	22300
24100	24600	22100	24350	25600	24100
25150	23000	26100	26400	25200	21200
20000	21600	24450	25200	24200	20900
23100	25000	24600	23200	26100	20200
25450	24800	22100	26400	26000	24500
23400	26300	23150	21150	25200	25450
26350	26200	20300	23250	25100	24200

პასუხები

1.

6	80, 84, 90
7	0, 0, 10, 40, 80
8*	0, 15, 30, 35
a	50, 50, 52, 54
8^{+}	60, 70, 73, 90
9*	20, 20, 20, 25
9^{+}	30, 30, 70
10	10, 15, 25

2.

28	50, 00, 00, 50, 00
29	50, 00, 00, 00, 00, 50, 50
30	00, 00, 00, 50, 50, 00, 50, 00
31	50, 00, 50, 50, 00, 00
32	00, 50, 00

3.

20	0, 200, 300, 900
21	150, 200, 600
22	100, 100, 300, 350, 700
23	0, 100, 150, 200, 200, 250, 300, 400
24*	100, 100, 200, 200, 350, 450
24^{+}	500, 600, 600, 750, 800
25	0, 100, 150, 200, 200, 200, 450, 450, 600
26	0, 100, 100, 200, 300, 350, 400, 400

§ 7. ინტერვალური მონაცემების დისკრეტიზირება და დისკრეტული მონაცემების ინტერვალებად დაყოფა

უწყვეტი ტიპის მონაცემებისათვის მე-5 პარაგრაფში განვიხილეთ მონაცემთა ინტერვალებად დაყოფის საშუალებით სიხშირეთა და ფარდობით სიხშირეთა ჰისტოგრამების აგება. ხშირ შემთხვევაში მიმართავენ ამ მონაცემების დისკრეტიზირებას და მათთვის ისეთივე გრაფიკულ წამოდგენებს იღებენ, როგორიც დისკრეტული მონაცემებისათვის გვქონდა.

შევჩერდეთ ამ საკითხზე და განვიხილოთ უწყვეტი მონაცემების სიხშირეთა ინტერვალური განაწილების ცხრილი (ცხრ. 7.1). განვიხილოთ ტოლი Δ სიგრძის ინტერვალებად დაყოფის შემთხვევა

ინტერვალი	$[a_0\ a_1]$	$[a_1\ a_2)$	…	$[a_{k\text{-}1}\quad a_k]$
სიხშირე v_i	v_1	v_2	…	v_k

ცხრ. 7.1. სიხშირეთა ინტერვალური განაწილების ცხრილი

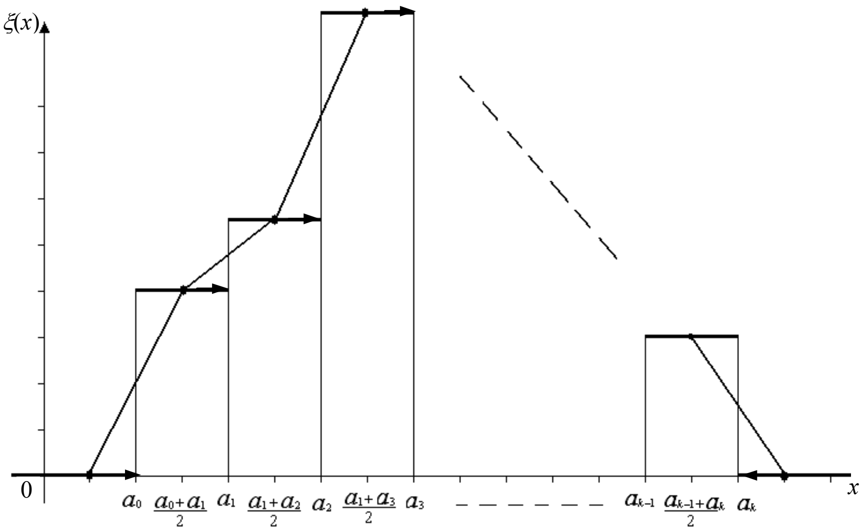
ამ ცხრილის მიხედვით შევადგინოთ სიხშირეთა ჰისტოგრამის ასაგებად საჭირო ცხრილი. თითოეული *i*-ური ინტერვალის შუა წერტილს შევუსაბამოთ შესაბამისი სიხშირის სიმკვრივე $\frac{v_i}{\Delta}$ (ანუ ჰისტოგრამის მნიშვნელობა ამ წერტილზე). მივიღებთ სიხშირეთა განაწილების ცხრილს დისკრეტიზირებული უწყვეტი ტიპის მონაცემებისათვის (ცხრ. 7.2)

x_i	$\frac{a_0+a_1}{2}$	$\frac{a_1+a_2}{2}$	…	$\frac{a_{k-1}+a_k}{2}$
$\frac{v_i}{\Delta}$	$\frac{v_1}{\Delta}$	$\frac{v_2}{\Delta}$	…	$\frac{v_k}{\Delta}$

ცხრ. 7.2. სიხშირეთა განაწილების ცხრილი დისკრეტიზირებული უწყვეტი ტიპის მონაცემებისათვის

თუ ამ ცხრილის მონაცემებით ავაგებთ წერტილებს, მივიღებთ სიხშირეთა წერტილოვან დიაგრამას უწყვეტი ტიპის მონაცემებისათვის.

ახლა პირველი ინტერვალის მარცხენა ბოლოს მარცხნივ, ხოლო ბოლო ინტერვალის მარჯვენა ბოლოს მარჯვნივ აბსცისათა ღერძზე მოვინიშნოთ შესაბამისი წერტილებიდან $\frac{\Delta}{2}$ მანძილით დაშორებული წერტილები და მიღებული წერტილები მიმდევრობით შევაერთოთ. ასე მიღებულ ტეხილს უწყვეტი **მონაცემების პოლიგონი** ეწოდება. ნახაზზე (ნახ. 7.1) აგებულია სიხშირეთა ჰისტოგრამა და პოლიგონი უწყვეტი ტიპის მონაცემებისათვის.



ნახ. 7.1. სიხშირეთა $\xi(x)$ ჰისტოგრამა და პოლიგონი უწყვეტი ტიპის მონაცემებისათვის

ასევე შეგვიძლია ავაგოთ ფარდობით სიხშირეთა პოლიგონიც უწყვეტი განაწილებისათვის. იმ შემთხვევაში, როცა ჰისტოგრამები აგებულია ტოლი სიდიდის ინტერვალებად დაყოფით სიხშირეთა პოლიგონსა და OX ღერძს შორის

მოთავსებული სიბრტყის ნაწილის ფართობი შერჩევის n მოცულობის ტოლია. ფარდობით სიხშირეთა პოლიგონსა და OX ღერძს შორის მოთავსებული სიბრტყის ნაწილის ფართობი კი 1-ის ტოლია.

მაგალითი 7.1. განვიხილოთ $n=30$ მოცულობის შერჩევა ნაკრძალში არსებული გარეული თხების წონების (კგ-ებში) პოპულაციიდან.

9,5	16,2	17	12,5	14,8	8,4	10,1	11,5	15,2	17,6
17,4	9	10,2	14,3	13,1	9,3	16,4	11	14,9	16,2
8,5	17,1	17,8	9,7	14,2	8,1	16,3	15,2	17,4	9,2

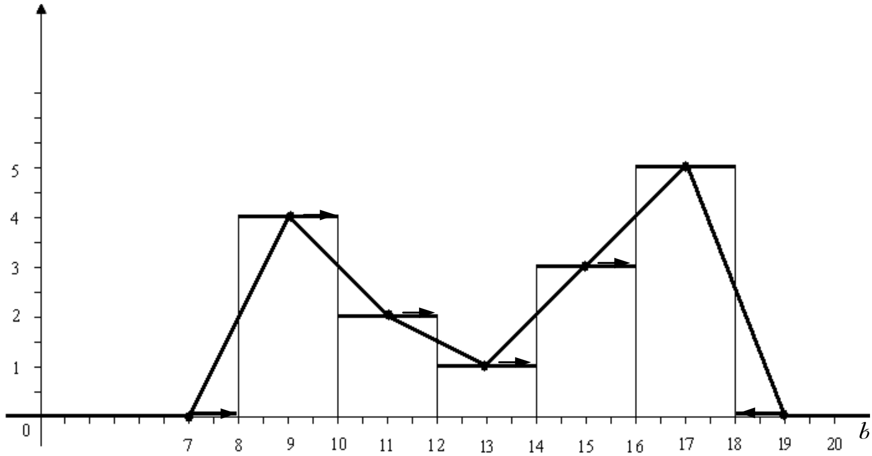
ავაგოთ სიხშირეთა ჰისტოგრამა და პოლიგონი.

ამოხსნა. უმცირესი მონაცემია 8,1 და უდიდესი 17,8. განვიხილოთ [8, 18] შუალედი და დავყოთ $k=5$ ინტერვალად. შესაბამისად ინტერვალების სიგრძეები იქნება $\Delta=\frac{18-8}{5}=2$. შევადგინოთ სიხშირეთა განაწილების ცხრილი გვერდით მივუწეროთ ინტერვალების შუა წერტილები და სიხშირეთა ჰისტოგრამის მნიშვნელობები (ცხრ. 7.3).

ინტერ-ვალი	სიხშირეთა დათვლა	ინტერვალის სიხშირე	ჰისტოგრამის მნიშვნელობა	ინტერვალის შუა წერტილი
[8 10)	卌 |||	8	4	9
[10 12)	||||	4	2	11
[12 14)	||	2	1	13
[14 16)	卌 |	6	3	15
[16 18]	卌 卌	10	5	17
$(-\infty\ 8)$	0	0	0	7
$(18\ +\infty)$	0	0	0	19

ცხრ. 7.3. პოლიგონის ცხრილის მიღება 7.1 მაგალითისათვის

ავაგოთ ერთი და იგივე ნახაზზე შესაბამისი ჰისტოგრამა და პოლიგონი (ნახ. 7.2).



ნახ. 7.2. სიხშირეთა ჰისტოგრამა და პოლიგონი 7.1 მაგალითისათვის

ზოგჯერ დისკრეტული მონაცემების პოლიგონის ასაგებად საჭიროა მათი ინტერვალებად დაჯგუფება, ჰისტოგრამის აგება და შემდეგ დისკრეტიზირების საშუალებით პოლიგონის მიღება. ამ გზას მივმართავთ ისეთ შემთხვევაში, როდესაც დისკრეტული მონაცემები თითქმის არ მეორდებიან ან ძალიან ბევრ განსხვავებულ მნიშვნელობებს იღებენ. განვიხილოთ მაგალითი.

მაგალითი 7.2. გვაქვს ფერმერის მიერ საკარმიდამო ნაკვეთში სხვადასხვა წლებში მოყვანილი სიმინდის მოსავლის (კგ-ებში) რაოდენობების $n=20$ მოცულობის შერჩევა.

245	288	275	246	310	295	302	267	292	305
310	257	274	302	311	298	277	285	315	304

ავაგოთ სიხშირეთა პოლიგონი.

ამოხსნა. როგორც ჩანს მონაცემები თითქმის არ მეორდება, ამიტომ სიხშირეთა წერტილოვანი დიაგრამის

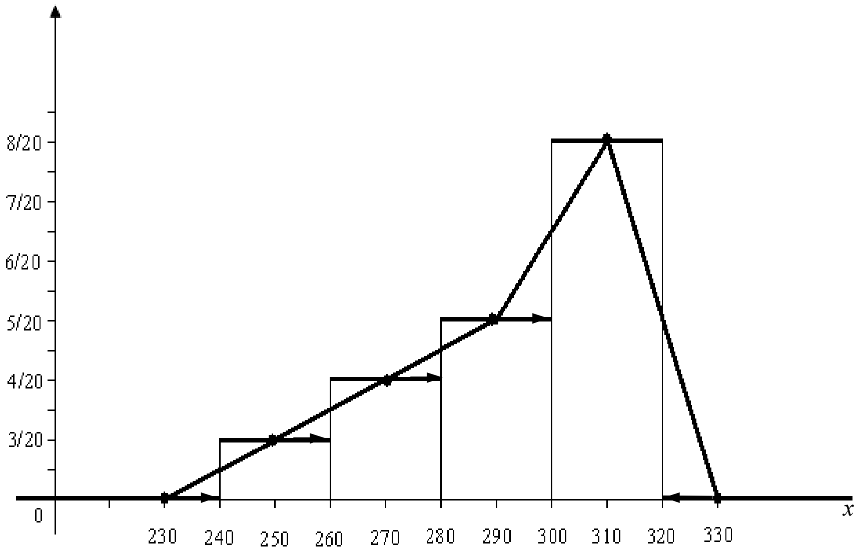
აგებას (რომლის წერტილების შეერთებითაც მიიღება სიხშირეთა პოლიგონი) აზრი არ აქვს.

ჯერ მოვახდინოთ მონაცემების ინტერვალებად დაჯგუფება. უმცირესი მონაცემია 245 უდიდესი კი 315. განვიხილოთ [240, 320] შუალედი და დავყოთ $k=4$ რაოდენობის ტოლი სიგრძის შუალედად. თითოეულის სიგრძე იქნება $\Delta=\frac{320-240}{4}=20$. ავიღოთ თითოეული ინტერვალის შუა წერტილი, ხოლო კიდურა წერტილებისაგან მარცხნივ და მარჯვნივ [240, 320] ინტერვალის გარეთ $\frac{\Delta}{2}=10$ მანძილით დაშორებული კიდევ თითო წერტილი 230 და 330. თითოეულ წერტილს შევუსაბამოთ სიხშირის სიმკვრივე შესაბამის ინტერვალზე (ცხრ. 7.4).

ინტერვალი	სიხშირეთა დათვლა	სიხშირე v_i	სიხშირის სიმკვრივე	ინტერვალის წერტილი
$(-\infty\ \ 240)$	0	0	0	230
[240 260)	|||	3	3/20	250
[260 280)	||||	4	4/20	270
[280 300)	$\cancel{\|\|\|\|}$	5	5/20	290
[300 320]	$\cancel{\|\|\|\|}$ |||	8	8/20	310
$(320\ \ \infty)$	0	0	0	330

ცხრ. 7.4. სიხშირეთა პოლიგონის ცხრილის მიღება 7.2 მაგალითისათვის

ცხადია ამ ცხრილის საშუალებით შეიძლება ავაგოთ მოცემული დისკრეტული მონაცემებისათვის ჰისტოგრამაც და პოლიგონიც (იხ. ნახ. 7.3).



ნახ. 7.3. სიხშირეთა ჰისტოგრამა და პოლიგონი 7.2 მაგალითისათვის

ამოცანები

1. მოცემულია $n=30$ მოცულობის შერჩევა დისკრეტული მონაცემებიდან. შეადგინეთ ინტერვალური განაწილების ცხრილი და ააგეთ სიხშირეთა ჰისტოგრამა. (**მითითება**: მოახდინეთ 5 ტოლ ინტერვალად დაყოფა)

20, 18, 12, 13, 28, 11, 17, 12, 16, 21, 21, 12, 14, 15, 22,
16, 16, 20, 14, 23, 10, 12, 24, 24, 19, 10, 18, 19, 25, 26.

2. შედგინეთ სიხშირეთა ინტერვალური განაწილების ცხრილი და ააგეთ პოლიგონი §5-ის ამოცანების № 2 ამოცანის მოცემულობისათვის. (მითითება: მოახდინეთ მონაცემების დაჯგუფება 7 ტოლ ინტერვალად).

3. სუპერმარკეტის მენეჯერს აქვს ძეხვეულისა და ხორცპროდუქტების განყოფილების დღიური ნავაჭრის რაოდენობის მონაცემები ორი თვის განმავლობაში (ასეულ

ლარებში). შედგინეთ სიხშირეთა ინტერვალური განაწილების ცხრილი და ააგეთ სიხშირეთა პოლიგონი (მითითება: მოახდინეთ მონაცემების დაჯგუფება 6 ტოლ ინტერვალად).

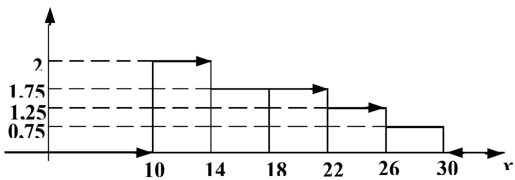
2,897	3,3	3,548	2,583	2,665	2,687	2,972	2,986	3,071	3,157
3,269	2,512	2,741	3,021	2,838	2,804	2,735	2,91	3,148	2,902
2,754	3,428	2,62	3,407	3,368	3,16	2,713	2,941	2,771	3,245
2,561	2,8	3,397	3,61	3,025	2,961	3,083	2,781	3,193	3,213
3,08	2,983	3,312	3,214	2,507	3,2	2,826	2,9	2,823	2,973
3,181	2,82	2,774	3,579	2,584	2,885	2,8	2,618	3,421	3,204

4. თქვენ თავად მოიპოვეთ მონაცემები (ან დაკვირვებით, ან მეორადი მონაცემები) დოლარის კურსის ლართან მიმართებაში. დაახლოებით 50 მონაცემის შემთხვევაში ააგეთ სიხშირეთა ჰისტოგრამა და პოლიგონი.

პასუხები

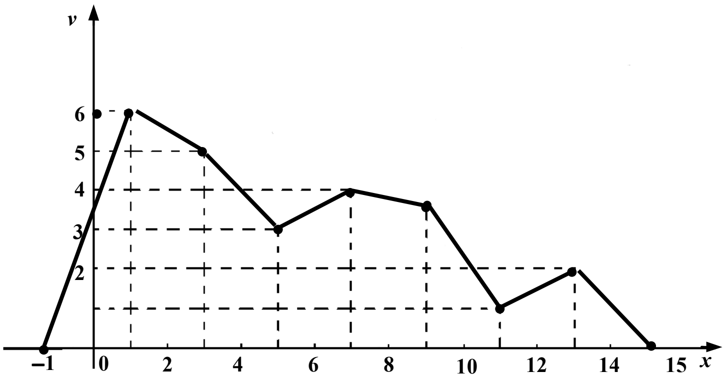
1)

ინტერ-ვალი	სიხ-შირე	სიხშირის სიმკვრივე
[10 14)	8	2
[14 18)	7	1,75
[18 22)	7	1,75
[22 26)	5	1,25
[26 30]	3	0,75



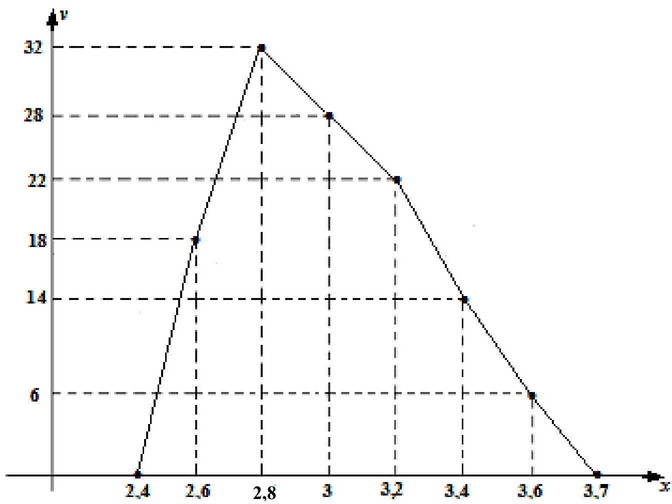
2)

ინტერვალი	სიხშირე v_i	სიხშირის სიმკვრივე	ინტერვალის შუაწერტილი
[0 2)	12	6	1
[2 4)	10	5	3
[4 6)	6	3	5
[6 8)	8	4	7
[8 10)	7	3,5	9
[10 12)	3	1,5	11
[12 14]	4	2	13



3.

ინტერვალი	[2,5 2,7)	[2,7 2,9)	[2,9 3,1)	[3,1 3,3)	[3,3 3,5)	[3,5 3,7]
სიხშირე	9	16	14	11	7	3
სიხშირის სიმკვრივე	18	32	28	22	14	6
შუა წერტილი	2,6	2,8	3	3,2	3,4	3,6



§ 8. თვისებრივი მონაცემების განაწილების ცხრილი

ამ პარაგრაფში ჩვენ შევეხებით თვისებრივ მახასიათებლებს, მათი სიხშირეებისა და ფარდობითი სიხშირეების ცხრილებს და შემდგომში განვიხილავთ მათ გრაფიკულ წარმოდგენებს.

ვთქვათ, ვაკვირდებით ობიექტთა ერთობლიობის რაიმე A თვისებრივ მახასიათებელს, რომელსაც გააჩნია K დონე (ატრიბუტი) $A_1, A_2, \ldots, A_k$. განვიხილოთ A მახასიათებლის პოპულაციიდან n მოცულობის შერჩევა.

შევადგინოთ ცხრილი, რომლის პირველ სტრიქონში ჩავწეროთ A მახასიათებლის ატრიბუტები. თითოეულ მათგანს ქვემოთ მივუწეროთ შერჩევის იმ ერთეულთა რაოდენობა, რომელთაც ეს ატრიბუტი გააჩნიათ, ანუ ამ ატრიბუტის სიხშირე. მიღებულ ცხრილს A **თვისებრივი მახასიათებლის სიხშირეთა განაწილების ცხრილი** ეწოდება, თუ სიხშირეების ნაცვლად მივუწერთ შესაბამის W_i ფარდობით სიხშირეებს მივიღებთ A **თვისებრივი მახასიათებლის ფარდობით სიხშირეთა განაწილების ცხრილს.** შეიძლება ერთ ცხრილში ორივე განაწილება გვქონდეს მოცემული (იხ. ცხრ. 8.1).

ატრიბუტები (დონეები)	A_1	A_2	...	A_k
სიხშირე v_i	v_1	v_2	...	v_k
ფარდობითი სიხშირე W_i	W_1	W_2	...	W_k

ცხრ. 8.1. თვისებრივი მახასიათებლის სიხშირეთა და ფარდობით სიხშირეთა განაწილების ცხრილი

მაგალითი 8.1. გვაქვს $n = 400$ მოცულობის შერჩევა 30-35 წლის მოსახლეობის განათლების პოპულაციიდან ატრიბუტებით: არასრული საშუალო, საშუალო, პროფესიული, არასრული უმაღლესი, უმაღლესი. სიხშირეთა განაწილების ცხრილს შეიძლება ჰქონდეს სახე.

ატრიბუტები (დონეები)	არასრული საშუალო	საშუალო	პროფე-სიული	არასრული უმაღლესი	უმაღლესი
სიხშირეები ν_i	37	62	135	74	92

ცხრ. 8.2. სიხშირეთა განაწილების ცხრილი განათლების მიხედვით

განსახილველ A მახასიათებელს შეიძლება ჰქონდეს მხოლოდ ორი დონე ანუ თითოეულ ერთეულს შეიძლება ჰქონდეს A დონე ან არ ჰქონდეს, ანუ ჰქონდეს მისი უარყოფა $\overline{A}$. სიხშირეთა და ფარდობით სიხშირეთა განაწილების ცხრილს ექნება სახე

დონე	A	$\overline{A}$
სიხშირე	ν_A	$\nu_{\overline{A}} = n - \nu_A$
ფარდ. სიხშირე	W_A	$W_{\overline{A}} = 1 - W_A$

ცხრ. 8.3. სიხშირეთა განაწილების ცხრილი ორდონიანი მახასიათებლისათვის

მაგალითი 8.2. გამოკითხულია 100 ადამიანი. მახასიათებელია ცურვის ცოდნა. გვაქვს სიხშირეთა განაწილების შემდეგი ცხრილი.

დონე	იცის ცურვა	არ იცის
სიხშირე	43	57

ცხრ. 8.4. სიხშირეთა განაწილების ცხრილი ცურვის ცოდნის მიხედვით

ერთი ნიშნის მიხედვით დაჯგუფებას პირველი რიგის დაჯგუფებას უწოდებენ. ზემოთ მოყვანილი ცხრილები ყველა პირველი რიგის დაჯგუფების ცხრილებია.

ახლა განვიხილოთ ორი A და B თვისებრივი ნიშნის მიხედვით მონაცემების სიხშირეთა განაწილების ცხრილი. ასეთ შემთხვევაში ამბობენ, რომ გვაქვს მეორე რიგის დაჯგუფება. დავუშვათ A ნიშანს აქვს k დონე $A_1, A_2, \ldots, A_k$

ხოლო B ნიშანს კი l დონე $B_1, B_2, \ldots, B_l$. სიხშირეთა განაწილების ცხრილს ექნება სახე: (ცხრ. 8.5), სადაც ν_{ij} ($1 \le i \le l$, $1 \le l \le k$) არის პოპულაციის იმ ერთეულთა რაოდენობა, რომელთაც გააჩნიათ B მახასიათებლის B_i ატრიბუტი და A მახასიათებლის A_j ატრიბუტი.

$B \backslash A$	A_1	A_2	$\ldots$	A_k	$\sum$
B_1	ν_{11}	ν_{12}	$\ldots$	ν_{1k}	ν_{B_1}
B_2	ν_{21}	ν_{22}	$\ldots$	ν_{2k}	ν_{B_2}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
B_l	ν_{l1}	ν_{l2}	$\ldots$	ν_{lk}	ν_{B_l}
$\sum$	ν_{A_1}	ν_{A_2}	$\ldots$	ν_{A_k}	n

$$\nu_{i1} + \nu_{i2} + \ldots + \nu_{ik} = \nu_{Bi} \qquad 1 \le i \le l$$

$$\nu_{1j} + \nu_{2j} + \ldots + \nu_{lj} = \nu_{Aj} \qquad 1 \le j \le k$$

ცხრ. 8.5. სიხშირეთა განაწილების ცხრილი ორი მახასიათებლის მიხედვით

ასეთ ცხრილს **ორი ნიშნის მიხედვით შეუღლების** ცხრილი ეწოდება. ცხრილის ბოლო სვეტში მოცემულია სტრიქონებში შემავალი სიხშირეების ჯამები, რომლებიც წარმოადგენენ B მახასიათებლის სიხშირეთა განაწილებას, ხოლო ბოლო სტრიქონში კი მოცემულია იმ სიხშირეების ჯამები, რომლებიც სვეტებში გვხვდება და ისინი წარმოადგენენ A მახასიათებლის სიხშირეთა განაწილებას. ცხადია

$$\sum_{i=1}^{l} \sum_{j=1}^{k} \nu_{ij} = n .$$

მაგალითი 8.3. ჩატარდა გამოკითხვა სხვადასხვა კოლეჯების სტუდენტებს შორის სპორტულ სექციებში მონაწილეობისა და უცხო ენების სწავლების კუთხით. უცხო ენების შესახებ კითხვარში იყო 4 ატრიბუტი: ინგლისური; გერმანული; ფრანგული და რომელიმე სხვა ენა, ხოლო სპორტულ სექციებში მონაწილეობის შესახებ 7

ატრიბუტი: კალათბურთი ფრენბურთი, ცურვა, ჭადრაკი, ჩოგბურთი, სპორტის სხვა რომელიმე სახეობა და სპორტის არც ერთი სახეობა. გამოკითხული იქნა 300 ახალგაზრდა. მას შემდეგ, რაც მონაცემების კლასიფიკაცია მოხდა (ზემოთ ჩამოთვლილი ატრიბუტების მიხედვით) აღმოჩნდა, რომ ორი ნიშნის მიხედვით შეუღლების ცხრილს ჰქონდა სახე:

	კალათბურთი	ფრენბურთი	ცურვა	ჭადრაკი	ჩოგბურთი	რომელიმე სხვა სახეობა	არც ერთი	ჯამი
ინგლისური	21	0	28	4	12	10	39	114
გერმანული	0	10	17	12	0	13	27	79
ფრანგული	16	5	3	0	11	0	18	53
სხვა ენა	18	5	6	11	0	4	10	54
ჯამი	55	20	54	27	23	27	94	300

ცხრ 8.6. ორი ნიშნის მიხედვით შეუღლების ცხრილი 8.3 მაგალითისათვის

§ 9. თვისებრივი მონაცემების გრაფიკული წარმოდგენის მეთოდები. მართკუთხედებიანი და წრიული დიაგრამები

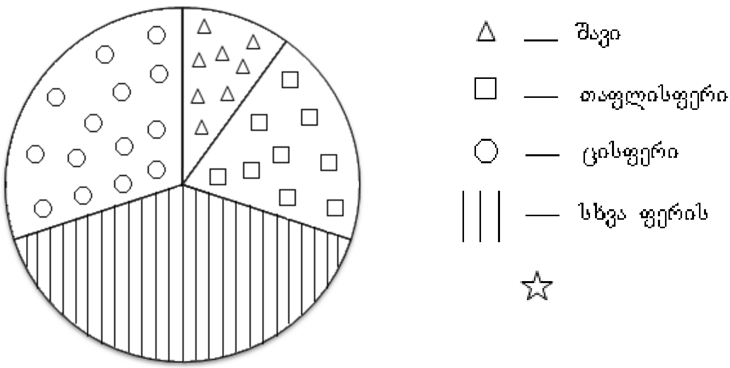
თვისებრივი მონაცემების გრაფიკულად წარმოდგენა ხდება **მართკუთხედებიანი** (ხანდახან მას სვეტოვანსაც უწოდებენ) **დიაგრამით** ან **წრიული დიაგრამით**. განვიხილოთ ორივე შემთხვევა მაგალითზე.

მაგალითი 9.1. მოცემული გვაქვს ერთ-ერთი სამოდელო სახლის მოდელი გოგონების თვალის ფერის სიხშირეთა განაწილების ცხრილი

თვალის ფერი	შავი	თაფლისფერი	ცისფერი	სხვა ფერის
სიხშირე v	5	10	20	15

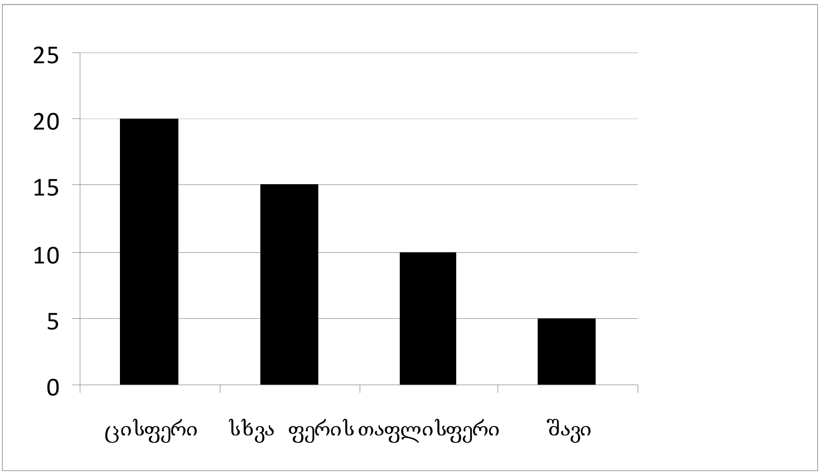
ავაგოთ სიხშირეთა განაწილების წრიული დიაგრამა.

ამოხსნა. წრიული დიაგრამის ასაგებად საჭიროა გამოვთვალოთ თითოეული სიხშირის პროცენტული მაჩვე-ნებელი. შავი ფერის არის $\frac{5}{50}$ ნაწილი ანუ 10%, თაფლის-ფერი $\frac{10}{50}$ ანუ 20%, ცისფერი $\frac{20}{50}$ ანუ 40%, ხოლო სხვა ფერის $\frac{15}{50}$ ანუ 30%. წრე დავყოთ ამ პროცენტების შესა-ბამის პროპორციულ სექტორებად. ამისთვის ცენტრალური 360° კუთხე დავყოთ პროცენტების მიხედვით. შავი – $360 \cdot \frac{10}{100} = 36°$, თაფლისფერი – $360 \cdot \frac{20}{100} = 72°$, ცისფერი – $360 \cdot \frac{40}{100} = 144°$, სხვა ფერები – $360 \cdot \frac{30}{100} = 108°$. (იხ. ნახ. 9.1).



ნახ. 9.1. წრიული დიაგრამა 9.1 მაგალითისათვის

ამავე მაგალითისათვის ავაგოთ მართკუთხედებიანი დიაგრამაც. ამისათვის აბსცისათა ღერძზე აღვნიშნოთ თვალის ფერის სახეობები და მათზე დავადგათ ტოლი სიგანის შესაბამისი სიხშირის სიმაღლის მართკუთხედები. შეგვიძლია ისინი დავალაგოთ სიმაღლეების კლების (ან ზრდის) მიხედვით (იხ. ნახ. 9.2).



ნახ. 9.2. მართკუთხედებიანი დიაგრამა 9.1 მაგალითისათვის

წრიული დიაგრამის აგების დროს სიხშირეთა პროცენტული მაჩვენებლის გამოთვლის ნაცვლად შეიძლება ფარდობით სიხშირეთა განაწილების ცხრილის გამოყენება. 9.1 მაგალითისათვის ფარდობით სიხშირეთა განაწილების ცხრილს ექნება სახე.

თვალის ფერი	შავი	თაფლისფერი	ცისფერი	სხვა ფერი
ფარდობითი სიხშირე W	$\frac{5}{50}$	$\frac{10}{50}$	$\frac{20}{50}$	$\frac{15}{50}$

წრიული დიაგრამის ასაგებად თითოეული ფერის შესაბამისი ცენტრალური კუთხე უნდა იყოს ფარდობითი სიხშირის პროპორციული. შავი ფერის თვალებისთვის $360\cdot\frac{5}{50}=36°$ და ა.შ. მივიღებთ ზუსტად 9.1 ნახაზზე გამოსახულ წრიულ დიაგრამას.

მართკუთხედებიან დიაგრამაში მართკუთხედები შეიძლება აიგოს OX ღერძის პარალელურად. ამ დროს სიხშირეები ჰორიზონტალურად განლაგდება.

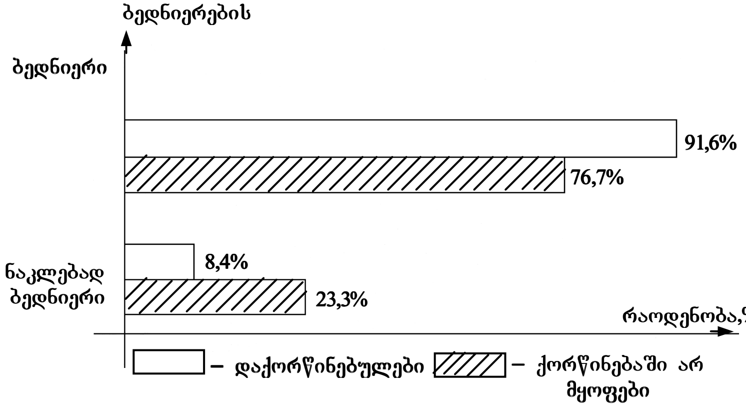
თვისებრივი მახასიათებლის მართკუთხედებიანი დიაგრამა შეიძლება აიგოს ელემენტარულ ერთეულთა ორი ან მეტი სხვადასხვა ჯგუფისათვის.

მაგალითი 9.2. იმისათვის რომ დაედგინათ არის თუ არა კავშირი ქორწინებასა და ზოგად ბედნიერებას შორის აშშ-ში მოხდა 1497 რესპონდენტის გამოკითხვა (სოციალური კვლევა: General Social Survey 2008, იხ. [21] გვ. 437). 969 დაქორ-წინებულიდან ბედნიერად თავი ჩათვალა 888-მ, ხოლო ნაკლებად ბედნიერად 81-მა. ანალოგიურ კითხვაზე ქორწი-ნებაში არ მყოფი 528 რესპონდენტიდან პასუხები იყო 405 და 123 შესაბამისად. ავაგოთ მართკუთხედებიანი დიაგრამა.

ამოხსნა. გვაქვს ელემენტარულ ერთეულთა ორი ჯგუფი დაქორწინებულები და ქორწინებაში არ მყოფები. თვისებრივი მახასიათებელია ბედნიერება ორი ატრიბუტით ბედნიერი და ნაკლებად ბედნიერი.

გამოთვლებით ვღებულობთ პროცენტულ მონაცემებს. 969 დაქორწინებულიდან ბედნიერად თავს თვლის 91,6%, ხოლო ნაკლებად ბედნიერად 8,4%. ქორწინებაში არ მყოფთათვის ეს მაჩვენებლები შესაბამისად არის 76,7% და 23,3%. *Oy* ღერძზე დავიტანოთ ატრიბურები, ხოლო პროცენტების შესაბამისი სიგრძის მართკუთხედები ავაგოთ *Ox* ღერძის პარალელურად. მივიღებთ დიაგრამას (ნახ. 9.3).

ნახ. 9.3. ბედნიერების დიაგრამა ქორწინების მიხედვით

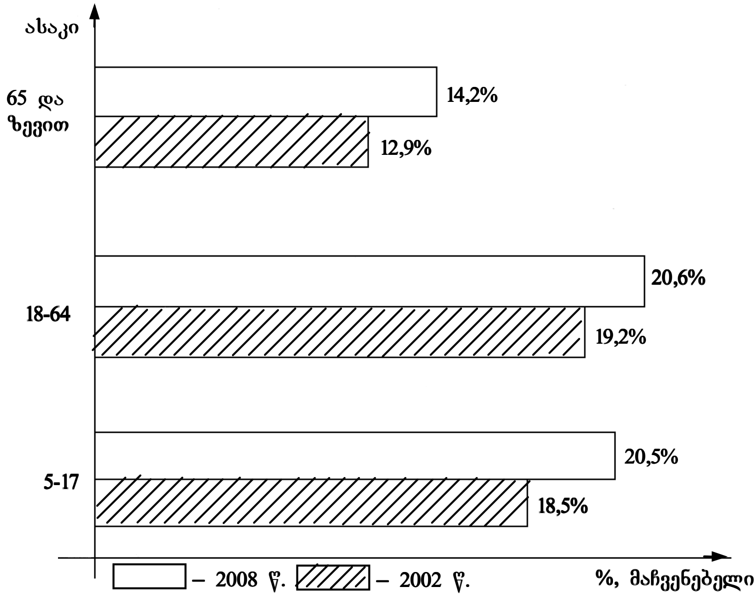


9.2 მაგალითისათვის

შეიძლება ერთი და იგივე მართკუთხედებიან დიაგრამაზე წარმოდგენილი იყოს რაიმე თვისებრივი მახასიათებელი ერთდროულად სხვადასხვა დროშიც და ელემენტარულ ერთეულთა სხვადასხვა ჯგუფებისათვისაც. ქვემოთ (ნახ. 9.4) მოყვანილ დიაგრამაზე წარმოდგენილია არა ინგლისურ ენაზე მოსაუბრე ამერიკელი მოსახლეობის ხვედრითი წილი 5 წლის ზემოთ სამ ასაკობრივ ქვეჯგუფში და დროის ორ შუალედში 2002 და 2008 წლებში (კვლევა: Community survey 2002 and 2008, იხ. [21] გვ. 104).

დიაგრამაზე თვისებრივი მახასიათებელია სასაუბრო ენა. ორი ატრიბუტით არა ინგლისური და ინგლისური (ცხადია არა ინგლისურის დაფიქსირება ინგლისურის პროცენტულ მაჩვენებელს განსაზღვრავს და ამიტომ მხოლოდ ის არის აღნიშნული). დროის შუალედი არის ორი 2002 წ.

და 2008 წ. ელემენტარულ ერთეულთა სიმრავლე დაყოფილია სამ ასაკობრივ ჯგუფად.



ნახ. 9.4. არაინგლისურ ენაზე მოსაუბრე ამერიკელთა პროცენტული მაჩვენებელი

ამოცანები:

1. ჩატარდა 50 სტუდენტის გამოკითხვა სპორტულ სექციებში მონაწილეობის კუთხით. აღმოჩნდა, რომ 5 დადის კალათბურთზე, 7 ცურვაზე, 6 – მძლეოსნობაზე, 10 სპორტის სხვა სახეობებიდან რომელიმეზე, ხოლო 22 სტუდენტი არაა სპორტით დაკავებული. ააგეთ სიხშირეთა მართკუთხედებიანი დიაგრამა.

2. მოცემულია ერთ-ერთი სადაზღვევო კომპანიის მიერ ერთი წლის განმავლობაში დაზღვეულ ავტომანქანებზე გაცემული სადაზღვევო კომპენსაციების სიხშირეთა ცხრილი.

ავტომობილის მარკა	მიცუბიში	BMW	მერსედესი	აუდი	სხვა მარკის
სიხშირე	5	12	4	8	21

ააგეთ სიხშირეთა განაწილების წრიული დიაგრამა.

3. ერთი სამუშაო კვირის განმავლობაში ელექტროტექნიკით მოვაჭრე სალონში გაიყიდა 30 ტელევიზორი, 20 მაცივარი, 15 მობილური ტელეფონი და 35 ელექტროგამათბობელი. მონაცემები წარმოადგინეთ წრიული დიაგრამის სახით.

4. მოცემულია ერთი კვირის განმავლობაში ერთ-ერთი ავტოსახელოსნოს მომსახურებით მოსარგებლე ავტომობილების სიხშირეთა ცხრილი.

ავტომობილი	BMW	ოპელი	ფორდი	მერსედესი	ჰონდა	სხვა ფირმები
სიხშირე	25	18	13	n	12	30

გამოთვალეთ n და ააგეთ მოცემული ცხრილით სიხშირეთა მართკუთხედებიანი დიაგრამა თუ ცნობილია, რომ ავტოსახელოსნოს მომსახურებით მოსარგებლე ოპელის ფირმის ავტომობილების ფარდობითი სიხშირე $\frac{1}{20}$-ით მეტია, ვიდრე ჰონდას ფირმის ავტომობილების.

5. ერთ-ერთი რესტორნების ქსელის მომსახურე ოფიციანტებიდან 10-ს აქვს შავი ფერის თმა, 4-ს თეთრი, 12-ს ყვითელი, 8-ს ჩალისფერი, ხოლო 5-ს წითელი. ააგეთ თმის ფერის შესაბამის სიხშირეთა მართკუთხედებიანი დიაგრამა.

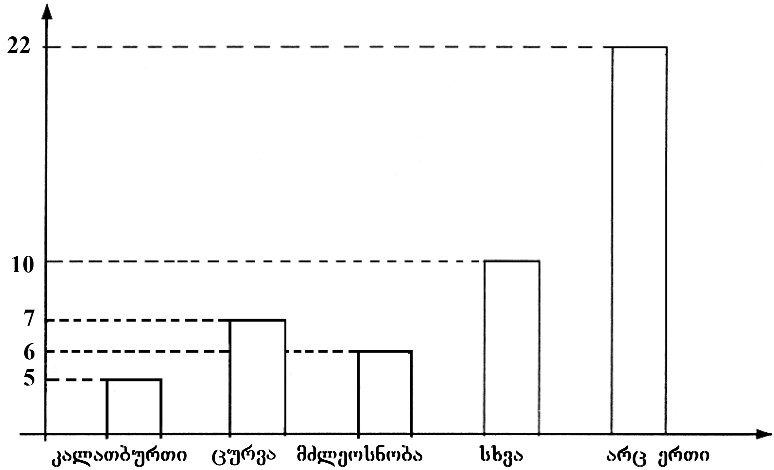
6. ჩატარდა 40 წელზე მეტი ასაკის 1341 ამერიკელის გამოკითხვა ჯანმრთელობის მდგომარეობის შესახებ. კლასიფიკაცია მოხდა ორი ატრიბუტით. I – საშუალო ან ცუდი და II – კარგი ან შესანიშნავი. გამოკითხულთა შორის 688 იყო დაბალი და მშრომელი კლასის წარმომადგენელი (რომლებიც I და II კატეგორიაში განაწილდნენ შესაბამისად 281 და 407) და 653 საშუალო და მაღალი კლასის წარმომადგენელი (რომლებიც განაწილდნენ შესაბამისად 134 და

519) (კვლევა: General Social Survey 2008). ([21] გვ. 437). ააგეთ მართკუთხედებიანი დიაგრამა ორივე კლასის წარმომად-გენლებისათვის (პროცენტული მაჩვენებლებით).

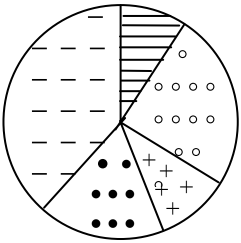
7. გამოკითხეთ თქვენი ნაცნობი პიროვნებებიდან რომელ უცხო ენას რამდენი სწავლობს (ან ნასწავლი აქვს) და ააგეთ სიხშირეთა და ფარდობით სიხშირეთა წრიული დიაგრამები (შეიძლება ჩატარდეს ჯგუფში დამოუკიდებელი სამუშაო).

პასუხები

1.

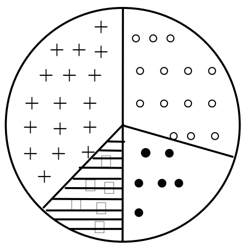


2.



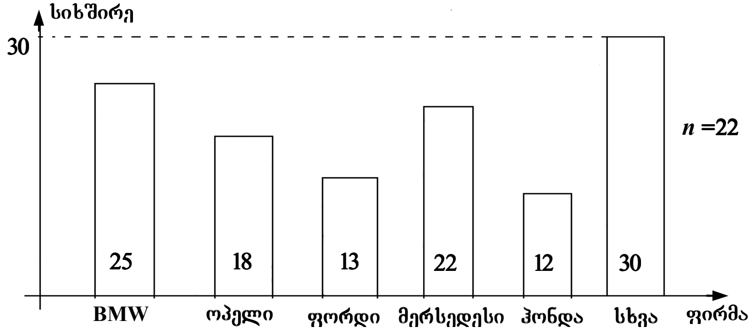
≡ – მიცუბიში
◦ – BMW
\+ – მერსედესი
• – აუდი
— – სხვა მარკის

3.

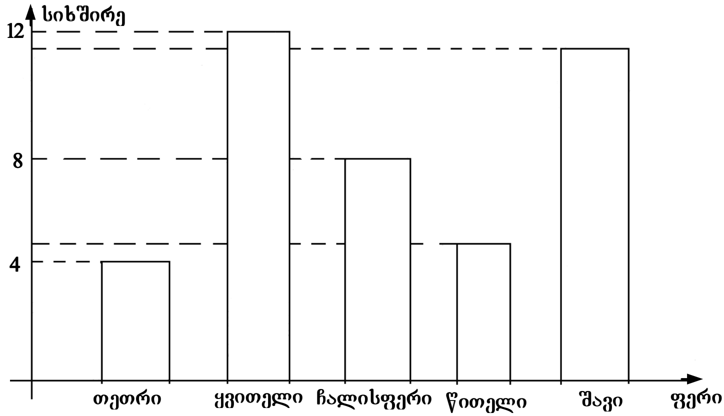


◦ – ტელევიზორი
• – მაცივარი
≡ – მობილური
\+ – გამათბობელი

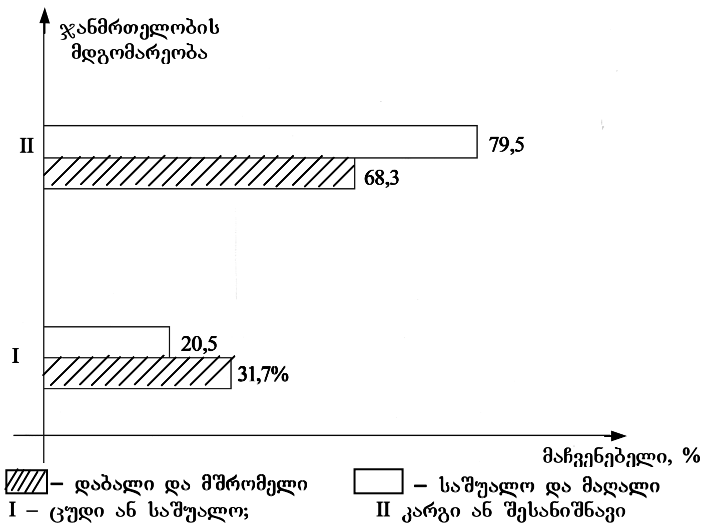
4.


სიხშირე
30
n =22
25
18
13
22
12
30
BMW
ოპელი
ფორდი
მერსედესი
ჰონდა
სხვა
ფირმა

5.


სიხშირე
12
8
4
თეთრი
ყვითელი
ჩალისფერი
წითელი
შავი
ფერი

6.


ჯანმრთელობის მდგომარეობა
II
79,5
68,3
I
20,5
31,7%
მაჩვენებელი, %
– დაბალი და მშრომელი
– საშუალო და მაღალი
I – ცუდი ან საშუალო;
II კარგი ან შესანიშნავი

§ 10. შერჩევის რიცხვითი მახასიათებლები. შერჩევითი საშუალო და შერჩევითი დისპერსია

აღწერით სტატისტიკაში მიღებული ნედლი მონაცემების დამუშავებისას მათი გრაფიკული წარმოდგენების გარდა ხდება ამ მონაცემებისათვის გარკვეული რიცხვითი მახასიათებლების გამოთვლა, რომლებიც უმარტივესი სტატისტიკური დასკვნების გაკეთების საშუალებას იძლევიან.

განსაზღვრება 10.1. $x_1, x_2, \ldots, x_n$ შერჩევის მონაცემების **შერჩევითი საშუალო** აღინიშნება $\overline{X}_n$ სიმბოლოთი და ეწოდება მის არითმეტიკულ საშუალოს

$$\overline{X}_n = \frac{x_1 + x_2 + \ldots + x_n}{n}. \tag{10.1}$$

თუ მოცემულია მონაცემების სიხშირეთა და ფარდობით სიხშირეთა განაწილების ცხრილები (იხ. ცხრ. 4.1), მაშინ

$$\overline{X}_n = \frac{v_1 x_{(1)} + v_2 x_{(2)} + \ldots + v_k x_{(k)}}{n} = w_1 x_{(1)} + w_2 x_{(2)} + \ldots + w_k x_{(k)}.$$

შერჩევითი საშუალო გვიჩვენებს მონაცემების „ცენტრს“ ანუ იმ წერტილს, რომლის ირგვლივაც სხვა წერტილებთან შედარებით ყველაზე “ახლოს” არიან თავმოყრილი მთელი მონაცემები. შემდგომში ჩვენ ინდექსს n აღარ მივუთითებთ თუ განსხვავებული მოცულობების შერჩევებზე არ გვექნება საუბარი.

მაგალითი 10.1. მოცემულია ბენზინგასამართი სადგურის შემოსავლები 10 დღის განმავლობაში (ასეული ლარებით) 6, 12, 5, 8, 9, 5, 9, 8, 9, 9. ვიპოვოთ საშუალო დღიური შემოსავალი ამ 10 დღეში ანუ შერჩევითი საშუალო.

$$\overline{x} = \frac{5\cdot 2 + 6\cdot 1 + 8\cdot 2 + 9\cdot 4 + 12\cdot 1}{10} = 8.$$

როდესაც მოცემული გვაქვს ინტერვალებად დაჯგუფებული მონაცემების შემთხვევაში ინტერვალების შუა წერტილებს მიეწერება ამ ინტერვალების სიხშირეები და შემდეგ ითვლება შერჩევითი საშუალო. ცხადია შედეგი მიახლოებითი იქნება ვინაიდან ინტერვალების შეცვლა შუა წერტილებით იწვევს შეცდომების დაგროვებას.

მაგალითი 10.2. მოცემული გვაქვს მსხვილფეხა რქოსანი პირუტყვის ფერმების სიხშირეთა განაწილების ცხრილი პირუტყვის სულადობის მიხედვით (იხ. ცხრ. 10.1). ვიპოვოთ ფერმაში პირუტყვის სულადობის საშუალო მაჩვენებელი.

სულადობა	ნაკლები 50-ზე	[50 70)	[70 100)	[100 150]	მეტი 150-ზე
სიხშირე v_i	5	2	10	5	3

ცხრ. 10.1.

ამოხსნა. ყოველ ინტერვალს შევუსაბამოთ შუა წერტილი. კიდურა ინტერვალების სიგრძეებად ავიღოთ მათი მოსაზღვრე უახლოესი ინტერვალის სიგრძეები და შემდეგ დავაფიქსიროთ მათი შუა წერტილები. მარცხენა კიდურა ინტერვალის სიგრძედ ავიღებთ მისი მეზობელი მარჯვენა ინტერვალის სიგრძეს (70-50=20) 20-ს და ინტერვალი იქნება [30 50), ხოლო შუა წერტილი კი 40. მარჯვენა კიდურა ინტერვალის სიგრძედ ავიღოთ მისი მეზობელი მარცხენა ინტერვალის სიგრძე 150-100=50 და ბოლო კიდურა ინტერვალი იქნება [150 200], ხოლო შუა წერტილი 175. საბოლოოდ გვექნება ცხრილი.

სულადობა	40	60	85	125	175
სიხშირე	5	2	10	5	3

საიდანაც

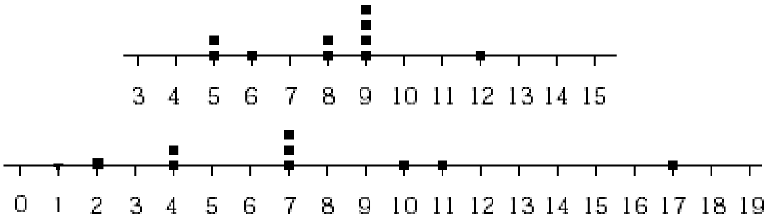
$$\bar{x} = \frac{40 \cdot 5 + 60 \cdot 2 + 85 \cdot 10 + 125 \cdot 5 + 175 \cdot 3}{25} = 92{,}8 .$$

შერჩევითი საშუალო ერთ-ერთი მნიშვნელოვანი რიცხვითი მახასიათებელია, მაგრამ ვერ გვაძლევს სრულ ინფორმაციას შერჩევის მონაცემების შესახებ. 10.1 მაგალითში თუ მეორე შერჩევა იქნება

$$2,\ 7,\ 11,\ 17,\ 4,\ 4,\ 11,\ 7,\ 7,\ 10$$

შერჩევითი საშუალო იგივე დარჩება $\overline{X} = 8$, მაგრამ აშკარაა რომ ეს უკანასკნელი მონაცემები განსხვავდებიან მაგალითში განხილული შერჩევის მონაცემებისაგან. ეს განსხვავება რომ კარგად გამოჩნდეს შერჩევის მონაცემები

დავიტანოთ რიცხვით ღერძზე ერთმანეთის ქვეშ ისე, რომ ღერძის თითოეულ დანაყოფზე ზემოდან იმდენი წერტილი დავსვათ რისი ტოლიცაა ამ რიცხვის შესაბამისი სიხშირე მონაცემებში. შევადაროთ ორივე შემთხვევისათვის მიღებული წარმოდგენები (იხ. [1]):



პირველი რაც თვალში გვხვდება ესაა სხვადასხვა “გაფანტულობა” ანუ გაბნევა შერჩევითი საშუალოს ირგვლივ, ამიტომ ბუნებრივია შემოვიტანოთ ამ “გაფანტულობის” მახასიათებელი.

განსაზღვრება 10.2. $x_1, x_2, \ldots, x_n$ **შერჩევის გაბნევის დიაპაზონი** აღინიშნება d სიმბოლოთი და ეწოდება შერჩევის უდიდესი და უმცირესი ელემენტის სხვაობას, ანუ შესაბამისი ვარიაციული მწკრივის “სიგრძეს” მას ხშირად **ვარიაციის ამპლიტუდასაც** უწოდებენ.

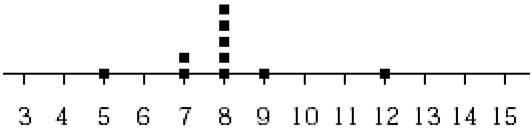
$$d = \max_{1\le i\le n} x_i - \min_{1\le i\le n} x_i = x_{(n)} - x_{(1)} .$$

გაბნევის დიაპაზონი (ჩვენს მიერ შესადარებელ ორ შემთხვევაში) აშკარად მეტია მეორე ვარიანტში.

უნდა აღინიშნოს გაბნევის დიაპაზონის ერთი ნაკლიც. ის ვერ ასახავს შუალედური (უდიდეს და უმცირეს მონაცემებს შორის მოქცეული) მონაცემების “ყოფაქცევას”, ანუ იმას, თუ სად კონცენტრირდებიან ეს მონაცემები.

მაგალითად, განვიხილოთ ისევ 10.1 მაგალითის შემთხვევაში სხვა დღეებში ჩატარებული დაკვირვებებით მიღებული იგივე $n = 10$ მოცულობის შერჩევა (ანუ მესამე შერჩევა) 5, 12, 9, 7, 8, 8, 8, 8, 7, 8.

ცხადია, ამ შემთხვებაშიც, $\bar{x} = 8$ და $d = 12 - 5 = 7$ ისევე როგორც პირველ შემთხვევაში, მაგრამ მესამე შერჩევის მონაცემები უფრო ახლოსაა თავმოყრილი მისი $\bar{X}$ საშუალოს ირგვლივ ვიდრე პირველი შერჩევის მონაცემები.



საშუალოს ირგვლივ გაფანტულობის ერთ-ერთი კარგი მაჩვენებელია შერჩევითი დისპერსია.

განსაზღვრება 10.3. $x_1, x_2, \ldots, x_n$ შერჩევის **შერჩევითი დისპერსია** აღინიშნება S_n^2-ით და წარმოადგენს შერჩევის მონაცემების საშუალოდან გადახრების კვადრატების საშუალოს

$$S_n^2 = \frac{1}{n}\sum_{k=1}^{n}(x_i - \overline{x})^2 = \frac{(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \ldots + (x_n - \overline{x})^2}{n}. \quad (10.2)$$

შერჩევითი დისპერსიიდან კვადრატულ ფესვს S_n-ს უწოდებენ **სტანდარტულ გადახრას** ან **შერჩევით საშუალო კვადრატულ გადახრას**.

მარტივად შეიძლება იმის ჩვენება, რომ

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}x_i^2 - \overline{x}^2 .$$

ცხადია, თუ გავითვალისწინებთ შერჩევის მონაცემების სიხშირეთა და ფარდობით სიხშირეთა განაწილების ცხრილს, გვექნება

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{k}(x_{(i)} - \overline{x})^2 v_i = \sum_{i=1}^{k}(x_{(i)} - \overline{x})^2 W_i .$$

გამოვთვალოთ 10.1 მაგალითში შერჩევითი დისპერსია სამივე შერჩევისათვის

I შერჩევისათვის:

$$S_{10}^2 = \frac{1}{10}[(6-8)^2 + (12-8)^2 + (5-8)^2 + (8-8)^2 + (9-8)^2 +$$

$$+(5-8)^2 + (9-8)^2 + (8-8)^2 + (9-8)^2 + (9-8)^2] = 4,2 .$$

II შერჩევისათვის:

$$S_{10}^2 = \frac{1}{10}[(2-8)^2 + (7-8)^2 + (11-8)^2 + (17-8)^2 + (4-8)^2 +$$

$$+ (4-8)^2 + (11-8)^2 + (7-8)^2 + (7-8)^2 + (10-8)^2] = 17,4 .$$

III შერჩევისათვის:

$$S_{10}^2=\frac{1}{10}[(5-8)^2+(12-8)^2+(9-8)^2+(7-8)^2+(8-8)^2+$$
$$+(8-8)^2+(8-8)^2+(8-8)^2+(7-8)^2+(8-8)^2]=2{,}8\,.$$

ყველაზე მცირე შერჩევითი დისპერსია გააჩნია მესამე შერჩევას (როგორც ჩანდა კიდეც რიცხვით ღერძზე წარმოდგენილი მონაცემებიდან).

თუ მთელი პოპულაციაა $X_1,X_2,\ldots,X_N$, ხოლო შერჩევა $x_1,x_2,\ldots,x_n$, მაშინ სიდიდეებს.

$$\mu=\frac{1}{N}\sum_{i=1}^{N}X_i\,,\quad \sigma^2=\frac{1}{N}\sum_{i=1}^{N}(X_i-\mu)^2 \qquad (10.3)$$

ეწოდებათ **პოპულაციის საშუალო** და **პოპულაციის დისპერსია** შესაბამისად.

როდესაც პოპულაციის საშუალო μ ცნობილია, მაშინ **შერჩევითი დისპერსია**

$$\bar{S}_n^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$$

ფორმულით გამოითვლება. შევნიშნოთ, რომ როდესაც n უახლოვდება N-ს, მაშინ შერჩევითი საშუალო და შერჩევითი დისპერსია პოპულაციის საშუალოსა და პოპულაციის დისპერსიას უახლოვდებიან, ამიტომ სტატისტიკური კვლევების ჩატარების დროს შერჩევითი საშუალო და შერჩევითი დისპერსია პოპულაციის საშუალოსა და დისპერსიის „შეფასებებად" გამოდგებიან.

შევნიშნოთ, რომ (10.2) და (10.3) ფორმულებით მოცემული შერჩევითი დისპერსიისა და პოპულაციის დისპერსიის ნაცვლად სტატისტიკურ კვლევებში სარგებლობენ ე.წ. **შესწორებული შერჩევითი დისპერსიისა** და **შესწორებული პოპულაციის დისპერსიის** ფორმულებით (ამ საკითხს ჩვენ § 27-ში დავუბრუნდებით)

$$\begin{aligned} S_n'^2&=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2\equiv\frac{n}{n-1}S_n^2,\\ \sigma'^2&=\frac{1}{N-1}\sum_{i=1}^{N}(X_i-\mu)^2\equiv\frac{N}{N-1}\sigma^2. \end{aligned} \qquad (10.4)$$

ჩამოვთვალოთ შერჩევითი საშუალოს თვისებები (რომელთა დამტკიცებაც მკითხველს თავად შეუძლია ჩაატაროს):

1) შერჩევითი საშუალოდან შერჩევის მონაცემების გადახრების ჯამი ნულის ტოლია. ე. ი.
$$\sum_{i=1}^{n}(x_i-\bar{x})=0 .$$
2) შერჩევითი საშუალოდან შერჩევის მონაცემების გადახრების კვადრატების ჯამი მინიმალურია. ნებისმიერი a რიცხვისათვის
$$\sum_{i=1}^{n}(x_i-\bar{x})^2\le\sum_{i=1}^{n}(x_i-a)^2 .$$
3) თუ $x_1,x_2,\ldots,x_n$ მონაცემებიდან მივიღებთ ახალ $y_i=ax_i+b$; $i=\overline{1,n}$ მონაცემებს, მაშინ
$$\bar{y}=a\bar{x}+b .$$
4) თუ ორი სხვადასხვა მოცულობის $x_1,x_2,\ldots,x_n$ და $y_1,y_2,\ldots,y_m$ შერჩევების საშუალოებია შესაბამისად $\bar{x}$ და $\bar{y}$, მაშინ $x_1,x_2,\ldots,x_n$, $y_1,y_2,\ldots,y_m$ გაერთიანებული შერჩევის საშუალო იქნება ამ საშუალოების შეწონილი ჯამი
$$\frac{m}{n+m}\bar{x}+\frac{m}{n+m}\bar{y} .$$
5) თუ $x_1,x_2,\ldots,x_n$ და $y_1,y_2,\ldots,y_n$ ორი ერთნაირი მოცულობის შერჩევაა და შევქმნით ახალ $z_i=x_i\pm y_i$ $i=\overline{1,n}$ შერჩევას, მაშინ
$$\bar{z}=\bar{x}\pm\bar{y} .$$

მარტივად შეიძლება შევამოწმოთ შეჩევითი დისპერსიის შემდეგი თვისებებიც:

1. თუ $y_i=ax_i+b$ $(i=\overline{1,n})$, ხოლო S_{nx}^2, S_{ny}^2, $S_{nx}'^2$ და $S_{ny}'^2$ კი x_i, y_i $i=\overline{1,n}$ შერჩევების დისპერსიები და შესწორებული დისპერსიებია შესაბამისად, მაშინ
$$S_{ny}^2=a^2S_{nx}^2 ,\quad S_{ny}'^2=a^2S_{nx}'^2 .$$

2. თუ $x_1, x_2, \ldots, x_n$ და $y_1, y_2, \ldots, y_m$ ორი შერჩევაა მაშინ გართიანებული შერჩევისათვის

$$S^2_{(n+m)xy} = \frac{n}{n+m} S^2_{nx} + \frac{m}{n+m} S^2_{my} + \frac{mn}{(n+m)^2} (\overline{x}_n - \overline{y}_m)^2,$$

$$S'^2_{(n+m)xy} = \frac{n-1}{n+m} S'^2_{nx} + \frac{m-1}{n+m} S'^2_{my} + \frac{mn}{(n+m-1)(n+m)} (\overline{x}_n - \overline{y}_m)^2, \qquad 10.5$$

სადაც $S^2_{(n+m)xy}$ და $S'^2_{(n+m)xy}$ გაერთიანებული შერჩევის შერჩევითი დისპერია და შესწორებული შერჩევითი დისპერსიებია შესაბამისად.

შემოვიღოთ შერჩევის კიდევ რამოდენიმე რიცხვითი მახასიათებელი, რომელთაც შემდეგ პარაგრაფში შერჩევის სტატისტიკური განაწილების ტიპების დადგენისას გამოვიყენებთ.

განსაზღვრება 10.4. შერჩევითი ვარიაციის (შესწორებული შერჩევითი ვარიაციის) კოეფიციენტი უგანზომილებო სიდიდეა, აღინიშნება CV (შესაბამისად, CV') სიმბოლოთი და განისაზღვრება როდესაც $\overline{x} \neq 0$ ფორმულით

$$CV = \frac{s}{\overline{x}} \quad \left(CV' = \frac{s'}{\overline{x}} \right). \qquad (10.6)$$

შერჩევითი ვარიაციის კოეფიციენტი გამოიყენება მონაცემთა ორი სიმრავლის გაფანტულობის შესადარებლად. იგი საშუალოთა განსხვავების გამორიცხვის საშუალებას იძლევა. განვიხილოთ ორი შერჩევა თითოეული $n = 8$ მოცულობის:

A – 7, 10, 10, 7, 8, 12, 13, 13

B – 100, 110, 108, 114, 119, 109, 120, 100.

ცხადია, გვაქვს

$$d_A = 13 - 7 = 6; \;\; \overline{x}_A = 10;$$

$$S_A^2 = \frac{1}{8}(7-10)^2 + (10-10)^2 + (10-10)^2 + (7-10)^2 + (8-10)^2 +$$
$$+(12-10)^2 + (13-10)^2 + (13-10)^2] = 5{,}5;$$

$$S_A \approx 2{,}345\ ;\ \ CV_A = \frac{s_A}{\overline{x}_A} \approx 0{,}2345\,,$$

$$d_B = 120 - 100 = 20\,; \qquad \overline{x}_B = 110\,.$$

$$S_B^2 = \frac{1}{8}[(100-110)^2 + (110-110)^2 + (108-110)^2 + (114-110)^2 +$$
$$+ (119-110)^2 + (109-110)^2 + (120-110)^2 + (100-110)] = 50{,}25\,;$$

$$S_B \approx 7{,}088\ ;\ \ CV_B = \frac{s_B}{\overline{x}_B} = \frac{7{,}088}{110} \approx 0{,}064\,.$$

როგორც ვხედავთ, B შერჩევის სტანდარტული გადახრა მეტია, ვიდრე A შერჩევის, ხოლო ვარიაციის კოეფიციენტი კი, პირიქით, ნაკლებია. ე. ი.

$$CV_B = 0{,}064 < 0{,}23 = CV_A\,.$$

ინტუიციური მოსაზრებიდან გამომდინარე ეს ასეც უნდა იყოს, ვინაიდან 7-8 ერთეულით სტანდარტული გადახრა ასეულებით გამოსახული რიცხვებისათვის უფრო უმნიშვნელოა, ვიდრე 2-3 ერთეულით სტანდარტული გადახრა ათეულებით გამოსახული რიცხვებისათვის, ისევე როგორც ფარდობითი ცდომილების შემთხვევაში. მაგალითად 500კმ სიგრძის რკინიგზის ლიანდაგის სიგრძის გაზომვისას 100-200 მეტრი ცდომილება უფრო უმნიშვნელოა ვიდრე 2 მეტრი სიგრძის დეტალის გაზომვისას 10-20 სანტიმეტრი სიგრძის ცდომილება.

განსაზღვრება 10.5. **ასიმეტრიის შერჩევითი კოეფიციენტი** ეწოდება სიდიდეს

$$\hat{a}_n = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^3}{S_n^3}\,, \tag{10.7}$$

ხოლო **ექსცესის შერჩევითი კოეფიციენტი** სიდიდეს

$$\hat{e}_n = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^4}{S_n^4} - 3\,. \tag{10.8}$$

ამ კოეფიციენტების მნიშვნელობას შემდეგ პარაგრაფში გავარკვევთ.

ამოცანები

1. მოცემულია $n=10$ მოცულობის შერჩევა

12, 15, 11, 10, 11, 15, 17, 11, 12, 9

გამოთვალეთ $\overline{X}$, S^2, S'^2, CV და CV' რიცხვითი მახასიათებლები.

2. ცნობილია, რომ „GILBERT“-ის ფირმის გამოშვებული რაგბის ბურთის სტანდარტული წონაა 435 გ. მოცემულია ფირმის გამოშვებული 16 ბურთის წონა. გამოთვალეთ $\overline{X}$, $\overline{S}^2$ და CV, შერჩევითი საშუალო, შერჩევითი დისპერსია და შერჩევითი ვარიაციის კოეფიციენტი

424	417	440	439	445	441	418	453
437	428	434	427	439	447	435	432

3. ცნობილია ერთი და იგივე პოპულაციიდან სხვადასხვა დროს ჩატარებული ორი შერჩევის შერჩევითი საშუალოები $\overline{X}_{10,x}=8{,}5$; $\overline{X}_{20,y}=9{,}7$ და შერჩევითი დისპერსიები $S^2_{10,x}=4$ და $S^2_{20,y}=3{,}2$. გამოთვალეთ ამ ორი შერჩევის გაერთიანებით მიღებული შერჩევის შერჩევითი საშუალო და შერჩევითი დისპერსია $\overline{X}_{30,x,y}$ და $S^2_{30,x,y}$.

4. მენეჯერს აქვს 30 დღის განმავლობაში სუპერმარკეტში ნავაჭრი თანხების (ლარებში) დღიური მონაცემები. გამოთვალეთ გაბნევის დიაპაზონი, შერჩევითი საშუალო, შესწორებული შერჩევითი დისპერსია და შესწორებული შერჩევითი ვარიაციის კოეფიციენტი.

1730	1447	1480	1610	1578	1581	1679	1623
1518	1635	1709	1683	1701	1647	1609	1597
1483	1517	1541	1618	1644	1638	1615	1703
1681	1658	1663	1648	1639	1571		

5. გამოთვალეთ შერჩევითი საშუალო, შესწორებული სტანდარტული გადახრა, შესწორებული შერჩევითი ვარიაციის კოეფიციენტი, შერჩევითი ასიმეტრიისა და შერჩევითი ექსცესის კოეფიციენტები §6-ის ამოცანებიდან №2 ამოცანაში მოყვანილი გაყიდული ყავის წონებისათვის.

6. გამოთვალეთ მიკროსაფინანსო ორგანიზაციის მიერ ყოველთვიურად გაცემული სესხებისათვის იგივე მახასიათებლები, რაც წინა მაგალითში. გამოიყენეთ შერჩევა § 6-ის ამოცანების №3 ამოცანიდან.

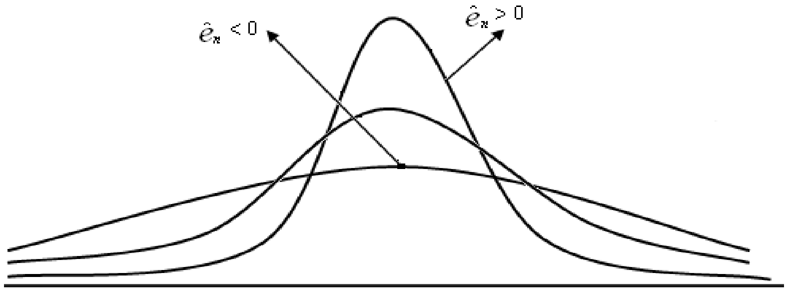
პასუხები

1) 12,3; 5,81; 6,4556; 0,196; 0,2066. 2) 434,75; 95,125; 0,0224.
3) 9,3; 3,7866. 4) 283; 1614,8667; 5247,223; 0,0449.
5) 3015; 134,6451; 0,0447; 1,355; -1,0141. 6) 23923,168; 1799,1721; 0,0752; 1,4512; -0,626.

§ 11. განაწილებათა ტიპები

შერჩევის განაწილების სახის დადგენა შეგვიძლია მისი ჰისტოგრამის ან პოლიგონის მიხედვით. რაც უფრო დიდია შერჩევის მოცულობა და მცირეა დაჯგუფების ინტერვალების სიგრძეები მით უფრო ემსგავსება ჰისტოგრამა გლუვ წირს. შერჩევის განაწილების მიხედვით შეგვიძლია წარმოდგენა ვიქონიოთ პოპულაციის განაწილებაზე. გავეცნოთ განაწილებათა რამოდენიმე ძირითად ტიპს.

I. სიმეტრიული განაწილება – სიმეტრიული განაწილება ისეთი განაწილებაა, რომლის კიდურა მნიშვნელობები უფრო იშვიათად გვხვდება, ვიდრე შუა, თანაც ცენტრის (ანუ $\overline{X}$-ის) მიმართ ორივე მხარეს მონაცემები სიმეტრიულადაა განლაგებული (ცხადია გარკვეული მცირე გადახრებით ვინაიდან აბსოლუტურად სიმეტრიული განაწილების მოძებნა ძნელია). მათ ჰისტოგრამებს დაახლოებით აქვთ სახე.

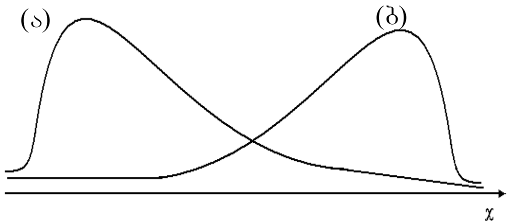


ასეთ განაწილებებზე როდესაც ისინი ძალიან არ არიან გაწელილი წვეროს მხარეს, ამბობენ რომ მათ **ზარისებური** ფორმა აქვთ.

ექსცესის კოეფიციენტი $\hat{e}_n$ არის სიხშირეთა ჰისტოგრამის წვეროს ზემოთ გაწელილობის მაჩვენებელი. როცა $\hat{e}_n > 0$, მაშინ წვერო ზემოთაა აწეული და $\hat{e}_n$ ზრდასთან ერთად წვერო იწევს ზემოთ. ხოლო როცა $\hat{e}_n < 0$ წვერო

არაა ზემოთ აწეული და მონაცემები მდორედ იცვლებიან ცენტრის ($\overline{X}$-ის) ირგვლივ.

II. უმნიშვნელოდ ასიმეტრიული განაწილებები. უმნიშვნელოდ ასიმეტრიული განაწილება ეწოდება ისეთ განაწილებას, რომელიც ასიმეტრიულია წვეროს შესაბამისი ინტერვალის ირგვლივ. **მარჯვენა ასიმეტრიული** ეწოდება თუ მაქსიმალური სიხშირის ინტერვალის მარცხნივ მნიშვნელობების სიხშირეები უფრო მცირეა, ვიდრე შესაბამისი მანძილის მოშორებით მარჯვნივ აღებული მნიშვნელობების სიხშირეები და ამის გამო მარჯვენა მხარეს განაწილებას აქვს კუდი (იხ. ნახ. 11.1, ა). შესაბამისი ასიმეტრიის შერჩევითი კოეფიციენტი $\hat{a}_n > 0$. **განაწილებას მარცხენა ასიმეტრიული** ეწოდება თუ პირიქით ხდება და მარჯვენა მხარეს კლებულობს შესაბამისი მონაცემების სიხშირეები უფრო სწრაფად, ვიდრე მარცხენა მხარეს. შესაბამისად განაწილებას კუდი აქვს მარცხნივ (იხ. ნახ. 11.1, ბ). შესაბამისი ასიმეტრიის კოეფიციენტი $\hat{a}_n < 0$.



ნახ. 11.1. უმნიშვნელოდ მარჯვენა (ა), მარცხენა (ბ) ასიმეტრიული განაწილებები

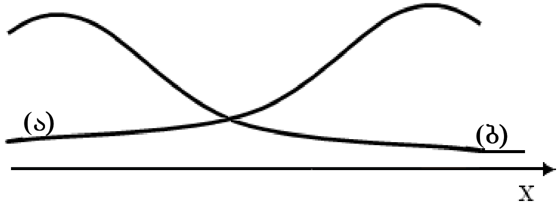
დავუბრუნდეთ მე-5 პარაგრაფში მოყვანილ 5.1 მაგალითს $k = 10$ რაოდენობის ინტერვალებად დაყოფის შემთხვევაში (იხ. ნახ. 5.2). ცხადია, ეს განაწილება არის უმნიშვნელოდ მარჯვენა ასიმეტრიული.

III. მნიშვნელოვნად ასიმეტრიული განაწილებები. განაწილებას ეწოდება **მნიშვნელოვნად ასიმეტრიული** თუ მისი ინტერვალური სიხშირეები მნიშვნელოვნად იზრდება მხოლოდ ერთ მხარეს. ე.ი. შესაბამისად კლებულობს მხოლოდ ერთ მხარეს. ასე მაგალითად თუ ინტერვალური

სიხშირეები იზრდება მარცხნიდან მარჯვნივ გვაქვს მნიშვნელოვნად მარცხენა ასიმეტრია (ნახ. 11.2, ა), ხოლო თუ სიხშირეები იზრდებიან მარჯვნიდან მარცხნივ, მაშინ გვაქვს მნიშვნელოვნად მარჯვენა ასიმეტრია (ნახ. 11.2, ბ).

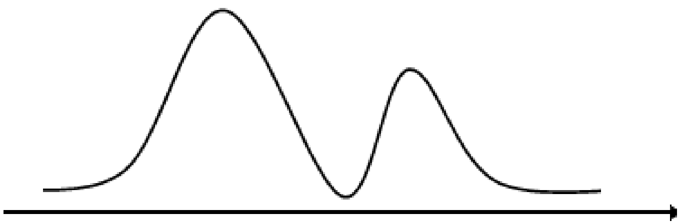
მნიშვნელოვნად მარცხენა ასიმეტრიული განაწილების მაგალითია § 7-ში განხილული 7.2 მაგალითის შესაბამისი ჰისტოგრამა (იხ. ნახ. 7.3).

შევნიშნოთ, რომ ზოგიერთ ლიტერატურაში მარჯვენა ასიმეტრიულ განაწილებას **დადებითად ასიმეტრიულს** უწოდებენ, ხოლო მარცხენა ასიმეტრიულს კი **უარყოფითად ასიმეტრიულს.**



ნახ. 11.2. მნიშვნელოვნად მარცხენა (ა) მარჯვენა (ბ) ასიმეტრიული განაწილების ფორმები

IV. განაწილებები, რომელთაც რამოდენიმე წვერო აქვთ. ასეთი განაწილებები ისეთი პოპულაციის შერჩევებისთვის მიიღება რომელთა ელემენტარული ერთეულები თვისებრივად განსხვავებული ორი ან მეტი ჯგუფისაგან შედგება. (იხ. ნახ. 11.3). მაგალითად, როდესაც ვიხილავთ რომელიმე სოფლის მსხვილფეხა რქოსანი პირუტყვის წონების პოპულაციას მრუდს ექნება ორი წვერი, ვინაიდან მდედრი და მამრი წარმომადგენლების წონები მნიშვნელოვნადაა განსხვავებული და მათ სხვადასხვა საშუალოები გააჩნიათ.



ნახ. 11.3. ორი წვეროს მქონე განაწილების მრუდის მაგალითი

ასეთი ტიპის განაწილება აქვს 7.1 მაგალითის მონაცემებს (იხ. ნახ. 7.2)

V. განაწილებები, რომელთა შესაბამის მრუდებს U-სებრი ფორმა აქვთ. ასეთი განაწილებები სიმეტრიულია და კიდურა მნიშვნელობები უფრო ხშირად გვხვდება ვიდრე შუა მნიშვნელობები.

ამოცანები

1. როგორი განაწილება აქვს § 10-ის ამოცანების № 2 ამოცანაში მოყვანილ შერჩევას?
2. როგორი განაწილება აქვს § 7-ის ამოცანების № 3 ამოცანაში მოყვანილ შერჩევას?
3. როგორი განაწილება აქვს § 5-ის ამოცანების № 1 ამოცანაში მოყვანილ შერჩევას?
4. როგორი განაწილება აქვს § 5-ის ამოცანების № 2 ამოცანაში მოყვანილ შერჩევას?
5. როგორი სახის განაწილება აქვს შერჩევას?

14,7	9,4	8,3	18,9	10,4	12,3	19,8	20,7	21,3	17,4
8,4	7,6	17,4	19,2	21,6	9,2	8,1	15,2	10,7	11,5
20	19,6	9,5	7,4	11,6	19,5	17,6	12,9	9,1	17,4
20,5	11,1	7,6	10,2	7	18,3	16,2	19,4	14,5	7,4

(მითითება: გამოიყენეთ მონაცემების დაჯგუფება $K=5$ ინტერვალად).

პასუხები

1) ზარისებური 2) უმნიშვნელოდ მარჯვნიდან ასიმეტრიული. 3) ზარისებური 4) მნიშვნელოვნად მარჯვნიდან ასიმეტრიული. 5) U-სებრი.

§ 12. მედიანა. მოდა

ჩვენ უკვე დავიწყეთ შერჩევის რიცხვითი მახასიათებლების განხილვა. ისინი პირობითად შეგვიძლია ორ ჯგუფად დავყოთ. პირველს მიეკუთვნებიან მონაცემთა ცენტრის საზომები: შერჩევითი საშუალო (რომელიც წინა პარაგრაფში განვიხილეთ), მედიანა და მოდა (რომლებსაც ეხლა შევეხებით). მეორეს კი ამ ცენტრის ირგვლივ მონაცემთა გაბნევის ზომის მახასიათებლები (როგორებიცაა ჩვენთვის ცნობილი სიდიდეები): გაბნევის დიაპაზონი, ვარიაციის კოეფიციენტი, შერჩევითი დისპერსია, შერჩევითი საშუალო კვადრატული გადახრა და ასევე პროცენტილები, რომელთაც მომდევნო პარაგრაფში განვიხილავთ.

როდესაც შერჩევის რაიმე რიცხვითი მახასიათებელი შემოგვაქვს სასურველია, რომ შერჩევის მცირე რაოდენობა მონაცემების ცვლილება მაინცდამაინც დიდ გავლენას არ ახდენდეს ამ მახასიათებელზე. ცხადია შერჩევითი საშუალო $\overline{X}$ შერჩევის ცენტრის ერთ-ერთი საუკეთესო მახასიათებელია, მაგრამ აქვე უნდა შევნიშნოთ რომ თვით 1-2 მონაცემის მკვეთრმა ცვლილებამაც კი შეძლება მასზე გავლენა მოახდინოს.

მაგალითი 12.1. განვიხილოთ თევზსაჭერი ტრაილერის ეკიპაჟის მიერ კვირის განმავლობაში დაჭერილი თევზის რაოდენობები (ტონებში)

12, 10, 102, 13, 9, 17, 14.

ამ მონაცემების საშუალო ტოლია

$$\overline{x}_7 = \frac{1}{7}(12+10+102+13+9+17+14) = 25{,}3\,.$$

საშუალო დღიური მაჩვენებლის როლში ამ რიცხვის განხილვა რეალურ სურათს არ მოგვცემს. მესამე დღეს მკვეთრად წარმატებული თევზჭერის მაჩვენებელმა (102 ტ) აშკარად დიდი გავლენა იქონია საშუალოზე (მონაცემთა სიმცირის გამო) და მისი შედეგი დააცილა რეალურ სურათს. თუ ამ დღის შედეგს არ ჩავთვლით გვექნება შერჩევა

12, 10, 13, 9, 17, 14

რომლის საშუალო ტოლია

$$\bar{x}_6 = \frac{1}{6}(12+10+13+9+17+14) = 12{,}5 \,,$$

რომელიც, ფაქტობრივად, რეალურ წარმოდგენას გვაძლევს თევზჭერის საშუალო დღიურ მაჩვენებელზე.

შემოვიღოთ მახასიათებელი, რომელიც მცირე რაოდენობის მონაცემების ცვლილებაზე არ რეაგირებს და იჩენს გარკვეული აზრით მდგრადობას.

განსაზღვრება 12.1. $x_1, x_2, \ldots, x_n$ შერჩევის **მედიანა** (რომელიც აღინიშნება $\tilde{X}$ სიმბოლოთი) წარმოადგენს $x_{(1)}, x_{(2)}, \ldots, x_{(n)_n}$ ვარიაციული მწკრივის შუა ელემენტს (თუ n კენტია), ან ორი შუა მეზობლის საშუალო არითმეტიკულს (თუ n ლუწია).

$$\tilde{X} = \begin{cases} x_{\left(\frac{n+1}{2}\right)}, & \text{თუ } n \text{ კენტია} \\ \dfrac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2}, & \text{თუ } n \text{ ლუწია} \end{cases}$$

მედიანა შეიძლება არ ემთხვეოდეს შერჩევის არც ერთ მონაცემს (როდესაც შერჩევის მოცულობა n ლუწია). თუ 12.1 მაგალითის მონაცემებს დავალაგებთ ვარიაციულ მწკრივად გვექნება

9, 10, 12, 13, 14, 17, 102.

ცხადია $\tilde{x} = x_{(4)} = 13$, ხოლო თუ მოვაცილებთ $x_{(7)} = 102$ მონაცემს მივიღებთ 9, 10, 12, 13, 14, 17 ვარიაციულ მწკრივს რომლისგანაც

$$\tilde{x} = \frac{x_{(3)} + x_{(4)}}{2} = \frac{12+13}{2} = 12{,}5 \,.$$

როგორც ვხედავთ, ორივე შემთხვევაში მედიანა რეალურ სურათს გვიჩვენებს და ერთი მონაცემის მოცილება მასზე დიდ გავლენას არ ახდენს. მედიანას აქვს იმ თვისებების ნაწილი, რომლებიც გააჩნია $\bar{X}$ საშუალოს.

1) $\sum_{i=1}^{n}(x_i - \tilde{X}) = 0$,

2) $\sum_{i=1}^{n}\left|(x_i - \tilde{X})\right| \le \sum_{i=1}^{n}\left|x_i - a\right|$

ნებისმიერი a რიცხვისათვის

3) თუ $y_i = ax_i + b \quad (i = \overline{1,n})$, მაშინ $\tilde{y} = a\tilde{x} + b$.

თუ მედიანის მარცხენა მხარეს მოთავსებული მონაცემებიდან რომელიმეს შევამცირებთ ან მარჯვენა მხარეს მოთავსებული მნიშვნელობებიდან რომელიმეს გავზრდით, ამით მედიანა უცვლელი დარჩება. მიუხედავად იმისა, რომ ზოგიერთ შემთხვევაში მედიანა უფრო კარგად ახასიათებს შერჩევას და მდგრადია მონაცემების მცირე ნაწილის ცვლილებებისას, შერჩევითი საშუალო მაინც უფრო მეტად ინფორმაციულია, ვინაიდან ის ყველა მონაცემს ითვალისწინებს. ამიტომ თუ ორი შერჩევის ერთმანეთთან შედარება ხდება საჭირო, პირველ რიგში საშუალოების შედარებას ახდენენ.

თუ მონაცემთა განაწილების ფორმა სიმეტრიულია, მაშინ $\tilde{X} = \bar{X}$, ხოლო თუ ასიმეტრიულია $\tilde{X} \ne \bar{X}$. მარჯვენა ასიმეტრიული განაწილებისათვის $\tilde{X} > \bar{X}$, ხოლო მარცხენა ასიმეტრიული განაწილებისათვის $\tilde{X} < \bar{X}$.

შემოვიღოთ მონაცემების ცენტრალური ტენდენციის კიდევ ერთი მახასიათებელი.

განსაზღვრება 12.2. მოდა აღინიშნება m სიმბოლოთი და ეწოდება იმ მნიშვნელობას რომელიც ყველაზე ხშირად გვხვდება შერჩევაში. თუ შერჩევაში ასეთი მნიშვნელობა ერთია, რაც ნიშნავს რომ შერჩევის პოლიგონს ერთი პიკი აქვს, მაშინ შერჩევას **უნიმოდალური** ეწოდება, თუ ორია – **ბიმოდალური**. თუ შერჩევას სამი ან მეტი მოდა აქვს მას **პოლიმოდალურს** (ხშირად **მულტიმოდალურსაც**) უწოდებენ.

მაგალითად, 7, 10, 5, 3, 17, 4, 3, 5, 5, 10, 7, 9, 12. შერჩევისათვის $m = 5$.

შენიშვნა. თუ შერჩევაში ორი ან სამი მონაცემია, რომელთა სიხშირეები მკვეთრად აღემატება დანარჩენების

სიხშირეებს, მათ ყველას თვლიან მოდად ამ სიხშირეების განსხვავების დროსაც კი (ასეთი მიდგომა ხშირად მოსახერხებელია სოციოლოგიურ კვლევებში მონაცემების დამუშავების დროს).

თუ შერჩევის ყველა მონაცემის სიხშირე ტოლია ამბობენ, რომ მას მოდა არა აქვს. ასეთ დროს ახდენენ ინტერვალებად დაჯგუფებას და მოდად ითვლება ის ინტერვალი რომელსაც ყველაზე დიდი სიხშირე ექნება. ამასთანავე აღვნიშნოთ, რომ ეს დაჯგუფება შეიძლება სხვადასხვა რაოდენობის ინტერვალებად მოხდეს და შესაბამისად **მოდალური** ინტერვალი შეიძლება სხვადასხვა იყოს. ამიტომ ამ პროცესს ახლავს ანალოგიური სირთულე, რაც ჰისტოგრამის აგებასთანაა დაკავშირებული.

დავალაგოთ ვარიაციულ მწკრივად ზემოთ მოყვანილი მონაცემაბი.

3, 3, 4, 5, 5, 5, 7, 7, 9, 10, 10, 12, 17.

ამ შერჩევისთვის $\bar{x} \approx 7{,}46$, $\tilde{x} = 7$, $m = 5$.

თუ ბოლო მონაცემი 17 შეიცვლება 50-ით, მაშინ $\bar{x} = 10$, $\tilde{x} = 7$ და $m = 5$. როგორც ვხედავთ მოდაზეც, ისევე როგორც მედინაზე, ამ ერთი მონაცემის ცვლილებამ არ იმოქმედა, საშუალო კი შეიცვალა.

ამოცანები

1. იპოვეთ მედიანა და მოდა § 4-ის ამოცანების № 1 ამოცანისათვის.

2. იპოვეთ $\tilde{x}$ და m §4-ის ამოცანების №4 ამოცანისათვის.

3. იპოვეთ მედიანა და მოდალური ინტერვალი § 5-ის ამოცანების №1 ამოცანისათვის. (**მითითება**: გამოიყენეთ ამოცანის ამოხსნისას განხილული ინტერვალური დაყოფა).

4. იპოვეთ რაგბის ბურთის წონის მედიანა და მოდალური ინტერვალი § 10-ის ამოცანების № 2 ამოცანაში მოყვანილი შერჩევისათვის (**მითითება**: მოახდინეთ დაყოფა 5 ტოლ ინტერვალად).

5. გამოთვალეთ მედიანა და მოდალური ინტერვალი სუპერმარკეტის ერთ დღეში ნავაჭრი თანხის რაოდენობისათვის, § 10-ის ამოცანების № 4 ამოცანაში მოყვანილი შერჩევისათვის (**მითითება**: დაყავით 6 ტოლ ინტერვალად).

პასუხები

1) 6; 2 და 7. 2) 4; 3. 3) 14; [12, 17). 4) 436; [430, 440). 5) 1629; [1590, 1640).

§ 13. პროცენტილები, კვარტილები და დეცილები

განსაზღვრება 13.1. მონაცემთა *p* რიგის პროცენტილი $(0 < p < 100)$ ეწოდება იმ $\widetilde{X}_p$ რიცხვს, რომელსაც არ აღემატება მონაცემთა არაუმეტეს $p\%$ და რომელიც თავად არ აღემატება მონაცემთა არაუმეტეს $(100 - p)\%$ -ს.

$$N\left(x \middle| x \le \tilde{x}_p\right) \le \frac{np}{100},$$
$$N\left(x \middle| \tilde{x}_p \le x\right) \le \frac{n(100-p)}{100}, \qquad (13.1)$$

სადაც n მონაცემთა რაოდენობაა, ხოლო N აღნიშნავს იმ მონაცემების რაოდენობას, რომლებიც აკმაყოფილებენ მის შემდეგ ფრჩხილებში ჩაწერილ პირობას.

თუ აღვნიშნავთ $\alpha = \dfrac{p}{100}$, მაშინ $\widetilde{x}_p$ გამოითვლება შემდეგნაირად

$$\tilde{x}_p = \begin{cases} x_{([n\alpha]+1)}, & \text{როცა } n\alpha \text{ არაა მთელი რიცხვი} \\ \dfrac{x_{(n\alpha)} + x_{(n\alpha+1)}}{2}, & \text{როცა } n\alpha \text{ მთელი რიცხვია} \end{cases}$$

სადაც კვადრატული ფრჩხილები აღნიშნავს მთელ ნაწილს. ცხადია, რომ თუ $n\alpha$ მთელი რიცხვია, მაშინ $\widetilde{x}_p$ შეიძლება არ ემთხვეოდეს შერჩევის არც ერთ მონაცემს.

შევნიშნოთ, რომ ზოგიერთ უცხოურ ლიტერატურაში, განსაკუთრებით სოციოლოგიური კვლევებისადმი განკუთვნილში, პროცენტილებს **პერცენტილებად** მოიხსენიებენ.

მაგალითი 13.1. მოცემულია ვარიაციული მწკრივი $(n = 10)$

$$3,\ 5,\ 5,\ 12,\ 14,\ 15,\ 16,\ 18,\ 18,\ 18.$$

ვიპოვოთ $\widetilde{x}_{30}$ და $\widetilde{x}_{42}$.

ამოხსნა. $p_1 = 30$, $\alpha_1 = \frac{30}{100} = 0{,}3$, $n\alpha_1 = 10 \cdot 0{,}3 = 3$, ამიტომ

$$\widetilde{x}_{30} = \frac{x_{(3)} + x_{(4)}}{2} = \frac{5+12}{2} = 8{,}5$$

$p_2 = 42$, $\alpha_2 = 0{,}42$, $n\alpha_2 = 4{,}2$ და $[n\alpha_2] = 4$ რის გამოც

$$\widetilde{x}_{42} = x_{(4+1)} = x_{(5)} = 14.$$

პროცენტილები გამოიყენებიან მონაცემთა როგორც განლაგების ასევე გაფანტულობის საზომადაც. განსაკუთრებით 25, 50, 75 და ყველა 10-ის ჯერადი პროცენტილები.

განსაზღვრება 13.2. მონაცემთა 25 პროცენტილს ეწოდება **პირველი კვარტილი** და აღინიშნება Q_1-ით, 75 პროცენტილს **მესამე კვარტილი** და აღინიშნება Q_3-ით, ხოლო 50 პროცენტილს კი **მეორე კვარტალი**, აღინიშნება Q_2-ით და ის, ცხადია, მედინაა.

$$Q_1 = \widetilde{X}_{25},\ Q_2 = \widetilde{X} = \widetilde{X}_{50},\ Q_3 = \widetilde{X}_{75}. \tag{13.2}$$

ხშირად პირველ კვარტილს **ქვედა** კვარტილს უწოდებენ, ხოლო მესამე კვარტილს **ზედა** კვარტილს.

განსაზღვრება 13.3. 10-ის ჯერად პროცენტილებს **დეცილები** ეწოდებათ. $\widetilde{X}_{10} = q_1$ არის **პირველი დეცილი**, $\widetilde{X}_{20} = q_2$ – მეორე **დეცილი** და ა.შ. $\widetilde{X}_{90} = q_9$ – **მეცხრე დეცილი**.

განსაზღვრება 13.4 სიდიდე $Q_3 - Q_1$ წარმოადგენს ვარიაციული მწკრივის შუაში თავმოყრილი 50% მონაცემების გაფანტულობის საზომს. მას **კვარტილთშორისი გაბნევის დიაპაზონი** ეწოდება და აღინიშნება *IQR*-ით.

$$IQR = Q_3 - Q_1.$$

კვარტილთშორისი გაბნევის დიაპაზონი მდგრადი მახასიათებელია და არ რეაგირებს Q_1-ის მარცხნივ და

Q_3-ის მარჯვნივ მოთავსებული მონაცემების შემცირებაზე ან გაზრდაზე.

ცხადია Q_1, $\widetilde{X}$ და Q_3 სიდიდეები ახასიათებენ მონაცემთა ძირითადი, შუა 50%-ის გაფანტულობას და ცენტრს. მაგრამ ისინი არავითარ ინფორმაციას არ იძლევიან განაწილების ფორმის (ანუ გაგლუვებული ჰისტოგრამის) კიდურა მნიშვნელობების შესახებ. ამიტომ განაწილების ფორმის უფრო სრულად დასახასიათებლად იყენებენ ხუთეულს q_1, Q_1, $\widetilde{X}$, Q_3, q_9, ხოლო მთელ ინფორმაციას იძლევა ხუთეული $X_{\min}$, Q_1, $\widetilde{X}$, Q_3, $X_{\max}$.

თუ $Q_3 - \widetilde{X} > \widetilde{X} - Q_1$ განაწილება არის მარჯვნივ ასიმეტრიული ანუ $\widetilde{X}$-დან მარჯვენა მხარე უფრო გრძელია ვიდრე მარცხენა, ხოლო თუ $\widetilde{X} - Q_1 > Q_3 - \widetilde{X}$, მაშინ განაწილება არის მარცხნივ ასიმეტრიული.

დავუბრუნეთ 13.1 მაგალითს გვაქვს ვარიაციული მწკრივი
3, 5, 5, 12, 14, 15, 16, 18, 18, 18.

გამოვთვალოთ $\widetilde{X}$, Q_1 და Q_3

$$\widetilde{X} = \widetilde{X}_{50} = \frac{x_{(5)} + x_{(6)}}{2} = \frac{14+15}{2} = 14{,}5 .$$

$Q_1 = \widetilde{X}_{25} = x_{(3)} = 5$, ხოლო $Q_3 = \widetilde{X}_{75} = x_{(8)} = 18$, რადგან $Q_3 - \widetilde{X} < \widetilde{X} - Q_1$, ამიტომ განაწილება მარცხნივ ასიმეტრიულია

$$IQR = Q_3 - Q_1 = 18 - 5 = 13 .$$

შევნიშნოთ, რომ უწყვეტი ტიპის მონაცემებისათვის ან ისეთი დისკრეტული მონაცემებისათვის, როდესაც მონაცემები თითქმის არ მეორდება, ხდება ინტერვალებად დაყოფა და მონაცემთა ხელოვნური დისკრეტიზირება (იხ. § 7), ინტერვალების შუა წერტილებს მიენიჭებათ შესაბამისი სიხშირეები და შემდეგ მიღებული დისკრეტიზირებული მონაცემებისათვის ითვლება პროცენტილებიც, მედიანაც და მოდაც.

მაგალითი 13.2. შოკოლადის მწარმოებელმა ფირმამ აითვისა ახალი ბაზარი, მენეჯერს აქვს შესაბამისი სადისტრიბუციო კომპანიის მიერ მოწოდებული მონაცემები პირველი 100 დღის განმავლობაში ამ ბაზარზე ფირმის წარმოებული შოკოლადის ნაწარმის დღიური გაყიდვის წონების შესახებ (წონა მოცემულია კგ-ებში).

12,5	13	18,3	19,2	22,2	24,8	15,2	16,4	13,6	18,8
22,6	22,6	24,2	25,4	26,3	27,2	28,4	28,9	28,8	19,8
26,7	29,3	28,1	29,7	24,2	18,5	17,3	30,5	31	32
29,3	31,4	30,7	33,2	33	33,5	28,5	31,5	32,4	34,4
34,1	34,9	34,9	31	35,3	34,7	34,2	31,8	35,2	35,8
34,2	35	36,3	33,4	36	35,4	39	39,5	39,8	40,5
40,7	36,7	31,2	41	40,8	40,3	35,5	39,3	40,2	43,2
44	40,8	41,2	41,5	44	39,2	44,8	44,9	45,2	43,5
40,5	40,8	44,5	47,9	47,2	45,5	45,9	32,8	47	35,8
29,5	47	45,5	25,2	22,8	21	19,2	13,2	16	13,7

შევადგინოთ მონაცემების ინტერვალური განაწილების ცხრილი და მოვახდინოთ მონაცემების დისკრეტიზირება. გამოვთვალოთ $\overline{X}$, m, Q_1, Q_3 და IQR სიდიდეების მნიშვნელობები.

ამოხსნა. უმცირესი მონაცემია 12,5 უდიდესი 47,9. განვიხილოთ [10 50] შუალედი. მონაცემების რაოდენობაა 100. ვისარგებლოთ § 5-ში მოყვანილი რეკომენდაციით და დავაჯგუფოთ მონაცემები 10 ინტერვალად. ინტერვალის სიგრძე იქნება $\Delta = \dfrac{50-10}{10} = 4$. შევადგინოთ ინტერვალური განაწილების ცხრილი და ყოველი ინტერვალი შევცვალოთ მისი შუა წერტილით.

Δ_i	v_i	$\dfrac{a_i+b_i}{2}$
[10 14)	5	12
[14 18)	4	16

[18 22)	7	20
[22 26)	9	24
[26 30)	12	28
[30 34)	15	32
[34 38)	18	36
[38 42)	16	40
[42 46)	11	44
[46 50]	3	48

მოდა იქნება $m = 36$, ხოლო შერჩევითი საშუალო $\bar{x} = 32{,}12$. გამოვთვალოთ Q_1 და Q_3.

$Q_1 = \widetilde{x}_{25}$; $p_1 = 25$; $\alpha_1 = \frac{25}{100}$; $n\alpha_1 = 25$; $\widetilde{x}_{25} = x_{(25)} = 24$.

$Q_3 = \widetilde{x}_{75}$; $p_2 = 75$; $\alpha_2 = \frac{75}{100}$; $n\alpha_2 = 75$; $\widetilde{x}_{75} = x_{(75)} = 40$.

ცხადია, $IQR = Q_3 - Q_1 = 16$.

ამოცანები

1. გამოთვალეთ I, II და III კვარტილები, 35 პროცენტილი, მე-7-ე და მე-8-ე დეცილები და კვარტილთშორისი გაბნევის IQR დიაპაზონი § 5-ის ამოცანების № 1 ამოცანის მონაცემებით სახელმწიფო სამედიცინო დაზღვევით მოსარგებლე ბავშვების რაოდენობებისათვის.

2. გამოთვალეთ მიკროსაფინანსო ორგანიზაციისაგან ყოველთვიურად გაცემული სესხების რაოდენობების მონაცემებისათვის I და III კვარტილი მე-2-ე დეცილი და კვარტილთშორისი გაბნევის დიაპაზონი §6-ის ამოცანების № 3 ამოცანის მონაცემებით.

3. გამოთვალეთ Q_1; Q_3; q_6 და IQR სიდიდეები გაყიდული ყავის წონების შერჩევისათვის § 6-ის ამოცანების №2 ამოცანიდან.

4. გამოთვალეთ Q_1; Q_2; Q_3; $\widetilde{X}_{47}$; q_4 და IQR სიდიდეები სუპერმარკეტში დღიურად ნავაჭრი თანხებისათვის § 10-ის ამოცანების № 4 ამოცანისათვის.

5. 1651 წელს დიდ ბრიტანეთში ჩატარდა კვლევა (Women's Measurements and Sizes, London, H.M.S.O. 1957, იხ. [16], გვ. 376). 4995 მანდილოსნის სიმაღლეები გაზომილია დიუმებში (1 დიუმი=2,54 სმ) და მოცემულია ცხრილის სახით (მოხდა ინტერვალებად დაჯგუფება და დისკრეტიზირება. სიხშირეები მიწერილი აქვს ინტერვალის შუა წერტილებს). გამოთვალეთ Q_1, Q_2, Q_3 და IQR სიდიდეების მნიშვნელობები.

სიმაღლე	54	56	58	60	62	64	66	68	70	72	74
„სიხშირე	5	33	254	813	1340	1454	750	275	56	11	4

პასუხები

1) 9,75; 14,5; 19,25; 12; 17; 21,5; 10,5. 2) 22850; 25200; 22300; 2350. 3) 2900; 3100; 3050; 200. 4) 1578; 1629; 1663; 1623; 1612,5; 85. 5) 60; 62; 64; 4.

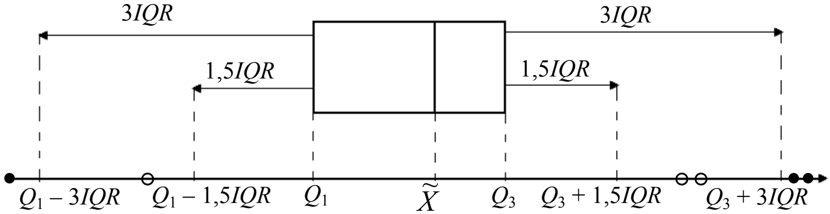
§ 14. მონაცემთა სიმრავლის კიდევ ერთი გრაფიკული წარმოდგენა, ბოქსპლოტი

სიხშირეთა ჰისტოგრამა იძლევა შერჩევის მონაცემების სიმრავლის ზოგად დახასიათებას. ფართოდაა გავრცელებული აგრეთვე $\widetilde{X}$, Q_1 და Q_3 სიდიდეების საშუალებით მონაცემთა ყოფაქცევის დახასიათების კიდევ ერთი გრაფიკული წარმოდგენა, რომელსაც **ბოქსპლოტი (boxplot)** ეწოდება. ბოქსპლოტი აიგება შემდეგნაირად:

1) გავავლოთ ჰორიზონტალური წრფე, რომელზეც მოვნიშნოთ $\widetilde{X}$, Q_1 და Q_3 წერტილები;
2) ამ წერტილებიდან აღვმართოთ ვერტიკალური მონაკვეთები, რომლებიც ჰორიზონტალური წრფის პარალელურ მონაკვეთებთან ერთად ქმნიან მართკუთხედს.
3) მიღებული მართკუთხედის გარეთ ორივე მხარეს გავავლოთ ორ-ორი, ჰორიზონტალური წრფის პარალელური წრფეები, რომელთაგან ზედა ორზე გადავზომოთ $3IQR$ სიგრძის მონაკვეთები, ხოლო ქვედა ორზე $1{,}5IQR$ სიგრძის მონაკვეთები.
4) მონაცემების შესაბამისი წერტილები, რომლებიც ჰორიზონტალურ ღერძზე მოთავსებული არიან $[Q_1 - 3IQR,\ \ Q_1 - 1{,}5IQR]$ ან $[Q_3 + 1{,}5IQR,\ \ Q_3 + 3IQR]$ შუალედებში შემოვხაზოთ წრეებით. მათ „**ზომიერი ამოვარდნები**" ეწოდებათ, ხოლო მონაცემების შესაბამისი ის წერტილები, რომლებიც ჰორიზონტალურ ღერძზე $Q_1 - 3IQR$ წერტილის მარცხნივ ან $Q_3 + 3IQR$ წერტილის მარჯვნივ არიან მოთავსებული, გავამუქოთ. მათ „**ექსტრემალური ამოვარდნები**" ეწოდებათ. საბოლოოდ ბოქსპლოტს შეიძლება ჰქონდეს სახე: (ნახ. 14.1)

შევნიშნოთ, რომ საზოგადოდ ამოვარდნების ექსტრემალურობა მარტო ბოქსპლოტით არ დგინდება, მას სხვა გამოკვლევაც სჭირდება (იხ. [1]).

რას გვიჩვენებს ბოქსპლოტი? ვინაიდან $[Q_1, Q_3]$ შუალედში მოთავსებულია მონაცემთა 50%, ამიტომ თუ $\widetilde{X}$ მედიანა რომელიმე ბოლოსთან უფრო ახლოა ეს ნიშნავს, რომ მონაცემთა ეს ნაწილი უფრო თავმოყრილია ამ ბოლოსკენ. სიმეტრიული განაწილების დროს მართკუთხედი $[X_{\min}, X_{\max}]$ დიაპაზონის შუაში იქნება მოთავსებული, ხოლო მედიანა კი დაახლოებით მართკუთხედის შუაში. თუ მართკუთხედიც წანაცვლებულია დიაპაზონის რომელიმე ბოლოსკენ და მედიანაც არაა მართკუთხედის შუაში, მაშინ განაწილება ასიმეტრიულია, თანაც იმ მხრიდან ასიმეტ-რიული მართკუთხედის რომელი მხრიდანაც უფრო მოშორებულია მედიანა, თუ პატარა მნიშვნელობები სჭარბობს, მართკუთხედი პირველ კვარტილთან იქნება ახლოს, თუ დიდი, მაშინ მესამე კვარტილთან.



ნახ. 14.1. ბოქსპლოტის ზოგადი სახე

მაგალითი 14.1. მოცემული გვაქვს მეთევზის მიერ 16 დღის განმავლობაში ყოველდღიურად დაჭერილი თევზების რაოდენობისაგან შედგენილი ვარიაციული მწკრივი. გამოვთვალოთ q_1, q_9, Q_1, Q_3, $\widetilde{X}$. ავაგოთ ბოქსპლოტი და დავახასიათოთ კიდურა მონაცემები ამოვარდნების მიხედვით.

3	7	15	19	19	19	20	20	21	21	21	21	23	29	34	45

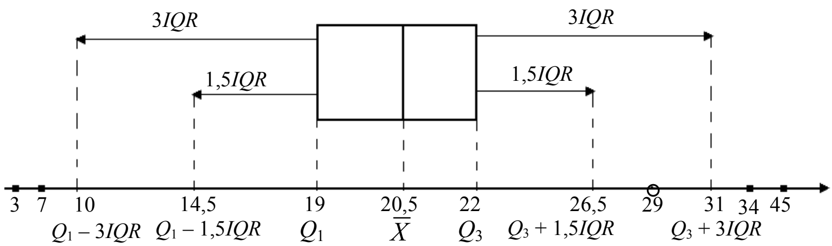
ამოხსნა. ცხადია რომ $n = 16$, $q_1 = \widetilde{X}_{10} = x_{([1,6]+1)} = x_{(2)} = 7$;

$$q_9 = \widetilde{X}_{90} = x_{([14,4]+1)} = x_{(15)} = 34\,; \qquad Q_1 = \widetilde{X}_{25} = \frac{x_{(4)} + x_{(5)}}{2} = 19\,;$$

$$Q_3 = \widetilde{X}_{75} = \frac{x_{(12)} + x_{(13)}}{2} = 22\,; \qquad \widetilde{X} = \frac{x_{(8)} + x_{(9)}}{2} = 20{,}5$$

$$IQR = Q_3 - Q_1 = 22 - 19 = 3\,.$$

ავაგოთ ბოქსპლოტი



ბოქსპლოტი გვიჩვენებს, რომ განაწილება სიმეტრიულია. ზომიერად ამოვარდნილი მონაცემია 29, ხოლო ექსტრემალურად ამოვარდნილი მონაცემებია 3, 7, 34 და 45.

ბოქსპლოტი შეიძლება გამოვიყენოთ ორი პოპულაციის განაწილების ერთმანეთთან შესადარებლად. ასეთ შემთხვევაში ერთ ნახაზზე ერთმანეთის ქვემოთ იგება ორივე პოპულაციის შერჩევების ბოქსპლოტები.

მაგალითი 14.2. მოცემულია შერჩევები ბაზარზე A და B ფირმების ჩამოსხმული სასმელი წყლის ყოველკვირეული გაყიდვებიდან (1000 ლ. სიზუსტით).

A –	18,4;	22,2;	22,8;	16,7;	19,4;	20;	24,2;	24,7;	19;2;
	13,1;	17,3;	24,1;	24,8;	26,3;	21,4;	19,4;	19,8	

B –	8;	10,3;	11,7;	14,3;	6,7;	19,4;	21,7;	18,8;	13,7;
	7,1;	5,2;	9,4;	14,1;	15,2;	16,1;	9,6;	10;	9,8;
	12,8;	22,7;	25,4;	13,4					

ორივე შერჩევისათვის გამოვთვალოთ შერჩევითი საშუალო და სტანდარტული გადახრა, მოდალური ინტერვალი, მედიანა, პირველი და მესამე კვარტილები და კვარტილთშორისი გაბნევის დიაპაზონი, ერთ ნახაზზე ავაგოთ ორივე ბოქსპლოტი და შევაფასოთ განაწილებათა სახეები.

ამოხსნა. ცხადია $n_A = 17$; $n_B = 22$. მარტივი გამოთვლებით შეგვიძლია მივიღოთ, რომ: $\bar{x}_A = 20{,}8118$; $S_A = 3{,}3647$; $\tilde{x}_A = 20$; $Q_1^A = 19{,}2$; $Q_3^A = 24{,}2$; $IQR^A = 5$; $\bar{x}_B = 13{,}4182$; $S_B = 5{,}317$; $\tilde{x}_B = 13{,}1$; $Q_1^B = 9{,}6$; $Q_3^B = 16{,}1$. $IQR^B = 6{,}5$. მოდალური ინტერვალების დასადგენად დავყოთ თითოეული შერჩევის დიაპაზონი 5-5 ტოლ ინტერვალად და შევადგინოთ ინტერვალური განაწილების ცხრილები. მივიღებთ

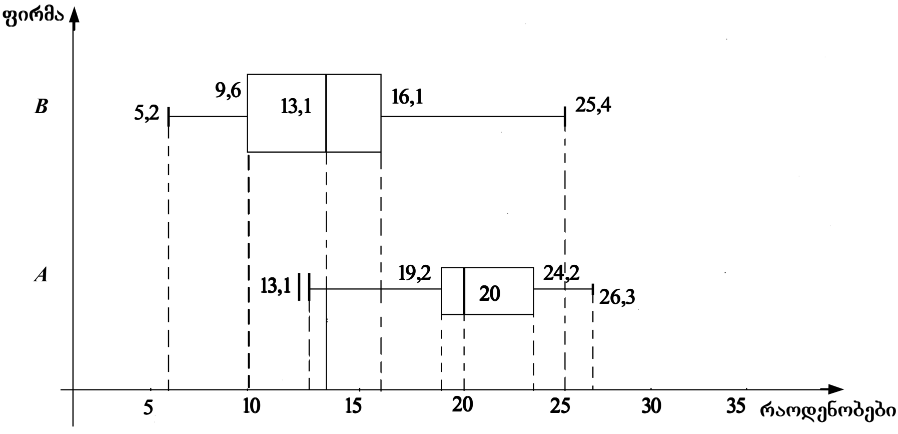
A

Δ	[13 16)	[16 19)	[19 22)	[22 25)	[25 28]
ν_i	1	3	6	6	1

B

Δ	[5 9,2)	[9,2 13,4)	[13,4 17,6)	[17,6 21,8)	[21,8 26]
ν_i	4	9	4	3	2

მაშასადამე *A* შერჩევას ორი მოდალური ინტერვალი აქვს [19 , 22) და [22 , 25), ხოლო *B* შერჩევის მოდალური ინტერვალია [9,2 13,4). ცხრილებიდან ჩანს, რომ *A* შერჩევა უმნიშვნელოდ მარცხნიდან ასიმეტრიულია, ხოლო *B* – უმნიშვნელოდ მარჯვნიდან ასიმეტრიული. ავაგოთ ბოქსპლოტები ერთ ნახაზზე. გვექნება:



A შერჩევის ბოქსპლოტზე ჩანს, რომ კვარტილთშორისი მართკუთხედი მარჯვენა ბოლოსთან უფრო ახლოა. მაშასადამე, განაწილება მარცხნიდან ასიმეტრიულია, ხოლო *B* შერჩევის ბოქსპლოტზე კვარტილთშორისი მართკუთხედი მარცხენა ბოლოსთან უფრო ახლოსაა. მაშასადამე, განაწილება მარჯვნიდან ასიმეტრიულია. ეს შედეგები ეთანხმება ჩვენს მიერ გამოთვლილ ინტერვალური განაწილების ცხრილებს. მაგრამ ბოქსპლოტები სხვა ინფორმაციასაც გვაძლევენ. კერძოდ ჩანს, რომ მონაცემთა 50% *A* შერჩევისათვის მარჯვნიდან ასიმეტრიულადაა განაწილებული, ვინაიდან მედიანა მარცხენა Q_1^A ბოლოსთან უფრო ახლოა, ხოლო *B* შერჩევისათვის კი მონაცემთა 50% სიმეტრიულადაა განაწილებული, ვინაიდან მედიანა თითქმის მართკუთხედის შუაშია.

ამოცანები

1. ააგეთ ბოქსპლოტი გაყიდული ყავის წონებისათვის §6-ის ამოცანებიდან № 2 ამოცანის შერჩევისათვის. გააკეთეთ დასკვნები განაწილების შესახებ.

2. ააგეთ ბოქსპლოტი შერჩევისათვის

19, 19, 42, 23, 24, 31, 8, 20, 2, 23, 24, 3, 35, 23, 20

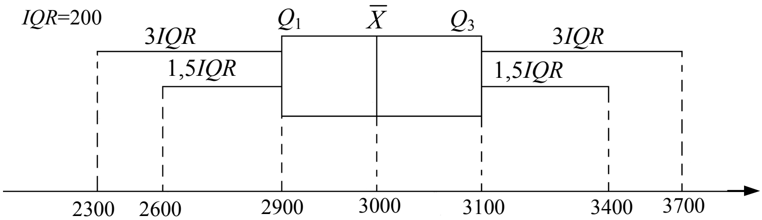
გააკეთეთ დასკვნები განაწილების შესახებ.

3. სოციოლოგიური კვლევისას (U Gensus Bureau American Community Survey 2008. [21]), მიიღეს ორი შერჩევა თითოეული 12 მოცულობის. კვლევა ეხებოდა ბავშვების რაოდენობას შემთხვევითად შერჩეულ ოჯახებში. ცხრილის მიხედვით გამოთვალეთ გაბნევის დიაპაზონი თითოეული მათგანისათვის. წარმოადგინეთ გრაფიკულად და გააკეთეთ დასკვნები (*x* არის ბავშვების რაოდენობა *ν* კი შესაბამისი ოჯახების სიხშირე). ააგეთ ბოქსპლოტები.

x	0	1	2	3	4	5	6	7	8	9	10	
ν_{I}	1	3	3	3	–	–	–	–	–	–	2	I შერჩევა
ν_{II}	1	3	2	2	–	–	1	–	2	–	1	II შერჩევა

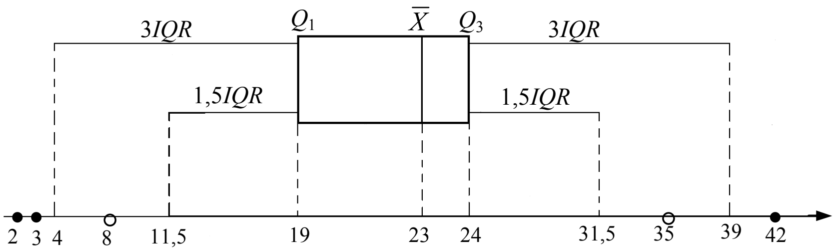
პასუხები

1)



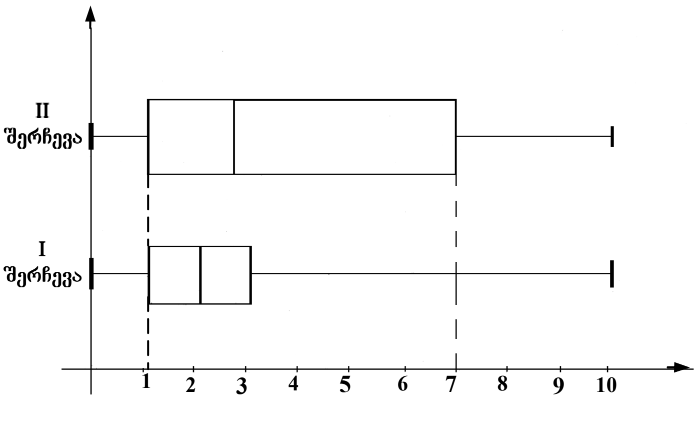
არც ზომიერი ამოვარდნები გვაქვს და არც ექსტრემალური. განაწილება სიმეტრიულია.

2) $n = 15$; $Q_1 = X_{(4)} = 19$; $\widetilde{X} = X_{(8)} = 23$; $Q_3 = X_{(12)} = 24$; $IQR = 6$.



მთლიანად განაწილება სიმეტრიულია, მაგრამ მონაცემთა შუა 50% მარცხნიდან ასიმეტრიულია. ზომიერ ამოვარდნებად შეიძლება ვიგულისხმოთ 8 და 35, ხოლო ექსტრემალურ ამოვარდნებად 2, 3 და 42.

3).



გაბნევის დიაპაზონები ორივე შერჩევისათვის ტოლია, ხოლო კვარტილთშორისი გაბნევის დიაპაზონი II შერჩევისათვის მეტია I შერჩევაზე (6>2) მიზეზი ისაა, რომ II შემთხვევაში მარჯვენა მხრისკენ მეტი მონაცემებია, ვიდრე პირველში. შუა 50% მონაცემები I შერჩევაში სიმეტრიულადაა განაწილებული, ხოლო II შერჩევაში მარჯვნიდან ასიმეტრიულად.

§ 15. შერჩევის საშუალებით პოპულაციის ინტერვალური საზღვრების დადგენა

პრაქტიკულ საქმიანობაში ხშირად გვხვდება ამოცანები, რომელთა გადაჭრისათვის საჭიროა პოპულაციის ძირითადი საზღვრების დადგენა ანუ იმ შუალედის პოვნა, რომელშიც მოხვდება პოპულაციის ელემენტთა რაოდენობის, მაგალითად, არანაკლებ 75%-ის ან არანაკლებ 80%-ის ან არანაკლებ 95%-ის და ა.შ. ცხადია ამ საზღვრების დადგენა შერჩევის საშუალებით უნდა მოხდეს.

ცნობილია, რომ თუ $X_1, X_2, \ldots, X_n$ რაიმე შერჩევაა ($\overline{X}_n$ და S_n მახასიათებლებით), მაშინ ჩეშიბევის უტოლობა ([1], [6], [9], [12]) საშუალებას გვაძლევს დავადგინოთ, რომ $[\overline{X}-KS,\ \overline{X}+KS]$ შუალედში მოხვედრილ შერჩევის მონაცემთა ფარდობითი სიხშირე არანაკლებია $1-\dfrac{1}{k^2}$-ზე. k-ს სხვადასხვა მნიშვნელობისათვის შეგვიძლია მოვიყვანოთ შესაბამის ინტერვალებში მოხვედრილი მონაცემების ფარდობითი სიხშირეთა და მაშასადამე მთელი მონაცემების პროცენტულ რაოდენობათა ქვედა საზღვები (იხ. ცხრ. 15.1).

ინტერვალი	ფარდობითი სიხშირე	პროცენტული რაოდენობა
$[\overline{X}-S,\ \overline{X}+S]$	≥ 0	$\geq$ 0%
$[\overline{X}-2S,\ \overline{X}+2S]$	$\geq \frac{3}{4}$	$\geq$ 75%
$[\overline{X}-3S,\ \overline{X}+3S]$	$\geq \frac{8}{9}$	$\geq$ 89,9%
$[\overline{X}-4S,\ \overline{X}+4S]$	$\geq \frac{15}{16}$	$\geq$ 93,7

ცხრ. 15.1. ინტერვალში მოხვედრილი შერჩევის მონაცემების ფარდობით სიხშირეთა ქვედა საზღვრების ცხრილი

იგივე სამართლიანია მთელი პოპულაციისათვის, თუ მისი საშუალო და დისპერსიაა μ და σ^2 და შესაბამისი ინტერვალებია $[\mu - k\sigma, \ \mu + k\sigma]$, $k = 1,2,3,4$.

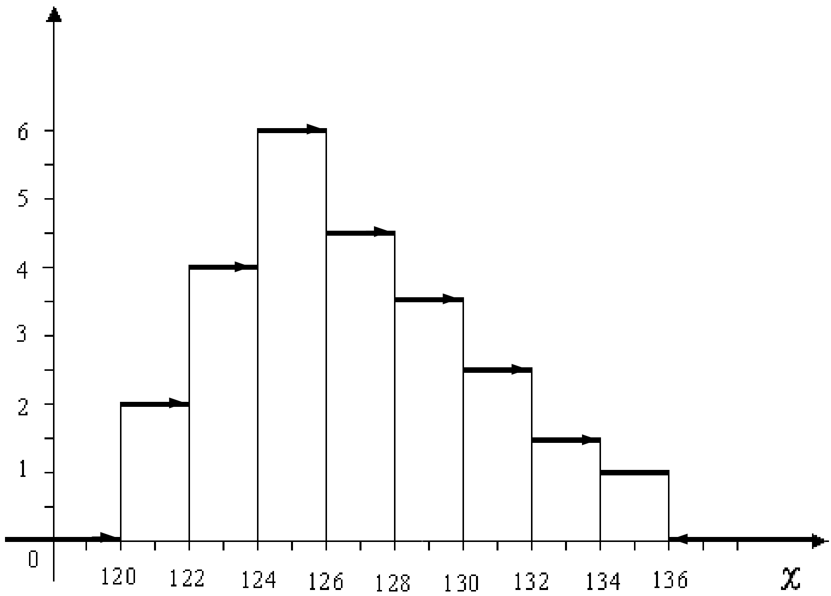
მაგალითი 15.1. A და B ქალაქების დამაკავშირებელი გზის მომიჯნავე კლდოვანი სიმაღლიდან ხდება ქვების ცვენა გარკვეულ მონაკვეთზე. საჭიროა განისაზღვროს შემაკავებელი ჯებირის მშენებლობის ოპტიმალური მონაკვეთი. გვაქვს 50 დაკვირვება ბოლო 2-3 წლის განმავლობაში ქვების ჩამოცვენის ადგილების შესახებ (მანძილი გადათვლილია კილომეტრებში A ქალაქიდან). ავაგოთ სიხშირეთა ჰისტოგრამა და გამოვთვალოთ მონაცემთა გაბნევის დიაპაზონი.

127,1	123,2	130	128,4	125,2	127,6	133,1	125,6	123,2	127
124,2	133	130,4	125	121	127,5	125,2	135,9	123,7	129,1
130,6	124,5	128,8	132,1	122,6	127,6	127,8	126	132	135,2
123,5	121	125	129,6	125,2	130,6	129,4	127	127,1	123,5
121,8	120,2	130	122,5	125,7	129,3	127,9	125,9	124	125,8

ამოხსნა. გამოთვლების შედეგად ვღებულობთ $\overline{X}_{50} = 126,952$; $S_{50}^2 = 13,063$; $S_{50} = 3,614$; გრაფიკული წარმოდგენისთვის მონაცემები დავაჯგუფოდ ინტერვალებად. ცხადია $X_{\min} = 120.2$, $X_{\max} = 135.9$, $d = 135.9 - 120.2 = 15.7$. მონაცემების მინიმალური მნიშვნელობა მივიყვანოთ უახლოეს ნაკლებ მთელ რიცხვამდე ე.ი. 120-მდე, მაქსიმალური-უახლოეს მთელ რიცხვამდე მეტობით ე.ი. 136-მდე. შემდეგ სტეიჯსის წესის (იხ. § 5) თანახმად მონაცემთა ინტერვალებად დაყოფისათვის, რადგან $n \in [50;100]$, შეგვიძლია ავიღოთ ინტერვალების $k = 8$ რაოდენობა, ამიტომ ინტერვალის სიგანე იქნება $\Delta = \dfrac{136 - 120}{8} = 2$. გვექნება სიხშირეთა განაწილების შემდეგი ცხრილი.

ინტერვალი	ν_i
[120 122)	4
[122 124)	8
[124 126)	12
[126 128)	9
[128 130)	7
[130 132)	5
[132 134)	3
[134 136]	2

სიხშირეთა ჰისტოგრამას ექნება სახე



ნახ. 15.1. სიხშირეთა ჰისტოგრამა 15.1 მაგალითისათვის

როგორც ვხედავთ განაწილება უმნიშვნელოდ მარჯვნიდან ასიმეტრიულია.

15.1 ცხრილის თანახმად $[\overline{X}-3S \quad \overline{X}+3S]$ შუალედში, რომელიც [116,11 137,794] ინტერვალს წარმოადგენს მოხვდება მთელი შერჩევის მონაცემთა არანაკლებ 89%. ეს ფაქტი შეიძლება განვაზოგადოთ მთელი პოპულაციისთვის რომელიც

ამ მონაკვეთზე ყველა შესაძლო ქვების ცვენის ადგილების სიმრავლეს ნიშნავს. ე.ი. [116,11 , 137,794] შუალედი არის ოპტიმალურად კარგი მონაკვეთი რომლის შეკეთების შემდეგაც გზის ამ ნაწილის არანაკლებ 90%-ის იქნება დაცული.

თუ შერჩევის სიხშირეთა განაწილებას აქვს სიმეტ-რიული ფორმა, არ გააჩნია გრძელი კუდები და არ არის ზემოთ ძალიან გაწელილი, მაშინ მას აქვს ზარისებური ფორმა. ის დაახლოებით $\frac{1}{\sqrt{2\pi}s}e^{-\frac{(\bar{X}-X)^2}{2s^2}}$ ფუნქციის გრაფიკის მსგავსია. ამ შემთხვევაში $[X-KS\,,X+KS]$ ინტერვალში მოხვედრილი შერჩევის მონაცემთა დაახლოებითი პროცენტული მაჩვენე-ბელი K-ს სხვადასხვა მნიშვნელობისათვის მოცემულია 15.2 ცხრილში

ინტერვალი	პროცენტული მაჩვენებელი
$[\bar{X}-S\,,\ \bar{X}+S]$	68%
$[\bar{X}-2S\,,\ \bar{X}+2S]$	95%
$[\bar{X}-3S\,,\ \bar{X}+3S]$	99%
$[\bar{X}-4S\,,\ \bar{X}+4S]$	თითქმის ყველა მონაცემი

ცხრ. 15.2. ინტერვალში მოხვედრილ მონაცემთა დაახლოებითი პროცენტული მაჩვენებელი სიხშირეთა ჰისტოგრამის ზარისებური ფორმისათვის.

ცნობილია ასევე, რომ როდესაც მონაცემების სიხშირეთა ჰისტოგრამას აქვს ზარისებრი ფორმა, მაშინ ადგილი აქვს მიახლოებით ტოლობას

$$S\approx\frac{d}{4}\,.$$

ამიტომ თუ ცნობილია, რომ ჰისტოგრამას ზარისებრი ფორმა აქვს, მაშინ s და d სიდიდეებიდან ერთ-ერთის ცოდნის შემთხვევაში მეორე შეიძლება ასეთი “უხეში” მიახლოებით გამოვთვალოთ.

პირიქით, თუ d და s სიდიდეებია ცნობილია, მაშინ მოყვანილი მიახლოებითი ტოლობა საშუალებას გვაძლევს გამოვთქვათ ვარაუდი იმის შესახებ, რომ მონაცემების სიხშირეთა განაწილების ჰისტოგრამას აქვს ზარისებური ფორმა.

ამოცანები

1. გამოთვალეთ §6-ის ამოცანებიდან №2 ამოცანის შერჩევისათვის 50%-იანი, 68%-იანი და 95%-იანი ინტერვალები.
2. გამოთვალეთ §6-ის ამოცანებიდან №3 ამოცანის შერჩევისათვის 50%-იანი ინტერვალი.
3. გამოთვალეთ §5-ის ამოცანებიდან №1 ამოცანის შერჩევისათვის 50%-იანი და 68%-იანი ინტერვალები.
4. §4-ის ამოცანებიდან №4 ამოცანის შერჩევისათვის გამოთვალეთ 50%-იანი, 68%-იანი და 95%-იანი ინტერვალები.
5. მოცემული შერჩევისათვის გამოთვალეთ 50%-იანი, 68%-იანი და 95%-იანი ინტერვალები.

7; 12; 23; 31; 18; 10; 12; 12; 18; 5; 25; 11,

პასუხები

1) $[Q_1, Q_3]=[2900, 3100]$; $[\overline{X}-S, \ \overline{X}+S]=[2880{,}4 \ 3149{,}6]$; $[\overline{X}-2S, \ \overline{X}+2S]=[2745{,}7 \ \ 3284{,}3]$ .2)$[22850; \ 25200]$ .3)$[9{,}75, 19{,}25]$; [8,13; 21,34]. 4) [3, 5]; [2,557, 5,177]; [1,247, 6,487]. 5) [10,5, 20,5]; [7,866, 22,794]; [0,402, 30,258].

I თავი

ალბათობის თეორია

§16. ცდა, ხდომილობა, ელემენტარულ ხდომილობათა სივრცე

ალბათობის თეორიაში ძირითადად შეისწავლება ისეთი მოვლენები, რომელთა შედეგების უტყუარი წინასწარმეტყველება შეუძლებელია. ეს ხდება ამ მოვლენების გამომწვევი უამრავი სხვადასხვა, თუნდაც უმნიშვნელო ფაქტორების გაუთვალისწინებელი და მოუწესრიგებელი ცვალებადობის გამო. ასეთ მოვლენებს შემთხვევით მოვლენებს უწოდებენ და ამბობენ, რომ ისინი აღიწერებიან შემთხვევითი ანუ სტოქასტური სქემით. ასეთი მოვლენებია, მაგალითად კვირის განმავლობაში რომელიმე კინოთეატრში მისულ მაყურებელთა რაოდენობა, დღე-ღამის განმავლობაში საქართველოში დაბადებულ ბავშვთა რაოდენობა და ა.შ. თუმცა, ამასთანავე, არსებობენ ისეთი მოვლენებიც, რომელთა წინასწარმეტყველებაც შესაძლებელია და მათ საბუნებისმეტყველო მეცნიერებები შეისწავლიან (მათზე ამბობენ, რომ ისინი აღიწერებიან არაშემთხვევითი ანუ დეტერმინისტული სქემით). მაგალითად სხეულის ვარდნისას შესაძლებელია სიჩქარის განსაზღვრა, ჰაერის ტენიანობის მონაცემებზე დაკვირვებით წვიმის მოსვლის წინასწარმეტყველება და ა.შ.

შემთხვევით მოვლენას შეიძლება სასრული რაოდენობის სხვადასხვა შესაძლო შედეგებიც გააჩნდეს (მაგალითად, სტენდზე ტყვიის სროლაში ერთი გასროლისას მიღებული შესაძლო ქულების რაოდენობა) და შეიძლება უსასრულო რაოდენობისაც (მაგალითად, რომელიმე ადამიანს ვთხოვოთ რაიმე რიცხვის დასახელება). იმისათვის, რომ შემთხვევითი მოვლენის რაიმე შესაძლო შედეგის განხორციელების „შანსი“ ანუ „პროცენტული შესაძლებლობა“ შევაფასოთ, საჭიროა მოვლენა აღვწეროთ მათემატიკური მოდელით. ამისათვის უნდა დავადგინოთ რა შესაძლო შედეგებია მოსალოდნელი ამ მოვლენის განხორციელებისას.

ლითონის ფულის აგდებისას მის ზემოთა მხარეს შეიძლება აღმოჩნდეს ან „გერბი“ ან „საფასური“. კამათლის გაგორებისას შეიძლება „მოვიდეს“. 1, 2, 3, 4, 5 ან 6 ციფრებიდან ერთ-ერთი. ასეთი ცდები რეალურად შეგვიძლია ჩავატაროთ და მათზე დაკვირვებით ლოგიკური დასკვნები გამოვიტანოთ.

შემთხვევით მოვლენაზე დაკვირვების მიზნით ჩატარებულ ცდებს შემთხვევითი ცდები ეწოდებათ. ისინი თავისი ბუნებით განსხვავდებიან ფიზიკის ან ქიმიის გაკვეთილზე სკოლის მოსწავლეებთან ჩატარებული ისეთი ცდებისაგან, რომელთა დროსაც შეიძლება რაიმე სიდიდის გაზომვა ან რაიმე რეაქციაზე დაკვირვება (მაგალითად, სხეულის თავისუფალი ვარდნის აჩქარების გასაზომად, ან ზამბარის სიხისტის კოეფიციენტის დასადგენად ჩატარებული ცდები).

მომავალში შემთხვევით ცდას უბრალოდ ცდას ვუწოდებთ. ჩავატაროთ ერთი და იგივე ცდა n-ჯერ. დავუშვათ ამ ცდების დროს ვაკვირდებით რაიმე A-შედეგს და იგი მოხდა m-ჯერ. მაშინ ცხადია $W(A)=\dfrac{m}{n}$ წარმოადგენს A-ს ფარდობით სიხშირეს. თუ ჩავატარებთ იგივე ცდას სხვა დროს იგივე პირობებში n_1-ჯერ და A მოხდება m_1-ჯერ მაშინ $W_1(A)=\dfrac{m_1}{n_1}$. ალბათობის თეორია შეისწავლის ისეთი ცდების მათემატიკურ მოდელებს, რომლებიც მდგრადი ფარდობითი სიხშირით ხასიათდებიან. რაც ნიშნავს რომ დიდი n და n_1, რიცხვებისათვის $W(A)$ და $W_1(A)$ ახლოს უნდა იყვნენ ერთი და იგივე რიცხვთან.

ცნობილია რომ კ. **პირსონმა** (1857-1936) 24000-ჯერ ააგდო ლითონის ფული და გერბი მოვიდა 12012-ჯერ. მისი ფარდობითი სიხშირე $W=\dfrac{12012}{24000}$ ახლოსაა $\dfrac{1}{2}$-თან. პირსონმა ასეთი ცდების სერია რამოდენიმეჯერ გაიმეორა და აღმოაჩინა რომ ფარდობითი სიხშირე ყოველთვის ახლოს იყო $\dfrac{1}{2}$-თან.

თითოეულ ჩვენთაგანს ადვილად შეუძლია კამათლის (რა თქმა უნდა, სიმეტრიულის) ბევრჯერ გაგორებისას დაითვალოს რაიმე ერთი და იგივე რიცხვის მოსვლათა რაოდენობა. შევამჩნევთ, რომ ამ რიცხვის მოსვლის ფარდობითი სიხშირე დაახლოებით $\frac{1}{6}$-ის ტოლი იქნება.

მოვიყვანოთ მაგალითი, რომელიც გვიჩვენებს თუ როგორ ახერხებენ ბიოლოგები ტბაში თევზის მარაგის დათვლას ფარდობითი სიხშირის გამოყენებით. სპეციალური, ბიოლოგიური კვლევებისათვის განკუთვნილი გემით ხდება თევზის დაჭერა, მათი დანიშვნა ამ მიზნისათვის განკუთვნილი მსუბუქი ნიშნებით და ტბაში ჩაბრუნება. დავუშვათ დაიჭირეს n_1 ცალი თევზი. თუ მთელი თევზების რაოდენობა ტბაში არის N-ცალი, მაშინ დაჭერილი თევზების გაშვების შემდეგ დანიშნული თევზების ფარდობითი სიხშირე ტბაში იქნება $W = \frac{n_1}{N}$. აცდიან რამოდენიმე დღეს რათა მოხდეს თევზების გაბნევა ბუნებრივი შემთხვევითი გადაადგილებებით და მათი შერევა ტბის დანარჩენ თევზებთან. შემდეგ კვლავ წარმოებს თევზჭერა. დავუშვათ მეორე ჯერზე დაიჭერს n_2 ცალი რომელშიც K ცალი ძველი დანიშნული თევზი იყო. რაკი ტბაში თევზების შემთხვევითი შერევა ხდება ბუნებრივია ვგულისხმობთ რომ ამ ახლად დაჭერილ თევზებში დანიშნული თევზების ხვედრითი წილი ანუ ფარდობითი სიხშირე იგივე უნდა იყოს რაც მთელ ტბაში. ამოტომ $\frac{K_1}{n_2} \approx \frac{n_1}{N}$ ანუ $N \approx \frac{n_1 n_2}{K}$.

მაგალითად, თუ პირველ დაჭერაზე დანიშნული 5000 თევზიდან მეორედ დაჭერილ 3000 ცალში ძველი დანიშნული აღმოჩნდა 21 ცალი, მაშინ $\frac{21}{5000} \approx \frac{5000}{N}$ ანუ

$$N \approx \frac{15000000}{21} \approx 710000\,.$$

დავუბრუნდეთ ცდას და მასთან დაკავშირებულ შედეგებს და მოვიყვანოთ რამოდენიმე განსაზღვრება.

განსაზღვრება 16.1. **ცდა** ვუწოდოთ პირობათა რაიმე ერთობლიობის შესრულებას, რომლის შედეგიც არაა ცალსახად განსაზღვრული და შემთხვევითია. ნებისმიერ გამონათქვამს ცდის შედეგების შესახებ ვუწოდოთ **ხდომილობა**. ხდომილობები აღინიშნებიან ლათინური ანბანის დიდი ასოებით $A, B, C, \ldots$. ცდის იმ ელემენტარული შედეგს, რომლის წარმოდგენა სხვა შედეგების საშუალებით შეუძლებელია **ელემენტარული ხდომილობა** ეწოდება. ხოლო ცდის ყველა შესაძლო ელემენტარულ ხდომილობათა სიმრავლეს **ელემენტარულ ხდომილობათა სივრცე** ეწოდება და Ω (ომეგა) სიმბოლოთი აღინიშნება. ყოველი ხდომილობა შეიძლება გავაიგივეოთ მის ხელშემწყობ ელემენტარულ ხდომილობათა სიმრავლესთან.

მაგალითი 16.1. ცდა მდგომარეობს ერთი კამათლის ერთხელ გაგორებაში.

ელემენტარული ხდომილობებია მოვიდა„1“, მოვიდა„2“, მოვიდა„3“,…,მოვიდა„6“. ელემენტარულ ხდომილობათა სივრცეა $\Omega = \{1,2,3,4,5,6\}$.

ჩამოვთვალოთ ზოგიერთი ხდომილობები ამ ცდის შესახებ და გავაიგივოთ ისინი მათ ხელშემწყობ ელემენტარულ ხდომილობათა სიმრავლეებთან. თუ ხელშემწყობი ელემენტალური ხდომილობები არ არსებობენ მივუწეროთ ცარიელი სიმრავლის სიმბოლო $\varnothing$.

A – მოვიდა ლუწი რიცხვი = {2, 4, 6}

B – მოვიდა რიცხვი 5 = {5}

C – მოვიდა 4-ზე მეტი რიცხვი = {5, 6}

D – მოვიდა რიცხვი რომელიც არ აღემატება 4-ს ={1,2, 3, 4}

E – მოვიდა 10-ზე მეტი რიცხვი = $\varnothing$

F – მოვიდა დადებითი რიცხვი = {1, 2, 3, 4, 5, 6}.

ცხადია, რომ B ელემენტარული ხდომილობაა. E ხდომილობა არასდროს არ განხორციელდება, ხოლო F ხდომილობა ცდის ჩატარებისას აუცილებლად განხორციელდება.

განსაზღვრება 16.2. ხდომილობას, რომელიც ცდის ჩატარებისას აუცილებლად მოხდება **აუცილებელი ხდომილობა** ეწოდება და აღინიშნება Ω სიმბოლოთი. ხდომილობას, რომელიც ცდის ჩატარებისას არასდროს არ შეიძლება მოხდეს ეწოდება **შეუძლებელი ხდომილობა** და აღინიშნება $\varnothing$ სიმბოლოთი. A **ხდომილობის საწინააღმდეგო ხდომილობა** (ანუ A-ს დამატება) აღინიშნება $\overline{A}$ სიმბოლოთი და ეწოდება $\Omega - A$ ხდომილობას $(\overline{A} = \Omega - A)$. ცხადია, რომ $\overline{\overline{A}} = A$.

ჩვენ ხდომილობა გავაიგივეთ მის ხელშემწყობ ელემენტარულ ხდომილობათა სიმრავლესთან (ამით უკვე გასაგები გახდა აუცილებელი და შეუძლებელი ხდომილობების აღნიშვნები). მათზე შეგვიძლია ჩავატაროთ ისეთივე მოქმედებები როგორიც სიმრავლეებზე.

განსაზღვრება 16.3. 1) თუ M ხდომილობის მოხდენისას ხდება N ხდომილობა ამბობენ, რომ M იწვევს N-ს (ან M-დან გამომდინარეობს N) და ჩაიწერება $M \subset N$ ან $N \supset M$.

2) M და N ხდომილობების **ნამრავლი (ანუ თანაკვეთა)** აღინიშნება MN (ან $M \bigcap N$) სიმბოლოთი და წარმოადგენს ხდომილობას, რომელიც ხდება მაშინ და მხოლოდ მაშინ, როდესაც ხდება M და N ორივე ხდომილობა ერთდროულად. როდესაც $M \bigcap N = \varnothing$, მაშინ M და N ხდომილობებს **არათავსებადი (უთავსადი) ხდომილობები** ეწოდებათ.

3) M და N ხდომილებების **გაერთიანება** აღინიშნება $M \bigcup N$ სიმბოლოთი და წარმოადგენს ხდომილობას, რომელიც ხდება მაშინ და მხოლოდ მაშინ, როდესაც ხდება ერთი მაინც M და N ხდომილობებიდან. არათავსებად M და N ხდომილობათა გაერთიანებას მათ ჯამს უწოდებენ და ხშირად ჩაწერენ $M + N$ სიმბოლოთიც.

4) M და N ხდომილობების სხვაობა აღინიშნება $M - N$ სიმბოლოთი (ხშირად იხმარება $M \setminus N$ ჩანაწერიც) და წარმოადგენს ხდომილობას, რომელიც ხდება მაშინ და მხოლოდ მაშინ, როდესაც ხდება M და არ ხდება N.

დავუბრუნდეთ **16.1** მაგალითს. ზემოთ თქმულიდან გამომდინარე ცხადია, რომ $B \subset C$;

$A \bigcap B$ ნიშნავს – მოვიდა ლუწი რიცხვი, რომელიც 5-ის ტოლია. $A \bigcap B = \varnothing$;

$A \bigcap C$ ნიშნავს – მოვიდა ლუწი რიცხვი, რომელიც მეტია 4-ზე. $A \bigcap C = \{6\}$;

$\overline{C}$ ნიშნავს – არ მოვიდა 4-ზე მეტი რიცხვი. $\overline{C} = \{1,2,3,4\} = D$;

$A \setminus C$ – ნიშნავს: მოვიდა ლუწი რიცხვი, რომელიც არაა მეტი 4-ზე. $A \setminus C = \{2,4\}$

$C \setminus A$ ნიშნავს: მოვიდა 4-ზე მეტი რიცხვი რომელიც არაა ლუწი $C \setminus A = \{5\}$

$\overline{A}$ – ნიშნავს: არ მოვიდა ლუწი რიცხვი $\overline{A} = \{1,3,5\}$

$C \setminus F$ ნიშნავს: მოვიდა 4-ზე მეტი რიცხვი, რომელიც არაა დადებითი $C \setminus F = \varnothing$.

პრაქტიკაში ხშირად გვხვდება ამოცანები, რომლებშიც საჭიროა მოცემული სიმრავლიდან ამოვირჩიოთ რაიმე თვისების მქონე ქვესიმრავლეები, ერთი ან რამდენიმე სიმრავლის ელემენტები დავალაგოთ გარკვეული თანმიმდევრობით და ა.შ. მაგალითად, ადამიანთა მოცემული ჯგუფიდან ავირჩიოთ წარმომადგენლები კონკრეტული მისიის შესასრულებლად, ციფრთა მოცემული სიმრავლისაგან შევადგინოთ სეიფის საიდუმლო კოდი და ა.შ.

რადგან ასეთი სახის ამოცანებში საქმე გვაქვს ობიექტების ამა თუ იმ კომბინაციებთან, ამიტომ ასეთი სახის ამოცანებს კომბინატორული ამოცანები ეწოდებათ. ხოლო მათემატიკის მიმართულებას, რომელიც სწავლობს კომბინატორულ ამოცანებს კომბინატორიკა ეწოდება. კომბინატორიკა შეიძლება განხილულ იქნას, როგორც სიმრავლეთა თეორიის ნაწილი იმ აზრით, რომ კომბინატორიკის ყოველი ამოცანა შესაძლებელია დაყვანილ იქნას სასრულ სიმრავლეებზე და ასეთი სიმრავლეებს შორის გადასახვებზე.

სიმრავლეთა თეორიის საფუძვლების გაცნობა წარმოადგენს კომბინატორული ანალიზის მეთოდების გაცნობის მყარ ბაზას. კომბინატორული ამოცანების რაოდენობა და

მათი მრავალფეროვნება დღითიდღე იზრდება, რაც განპირობებულია კომბინატორიკის გამოყენებით სხვადასხვა პრაქტიკულ ამოცანებში. ამიტომ კომბინატორული ამოცანების ჩამოყალიბება ხშირად ნათელი და მარტივია, მაგრამ მათი ამოხსნა დაკავშირებულია გარკვეულ სირთულეებთან. კომბინატორული ანალიზის ამოცანებში გამოკვლევის ძირითად ობიექტებს წარმოადგენენ დისკრეტული სიმრავლეები, ე.ი. სიმრავლეები, რომელთა ელემენტები არიან განცალკევებულნი. უმეტეს შემთხვევებში კვლევის ობიექტები სასრული სიმრავლეებია, თუმცა ბევრი ამოცანა დაკავშირებულია უსასრულო სიმრავლეების განხილვასთანაც. კომბინატორული ამოცანების თავისებურება მდგომარეობს იმაში, რომ უპირატესი ყურადღება ექცევა მოცემული სიმრავლიდან ქვესიმრავლეთა ამორჩევას და ელემენტების დალაგებას. ეს ორი ოპერაცია არის ძირითადი კომბინატორული ოპერაცია.

კომბინატორული ამოცანების განხილვისას არსებითია მოიძებნოს თუნდაც ერთი განლაგება ობიექტებისა, რომლებიც ფლობენ სასურველ თვისებას (მაგალითად, ჭადრაკის დაფაზე განლაგდეს რვა დედოფალი ისე, რომ ისინი ერთმანეთს არ კლავდნენ). ზოგიერთ შემთხვევაში შესაძლებელია ვაჩვენოთ, რომ დასმულ ამოცანას ამონახსნი არ აქვს (მაგალითად, შეუძლებელია ხუთი კურდღელი ჩავსვათ ოთხ გალიაში ისე, რომ ერთ გალიაში არ იჯდეს ერთ კურდღელზე მეტი). თუ ამოცანას გააჩნია რამდენიმე ამოხსნა, მაშინ დგება საკითხი ამონახსნების რაოდენობის დადგენაში და ყველა ამონახსნის გაანალიზებაში.

წარმოდგენილ წიგნში განვიხილავთ კომბინატორული ამოცანის ამოხსნის რაოდენობების დადგენას. სწორედ კომბინატორიკის ეს ნაწილია მჭიდროდ დაკავშირებული ალბათობის თეორიასთან.

$N(A)$ სიმბოლოთი აღვნიშნოთ სასრული A სიმრავლის ელემენტთა რაოდენობა. თუ მოცემულია ორი სასრული A და B სიმრავლე და $A \bigcap B = \varnothing$, მაშინ მართებულია შემდეგი ტოლობა:

$$N(A \bigcup B) = N(A) + N(B)\,.$$

კომბინატორიკაში ბოლო ტოლობას ეწოდება ჯამის წესი. ის შეიძლება ჩამოვაყალიბოთ შემდეგნაირად: თუ $N(A)=k$, $N(B)=m$ და A სიმრავლიდან ელემენტების ნებისმიერი ამორჩევა არ არის დამოკიდებული B სიმრავლიდან ელემენტების ნებისმიერ ამორჩევაზე, მაშინ ელემენტის ამორჩევა A ან B სიმრავლიდან შესაძლებელია $k+m$ რაოდენობა ხერხით.

მაგალითად, თუ კალათაში მოთავსებულია 10 წითელი და 8 შავი ბურთულა, მაშინ კალათიდან ერთი ბურთულა შეიძლება ამოვირჩიოთ $10+8=18$ ხერხით.

იმ შემთხვევაში, როცა $A\bigcap B\neq\emptyset$, მაშინ ცხადია, რომ $N(A\bigcup B)\neq N(A)+N(B)$. ეს განპირობებულია იმით, რომ A და B სიმრავლეებს გააჩნიათ საერთო ელემენტები, რომლებიც $A\bigcup B$ გაერთიანებაში შედიან მხოლოდ ერთჯერ. ამიტომ საჭიროა $N(A)+N(B)$ ჯამს გამოვაკლოთ $A\bigcap B$ თანაკვეთის ელემენტთა რაოდენობა. ამრიგად, მივიღებთ

$$N(A\bigcup B(=N(A)+N(B)-N(A\bigcap B).$$

ორი სიმრავლის გაერთიანების ელემენტების რაოდენობის გამოსათვლელ მითითებულ ფორმულას ჩართვისა და გამორიცხვის მეთოდი ეწოდება. ეს მეთოდი შეიძლება განზოგადდეს სასრული რაოდენობის სასრულ სიმრავლეთა ოჯახისათვის. კერძოდ, თუ მოცემულია სასრულ სიმრავლეთა $A_1,A_2,A_3,\ldots,A_m$ სასრული ოჯახი, შესაბამისად

$$N(A_1),N(A_2),N(A_3),\ldots,N(A_m)$$

ელემენტთა რაოდენობებით, მაშინ ადგილი აქვს შემდეგ ფორმულას:

$$N(A_1\bigcup A_2\bigcup A_3\bigcup\ldots\bigcup A_m)=N(A_1)+N(A_2)+N(A_3)+\ldots+N(A_m)-$$
$$-\{N(A_1\bigcap A_2)+N(A_1\bigcap A_3)+\cdots+N(A_{m-1}\bigcap A_m)\}+$$
$$+\{N(A_1\bigcap A_2\bigcap A_3)+N(A_1\bigcap A_2\bigcap A_4)+\cdots+N(A_{m-2}\bigcap A_{m-1}\bigcap A_m)\}+\cdots$$
$$\cdots+(-1)^{m-1}N(A_1\bigcap A_2\bigcap A_3\bigcap\ldots\bigcap A_m).$$

ორი A და B სიმრავლის **დეკარტული ნამრავლი** აღინიშნება $A\times B$ სიმბოლოთი და ეწოდება (a,b) წყვილების სიმრავლეს, სადაც $a\in A$ და $b\in B$. თუ

$N(A)=k$ და $N(B)=m$, მაშინ $N(A\times B)=k\cdot m$. **მივიღებთ ე.წ. გამრავლების წესს**:

თუ A სიმრავლიდან რაიმე a ელემენტის ამორჩევა შესაძლებელია k ხერხით, ხოლო b ელემენტის ამორჩევა შესაძლებელია m ხერხით, მაშინ (a, b) წყვილის ამორჩევა შესაძლებელი იქნება $k\cdot m$ ხერხით.

მაგალითად, ქართული ანბანის ასოებისაგან შედგენილი ორასოიანი მარცვლების რაოდენობა, რომლის პირველი ასოა თანხმოვანი, ხოლო მეორე ასო კი – ხმოვანი, არის $28\cdot 5=140$. აქვე შევნიშნოთ, რომ ანალოგიური ამოცანის შედეგი, ყველა შესაძლო ორასოიანი მარცვლებისათვის, იქნება $33^2=1089$.

თუ განვაზოგადებთ დეკარტული ნამრავლის ცნებას სასრული რაოდენობა სასრული სიმრავლეებისათვის და თუ $N(A)=m$, მაშინ $N(\underbrace{A\times A\times\ldots\times A}_{k-ჯერ})=m^k$. ამ ტოლობიდან მიიღება შემდეგი წესი:

m ელემენტიანი სიმრავლის ყველა ისეთი ქვესიმრავლეთა რაოდენობა, რომლებიც შედგებიან k ცალი ელემენტებისგან, ტოლია m^k.

მაგალითად, ყველა ოთხნიშნა რიცხვების რაოდენობა, რომლის პირველი ციფრი ლუწი რიცხვია, იქნება $4\cdot 10^3=4000$, რადგან პირველი ციფრი შეიძლება იყოს 2, 4, 6, 8 და დანარჩენი სამ ციფრიანი რიცხვი უნდა შევადგინოთ 10 ციფრისაგან.

ელემენტთა ამორჩევა მოცემული სიმრავლიდან შესაძლებელია განხორციელდეს ორი ხერხით: **ამორჩევა დაბრუნებით** და **ამორჩევა დაბრუნების** გარეშე. ცხადია, რომ პირველ შემთხვევაში ამორჩეული ელემენტები შეიძლება განმეორდეს, ხოლო მეორე შემთხვევაში კი – არა. დაბრუნების გარეშე ამორჩევის შემთხვევაში ამორჩეული ელემენტები შეიძლება დალაგდნენ გარკვეული რიგით. ასეთი გზით მიღებულ სიმრავლეს დალაგებული სიმრავლე ეწოდება. მაგალითად, ადამიანთა გარკვეული ჯგუფი შეიძლება დავალაგოთ ასაკის, წონის, გვარების პირველი ასოების ანბანში რიგის მიხედვით და სხვ.

ვთქვათ, $N(A)=m$ და $k \le m$. ამ შემთხვევაში შეიძლება შევადგინოთ A სიმრავლის k ელემენტიანი ქვესიმრავლეები. მაგალითად, $A=\{1,3,4,7\}$, მაშინ ყველა ორელემენტიან დალაგებული ქვესიმრავლეთა რაოდენობა იქნება 12. ესენია:
(1,3), (1,4), (1,7), (3,1), (3,4), (3,7), (4,1), (4,3), (4,7), (7,1), (7,3), (7,4).

m ელემენტიანი სიმრავლის k ელემენტიან დალაგებულ სიმრავლეს ეწოდება **წყობა** გამეორების გარეშე m ელემენტიდან k ელემენტად. მათი რაოდენობა აღინიშნება A_m^k სიმბოლოთი. მარტივია იმის ჩვენება, რომ

$$A_m^k = \frac{m!}{(m-k)!} \quad (m! = 1\cdot 2\cdot \ldots \cdot m).$$

ახლა განვიხილოთ მოცემული m ელემენტიანი სიმრავლის სხვადასხვა დალაგება. ამ შემთხვევაში დალაგებული სიმრავლეები ერთმანეთისაგან განსხვავდებიან მხოლოდ მათში შემავალი ელემენტების დალაგების რიგით. მათ ეწოდებათ გადანაცვლებები გამეორების გარეშე და მათი რაოდენობა აღინიშნება P_m სიმბოლოთი. მაგალითად, თუ $A=\{1,3,7\}$, მაშინ შეიძლება შევადგინოთ $P_3 = 3! = 6$ გადანაცვლება. ესენია:

(1,3,7), (1,7,3), (3,1,7), (3,7,1), (7,1,3), (7,3,1).

კომბინატორიკის ერთ-ერთ მნიშვნელოვან ამოცანას წარმოადგენს m ელემენტიანი სიმრავლიდან k ელემენტიანი ქვესიმრავლეების რაოდენობის დადგენა. ასეთ არა დალაგებელ ქვესიმრავლეს ეწოდება **ჯუფთება** გამეორების გარეშე m ელემენტიდან k ელემენტად და მათი რაოდენობა აღინიშნება C_m^k სიმბოლოთი. მაგალითად, თუ $A=\{1,3,4,7\}$, მაშინ ყველა ორელემენტიან ქვესიმრავლეთა რაოდენობა იქნება 6. ესენია:

(1,3), (1,4), (1,7), (3,4), (3,7), (4,7).

ჯუფთების შემთხვევაშიც ადვილია ჩვენება იმისა, რომ

$$C_m^k = \frac{m!}{k!\cdot(m-k)!} \quad \left(C_m^k = \frac{A_m^k}{P_k} \right).$$

მართებულია შემდეგი ტოლობები:

1. $C_m^0 = 1$;

2. $C_m^m = 1$;

3. $C_m^k = C_m^{m-k} \qquad (0 \le k \le m)$;

4. $C_m^0 + C_m^k + \cdots + C_m^k + \cdots + C_m^m = 2^m \qquad (m \ge 0)$;

5. $C_m^k = C_{m-1}^{m-k} + C_{m-1}^k \ (0 \le k \le m)$.

ამოცანები

1. კამათელს აგორებენ ორჯერ: ა) რამდენი ელემენტისაგან შედგება ელემენტარულ ხდომილობათა სივრცე? ბ) რამდენი ელემენტარული ხდომილობაა იმ ფაქტის ხელშეწყობი, რომ პირველ გაგორებაზე მოვიდა უფრო დიდი ციფრი, ვიდრე მეორეზე?

2. ერთდროულად აგორებენ ორ კამათელს. ელემენტარულია თუ არა ხდომილობები: A-მოსული ციფრების ნამრავლი 36-ის ტოლია; B-მოსული ციფრების ნამრავლი 25-ის ტოლია; C-მოსული ციფრების ნამრავლი 20-ის ტოლია; D-მოსული ციფრების ჯამია 2; E-მოსული ციფრების ჯამია 11; F-მოსული ციფრების ჯამია 12.

3. ერთდროულად აგორებენ ორ კამათელს. ჩამოწერეთ შემდეგი ხდომილობების ხელშემწყობი ელემენტარული ხდომილობები: A-მოსული ციფრების ჯამია 5; B-მოსული ციფრების სხვაობის სიდიდეა 4.

4. ერთდროულად აგორებენ ოთხ კამათელს. რამდენი ელემენტარული ხდომილობისაგან შესდგება ელემენტარულ ხდომილობათა სივრცე და ხდომილობები: A-ყველა კამათელზე მოვიდა კენტი რიცხვი; B-ყველაზე მოვიდა მარტივი რიცხვი. D-ორზე მოვიდა კენტი რიცხვი, ორზე ლუწი.

5. 36 ქაღალდიანი ბანქოს დასტიდან იღებენ ერთ ცალს. რამდენი ელემენტარული ხდომილობისაგან შესდგებიან ხდომილობები: A-ამოვიდა წითელი ფერის; B-ამოვიდა ნახატი; C-ამოვიდა გული, $A-B$; $B-A$; $C-B$; $B-C$.

6. მე-5 ამოცანის პირობებში ჩაწერეთ შემდეგი ხდომილობების ხელშემწყობი ელემენტარული ხდომილობები: A-ამოვიდა შავი ფერის ტუზი; B-ამოვიდა 8-ზე ნაკლები

რიცხვი; C-ამოვიდა წითელი ფერის 6-ზე მეტი რიცხვი; $B \bigcap C$; $B-C$; $C-B$.

7. ყუთში 10 ცალი თეთრი და 5 ცალი შავი ერთნაირი ზომის ბურთია. თეთრებს აწერია რიცხვები 1-დან 10-ის ჩათვლით, ხოლო შავებს – 11-დან 15-ის ჩათვლით. ყუთიდან ჩაუხედავად იღებენ ერთ ცალ ბურთს. რამდენი ელემენტარული ხდომილობისაგან შესდგებიან ხდომილობები: A-ამოვიდა ლუწი ნომრის ბურთი; B-ამოვიდა შავი ბურთი; C-ამოვიდა თეთრი ბურთი 3-ის ჯერადი ნომრით; D-ამოვიდა ბურთი, რომლის ნომერი ნაკლებია 14-ზე; $A \bigcap B$; $C-A$; $D-C$.

8. ლითონის მონეტას აგდებენ სამჯერ. ჩამოწერეთ ცდის ყველა შესაძლო ელემენტარულ ხდომილობათა სიმრავლე და ელემენტარული ხდომილობებით გამოსახეთ შემდეგი ხდომილობები:

$A-$ გერბი მოვიდა ერთხელ;

$B-$ გერბი მოვიდა ერთხელ მაინც;

$C-$ გერბი არ მოვიდა არც ერთხელ.

ამ ხდომილობებიდან არის თუ არა რომელიმე ელემენტარული ხდომილობა?

9. ლითონის მონეტას აგდებენ ოთხჯერ. რამდენი ელემენტარული შედეგი არსებობს, რომ: ა) გერბის მოსვლათა რაოდენობა მეტი იყოს საფასურის მოსვლათა რაოდენობაზე; ბ) გერბისა და საფასურის მოსვლათა რაოდენობები იყოს ტოლი?

10. ცდის პირი ასახელებს რაიმე ორნიშნა ნატურალურ რიცხვს ა) რამდენი ელემენტარული ხდომილობისაგან შედგებიან ხდომილობები, რომ დასახელებული რიცხვი: A – მეტია 20-ზე; B – ნაკლებია 50-ზე და 7-ის ჯერადია; C – მარტივია და ნაკლებია 27-ზე; D – 15-ზე მეტია და 5-ის ჯერადია; $A \bigcap B$; $B-C$; $C-D$; $D-B$. ბ) რომელი წყვილია არათავსებადი A, B, C და D ხდომილობებიდან?

11. მაგალითი 2-ის პირობებში წარმოადგინეთ ელემენტარულ ხდომილობათა სივრცე, როგორც ორი სიმრავლის დეკარტული ნამრავლი.

12. ორი პიროვნება შეთანხმდა შეხვედრაზე 10^{00} საათიდან 12^{00} საათამდე. თითოეული მიდის შეხვედრის ადგილას დროის ამ შუალედში. რა სიმრავლეა ცდის ყველა ელემენტარულ ხდომილობათა Ω სივრცე? გამოსახეთ ის სიბრტყეზე. წარმოადგინეთ Ω როგორც ორი სიმრავლის დეკარტული ნამრავლი (მითითება: OXY კოორდინატთა სისტემის ერთ ღერძზე გადაზომეთ პირველი პიროვნების მისვლის დრო, მეორეზე კი მეორის).

13. სამიზნეს ერთდროულად ესროლა სამმა ტყვიის მსროლელმა. A_i ხდომილობა ნიშნავდეს რომ i-ურმა $(i=1,2,3)$ მსროლელმა დააზიანა სამიზნე, ხოლო $\overline{A}_i$ – ააცილა. გამოსახეთ A_i და $\overline{A}_i$ ხდომილობების საშუალებით შემდეგი ხდომილობები:

B – სამიზნეს მოხვდა სამივე ტყვია;

C – სამიზნეს მოხვდა ორი ტყვია;

D – სამიზნეს არ მოხვდა არც ერთი ტყვია.

14. ცდა მდგომარეობს ლითონის მონეტის აგდებაში მანამდე, სანამ არ მოვა საფასური. აღნიშნეთ საფასურის მოსვლა 1-ით, ხოლო არ მოსვლა – 0-ით. ჩაწერეთ შემდეგი ხდომილობები

A – ცდა დამთავრდა 2 აგდებაში;

B – ცდა დამთავრდა 5 აგდებაში;

C – ცდა დამთავრდა 3-ზე ნაკლებ აგდებაში.

სასრულია თუ არა ცდის ელემენტარულ ხდომილობათა Ω სივრცე?

15. ყუთში მოთავსებულია 1-დან 12-მდე რიცხვებით გადანომრილი 4 ცალი წითელი, 2 ცალი შავი და 6 ცალი თეთრი ერთნაირი ზომის ბურთულა. წითლები ნომრებით 1, 2, 3, 4; შავები – 5, 6. ალალბედზე იღებენ ერთ ცალს. ბურთულების ნომრების საშუალებით ჩაწერეთ შემდეგი ხდომილობები:

A – არ ამოვიდა თეთრი ბურთულა;

B – ამოვიდა შავი ბურთულა;

C – ამოვიდა თეთრი ბურთულა 3-ის ჯერადი ნომრით;

D – ამოვიდა წითელი ბურთულა ლუწი ნომრით.

16. მოცემულია ხდომილობების უსასრულო $\{A_k\}_{k\in N}$ სიმრავლე, სადაც $A_k = \left[\frac{2k+1}{k}, \frac{5k-1}{k}\right]$. ჩაწერეთ შემდეგი ხდომილობები ინტერვალების სახით

$$\bigcup_{k=1}^{10} A_k\ ;\ \bigcap_{k=1}^{10} A_k\ ;\ \bigcup_{k=1}^{\infty} A_k\ ;\ \bigcap_{k=1}^{\infty} A_k\ .$$

17. მოცემულია ხდომილობების უსასრულო $\{A_k\}_{k\in N}$ სიმრავლე, სადაც $A_k = \left[5-\frac{1}{k}, 8-\frac{2}{k}\right)$. ჩაწერეთ შემდეგი ხდომილობები ინტერვალების სახით

$$A_1 - A_2\ ;\ A_2 - A_1\ ;\ \bigcup_{k=1}^{10} A_k\ ;\ \bigcap_{k=1}^{10} A_k\ ;\ \bigcup_{k=1}^{\infty} A_k\ ;\ \bigcap_{k=1}^{\infty} A_k\ .$$

18. ლითონის მონეტას აგდებენ სამჯერ. გერბის მოსვლა აღნიშნეთ 0-ით, საფასურის 1-ით და ჩაწერეთ ელემენ-ტარულ ხდომილობათა სივრცე და ხდომილობები:

A – პირველ და მეორე აგდებაზე არ მოვიდა გერბი;

B – მეორე აგდებაზე მოვიდა საფასური;

C – სამივე აგდებაზე მოვიდა მონეტის ერთი და იგივე მხარე;

D – მესამე აგდებაზე მოვიდა საფასური.

$A \bigcap B$; $A - C$; $C - D$; $A \bigcap D$; $B \bigcap C$; $B - C$.

19. გამოთვალეთ $\frac{P_5}{A_5^2}$; $\frac{A_6^2}{C_6^2}$; $A_6^2 + A_6^3$.

20. რისი ტოლია $\frac{A_5^3}{P_3}$ და $C_{10}^9 \cdot P_9$?

21. რამდენი სხვადასხვა ხუთკაციანი კალათბურთის გუნდი შეიძლება შევადგინოთ 8 კალათბურთელისაგან?

22. რამდენნაირად შეიძლება 8 ადამიანისგან შევადგინოთ 5 სხვადასხვა თანამდებობაზე დანიშნული ადამიანების ხუთეული?

23. რამდენი ელემენტისაგან შესდგება ხუთელემენტიანი სიმრავლის ყველა ქვესიმრავლეთა სიმრავლე?

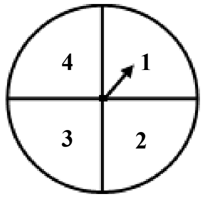
24. რამდენნაირად შეიძლება რიგში დავაყენოთ 4 ადამიანი?

პასუხები

1) 36; 15. 2) კი, კი, არა, კი, არა, კი. 3) $\{2,3);(3;2)\}$; $\{6,2);(2,6);(5,1);(1,5)\}$. 4) 6^4; 81; 81; 486. 5) 18; 16; 9; 10; 8; 5; 12. 7) 7; 5; 3; 13; 2; 2; 10. 8) *C*. 9) 5; 6. 10) 79; 7; 9; 16; 5; 6; 9; 15; *C* და *D*. 12) მართკუთხედია. 13) $B = A_1 \cap A_2 \cap A_3$; $C = A_1 \cap A_2 \cap \overline{A}_3 + A_1 \cap \overline{A}_2 \cap A_3 + \overline{A}_1 \cap A_2 \cap A_3$; $D = \overline{A}_1 \cap \overline{A}_2 \cap \overline{A}_3$. 14) $A = \{0,1\}$; $B = \{0,0,0,0,1\}$; $C = \{$„1“; „0,1“$\}$; არა. 15) $A = \{1,2,3,4,5,6\}$; $B = \{5,6\}$; $C = \{9,12\}$; $D = \{2,4\}$. 16) [2,1; 4,9]; [3, 4]; (2, 5); [3, 4]. 17) [4, 4,5); [6, 7]; [4; 7,8); [4,9; 6); [4, 8); (5, 6). 19) 6; 2; 150. 20) 10; A_{50}^{9}. 21) C_8^5. 22) A_8^5. 23) 32. 24) $P_4 = 4!$.

§ 17. ხდომილობის ალბათობა. ჯამისა და ნამრავლის ალბათობა

ყოველ ხდომილობას თავისი განხორციელების გარკვეული შანსი გააჩნია. მაგალითად, ცხადია, რომ კამათლის ერთხელ გაგორებისას ლუწი რიცხვის (ისევე როგორც-კენტის) მოსვლის პროცენტული შესაძლებლობა 50-ის ტოლია. თუ ამ პროცენტულ მაჩვენებლის 100-ზე გავყოფთ მივიღებთ $\frac{1}{2}$-ს. ასე, რომ ლუწი რიცხვის მოსვლის შანსი შეიძლება $\frac{1}{2}$-ის ტოლად ჩავთვალოთ.



ნახ. 17.1

აუცილებელი ხდომილობა ცდის ყოველ ჩატარებას მოჰყვება. მისი განხორციელების პროცენტული შესაძლებლობაა 100%, ამიტომ მისი განხორციელების შანსია 1. ბუნებრივია შეუძლებელი ხდომილობის განხორციელების შანსი არის 0.

ხდომილობის ალბათობა „გამოხატავს" ამ ხდომილობის განხორციელების „შანსს". ალბათობას აღნიშნავენ p ასოთი. A ხდომილობის ალბათობა აღინიშნება $P(A)$ სიმბოლოთი.

ნებისმიერი A ხდომილობის ალბათობა აკმაყოფილებს $0 \le P(A) \le 1$ პირობას, ამასთან $P(\varnothing)=0$ და $P(\Omega)=1$.

მაგალითი 17.1. ცდა მდგომარეობს ერთი კამათლის ერთხელ გაგორებაში. მოსალოდნელია „1", „2", „3", „4", „5" და „6" ელემენტარული ხდომილობების განხორციელება. ცხადია თუ კამათელი სიმეტრიულია (იდეალურია), მაშინ თითოეული რიცხვის მოსვლას ტოლი შანსი აქვს. რადგან მათი რაოდენობაა 6 და ისინი ურთიერთგამომრიცხავია ამიტომ თითოეული მათგანის მოსვლის „შანსი" ანუ ალბათობა იქნება $\frac{1}{6}$.

მაგალითი 17.2. ავიღოთ ოთხ ტოლ სექტორად დაყოფილი დისკო, რომლის ცენტრშიც ბრუნავს ისარი (ნახ. 17.1) (მას „რულეტკას“ უწოდებენ). ცდა მდგომარეობს ისრის დაბრუნებაში და შედეგად ითვლება ის დანაყოფი ანუ ის „ციფრი“, რომელზედაც ისარი გაჩერდება. ვინაიდან დანაყოფების სიდიდეები ტოლია, ამიტომ თითოეული ციფრის მოსვლის ალბათობა არის $\frac{1}{4}$.

ამ მაგალითებში ყველა შესაძლო ელემენტარული შედეგი იყო თანაბრად მოსალოდნელი. ფრანგმა მათემატიკოსმა **პ. ლაპლასმა** (1749-1827) აჩვენა, რომ თუ ცდის შედეგთა სრული ჩამონათვალი შეიცავს n თანაბრადმოსალოდნელ ელემენტარულ ხდომილობას, მაშინ თითოეული ასეთი ელემენტარული ხდომილობის ალბათობა $\frac{1}{n}$-ის ტოლია.

მოვიყვანოთ ალბათობის რამდენიმე თვისება:

1) $P(\varnothing)=0$

2) თუ A და B ხდომილობებია, მაშინ

$$P(A\bigcup B)=P(A)+P(B)-P(A\bigcap B).$$

დამტკიცება. ცხადია, რომ $A\bigcup B=A\bigcup(B-A)$ და $B=(B-A)\bigcup(A\bigcap B)$. ამიტომ

$$P(A\bigcup B)=P(A)+P(B-A); \quad P(B)=P(B-A)+P(A\bigcap B).$$

აქედან მარტივად მიიღება დასამტკიცებელი ტოლობა.

შედეგი 1. თუ A და B არათავსებადი ხდომილობებია $(AB=\varnothing)$, მაშინ

$$P(A\bigcup B)=P(A)+P(B).$$

შედეგი 2. თუ $A\subset B$, მაშინ

$$P(B-A)=P(B)-P(A).$$

შედეგი 3. $P(\overline{A})=1-P(A)$.

ცხადია მეორე თვისება შეიძლება ორზე მეტი სიმრავლისთვისაც ჩავწეროთ. მაგალითად, სამი A, B და C სიმრავლისათვის ჩაიწეროს შემდეგნაირად

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) -$$
$$- P(B \cap C) + P(A \cap B \cap C),$$

ხოლო შედეგი 1 შეიძლება განვაზოგადოთ ნებისმიერი სასრული n რაოდენობის $A_1, A_2, A_3, \dots, A_n$ წყვილ-წყვილად არათავსებადი ხდომილობებისთვის

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i).$$

მაგალითი 17.3. ფირმის მენეჯერს ახალი საქონლის შესაძენად სხვადასხვა კონტრაქტების გარდა 0,4 ალბათობით შეუძლია ისარგებლოს სასაქონლო ბირჟის მომსახურებით, ან 0,5 ალბათობით გააფორმოს საჭირო კომპანიასთან სალიზინგო ხელშეკრულება. ვიპოვოთ ალბათობა იმისა, რომ მოცემულ მომენტში ის არც ერთ გადაწყვეტილებას არ მიიღებს, თუ ცნობილია, რომ ერთდროულად ორივეთი სარგებლობს 0,2 ალბათობით.

ამოხსნა: შემოვიღოთ ხდომილობები:

A – მენეჯერი სარგებლობს სასაქონლო ბირჟით;

B – აფორმებს სალიზინგო ხელშეკრულებას;

C – არც ბირჟით სარგებლობს და არც ხელშეკრულებას აფორმებს.

ამოცანის პირობის თანახმად $P(A) = 0,4$; $P(B) = 0,5$; $P(A \cap B) = 0,2$ გვეკითხებიან $P(C)$-ს გამოთვლას.

ცხადია

$$C = \overline{A \cup B};$$

ხოლო

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0,4 + 0,5 - 0,2 = 0,7,$$

ამიტომ

$$P(C) = 1 - P(A \cup B) = 0,3.$$

ამოცანები

1. ცნობილია, რომ $P(A) = 0,5$, $P(B) = 0,3$ და $P(A \cap B) = 0,1$ გამოთვალეთ შემდეგი ხდომილებების ალბათობები $A \cup B$; $A - B$; $B - A$.

2. მოცემულია: $P(A)=0{,}6$; $P(A-B)=0{,}5$; $P(\overline{B})=0{,}8$. გამოთვალეთ ალბათობები: $P(A\bigcap B)$; $P(B-A)$.

3. A, B და C ერთი და იგივე ცდასთან დაკავშირებული ხდომილობებია. აუცილებლად სამართლიანია თუ არა ტოლობები?

$$P(A-B)=P(A)-P(B);$$
$$P(A\bigcup B)=P(A)+P(B);$$
$$P(A-B\bigcap C)=P(A)-P(A\bigcap B\bigcap C);$$
$$P(A\bigcup C)=P(A)+P(C-A).$$

4. მოცემულია $P(\overline{B})=0{,}4$, $P(\overline{C})=0{,}5$, $P(\overline{B\bigcap C})=0{,}7$. გამოთვალეთ ალბათობები $P(B)$; $P(B\bigcup C)$; $P(B\bigcap C)$.

5. ცდა მდგომარეობს ლითონის ფულის სამჯერ აგდებაში. რამდენი ელემენტისაგან შესდგება ელემენტარულ ხდომილებათა სივრცე? ყველა ქვესიმრავლეთა სიმრავლე?

6. A და B არათავსებადი ხდომილობებია. გამოთვალეთ ალბათობები $P(A\bigcup B)$ და $P(\overline{A}\bigcap \overline{B})$. თუ ცნობილია, რომ $P(A)=0{,}5$; $P(B)=0{,}2$.

7. A და B დამოუკიდებელი ხდომილობებია. გამოთვალეთ $P(A)$ და $P(A\bigcap B)$ ალბათობები, თუ $P(B)=0{,}7$, ხოლო $P(A\bigcup B)=0{,}9$.

8. მოცემულია $P(\overline{A})=0{,}4$; $P(\overline{A}\bigcap \overline{B})=0{,}1$ და $P(A-B)=0{,}45$. გამოთვალეთ ალბათობები $P(A)$, $P(B)$, $P(A\bigcap B)$, $P(B-A)$.

9. ცნობილია $A\subset B$ და $P(A)=0{,}5$, $P(B)=0{,}7$. გამოთვალეთ $P(B-A)$. დამოუკიდებელია თუ არა A და B ხდომილობები?

10. ცდა მდგომარეობს კამათლის ერთხელ გაგორებაში. მოცემულია ხდომილობები A-მოვიდა ლუწი რიცხვი, B-მოვიდა 4-ზე მეტი რიცხვი, C-მოვიდა 5-ის ჯერადი რიცხვი, D მოვიდა 5-ზე ნაკლები რიცხვი. ხდომილებათა რომელი წყვილებია არათავსებადი?

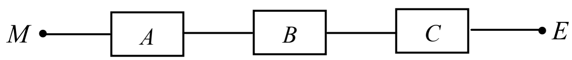
11. სასწრაფო სამედიცინო დახმარების ჯგუფის ღამის ცვლაში მორიგეობის დროს გამოძახება A რაიონიდან

შემოდის 0,3, ხოლო B რაიონიდან – 0,4 ალბათობით. ცნობილია, რომ ერთდროულად ორივე რაიონიდან გამოძახების შემოსვლის ალბათობაა 0,12. გამოთვალეთ ალბათობა იმისა, რომ ა) გამოძახება შემოვა A და B რაიონებიდან ერთიდან მაინც. ბ) გამოძახება შემოვა A და B რაიონებიდან მხოლოდ ერთიდან.

12. სტატიკურ მონაცემებზე დაკვირვებით დაადგინეს, სტატიკური ალბათობა იმისა, რომ რაჭაში ივლისის თვეში ნებისმიერი წინასწარ დაფიქსირებული დღე იქნება: ქარიანი – ტოლია 0,18-ის; წვიმიანი – 0,21; არც წვიმიანი და არც ქარიანი 0,66. იპოვეთ იმის ალბათობა, რომ 15 ივლისს წვიმაც მოვა და ქარიც იქნება.

13. ტყვიის მსროლელის მიერ ოთხი გასროლიდან სამიზნეში ერთხელ მაინც მოხვედრის ალბათობაა 0,9984. იპოვეთ შემდეგი ხდომილობების ალბათობები: A – ერთხელ გასროლით სამიზნეში მოხვედრა; B – ორი გასროლიდან სამიზნეში ერთხელ მაინც მოხვედრა. C – სამი გასროლიდან სამიზნეში ორჯერ მოხვედრა.

14. A, B და C სამი ხელსაწყო შეერთებულია მიმდევრობითი სქემით



ხელსაწყოები დენს ატარებენ შესაბამისად $P(A)=0{,}9$; $P(B)=0{,}95$; $P(C)=0{,}8$ ალბათობით. იპოვეთ იმის ალბათობა, რომ ME მონაკვეთზე დენი არ გაივლის.

15. ხელსაწყო შედგება ორი ნაწილისაგან. მათი მწყობრიდან გამოსვლის ალბათობებია შესაბამისად 0,05 და 0,02. იპოვეთ ხელსაწყოს მწყობრიდან გამოსვლის ალბათობა, თუ ის მუშაობს, როდესაც ორივე ნაწილი მწყობრშია.

16. ალბათობა იმისა, რომ სადისტრიბუციო კომპანია ფასდაკლების აქციის პერიოდში სავაჭრო ქსელში გაიტანს ტელევიზორების ახალ პარტიას ტოლია 0,6-ის, ხოლო ალბათობა იმისა, რომ გაიტანს პერსონალური კომპიუტერების ახალ პარტიას – 0,7-ის. იპოვეთ ალბათობა იმისა, რომ ფასდაკლების აქციის პერიოდში სადისტრიბუციო კომპანია

არ განაახლებს ქსელში არც ტელევიზორების და არც კომპიუტერების პარტიას. ცნობილია, რომ აქციის პერიოდში კომპანიის მიერ ერთდროულად ამ ორი დასახელების პროდუქციის განახლების ალბათობა 0,5-ის ტოლია.

17. ალბათობები იმისა, რომ კომპანია მოცემულ რეგიონში ჩადებს ინვესტიციის ჩაის კულტურის წარმოებაში, სწრაფი კვების ობიექტების ქსელში ან პურ-ფუნთუშეულის წარმოებაში შესაბამისად 0,25; 0,1 და 0,5-ის ტოლია. იპოვეთ იმის ალბათობა, რომ ხსენებული კომპანია ამ რეგიონში არც ერთ ინვესტიციას არ განახორციელებს თუ ცნობილია, რომ კომპანია ერთზე მეტ პროექტში ინვესტიციას არ დებს.

18. ცნობილია, რომ ფერმერი ვალუტის კურსის მკვეთრი მერყეობს პერიოდში ახალი ტექნიკის შეძენას ანხორციელებს 0,2 ალბათობით, ხოლო ყამირი მიწის ათვისებას 0,15 ალბათობით. იპოვეთ იმის ალბათობა, რომ ფერმერი ამ პერიოდში ან ახალ ტექნიკას შეიძენს ან ახალ მიწებს აითვისებს თუ მის მიერ ერთდროულად ამ ორი გადაწყვეტილების მიღების ალბათობა 0,1-ის ტოლია.

პასუხები

1) 0,7; 0,4; 0,2. 2) 0,1; 0,1. 3) არა; არა; კი; კი. 4) 0,6; 0,8; 0,3. 5) 8; 2^8. 6) 0,7; 0,3. 7) 2/3; 7/15. 8) 0,6; 0,45; 0,15; 0,3. 9) 0,2; არა. 10) *A* და *C*; *B* და *D*; *C* და *D*. 11) 0,58; 0,46. 12) 0,05. 13) 0,8; 0,96; 0,384. 14) 0,316. 15) 0,069. 16) 0,2. 17) 0,15. 18. 0,25.

§ 18. კლასიკური სქემა. ალბათობის კლასიკური განსაზღვრება

ავიღოთ კამათელი, რომლის წახნაგებზეც აწერია ციფრები: 1, 2, 3, 4, 5, 5. მისი გაგორების შედეგად მოსალოდნელია ხუთი ელემენტარული შედეგი: 1, 2, 3, 4 ან 5 მაგრამ ისინი ტოლშესაძლებელი არაა. ციფრი 5-ის მოსვლას ორი წახნაგი უწყობს ხელს სხვა ციფრებისას კი თითო. ჩვენ ასეთ ცდებს არ განვიხილავთ. შემდგომში შევისწავლით ისეთ ცდებთან დაკავშირებულ ხდომილობებს, რომელთაც გააჩნიათ სასრული რაოდენობის, ტოლშესაძლებელი, ურთიერთგამომრიცხავი ელემენტარული შედეგები. ასეთ ცდებზე ამბობენ, რომ ისინი აღიწერებიან **კლასიკური სქემით**. ეს ტერმინი უკავშირდება იმ ფაქტს, რომ აზარტულ თამაშებთან დაკავშირებული კლასიკური ამოცანები რომლებმაც ისტორიულად ალბათობის თეორიის საწყისებამდე მიგვიყვანეს სწორედ ასეთი სქემით აღიწერებიან.

განსაზღვრება 18.1. როდესაც ცდა აღიწერება **კლასიკური სქემით**, ნებისმიერი A ხდომილობის ალბათობა წარმოადგენს წილადს, რომლის მრიცხველია A-ს ხელშემწყობი ყველა ელემენტარული ხდომილებების რაოდენობა $N(A)$, ხოლო მნიშვნელი ცდის ყველა შესაძლო ელემენტარული ხდომილებების რაოდენობა $N(\Omega)$

$$P(A)=\frac{N(A)}{N(\Omega)}.$$

მაგალითი 18.1. ყუთში მოთავსებულია 4 შავი და 6 თეთრი ფერის ერთნაირი ზომის ბურთი. ალალბედზე იღებენ 5 ბურთს დაბრუნების გარეშე. ვიპოვოთ შემდეგი ხდომილობების ალბათობები:

A – ამოღებულია 2 შავი და 3 თეთრი ბურთი

B – ამოღებულია 5 შავი ბურთი

C – ამოღებულია 5 თეთრი ბურთი

D – ამოღებულ ბურთებში ერთი მაინც არის თეთრი

E – ამოღებულ ბურთებში ერთი მაინც არის შავი.

ამოხსნა. დავალაგოთ ჯერ შავი და შემდეგ თეთრი ბურთები და მიყოლებით გადავნომროთ რიცხვებით 1-დან 10-მდე

$$\underbrace{1,2,3,4}_{\text{შავები}} \qquad \underbrace{5,6,7,8,9,10}_{\text{თეთრები}}$$

ამ 10 ცალი ბურთიდან 5 ცალი ბურთის შესაძლო განსხვავებულ არჩევანთა რაოდენობა იქნება ჯუფთებათა რაოდენობა 10 დან 5-ად ე.ი. C_{10}^{5}. მაშასადამე

$$N(\Omega)=C_{10}^{5}.$$

A ხდომილობის განხორციელებისათვის 2 შავი ბურთი არჩეული უნდა იყოს პირველი 4 ცალიდან, ყველა განსხვავებულ არჩევათა რაოდენობაა C_4^2. 3 თეთრი ბურთი ბოლო ექვსიდან შეიძლება არჩეულ იქნეს C_6^3 ხერხით. სულ განსხვავებულ არჩევათა რაოდენობაა $N(A)=C_4^2\cdot C_6^3$, მაშასადამე,

$$P(A)=\frac{N(A)}{N(\Omega)}=\frac{C_4^2\cdot C_6^3}{C_{10}^5}=\frac{120}{252}=\frac{10}{21}.$$

ყუთში სულ 4 შავი ბურთია, რის გამოც 5 შავი ბურთის ამოღება შეუძლებელია. მაშასადამე B შეუძლებელი ხდომილობაა. ე.ი.

$$P(B)=0.$$

c-ხდომილობა განხორციელდება თუ ხუთივე ბურთი არჩეული იქნება ბოლო 6 ბურთიდან ასეთ განსხვავებულ არჩევათა რაოდენობაა $C_6^5=C_6^{6-5}=C_6^1=6$, ამიტომ

$$P(C)=\frac{N(C)}{N(\Omega)}=\frac{6}{252}=\frac{1}{42}.$$

ცხადია D აუცილებელი ხდომილობა, ვინაიდან ყუთში მხოლოდ 4 შავი ბურთია ამორჩეულ 5 ცალში აუცილებლად იქნება 1 თეთრი მაინც.

$$P(D)=1.$$

$\overline{E}$ ნიშნავს, რომ ამოღებულ 5 ბურთში არც ერთი არაა შავი. მაშასადამე $\overline{E}=C$, ცხადია რომ

$$P(E)=1-P(\overline{E})=1-P(C)=1-\frac{1}{42}=\frac{41}{42}.$$

მაგალითი 18.2. ვაგორებთ ორ კამათელს ერთდროულად. გამოვთვალოთ შემდეგ ხდომილობათა ალბათობები:

A – მოვიდა წყვილი

B – მოსული ციფრების ჯამია 7

C – მოსული ციფრების ჯამია 9

D – მოსული ციფრების ჯამი ნაკლებია 5-ზე

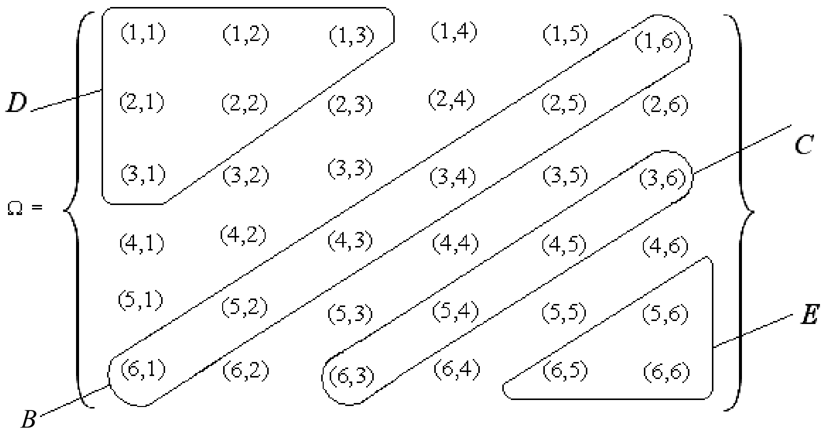
E – მოსული ციფრების ჯამი მეტია 10-ზე.

ამოხსნა. ცდის შედეგი იქნება რიცხვთა წყვილი (x, y), რომლის პირველი ელემენტი x არის პირველ კამათელზე მოსული ციფრი, ხოლო მეორე ელემენტი y კი მეორე კამათელზე მოსული ციფრი. სიმარტივისათვის ერთი კამათელი შევღებოთ მაგალითად წითლად და მას პირველი კამათელი დავარქვათ.

ვინაიდან თითოეულ კამათელზე შეიძლება 6 განსხვავებული შედეგი მოვიდეს, ამიტომ ცდის ყველა შესაძლო ელემენტარულ შედეგთა სიმრავლე შესდგება 36 წყვილისგან.

$$\Omega = \{(x, y) \mid 1 \le x, y \le 6,\ x \in N,\ y \in N\}.$$

ჩამოვწეროთ Ω სიმრავლე ცხრილის სახით და მოვნიშნოთ მასზე ამოცანაში ნახსენები ხდომილობების ხელშემწყობი ვარიანტები



ცხრ. 18.1.

შევნიშნოთ, რომ მაგალითად (2,3) და (3,2) ცდის სხვადასხვა შედეგებია (ამაში მარტივად დავრწმუნდებით თუ თამაშის დროს მოვითხოვთ რომ ჯერ წითელ კამათელზე მოსული რიცხვი ვითამაშოთ).

ცხადია A ხდომილობის ხელშემწყობი ვარიანტებია (1,1) (2,2) (3,3) (4,4) (5,5) და (6,6) დანარჩენი ხდომილობების ხელშემწყობი ვარიანტები შეიძლება დავთვალოთ 18.1 ცხრილიდან

$$P(A)=\frac{6}{36};\quad P(B)=\frac{6}{36};\quad P(C)=\frac{4}{36};\quad P(D)=\frac{6}{36};\quad P(E)=\frac{3}{36}.$$

ამოცანები

1. ყუთში 12 ერთნაირი ბურთულაა, რომელთაგან 7 შავია, 3 წითელი და 2 თეთრი. ყუთიდან ჩაუხედავად ალალბედზე იღებენ 1 ცალ ბურთულას. იპოვეთ ალბათობა იმისა, რომ ამოღებული ბურთულა: ა) შავი ფერისაა; ბ) ან შავი ან თეთრი ფერისაა; გ) არაა თეთრი ფერის; დ) მწვანე ფერისაა; ე) არაა ყვითელი ფერის.

2. ყუთში 16 ერთნაირი ბურთულაა, რომელთაგან 6 შავია, 4 წითელი, დანარჩენი ბურთულებიდან თითოეული ნახევრად შავი და ნახევრად წითელია. შემთხვევითად იღებენ ერთ ცალ ბურთულას. იპოვეთ ალბათობა იმისა, რომ: ა) ამოღებულია წითელი ბურთულა; ბ) ამოღებულ ბურთულაზე წითელი ფერია; გ) ამოღებულია ან შავი ან წითელი ბურთულა; დ) ამოღებულ ბურთულაზე შავი ან წითელი ფერია.

3. კლასში მოსწავლეებიდან 15 გოგონა და 10 ბიჭია. ცისფერი თვალები აქვთ ბიჭების 20 პროცენტს და გოგონების 40 პროცენტს. სიის მიხედვით შემთხვევითი პრინციპით ამოარჩიეს ერთი მოსწავლე. იპოვეთ ალბათობა იმისა, რომ ამორჩეული მოსწავლე: ა) ბიჭია; ბ) ცისფერთვალებაა; გ) ცისფერთვალება გოგონაა; დ) არაა ცისფერთვალება ბიჭი.

4. ყუთში 6 შავი და 4 თეთრი ერთნაირი ბურთულაა. ჩაუხედავად შემთხვევითად იღებენ 2 ცალს. იპოვეთ შემდეგი ხდომილობების ალბათობები A – ორივე თეთრია, B – ბურთულები სხვადასხვა ფერისაა; C – ორივე ერთი ფერისაა.

5. ყუთში 5 შავი, 4 თეთრი და 6 წითელი ფერის ერთნაირი ბურთულაა. ყუთიდან მასში ჩაუხედავად შემთხვევითად ირჩევენ 7 ცალს. იპოვეთ ალბათობა იმისა, რომ ამორჩეულ ბურთულებში: ა) ზუსტად 3 ცალია თეთრი ფერის; ბ) ორი შავი ფერისაა, დანარჩენი წითელი; გ) 5 თეთრი ფერის ბურთულაა; დ) ყველა ბურთულა ერთნაირი ფერის არ არის; ე) მხოლოდ ორი განსხვავებული ფერის ბურთულაა.

6. ბადრი და ბექა ერთდროულად აგორებენ თითო ცალ კამათელს. იპოვეთ შემდეგი ხდომილობების ალბათობები: A – ბადრის კამათელზე მოვიდა უფრო დიდი ციფრი, ვიდრე ბექასაზე; B – ორივეს მოუვიდა ერთნაირი ციფრები; C – ბექას კამათელზე მოსული ციფრი ორჯერ მეტია, ვიდრე ბადრის კამათელზე მოსული.

7. საწარმოს გამოშვებული დეტალებიდან 1 პროცენტი დეფექტურია. სავაჭრო ქსელში გატანამდე, შემოწმების შედეგად ხერხდება დეფექტური დეტალების 50 პროცენტის პოვნა და პარტიიდან ამოღება. იპოვეთ იმის ალბათობა, რომ ამის შემდეგ დეტალების პარტიიდან შემთხვევითად აღებული 1 ცალი იქნება დეფექტური.

8. ყუთში მოთავსებულია მუყაოს ქაღალდებზე სათითაოდ დაწერილი ათივე ციფრი. ჩაუხედავად ერთმანეთის მიყოლებით ორჯერ იღებენ თითო ქაღალდს. იპოვეთ იმის ალბათობა, რომ მეორედ ამოღებული ციფრი იქნება უფრო დიდი, ვიდრე პირველად ამოღებული. განიხილეთ ორი შემთხვევა: ა) ამოღება ხდება დაბრუნების გარეშე; ბ) ამოღება ხდება დაბრუნებით.

9. რვასართულიანი სახლის ლიფტში პირველ სართულზე შევიდა სამი ადამიანი. იპოვეთ შემდეგი ხდომილობების ალბათობები: A – სამივე გამოვა მეოთხე სართულზე. B – ამისვე გამოვა ერთი და იგივე სართულზე; C – სამივე გამოვა სხვადასხვა სართულზე.

10. განიხილეთ ისევ მე-9 ამოცანა იმ შემთხვევაში, როდესაც ლიფტი არ აჩერებს მეორე სართულზე.

11. ქართული ანბანის ასოები სათითაოდ დაწერილია ერთნაირი ზომის მუყაოს ქაღალდის ნაჭრებზე და ჩაყრი-

ლია ყუთში. ყუთიდან შემთხვევითად იღებენ 5 ცალს. იპოვეთ შემდეგი ხდომილობების ალბათობები: A – ამოღებულია მხოლოდ თანხმოვნები; B – ამოღებულია მხოლოდ ხმოვნები; C – ამოღებულ ასოებში ორი ხმოვანია.

12. 4 მგზავრი ერთმანეთისაგან დამოუკიდებლად აპირებს ჩასხდომას მატარებლის ექვს ვაგონიან შემადგენლობაში. იპოვეთ შემდეგი ხდომილობების ალბათობები: A – ოთხივე სხვადასხვა ვაგონში მოთავსდება, B – ყველა მოხვდება ერთ ვაგონში, C – ყველა მოხვდება მეორე ვაგონში, D – მხოლოდ ორი მოხვდება ერთ ვაგონში, დანარჩენი კი სხვადასხვაში.

13. ჯგუფში 15 სტუდენტიდან 10 ვაჟია და 5 გოგონაა. ცნობილია, რომ გამოცდა წარმატებით ჩააბარა შვიდმა სტუდენტმა. გამოთვალეთ შემდეგი ხდომილობების ალბათობები: A – ჩაბარებულთა შორის 4 გოგონაა და 3 ვაჟი; B – ყველა გოგონამ ჩააბარა გამოცდა; C – მხოლოდ გოგონებმა ჩააბარეს გამოცდა, ვაჟებმა კი ვერა; D – გამოცდა ჩააბარა მინიმუმ ორმა ვაჟმა.

14. ყუთში მოთავსებულია 9 ცალ მუყაოს ქაღალდზე სათითაოდ დაწერილი ციფრები 1-დან 9-ის ჩათვლით. მიმდევრობით სათითაოდ იღებენ 4 ცალ ქაღალდს და ამოღებულებს ალაგებენ რიგის მიხედვით. იპოვეთ ალბათობა იმისა, რომ მიღებული ოთხნიშნა რიცხვი: ა) შესდგება მხოლოდ კენტი ციფრებისაგან; ბ) ლუწია; გ) შესდგება ერთმანეთის მომდევნო ციფრებისაგან.

15. მე-14-ე ამოცანაში იგულისხმეთ, რომ ინიშნავენ ამოღებული ციფრს და მუყაოს ქაღალდს აბრუნებენ უკან ყუთში. გამოთვალეთ იგივე ხდომილობების ალბათობები.

16. ბადრი, ბექა, ნანა, ციალა და ია უნდა დასხდნენ მართკუთხოვანი მაგიდის ერთ მხარეს ხუთ ადგილზე ერთმანეთის გვერდით. ისინი ადგილს ირჩევენ შემთხვევითი პრინციპით. იპოვეთ ალბათობა იმისა, რომ:

ა) ბექა და ციალა მოხვდებიან ერთმანეთის გვერდით;

ბ) მანდილოსნები მოხვდებიან თითო ადგილის გამოტოვებით.

17. სხვადასხვა მისამართზე გასაგზავნი ოთხი წერილი ჩადეს წინასწარ დამისამართებულ კონვერტებში არეულად ისე, რომ მისამართებზე არ დაუხედავთ. იპოვეთ ხდომილობების ალბათობები: A – ოთხივე წერილი მივა ზუსტ მისამართზე; B – არც ერთი წერილი არ მივა ზუსტ მისამართზე; C – ერთი წერილი მაინც მივა ზუსტ მისამართზე.

პასუხები

1) $7/12$; $3/4$; $5/6$; 0; 1. 2) $1/4$; $5/8$; $5/8$; 1. 3) $2/5$; $8/25$; $6/25$; $23/25$. 4) $2/15$; $8/15$; $7/15$. 5) $C_4^3 C_{11}^4 / C_{15}^7$; $C_5^2 C_6^5 / C_{15}^7$; 0; 1; $(C_9^7 + C_{10}^7 + C_{11}^7)/C_{15}^7$. 6) $5/12$; $1/6$; $1/12$. 7) $1/199$. 8) $1/2$; $9/20$. 9) $1/343$; $1/49$; $A_7^3/7^3$. 10) $1/216$; $1/36$; $A_6^3/6^3$. 11) C_{28}^5/C_{33}^5; $1/C_{33}^5$; $C_{28}^3 C_5^2/C_{33}^5$. 12) $A_6^4/6^4$; $1/6^3$; $1/6^4$; $A_5^2/6^3$. 13) $C_5^4 C_{10}^3/C_{15}^7$; C_{10}^2/C_{15}^7; 0; 1. 14) A_5^4/A_9^4; $4A_8^3/A_9^4$; $6/A_9^4$. 15) $5^4/9^4$; $4/9$; $6/9^4$. 16) $8 \cdot 3!/5!$; $2!3!/5!$. 17) $1/24$; $3/8$; $5/8$.

§ 19. გეომეტრიული ალბათობა

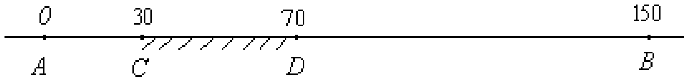
ზოგიერთ შემთხვევაში ცდის ყველა შესაძლო შედეგთა სიმრავლე ვერ დაითვლება რაოდენობრივად, მაგრამ იგი შეიძლება გარკვეულ გეომეტრიულ ფიგურასთან გავაიგივოთ. ჩვენთვის საინტერესო ხდომილობის ხელშემწყობ ელემენტარულ ხდომილობათა სიმრავლე აღმოჩნდება ამ ფიგურის რაიმე ნაწილი და მისი ალბათობა **„გეომეტრიული წესით“** გამოითვლება.

ცდის ყველა შესაძლო ელემენტარულ ხდომილობათა Ω სივრცე შეიძლება გავაიგივოთ რაიმე მონაკვეთთან (წრფეზე), ფიგურასთან (სიბრტყეზე) ან სხეულთან (სივრცეში), ხოლო A ხდომილობის ხელშემწყობი ელემენტარული შედეგების სიმრავლე იქნება Ω-ს რაიმე ნაწილი. ასეთ შემთხვევაში $P(A)$ ალბათობა გამოითვლება როგორც A და Ω-ს შესაბამისი სიგრძეების, ფართობების ან მოცულობების შეფარდება იმის მიხედვით თუ რას წარმოადგენდა Ω (მონაკვეთს, ფიგურას თუ სხეულს).

$$P(A) = \frac{|A|}{|\Omega|}.$$

მაგალითი 19.1. დაზიანდა A და B ქალაქების შემაერთებელი 150კმ სიგრძის სატელეფონო ხაზი მხოლოდ ერთ ადგილას. შემკეთებელთა ჯგუფის გაგზავნა მოხერხდა ხაზის მხოლოდ 30-ე და 70-ე კილომეტრებს შორის მონაკვეთზე. რა არის იმის ალბათობა, რომ კავშირი აღდგება?

ამოხსნა. x-ით აღვნიშნოთ დაზიანების ადგილის დაშორება A ქალაქიდან. ცხადია, $\Omega = \{x \mid 0 < x < 150\}$. შემოვიღოთ ხდომილობა E - „კავშირი აღდგა“. ცხადია, ეს ხდომილობა განხორციელდება მხოლოდ მაშინ, თუ დაზიანება მოხდა C და D წერტილებს შორის (იხ. ნახ. 19.1) ანუ იმ მონაკვეთზე, რომელიც შეამოწმა შემკეთებელთა ჯგუფმა.



ნახ. 19.1.

ე.ი.

$$P(E)=\frac{|CD|}{|\Omega|}=\frac{70-30}{150-0}=\frac{4}{15}\,.$$

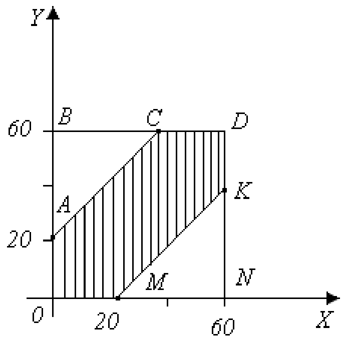
მაგალითი 19.2. (ამოცანა შეხვედრაზე). ორი პიროვნება შეთანხმდა შეხვედრაზე დათქმულ ადგილას 10-საათიდან 11-მდე. ამასთან, შეხვედრის ადგილზე მისული ელოდება მეორეს მხოლოდ 20 წუთს. ვიპოვოთ შეხვედრის ალბათობა, თუ ყოველი მათგანი მიდის შეხვედრის ადგილას დათქმული დროის განმავლობაში ალალბედზე და მათი მისვლის დროები ერთმანეთისაგან დამოუკიდებელია.

ამოხსნა. ავიღოთ OXY კოორდინატთა სიბრტყეზე $(x\ \ y)$ წყვილების სიმრავლე, სადაც x არის პირველი პიროვნების მისვლის დროის წუთების მაჩვენებელი 10-დან ათვლილი, ხოლო y-მეორის. მივიღებთ ამ პიროვნებების დათქმულ ადგილას მისვლის ყველა შესაძლო ვარიანტებს – $(x,\ y)$ კოორდინატების წყვილთა სიმრავლის სახით.

$$\Omega=\{(x,y)\,|\,0\le x\le 60\,,\ \ 0\le y\le 60\}\,.$$

E იყოს ხდომილობა - შეხვედრა შედგა. ცხადია

$$E=\{(x,y):|x-y|<15,\ \ (x,y)\in\Omega\}\,.$$



ნახ. 19.2

შეხვედრას ხელს უწყობს იმ $(x\ \ y)$ წერტილთა სიმრავლე Ω-დან, რომელთათვისაც

$$|x-y|\le 15 \quad \text{ანუ} \quad \begin{cases} y\le x+15\\ y\ge x-15\\ 0\le x\le 60\\ 0\le y\le 60\end{cases}$$

ნახაზზე (იხ. ნახ. 19.2) გამოსახულია Ω ელემენტარულ

ხდომილობათა სივრცე, რომელზეც დაშტრიხულია შეხვედრის E ხდომილობის შესაბამისი წერტილების სიმრავლე, რომელიც $OACDKM$ ექვსკუთხედია.

$$P(E)=\frac{S_{OACDKM}}{S_{OBDN}}=\frac{60^2-40^2}{60^2}=\frac{5}{9}.$$

აქვე გვინდა შევნიშნოთ, რომ გეომეტრიული ალბათობის გამოყენებით ამოცანების ამოხსნისას სიფრთხილე უნდა გამოვიჩინოთ ელემენტარულ ხდომილობათა სივრცის აგების დროს. წინააღმდეგ შემთხვევაში შეიძლება ისე მოხდეს, რომ ჩვენი ამოხსნა არ შეესაბამებოდეს დასმულ ამოცანას. მოვიყვანოთ მაგალითი.

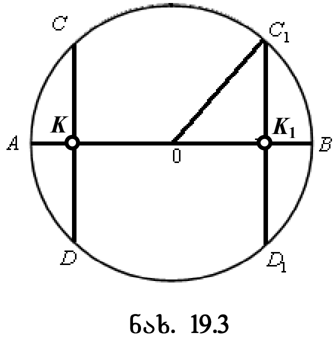
მაგალითი 19.3. ვიპოვოთ იმის ალბათობა, რომ წრეში მოცემული დიამეტრის მართობულად გავლებული ქორდის სიგრძე რადიუსზე მეტი იქნება.

მოვიყვანოთ ამ ამოცანის ორი განსხვავებული ამოხსნა და გავარკვიოთ რომელია (და რა მიზეზის გამო) მათგან მცდარი. M იყოს ხდომილობა: „მოცემული დიამეტრის მართობულად გავლებული ქორდის სიგრძე რადიუსზე მეტია“.

ამოხსნა 1. დავაფიქსიროთ R რადიუსიან წრეში AB დიამეტრი $|AB|=2R$. მის ნებისმიერ x წერტილზე შეიძლება დიამეტრის მართობულად ერთადერთი ქორდის გავლება. x იყოს ამ ქორდის „კოორდინატი“ 0 ცენტრის მიმართ.

$$\Omega=\{x\,|\,-R<x<R\}.$$

გავავლოთ დიამეტრის მართობული და რადიუსის სიგრძის ტოლი ორი ქორდა CD და C_1D_1. მათი გადაკვეთის წერტილები AB დიამეტრთან იყოს შესაბამისად K და K_1 (ნახ. 19.3).



ნახ. 19.3

$$OC_1=C_1D_1=CD=R\Rightarrow OK_1=OK=\frac{R\sqrt{3}}{2}.$$

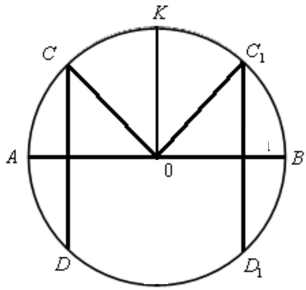
ცხადია იმ ქორდის სიგრძე,

რომელიც ცენტრიდან დაცილებულია $\dfrac{R\sqrt{3}}{2}$-ზე ნაკლები მანძილით, ანუ გავლებულია (K, K_1) ინტერვალის რომელიმე წერტილში მის მართობულად, მეტი იქნება რადიუსზე. მაშასადამე,

$$M = \left\{ x : -\frac{R\sqrt{3}}{2} < x < \frac{R\sqrt{3}}{2} \right\},$$

$$P(M) = \frac{|KK_1|}{|AB|} = \frac{\sqrt{3}}{2}.$$

ამოხსნა 2. R რადიუსიან წრეში დავაფიქსიროთ $2R$ სიგრძის AB დიამეტრი. მისი O ცენტრიდან აღვმართოთ მართობული რადიუსი OK (ნახ. 19.4). ნებისმიერი ქორდა, რომელსაც AB დიამეტრის მართობულად გავავლებთ, გადაიკვეთება AKB ნახევარწრეწირთან. გადაკვეთის წერტილის დაცილება K წერტილიდან რკალის გასწვრივ იყოს x. ცხადია x ცალსახად განსაზღვრავს ქორდის სიგრძეს. ვინაიდან $\ell_{\breve{AKB}} = \pi R$, ამიტომ



ნახ. 19.4

$$\Omega = \breve{AKB}, \quad \Omega = \left\{ x : -\frac{\pi R}{2} < x < \frac{\pi R}{2} \right\}.$$

გავავლოთ O ცენტრის ორივე მხარეს AB დიამეტრის მართობულად რადიუსის სიგრძის CD და C_1D_1 ქორდები. შევაერთოდ C და C_1 წერტილები O ცენტრთან. ნახევარწრეწირის CKC_1 რკალის წერტილებიდან დიამეტრის მართობულად გავლებული ყოველი ქორდის სიგრძე მეტი იქნება რადიუსზე. $\angle C_1OK = 60° \Rightarrow \angle COC_1 = 120°$.

ცხადია

$$\ell_{\breve{kc_1}} = \frac{\pi R}{3}.$$

ამიტომ

$$M=\left\{x:-\frac{\pi R}{3}<x<\frac{\pi R}{3}\right\},$$

$$P(M)=\frac{\ell_{C\breve{K}C_1}}{\ell_{A\breve{K}B}}=\frac{2}{3}.$$

მივიღეთ ორი სხვადასხვა ამონახსნი. ამოცანის პირობიდან გამომდინარე პირველი ამონახსნი ზუსტია. მეორე ამონახსნი $P=\frac{2}{3}$ არაა ზუსტი, ვინაიდან მეორე ამოხსნაში Ω სივრცის აგება არ არის ჩატარებული ამოცანის პირობის შესაბამისად. საქმე იმაშია რომ მართალია AB დიამეტრსა და AKB რკალის წერტილებს შორის ურთიერთცალსახა შესაბამისობაა, მაგრამ ის მანძი-ლებს არ ინარჩუნებს. მარტივად შეგვიძლია შევამოწმოთ, რომ ამ შესაბამისობის დროს რადიუსის შუაზე გაყოფის შედეგად მის ორ ტოლ ნახევარს სხვადასვა სიგრძის რკალები შეესაბამება, ამიტომ დიამეტრის მართობულად გავლებული ქორდის სიგრძის დახასიათება დიამეტრის შესაბამისი ნახევარწრეწირის წერტილების მიხედვით არ პასუხობს ამოცანის პირობას.

მაგალითი 19.4. წრეწირის ფიქსირებული წერტილიდან გავლებულია ქორდა. ვიპოვოთ იმის ალბათობა, რომ ქორდის სიგრძე რადიუსზე მეტია.

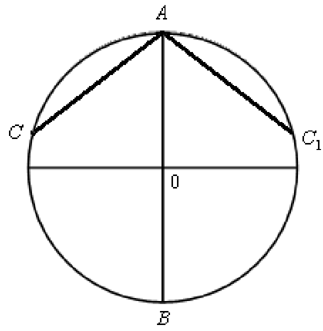
ამოხსნა. M იყოს ხდომილობა: R რადიუსიანი წრეწირის ფიქსირებული წერტილიდან გავლებული ქორდა რადიუსზე გრძელია. წრეწირზე ავიღოთ A წერტილი და გავალოთ $AB=2R$ სიგრძის დიამეტრი. შემდეგ A წერტილიდან გავლებულ ნებისმიერ ქორდას შევუსაბამოთ φ კუთხე AB დიამეტრსა და ამ ქორდას შორის.

$$\Omega=\{\varphi\,|-90°<\varphi<90°\}.$$

ეხლა A წერტილიდან AB დიამეტრის ორივე მხარეს გავავლოთ R რადიუსის ტოლი ქორდები AC და AC_1 (ნახ. 19.5). ცხადია ყოველი ქორდა რომელიც CAC_1 კუთხეში

გაივლება რადიუსზე მეტი სიგრძის იქნება $\angle CAC_1 = 120°$, ამიტომ

$$P(M) = \frac{120}{180} = \frac{2}{3}.$$



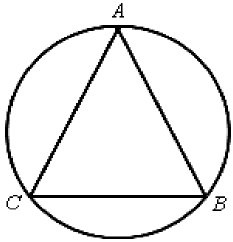
ნახ. 19.5

გეომეტრიული ალბათობით ამოცანების ამოხსნისას შემთხვევითი სიდიდეების შესაძლო მნიშვნელობები გარკვეული გეომეტრიული ობიექტებია (წერტილები, წრფეები და ა.შ.). მათთვის ალბათობის „მიწერა“ მოითხოვს გარკვეულ სიზუსტეს. რაიმე ხდომილობის შესაბამისი არასწორად განსაზღვრული ასეთი ობიექტების სიმრავლისათვის ალბათობის გამოთვლამ შეიძლება „პარადოქსამდეც“ კი მიგვიყვანოს. სინამდვილეში ეს „პარადოქსი“ სხვა არაფერია თუ არა ამოცანის პირობის არა მკაფიოდ და არა ცალსახად დასმის შედეგი. მოვიყვანოთ ერთი ასეთი ისტორიულად ცნობილი „პარადოქსი“, რომელიც ჟ. ბერტრანმა (1907) განიხილა და მისი სახელითაა ცნობილი.

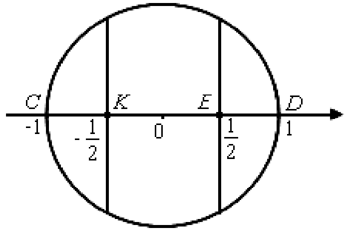
ბერტრანის პარადოქსი (იხ. [14]). ვიპოვოთ ალბათობა იმისა, რომ ერთეულრადიუსიან წრეში შემთხვევითად გავლებული ქორდის სიგრძე მეტი იქნება $\sqrt{3}$-ზე. განვიხილოთ ამ ამოცანის სამი ამოხსნა. A იყოს ხდომილობა: ქორდის სიგრძე მეტია $\sqrt{3}$-ზე.

ამოხსნა 1. ყოველი ქორდა ხასიათდება წრეწირთან გადაკვეთის ორი წერტილით. ვგულისხმობთ, რომ ერთ-ერთი მათგანი არის წრეში ჩახაზული წესიერი ABC სამკუთხედის A წვერო (ნახ. 19.6), მაშინ მეორე წვერომ შეიძლება წრეწირზე ნებისმიერი სხვა ადგილი დაიკავოს. ცხადია, რომ ქორდის სიგრძე მეტი იქნება $\sqrt{3}$-ზე თუ მეორე წვერო მოთავსდება მცირე CB რკალზე. მაშასადამე

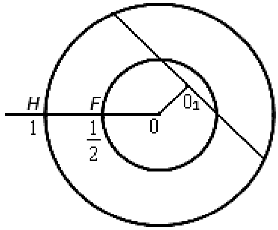
$$P(M) = \frac{\ell_{\breve{CB}}}{\ell_{წრ}} = \frac{1}{3}.$$



ნახ. 19.6



ნახ. 19.7



ნახ. 19.8.

ამოხსნა 2. ქორდის სიგრძე დამოკიდებულია წრის ცენტრიდან მის დაშორებაზე და არაა დამოკიდებული ქორდის მიმართულებაზე, შეგვიძლია ჩავთვალოთ, რომ ქორდას აქვს რაიმე ფიქსირებული მიმართულება და მისი დახასიათება მოხდება ამ მიმართულების მართობულ დიამეტრთან თანაკვეთის წერტილით. ეს წერტილი CD დიამეტრზე მდებარეობს (ნახ. 19.7). იმისათვის რომ ქორდის სიგრძე მეტი იყოს $\sqrt{3}$-ზე საჭიროა ქორდა O ცენტრიდან დაშორებული იყოს $\frac{1}{2}$-ზე ნაკლები მანძილით ე.ი. $P(M) = \frac{|KE|}{|CD|} = \frac{1}{2}$.

ამოხსნა 3. ყოველი ქორდის სიგრძე ცალსახად განისაზღვრება წრეწირის O ცენტრიდან მასზე დაშვებული მართობის O_1 ფუძის O ცენტრიდან დაშორებით. ეს O_1 წერტილი მოცემულ OH რადიუსთან წრეში ნებისმიერ ადგილას შეიძლება მდებარეობდეს, ხოლო ქორდის სიგრძე

მეტი იქნება $\sqrt{3}$-ზე, თუ ეს წერტილი მოხვდა მოცემული წრეწირის კონცენტრული $OF=\frac{1}{2}$ რადიუსის მქონე წრეწირის შიგნით (ნახ. 19.8). მაშასადამე

$$P(M)=\frac{S_{(O,OF)}}{S_{(O,OH)}}=\frac{1}{4}.$$

სამივე ეს ამოხსნა სწორია, მაგრამ სამივე მათგანი ამოცანის სხვადასხვა დასმას შეესაბამება. საქმე იმაშია, რომ ჩვენ მკაცრად უნდა განვსაზღვროთ რა გვესმის მოცემული ამოცანის პირობაში ნათქვამი „შემთხვევითად გავლებული ქორდის" ქვეშ.

პირველი ამოხსნა ეხება: წრეწირზე შემთხვევითად აღებულ ორ წერტილზე გავლებულ ქორდას. მეორე წრეწირის მოცემული დიამეტრის მართობულად შემთხვევითად გავლებულ ქორდას შევადაროთ (მაგალითი 19.4)-ს. მესამე კი წრეწირის ცენტრიდან შემთხვევითი მანძილით დაშორებულ ქორდას.

ერთი ანალოგიური ტიპის ამოცანის ამოხსნას მკითხველს მივანდობთ: წრეში შემთხვევითად აღებულ წერტილზე ავლებენ ქორდას. გამოთვალეთ ალბათობა იმისა, რომ ქორდის სიგრძე რადიუსზე მეტია (პასუხი $\frac{3}{4}$).

ბოლოს გავაკეთოთ ერთი შენიშვნა, რომელიც ეხება განსხვავებას კლასიკური ალბათობის ფორმულით გამოთვლის მეთოდსა და გეომეტრიული ალბათობის ფორმულით გამოთვლის მეთოდებს შორის.

კლასიკური ალბათობის ფორმულით გამოთვლის დროს ვიცით, რომ აუცილებელი ხდომილობის ალბათობა არის 1, ხოლო შეუძლებლის კი – 0. ამ დროს სამართლიანია შებრუნებული დებულებაც: თუ ალბათობაა 0, ხდომილობა შეუძლებელია, ხოლო თუ 1 – აუცილებელი. მართლაც: ვინაიდან

$$P(A)=\frac{N(A)}{N(\Omega)}.$$

ამიტომ თუ $P(A)=0$, მაშინ $N(A)=0$ და A ხდომილობის მოხდენას ცდის არც ერთი ელემენტარული შედეგი არ უწყობს ხელს ე.ი. A შეუძლებელი ხდომილობაა.

ანალოგიურად თუ $P(A)=1$, ე.ი. $N(A)=N(\Omega)$ და ეს ნიშნავს, რომ A-ს ხელშემწყობ ელემენტარულ ხდომილობათა რაოდენობა ტოლია ცდის ყველა შესაძლო ელემენტარულ ხდომილობათა რაოდენობის. ე.ი. არ არსებობს ელემენტარული ხდომილობა, რომელიც A-ს ხელს არ უწყობს. ე.ი. A აუცილებელი ხდომილობაა.

გეომეტრიული ალბათობის ფორმულით გამოთვლის დროს ეს შებრუნებული დებულებები სამართლიანი არაა. მაგალითად, თუ წერტილი შემთხვევითად ვარდება წრეში, ის შეიძლება მოხვდეს წრის რაიმე $A(x_0,y_0)$ წერტილში. მეორეს მხრივ ამ წერტილში (ისევე, როგორც ნებისმიერ სხვა წერტილში) მოხვედრის გეომეტრიული ალბათობა ნულის ტოლია.

იგივე ხდება აუცილებელი ხდომილობის შემთხვევაშიც.

ალბათობა იმ ხდომილობისა, რომ წრეში ჩავარდნილი წერტილი წრეწირის შიგნით მოხვდება არის 1. მაგრამ ეს ხდომილობა არ არის აუცილებელი, ვინაიდან წერტილი შეიძლება ზედ წრეწირზე მოხვდეს.

მკითხველს ვაძლევთ ერთ მარტივ რჩევას: გეომეტრიული ალბათობის გამოყენებით ამოცანის ამოხსნისას, საჭიროა გავითვალისწინოთ თუ რამდენი შემთხვევითი მომენტია ამოცანის პირობაში ნაგულისხმევი. თუ შემთხვევითი მომენტი ერთია (იხ. მაგალითი 1), მაშინ Ω წარმოადგენს წრფის ნაწილს (მონაკვეთს), თუ ორია (მაგალითები 2–5), მაშინ Ω არის სიბრტყის ნაწილი (გარკვეული ბრტყელი ფიგურა), ხოლო თუ სამია – სივრცის ნაწილი, რაიმე „მოცულობითი ფიგურა“.

ამოცანები

1. კვადრატის წვეროები მდებარეობენ R – რადიუსიან წრეწირზე. ცნობილია, რომ წრეში შემთხვევით დასვეს წერტილი. რას უდრის იმის ალბათობა, რომ წერტილი კვადრატში მოხვდება?

2. სამკუთხედში შემთხვევით დასვეს წერტილი. რას უდრის იმის ალბათობა, რომ წერტილი მოხვდება ამ სამკუთხედის გვერდების შუა წერტილების შეერთებით მიღებული სამკუთხედის შიგნით?

3. მოცემულია ორი კონცენტრული წრეწირი რადიუსებით: 3 და 5სმ. დიდ წრეში შემთხვევით დასვეს წერტილი. გამოთვალეთ იმის ალბათობა, რომ ეს წერტილი არ მოხვდება პატარა წრეში.

4. ბაქტერია მოხვდა მუხუდოს მარცვალში. რას უდრის იმის ალბათობა, რომ ბაქტერია ცენტრთან უფრო ახლოსაა, ვიდრე ზედაპირთან (მითითება: ჩათვალეთ, რომ მუხუდოს მარცვალს სფეროს ფორმა აქვს).

5. შაქრო და ცირა შეთანხმდნენ შეხვედრაზე 14^{00} საათიდან 15^{00} საათამდე შემდეგი პირობით: დათქმულ ადგილას მისული შაქრო იცდის 20 წუთს, ხოლო ცირა - 10 წუთს. გამოთვალეთ შეხვედრის ალბათობა, თუ ცნობილია, რომ თითოეულის მისვლის დრო შემთხვევითია 14^{00}-დან 15^{00}-მდე.

6. სახაზავი გაჭრეს ორად. გამოთვალეთ ალბათობა იმისა, რომ ამ ნაჭრების სხვაობის მოდული არ აღემატება:

ა) სახაზავის სიგრძის ნახევარს.

ბ) სახაზავის სიგრძის მეოთხედს.

7. სახაზავი გაჭრეს ორ ადგილას. გამოთვალეთ ალბათობა იმისა, რომ მიღებული ნაჭრებისაგან შეიძლება სამკუთხედის აგება.

8. სიბრტყეზე დატანილია პარალელური ხაზები, რომელთა შორის მანძილი 10 სმ-ია. სიბრტყეზე შემთხვევით

ვარდება მონეტა, რომლის რადიუსი 3 სმ-ია. იპოვეთ იმის ალბათობა, რომ მონეტა არ გადაკვეთს არც ერთ ხაზს.

9. 20 სმ. დიამეტრის წითელი ფერის წრეში მოთავსებულია 8 სმ. დიამეტრის შავი ფერის ფირფიტა, რომელთა ცენტრებიც ერთმანეთს ემთხვევა. 4 სმ. დიამეტრის მონეტა მთლიანად დადეს დიდი წრის შიგნით. იპოვეთ იმის ალბათობა, რომ მონეტა:

ა) მთლიანად შავი ფერის ფირფიტაზე მოხვდება;

ბ) მხოლოდ წითელი ფერის ნაწილზე მოხვდება;

გ) ორივე ფერის ნაწილზე მოხვდება.

10. სივრცეში მოცემულია ორი ბირთვი ცენტრით ერთსა და იმავე წერტილში და რადიუსებით 4 და 10სმ. წერტილი შემთხვევით ვარდება დიდ ბირთვში. იპოვეთ ალბათობა იმისა, რომ:

ა) წერტილი მოხვდება მცირე ბირთვში;

ბ) წერტილი არ მოხვდება მცირე ბირთვში.

პასუხები:

1) 2/π. **2**) 1/4. **3**) 16/25. **4**)18. **5**) 31/72. **6**) 1/2; 1/4. **7**) 1/4. **8**) 2/5. **9**) 1/16; 7/16; 1/2. **10**) 8/125; 117/125.

§ 20. პირობითი ალბათობა. ხდომილობათა ნამრავლის ალბათობა. ხდომილობათა დამოუკიდებლობა

ზოგიერთ შემთხვევაში A ხდომილობის ალბათობის მნიშვნელობაზე შეიძლება გავლენა მოახდინოს ამავე ცდასთან დაკავშირებული სხვა ხდომილობის შესახებ ინფორმაციის (მოხდენის ან არ მოხდენის ფაქტის) ცოდნამ.

მაგალითი 20.1. დავუშვათ, ერთი მოთამაშე აგორებს კამათელს, ხოლო მეორემ უნდა გამოიცნოს მასზე მოსული ციფრი. A იყოს ხდომილობა – ციფრი გამოცნობილია. გამოვთვალოთ $P(A)$.

ამოხსნა. ცხადია $\Omega=\{1,2,3,4,5,6\}$, $N(\Omega)=6$, ხოლო A ხდომილობის განხორციელებას ხელს უწყობს ის ერთად-ერთი რიცხვი, რომელიც მოვიდა კამათელზე $N(A)=1$

$$P(A)=\frac{N(A)}{N(\Omega)}=\frac{1}{6}.$$

მაგალითი 20.2. დავუშვათ ცდა იგივეა როგორც 20.1 მაგალითის დროს, ოღონდ კამათლის ლუწი ციფრები შეღებილი იყოს წითლად. დავუშვათ გამომცნობელმა მოთამაშემ შემაჩნია, რომ კამათელზე მოვიდა წითელი გვერდი. ვიპოვოთ ისევ $P(A)$.

ამოხსნა. შემოვიღოთ ხდომილობა B – მოვიდა ლუწი რიცხვი. ცხადია, რომ რადგან გამომცნობელმა შეამჩნია წითელი გვერდის მოსვლა, ამიტომ გამოცნობის დროს იგი არ დაასახელებს კენტს. ე.ი. არჩევანს გააკეთებს სამი შესაძლო ვარიანტიდან. $\Omega'=\{2,4,6\}$, ხოლო A ხდომილობას ამათგან ხელს უწყობს მხოლოდ ერთი, ის რომელიც მოსულია. ე.ი. ჩვენ ამ ცდის შესაბამისი Ω ელემენტარულ ხდომილო-ბათა სივრცის ნაცვლად ვიხილავთ ახალ $\Omega'=\{2,4,6\}$ სივრცეს.

$$P(A)=\frac{N(A)}{N(\Omega')}=\frac{1}{3}.$$

როგორც ვხედავთ B ხდომილობის მოხდენის ფაქტის ცოდნამ შეცვალა A-ს ალბათობა.

განსაზღვრება 20.1. **თუ $P(B)>0$, მაშინ პირობითი ალბათობა A ხდომილობისა B-ს პირობით** ეწოდება A-ს ალბათობას იმ დროს როდესაც ცნობილია, რომ B უკვე მოხდა. პირობითი ალბათობა ჩაიწერება $P(A|B)$ ან $P_B(A)$ სახით და გამოითვლება ფორმულით

$$P(A|B)=\frac{P(A\bigcap B)}{P(B)}. \tag{20.1}$$

ამ ტოლობიდან მიიღება ხდომილობათა ნამრავლის ალბათობის გამოსათვლელი ფორმულა

$$P(A\bigcap B)=P(B)P(A|B). \tag{20.2}$$

ეს უკანასკნელი ტოლობა შეიძლება განვაზოგადოთ ნებისმიერი სასრული რაოდენობა $A_1,A_2,A_3,\ldots,A_n$ ხდომილობებისათვის.

$$P(A_1\bigcap A_2\bigcap A_3\bigcap\ldots\bigcap A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1\bigcap A_2)\cdot\ldots$$
$$\ldots\cdot P(A_n|A_1\bigcap\ldots\bigcap A_{n-1}). \tag{20.3}$$

მაგალითი 20.3. ბანქოს სათამაშო 36 ქაღალდიანი დასტიდან სათითაოდ იღებენ 3 ცალ ქაღალდს დაბრუნების გარეშე. ვიპოვოთ იმის ალბათობა, რომ სამივე ამოღებულზე იქნება ციფრი 7.

ამოხსნა. შემოვიტანოთ ხდომილობები:

A – სამივე ამოღებული არის შვიდიანი

A_1 – პირველად ამოღებული არის შვიდიანი

A_2 – მეორედ ამოღებული არის შვიდიანი

A_3 – მესამედ ამოღებული არის შვიდიანი.

ცხადია რომ $A=A_1\bigcap A_2\bigcap A_3$. (20.3) ფორმულის თანახმად გვექნება

$$P(A)=P(A_1)P(A_2|A_1)P(A_3|A_1\bigcap A_2).$$

$A_2|A_1$ წარმოადგენს ხდომილობას – „მეორედ ამოღებისას ამოვიდა შვიდიანი მაშინ როდესაც პირველად ამოღებული იყო შვიდიანი“. ვინაიდან პირველად შვიდიანი ამოიღეს, ამიტომ მეორედ ამოღების წინ ბანქოს დასტაში 35 ქაღალდი იყო, რომელთაგან 3 ცალი იყო შვიდიანი. კლასიკური ალბათობის ფორმულით გვექნება

$$P(A_2 \mid A_1) = \frac{3}{35}.$$

$A_3 \mid A_1 A_2$ ხდომილობა ნიშნავს – „მესამედ ამოღებისას ამოვიდა შვიდიანი, მაშინ როდესაც პირველი და მეორე ამოღების დროს ორივეჯერ შვიდიანი ამოვიდა“. პირველი ორი ამოღების დროს შვიდიანების ამოღების გამო მესამე ამოღების წინ ყუთში 34 ქაღალდი იყო, რომელშიც სულ ორი შვიდიანიღა იყო დარჩენილი. ამიტომ

$$P(A_3 \mid A_1 \cap A_2) = \frac{2}{34}.$$

ვინაიდან $P(A_1) = \frac{4}{36}$, ამიტომ მივიღებთ

$$P(A) = \frac{4}{36} \times \frac{3}{35} \times \frac{2}{34} = \frac{1}{1785}.$$

დავუბრუნდეთ 20.2 მაგალითს. ამოვხსნათ ის (20.1) ფორმულის გამოყენებით კამათლის ერთხელ გაგორების დროს ცდის ყველა შესაძლო ელემენტარულ ხდომილობათა სიმრავლეა

$$\Omega = \{1,2,3,4,5,6\}$$

ამათგან ხდომილებას რომელიც მოსული ციფრის გამოცნობას ნიშნავს ხელს უწყობს მხოლოდ ერთი მათგანი ω_0, ის რომელიც მოსულია. ვინაიდან $B = \{2,4,6\}$, ამიტომ $AB = \{\omega_0\}$ ცხადია $P(B) = \frac{3}{6}$ და $P(AB) = \frac{1}{6}$. მაშასადამე

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1}{6} : \frac{3}{6} = \frac{1}{3}.$$

თუ A ხდომილობა არაა დამოკიდებული B-ზე, ცხადია B-ს შესახებ ინფორმაციის ცოდნა (ანუ B-ს მოხდენა ან არ მოხდენა) არ ახდენს გავლენას A-ს ალბათობაზე. ე.ი. $P(A \mid B) = P(A)$. თუ გავითვალისწინებთ ამ ფაქტს (20.2) ტოლობაში მივიღებთ

$$P(A \cap B) = P(B)P(A),$$

ამიტომ შემდეგ ტოლობათა ჯაჭვიდან ცხადია, რომ არც B ყოფილა დამოკიდებული A-ზე.

$$P(B \mid A)=\frac{P(B \bigcap A)}{P(A)}=\frac{P(B)P(A)}{P(A)}=P(B).$$

ბუნებრივია შემდეგი განსაზღვრება.

განსაზღვრება 20.2. A და B ხდომილობებს ეწოდებათ **დამოუკიდებლები**, თუ

$$P(A \bigcap B)=P(A)P(B).$$

მაგალითი 20.4. ცდა მდგომარეობს ერთი კამათლის ერთხელ გაგორებაში. განვიხილოთ ხდომილობები: A–მოვიდა ლუწი რიცხვი; B – მოვიდა 3-ის ჯერადი რიცხვი; C – მოვიდა 5-ზე ნაკლები რიცხვი.

ვაჩვენოთ, რომ A და B ხდომილობათა წყვილი დამოუკიდებელია, ისევე როგორც A და C ხდომილობათა წყვილი, ხოლო B და C ხდომილობები კი დამოკიდებულია.

ამოხსნა. $\Omega=\{1,2,3,4,5,6\}$; $A=\{2,4,6\}$; $B=\{3,6\}$; $C=\{1,2,3,4\}$;

$$AB=\{6\};\ AC=\{2,4\};\ BC=\{3\}.$$

მარტივად გამოითვლება, რომ

$$P(A)=\frac{1}{2};\ P(B)=\frac{1}{3};\ P(C)=\frac{2}{3};$$

$$P(A \bigcap B)=\frac{1}{6};\ P(A \bigcap C)=\frac{1}{3};\ P(B \bigcap C)=\frac{1}{6}.$$

მაშასადამე,

$$P(A \bigcap B)=P(A)P(B);\ P(A \bigcap C)=P(A)P(C);$$

$$P(B \bigcap C) \neq P(B)P(C).$$

რაც ნიშნავს, რომ ხდომილობათა A, B და A, C წყვილები დამოუკიდებლებია, ხოლო ხდომილობათა B და C წყვილი კი არა.

განვიხილოთ $A_1, A_2, A_3, \ldots, A_n$ ხდომილობათა სიმრავლე და შემოვიღოთ შემდეგი განსაზღვრებები:

განსაზღვრება 20.3. $A_1, A_2, A_3, \ldots, A_n$ ხდომილობებს ეწოდებათ **წყვილ-წყვილად დამოუკიდებლები** თუ მათგან აღებული ნებისმიერი წყვილი დამოუკიდებელია, ანუ

$$P(A_i \cap A_k) = P(A_i)P(A_k),\ 1 \le i,\ \ k \le n \quad i \ne k .$$

განსაზღვრება 20.4. $A_1, A_2, A_3, \ldots, A_n$ ხდომილობებს ეწოდებათ **ერთობლივად დამოუკიდებლები**, თუ მათგან აღებული ნებისმიერი k რაოდენობის ($k \le n$) $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$ ხდომილობისათვის სრულდება ტოლობა

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2})\ldots P(A_{i_k}) .$$

ცხადია, ხდომილობათა ერთობლივად დამოუკიდებლობა ნიშნავს წყვილ-წყვილად დამოუკიდებლობასაც. ახლა მოვიყვანოთ მაგალითი, რომელიც გვიჩვენებს რომ ხდომილობები შეიძლება იყვნენ წყვილ-წყვილად დამოუკიდებლები, მაგრამ არ იყვნენ ერთობლივად დამოუკიდებლები. ანუ ერთობლივად დამოუკიდებლობა უფრო დიდი „მოთხოვნაა", ვიდრე წყვილ-წყვილად დამოუკიდებლობა.

მაგალითი 20.5. **(ბერნშტეინი)** ავიღოთ ტეტრაედრი (სამკუთხა პირამიდა რომლის ყველა წიბო ტოლია), რომლის 4 წახნაგიდან სამზე სათითაოდ აწერია ციფრები „1", „2", „3", ხოლო მეოთხე წახნაგზე რიცხვი „123". პირამიდას გადააგდებენ და მოსულად ითვლება ის წახნაგი, რომელიც მოექცევა ქვემოთა მხარეს როდესაც ის გაჩერდება. შემოვიღოთ ხდომილობები: A_1-მოსულ გვერდზე მოჩანს ციფრი 1; A_2-ციფრი 2; A_3-ციფრი 3. ვაჩვენოთ, რომ A_1, A_2, A_3 წყვილ-წყვილად დამოუკიდებელი ხდომილობებია, მაგრამ არ არიან ერთობლივად დამოუკიდებლები.

$$\Omega = \{1; 2; 3; 123\}$$

$$A_1 = \{1;\ 123\};\ \ A_2 = \{2;\ 123\};\ \ A_3 = \{3;\ 123\};$$

$$A_1 \cap A_2 = A_1 \cap A_3 = A_2 \cap A_3 = \{123\}.$$

ცხადია რომ

$$P(A_1) = P(A_2) = P(A_3) = \frac{2}{4} = \frac{1}{2};$$

$$P(A_1 \cap A_2) = P(A_1 \cap A_3) = P(A_2 \cap A_3) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} =$$
$$= P(A_1)P(A_2) = P(A_1)P(A_3) = P(A_2)P(A_3).$$

ე.ი. A_1, A_2, A_3 წყვილ-წყვილად დამოუკიდებელი ხდომილობებია. ახლა განვიხილოთ ხდომილობა $A_1 \cap A_2 \cap A_3 = \{123\}$. გვაქვს

$$P(A_1 \cap A_2 \cap A_3) = \frac{1}{4} \neq P(A_1)P(A_2)P(A_3).$$

რაც ნიშნავს რომ A_1, A_2 და A_3 არ არიან ერთობლივად დამოუკიდებელი ხდომილობები.

მივიყვანოთ მაგალითი შესაძლო მოგების განაწილებაზე დასრულებამდე შეწყვეტილ თამაში.

მაგალითი 20.6. ორი A და B პიროვნება რიგრიგობით აგდებს ლითონის მონეტას. თუ მოვიდა გერბი ქულას იწერს A მოთამაშე, თუ მოვიდა საფასური – B. თამაში გრძელდება 3 გამარჯვებამდე. გარკვეული მიზეზის გამო თამაში შეწყდა, როდესაც ანგარიში იყო 2:0 A მოთამაშის სასარგებლოდ. როგორ შეიძლება განაწილდეს თანხა, რომელსაც გამარჯვებული მოიგებდა თამაში რომ დამთავრებულიყო.

ამოხსნა. ბუნებრივად გვეჩვენება ასეთი მსჯელობა: რადგან თამაშში წარმატების მისაღწევად A-ს სჭირდება 1 გამარჯვება, ხოლო B-ს კი 3, ამიტომ თანხა უნდა განაწილდეს 1 და 3-ის უკუპროპორციულ ნაწილებად ე.ი. A მოთამაშემ მიიღოს თანხის $\frac{3}{4}$ ნაწილი, ხოლო B-მ კი $\frac{1}{4}$.

ახლა ვაჩვენოთ, რომ ასეთი განაწილება არ შეესაბამება რეალობას.

დავუშვათ, რომ თამაში შეწყდა დროის t_0 მომენტში, როდესაც ანგარიში იყო 2:0 A მოთამაშის სასარგებლოდ და გავაგრძელოთ თამაში t_0 მომენტიდან. შემოვიღოთ ხდომილობები:

H – ანგარიშია 2:0 A მოთამაშის სასარგებლოდ;
A – გაიმარჯვა A მოთამაშემ;

B – გაიმარჯვა B მოთამაშემ;

C_1 – დროის t_0 მომენტის შემდეგ მონეტის პირველ აგდებაზე მოვიდა საფასური;

C_2 – დროის t_0 მომენტის შემდეგ მონეტის მეორე აგდებაზე მოვიდა საფასური;

C_3 – დროის t_0 მომენტის შემდეგ მონეტის მესამე აგდებაზე მოვიდა საფასური.

ცხადია, B მოთამაშის გამარჯვებას t_0 მომენტის შემდეგ (როდესაც 0:2 აგებდა) ხელს უწყობს მხოლოდ $C_1 \bigcap C_2 \bigcap C_3$ ხდომილობა. მაშასადამე

$$B \mid H = C_1 \bigcap C_2 \bigcap C_3 .$$

ვინაიდან მონეტის აგდების ცდები დამოუკიდებლებია, ამიტომ ამ ცდებთან დაკავშირებული C_1, C_2 და C_3 ხდომილობებიც დამოუკიდებლებია. რადგან $P(C_1) = P(C_2) = P(C_3) = \dfrac{1}{2}$, ამიტომ გვექნება

$$P(B \mid H) = P(C_1 \bigcap C_2 \bigcap C_3) = P(C_1) \cdot P(C_2) \cdot P(C_3) = \frac{1}{8} .$$

მივიღეთ, რომ B მოთამაშის გამარჯვების ალბათობა, იმ მომენტიდან, როდესაც A მოთამაშე იგებდა ანგარიშით 2:0, არის $\dfrac{1}{8}$. შესაბამისად A მოთამაშის გამარჯვების ალბათობა ამ t_0 მომენტიდან იქნება

$$P(A \mid H) = 1 - P(B \mid H) = \frac{7}{8} .$$

ბუნებრივია, რომ თანხა რომელსაც გამარჯვებული მოიგებდა უნდა განაწილდეს შემდეგნაირად: A მოთამაშე მიიღებს თანხის $\dfrac{7}{8}$ ნაწილს, B კი - $\dfrac{1}{8}$ ნაწილს.

ამოცანები

1. ალბათობა იმისა, რომ ავტომრბოლელს ტრასაზე დასჭირდება ტექნიკური დახმარება 0,15-ის ტოლია. თუ მრბოლელი ტრასას გაივლის შეფერხების გარეშე, მაშინ ის 0,8 ალბათობით ახერხებს ჩასათვლელი ქულების მოპოვებას. გამოთვალეთ ალბათობა იმისა, რომ ერთი შემთხვევითად აღებული ავტომრბოლელი მოიპოვებს ჩასათვლელ ქულებს.

2. ფირმის მენეჯერი 0,7 ალბათობი იღებს საქართველოს ბანკში საკრედიტო ხაზის გახსნის გადაწყვეტილებას. ალბათობა იმისა, რომ კრედიტ-ოფიცერი განაცხადზე დადებით პასუხს გასცემს 0,8-ის ტოლია. იპოვეთ იმის ალბათობა, რომ ფირმა მიიღებს საქართველოს ბანკის საკრედიტო ხაზს.

3. ყუთში 4 თეთრი და 6 შავი ერთნაირი ზომის ბურთულაა. ჩაუხედავად სათითაოდ ერთიმეორის მიყოლებით იღებენ ორ ცალს. იპოვეთ იმის ალბათობა, რომ ა) ორივე ამოღებული ბურთულა შავია; 2) ორივე ამოღებული ბურთულა თეთრია.

4. ელექტროენერგიის მიმწოდებელ კომპანიას რეგიონში თოვლის მოსვლის დროს აქვს დანაკარგები, ხოლო თოვლის არ მოსვლის შემთხვევაში კომპანია 0,8 ალბათობით ახორციელებს უდანაკარგო დისტრიბუციას. იპოვეთ იმის ალბათობა, რომ მოცემულ რეგიონში თებერვლის თვეში მიმწოდებელ კომპანიას არ ექნება დანაკარგი, თუ ცნობილია, რომ ამ რეგიონში თებერვალში თოვლის მოსვლის ალბათობა 0,75.

5. ხარატი წუნდებულ დეტალს უშვებს 0,08 ალბათობით, ხოლო ზეინკალი სტანდარტული დეტალებიდან ვარგისიან ხელსაწყოს აწყობს 0,95 ალბათობით. გამოთვალეთ იმის ალბათობა, რომ ზეინკლის მიერ გამოშვებული ხელსაწყო ვარგისიანია.

6. ორ კამათელს აგორებენ ერთდროულად. ცნობილია, რომ ორივეზე მოვიდა კენტი რიცხვი. გამოთვალეთ ალბათობა იმისა, რომ:

ა) მოვიდა ერთნაირი რიცხვები;

ბ) მოსული ციფრების ჯამი 8-ზე მეტია.

7. შუა ამერიკულ სასაქონლო ბირჟაზე („MidAm"-Mid America Commodity Exchange) ვაშლის საბითუმო ფასი თოვლის მოსვლის დროს 0,8 ალბათობით მატულობს. იპოვეთ იმის ალბათობა, რომ ბირჟაზე ვაშლის ფასი გაიზრდება, თუ ცნობილია, რომ თებერვლის თვეში თოვლის მოსვლის ალბათობა 0,75.

პასუხები

1) 0,68. 2) 0,56. 3) 1/3; $2/15$. 4) 0,2. 5) 0,874. 6) 1/3; 1/9. 7) 0,6.

§ 21. ხდომილობათა სრული ჯგუფი. სრული ალბათობისა და ბაიესის ფორმულები

განსაზღვრება 21.1. ერთი და იგივე ცდასთან დაკავშირებულ რაიმე $H_1, H_2, \ldots, H_n$ ხდომილობებს ეწოდებათ **ხდომილობათა სრული ჯგუფი** (ან ხდომილობათა **სრული სისტემა**) თუ ეს ხდომილობები წყვილ წყვილად არათავსებადია და მათი გაერთიანება არის Ω. ე.ი

$$H_i \bigcap H_k = \varnothing \text{ თუ } i \neq k,$$

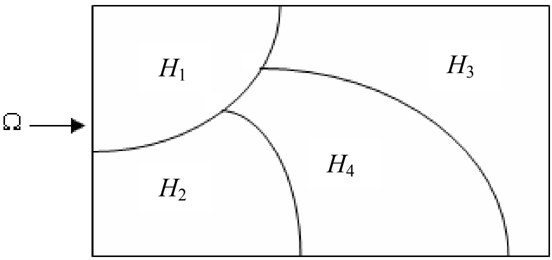
$$H_1 \bigcup H_2 \bigcup \ldots \bigcup H_n = \Omega.$$

მაგალითი 21.1. ერთი ცალი კამათლის გაგორების დროს სრული ჯგუფი იქნება ხდომილობები: H_1 – მოვიდა ლუწი რიცხვი; H_2 – მოვიდა 3 ან 5; H_3 - მოვიდა 1.

ცხადია

$$H_1 = \{2,4,6\}; \; H_2 = \{3,5\}; \; H_3 = \{1\}.$$

გეომეტრიულ წარმოდგენას მივიღებთ თუ Ω იქნება რაიმე ბრტყელი ფიგურა და მას დავყოფთ მაგალითად 4 თანაუკვეთ ნაწილად H_1; H_2; H_3 და H_4.



განსაზღვრება 21.2. თუ $H_1, H_2, \ldots, H_n$ ხდომილობათა სრული ჯგუფია, ხოლო A იგივე ცდასთან დაკავშირებული რაიმე ხდომილობა, მაშინ $P(A)$ ალბათობა შეიძლება გამოვთვალოთ **სრული ალბათობის** ფორმულით

$$P(A) = \sum_{k=1}^{n} P(H_k) P(A \mid H_k) = P(H_1) P(A \mid H_1) + \ldots + P(H_n) P(A \mid H_n).$$

ამ ფორმულაში შემავალ $H_1, H_2, \ldots, H_n$ ხდომილობებს ხშირად **ჰიპოთეზებსაც** უწოდებენ.

მაგალითი 21.2. ტელევიზორების მწარმოებელმა ფირმამ თავის სამ ფილიალში გამოშვებული 1000 ტელევიზორი გაიტანა საცალო სავაჭრო ქსელში გასაყიდად. I ფილიალიდან 250 ცალი, II-დან 350, ხოლო III-დან 400. რა არის იმის ალბათობა, რომ მყიდველის მიერ ნაყიდი ტელევიზორი აღმოჩნდება სტანდარტის შესაბამისი თუ ცნობილია, რომ I ფილიალი უშვებს 99% სტანდარტულ პროდუქციას, II ფილიალი 97%-ს, ხოლო III კი 96%-ს.

ამოხსნა. შემოვიღოთ ხდომილობები A – მყიდველის მიერ ნაყიდი ტელევიზორი სტანდარტულია. H_1 – გასაყიდი პარტიიდან შემთხვევითად არჩეული ტელევიზორი გამოშვებულია I ფილიალის მიერ, H_2 – მეორის მიერ, ხოლო H_3-მესამის მიერ. ცხადია H_1, H_2 და H_3 ხდომილობათა სრული ჯგუფია, რომელთა ალბათობებიც გამოითვლება ალბათობის კლასიკური ფორმულით

$$P(H_1) = \frac{250}{1000} = 0{,}25\,; \;\; P(H_2) = \frac{350}{1000} = 0{,}35\,; \;\; P(H_3) = \frac{400}{1000} = 0{,}4\,.$$

$P(A|H_1)$ არის ალბათობა იმისა, რომ ტელევიზორი სტანდარტულია, მაშინ როდესაც ცნობილია, რომ ის გამოშვებულია I ფილიალის მიერ. ამოცანის პირობით $P(A|H_1) = 0{,}99$, ასევე $P(A|H_2) = 0{,}97$ და $P(A|H_3) = 0{,}96$. სრული ალბათობის ფორმულით გვექნება

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + P(H_3)P(A|H_3) =$$
$$= 0{,}25 \cdot 0{,}99 + 0{,}35 \cdot 0{,}97 + 0{,}4 \cdot 0{,}96 = 0{,}971\,.$$

ზოგიერთ შემთხვევაში ისმის გარკვეულწილად შებრუნებული ამოცანა: რაიმე A ხდომილობა უკვე მომხდარია და აინტერესებთ რა ალბათობით შეიძლება ვივარაუდოთ, რომ მისი გამომწვევი მიზეზი იყო ხდომილობათა სრული ჯგუფის ესა თუ ის H_i $(i = \overline{1,n})$ ხდომილობა. ამ საკითხის გადასაჭრელად სარგებლობენ **ბაიესის** ფორმულებით:

$$P(A|H_k) = \frac{P(H_k)P(A|H_k)}{P(A)}, \quad k = 1, \ldots, n\,. \qquad (21.1)$$

მაგალითი 21.3. 21.2 ამოცანის პირობებში დავუშვათ, რომ ტელევიზორების პარტია ფირმისთვის ხელსაყრელ დროში გაიყიდა და ფირმის ხელმძღვანელობამ გადაწყვიტა მოგებიდან 10000 აშშ დოლარი გაანაწილოს პრემიის სახით ფილიალებში. როგორ უნდა განაწილდეს ეს თანხა?

ამოხსნა. ისევე როგორც 21.2 ამოცანის ამოხსნისას A, H_1, H_2 და H_3 იყოს იგივე ხდომილობები. გამოვითვალოთ $P(H_1 | A)$, $P(H_2 | A)$ და $P(H_3 | A)$ ალბათობები და თანხა გავანაწილოთ მათ პროპორციულ ნაწილებად.

$P(H_1 | A)$ წარმოადგენს იმის ალბათობას, რომ შემთხვევითად არჩეული (ანუ მყიდველის მიერ ნაყიდი) ტელევიზორი გამოშვებულია I ფილიალის მიერ, მაშინ როცა ცნობილია რომ ეს ტელევიზორი სტანდარტულია. ანალოგიურად ცხადია, რას წარმოადგენენ $P(H_2 | A)$ და $P(H_3 | A)$ ალბათობები.

ბაიესის (21.1) ფორმულების თანახმად გვაქვს

$$P(H_1 | A) = \frac{P(H_1)P(A | H_1)}{P(A)} = \frac{0{,}25 \cdot 0{,}99}{0{,}971} = \frac{495}{1942};$$

$$P(H_2 | A) = \frac{P(H_2)P(A | H_2)}{P(A)} = \frac{0{,}35 \cdot 0{,}97}{0{,}971} = \frac{679}{1942};$$

$$P(H_3 | A) = \frac{P(H_3)P(A | H_3)}{P(A)} = \frac{0{,}4 \cdot 0{,}96}{0{,}971} = \frac{768}{1942}.$$

მაშასადამე I ფილიალი მიიღებს $\frac{495}{1942} \cdot 10000$, II- $\frac{679}{1942} \cdot 10000$, ხოლო III-$\frac{768}{1942} \cdot 10000$ აშშ დოლარს.

ახლა განვიხილოთ მაგალითი, რომელიც გვიჩვენებს, რომ გამოცდაზე მისულ მოსწავლეს ერთი და იგივე ალბათობა აქვს გამოცდის ჩაბარების იმისდა მიუხედავად, თუ რომელი გავა (რიგის მიხედვით) იგი გამოცდაზე.

მაგალითი 21.4. გამოცდაზე მისულ სტუდენტს n ბილეთიდან მომზადებული აქვს m ბილეთი ($m < n$). როდის უფრო დიდი აქვს გამოცდის ჩაბარების ალბათობა, როდესაც გამოცდაზე გავა რიგით პირველი თუ როდესაც გავა რიგით მეორე?

ამოხსნა. თუ სტუდენტი გამოცდაზე გავა პირველი, მაშინ ცხადია გამოცდის ჩაბარების (ანუ მომზადებული ბილეთის შეხვედრის) ალბათობა არის $\frac{m}{n}$.

დავუშვათ, სტუდენტი გამოცდაზე გადის რიგით მეორე. შემოვიღოთ ხდომილობები: H_1 მის წინ გასულ სტუდენტს შეხვდა ბილეთი სწორედ იმ m ცალიდან, რომელიც მას მომზადებული აქვს. H_2- იმ $n-m$ ცალიდან, რომელიც მას არა აქვს მომზადებული. A იყოს ხდომილობა მეორედ გასულ სტუდენტს შეხვდა მომზადებული ბილეთი.

ცხადია H_1 და H_2 ხდომილობათა სრული სისტემაა და

$$P(H_1)=\frac{m}{n}; \qquad P(H_2)=\frac{n-m}{n}.$$

ხდომილობა $A|H_1$ ნიშნავს, რომ მეორედ გასულ სტუდენტს შეხვდა მომზადებული ბილეთი, მაშინ, როდესაც წინ გასულმა სტუდენტმა აიღო მის მიერ მომზადებული ბილეთებიდან ერთ-ერთი. ამ დროს მთელი დარჩენილი $n-1$ ბილეთიდან სტუდენტის მომზადებული ბილეთების რაოდენობაა $m-1$, ამიტომ $P(A|H_1)=\frac{m-1}{n-1}$. ხდომილობა $A|H_2$ ნიშნავს რომ მეორედ გასულ სტუდენტს შეხვდა მომზადებული ბილეთი, მაშინ როდესაც მის წინ გასულმა აიღო ერთ-ერთი ის ბილეთი, რომელიც მას არ ჰქონდა მომზადებული. ასეთ შემთხვევაში მთელი დარჩენილი $n-1$ ბილეთიდან სტუდენტს მომზადებული ექნება ისევ m ბილეთი მაშასადამე $P(A|H_2)=\frac{m}{n-1}$. სრული ალბათობის ფორმულით მივიღებთ:

$$P(A)=P(H_1)P(A|H_1)+P(H_2)P(A|H_2)=$$
$$=\frac{m}{n}\cdot\frac{m-1}{n-1}+\frac{n-m}{n}\cdot\frac{m}{n-1}=\frac{m}{n}.$$

მაშასადამე სტუდენტს გამოცდის ჩაბარების ერთი და იგივე შანსი აქვს გინდა რიგით პირველი გავიდეს და გინდა მეორე.

მოვიყვანოთ მაგალითი, რომელიც გვიჩვენებს თუ როგორ შეიძლება გამოვიყენოთ სრული ალბათობის ფორმულა მარტივი მარკეტინგული ამოცანის გადაჭრისას.

მაგალითი 21.5. საყოფაცხოვრებო ტექნიკის მწარმოებელმა M კომპანიამ გადაწყვიტა ერთი წლის შემდეგ მეზობელი ქვეყნის ბაზრის ათვისება. კონკრეტულად ამ ბაზარზე მის მიერ წარმოებული მაცივრების პარტიის გატანა. ფირმის წარმომადგენლებმა შეისწავლეს ბაზარი და აღმოჩნდა, რომ საყოფაცხოვრებო ტექნიკით აქ ვაჭრობს ორი A და B კომპანია. დაადგინეს, რომ M კომპანიის მაცივრების მიწოდება მოთხოვნას გადააჭარბებს 10%-ით თუ ბაზარზე მხოლოდ A კომპანიას აქვს გამოტანილი მაცივრები, 5%-ით თუ მხოლოდ B-ს და 20%-ით თუ ორივე A და B კომპანიას ერთდროულად აქვთ გამოტანილი მაცივრები. უხეში მიახლოებით დავუშვათ, რომ პროდუქციის მთლიანი პარტიის გაყიდვის ალბათობა იმდენივე პროცენტით იკლებს რამდენითაც მიწოდება აჭარბებს მოთხოვნას.

ვიპოვოთ, თუ რა ალბათობით შეიძლება ელოდონ M ფირმის წარმომადგენლები ერთი წლის შემდეგ მოცემულ ბაზარზე გატანილი თავიანთი მაცივრების პარტიის მთლიანად გაყიდვას. ჩავთვალოთ, რომ A და B კომპანიები მაცივრების გატანას ბაზარზე ანხორციელებენ ერთმანეთისაგან დამოუკიდებლად შესაბამისად 0,9 და 0,8 ალბათობებით. ამასთანავე ვგულისხმობთ, რომ ბაზარზე სხვა ფირმის მიერ წარმოებული მაცივრების არ არსებობის შემთხვევაში M კომპანიის პროდუქცია მთლიანად იყიდება.

ამოხსნა. შემოვიღოთ ხდომილობები:

A – ერთი წლის შემდეგ ბაზარზე მაცივრები გატანილი აქვს A კომპანიას. $P(A)=0{,}9$.

B – ერთი წლის შემდეგ ბაზარზე მაცივრები გატანილი აქვს B კომპანიას. $P(B)=0{,}8$.

M – მაცივრების პარტია, რომელიც M კომპანიამ გაიტანა მთლიანად გაიყიდა.

$H_1=\overline{A}\bigcap\overline{B}$ ნიშნავს, რომ ბაზარზე არც A და არც B კომპანიას მაცივრები არ აქვთ გატანილი.

$H_2 = A \bigcap \overline{B}$ ნიშნავს, რომ ბაზარზე მხოლოდ A კომპანიის მაცივრებია გატანილი.

$H_3 = \overline{A} \bigcap B$ ნიშნავს, რომ ბაზარზე მხოლოდ B კომპანიის მაცივრებია გატანილი.

$H_4 = A \bigcap B$ ნიშნავს, რომ ბაზარზე ორივე A და B კომპანიის მაცივრებია გატანილი.

ცხადია, რომ H_1, H_2, H_3, H_4 ხდომილობათა სრული ჯგუფია. ვინაიდან ამოცანის პირობის თანახმად A და B კომპანიები ბაზარზე მუშაობენ ერთმანეთისაგან დამოუკიდებლად, ამიტომ:

$$P(H_1) = P(\overline{A} \bigcap \overline{B}) = P(\overline{A})P(\overline{B}) = (1-0,9)(1-0,8) = 0,02\,;$$

$$P(H_2) = P(A \bigcap \overline{B}) = P(A)P(\overline{B}) = 0,9(1-0,8) = 0,18\,;$$

$$P(H_3) = P(\overline{A} \bigcap B) = P(\overline{A})P(B) = (1-0,9)0,8 = 0,08\,;$$

$$P(H_4) = P(A \bigcap B) = P(A)P(B) = 0,9 \cdot 0,8 = 0,72\,.$$

$M \mid H_1$ არის ხდომილობა, რომ M კომპანიის მაცივრების პარტია მთლიანად გაიყიდება, როდესაც ბაზარზე კონკურენტი A და B ფირმების მაცივრები არაა გამოტანილი. ამოცანის პირობის თანახმად $P(M \mid H_1) = 1$.

$M \mid H_2$ არის ხდომილობა, რომ M კომპანიის მაცივრების პარტია მთლიანად გაიყიდება თუ ბაზარზე მხოლოდ A კომპანიას აქვს მაცივრები გამოტანილი. ამ დროს ამოცანის პირობის თანახმად მიწოდება აჭარბებს მოთხოვნას 10%-ით ე.ი მთლიანი პარტია 10%-ით ნაკლები „ალბათობით“, ანუ 0,1-ით ნაკლები ალბათობით, ე.ი. 0,9 ალბათობით გაიყიდება.

ანალოგიურად ცხადია, რომ

$$P(M \mid H_3) = 0,95\,,$$

$$P(M \mid H_4) = 0,8\,.$$

სრული ალბათობის ფორმულის გამოყენებით მივიღებთ

$$P(M) = P(H_1)P(M \mid H_1) + P(H_2)P(M \mid H_2) + P(H_3)P(M \mid H_3) +$$
$$+ P(H_4)P(M \mid H_4) = 0,02 \cdot 1 + 0,18 \cdot 0,9 + 0,08 \cdot 0,95 + 0,72 \cdot 0,8 = 0,834\,.$$

მაშასადამე M ფირმის წარმომადგენლები ერთი წლის შემდეგ ახალ ბაზარზე გატანილი მაცივრების მთლიანი პარტიის გაყიდვას უნდა ელოდნენ 0,834 ალბათობით.

ამოცანები

1. ხელსაწყოს მუშაობისას 90% შემთხვევაში ძაბვა ნორმის ფარგლებშია, ხოლო 10% შემთხვევაში – არა. როდესაც ძაბვა ნორმის ფარგლებშია ხელსაწყოს მწყობრიდან გამოსვლის ალბათობაა 0,05, როდესაც არაა ნორმის ფარგლებში – 0,8. გამოთვალეთ ერთ, შემთხვევით აღებულ მომენტში ხელსაწყოს მწყობრიდან გამოსვლის ალბათობა.

2. ფირმის ფინანსური მდგომარეობის გამოსასწორებლად მენეჯერი 0,6 ალბათობით მიმართავს ბანკს საკრედიტო ხაზის ლიმიტის გასაზრდელად, ხოლო 0,4 ალბათობით აპირებს ბირჟაზე გაყიდოს აქციათა პაკეტის ნაწილი. ალბათობა იმისა, რომ ბანკი დაამტკიცებს განაცხადს საკრედიტო ხაზის ლიმიტის გაზრდაზე არის 0,8; ხოლო ალბათობა იმისა, რომ აქციათა პაკეტი ბირჟაზე გაიყიდება დროულად და ხელსაყრელ ფასად ტოლია 0,75-ის. გამოთვალეთ იმის ალბათობა, რომ მენეჯერის მცდელობით ფირმის ფინანსური მდგომარეობა გამოსწორდება.

3. გვაქვს ორი ერთნაირი ყუთი. პირველში 6 თეთრი და 4 შავი ბურთულაა, ხოლო მეორეში 3 თეთრი და 7 შავი ისეთივე ბურთულა. შემთხვევით ირჩევენ ერთ ყუთს და იქიდან ჩაუხედავად იღებენ ერთ ცალ ბურთულას. იპოვეთ იმის ალბათობა, რომ ამოღებული ბურთულა თეთრია.

4. გვაქვს ორი ერთნაირი ყუთი. პირველ ყუთში 3 თეთრი და 5 შავი ერთი ზომის ბურთულაა, ხოლო მეორეში – ისეთივე ზომის 6 თეთრი და 3 შავი ბურთულა. პირველი ყუთიდან შემთხვევით (ჩაუხედავად) იღებენ ერთ ცალს და ფერის დაუნახავად დებენ მეორე ყუთში. შემდეგ მეორე ყუთიდან შემთხვევით იღებენ ერთ ცალ ბურთულას. იპოვეთ ალბათობა იმისა, რომ მეორედ ამოღებული ბურთულა თეთრია.

5. ჯგუფში 12 გოგონა და 8 ბიჭია. ბიჭი გამოცდას აბარებს 0,8, ხოლო გოგონა 0,6 ალბათობით. ცნობილია, რომ გამოცდაზე გავიდა მხოლოდ ერთი მათგანი და ჩააბარა. რას უდრის იმის ალბათობა, რომ გამოცდაზე გოგონა გავიდა?

6. მობილური ტელეფონიდან A ქალაქში დარეკვის დროს კავშირის დამყარება ნორმალურ პირობებში შესაძლებელია 0,95, ხოლო ძლიერ გაუარესებული ამინდის დროს – 0,4 ალბათობით. მომდევნო დღისთვის პროგნოზით ამინდის მკვეთრი გაუარესება მოსალოდნელია 0,8 ალბათობით. პიროვნება დარეკვას ორჯერ ცდის. გამოთვალეთ ალბათობები:

ა) დაკავშირება მოხდა პირველივე ცდაზე.

ბ) დაკავშირება მოხდა.

7. ალბათობა იმისა, რომ სასწრაფო დახმარების ბრიგადა თავისუფალია, როდესაც დახმარების ცენტრში შემოვა გამოძახება არის 0,9. ალბათობა იმისა, რომ პაციენტთან მისული სასწრაფო დახმარების ბრიგადა გადაარჩენს მას არის 0,95, ხოლო სასწრაფო დახმარების დაგვიანებისას ავადმყოფის თვითგამოკეთების ალბათობაა – 0,6. იპოვეთ იმის ალბათობა, რომ ავადმყოფი გადარჩება. (ჩათვალეთ, რომ თავისუფალი ბრიგადა პაციენტთან არ იგვიანებს).

8. საწარმოს გამოშვებული დეტალებიდან 40% დამზადებულია ღამის ცვლაში. ღამის ცვლაში დამზადებული დეტალი დეფექტურია ალბათობით 0,05, ხოლო დღის ცვლაში დამზადებული– 0,01 ალბათობით. იპოვეთ ალბათობა იმისა, რომ მთელი გამოშვებული დეტალებიდან აღებული ერთი დეტალი დეფექტურია.

9. ალბათობები იმისა, რომ პერსონალური კომპიუტერის მწყობრიდან გამოსვლა გამოიწვია ოპერატიულ მეხსიერებაში მომხდარმა დაზიანებამ, ეკრანზე მომხდარმა დაზიანებამ ან დანარჩენ მოწყობილობებში მომხდარმა დაზიანებამ ისე შეეფარდება ერთმანეთს, როგორც 3:2:5. ცნობილია, რომ მომსახურე პერსონალს თავისი ძალებით შეუძლია ოპერატიულ მეხსიერებაში მომხდარი დაზიანების გასწორება 0,7 ალბათობით, ეკრანზე მომხდარის – 0,9,

ხოლო სხვა დაზიანების – 0,8 ალბათობით. იპოვეთ იმის ალბათობა, რომ პერსონალი თავად შეაკეთებს დაზიანებულ კომპიუტერს.

10. გასაყიდად გამზადებულ 200 ერთი და იგივე ტიპის ავტომობილიდან 6-ს დაზიანებული აქვს ძრავი, ხოლო 10-ს – სამუხრუჭე სისტემა. საგარანტიო ვადაში სერვისცენტრს ძრავის შეკეთება შეუძლია 0,7, ხოლო სამუხრუჭე სისტემის – 0,9 ალბათობით. თუ ახლად შეძენილი ავტომანქანა ვერ დაექვემდებარება შეკეთებას, ფირმა მას ცვლის ახლით. იპოვეთ იმის ალბათობა, რომ მყიდველს დასჭირდება ავტომანქანის შეცვლა.

11. ერთი და იგივე დასახელების ელექტრულ ხელსაწყოს უშვებს ორი საწარმო. პირველმა დაამზადა მთელი ხელსაწყოების 60%, ხოლო მეორემ – 40%. პირველის მიერ გამოშვებული ხელსაწყოს საიმედოდ მუშაობის ალბათობაა 0,9, ხოლო მეორის – 0,95. გამოთვალეთ იმის ალბათობა, რომ მთელი პარტიიდან არჩეული ერთი ხელსაწყო იმუშავებს შეფერხების გარეშე.

12. წინა ამოცანის პირობებში ხელსაწყოების მთლიანი პარტიიდან ამოარჩიეს ერთი, რომელმაც ვერ იმუშავა საჭირო რეჟიმში. გამოთვალეთ იმის ალბათობა, რომ ეს ხელსაწყო დამზადებულია მეორე საწარმოში.

13. საავადმყოფოს რეანიმაციულ განყოფილებაში საშუალოდ 40% მოჰყავთ პირდაპირ შემთხვევის ადგილიდან, ხოლო 60% გადაჰყავთ – საავადმყოფოში მდგომარეობის გართულების შემდეგ. ცნობილია, რომ სიკვდილიანობის საშუალო მაჩვენებელი პირდაპირ შემთხვევის ადგილიდან მოყვანილი პაციენტებისათვის არის 25%, მკურნალობის გართულების შემდეგ მოყვანილთათვის – 10%. ერთი პაციენტი რეანიმაციის განყოფილებაში დაიღუპა. იპოვეთ იმის ალბათობა, რომ ეს პაციენტი პირდაპირ შემთხვევის ადგილიდან იყო მოყვანილი.

14. ზამთრის სეზონში ჩვეულებრივი საბურავებით, ავარიული ვითარების დროს, შეჯახების თავიდან აცილება ხდება 0,7 ალბათობით, ხოლო ზამთრის საბურავებით – 0,95 ალბათობით. სტატისტიკური დაკვირვებით ცნობილია,

რომ ზამთრის საბურავებზე გადასვლას ახერხებს მძღოლების 80%. იპოვეთ იმის ალბათობა, რომ მძღოლს, რომელიც ფეხით მოსიარულეს შეეჯახა, არ ეყენა ზამთრის საბურავები.

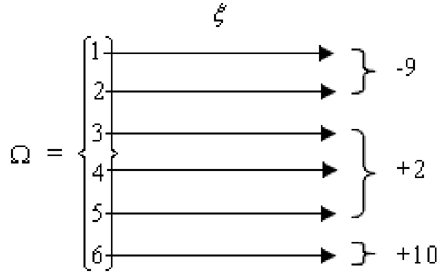
15. ფირმამ გასაყიდად გაიტანა ტელევიზორების პარტია, რომელთაგან 20% დამზადებული იყო I საწარმოში, 35% – II, ხოლო 45% – III საწარმოში. ტელევიზორები გაიყიდა მოკლე დროში და ფირმის ხელმძღვანელობამ გადაწყვიტა N ლარით დააჯილდოვოს საწარმოები. როგორ უნდა განაწილდეს ეს თანხა ამ სამ საწარმოს შორის, თუ ცნობილია, რომ I საწარმო დეფექტურ დეტალებს უშვებს 0,2; II – 0,1, ხოლო III – 0,3 ალბათობით (მითითება: გამოიყენეთ ბაიესის ფორმულები).

პასუხები

1) 0,125. **2**) 0,78. **3**) $\frac{9}{20}$. **4**) $\frac{51}{80}$. **5**) $\frac{9}{17}$. **6**) 0,51; 0,7599. **7**) 0,915. **8**) 0,026. **9**) 0,79. **10**) 0,014. **11**) 0,92. **12**) 0,25. **13**) 0,625. **14**) 0,6. **15**) $16N/79$; $63N/158$; $63N/158$.

§ 22. შემთხვევითი სიდიდე. განაწილება და განაწილების ფუნქცია

მაგალითი 22.1. განვიხილოთ ასეთი თამაში. მოთამაშე აგორებს კამათელს და თუ მოვიდა რიცხვი „6“ იგებს 10 ლარს, თუ მოვიდა „3“, „4“ ან „5“ – 2 ლარს, ხოლო „1“ ან „2“-ის მოსვლაზე იგი იხდის 9 ლარს. მოთამაშის შემოსავალი აღვნიშნოთ „+“, ხოლო დანაკარგი „-“ სიმბოლოთი. მივიღებთ ელემენტარულ ხდომილობათა სივრცეზე განსაზღვრულ რიცხვით ξ ფუნქციას.



ცხადია ξ ფუნქცია ყოველ მნიშვნელობას იღებს იმაზე დამოკიდებულებით თუ ცდის შედეგად რომელი ელემენტა-რული ω ხდომილობა მოხდა. რის გამოც ეს მნიშვნელობა შემთხვევითია და მას ღებულობს გარკვეული ალბათობით. ამიტომ ξ ფუნქციას „**შემთხვევით სიდიდეს**“ უწოდებენ.

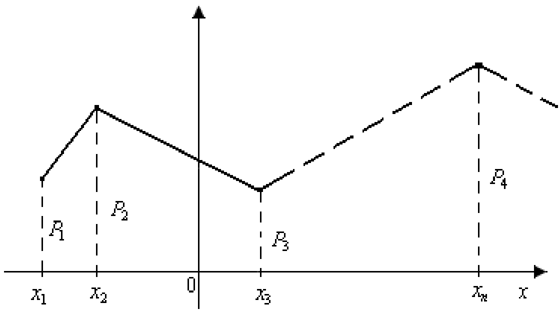
განსაზღვრება 22.1. ξ შემთხვევით სიდიდეს ეწოდება **დისკრეტული ტიპის** თუ იგი იღებს არაუმეტეს თვლადი რაოდენობის მნიშვნელობებს $x_1, x_2, x_3, \ldots, x_i \ldots$ შესაბამისი ალბათობებით $P_i = P(\xi = x_i)$ $i = 1,2,\ldots$ ამ ალბათობების ერთობლიობას (რომლებიც აკმაყოფილებენ პირობას $\sum_i P_i = 1$) შემთხვევითი სიდიდის **განაწილება** ეწოდება. განაწილება შეიძლება მოცემული იყოს ცხრილის სახით.

ξ	x_1	x_2	x_3	x_4	...
P	P_1	P_2	P_3	P_4	...

ცხრ. 22.1. დისკრეტული შემთხვევითი სიდიდის განაწილების ცხრილი

როდესაც დისკრეტული შემთხვევითი სიდიდის მნიშვნელობათა სიმრავლე სასრულია, მას მარტივ შემთხვევით სიდიდეს უწოდებენ.

შეიძლება გრაფიკულად წარმოვადგინოთ 22.1 ცხრილით მოცემული დისკრეტული შემთხვევითი სიდიდის განაწილება. ამისათვის OX ღერძზე გადავზომოთ x_i მნიშვნელობები, ხოლო OY-ზე შესაბამისი P_i ალბათობები. მიღებული $(x_i\ P_i)$ წერტილები შევაერთოთ მიმდევრობით, მიღებულ ტეხილს **დისკრეტული შემთხვევითი სიდიდის მრავალგვერდას** უწოდებენ. (იხ. ნახ. 22.1).



ნახ. 22.1. 22.1 ცხრილით მოცემული დისკრეტული შემთხვევითი სიდიდის მრავალგვერდა

განსაზღვრება 22.2. ფუნქციას

$$F_\xi(x): R \to [0,1]$$

ეწოდება ξ შემთხვევითი სიდიდის **განაწილების ფუნქცია,** თუ $x \in R$ წერტილში მისი მნიშვნელობა გამოითვლება, როგორც Ω სივრცის იმ ω წერტილებისაგან შედგენილი სიმრავლის ალბათობა, რომელშიც $\xi(\omega)$ ნაკლებია x-ზე. ე. ი.

$$F_\xi(x) = P(\omega : \xi(\omega) < x).$$

განაწილების ფუნქციას აქვს შემდეგი თვისებები:

1) $F_\xi(\infty) = \lim\limits_{x \to \infty} F_\xi(x) = 1;$

2) $F_\xi(-\infty) = \lim\limits_{x \to -\infty} F_\xi(x) = -1;$

3) განაწილების ფუნქცია მარცხნიდან უწყვეტი არაკლებადი ფუნქციაა.

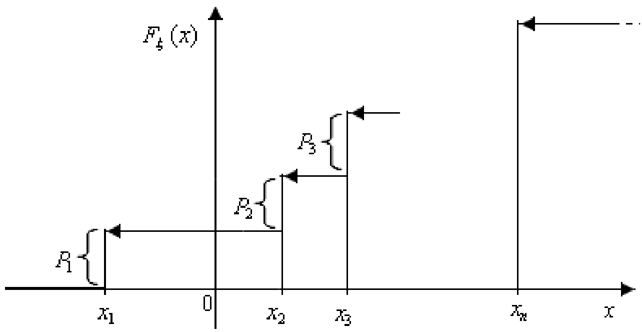
ცხადია, რომ ნებისმიერი $[a,b)$ ინტერვალისათვის

$$P\{\xi \in [a,b)\} = F_\xi(b) - F_\xi(a).$$

ავაგოთ დისკრეტული სიდიდის განაწილების ფუნქციის გრაფიკი. თუ განაწილება მოცემულია ცხრილით.

ξ	x_1	x_2	x_3	...	x_n
P	p_1	p_2	p_3	...	p_n

მისი გრაფიკი წარმოადგენს მარცხნიდან უწყვეტ საფეხურა ფუნქციას, რომელსაც x_i $i=\overline{1,n}$ წერტილებში გააჩნია p_i $i=\overline{1,n}$ სიდიდის ნახტომები (იხ. ნახ. 22.2).



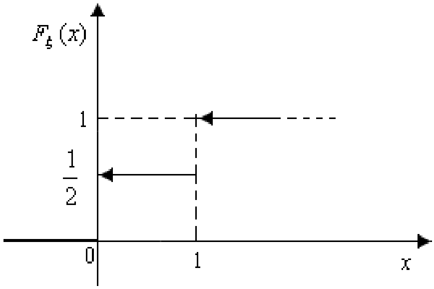
ნახ. 22.2. დისკრეტული შემთხვევითი სიდიდის განაწილების ფუნქციის გრაფიკი

მაგალითი 22.2. ლითონის მონეტას აგდებენ ერთხელ ξ იყოს შემთხვევითი სიდიდე, რომელიც ხდება 1-ის ტოლი თუ მოვიდა "გერბი" და 0-ის ტოლი თუ მოვიდა "საფასური". შევადგინოთ ξ შემთხვევითი სიდიდის განაწილების ცხრილი და ავაგოთ მისი განაწილების ფუნქციის გრაფიკი.

ამოხსნა. ცხადია $P(\xi=1)=\dfrac{1}{2}$ და $P(\xi=0)=\dfrac{1}{2}$, ამიტომ გვექნება ξ შემთხვევითი სიდიდის განაწილების შედეგი ცხრილი

ξ	0	1
P	1/2	1/2

მისი განაწილების ფუნქციის გრაფიკს ექნება სახე



მაგალითი 22.3. A და B მიკროსაფინანსო ორგანიზაციები ერთმანეთისაგან დამოუკიდებლად მუშაობენ საინვესტიციო ბაზარზე. ξ და η შემთხვევითი სიდიდეები აღნიშნავენ დროის ერთი და იგივე ფიქსირებულ შუალედში თითოეული მათგანის მოგებას შესაბამისად (თანხის პირობითი ერთეულებით). მათი განაწილებები მოცემულია ცხრილებით.

ξ	1	2	4
η	0,2	0,5	0,3

η	2	3
P	0,3	0,7

დავადგინოთ დროის ამავე შუალედში მათი საერთო მოგების შესაძლო განაწილება.

ამოხსნა. ცხადია, დასადგენია $\xi+\eta$ შემთხვევითი სიდიდის განაწილება. ამ შემთხვევითი სიდიდის შესაძლო მნიშვნელობებია 3, 4, 5, 6, 7. მაგალითად, 4-ის ტოლი მნიშვნელობა $\xi+\eta$ შემთხვევითმა სიდიდემ შეიძლება მიიღოს ორ შემთხვევაში: ან როდესაც $\xi=2$ და $\eta=2$, ან როდესაც $\xi=1$ და $\eta=3$. ამიტომ ξ და η-ს დამოუკიდებლობის გამო (ვინაიდან A და B კომპანიები ერთმანეთისგან დამოუკიდებლად მუშაობენ) გვაქვს:

$$P(\xi+\eta=4)=P(\xi=2)P(\eta=2)+P(\xi=1)P(\eta=3)=$$
$$=0,5\cdot 0,3+0,2\cdot 0,7=0,29\,.$$

ანალოგიურად,

$$P(\xi+\eta=3)=P(\xi=1)P(\eta=2)=0,2\cdot 0,3=0,06$$
$$P(\xi+\eta=5)=P(\xi=2)P(\eta=3)=0,5\cdot 0,7=0,35$$
$$P(\xi+\eta=6)=P(\xi=4)P(\eta=2)=0,3\cdot 0,3=0,09$$

$P(\xi+\eta=7)=P(\xi=4)P(\eta=3)=0{,}3\cdot 0{,}7=0{,}21$.

მაშასადამე გვექნება განაწილება

$\xi+\eta$	3	4	5	6	7
P	0,06	0,29	0,35	0,09	0,21

განსაზღვრება 22.3. ξ შემთხვევით სიდიდეს ეწოდება **აბსოლუტურად უწყვეტი ტიპის** თუ არსებობს ნამდვილ რიცხვთა ღერძზე განსაზღვრული ისეთი არაუარყოფითი $f_\xi(x)$ ფუნქცია, რომ

$$F_\xi(x)=\int_{-\infty}^{x} f_\xi(t)dt \ . \tag{22.1}$$

ამ $f_\xi(x)$ ფუნქციას ξ **შემთხვევითი სიდიდის განაწილების სიმკვრივე** ეწოდება.

ცხადია, რომ მას გააჩნია შემდეგი თვისებები:

ა) $f(t)\geq 0$

ბ) $\int_{-\infty}^{\infty} f_\xi(t)dt=1$

გ) $f_\xi(t)=F'_\xi(t)$ ყოველ t წერტილში, სადაც $f_\xi(x)$ უწყვეტია.

განსაზღვრება 22.4. თუ ξ შემთხვევითი სიდიდის $F_\xi(x)$ განაწილების ფუნქცია უწყვეტი და მკაცრად ზრდადია, მაშინ მისი ***p*** **რიგის კვანტილი** აღინიშნება x_p სიმბოლოთი და ეწოდება

$$F_\xi(x)=p$$

განტოლების ამონახსნს, ხოლო $x_{1-\alpha}$ კვანტილს უწოდებენ განაწილების **ზედა α კრიტიკულ წერტილს**.

განსაზღვრებიდან ცხადია, რომ

$$P\{\omega:\xi(\omega)<x_p\}=p$$

$$P\{\omega:\xi(\omega)>x_{1-\alpha}\}=\alpha \ .$$

ჩვენ კვანტილისა და ზედა α კრიტიკული წერტილის ცნებები დაგვჭირდება მხოლოდ აბსოლუტურად უწყვეტი განაწილების შემთხვევაში (§ 25), ამიტომ არ ვისახავთ მიზნად ამ საკითხის განხილვას განსაზღვრებაში მოყვანილის გარდა სხვა ტიპის განაწილებებისათვის (იხ. [1]).

ამოცანები

1. ξ შემთხვევითი სიდიდის განაწილება მოცემულია ცხრილით

x_i	-2	7	10	11	15
P_i	1/10	3/10	1/10	1/10	2/5

გამოთვალეთ: $P(\xi < -2)$; $P(\xi \leq 10)$; $P(\xi > 9)$; $P(4 < \xi < 10{,}2)$.

2. შემთხვევითი სიდიდე წარმოადგენს კამათლის ერთხელ გაგორებისას მოსულ რიცხვს. დაწერეთ მისი განაწილება და განაწილების ფუნქცია.

3. დისკრეტული η შემთხვევითი სიდიდის განაწილების ცხრილიდან იპოვეთ a და b თუ ცნობილია, რომ $F_\eta(5{,}2) = 1/2$. გამოთვალეთ $F_\eta(7)$ და $P(\eta > 3)$.

η_i	-4	3	4	7	9	10	15
P_i	1/12	1/6	a	1/6	1/24	b	1/6

4. მოცემულია დისკრეტული ξ შემთხვევითი სიდიდის განაწილება

ξ_i	3	7	9	10	11	17
P_i	1/15	1/3	a	1/5	1/30	1/30

გამოთვალეთ: a; $F_\xi(7)$; $F_\xi(9{,}2)$; $P(7{,}1 \leq \xi < 11{,}3)$; $P(\xi > 10)$.

5. მოცემულია დისკრეტული ξ შემთხვევითი სიდიდის განაწილება

ξ_i	-2	-1	0	1	3
P_i	1/10	1/10	1/15	3/10	3/10

იპოვეთ $\eta = 2\xi - 1$ შემთხვევითი სიდიდის განაწილება და გამოთვალეთ ალბათობები $P(\eta = -2)$; $P(\eta < 1)$; $P(\eta > 0{,}2)$.

6. მოცემულია დისკრეტული ξ შემთხვევითი სიდიდის განაწილება

ξ_i	-2	-1	0	1	3
P_i	1/9	2/9	2/9	1/3	1/9

იპოვეთ $\eta = \xi^2 + 1$ შემთხვევითი სიდიდის განაწილება და გამოთვალეთ $F_\eta(1{,}2)$; $F_\eta(3{,}1)$; $F_\eta(5)$.

7. მოცემულია დამოუკიდებელი ξ და η დისკრეტული შემთხვევითი სიდიდეების განაწილებები

ξ_i	-1	0	2
P_i	1/8	1/4	5/8

η_i	-1	0	1
P_i	1/4	1/2	1/4

იპოვეთ $\tau = \xi + \eta$ შემთხვევითი სიდიდის განაწილება.

8. მოცემულია დამოუკიდებელი ξ და η შემთხვევითი სიდიდეების განაწილებები

ξ_i	0	-1	2
P_i	1/6	1/3	1/2

ξ_i	-2	1
P_i	1/4	3/4

იპოვეთ $\tau = \xi \cdot \eta + 1$ შემთხვევითი სიდიდის განაწილების კანონი.

პასუხები:

1) 0; 1/2; 3/5; 2/5.

2)

ξ_i	1	2	3	4	5	6
P_i	1/6	1/6	1/6	1/6	1/6	1/6

$F_\xi(x) = 0$ თუ $x \le 1$; $F_\xi(x) = 1/6$ თუ $1 < x \le 2$; $F_\xi(x) = 2/6$ თუ $2 < x \le 3$; $F_\xi(x) = 3/6$ თუ $3 < x \le 4$; $F_\xi(x) = 4/6$

თუ $4 < x \le 5$; $F_{\xi}(x) = 5/6$ თუ $5 < x \le 6$; $F_{\xi}(x) = 1$ თუ $x > 6$.
3) $a = 1/4$; $b = 1/8$; $1/2$; $3/4$. **4)** $1/3$; $1/15$; $11/15$; $17/30$; $1/15$.
5)

η	1	2	5	10
P	2/9	5/9	1/9	1/9

0; 2/5; 3/5.

6)

η	-5	-3	-1	1	5
P	1/10	1/10	1/5	3/10	3/10

2/9; $7/9$; $7/9$.

7)

τ	-2	-1	0	1	2	3
P	1/32	1/8	5/32	7/32	5/16	5/32

8)

τ	-3	0	1	3
P	1/8	1/4	1/6	11/24

§ 23. შემთხვევითი სიდიდის რიცხვითი მახასიათებლები. მათემატიკური ლოდინი და დისპერსია

შემთხვევითი სიდიდის განაწილების ფუნქცია გვაძლევს სრულ თუმცა რთულად გასააზრებელ ინფორმაციას შემთხვევითი სიდიდის შესახებ. ხშირად საკმარისია ვიცოდეთ შემთხვევითი სიდიდის ერთი ან რამოდენიმე რიცხვითი მახასიათებელი (რომლებიც ისევ და ისევ განაწილების ფუნქციის საშუალებით გამოითვლებიან) რათა მის შესახებ საკმაოდ ზუსტი წარმოდგენა ვიქონიოთ. ამ რიცხვითი მახასიათებლებიდან პირველ რიგში უნდა გამოიყოს ისინი, რომლებიც წარმოდგენას გვაძლევენ მისი მნიშვნელობების თავმოყრის ადგილის შესახებ და ამ თავმოყრის ადგილის ირგვლივ გაფანტულობის ხარისხზე. ხანდახან სასარგებლოა ამ მახასიათებლებთან ერთად განვიხილოთ ალბათური განაწილების ფორმის ზოგიერთი მახასიათებელიც (უწყვეტი განაწილებების დროს სიმკვრივის ფორმის, ხოლო დისკრეტულის შემთხვევაში განაწილების მრავალგვერდას ფორმის მახასიათებლები).

განვიხილოთ ξ შემთხვევითი სიდიდის ცენტრალური ტენდენციის ერთ-ერთი მახასიათებელი, **მათემატიკური ლოდინი** $\boldsymbol{E\xi}$ (E პირველი ასოა ინგლისური სიტყვისა Expectation-ლოდინი).

ვთქვათ ξ დისკრეტული ტიპის შემთხვევითი სიდიდეა განაწილების კანონით (ცხრ. 23.1)

ξ	x_1	x_2	…	x_n	…
P	P_1	P_2	…	P_n	…

ცხრ. 23.1

განსაზღვრება 23.1. თუ აბსოლუტურად კრებადია მწკრივი

$$\sum_i x_i p_i = x_1 p_1 + x_2 p_2 + \cdots + x_n p_n ,$$

მაშინ მის ჯამს ეწოდება ξ შემთხვევითი სიდიდის **მათემატიკური ლოდინი** და აღინიშნება $E\xi$ სიმბოლოთი. ხშირად მას ξ შემთხვევითი სიდიდის **საშუალოსაც** უწოდებენ.

$$E\xi = \sum_i x_i p_i\,. \qquad (23.1)$$

მაგალითი 23.1. ნებისმიერი A ხდომილობისათვის განვიხილოთ ინდიკატორ ფუნქცია $I_{(A)}$

$$I_{(A)}(\omega) = f(x) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}.$$

გამოვთვალოთ მისი მათემატიკური ლოდინი.

ამოხსნა. ცხადია რომ $I_{(A)}$ შემთხვევით სიდიდეს აქვს შემდეგი დისკრეტული განაწილება

$I_{(A)}$	1	0
P	$P(A)$	$1-P(A)$

ამიტომ

$$EI_{(A)} = 1\cdot P(A) + 0(1-P(A)) = P(A)\,.$$

მაგალითი 23.2. დავუბრუნდეთ 22.1. მაგალითს. ცხადია, მოთამაშის „მოგების“ ფუნქცია არის ξ შემთხვევითი სიდიდე, რომლის განაწილებაც მოცემულია ცხრილით

ξ	-9	2	10
P	1/3	1/2	1/6

მისი მათემატიკური ლოდინი იქნება

$$E\xi = -9\cdot\frac{1}{3} + 2\cdot\frac{1}{2} + 10\cdot\frac{1}{6} = -\frac{1}{3}\,.$$

მაშასადამე, თამაში წამგებიანია მოთამაშისათვის საშუალოდ 1 თამაშში $\frac{1}{3}$ ლარით.

მათემატიკურ ლოდინს გააჩნია მარტივი ფიზიკური შინაარსი. თუ რიცხვითი ღერძის $x_1, x_2, \ldots, x_n$ წერტილებში მოვათავსებთ $m_1, m_2, \ldots, m_n$ მასებს შესაბამისად, მაშინ მათემატიკური ლოდინის საშუალებით შეგვიძლია ვიპოვოთ ამ მასათა სისტემის სიმძიმის ცენტრის კოორდინატი. მოვახდინოთ მასათა სისტემის ნორმირება. დავთვალოთ ყველა

მასების ჯამი $M=\sum_{i=1}^{n} m_n$ და განვიხილოთ შემთხვევითი სიდიდე ξ რომლის განაწილება მოცემულია ცხრილით

ξ	x_1	x_2	…	x_n
P	$\frac{m_1}{M}$	$\frac{m_2}{M}$	…	$\frac{m_n}{M}$

სიმძიმის ცენტრის x კოორდინატი იქნება $E\xi$

$$X=E\xi=\sum_{i=1}^{n} x_i \frac{m_i}{M}.$$

ანალოგიურად სივრცეში განთავსებული მასათა სისტემის ცენტრის კოორდინატები შეგვიძლია გამოვთვალოთ OX, OY და OZ ღერძებზე დაგეგმილების შედეგად.

განსაზღვრება 23.2. თუ ξ აბსოლუტურად უწყვეტი ტიპის შემთხვევითი სიდიდეა $f_\xi(x)$ სიმკვრივით, მისი **მათემატიკური ლოდინი** ეწოდება

$$E\xi=\int_{-\infty}^{\infty} x f_\xi(x)dx$$

სიდიდეს, თუ ცნობილია, რომ მარჯვენა მხარეში ინტეგრალი აბსოლუტურად კრებადია.

დამტკიცების გარეშე მოვიყვანოთ მათემატიკური ლოდინის თვისებები, ვიგულისხმოთ რომ ξ და η შემთხვევით სიდიდეებს გააჩნიათ მათემატიკური ლოდინები

1) მუდმივი შემთხვევითი სიდიდის მათემატიკური ლოდინი თვით ამ მუდმივის ტოლია $EC=C$;
2) $E(\xi+\eta)=E\xi+E\eta$;
3) $E(C\xi)=CE\xi$;
4) თუ ξ და η დამოუკიდებელი შემთხვევითი სიდიდეებია, მაშინ $E(\xi\eta)=E\xi\cdot E\eta$.

შევნიშნოთ, რომ მათემატიკური ლოდინი გვაძლევს შემთხვევითი სიდიდის „ცენტრის“ ანუ იმ „ადგილის“ პოვნის საშუალებას, რომლის ირგვლივაც გაბნეულია შემთხვევითი სიდიდე, თუმცა ამ გაბნევის „სიდიდეს“ ვერ

ახასიათებს. ამის თვალსაჩინო წარმოდგენას იძლევა შემდეგი მაგალითი.

მაგალითი 23.3. გვაქვს ξ და η დისკრეტული შემთხვევითი სიდიდეები

ξ	-1	1
P	1/2	1/2

η	-10	10
P	1/2	1/2

ცხადია, ამ ორ სიდიდეს ტოლი მათემატიკური ლოდინები აქვთ $E\xi = E\eta = 0$, თუმცა ამ საერთო ცენტრის ირგვლივ η სიდიდე უფრო გაფანტულია ვიდრე ξ. შემთხვევითი სიდიდის „გაფანტულობის სიდიდის“ დასახასიათებლად შემოვიტანოთ **დისპერსიის** ცნება.

განსაზღვრება 23.3. ξ შემთხვევითი სიდიდის **დისპერსია** აღინიშნება $D\xi$ სიმბოლოთი (პირველი ასო ინგლისური სიტყვიდან Dispersion) და ეწოდება $\xi - E\xi$ შემთხვევითი სიდიდის კვადრატის მათემატიკურ ლოდინს

$$D\xi = E[\xi - E\xi]^2 . \qquad (23.2)$$

თუ ξ დისკრეტული ტიპისაა (23.1) ცხრილით მოცემული განაწილებით, მაშინ

$$D\xi = \sum_{i=1}^{\infty} (x_i - E\xi)^2 P_i .$$

ხოლო თუ ξ აბსოლუტურად უწყვეტი ტიპისაა $f_\xi(x)$ სიმკვრივით, მაშინ

$$D\xi = \int_{-\infty}^{\infty} (x - E\xi)^2 f_\xi(x)dx .$$

მათემატიკური ლოდინის თვისებების გამოყენებით მარტივად მიიღება რომ

$$D\xi = E\xi^2 - (E\xi)^2 . \qquad (23.3)$$

მოვიყვანოთ დამტკიცების გარეშე დისპერსიის რამოდენიმე თვისება. ვთქვათ C და a რაიმე მუდმივებია:

1) თუ $\xi = C$, მაშინ $D\xi = 0$;

2) $D(C\xi) = C^2 D\xi$;

3) $D(\xi + a) = D\xi$;

4) თუ ξ და η დამოუკიდებელი შემთხვევითი სიდიდეებია

$$D(\xi + \eta) = D\xi + D\eta ,$$

$$D(\xi \ \eta) = D\xi \cdot D\eta + (E\xi)^2 D\eta + (E\eta)^2 D\xi .$$

განსაზღვრება 23.4. ξ შემთხვევითი სიდიდის **სტანდარტული გადახრა** (ანდა **საშუალო კვადრატული გადახრა**) აღინიშნება $\sigma(\xi)$ სიმბოლოთი და ეწოდება არითმეტიკულ ფესვს მისი დისპერსიიდან

$$\sigma(\xi) = \sqrt{D\xi} . \qquad (23.4)$$

მაგალითი 23.4 დავუბრუნდეთ 23.3 მაგალითს. მარტივად გამოვთვლით, რომ

$$D\xi = (-1-0)^2 \cdot \frac{1}{2} + (1-0)^2 \cdot \frac{1}{2} = 1 ; \qquad \sigma(\xi) = \sqrt{1} = 1$$

$$D\eta = (-10-0)^2 \cdot \frac{1}{2} + (10-0)^2 \cdot \frac{1}{2} = 100 ; \qquad \sigma(\eta) = \sqrt{100} = 10 .$$

როგორც ვხედავთ საშუალო კვადრატული გადახრა დისპერსიასთან შედარებით უფრო ზომიერად ახასიათებს შემთხვევითი სიდიდის საშუალოს ირგვლივ გაფანტულობის „ხარისხს“.

შევნიშნოთ რომ დისპერსიას ხშირად σ^2-ით აღნიშნავენ, გამომდინარე საშუალო კვადრატული გადახრის σ აღნიშვნიდან.

განსაზღვრება 23.5. ξ შემთხვევითი სიდიდისაგან მიღებული **სტანდარტიზირებული შემთხვევითი სიდიდე** ეწოდება.

$$\overset{0}{\xi} = \frac{\xi - E\xi}{\sigma(\xi)} ,$$

შემთხვევით სიდიდეს. ცხადია $E\overset{0}{\xi} = 0$ და $D\overset{0}{\xi} = 1$.

შევნიშნოთ რომ შემთხვევითი სიდიდისთვის მისი ლოდინის დაკლებას **ცენტრირება** ეწოდება, შემთხვევითი სიდიდის გაყოფას საშუალო კვადრატულ გადახრაზე – **ნორმირება**, ხოლო ორივე მოქმედების ერთდროულად ჩატარებას შემთხვევითი სიდიდის **სტანდარტიზირება** ეწოდება.

მოვიყვანოთ მაგალითზე დისკრეტული შემთხვევითი სიდიდის დისპერსიის გამოთვლის გამარტივებული სქემა.

მაგალითი 23.5. ცდა მდგომარეობს კამათლის ერთხელ გაგორებაში. ξ შემთხვევითი სიდიდე იყოს კამათელზე მოსული ციფრი. გამოვთვალოთ მათემატიკური ლოდინი და დისპერსია.

ამოხსნა. ცხადია რომ თითოეული ციფრის მოსვლის ალბათობაა $\frac{1}{6}$, ამიტომ ξ შეთხვევით სიდიდეს ექნება შემდეგი განაწილება

ξ	1	2	3	4	5	6
P	1/6	1/6	1/6	1/6	1/6	1/6

დისპერსიის გამოთვლა ჩავატაროთ შემდეგნაირად.

i	x_i	P_i	x_iP_i	x_i^2	$x_i^2P_i$
1	1	1/6	1/6	1	1/6
2	2	1/6	2/6	4	4/6
3	3	1/6	3/6	9	9/6
4	4	1/6	4/6	16	16/6
5	5	1/6	5/6	25	25/6
6	6	1/6	6/6	36	36/6
			$E\xi=\sum_{i=1}^{6}x_ip_i=\frac{7}{2}$		$E\xi^2=\sum_{i=1}^{6}x_i^2p_i=\frac{91}{6}$
$D\xi=E\xi^2-(E\xi)^2=\frac{35}{12}$					

ამოცანები

1. ცნობილია, რომ $E\xi = 2$, $E\eta = 3$. გამოთვალეთ $E(\xi - \eta)$; $E(\xi + 2\eta)$; $E(3\xi + \eta - 5)$.

2. ცნობილია, რომ $D\xi = 3$. გამოთვალეთ $D(2\xi - 1)$; $D(\xi + 4)$; $\sigma(\xi + 1)$; $\sigma(2\xi)$.

3. ξ და η დამოუკიდებელი შემთხვევითი სიდიდეებია. ცნობილია, რომ $E\xi = 10$, $E\eta = 25$, $D\xi = 4$, $D\eta = 9$. გამოთვალეთ $E(\xi\eta)$; $E(2\xi\eta - 5)$; $D(\xi - \eta)$; $D(\xi + \eta)$; $D(2\xi + 3\eta)$.

4. მოცემულია ξ დისკრეტული შემთხვევითი სიდიდის განაწილების ცხრილი

ξ_i	–2	–1	2	4
P_i	1/6	1/3	1/6	1/3

გამოთვალეთ $E\xi$, $D\xi$, $\sigma\xi$.

5. მოცემულია ξ შემთხვევითი სიდიდის განაწილება

ξ_i	–6	0	1	a
P_i	1/6	1/6	1/3	1/3

იპოვეთ a-ს მნიშვნელობა და გამოთვალეთ $D\xi$ თუ $E\xi = 1$.

6. ერთდროულად აგორებენ ორ კამათელს. ξ შემთხვევითი სიდიდე იყოს კამათლებზე მოსული ციფრების ჯამი. დაწერეთ ξ-ს განაწილება და გამოთვალეთ მისი მათემატიკური ლოდინი და დისპერსია.

7. დამოუკიდებელი ξ და η შემთხვევითი სიდიდეები იღებენ მხოლოდ ნულის და ერთის ტოლ მნიშვნელობებს $P(\xi = 1) = 0{,}3$; $P(\eta = 1) = 0{,}1$. გამოთვალეთ $E(\xi - \eta)^2$.

8. ξ დისკრეტული შემთხვევითი სიდიდე იღებს მხოლოდ ორ მნიშვნელობას x_1 და x_2. ცნობილია, რომ $x_2 > x_1$; $P(\xi = x_1) = 0{,}6$; $E\xi = 1{,}4$; $D\xi = 0{,}24$. იპოვეთ ξ შემთხვევითი სიდიდის განაწილება.

პასუხები

1) -1; 8; 4. **2**) 12; 3; $\sqrt{3}$; $2\sqrt{3}$. **3**) 250; 495; 13; 13; 97. **4**) 1; 6; $\sqrt{6}$. **5**) 5; 41/3.

6)

ξ_i	2	3	4	5	6	7	8	9	10	11	12
P_i	1/36	1/18	1/12	1/9	5/36	1/6	5/36	1/9	1/12	1/18	1/36

7; 35/6. **7**) 0,34. **8)** $P(\xi = 1) = 0{,}6$; $P(\xi = 2) = 0{,}4$.

§ 24. ზოგიერთი დისკრეტული განაწილება

ბინომური განაწილება (binomial distribution): ξ შემთხვევით სიდიდეს ეწოდება **ბინომური კანონით განაწილებული**, პარამეტრებით n და p და ჩაიწერება $\xi \sim \mathrm{bi}(n,p)$ თუ იგი იღებს მნიშვნელობებს $0,1,2,\ldots,n$ ალბათობებით $P_n(k)=\mathrm{bi}(k,n,p)=P(\xi=k)=C_n^k P^k(1-P)^{n-k}$. მისი განაწილების ცხრილია

ξ	0	1	...	n
P	$P_n(0)$	$P_n(1)$	...	$P_n(\mathrm{n})$

იმისათვის, რომ კარგად გავიგოთ ბინომური განაწილების არსი განვიხილოთ ბერნულის ცდათა სქემა: ვატარებთ n დამოუკიდებელ ორშედეგიან ცდას. თითოეულში ვაკვირდებით ერთი და იგივე A ხდომილობას, რომლის მოხდენის ალბათობაა p, ხოლო არ მოხდენის კი $q=1-p$. ყოველ i-ურ ცდაში A ხდომილობის მოხდენას ვუწოდოთ „წარმატება“ და აღვნიშნოთ 1-ით, ხოლო არ მოხდენას კი „მარცხი“ და აღვნიშნოთ 0-ით. მივიღებთ v_i შემთხვევითი სიდიდეს $(i=\overline{1,n})$ შემდეგი განაწილებით

v_i	1	0
p_i	p	q

რომელსაც **ბერნულის კანონით** განაწილებული შემთხვევითი სიდიდე ეწოდება.

განვიხილოთ ჯამი $\xi=v_1+v_2+,\ldots,+v_n$. ცხადია, რომ ξ იღებს მნიშვნელობებს $0,1,2,\ldots,n$. $\xi=k$ ნიშნავს რომ n ცდაში ზუსტად k-ჯერ მოხდა A ხდომილობა, ხოლო $n-k$-ჯერ $\overline{A}$, ანუ $v_1,v_2,\ldots,v_n$ შემთხვევითი სიდიდეებიდან ზუსტად k ცალმა მიიღო მნიშვნელობა 1, ხოლო $n-k$-ცალმა მნიშვნელობა 0. ასეთ ელემენტარულ ხდომილობათა რაოდენობაა C_n^k, ხოლო თითოეული მათგანის ალბათობა

($v_1, v_2, \ldots, v_n$ შემთხვევითი სიდიდეების დამოუკიდებლობის გამო) $p^k q^{n-k}$-ს ტოლია, ამიტომ

$$P(\xi = k) = C_n^k p^k q^{n-k} .$$

მარტივად შეგვიძლია ვაჩვენოთ, რომ $Ev_i = p$, $Dv_i = pq$, ხოლო $v_1, v_2, \ldots, v_n$ შემთხვევითი სიდიდეების დამოუკიდებლობის გამო

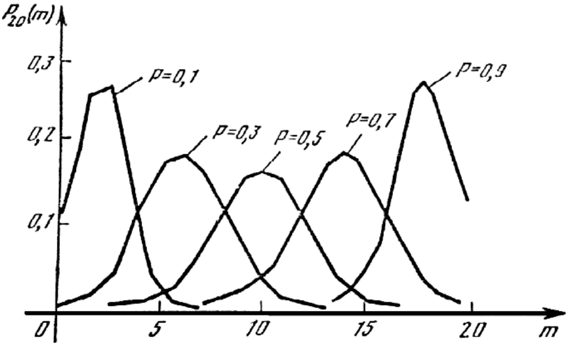
$$E\xi = np , \ D\xi = npq , \ \sigma(\xi) = \sqrt{npq} .$$

ბინომური შემთხვევითი სიდიდის განაწილების ფუნქცია შეგვიძლია ჩავწეროთ $F_n(x, p)$ ან $\mathrm{Bi} = (x, n, p)$ სახით

$$\mathrm{Bi}(x, n, p) = F_n(x\ p) = \sum_{0 \le k < x} C_n^k p^k q^{n-k} .$$

უალბათესი რიცხვი ეწოდება იმ მნიშვნელობას, რომლის ალბათობაც მაქსიმალურია. ბინომურ განაწილებას შეიძლება ჰქონდეს ერთი ან ორი. თუ $(n+1)p$ არ არის მთელი რიცხვი, მაშინ უალბათესი რიცხვი ერთია $k^* = [(n+1)p]$ და მაშასადამე ამ დროს განაწილება უნიმოდალურია. თუ $(n+1)p$ მთელი რიცხვია, მაშინ უალბათესი რიცხვი ორია $k^* = (n+1)p$ და $k^* = (n+1)p - 1$ და მაშასადამე განაწილება ბიმოდალურია.

როგორც ცნობილია *n*-ის ზრდასთან ერთად ბინომური განაწილება თანდათან უფრო სიმეტრიული ხდება თავისი *np* საშუალოს მიმართ. ნებისმიერი ფიქსირებული *n* რიცხვისათვის თუ $p = q = 0{,}5$, ბინომური განაწილების მრავალგვერდა სიმეტრიულია საშუალოს მიმართ. თუ $p > 0{,}5$ მაშინ $\beta < 0$, განაწილებას მარცხენა კუდი აქვს და მარცხნივ ასიმეტრიულია. თუ $p < 0{,}5$, მაშინ $\beta > 0$, განაწილებას მარჯვენა კუდი აქვს და მარჯვნივ ასიმეტრიულია. 24.1 ნახაზზე მოყვანილია $n = 20$ შემთხვევისათვის ბინომური განაწილების მრავალგვერდები *p*-ს რამოდენიმე მნიშვნელობისათვის.



ნახ. 24.1. ბინომური განაწილების მრავალგვერდები $n=20$ შემთხვევისათვის p-ს ზოგიერთი მნიშვნელობისათვის

მაგალითი 24.1. ფირმამ დაამზადა 200-200 ერთნაირი ელექტრული ხელსაწყოსაგან შედგენილი 30 პარტია. თითოეული პარტიის თვითღირებულებაა 5000 ევრო. რეალიზაციამდე ხდება ყოველი პარტიის შემოწმება შემდეგნაირად. ირჩევენ შემთხვევითი პრინციპით 10 დეტალს პარტიიდან და თუ ერთი მაინც აღმოჩნდება არასტანდარტული, მთელ პარტიას აბრუნებენ წარმოებაში დასარეგულირებლად, რისი ღირებულებაც თითოეული პარტიისათვის შეადგენს 1000 ევროს. ვიპოვოთ მთელი პროდუქციის თვითღირებულება საბოლოოდ და მისი საშუალო კვადრატული გადახრა თუ ცნობილია რომ თითოეული დეტალი სტანდარტულია 0,95 ალბათობით და ის არაა დამოკიდებული სხვა დეტალების მდგომარეობაზე.

ამოხსნა. ჯერ ვიპოვოთ იმის ალბათობა, რომ პარტიას შემოწმებისას დაიწუნებენ და დააბრუნებენ წარმოებაში.

ვამოწმებთ $n=10$ დეტალს. τ იყოს არასტანდარტული დეტალების რაოდენობა. „წარმატების“, ანუ დეტალის არასტანდარტულობის ალბათობა იყოს $p=1-0,95=0,05$. ცხადია, τ-ს აქვს ბინომური განაწილება $(\tau \sim \mathrm{bi}(10,\ 0,05))$. A იყოს ხდომილობა-„პარტია დაწუნებულია“.

$$P(A)=P(\tau \ge 1)=1-P(\tau=0)=1-C_{10}^{0}p^{0}(1-p)^{10-0}=1-0,95^{10}\approx 0,4\,.$$

დაწუნებული პარტიების რაოდენობა იყოს η შემთხვევითი სიდიდე. განვიხილოთ ახალი ბერნულის ცდათა სქემა. ცდა იყოს პროდუქციის ერთი პარტიის შემოწმება, წარმატება კი პარტიის დაწუნება. ცდათა რიცხვი $n=30$, ხოლო წარმატების $p=P(A)=0{,}4$ ალბათობა იქნება პარტიის დაწუნების ალბათობა. ცხადია $(\eta \sim \mathrm{bi}(30,\ 0{,}4))$.

$$E\eta = np = 30 \cdot 0{,}4 = 12,$$

$$\sigma\eta = \sqrt{npq} = \sqrt{30 \cdot 0{,}4 \cdot (1-0{,}4)} \approx 2{,}68.$$

მთელი პროდუქციის საბოლოო თვითღირებულება იყოს ξ შემთხვევითი სიდიდე. ცხადია $\xi = C + 1000\eta$, სადაც $C = 30 \cdot 5000 = 150000$ პროდუქციის თვითღირებულებაა შემოწმების პროცედურამდე.

$$E\xi = E(C+1000\eta) = C + 1000E\eta \approx 150000 + 1000 \cdot 12 = 162000,$$

$$\sigma\xi = \sigma(C+1000\eta) = 1000\sigma\eta \approx 1000 \cdot 2{,}68 = 2680.$$

მაშასადამე, პროდუქციის საბოლოო თვითღირებულებაა 162000 ევრო, ხოლო სტანდარტული გადახრა 2680 ევრო.

ამოცანები

1. ξ შემთხვევით სიდიდეს აქვს ბინომური განაწილება პარამეტრით $n=5$; $p=\dfrac{1}{2}$. გამოთვალეთ ალბათობები $p(\xi=2)$; $p(\xi<2)$.

2. ალბათობა იმისა, რომ აუქციონზე გამოტანილი ლოტი გაიყიდება არის 0,6-ის ტოლი. გამოთვალეთ იმის ალბათობა, რომ 4 გამოტანილი ლოტიდან გაიყიდება ზუსტად 3 ცალი.

3. მსროლელი ერთი გასროლით სამიზნეს აზიანებს 0,8 ალბათობით. გამოთვალეთ იმის ალბათობა, რომ მსროლელი სამი გასროლიდან სამიზნეს ვერც ერთხელ ვერ დააზიანებს.

4. ალბათობა იმისა, რომ სავაჭრო ცენტრში შესული მყიდველი შეიძენს 20 ლარზე მეტი ღირებულების ნივთს ტოლია 0,6-ის. გამოთვალეთ იმის ალბათობა, რომ 10 მყიდველიდან 7 შეიძენს 20 ლარზე მეტი ღირებულების ნივთს.

5. ფერმერი კრედიტის მისაღებად ბანკს მიმართავს 0,3 ალბათობით. გამოთვალეთ იმის ალბათობა, რომ 6 ფერმე-რიდან ბანკს კრედიტისთვის მიმართავს: ა) მხოლოდ ერთი მათგანი; ბ) ერთი მათგანი მაინც.

პასუხები

1) 5/16; 3/16. **2**) 0,3456. **3)** 0,008. **4**) $C_{10}^{7}\cdot 0{,}6^{7}\cdot 0{,}4^{3}$. **5**) $C_{6}^{1}\cdot 0{,}3\cdot 0{,}7^{5}$; $1-0{,}7^{6}$.

§ 25. აბსოლუტურად უწყვეტი ტიპის ზოგიერთი განაწილება

ნორმალური განაწილება. ξ შემთხვევით სიდიდეს ეწოდება **ნორმალურად განაწილებული** პარამეტრებით (a,σ^2) და ჩაიწერება $\xi \sim N\left(a,\sigma^2\right)$, თუ მისი განაწილების ფუნქციაა

$$F_\xi(x) = \Phi_{a,\sigma^2}(x) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} e^{\frac{-(t-a)^2}{2\sigma^2}}dt$$

ხოლო სიმკვრივეა

$$f_\xi(x) = \varphi_{a,\sigma^2}\left(x\right) = \frac{1}{\sqrt{2\pi e}} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

მისი რიცხვითი მახასიათებლებია

$$E\xi = a; \quad D\xi = \sigma^2.$$

განვიხილოთ $\eta = \dfrac{\xi - a}{\sigma}$ შემთხვევითი სიდიდე. ცხადია $E\eta = 0$, $D\dot{\eta} = 1$. ნორმალურ განაწილებას პარამეტრებით $a = 0$ და $\sigma^2 = 1$ ეწოდება **სტანდარტული ნორმალური განაწილება.** მისი სიმკვრივე და განაწილების ფუნქცია აღინიშნება შესაბამისად $\varphi(x)$ და $\Phi(x)$ სიმბოლოებით. ე. ი.

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad \Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt.$$

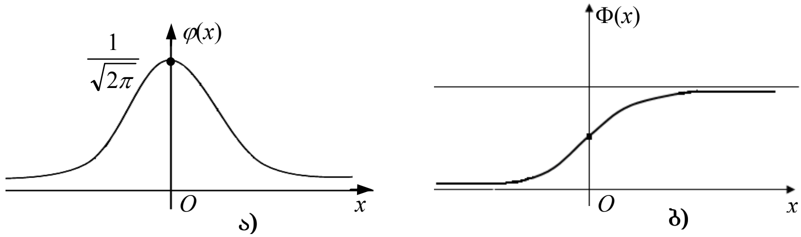
როდესაც ξ განაწილებულია ნორმალურად პარამეტრებით a და σ^2 იყენებენ ჩანაწერს $\xi \sim N(a,\sigma^2)$, ხოლო თუ ξ-ს აქვს სტანდარტული ნორმალური განაწილება, მაშინ ჩაწერენ ასე $\xi \sim N(0,1)$.

სტანდარტული ნორმალური განაწილების ფუნქციისა და მისი სიმკვრივის გრაფიკებს აქვთ სახე (ნახ. 25.1).

სტანდარტული ნორმალური განაწილების p რიგის კვანტილი აღინიშნება x_p-თი, ხოლო ზედა α კრიტიკული

წერტილი z_α-თი, $z_\alpha = x_{1-\alpha}$. ვინაიდან $\Phi(0) = 0,5$, ამიტომ $x_{0,5} = 0$. თუ $p < 0,5$, მაშინ $x_p < 0$, ხოლო თუ $p > 0,5$, მაშინ $x_p > 0$. $\Phi(-x) = 1 - \Phi(x)$ ტოლობის გამო

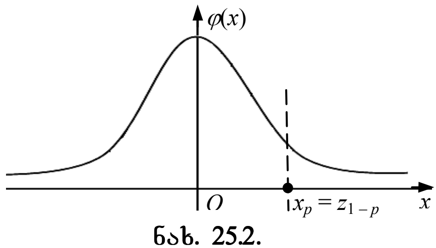
$$x_p = -x_{1-p}.$$



ნახ. 25.1. *N*(0, 1) განაწილების სიმკვრივის (ა) და განაწილების ფუნქციის (ბ) გრაფიკები

ამიტომ სტანდარტული ნორმალური განაწილების კვანტილების ცხრილები მოცემულია მხოლოდ $p > 0,5$ მნიშვნელობებისათვის (იხ. დანართი ცხრ. 2). იქვე მოყვანილია ზოგიერთი z_α ზედა α კრიტიკული წერტილების მნიშვნელობები.

ნახაზზე (ნახ. 25.2) ნაჩვენებია $p > 0,5$ წერტილის x_p კვანტილის შესაძლო ადგილმდებარეობა



ნახ. 25.2.

ცხადია x_p ვერტიკალური ხაზის მარცხნივ $\varphi(x)$-ის გრაფიკსა და ox ღერძს შორის მოთავსებული არის

ფართობია p, ხოლო x_p-ს მარჯვნივ $\varphi(x)$-სა და ox ღერძს შორის მოთავსებული არის ფართობი კი $1-p$.

მაგალითი 25.1. ξ შემთხვევითი სიდიდე განაწილებულია $N(0,1)$ კანონით. იპოვეთ: ა) $z_{0,19}$ ბ) კოორდინატთა სათავის მიმართ სიმეტრიული ინტერვალი, რომელშიც ξ მოხვდება 0,9 ალბათობით.

ამოხსნა. ა) ცხადია $z_{0,19}=x_{1-0,19}=x_{0,81}$ და $\Phi(x)$ განაწილების მნიშვნელობების ცხრილში ვპოულობთ კვანტილს $x_{0,81}$ ანუ იმ წერტილს, რომელშიც $\Phi(x)$-ის მნიშვნელობა 0,81-ის ტოლია

$$z_{0,19}=x_{0,81}=0,88\,.$$

ბ) უნდა ვიპოვოთ ისეთი $[-a,a]$ შუალედი რომელშიც ξ მოხვდება 0,9 ალბათობით ანუ

$$0,9=P(\xi\in[-a,a]=\Phi(a)-\Phi(-a)=$$
$$=\Phi(a)-(1-\Phi(a))=2\Phi(a)-1$$

ანუ

$$\Phi(a)=0,95\,,\ \ a=x_{0,95}\,.$$

ცხრილებში ვპოულობთ, რომ $x_{0,95}=1,65$.

მაშასადამე საძიებელი ინტერვალი ყოფილა $[-1,65\ \ 1,65]$.

სტიუდენტის *t*(*n*) განაწილება განვიხილოთ $\xi_i\sim N(0,\sigma^2)$ $i=0,1,2,\dots,n$ დამოუკიდებელი შემთხვევითი სიდიდეები. შემთხვევით სიდიდეს

$$T(n)=\frac{\xi_0}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\xi_i^2}} \tag{25.1}$$

აქვს **სტიუდენტის განაწილება *n* თავისუფლების ხარისხით** და ჩაიწერება $T(n)\sim t(n)$.

სახელმძღვანელოს დანართში მოცემულია სტიუდენტის $t(n)$ განაწილების $t_{n,\alpha}$ ზედა α კრიტიკული წერტილების მნიშვნელობების ცხრილები სხვადასხვა n-სთვის. $P\{T(n)>t_{n,\alpha}\}=\alpha$ (იხ. დანართი ცხრ. 3). ჩვენ α დონის კვანტილებს აღვნიშნავთ $t_\alpha(n)$-ით, ხოლო ზედა α

კრიტიკულ წერტილებს $t_{n,\alpha}$-თი. $P\{T(n)\le t_{\alpha}(n)\}=\alpha$, ცხადია $t_{n,\alpha}=t_{1-\alpha}(n)$. რადგან $t(n)$ განაწილება სიმეტრიულია 0-ის მიმართ, ამიტომ $t_{\alpha}(n)=-t_{1-\alpha}(n)$.

ცნობილია, რომ $T_{(n)}$ შემთხვევითი სიდიდის სიმკვრივეს, $f_{T_{(n)}}(x)$-ს, ისეთივე ფორმა აქვს როგორიც სტანდარტული ნორმალური განაწილების $\varphi(x)$ სიმკვრივეს, ისიც სიმეტრიულია კოორდინატთა სათავის მიმართ ოღონდ $\varphi(x)$-თან შედარებით უფრო დაბალი გლუვი წვერო აქვს. ცნობილია, რომ

$$ET(n)=0 \qquad DT(n)=\begin{cases}\dfrac{n}{n-2}, & \text{თუ} \quad n>2\\ \infty & \text{თუ} \quad 0<n\le 2\end{cases}.$$

მაგალითი 25.2. ξ შემთხვევით სიდიდეს აქვს სტიუდენტის განაწილება $n=9$ თავისუფლების ხარისხით. ვიპოვოთ ა) $P(\xi<3{,}25)$, ბ) ნულის მიმართ სიმეტრიული ის შუალედი, რომელშიც ξ მოხვდება 0,95 ალბათობით.

ამოხსნა. ა) სტიუდენტის განაწილების ზედა α კრიტიკული წერტილების მნიშვნელობების ცხრილებში $n=9$ სტრიქონის გასწვრივ ვპოულობთ 3,25-ს. ამოვწეროთ მისი შესაბამისი სვეტის თავზე მდგომი α-ს მნიშვნელობა $\alpha=0{,}005$. ის წარმოადგენს ზედა კრიტიკული დონის მაჩვენებელს, ანუ $t_{9,\,0{,}005}=3{,}25$. რაც ნიშნავს, რომ

$$P(\xi>3{,}25)=0{,}005.$$

მაშასადამე

$$P(\xi<3{,}25)=1-0{,}005=0{,}995.$$

ბ) უნდა ვიპოვოთ ისეთი $[-a,a]$ შუალედი, რომ

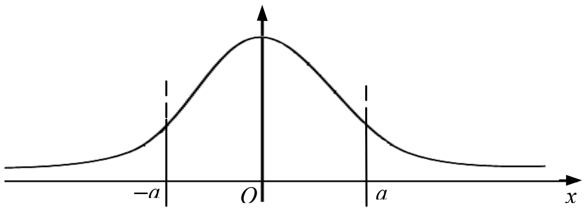
$$P(\xi\in[-a,a])=P(-a<\xi<a)=0{,}95$$

სტიუდენტის განაწილების სიმკვრივის 0-ის მიმართ სიმეტ-რიულობის გამო ნახაზზე (ნახ. 25.3) ჩანს, რომ $-a$-ს მარცხნივ და a-ს მარჯვნივ მოთავსებული ნაწილების

ფართობები ტოლია. მაშასადამე თითოეული მათგანის ფართობი იქნება $\frac{1-0{,}95}{2}=0{,}025$. ეს კი ნიშნავს, რომ a წარმოადგენს ზედა 0,025 კრიტიკულ წერტილს $n=9$ თავისუფლების ხარისხისათვის $a=t_{9;0,025}$. ცხრილებში ვიპოვით, რომ

$$t_{9;0,025}=2{,}262\,.$$

მაშასადამე $a=2{,}262$ და ე.ი. საძიებელი სიმეტრიული შუალედი ყოფილა [–2,262   2,262].



ნახ. 25.3

როდესაც n დიდია ($n\geq 30$) $T(n)$ შემთხვევით სიდიდეს დაახლოებით აქვს $\Phi(x)$ განაწილება $P(T(n)\leq x)\approx\Phi(x)$, ამიტომ შეგვიძლია ვისარგებლოთ სტანდარტული ნორმალური განაწილების ცხრილებით.

ამოცანები

1. ξ შემთხვევით სიდიდეს აქვს სტანდარტული ნორმალური განაწილება. გამოთვალეთ $P(\xi<0{,}5)$ და $P(\xi<0{,}4)$ ალბათობები.

2. ξ შემთხვევითი სიდიდე წარმოადგენს ღუმელის სავენტილაციო გამწოვში ნალექის წონას. დავუშვათ, რომ ის განაწილებულია სტანდარტული ნორმალური კანონით და გამოთვალეთ ალბათობები $P(\xi<1)$ და $P(\xi>1{,}4)$.

3. იპოვეთ ნულის მიმართ სიმეტრიული ინტერვალი, რომელშიც სტანდარტული განაწილების მქონე შემთხვევითი სიდიდე მოხვდება ა) 0,8; ბ) 0,5 ალბათობით.

4. იპოვეთ სტანდარტული ნორმალური განაწილების ზედა 0,3 კრიტიკული წერტილის მნიშვნელობა.

5. ξ შემთხვევით სიდიდეს აქვს სტიუდენტის განაწილება 10 თავისუფლების ხარისხით. იპოვეთ $P(\xi < 2{,}228)$ და $P(\xi > 3{,}581)$ ალბათობები.

6. ξ შემთხვევით სიდიდეს აქვს სტიუდენტის განაწილება 7 თავისუფლების ხარისხით. რისი ტოლია ზედა 0,05 კრიტიკული წერტილის მნიშვნელობა?

7. ξ შემთხვევით სიდიდეს აქვს სტიუდენტის განაწილება 20 თავისუფლების ხარისხით. იპოვეთ ნულის მიმართ სიმეტრიული შუალედი, რომელშიც ის მოხვდება: ა) 0,9 ალბათობით; ბ) 0,98 ალბათობით.

პასუხები

1) 0,691; 0,655. **2**) 0,841; 0,081. **3**) (−1,28, 1,28); (−0,68, 0,68). **4**) 0,53. **5**) 0,975; 0,0025. **6**) 1,895. **7**) (−1,725, 1,725); (−2,528, 2,528).

§ 26. შემთხვევით სიდიდეთა ჯამების კრებადობის საკითხები. დიდ რიცხვთა კანონი

ალბათობის თეორიაში ხშირად განიხილება შემთხვევით სიდიდეთა ჯამები და საჭიროა მათი განაწილების დადგენა, როდესაც შესაკრებთა რაოდენობა უსასრულოდ იზრდება.

ალბათობის თეორიის ერთ-ერთი შემეცნებითი ღირებულება იმითაა განპირობებული, რომ ერთი და იგივე მოვლენის მრავალჯერ განმეორებისას თავს იჩენს ის ძირითადი კანონზომიერება, რომელიც ახასიათებს ამ მოვლენას. მიუხედავად იმისა, რომ მრავალი შემთხვევითი ფაქტორები სხვადასხვა ზეგავლენას ახდენენ ამა თუ იმ ექსპერიმენტის შედეგზე, მაინც ექსპერიმენტის მრავალჯერ ჩატარებისას სხვადასხვა ქაოტური „ზემოქმედებები“ ასე ვთქვათ ერთმანეთს „აქრობენ“ და რჩება ის ძირითადი, რომელიც ამ შედეგს ახასიათებს.

კურსის დასაწყისში ჩვენ აღვნიშნეთ, რომ ლითონის ფულის აგდებისას გერბის მოსვლის ფარდობითი სიხშირე გარკვეული აზრით მდგრადობას იჩენს და ახლოსაა $\frac{1}{2}$-თან. ანალოგიურად, თუ კამათელს ვაგორებთ და 10 გაგორებაში რიცხვი 5 მოვიდა 4-ჯერ ეს ყურადღების მიღმა შეიძლება დაგვრჩეს, მაგრამ თუ 1000-ჯერ გაგორებისას რიცხვი 5 მოვიდა 400-ჯერ ეს უკვე აღარაა შემთხვევითი. თითოეული ჩვენგანი გააკეთებს დასკვნას ამ კამათლის არასიმეტრიულობის შესახებ, ვინაიდან სიმეტრიულობის შემთხვევაში რიცხვი 5 უნდა მოსულიყო დაახლოებით $\frac{1000}{6}\approx 166$-ჯერ. დასაშვებია ამ რიცხვიდან გადახრები, მაგრამ არა იმდენად დიდი, რომ ამ რაოდენობა ცდებისათვის მან 400-ს მიაღწიოს.

ასე, რომ ერთი და იგივე მოვლენის დიდი რაოდენობით მოხდენისას ჩვენ გარკვეულ კანონზომიერებებს ვადგენთ. რაც თავისებურად დიდ რიცხვთა (ანუ ჩატარებული

ცდების დიდ რაოდენობასთან, ან მოვლენის ბევრჯერ მოახდენასთან) ეფექტთან შეიძლება ასოცირდეს.

დიდ რიცხვთა კანონი პირველად ბერნულიმ ჩამოაყალიბა. განვიხილოთ ბერნულის ცდათა მიმდევრობა წარმატების p ალბათობით. ν_n იყოს n-ჯერ ჩატარებულ ცდაში წარმატებათა რაოდენობა.

მოვიყვანოთ დიდ რიცხვთა კანონი ბერნულის ფორმით.

თეორემა. დიდ რიცხვთა კანონი (ბერნულის). $\forall \varepsilon > 0$ რიცხვისათვის

$$P\left\{\left|\frac{\nu_n}{n} - p\right| \geq \varepsilon\right\} \to 0\,, \quad n \to \infty\,.$$

ეს თეორემა ამტკიცებს, რომ ექსპერიმენტის დიდი რაოდენობით ჩატარების დროს ფარდობითი სიხშირე უახლოვდება წარმატების p ალბათობას ანუ ფარდობითი სიხშირე თეორიული უცნობი p ალბათობის შეფასებად გამოგვადგება დიდი n-ის დროს.

III თავი

სტატისტიკური დასკვნების თეორია

I თავში უკვე აღვნიშნეთ, რომ მათემატიკური სტატისტიკის მეთოდებით ხდება სხვადასხვა პროცესებზე და მოვლენებზე დაკვირვებების (ან ექსპერიმენტების) შედეგად მიღებული მონაცემების დამუშავება და ამ მოვლენის შესახებ დასკვნების გაკეთება. ეს დასკვნები შეიძლება იყოს მაგალითად საშუალო მნიშვნელობის შესახებ, განაწილების ან მისი პარამეტრების შესახებ და ა.შ. დასკვნები ეხება ამ მოვლენის მახასიათებელს ზოგადად და არა ერთ კონკრეტულ შედეგს.

ჩვენ შემოტანილი გვაქვს (§3) პოპულაციისა და შერჩევის ცნებები. შერჩევის მონაცემების საფუძველზე ვღებულობთ დასკვნებს, რომლებიც შემდეგ ვრცელდება მთელს პოპულაციაზე (ანუ გენერალურ ერთობლიობაზე). თუ მაგალითად დასკვნა გაკეთდა პოპულაციის ერთეულის საშუალო წონაზე (შემდგომში ასეთ დასკვნებს შეფასებებს დავარქმევთ), ეს არ ნიშნავს, რომ შემთხვევითი შერჩევის ერთ-ერთი მონაცემი რომ ავიღოთ იგი ამ წონის იქნება. (ჩვენი მიღებული დასკვნა ხომ საზოგადოდ მთელ პოპულა-ციას ეხება). თუმცა თუ დიდი მოცულობის შერჩევას ჩავატარებთ, თურმე მიღებული მონაცემების არითმეტი-კული საშუალო ანუ შერჩევითი საშუალო ჩვენი დასკვნის შედეგთან ახლოს იქნება (დიდ რიცხვთან კანონის თანახმად).

განვიხილოთ **მაგალითი.** საჭიროა გავაკეთოთ დასკვნა საქართველოში გავრცელებული მგლების წონის შესახებ. პოპულაციის ელემენტარული ერთეულები ამ შემთხვევაში არიან მგლები, ხოლო პოპულაცია იქნება საქართველოში არსებული ყველა მგელის წონების სიმრავლე. პოპულაციის ელემენტარული ერთეულის (ანუ, ზოგადად, მგელის) წონა აღვნიშნოთ X-ით. ცხადია, X სიდიდეს რაღაც საშუალოც გააჩნია, დისპერსიაც და საზოგადოდ, რომ ვთქვათ ის გამოსაკვლევი შემთხვევითი სიდიდეა. დავუშვათ მკვლე-ვარმა გადაწყვიტა შეიმუშაოს მეთოდი, რომლითაც

შერჩევის მიღების შემდეგ შეძლებს წონების პოპულაციის საშუალოს დადგენას და ამ საშუალოს განაზოგადებს მთელ პოპულაციაზე. ვთქვათ იგი ფიქრობს $n=10$ მოცულობის შერჩევის ჩატარებას. ე.ი. საჭიროა მოპოვებულ იქნეს 10 ცალი მგლის წონის მონაცემი. X_1 იყოს პირველად მოპოვებული მგლის წონის სიდიდე, X_2 მეორედ მოპოვებულის და ა.შ. X_{10} მეათედ მოპოვებულის. სანამ შერჩევა ჩაგვიტარებია $X_1, X_2, \ldots, X_{10}$ შემთხვევითი სიდიდეებია, რომლებიც დამოუკიდებლებია (მათი შემთხვევითი შერჩევის გამო) და თითოეულს აქვს ისეთივე განაწილება როგორიც X-ს (პოპულაციის ზოგად ერთეულს). შეიძლება მკვლევარ სტატისტიკოსს ჯერ შერჩევის მონაცემები არც ჰქონდეს და მათ გარეშე წინასწარ შეიმუშაოს გეგმა, რომლის მიხედვითაც ის მოახდენს მის წინაშე დასმული ამოცანის გადაჭრას. ამ დროს საჭირო ფორმულებში ის რიცხვითი მონაცემების ნაცვლად იყენებს $X_1, X_2, \ldots, X_{10}$ შემთხვევით სიდიდეებს.

დავუშვათ, მკვლევარი ფიქრობს რომ მათი საშუალო წონა

$$\overline{X}_{10} = \frac{X_1 + X_2 + \cdots + X_{10}}{10}. \qquad (*1)$$

ჩათვალოს მგლების წონების პოპულაციის საშუალო მნიშვნელობად. სანამ, შერჩევა ჩატარდება $(X_1, X_2, \ldots, X_{10})$ შემთხვევითი ვექტორია. როგორც კი ამ ათივე მონაცემს მოიპოვებენ მიიღებენ $(x_1, x_2, \ldots, x_{10})$ რიცხვით ათეულს, რომელსაც ეწოდება $(X_1, X_2, \ldots, X_{10})$ ვექტორის რეალიზაცია ანუ კონკრეტული მნიშვნელობა. ხდება მათი ჩასმა (*1) ფორმულაში და მიიღება $\bar{x}_{10}$ კონკრეტული რიცხვი, რომელსაც მკვლევარი მიიჩნევს წონების პოპულაციის საშუალო მნიშვნელობად. თუ სხვა დროს კიდევ ჩატარდება $n=10$ მოცულობის შერჩევა, მიიღებენ $(x_1', x_2', \ldots, x_{10}')$ ახალ რეალიზაციას, რომელიც, საზოგადოდ რომ ვთქვათ, განსხ-

ვავებული იქნება ძველისაგან. საშუალო წონის შესახებ დასკვნა მიიღება მათი ჩასმით (*1) ფორმულაში

$$\bar{x}'_{10} = \frac{x'_1 + x'_2 + \cdots + x'_{10}}{10}.$$

ახლა ჩვენ უკვე ნათლად ვხედავთ განსხვავებას $(x_1, x_2, \ldots, x_{10})$ პრაქტიკულად მოპოვებულ მონაცემებსა და $(X_1, X_2, \ldots, X_{10})$ შემთხვევით ვექტორს შორის. ცხადია პირველი არის ამ უკანასკნელის რეალიზაცია ანუ მისი მნიშვნელობა მოცემული „შერჩევის მომენტში".

როგორც მაგალითზე ვნახეთ პოპულაციას გარკვეული განაწილება გააჩნია. ამიტომ თუ ეს განაწილება აქვს X შემთხვევით სიდიდეს, მაშინ ჩაწერენ რომ პოპულაციას აქვს $\mathsf{L}(X)$ განაწილება და n მოცულობის $X_1, X_2, \ldots, X_n$ შერჩევა შეიძლება განვიხილოთ როგორც $\mathsf{L}(X)$ კანონით განაწილებული შემთხვევითი სიდიდეების სასრული მიმდევრობა. ჩვენ გვინდა რომ ეს $X_1, X_2, \ldots, X_n$ შემთხვევითი სიდიდეები იყოს დამოუკიდებლები (რაც არსებით როლს თამაშობს სტატისტიკურ კვლევებში). თუ პოპულაცია უსასრულოა შერჩევის შემთხვევითობის გამო $X_1, X_2, \ldots, X_n$ სიდიდეები იქნებიან დამოუკიდებლები. სასრული პოპულაციის შემთხვევაში თუ შერჩევა ხდება დაბრუნების გარეშე, შერჩევის თანაბარშესაძლებლობის გამო ვღებულობთ დამოკიდებულ შემთხვევით სიდიდეებს, ამიტომ დამოუკიდებული შემთხვევითი სიდიდეების მისაღებად სასრული პოპულაციის შემთხვევაში საჭიროა შერჩევა ვაწარმოოთ დაბრუნებით. თუმცა სასრული პოპულაციის შემთხვევაშიც კი შემთხვევითი შერჩევის ელემენტები დამოუკიდებლები იქნება იმისდა მიუხედავად თუ როგორ ვაწარმოებთ შერჩევას, დაბრუნებით თუ დაუბრუნებლად, თუ შერჩევის n მოცულობა ძალიან მცირეა პოპულაციის N მოცულობასთან შედარებით. ($\frac{n}{N} < 0{,}05$ იხ. [1]).

მაშასადამე, როდესაც შერჩევას ვატარებთ ვიგულისხმოთ, რომ შერჩევის $X_1, X_2, \ldots, X_n$ ელემენტები არიან

დამოუკიდებელი და ერთნაირი $\mathsf{L}(X)$ კანონით განაწილებული შემთხვევითი სიდიდეები.

მათემატიკური სტატისტიკის ძირითადი ამოცანაა პოპულაციის ზოგადი მახასიათებლების შესაფასებლად სხვადასხვა მეთოდების შემუშავება, სტატისტიკური ჰიპოთეზების შემოწმება ამა თუ იმ მახასიათებლის ან მთლიანად განაწილების შესახებ, სხვადასხვა მახასიათებლებს შორის კავშირის გამოვლენა, ექსპერიმენტის მიმდინარეობის დაგეგმვა და ა.შ.

შევნიშნოთ, რომ ამოცანები რომლებიც სტატისტიკის წინაშე დგას იყოფა ორ ჯგუფად. პარამეტრული და არაპარამეტრული ამოცანები. თუ ცნობილია პოპულაციის $\mathsf{L}(X)$ განაწილების კანონი და მისი უცნობი პარამეტრის ცვლილების რაიმე არე, მაშინ ამოცანა შეიძლება მივაკუთნოთ პარამეტრული ტიპის ამოცანათა კლასს, ხოლო თუ უცნობია $\mathsf{L}(X)$ განაწილების სახე, რომ აღარაფერი ვთქვათ მისი პარამეტრის შესაძლო მნიშვნელობების სიმრავლის შესახებ, მაშინ არაპარამეტრულ ამოცანათა კლასს.

§ 27. შეფასებათა თეორიის ელემენტები. წერტილოვანი შეფასებები

პრაქტიკული ამოცანების გადაჭრისას ხშირად საჭირო ხდება $\mathsf{L}(X)$ განაწილების პოპულაციიდან მიღებული $X_1, X_2, \ldots, X_n$ შერჩევის საშუალებით გავაკეთოთ დასკვნა განაწილების რაიმე უცნობი θ პარამეტრის შესახებ. ანუ „შევაფასოთ" θ პარამეტრი. დავუშვათ ცნობილია Θ სიმრავლე, რომელშიც უნდა ვეძებოთ θ პარამეტრი.

განსაზღვრება 27.1. $X_1, X_2, \ldots, X_n$ შერჩევის ნებისმიერ $T_n = T_n(X_1, X_2, \ldots, X_n)$ ფუნქციას ეწოდება **სტატისტიკა.**

რადგან T_n შემთხვევითი სიდიდეების ფუნქციაა, ამიტომ თავადაც შემთხვევით სიდიდეს წარმოადგენს, როდესაც შერჩევა უკვე ჩატარებულია და მიღებული გვაქვს $X_1, X_2, \ldots, X_n$ შემთხვევითი სიდიდეების $(x_1, x_2, \ldots, x_n)$ რეალიზაცია, მაშინ T_n ღებულობს $T_n = T_n(x_1, x_2, \ldots, x_n)$ რიცხვით მნიშვნელობას.

განსაზღვრება 27.2. ნებისმიერ T_n სტატისტიკას მნიშვნელობებით Θ-ში ეწოდება θ პარამეტრის **შეფასება** და ხშირად აღინიშნება $\hat{\theta}_n$ სიმბოლოთი. $\hat{\theta}_n = \hat{\theta}_n(x_1, x_2, \ldots, x_n)$.

შეფასებას უწოდებენ **წერტილოვანს,** თუ იგი ერთ მნიშვნელობას ღებულობს.

მაგალითი 27.1. კალათბურთის გუნდის მწვრთნელი აკვირდება ახალ მოთამაშეს, რათა დაადგინოს მისი შედეგიანი სროლების პროცენტული მაჩვენებელი. სამ დღეში მას დაუგროვდა მონაცემები.

დღეები	I	II	III
სროლები	30	40	28
შედეგიანი სროლები	20	30	22

შევაფასოთ ბურთის კალათში ტყორცნის სიზუსტე, ანუ ზუსტი სროლის ალბათობა.

ამოხსნა. პოპულაცია (კალათბურთელის ყველა შესაძლო სროლების სიმრავლე) ცხადია, წარმოსახვითია და დიდი მოცულობისაა. თითოეული სროლის შედეგი იყოს X შემთხვევითი სიდიდე, რომელიც იღებს მნიშვნელობას 1 თუ ბურთი კალათში მოთავსდა და 0 თუ არ მოთავსდა. თუ ზუსტი სროლის ალბათობაა p მაშინ X-ს აქვს ბერნულის განაწილება p პარამეტრით. $X \sim \text{bi}(1,P)$, ცხადია ეს არის პოპულაციის განაწილება $\mathsf{L}(X) \in \text{bi}(1,P)$. მწვრთნელის მიზანია შეაფასოს P ალბათობა.

როგორც ჩვენთვის ცნობილია (§ 24) $EX = p$; $D(X) = pq$. შეფასებად გამოვიყენოთ სტატისტიკა $T_n = \overline{X}_n = \dfrac{1}{n}\sum_{i=1}^{n} X_i$.

პირველ დღეს ჩატარებული $n = 30$ მოცულობის შერჩევისათვის $\hat{p}_1 = \overline{x}_{30} = \dfrac{20}{30} = \dfrac{2}{3}$.

მეორე დღეს ჩატარებული $n = 40$ მოცულობის შერჩევისათვის $\hat{p}_2 = \overline{x}_{40} = \dfrac{3}{4}$.

მესამე დღეს – $\hat{p}_3 = \overline{x}_{28} = \dfrac{22}{28} = \dfrac{11}{14}$.

საბოლოოდ მწრთვნელი მოთამაშის სროლების შედეგიანობას შეაფასებს $\hat{p}_1$, $\hat{p}_2$ და $\hat{p}_3$ რიცხვების საშუალო არითმეტიკულით

$$\hat{p}_4 = \frac{1}{3}(\hat{p}_1 + \hat{p}_2 + \hat{p}_3) = 0{,}7341.$$

მას შეეძლო ყველა განხორციელებული 98 სროლისათვის გამოეთვალა ფარდობითი სიხშირე.

$$\hat{p}_5 = \frac{72}{98} = 0{,}7347.$$

მაშასადამე, მიღებულია სროლის სიზუსტის ორი შეფასება $\hat{p}_4$ და $\hat{p}$, რომლებიც უმნიშვნელოდ განსხვავდებიან ერთმანეთისაგან. ნებისმიერი მათგანის მიხედვით მწვრთნელი თვლის, რომ ამ კალათბურთელის სროლათა 73% შედეგიანია.

ყველა სტატისტიკა კარგ შედეგს ვერ მოგვცემს. ზოგიერთი მათგანის საშუალებით მიღებული შედეგები რეალურ სურათს აშკარად ვერ ასახავს. მოვიყვანოთ მაგალითი.

მაგალითი 27.2. ხორცის მწარმოებელმა კომპანიამ გადაწყვიტა შეესწავლა ქალაქის ერთ-ერთ სუპერმარკეტში საქონლის ხორცზე მოთხოვნა. ფირმამ საბუღალტრო ანგარიშებიდან ამოწერა მონაცემები ამ ობიექტზე საქონლის ხორცის ყოველდღიური გაყიდვების შესახებ. (იხ. ცხრ. 27.1) (გავიხსენოთ, რომ ასე მოპოვებულ მონაცემებს § 1-ში მეორადი მონაცემები ვუწოდეთ).

i	1	2	3	4	5	6	7	8	9	10
x_i	35	34	28	22	42	26	34	25	37	34
i	11	12	13	14	15	16	17	18	19	20
x_i	38	41	42	30	32	27	26	40	42	36

ცხრ.27.1

როგორ შევაფასოთ საშუალო დღიური მოთხოვნა საქონლის ხორცზე ამ მაღაზიაში?

ამოხსნა. ცხადია პოპულაციას წარმოადგენს ყოველდღე საქონლის ხორცზე მოთხოვნათა რაოდენობა. ჩვენ კი გვაქვს გაყიდული ხორცის წონები. ბუნებრივია მოთხოვნა და გაყიდული რაოდენობა ერთმანეთს შეიძლება არ ემთხვეოდეს, იმ მიზეზით, რომ შეიძლება გასაყიდი ხორცი დამთავრდეს და იმ დღეს ზოგიერთმა მყიდველმა ვერ შეძლოს მისი ყიდვა. მაგრამ ჩვენ ჩავთვალოთ, რომ ასეთი რამ არ ხდებოდა და ყოველდღე დარჩენილ ხორცს ინახავდნენ მეორე დღისათვის. მაშასადამე, გვაქვს $n=20$ მოცულობის შერჩევა ე.ი. გვაქვს $X_1, X_2, \ldots, X_{20}$ დამოუკიდებელი ერთნაირად განაწილებული შემთხვევითი სიდიდე, რომელთაგან თითოეული წარმოადგენს დღის განმავლობაში გაყიდული ხორცის წონას. განვიხილოთ სტატისტიკა

$$T_{20} = \frac{X_1, X_2, \ldots, X_{20}}{20} = \overline{X}_{20}.$$

როგორც §10-ში ვაჩვენეთ შერჩევითი საშუალო $\overline{X}$ შეიძლება მივიჩნიოთ გაყიდული ხორცის დღიური წონის

სიდიდის (და მაშასადამე მოთხოვნის) საშუალოდ. მოცემული $x_1, x_2, \dots, x_{20}$ რეალიზაციისათვის გამოთვლით ვღებულობთ

$$\bar{x} = \frac{1}{20}\sum_{i=1}^{20} x_i = 33{,}55 .$$

გამოვთვალოთ შერჩევითი დისპერსიაც

$$S^2 = \frac{1}{20}\sum_{i=1}^{20}(x_i - \bar{x})^2 = 38{,}0475,$$

$$S = 6{,}17 .$$

მაშასადამე, შეგვიძლია ჩავთვალოთ, რომ საქონლის ხორცზე ამ რაიონში დღიური მოთხოვნაა საშუალოდ 33,55 კგ, ხოლო სტანდარტული გადახრა 6,17კგ. შეფასებისათვის რომ გამოგვაყენებინა სტატისტიკები

$$T'_{20} = \frac{X_1^2 + X_2^2 + \cdots + X_{20}^2}{20}; \quad T''_{20} = \frac{X_1^2 + X_2^2 + \cdots + X_{20}^2}{20^2}$$

ან

$$T'''_{20} = \frac{X_1X_2 + X_3X_4 + \cdots + X_{19}X_{20}}{20^2}$$

მივიღებდით შეფასებებს:

$$T'_{20} = 1163{,}65 ; \quad T''_{20} = 58{,}18 ; \quad T'''_{20} = 28{,}1$$

რომელთაგან არც ერთი არ გვაძლევს ზუსტ წარმოდგენას ამ სუპერმარკეტით მოსარგებლე მოსახლეობის საქონლის ხორცის მოთხოვნაზე.

ახლა შევეცდებით მოვიყვანოთ პირობები, რომლებსაც უნდა აკმაყოფილებდეს სტატისტიკა, რომ ის „კარგ“ (თუ რა აზრით „კარგ“ ამ საკითხს ქვემოთ შევეხებით) შეფასებად მივიჩნიოთ.

განსაზღვრება 27.3. $T_n(X) = T_n(X_1, X_2, \dots, X_n)$ შეფასებას ეწოდება **ჩაუნაცვლებელი**, თუ $\forall \theta \in \Theta$ პარამეტრისათვის

$$E_\theta T_n(X) = \theta . \qquad (27.1)$$

ვინაიდან X_i, $i = \overline{1,n}$ შემთხვევითი სიდიდეებს აქვთ $\mathsf{L}(X)$ განაწილება (რომლის უცნობ θ პარამეტრსაც აფასებს T_n სტატისტიკა), ამიტომ მათემატიკური ლოდინი გამოითვლება $\mathsf{L}(X)$ განაწილების საშუალებით. თუ უცნობი θ

პარამეტრის მნიშვნელობად ავიღებთ რაიმე θ_0 რიცხვს, მაშინ $E_{\theta_0}T_n(X)$ სიდიდემ უნდა მიიღოს θ_0 მნიშვნელობა.

თუ 27.1 ტოლობა არ სრულდება, მაშინ ასეთი $T_n(X)$ შეფასებისათვის განისაზღვრება

$$b_n(\theta) = E_\theta T_n(X) - \theta$$

სიდიდე, რომელსაც **ჩანაცვლებას** უწოდებენ.

ვთქვათ, გვაქვს n მოცულობის $X_1, X_2, \ldots, X_n$ შერჩევა μ საშუალოსა და σ^2 დისპერსიის მქონე რაიმე პოპულაციიდან. განვიხილოთ ორი სტატისტიკა: $\overline{X}_n$ შერჩევითი საშუალო და S_n^2 შერჩევითი დისპერსია

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i\ ; \quad S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2\ .$$

როდესაც პოპულაციის საშუალო θ_1 უცნობია, მის შეფასებად მივიღოთ $\overline{X}_n$. როდესაც θ_2^2 დისპერსია უცნობია მის შეფასებად, თუ საშუალოც უცნობია, განვიხილოთ S_n^2 შერჩევითი დისპერსია, ან **შესწორებული შერჩევითი დისპერსია**

$$S_n'^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{X})^2\ ,$$

ხოლო თუ საშუალო ცნობილია განვიხილოთ სტატისტიკა

$$\overline{S}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2\ .$$

ამოცანები

1. საცდელ გაფრენებზე მოხდა Boeing 707-300 თვითმფრინავის მიერ 1 საათის ფრენაზე დახარჯული საწვავის რაოდენობის (ლიტრებში) გამოთვლა. მოყვანილი მონაცემების საშუალებით შეაფასეთ ამ მარკის თვითმფრინავის მიერ 1 საათის ფრენაზე დახარჯული საწვავის საშუალო რაოდენობა და დისპერსია.

4373; 4620; 4437; 4563; 4554; 4610; 4512; 4427; 4360; 4545.

2. ბიოლოგებმა გამოიკვლიეს ჩრდილოეთ ამერიკაში გავრცელებული შევარდენის ერთ-ერთი ჯიშის (A. coopori) 20 ზრდასრული ერთეული. მათი წონების მონაცემებით (წონა გრამებში) შეაფასეთ შევარდნების ამ პოპულაციის საშუალო წონა და დისპერსია (მითითება: დისპერსია შეაფასეთ შესწორებული შერჩევითი დისპერსიით).

575;	542;	537;	589;	556;	549;	571;	582;	558;	564
537;	580;	561;	563;	548;	552;	572;	564;	542;	548

3. ტექნიკური მომსახურების გავლის შემდეგ შემოწმდა 10 ცალი რუსული წარმოების МИГ-21-93-ის წონა აღჭურვილობის გარეშე. მიღებული მონაცემებით (კგ-ებში). შეაფასეთ ამ მარკის თვითმფრინავის საშუალო წონა (აღჭურვილობის გარეშე) და დისპერსია

5227	5158	5375	5432	5512
5425	5248	5320	5425	5335

4. ოქროს საბითუმო ყიდვის ფასი (1 ტროიანული უნცია)* 2013 წლის 1 დეკემბრიდან 10 დეკემბერის დღეებში ნიუ-იორკის ძვირფასი ლითონების ბირჟის მონაცემებით („NYSE”-„New York Stock Exchange"), რომელიც ამთავრებს მსოფლიოში დღის განმავლობაში მიმდინარე ვაჭრობას, შეადგენდა შემდეგ სიდიდეებს (დოლარებში).

1261,42	1261,53	1261,48	1261,57	1261,63	1261,60
1261,52	1261,58	1261,81	1261,67	1261,72	1261,64

შეაფასეთ ამ დღეებში ოქროს ფასის საშუალო მნიშვნელობა და დისპერსია (შესწორებული შერჩევითი დისპერსიით).

5. მოცემულია $n = 25$ მოცულობის შერჩევა გამოთვალეთ შერჩევითი საშუალო და შერჩევითი დისპერსია

x_i	5	7	10	11	13	17
n_i	4	8	4	2	1	6

* 1 ტროიანული უნცია = 31,1034768 გრამს.

6. 10 დღის განმავლობაში სუპერმარკეტში გაყიდული სიგარეტის რაოდენობებია (კოლოფებით) 45; 38; 40; 43; 45; 42; 50; 51; 48; 45. შეაფასეთ დღეში გაყიდული კოლოფების საშუალო რაოდენობა და შერჩევითი დისპერსიის საშუალებით შეაფასეთ გაყიდული კოლოფების რაოდენობის დისპერსია.

7. ფერმერს აქვს 16 წლის განმავლობაში მიწის 1 მ^2 ფართობზე მიღებული ლობიოს მოსავლის რაოდენობების მონაცემები (კგ-ში). შეაფასეთ ლობიოს საშუალო მოსავლია-ნობა და მისი დისპერსია

0,4; 0,45; 0,5; 0,35; 0,42; 0,45; 0,41; 0,38;
0,42; 0,39; 0,38; 0,34; 0,39; 0,41; 0,42; 0,4.

პასუხები

1) 4500,1; 8058,09; **2**) 559,5; 234,263. **3**) 5345,7; 10818. **4**) 1261,5975; 0,011. **5**) 10,12; 19,0656. **6**) 44,7; 15,61, **7**) 0,4069; 0,0019.

§ 28. ინტერვალური შეფასებები. ნდობის ინტერვალი

ჩვენ გავეცანით $X=(X_1,X_2,\ldots,X_n)$ შერჩევის საშუალებით უცნობი θ პარამეტრს θ_0 ჭეშმარიტი მნიშვნელობის $\hat{\theta}_n$ „კარგი" შეფასების აგების მეთოდებს. როდესაც შერჩევის n მოცულობა საკმაოდ დიდია, მაშინ $\hat{\theta}_n$ შეფასება „ძალიან ახლოსაა" პარამეტრის θ_0 მნიშვნელობასთან. მაგრამ თვით ეს ფაქტიც კი არ ნიშნავს იმას, რომ შერჩევის ფიქსირებული $(x_1,x_2,\ldots,x_n)$ რეალიზაციისათვის გამოთვლილი $\hat{\theta}_n(x_1,x_2,\ldots,x_n)$ მნიშვნელობა შერჩევის დიდი n მოცულობის დროსაც კი დააკმაყოფილებს $|\hat{\theta}_n(x_1,x_2,\ldots,x_n)-\theta_0|<\varepsilon$ უტოლობას.

ბუნებრივია უფრო მოსახერხებელია ისეთი $(\hat{\theta}_1,\hat{\theta}_2)$ ინტერვალის აგება, რომელიც დიდი ალბათობით მოგვცემს იმის თქმის საშუალებას, რომ პარამეტრის შესაფასებელი θ_0 მნიშვნელობა მოთავსებულია ამ ინტერვალში. ფაქტობრივად ვეძებთ ისეთ ორ $T'(X)<T''(X)$ სტატისტიკას, რომ $(T'(X),\ T''(X))$ შემთხვევითი ინტერვალი, რაღაც γ ალბათობით ფარავდეს შესაფასებელი პარამეტრის θ_0 მნიშვნელობას

$$P(T'(X)<\theta_0<T''(X))=\gamma .$$

X შერჩევის $(x_1,x_2,\ldots,x_n)$ რეალიზაციისათვის გვექნება ფიქსირებული $(\hat{\theta}_1,\hat{\theta}_2)$ ინტერვალი, სადაც

$$\hat{\theta}_1=T'(x_1,x_2,\ldots,x_n)\,,\quad \hat{\theta}_2=T''(x_1,x_2,\ldots,x_n)\,.$$

ჩვენი სურვილია, რომ $(\hat{\theta}_1,\hat{\theta}_2)$ ინტერვალი იყოს რაც შეიძლება მცირე სიგრძის, ხოლო γ ალბათობა რაც შეიძლება დიდი. ვინაიდან ეს ორი ფაქტი ერთმანეთთან უკუკავშირშია, ამიტომ ვიქცევით შემდეგნაირად. წინასწარ ვაფიქსირებთ γ ალბათობას, რაც შეიძლება ახლოს 1-თან. შემდეგ ვეძებთ ისეთი უმცირესი სიგრძის $(T_n'(X),T_n''(X))$

ინტერვალს რომლისთვისაც ყოველი θ-სათვის $\left(\theta \in \Theta\right)$ შესრულებულია პირობა

$$P_\theta(T_n'(X) < \theta_\rho < T_n''(X)) \geq \gamma \, . \qquad (28.1)$$

ეს ნიშნავს, რომ თუ ჩვენ ამ ინტერვალს გამოვიყენებთ θ_0 პარამეტრის შეფასებად n შემთხვევაში, საშუალოდ არანაკლებ $\gamma \cdot n$-ჯერ ჩვენი დასკვნა იქნება ჭეშმარიტი.

განსაზღვრება 28.1. $(T_n'(X),\ T_n''(X))$ შემთხვევით ინტერვალს ეწოდება θ პარამეტრის **γ საიმედობის ნდობის ინტერვალი** (ან γ ნდობის ინტერვალი), თუ იგი არანაკლებ γ ალბათობით ფარავს შესაფასებელი θ პარამეტრის ჭეშმარიტ მნიშვნელობას. ე.ი. სრულდება პირობა

$$P_\theta(T_n'(X) < \theta < T_n''(X)) \geq \gamma \, , \qquad \forall \theta \in \Theta$$

γ რიცხვს უწოდებენ **ნდობის ალბათობას**.

ნდობის ალბათობას ხშირად წარმოადგენენ $\gamma = 1 - \alpha$ სახით. მოცემული ნდობის ინტერვალით უცნობი პარამეტრის შეფასების სიზუსტეს. სიზუსტეს უწოდებენ ნდობის ინტერვალის სიგრძის ნახევარს, ანუ $\dfrac{T_n''(X) - T_n'(X)}{2}$ სიდიდეს.

შენიშვნა: $\gamma = 1 - \alpha$ ტოლობის გამო ბუნებრივია, რომ γ საიმედობის ნდობის ინტერვალს ხშირად **$1 - \alpha$ საიმედობის** ან $(1-\alpha)\cdot 100\%$-იან ნდობის ინტერვალს უწოდებენ.

§ 29. ნდობის ინტერვალის აგება ნორმალურად განაწილებული პოპულაციის უცნობი საშუალოსათვის

შევთანხმდეთ და ქვემოთ ყველგან ვიგულისხმოთ, რომ გენერალური ერთობლიობიდან მიღებული გვაქვს n მოცულობის შერჩევა და როდესაც $\overline{X}_n$ შერჩევით საშუალოს, S_n^2 და $S_n'^2$ შერჩევით დისპერსიებს, Z_n და T_n სტატისტიკებს გამოვიყენებთ ინდექსი n გამოვტოვოთ ჩანაწერის სიმარტივისათვის.

ჯერ განვიხილოთ $N(\theta,\sigma^2)$ მოდელი, როდესაც **პოპულაციის** σ^2 **დისპერსია ცნობილია**. $1-\alpha$ საიმედობის ნდობის ინტერვალი θ უცნობი საშუალოსათვის არის ინტერვალი

$$\left(\overline{X}-x_{1-\alpha/2}\frac{\sigma}{\sqrt{n}},\quad \overline{X}+x_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right). \tag{29.1}$$

$x_{1-\alpha/2}=z_{\alpha/2}$ და მისი პოვნა შესაძლებელია დანართის ცხრილებში (ცხრ. 2).

მაგალითი 29.1 ცნობილია, რომ ერთ-ერთი სასურსათო მაღაზიის დღიური მოგება ნორმალურად განაწილებული შემთხვევითი სიდიდეა 20 ლარის ტოლი სტანდარტული გადახრით. მენეჯერს სურს რომ შეაფასოს საშუალო დღიური მოგება. მან მოახდინა დაკვირვება 16 დღის განმავლობაში და მიიღო შერჩევა.

215; 185; 210; 220; 196; 198; 202; 182; 188; 194; 221; 207;
191; 195; 197; 203.

ავაგოთ ნდობის ინტერვალი საშუალო დღიური მოგებისათვის 0,99 საიმედობის ალბათობით.

ამოხსნა. $n=16$, $\sigma=20$ გამოვთვალოთ შერჩევითი საშუალო $\bar{x}=200{,}25$, $1-\alpha=0{,}99$, ამიტომ $z_{\alpha/2}=z_{0,005}=2{,}57$;

$$\frac{\sigma}{\sqrt{n}}=5;\quad z_{\alpha/2}\frac{\sigma}{\sqrt{n}}=2{,}57\cdot 5=12{,}85.$$

შევიტანოთ ეს მნიშვნელობები (29.1) ინტერვალში. მივიღებთ

$$(200,25-12,85;\quad 200,25+12,85)=(187,4\ ;\ 213,1)\ .$$

განვიხილოთ შემთხვევა, როდესაც პოპულაციის დისპერსია უცნობია. მაშასადამე ვიხილავთ $N(\theta_1,\theta_2^2)$ მოდელს, ასეთ დროს (29.1) ინტერვალი აღარ გამოგვადგება ვინაიდან σ^2 უცნობია. შევცვალოთ იგი S'^2 შერჩევითი დისპერსიით. განვიხილოთ სტატისტიკა.

$$T=\frac{\overline{X}-\theta}{\frac{S'}{\sqrt{n}}}\,,$$

რომელსაც ცნობილია, რომ აქვს სტიუდენტის განაწილება $n-1$ თავისუფლების ხარისხით. დავაფიქსიროთ $1-\alpha$ საიმედობის ალბათობა და ვიპოვოთ

$$P\{|T|<x\}=1-\alpha$$

ტოლობიდან x. გავიხსენოთ, რომ სტიუდენტის განაწილების სიმკვრივე სიმეტრიულია ნული წერტილის მიმართ და აქვს თითქმის ისეთივე ფორმა როგორც $\varphi(x)$ ფუნქციას. ამიტომ ამ ტოლობიდან მივიღებთ ტოლობას

$$P\{T<x\}=1-\frac{\alpha}{2}\,.$$

მაშასადამე,

$$x=t_{n-1}(1-\frac{\alpha}{2})\,,$$

რომელიც, თუ გავიხსენებთ სტიუდენტის $t(n-1)$ განაწილების ზედა α კრიტიკული წერტილის განსაზღვრებას, ტოლია $t_{n-1;\frac{\alpha}{2}}$ ზედა $\dfrac{\alpha}{2}$ კრიტიკული წერტილის. ე.ი.

$$1-\alpha=P\Big(|T|<t_{n-1;\,\alpha/2}\Big)=P\left\{t_{n-1;\,\alpha/2}<\frac{\overline{X}-\theta}{\frac{S'}{\sqrt{n}}}<t_{n-1;\,\alpha/2}\right\}=$$

$$= P\left\{\overline{X} - \frac{S'}{\sqrt{n}} t_{n-1;\,\alpha/2} < \theta < \overline{X} + \frac{S'}{\sqrt{n}} t_{n-1;\,\alpha/2}\right\}.$$

მივიღეთ $1-\alpha$ საიმედობის ნდობის ინტერვალი ნორმალური განაწილების საშუალოსათვის, როდესაც დისპერსია უცნობია.

$$\left(\overline{X} - \frac{S'}{\sqrt{n}} t_{n-1;\,\alpha/2};\ \ \overline{X} + \frac{S'}{\sqrt{n}} t_{n-1;\,\alpha/2}\right). \tag{29.2}$$

შეგვეძლო გამოგვეყენებინა იგივე T სტატისტიკა

$$T = \frac{\overline{X} - \theta}{\dfrac{S}{\sqrt{n-1}}}$$

სახით და მივიღებდით ნდობის ინტერვალს

$$\left(\overline{X} - \frac{S}{\sqrt{n-1}} t_{n-1;\ \alpha/2};\ \ \overline{X} + \frac{S}{\sqrt{n-1}} t_{n-1;\ \alpha/2}\right). \tag{29.3}$$

მოვიყვანოთ სტიუდენტის განაწილების ზედა α კრიტიკული წერტილების ცხრილის ფრაგმენტი (იხ. ცხრ. 29.1).

n \ α	0,1	0,05	0,025	0,01	0,005	0,0025	0,001
$\vdots$	…	…	…	…	…	…	…
$n=20$	1,325	1,725	2,086	2,528	2,845	3,153	3,552
$\vdots$	…	…	…	…	…	…	…

ცხრ. 29.1. სტიუდენტის განაწილების ზედა α კრიტიკული წერტილების მნიშვნელობების ცხრილის ნაწილი (n =20)

მაგალითი 29.2. ვიპოვოთ ინტერვალი, რომელშიც $\xi \sim t(20)$ შემთხვევითი სიდიდე მოხვდება 0,95 ალბათობით.

ამოხსნა. $1-\alpha = 0{,}95$; $\alpha = 0{,}05$; $\dfrac{\alpha}{2} = 0{,}025$. 29.1 ცხრილში მოვძებნოთ $n=20$ თავისუფლების ხარისხისათვის ზედა 0,025 კრიტიკული წერტილი $t_{20;0,025} = 2{,}086$. მაშასადამე საძიებელი ინტერვალია (–2,086 2,086).

მაგალითი 29.3. დავუბრუნდეთ 15.1 მაგალითს. პოპულაციას წარმოადგენს ქვის ცვენის ადგილების დაშორებები A ქალაქიდან (B ქალაქისაკენ მიმავალ გზაზე) გაზომილი კილომეტრებში. პოპულაციის საშუალოც და დისპერსიაც უცნობია. გვაქვს $n=50$ მოცულობის შერჩევა. ავაგოთ ორი ნდობის ინტერვალი უცნობი საშუალოსათვის 0,9 და 0,99 ნდობის ალბათობებით.

ამოხსნა. გამოვიყენოთ 15.1 მაგალითის ამოხსნისას ჩატარებული გამოთვლები

$$\bar{x}_{50} = 126{,}952, \quad S_{50}^2 = 13{,}063, \qquad S_{50} = 3{,}614 .$$

თუ დავაკვირდებით სიხშირეთა ჰისტოგრამას (ნახ. 15.1), ჩანს რომ მას ნორმალური განაწილების სიმკვრივესთან მიახლოებული ფორმა აქვს. შეგვიძლია ვიგულისხმოთ, რომ პოპულაციას ნორმალური განაწილება აქვს. მაშასადამე ვიხილავთ $N(\theta_1,\theta_2^2)$ მოდელს. ნდობის ინტერვალი $1-\alpha$ საიმედობით (29.2) აიგება $T = \dfrac{X-\theta}{\dfrac{S}{\sqrt{n-1}}}$ სტატისტიკით, რომელსაც სტიუდენტის განაწილება აქვს $n-1$ თავისუფლების ხარისხით. მივიღებთ

$$\left[\overline{X} - \frac{S}{\sqrt{n-1}} t_{n-1;\,\alpha/2},\ \overline{X} + \frac{S}{\sqrt{n-1}} t_{n-1;\,\alpha/2} \right].$$

რადგან მონაცემთა რაოდენობა დიდია $n>30$ შეგვიძლია T ჩავთვალოთ სტანდარტული ნორმალური კანონით განაწილებულად და სტიუდენტის ზედა $\dfrac{\alpha}{2}$ კრიტიკული წერტილი შევცვალოთ ნორმალური განაწილების ზედა $\dfrac{\alpha}{2}$ კრიტიკული წერტილით, $t_{n-1;\alpha/2} = z_{\alpha/2}$.

ჯერ განვიხილოთ შემთხვევა $1-\alpha = 0{,}9 \Rightarrow \alpha = 0{,}1$, მაშასადამე $t_{n-1;\,\alpha/2} = Z_{\alpha/2} = Z_{0,05} = 1{,}64$, ამიტომ $\dfrac{S}{\sqrt{n-1}} z_{0,05} =$

$=\frac{3,614}{7}\cdot 1,64=0,847$. ე.ი. ნდობის ინტერვალი 0,9 საიმედობის ალბათობით იქნება

$$[126,105\ ,\ 127,799].$$

მეორე ნდობის ინტერვალი იგება საიმედობით $1-\alpha=0,99\Rightarrow\alpha=0,01$, მაშასადამე $z_{0,005}=2,57$, ამიტომ $\frac{S}{\sqrt{n-1}}z_{0,005}=1,327$, ე.ი. 0,99 საიმედობის ნდობის ინტერვალი იქნება

$$[125,625\ ,\ 128,279].$$

მაგალითი 29.4. სამშენებლო ფირმას დასაქმებული პერსონალის კვების უზრუნველსაყოფად (სხვა პროდუქტებთან ერთად) სჭირდება თვეში 500 ერთეული გაყინული ქათამი. ფირმის წარმომადგენელმა გასასინჯად შეიძინა 20 ერთეულისაგან შედგენილი პარტია, რომელთა წონები (მოყვანილია 0,05 კგ სიზუსტით) აღმოჩნდა შემდეგი რიცხვები:

0,9	1	0,95	1,1	1,05	1,25	0,85	1,1	1	1
1,05	0,95	1,2	0,95	1,05	1,1	1,2	0,9	0,9	0,95

ა) ავაგოთ 0,95 საიმედობის ნდობის ინტერვალი უცნობი საშუალოსათვის;

ბ) რა თანხის ფარგლებში უნდა ელოდოს ფირმის ბუღალტერი 0,95 ალბათობით 1 თვის განმავლობაში ქათმების საყიდლად გაწეულ დანახარჯს თუ 1 კგ ქათმის ხორცის გასაყიდი ფასია 3€.

ამოხსნა. ვინაიდან გაყინული დაკლული ქათმის წონა გაყიდვის შემდეგ დამოკიდებულია თავად ქათმის წონაზეც, რომელიც მრავალი სხვადასხვა ფაქტორების ზეგავლენას განიცდის. გასუფთავებისა და შეფუთვის პროცედურაზეც, რომელიც ინდივიდუალურია თითოეული ქათმისათვის და გაყინვის პროცედურაზეც, რომელიც ასევე სხვადასხვა ფორმით მიმდინარეობს თითოეული ეგზემპლარისათვის, ამიტომ შეგვიძლია ჩავთვალოთ, რომ გაყინული ქათმის წონა ნორმალური კანონითაა განაწილებული.

ა) ქათმების წონის დისპერსია უცნობია, ამიტომ უცნობი საშუალოს ნდობის ინტერვალი იქნება (29.3) ინტერვალი

$$\left(\overline{X}-\frac{S}{\sqrt{n-1}}t_{n-1;\,\alpha/2},\ \ \overline{X}+\frac{S}{\sqrt{n-1}}t_{n-1;\,\alpha/2}\right).$$

ჩვენს ამოცანაში $1-\alpha=0{,}95$, ე.ი. $\frac{\alpha}{2}=0{,}025$, $n=20$.

გამოთვლების შედეგად ვღებულობთ

$$\overline{x}=1{,}0225;\quad S^2=\frac{1}{20}\sum_{i=1}^{20}(x_i-\overline{x})^2\approx 0{,}0114;\quad S=0{,}1068.$$

სტიუდენტის განაწილების ცხრილში მოვძებნით $t_{19,\ 0{,}025}=2{,}093$ მაშასადამე $\frac{S}{\sqrt{n-1}}t_{n-1;\,\alpha/2}=0{,}0513$. ე.ი. 0,95 საიმე-დობის ნდობის ინტერვალი 1 ცალი გაყინული ქათმის საშუალო წონისათვის იქნება

(0,9712; 1,0738).

ბ) $500\cdot 0{,}9712=485{,}6$; $500\cdot 1{,}0738=536{,}9$, ამიტომ 500 ცალი გაყინული ქათმის საშუალო წონა მოთავსებული იქნება 485,6 და 536,9 კილოგრამებს შორის. იმის გათვალისწინე-ბით, რომ 1კგ გაყინული ქათმის ხორცის ფასია 3 ევრო. ფირმის მოსალოდნელი დანახარჯი მოთავსებული იქნება 1456,8 და 1610,7 ევროს შორის.

მაგალითი 29.5. აშშ-ში მოსახლე ლათინოამერიკელი მამა-კაცების შემოსავლებისა და მიგრაციის საკითხების კვლე-ვისას 2000 წლის აღწერიდან (The U.S. Census 2000, IPUMS. Integrated Public Use Micro data Series, იხ. [21]) მიღებულია კუბელი მამაკაცების შემოსავლების მონაცემების $n=29233$ მოცულობის შერჩევა. მათთვის გამოთვლილია $\overline{X}=\$24018$; $S'=\$36298$. ავაგოთ 0,02 მნიშვნელოვნობის დონის ნდობის ინტერვალი აშშ-ს კუბელი წარმოშობის მამაკაცების საშუალო შემოსავლისათვის.

ამოხსნა. $\overline{X}=\$24018$; $n=29233$, $S'=\$36298$, $\alpha=0{,}02$.

გამოვიყენოთ (29.2) ინტერვალი. ვინაიდან n დიდია $t_{n-1,\,\alpha/2}$ შეგვიძლია შევცვალოთ $z_{\alpha/2}=z_{0{,}01}=2{,}33$-ით

$$Z_{0{,}01}\frac{S'}{\sqrt{n}}\approx \$494{,}65\,.$$

მაშასადამე, ჩვენი საძიებელი 0,02 მნიშვნელოვნობის ნდობის ინტერვალი იქნება

($23532,35 , $24512,65).

ამოცანები

1. §27-ის ამოცანების № 1 ამოცანის პირობებში ააგეთ 0,95 საიმედობის ნდობის ინტერვალი Boeing 707-300 თვითმფრინავის 1 საათის ფრენისათვის საჭირო საწვავის რაოდენობის საშუალოსათვის.

2. §27-ის ამოცანების № 2 ამოცანის პირობებში ააგეთ 0,99 საიმედობის ნდობის ინტერვალი ჩრდილოეთ ამერიკაში გავრცელებული შევარდნების (A coopori ჯიშის) პოპულაციის საშუალო წონისათვის.

3. ააგეთ 0,998 საიმედობის ნდობის ინტერვალი რუსული წარმოების ავიაგამანადგურებელ МИГ 21-93-ის საშუალო წონისათვის (აღჭურვილობის გარეშე) გამოიყენეთ § 27-ის ამოცანებიდან № 3 ამოცანის მონაცემები.

4. ააგეთ 0,998 საიმედობის ნდობის ინტერვალი ერთ ტროიანული უნცია ოქროს საშუალო საბითუმო ფასისათვის 2013 წლის დეკემბრის 1-10 რიცხვებში ნიუ-იორკის ძვირფასი ლითონების ბირჟის მონაცემებით. გამოიყენეთ § 27-ის ამოცანებიდან № 4 ამოცანის მონაცემები.

5. საწარმო უშვებს ფეხბურთის ბურთებს დაშვებული სტანდარტული წონიდან 280 გრამი დისპერსიით. შეამოწმეს გამოშვებული 10 ბურთი, რომელთა საშუალო წონა აღმოჩნდა $\overline{X} = 428$ გრ. ააგეთ 0,99 საიმედობის ნდობის ინტერვალი ბურთის საშუალო წონისთვის.

პასუხები

1) [4421,7, 4578,5]. **2**) [549,71, 569,29]. **3**) [5196,7, 5494,7]. **4**) [1261,4755, 1261,7195]. **5**) [414,4, 441,6].

§ 30. ამოცანის ზოგადი დასმა სტატისტიკური ჰიპოთეზების შესახებ

პოპულაციიდან მიღებული შერჩევის გამოყენებით ჩვენ შეგვიძლია პოპულაციის რაიმე უცნობი პარამეტრის წერტილოვანი ან ინტერვალური შეფასებების მიღება. მაგრამ ამასთანავე შეიძლება ჩვენს წინაშე სულ სხვა ტიპის ამოცანებიც დაისვას. მაგალითად:

1) შერჩევის მონაცემების საშუალებით შევამოწმოთ ვარაუდი იმის შესახებ, რომ პოპულაციის განაწილება ნორმალურია.
2) ცნობილია, რომ პოპულაციის განაწილება ნორმალურია და გვაინტერესებს შევამოწმოთ ვარაუდი იმის შესახებ, რომ დისპერსია 7-ის ტოლია.
3) ცნობილია, რომ პოპულაციის განაწილება ნორმალურია და გვაინტერესებს შევამოწმოთ ვარაუდი იმის შესახებ, რომ საშუალო მეტია 12-ზე.
4) შევამოწმოთ ვარაუდი იმის შესახებ, რომ პოპულაცია განაწილებულია exp(0,4) მაჩვენებლიანი კანონით.
5) მოცემული გვაქვს ერთი და იგივე სიდიდეზე დაკვირვებები $X_1, X_2, \ldots, X_n$ და $Y_1, Y_2, \ldots, Y_m$ ჩატარებული სხვადასხვა დროის ინტერვალებში. გვაინტერესებს პირველი და მეორე შერჩევა (ანუ პირველი და მეორე პოპულაცია, რომლებიდანაც ეს შერჩევებია აღებული) ერთნაირადაა განაწილებული თუ არა? მაგალითად პირველი შერჩევა იყოს მიწისძვრის მოხდენის t_0 დრომდე ქანებში მიმდინარე ტექტონიკური პროცესის შედეგად მიღებული ელექტრომაგნიტური გამოსხივების (ე.მ.გ) რიცხობრივი მაჩვენებლები გაზომილი $(t_0 - a, t_0)$ დროის შუალედში, ხოლო მეორე შერჩევა იგივე მაჩვენებლები გაზომილი $(t_0, t_0 + b)$ შუალედში.

6) მოცემულია ორგანზომილებიანი $(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$ შერჩევა ელემენტალური ერთეულების ორი სხვადასხვა პოპულაციიდან. მაგალითად, გვაინტერესებს რომელიმე კოლეჯის სპორტულ სექციებში დაკავებული მოსწავლეების X_i აკადემიური მოსწრებების და სპორტულ სექციებში გატარებული Y_i დროის ხანგრძლივობების სიდიდეები დამოუკიდებლები არიან თუ არა?

განსაზღვრება 30.1. რაიმე გამონათქვამს, ანუ მოსაზრებას პოპულაციის განაწილების კანონის ან მისი რიცხვითი მახასიათებლების შესახებ **სტატისტიკური ჰიპოთეზა** ეწოდება. ჰიპოთეზას აღნიშნავენ H სიმბოლოთი.

განსაზღვრება 30.2. სტატისტიკურ ჰიპოთეზას ეწოდება **პარამეტრული,** თუ ის ეხება პოპულაციის განაწილების რაიმე პარამეტრს (ამოცანები 2) და 3)), ხოლო თუ ჰიპოთეზა ეხება განაწილების სახეს მას **არაპარამეტრული** ეწოდება.

განსაზღვრება 30.3. სტატისტიკურ ჰიპოთეზას ეწოდება **მარტივი,** თუ ის ცალსახად განსაზღვრავს პოპულაციის განაწილების კანონს (ამოცანა 4)) ან პარამეტრს, როდესაც პარამეტრულ ჰიპოთეზასთან გვაქვს საქმე (ამოცანა 2)). წინააღმდეგ შემთხვევაში, სტატისტიკურ ჰიპოთეზას ეწოდება **რთული** (ამოცანები 1) და 3)).

პარამეტრული ჰიპოთეზის შემოწმების დროს მასთან ერთად განიხილავენ **ალტერნატიულ (საწინააღმდეგო)** ჰიპოთეზას. მაგალითად თუ ძირითადი ჰიპოთეზაა $H_0:\sigma^2=4$ ალტერნატიული შეიძლება იყოს $H_1:\sigma^2>4$, $H_2:\sigma^2<4$, $H_3:\sigma^2\neq 4$, $H_4:\sigma^2=4$. პირველს ეწოდება მარჯვენა ცალმხრივი ალტერნატივა, მეორეს მარცხენა ცალმხრივი ალტერნატივა, მესამეს ორმხრივი ალტერნატივა, ხოლო მეოთხეს კი მარტივი ალტერნატივა.

პოპულაციის უცნობი პარამეტრი იყოს θ. განვიხილოთ ძირითადი $H_0:\theta=\theta_0$ პარამეტრული ჰიპოთეზა და მისი მარტივი ალტერნატივა $H_1:\theta=\theta_1$.

ჩვენ უნდა შევამოწმოთ ჭეშმარიტია H_0 ჰიპოთეზა, თუ არა H_1 ალტერნატივის წინააღმდეგ. ეს შემოწმება უნდა ჩატარდეს n მოცულობის $X=(X_1,X_2,\ldots,X_n)$ შერჩევის საშუალებით.

უნდა მოვიფიქროთ გადაწყვეტილების მიღების ისეთი წესი, ანუ კრიტერიუმი, რომელიც გვიჩვენებს თუ როგორი $(x_1,x_2,\ldots,x_n)$ დაკვირვებული მნიშვნელობებისათვის უნდა მივიღოთ H_0 ჰიპოთეზა და როგორისათვის უარვყოთ. ამისათვის ვიყენებთ $T_n(X)=T(X_1,X_2,\ldots,X_n)$ სტატისტიკას, რომლის განაწილება წინასწარ ცნობილია და ვაგებთ ისეთ U არეს (ანუ ტერიტორიას), რომელშიც მოცემული სტატისტიკის გამოთვლილი $T_n(x)$ მნიშვნელობის მოხვედრის შემდეგ H_0 ჰიპოთეზას უარვყოფთ და სამართლიანად ვთვლით H_1 ჰიპოთეზას. ამ U არეს **კრიტიკულ არეს** უწოდებენ, ხოლო T_n სტატისტიკას კი **კრიტერიუმის სტატისტიკას** ანდა უბრალოდ **კრიტერიუმსაც** უწოდებენ. მოცემული $(x_1,x_2,\ldots,x_n)$ შერჩევისათვის გამოთვლილ $T_n(x)$ სტატისტიკის მნიშვნელობას ხშირად **სტატისტიკის დაკვირვებულ მნიშვნელობას** უწოდებენ.

შეიძლება გადაწყვეტილების მიღებისას დავუშვათ ორი სახის შეცდომა:

1) შეცდომა, როდესაც H_0 ჰიპოთეზას უკუვაგდებთ, ანუ არ მივიღებთ მაშინ, როდესაც ის ჭეშმარიტია და ამ დროს ვღებულობთ მცდარ ალტერნატიულ H_1 ჰიპოთეზას. ამ შეცდომას **პირველი გვარის შეცდომას** უწოდებენ;

2) შეცდომა როდესაც H_0 ჰიპოთეზას მივიღებთ, მაშინ როდესაც ის არაა ჭეშმარიტი და ე.ი. უკუვაგდებთ ჭეშმარიტ ალტერნატიულ H_1 ჰიპოთეზას. ამ შეცდომას **მეორე გვარის შეცდომა** ეწოდება.

ეს ორივე შეცდომა შეიძლება ცხრილის სახით ჩავწეროთ (იხ. ცხრ. 30.1).

	გ ა დ ა წ ყ ვ ე ტ ი ლ ე ბ ა	
	არ უარვყოფთ H_0-ს	უარვყოფთ H_0-ს
ჭეშმარიტია H_0	მიღებული გადაწყვეტილება სწორია	მიღებული გადაწყვეტილება მცდარია **I გვარის შეცდომა**
ჭეშმარიტია H_1	მიღებული გადაწყვეტილება მცდარია **II გვარის შეცდომა**	მიღებული გადაწყვეტილება სწორია

ცხრ. 30.1

საერთოდ ცდილობენ, რომ პირველი გვარის შეცდომის ალბათობა იყოს რაც შეიძლება მცირე.

წამოვიდგინოთ ასეთი მაგალითი. საჭიროა შეირჩეს ნათურა ოთახისათვის, რომელიც დაბალი ხარისხის დენის წყაროთი მარაგდება. გვაქვს ყუთით ნათურები, რომლებ-შიც ერთი ცალი არის ისეთი, რომელც გამოირჩევა ძაბვის ცვალებადობის მიმართ გამძლეობით. პიროვნება მიდის ყუთთან, იღებს 1 ცალს ამოწმებს მას სპეციალურ სტენდზე, სადაც ძაბვის ცვალებადობის რეჟიმში ამუშავებენ ნათურებს გარკვეული დროის განმავლობაში. თუ მოეწონება მისი ნათება მიაქვს, ოთახის გასანათებლად. თუ არა და ცვლის მას და ახალ ნათურას არჩევს. შეიძლება მან დაუშვას ორნაირი შეცდომა. პირველი – თუ იგი კარგ ნათურას ჩათვლის უვარგისად და გადააგდებს; მეორე – თუ იგი ცუდ ნათურას ჩათვლის ხარისხიანად და წაიღებს მას.

დავაკვირდეთ რა შედეგი მოყვება თითოეულ შეცდომას. პირველი შეცდომის შემდეგ კარგი ნათურა უკვე აღარ არსებობს და ამიტომ პიროვნებამ რამდენი ნათურაც არ უნდა გამოსცადოს ან მაშინვე დაიწუნებს თითოეულს ან ჩათვლის ხარისხიანად და წაღების. შემდეგ რამოდენიმე დღეში ისევ მივა ახლის წასაღებად.

მეორე შეცდომის შემდეგ მართალია მას ცუდი ნათურა შერჩება ხელთ, მაგრამ როდესაც ის გაფუჭდება, ახლის შერჩევის დროს პიროვნებას ისევ აქვს შანსი კარგი ნათურა შეარჩიოს.

მკითხველი უკვე დარწმუნდა, რომ I გვარის შეცდომის შანსი რაც შეიძლება მცირე ჯობია რომ იყოს.

ახლა დაუბრუნდეთ ზემოთ განხილულ ცხრილს. რადგან U კრიტიკული არეა და H_0 ჰიპოთეზას უკუვაგდებთ, როდესაც T სტატისტიკა ხვდება U არეში, ამიტომ პირველი გვარის შეცდომის ალბათობაა

$$P\{T \in U \mid H_0\}.$$

რაც წარმოადგენს იმის ალბათობას, რომ T სტატისტიკა იღებს მნიშვნელობას U არიდან, მაშინ როდესაც ცნობილია, რომ H_0 ჰიპოთეზა სამართლიანია. ანალოგიური მოსაზრებიდან გამომდინარე ცხადია, რომ მეორე გვარის შეცდომის ალბათობაა

$$P\{T \notin U \mid H_1\}.$$

ამ დროს ვინაიდან $T(X) \notin U$ ჩვენ არ უარვყოფთ H_0-ს, ვიღებთ მას და ვამბობთ, რომ H_0 ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს.

განსაზღვრება 30.4. კრიტერიუმის მნიშვნელოვნობის დონე აღინიშნება α სიმბოლოთი და ეწოდება პირველი გვარის შეცდომის ალბათობას

$$P\{T \in U \mid H_0\} = \alpha.$$

განსაზღვრება 30.5. მეორე გვარის შეცდომის ალბათობას აღნიშნავენ β სიმბოლოთი

$$P\{T \notin U \mid H_1\} = \beta,$$

ხოლო $1-\beta$ სიდიდეს T **კრიტერიუმის სიმძლავრეს** უწოდებენ. ის წარმოადგენს H_0 ჰიპოთეზის უარყოფის ალბათობას, როდესაც სწორია H_1.

$$1-\beta = 1 - P\{T \notin U \mid H_1\} = P(T \in U \mid H_1).$$

პრაქტიკაში სასარგებლოა, რომ ორივე გვარის შეცდომის ალბათობა იყოს რაც შეიძლება მცირე. ე.ი. კრიტე-

რიუმის მნიშვნელოვნობის დონე α იყოს რაც შეიძლება მცირე, ხოლო სიმძლავრე $1-\beta$ რაც შეიძლება დიდი. ისეთი U არის აგება, რომ α და β ორივე იყოს უმცირესი პრაქტიკულად შეუძლებელია. ამიტომ აფიქსირებენ α სიდიდეს, რაც შეიძლება მცირეს მაგალითად 0,05 ან 0,01 და შემდეგ ცდილობენ ისეთი U არის აგებას, რომ შესრულდეს

$$P\{T \in U \mid H_0\} \le \alpha$$

და $1-\beta$ შეძლებისდაგვარად დიდი გამოვიდეს.

მაგალითად თუ მიიღეს კრიტერიუმი, რომლისთვისაც $\alpha = 0{,}01$, ხოლო $1-\beta = 0{,}8$ ეს ნიშნავს, რომ ამ კრიტერიუმის მიხედვით ჭეშმარიტი H_0 ჰიპოთეზის უგულვებელყოფის ალბათობაა 0,01, ხოლო სიმძლავრეა 0,8 ე.ი. H_1 ალტერნატივის სამართლიანობის დროს საშუალოდ 80 შემთხვევაში 100-დან H_0 ჰიპოთეზას „გადავაგდებთ“ ანუ მივიღებთ სწორ გადაწყვეტილებას, ხოლო 20 შემთხვევაში H_0-ს ჩავთვლით ჭეშმარიტად და მივიღებთ მცდარ გადაწყვეტილებას.

I და II გვარის შეცდომების და კრიტერიუმის სიმძლავრის გეომეტრიული ინტერპრეტაციისათვის განვიხილოთ θ პარამეტრის შესახებ რაიმე მარტივი H_0 ჰიპოთეზა და მისი მარტივი ალტერნატივა H_1

$$H_0 : \theta = \theta_0 ,$$

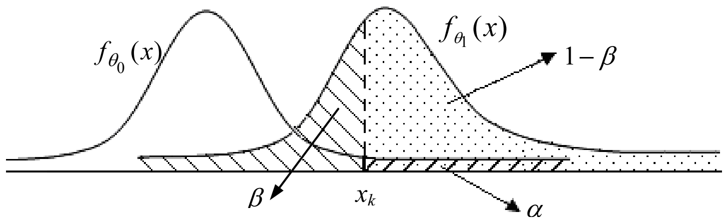
$$H_0 : \theta = \theta_1 .$$

კრიტერიუმის სტატისტიკა იყოს $T(X,\theta)$ სტატისტიკა განაწილების $F_{T(\theta)}(x)$ ფუნქციით და $f_{T(\theta)}(x)$ სიმკვრივით, რომელიც $\theta = \theta^*$ მნიშვნელობისათვის აღვნიშნოთ $f_{\theta^*}(x)$ სახით $(f_{\theta^*}(x) = f_{T(\theta^*)}(x))$.

დავაფიქსიროთ კრიტერიუმის მნიშვნელოვნობის დონე α. განვიხილოთ შემთხვევა, როდესაც $\theta_1 > \theta_0$ და $f_{\theta_1}(x)$ სიმკვრივის გრაფიკი მოთავსებულია $f_{\theta_0}(x)$ სიმკვრივის გრაფიკის მარჯვნივ (ნახ. 30.1). კრიტიკული არე იყოს, კრი-

ტერიუმის სტატისტიკის გამოყენებით მიღებული $(x_k,+\infty)$ მარჯვნიდან უსასრულო ინტერვალი.

შემდეგ $x=(x_1,x_2,\ldots,x_n)$ შერჩევისათვის გამოვთვალოთ $T(x)$ სტატისტიკის მნიშვნელობა.



ნახ. 30.1

1) თუ მივიღეთ $T(x)>x_k$ ამ დროს ძირითად H_0 ჰიპოთეზას უკუვაგდებთ და თუ იგი სამართლიანი იყო, მაშინ პირველი გვარის შეცდომის ალბათობა

$$\alpha=P(T(x)>x_k\mid H_0)=1-P_{\theta_0}(T(x)<x_k)=1-F_{T(\theta_0)}(x_k)$$

წარმოადგენს x_k წერტილის მარჯვნივ OX ღერძსა და $f_{\theta_0}(x)$ ფუნქციის გრაფიკს შორის მოთავსებული უსასრულო მრუდწირული ტრაპეციის ფართობს.

2) თუ მივიღეთ $T(x)\le x_k$ ამ დროს ძირითად H_0 ჰიპოთე-ზას ვღებულობთ და თუ იგი მცდარი იყო, ანუ სამართ-ლიანი იყო H_1, მაშინ II გვარის შეცდომის ალბათობა

$$\beta=P(T(x)\le x_k\mid H_1)=P_{\theta_1}(T(x)\le x_k)=F_{T(\theta_1)}(x_k)$$

წარმოადგენს x_k წერტილის მარცხნივ OX ღერძსა და $f_{\theta_1}(x)$ ფუნქციის გრაფიკს შორის მოთავსებული უსასრულო მრუდწირულ ტრაპეციის ფართობს.

კრიტერიუმის სიმძლავრე $1-\beta$, წარმოადგენს x_k წერ-ტილის მარჯვნივ OX ღერძსა და $f_{\theta_1}(x)$ ფუნქციის გრაფიკს შორის მოთავსებული არის ფართობს.

§ 31. მარტივი ჰიპოთეზის შემოწმება ნორმალური პოპულაციის საშუალოს შესახებ

განვიხილოთ ნორმალური კანონით განაწილებული პოპულაცია $\mathsf{L}(X) \in N(a,\sigma^2)$. ვიგულისხმოთ, რომ σ^2 დისპერსია ცნობილია. შევამოწმოთ ჰიპოთეზა საშუალოს შესახებ.

შევამოწმოთ $H_0 : a = a_0$ ძირითადი ჰიპოთეზა $H_1 : a = a_1$ ალტერნატივის წინააღმდეგ. მოცემული გვაქვს პოპულაციიდან n მოცულობის $X_1, X_2, \ldots, X_n$ შერჩევა. განვიხილოთ ორი შემთხვევა

1) $H_0 : a = a_0$;

$H_1 : a = a_1$; $a_1 < a_0$.

როგორც ვიცით (§ 27) $\overline{X}_n$ წარმოადგენს უცნობი a_0 საშუალოს შეფასებას. როდესაც პოპულაცია ნორმალურადაა განაწილებული $\overline{X}_n$ საშუალოსაც ნორმალური განაწილება აქვს და

$$E\overline{X}_n = a_0; \quad D\overline{X}_n = \frac{\sigma^2}{n}.$$

განვიხილოთ სტატისტიკა

$$T_n(X) = Z_n(X) = \frac{\overline{X}_n - a_0}{\dfrac{\sigma}{\sqrt{n}}}. \qquad (31.1)$$

ცხადია, $Z_n(X) \sim N(0,1)$. როდესაც პოპულაციის საშუალოა a_0, მაშინ $Z_n == 0$, ხოლო თუ პოპულაციის საშუალოა a_1, მაშინ

$$Z_n = \frac{a_1 - a_0}{\dfrac{\sigma}{\sqrt{n}}}.$$

ავარჩიოთ კრიტიკული არე. როდესაც Z_n სტატისტიკა მცირე მნიშვნელობას იღებს, ინტუიციურად უნდა ვიგულისხმოთ, რომ H_1 ჰიპოთეზას უფრო ექნება ადგილი

ვიდრე H_0 ჰიპოთეზას, ამიტომ კრიტიკული არე ავირჩიოთ ნორმალური განაწილების მარცხენა კუდის მხრიდან.

დავაფიქსიროთ I გვარის შეცდომის ალბათობა α, ანუ კრიტერიუმის მნიშვნელოვნობის დონე და ვიპოვოთ მისი შესაბამისი კვანტილი x_α, რომელიც როგორც ვიცით ნორმალური განაწილების ზედა α კრიტიკული წერტილის z_α-ს მოპირდაპირე რიცხვია, $x_\alpha = -z_\alpha$. კრიტიკული არე Z_n სტატისტიკისათვის იქნება $U = (-\infty, -z_\alpha)$ ინტერვალი.

გამოვთვალოთ $Z_n(X)$ სტატისტიკის მნიშვნელობა მოცემული შერჩევისათვის, $Z_n = Z_n(x_1, x_2, \ldots, x_n) = \dfrac{\overline{X}_n - a_0}{\frac{\sigma}{\sqrt{n}}}$ და თუ აღმოჩნდა, რომ ის არ აღემატება $-z_\alpha$ რიცხვს მაშინ H_0 ჰიპოთეზას უარვყოფთ და მივიღებთ H_1 ალტერნატივას, ხოლო თუ სტატისტიკის გამოთვლილი მნიშვნელობა მეტი აღმოჩნდა $-z_\alpha$ რიცხვზე, მაშინ H_0 ჰიპოთეზას უარყოფის საფუძველი არა გვაქვს. მაშასდამე I გვარის შეცდომის ალბათობაა

$$P\left(\frac{\overline{X}_n - a_0}{\frac{\sigma}{\sqrt{n}}} < -z_\alpha \mid H_0 \right) = \Phi(-z_\alpha) = \Phi(x_\alpha) = \alpha \,.$$

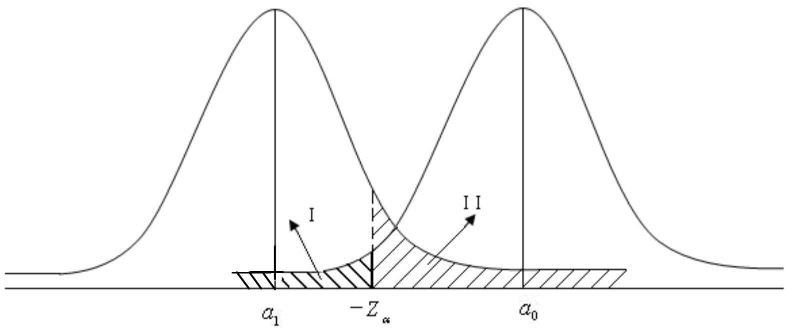
მეორე გვარის შეცდომის ალბათობა იქნება

$$\beta = P\left(\frac{\overline{X}_n - a_0}{\frac{\sigma}{\sqrt{n}}} \geq -z_\alpha \mid H_1 \right).$$

ნახაზზე (ნახ. 31.1) აღნიშნულია ორივე გვარის შეცდომის შესაბამისი არეები.

მაშასადამე, Z_n სტატისტიკისათვის კრიტიკული არეა $(-\infty; -z_\alpha)$. ის არაა დამოკიდებული a_1 რიცხვზე, გამოვი-

ყენეთ მხოლოდ ის ფაქტი, რომ $a_1 < a_0$ და ავაგეთ მარცხენა კრიტიკული არე.



ნახ. 31.1

საერთოდ დიდი მნიშვნელობა აქვს ძირითადი და ალტერნატიული ჰიპოთეზების ზუსტად ჩამოყალიბებას. ამის საილუსტრაციოდ გთავაზობთ 31.1 მაგალითს, რომელსაც ახლა და პარაგრაფის ბოლოს ამოვხსნით სხვადასხვა გზით.

მაგალითი 31.1. ცნობილია რომ ნათესი ფართობის ერთ კვადრატულ მეტრზე ხორბლის საშუალო მოსავლიანობა ნორმალურად განაწილებული შემთხვევითი სიდიდეა 50გრ. დისპერსიით. აგრონომების ნაწილმა გამოიყენა საცდელ ნაკვეთებზე ახალი ტიპის სასუქი და ელიან ხორბლის მოსავლიანობის გაუმჯობესებას ერთ კვადრატულ მეტრზე 380 გრამს. თუმცა მათი მოწინააღმდეგე ნაწილი თვლის რომ მოსავლიანობა ისევ ძველ საშუალო დონეზე 340 გრამზე (ერთ კვადრატულ მეტრიდან) დარჩება.

10 საცდელი ნაკვეთიდან მოსავლის აღების შემდეგ მიღებულია ხორბლის საშუალო მოსავლიანობის მაჩვენებლები თითოეული ნაკვეთის ერთ კვადრატულ მეტრზე

390; 380; 390; 320; 360; 340; 385; 365; 390; 410.

შევამოწმოთ $\alpha = 0{,}01$ მნიშვნელოვნობის დონით შემდეგი ჰიპოთეზა

$$H_0 : m = 380$$

$$H_1 : m = 340 .$$

ხორბლის წონა ერთ კვადრატულ მეტრზე აღნიშნულია m სიმბოლოთი.

ამოხსნა. როგორც ვხედავთ ალტერნატიული H_1 ჰიპოთეზა გვიჩვენებს ძირითად H_0 ჰიპოთეზაზე ნაკლებ მნიშვნელობას. ამიტომ კრიტიკულ არედ ავიღოთ $(-\infty; -z_{0,01})$ ინტერვალი. ცხრილებიდან მოვძებნით, რომ $z_{0,01} = 2,33$. გამოთვლებით მივიღებთ $\bar{x}_{10} = 373$. ცნობილია $\sigma^2 = 50$. განვიხილოთ

$$Z_{10}(X) = \frac{\overline{X}_{10} - 380}{\frac{\sqrt{50}}{\sqrt{10}}}$$

სტატისტიკის მნიშვნელობა მოცემული შერჩევისათვის. მივიღებთ $Z_{10}(x) = -3,13$.

რადგან $-3,13 < -2,33$, ე.ი. $Z_{10}(X) < -z_{0,01}$ ე.ი. კრიტერიუმის სტატისტიკა მოხვდა კრიტიკულ არეში. მაშასადამე ჰიპოთეზა, რომ ახალი ტიპის სასუქმა მოსავლიანობა ერთ კვადრატულ მეტრზე გახადა 380 გრ უნდა უარვყოთ 0,01 მნიშვნელოვნობის დონით.

2. განვიხილოთ შემთხვევა, როდესაც $a_1 > a_0$

$$H_0 : a = a_0$$

$$H_0 : a = a_1 ; \quad a_1 > a_0 .$$

ისევ გამოვიყენოთ სტატისტიკა

$$Z_n(X) = \frac{\overline{X}_n - a_0}{\frac{\sigma}{\sqrt{n}}} .$$

გავითვალისწინოთ, რომ $\overline{X}_n$ სიდიდის ზრდა a_1 ალტერნატივის უპირატესობაზე მეტყველებს და ამიტომ კრიტიკული არე ავარჩიოთ $Z_n(X)$ სტატისტიკის განაწილების, ანუ სტანდარტული ნორმალური განაწილების მარჯვენა კუდის მხრიდან.

ფიქსირებული α მნიშვნელოვნობის დონისათვის ვიპოვოთ სტანდარტული ნორმალური განაწილების ზედა α კრიტი-

კული წერტილი z_α და $Z_n(X)$ სტატისტიკის შესაბამის კრიტიკულ არედ ავარჩიოთ $(z_\alpha;\ +\infty)$ ინტერვალი.

ცხადია პირველი გვარის შეცდომის ალბათობაა

$$P\left(\frac{\overline{X}_n - a_0}{\sigma/\sqrt{n}} \geq z_\alpha \mid H_0\right) = 1 - \Phi(z_\alpha) = 1 - (1-\alpha) = \alpha .$$

კრიტიკული არეა: $(z_\alpha; +\infty)$.

მაგალითი 31.2 უნივერსიტეტის რექტორმა გადაწყვიტა წარჩინებულ სტუდენტებს აუნაზღაუროს ლექციებზე დასწრებისათვის საჭირო მგზავრობის მატერიალური დანახარჯები. ცნობილია, რომ საშუალოდ ეს დანახარჯი ნორმალურად განაწილებული შემთხვევითი სიდიდეა თვეში 4 ლარი დისპერსიით. უნივერსიტეტის წარმომადგენლებმა შეისწავლეს ერთი თვის განმავლობაში 16 წარჩინებული სტუდენტის ლექციებზე სიარულისათვის საჭირო სამგზავრო ხარჯები და მიიღეს მონაცემები

49	55	46	51	52	48	47	46
52	51	49	47	49	56	53	55

რექტორმა გადაწყვიტა სამგზავროდ გამოყოს დანამატი 50 ლარი, მაგრამ არის ალტერნატიული წინადადებაც, რომ ეს თანხა იყოს 70 ლარი. შევამოწმოთ $\alpha = 0{,}05$ მნიშვნელოვნობის დონით შემდეგი ჰიპოთეზა (თანხა აღვნიშნოთ S ით).

$H_0:\ S = 50$;

$H_0:\ S = 70$.

ამოხსნა. $n = 16$, $\sigma^2 = 4$, $\alpha = 0{,}05$, რადგან $70 > 50$, ამიტომ კრიტიკულ არედ ავიღოთ სტანდარტული ნორმალური განაწილების მარჯვენა კუდის მხარე. ავიღოთ ზედა α კრიტიკული z_α წერტილი და კრიტერიუმის კრიტიკული არე T_n სტატისტიკისათვის იქნება

$$(z_\alpha;\ +\infty).$$

$z_{0,05}=1{,}64$ გამოთვლით მივიღებთ $\overline{x}_{16}=51{,}375$. გამოვიყენოთ იგივე სტატისტიკა რაც 31.1 მაგალითში და გამოვთვალოთ მისი მნიშვნელობა

$$Z_{16}(x)=\frac{\overline{x}_{16}-30}{\frac{\sqrt{4}}{\sqrt{16}}}=\frac{51{,}375-50}{\frac{1}{2}}=2{,}75\ .$$

რადგან $Z_{16}(x)=2{,}75\in(1{,}64;\ \infty)$, ამიტომ H_0 ჰიპოთეზას უარვყოფთ 0,05 მნიშვნელოვნობის დონით. ე. ი. წარჩინებულ სტუდენტებს დაენიშნებათ დანამატი სტიპენდიაზე 70 ლარი.

დავუბრუნდეთ ძირითადი და ალტერნატიული ჰიპოთეზების შერჩევის ფაქტს, რომელი წინადადება მივიღოთ ძირითადად და რომელი ალტერნატივად? შევნიშნოთ, რომ უფრო მართებული იქნება თუ ძირითად ჰიპოთეზად განვიხილავთ ძველ ფაქტს, რომელსაც ადგილი ჰქონდა ჰიპოთეზის დასმის მომენტამდე და შევამოწმებთ ჰიპოთეზას იმის შესახებ, გვაქვს თუ არა რაიმე საიმედოობით ამ ფაქტის უგულვებელყოფის საფუძველი. ამიტომ იგივე საკითხის გადასაჭრელად რაც 31.1 ამოცანაშია მოვიქცეთ შემდეგნაირად.

ახლა განვიხილოთ 0,01 მნიშვნელოვნობის დონით შემდეგი ჰიპო-თეზა:

$H_0:\ m=340$

$H_0:\ m=380$.

ამოხსნა: ალტერნატიული ჰიპოთეზა გვიჩვენებს ძირითად ჰიპოთეზაზე მეტ მნიშვნელობას, ამიტომ კრიტიკულ არედ განვიხილოთ მარჯვენა ინტერვალი $(z_{0,01}\ \infty)$. გამოვიყენოთ ძველი გამოთვლების შედეგები $\overline{X}_{10}=373$; $z_{0,01}=2{,}33$. კრიტერიუმის სტატისტიკა იქნება

$$Z_{10}(X)=\frac{\overline{X}_{10}-340}{\frac{\sqrt{50}}{\sqrt{10}}}\ .$$

მოცემული შერჩევისათვის მივიღებთ $Z_{10}(x)=14,76$ და რადგან $Z_{10}(x)\in U$. ე.ი. H_0 ჰიპოთეზას უარყოფთ 0,01 მნიშვნელოვნობის დონით და ვიღებთ H_1-ს. ე.ი. მართალია იმ მეცნიერთა ვარაუდი, რომლებიც თვლიან, რომ მოსავლიანობა გაიზრდება.

როგორც ვხედავთ, მივიღეთ საწინააღმდეგო მოსაზ-რება ხორბლის მოსავლიანობის შესახებ. ვიდრე გვქონდა მიღებული ზემოთ. ამიტომ უნდა შევეცადოთ ზუსტად შევარჩიოთ რომელი წინადადება იქნება ძირითადი ჰიპოთეზა. ამ მაგალითის შემთხვევაში სწორი ამოხსნის გზაა $H_0:\ m=340$ ძირითადი ჰიპოთეზის განხილვა.

ეს საკითხი არ დგას 31.2 მაგალითში გამომდინარე იქიდან, რომ იქ ხდება არა ძველი დანამატის გაუმჯობესების საკითხის გადაჭრა, არამედ დანამატის დანიშვნის სიდიდის განსაზღვრა, ამიტომ ძირითად ჰიპოთეზად უნდა მივიღოთ რექტორის (ინსტიტუციონალური დაწესებულებების უმაღლესი პირის) გადაწყვეტილება ალტერნატიულად კი სხვა მოსაზრება.

იმ შემთხვევაში როდესაც პოპულაციის σ^2 დისპერსია უცნობია მის მაგივრად კრიტერიუმის სტატისტიკაში იღებენ შერჩევით დისპერსიას S^2-ს (ან შესწორებულ S'^2) და განიხილავენ სტატისტიკას

$$T_n(x)=\frac{\overline{X}_n-a_0}{S/\sqrt{n-1}}\qquad\left(\text{ან } T_n(x)=\frac{\overline{X}_n-a_0}{S'/\sqrt{n}}\right)$$

რომელსაც აქვს სტუდენტის განაწილება $n-1$ თავისუფლების ხარისხით, $t(n-1)$.

კრიტიკული არე გამოისახება ზედა α კრიტიკული წერტილების საშუალებით მარცხენა ან მარჯვენა უსასრულო ინტერვალით იმის მიხედვით, თუ როგორი მნიშვნელობა აქვს ალტერნატივას, ნაკლები თუ მეტი (შესაბამისად) ძირითადი ჰიპოთეზის მნიშვნელობაზე. გვექნება ორი შემთხვევა

1) $H_0:\ a=a_0$

$H_0:\ a=a_1 \qquad a_1<a_0 \qquad U=(-\infty\ \ -t_{n-1,\alpha})$;

2) $H_0 : a = a_0$

$H_0 : a = a_1 \qquad a_1 > a_0 \qquad U = (t_{n-1,\alpha} \ \ \infty)$,

სადაც α წინასწარ დასახელებული მნიშვნელოვნობის დონეა.

მაგალითი 31.3. ფერმერს ბრინჯის მოსავალი მოჰყავდა საშუალოდ 350 გრ. 1 მ²-ზე. აგრონომებმა გამოიყვანეს ბრინჯის ახალი ჯიში და ელოდებიან 1 მ²-ზე 400 გრ. მოსავალს. 10 საცდელ ნაკვეთზე მათ მიიღეს შემდეგი მოსავალი (1 მ²-ზე)

360 385 412 425 352 340 418 395 410 385

ამ მონაცემების საფუძველზე შეუძლიათ თუ არა აგრონომებს 0,01 რისკის ალბათობით ამტკიცონ, რომ საშუალო მოსავლიანობა გახდა 400 გრ. მ²-ზე?

ამოხსნა. დავსვათ შემდეგი ჰიპოთეზა

$H_0 : a = 350$

$H_1 : a = 400$.

ვინაიდან $400 > 350$ გვაქვს მარჯვენა ალტერნატივა. განვიხილოთ სტატისტიკა

$$T_{10}(x) = \frac{\overline{X}_n - 350}{S' / \sqrt{10}},$$

სადაც S'^2 შესწორებული შერჩევითი დისპერსია. კრიტიკული არე იქნება

$$U = (t_{9;\,0,01} \ \ \infty).$$

სტუდენტის განაწილების ცხრილებში ვპოულობთ ზედა α კრიტიკულ წერტილს $Z_{9;\,0,01} = 2,821$.

გამოთვლებით ვღებულობთ $\overline{X}_{10} = 388,2$; $S'_{10} = 864,4$ და სტატისტიკის მნიშვნელობა იქნება

$$T(x) = 4,1$$

რომელიც ცხადია შედის კრიტიკულ არეში. მაშასადამე H_0 ჰიპოთეზას უარვყოფთ 0,01 მნიშვნელოვნობის დონით. ე.ი. აგრონომებს შეუძლიათ 0,01 რისკის ალბათობით ამტკიცონ, რომ ახალი ჯიშის მოსავლიანობა 1 მ²-ზე შეადგენს 400 გრამს.

ამოცანები

1. ტურისტული სააგენტოს ხელმძღვანელობას სურს ჯგუფი გაუშვას სამოგზაუროდ აშშ-ს რამოდენიმე რეგიონში. მათ შეისწავლეს ორადგილიანი ნომრების ფასები ამ რეგიონების რამოდენიმე სასტუმროში. მონაცემები მოყვანილია ცხრილის სახით (1 დღე-ღამის ღირებულება 2013 წ. 25 დეკემბრის მონაცემებით)

Best Western Payson lnn	$ 75
Days lnn and Suites of Payson	$ 72
Majestic Mountain lnn	$ 78
Comfort lnn Payson	$ 88
Dest Western Plus Barsana Hotel &	$ 88
Quality lnn Soutwest	$ 72
Comfort lnn	$ 89
American Best Value lnn	$ 78

მენეჯერი ფიქრობს, რომ ერთი ნომრის სადღეღამისო ღირებულებაზე ფირმამ უნდა გაითვალისწინოს საშუალოდ $ 85, ხოლო ფირმის დირექტორს აქვს $80 – მოსაზრება, შეამოწმოთ $\alpha = 0{,}01$ მნიშვნელოვნობის დონით ჰიპოთეზა : (საშუალო ღირებულებაა a).

$H_0 : a = 80$

$H_1 : a = 85$.

2. პერიარის ნაკრძალში (ჩრდ. ინდოეთი. კეპალი) შეამოწმეს 10 ზრდასრული ინდური სპილოს წონა. მიღებული მონაცემებით (გაზომილია კგ-ებში) 0,05 რისკის ალბათობით შეუძლიათ თუ არა იმის მტკიცება, რომ საშუალო წონაა 5400 კგ. ალტერნატივის წინააღმდეგ, რომ საშუალო წონაა 5100 კგ.

5215	5416	5320	5420	5642
5338	5327	5378	5520	5356

შეამოწმეთ $\alpha = 0{,}01$ მნიშვნელოვნობის დონით ჰიპოთეზა ამ ნაკრძალში ინდური სპილოების პოპულაციის საშუალოს შესახებ

$H_0 : a = 5400$

$H_1 : a_1 = 5100$.

3. ცხრილში მოყვანილია ტაილანდის ქალაქ პატონგის (Patong) საუკეთესო რესტორნებში ერთი პერსონის სადილისათვის საჭირო მინიმალური თანხები (ლარებში)

Baan Rim Pa	52
La Piazzetta di capri Patong	45
Kantok Restaurant of Burasari Resort	43
Acqua Restaurant	43
Da Maurizio Ri storante	43
Old Siam Restaurant	39
Brush Restaurant & Luunge	36
Starfish Bar and Restaurant	34
Vhite Box Restaurant	34

შეამოწმეთ 0,01 მნიშვნელოვნობის დონით ჰიპოთეზა საშუალო ფასის შესახებ

$H_0 : a = 43$

$H_1 : a = 36$.

4. სუპერმარკეტის მენეჯერმა შეცვალა ადგილზე დაფქვილი ყავის შეფუთვის სახე. თუ ადრე დღეში საშუალოდ ყიდდა 18 კგ ყავას, ახალი გაყიდვის მონაცემების მიხედვით ვარაუდობს, რომ დღეში გაიყიდება 20 კგ-ს ფარგლებში. აქვს თუ არა მონაცემების მიხედვით 0,01 რისკის ალბათობით ამ ფაქტის მტკიცების საფუძველი? (**მითითება**: შეამოწმეთ ჰიპოთეზა $H_0 : a = 18$

$H_1 : a = 20$)

წონები მოცემულია კილოგრამების სიზუსტით
16,2; 19,5; 20,6; 21,4; 18,6; 17,2; 19,5; 21,7; 20,3; 17,7.

პასუხები:

1) $H_0 : a = \$80$ -ის უარყოფის საფუძველი არა გვაქვს: 2) შეუძლიათ. 3) $H_0 : a = \$43$-ის უარყოფის საფუძველი არა გვაქვს. 4) არა აქვს, ვინაიდან H_0 ჰიპოთეზის უგულვებელყოფის საფუძველი არა გვაქვს.

§ 32. რთული ჰიპოთეზის შემოწმება ნორმალური პოპულაციის საშუალოს შესახებ

დავუშვათ, ცნობილია რომ პოპულაციას ნორმალური განაწილება აქვს და მისი **დისპერსია** σ^2 **ცნობილია**. განვიხილოთ ჰიპოთეზები პოპულაციის საშუალოს შესახებ. ვიგულისხმოთ, რომ გვაქვს n მოცულობის შემთხვევითი შერჩევა $x_1, x_2, \ldots, x_n$. ჰიპოთეზის მნიშვნელოვნობის დონე იყოს α. განვიხილოთ სამი შემთხვევა.

1) $H_0 : a = a_0$;

$H_1 : a < a_0$.

ამ შემთხვევაში ძირითადი H_0 ჰიპოთეზა მარტივია, ხოლო მარცხენა ალტერნატივა H_1 კი რთული. კრიტერიუმის სტატისტიკად განვიხილოთ სტატისტიკა

$$Z_n(X) = \frac{\overline{X}_n - a_0}{\dfrac{\sigma}{\sqrt{n}}}. \qquad (32.1)$$

გავიხსენოთ, რომ წინა პარაგრაფში მარტივი ჰიპოთეზების შემოწმების დროს აგებული კრიტიკული არე არ იყო დამოკიდებული $H_0 : a = a_1$ ალტერნატიული მარტივი ჰიპოთეზის ზუსტ მნიშვნელობაზე. ითვალისწინებდა მხოლოდ იმას, ეს a_1 მნიშვნელობა ძირითადი ჰიპოთეზის a_0 მნიშვ-ნელობაზე მეტი იყო თუ ნაკლები, რის მიხედვითაც კრიტიკულ არედ ვიღებდით ნორმალური განაწილების მარჯვენა ან მარცხენა კუდებს შესაბამისად.

კრიტიკულ არედ $Z_n(X)$ სტატისტიკისათვის ავიღოთ სტანდარტული ნორმალური განაწილების მარცხენა α ზომის არე ანუ

$$(-\infty;\ -z_\alpha). \qquad (32.2)$$

უსასრულო ინტერვალი. თუ $x_1, x_2, \ldots, x_n$ რეალიზაციისათვის გამოთვლილი $Z_n(X)$ სტატისტიკის რიცხვითი მნიშვნელობა დააკმაყოფილებს

$$Z_n(X) < -z_\alpha$$

უტოლობას მაშინ H_0 ჰიპოთეზას უარვყოფთ α მნიშვნელოვნობის დონით, წინააღმდეგ შემთხვევაში მისი უარყოფის საფუძველი არა გვაქვს.

მაგალითი 32.1. განვიხილოთ მაგალითი 10.1. მოცემული გვაქვს $n=10$ დღის განმავლობაში ბენზინგასამართი სადგურის შემოსავლები (ასეულ ლარებში)

6; 12; 5; 8; 9; 5; 9; 8; 9; 9

ჩავთვალოთ, რომ შემოსავლების დისპერსიაა 150 ლარი და შევამოწმოთ 0,01 მნიშვნელოვნობის დონით H_0 ჰიპოთეზა იმის შესახებ, რომ ყოველდღიური საშუალო შემოსავალია 900 ლარი. ალტენატივის წინააღმდეგ, რომ შემოსავალი 900 ლარზე ნაკლებია.

ამოხსნა.

$H_0:\ a=9$

$H_0:\ a<9$.

ალტერნატიული ჰიპოთეზა მარცხენა მხარისაა, ამიტომ კრიტიკულ არედ ავიღოთ (32.2) უსასრულო ინტერვალი

$$(\infty;\ -z_{0,01}).$$

ცხრილებიდან მოიძებნება, რომ $z_{0,01}=2{,}33$. ვიცით $n=10$; $\sigma^2=1{,}5$. გამოთვლებით მივიღებთ $\bar{x}_{10}=8$

$$Z_{10}(x)=\frac{8-9}{\frac{\sqrt{1{,}5}}{\sqrt{10}}}=-2{,}58.$$

რადგან $Z_{10}(x)=-2{,}58<-2{,}33$, ამიტომ $Z_{10}(x)$ სტატისტიკა ხვდება კრიტიკულ არეში. ე.ი. H_0 ჰიპოთეზას უარვყოფთ 0,01 მნიშვნელოვნობის დონით და ჩავთვლით, რომ ყოველდღიური შემოსავალი 900 ლარზე ნაკლებია.

იგივე გზით შეგვიძლია შევამოწმოთ ჰიპოთეზის სამართლიანობა იმ შემთხვევაშიც როდესაც ძირითადი და ალტერნატიული ჰიპოთეზა ორივე რთულია და გვაქვს მარცხენა ალტერნატივა

$H_0:\ a\ge a_0$

$H_0:\ a<a_0$.

$Z_n(x)$ სტატისტიკისათვის α მნიშვნელოვნობის დონის კრიტიკული არე იქნება ისევ (32.2) მარცხნიდან უსასრულო ინტერვალი $(-\infty;\ -z_\alpha)$.

მაგალითი 32.2. საჩრდილობელი ქოლგების მწარმოებელი ფირმის მენეჯერს აინტერესებს ერთ-ერთ ზღვისპირა ქალაქში ზაფხულის სეზონზე მის პროდუქციაზე მოთხოვნა. მან შეისწავლა 15 წლის განმავლობაში ზაფხულის სეზონზე გაყიდული ქოლგების რაოდენობები და მოიპოვა მონაცემები:

38; 56; 45; 55; 52; 56; 62; 60; 54; 48; 51; 47; 54; 53; 52.

შეუძლია თუ არა მას 0,99 ალბათობით ივარაუდოს, რომ სეზონზე მოთხოვნა მის პროდუქციაზე აღემატება 50- ს? თუ ცნობილია, რომ გაყიდული პროდუქციის დისპერსიაა 9.

ამოხსნა. პროდუქციაზე მოთხოვნის რაოდენობა ნორმალურად განაწილებული შემთხვევითი სიდიდეა მრავალი სხვადასხვა დამოუკიდებელი ფაქტორების გამო, რომლებიც ამ მოთხოვნას განაპირობებენ. განვიხილოთ ჰიპოთეზა

$$H_0:\ a \geq 50$$

$$H_0:\ a < 50$$

$a_0 = 50$; $n = 15$; $\alpha = 0{,}01$; $\sigma = \sqrt{9} = 3$, გამოთვლებით მივიღებთ $\overline{x}_{15} = 52{,}2$. ცხრილში ვპოულობთ $z_\alpha = z_{0{,}01} = 2{,}33$, ე.ი. $-z_\alpha = -2{,}33$

$$Z_{15}(x) = \frac{\overline{x}_{15} - 50}{\dfrac{3}{\sqrt{15}}} = \frac{52{,}2 - 50}{\dfrac{3}{\sqrt{15}}} = 2{,}8\,.$$

ვინაიდან $2{,}8 \nless -2{,}33$, ამიტომ H_0 ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს. მაშასადამე მენეჯერს შეუძლია ივარაუდოს, რომ მის პროდუქციაზე მოთხოვნა სეზონის განმავლობაში არაა ნაკლები 50-ზე.

განვიხილოთ მარჯვენა რთული ალტერნატივის შემთხვევა

2) $H_0:\ a = a_0$

$H_1:\ a_1 > a_0$

H_1 მარჯვენა ალტერნატივა რთული ჰიპოთეზაა. კვლავ ვიხილავთ

$$Z_n(X) = \frac{\overline{X}_n - a_0}{\frac{\sigma}{\sqrt{n}}}$$

სტატისტიკას და ალტერნატივიდან გამომდინარე კრიტიკული არე იყოს სტანდარტული ნორმალური განაწილების მარჯვენა α ზომის არე, ანუ ზედა α კრიტიკული არე.

$$(z_\alpha \ +\infty). \tag{32.3}$$

მაშასადამე, თუ შესრულდა

$$Z_n(x) > z_\alpha \tag{32.4}$$

მაშინ H_0 ჰიპოთეზას უკუვაგდებთ α მნიშვნელოვნობის დონით.

მაგალითი 32.3. ლუდის მწარმოებელი ფირმის მიერ კვების ობიექტზე 10 დღის განმავლობაში ყოველდღიურად გაყიდული 0,5 ლ. მოცულობის ბოთლების რაოდენობებია

210; 176; 192; 200; 169; 197; 208; 213; 178; 188

შეუძლია თუ არა ფირმის მენეჯერს, 0,05 მნიშვნელოვნობის დონით, სამართლიანად ჩათვალოს ჰიპოთეზა, რომ საშუალო დღიური გაყიდვების რაოდენობაა 180 ბოთლი, ალტერნატიული ჰიპოთეზის წინააღმდეგ - „180 ბოთლზე მეტი?“

ჩავთვალოთ, რომ საშუალო კვადრატული გადახრაა 15 ბოთლი.

ამოხსნა. $n = 10$, $\alpha = 0{,}05$, $\sigma = 15$

$H_0: a = 180$

$H_1: a > 180$

გვაქვს მარჯვენა ალტერნატივა, ამიტომ კრიტიკული არეა (32.3) ინტერვალი

$$(z_\alpha; +\infty)$$

$z_{0,05} = 1{,}64$. გამოვიყენოთ (32.1) სტატისტიკა

$$Z_n(X)=\frac{\overline{X}_n-a_0}{\frac{\sigma}{\sqrt{n}}}.$$

გამოთვლებით მივიღებთ $\overline{x}_{10}=193{,}1$

$$Z_{10}(x)=\frac{\overline{x}_{10}-180}{\frac{15}{\sqrt{10}}}=2{,}761.$$

ვინაიდან $Z_{10}(x)=2{,}761>1{,}64=z_{0,05}$, ამიტომ $Z_n(x)$ მოხვდა კრიტიკულ არეში, რის გამოც H_0 ჰიპოთეზას უარვყოფთ 0,05 მნიშვნელოვნობის დონით და ჩავთვლით, რომ ყოველდღიურად გაყიდული ბოთლების რაოდენობა მეტია 180-ზე $\alpha=0{,}05$ რისკის ალბათობით.

ანალოგიურად შეიძლება შევამოწმოთ რთული ჰიპოთეზა მარჯვენა რთული ალტერნატივით

$H_0:\ a\le a_0$

$H_1:\ a>a_0$

კრიტიკული არე იქნება ისევ (32.3)

$$(z_\alpha;\ +\infty)$$

უსასრულო შუალედი. თუ $Z_n(x)$ სტატისტიკის მნიშვნელობა $x=(x_1,x_2,\ldots,x_n)$ წერტილში $Z_n(x)$ აკმაყოფილებს პირობას

$$Z_n(x)>z_\alpha$$

მაშინ H_0 ჰიპოთეზას უარვყოფთ α მნიშვნელოვნობის დონით. წინააღმდეგ შემთხვევაში მისი უარყოფის საფუძველი არა გვაქვს.

მაგალითი 32.4. სადაზღვევო კომპანიას უძრავი ქონების დაზღვევის სექტორში თითო უბედურ შემთხვევაზე კლიენტისათვის 5000 ევროს გადახდა უწევს. უკანასკნელი 10 წლის განმავლობაში უბედურ შემთხვევათა სტატისტიკა იმ კლიენტებისათვის, რომელთაც უძრავი ქონება ამ კომპანიაში დააზღვიეს ასეთია:

18; 15; 20; 25; 17; 24; 23; 19; 26; 22.

შეუძლია თუ არა კომპანიის მენეჯერს 0,95 ალბათობით ელოდოს, რომ კომპანიის მიერ წლის განმავლობაში გასაცემი თანხა არ აღემატება 100.000 ევროს? თუ ცნობილია, რომ უბედურ შემთხვევათა დისპერსიაა 8.

ამოხსნა. ცხადია 100.000 ევროს გაცემა მოხდება, თუ წლის განმავლობაში მოხდა 20 უბედური შემთხვევა, ამიტომ ამოცანა შეგვიძლია ასე ჩამოვაყალიბოთ.

შეუძლია თუ არა კომპანიის მენეჯერს 0,95 ალბათობით ელოდოს, რომ კომპანიაში დაზღვეული უძრავი ქონების უბედურ შემთხვევათა რაოდენობა არ გადააჭარბებს 20-ს?

ვიგულისხმოთ, რომ უბედური შემთხვევები უამრავი სხვადასხვა დამოუკიდებელი მიზეზებითაა გამოწვეული და ამიტომ ნორმალურად განაწილებულად შეგვიძლია მივიჩნიოთ. მაშასადამე გვაქვს ნორმალურად განაწილებული პოპულაცია. ცნობილი $\sigma^2 = 8$ დისპერსიით და უნდა შევამოწმოთ $\alpha = 1 - 0{,}95 = 0{,}05$ მნიშვნელოვნობის დონით შემდეგი ჰიპოთეზა

$$H_0:\ a \le 20$$

$$H_1:\ a > 20$$

$n = 10$; $\alpha = 0{,}05$; $\sigma_0 = \sqrt{8}$. მოვძებნოთ სტანდარტული ნორმალური განაწილების ზედა α კრიტიკული წერტილი. $z_{0,05} = 1{,}64$. გამოთვლებით მივიღებთ $\overline{x}_{10} = 20{,}9$

$$Z_{10}(x) = \frac{\overline{x}_{10} - 20}{\frac{\sqrt{8}}{\sqrt{10}}} = 1{,}006\ .$$

კრიტიკული არე ავიღოთ ($z_{0,05}$; $+\infty$) ანუ (1,64; +∞) ინტერვალი

$$Z_{10}(x) \not> 1{,}64\ .$$

მაშასადამე 0,05 მნიშვნელოვნობის დონით ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს. ე.ი. მენეჯერს შეუძლია 0,95 ალბათობით ივარაუდოს, რომ უბედურ შემთხვევათა რიცხვი დაზღვეული უძრავი ქონებისათვის არ აღემატება 20-ს და შესაბამისად კომპანიის მიერ წლის განმავლობაში გასაცემი სადაზღვევო თანხა არ აღემატება 100000 ევროს.

განვიხილოთ ორმხრივი რთული ალტერნატივის შემთხვევა

3) $H_0: a = a_0$

$H_1: a \neq a_0$

H_1 ორმხრივი ალტერნატიული რთული ჰიპოთეზაა. ამ შემთხვევაში H_0 ჰიპოთეზას უარვყოფთ და H_1 ალტერნა-ტივას მივიღებთ იმ $x_1, x_2, \ldots, x_n$ შერჩევისათვის რომლისთვი-საც $\overline{X}_n$ სტატისტიკის მნიშვნელობა a_0 რიცხვს მოსცილ-დება (ნებისმიერ მხარეს) გარკვეულ მანძილზე მეტად. ეს მანძილი შევარჩიოთ შემდეგნაირად: განვიხილოთ იგივე $Z_n(x)$ სტატისტიკა, რომელსაც ზემოთ ვიყენებდით და მოცემული α მნიშვნელოვნობის დონისათვის ავიღოთ სტანდარტული ნორმალური განაწილების ზედა $\dfrac{\alpha}{2}$ და ქვედა $\dfrac{\alpha}{2}$ კრიტიკული წერტილები, შესაბამისად $z_{\alpha/2}$ და $-z_{\alpha/2}$. H_0 ჰიპოთეზას უარვყოფთ თუ

$$Z_n(x) \in (-\infty \;\; -z_{\alpha/2}) \cup (z_{\alpha/2} \;\; \infty)$$

ანუ

$$|Z_n(x)| > z_{\alpha/2}. \tag{32.5}$$

მაგალითი 32.5. დავუბრუნდეთ 29.1 მაგალითს. შევამოწ-მოთ ჰიპოთეზა სასურსათო მაღაზიის საშუალო დღიური ნავაჭრის შესახებ. შევამოწმოთ $\alpha = 0{,}1$ მნიშვნელოვნობის დონით მარტივი ჰიპოთეზა ორმხრივი ალტერნატივით

$H_0: a = 205$

$H_1: a \neq 205$

ამოხსნა. 29.1. მაგალითიდან ცნობილია $\sigma = 20$; $\bar{x} = 200{,}25$; $n = 16$; $\alpha = 0{,}1$ და რადგან ალტერნატივა ორმხრივია მოვძებნოთ სტანდარტული ნორმალური განაწილების ზედა $\dfrac{\alpha}{2} = 0{,}05$ კრიტიკული წერტილი. $z_{0,05} = 1{,}64$

$$Z_n(X)=\frac{\bar{X}_n-a}{\frac{\sigma}{\sqrt{n}}}=\frac{200,25-205}{\frac{20}{\sqrt{15}}}=-0,92\,.$$

ვინაიდან მივიღეთ, რომ

$$\left|Z_n(x)\right|\not> z_{0,05}\,,$$

ამიტომ 0,1 მნიშვნელოვნობის დონით H_0 ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს. ე.ი. მაღაზიის მენეჯერს 0,1 მნიშვნელოვნობის დონით შეუძლია ჩათვალოს, რომ მაღაზიის დღიური ნავაჭრი 205 ლარის ტოლია.

ახლა განვიხილოთ შემთხვევა, როდესაც პოპულაციის დისპერსია უცნობია. ამ დროს პოპულაციის საშუალოს შესახებ ჰიპოთეზის შემოწმებისას კრიტერიუმის (32.1) სტატისტიკაში σ^2 დისპერსია შევცვალოთ S^2 შერჩევითი დისპერსიით და განვიხილოთ

$$Z_n(X)=T_n=\frac{\bar{X}_n-a_0}{\frac{S'}{\sqrt{n}}}=\frac{\bar{x}_n-a_0}{\frac{S}{\sqrt{n-1}}} \qquad (32.6)$$

სტატისტიკა, რომელსაც აქვს სტიუდენტის განაწილება $n-1$ თავისუფლების ხარისხით $t(n-1)$. მოვიყვანოთ U კრიტიკული არე განხილული სამივე შემთხვევისათვის. დავაფიქსიროთ ჰიპოთეზის მნიშვნელოვნობის α დონე.

1) $H_0: a=a_0$

$H_1: a<a_0 \qquad U=(-\infty;\ -t_{n-1,\alpha})$ (32.7)

2) $H_0: a=a_0$

$H_1: a>a_0 \qquad U=(t_{n-1,\alpha};\ +\infty)$ (32.8)

3) $H_0: a=a_0$

$H_1: a\neq a_0 \qquad U=(-\infty;\ -t_{n-1,\alpha/2})\cup(t_{n-1,\alpha/2};\ +\infty)\,.$ (32.9)

მაგალითი 32.6. დავუბრუნდეთ 32.1 მაგალითს. მოცემული გვაქვს 10 დღის განმავლობაში ბენზინგასამართი სადგურის შემოსავლები (ასეული ლარებში)

6; 12; 5; 8; 9; 5; 9; 8; 9; 9

ბენზინგასამართი სადგურის შემოსავლების დღიური რაოდენობის დისპერსია ჩავთვალოთ უცნობად.

შევამოწმოთ ჰიპოთეზა იმის შესახებ, რომ საშუალო დღიური შემოსავალია 900 ლარი. ალტერნატიული ჰიპოთეზის წინააღმდეგ, რომ შემოსავალი ნაკლებია 900 ლარზე. მნიშვნელოვნობის დონე α იყოს 0,01.

ამოხსნა.

$H_0: a=9$

$H_1: a<9$

დისპერსია უცნობია $n=10$. $\alpha=0{,}01$. გვაქვს მარცხენა ალტერნატიული ჰიპოთეზა, ამიტომ კრიტიკული არეა (32.7) მარცხნიდან უსასრულო ინტერვალი

$$(-\infty;\ -t_{n-1,\alpha}).$$

$t_{n-1,\alpha}=t_{9;0,01}=2{,}821$. გამოვიყენოთ (32.6) სტატისტიკა. ვღებულობთ

$$\bar{x}_{10}=8;\quad S_{10}^2=4{,}2;$$

$$T_{10}(x)=\frac{8-9}{\dfrac{\sqrt{4{,}2}}{\sqrt{9}}}\approx-1{,}464$$

ვინაიდან $T_{10}(x)=-1{,}464\not<-2{,}821=t_{9;0,01}$, ამიტომ კრიტერიუმის სტატისტიკა არ შედის კრიტიკულ არეში. ე.ი. H_0 ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს.

შეგვიძლია ჩავთვალოთ, რომ ბენზინგასამართი სადგურის დღიური შემოსავალია 900 ლარი.

როგორც ვხედავთ, ამოცანის პასუხი განსხვავდება იმ შემთხვევაში პასუხისაგან, როდესაც დისპერსია ცნობილი იყო.

ანალოგიურად შემოწმდება, (32.6) სტატისტიკის საშუალებით რთული ჰიპოთეზები საშუალოს შესახებ, როდესაც დისპერსია უცნობია.

მოვიყვანოთ მათი შესაბამისი კრიტიკული არეები:

ა) $H_0: a\ge a_0$

$H_1: a<a_0 \qquad U=(-\infty;\ -t_{n-1,\alpha})$ (32.10)

ბ) $H_0: a \le a_0$

$H_1: a > a_0 \qquad U = (t_{n-1,\alpha};\ +\infty)$. (32.11)

მაგალითი 32.7 10 დღის განმავლობაში გაყიდული ფოტოაპარატების რაოდენობებია

11; 19; 18; 16; 20; 12; 19; 14; 13; 19

ვიგულისხმოთ, რომ ფოტოაპარატებზე მოთხოვნა ნორმალურად განაწილებული შემთხვევითი სიდიდეა და 0,005 მნიშვნელოვნობის დონით შევამოწმოთ ჰიპოთეზა რომ დღეში საშუალოდ გაყიდული ფოტოაპარატების რაოდენობა არაა ნაკლები 16-ზე ალტერნატივით, რომ 16-ზე ნაკლებია

$H_0: a \ge 16$

$H_1: a < 16$

ამოხსნა. რადგან დისპერსია უცნობია, ვიყენებთ (32.6) სტატისტიკას. $n = 10$ გამოთვლებით ვღებულობთ $\bar{x}_{10} = 16{,}1$; $S_{10}^2 = 10{,}09$.

კრიტიკული არე იქნება (32.10) ინტერვალი. $t_{n-1,\alpha} = t_{9;0,005} = 3{,}250$.

გამოვთვალოთ სტატისტიკის მნიშვნელობა მოცემული შერჩევისათვის

$$T_{10}(x) = \frac{\bar{x}_{10} - 16}{\frac{S_{10}}{\sqrt{9}}} = 0{,}094 .$$

ვინაიდან $T_{10}(x) = 0{,}094 \not< -3{,}25 = -t_{9;0,005}$, ამიტომ სტატისტიკის მნიშვნელობა მოცემული შერჩევისათვის (რეალიზაციისათვის) არ შედის კრიტიკულ არეში. მაშასადამე H_0 ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს. ამიტომ 0,05 რისკის ალბათობით შეგვიძლია ჩავთვალოთ, რომ საშუალოდ დღეში გაყიდული ფოტოაპარატების რაოდენობა არაა ნაკლები 16-ზე.

ამოცანები

1. ცნობილია 15 საცდელ ნაკვეთზე მოყვანილი ხორბლის ერთ-ერთი ჯიშის (ЯЛЬЯНС – производство ин-т растениеводства им. Я.Юриева. Укр.) მოსავლიანობის მაჩვენებლები 1 $მ^2$-ზე (კგ-ებში).

1,03; 0,94; 0,98; 1,07; 1,11; 1,2; 0,93; 0,97; 1,05; 1,1; 0,99; 1,07; 1,15; 1,1; 1,02

ჩათვალეთ, რომ დისპერსია 1 $მ^2$-ზე მოყვანილი ხორბლის ამ ჯიშისათვის 0,2-ის ტოლია და შეამოწმეთ 0,001 მნიშვნელოვნობის დონით ჰიპოთეზა იმის შესახებ, რომ საშუალო მოსავლიანობა 1 კვ.მ-ზე არაა ნაკლები 1,05 კგ-ზე ალტერნატივის წინააღმდეგ, რომ საშუალო ნაკლებია 1,05 კგ-ზე.

2. ცხრილში მოყვანილია კამბოჯის ხუთვარსკვლავიან რესტორნებში ერთი ადამიანის სადილისათვის საჭირო მინიმალური ფასები (ლარებში) (2013 წ. დეკემბერის მონაცემებით).

Topaz	69
MIRO Restaurant & Lounge	43
Dino in the Dark	31
Sonoma Oyster	26
Kanii Japanese Pestaurant	21
Destination Dining	103
Circle	62
PHUMKREUS KHMER – Fine Cambodjian Cuisine	43
Cuisine Wat Damnak	38
Heritage Suites Restaurant	34
Takezono Restaurant	26
Abacus Restaurant, Garden & Bar	22

შეამოწმეთ 0,005 მნიშვნელოვნობის დონით ჰიპოთეზა საშუალო ფასის შესახებ

$H_0 : a \le 40$

$H_1 : a > 40$.

3. კომპანია Midgulf International Ltd-ს წარმომადგენლები 2007 წლის იანვრიდან 2013 წლის მარტის ფასების კვოტირებამდე ჩიკაგოს საფონდო ბირჟაზე „CME“ (Chicago Mercantile Exchange) ახორციელებდნენ ხორბლის შესყიდვას 1 ტონაზე შემდეგი ფასებით (აშშ დოლარებში)

293;	291,2;	297,1;	295,11;	298,05;	301,1;	303,05;	307,1;
295,03;	297;	299,1;	302,1;	302,12;	307,1;	309,2;	311,1;
309,4;	310,1.						

ჩათვალეთ, რომ ფასის დისპერსია 100-ის ტოლია და შეამოწმეთ 0,01 მნიშვნელობის დონით ჰიპოთეზა იმის შესახებ, რომ საშუალო ფასი ამ წლების განმავლობაში ხორბლის ერთ ტონაზე იყო $294 ალტერნატივის წინააღმდეგ, რომ ფასი არ იყო $294.

4. ცხოველთა დაცვის მსოფლიო ორგანიზაციის (WSPA-World Society for the Protection of Animals) ბალტიის ზღვის აუზის ფილიალში გამოიკვლიეს მოგზაური ალბატროსის (Diomedea exulans) 20 ზრდასრული ერთეულის ფრთების გაშლის სიდიდე. ჩათვალეთ, რომ ფრთების გაშლილობის დისპერსია 0,6 მეტრის ტოლია და მიღებული მონაცემებისათვის 0,005 მნიშვნელოვნობის დონით შეამოწმეთ შემდეგი ჰიპოთეზა

$H_0 : l = 2,9$

$H_0 : l > 2,9$.

2,8;	2,7;	3;	3,12;	2,85;	2,93;	2,9;	3,12;	3,15;	3,2;
3,1;	3,05;	2,9;	2,72;	2,75;	3,10;	3,13;	2,9;	2,82;	2,78

5. ბელორუსიის უნივერსალურ საფონდო ბირჟაზე (ОАО „БВФБ” – Открытое Акционерное Общество „Белорусская Валютно фондовая Биржа”) 2013 წლის იანვრის თვის გარიგებებში დაფიქსირდა ერთი კილოგრამი ღორის ხორცის შემდეგი ფასები (რუსულ რუბლში)

102; 100; 104; 98; 97; 100; 101; 99; 96;
105; 100; 97; 100; 101; 98; 99.

შეიძლება თუ არა ამ მონაცემებზე დაყრდნობით 0,001 რისკის ალბათობით ვამტკიცოთ, რომ დროის ამ შუალედში ღორის ხორცის შესყიდვის საშუალო ფასი იყო 100 რუბლი იმ ალტერნატივის წინააღმდეგ, რომ ნაკლები იყო 100 რუბლზე.

პასუხები

1) 0,001 რისკის ალბათობით ძირითადი ჰიპოთეზის უგულვებელყოფის საფუძველი არ არსებობს. შეიძლება ვამტკიცოთ, რომ 1 $მ^2$-ზე ხორბლის ამ ჯიშის მოსავალი არაა ნაკლები 1,05-ზე.

2) 0,005 რისკის ალბათობით ძირითადი ჰიპოთეზის უგულვებელყოფის საფუძველი არ არსებობს. ვთვლით, რომ რესტორანში ერთი ადამიანის სადილის საშუალო ფასი არ აღემატება 40 ლარს.

3) H_0 ჰიპოთეზას უარვყოფთ და 0,01 რისკის ალბათობით ვთვლით, რომ 1 ტონა ხორბლის შესყიდვის ფასი არ არის 294 აშშ დოლარის ტოლი.

4) 0,001 რისკის ალბათობით ძირითადი ჰიპოთეზის უარყოფის საფუძველი არ არსებობს. ვთვლით, რომ ალბატროსის ფრთების გაშლილობაა 2,9 მ.

5) 0,001 რისკის ალბათობით ჰიპოთეზის უარყოფის საფუძველი არ არსებობს. შეიძლება ჩავთვალოთ, რომ 1 კგ ხორცის ფასი იყო 100 რუსული რუბლი ალტერნატივის წინააღმდეგ, რომ ფასი ნაკლებია 100 რუსულ რუბლზე.

დანართი

ცხრილი 1. თანაბრად განაწილებულინ შემთხვევითი რიცხვები

10097	32533	76520	13586	34673	54876	80959	09117	39292	74945
37542	04805	64894	74296	24805	24037	20636	10402	00822	91665
08422	68953	19645	09303	23209	02560	15953	34764	35080	33606
99019	02529	09376	70715	38311	31165	88676	74397	04436	27659
12807	99970	80157	36147	64032	36653	98951	16877	12171	76833
66065	74717	34072	76850	36697	36170	65813	39885	11199	29170
31060	10805	45571	82406	35303	42614	86799	07439	23403	09732
85269	77602	02051	65692	68665	74818	73053	85247	18623	88579
63573	32135	05325	47048	90553	57548	28468	28709	83491	25624
73796	45753	03529	64778	35808	34282	60935	20344	35273	88435
98520	17767	14905	68607	22109	40558	60970	93433	50500	73998
11805	05431	39808	27732	50725	68248	29405	24201	52775	67851
83452	99634	06288	98083	13746	70078	18475	40610	68711	77817
88685	40200	86507	58401	36766	67951	90364	76493	29609	11062
99594	67348	87517	64969	91826	08928	93785	61368	23478	34113
65481	17674	17468	50950	58047	76974	73039	57186	40218	16544
80124	35635	17727	08015	45318	22374	21115	78253	14385	53763
74350	99817	77402	77214	43236	00210	45521	64237	96286	02655
69916	26803	66252	29148	36936	87203	76621	13990	94400	56418
09893	20505	14225	68514	46427	56788	96297	78822	54382	14598
91499	14523	68479	27686	46162	83554	94750	89923	37089	20048
80336	94598	26940	36858	70297	34135	53140	33340	42050	82341
44104	81949	85157	47954	32979	26575	57600	40881	22222	06413
12550	73742	11100	02040	12860	74697	96644	89439	28707	25815
63606	49329	16505	34484	40219	52563	43651	77082	07207	31790
61196	90446	26457	47774	51924	33729	65394	59593	42582	60527
15474	45266	95270	79953	59367	83848	82396	10118	33211	59466
94557	28573	67897	54387	54622	44431	91190	42592	92927	45973
42481	16213	97344	08721	16868	48767	03071	12059	25701	46670
23523	78317	73208	89837	68935	91416	26252	29663	05522	82562

ცხრილი 2

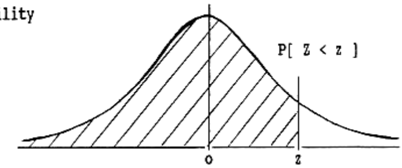
სტანდარტული ნორმალური განაწილების ფუნქციის მნიშვნელობების ცხრილი

STANDARD STATISTICAL TABLES

1. Areas under the Normal Distribution

The table gives the cumulative probability up to the standardised normal value z i.e.

$$P[\ Z < z\ ] = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}Z^2)\ dZ$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5159	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7854
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8804	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9773	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9865	0.9868	0.9871	0.9874	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9924	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9980	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986

z	3.00	3.10	3.20	3.30	3.40	3.50	3.60	3.70	3.80	3.90
P	0.9986	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000

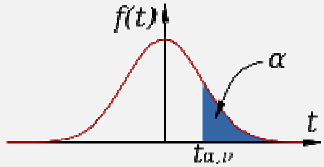
სტანდარტული ნორმალური განაწილების ზედა α კრიტიკული (Z_α) წერტილების მნიშვნელობები

α	0.1	0.05	0.025	0.0125	0.01	0.005	0.0025	0.001
Z_α	1.28	1.64	1.96	2.24	2.33	2.57	2.81	3.08

ცხრილი 3

ცხრილი 3. სტიუდენტის $t_{n,\alpha}$ (ზედა α კრიტიკული) წერტილების ცხრილი

Table of Critical Values, $t_{n,\alpha}$, in a Student T-Distribution with n degrees of freedom and a confidence limit $p$where $\alpha=1-p$.



n	Confidence Limits (top) and *α* (bottom) for a One-Tailed Test.												
	60%	**75%**	**80%**	**85%**	**90%**	**95%**	**97.5%**	**98%**	**99%**	**99.5%**	**99.75%**	**99.9%**	**99.95%**
	0.4	**0.25**	**0.2**	**0.15**	**0.1**	**0.05**	**0.025**	**0.02**	**0.01**	**0.005**	**0.0025**	**0.001**	**0.0005**
1	0.32492	1.00000	1.37638	1.96261	3.07768	6.31375	12.70620	15.89454	31.82052	63.65674	127.32134	318.30884	636.61925
2	0.28868	0.81650	1.06066	1.38621	1.88562	2.91999	4.30265	4.84873	6.96456	9.92484	14.08905	22.32712	31.59905
3	0.27667	0.76489	0.97847	1.24978	1.63774	2.35336	3.18245	3.48191	4.54070	5.84091	7.45332	10.21453	12.92398
4	0.27072	0.74070	0.94096	1.18957	1.53321	2.13185	2.77645	2.99853	3.74695	4.60409	5.59757	7.17318	8.61030
5	0.26718	0.72669	0.91954	1.15577	1.47588	2.01505	2.57058	2.75651	3.36493	4.03214	4.77334	5.89343	6.86883
6	0.26483	0.71756	0.90570	1.13416	1.43976	1.94318	2.44691	2.61224	3.14267	3.70743	4.31683	5.20763	5.95882
7	0.26317	0.71114	0.89603	1.11916	1.41492	1.89458	2.36462	2.51675	2.99795	3.49948	4.02934	4.78529	5.40788
8	0.26192	0.70639	0.88889	1.10815	1.39682	1.85955	2.30600	2.44898	2.89646	3.35539	3.83252	4.50079	5.04131
9	0.26096	0.70272	0.88340	1.09972	1.38303	1.83311	2.26216	2.39844	2.82144	3.24984	3.68966	4.29681	4.78091
10	0.26018	0.69981	0.87906	1.09306	1.37218	1.81246	2.22814	2.35931	2.76377	3.16927	3.58141	4.14370	4.58689
11	0.25956	0.69745	0.87553	1.08767	1.36343	1.79588	2.20099	2.32814	2.71808	3.10581	3.49661	4.02470	4.43698

n	Confidence Limits (top) and α (bottom) for a One-Tailed Test.												
	60%	**75%**	**80%**	**85%**	**90%**	**95%**	**97.5%**	**98%**	**99%**	**99.5%**	**99.75%**	**99.9%**	**99.95%**
	0.4	**0.25**	**0.2**	**0.15**	**0.1**	**0.05**	**0.025**	**0.02**	**0.01**	**0.005**	**0.0025**	**0.001**	**0.0005**
12	0.25903	0.69548	0.87261	1.08321	1.35622	1.78229	2.17881	2.30272	2.68100	3.05454	3.42844	3.92963	4.31779
13	0.25859	0.69383	0.87015	1.07947	1.35017	1.77093	2.16037	2.28160	2.65031	3.01228	3.37247	3.85198	4.22083
14	0.25821	0.69242	0.86805	1.07628	1.34503	1.76131	2.14479	2.26378	2.62449	2.97684	3.32570	3.78739	4.14045
15	0.25789	0.69120	0.86624	1.07353	1.34061	1.75305	2.13145	2.24854	2.60248	2.94671	3.28604	3.73283	4.07277
16	0.25760	0.69013	0.86467	1.07114	1.33676	1.74588	2.11991	2.23536	2.58349	2.92078	3.25199	3.68615	4.01500
17	0.25735	0.68920	0.86328	1.06903	1.33338	1.73961	2.10982	2.22385	2.56693	2.89823	3.22245	3.64577	3.96513
18	0.25712	0.68836	0.86205	1.06717	1.33039	1.73406	2.10092	2.21370	2.55238	2.87844	3.19657	3.61048	3.92165
19	0.25692	0.68762	0.86095	1.06551	1.32773	1.72913	2.09302	2.20470	2.53948	2.86093	3.17372	3.57940	3.88341
20	0.25674	0.68695	0.85996	1.06402	1.32534	1.72472	2.08596	2.19666	2.52798	2.84534	3.15340	3.55181	3.84952
21	0.25658	0.68635	0.85907	1.06267	1.32319	1.72074	2.07961	2.18943	2.51765	2.83136	3.13521	3.52715	3.81928
22	0.25643	0.68581	0.85827	1.06145	1.32124	1.71714	2.07387	2.18289	2.50832	2.81876	3.11882	3.50499	3.79213
23	0.25630	0.68531	0.85753	1.06034	1.31946	1.71387	2.06866	2.17696	2.49987	2.80734	3.10400	3.48496	3.76763
24	0.25617	0.68485	0.85686	1.05932	1.31784	1.71088	2.06390	2.17154	2.49216	2.79694	3.09051	3.46678	3.74540
25	0.25606	0.68443	0.85624	1.05838	1.31635	1.70814	2.05954	2.16659	2.48511	2.78744	3.07820	3.45019	3.72514
26	0.25595	0.68404	0.85567	1.05752	1.31497	1.70562	2.05553	2.16203	2.47863	2.77871	3.06691	3.43500	3.70661
27	0.25586	0.68368	0.85514	1.05673	1.31370	1.70329	2.05183	2.15782	2.47266	2.77068	3.05652	3.42103	3.68959
28	0.25577	0.68335	0.85465	1.05599	1.31253	1.70113	2.04841	2.15393	2.46714	2.76326	3.04693	3.40816	3.67391
29	0.25568	0.68304	0.85419	1.05530	1.31143	1.69913	2.04523	2.15033	2.46202	2.75639	3.03805	3.39624	3.65941

Confidence Limits (top) and α (bottom) for a One-Tailed Test.

n	60%	75%	80%	85%	90%	95%	97.5%	98%	99%	99.5%	99.75%	99.9%	99.95%
	0.4	**0.25**	**0.2**	**0.15**	**0.1**	**0.05**	**0.025**	**0.02**	**0.01**	**0.005**	**0.0025**	**0.001**	**0.0005**
30	0.25561	0.68276	0.85377	1.05466	1.31042	1.69726	2.04227	2.14697	2.45726	2.75000	3.02980	3.38518	3.64596
40	0.25504	0.68067	0.85070	1.05005	1.30308	1.68385	2.02108	2.12291	2.42326	2.70446	2.97117	3.30688	3.55097
50	0.25470	0.67943	0.84887	1.04729	1.29871	1.67591	2.00856	2.10872	2.40327	2.67779	2.93696	3.26141	3.49601
60	0.25447	0.67860	0.84765	1.04547	1.29582	1.67065	2.00030	2.09936	2.39012	2.66028	2.91455	3.23171	3.46020
70	0.25431	0.67801	0.84679	1.04417	1.29376	1.66691	1.99444	2.09273	2.38081	2.64790	2.89873	3.21079	3.43501
80	0.25419	0.67757	0.84614	1.04320	1.29222	1.66412	1.99006	2.08778	2.37387	2.63869	2.88697	3.19526	3.41634
90	0.25410	0.67723	0.84563	1.04244	1.29103	1.66196	1.98667	2.08394	2.36850	2.63157	2.87788	3.18327	3.40194
100	0.25402	0.67695	0.84523	1.04184	1.29007	1.66023	1.98397	2.08088	2.36422	2.62589	2.87065	3.17374	3.39049
500	0.25348	0.67498	0.84234	1.03751	1.28325	1.64791	1.96472	2.05912	2.33383	2.58570	2.81955	3.10661	3.31009
1000	0.25341	0.67474	0.84198	1.03697	1.28240	1.64638	1.96234	2.05643	2.33008	2.58075	2.81328	3.09840	3.30028
∞	0.25335	0.67449	0.84162	1.03643	1.28155	1.64485	1.95996	2.05375	2.32635	2.57583	2.80703	3.09023	3.29053

ლიტერატურა

1. ალბათობის თეორია და მათემატიკური სტატისტიკა ეკონომისტებისათვის. ა. რაზმაძის სახ. მათემ. ინსტიტუტი, უმაღლესი სასწავლებელი „ESM-თბილისი“, ფონდი „ევრაზია“. თბ. 2000. 664 გვ.
2. ა. რაზმაძის ლექციების კურსი ალბათობის თეორიაში. თბილისი. მეცნიერება. 1991. 223 გვ.
3. მათემატიკური სტატისტიკის ზოგიერთი მეთოდი. საქ. სსრ მეცნ. აკად. გამომც. თბილისი. 1963. 168 გვ.
4. გ. მანია. ალბათობის თეორია და მათემატიკური სტატისტიკა. სახელმძღვანელო ეკონომიკური ფაკულ-ტეტის სტუდენტებისათვის. თსუ. თბილისი. 1976. 244 გვ.
5. გ. მანია, ნ. ანთელავა, ა. ედიბერიძე. ალბათობის თეორიისა და მათემატიკური სტატისტიკის ამოცანათა კრებული. თსუ. თბილისი. 1980. 132 გვ.
6. თ. შერვაშიძე. ალბათობის თეორია (ლექციათა კურსი). თსუ. თბილისი. 1980.
7. ბ. დოჭვირი. ალბათობის თეორია და მათემატიკური სტატისტიკა. ლექციების კურსი ეკონომიკის ფაკულტეტის სტუდენტებისათვის. ნაწილი I და II. თსუ. თბილისი. 1984. 280 გვ.
8. გ. ფანცულაია. ალბათობის თეორია და მათემატიკური სტატისტიკა. ნაწ. I. თსუ. თბილისი. 1998. 136 გვ.
9. გ. ფანცულაია, ზ. ქვათაძე, გ. გიორგაძე. ალბათობის თეორია და მათემატიკური სტატისტიკა. სტუ. გამომცემ-ლობა. თბილისი. 2007. 228 გვ.
10. ნ. დურგლიშვილი. სოციოლოგიურ მონაცემთა ანალიზი. სალექციო კურსი სოციალური მეცნიერებების მაგისტ-რებისათვის – სოციალურ მეცნიერებათა ცენტრი CSS. პროგრამა „HESP“. თბილისი. 2006. 186 გვ.
11. ი. სხირტლაძე, თ. ტუღუში, ა. ცივაძე, მ. ნადარეიშვილი. ალბათობის თეორია და მათემატიკური სტატისტიკა. „განათლება“. თბილისი. 1990. 280 გვ.
12. ო. ფურთუხია, ალბათობის თეორია და მათემატიკური სტატისტიკა ეკონომისტებისათვის, თსუ, 2011, 210 გვ.

13. Гихман И.И., Скороход А.В., Ядренко М.И. Теория вероятностей и математическая статистика. Киев. «Выща школа». 1988. 440 с.
14. Кендалл М., Моран П. Геометрические вероятности. М. „Наука”. 1972. 192 с.
15. Кендалл М.Дж., Стьюарт А. Многомерный статистический анализ и временные ряды. М. „Наука”. 1976. 736 с.
16. Кендалл М.Дж., Стьюарт А. Статистические выводы и связи. М.: „Наука”. 1972. 900 с.
17. Прохоров Ю. В., Розанов Ю. А., Теория вероятностей, М. „Наука”, 1973, 496 с.
18. Маркентинговые исследования. Нерош К. Малхотра. Москва. Санкт-Петербург. Киев, 2003. 656 с.
19. SpSS – Искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей. Москва. Санкт-Петербург. Киев. 2005. Diasoft. 756 с.
20. Гильберт Черчиль, Дон Якобуччи. Маркентинговые исследования. Сантк-Петербург. 2004. Изд. «Нева». 532 с.
21. ჩავა ფრანკფორტ ნაჩმიასი, ანა ლეონ გერერო. სოციალური სტატისტიკა მრავალფეროვანი საზოგადოებისათვის. მეექვსე გამოცემა. თბილისის უნივერსიტეტის გამომცემლობა „Funded through the Book“. Translation Program, U.S. Embassy is Georgia. 2012. 634 გვ.
22. Paul Newbold, William L. Carlson, Betty Thorne. Statistics for Business and Economics. Fifth Edition. Pearson Education, Inc.,New Jersey, 2003, 850 p.

საგნობრივი საძიებელი

ჩ

ც

წ

ხ

ჯ

ჰ