2024

ალბათობის თეორიის და მათემატიკური სტატისტიკის საფუძვლები

დალი მაგრაქველიძე

რედაქტორი: ვახტანგ კვარაცხელია, ნიკო მუსხელიშვილის სახელობის გამოთვლითი მათემატიკის ინსტიტუტის დირექტორი, სქართველოს ტექნიკური უნივერსიტეტის პროფესორი

რეცენზენტი: ვაჟა ტარიელაძე, ფიზიკა-მათემატიკის მეცნიერებათა დოქტორი, საქართველოს ტექნიკური უნივერსიტეტის პროფესორი.

სახელმძღვანელოში მოცემული ალბათობისა და მათემატიკური სტატისტიკის ძირითადი საკითხები. წიგნში თემები წარმოდგენილია მკითხველისათვის ადვილად გასაგები ფორმით, ამიტომ მისი გამოყენება შესაძლებელია როგორც ტექნიკური სპეციალობის სტუდენტებისათვის, ასევე ნებისმიერი ამ საკითხებით დაინტერესებული მკითხველისათვის.

ISBN 978-9941-8-7199-3

შინაარსი

შესავალი

სტატისტიკა არის დარგი, რომელიც აწარმოებს მონაცემთა შეგროვებას, მათ ორგანიზებას, ანალიზს, ინტერპრეტაციას და პრეზენტაციას. მეცნიერული, ინდუსტრიული ან სოციალური პრობლემების გადასაჭრელად სტატისტიკის გამოყენება ჩვეულებრივ იწყება სტატისტიკური პოპულაციის შეგროვებით ან შესასწავლი სტატისტიკური მოდელის აგებით. პოპულაციები შეიძლება იყოს ადამიანთა სხვადასხვა ჯგუფები ან ობიექტები, როგორიცაა "ყველა ადამიანი, რომელიც ცხოვრობს ქვეყანაში" ან "თითოეული ატომი, რომელიც ქმნის კრისტალს". სტატისტიკა ეხება მონაცემთა ყველა ასპექტს, მათ შორის მონაცემთა შეგროვების დაგეგმვას კვლევებისა და ექსპერიმენტების ჩატარებისათვის.

ელემენტარული სტატისტიკური გაანგარიშებები ხორციელდებოდა ათასობით წლის წინაც. მაგალითად, გამოიყენებოდა მთელი რიგი გაზომვების ან დაკვირვებადი მონაცემების არითმეტიკული საშუალოს გაანგარიშების მეთოდი.

პირველი აღწერითი სტატისტიკა წარმოიშვა მონაცემთა შეგროვების შედეგად, მაგალითად, მოსახლეობის აღწერის ან ჯანმრთელობის ჩანაწერების რეესტრებში, შემდეგ კი დაკომპლექტდა გრაფიკული სახით.

მათემატიკური სტატისტიკის ალბათობის თეორიის საფუძველზე აგება პირველად მე-19 საუკუნის ბოლოდან, მე-20 საუკუნის დასაწყისში მოხდა. კარლ პირსონი და რონალდ ეილმერ ფიშერი ამ დარგის პიონერები იყვნენ.

ფიშერის წიგნი (1925) იყო მნიშვნელოვანი ეტაპი, რომელმაც ექსპერიმენტატორებს სტატისტიკურ კვლევებში ალბათობის თეორიის ძირითადი ცნებების გამოყენების საშუალება მისცა.

შეგვიძლია ვთქვათ, რომ სტატისტიკა ორი ნაწილისგან შედგება. პირველი ნაწილი შეეხება იმას, თუ როგორ უნდა ჩამოვაყალიბოთ და შევაფასოთ პრობლემის სავარაუდო მოდელები. მეორე ნაწილი კი მოიცავს პასუხის მიღების საკითხს გარკვეული მოდელის შერჩევის შემდეგ. სტატისტიკის ეს ნაწილი, ფაქტობრივად, ალბათობის თეორიის საგანია და პრაქტიკაში ასევე შეიცავს რიცხვითი ანალიზის

მნიშვნელოვან საკითხებსაც. ამრიგად, ალბათობის თეორია საშუალებას გვაძლევს ვიპოვოთ მოცემული იდეალური სამყაროს შედეგები, ხოლო სტატისტიკური თეორია საშუალებას გვაძლევს გავზომოთ ჩვენი სამყაროს იდეალურობის ხარისხი.

ალბათობას მკაცრად განსაზღვრული მათემატიკური მნიშვნელობა აქვს, ხოლო სტატისტიკა არის მეცნიერება მონაცემების შესახებ. ისინი ერთმანეთთან დაკავშირებულია, რადგან ალბათობა წარმოადგენს სტატისტიკის ერთგვარ საფუძველს, რომელიც უზრუნველყოფს ძირითად იდეებს.

ალბათობა ეხება განუსაზღვრელობის რაოდენობრივ განსაზღვრას, მაშინ როცა სტატისტიკა ხსნის განსაზღვრულობის გარკვეულ საზომში არსებულ განსხვავებას (მაგ., რატომ იცვლება შემოსავლის დონე?), რასაც ჩვენ ვხედავთ რეალურ სამყაროში.

სტატისტიკით ხდება მონაცემების შესწავლა: მათი ჩვენება (ინსტრუმენტების გამოყენებით, როგორიცაა დიაგრამები), შეჯამება (საშუალოების გამოყენება და სტანდარტული გადახრები და ა.შ). - ეს არის მთავარი რაოდენობრივი შეფასება იმისა, თუ რამდენად დარწმუნებული ვართ ჩვენს დასკვნებში. მაგრამ იმის დასადგენად, თუ რამდენად შეგვიძლია დარწმუნებულები ვიყოთ ჩვენს აღმოჩენებში, უნდა გამოვიყენოთ ალბათობა. ასე რომ, შეიძლება ითქვას, ალბათობა არის სტატისტიკური თეორიის მათემატიკური საფუძველი.

ალბათობის თეორიის საგანი. ცხოვრება უფრო მარტივი იქნებოდა, თუ უეჭველად გვეცოდინება რა მოხდება მომავალში. ასეთ შემთხვევაში ნებისმიერი გადაწყვეტილების შედეგი დამოკიდებულია მხოლოდ იმაზე, თუ რამდენად ლოგიკური და რაციონალურია გადაწყვეტილება. თუ ფული დაკარგეთ საფონდო ბირჟაზე, ეს იმის გამო, რომ თქვენ ვერ გაითვალისწინეთ ყველა ინფორმაცია ან არ მიიღეთ ლოგიკური გადაწყვეტილება. თუ წვიმაში დასველდით, ეს იმიტომ, რომ უბრალოდ დაგავიწყდათ ქოლგა. ყოველთვის შეგიძლიათ თავიდან აიცილოთ ზედმეტად დიდი ქარხნის აშენება, ინვესტირება კომპანიაში, რომელიც დაკარგავს ფულს, თუ ამოეწურება მარაგი ან დაკარგავს მოსავალს ცუდი ამინდის გამო. არცერთი ინვესტიცია აღარ იქნებოდა სარისკო. ასეთი ცხოვრება უფრო მარტივია, მაგრამ -მოსაწყენი.

ალბათობა არის რიცხვითი წარმოდგენა მოვლენის მოხდენის შესახებ. ამ წიგნში განვიხილავთ ალბათობის იმ ძირითად ცნებებს, რომლებიც სასარგებლოა მრავალი რაოდენობრივი ანალიზის პრობლემის გადასაჭრელად.

ხდომილებები შეიძლება იყოს აუცილებელი, შეუძლებელი და შემთხვევითი.

აუცილებელი ეწოდება ხდომილებას, რომელიც მოცემული ექსპერიმენტის პირობებში აუცილებლად ხდება. მაგალითად, თუ ჭურჭელში წყალი ასხია ოთახის 20 გრადუს ტემპერატურის და ნორმალური ატმოსფერული წნევის პირობებში, მაშინ ხდომილება „ჭურჭელში წყალი თხევად მდგომარეობაშია“ სარწმუნოა.

შეუძლებელი ეწოდება ხდომილებას, რომელიც მოცემული ექსპერიმენტის პირობებში არასდროს ხდება. მაგალითად, ხდომილება „ჭურჭელში წყალი მყარ მდგომარეობაშია“ არასოდეს მოხდება, თუ ოთახში ტემპერატურა არანაკლებ 20 გრადუსია.

შემთხვევითს უწოდებენ ხდომილებებს, რომლებიც მოცემული ექსპერიმენტის პირობებში ან მოხდება ან არა. მაგალითად, აგდებული მონეტა შეიძლება ისე დაეცეს, რომ ზემოთ იყოს გერბი ან საფასური. ამიტომ ხდომილება „მონეტის აგდების დროს გერბი მოვიდა“ - შემთხვევითია. თითოეული შემთხვევითი ხდომილება წარმოადგენს მრავალი შემთხვევითი მიზეზის მოქმედების შედეგს. შეუძლებელია შედეგზე ყველა ამ მიზეზის გავლენის გათვალისწინება, რადგან მათი რაოდენობა ძალიან ბევრია და მათი მოქმედების კანონები კი უცნობი. ამიტომ ალბათობის თეორიის წინაშე არ დგას ცალკეული ხდომილების მოხდენის წინასწარმეტყველების ამოცანა, ეს მას უბრალოდ არ შეუძლია.

სხვაგვარადაა საქმე, თუ განვიხილავთ შემთხვევით ხდომილებებს, რომელთა დაკვირვება მრავალჯერ ხდება ექსპერიმენტის ერთიდაიმავე პირობების განხორციელების დროს. საკმარისად დიდი რაოდენობით ერთგვაროვანი შემთხვევითი ხდომილებები მათი სპეციფიკური ხასიათის მიუხედავად ემორჩილება გარკვეულ კანონზომიერებებს. სწორედ ამ კანონზომიერებების დადგენას სწავლობს ალბათობის თეორია.

ამრიგად, ალბათობის თეორიის საგანს წარმოადგენს მასობრივი შემთხვევითი ხდომილებების ალბათური კანონზომიერებების შესწავლა.

მასობრივი შემთხვევითი მოვლენების მარეგულირებელი კანონების ცოდნა შესაძლებელს ხდის წინასწარ განისაზღვროს, თუ როგორ ხდება ამ მოვლენების გაგრძელება. მაგალითად, მართალია, შეუძლებელია წინასწარ განისაზღვროს ერთი მონეტის აგდების შედეგი, მაგრამ შესაძლებელია წინასწარ განვსაზღვროთ, და ამასთან მცირე ცდომილებით, „გერბის“ მოსვლის შემთხვევების რაოდენობა, თუ მონეტას ავაგდებთ საკმარისად ბევრჯერ (რა თქმა უნდა იგულისხმება, რომ მონეტის აგდება ხდება ერთსა და იმავე პირობებში). ალბათობის თეორიის მეთოდები ფართოდ გამოიყენება საბუნებისმეტყველო მეცნიერებებისა და ტექნოლოგიების სხვადასხვა დარგში: საიმედოობის თეორიაში, მასობრივი მომსახურების (რიგის) თეორიაში, თეორიულ ფიზიკაში, გეოდეზიაში, ასტრონომიაში, ბალისტიკაში, ცდომილებათა თეორიაში, ავტომატური მართვის თეორიაში, და ბევრ სხვა თეორიულ თუ გამოყენებით მეცნიერებებში. ალბათობის თეორია წარმოადგენს მათემატიკური და გამოყენებითი სტატისტიკის საფუძველს, რომლებიც, თავის მხრივ, გამოიყენება წარმოების დაგეგმვასა და ორგანიზებაში, ტექნოლოგიური პროცესების ანალიზში, პროდუქციის ხარისხის პრევენციულ კონტროლში და ასევე მრავალი სხვა მიზნისთვის. ბოლო პერიოდში ალბათობის თეორიის მეთოდებმა უფრო ღრმად შეაღწია მეცნიერებისა და ტექნოლოგიების სხვადასხვა დარგში, რამაც ხელი შეუწყო მათ პროგრესს.

ალბათობის თეორიის პირველ ნაშრომებში იყო აზარტული თამაშების თეორიის შექმნის მცდელობები (კარდანო, ჰიუგენსი, პასკალი, ფერმა და სხვები XVI-XVII სს.). ალბათობის თეორიის განვითარების შემდეგი ეტაპი ასოცირდება იაკობ ბერნულის (1654-1705) სახელთან. მის მიერ დამტკიცებული თეორემა, რომელსაც მოგვიანებით ეწოდა "დიდ რიცხვთა კანონი", იყო ამ კანონის პირველი მკაცრი მათემატიკური დამტკიცება.

თავი 1. ალბათობის თეორიის ძირითადი ცნებები

1.1. ცდა და ხდომილება

ხდომილებას ეწოდება შემთხვევითი, თუ განსაზღვრული პირობების რაიმე S ერთობლიობის განხორციელების დროს ის შეიძლება მოხდეს ან არ მოხდეს. შემდგომში იმის მაგივრად, რომ ითქვას „პირობების S ერთობლიობის განხორციელება", მოკლედ ვიტყვით „ჩატარებულია შემთხვევითი ცდა", ანუ ცდა, რომლის შედეგს წინასწარ ვერ განვსაზღვრავთ. ამრიგად, ხდომილება ჩაითვლება როგორც (შემთხვევითი) ცდის შედეგი.

მაგალითი 1. მსროლელი ესვრის მიზანს, რომელიც დაყოფილია არეებად. სროლა ეს არის ცდა. მიზნის გარკვეულ არეში მოხვედრა კი - ხდომილება.

მაგალითი 2. ურნაში ყრია ფერადი ბურთები. ურნიდან შემთხვევითი შერჩევით იღებენ ერთ ბურთს. ურნიდან ბურთის ამოღება ცდაა. გარკვეული ფერის ბურთის ამოღება - ხდომილება.

1.2 შემთხვევითი ხდომილებების სახეები

ორ ხდომილებას უწოდებენ *უთავსად (არათავსებად) ხდომილებებს*, თუ ერთ-ერთის მოხდენა გამორიცხავს იმავე ცდაში მეორე ხდომილების მოხდენას.

მაგალითი 1. დეტალების ყუთიდან შემთხვევითი შერჩევით იღებენ დეტალებს. სტანდარტული დეტალის ამოღება გამორიცხავს არასტანდარტული დეტალის ამოღებას. ხდომილებები „ამოღებულია სტანდარტული დეტალი" და „ამოღებულია არასტანდარტული დეტალი" - უთავსადი ხდომილებებია.

მაგალითი 2. ვაგდებთ მონეტას. „გერბის" მოხდენა გამორიცხავს „საფასურის" მოხდენას. ხდომილებები „გერბის მოხდენა და „საფასურის მოხდენა" უთავსი ხდომილებებია.

ნათქვამია, რომ რამდენიმე ხდომილება ქმნის ხდომილებათა სრულ სისტემას, თუ ისინი არიან წყვილ-წყვილად უთავსადი და ცდის შედეგად ერთ-ერთი მათგანი აუცილებლად ხდება.

მაგალითი 3. ნაყიდია ლატარიის ორი ბილეთი. აუცილებლად მოხდება შემდეგი ხდომილებებიდან ერთ-ერთი: „მოიგო პირველმა ბილეთმა და არ მოიგო მეორემ“, „არ მოიგო პირველმა ბილეთმა და მოიგო მეორემ“, „ორივე ბილეთმა მოიგო“, „არც ერთმა ბილეთმა მოიგო“. ეს ხდომილებები ქმნიან ხდომილებათა სრულ სისტემას.

მაგალითი 4. მსროლელმა ისროლა მიზანში. აუცილებლად მოხდება შემდეგი ხდომილებიდან ერთ-ერთი: მოხვედრა, აცდენა. ეს ორი უთავსადი ხდომილება ქმნის ხდომილებათა სრულ სისტემას.

მოცემულ ცდაში რამდენიმე ხდომილებას ეწოდებათ თანაბრად მოსალოდნელი (ან თანაბრად ალბათური), თუ ყველა ხდომილებას გააჩნია მოხდენის ერთნაირი (თანაბარი) „შანსი“.

მაგალითი 5. კამათელის აგდებისას ამა თუ იმ რიცხვის მოხდენა თანაბრად მოსალოდნელი ხდომილებებია. მართლაც, ნავარაუდევია, რომ კამათელი დამზადებულია ერთგვაროვანი მასალისაგან და აქვს წესიერი მრავალწახნაგას ფორმა, ამიტომ ნებისმიერი წახნაგის მოხდენას თანაბარი შანსი აქვს.

1.3 ალბათობის კლასიკური განსაზღვრება

ალბათობის თეორიის ძირითადი ცნებაა ალბათობა.

განვიხილოთ მაგალითი. ვთქვათ, ურნაში 6 ერთნაირი ბურთია. ამასთან 2 წითელია, 3 ლურჯია და 1 თეთრი. ცხადია, ფერადის ამოღების შესაძლებლობა უფრო მეტია ვიდრე თეთრის. შესაძლებელია ამ შესაძლებლობის რიცხვით დახასიათება. სწორედ ამ რიცხვს უწოდებენ ალბათობას. სხვა სიტყვებით რომ ვთქვათ, ალბათობა არის ხდომილების მოხდენის რიცხვითი ზომა.

ზემოთ განხილულ მაგალითში რიცხობრივად შევაფასოთ შემთხვევით ფერადი ბურთის ამოღების შანსი. ფერადი ბურთის ამოღება აღვნიშნოთ A ხდომილებით. ცდის ნებისმიერ შესაძლო შედეგს ვუწოდოთ ელემენტარული ხდომილება. ელემენტარული ხდომილებები აღვნიშნოთ ω_1 , ω_2 , ω_3 და ა.შ. ჩვენს მაგალითში შესაძლებელია 6 ელემენტარული ხდომილება: ω_1 - თეთრი ბურთის ამოღება; ω_2 , ω_3 , - წითელი ბურთის; $\omega_4, \omega_5, \omega_6$ კი ლურჯის. ადვილი შესამჩნევია, რომ ეს ხდომილებები ქმნიან ხდომილებათა სრულ სისტემას და ისინი არიან თანაბრად მოსალოდნელები.

იმ ელემენტარულ ხდომილებებს, რომელთა მოხდენა იწვევს ჩვენთვის საინტერესო ხდომილების მოხდენას, უწოდებენ ამ ხდომილების ხელშემწყობ შემთხვევებს. ჩვენ შემთხვევაში A ხდომილების ხელშემწყობი შემთხვევებია ω_2,ω_3,ω_4,ω_5,ω_6 ხდომილებები.

ამგვარად, A ხდომილება ხდება, თუ ხდება A ხდომილების ხელშემწყობი შემთხვევებიდან ერთ-ერთი მაინც; მაშასადამე, ჩვენს მაგალითში A ხდომილება ხდება, თუ ხდება ან $\omega_{2,}$ ან ω_3 , ან ω_4 , ან ω_5 , ან ω_6 ელემენტარული ხდომილება. ესე იგი A ხდომილება იშლება რამდენიმე ელემენტარულ ხდომილებად (ω_2,ω_3,ω_4,ω_5,ω_6); თვით ელემენტერული ხდომილებები კი არ იშლება სხვა ხდომილებებად. ამით განსხვავდებიან A ხდომილება და ელემენტერული ხდომილებები. A ხდომილებას ზოგჯერ რთულ, შედგენილ ხდომილებას უწოდებენ.

A ხდომილების ხელშემწყობ შემთხვევათა რაოდენობის ფარდობას ყველა შესაძლო შემთხვევათა რაოდენობასთან (ანუ ელემენტარულ ხდომილებათა საერთო რაოდენობასთან) უწოდებენ A ხდომილების ალბათობას და აღნიშნავენ P(A) სიმბოლოთი. აქედან გამომდინარე, ფერადი ბურთის ამოღების ალბათობა იქნება $P(A) = {}^{5}\!/_{6}$. ეს რიცხვი იძლევა ფერადი ბურთის ამოღების შანსის იმ რაოდენობრივ შეფასებას, რომელსაც ვეძებდით. ახლა შემოვიტანოთ ალბათობის კლასიკური განსაზღვრება.

ვთქვათ, ცდის ყველა შესაძლო შედეგთა ერთობლიობის რაოდენობა სასრული n რიცხვია და დავუშვათ, რომ თითოეული შედეგის მოხდენა ერთნაირად არის

მოსალოდნელი. მაშინ, ალბათობის კლასიკური განსაზღვრების თანახმად A *ხდომილების ალბათობა ტოლია* ამ ხდომილების ხელშემწყობ შემთხვევათა რაოდენობის შეფარდებისა ყველა შესაძლო შედეგთა რაოდენობასთან. ამრიგად, A ხდომილების ალბათობა განისაზღვრება ფორმულით:

$$P(A) = {}^{m}/_{n}$$

სადაც m არის A ხდომილების ხელშემწყობ შემთხვევათა რაოდენობა, ხოლო n კი ცდის ყველა შესაძლო შედეგთა რაოდენობაა.

თვისება 1. აუცილებელი ხდომილების ალბათობა ერთის ტოლია.

მართლაც, თუ ხდომილება აუცილებელია, მაშინ ცდის ნებისმიერი ელემენტარული ხდომილება იქნება ამ ხდომილების ხელშემწყობი. ე.ი. ამ შემთხვევაში $m = n$, შესაბამისად

$$P(A) = {}^{n}/_{n} = 1.$$

თვისება 2. შეუძლებელ ხდომილების ალბათობა ნულის ტოლია.

მართლაც, თუ ხდომილება შეუძლებელია, მაშინ ცდის ნებისმიერი ელემენტარული ხდომილება არ არის ხდომილების ხელშემწყობი. ამ შემთხვევაში m=0, შესაბამისად

$$P(A) = {}^{0}/_{n} = 0.$$

თვისება 3. შემთხვევითი ხდომილების ალბათობა არის ნულსა და ერთს შორის მოთავსებული არაუარყოფითი რიცხვი.

მართლაც, შემთხვევით ხდომილების ხელშემწყობია მხოლოდ ცდის ელემენტარული ხდომილებების საერთო რაოდენობის ნაწილი. ამ შემთხვევაში $0 \leq m \leq n$, ე.ი. $0 \leq m/n \leq 1$, შესაბამისად,

$$0 \leq P(A) \leq 1.$$

ამრიგად, ნებისმიერი ხდომილების ალბათობა აკმაყოფილებს ორმაგ უტოლობას

$$0 \leq P(A) \leq 1.$$

შემდგომში მოვიყვანთ თეორემებს, რომლებიც ზოგიერთი ხდომილების ცნობილი ალბათობების მეშვეობით სხვა ხდომილებების ალბათობის პოვნის საშუალებას იძლევიან.

1.4. კომბინატორიკის ძირითადი ფორმულები

კომბინატორიკა არის მათემატიკის ნაწილი, რომელიც შეისწავლის მოცემული სიმრავლიდან ელემენტთა ამორჩევის ამოცანებს და მოცემული წესების მიხედვით მათ განლაგებას ჯგუფებში. კერძოდ, ასეთ ამოცანათა რიგს მიეკუთვნება მოცემული სასრულო სიმრავლის ელემენტებისგან მიღებული კომბინაციების (ამორჩევების) რაოდენობების გამოთვლა. კლასიკური განსაზღვრების საშუალებით ალბათობის გამოსათვლელად ხშირად საჭირო ხდება კომბინატორიკის ზოგიერთი მეთოდის და შედეგის გამოყენება.

n ელემენტიანი სიმრავლის ნებისმიერ დალაგებას n ელემენტიანი გადანაცვლება ეწოდება, რომელთა რაოდენობა აღვნიშნოთ P_n-ით.

$$P_n = n!,$$

სადაც $n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$.

შევნიშნოთ, რომ $0! = 1$.

მაგალითი 1. რამდენი სამნიშნა რიცხვი შეიძლება შევადგინოთ ციფრებისაგან 1, 2, 3, თუ თითოეული ციფრი გამოსახულებაში შედის მხოლოდ ერთხელ?

ამოხსნა: მოსაძებნი სამნიშნა რიცხვების რაოდენობა ტოლია

$$P_3 = 3! = 1 \cdot 2 \cdot 3 = 6.$$

n ელემენტიანი სიმრავლის ყოველ m ელემენტიან დალაგებულ ქვესიმრავლეს ეწოდება n ელემენტიდან m ელემენტიანი წყობა. განსაზღვრებიდან გამომდინარეობს, რომ წყობა ეს არის m ელემენტისგან შემდგარი ამორჩევები (კომბინაციები), რომლებიც ერთმანეთისაგან განსხვავდებიან ან ელემენტთა შემადგენლობით, ან მათი დალაგების რიგით. n ელემენტიანი სიმრავლის m ელემენტიანი წყობის რაოდენობა აღინიშნება A_n^m სიმბოლოთი (იკითხება: A n-დან m-ად) და გამოითვლება ფორმულით:

$$A_n^m = n(n-1)(n-2)\dots(n-m+1).$$

მაგალითი 2. რამდენი სიგნალი შეიძლება შევადგინოთ ორ-ორად აღებული სხვადასხვა ფერის 6 ალმისაგან?

ამოხსნა: სიგნალების საძიებელი რაოდენობა ტოლია

$$A_6^2 = 6 \cdot 5 = 30.$$

n ელემენტიანი სიმრავლის ყოველ m ელემენტიან ქვესიმრავლეს ეწოდება n ელემენტიდან m ელემენტიანი **ჯუფთება**. განსაზღვრებიდან გამომდინარეობს, რომ ჯუფთება არის ამორჩევები (კომბინაციები), რომელთაგან თითოეული შედგება მოცემული n ელემენტიანი სიმრავლიდან აღებული m ელემენტისაგან, რომლებიც ერთმანეთისაგან განსხვავდებიან ელემენტთა შემადგენლობით. n ელემენტიანი სიმრავლის m ელემენტიანი ჯუფთებების რაოდენობა აღინიშნება C_n^m სიმბოლოთი (იკითხება: C n-დან m-ად) და გამოითვლება ფორმულით:

$$C_n^m = n!/(m!\,(n-m)!).$$

მაგალითი 3. რამდენაირად შეიძლება ამოვარჩიოთ ორი დეტალი ყუთიდან რომელშიც 10 დეტალია?

ამოხსნა: შერჩევის ხერხების რაოდენობა ტოლია

$C_{10}^2 = 10!/(2!\,8!) = 45.$

შევნიშნოთ, რომ გადანაცვლების, წყობის და ჯუფთების რაოდენობები ერთმანეთთან შემდეგი ფორმულითაა დაკავშირებული

$$A_n^m = P_m C_n^m.$$

მრავალი კომბინატორული ამოცანა შეიძლება ამოიხსნას შემდეგი ორი მნიშვნელოვანი წესის გამოყენებით, რომელთაც, შესაბამისად, გამრავლების და შეკრების წესს უწოდებენ.

გამრავლების წესი: ვთქვათ, A სიმრავლე შეიცავს m ცალ ელემენტს $A = \{a_1, a_2, \cdots, a_m\}$ და B სიმრავლე შეიცავს n-ცალ ელემენტს $B = \{b_1, b_2, \cdots, b_n\}$. მაშინ ამ ორი სიმრავლის ელემენტებიდან შეიძლება შედგეს $m \cdot n$ -ცალი შემდეგი სახის წყვილი (a_i, b_k), $i = 1,2,\cdots,m;\ k = 1,2,\cdots,n.$

ეს წესი, ცხადია, ვრცელდება სამი და მეტი რაოდენობის სიმრავლეებზეც.

შეკრების წესი. ვთქვათ, A სიმრავლე შეიცავს mცალ ელემენტს $A = \{a_1, a_2, \cdots, a_m\}$, B სიმრავლე შეიცავს n-ცალ ელემენტს $B = \{b_1, b_2, \cdots, b_n\}$ და $AB = \emptyset$. მაშინ ამ ორი სიმრავლის ელემენტებიდან ერთი ელემენტი შეიძლება ამოირჩეს $m + n$ სხვადასხვა ხერხით.

ეს წესი, ცხადია, ვრცელდება სამი და მეტი რაოდენობის წყვილ-წყვილად თანაუკვეთ სიმრავლეებზე.

1.5. ფარდობითი სიხშირე. ფარდობითი სიხშირის მდგრადობა

ფარდობითი სიხშირე ალბათობასთან ერთად წარმოადგენს ალბათობის თეორიის ერთ-ერთ ძირითად ცნებას.

ვთქვათ, A არის რაიმე შემთხვევითი ცდის ერთ-ერთი შესაძლო შედეგი (ხდომილება). ეს ცდა ჩავატაროთ n-ჯერ და m-ით აღვნიშნოთ A შედეგის მოხდენის რაოდენობა ამ n-ჯერ ჩატარებულ ცდაში. m რიცხვს უწოდებენ ცდათა მოცემულ მიმდევრობაში A ხდომილების აბსოლუტურ სიხშირეს, ხოლო შეფარდებას $W(A) = m/n$, ეწოდება A ხდომილების **სიხშირე (ფარდობითი სიხშირე)**. ის, ცხადია, დამოკიდებულია ცდათა რაოდენობაზე (ანუ n-ზე). გარდა ამისა, შევნიშნოთ, რომ თუ ხელმეორედ ჩავატარებთ იმავე რაოდენობის ცდებს, სიხშირის მნიშვნელობა, საზოგადოდ, შეიცვლება.

მაგალითი 1. განხორციელდა მიზანში 24 გასროლა, ამასთან დარეგისტრირდა მიზნში 19 მოხვედრა. მიზანში მოხვედრის ფარდობითი სიხშირე ტოლია

$$W(A) = 19/24.$$

$W(A)$ სიხშირის მნიშვნელობა შემთხვევითია, რადგან შეუძლებელია n ცდისგან შემდგარი ცდათა სერიის ჩატარებამდე ვიწინასწარმეტყველოთ მისი ზუსტი მნიშვნელობა. მაგრამ პრაქტიკაში შეინიშნება, რომ სიხშირეს გააჩნია შემდეგი ფუნდამენტური თვისება, რომელსაც სტატისტიკური მდგრადობის თვისება ეწოდება და მდგომარეობს შემდეგში: დიდი n-ებისათვის A ხდომილების სიხშირე n-ის ცვლილებისას საკმარისად

მცირედ იცვლება (ანუ ჯგუფდება) რომელიღაც არაშემთხვევითი რიცხვის მახლობლობაში. უფრო მეტიც, თუ ჩავატარებთ ექსპერიმენტთა რამდენიმე სერიას, მაშინ სიხშირეები ყოველ სერიაში იქნება ერთმანეთთან საკმაოდ ახლოს, როგორც კი ჩატარებული ცდების რაოდენობები თითოეულ სერიაში იქნება საკმარისად დიდი რიცხვი.

1.6. სტატისტიკური ალბათობა

ალბათობის კლასიკური განსაზღვრება გულისხმობს, რომ ელემენტარული ხდომილებების რაოდენობა სასრულია. პრაქტიკაში კი ხშირად გვხდება ისეთი ცდები, რომლებშიც შესაძლო შედეგები უსასრულოა. ასეთ შემთხვევაში, ცხადია, ალბათობის კლასიკურ განსაზღვრებას ვერ გამოვიყენებთ.

კლასიკური განსაზღვრების სუსტი მხარე მდგომარეობს იმაში, რომ ხშირად შეუძლებელია ცდის შედეგის ელემენტარული ხდომილებების ერთობლიობის სახით წარმოდგენა. უფრო რთულია ელემენტარული ხდომილებების თანაბრად მოსალოდნელობის დასაბუთება. ჩვეულებრივ თანაბრად მოსალოდნელობას განსაზღვრავენ სიმეტრიის მოსაზრებებიდან გამომდინარე. ასე მაგალითად, სიმეტრიის მოსაზრებებიდან გამომდინარე ითვლება, რომ სათამაშო კამათელს აქვს წესიერი მრავალწახნაგას (კუბის) ფორმა და ის დამზადებული ერთგვაროვანი მასალისაგან. მაგრამ პრაქტიკაში საკმაოდ იშვიათია ამოცანები, რომლებშიც შეიძლება გამოყენებული იქნეს სიმეტრიის მოსაზრებები. ამის გამო კლასიკურ განსაზღვრებასთან ერთად გამოიყენება სხვა განსაზღვრებებიც, რომლის მაგალითსაც კერძოდ, წარმოადგენს სტატისტიკური განსაზღვრება: *ხდომილების სტატისტიკურ ალბათობად მიღებულია მისი ფარდობითი სიხშირე ან მასთან მიახლოებული რიცხვი.* მაგალითად, თუ საკმარისად დიდი რაოდენობის ცდის შედეგად აღმოჩნდა, რომ ფარდობითი სიხშირე 0,4-თან საკმაოდ ახლოსაა, მაშინ ეს რიცხვი შეიძლება მივიღოთ ხდომილების სტატისტიკურ ალბათობად.

ადვილი შესამოწმებელია, რომ ალბათობის კლასიკური განსაზღვრებიდან გამომდინარე ალბათობის თვისებები შენარჩუნებულია ალბათობის სტატისტიკური განსაზღვრების დროსაც. მართლაც, თუ ხდომილება აუცილებელია, მაშინ $m = n$ და ფარდობითი სიხშირე

$$m/n = n/n = 1.$$

ე.ი. აუცილებელი ხდომილების სტატისტიკური ალბათობა (ისევე როგორც კლასიკური განსაზღვრების დროს) ერთის ტოლია.

თუ ხდომილება შეუძლებელია მაშინ $m = 0$ და, შესაბამისად, ფარდობითი სიხშირე ტოლია

$$0/n = 0.$$

ე.ი. შეუძლებელი ხდომილების სტატისტიკური ალბათობა ნულის ტოლია.

ნებისმიერი ხდომილებისათვის $0 \le m \le n$, და შესაბამისად ფარდობითი სიხშირესათვის სამართლიანია უტოლობა

$$0 \le m/n \le 1,$$

ე.ი. ნებისმიერი ხდომილების სტატისტიკური ალბათობა მოთავსებულია ნულსა და ერთს შორის.

იმისათვის, რომ არსებობდეს A ხდომილების სტატისტიკური ალბათობა საჭიროა:

ა) პრინციპულად მაინც იყოს შესაძლებელი განხორციელდეს ცდების შეუზღუდავი რაოდენობა, რომელთაგან თითოეულში A ხდება ან არ ხდება;

ბ) A-ს მოხდენის ფარდობითი სიხშირეების მდგრადობა ცდების საკმაოდ დიდი რაოდენობის სხვადასხვა სერიაში.

სტატისტიკური განსაზღვრების ნაკლს წარმოადგენს სტატისტიკური ალბათობის არაცალსახობა; ზევით ნახსენებ მაგალითში ხდომილების ალბათობად შეიძლება მიღებული იქნას არა მარტო 0,4, არამედ 0,39; 0,41 და ა.შ.

1.7. გეომეტრიული ალბათობა

ვთქვათ, l მონაკვეთი წარმოადგენს L მონაკვეთის ნაწილს. L მონაკვეთზე ნებისმი-ერად აღებულია წერტილი. აქ ჩვენ ვგულისხმობთ, რომ შესრულებულია შემდეგი

დაშვებები: აღნიშნული წერტილი შეიძლება აღმოჩნდეს L მონაკვეთის ნებისმიერ წერტილში და l მონაკვეთზე წერტილის მოხვედრის ალბათობა ამ მონაკვეთის სიგრძის პროპორციულია და არაა დამოკიდებული L მონაკვეთის მიმართ l მონაკვეთის მდებარეობაზე. ამ დაშვებებში წერტილის l მონაკვეთზე მოხვედრის ალბათობა განისაზღვრება ტოლობით

$$P = l \text{ სიგრძე}/L \text{ სიგრძე}.$$

ვთქვათ ბრტყელი ფიგურა g შეადგენს G ფიგურის შემადგენელ ნაწილს. ფიგურა G -ზე ნებისმიერად შერჩეულია წერტილი. ეს ნიშნავს, რომ შერჩეული წერტილი შეიძლება აღმოჩნდეს G ფიგურის ნებისმიერ წერტილში და აღნიშნული წერტილის ფიგურა g-ზე მოხვედრის ალბათობა პროპორციულია ამ ფიგურის ფართობის და არაა დამოკიდებული არც ამ ფიგურის G -ს მიმართ მდებარეობაზე და არც g-ს ფორმაზე. ამ დაშვებებში წერტილის g ფიგურაში მოხვედრის ალბათობა განისაზღვრება შემდეგი ტოლობით

$$P = g \text{ ფართობი}/G \text{ ფართობი}$$

1.8 ალბათობის სახეები

უნდა აღინიშნოს, რომ ალბათობის განსაზღვრის ორი განსხვავებული მიდგომა არსებობს: ობიექტური მიდგომა და სუბიექტური მიდგომა.

ობიექტური ალბათობა. ობიექტური ალბათობა შეიძლება გამოითვალოს შემდეგი ფორმულით:

$$P(\text{ხდომილება}) = \frac{\text{ხდომილების მოხდენის რიცხვი}}{\text{ცდების და შედეგების მთლიანი რაოდენობა}}$$

ობიექტური ალბათობა ასევე შეიძლება დადგინდეს კლასიკური ან ლოგიკური მეთოდის გამოყენებით. მთელი რიგი ცდების ჩატარების გარეშე, ჩვენ ხშირად შეგვიძლია ლოგიკურად განვსაზღვროთ როგორი უნდა იყოს სხვადასხვა ხდომილების ალბათობა. მაგალითად, ალბათობა იმისა, რომ წესიერი მონეტის ერთხელ აგდების შედეგად მიიღოთ საფასური, არის

$$P(\text{საფასური}) = \frac{1}{2}$$

საფასურის მღების გზების რაოდენობა

შესაძლო შედეგების რაოდენობა

ანალოგიურად, 52 სათამაშო ბანქოს დასტიდან ყვავის ამოღების ალბათობა ლოგიკურად შეიძლება დადგინდეს შემდეგნაირად:

$$P(\text{ყვავი}) = \frac{13}{52}$$

ყვავის მოსვლის შემთხვევების რიცხვი

შესაძლო შედეგების რაოდენობა

$= \frac{1}{4} = 0.25 = 25\%$

სუბიექტური ალბათობა. როდესაც ლოგიკა და წარსული ისტორია არ არის შესაბამისობაში, ალბათობის მნიშვნელობები შეიძლება შეფასდეს სუბიექტურად. სუბიექტური ალბათობების სიზუსტე დამოკიდებულია შემფასებლის გამოცდილებაზე და მოსაზრებაზე. რიგი ალბათობის მნიშვნელობები ვერ განისაზღვრება, თუ სუბიექტური მიდგომა არ იქნება გამოყენებული. რა არის ალბათობა იმისა, რომ მომდევნო წლებში 1 ლიტრი ბენზინის ფასი 4 ლარზე მეტი იქნება? რა არის იმის ალბათობა, რომ ჩვენი ეკონომიკა მძიმე დეპრესიაში იქნება 2025 წელს? რა არის იმის ალბათობა, რომ თქვენ გახდებით მსხვილი კორპორაციის პრეზიდენტი 20 წლის შემდეგ?

არსებობს სუბიექტური ალბათობის შეფასების რამდენიმე მეთოდი. საზოგადოებრივი აზრის გამოკითხვა შეიძლება გამოყენებული იქნას საარჩევნო სუბიექტური ალბათობის შესაძლო შედეგების დასადგენად. ალბათობის მნიშვნელობების შესაფასებლად, ზოგ შემთხვევაში, აუცილებელია გამოვიყენოთ გარკვეული სუბიექტური მოსაზრებები, გამოცდილება და განსჯა. მაგალითად, წარმოების მენეჯერს შეიძლება სჯეროდეს, რომ ახალი პროდუქტის წარმოების ალბათობა ერთი დეფექტის გარეშე არის 0.85.

თავი 2. ალბათობათა შეკრების თეორემა

2.1 უთავსადი ხდომილებების ალბათობათა შეკრება

ორი A და B ხდომილებათა A+B ჯამს უწოდებენ ხდომილებას, რომლის დროსაც ხდება ან A ხდომილება ან B ხდომილება ან ორივე ხდომილება ერთდროულად. მაგალითად, თუ ქვემეხით ორი გასროლაა განხორციელებული და A არის პირველი გასროლით მოხვედრა, B - მეორე გასროლით მოხვედრა, მაშინ A+B არის პირველი გასროლით მოხვედრა, ან მეორე გასროლით მოხვედრა, ან ორივე გასროლით მოხვედრა.

თუ ორი ხდომილება A და B უთავსადია, მაშინ A+B არის ხდომილება, რომელიც მდგომარეობს ამ ხდომილებებიდან მხოლოდ ერთ-ერთის მოხდენაში.

რამდენიმე ხდომილების ჯამს უწოდებენ ხდომილებას, როდესაც ერთერთი მათგანი მაინც ხდება. მაგალითად ხდომილება $A + B + \cdots + C$ მდგომარეობს შემდეგი A, B,⋯,C ხდომილებებიდან ერთ-ერთის მოხდენაში,.

ვთქვათ A და B არათავსებადი ხდომილებებია, ამასთან ამ ხდომილებების ალბათობები ცნობილია. როგორ ვიპოვოთ იმის ალბათობა, რომ მოხდება ან A ხდომილება ან B ხდომილება? ამაზე პასუხს შეკრების თეორემა იძლევა.

თეორემა. *ორი უთავსადი ხდომილებიდან ერთ-ერთის მოხდენის ალბათობა ტოლია ამ ხდომილებების ალბათობების ჯამის:*

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

დამტკიცება. შემოვიღოთ აღნიშვნები: n იყოს ცდის ყველა შესაძლო შედეგთა რაოდენობა; m_1 - იყოს A-ს ხელშემწყობ შემთხვევათა რაოდენობა; m_2 კი იყოს B-ს ხელშემწყობ შემთხვევათა რაოდენობა. ან A ან B ხდომილების მოხდენის ხელშემწყობ შემთხვევათა რაოდენობა, ცხადია, ტოლია $m_1 + m_2$ის. შესაბამისად,

$$P(A + B) = (m_1 + m_2)/n = m_1/n + m_2/n.$$

თუ გავითვალისწინებთ, რომ $m_1/n = P(A)$ და $m_2/n. = P(B)$, საბოლოოდ მივიღებთ

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

შედეგი. *რამდენიმე წყვილ-წყვილად უთავსადი ხდომილებიდან ერთ-ერთის მოხდენის ალბათობა ტოლია ამ ხდომილებების ალბათობების ჯამის:*

$$P(A_1 + A_2 + \cdots + A_n) = P(A_1) + P(A_2) + \cdots + P(A_n).$$

შენიშვნა: დამტკიცებისათვის შეგვიძლია გამოვიყენოთ მათემატიკური ინდუქციის მეთოდი.

2.2. ხდომილებათა სრული სისტემა

თეორემა. $A_1, A_2, \dots, A_n$ ხდომილებათა ალბათობების ჯამი, რომლებიც ხდომილებათა სრულ სისტემას ქმნიან, ერთის ტოლია:

$$P(A_1) + P(A_2) + \cdots + P(A_n) = 1.$$

დამტკიცება. რამდენადაც ხდომილებათა სრული სისტემის ხდომილებებიდან ერთ-ერთის მოხდენა აუცილებელია, ხოლო აუცილებელი ხდომილების ალბათობა ერთის ტოლია, ამიტომ

$$P(A_1 + A_2 + \cdots + A_n) = 1 \tag{2.1}$$

ხდომილებათა სრული სისტემის ნებისმიერი ორი ხდომილება უთავსადია, ამიტომ შეგვიძლია გამოვიყენოთ ალბათობათა შეკრების თეორემა:

$$P(A_1 + A_2 + \cdots + A_n) = P(A_1) + P(A_2) + \cdots + P(A_n). \tag{2.2}$$

თუ (2.1) და (2.2) ტოლობებს შევადარებთ, მივიღებთ

$$P(A_1) + P(A_2) + \cdots + P(A_n) = 1.$$

2.3. მოპირდაპირე ხდომილებები

ორ ხდომილებას უწოდებენ მოპირდაპირე ხდომილებებს, თუ ისინი ქმნიან ხდომილებათა სრულ სისტემას. თუ მოპირდაპირე ხდომილებებიდან ერთ-ერთს A-თი აღვნიშნავთ, მაშინ მიღებულია მეორის $\bar{A}$-ით აღნიშვნა.

თეორემა. მოპირდაპირე ხდომილებების ალბათობათა ჯამი ერთის ტოლია

$$P(A) + P(\bar{A}) = 1.$$

დამტკიცება. მოპირდაპირე ხდომილებები ქმნიან ხდომილებათა სრულ სისტემას, ხოლო, როგორც ზემოთ აღვნიშნეთ, იმ ხდომილებების ალბათობათა ჯამი, რომლებიც ხდომილებათა სრულ სისტემას ქმნიან, ერთის ტოლია.

შენიშვნა 1. თუ მოპირდაპირე ხდომილებებიდან ერთ-ერთის ალბათობას ავღნიშნავთ p-თი,ხოლო მეორისას q-თი, მაშინ წინა თეორემიდან გამომდინარე

$$p + q = 1.$$

შენიშვნა 2. A ხდომილების ალბათობის მოსაძებნად ხშირად მოსახერხებელია ჯერ გამოვთვალოთ $\bar{A}$ ხდომილების ალბათობა და შემდეგ გამოვთვალოთ საძიებელი ალბათობა შემდეგი ფორმულის მეშვეობით

$$P(A) = 1 - P(\bar{A}).$$

2.4. ნაკლებალბათური ხდომილებების პრაქტიკულად შეუძლებლობის პრინციპი

მრავალი პრაქტიკული ამოცანის დროს საქმე გვაქვს ხდომილებებთან, რომელთა ალბათობები ძალიან მცირეა, ე.ი. ნულთან ახლოსაა. შეიძლება თუ არა ჩავთვალოთ, რომ მცირე ალბათობის მქონე A ხდომილება კონკრეტულ ცდაში არ მოხდება? ასეთი დასკვნის გაკეთება არ შეიძლება, რამდენადაც გამორიცხული არაფერია, თუმცა ნაკლებად მოსალოდნელია, რომ A ხდომილება მოხდება.

თითქოს ცალკეული ცდის ჩატარების დროს მცირე ალბათობის მქონე ხდომილების მოხდენა-არმოხდენის წინასწარმეტყველება შეუძლებელია. მაგრამ ხანგრძლივი გამოცდილება გვიჩვენებს, რომ მცირე ალბათობის მქონე ხდომილება ერთეული ცდების უმეტესობაში არ ხდება. ამ ფაქტზე დაყრდნობით ჩამოყალიბებულია შემდეგი „მცირე ალბათობის მქონე ხდომილებების პრაქტიკულად შეუძლებლობის პრინციპი“: *თუ შემთხვევით ხდომილებას გააჩნია ძალიან მცირე ალბათობა, მაშინ შეიძლება ჩაითვალოს, რომ პრაქტიკულად ერთეულ ცდაში ეს ხდომილება არ ხდება.*

ბუნებრივია ჩნდება შეკითხვა: რამდენად მცირე უნდა იყოს ხდომილების ალბათობა, რომ ერთ ცდაში მისი მოხდენა ჩაითვალოს შეუძლებლად? ამ კითხვაზე ცალსახა პასუხის გაცემა შეუძლებელია. არსებითად განსხვავებულ ამოცანებზე პასუხი

სხვადასხვაა. მაგალითად, თუ იმის ალბათობა, რომ გადახტომისას პარაშუტი არ გაიხსნება ტოლია 0,01, მაშინ დაუშვებელია ასეთი პარაშუტების გამოყენება. თუკი ალბათობა იმისა, რომ საქალაქთაშორისო მატარებელი დააგვიანებს, ტოლია 0,01-ის, მაშინ პრაქტიკულად შეიძლება ვიყოთ დარწმუნებული, რომ მატარებელი დროზე მოვა.

საკმარისად მცირე ალბათობას, რომლის დროსაც (განსაზღვულ ამოცანაში) ხდომილება შეიძლება პრაქტიკულად შეუძლებლად ჩაითვალოს, ეწოდება *მნიშვნელოვნობის დონე.* ხშირ შემთხვევაში პრაქტიკაში მნიშვნელოვნობის დონე მოთავსებულია 0,01-სა და 0,05-ს შორის.

შევნიშნოთ, რომ აქ განხილული პრინციპი საშუალებას იძლევა გაკეთდეს პროგნოზები არა მარტო მცირე ალბათობის მქონე ხდომილებების შესახებ, არამედ ისეთ ხდომილებებზეც, რომელთა ალბათობა ერთთან ახლოსაა. მართლაც თუ A ხდომილების ალბათობა ნულთან ახლოსაა, მაშინ მოპირდაპირე $\overline{A}$ ხდომილების ალბათობა არის ერთთან ახლოს. მერეს მხრივ, A ხდომილების არმოხდენა ნიშნავს მოპირდაპირე $\bar{A}$ ხდომილების მოხდენას. ამგავარად, მცირე ალბათობის მქონე ხდომილებების შეუძლებლობის პრინციპიდან გამომდინარეობს შემდეგი დასკვნა: *თუ შემთხვევით ხდომილების ალბათობა ერთთან ძალიან ახლოსა, მაშინ პრაქტიკულად შეიძლება ჩაითვალოს, რომ ერთეულ ცდაში ეს ხდომილება მოხდება.* რა თქმა უნდა, აქაც პასუხი შეკითხვაზე, თუ რომელი ალბთობა უნდა ჩაითვალოს ერთთან ახლოს, დამოკიდებულია თვით ამოცანის შინაარსზე.

თავი 3. ალბათობათა გამრავლების თეორემა

3.1. ხდომილებათა ნამრავლი

ორი A და B ხდომილების ნამრავლს უწოდებენ AB ხდომილებას, რომელიც მდგომარეობს ამ ხდომილებათა ერთდროულად მოხდენაში. მაგალითად, თუ A ვარგისი დეტალია, B - შეღებილი დეტალი, მაშინ AB - ვარგისი და შეღებილი დეტალია.

რამდენიმე ხდომილების ნამრავლს უწოდებენ ხდომილებას, რომელიც მდგომარეობს ყველა ამ ხდომილების ერთდროულ მოხდენაში. მაგალიად, თუ A, B, C „გერბის" მოსვლაა შესაბამისად მონეტის პირველ, მეორე და მესამე აგდებისას, მაშინ ABC „გერბის" მოსვლაა სამივე ცდაში.

3.2. პირობითი ალბათობა

შემთხვევითი ხდომილება ჩვენ განვსაზღვრეთ როგორც ხდომილება, რომელიც S პირობების ერთობლიობის შესრულებისას შეიძლება მოხდეს ან არ მოხდეს. თუ ხდომილების ალბათობის გამოთვლისას S პირობების გარდა სხვა შეზღუდვა არაა მოცემული, მაშინ ასეთ ალბათობას *უპირობოს* უწოდებენ; თუ რაიმე სხვა დამატებითი პირობებიც მოცემულია, მაშინ ხდომილების ალბათობას *პირობითს* უწოდებენ. მაგალითად, ხშირად ითვლიან B ხდომილების ალბათობას A ხდომილების მოხდენის დამატებით პირობებში. შევნიშნოთ, რომ უპირობო ალბათობაც, მკაცრად რომ ვთქვათ, პირობითია, რამდენადაც ნავარაუდევია S პირობების განხორციელება.

B ხდომილების პირობით ალბათობას A პირობით აღნიშნავენ $P_A(B)$ სიმბოლოთი და უწოდებენ ხდომილებას, რომელიც გამოითვლება იმის დაშვებით, რომ A ხდომილება უკვე დადგა.

მაგალითი. ურნაში 3 თეთრი და 3 შავი ბურთია. ურნიდან ორჯერ იღებენ ერთ ბურთს და უკან არ აბრუნებენ. იპოვეთ თეთრი ბურთის ამოღების ალბათობა მეორე

ცდაზე (B ხდომილება), თუ პირველ ცდაზე შავი ბურთი იქნა ამოღებული (A ხდომილება).

ამოხსნა. პირველი ცდის შემდეგ ურნაში დარჩა 5 ბურთი, მათგან 3 თეთრი. საძებნი პირობითი ალბთობა ტოლია

$$P_A(B) = 3/\,5.$$

ეს შედეგი შესაძლებელია მივიღოთ შემდეგი ფორმულით

$$P_A(B) = P(AB)/P(A) \quad (P(A) > 0). \tag{3.2.1}$$

მართლაც, პირველ ცდაში თეთრი ბურთის ამოღების ალბათობა ტოლია

$$P(A) = 3/6 = 1/2.$$

მოვძებნოთ ალბათობა $P(AB)$, რომ პირველ ცდაში გამოჩნდება შავი ბურთი, ხოლო მეორეში თეთრი. ორი ბურთის, არა აქვს მნიშვნელობა რა ფერის, ერთად ამოღების რაოდენობა შემდეგი წყობის ტოლია $A_6^2 = 6 \cdot 5 = 30$. ამ რაოდენობიდან AB ხდომილების ხელშემწყობია $3 \cdot 3 = 9$ შედეგი. შესაბამისად,

$$P(AB) = 9/30 = 3/10.$$

საძიებელი პირობითი ალბათობა ტოლია

$$P_A(B) = P(AB)/P(A) = (3/10)/(1/2) = 3/5\,.$$

როგორც ვხედავთ იგივე შედეგი მივიღეთ.

ალბათობის კლასიკური განსაზღვრებიდან გამომდინარე შესაძლებელია (3.2.1) ფორმულის დამტკიცება. ეს გარემოება შემდეგი ზოგადი განსაზღვრების საფუძველია (რომელიც მისაღებია არა მარტო კლასიკური განსაზღვრებისათვის).

B ხდომილების პირობითი ალბათობა იმ პირობებში, როდესაც A ხდომილება უკვე მოხდა, განსაზღვრების თანახმად ტოლია

$$P_A(B) = P(AB)/P(A), \quad (P(A) > 0).$$

3.3 ალბათობათა გამრავლების თეორემა

განვიხილოთ ორი ხდომილება: A და B; ვთქვათ ალბათობები $P(A)$ და $P_A(B)$ ცნობილია. როგორ მოვძებნოთ ამ ხდომილებების შეთავსების ალბათობა, ე.ი. იმის

ალბათობა, რომ გამოჩნდება როგორც A ხდომილება ისე B ხდომილება? ამაზე პასუხს იძლევა გამრავლების თეორემა.

თეორემა. *ორი ხდომილების ერთდროულად მოხდენის ალბათობა ერთ-ერთი მათგანის ალბათობის მეორის პირობით ალბათობაზე ნამრავლის ტოლია:*

$$P(AB) = P(A)P_A(B).$$

დამტკიცება. პირობითი ალბათობის განსაზღვრების თანახმად,

$$P_A(B) = P(AB)/P(A)$$

საიდანაც,

$$P(AB) = P(A)P_A(B). \qquad (3.3.1)$$

შენიშვნა. თუ (3.3.1) ფორმულას გამოვიყენებთ BA ხდომილებისათვის, მივიღებთ

$$P(BA) = P(B)P_B(A).$$

ან, რამდენადაც BA ხდომილება არ განსხვავდება AB ხდომილებისაგან

$$P(AB) = P(B)P_B(A). \qquad (3.3.2)$$

(3.3.1) და (3.3.2) ფორმულების შედარებით ვასკვნით შემდეგი ტოლობის სამართლიანობას

$$P(A)P_A(B) = P(B)P_B(A). \qquad (3.3.3)$$

შედეგი. *რამდენიმე ხდომილების ერთდროულად მოხდენის ალბათობა ერთ-ერთი მათგანის ალბათობის ყველა დანარჩენის პირობით ალბათობებზე ნამრავლის ტოლია, ამასთან ყველა მომდევნო ხდომილების ალბათობა გამოთვლილია იმ ვარაუდით, რომ ყველა წინა ხდომილება უკვე მოხდა:*

$$P(A_1A_2A_3 \dots A_n) = P(A_1)P_{A_1}(A_2)P_{A_1A_2}(A_3) \dots P_{A_1A_2\dots A_{n-1}}(A_n),$$

სადაც, $P_{A_1A_2\dots A_{n-1}}$ არის A_n ხდომილების ალბათობა, რომელიც გამოთვლილია იმის გათვალისწინებით, რომ $A_1, A_{2,}A_3, \dots, A_{n-1}$ ხდომილებები უკვე მოხდა. კერძოდ, სამი ხდომილებისათვის გვექნება

$$P(ABC) = P(A)P_A(B)P_{AB}(C).$$

შევნიშნოთ, რომ ხდომილებების განლაგების მიმდევრობა ნებისმიერად შეიძლება იყოს შერჩეული, ე.ი. სულ ერთია რომელს ჩავთვლით პირველად, რომელს მეორედ და ა.შ.

3.4. დამოუკიდებელი ხდომილებები. გამრავლების თეორემა დამოუკიდებელი ხდომილებებისათვის

ვთქვათ, B ხდომილება არაა დამოკიდებული A ხდომილებაზე.

B ხდომილებას უწოდებენ A ხდომილებისაგან დამოუკიდებელს, თუ A ხდომილების მოხდენა არ ცვლის B ხდომილების ალბათობას, ე.ი. თუ B ხდომილების პირობითი ალბათობა ტოლია მისი უპირობო ალბათობის:

$$P_A(B) = P(B). \tag{3.4.1}$$

თუ (3.3.3) ტოლობაში ჩავსვამთ (3.4.1) ტოლობას, მივიღებთ

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

აქედან

$$P_B(A) = P(A),$$

ე.ი. A ხდომილების პირობითი ალბათობა იმის გათვალისწინებით, რომ B ხდომილება უკვე მოხდა, ტოლია მისი უპირობო ალბათობის. სხვა სიტყვებით, A ხდომილება არაა დამოკიდებული B ხდომილებაზე.

ამგვარად, თუ B ხდომილება არაა დამოკიდებული A ხდომილებაზე, მაშინ არც A ხდომილებაა დამოკიდებული B ხდომილებაზე; ეს ნიშნავს, რომ ხდომილებათა *დამოუკიდებლობის თვისება ორმხრივია.*

დამოუკიდებელი ხდომილებებისათვის ხდომილებათა ნამრავლის თეორემას $P(AB) = P(A)P_A(B)$ აქვს შემდეგი სახე

$$P(AB) = P(A)P(B), \tag{3.4.2}$$

ე.ი. ორი დამოუკიდებლი ხდომილების ერთად მოხდენის ალბათობა ტოლია ამ ხდომილებათა ალბათობების ნამრავლის.

(3.4.2) ტოლობა მიიღება როგორც დამოუკიდებელი ხდომილებების განსაზღვრება.

ორ ხდომილებას უწოდებენ დამოუკიდებელს, თუ მათი ერთად მოხდენის ალბათობა ტოლია ამ ხდომილებების ალბათობათა ნამრავლის; წინააღმდეგ შემთხვევაში ხდომილებებს *დამოკიდებულებს* უწოდებენ.

პრაქტიკაში ხდომილებების დამოუკიდებლობას ამოცანის არსის მიხედვით ადგენენ. მაგალითად, ორი ქვემეხიდან თითოეულის მიზანში მოხვედრის ალბათობა

არაა დამოკიდებული მეორე ქვემეხიდან მიზანს მოხვდა თუ არა, ამიტომ ხდომილებები „პირველი ქვემეხიდან მიზანს მოხვდა“ და „მეორე ქვემეხიდან მიზანს მოხვდა“ დამოუკიდებლებია.

შენიშვნა 1. თუ ხდომილებები A და B დამოუკიდებლებია, მაშინ ასევე დამოუკიდებელია A და$\bar{B}$, $\bar{A}$და B, $\bar{A}$და$\bar{B}$. მართლაც

$$A = A\bar{B} + AB.$$

ამგვარად,

$P(A) = P(A\bar{B}) + P(AB)$, ან $P(A) = P(A\bar{B}) + P(A)P(\mathrm{B})$,

აქედან

$$P(A\bar{B}) = P(A)[1 - P(B)] \text{ ან } P(A\bar{B}) = P(A)P(\bar{B})$$

ე.ი. A და $\bar{B}$ დამოუკიდებლებია. $\bar{B}$, $\bar{A}$ და B, $\bar{A}$ ხდომილებების დამოუკიდებლობა - დამტკიცებული დებულების შედეგებია.

რამდენიმე ხდომილებას უწოდებენ წყვილ-წყვილად დამოუკიდებლს, თუ მათგან ყოველი განსხვავებული ორი ხდომილება დამოუკიდებელია. მაგალითად, A, B, C წყვილ-წყვილად დამოუკიდებელია, თუ ხდომილებები A და B, A და C , B და C დამოუკიდებლები არიან.

რამდენიმე ხდომილებისათვის ნამრავლის თეორემის განზოგადებისათვის შემოვიღოთ ერთობლიობაში ხდომილებათა დამოუკიდებლობის ცნება.

რამდენიმე ხდომილებას უწოდებენ დამოუკიდებელს ერთობლიობაში (ან უბრალოდ დამოუკიდებლებს), თუ მათგან ყოველი ორი დამოუკიდებელია და ასევე დამოუკიდებელია თითოეული ხდომილება და დანარჩენების ყველა შესაძლო ნამრავლი. მაგალითად, თუ A_1, A_2, A_3 დამოუკიდებლებია ერთობლიობაში, მაშინ დამოუკიდებლებია ხდომილებები A_1 და A_2 , A_1 და A_3 , A_2 და A_3 ; A_1 და A_2A_3, A_2 და $A_1A_{3,}, A_3$ და A_1A_2 ნათქვამიდან გამომდინარეობს, რომ თუ ხდომილებები ერთობლიობაში დამოუკიდებლები არიან, მაშინ მათგან ნებისმიერის მოხდენის პირობითი ალბათობა, რომელიც გამოთვლილია იმის ვარაუდით, რომ უკვე მოხდა რომელიმე სხვა ხდომილება დარჩენილებიდან, ტოლია მისი უპირობო ალბათობის.

შევნიშნოთ, რომ თუ რამდენიმე ხდომილება წყვილ-წყვილად დამოუკიდებელია, მაშინ აქედან არ გამომდინარეობს მათი ერთობლიობაში დამოუკიდებლობა. ამ აზრით ერთობლიობაში დამოუკიდებლობის მოთხოვნა უფრო ძლიერია წყვილ-წყვილად დამოუკიდებლობის მოთხოვნასთან შედარებით.

ზემოთ ნათქვამი ავხსნათ მაგალითის მეშვეობით. ვთქვათ, ურნაში 4 ბურთია, რომლებიდან: ერთი წითლადაა შეღებილი (A), ერთი - ლურჯად (B), ერთი - შავად (C) და ერთის შესაღებად სამივე ფერია გამოყენებული (ABC). რას უდრის იმის ალბათობა, რომ ურნიდან ამოღებულ ბურთს წითელი ფერი აქვს?

რამდენადაც ოთხი ბურთიდან ორს აქვს წითელი ფერი, ამდენად $P(A) = 2/4 = 1/2$. ახლა ვთქვათ ამოღებულ ბურთს აქვს ლურჯი ფერი, ე.ი. B ხდომილება უკვე მოხდა. შეიცვლება თუ არა იმის ალბათობა, რომ ამოღებულ ბურთს ექნება წითელი ფერი, ე.ი. შეიცვლება თუ არა A ხდომილების ალბათობა? ორი ბურთიდან, რომელსაც ლურჯი ფერი აქვს, ერთს წითელი ფერიც გააჩია, ამიტომ A ხდომილების ალბათობა ისევ $1/2$-ის ტოლია. სხვა სიტყვებით, A ხდომილების პირობითი ალბათობა, რომელიც გამოთვლილია იმის დაშვებით, რომ B ხდომილება უკვე მოხდა, ტოლია მისი უპირობო ალბათობის. აქედან გამომდინარე, A და B ხდომილებები დამოუკიდებელნი არიან. ანალოგიურად ვასკვნით, რომ A და C, B და C ხდომილებები დამოუკიდებლები არიან. ამგვარად, A, B, C წყვილ-წყვილად დამოუკიდებლებია.

ეს ხდომილებები ერთობლიობაში დამოუკიდებლები არიან? აღმოჩნდა რომ არა. მართლაც, ვთქვათ, აღნიშნულ ბურთს ორი ფერი აქვს, მაგალითად ლურჯი და შავი. როგორია იმის ალბათობა რომ ამოღებულ ბურთს წითელი ფერიც აქვს? მხოლოდ ერთი ბურთია სამ ფერად შეღებილი. ამგვარად, თუ დავუშვებთ, რომ B და C ხდომილებები უკვე მოხდა, მივდივართ დასკვნამდე, რომ A ხდომილება აუცილებლად დადგება. აქდან გამომდინარე ეს ხდომილება აუცილებელი ხდომილებაა და მისი ალბათობა ერთის ტოლია. სხვა სიტყვებით, A ხდომილების პირობითი ალბათობა $P_{BC}(A) = 1$ და, მაშასადამე, არ უდრის მის უპირობო ალბათობას $P(A) = 1/2$. ამგვარად, წყვილ-წყვილად დამოუკიდებელი A, B, C ხდომილებები არ წარმოადგენენ ერთობლიობაში დამოუკიდებლებს.

მოვიყვანოთ ალბათობათა ნამრავლის თეორემის შედეგი

შედეგი. *ერთობლიობაში დამოუკიდებელი რამდენიმე ხდომილების ერთად მოხდენის ალბათობა ამ ხდომილებების ალბათობათა ნამრავლის ტოლია*:

$$P(A_1A_2 \dots A_n) = P(A_1)P(A_2) \dots P(A_n).$$

დამტკიცება. განვიხილოთ სამი ხდომილება A, B და C. ცხადია შემდეგი ტოლობის სამართლიანობა

$$P(ABC) = P(AB \cdot C).$$

რამდენადაც A, B, C ხდომილებები ერთობლიობაში დამოუკიდებლებია, ამიტომ დამოუკიდებლებია კერძოდ, AB და C ხდომილებები, ასევე A და B ხდომილებებიც. ორი დამოუკიდებელი ხდომილებისათვის ალბათობათა ნამრავლის თეორემის მიხედვით გვაქვს:

$$P(AB \cdot C) = P(AB)P(C) \text{ და } P(AB) = P(A)P(B).$$

ამგვარად, საბოლოოდ მივიღებთ

$$P(ABC) = P(A)P(B)P(C).$$

ნებისმიერი n -სთვის დამტკიცება ხდება მათემატიკური ინდუქციის მეთოდის გამოყენებით.

შენიშვნა 2. თუ ხდომილებები $A_1, A_2, \dots, A_n$ დამოუკიდებლებია ერთობლიობაში, მაშინ მათი მოპირდაპირე ხდომილებებიც $\overline{A_1}, \overline{A_2}, \dots, \overline{A_n}$ ასევე ერთობლიობაში დამოუკიდებლებია.

3.5. თუნდაც ერთი ხდომილების მოხდენის ალბათობა

ვთქვათ, ცდის შედეგად შესაძლებლია ერთობლიობაში დამოუკიდებელი n ხდომილების ან ზოგიერთი მათგანის (კეძოდ, მხოლოდ ერთის ან არცერთის) მოხდენა, ამასთან, თითოეული ხდომილების მოხდენის ალბათობა ცნობილია. როგორ ვიპოვოთ იმის ალბათობა, რომ ამ ხდომილებებიდან ერთი მაინც მოხდება? მაგალითად, თუ ცდის შედეგად შესაძლებელია სამი ხდომილების მოხდენა, მაშინ ამ ხდომილებებიდან

თუნდაც ერთის მოხდენა ნიშნავს თუ არა ან ერთის, ან ორის, ან სამი ხდომილების მოხდენას? ამაზე პასუხს იძლევა შემდეგი თეორემა.

თეორემა. ერთობლიობაში დამოუკიდებელი $A_1, A_2, \dots, A_n$ ხდომილებებიდან თუნდაც ერთის მოხდენის ალბათობა 1-ის და მოპირდაპირე ხდომილებების $\overline{A_1}, \overline{A_2}, \dots, \overline{A_n}$ ალბათობათა ნამრავლის სხვაობის ტოლია:

$$P(A) = 1 - q_1 q_2 \dots q_n. \qquad (3.5.1)$$

დამტკიცება. შემოვიტანოთ აღნიშვნა:

$$A = A_1 + A_2 + \dots + A_n.$$

რადგან $\bar{A} = \overline{(A_1 + A_2 + \dots + A_n)} = \overline{A_1}\,\overline{A_2} \dots \overline{A_n}$, ამიტომ A და $\overline{A_1}\,\overline{A_2} \dots \overline{A_n}$ ხდომილებები მოპირდაპირე ხდომილებებია. აქედან გამომდინარე მათი ალბათობების ჯამი ერთის ტოლია:

$$P(A) + P(\overline{A_1}\,\overline{A_2} \dots \overline{A_n}) = 1.$$

აქედან, ნამრავლის თეორემის გამოყენებით, მივიღებთ

$$P(A) = 1 - P(\overline{A_1 A_2} \dots \overline{A_n}) = 1 - P(\bar{A}_1)P(\bar{A}_2) \dots P(\bar{A}_n),$$

ანუ

$$P(A) = 1 - q_1 q_2 \dots q_n.$$

კერძო შემთხვევა. თუ $A_1, A_2, \dots, A_n$ ხდომილებებს ერთი და იგივე q -ს ტოლი ალბათობები გააჩნიათ, მაშინ ალბთობა იმისა, რომ ამ ხდომილებებიდან თუნდაც ერთი მოხდება, ტოლია

$$P(A) = 1 - q^n. \qquad (3.5.2)$$

თავი 4. ჯამის და გამრავლების თეორემების შედეგი

4.1 თავსებადი ხდომილებების ალბათობათა შეკრების თეორემა

აქამდე ალბათობათა ჯამის თეორემა განვიხილეთ უთავსი ხდომილებებისათვის. ახლა განვიხილოთ ეს თეორემა თავსებადი ხდომილებებისათვის.

ცხადია, ორი ხდომილებაn თავსებადია, თუ ერთი და იმავე ცდაში ერთ-ერთის მოხდენა არ გამორიცხავს მეორის მოხდენას.

მაგალითი 1. A - კამათელის აგდებისას ოთხიანის მოსვლა; B- კამათლის აგდებისას ლუწი რიცხვის მოსვლა. ცხადია, A და B - თავსებადი ხდომილებებია.

ვთქვათ, მოცემულია A და B ხდომილებების ალბათობები და მათი ერთობლივი მოხდენის ალბათობა. როგორ ვიპოვოთ $A+B$ ხდომილების ალბათობა, ანუ A და B ხდომილებებიდან ერთ-ერთის მოხდენის ალბათობა? ამაზე პასუხს იძლევა ხდომილებათა ალბათობების ჯამის თეორემა.

თეორემა. *ორი ხდომილებიდან ერთ-ერთის მოხდენის ალბათობა ტოლია ამ ხდომილებათა ალბათობების ჯამს გამოკლებული მათი ერთობლივი მოხდენის ალბათობა:*

დამტკიცება. ცხადია, რომ $A+B$ ხდომილება ხდება მაშინ, როდესაც ხდება შემდეგი სამი უთავსადი ხდომილებიდან ერთ-ერთი: $A\bar{B}$, $\bar{A}B$ ან AB. უთავსი ხდომილებების ალბათობათა შეკრების თეორემის თანახმად გვაქვს,

$$P(A+B)=P(A\bar{B})+P(\bar{A}B)+P(AB). \tag{4.1.1}$$

A ხდომილება მოხდება, თუ მოხდება ორი უთავსადი ხდომილებიდან ერთ-ერთი: $A\bar{B}$ ან AB . უთავსადი ხდომილებების შეკრების თეორემიდან გამომდინარე გვაქვს

$$P(A)=P(A\bar{B})+P(AB).$$

აქედან

$$P(A\bar{B})=P(A)-P(AB). \tag{4.1.2}$$

ანალოგიურად გვექნება

$$P(B)=P(\bar{A}B)+P(AB).$$

საიდანაც

$$P(\bar{A}B)=P(B)-P(AB). \tag{4.1.3}$$

თუ (4,1,2) და(4.1.3) ტოლობებს ჩავსმათ (4.1.1)-ში, საბოლოოდ მივიღებთ

$$P(A+B)=P(A)+P(B)-P(AB). \qquad (4.1.4)$$

შენიშვნა 1. მიღებული ფორმულის გამოყენებისას მხდველობაში უნდა მივიღოთ, რომ A და B ხდომილებები შეიძლება იყოს როგორც დამოუკიდებლები, ისე დამოკიდებულებიც.

დამოუკიდებელი ხდომილებისათვის

$$P(A+B)=P(A)+P(B)-P(A)P(B);$$

დამოკიდებული ხდომილებებისათვის კი

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

შენიშვნა 2. თუ A და B ხდომილებები უთავსადია, მაშინ მათი ნამრავლი შეუძლებელი ხდომილებაა და აქედან გამომდინარე, $P(AB)=0$.

ფორმულა (4.1.4) უთავსი ხდომილებებისათვის მიიღებს შემდეგ სახეს

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

ისევ მივიღეთ შეკრების თეორემა უთავსადი ხდომილებებისათვის. ამგვარად, ფორმულა (4.1.4) სამართლიანია როგორც თავსებადი, ისე უთავსადი ხდომილებებისათვის.

4.2. სრული ალბათობის ფორმულა

ვთქვათ, მოცემულია A ხდომილება და ხდომილებათა სრული სისტემა $B_1, B_2, \dots, B_n$. ვთქვათ, ცნობილია ამ ხდომილებების ალბათობები და A ხდომილების პირობითი ალბათობები $P_{B_1}(A), P_{B_2,}(A), \dots, P_{B_n}(A)$. როგორ ვიპოვოთ A ხდომილების ალბათობა? ამ კითხვას პასუხობს შემდეგი თეორემა.

თეორემა. *A ხდომილების ალბათობა ტოლია:*

$$P(A)=P(B_1)P_{B_1}(A)+P(B_2)P_{B_2}(A)+\cdots+P(B_n)P_{B_n}(A).$$

ამ ფორმულას უწოდებენ „სრული ალბათობის ფორმულას“.

დამტკიცება. რადგან ხდომილებები $B_1, B_2, \dots, B_n$ წარმოადგენენ ხდომილებათა სრულ სისტემას, ამიტომ, ცხადია, სამართლიანია შემდეგი ტოლობა:

$$A = B_1A + B_2A + \cdots + B_nA.$$

ახლა გამოვიყენოთ ხდომილებათა ჯამის თეორემა. მივიღებთ

$$P(A) = P(B_1A) + P(B_2A) + \cdots + P(B_nA). \tag{4.2.1}$$

გვრჩება გამოვთვალოთ თითოეული შესაკრები. ხდომილებათა ნამრავლის ალბათობების თეორემის თანახმად გვაქვს

$$P(B_1A) = P(B_1)P_{B_1}(A); \quad P(B_2A) = P(B_2)P_{B_2}(A); \ldots; P(B_nA) = P(B_n)P_{B_n}(A).$$

თუ ამ ტოლობების პირველ ნაწილს ჩავსამთ (4.2.1) ტოლობაში, მივიღებთ სრული ალბათობის ფორმულას

$$P(A) = P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A) + \cdots + P(B_n)P_{B_n}(A).$$

4.3. ჰიპოთეზების ალბათობა. ბაიესის ფორმულები

ვთქვათ, მოცემულია A ხდომილება და ხდომილებათა სრული სისტემა $B_1, B_2, \ldots, B_n$. ხდომილებებს $B_1, B_2, \ldots, B_n$ ხშირად ჰიპოთეზებსაც უწოდებენ. სრული ალბათობის ფორმულის მიხედვით A ხდომილების ალბათობა გამოითვლება შემდეგნაირად:

$$P(A) = P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A) + \cdots + P(B_n)P_{B_n}(A). \tag{4.3.1}$$

ვთქვათ, ჩატარდა ცდა, რომლის შედეგადაც მოხდა A ხდომილება. განვსაზღვროთ, როგორ შეიცვალა (იმასთან დაკავშირებით, რომ A ხდომილება უკვე მოხდა) ჰიპოთეზების ალბათობები. ანუ, მოვძებნოთ პირობითი ალბათობები

$$P_A(B_1), P_A(B_2), \ldots, P_A(B_n).$$

თავდაპირველად მოვძებნოთ $P_A(B_1)$ პირობითი ალბათობა. ალბათობათა ნამრავლის თეორემის თანახმად გვაქვს

$$P(AB_1) = P(A)P_A(B_1) = P(B_1)P_{B_1}(A).$$

აქედან

$$P_A(B_1) = \frac{P(B_1)P_{B_1}(A)}{P(A)}$$

თუ $P(A)$-ს შევცვლით (4.3.1) ფორმულით, მივიღებთ

$$P_A(B_1) = \frac{P(B_1)P_{B_1}(A)}{P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A) + \cdots + P(B_n)P_{B_n}(A)}.$$

ანალოგიურად გამოიყვანება ფორმულები, რომლებიც განსაზღვრავენ დანარჩენი ჰიპოთეზების ალბათობებს, ე.ი. ნებისმიერი B_i $(i = 1,2, \dots, n)$ ჰიპოთეზის პირობითი ალბათობა შიძლება გამოითვალოს შემდეგი ფორმულით

$$P_A(B_i) = \frac{P(B_i)P_{B_i}(A)}{P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A) + \dots + P(B_n)P_{B_n}(A)}.$$

მიღებულ ფორმულებს უწოდებენ *ბაიესის ფორმულებს,* *ბაიესის ფორმულები საშუალებას იძლევა გადავაფასოთ ჰიპოთეზების ალბათობები იმის შემდეგ, რაც ცნობილი ხდება ცდის შედეგი, რომლის შედეგადაც მოხდა* A *ხდომილება.*

თავი 5. განმეორებითი ცდები

5.1 ბერნულის ფორმულა

თუ ტარდება რამდენიმე ცდა, ამასთან A ხდომილების ალბათობა თითოეულ ცდაში არაა დამოკიდებული სხვა ცდების შედეგებზე, ასეთ ცდებს უწოდებენ A ხდომილების მიმართ დამოუკიდებელს.

სხვადასხვა დამოუკიდებლ ცდაში A ხდომილებას შეიძლება ჰქონდეს ან განსხვავებული, ან ერთი და იგივე ალბათობა. შემდგომში განვიხილავთ მხოლოდ ისეთ დამოუკიდებელ ცდებს, რომლებშიც A ხდომილებას აქვს ერთი და იგივე ალბათობა.

ვთქვათ ტარდება n დამოუკიდებელი ცდა, რომელთაგან თითოეულში A ხდომილება ან ხდება ან არა. ჩათვალოთ, რომ A ხდომილების ალბათობა თითოეულ ცდაში ერთი და იგივეა და p-ს ტოლია. აქედან გამომდინარე, A ხდომილების არმოხდენის ალბათობა თითოეულ ცდაში მუდმივია და $q = 1 - p$-ს ტოლია.

გამოვთვალოთ იმის ალბათობა, რომ n ცდის ჩატარების დროს A ხდომილება მოხდება ზუსტად k-ჯერ და შესაბამისად არ მოხდება $n - k$-ჯერ. მნიშვნელივანია აღინიშნოს, რომ არ მოითხოვება A ხდომილება მეორდებოდეს ზუსტად k-ჯერ განსაზღვრული მიმდევრობით. მაგალითად, თუ საუბარია ოთხ ცდაში A ხდომილების სამჯერ მოხდენაზე, შესაძლებელია A ხდომილება მოხდეს შემდეგი თანმიმდევრობებით: $AAA\bar{A}$, $AA\bar{A}A$, $A\bar{A}AA$, $\bar{A}AAA$. ჩანაწერი $AAA\bar{A}$ ნიშნავს, რომ პირველ, მეორე და მესამე ცდაში A ხდომილება მოხდა, ხოლო მეოთხე ცდაში ის არ მოხდა, ანუ მოხდა მოპირდაპირე $\bar{A}$ ხდომილება; შესაბამისი შინაარსი გააჩნიათ სხვა ჩანაწერებსაც.

საძიებელი ალბათობა აღვნიშნოთ $P_n(k)$-ით. მაგალითად, $P_5(3)$ ნიშნავს, რომ ხუთ ჩატარებულ ცდაში ხდომილება A მოხდა 3-ჯერ და, შესაბამისად, არ მოხდა 2-ჯერ.

დასმული ამოცანის ამოხსნა შესაძლებელია ე.წ. ბერნულის ფორმულის მეშვეობით.

ბერნულის ფორმულის გამოყვანა. იმ ხდომილების ალბათობა, როდესაც n ჩატარებული ცდის დროს A ხდომილება მოხდება k-ჯერ და არ მოხდება $n - k$-ჯერ, დამოუკიდებელი ხდომილებების ალბათობების გამრავლების თეორემის თანახმად

ტოლია $p^k q^{n-k}$. ასეთი ხდომილება შეიძლება იყოს იმდენი, რამდენი წყობაც შეიძლება შევადგინოთ n ელემენტისაგან k ელემენტად, ე.ი. C_n^k. რამდენადაც ეს ხდომილებები უთავსადია, ამდენად, უთავსადი ხდომილებათა ალბათობების შეკრების თეორემის თანახმად, საძიებელი ალბათობა ტოლია ყველა ასეთი ხდომილებების ალბათობათა ჯამის. რადგან ამ ხდომილებების ალბათობები ერთი და იგივეა, ამიტომ საძიებელი ალბათობა (n ცდის დროს A ხდომილების k-ჯერ მოხდენა) ტოლია:

$$P_n(k) = C_n^k p^k q^{n-k}.$$

ან

$$P_n(k) = \frac{n!}{k!\,(n-k)!} p^k q^{n-k}.$$

მიღებულ ფორმულას *ბერნულის ფორმულას* უწოდებენ.

5.2. ლაპლასის ლოკალური თეორემა

ზემოთ გამოვიყვანეთ ბერნულის ფორმულა, რომელიც საშუალებას იძლევა გამოვთვალოთ იმის ალბათობა, რომ ხდომილება n ცდის ჩატარების დროს ზუსტად k-ჯერ მოხდება. გამოყვანის დროს ჩვენ ვვარაუდობდით, რომ თითოეულ ჩატარებულ ცდაში ხდომილების მოხდენის ალბათობა ერთი და იგივეა. ადვილი დასანახია, რომ ბერნულის ფორმულით სარგებლობა n ის დიდი მნიშვნელობებისათვის საკმაოდ რთულია, რამდენადაც ფორმულა მოითხოვს დიდ რიცხვებზე მოქმედებებს. მართალია, შესაძლებელია რამდენადმე გამარტივდეს გამოთვლები ცხრილების გამოყენებით, მაგრამ ეს გზა მაინც რჩება მოუქნელი.

ბუნებრივად ისმება შეკითხვა: შესაძლებელია თუ არა ჩვენთვის საინტერესო ალბათობის გამოთვლა ბერნულის ფორმულის გარეშე? აღმოჩნდა, რომ შეიძლება. ლაპლასის ლოკალური თეორემა გვაძლევს ასიმპტოტურ [1] ფორმულას, რომელიც

[1] $\varphi(x)$ ფუნქციას უწოდებენ $f(x)$ ფუნქციასთან ასიმპტოტურად მიახლოებულს, თუ $\lim_{x\to\infty}\frac{f(x)}{\varphi(x)} = 1$.

საშუალებას იძლევა მიახლოებით ვიპოვოთ ხდომილების ზუსტად k-ჯერ მოხდენის ალბათობა n ცდის ჩატარების დროს, თუ ცდების რაოდენობა n საკმაოდ დიდი რიცხვია.

ჩამოვაყალიბოთ ლაპლასის ლოკალური თეორემა დამტკიცების გარეშე.

ლაპლასის ლოკალური თეორემა. *თუ A ხდომილების მოხდენის ალბათობა თითოეულ ცდაში მუდმივია და განსხვავდება ნულისგან და ერთისაგან, მაშინ $P_n(k)$ ალბათობა იმისა, რომ A ხდომილება n ცდის ჩატარების დროს ზუსტად k-ჯერ მოხდება, მიახლოებით (რაც უფრო დიდი n, მით უფრო მეტი სიზუსტით) გამოითვლება შემდეგი ფორმულით:*

$$P_n(k) \approx \frac{1}{\sqrt{npq}} \cdot \varphi(x),$$

სადაც $x = (k - np)/\sqrt{npq}$, ხოლო $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

არსებობს ცხრილები, რომლებშიც მოცემულია $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ ფუნქციის მნიშვნელობები, რომლებიც შეესაბამება x არგუმენტის დადებით მნიშვნელობებს. არგუმენტის უარყოფითი მნიშვნელობებისათვის იმავე ცხრილებით სარგებლობენ, რამდენადაც $\varphi(x)$ ფუნქცია ლუწია, ე.ი. $\varphi(-x) = \varphi(x)$.

შემთხვევითი სიდიდეები

თავი 6. შემთხვევითი სიდიდეების სახეები. დისკრეტული შემთხვევითი სიდიდის მოცემა

6.1. შემთხვევითი სიდიდე

ვთქვათ, ვაგორებთ კამათელს. გაგორების შედეგად გამოჩნდება 1, 2, 3, 4, 5 და 6 ცოფრებიდან ერთ-ერთი, თუ კამათელს რამდენიმეჯერ გავაგორებთ, მოსული ციფრების რაოდენობას წინასწარ ვერ განვსაზღვრავთ, რადგან ის დამოკიდებულია მრავალ შემთხვევით მიზეზე, რომელთა სრულად გათვალისწინება შეუძლებელია. ამ თავლსაზრისით კამათლის გაგორებისას მოსულიციფრების რაოდენობა შემთხვევითი სიდიდეა; ციფრები 1, 2, 3, 4, 5 და 6 ამ შემთხვევითი სიდიდის შესაძლო მნიშვნელობებია.

შემთხვევითს უწოდებენ სიდიდეს, რომელიც შემთხვევითი ცდის შედეგებზე დაყრდნობით ღებულობს სხვადასხვა რიცხვით მნიშვნელობებს.

მაგალითი 1. ას ახალშობილს შორის დაბადებული ვაჟების რაოდენობა არის შემთხვევითი სიდიდე, რომელსაც გააჩნია შემდეგი შესაძლო მნიშვნელობები: 0, 1, 2, ..., 100.

მაგალითი 2. მანძილი რომელსაც გაიფრენს ჭურვი ქვემეხიდან გასროლისას არის შემთხვევითი სიდიდე. მართლაც, მანძილი დამოკიდებულია არა მარტო მიზნის დაყენებაზე, არამედ სხვა მრავალ მიზეზეც (ქარის ძალაზე და მიმართულებაზე, ტემპერატურაზე და ა.შ.), რომელთა სრულად გათვალისწინება შეუძლებელია. ასეთი სიდიდის შესაძლო მნიშვნელობები მიეკუთვნება რომელიღაც (a, b) რიცხვით შუალედს.

შემდგომში შემთხვევით სიდიდეებს დიდი ლათინური ასოებით $X, Y, Z,$ აღვნიშნავთ, ხოლო მათ შესაძლო მნიშვნელობებს - შესაბამისი პატარა ასოებით $x, y\, z$. მაგალითად, თუ X შემთხვევით სიდიდეს გააჩნია სამი შესაძლო შემთხვევა, მაშინ მათ ასე ავღნიშნავთ: x_1, x_2, x_3.

6.2 დისკრეტული და უწყვეტი შემთხვევითი სიდიდეები

დავუბრუნდეთ ზემოთ მოყვანილ მაგალითებს. პირველ შემთხვევაში X შემთხვევით სიდიდეს შეეძლო მიეღო შემდეგი შესაძლო მნიშვნელობებიდან ერთ-ერთი: 0, 1, 2, ..., 100. ეს სიდიდეები ერთმანეთისაგან შუალედებითაა გამოყოფილი, რომელშიც არაა X-ის შესაძლო მნიშვნელობები. ამგვარად, ამ მაგალითში შემთხვევითი სიდიდე იღებს ცალკეულ იზოლირებულ შესაძლო მნიშვნელობებს. მეორე მაგალითში შემთხვევით სიდიდეს შეეძლო მიეღო (a, b) შუალედიდან ნებისმიერი მნიშვნელობა. აქ ვერ გამოვყოფთ ერთ შესაძლო მნიშვნელობას მეორესაგან შუალდით, რომელიც არ შეიცავს შემთხვევითი სიდიდის შესაძლო მნიშვნელობას.

ნათქვამიდან უკვე შეიძლება დავასკვნათ, რომ მიზანშეწონილია განვასხვავოთ ის შემთხვევითი სიდიდეები, რომლებიც იღებენ ცალკეულ იზოლირებულ მნიშვნელობას, და ის შეთხვევითი სიდიდეები, რომელთა შესაძლო მნიშვნელობები მთლიანად ავსებენ რაიმე შუალედს.

დისკრეტული ეწოდება შემთხვევით სიდიდეს, რომელიც იღებს ცალკეულ, იზოლირებულ შესაძლო მნიშვნელობებს განსაზღვრული ალბათობებით. დისკრეტული სიდიდის შესაძლო მნიშვნელობათა რაოდენობა შეიძლება იყოს სასრული და უსარულო.

უწყვეტი ეწოდება შემთხვევით სიდიდეს, რომელსაც შეუძლია მიიღოს რაიმე სასრული ან უსასრულო შუალედიდან ყველა მნიშვნელობა. ცხადია, უწყვეტი სიდიდის შესაძლო მნიშვნელობათა რაოდენობა უსასრულოა.

შენიშვნა. უწყვეტი შემთხვევითი სიდიდის ეს განსაზღვრება არ წარმოადგენს ზუსტს.

6.3. დისკრეტული შემთხვევითი სიდიდის განაწილების კანონი

ერთი შეხედვით შეიძლება მოგვეჩვენოს, რომ დისკრეტული სიდიდის მოცემისათვის საკმარისია ჩამოვთვალოთ ყველა მისი შესაძლო მნიშვნელობა.

სინამდვილეში ეს ასე არაა: შემთხვევით სიდიდეებს შეიძლება ჰქონდეთ შესაძლო შემთხვევების *ერთნაირი* ჩამონათვალი, ალბათობები კი - *სხვადასხვა.* ამიტომ დისკრეტული შემთხვევითი სიდიდის მოცემისათვის არაა საკმარისი მისი ყველა შესაძლო მნიშვნელობების ჩამოთვლა, საჭიროა ასევე მივუთითოთ მათი ალბათობები.

დისკრეტული შემთხვევითი სიდიდის განაწილების კანონს უწოდებენ შესაბამისობას მის შესაძლო მნიშვნელობებსა და მათ ალბთობებს შორის; ის შეიძლება მოიცეს ცხრილურად, ანალიზურად (ფორმულის სახით) და გრაფიკულად.

დისკრეტული შემთხვევითი სიდიდის განაწილების კანონის ცხრილურად მოცემის დროს ცხრილის პირველი ხაზი შეიცავს შესაძლო მნიშვნელობებს, ხოლო მეორე - მათ ალბათობებს:

$$X \quad x_1 x_2 \ \ \ldots. \ x_n$$
$$p \quad p_1 p_2 \ \cdots \ p_n.$$

მივიღებთ რა მხედველობაში, რომ ერთ ცდაში შემთხვევითი სიდიდე იღებს ერთ და მხოლოდ ერთ შესაძლო მნიშვნელობას, ვასკვნით, რომ ხდომილებები $X = x_1$, $X = = x_2$, ... , $X = x_n$ ქმნიან ხდომილებათა სრულ სისტემას; შესაბამისად, ამ ხდომილებათა ალბათობების ჯამი, ე.ი. ცხრილის მეორე ხაზის ალბათობების ჯამი, ერთის ტოლია:

$$p_1 + p_2 + \cdots + p_n = 1.$$

თუ X-ის შესაძლო მნიშვნელობათა სიმრავლე უსასრულოა, მაშინ მწკრივი $p_1 + p_2 + \cdots$ კრებადია და მისი ჯამი ერთის ტოლია.

დისკრეტული შემთხვევითი სიდიდის განაწილების კანონის თავლსაჩი-ნოებისთვის ის შემლება გამოვსახთ გრაფიკულადაც, ამისათვის კოორდინატთა მართკუთხა სისტემაში აგებენ $(x_i, \ p_i)$ წერტილებს, ხოლო შემდეგ მათ აერთებენ წრფის მონაკვეთებით. მიღებულ ფიგურას უწოდებენ *განაწილების მრავალკუთხედს.*

6.4. ბინომური განაწილება

ვთქვათ ხორციელდება n დამოუკიდებელი ცდა, რომელთაგან თითოეულში A ხდომილება შეიძლება მოხდეს ან არ მოხდეს. ყველა ცდაში ხდომილების მოხდენის

ალბათობა მუდმივია და p-ს უდრის (შესაბამისად არ მოხდენის ალბათობა ტოლია $q = 1 - p$). X დისკრეტული შემთხვევითი სიდიდის სახით განვიხილოთ ამ ცდებში A ხდომილების მოხდენის რაოდენობა.

ამოცანად დავისახოთ X სიდიდის განაწილების კანონის მოძებნა. ამ ამოცანის გადასაწყვეტად საჭიროა განვსაზღვროთ X -ის შესაძლო მნიშვნელობები და მათი ალბათობები. ცხადია, A ხდომილება n ცდაში შეიძლება არ მოხდეს, ან მოხდეს 1-ჯერ, ან 2-ჯერ, ..., ან n-ჯერ. ამგვარად, X -ის შესაძლო მნიშვნელობები ასეთია: $x_1 == 0$, $x_2 = 1$, $x_3 = 2, \dots, x_{n+1} = n$. რჩება ვიპოვოთ ამ შესაძლო მნიშვნელობების ალბათობები, ამისათვის საკმარისია გამოვიყენოთ ბერნულის ფორმულა:

$$P_n(k) = C_n^k p^k q^{n-k}, \tag{6.4.1}$$

სადაც $k = 0, 1, 2, \dots, n$.

სწორედ (6.4.1) ფორმულა წარმოადგენს საძიებელი განაწილების კანონის ანალიზურ გამოსახულებას.

ბინომური ეწოდება ბერნულის ფორმულით განსაზღვრულ ალბათობათა განაწილებას. კანონს ეწოდება „ბინომური“ იმოტომ, რომ (6.4.1) ტოლობის მარჯვენა მხარე შეიძლება განვიხილოთ, როგორც ნიუტონის ბინომის გაშლის ზოგადი წევრი:

$$(p+q)^n = C_n^n p^n + C_n^{n-1} p^{n-1} q + \dots + C_n^k p^k q^{n-k} + \dots + C_n^0 q^n.$$

ამგვარად, გაშლის პირველი წევრი p^n განსაზღვრავს განხილული ხდომილების n დამოუკიდებელ ცდაში n -ჯერ დადგომის ალბათობას; მეორე წევრი $np^{n-1}q$ განსაზღვრავს ხდომილების $n-1$ -ჯერ დადგომის ალბათობას; ... ; ბოლო q^n წევრი განსაზღვრავს იმის ალბათობას, რომ ხდომილება არც ერთხელ არ დადგება.

$$\begin{array}{llllll} X & n & n-1 & \dots & k & \dots \quad 0 \\ P & p^n & np^{n-1}q & \dots & C_n^k p^k q^{n-k} & \dots q^n. \end{array}$$

6.5. პუასონის განაწილება

ვთქვათ, ტარდება n დამოუკიდებელი ცდა, რომელთაგან თითოეულში Aხდომილების მოხდენის ალბათობა p -ს ტოლია. იმისათვის, რომ განვსაზღვროთ ამ ცდებში A ხდომილების k -ჯერ მოხდენის ალბათობა, გამოიყენება ბერნულის

ფორმულა. თუ n დიდია, მაშინ იყენებენ ლაპლასის ასიმპტოტურ ფორმულას. მაგრამ ეს ფორმულა გამოუსადეგარია, თუ ხდომილების ალბათობა მცირეა ($p \leq 0{,}1$). ამ შემთხვევაში (n დიდია, p მცირეა) იყენებენ პუასონის ასიმპტოტურ ფორმულას.

ამგვარად, დავსვათ ამოცანა, ვიპოვოთ იმის ალბათობა, რომ ცდების დიდი რაოდენობის შემთხვევაში, რომელთაგან თითოეულში ხდომილების ალბათობა ძალიან მცირეა, ხდომილება ზუსტად k-ჯერ დგება. გავაკეთოთ მნიშვნელოვანი დაშვება: np ნამრავლი ინარჩუნებს მუდმივ მნიშვნელობას, კერძოდ, $np = \lambda$. როგორც შემდგომში აღვნიშნავთ, ეს ნიშნავს, რომ ხდომილების მოხდენის საშუალო რაოდენობა ცდების სხვადასხვა სერიებში, ე.ი. n-ის სხვადასხვა მნიშვნელობისათვის, უცვლელი რჩება.

გამოვიყენოთ ბერნულის ფორმულა ჩვენთვის საინტერესო ალბათობის გამოსათვლელად:

$$P_n(k) = \frac{n(n-1)(n-2)\ldots[n-(k-1)]}{k!}p^k(1-p)^{n-k}$$

რამდენადაც, $pn = \lambda$, მაშინ $p = \lambda/n$. ამგვარად,

$$P_n(k) = \frac{n(n-1)(n-2)\ldots[n-(k-1)]}{k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}.$$

თუ მხედველობაში მივიღებთ, რომ n-ის მნიშვნელობა ძალიან დიდია, $P_n(k)$-ის ნაცვლად ვიპოვოთ ზღვარი $\lim\limits_{n\to\infty} P_n(k)$. შევნიშნოთ, რომ რამდენადაც np ნამრავლი ინარჩუნებს მუდმივ მნიშვნელობას, ამიტომ $n \to \infty$ -სთვის, ცხადია, ალბათობა $p \to 0$.

ამრიგად,

$$\begin{aligned}P_n(k) \cong \lim_{n\to\infty}\frac{n(n-1)(n-2)\ldots[n-(k-1)]}{k!}\cdot\frac{\lambda^k}{n^k}\left(1-\frac{\lambda}{n}\right)^{n-k} &= \\ = \frac{\lambda^k}{k!}\lim_{n\to\infty}\left[1\cdot\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{k-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n-k}\right] &= \\ = \frac{\lambda^k}{k!}\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n\cdot\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-k} = \frac{\lambda^k}{k!}e^{-\lambda}\cdot 1.&\end{aligned}$$

ამგვარად (ჩანაწერის სიმარტივისათვის მიახლოების ნიშანი გამოტოვებულია)

$$P_n(k) = \lambda^k e^{-\lambda}/k!.$$

ეს ფორმულა გამოხატავს მასაბრივი (n ძალიან დიდია) და იშვიათი (p მცირეა) ხდომილობების ალბათობების პუასონის განაწილების კანონს.

6.6. ხდომილებების უმარტივესი ნაკადი

განვიხილოთ ხდომილებები, რომლებიც დგება დროის შემთხვევით მომენტებში.

ხდომილებათა ნაკადი ეწოდება ხდომილებათა მიმდევრობებს, რომლებიც ხდებიან დროის შემთხვევით მომენტებში. ნაკადების მაგალითებია: სატელეფონო სადგურში დროის გარკვეულ მონაკვეთში შესული სატელეფონო გამოძახებების რაოდენობა, სასწრაფო-სამედიცინო დახმარების პუნქტში შესული გამოძახებების რაოდენობა, აეროპორტში თვითმფრინავების დაფრენა, საყოფაცხოვრებო მომსახურების საწარმოებში კლიენტების მოსვლა და სხვა მრავალი.

იმ თვისებებიდან, რომლებიც შეიძლება გააჩნდეს ნაკადს, გამოვყოთ სტაციონარობის, შედეგიის და ორდინარულობის არქონის თვისება.

სტაციონარობის თვისება იმით ხასიათდება, რომ k რაოდენობის ხდომილებების მოხდენის ალბათობა დროის ნებისმიერ შუალედში დამოკიდებულია მხოლოდ k რაოდენობაზე და t შუალედის ხანგრძლივობაზე და არაა დამოკიდებული მის დაწყებაზე; ამასთან ნავარაუდევია, რომ დროის განსხვავებული შუალედები არ იკვეთებიან. მაგალითად, k ხდომილებების მოხდენის ალბათობა დროის შემდეგ შუალედებში (1; 7), (10, 16) (ერთნაირი $t = 6$ დრ. ერთეული) ერთმანეთის ტოლია.

ამგრიგად, *თუ ნაკადს გააჩნია სტაციონარობის თვისება, მაშინ k რაოდენობის ხდომილებების მოხდენის ალბათობა დროის t ხანგრძლივობის შუალედში არის ფუნქცია, რომელიც მხოლოდ k და t-ზეა დამოკიდებული.*

შედეგის არქონის თვისება იმით ხასიათდება, რომ k ხდომილებების მოხდენის ალბათობა დროის ნებისმიერ შუალედში არაა დამოკიდებული იმაზე, მოხდნენ თუ არ მოხდნენ ხდომილებები განხილვის შუალედის დაწყებამდე დროის მომენტში. სხვა სიტყვებით, k ხდომილებების მოხდენის პირობითი ალბათობა დროის ნებისმიერ შუალედში, რომელიც გამოთვლილია განხილული შუალედის დაწყებამდე რაიმეს მოხდენის ნებისმიერი დაშვების გარეშე (რამდენი ხდომილება გამოჩნდა, როგორი მიმდევრობით), ტოლია უპირობო ალბათობის. ამრიგად, ნაკადის პრეისტორია არ აისახება უახლოეს მომავალში ხდომილების მოხდენის ალბათობაზე.

ამგვარად, თუ ნაკადს გააჩნია *შედეგის არ ქონის* თვისება, მაშინ ადგილი აქვს დროის არაგადამკვეთ შუალედებში ამა თუ იმ რაოდენობის ხდომილებების მოხდენის ურთიერთდამოუკიდებლობას.

ორდინარულობის თვისება ხასითდება იმით, რომ ორი ან მეტი ხდომილების მოხდენა დროის მცირე შუალედში პრქტიკულად შეუძლებელია. სხვა სიტყვებით, ერთზე მეტი მოვლენის მოხდენის ალბათობა უმნიშვნელოა მხოლოდ ერთი მოვლენის მოხდენის ალბათობასთან შედარებით.

ამრიგად, თუ ნაკადს გააჩნია ორდინარულობის თვისება, მაშინ დროის უსასრულოდ მცირე შუალედში შეიძლება გამოჩნდეს არა უმეტეს ერთი ხდომილებისა.

უმარტივეს (პუასონურს) უწოდებენ სტაციონალურ და ორდინარულ ხდომილებათა ნაკადს, რომელსაც გააჩნია შედეგის არ ქონის თვისება.

შენიშვნა. ხშირად პრაქტიკაში ძნელია იმის დადგენა, გააჩნია თუ არა ნაკადს ზემოთ აღნიშნული თვისებები. ამიტომ ნაპოვნია სხვა პირობები, რომელთა დაცვით ნაკადი შეიძლება ჩაითვალოს უმარტივესად ან უმარტივესთან ახლოს. კერძოდ, დადგენილია, თუ ნაკადი წარმოადგენს დამოუკიდებელ სტაციონალურ ნაკადების ძალიან დიდი რაოდენობის ჯამს, რომელთაგან თითოეულის გავლენა მთელ ჯამზე (ჯამურ ნაკადზე) ძალიან მცირეა, მაშინ ჯამური ნაკადი (მისი ორდინარობის პირობებში) ახლოსაა უმარტივესთან.

λ ნაკადის ინტენსივობას უწოდებენ ხდომილებათა რაოდენობის საშუალოს, რომლებიც ჩნდებიან დროის ერთეულში

შეიძლება დამტკიცდეს, რომ თუ ნაკადის მუდმივი ინტენსივობა ცნობილია, მაშინ უმარტივესი ნაკადის k ხდომილებების მოხდენის ალბათობა დროის t ხანგრძლივობისას განისაზღვრება პუასონის ფორმულით

$$P_t(k) = (\lambda t)^k \cdot e^{-\lambda t}/k!$$

ეს ფორმულა გამოსახავს უმარტივესი ნაკადის ყველა თვისებას.

მართლაც, ფორმულიდან ჩანს, რომ t დროის განმავლობაში k ხდომილებების მოხდენის ალბათობა მოცემული ინტენსივობის დროს წარმოადგენს k და t-ს ფუნქციას, რაც ახასიათებს სტაციონალურობის თვისებას.

ფორმულა არ იყენებს ინფორმაციას განხილულ შუალედამდე ხდომილების მოხდენის შესახებ, რაც ახასიათებს შედეგის არ ქონის თვისებას.

დავრწმუნდეთ, რომ ფორმულა ასახავს ორდინარობის თვისებას. მივანიჭოთ მნიშვნელობები $\mathrm{k} = 0$ და $\mathrm{k} = 1$, და ვიპოვოთ შესაბამისად ხდომილების არმოხდენის და ერთი ხდომილების მოხდენის ალბათობა:

$P_t(0) = e^{-\lambda t}, P_t(1) = \lambda t e^{-\lambda t}$.

აქდან გამომდინარე, ერთზე მეტი ხდომილების მოხდენის ალბათობა ტოლია

$P_t(k > 1) = 1 - [P_t(0) + P_t(1)] = 1 - [e^{-\lambda t} + \lambda t e^{-\lambda t}]$.

მივიღებთ

$$e^{-\lambda t} = 1 - \lambda t + (\lambda t)^2/2! - \cdots.$$

ელემენტარული გარდაქმნებით მივიღებთ

$$P_t(k > 1) = (\lambda t)^2/2 + \cdots.$$

$P_t(1)$ და $P_t(k > 1)$-ის შედარებით ვასკვნით, რომ t-ს მცირე მნიშვნელობებისათვის ერთზე მეტი მოვლენის მოხდენის ალბათობა უმნიშვნელოა მხოლოდ ერთი მოვლენის მოხდენის ალბათობასთან შედარებით, რაც ორდინარობის თვისებას ახასითებს.

ამგვარად, პუასონის ფორმულა შეიძლება ჩაითვალოს ხდომილებების უმარტივესი ნაკადის მათემატიკურ მოდელად.

6.7. გეომეტრიული განაწილება

ვთქვათ, ტარდება დამოუკიდებელი ცდები, რომელთაგან თითოეულში A ხდომილების მოხდენის ალბათობა p-ს ტოლია $(0 < p < 1)$ და, აქედან გამომდინარე, მისი არ მოხდენის ალბათობა $q = 1 - p$. ცდა მაშინვე მთავრდება როგორც კი A ხდომილება მოხდება. ამგვარად, თუ A ხდომილება მოხდა k-ურ ცდაში, მაშინ მის წინა $\mathrm{k} - 1$ ცდაში ის არ მომხდარა.

X-ით აღვნიშნოთ დისკრეტული შემთხვევითი სიდიდე - ცდების რაოდენობა, რომელიც უნდა ჩატარდეს A ხდომილების პირველ მოხდენამდე. ცხადია, რომ X-ის შესაძლო მნიშვნელობები წარმოადგენს ნატურალურ რიცხვებს: $x_1 = 1, x_2 = 2, \ldots$

ვთქვათ პირველ $\mathrm{k}-1$ ცდაში A ხდომილება არ მოხდა, ხოლო k-ურ ცდაში მოხდა. ამ „რთული ხდომილების“ ალბათობა დამოუკიდებელ ხდომილებათა ალბათობების გამრავლების თეორემის თანახმად, ტოლია

$$P(X=k)=q^{k-1}p. \tag{6.7.1}$$

თუ გავითვალისწინებთ, რომ (6.7.1) ფორმულაში $k=1,2,\ldots$, მივიღეთ გეომეტრიულ პროგრესიას პირველი p წევრით და q $(0<q<1)$ მნიშვნელით :

$$p, qp, q^2p, \ldots, q^{k-1}p, \ldots. \tag{6.7.2}$$

ამის გამო (6.7.1) განაწილებას უწოდებენ *გეომეტრიულს.*

თუ (6.7.2) მიმდევრობის ელემენტებისაგან შევადგენთ მწკრივს, მაშინ ადვილად დავრწმუნდებით, რომ შედგენილი მწკრივი კრებადია და მისი ჯამი ერთის ტოლია. მართლაც, (6.7.2) მწკრივის ჯამი ტოლია

$$p+qp+q^2p+\cdots+q^{k-1}p+\cdots=\frac{p}{1-q}=\frac{p}{p}=1.$$

თავი 7. დისკრეტული შემთხვევითი სიდიდის მათემატიკური ლოდინი

7.1 დისკრეტული შემთხვევითი სიდიდის რიცხვითი მახასიათებლები

როგორც ცნობილია, განაწილების კანონი სრულად ახასიათებს შემთხვევით სიდიდეს. მაგრამ ხშირად განაწილების კანონი უცნობია და გვიწევს შემოვიფარგლოთ მცირე ინფორმაციით. ხანდახან მომგებიანია ვისარგებლოთ რიცხვებით, რომლებიც აღწერენ შემთხვევით სიდიდეს ჯამურად; ასეთ რიცხვებს უწოდებენ *შემთხვევითი სიდიდის რიცხვით მახასიათებლებს*. მნიშვნელოვან რიცხვით მახასიათებელს მიეკუთვნება მათემატიკური ლოდინი.

როგორც შემდგომში იქნება ნაჩვენები, მათემატიკური ლოდინი მიახლოებით შემთხვევითი სიდიდის საშუალო მნიშვნელობის ტოლია. მრავალი ამოცანის ამოსახსნელად საკმარისია მათემატიკური ლოდინის ცოდნა. მაგალითად, თუ ცნობილია, რომ პირველი მსროლელის მიერ დაგროვილი ქულების მათემატიკური ლოდინი მეტია ვიდრე მეორე მსროლელის მიერ დაგროვილი ქულების მათემატიკურ ლოდინზე, მაშინ პირველი მსროლელი საშუალოდ უფრო მეტ ქულას აგროვებს ვიდრე მეორე და, შესაბამისად, მეორეზე უკეთ ისვრის. თუმცა მათემატიკური ლოდინი შემთხვევით სიდიდეზე ნაკლებ ცნობებს იძლევა ვიდრე მისი განაწილების კანონი, თუმცა მათემატიკური ლოდინის ცოდნა საკმარისია ხოლმე ზოგიერთი ამოცანის ამოსახსნელად.

7.2. დისკრეტული შემთხვევითი სიდიდის მათემატიკური ლოდინი

დისკრეტული შემთხვევითი სიდიდის მათემატიკურ ლოდინი ეწოდება მისი ყველა შესაძლო მნიშვნელობის მათ ალბათობაზე ნამრავლთა ჯამს.

ვთქვათ, X შემთხვევთ სიდიდეს შეუძლია მიიღოს მხოლოდ შემდეგი მნშვნელობები $x_1, x_2, \dots, x_n$, რომელთა ალბათობები შესაბამისად ტოლია $p_1, p_2, \dots, p_n$ -ის. მაშინ X შემთხვევთი სიდიდის $M(X)$ მათემატიკური ლოდინი განისაზღვრება ტოლობით

$$M(X) = x_1 p_1 + x_2 p_2 + \cdots + x_n p_n.$$

თუ დისკრეტული შემთხვევითი სიდიდის მიერ მიღებული მნიშვნელობები თვლადია, მაშინ

$$M(X)=\sum_{i=1}^{\infty} x_i p_i.$$

ამასთან მათემატიკური ლოდინი არსებობს, თუ ტოლობის მარჯვენა ნაწილში მწკრივი აბსოლუტურად კრებადია (ანუ თუ სრულდება პირობა $\sum_{i=1}^{\infty}|x_i p_i|<\infty$).

შენიშვნა. განსაზღვრებიდან გამომდინარეობს, რომ დისკრეტული შემთხვევითი სიდიდის მათემატიკური ლოდინი მუდმივი სიდიდეა. შემდგომში ნაჩვენები იქნება, რომ უწყვეტი შემთხვევითი სიდიდის მათემატიკური ლოდინი ასევე მუდმივი სიდიდეა.

მაგალითი 1. ვიპოვოთ ერთ ცდაში A ხდომილების მოხდენის მათემატიკური ლოდინი, თუ A ხდომილების ალბათობა p-ს ტოლია.

ამოხსნა. X შემთხვევითი სიდიდეს - A ხდომილების მოხდენა ერთ ცდაში - შეუძლია მიიღოს მხოლოდ ორი მნიშვნელობა: $x_1=1$ (A ხდომილება მოხდა) p ალბთობით და $x_2=0$ (A ხდომილება არ მოხდა) $q=1-p$ ალბათობით. საძებნი მათემატიკური ლოდინი იქნება

$$M(X)=1\cdot p+0\cdot q=p.$$

7.3. მათემატიკური ლოდინის ალბათური შინაარსი

ვთქვათ, ტარდება n ცდა, რომლებშიც X შემთხვევთმა სიდიდემ m_1-ჯერ მიიღო x_1 მნიშვნელობა, m_2-ჯერ x_2 მნიშვნელობა, ..., m_k-ჯერ x_k მნიშვნელობა, ამასთან $m_1+m_2+\cdots+m_k=n$ მაშინ ყველა მნიშვნელობების ჯამი, რომელიც X-მა მიიღო ტოლია

$$x_1m_1+x_2m_2+\cdots+x_km_{k.}$$

ვიპოვოთ ყველა მნიშვნელობის, რომელიც შემთხვევითმა სიდიდემ მიიღო, $\bar{X}$ საშუალო არითმეტიკული, რისთვისაც მოძებნილი ჯამი უნდა გავყოთ ცდების საერთო რაოდენობაზე

$$\bar{X} = (x_1 m_1 + x_2 m_2 + \cdots + x_k m_{k.})/n,$$

ანუ

$$\bar{X} = x_1(m_1/n) + x_2(m_2/n) + \cdots + x_k(m_k/n) \qquad (7.3.1)$$

შევნიშნოთ, რომ შეფარდება m_1/n - x_1 მნიშვნელობის W_1 ფარდობითი სიხშირეა, m_2/n – x_2 მნიშვნელობის W_2 ფარდობითი სიხშირეა და ა.შ. ჩავწეროთ (7.3.1) ტოლობა ამგვარად:

$$\bar{X} = x_1 W_1 + x_2 W_2 + \cdots + x_k W_k \qquad (7.3.2)$$

დავუშვათ, რომ ცდების რაოდენობა საკმაოდ დიდია. მაშინ ფარდობითი სიხშირე მიახლოებთ უდრის ხდომილების მოხდენის ალბათობას (ამას მოგვიანებით დავამტკიცებთ):

$$W_1 \approx p_1, W_2 \approx p_2, \ldots, W_k \approx p_k.$$

(7.3.2) დამოკიდებულებაში ფარდობითი სიხშირეები შევცვალოთ შესაბამისი ალბათობებით, მივიღებთ

$$\bar{X} = x_1 p_1 + x_2 p_2 + \cdots + x_k p_k .$$

ამ მიახლოებითი ტოლობის მარჯვენა მხარე $M(X)$-ია.

ამგვარად

$$\bar{X} \approx M(X) .$$

მიღებული შედეგის ალბათური შინაარსი ასეთია: *მათემატიკური ლოდინი მიახლოებით ტოლია (რაც უფრო დიდია ცდების რაოდენობა მით უფრო ზუსტად) შემთხვევითი სიდიდის დაკვირვებადი მნიშვნელობების საშუალო არითმეტიკულის.*

7.4. მათემატიკური ლოდინის თვისებები

თვისება 1. მუდმივი სიდიდის მათემატიკური ლოდინი თვით ამ მუდმივის ტოლია:

$$M(C) = C.$$

დამტკიცება. C მუდმივი შეგვიძლია განვიხილოთ როგორც დისკრეტული შემთხვევითი სიდიდე, რომელსაც გაჩნია ერთი შესაძლო მნიშვნელობა (C) და მას იღებს $p = 1$ ალბათობით. აქედან გამომდინარე,

$$M(C) = C \cdot 1 = C.$$

შენიშვნა 1. განვსაზღვროთ C მუდმივი სიდიდის X დისკრეტულ შემთხვევით სიდიდეზე ნამრავლი როგორც დისკრეტული შემთხვევითი CX, რომლის შესაძლო მნიშვნელობები ტოლია მუდმივი C-ს ნამრავლის X -ის შესაძლო მნიშვნელობებზე; CX-ის შესაძლო მნიშვნელობების ალბათობები ტოლია შესაბამის X -ის შესაძლო მნიშვნელობათა ალბათობების. მაგალითად, თუ x_1 შესაძლო მნიშვნელობის ალბათობა ტოლია p_1-ის, მაშინ იმის ალბათობაც, რომ CX სიდიდე მიიღებს Cx_1 მნიშვნელობას, ასევე p_1-ს ტოლია.

თვისება 2. მუდმივი მამრავლი შეიძლება გატანილი იქნას მათემატიკური ლოდინის ნიშნის გარეთ:

$$M(CX) = CM(X).$$

დამტკიცება. ვთქვათ, X შემთხვევთი სიდიდე მოცემულია ალბათობათა განაწილების შემდეგი კანონით:

$$X \quad x_1 x_2 \;\; \ldots . \;\; x_n$$

$$p \quad p_1 p_2 \;\; \cdots \;\; p_n.$$

გავითვალისწინოთ შენიშვნა 1, ჩავწეროთ CX შემთხვევითი სიდიდის განაწილების კანონი:

$$CX \quad Cx_1 \quad Cx_2 \;\; \ldots . \;\; Cx_n$$

$$p \quad p_1 \quad p_2 \;\; \cdots \;\; p_n.$$

CX შემთხვევითი სიდიდის მათემატიკური ლოდინი ტოლია:

$$M(CX) = Cx_1p_1 + Cx_2p_2 + \cdots + Cx_np_n = C(x_1p_1 + x_2p_2 + \cdots + x_np_n) = CM(X).$$

შენიშვნა 2. ვიდრე შემდეგ თვისებაზე გადავიდოდეთ, მივუთითოთ, რომ ორ შემთხვევით სიდიდეს ვუწოდებთ *დამოუკიდებელს,* თუ ერთ-ერთის განაწილების კანონი არაა დამოკიდებული იმაზე, თუ რა მნიშვნელობები მიიღო მეორე შემთხვევითმა სიდიდემ. წინააღმდეგ შემთხვევაში შემთხვევითი სიდიდეები *დამოკიდებულები* არიან. რამდენიმე შემთხვევით სიდიდეს უწოდებენ ურთიერთდამოუკიდებელს, თუ მათი რაოდენობიდან ნებისმიერის განაწილების კანონი არაა დამოკიდებული იმაზე, თუ რა შესაძლო მნიშვნელობებს იღებენ დანარჩენი სიდიდეები.

შენიშვნა 3. განვსაზღვროთ დამოუკიდებელი შემთხვევითი სიდიდეების X და Y ნამრავლი, როგორც შემთხვევითი სიდიდე, რომლის შესაძლო მნიშვნელობები ტოლია X -ის თითოეული შესაძლო მნიშვნელობის Y-ის თითოეული შესაძლო მნიშვნელობაზე ნამრავლის; XY ნამრავლის შესაძლო მნიშვნელობათა ალბათობა ტოლია თითოეული თანამამრავლის შესაძლო მნიშვნელობათა ალბათობების ნამრავლის. მაგალითად, თუ შესაძლო x_1 მნიშვნელობის ალბათობა p_1 -ის ტოლია, y_1 შესაძლო მნიშვნელობის ალბათობა g_1ის ტოლია, მაშინ x_1y_1 შესაძლო მნიშვნელობის ალბათობა p_1g_1-ის ტოლია.

შევნიშნოთ, რომ ზოგიერთი x_iy_j ნამრავლი შეიძლება ერთმანეთის ტოლი იყოს. ამ შემთხვევაში შესაძლო მნიშვნელობათა ნამრავლების ალბათობა შესაბამისი ალბათობათა ჯამის ტოლია. მაგალითად, თუ $x_1y_2 = x_3y_5$, მაშინ x_1y_2 -ის (ან x_3y_5 -ის) ალბათობა $p_1g_2 + p_3g_5$-ის ტოლია.

თვისება 3. ორი დამოუკიდებელი შემთხვევითი სიდიდის ნამრავლის მათემატიკური ლოდინი მათი მათემატიკური ლოდინების ნამრავლის ტოლია:

$$M(XY) = M(X)M(Y).$$

დამტკიცება. ვთქვათ X და Y შემთხვევითი სიდიდეები თავიანთი განაწილების კანონებითაა მოცემული[2]:

$$\begin{array}{cccc} X & x_1x_2 & Y & y_1y_2 \\ P & p_1p_2 & g & g_1g_2. \end{array}$$

შევადგინოთ ყველა მნიშვნელობა, რომელთა მიღებაც შეუძლია XY შემთხვევით სიდიდეს. ამისათვის გადავამრავლოთ X -ის ყველა შესაძლო მნიშვნელობა Y -ის ყველა შესაძლო მნიშვნელობაზე; შედეგად მივიღებთ $x_1y_1,\ x_2y_1,\ x_1y_2$ და x_2y_2. შენიშვნა 3-ის გათვალისწინებით, დავწეროთ XY ის განაწილების კანონი, სიმარტივისათვის ვივარაუდოთ, რომ ნამრავლების ყველა შესაძლო შემთხვევა განსხვავებულია (თუ ეს ასე არაა, დამტკიცება ანალოგიურად სრულდება):

[2] გაანგარიშებების გასამარტივებლად ჩვენ შესაძლო მნიშვნელობების მცირე რაოდენობით შემოვიფარგლეთ. ზოგად შემთხვევაში დამტკიცება ანალოგიურია.

$$XY \quad x_1y_1 \quad x_2y_1 \quad x_1y_2 \; x_2y_2$$
$$p \quad\; p_1g_1 \quad p_2g_1 \quad p_1g_2 \quad p_2g_2.$$

მათემატიკური ლოდინი ტოლია ყველა შესაძლო შემთხვევის მათ ალბათობებზე ნამრავლების ჯამის:

$$M(XY) = x_1y_1 \cdot p_1g_1 + x_2y_1 \cdot p_2g_1 + x_1y_2 \cdot p_1g_2 + x_2y_2 \cdot p_2g_2.$$

ანუ

$$M(XY) = y_1g_1(x_1p_1 + x_2p_2) + y_2g_2(x_1p_1 + y_2g_2) = (x_1p_1 + x_2p_2)(y_1g_1 + y_2g_2) =$$
$$= M(X) \cdot M(Y).$$

ამგვარად, $M(XY) = M(X) \cdot M(Y)$.

შედეგი. რამდენიმე ურთიერთდამოუკიდებელი შემთხვევითი სიდიდის ნამრავლის მათემატიკური ლოდინი მათი მათემატიკური ლოდინების ნამრავლის ტოლია.

შემთხვევითი სიდიდეების ნებისმიერი რაოდენობისათვის დამტკიცება ხდება მათემატიკური ინდუქციის მეთოდის გამოყენებით.

ქვემოთ მოყვანილი თვისება სამართლიანია როგორც დამოუკიდებელი, ისე დამოკიდებული შემთხვევითი სიდიდეებისათვის.

თვისება 4. ორი შემთხვევითი სიდიდის ჯამის მათემატიკური ლოდინი ტოლია შესაკრებთა მათემატიკური ლოდინების ჯამის:

$$M(X+Y) = M(X) + M(Y)$$

დამტკიცება. ვთქვათ X და Y შემთხვევითი სიდიდეები მოცემულია შემდეგი განაწილების კანონებით[3]:

$$X \quad x_1x_2 \quad\; Y \quad y_1y_2$$
$$P \quad p_1p_2 \quad g \quad\; g_1g_2.$$

შევადგინოთ $X+Y$ სიდიდის ყველა შესაძლო მნიშვნელობა. ამისათვის X -ის ყველა შესაძლო შემთხვევას მივუმატოთ Y-ის თითოეული შესაძლო მნიშვნელობა; მივიღებთ x_1+y_1, x_1+y_2, x_2+y_1 და x_2+y_2. სიმარტივისათვის დავუშვათ, რომ ეს შესაძლო

[3] გამოყვანის გასამარტივებლად, შემოვიფარგლეთ თითოეული სიდიდის მხოლოდ ორი შესაძლო მნიშვნელობით. ზოგად შემთხვევაში დამტკიცება ანალოგიურია.

მნიშვნელობები სხვადასხვაა (თუ ეს ასე არაა, დამტკიცება ანალოგიურად წარიმართება) და მათი ალბათობები შესაბამისად აღვნიშნოთ p_{11}, p_{12}, p_{21} და p_{22}-ით.

$X+Y$ სიდიდის მათემატიკური ლოდინი ტოლია შესაძლო მნიშვნელობების მათ ალბათობებზე ნამრავლების ჯამის:

$$M(X+Y) = (x_1+y_1)p_{11} + (x_1+y_2)p_{12} + (x_2+y_2)p_{21} + (x_2+y_2)p_{22},$$

ანუ

$$M(X+Y) = x_1(p_{11}+p_{12}) + x_2(p_{21}+p_{22}) + y_1(p_{11}+p_{21}) + y_2(p_{12}+p_{22}). \qquad (7.4.1)$$

დავამტკიცოთ, რომ $p_{11}+p_{12}=p_1$. ის ხდომილება, რომ X იღებს x_1 მნიშვნელობას (ამ ხდომილების ალბათობა p_1-ის ტოლია), იწვევს ხდომილებას, რომელიც იმაში მდგომარეობს, რომ $X+Y$ იღებს x_1+y_1 ან x_1+y_2 მნიშვნელობას, (ამ ხდომილების ალბათობა შეკრების თეორემის თანახმად $p_{11}+p_{12}$-ის ტოლია) და პირიქით. აქედან კი გამომდინარეობს, რომ $p_{11}+p_{12}=p_1$. ანალოგიურად მტკიცდება ტოლობები

$$p_{21}+p_{22}=p_2, \qquad p_{11}+p_{21}=g_1, \qquad \text{და} \qquad p_{12}+p_{22}=g_2.$$

ამ ტოლობების მარჯვენა მხარეებს თუ ჩავსვამთ (7.4.1) გამოსახულებაში, მივიღებთ

$$M(X+Y) = (x_1p_1 + x_2p_2) + (y_1g_1 + y_2g_2),$$

ანუ საბოლოოდ

$$M(X+Y) = M(X) + M(Y).$$

შედეგი. *რამდენიმე შემთხვევითი სიდიდის ჯამის მათემატიკური ლოდინი შესაკრებთა მათემატიკური ლოდინების ჯამის ტოლია.*

შესაკრები სიდიდეების ნებისმიერი რაოდენობისათვის დამტკიცება ხდება მათემატიკური ინდუქციის მეთოდის გამოყენებით.

7.5 დამოუკიდებელ ცდებში ხდომილების მოხდენის რაოდენობის მათემატიკური ლოდინი

ვთქვათ, ტარდება n დამოუკიდებელი ცდა, რომელთაგან თითოეულში A ხდომილების მოხდენის ალბათობა მუდმივია და p-ს უდრის. რას უდრის ამ ცდებში A ხდომილების მოხდენის საშუალო რაოდენობა? ამ კითხვაზე პასუხს იძლევა შემდეგი თეორემა.

თეორემა. n დამოუკიდებელ ცდაში A ხდომილების მოხდენის რაოდენობის მათემატიკური ლოდინი $M(X)$ ტოლია ცდების რაოდენობის ნამრავლის თითოეულ ცდაში ხდომილების მოხდენის ალბათობაზე:

$$M(X) = np.$$

დამტკიცება. X შემთხვევითი სიდიდის სახით განვიხილავთ n დამოუკიდებელ ცდაში A ხდომილების მოხდენის რაოდენობას. აღვნიშნოთ X_1-ით პირველ ცდაში A ხდომილების მოხდენის რაოდენობა, X_2-ით - მეორე ცდაში A ხდომილების მოხდენის რაოდენობა, ..., $X_n - n$-ურ ცდაში A ხდომილების მოხდენის რაოდენობა. მაშინ A ხდომილების მოხდენის საერთო რაოდენობა იქნება $X = X_1 + X_2 + \cdots + X_n$.

მათემატიკური ლოდინის მეოთხე თვისების თანახმად,

$$M(X) = M(X_1) + M(X_2) + \cdots + M(X_n). \tag{7.5.1}$$

ტოლობის მარჯვენა ნაწილში შესაკრებიდან თითოეული წარმოადგენს ერთ ცდაში ხდომილების მოხდენის რაოდენობას: $M(X_1)$ - პირველში, $M(X_2)$ - მეორეში, და ა.შ. რამდენადაც ერთ ცდაში ხდომილების მოხდენის მათემატიკური ლოდინი ხდომილების ალბათობის ტოლია, ამდენად $M(X_1) = M(X_2) = \cdots = M(X_n) = p$. თუ ჩავსვამთ (7.5.1) ტოლობის მარჯვენა მხარეში თითოეული შესაკრების ნაცვლად p -ს, მივიღებთ

$$M(X) = np.$$

შენიშვნა, რამდენადაც X შემთხვევითი სიდიდე განაწილებულია ბინომური კანონის მიხედვით, ამდენად დამტკიცებული თეორემა ასეც შეიძლება ჩამოვაყალიბოთ: ბინომური განაწილების მათემატიკური ლოდინი n და p პარამეტრებით np ნამრავლის ტოლია.

თავი 8. დისკრეტული შემთხვევითი სიდიდის დისპერსია

8.1 შემთხვევითი სიდიდის გაბნევის რიცხვითი მახასიათებელის შემოყვანის მიზანშეწონილობა

ადვილი მოსაძებნია ისეთი შემთხვევითი სიდიდეები, რომელთაც ერთნაირი მათემატიკური ლოდინი, მაგრამ სხვადასხვა შესაძლო მნიშვნელობები აქვთ. განვიხილოთ, მაგალითად, X და Y დისკრეტული შემთხვევითი სიდიდეები, რომლებიც შემდეგი განაწილების კანონებითაა მოცემული:

$$\begin{array}{ccc} X & -0{,}01 & 0{,}01 \\ P & 0{,}5 & 0{,}5 \end{array} \qquad \begin{array}{ccc} Y & -100 & 100 \\ P & 0{,}5 & 0{,}5. \end{array}$$

გამოვითვალოთ ამ სიდიდეების მათემატიკური ლოდინები:

$$M(X) = -0{,}01 \cdot 0{,}5 + 0{,}01 \cdot 0{,}5 = 0,$$

$$M(\mathrm{Y}) = -100 \cdot 0{,}5 + 100 \cdot 0{,}5 = 0.$$

აქ ორივე სიდიდის მათემატიკური ლოდინები ერთნაირია, ხოლო შესაძლო მნიშვნელობები განსხვავებული, ამასთან, X -ს გააჩნია მათემატიკურ ლოდინთან მიახლოებული შესაძლო მნიშვნელობები, ხოლო Y-ს თავის მათემატიკურ ლოდინთან დაშორებული. ამგვარად, როცა ვიცით შემთხვევითი სიდიდის მათემატიკური ლოდინი, ჯერ არ შეიძლება იმაზე მსჯელობა, თუ როგორი შესაძლო მნიშვნელობები შეუძლია მან მიიღოს, და არც იმაზე, თუ როგორაა ისინი გაბნეული მათემატიკური ლოდინის ირგვლივ. სხვა სიტყვებით, მათემატიკური ლოდინი შემთხვევით სიდიდეს სრულად ვერ ახასითებს.

ამ მიზეზით მათემატიკურ ლოდინთან ერთად შემოღებულია სხვა რიცხვითი მახასიათებლებიც. ასე მაგალითად, იმისათვის, რომ შევაფასოთ, თუ როგორაა გაბნეული შემთხვევითი სიდიდის შესაძლო მნიშვნელობები მისი მათემატიკური ლოდინის ირგვლივ, სარგებლობენ რიცხვითი მასახიათებლით, რომელსაც დისპერსიას უწოდებენ.

სანამ გადავალთ დისპერსიის და მისი თვისებებების განსაზღვრებამდე, შემოვიღოთ შემთხვევითი სიდიდის მისი მათემატიკური ლოდინიდან გადახრის ცნება.

8.2 შემთხვევითი სიდიდის მისი მათემატიკური ლოდინიდან გადახრა

ვთქვათ, X შემთხვევითი სიდიდეა და $M(X)$ - მისი მათემატიკური ლოდინი. ახალი შემთხვევითი სიდიდის სახით განვიხილოთ სხვაობა $X - M(X)$.

გადახრას უწოდებენ სხვაობას შემთხვევით სიდიდესა და მის მათემატიკურ ლოდინს შორის.

ვთქვათ X -ის განაწილების კანონი ცნობილია:

$$\begin{array}{ccccc} X & x_1 & x_2 & \ldots. & x_n \\ p & p_1 & p_2 & \cdots & p_n. \end{array}$$

დავწეროთ გადახრის განაწილების კანონი. იმისათვის, რომ გადახრამ მიიღოს $x_1 - M(X)$ მნიშვნელობა, საკმარისია, რომ შემთხვევითმა სიდიდემ მიიღოს x_1 მნიშვნელობა, ამის ალბათობა კი p_1-ის ტოლია; აქედან გამომდინარე იმის ალბათობაც, რომ გადახრამ მიიღოს $x_1 - M(X)$ მნიშვნელობა p_1-ის ტოლია. ვითარება ანალოგიურია დანარჩენი გადახრების შესაძლო მნიშვნელობებისათვის.

ამგვარად, გადახრას გააჩნია შემდეგი განაწილების კანონი:

$$\begin{array}{cccc} X - M(X) & x_1 - M(X) & x_2 - M(X) \ldots & x_n - M(X) \\ p & p_1 & p_2 \quad \ldots.. & p_n. \end{array}$$

მოვიყვანოთ გადახრის მნიშვნელოვანი თვისება, რომელსაც შემდგომში გამოვიყენებთ.

თეორემა. გადახრის მათემატიკური ლოდინი ნულის ტოლია:

$$M[X - M(X)] = 0.$$

დამტკიცება. მათემატიკური ლოდინის თვისებების გამოყენებით (სხვაობის მათემატიკური ლოდინი მათემატიკური ლოდინების სხვაობების ტოლია, მუდმივის მათემატიკური ლოდინი თვით ამ მუდმივის ტოლია) და თუ მხედველობაში მივიღებთ, რომ $M(X)$ მუდმივი სიდიდეა, გვექნება

$$M[X - M(X)] = M(X) - M[M(X)] = M(X) - M(X) = 0.$$

8.3 დისკრეტული შემთხვევითი სიდიდის დისპერსია

პრაქტიკაში ხშირად მოითხოვება შეფასდეს შემთხვევითი სიდიდის შესაძლო მნიშვნელობების გაბნევა მისი საშუალო მნიშვნელობის ირგვლივ. მაგალითად, არტილერიაში მნიშვნელოვანია იმის ცოდნა, რამდენად შეჯგუფებულად განლაგდება ჭურვები დასაზიანებელი მიზნის ირგვლივ.

ერთი შეხედვით შეიძლება მოგვეჩვენოს, რომ გაბნევის შეფასებისათვის უფრო მარტივია გამოვთვალოთ შემთხვევითი სიდიდის გადახრის ყველა შესაძლო მნიშვნელობა და შემდეგ ვიპოვოთ მისი საშუალო მნიშვნელობა. მაგრამ ეს გზა არაფერს მოგვცემს, რამდენადაც გადახრის საშუალო მნიშვნელობა, ე.ი. $M[X--M(X)]$, ნებისმიერი შემთხვევითი სიდიდისთვის ნულის ტოლია. ეს თვისება უკვე დამტკიცებული იყო წინა პარაგრაფში და იმით აიხსნება, რომ ზოგი შესაძლო შემთხვევა დადებითებია და ზოგიც უარყოფითი; მათი ურთიერთ გაბათილებით გადახრის საშუალო მნიშვენელობა ნულის ტოლია. ეს მოსაზრება მიუთითებს გადახრების შესაძლო მნიშვნელობების მათი აბსოლუტური მნიშვნელობებით ან კვადრატებით შეცვლის მიზანშეწონილობაზე. სწრედ ასე იქცევიან. მართალია იმ შემთხვევაში, როდესაც გადახრების შესაძლო მნიშვნელობებს მათი აბსოლუტური მნიშვნელობებით ცვლიან, გვიხდება აბსოლუტური სიდიდეებით ოპერირება, რასაც ზოგჯერ სერიოზულ სიძნელეებამდე მივყავართ. ამიტომ უფრო ხშირად ითვლიან გადახრის კვადრატის საშუალო მნიშვნელობას, რომელსაც დისპერსიას უწოდებენ.

დისკრეტული შემთხვევითი სიდიდის *დისპერსიას* უწოდებენ შემთხვევითი სიდიდის მისი მათემატიკური ლოდინიდან გადახრის კვადრატის მათემატიკურ ლოდინს:

$$D(X) = M[X - M(X)]^2.$$

ვთქვათ, შემთხვევითი სიდიდე მოცემულია განაწილების კანონით

$$\begin{array}{ccccc} X & x_1 & x_2 & \ldots. & x_n \\ p & p_1 & p_2 & \ldots & p_n. \end{array}$$

მაშინ გადახრის კვადრატს აქვს შემდეგი განაწილების კანონი

$$[x-M(X)]^2 [x_1-M(X)]^2 [x_2-M(X)]^2 \ldots [x_n-M(X)]^2$$

$$p \qquad p_1 \qquad p_2 \quad \ldots \quad p_n.$$

დისპერსიის განსაზღვრების თანახმად,

$$D(X)=M[X-M(X)]^2=[x_1-M(X)]^2 p_1+[x_2-M(X)]^2 p_2+\cdots+[x_n-M(X)]^2 p_n.$$

ამგვარად, იმისათვის, რომ ვიპოვოთ დისპერსია, საკმარისია გამოვთვალოთ გადახრის შესაძლო მნიშვნელობის მათ ალბათობებზე ნამრავლების ჯამი.

შენიშვნა. განსაზღვრებიდან გამომდინარეობს, რომ დისკრეტული შემთხვევითი სიდიდის დისპერსია არაშემთხვევითი (მუდმივი) სიდიდეა. შემდგომში გავიგებთ, რომ უწყვეტი შემთხვევითი სიდიდის დისპრესიაც ასევე მუდმივი სიდიდეა.

8.4. დისპერსიის გამოსათვლელი ფორმულა

დისპერსიის გამოსათვლელად ხშირად მოსახერხებელია შემდეგი თეორემით სარგებლობა.

თეორემა. *X შემთხვევითი სიდიდის დისპერსია ტოლია X ის კვადრატის მათემატიკური ლოდინისა და მისი მათემატიკური ლოდინის კვადრატს შორის სხვაობის:*

$$D(X)=M(X^2)-[M(X)]^2.$$

დამტკიცება. $M(X)$ მათემატიკური ლოდინი მუდმივი სიდიდეა, აქედან, $2M(X)$ და $M^2(X)$ ასევე მუდმივი სიდიდეებია. მივიღოთ ეს მხედველობაში და გამოვიყენოთ მათემატიკური ლოდინის თვისებები (მუდმივი მამრავლი შეიძლება გავიტანოთ მათემატიკური ლოდინის ნიშნის გარეთ, ჯამის მათემატიკური ლოდინი შესაკრებთა მათემატიკური ლოდინების ჯამის ტოლია), გავამარტივოთ დისპერსიის განმსაზღვრელი ფორმულა:

$$D(X)=M[X-M(X)]^2=M[X^2-2XM(X)+M^2(X)]=M(X^2)-2M(X)M(X)+M^2(X)=$$
$$=M(X^2)-2M^2(X)+M^2(X)=M(X^2)-M^2(X).$$

ამრიგად,

$$D(X)=M(X^2)-[M(X)]^2.$$

ფორმულაში კვადრატული ფრჩხილი შემოღებულია ფორმულის ადვილად დამახსოვრების მიზნით.

8.5. დისპერსიის თვისებები

თვისება 1. მუმივი C სიდიდის დისპერსია ნულის ტოლია:

$$D(C) = 0.$$

დამტკიცება. დისპერსიის განსაზღვრების თანახმად,

$$D(C) = M\{[C - M(C)]^2\}.$$

მათემატიკური ლოდინის პირველი თვისების გამოყენებით (მუდმივის მათემატიკური ლოდინი თვით ამ მუდმივის ტოლია) მივიღებთ

$$D(C) = M[(C - C)^2] = M(0) = 0.$$

ამგვარად,

$$D(C) = 0.$$

თვისება ცხადი ხდება, თუ გავითვალისწინებთ, რომ მუდმივი სიდიდე ინარჩუნებს ერთი და იმავე მნიშვნელობას და გაბნევა, რა თქმა უნდა არა აქვს.

თვისება 2. მუდმივი მამრავლი შეიძლება გავიტანოთ დისპერსიის ნიშნის გარეთ მისი კვადრატში აყვანით:

$$D(CX) = C^2D(X).$$

დამტკიცება. დისპერსიის განსაზღვრების თნხმად, გვაქვს

$$D(C) = M\{[CX - M(CX)]^2\}.$$

მათემატიკური ლოდინის მეორე თვისების გამოყენებით (მუდმივი მამრავლი შეიძლება გატანილი იქნას მათემატიკური ლოდინის ნიშნის გარეთ) მივიღებთ

$$D(CX) = M\{[CX - CM(X)]^2\} = M\{C^2[X - M(X)]^2\} = C^2M\{[X - M(X)]^2\} = C^2D(X)$$

ამგვარად,

$$D(CX) = C^2D(X).$$

თვისება 3. *ორი დამოუკიდებელი შემთხვევითი სიდიდის ჯამის დისპერსია ტოლია ამ სიდიდეების დისპერსიების ჯამისა:*

$$D(X + Y) = D(X) + D(Y).$$

დამტკიცება. დისპერსიის გამოსათვლელი ფორმულის მიხედვით, გვაქვს

$$D(X+Y) = M[(X+Y)^2] - [M(X+Y)]^2$$

თუ ფრჩხილებს გავხსნით და გამოვიყენებთ რამდენიმე სიდიდის ჯამის და ორი დამოუკიდებელი შემთხვევითი სიდიდის ნამრავლის მათემატიკური ლოდინის თვისებებს, მივიღებთ

$$\begin{aligned} D(X+Y) &= M[X^2 + 2XY + Y^2] - [M(X) + M(Y)]^2 = \\ &= M(X^2) + 2M(X)\cdot M(Y) + M(Y^2) - M^2(X) - 2M(X)\cdot M(Y) - M^2(Y) = \\ &= \{M(X^2) - [M(X)]^2\} + \{M(Y^2) - [M(Y)]^2\} = D(X) + D(Y) \end{aligned}$$

ამგვარად,

$$D(X+Y) = D(X) + D(Y)$$

შედეგი 1. რამდენიმე წყვილ-წყვილად დამოუკიდებელი შემთხვევითი სიდიდეების ჯამის დისპერია ამ სიდიდეების დისპერიების ჯამის ტოლია.

მაგალითად, სამი შესაკრებისათვის გვაქვს:

$$D(X+Y+Z) = D[X + (Y+Z)] = D(X) + D(Y+Z) = D(X) + D(Y) + D(Z).$$

შესაკრების ნებისმიერი რაოდენობისათვის დამტკიცება ხდება მათემატიკური ინდუქციის მეთოდის გამოყენებით.

შედეგი 2. *მუდმივი და შემთხვევითი სიდიდების ჯამის დისპერსია შემთხვევითი სიდიდის დისპერსიის ტოლია:*

$$D(C+X) = D(X).$$

დამტკიცება. C და X სიდიდეები დამოუკიდებლები არიან, ამიტომ მე-4 თვისების თანახმად

$$D(C+X) = D(C) + D(X).$$

პირველი თვისების ძალით $D(C) = 0.$ აქედან

$$D(C+X) = D(X).$$

თვისება 4. *ორი დამოუკიდებელი შემთხვევითი სიდიდის სხვაობის დისპერსია მათი დისპერსიების ჯამის ტოლია:*

$$D(X-Y) = D(X) + D(Y).$$

დამტკიცება. მე-4 თვისების თანახმად

$$D(X-Y)=D(X)+D(-Y).$$

მეორე თვისებით

$$D(X-Y)=D(X)+(-1)^2D(Y),$$

ანუ

$$D(X-Y)=D(X)+D(Y).$$

8.6. დამოუკიდებელ ცდებში ხდომილების მოხდენის რაოდენობის დისპერსია

ვთქვათ, ხორციელდება n დამოუკიდებელი ცდა, რომელთაგან თითოეულში A ხდომილების მოხდენის ალბათობა მუდმივია. რას უდრის ამ ცდებში ხდომილების მოხდენის რაოდენობის დისპერსია? ამ კითხვაზე პასუხს იძლევა შემდეგი თეორემა.

თეორემა. A ხდომილების მოხდენის ალბათობის დისპერსია n დამოუკიდებელი ცდაში, რომელთაგან თითოეულში ხდომილების მოხდენის p ალბათობა მუდმივია, ტოლია ცდების რაოდენობის და ერთ ცდაში ხდომილების მოხდენის და არ მოხდენის ალბათობებზე ნამრავლის:

$$D(X)=npq.$$

დამტკიცება. განვიხილოთ X შემთხვევითი სიდიდე - A ხდომილების დამოუკიდებლ ცდებში მოხდენის რაოდენობა. ცხადია, ამ ცდებში ხდომილების მოხდენის საერთო რაოდენობა ტოლია ცალკეულ ცდაში ხდომილების მოხდენათა ჯამის:

$$X=X_1+X_2+\cdots+X_n,$$

სადაც, X_1 - პირველ ცდაში A ხდომილების მოხდენის რაოდენობაა, X_2- მეორეში , ..., X_n - n -ურში.

$X_1,\ X_2,\dots,X_n$ სიდიდეები ურთიერთდამოუკიდებლები არიან, რამდენადაც თითოეული ცდა არაა დამოკიდებული დანარჩენების შედეგზე, ამიტომ უფლება გვაქვს გამოვიყენოთ შედეგი 1 (იხ. პარაგრაფი 8.5):

$$D(X)=D(X_1)+D(X_2)+\cdots+D(X_n). \qquad (8.6.1)$$

X_1-ის დისპერსია გამოვთვალოთ ფორმულით:

$$D(X_1) = M(X_1^2) - [M(X_1)]^2. \qquad (8.6.2)$$

X_1 შემთხვევითი სიდიდე არის A ხდომილების მოხდენის რაოდენობა პირველ ცდაში, ამიტომ (იხ. მაგალითი 1 პარაგრაფი 7.2) $M(X_1) = p$.

ვიპოვოთ X_1^2 შემთხვევითი სიდიდის მათემატიკური ლოდინი, რომელსაც მხოლოდ ორი მნიშვნელობის მიღება შეუძლია, კერძოდ: 1^2 -სა p ალბათობით და 0-სა q ალბათობით:

$$M(X_1^2) = 1^2 \cdot p + 0^2 \cdot q = p.$$

თუ მიღებულ შედეგს (8.6.2) დამოკიდებულებაში ჩავსვამთ, მივიღებთ

$$D(X_1) = p - p^2 = p(1-p) = pq.$$

ცხადია, დარჩენილი შემთხვევითი სიდიდეებიდან თითოეულის დისპერსიაც ასევე pq -ს ტოლია. თუ (8.6.1) ტოლობის მარჯვენა მხარეში თითოეულ შესაკრებს შევცვლით pq-თი, საბოლოოდ მივიღებთ:

$$D(X) = npq.$$

შენიშვნა. რამდენადაც შემთხვევითი სიდიდე განაწილებულია ბინომური კანონით, ამდენად დამტკიცებული თეორემა შეიძლება ასეც ჩამოვაყალიბოთ: *ბინომური განაწილების დისპერსია n და p პარამეტრებით ტოლია npq ნამრავლის.*

8.7. საშუალო კვადრატული გადახრა

შემთხვევითი სიდიდის შესაძლო მნიშვნელობების მისი საშუალო მნიშვნელობების გარშემო გაბნევის შესაფასებლად დისპერსიის გარდა ზოგიერთი სხვა მახასიათებლებიც გამოიყენება. მათ რიცხვს მიეკუთვნება საშუალო კვადრატული გადახრა.

X შემთხვევითი სიდიდის *საშუალო კვადრატულ გადახრას* უწოდებენ დისპერსიიდან ფესვს:

$$\sigma(X) = \sqrt{D(X)}.$$

ადვილი საჩვენებელია, რომ დისპერსიას გააჩნია განზომილება, რომელიც შემთხვევითი სიდიდის განზომილების კვადრატის ტოლია. რამდენადაც საშუალო კვადრატული გადახრა ტოლია დისპერიისდან ფესვის, ამდენად $\sigma(X)$-ის განზომილება ემთხვევა X -ის განზომილებას. ამიტომ იმ შემთხვევებში, როცა სასურველია, რომ

გაბნევას ჰქონდეს შეთხვევითი სიდიდის განზომილება, იყენებენ საშუალო კვადრატულ გადახრას და არა დისპერსიას. მაგალითად, თუ X გამოსახულია წრფივ განზომილებებში, მაშინ $\sigma(X)$-ც ასევ გამოისახება წრფივ განზომილებებში, ხოლო $D(X)$ კვადრატულ განზომილებებში.

8.8 ურთიერთდამოუკიდებელი შემთხვევითი სიდიდეების ჯამის საშუალო კვადრატული გადახრა

ვთქვათ, ცნობილია რამდენიმე წყვილ-წყვილად დამოუკიდებელი *შემთხვევითი* სიდიდის საშუალო კვადრატული გადახრა. როგორ ვიპოვოთ ამ სიდიდეების ჯამის საშუალო კვადრატული გადახრა? ამ კითხვაზე პასუხს იძლევა შემდეგი თეორემა.

თეორემა. *სასრული რაოდენობის* წყვილ-წყვილად დამოუკიდებელი *შემთხვევითი სიდიდეების ჯამის საშუალო კვადრატული გადახრა ამ სიდიდეების საშუალო კვადრატული გადახრების კვადრატების ჯამიდან კვადრატული ფესვის ტოლია:*

$$\sigma(X_1+X_2+\cdots+X_n)=\sqrt{\sigma^2(X_1)+\sigma^2(X_2)+\cdots+\sigma^2(X_n)}.$$

დამტკიცება. X-ით აღვნიშნოთ განსახილველი წყვილ-წყვილად დამოუკიდებელი *შემთხვევითი* სიდიდეების ჯამი:

$$X=X_1+X_2+\cdots+X_n.$$

რამდენიმე წყვილ-წყვილად დამოუკიდებელი შემთხვევითი სიდიდის ჯამის დისპერსია ტოლია შესაკრებთა დისპერსიების ჯამის (იხ. პარაგრაფი **8.5** შედეგი 1), ამიტომ

$$D(X)=D(X_1)+D(X_2)+\cdots+D(X_n).$$

აქედან

$$\sqrt{D(X)}=\sqrt{D(X_1)+D(X_2)+\cdots+D(X_n)},$$

ანუ საბოლოოდ

$$\sigma(X)=\sqrt{\sigma^2(X_1)+\sigma^2(X_2)+\cdots+\sigma^2(X_n)}.$$

8.9. ერთნაირად განაწილებული ურთიერთდამოუკიდებელი შემთხვევითი სიდიდეები

უკვე ცნობილია, რომ განაწილების კანონის მიხედვით შესაძლებელია შემთხვევითი სიდიდის რიცხვითი მახასითებლების მოძებნა. აქედან გამომდინარეობს, რომ თუ რამდენიმე შემთხვევით სიდიდეს გააჩნიათ ერთნაირი განაწილება, მაშინ მათი რიცხვითი მახასითებლები ერთნაირია.

განვიხილოთ n წყვილ-წყვილად დამოუკიდებელი შემთხვევითი სიდიდე $X_2, \dots, , X_n$, რომელთაც გააჩნიათ ერთნაირი განაწილებები, და მაშასადამე, ერთნაირი მახასიათებლები (მათემატიკური ლოდინი, დისპერსია და სხვა). ამ მახასითებლებიდან ჩვენთვის საინტერესოა ამ სიდიდეების საშუალო არითმეტიკული, რომელსაც აქ განვიხილავთ.

განხილული შემთხვევითი სიდიდეების საშუალო არითმეტიკული აღვნიშნოთ $\bar{X}$-ით:

$$\bar{X} = (X_1 + X_2 + \cdots + X_n)/n.$$

შემდეგი სამი პოზიცია ამყარებს კავშირს $\overline{X}$ საშუალო არითმეტიკულის რიცხვით მახასიათებლებსა და ცალკეული სიდიდის შესაბამის მახასითებლებს შორის.

1. ერთნაირად განაწილებული წყვილ-წყვილად დამოუკიდებელი *შემთხვევითი სიდიდეების საშუალო არითმეტიკულის მათმატიკური ლოდინი ტოლია თითოეული სიდიდის მათემატიკური ლოდინის:*

$$M(\bar{X}) = M(X_1).$$

დამტკიცება. წერის სიმოკლისათვის შემოვიტანოთ აღნიშვნა $M(X_1) = a$. მათემატიკური ლოდინის თვისებების გამოყენებით (მუდმივი შეიძლება გატანილ იქნას მათემატიკური ლოდინის ნიშნის გარეთ; ჯამის მათემატიკური ლოდინი შესაკრებთა მათემატიკური ლოდინების ჯამის ტოლია), გვაქვს

$$M(\bar{X}) = M\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{M(X_1) + M(X_2) + \cdots + M(X_n)}{n}.$$

თუ მხედველობაში მივიღებთ, რომ, პირობის თანახმად, თითოეულის მათემატიკური ლოდინი a -ს ტოლია, მივიღებთ

$$M(\bar{X}) = na/n = a.$$

2. n ერთნაირად განაწილებული წყვილ-წყვილად დამოუკიდებელი *შემთხვევითი სიდიდეების საშუალო არითმეტიკულის დისპერსია n -ჯერ ნაკლებია თითოეული სიდიდის D დისპერსიაზე:*

$$D(\bar{X}) = D/n. \tag{8.9.1}$$

დამტკიცება. დისპერსიის თვისებების გამოყენებით (მუდმივი სიდიდე შეიძლება გამოტანილ იქნას დისპერსიის ნიშნის გარეთ მისი კვადრატში აყვანით; დამოუკიდებელი სიდიდეების ჯამი შესაკრებთა დისპერსიების ჯამის ტოლია), გვაქვს:

$$D(\bar{X}) = D\left(\frac{X_1 + X_2 + \cdots + X_n}{n^2}\right) = \frac{D(X_1) + D(X_2) + \cdots + D(X_n)}{n^2}.$$

თუ მხედველობაში მივიღებთ, რომ პირობის თანახმად თითოეული სიდიდის დისპერსია D-ს ტოლია, მივიღებთ

$$\mathrm{D}(\bar{\mathrm{X}}) = \mathrm{nD}/\mathrm{n}^2 = \mathrm{D}/\mathrm{n}.$$

3. n ერთნაირად განაწილებული წყვილ-წყვილად დამოუკიდებელი *შემთხვევითი სიდიდეების საშუალო არითმეტიკულის საშუალო კვადრატული გადახრა $\sqrt{n}$ -ჯერ ნაკლებია თითოეული სიდიდის σ საშუალო კვადრატულ გადახრაზე:*

$$\sigma(\bar{X}) = \sigma/\sqrt{\mathrm{n}}. \tag{8.9.2}$$

დამტკიცება. რამდენადაც $D(\bar{X}) = D/n$, ამდენად $\sigma(\bar{X})$ საშუალო კვადრატული გადახრა ტოლია

$$\sigma(\bar{X}) = \sqrt{\mathrm{D}(\bar{\mathrm{X}})} = \sqrt{\mathrm{D}/\mathrm{n}} = \sqrt{D}/\sqrt{n}.$$

ფორმულებიდან (8.9.1) და (8.9.2)-დან მიიღება ზოგადი შედეგი: თუ გავიხსენებთ, რომ დისპერსია და საშუალო კვადრატული გადახრა წარმოადგენენ შემთხვევითი სიდიდის გაბნევის მახასიათებლებს, ვასკვნით, რომ საკმაოდ დიდი რაოდენობის წყვილ-წყვილად დამოუკიდებელი შემთხვევით სიდიდეებს ცალკეულ სიდიდეზე უფრო ნაკლები გაბნევა აქვთ.

8.10. საწყისი და ცენტრალური თეორიული მომენტები

განვიხილოთ X დისკრეტული შემთხვევითი სიდიდე, რომელიც შემდეგი განაწილების კანონითაა მოცემული:

X	1	2	5	100
p	0,6	0,2	0,19	0,01.

ვიპოვოთ X -ის მათემატიკური ლოდინი:

$$M(X) = 1 \cdot 0{,}6 + 2 \cdot 0{,}2 + 5 \cdot 0{,}19 + 100 \cdot 0{,}01 = 2{,}95.$$

დავწეროთ X^2-ის განაწილების კანონი:

X^2	1	4	25	10000
p	0,6	0,2	0,19	0,01.

ვიპოვოთ X^2-ის მათემატიკური ლოდინი:

$$M(X^2) = 1 \cdot 0{,}6 + 4 \cdot 0{,}2 + 25 \cdot 0{,}19 + 10\,000 \cdot 0{,}01 = 106{,}15.$$

ვხედავთ, რომ $M(X^2)$ გაცილებით მეტია $M(X)$-ზე. ეს იმით აიხსნება, რომ კვადრატში აყვანის შემდეგ X^2 სიდიდის შესაძლო მნიშვნელიბები, რომლებიც შეესაბამება X სიდიდის $x = 100$ მნიშვნელობას 10 000-ის ტოლი გახდა, ე.ი. მნიშვნელოვნად გაიზარდა; ამ მნიშვნელობის ალბათობა კი მცირეა (0,01).

ამგვარად, $M(X)$ დან $M(X^2)$ -ზე გადასვლამ საშუალება მოგვცა უკეთ გავითვალისწინოთ იმ შესაძლო მნიშვნელობის გავლენა მათემატიკურ ლოდინზე, რომელიც დიდია და გააჩნია მცირე ალბათობა. რა თქმა უნდა, თუ X სიდიდეს ექნებოდა შედარებით დიდი და ნაკლებ მოსალოდნელი მნიშვნელობა, მაშინ X^2 სიდიდეზე გადასვლა, მით უფრო X^3 , X^4 და ა.შ. სიდიდეებზე გადასვლა შესაძლებლობას მოგვცემდა უფრო „გაგვეძლიერებინა“ ამ დიდი, მაგრამ ნაკლებ მოსალოდნელი შესაძლო მნიშვნელობების „როლი“. აი რატომაა მიზანშეწონილი განვიხილოთ შემთხვევითი სიდიდის მთელი დადებითი ხარისხის მათემატიკური ლოდინი (არა მარტო დისკრეტულის, არამედ უწყვეტისაც).

X შემთხვევითი სიდიდის k *რიგის საწყის მომენტს* უწოდებენ X^k სიდიდის მათემატიკურ ლოდინს:

$$\gamma_k = M(X^k).$$

კერძოდ,

$$\gamma_1 = M(X),\ \gamma_2 = M(X^2).$$

ამ მომენტების გამოყენებით, დიპერსიის გამოსათვლელი ფორმულა $D(X) = M(X^2) - [M(X)]^2$ შეიძკება ასე ჩაიწეროს: $D(X) = \gamma_1 - \gamma_1^2$. (8. 10.1)

X შემთხვევითი სიდიდის მომენტების გარდა მიზანშეწონილია განვიხილოთ $X - M(X)$ გადახრის მომენტებიც.

X შემთხვევითი სიდიდის k *რიგის ცენტრალურ მომენტს* უწოდებენ $(X - M(X))^k$ სიდიდის მათემატიკურ ლოდინს:

$$\mu_k = M\left[\left(X - M(X)\right)^k\right].$$

კერძოდ,

$$\mu_1 = M[X - M(X)] = 0, \quad (8.10.2)$$

$$\mu_2 = M\left[\left(X - M(X)\right)^2\right] = D(X). \quad (8.10.3)$$

ადვილი გამოსაყვანია დამოკიდებულება, რომელიც საწყის და ცენტრალურ მომენტებს ერთმანეთთან აკავშირებს. მაგალითად, (8.10.1)-ის და (8.10.3)-ის შედარებით მივიღებთ

$$\mu_2 = \nu_2 - \nu_1^2.$$

ცენტრალური და საწყისი მომენტების განსაზღვრების გამოყენებით რთული არაა შემდეგი ფორმულების მიღება:

$$\mu_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3,$$
$$\mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4.$$

უფრო მაღალი რიგის მომენტები იშვიათად გამოიყენება.

შენიშვნა. აქ განხილულ მომენტებს უწოდებენ *თეორიულს*. თეორიული მომენტებისაგან განსხვავებით, იმ მომენტებს, რომლებიც დაკვირვების შედეგად მიღებული მონაცემებით გამოითვლება, უწოდებენ *ემპირიულს*.

თავი 9. დიდ რიცხვთა კანონი

9.1. წინასწარი შენიშვნები

როგორც უკვე აღინიშნა, წინასწარ შეუძლებელია დანამდვილებით ითქვას ცდის შედეგად თუ რომელ შესაძლო მნიშვნელობას მიიღებს შემთხვევითი სიდიდე; ეს მრავალ შემთხვევით თუ არაშემთხვევით მიზეზზეა დამოკიდებული, რომელთა გათვალისწინება შეუძლებელია. შეიძლება მოგვეჩვენოს, რომ რამდენადაც თითოეულ ამ შემთხვევით სიდიდეზე ჩვენ საკმაოდ მწირ ინფორმაციას ვფლობთ, საკმარისად დიდი რაოდენობის შემთხვევითი სიდიდების ჯამის და ქცევის კანონზომიერების დადგენას ვერ შევძლებთ. მაგრამ ეს ასე არაა. აღმოჩნდა, რომ გარკვეულ, შედარებით გაფართოებულ პირობებში, საკმაოდ დიდი რაოდენობის შემთხვევითი სიდიდეების ჯამური ქცევა თითქმის კარგავს თავის შემთხვევით ხასიათს და ხდება კანონზომიერი.

პრაქტიკაში ძალიან მნიშვნელოვანია იმ პირობების ცოდნა, რომლის შესრულების დროს ძალიან ბევრი შემთხვევითი მიზეზის ერთობლივ მოქმედებას მივყავართ შედეგამდე, რომელიც თითქმის არაა დამოკიდებული შემთხვევაზე, რამდენადაც საშუალებას იძლევა წინასწარ განვჭვრიტოთ მოვლენის განვითარება. ეს პირობები მითითებულია თეორემებში, რომელთა საერთო სახელია დიდ რიცხთა კანონი. მათ მიეკუთვნება ჩებიშევის და ბერნულის თეორემები (სხვა თეორმებს აქ არ განვიხილავთ). ამ თეორემების დასამტკიცებლად ვიყენებთ ჩებიშევის უტოლობას.

9.2. ჩებიშევის უტოლობა

ჩებიშევის უტოლობა სამართლიანია დისკრეტული და უწყვეტი შემთხვევითი სიდიდეებისათვის. სიმარტივისათვის შემოვიფარგლებით ამ თეორემის დისკრეტული შემთხვევითი სიდიდეებისათვის დამტკიცებით.

განვიხილოთ შემდეგი განაწილების ცხრილით მოცემული X დისკრეტული შემთხვევითი სიდიდე:

$$X \quad x_1 \; x_2 \; \dots . \; x_n$$

$$p \quad p_1 \, p_2 \; \dots \; p_n .$$

მიზნად დავისახოთ, შევაფასოთ იმის ალბათობა, რომ შემთხვევითი სიდიდის გადახრა თავისი მათემატიკური ლოდინიდან აბსოლუტური სიდიდით არ აღემატება დადებით ε რიცხვს. თუ ε საკმარისად მცირეა, მაშინ ჩვენ ვაფასებთ იმის ალბათობას, რომ X იღებს თავის მათემატიკურ ლოდინთან საკმაოდ ახლო მნიშვნელობას. ჩებიშევმა დაამტკიცა უტოლობა, რომელიც საშუალებას იძლევა ამას მივცეთ ჩვენთვის საინტერესო შეფასება.

ჩებიშევის უტოლობა. *იმის ალბათობა, რომ X შემთხვევითი სიდიდის გადახრა თავისი მათემატიკური ლოდინიდან აბსოლუტური სიდიდით ნაკლებია დადებით ε რიცხვზე, არაა ნაკლები $1 - D(X)/\varepsilon^2$ -ზე;*

$$P(|X - M(X)| < \varepsilon) \geq 1 - D\,(X)/\varepsilon^2.$$

დამტკიცება. რამდენადაც უტოლობაში შემავალი ხდომილებები $|X - M(X)| < \varepsilon$ და $|X - M(X)| \geq \varepsilon$ მოპირდაპირე ხდომილებებია, ამდენად, მათი ალბათობების ჯამი ერთის ტოლია, ე.ი.

$$P(|X - M(X) < \varepsilon|) + P(|X - M(X)| \geq \varepsilon) = 1$$

აქედან გამომდინარეობს

$$P(|X - M(X)| < \varepsilon) = 1 - P(|X - M(X)| \geq \varepsilon) \tag{9.2.1}$$

ამგვარად, ამოცანა დადის $P(|X - M(X)| \geq \varepsilon)$ ალბათობის გამოთვლაზე.

დავწეროთ X შემთხვევითი სიდიდის დისპერსიის გამოსახულება:

$$D(X) = [X_1 - M(X)]^2 p_1 + [X_2 - M(X)]^2 p_2 + \cdots + [X_n - M(X)]^2 p_n$$

ცხადია ამ ჯამის ყველა შესაკრები არაუარყოფითია.

გადავაგდოთ ის შესაკრებები, რომელთათვისაც $|x_i - M(X)| < \varepsilon$ (დანარჩენი შესაკრებებისათვის $|x_i - M(X)| \geq \varepsilon$), რის შედეგადაც ჯამი შეიძლება მხოლოდ შემცირდეს. სიცხადისათვის შევთანხმდეთ, რომ პირველი k შესაკრები იგულისხმება გადაგდებულად (ზოგადობის დარღვევის გარეშე, შეიძლება ჩავთვალოთ, რომ განაწილების ცხრილში შემთხვევითი სიდიდეები სწორედ ასეთი რიგითაა დანომრილი). ამგვარად,

$$D(X) \geq [x_{k+1} - M(X)]^2 p_{k+1} + [x_{k+2} - M(X)]^2 p_{k+2} + \cdots + [x_n - M(X)]^2 p_n.$$

შევნიშნოთ, რომ უტოლობის ორივე მხარე $\left|x_j - M(X)\right| \geq \varepsilon$ $(j = k+1, k+2, \dots, n)$ დადებითია, ამიტომ მათი კვადრატში აყვანით მივიღებთ ტოლფას უტოლობას $|x_j - M(X)|^2 \geq \varepsilon^2$. ვისარგებლოთ ამ შენიშვნით და დარჩენილ ჯამში $|x_j - M(X)|^2$ მამრავლებიდან თითოეული შევცვალოთ ε^2 რიცხვით (ამით უტოლობა მხოლოდ უფრო გაძლიერდება), მივიღებთ

$$D(X) \geq \varepsilon^2(p_{k+1} + p_{k+2} + \cdots + p_n). \quad (9.2.2)$$

შეკრების თეორემის თანახმად, $p_{k+1} + p_{k+2} + \cdots + p_n$ ალბათობათა ჯამი არის იმის ალბათობა, რომ X იღებს $x_{k+1}, x_{k+1}, \dots, x_n$ მნიშვნელობებიდან ერთ-ერთს, არა აქვს მნიშვნელობა რომელს, ხოლო ნებისმიერი მათგანის დროს გადახრა აკმაყოფილებს უტოლობას $\left|x_j - M(X)\right| \geq \varepsilon$. აქედან გამომდინარეობს, რომ $p_{k+1} + p_{k+2} + \cdots + p_n$ ჯამი გამოხატავს შემდეგ ალბათობას:

$$P(|X - M(X)| \geq \varepsilon).$$

ეს მოსაზრება საშუალებას იძლევა (9.2.2) უტოლობა ასე გადავწეროთ:

$$D(X) \geq \varepsilon^2 P(|X - M(X)| \geq \varepsilon),$$

ან

$$P(|X - M(X)| \geq \varepsilon) \leq D(X)/\varepsilon^2 \quad (9.2.3)$$

თუ (9.2.3)-ს ჩავსვამთ (9.2.1)-ში, საბოლოოდ მივიღებთ

$$P(|X - M(X)| < \varepsilon) \geq 1 - D(X)/\varepsilon^2,$$

რაც უნდა დაგვემტკიცებინა.

შენიშვნა. პრაქტიკისათვის ჩებიშევის უტოლობას შეზღუდული მნიშვნელობა აქვს, რამდენადაც მისი ნაწილი მოითხოვს უხეშ, ზოგჯერ ტრივიალურ (უინტერესო) შეფასებას. მაგალითად, თუ $D(X) > \varepsilon^2$, და შესაბამისად, $D(X)/\varepsilon^2 > 1$, მაშინ $1 - D(X)/\varepsilon^2 < 0$; ამგვარად ამ შემთხვევაში ჩებიშევის უტოლობა მიუთითებს მხოლოდ იმას, რომ გადახრის ალბათობა არაუარყოფითია, ეს კი ისედაც ცხადია, რამდენადაც ნებისმიერი ხდომილების ალბათობა არაუარყოფითი რიცხვია.

თეორიულად კი ჩებიშევის უტოლობის მნიშვნელობა ძალიან დიდია. ქვემოთ ვისარგებლებთ ამ უტოლობით ჩებიშევის თეორემის დასამტკიცებლად.

9.3 ჩებიშევის თეორემა

ჩებიშევის ***თეორემა.*** *თუ* $X_1, X_2, \dots, X_n \dots$ *წყვილწყვილად დამოუკიდებელი შემთხვევითი სიდიდეებია, ამასთან მათი დისპერსია თანაბრად შემოსაზღვრულია (არ აღემატება მუდმივ C რიცხვს), მაშინ როგორი მცირეც არ უნდა იყოს* ε *დადებითი რიცხვი, შემდეგი უტოლობის* $\left|\frac{X_1+X_2+\cdots+X_n}{n} - \frac{M(X_1)+M(X_2)+\cdots+M(X_n)}{n}\right| < \varepsilon$ *ალბათობა იქნება რაგინდ ახლოს ერთთან, თუ შემთხვევითი სიდიდეების რაოდენობა საკმაოდ დიდია; ანუ*

$$\lim_{n\to\infty} P\left(\left|\frac{X_1+X_2+\cdots+X_n}{n} - \frac{M(X_1)+M(X_2)+\cdots+M(X_n)}{n}\right| < \varepsilon\right) = 1.$$

ამგვარად, ჩებიშევის თეორემა ამტკიცებს, რომ თუ დამოუკიდებლი შემთხვევითი სიდიდეების საკმაო რაოდენობა განიხილება, რომელთაც შემოსაზღვრული დისპერსიები გააჩნიათ, მაშინ ხდომილება, რომ შემთხვევითი სიდიდეების საშუალო მნიშვნელობის გადახრა მათი მათემატიკური ლოდინის საშუალო არითმეტიკულიდან იქნება აბსოლუტური მნიშვნელობით რაგინდ მცირე, შეიძლება ჩაითვალოს ჭეშმარიტად.

დამტკიცება. შემოვიტანოთ ახალი შემთხვევთი სიდიდე - შემდეგი შემთხვევითი სიდიდეების საშუალო არითმეტიკული

$$\bar{X} = (X_1 + X_2 + \cdots + X_n)/n.$$

ვიპოვოთ $\bar{X}$-ის მათემატიკური ლოდინი. მათემატიკური ლოდინის თვისებების გამოყენებით (მუდმივი მამრავლი შესაძლოა გატანილი იქნას მათემატიკური ლოდინის ნიშნის გარეთ, ჯამის მათემატიკური ლოდინი შესაკრებთა მათემატიკური ლოდინების ჯამის ტოლია), მივიღებთ

$$M\left(\frac{X_1+X_2+\cdots+X_n}{n}\right) = \frac{M(X_1)+M(X_2)+\cdots+M(X_n)}{n}. \qquad (9.3.1)$$

თუ $\bar{X}$ სიდიდის მიმართ ჩებიშევის უტოლობას გამოვიყენებთ, მივიღებთ

$$P\left(\left|\frac{X_1+X_2+\cdots+X_n}{n} - M\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)\right| < \varepsilon\right) \geq 1 - \frac{D\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)}{\varepsilon^2}$$

თუ გავითვალისწინებთ (9.3.1) დამოკიდებულებას

$$P\left(\left|\frac{X_1+X_2+\cdots+X_n}{n}-\frac{M(X_1)+M(X_2)+\cdots+M(X_n)}{n}\right|<\varepsilon\right)\geq$$

$$\geq 1-\frac{D\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)}{\varepsilon^2}, \qquad (9.3.2)$$

დისპერსიის თვისებების გამოყენებით (მუდმივი სიდიდე შეიძლება გამოტანილ იქნას დისპერსიის ნიშნის გარეთ მისი კვადრატში აყვანით; დამოუკიდებელი სიდიდეების ჯამი შესაკრებთა დისპერსიების ჯამის ტოლია), მივიღებთ

$$D\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)=\frac{D(X_1)+D(X_2)+\cdots+D(X_n)}{n^2}$$

პირობის თანახმად ყველა შემთხვევითი სიდიდის დისპერსიები შემოსაზღვრულია C მუდმივი რიცხვით, ე.ი. ადგილი აქვს უტოლობებს: $D(X_1)\leq C$; $D(X_2)\leq C$;...; $D(X_n)\leq C$, ამიტომ $(D(X_1)+D(X_2)+\cdots++D\,(X_n)/n^2)\leq(C+C+\cdots+C)/n^2=n\,C/n^2=C/n$

ამგვარად,

$$D\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)\leq\frac{C}{n} \qquad (9.3.3)$$

თუ (9.3.3)-ის მარჯვენა მხარეს ჩავსვამთ (9.3.2) უტოლობაში (რის შედეგადაც უკანასკნელი მხოლოდ გაძლიერდება), გვექნება

$$P\left(\left|\frac{X_1+X_2+\cdots+X_n}{n}-\frac{M(X_1)+M(X_2)+\cdots+M(X_n)}{n}\right|<\varepsilon\right)\geq 1-\frac{C}{n\varepsilon^2}$$

აქედან, ზღვარზე გადასვლით, როცა $n\to\infty$, მივიღებთ

$$\lim_{n\to\infty}P\left(\left|\frac{X_1+X_2+\cdots+X_n}{n}-\frac{M(X_1)+M(X_2)+\cdots+M(X_n)}{n}\right|<\varepsilon\right)\geq 1$$

და ბოლოს, გავითვალისწინებთ რა, რომ ალბათობა ერთს ვერ გადააჭარბებს, საბოლოოდ შეგვიძლია დავწეროთ

$$\lim_{n\to\infty}P\left(\left|\frac{X_1+X_2+\cdots+X_n}{n}-\frac{M(X_1)+M(X_2)+\cdots+M(X_n)}{n}\right|<\varepsilon\right)=1.$$

თეორემა დამტკიცებულია.

როდესაც ზემოთ ჩამოვაყალიბეთ ჩებიშევის თეორემა, ჩვენ ვვარაუდობდით, რომ შემთხვევით სიდიდეებს ჰქონდათ სხვადასხვა მათემატიკური ლოდინები. პრაქტიკაში ხშირად ხდება, როცა შემთხვევით სიდიდეებს ერთიდაიგივე მათემატიკური ლოდინი აქვთ. ცხადია, თუ ისევ დავუშვებთ, რომ ამ სიდიდეების დისპერსიები შემოსაზღვ-რულია, მათ მიმართ ჩებიშევის თეორემა მისაღებია.

თითოეული შემთხვევითი სიდიდის მათემატიკუერი ლოდინი აღვნიშნოთ a-თი; განსახილველ შემთხვევაში, როგორც ადვილი დასანახია, საშუალო არითმეტიკულის მათემატიკური ლოდინიც ასევე a-ს ტოლია. ჩვენ შეგვიძლია ჩამოვაყალიბოთ ჩებიშევის თეორემა განხილული კერძო შემთხვევისათვის.

თუ $X_1, X_2, \dots, X_n, \dots$ წყვილწყვილად დამოუკიდებელი შემთხვევითი სიდიდეებია, რომელთაც ერთი და იგივე a მათემატიკური ლოდინები აქვთ, ამასთან მათი დისპერსია თანაბრად შემოსაზღვრულია, მაშინ როგორი მცირეც არ უნდა იყოს $\varepsilon > 0$, შემდეგი უტოლობის ალბათობა

$$\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - a\right| < \varepsilon$$

იქნება რაგინდ ახლოს ერთთან, თუ შემთხვევითი სიდიდეების რაოდენობა საკმარისად დიდია.

სხვა სიტყვებით, თეორემის პირობებში ადგილი ექნება ტოლობას

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - a\right| < \varepsilon\right) = 1.$$

9.4. ჩებიშევის თეორემის არსი

დამტკიცებული თეორემის არსი ასეთია: თუმცა ცალკეულ დამოუკიდებელ შემთხვევით სიდიდეებს შეუძლიათ თავის მათემატიკური ლოდინისაგან საკმაოდ განსხვავებული მნიშვნელობების მიღება, მაგრამ საკმარისად დიდი რაოდენობის შემთხვევითი სიდიდეების საშუალო არითმეტიკული დიდი ალბათობით იღებს მნიშვნელობას გარკვეული მუდმივი რიცხვის მახლობლობაში, კერძოდ, $(M(X_1) + M(X_2) + \cdots + M(X_n))/n$ რიცხვის (ან a რიცხვის მახლობლობაში კერძო შემთხვევაში). სხვა სიტყვებით რომ ვთქვათ, ცალკეულ შემთხვევით სიდიდეებს

შეიძლება ჰქონდეთ მნიშვნელოვანი გაფანტულობა, მაგრამ მათი საშუალო არითმეტიკული მცირედით არის გაფანტული.

ამგვარად, არ შეიძლება დანამდვილებით ვიწინასწარმეტყველოთ თითოეული შემთხვევითი სიდიდე თუ რომელ შესაძლო მნიშვნელობას მიიღებს, მაგრამ შეიძლება ვიწინასწარმეტყველოთ, თუ რა მნიშვნელობებს მიიღებენ მათი საშუალო არითმეტიკულები.

ამრიგად, *საკმარისად დიდი რაოდენობის დამოუკიდებელი შემთხვევითი სიდიდეების საშუალო არითმეტიკული* (რომელთა დისპერსიები თანაბრად შემოსაზღვრულია) *კარგავს შემთხვევით ხასიათს.* ეს იმით აიხსნება, რომ თითოეული სიდიდის საკუთარი მათემატიკური ლოდინიდან გადახრა შეიძლება იყოს როგორც დადებითი ისე უარყოფითი, ხოლო საშუალო არითმეტიკულში ხდება მათი ურთიერთ გაბათილება.

ჩებიშევის თეორემა სამართლიანია არა მარტო დისკრეტული, არამედ უწყვეტი შემთხვევითი სიდიდეებისათვის.

9.5 ბერნულის თეორემა

ვთქვათ, ტარდება n დამოუკიდებელი ცდა, რომელთაგან თითოეულში A ხდომილების მოხდენის ალბათობა p-ს ტოლია. შესაძლებელია განვსაზღვროთ როგორი იქნება A ხდომილების მოხდენის ფარდობითი სიხშირე? ამ კითხვაზე დადებით პასუხს იძლევა ბერნულის თეორემა.

ბერნულის თეორემა. *თუ თითოეულ n დამოუკიდებელი ცდაში A ხდომილების მოხდენის ალბათობა მუდმივია და p-ს ტოლია, მაშინ რაგინდ ახლოსაა ერთთან იმის ალბათობა, რომ ფარდობითი სიხშირის და p ალბათობის სხვაობის აბსოლუტური მნიშვნელობა იქნება რაგინდ მცირე, თუ ცდების რაოდენობა საკმარისად დიდია.*

სხვა სიტყვებით რომ ვთქვათ, თუ ε რაგინდ მცირე დადებითი რიცხვია, მაშინ თეორემის პირობებში ადგილი აქვს ტოლობას

$$\lim_{n\to\infty} P(|m/n - p| < \varepsilon) = 1.$$

დამტკიცება. აღვნიშნოთ X_1-ით დისკრეტული შემთხვევითი სიდიდე, რომელიც პირველ ცდაში ხდომილების მოხდენის რაოდენობაა, X_2-ით მეორეში, ..., X_n -ით n -ურში. ცხადია, რომ თითოეულ შემთხვევით სიდიდეს შეუძლია მიიღოს მხოლოდ ორი მნიშვნელობა: 1 (A ხდომილება მოხდა) p ალბათობით და 0 (A ხდომილება არ მოხდა) $1 - p = q$ ალბათობით.

თუ შემთხვევითი სიდიდეები წყვილ-წყვილად დამოუკიდებლებია და დისპერსიები შემოსაზღვრული, შესაძლებელია გამოვიყენოთ ჩებიშევის თეორემა. მართლაც, $X_1, X_2, \dots, X_n$ სიდიდეების წყვილ-წყვილად დამოუკიდებლობა გამომდინარეობს იქიდან, რომ ცდები დამოუკიდებლებია. ნებისმიერი $X_i (i = 1,2, \dots n)$ სიდიდის დისპერსია pq ნამრავლის ტოლია, რამდენედაც $p + q = 1$, pq ნამრავლი არ აღემატება $1/4$-ს და, აქედან გამომდინარე შემთხვევითი სიდიდეების დისპერსიები შემოსაზღვრულია, მაგალითად $C = 1/4$ რიცხვით.

თუ გამოვიყენებთ ჩებიშევის თეორემას (კერძო შემთხვევა) განხილული შემთხვევითი სიდიდეების მიმართ, გვექნება

$$\lim_{n\to\infty} P(|X_1 + X_2 + \cdots + X_n/n - a| < \varepsilon) = 1.$$

თუ მხედველობაში მივიღებთ, რომ თითოეული X_i სიდიდის a მათემატიკური ლოდინი (ერთ ცდაში ხდომილების მოხდენის რაოდენობის მათემატიკური ლოდინი) ტოლია ხდომილების მოხდენის p ალბათობის (იხ. პარაგრაფი 7.2. მაგალითი 2), მივიღებთ:

$$\lim_{n\to\infty} P(|X_1 + X_2 + \cdots + X_n/n - p| < \varepsilon) = 1.$$

ისღა დაგვრჩენია ვაჩვენოთ, რომ $X_1 + X_2 + \cdots + X_n/n$ წილადი ცდებში A ხდომილების მოხდენის ფარდობითი სიხშირის ტოლია. მართლაც, $X_1, X_2, \dots, X_n$ სიდიდეებიდან შესაბამის ცდაში A ხდომილების მოხდენისას ერთის ტოლ

მნიშვნელობას იღებს, აქედან $X_1 + X_2 + \cdots + X_n$ ჯამი ტოლია n ცდაში ხდომილების მოხდენის m რაოდენობის ე.ი.

$$(X_1 + X_2 + \cdots + X_n/n) = m/n.$$

ამ ტოლობის გათვალისწინებით, საბოლოოდ მივიღებთ

$$\lim_{n\to\infty} P(|m/n - p| < \varepsilon) = 1.$$

თავი 10. მარტივი წრფივი რეგრესია და კორელაცია

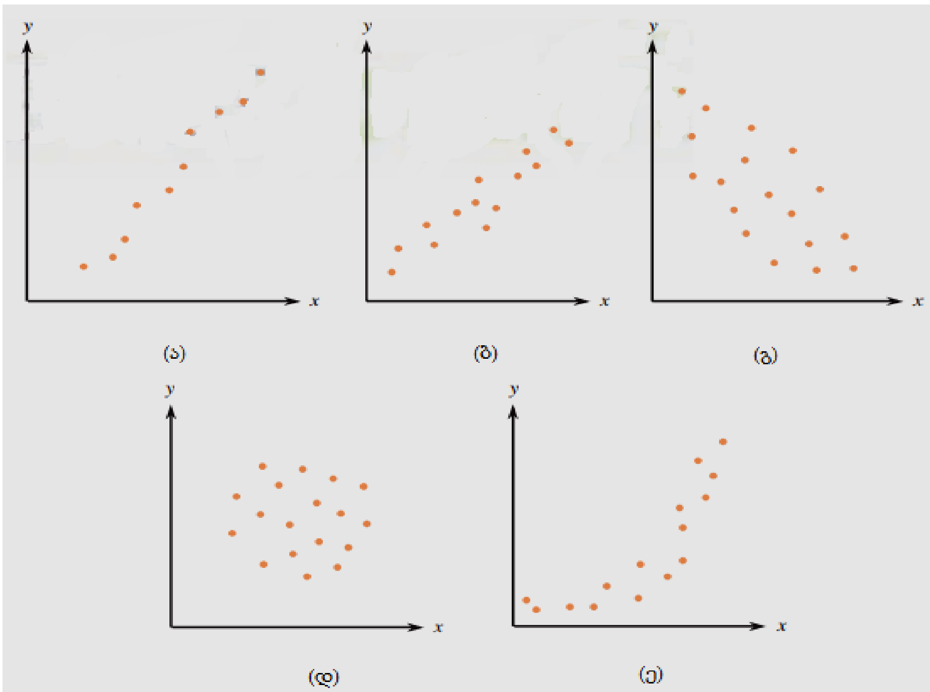
10.1 კორელაცია

მკვლევარს ხშირად აინტერესებს, თუ როგორ არის დაკავშირებული პოპულაციის ინდივიდების ან ობიექტების ორი ან მეტი თვისება ერთმანეთთან. მაგალითად, გარემოს დაცვის მკვლევარს შეიძლება სურდეს იცოდეს, როგორ იცვლება ნიადაგში ტყვიის შემცველობა ძირითადი მაგისტრალიდან დაშორების მიხედვით. სკოლამდელი განათლების მკვლევარებმა შეიძლება გამოიკვლიონ, თუ როგორ არის დაკავშირებული ლექსიკის ზომა ასაკთან.

ბივარიაციული რიცხვითი მონაცემების გაბნევა იძლევა ვიზუალურ წარმოდგენას რამდენად მჭიდროდ არის დაკავშირებული x და y მნიშვნელობები. თუმცა, ზუსტი განცხადებების გასაკეთებლად და მონაცემებიდან დასკვნების გამოსატანად, ჩვენ უნდა შემოვიტანოთ ახალი ცნება, კერძოდ **კორელაციის კოეფიციენტი.** იგი წარმოადგენს (x, y) წყვილების სიმრავლეში x და y მნიშვნელობებს შორის ურთიერთობის სიძლიერის რიცხვითი შეფასებას. ამ თავში წარმოგიდგენთ ყველაზე ხშირად გამოყენებულ კორელაციის კოეფიციენტს.

ნახაზ 1-ზე მოცემულია გაბნევის რამდენიმე დიაგრამა, რომლებიც აჩვენებენ განსხვავებულ ურთიერთობას x და y მნიშვნელობებს შორის. ნახაზზე 1(ა) დიაგრამა მიუთითებს ძლიერ დადებით კავშირზე x და y-ს შორის; გრაფიკის წერტილების ყველა წყვილითვის x-ის უფრო დიდი მნიშვნელობის მქონე წერტილს ასევე აქვს y-ის უფრო დიდი მნიშვნელობა, ანუ x-ის ზრდასთან ერთად იზრდება y-ის მნიშვნელობა. დიაგრამა 1(ბ) ნახაზზე გვიჩვენებს y-ის ძლიერ ტენდენციას გაიზარდოს x-ის მსგავსად, მაგრამ არის რამდენიმე გამონაკლისი. მაგალითად, ნახაზის უკიდურეს ზედა მარჯვენა კუთხეში ორი წერტილის x და y მნიშვნელობები მოპირდაპირე მიმართულებით მიდიან (x იზრდება, მაგრამ y მცირდება). მიუხედავად ამისა, ასეთი სიუჟეტი კვლავ მიუთითებს საკმაოდ ძლიერ დამოკიდებულებაზე. ნახაზი 1 (გ) გვიჩვენებს რომ x და y უარყოფითად არიან დაკავშირებულნი - რამდენადაც იზრდება x , იმდენად მცირდება y. ამ გრაფიკზე უარყოფითი კავშირი ისე ძლიერი არ არის, როგორც დადებითი კავშირი ნახაზ 1 (ბ)-ზე, თუმცა ორივე გრაფიკი გამოხატავს მკვეთრად გამოხატულ წრფივ

კანონზომიერებას. გრაფიკი ნახაზ 1 (დ) გამოსახავს სუსტ კავშირს x და y-ს შორის; ამ შემთხვევაში *y* არ იზრდება ან მცირდება *x* -ის ზრდასთან ერთად. და ბოლოს, როგორც ნაჩვენებია ნახაზ 1(ე)-ზე, გაფანტვის დიაგრამა შეიძლება გვიჩვენებდეს ძლიერ დამოკიდებულებას, რომელსაც გააჩნია მრუდწირული და არა წრფივი ხასიათი.



ნახ. 1 გაფანტვის დიაგრამები, რომლებიც ასახავენ სხვადასხვა სახის მიმართებებს: (ა) დადებითი წრფივი მიმართება; (ბ) კიდევ ერთი დადებითი წრფივი მიმართება; (გ) უარყოფითი წრფივი მიმართება; (დ) არანაირი კავშირი; (ე) მრუდწირული მიმართება.

სურათი 1(გ) ვარაუდობს, რომ x და y უარყოფითად არის დაკავშირებული ერთმანეთთან - როცა x იზრდება, y კლებულობს. უარყოფითი კავშირი ამ ნახაზში არ არის ისეთი ძლიერი, როგორც დადებითი კავშირი ნახაზზე 1(ბ), თუმცა ორივე ნახაზი აჩვენებს კარგად განსაზღვრულ წრფივ დამოკიდეულებას. ნახაზ 1(დ)-ის დიაგრამა მიუთითებს x-სა და y-ს შორის კავშირის არარსებობაზე; არ არსებობს ტენდენცია, რომ y არც გაიზარდოს და არც შემცირდეს x გაზრდისას. და ბოლოს, როგორც ეს ილუსტრირებულია ნახაზ 1(ე)-ზე, გაბნევის დიაგრამამ შეიძლება აჩვენოს ძლიერი მრუდწირული და არა წრფივი დამოკიდებულება.

პირსონის შერჩევის კორელაციის კოეფიციენტი

ვთქვათ $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ აღნიშნავს (x, y) წყვილის შერჩევას. განვიხილოთ ყოველი x მნიშვნელობა შესაბამისი z მაჩვენებლით, z_x ($\bar{x}$ -ის გამოკლებით და s_x- ზე გაყოფით) და იგივენაირად ყოველი y მნიშვნელობა ჩავანაცვლოთ მისი z მაჩვენებლით. ყურადღება მიაქციეთ, რომ x სიდიდეებს რომელიც უფრო დიდია ვიდრე $\bar{x}$, ექნებათ დადებითი z მაჩვნებელი და ხოლო $\bar{x}$-ზე ნაკლებებს - უარყოფითი z მაჩვენებელი. წრფივი დამოკიდებულების სიძლიერის საერთო საზომს ეწოდება პირსონის შერჩევის კორელაციის კოეფიციენტი, იგი ეფუძნება z_x და z_y ნამრავლების ჯამს ყოველი დაკვირვებისათვის ორმხრივ მონაცემთა $\sum z_x z_y$ სიმრავლეში.

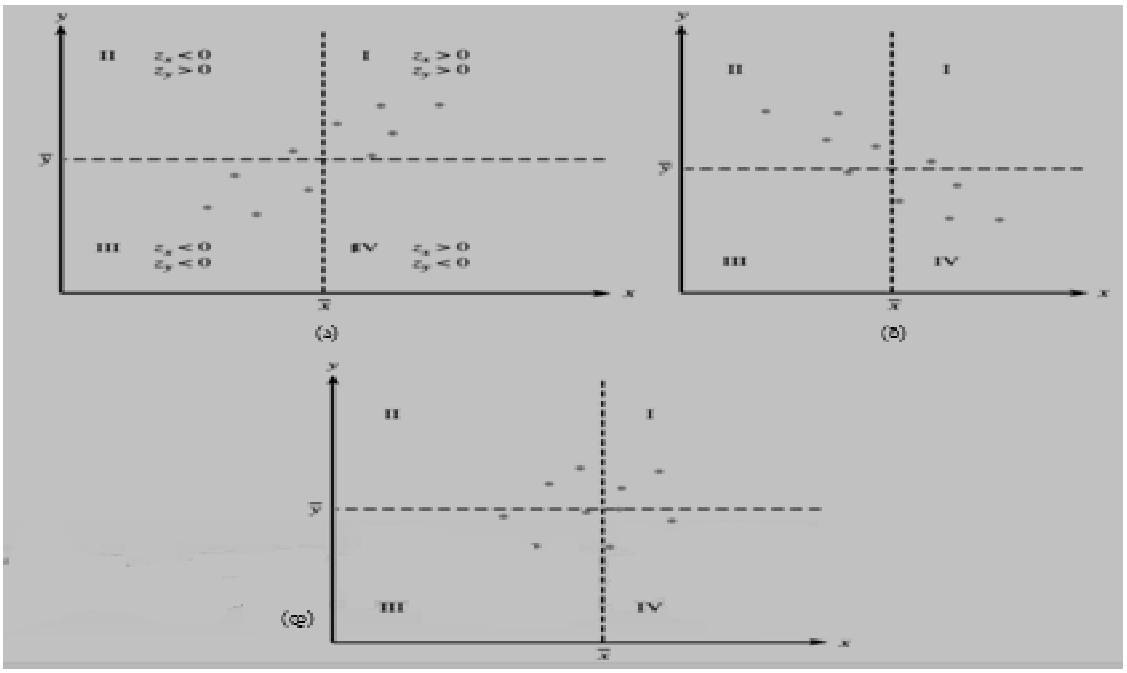
z მაჩვენებელი, რომელიც შეესაბამება რომელიმე სიდიდეს გამოითვლება შემდეგნაირად:

$$\text{z } მაჩვენებელი = \frac{სიდიდე - საშუალო}{სტანდარტული\ გადახრა}$$

z მაჩვენებელი გვეუბნება რამდენი სტანტარტული გადახრითაა განსხვავებული სიდიდე საშუალოსგან. იგი დადებითია ან უაყოფითი იქიდან გამომდინარე, სიდიდე საშუალოზე მეტია თუ ნაკლებია. საშუალოს გამოკლების და სტანდარტულ გადახრაზე გაყოფის პროცესს ზოგჯერ უწოდებენ სტანდარტიზაციას და z მაჩვენებელი წარმოადგენს სტანდარტიზაციის მაჩვენებლის მაგალითს.

ნახ. 2 (ა)-ზე მოცემული გაფანტვის დიაგრამა გვიჩვენებს ძლიერ დადებით ურთიერთკავშირს. ვერტიკალური ხაზი $\bar{x}$-ის გასწვრივ და ჰორიზონტალური ხაზი $\bar{y}$–ის გასწვრივ დიაგრამას ყოფს ოთხ ნაწილად. x და y აჭარბებს თავის საშუალო მნიშვნელობას I ნაწილში, ასე რომ z მაჩვენებელი x -სთვის და z მაჩვენებელი y-სთვის ორივე დადებითი რიცხვია. აქედან გამომდინარეობს, რომ $z_x z_y$ დადებითია. z მაჩვენებლების ნამრავლი ასევე დადებითია III ნაწილის ყველა წერტილისთვის, ვინაიდან ორივე z მაჩვენებელი უარყოფითია და მათი ნამრავლი დადებითი იქნება. დანარჩენ ორ ნაწილში, ერთი z მაჩვენებელი დადებითია და მეორე ურაყოფითი,

ამიტომ $z_x z_y$ უარყოფითია. მაგრამ ვინაიდან წერტილების უმეტესობა მოხვდა ნაწილ I და III-ში , z მაჩვენებლის ნამრავლს აქვს ტენდენცია იყოს დადებითი. ამგვარად, ნამრავლების ჯამი არის დიდი დადებითი რიცხვი.



ნახ.2. გაფანტვის დიაგრამის დათვალიერება z_x და z_y-ის ნიშნების მიხედვით (ა) დადებითი კავშირი; (ბ) უარყოფითი კავშირი, (გ) სუსტი კავშირი.

მონაცემთა შესახებ იგივე მსჯელობა მოცემულია ნახ. 2 (ბ)-ზე, რომელიც გამოსახავს ძლიერ უარყოფით კავშირს და გულისხმობს, რომ $\sum z_x z_y$ დიდი უარყოფითი რიცხვია. როდესაც არ არის ძლიერი კავშირი, როგორც ნახ. 2. (გ)-ში, დადებით და უარყოფით ნამრავლები აბათილებენ ერთმანეთს და წარმოქმნიან მნიშვნელობას $\sum z_x z_y$, რომელიც ახლოსაა ნულთან. საბოლოოდ,როგორც ჩანს $\sum z_x z_y$ არის x და y შორის კავშირის ხარისხის მნიშვნელოვანი საზომი; იგი შეიძლება იყოს დიდი დადებითი რიცხვი, დიდი უარყოფითი რიცხვი ან ნულთან მიახლოებული რიცხვი, იმაზე დამოკიდებულებით არის ძლიერად დადებითი, ძლიერად უარყოფითი თუ სუსტი წრფივი დამოკიდებულება.

პირსონის შერჩევის კორელაციის კოეფიციენტი აღინიშნება r-ით და მიიღება $\sum z_x\, z_y$-ის $(n-1)$-ზე გაყოფით.

განსაზღვრება: პირსონის შერჩევის კორელაციის კოეფიციენტი r მოიცემა შედეგნაირად:

$$r = \frac{\sum z_x z_y}{n-1}$$

მიუხედავად იმისა, რომ არსებობს რამდენიმე განსხვავებული კორელაციის კოეფიციენტი, პირსონის კორელაციის კოეფიციენტი ყველაზე ხშირად გამოიყენება, ამიტომ სახელწოდება "პირსონის" ხშირად გამოტოვებულია და მას უბრალოდ **კორელაციის კოეფიციენტად** მოიხსენიებენ.

r-ის თვისებები:

1. r -ის მნიშვნელობა არ არის დამოკიდებული თითოეული ცვლადის საზომ ერთეულზე. მაგალითად, თუ x სიმაღლეა და მისი შესაბამისი z მაჩვენებელი იგივეა, იმის მიუხედავად სიმაღლე გამოისახება ინჩებით, მეტრებით თუ დიუმებით, კორელაციის კოეფიციენტი არ იცვლება. კორელაციის კოეფიციენტი ზომავს ორ რიცხვით ცვლადს შორის წრფივი დამოკიდებულების შინაგან ძალას.

2. r -ის სიდიდე არ არის დამოკიდებული იმაზე, თუ რომელ ორ ცვლადს გამოსახავს x.

3. r-ის სიდიდე -1 და $+1$ შორისაა. სიდიდე ზედა ზღვართან ახლოს (+1) გამოსახავს არსებით დადებით კავშირს, მაშინ როცა r ახლოსაა ქვედა ზღვართან (-1), გამოსახავს არსებით უარყოფით კავშირს.

4. კორელაციის კოეფიციენტი $r = 1$ მხოლოდ მაშინ, როცა ყველა წერტილი მონაცემთა გაბნევის დიაგრამაზე ზუსტად ძევს ზევით მიმართული დახრის წრფეზე. ანალოგიურად, r = -1 მხოლოდ მაშინ, როდესაც ყველა წერტილი მონაცემთა გაბნევის დიაგრამაზე ზუსტად ძევს ქვევით მიმართული დახრის წრფეზე. მხოლოდ და მხოლოდ მაშინ, როდესაც შერჩევაში არის x-სა და y-ს შორის სრულყოფილი წრფივი ურთიერთობა, r იღებს ერთ-ერთს მისი ორ შესაძლო უკიდურეს სიდიდეს შორის.

5. r-ის მნიშვნელობა არის საზომი იმისა, თუ რამდენად არის x და y ერთმანეთთან წრფივად დაკავშირებული - ანუ რამდენად ხვდება წერტილები გაბნევის ზოლში წრფესთნ ახლოს. 0-თან ახლოს r მნიშვნელობა არ არეგულირებს x-სა და y-ს შორის ურთიერთობას; შეიძლება არსებობდეს ძლიერი ურთიერთიერთკავშირი, მაგრამ ისეთი, რომელიც არ არის წრფივი.

10.2 . წრფივი რეგრესია: ორგანზომილებიანი მონაცემებისადმი წრფის მორგება

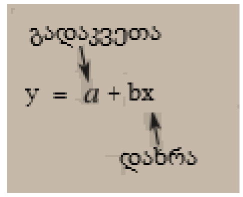
რეგრესიული ანალიზის მიზანია გამოიყენოს ინფორმაცია ერთი x ცვლადის შესახებ, რათა გამოიტანოს რაიმე სახის დასკვნა მეორე y ცვლადთან დაკავშირებით. ხშირად მკვლევარს სურს იწინასწარმეტყველოს y მნიშვნელობა, რომელიც მოჰყვება ერთი დაკვირვების განხორციელების შედეგად განსაზღვრულ x მნიშვნელობას - მაგალითად, იწინასწარმეტყველოს y = პროდუქტის გაყიდვები მოცემულ პერიოდში, როდესაც რეკლამაზე დახარჯული თანხა არის x = $10,000. რეგრესიის ანალიზში ორი ცვლადი თამაშობს განსხვავებულ როლს: y-ს ეწოდება **დამოკიდებული** ან საპასუხო ცვლადი, ხოლო x მოიხსენიება როგორც **დამოუკიდებელი**, **პროგნოზირებადი** ან **განმარტებითი** ცვლადი.

გაბნევის დიაგრამებს ხშირად წრფივი ხასიათი გააჩნიათ. ბუნებრივია, ასეთ შემთხვევებში ცვლადებს შორის ურთიერთობის დადგენა ხდება გრაფიკის წერტილებთან მაქსიმალურად ახლოს მყოფი წრფის მოძიებით. სანამ განვიხილავთ როგორ შეიძლება ამის გაკეთება, მოდით გადავხედოთ რამდენიმე ელემენტარულ ფაქტს წრფეებისა და წრფივი ურთიერთობების შესახებ.

x და y-ს შორის წრფივი მიმართების ზოგადი ფორმაა y = a + bx. კონკრეტული მიმართება განისაზღვრება a და b მნიშვნელობების არჩევით. ამრიგად, ერთ-ერთი ასეთი დამოკიდებულება არის y =10 + 2x; მეორე არის y = 100 - 5x. თუ ჩვენ ავირჩევთ რამდენიმე x მნიშვნელობას და გამოვითვლით y = a + bx-ს თითოეული მნიშვნელობისთვის, მიღებული (x, y) წყვილების შესაბამისი წერტილები გრაფიკზე ზუსტად წრფეზე მოხვდება.

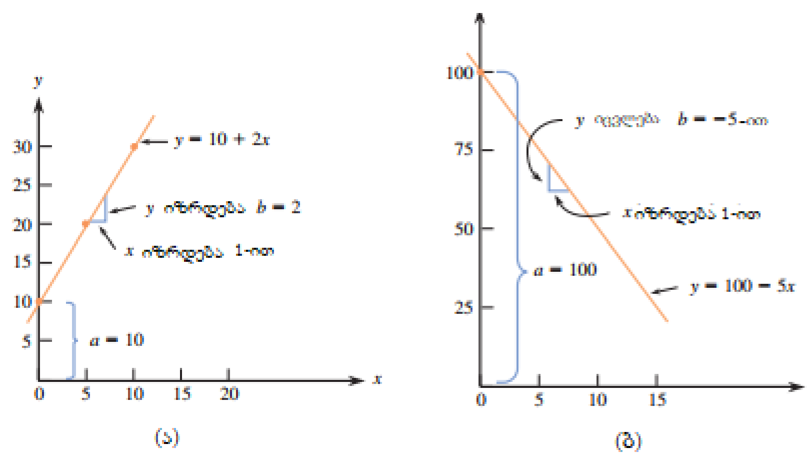
განსაზღვრება:

დამოკიდებულება



არის წრფის განტოლება. b-ს მნიშვნელობა, რომელსაც ეწოდება წრფის დახრილობა, არის ის სიდიდე, რომლითაც y იზრდება, როდესაც x იზრდება 1 ერთეულით. a-ს მნიშვნელობა, რომელსაც ეწოდება წრფის გადაკვეთის წერტილი (ზოგჯერ y ღერძის გადაკვეთა ან ვერტიკალური კვეთისწერტილი) ორდინატათა ღერძთან როცა $x = 0$.

განტოლებას $y = 10 + 2x$ აქვს დახრილობა $b = 2$, ამიტომ x-ის ყოველი ერთი ერთეულით ზრდა დაწყვილებულია y-ის 2-ჯერ გაზრდასთან. როცა $x = 0, y = 10$, ასე რომ, მონაკვეთი, რომელზეც წრფე კვეთს ორდინატათა ღერძს (სადაც $x = 0$) არის 10. ეს გამოსახულია ნახ. 3 (ა)-ზე. $y = 100 - 5x$ განსაზღვრული წრფის დახრილობა ტოლია -5 -ის, ასე რომ y იზრდება $-5 -$ □□ (ანუ, მცირდება 5-ით), როცა x იზრდება 1-ით. წრფე კვეთს ორდინატათა ღერძს, როცა $x = 0$-ს წერტილში $a = 100$-ს. მიღებული წრფე გამოსახულია ნახ. 5.7. (ბ)-ზე.



ნახ.3. ორი წრფის გრაფიკი: (ა) დახრა $b = 2$, გადაკვეთა $a = 10$; (ბ) დახრა $b = -5$, გადაკვეთა $a = 100$.

განსაზღვრება: წრფე $y = a + bx$ -ის ორმაგ მონაცემთან მორგების საუკეთესო ფართოდ გავრცელებული კრიტერიუმია წრფიდან კვადრატული გადახრების ჯამი:

$$\sum[y-(a+bx)]^2 = [y_1-(a+bx_1)]^2 + [y_2-(a+bx_2)]^2 + \cdots + [y_n-(a+bx_n)]^2.$$

უმცირეს კვადრატთა წრფე, რომელსაც ასევე უწოდებენ შერჩევის რეგრესიის წრფეს, არის წრფე რომელიც ამცირებს კვადრატული გადახრების ჯამს.

საბედნიეროდ უმცირეს კვადრატთა წრფის განტოლების ამოხსნა შესაძლებელია ყველა შესაძლო წრფის გადახრების გამოთვლის გარეშე.

უმცირეს კვადრატთა წრფის დახრილობა გამოითვლება ფორმულით

$$b = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2}$$

და y-თან თანაკვეთა არის

$$a = \bar{y} - b\bar{x}/$$

ჩვენ ვწერთ უმცირეს კვადრატთა წრფის განტოლებას შემდეგნაირად

$$\hat{y} = a + bx$$

სადაც, ˆ ნიშანი y-ის თავზე აღნიშნავს, რომ $\hat{y}$ (იკითხება y ქუდიანი) არის y -ის პროგნოზი განტოლებაში x-ის მნიშვნელობის ჩანაცვლებით.

სტატისტიკური პროგრამული პაკეტები ითვლიან უმცირეს კვადრატთა წრფის დახრილობას და თანაკვეთას. თუ დახრილობა და თანაკვეთის გამოთვლა ხდება ხელით, მაშინ გამოიყენება შემდეგი გამოსათვლელი ფორმულა

$$b = \frac{\sum xy - \frac{(\sum x)(\sum y)}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}$$

რეგრესია

უმცირეს კვადრატთა წრფეს ხშირად უწოდებენ **შერჩევის რეგრესიის წრფეს.** ეს ტერმინოლოგია მომდინარეობს უმცირეს კვადრატთა წრფისა და პირსონის კორელაციის კოეფიციენტის ურთიერთმიმართებიდან. ამ დამოკიდებულებაში გასარკვევად, პირველ რიგში საჭიროა b დახრილობის ალტერნატიული გამოსახვა და თავად წრფის განტოლება. მცირე ალგებრული გარდაქმნა იძლევა შემდეგს:

$$b = r\left(\frac{s_y}{s_x}\right)$$

$$\hat{y} = \bar{y} + r\left(\frac{s_y}{s_x}\right)(x - \bar{x})$$

სადაც s_x და s_y აღნიშნავენ x და y-ების შერჩევების სტანდარტულ გადახრებს.

არ არის საჭირო ამ ფორმულების გამოყენება რაიმე გამოთვლებში, მაგრამ მათი რამდენიმე მნიშვნელობა მნიშვნელოვანია იმის გასარკვევად, თუ როგორ იცვლება უმცირეს კვადრატთა წრფე.

1. როდესაც წრფის განტოლებაში $x = \bar{x}$ ჩანაცვლდება, მიიღება $\hat{y} = \bar{y}$. ანუ, უმცირეს კვადრატთა წრფე გადის საშუალო წერტილებზე.

2. ვთქვათ, $r = 1$. მაშინ ყველა ის წერტილი ძევს წრფეზე, რომლის განტოლებაა $\hat{y} = \bar{y} + \frac{s_y}{s_x}(x - \bar{x})$.ჩავსვათ მნიშვნელობა $x = \bar{x} + s_x$, რომელიც არის $\bar{x}$ -ზე 1 სტანდარტული გადახრით მეტი:

$$\hat{y} = \bar{y} + \frac{s_y}{s_x}(\bar{x} - s_x - \bar{x}) = \bar{y} + s_y$$

ანუ, თუ $r = 1$ და x არის მის საშუალოზე ერთი სტანდარტული გადახრით მეტი, ჩვენ ვახდენთ იმის პროგნოზსს, რომ y-ის შესაბამისი მნიშვნელობა იქნება მის საშუალოზე ერთი სტანდარტული გადახრით მეტი. ანალოგიურად, თუ $x = \bar{x} - 2s_x$, (საშუალოზე ორი სტანდარტული გადახრით ნაკლებია), მაშინ

$$\hat{y} = \bar{y} + \frac{s_y}{s_x}(\bar{x} - 2s_x - \bar{x}) = \bar{y} - 2s_y,$$

რაც ასევე ორი სტანდარტული გადახრით არის საშუალოზე ნაკლები. თუ $r = -1$ და $x = \bar{x} + s_x$, ნაშინ ვღებულობT $\hat{y} = \bar{y} - s_y$, ასე რომ, პროგნოზირებული y ასევე არის ერთი სტანდარტული გადახრით გადახრილი მისი საშუალოდან, მაგრამ $\bar{y}$-ის იმ მხარის საპირისპიროდ, სადაც x მდებარეობს $\overline{x}$ -თან მიმართებით. საზოგადოდ, თუ x და y არიან სრულყოფილად კორელირებულნი, პროგნოზირებული y მნიშვნელობა, რომელიც დაკავშირებულია მოცემულ x მნიშვნელობასთან, იქნება იგივე რაოდენობის სტანდარტული გადახრით გადახრილი y -ის საშუალოდან, როგორც არის x გადახრილი მისი საშუალო $\overline{x}$ - დან.

3. ახლა დავუშვათ, რომ x და y არ არის სრულყოფილად კორელირებულები. მაგალითად, დავუშვათ $r = 0.5$, ასე რომ, უმცირეს კვადრატთა წრფე მოიცემა შემდეგი განტოლებით

$$\hat{y} = \bar{y} + 0{,}5\left(\frac{s_y}{s_x}\right)(x - \bar{x})$$

$x = \bar{x} + s_x$, ჩასმით მივიღებთ

$$\hat{y} = \bar{y} + 0{,}5\left(\frac{s_y}{s_x}\right)(\bar{x} + s_x - \bar{x}) = \bar{y} + 0{,}5s_y,$$

ანუ $r = -.5$,-ისთვის, როდესაც x მის საშუალოზე ერთი სტანდარტული გადახრით მეტია, ჩვენ შეიძლება ვივარაუდოთ, რომ y იქნება მხოლოდ 0,5 სტანდარტული გადახრით მეტი მის საშუალოზე. ანალოგიურად, ჩვენ შეგვიძლია გავაკეთოთ y-ზე პროგნოზი, როდესაც r უარყოფითია. თუ $r = -0.5$, მაშინ y-ის პროგნოზირებული მნიშვნელობა იქნება მხოლოდ არსებული სტანდარტული გადახრების რაოდენობის ნახევარი, მაგრამ x და პროგნოზირებული y ახლა იქნება მათი შესაბამისი საშუალოების მოპირდაპირე მხარეს.

რეგრესიის (უმცირეს კვადრატთა) წრფის ალტერნატიული ფორმა ხაზს უსვამს, რომ x-ის ცოდნით y-ის პროგნოზირება არ არის იგივე პრობლემა, რაც x-ის პროგნოზირება y-ის ცოდნით. უმცირეს კვადრატთა წრფის დახრილობა x-ის პროგნოზირებისათვის არის $r(s_x/s_y)$ და არა $r(s_y/s_x)$. ხოლო წრფეების გადაკვეთის წერტილები თითქმის ყოველთვის განსხვავებულია. პროგნოზირების მიზნისათვის ვანსხვავებთ არის თუ არა y რეგრესირებული x-ზე, როგორც ჩვენ გავაკეთეთ, თუ x არის რეგრესირებული y-ზე. y-ის რეგრესიის წრფე x-ზე არ უნდა გამოვიყენოთ x-ის პროგნოზირებისათვის, რადგან ის არ არის წრფე, რომელიც ამცირებს კვადრატული გადახრების ჯამს x-ის მიმართულებით.

განსაზღვრება: პროგნოზირებული ან შერჩეული მნიშვნელობები მიიღება უმცირეს კვადრატთა წრფის განტოლებაში თითოეული შერჩევის x-ის მნიშვნელობის ჩასმით. ეს გვაძლევს

$$\hat{y}_1 = \textit{პირველად პროგნოზირებული მნიშვნელობა} = a + bx_1$$

$$\hat{y}_2 = \textit{მეორე პროგნოზირებული მნიშვნელობა} = a + bx_2$$

$$\cdots$$

$$\hat{y}_n = n - \textit{ური პროგნოზირებულიმნიშვნელობა} = a + bx_n$$

ნაშთი უმცირესი კვარდატების წრფიდან არის n რაოდენობის

$$y_1 - \hat{y}_1, y_2 - \hat{y}_2, \dots, y_n - \hat{y}_n$$

თითოეული ნაშთი არის სხვაობა დაკვირვებად y მნიშვნელობასა და შესაბამის პროგნოზირებად y მნიშვნელობას შორის.

10.3. მარტივი წრფივი რეგრესიის მოდელი

წინა პარაგრაფში რეგრესია და კორელაცია წარმოვადგინეთ როგორც (x, y) წყვილისგან შემდგარი ორგანზომილებიანი მონაცემის აღწერისა და განზოგადების მეთოდი. ამ პარაგრაფში ჩვენ განვიხილავთ ორგანზომილებიანი რიცხვითი მონაცემების დასკვნით მეთოდებს, მათ შორის ნდობის ინტერვალს (ინტერვალური შეფასება) საშუალო y მნიშვნელობისთვის, პროგნოზირების ინტერვალს ერთი y მნიშვნელობისთვის და ჰიპოთეზების შემოწმებას კორელაციის ხარისხთან დაკავშირებით (x, y) წყვილების მთელი სიმრავლისათვის.

დეტერმინისტულია ურთიერთობა, როდესაც y -ის მნიშვნელობა მთლიანად განისაზღვრება x დამოუკიდებელი ცვლადის მნიშვნელობით. ასეთი დამოკიდებულება შეიძლება აღიწეროს ტრადიციული მათემატიკური აღნიშვნის გამოყენებით, როგორიცაა $y = f(x)$, სადაც $f(x)$ არის x-ის განსაზღვრული ფუნქცია.

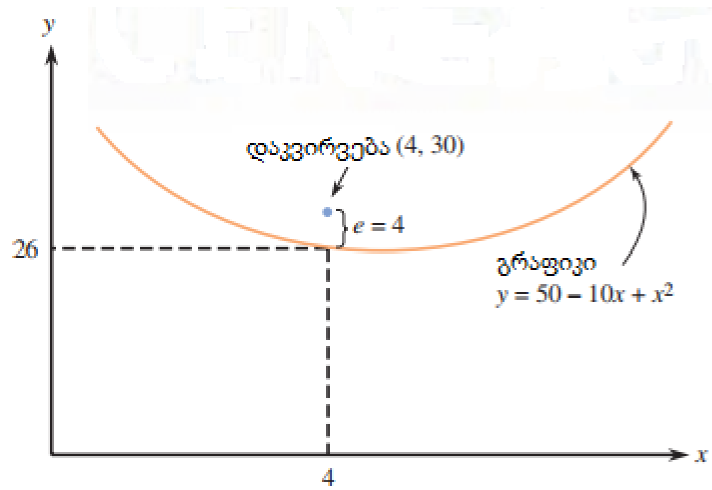
ორ x და y ცვლადს შორის, რომლებიც არ არიან ერთმანეთთან დეტერმინისტულად დაკავშირებული, კავშირის აღსაწერად გამოიყენება სპეციფიური ალბათური მოდელი. ადიტიური ალბათური მოდელის ზოგადი ფორმა საშუალებას აძლევს y-ს იყოს უფრო მცირე ვიდრე $f(x)$, შემთხვევითი e სიდიდით. მოდელის განტოლებას აქვს შემდეგი ფორმა

$$y = x - \textit{ის დეტერმინისტულ ფუნქცია} + \textit{შემთხვევითი გადახრა} = f(x) + e$$

ვთქვათ, x^* აღნიშნავს $x-$ის რაიმე კონკრეტულ მნიშვნელობას, და ასევე y-ზე დაკვირვება გაკეთებულია როდესაც $x = x^*$. მაშინ,

$$y > f(x^*) \ \text{თუ} \ e > 0$$
$$y < f(x^*) \ \text{თუ} \ e < 0$$
$$y = f(x^*) \ \text{თუ} \ e = 0$$

გეომეტრიულად თუ ვიფიქრებთ, თუ $e > 0$, მაშინ დაკვირვებადი წერტილი (e^*, y) მოხვდება გრაფიკი $y = f(x)$-ის მაღლა; $e < 0$ ნიშნავს, რომ ეს წერტილი ჩამოვა გრაფიკის დაბლა. ეს ნაჩვენებია ნახაზ 4-ზე. თუ $f(x)$ არის y და x-თან დაკავშირებული ალბათურ მოდელში გამოყენებული ფუნქცია და თუ y-ზე დაკვირვება კეთდება x-ის სხვადასხვა მნიშვნელობებით, შედეგად მიღებული (x, y) წერტილები განაწილებული იქნება $f(x)$ ფუნქციის გრაფიკით, ზოგი მოხვდება მის ზევით, ზოგიც - ქვევით.



ნახ. 4. ალბათობის მოდელის დეტერმინისტული ნაწილიდან გადახრა.

მარტივი წრფივი რეგრესიის მოდელი არის ზოგადი ალბათური მოდელის კერძო შემთხვევა, რომელშიც დეტერმინანტული ფუქცია $f(x)$ წრფივია (მისი გრაფიკი წრფეა).

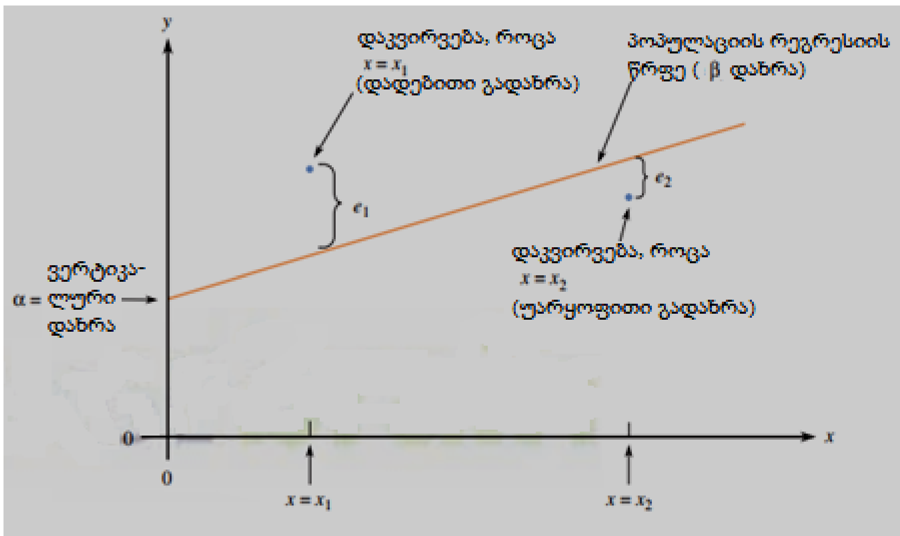
განსაზღვრება

მარტივი წრფივი რეგრესიის მოდელი ვარაუდობს, რომ ნამდვილი ან პოპულაციის რეგრესიის წრფე არის ვერტიკალური ან y ღერძს კვეთს α წერტილში β დახრილობით. როდესაც დამოუკიდებელი ცვლადის x-ის მნიშვნელობა ცნობილია და დაკვირვება მიმდინარეობს დამოკიდებულ y ცვლადზე , მაშინ

$$y = \propto + \beta x + e$$

e შემთხვევითი გადახრის გარეშე, ყველა დაკვირვებული (x, y) წერტილი მოხვდება ზუსტად პოპულაციის რეგრესიის წრფეზე. მოდელის განტოლებაში e -ს ჩართვა აღიარებს, რომ წერტილები გადახრილი იქნება წრფიდან შემთხვევითი რაოდენობით.

ნახაზი 5 გვიჩვენებს რამდენიმე დაკვირვებას პოპულაციის რეგრესიის წრფესთან მიმართებაში.



ნახ.5. რამოდენიმე დაკვირვება მიღებული მარტივი წრფივი რეგრესიის მოდელიდან.

მარტივი წრფივი რეგრესიის მოდელის ძირითადი დაშვებები

1. e -ს განაწილებას რომელიმე კონკრეტულ x მნიშვნელობაზე აქვს საშუალო მნიშვნელობა 0, ანუ $\mu_e = 0$.

2. e-ს სტანდარტული გადახრა (რომელიც აღწერს მისი განაწილების გავრცელებას) ერთი და იგივეა x-ის რომელიმე კონკრეტული მნიშვნელობისთვის. ეს სტანდარტული გადახრა აღინიშნება σ-ით.

3. e-ს განაწილება რომელიმე კონკრეტულ x მნიშვნელობაზე ნორმალურია.

4. შემთხვევითი გადახრები $e_1, e_{,2} \ldots, e_n$, რომლებიც დაკავშირებულია სხვადასხვა დაკვირვებებთან, ერთმანეთისგან დამოუკიდებელია.

თავი 11. შემთხვევითი სიდიდეების განაწილების ფუნქცია

11.1 განაწილების ფუნქციის განსაზღვრება

გავიხსენოთ, რომ დისკრეტული შემთხვევითი სიდიდე შეიძლება მოცემული იქნას ყველა მისი შესაძლო მნიშვნელობების და მათი ალბათობების ჩამონათვალით. მოცემის ასეთი ხერხი არ წარმოადგენს ზოგადს.

მართლაც, განვიხილოთ X შემთხვევითი სიდიდე, რომლის შესაძლო მნიშვნელობები სრულად ავსებს (a, b) ინტერვალს. შესაძლებელია ყველა შესაძლო შემთხვევის ჩამოთვლა? ალბათ არა. ეს მაგალითი მიუთითებს ნებისმიერი ტიპის შემთხვევითი სიდიდის მოცემის ზოგადი ხერხის მითითების მიზანშეწონილობაზე. ამ მიზნით შემოღებულია შემთხვევითი სიდიდეების განაწილების ფუნქცია.

ვთქვათ x ნამდვილი რიცხვია. ხდომილების ალბათობა იმაში მდგომარეობს, რომ X იღებს x-ზე ნაკლებ მნიშვნელობას, ე.ი. $X < x$ ხდომილების ალბათობა აღვნიშნოთ $F(x)$-ით. რა თქმა უნდა, თუ x იცვლება, მაშინ იცვლება $F(x)$-იც, ე.ი.$F(x)$ არის x-ის ფუნქცია.

განაწილების ფუნქციას უწოდებენ F ფუნქციას, რომელიც განსაზღვრავს იმის ალბათობას, რომ ცდის შედეგად X სიდიდე იღებს მნიშვნელობას, რომელიც x -ზე ნაკლებია, ე.ი.

$$F(x) = P(X < x).$$

გეომეტრიულად ეს ტოლობა შეიძლება შემდეგნაირად განვმარტოთ: $F(x)$ არის იმის ალბათობა, რომ შემთხვევითი სიდიდე მიიღებს მნიშვნელობას, რომელიც რიცხვით ღერძზე x წერტილის მარცხნივ მდებარეობს.

ხანდახან „განაწილების ფუნქციის" ნაცვლად გამოიყენება ტერმინი „ინტეგრალური ფუნქცია".

ახლა შესაძლებელია უწყვეტი შემთხვევითი სიდიდე უფრო ზუსტად განვმარტოთ: შემთხვევით სიდიდეს უწყვეტი ეწოდება, თუ მისი განაწილების ფუნქცია არის უწყვეტი, ცალმხრივი დიფერენცირებადი ფუნქცია უწვეტი წარმოებულით.

შენიშვნა. ფუნქცია ცალმხრივად დიფერენცირებადია ან ცალ-ცალკე უწყვეტი, თუ თითოეული ნაწილი დიფერენცირებადია მთელ ქვეარეში, თუნდაც მთელი ფუნქცია არ იყოს დიფერენცირებადი ნაწილებს შორის მდებარე წერტილებში.

11.2 განაწილების ფუნქციის თვისებები

თვისება 1. F განაწილების ფუნქციის მნიშვნელობები ეკუთვნის $[0,1]$ მონაკვეთს, ანუ x-ის ყოველი მნიშვნელობისათვის

$$0 \leq F(x) \leq 1.$$

დამტკიცება. თვისება გამომდინარეობს განაწილების ფუნქციის განსაზღვრებიდან როგორც ალბათობები: ალბათობა ყოველთვის არაუარყოფითი რიცხვია, რომლის სიდიდე ერთს არ აღემატებიან.

თვისება 2. განაწილების F ფუნქცია არაკლებადია, ე.ი.

$F(x_2) \geq F(x_1)$, თუ $x_2 > x_1$.

დამტკიცება. ვთქვათ $x_2 > x_1$. ხდომილება რომელიც იმაში მდგომარეობს, რომ X იღებს მნიშვნელობას, რომელიც x_2-ზე ნაკლებია, შეიძლება ორ უთავსად ხდომილებად დავყოთ: 1) X იღებს მნიშვნელობას, რომელიც x_1-ზე ნაკლებია, $P(X < x_1)$ალბათობით; 2) იღებს მნიშვნელობას, რომელიც აკმაყოფილებს $x_1 \leq X < x_2$ უტოლობას $P(x_1 \leq X < x_2)$ ალბათობით. შეკრების თეორემის თანახმად გვაქვს

$$P(X < x_2) = P(X < x_1) + P(x_1 \leq X < x_2).$$

აქედან

$$P(X < x_2) - P(X < x_1) = P(x_1 \leq X < x_2).$$

ანუ

$$F(x_2) - F(x_1) = P(x_1 \leq X < x_2). \qquad (10.2.1)$$

რამდენადაც ნებისმიერი ალბათობა არაუარყოფითი რიცხვია, ამდენად $F(x_2) - F(x_1) \geq 0$, ანუ $F(x_2) \geq F(x_1)$, რაც უნდა დაგვემტკიცებინა.

შედეგი 1. იმის ალბთობა, რომ შემთხვევითი სიდიდე მიიღებს მნიშვნელობას, რომელიც ინტერვალშია მოთავსებული, ტოლია ამ ინტერვალში ფუნქციის ნაზრდის:

$$P(a \le X < b) = F(b) - F(a). \qquad (10.2.2)$$

ეს მნიშვნელოვანი ფორმულა გამომდინარეობს (10.2.1) ფორმულიდან, თუ ჩავსვამთ $x_2 = b$–ს და $x_1 = a$ –ს.

მაგალითი. X შემთხვევითი სიდიდე მოცემულია შემდეგი განაწილების ფუნქციით

$$F(x) = \begin{cases} 0 & \textit{როცა} \quad x \le -1 \\ x/4 + 1/4 & \textit{როცა} \, -1 < x \le 3 \\ 1 & \textit{როცა} \quad x > 3 \end{cases}$$

ვიპოვოთ იმის ალბათობა, რომ ცდის შედეგად X მიიღებს მნიშვნელობას, რომელიც ეკუთვნის (0,2) ინტერვალს:

$P(0 < X < 2) = F(2) - F(0).$

ამოხსნა: რადგანაც (0,2) ინტერვალზე, პირობის თანახმად,

$$F(x) = x/4 + 1/4,$$

ამდენად

$$F(2) - F(0) = (2/4 + 1/4) - (0/4 + 1/4) = 1/2$$

ამრიგად,

$$P(0 < X < 2) = 1/2.$$

შედეგი 2. იმის ალბათობა, რომ X უწყვეტი შემთხვევითი სიდიდე იღებს ერთ განსაზღვრულ მნიშვნელობას, ნულის ტოლია.

ნამდვილად, (10.2.2) ფორმულაში $a = x_1$, $b = x_1 + \Delta x$, ჩასმით, მივიღებთ

$$P(x_1 \le X < x_1 + \Delta x) = F(x_1 + \Delta x) - F(x_1).$$

მივასწრაფოთ Δx ნულისაკენ. რამდენადაც X უწყვეტი შემთხვევითი სიდიდეა, ამდენად F უწყვეტია. F -ის უწყვეტობის ძალით x_1 წერტილში $F(x_1 + \Delta x) - F(x_1)$ ასევე მიისწრაფის ნულისაკენ; აქედან გამომდინარე, $P(X = x_1) = 0$. ამ მდგომარეობის გამოყენებით, ადვილად დავრწმუნდებით შემდეგი ტოლობების სამართლიანობაში

$$P(a \le X < b) = P(a < X < b) = P(a < X \le b) = P(a \le X \le b) \qquad (10.2.3)$$

მაგალითად, ტოლობა $P(a < X \le b) = P(a < X < b)$ ასე მტკიცდება:

$$P(a < X \le b) = P(a < X < b) + P(X = b) = P(a < X < b).$$

ამგვარად, საინტერესო არაა იმ ალბათობაზე საუბარი, რომ უწყვეტი შემთხვევითი სიდიდე იღებს ერთ განსაზღვრულ მნიშვნელობას, მაგრამ აზრი აქვს განხილული იქნას

ინტერვალში, თუნდაც ძალიან მცირეში, მისი მოხვედრის ალბათობა. ეს ფაქტი სავსებით შეესაბამება პრაქტიკული ამოცანების მოთხოვნებს.

შევნიშნოთ, რომ არასწორი იქნებოდა გვეფიქრა, რომ $P(X = x_1)$ ალბათობის ნულთან ტოლობა ნიშნავს, რომ $X = x_1$ ხდომილება შეუძლებელია (თუ რა თქმა უნდა არ შემოვიფარგლებით ალბათობის კლასიკური განმარტებით). სინამდვილეში, ცდის შედეგად შემთხვევითი სიდიდე აუცილებლად მიიღებს შესაძლო შემთხვევებიდან ერთ-ერთს; კერძოდ, ეს მნიშვნელობა შეიძლება x_1 -ის ტოლი აღმოჩნდეს.

თვისება 3. თუ შემთხვევითი სიდიდის შესაძლო მნიშვნელობა ეკუთვნის (a, b) ინტერვალს, მაშინ : 1) $F(x) = 0$ $x \leq a$-სთვის; 2) $F(x) = 1$ $x \geq b$-სთვის.

დამტკიცება. 1) ვთქვათ $x_1 \leq a$, მაშინ $X < x_1$ ხდომილება შეუძლებელია (რამდენადაც X სიდიდე x_1-ზე ნაკლებ მნიშვნელობას პირობის თანახმად არ იღებს) და აქედან გამომდინარე, მისი ალბათობა ნულის ტოლია.

2) ვთქვათ, $x_2 \geq b$. მაშინ $X < x_2$ ხდომილება აუცილებელია (რამდენადაც X -ის ყველა შესაძლო მნიშვნელობა x_2 -ზე ნაკლებია) და, აქედან გამომდინარე, მისი ალბათობა ერთის ტოლია.

შედეგი. თუ უწყვეტი შემთხვევითი სიდიდის შესაძლო მნიშვნელობები განთავსებულია მთელ ღერძზე, მაშინ სამართლიანია შემდეგი ზღვრული დამოკიდებულებები:

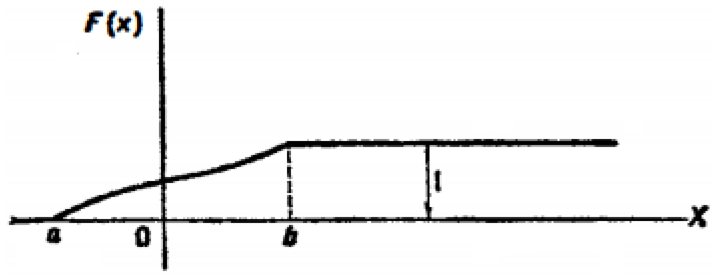
$$\lim_{x \to -\infty} F(x) = 0; \lim_{x \to +\infty} F(x) = 1.$$

11.3 განაწილების ფუნქციის გრაფიკი

დამტკიცებული თვისებები საშუალებას იძლევა წარმოვიდგინოთ, თუ როგორ გამოიყურება უწყვეტი შემთხვევითი სიდიდის განაწილების ფუნქციის გრაფიკი.

გრაფიკი განლაგებულია $y = 0$, $y = 1$ წრფეებით შემოსაზღვრულ ზოლში (პირველი თვისება), x -ის ზრდასთან ერთად (a, b) ინტერვალში, რომელშიც

მოთავსებულია შემთხვევითი სიდიდის ყველა შესაძლო მნიშვნელობა, გრაფიკი „მიიწევს ზევით“ (მეორე თვისება).



ნახ. 6

$x \leq a$ -სთვის გრაფიკის ორდინატა ნულის ტოლია; $x \geq b$ -სთვის გრაფიკის ორდინატები ერთის ტოლია (მესამე თვისება).

უწყვეტი შემთხვევითი სიდიდის განაწილების ფუნქციის გრაფიკი გამოსახულია ნახ. 6-ზე.

შენიშვნა. დისკრეტული შემთხვევითი სიდიდის გრაფიკს კიბისებური ფორმა აქვს.

თავი 12. უწყვეტი შემთხვევითი სიდიდის ალბათობების განაწილების სიმკვრივე

12.1 განაწილების სიმკვრივის განსაზღვრება

ზემოთ უწყვეტი შემთხვევითი სიდიდე მოიცემოდა განაწილების ფუნქციის მეშვეობით. მოცემის ეს ხერხი ერთადერთს არ წარმოადგენს. უწყვეტი შემთხვევითი სიდიდის მოცემა შეიძლება კიდევ ერთი ფუნქციის გამოყენებით, რომელსაც განაწილების სიმკვრივეს ან ალბათობის სიმკვრივეს უწოდებენ (ხანდახან მას დიფერენციალურ ფუნქციასაც უწოდებენ).

უწყვეტი შემთხვევითი X სიდიდის ალბათობების *განაწილების სიმკვრივეს* უწოდებენ f ფუნქციას, რომელიც არის განაწილების F ფუნქციის წარმოებული:

$$f(x) = F'(x).$$

განსაზღვრებიდან გამომდინარეობს, რომ განაწილების ფუნქცია წარმოადგენს განაწილების სიმკვრივის პირველად ფუნქციას.

12.2 უწყვეტი შემთხვევითი სიდიდის მოცემულ ინტერვალში მოხვედრის ალბათობა

თუ ვიცით განაწილების სიმკვრივე, შესაძლებელია გამოვთვალოთ ალბათობა იმისა, რომ შემთხვევითი სიდიდე მიიღებს მნიშვნელობას მოცემულ ინტერვალში. გამოთვლა ემყარება შემდეგ თეორემას.

თეორემა. *იმის ალბათობა, რომ უწყვეტი შემთხვევითი X სიდიდე მიიღებს მნიშვნელობებს (a, b) ინტერვალში, ტოლია განაწილების სიმკვრივიდან განსაზღვრული ინტეგრალისა a-დან b -მდე აღებულ საზღვრებში:*

$$P(a < X < b) = \int_a^b f(x)dx.$$

დამტკიცება. ვისარგებლოთ (10.2.2) დამოკიდებულებით

$$P(a \leq X < b) = F(b) - F(a).$$

ნიუტონ-ლაიბნიცის ფორმულის მიხედვით

$$F(b) - F(a) = \int_a^b F'(x)dx = \int_a^b f(x)\,dx.$$

ამგვარად,

$$P(a \leq X < b) = \int_a^b f(x)\,dx.$$

რამდენადაც უწყვეტი განაწილებისათვის $P(a \leq X < b) = P(a < X < b)$, ამდენად, საბოლოოდ მივიღებთ

$$P(a < X < b) = \int_a^b f(x)\,dx. \tag{11.2.1}$$

გეომეტრიულად მიღებული შედეგი ასე შეიძლება ჩამოვაყალიბოთ: შემთხვევითი სიდიდის მიერ (a, b) ინტერვალზე მდებარე მნიშვნელობის მიიღების ალბათობა ტოლია იმ მრუდწირული ტრაპეციის ფართობის, რომელიც შემოსაზღვრულია Ox ღერძით, $f(x)$ განაწილების სიმკვრივის მრუდით და $x = a$ და $x = b$ წრფეებით.

შენიშვნა. კერძოდ, თუ ფუნქცია ლუწია და ინტერვალის ბოლოები სიმერტიულია კოორდინატთა სათავის მიმართ, მაშინ:

$$P(-a < X < a) = P(|X| < a) = 2\int_0^a f(x)\,dx.$$

მაგალითი. მოცემულია X შემთხვევითი სიდიდის ალბათობების სიმკვრივე

$$f(x) = \begin{cases} 0 & \textit{როცა} \quad x \leq 0, \\ 2x & \textit{როცა} \quad 0 < x \leq 1, \\ 0 & \textit{როცა} \quad x > 1. \end{cases}$$

ვიპოვოთ იმის ალბათობა, რომ X ცდის შედეგად მიიღებს (0,5:1) ინტერვალში მდებარე მნიშვნელობას.

ამოხსნა. საძიებელი ალბათობა ტოლია

$$P(0{,}5 < X < 1) = 2\int_{0,5}^{1} xdx = x^2 \Big|_{0,5}^{1} = 1 - 0{,}25 = 0{,}75.$$

12.3. განაწილების ფუნქციის პოვნა ცნობილი განაწილების სიმკვრივის მიხედვით

თუ ცნობილია f განაწილების სიმკვრივე, შესაძლებელია F განაწილების ფუნქციის მოძებნა შემდეგი ფორმულის მიხედვით

$$F(x) = \int_{-\infty}^{x} f(x)\,dx.$$

მართლაც, როგორც ვიცით, $F(x)$-ით აღნიშნავს იმის ალბათობას, რომ შემთხვევითი სიდიდე X იღებს მნიშვნელობას, რომელიც x-ზე ნაკლებია, ე.ი. $F(x) = P(X < x)$.

ცხადია $X < x$ უტოლობა შეიძლება ჩავწეროთ $-\infty < X < x$ ორმაგი უტოლობის სახით, აქედან

$$F(x) = P(-\infty < X < x). \qquad (12.3.1)$$

თუ (12.2.1) ფორმულაში ვიგულისხმებთ $a = -\infty,\ b = x$, მივიღებთ

$$P(-\infty < X < x) = \int_{-\infty}^{x} f(x)dx.$$

საბოლოოდ, თუ $P(-\infty < X < x)$ -ს შევცვლით $F(x)$ -ით (12.3.1)-ის თანახმად მივიღებთ

$$F(x) = \int_{-\infty}^{x} f(t)\,dt.$$

ამრიგად, განაწილების სიმკვრივის ცოდნით შესაძლებელია განაწილების ფუნქციის პოვნა. რასაკვირველია ცნობილი განაწილების ფუნქციის მიხედვით შესაძლებელია განაწილების სიმკვრივის პოვნა, კერძოდ:

$$f(x) = F'(x).$$

მაგალითი. მოცემული განაწილების სიმკვრივის მიხედვით იპოვეთ განაწილების ფუნქცია:

$$f(x) = \begin{cases} 0 & \textit{როცა} \quad x \le a, \\ 1/(b-a) & \textit{როცა} \quad a < x \le b, \\ 0 & \textit{როცა} \quad x > b. \end{cases}$$

ააგეთ ნაპოვნი ფუნქციის გრაფიკი.

ამოხსნა. გამოვიყენოთ ფორმულა $F(x)=\int_{-\infty}^{x} f(x)\,dx$.

თუ, $x \leq a$, მაშინ $f(x)=0$, აქედან, $F(x)=0$. თუ $a < x \leq b$, მაშინ $f(x)=1/(b-a)$, აქედან გამომდინარე,

$$F(x)=\int_{-\infty}^{\infty} f(x)dx=\int_{-\infty}^{a} 0dx+\int_{a}^{x} \frac{1}{b-a}dx=\frac{x-a}{b-a}.$$

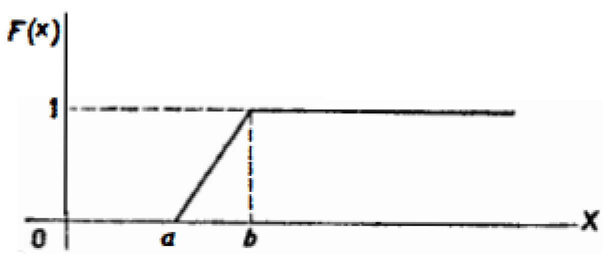
თუ $\mathrm{x} > b$, მაშინ

$$F(x)=\int_{-\infty}^{a} 0dx+\int_{a}^{b} \frac{dx}{b-a}+\int_{b}^{x} 0dx=\frac{b-a}{b-a}=1$$

ამგვარად, საძიებელი განაწილების ფუნცია

$$F(x)=\begin{cases} 0 & \textit{როცა} \quad x \leq a, \\ (x-a)/(b-a) & \textit{როცა} \quad a < x \leq b, \\ 1 & \textit{როცა} \quad x > b. \end{cases}$$

ამ ფუნქციის გრაფიკი ნახ.7-ზეა გამოსახული.



ნახ. 7

12.4 განაწილების სიმკვრივის თვისებები

თვისება 1.განაწილების სიმკვრივე არაუარყოფითი ფუნქციაა, ანუ x-ის ყოველი მნიშვნელობისათვის:

$$f(x) \geq 0$$

დამტკიცება. განაწილების ფუნქცია არაკლებადი ფუნქციაა, აქედან გამომდინარე მისი წარმოებული $F'(x)=f(x)$ არაუარყოფითია.

გეომეტრიულად ეს თვისება ნიშნავს, რომ განაწილების სიმკვრივის გრაფიკის წერტილები განთავსებულია ან Ox ღერძის ზემოთ ან მასზე.

განაწილების სიმკვრივის გრაფიკს *განაწილების მრუდს* უწოდებენ.

თვისება 2. *განაწილების სიმკვრივიდან არასაკუთრივი ინტეგრალი* $-\infty$*-დან* ∞*-მდე საზღვრებში ერთის ტოლია:*

$$\int_{-\infty}^{\infty} f(x)dx = 1.$$

დამტკიცება. ეს არასაკუთრივი ინტეგრალი გამოხატავს იმ ხდომილების ალბათობას, რომელიც მდგომარეობს იმაში, რომ შემთხვევითი სიდიდე ეკუთვნის ($-\infty$, ∞) შუალედს. რა თქმა უნდა ასეთი ხდომილება აუცილებელია, აქედან გამომდინარე მისი ალბათობა ერთის ტოლია.

გეომეტრიულად ეს ნიშნავს, რომ მრუწირული ტრაპეციის მთელი ფართობი, რომელიც Ox ღერძით და განაწილების მრუდითაა შემოსაზღვრული ერთის ტოლია.

კერძოდ, თუ თუ შემთხვევითი სიდიდის ყველა შესაძლო მნიშვნელობა ეკუთვნის (a, b) ინტერვალს, მაშინ

$$\int_{a}^{b} f(x)dx = 1.$$

12.5. განაწილების სიმკვრივის ალბათური შინაარსი

ვთქვათ, F არის უწყვეტი X შემთხვევითი სიდიდის განაწილების ფუნქცია. განაწილების სიმკვრივის განსაზღვრების თანახმად $f(x) = F'(x)$, ან სხვა ფორმით

$$f(x) = \lim_{\Delta x \to 0} \frac{F(x + \Delta x) - F(x)}{\Delta x}.$$

როგორც უკვე ცნობილია $F(x + \Delta x) - F(x)$ სხვაობა განსაზღვრავს იმის ალბათობას, რომ X მიიღებს მნიშვნელობას, რომელიც $(x, x + \Delta x)$ ინტერვალს ეკუთვნის. ამგვარად, ზღვარი შეფარდებისა უწყვეტი შემთხვევითი სიდიდის მიერ $(x, x + \Delta x)$ ინტერვალზე მდებარე მნიშვნელობის მიღების ალბათობისა ამ ინტერვალის სიგრძესთან, როცა ამ

ინტერვალის სიგრძე მიისწრაფვის ნულისაკენ ($\Delta x \to 0$), x წერტილში განაწილების სიმკვრივის ტოლია.

ამგვარად, $f(x)$ ფუნქცია განსაზღვრავს ალბთობების განაწილების სიმკვრივეს თითოეული x წერტილისათვის.

დიფერენციალური აღრიცხვიდან ცნობილია, რომ ფუნქციის ნაზრდი მიახლოებით ტოლია ფუნქციის დიფერენციალის, ე,ი,

$$F(x + \Delta x) - F(x) \approx dF(x),$$

ანუ

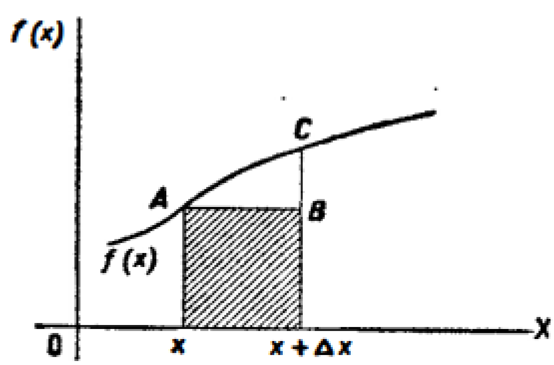
$$F(x + \Delta x) - F(x) \approx F'(x)dx.$$

რამდენადაც $F'(x) = f(x)$ და $dx = \Delta x$, ამდენად

$$F(x + \Delta x) - F(x) \approx f(x)\Delta x.$$

ამ ტოლობის ალბათური შინაარსი ასეთია: ალბათობა იმისა, რომ შემთხვევითი სიდიდე $(x, x + \Delta x)$ ინტერვალზე მდებარე მნიშვნელობას მიიღებს, მიახლოებით ტოლია (Δx მიმართ მაღალი რიგის უსასრულოდ მცირემდე სიზუსტით) x წერტილში ალბათობის სიმკვრივის Δx ინტერვალის სიგრძეზე ნამრავლის.

ეს შედეგი გეომეტრიულად ასე შეიძლება ჩამოვაყალიბოთ: ალბათობა იმისა, რომ შემთხვევითი სიდიდე $(x, x + \Delta x)$ ინტერვალზე მდებარე მნიშვნელობას მიიღებს, მიახლოებით ტოლია მართკუთხედის ფართობის Δx ფუძით და $f(x)$ სიმაღლით.



ნახ. 8

ნახ. 8-ზე ჩანს, რომ დაშტრიხული მართკუთხედის ფართობი, რომელიც $f(x)\Delta x$ ნამრავლის ტოლია, მხოლოდ მიახლოებით უდრის მრუდწირული ტრაპეციის

ფართობს (ჭეშმარიტი ალბათობა განსაზღვრულია $\int_x^{x+\Delta x} f(x)dx$ ინტეგრალით). ამ დროს დაშვებული ცდომილობა ტოლია ABC მრუდწირული სამკუთხედის ფართობის.

12.6 თანაბარი განაწილება

პრატიკული ამოცანების ამოხსნისას გვიწევს შეხვედრა უწყვეტი შემთხვევითი სიდიდეების სხვადასხვა განაწილებებთან. უწყვეტი შემთხვევითი სიდიდეების განაწილებების სიმკვრივეებს ასევე უწოდებენ განაწილების კანონებს. ხშირად გვხვდება, მაგალითად, თანაბარი, ნორმალური და მაჩვენებლიანი განაწილებები.

განაწლებას უწოდებენ თანაბარს, თუ იმ ინტერვალზე, რომელსაც ეკუთვნის შემთხვევითი სიდიდის ყველა შესაძლო მნიშვნელობა, განაწილების სიმკვრივე ინარჩუნებს მუდმივ მნიშვნელობას.

ვიტყვით, რომ შემთხვევითი სიდიდე განაწილებულია თანაბრად ინტერვალზე $[a, b]$, სადაც $a, b \in \mathbb{R}$, თუ მის განაწილების ფუნციას აქვს შემდეგი სახე:

$$F(x) = \begin{cases} 0\,, \text{როცა}\; x < a \\ \dfrac{x-a}{b-a}, \text{როცა}\; a \leq x < b \\ 1, \quad \text{როცა}\; x \geq b \end{cases}$$

თუ შემთხვევითი სიდიდე X განაწილებულია თანაბრად ინტერვალზე $[a, b]$, მას ასე აღნიშნავენ: $X \sim U[a, b]$. თუ $a = 0$ და $b = 1$, მაშინ ამბობენ, რომ გვაქვს სტანდარტული თანაბარი განაწილება.

თავი 13. ნორმალური განაწილება

13.1 უწყვეტი შემთხვევითი სიდიდის რიცხვითი მახასიათებლები

ვთქვათ, უწყვეტი X შემთხვევითი სიდიდე მოცემულია f განაწილების სიმკვრივით. დავუშვათ, რომ X -ის ყველა შესაძლო მნიშვნელობა მიეკუთვნება $[a, b]$ მონაკვეთს. დავყოთ ეს მონაკვეთი n ცალ $\Delta x_1, \Delta x_2, \dots, \Delta x_n$ სიგრძის მონაკვეთებად და თითოეული მათგანიდან ნებისმიერად ავირჩიოთ x_j წერტილი ($j = 1,2, \dots . n$). ჩვენ უნდა განვსაზღვროთ უწყვეტი X შემთხვევითი სიდიდის მათემატიკური ლოდინი; შევადგინოთ x_j შესაძლო მნიშვნელობების მათ Δx_j ინტერვალში მოხვედრის ალბათობებზე ნამრავლების ჯამი (შეგახსენებთ, რომ $f(x)\Delta x$ ნამრავლი მიახლოებით ტოლია X-ის Δx ინტერვალში მოხვედრის ალბათობის):

$$\sum x_i f(x_i) \Delta x_i.$$

თუ გადავალთ ზღვარზე, როცა არჩეული მონაკვეთებიდან უდიდესების სიგრძეები მიისწრაფვიან ნულისაკენ, მივიღებთ განსაზღვრულ ინტეგრალს:

$$\int_a^b x f(x) dx.$$

თუ უწყვეტი X შემთხვევითი სიდიდის შესაძლო მნიშვნელობები მიეკუთვნება $[a, b]$ მონაკვეთს, მაშინ მის *მათემატიკურ ლოდინს* უწოდებენ განსაზღვრულ ინტეგრალს

$$M(X) = \int_a^b x f(x) dx. \tag{13.1.1}$$

თუ შესაძლო მნიშვნელობები მიეკუთვნება მთელ Ox ღერძს, მაშინ

$$M(X) = \int_{-\infty}^{\infty} x f(x) dx.$$

იგულისხმება, რომ არასაკუთრივი ინტეგრალი აბსოლუტურად კრებადია, ე.ი. არსებობს $\int_{-\infty}^{-} |x| f(x) dx$ ინტეგრალი. თუ ეს მოთხოვნა არ შესრულდება, მაშინ ინტეგრალის მნიშვნელობა დამოკიდებული იქნება მისწრაფების სიჩქარეზე (ცალკ-

ცალკე) ქვედა ზღვრის $-\infty$-სკენ, ზედასი – ∞-სკენ და ამ შემთხვევაში ამბობენ, რომ შემთხვევით სიდიდეს არ გააჩნია მათემატიკური ლოდინი.

უწყვეტი შემთხვევითი სიდიდის დისპერსია ეწოდება მათემატიკური ლოდინიდან მისი გადახრის კვადრატის მათემატიკურ ლოდინს.

თუ X -ის შესაძლო მნიშვნელობები $[a, b]$მონაკვეთს მიეკუთვნება, მაშინ:

$$D(X) = \int_a^b [x - M(X)]^2 f(x) dx;$$

თუ შესაძლო მნიშვნელობები მიეკუთვნება მთელ Ox ღერძს, მაშინ:

$$D(X) = \int_{-\infty}^{\infty} [x - M(X)]^2 f(x) dx.$$

უწყვეტი შემთხვევითი სიდიდის საშუალო კვადრატული გადახრა განისაზღვრება, ისე, როგორც დისკრეტული სიდიდისათვის, შემდეგი ტოლობით:

$$\sigma(X) = \sqrt{D(X)}.$$

შენიშვნა 1. შესაძლებელია იმის დამტკიცება, რომ დისკრეტული სიდიდისათვის მათემატიკური ლოდინის და დისპერსიის თვისებები ნარჩუნდება უწყეტი სიდიდეებისათვისაც.

შენიშვნა 2. ადვილი მისაღებია დისპერსიის გამოსათვლელად უფრო მოსახერხებელი ფორმულები:

$$D(X) = \int_a^b x^2 f(x) dx - [M(X)]^2, \quad (13.1.2)$$

$$D(X) = \int_{-\infty}^{\infty} x^2 f(x) dx - [M(X)]^2.$$

13.2 ნორმალური განაწილება

ნორმალური ეწოდება უწყვეტი შემთხვევითი სიდიდის ალბთობების განაწილებებს, რომლებიც აღიწერებიან შემდეგი სიმკვრივით:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-a)^2/2\sigma^2}.$$

როგორც ვხედავთ, ნორმალური განაწილება განისაზღვრება ორი პარამეტრით: a და σ. საკმარისია ამ პარამეტრების ცოდნა იმისათვის, რომ მოიცეს ნორმალური განაწილება. ვაჩვენოთ, რომ ალბათური აზრი ამ პარამეტრებისა შემდეგია: a არის მათემატიკური ლოდონი და σ - საშუალო კვადრატული გადახრა.

ა) უწყვეტი შემთხვევითი სიდიდის მთემატიკური ლოდინის განსაზღვრების თანახმად,

$$M(X) = \int_{-\infty}^{\infty} xf(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x\, e^{-(x-a)^2/2\sigma^2} dx.$$

შემოვიტანოთ ახალი ცვლადი $z = (x-a)/\sigma$. აქედან $x = \sigma z + a,\ dx = \sigma dz$. ცხადია, რომ ინტეგრების ახალი საზღვრები ძველის ტოლია. მივიღებთ:

$$M(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma z + a)\, e^{\frac{-z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma z\, e^{\frac{-z^2}{2}} dz + \frac{a}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-z^2}{2}}\, dz.$$

პირველი შესაკრები ნულის ტოლია (ინტეგრალის ქვეშ კენტი ფუნქციაა; ინტეგრების საზღვრები კოორდინატთა სათავის მიმართაა სიმეტრიული). მეორე შესაკრები a -ს ტოლია (პუასონის ინტეგრალი $\int_{-\infty}^{\infty} e^{\frac{-z^2}{2}}\, dz = \sqrt{2\pi}$).

ამრიგად, $M(X) = a$, ე.ი. ნორმალური განაწილების მათემატიკური ლოდინი a პარამეტრის ტოლია.

ბ) უწყვეტი შემთხვევითი სიდიდის დისპერსიის განსაზღვრების თანახმად, იმის გათვალისწინებით, რომ $M(X) = a$, მივიღებთ:

$$D(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-a)^2 e^{-(x-a)^2/2\sigma^2} dx.$$

შემოვიტანოთ ახალი ცვლადი $z = (x-a)/\sigma$. აქედან, $x - a = \sigma z,\ dx = \sigma dz$. ცხადია, რომ ინტეგრირების ახალი საზღვრები ძველის ტოლია. მივიღებთ:

$$D(X) = \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z \cdot z e^{-z^2/2} dz.$$

ნაწილობრივი ინტეგრებით $u = z$, $dv = ze^{-z^2/2}dz$, ვიპოვით

$$D(X) = \sigma^2.$$

აქედან გამომდინარე,

$$\sigma(X) = \sqrt{D(X)} = \sqrt{\sigma^2} = \sigma$$

ამრიგად, ნორმალური განაწილების საშუალო კვადრატული გადახრა σ პარამეტრის ტოლია.

შენიშვნა. *ზოგადს* უწოდებენ ნორმალურ განაწილებას ნებისმიერი a და $\sigma(\sigma > 0)$ პარამეტრებით.

სტანდარტულს უწოდებენ ნორმალურ განაწილებას პარამეტრებით $a = 0$ და $\sigma = 1$. მაგალითად, თუ X ნორმალურად განაწილებული შემთხვევითი სიდიდეა a და σ პარამეტრებით, მაშინ $U = (X - a)/\sigma$ -სტანდარტული ნორმალურად განაწილებული შემთხვევითი სიდიდეა, ამასთან $M(U) = 0,\ \sigma(U) = 1$.

სტანდარტული ნორმალური განაწილების სიმკვრივე ტოლია

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

13.3 ნორმალური მრუდი

ნორმალური განაწილების გრაფიკს უწოდებენ ნორმალურ მრუდს (გაუსის მრუდს).

გამოვიკვლიოთ შემდეგი ფუნქცია: $y = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-a)^2/(2\sigma^2)}$, დიფერენციალური აღრიცხვის მეთოდებით.

1. ცხადია, ფუნქცია განსაზღვრულია მთელ Ox ღერძზე.

2. x-ის ყველა მნიშვნელობისათვის ფუნქცია იღებს დადებით მნიშვნელობას, ე.ი. ნორმალური მრუდი Ox ღერძის ზემოთ მდებარეობს.

3. ფუნქციის ზღვარი x -ის შეუზღუდავი ზრდისას (აბსოლუტური მნიშვნელობით) ნულის ტოლია: $\lim\limits_{|x|\to\infty} y = 0$, ე.ი. Ox ღერძი გრაფიკის ჰორიზონტალური ასიმპტოტის მოვალეობას ასრულებს.

4. გამოვიკვლიოთ ფუნქცია ექსტრემუმზე. ვიპოვოთ პირველი წარმოებული:

$$y' = -\frac{x-a}{\sigma^3\sqrt{2\pi}} e^{-(x-a)^2/(2\sigma^2)}.$$

ადვილი სანახავია, რომ $y' = 0$, როცა $x = a$, $y' > 0$, როცა $x < a$,$y' < 0$, როცა $x > a$.

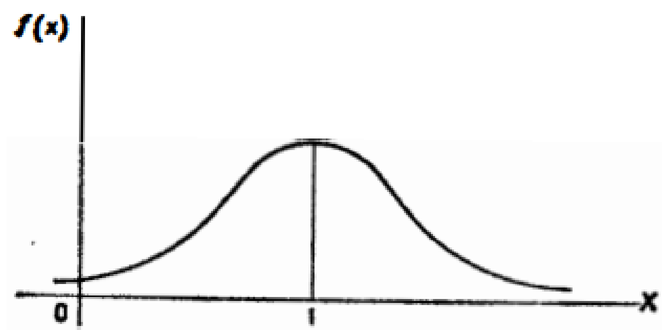
აქედან გამომდინარე, როცა $x = a$ ფუნქციას $1/(\sigma\sqrt{2\pi})$-ის ტოლი მაქსიმუმი გააჩნია.

5. $x - a$ სხვაობას ფუნქციის ანალიტიკური გამოსახულება კვადრატში აყვანილს შეიცავს, ე.ი. ფუნქციის გრაფიკი სიმეტრიულია $x = a$ წრფის მიმართ.

6. გამოვიკვლიოთ ფუნქცია გადაღუნვის წერტილზე. ვიპოვოთ მეორე წარმოებული:

$$y'' = -\frac{1}{\sigma^3\sqrt{2\pi}} e^{-(x-a)^2/(2\sigma^2)} \left[1 - \frac{(x-a)^2}{\sigma^2}\right].$$

ადვილი დასანახია, რომ როცა $x = a + \sigma$ და $x = a - \sigma$ მეორე წარმოებული ნულის ტოლია, ხოლო ამ წერტილების გავლის შემდეგ ის ნიშანს იცვლის (ორივე წერტილში ფუნქციის მნიშვნელობა $1/(\sigma\sqrt{2\pi e})$-ს ტოლია). ამრიგად, გრაფიკის წერტილები $(a - \sigma, 1/(\sigma\sqrt{2\pi e}))$ და $(a + \sigma, 1/(\sigma\sqrt{2\pi e}))$ წარმოადგენენ გადაღუნვის წერტილებს.



ნახ. 9

ნახ.9-ზე გამოსახულია ნორმალური მრუდი, როცა $a = 1$ და $\sigma = 2$-ს

13.4 ნორმალური განაწილების პარამეტრების გავლენა ნორმალური მრუდის ფორმაზე

გამოვარკვიოთ, როგორ გავლენას ახდენს a და σ პარამეტრების მნიშვნელობა ნორმალური მრუდის ფორმაზე და განლაგებაზე.

ცნობილია, რომ $f(x)$ და $f(x - a)$ ფუნქციების გრაფიკებს ერთნაირი ფორმა აქვთ; თუ $f(x)$ ის გრაფიკს გადავანაცვლებთ x ღერძზე დადებითი მიმართულებით a ერთეულით როცა $a > 0$, ან უარყოფითი მიმართულებით როცა $a < 0$, მივიღებთ

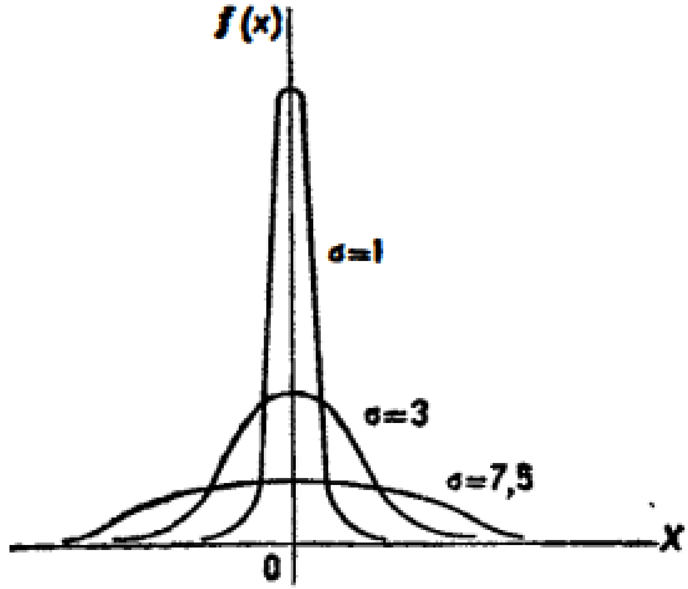
$f(x-a)$-ის გრაფიკს. აქედან გამომდინარეობს, *პარამეტრ* a*ს* (მათემატიკური ლოდინი) *ცვლილება არ იწვევს ნორმალური მრუდის ფორმის ცვლილებას, არამედ იწვევს მის წანაცვლებას* $0x$ *ღერძის გასწვრივ: მარჯვნივ, თუ* a *დადებითია, და მარცხნივ თუ* a *უარყოფითია.*

სხვაგვარადაა საქმე თუ იცვლება σ პარამეტრი (საშუალო კვადრატული გადახრა). რიგორც წინა პარაგრაფში იყო მითითებული, ნორმალური განაწილების დიფერენციალური ფუნქციის მაქსიმუმი $1/(\sigma\sqrt{2\pi})$-ის ტოლია. აქედან გამიმდნარეობს, რომ σ-ს გაზრდასთან ერთად ნორმალური მრუდის მაქსიმალური ორდინატა მცირდება, ხოლო თვითონ მრუდი უფრო დამრეცი ხდება, ე.ი. იწევს Ox ღერძისაკენ; σ -ს შემცირებასთან ერთად მრუდი ხდება უფრო „წვეტიანი“ და იჭიმება Oy ღერძის დადებითი მიმართულებით.

აღვნიშნოთ, რომ a და σ პარამეტრების ნებისმიერი მნიშვნელობისათვის ფართობი, რომელიც შემოსაზღვრულია ნორმალური მრუდით და Ox ღერძით ერთის ტოლი რჩება (განაწილების სიმკვრივის მეორე თვისება).

ნახ.10-ზე გამოსახულია ნორმალური მრუდები σ-ს სხვადასხვა მნიშვნელობის და $a=0$-სთვის. ნახაზი თვალსაჩინოდ ახდენს იმის ილუსტრირებას, თუ σ პარამეტრის ცვლილება როგორ აისახება ნორმალური მრუდის ფორმაზე.

შევნიშნოთ, რომ $a=0$ და $\sigma=1$ -სთვის $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ ნორმალურ მრუდს უწოდებენ სტანდარტულს.



ნახ.10

13.5 ნორმალური შემთხვევითი სიდიდის მოცემულ ინტერვალში მოხვედრის ალბათობა

უკვე ცნობილია, რომ თუ X შემთხვევითი სიდიდე მოცემულიაგანაწილების f სიმკვრივით, მაშინ იმის ალბათობა, რომ X მიიღებს მნიშვნელობას, რომელიც $(\propto, \beta)$ ინტერვალს ეკუთვნის, ასეთია:

$$P(\propto < X < \beta) = \int_{\alpha}^{\beta} f(x)dx.$$

ვთქვათ X შემთხვევითი სიდიდე მოცემულია ნორმალური კანონით. მაშინ იმის ალბათობა, რომ მიიღებს მნიშვნელობას, რომელიც $(\propto, \beta)$ ინტერვალს ეკუთვნის, ტოლია:

$$P(\propto < X < \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-(x-a)^2/(2\sigma^2)} dx.$$

ეს ფორმულა ისე გარდავქმნათ, რომ მზა ცხრილების გამოყენების საშუალება გვქონდეს. შემოვიღეთ ახალი ცვლადი $z = (x - a)/\sigma$. აქედან, $x = \sigma z + a$, $dx = \sigma dz$.

ვიპოვოთ ინტეგრების ახალი საზღვრები. თუ $x = \propto$, მაშინ $z = (\propto - a)/\sigma$, თუ $x = \beta$, მაშინ $z = (\beta - a)/\sigma$.

ამგვარად, გვექნება

$$P(\propto < X < \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int_{(\alpha-a)/\sigma}^{(\beta-a)/\sigma} e^{-z^2/2}(\sigma dz) =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{(\alpha-a)/\sigma}^{0} e^{-z^2/2} dz$$

$$+ \frac{1}{\sqrt{2\pi}} \int_{0}^{(\beta-a)/\sigma} e^{-z^2/2} dz = \frac{1}{\sqrt{2\pi}} \int_{0}^{(\beta-a)/\sigma} e^{-z^2/2} dz - \frac{1}{\sqrt{2\pi}} \int_{0}^{(\alpha-a)/\sigma} e^{-z^2/2} dz.$$

ლაპლასის ფუნქციის გამოყენებით

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-z^2/2} dz,$$

საბოლოოდ მივიღებთ:

$$P(\alpha < X < \beta) = \phi\left(\frac{\beta - a}{\sigma}\right) - \phi\left(\frac{\alpha - a}{\sigma}\right). \tag{13.5.1}$$

13.6. მოცემული გადახრის ალბათობის გამოთვლა

ხშირად საჭიროა გამოითვალოს იმის ალბათობა, რომ X შემთხვევითი სიდიდის ნორმალური განაწილების გადახრა აბსოლუტური მნიშვნელობით ნაკლები იყოს მოცემულ δ დადებით რიცხვზე, ე.ი. $|X - a| < \delta$ უტოლობის განხორციელების ალბათობა.

შევნიშნოთ, რომ ეს უტოლობა შემდეგი ორი უტოლობის ტოლფასია

$$-\delta < X - a < \delta, \text{ ან } a - \delta < X < a + \delta.$$

(13.5.1) ფორმულის გამოყენებით, მივიღებთ

$$P(|X - a| < \delta) = P(a - \delta < X < a + \delta) = \phi\left[\frac{(a + \delta) - a}{\sigma}\right] - \phi\left[\frac{(a - \delta) - a}{\sigma}\right] =$$

$$= \phi\left(\frac{\delta}{\sigma}\right) - \phi\left(-\frac{\delta}{\sigma}\right).$$

თუ მხედველობაში მივიღებთ შემდეგ ტოლობას

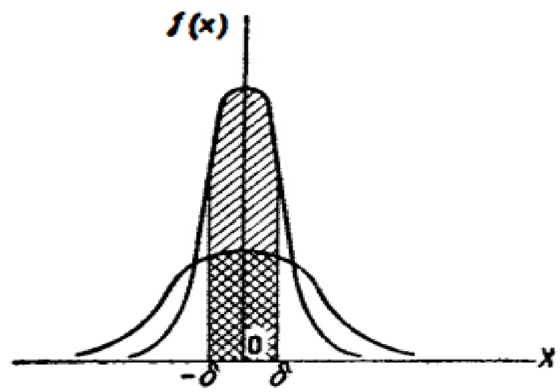
$$\phi(-\delta/\sigma) = -\phi(\delta/\sigma).$$

(ლაპლასის ფუნქცია კენტია), საბოლოოდ მივიღებთ

$$P(|X-a| < \delta) = 2\phi(\delta/\sigma).$$

კერძოდ, როცა $a = 0$

$$P(|X| < \delta) = 2\phi(\delta/\sigma).$$



ნახ. 11

ნახ.11-ზე თვალსაჩინოდაა ნაჩვენები, თუ ორი შემთვევითი სიდიდე ნორმალურადაა განაწილებული და $a = 0$, მაშინ იმ მნიშვნელობაზე მეტის მიღების ალბათობა, რომელიც $(-\delta, \delta)$ ინტერვალს ეკუთვნის აქვს იმ სიდიდეს, რომელსაც σ-ს ნაკლები მნიშვნელობა გააჩნია. ეს ფაქტი სრულად შეესაბამება σ პარამეტრის ალბათურ შინაარსს (σ საშუალო კვადრატული გადახრაა; ის ახასიათებს შემთხვევითი სიდიდის გაფანტვას მისი მათემატიკური ლოდინის ირგვლივ).

შენიშვნა. ცხადია $|X-a| < \delta$ და $|X-a| \geq \delta$ განხორციელების ხდომილებები მოპირდაპირეებია. ამიტომ თუ $|X-a| < \delta$ უტოლობის განხორციელების ალბათობა p-ს ტოლია, მაშინ $|X-a| \geq \delta$ უტოლობის ალბათობა $1-p$-სტოლია.

13.7. სამი სიგმას წესი

გარდავქმნათ შემდეგი ფორმულა:

$$P(|X-a| < \delta) = 2\phi(\delta/\sigma).$$

მივანიჭოთ $\delta = \sigma t.$ მივიღებთ,

$$P(|X - a| < \sigma t) = 2\phi(t).$$

თუ $t = 3$ და, შესაბამისად, $\sigma t = 3\sigma$, მაშინ

$$P(|X - a| < 3\sigma = 2\Phi(3) = 2 \cdot 0{,}49865 = 0{,}9973,$$

ე.ი. იმის ალბათობა, რომ გადახრა აბსოლუტური მნიშვნელობით გასამმაგებული საშუალო კვადრატულ გადახრაზე ნაკლები იქნება, ტოლია 0,9973-ის.

სხვა სიტყვებით, ალბათობა იმისა, რომ გადახრის აბსოლუტური მნიშვნელობა გასამმაგებული საშუალო კვადრატულ გადახრას გადააჭარბებს, ძალიან მცირეა, კერძოდ 0,0027-ის ტოლია. ეს ნიშნავს, რომ მხოლოდ შემთხვევების 0,27%-ში ეს შეიძლება მოხდეს. ხდომილების ასეთი შედეგი შეუძლებლობის პრინციპით ნაკლებადაა მოსალოდნელი და ის შეიძლება პრაქტიკულად შეუძლებლადაც ჩაითვალოს. სწორედ ამაში მდგომარეობს სამი სიგმას არსი: *თუ შემთხვევითი სიდიდე ნორმალურადაა განაწილებული, მაშინ მისი მათემატიკური ლოდინიდან გადახრის აბსოლუტური სიდიდე არ აღემატება გასამმაგებულ საშუალო კვადრატულ გადახრას.*

პრაქტიკაში სამი სიგმის წესი ასეა მიღებული: თუ შესასწავლი შემთხვევითი სიდიდის განაწილება უცნობია, მაგრამ წესში მოყვანილი პირობა სრულდება, გვაქვს საფუძველი ვიგულისხმოთ, რომ შესასწავლი სიდიდე განაწილებულია ნორმალურად; წინააღმდეგ შემთხვევაში ის არაა ნორმალურად განაწილებული.

13.8 ცენტრალური ზღვარითი თეორემის ფორმულირება

ცნობილია, რომ პრაქტიკაში ძალიან გავრცელებულია შემთხვევითი სიდიდის ნორმალური განაწილება. ეს რით აიხსნება? ამ კითხვაზე პასუხს იძლევა ლიაპუნოვის თეორემა (ცენტრალური ზღვარითი თეორემა): *თუ X შემთხვევითი სიდიდე წარმოადგენს ურთიერთდამოუკიდებელი შემთხვევითი სიდიდეების დიდი რაოდენობის ჯამს, რომელთაგან თითოეულის გავლენა მთელ ჯამზე უმნიშვნელოა, მაშინ X -ს გააჩნია განაწილება, რომელიც ნორმალურთან ახლოსაა.*

ჩამოვაყალიბოთ ცენტრალური ზღვარითი თეორემა, რომელიც ადგენს პირობას, რომლის დროსაც დიდი რაოდენობის დამოუკიდებელი შესაკრებების ჯამს გააჩნია განაწილება, რომელიც ნორმალურთან ახლოსაა.

ვთქვათ, $X_1, X_2, \dots, X_{n,} \dots$ დამოუკიდებელ შემთხვევით სიდიდეთა მიმდევრობაა, რომელთაგან თითოეულს გააჩნია სასრული მათემატიკური ლოდინი და დისპერსია:

$M(X_k) = a_{k\ ,}\quad D(X_k) = b_k^2.$

შემოვიტანოთ აღნიშვნები:

$S_n = X_1 + X_2 + \dots + X_n,\quad A_n = \sum_{k=1}^{n} a_k,\quad B^2 = \sum_{k=1}^{n} b_k^2.$

ნორმირებული ჯამის განაწილების ფუნქცია აღვნიშნოთ შემდეგნაირად

$$F_n(x) = P\left(\frac{S_n - A_n}{B_n} < x\right).$$

ამბობენ, რომ $X_1, X_2, \dots$ მიმდევრობა აკმაყოფილებს ზღვარით თეორემას, თუ ნებისმიერი x -სთვის ნორმირებული ჯამის განაწილების ფუნქცია $n \to \infty$ -სთვის მისწრაფვის ნორმალური განაწილების ფუნქციისაკენ:

$$\lim_{n\to\infty} P\left[\frac{S_n - A_n}{B_n} < x\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-z^2/2} dz.$$

კერძოდ, თუ ყველა $X_1, X_2, \dots$ შემთხვევითი სიდიდე ერთნაირადაა განაწილებული, მაშინ ასეთი მიმდევრობა აკმაყოფილებს ცენტრალურ ზღვარით თეორემას, თუ ყველა შემთხვევითი სიდიდის $X_i\ (i = 1,2 \dots)$ დისპერსია სასრულია და განსხვავებულია ნულისაგან.

13.9. თეორიული განაწილების ნორმალურისაგან გადახრის შეფასება. ასიმეტრია და ექსცესი

ემპირიულს უწოდებენ ფარდობითი სიხშირეების განაწილებას. ემპირიულ განაწილებას შეისწავლის მათემატიკური სტატისტიკა.

თეორიულს უწოდებენ ალბათობების განაწილებას. თეორიულ განაწილებებს შეისწავლის ალბათობის თეორია. აქ განვიხილავთ თეორიულ განაწილებებს.

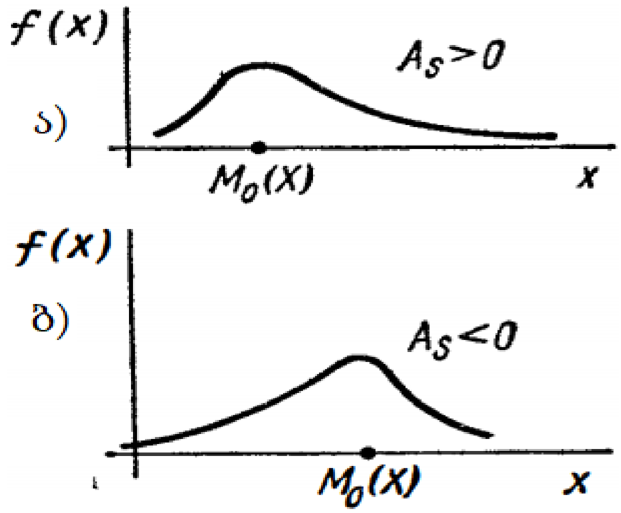
ნორმალურისაგან განსხვავებული განაწილებების შესწავლისას დგება ამ განსხვავების რაოდენობრივი შეფასების უცილებლობა. ამ მიზნით შემოტანილია სპეციალური მახასითებლები, კერძოდ, ასიმეტრია და ექსცესი. ნორმალური განაწილებისათვის ეს მახასიათებლები ნულის ტოლია. ამიტომ თუ შესასწავლი განაწილებისათვის ასიმეტრიას და ექსცესს აქვს მცირე მნიშვნელობა, მაშინ უნდა ვივარაუდოთ ამ განაწილების ნორმალურთან სიახლოვე. პირიქით, ასიმეტრიისა და ექსცესის დიდი მნიშვნელობა მიუთითებს ნორმალურიდან მნიშვნელოვან გადახრაზე.

როგორ შევაფასოთ ასიმეტრია? შეიძლება დამტკიცდეს, რომ სიმეტრიული განაწილებისათვის (ასეთი განაწილების გრაფიკი $x = M(X)$ წრფის მიმართ სიმეტრიულია) კენტი რიგის თითოეული მომენტი ნულის ტოლია. არასიმეტრიული განაწილებისათვის კენტი რიგის თითოეული მომენტი ნულისაგან განსხვავებულია. ამიტომ ამ მომენტებიდან ნებისმიერი (პირველი რიგის მომენტის გარდა, რომელიც ნებისმიერი განაწილებისათვის ნულის ტოლია) შეიძლება გამოდგეს ასიმეტრიის შესაფასებლად; ბუნებრივია მათ შორის უმარტივესის, ე.ი. μ_3 მესამე რიგის მომენტის არჩევა. მაგრამ ასიმეტრიის შესაფასებლად ამ მომენტის აღება მოუხერხებელია იმიტომ, რომ მისი მნიშვნელობა დამოკიდებულია ერთეულზე, რომელშიც იზომება შემთხვევთი სიდიდე. ამ ნაკლოვანების გამოსასწორებლად μ_3 -ს ყოფენ σ^3 -ზე და ამით იღებენ მახასიათებელს განზომილების გარეშე.

თეორიული განაწილების ასიმეტრიას უწოდებენ მესამე რიგის ცენტრალური მომენტის ფარდობას საშუალო კვადრატული გადახრის კუბთან:

$$A_s = \mu_3/\sigma^3.$$

ასიმეტრია დადებითია, თუ განაწილების მრუდის „გრძელი ნაწილი“ განლაგებულია მათემატიკური ლოდინიდან მარჯვნივ; ასიმეტრია უარყოფითია, თუ განაწილების მრუდის „გრძელი ნაწილი“ განლაგებულია მათემატიკური ლოდინიდან მარცხნივ. ასიმეტრიის ნიშანს პრაქტიკულად განსაზღვრავენ განაწილების მრუდის მოდის (დიფერენციალური ფუნქციის მაქსიმუმის წერტილი) მიმართ განლაგების მიხედვით: თუ მრუდის „გრძელი ნაწილი“ განლაგებულია მოდის მარჯვნივ, მაშინ ასიმეტრია დადებითა (ნახ. 12. ა)), თუ მარცხნივ უარყოფითი (ნახ.12. ბ)).



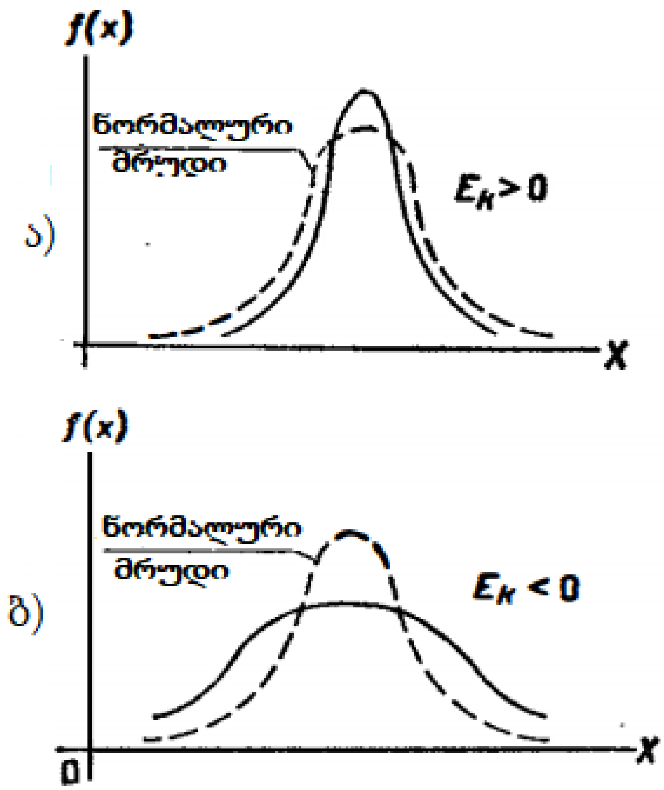
ნახ. 12.

„დამრეცობის“, ე.ი. ნორმალურ მრუდთან შედარებით თეორიული მრუდის მეტი თუ ნაკლები აწევის შესაფასებლად სარგებლობენ ექსცესით.

თეორიული განაწილების ექსცესს უწოდებენ მახასითებელს, რომელიც განისაზღვრება შემდეგი ტოლობით

$$E_k = (\mu_4/\sigma^4) - 3.$$

ნორმალური განაწილებისათვის $\mu_4/\sigma^4 = 3$; შესაბამისად, ექსცესი ნულის ტოლია. ამიტომ თუ რაიმე განაწილების ექსცესი ნულისაგან განსხვავებულია, ამ განაწილების მრუდი განსხვავდება ნორმალური მრუდისაგან: თუ ექსცესი დადებითია, მაშინ მრუდს აქვს შედარებით მაღალი და „მახვილი“ წვერო (ნახ.13.ა)), ვიდრე ნორმალურ მრუდს; თუ ექსცესი უარყოფითია, მაშინ მრუდს აქვს შედარებით შედარებით დაბალი და „გლუვი“ წვერო (ნახ.13.ბ)). ამასთან ნავარაუდევია, რომ ნორმალურ და თეორიულ განაწილებებს ერთნაირი მათემატიკური ლოდინი და დისპესია აქვთ.



ნახ.13.

13.10. ერთი შემთხვევითი არგუმენტის ფუნქცია და მისი განაწილება

წინასწარ შევნიშნოთ, რომ შემდგომში „ალბათობების განაწილების კანონის“ ნაცვლად მოკლედ ვიტყვით „განაწილებას“.

თუ X შემთხვევითი სიდიდის თითოეულ შესაძლო მნიშვნელობას შეესაბამება Y შემთხვევითი სიდიდის ერთი შესაძლო მნიშვნელობა, მაშინ Y -ს უწოდებენ X შემთხვევითი არგუმენტის ფუნქციას:

$$Y = \varphi(X).$$

შემდგომში ნაჩვენებია, როგორ ვიპოვოთ ფუქციის განაწილება დისკრეტული და უწყვეტი არგუმენტის ცნობილი განაწილების მიხედვით.

1. ვთქვათ X არგუმენტი დისკრეტული შემთხვევითი სიდიდეა.

ა) თუ X არგუმენტის სხვადასხვა შესაძლო მნიშვნელობას შეესაბამება Y ფუნქციის სხვადასხვა შესაძლო მნიშვნელობა, მაშინ X და Y შესაბამისი მნიშვნელობებისლბათობები ტოლია.

ბ) თუ X-ის სხვადასხვა შესაძლო მნიშვნელობას შეესაბამება Y -ის მნიშვნელობები, რომელთა შორის არის ერთმანეთის ტოლი მნიშვნელობები, მაშინ უნდა შევკრიბოთ Y-ის განმეორებადი მნიშვნელობების ალბათობები.

2) *ვთქვათ X არგუმენტი უწყვეტი შემთხვევითი სიდიდეა.* როგორ ვიპოვოთ $Y = \varphi(X)$ ფუნქციის განაწილება, თუ ვიცით X შემთხვევითი არგუმენტის განაწილების სიმკვრივე? მტკიცდება: თუ $Y = \varphi(X)$ დიფერენცირებადი მკაცრად ზრდადი ან მკაცრად კლებადი ფუნქციაა, რომლის შებრუნებული ფუნქციაა $x = \psi(y)$, მაშინ Y შემთვევითი სიდიდის განაწილების სიმკვრივე $g(y)$მოიძებნება შემდეგი ტოლობის მეშვეობით:

$$g(y) = f[\psi(y)]|\psi'(y)|. \tag{13.10.1}$$

13.11 ერთი შემთხვევითი არგუმენტის ფუნქციის მათემატიკური ლოდინი

მოცემულია X შემთხვევითი არგუმენტის $Y = \varphi(X)$ ფუნქცია, უნდა ვიპოვოთ ამ ფუნქციის მათემატიკური ლოდინი, თუ ვიცით არგუმენტის განაწილების კანონი.

1. *ვთქვათ X არგუმენტი დისკრეტული შემთხვევითი სიდიდეა* $x_1, x_2, \dots, x_n$ შესაძლო მნიშვნელობებით, რომელთა ალბათობები შესაბამისად ტოლია $p_1, p_2, \dots, p_n$-ის. რამდენადაც ხდომილება „X სიდიდემ მიიღო x_i მნიშვნელობა“ იწვევს ხდომილებას „Y სიდიდემ მიიღო $\varphi(x_i)$ მნიშვნელობა“, ამდენად, Y-ის შესაძლო მნიშვნელობათა ალბათობები შესაბამისად უდრის $p_1, p_2, \dots, p_n$ -ს. აქედან გამომდინარე ფუნქციის მათემატიკური ლოდინია:

$$M[\varphi(X)] = \sum_{i=1}^{n} \varphi(x_i) p_{i.} \tag{13.11.1}$$

2. *ვთქვათ X არგუმენტი უწყვეტი შემთხვევითი სიდიდეა*, რომელიც $f(x)$ განაწილების სიმკვრივითაა მოცემული. $Y = \varphi(X)$ ფუნქციის მათემატიკური ლოდინის მოსაძებნად ჯერ უნდა მოვძებნოთ Y სიდიდის განაწილების სიმკვრივე $g(y)$, ხოლო შემდეგ კი გამოვიყენოთ ფორმულა:

$$M(Y) = \int_{-\infty}^{\infty} yg(y)dy.$$

თუმცა, თუ $g(y)$ ფუნქციის მოძებნა გართულდა, მაშინ შეიძლება უშუალოდ მოიძებნოს $\varphi(X)$ ფუნქციის მათემატიკური ლოდინი შემდეგი ფორმულის მიხედვით:

$$M[\varphi(X)] = \int_{-\infty}^{\infty} \varphi(x)f(x)dx.$$

კერძოდ, თუ X-ის შესაძლო მნიშვნელობები ეკუთვნის (a, b) ინტერვალს, მაშინ

$$M[\varphi(X)] = \int_a^b \varphi(x)f(x)dx. \qquad (13.11.2)$$

13.12. ორი შემთხვევითი არგუმენტის ფუნქცია. დამოუკიდებელი შესაკრებების ჯამის განაწილება. ნორმალური განაწილების მდგრადობა

თუ X და Y შემთხვევითი სიდიდეების თითოეულ წყვილს შეესაბამება Z სიდიდის ერთი შესაძლო მნიშვნელობა, მაშინ Z-ს უწოდებენ ორი X და Y შემთხვევითი არგუმენტის ფუნქციას:

$$Z = \varphi(X, Y).$$

1. *ვათქვათ X და Y დისკრეტული დამოუკიდებელი შემთხვევითი სიდიდეებია.* იმისათვის, რომ შევადგინოთ $Z = X + Y$ ფუნქციის განაწილების კანონი, უნდა ვიპოვოთ Z-ის ყველა შესაძლო მნიშვნელობა და მათი ალბათობები.

2. *ვთქვათ X და Y უწყვეტი შემთხვევითი სიდიდეებია.* მტკიცდება, რომ თუ X და Y დამოუკიდებლები არიან, მაშინ $Z = X + Y$ ჯამის განაწილების სიმკვრივე $g(z)$ (იმ პირობით, რომ თუნდაც ერთი არგუმენტის სიმკვრივე მოცემულია $(-\infty, \infty)$ ინტერვალზე ერთი ფორმულით) შესაძლებელია მოიძებნოს შემდეგი ფორმულით:

$$g(z) = \int_{-\infty}^{\infty} f_1(x)f_2(z-x)\, dz \qquad (13.12.1)$$

ან ტოლფასი ტოლობით

$$g(z) = \int_{-\infty}^{\infty} f_1(z-y)f_2(y)\, dy \qquad (13.12.2)$$

სადაც, f_1 და f_2 არგუმენტების განაწილებების სიმკვრივეებია.

თუ არგუმენტების შესაძლო მნიშვნელობები არაუარყოფითებია, მაშინ $g(z)$ შემდეგი ფორმულით პოულობენ:

$$g(z) = \int_0^z f_1(x) f_2(z-x) dx \tag{13.12.3}$$

ან ტოლფასი ფორმულით

$$g(z) = \int_0^z f_1(z-y) f_2(y) dy \tag{13.12.4}$$

დამოუკიდებელი შემთხვევითი სიდიდეების ჯამის განაწილების სიმკვრივეს უწოდებენ *კომპოზიციას*.

ალბათობების განაწილების კანონს უწოდებენ მდგრადს, თუ ასეთი კანონების კომპოზიცია ისეთივე კანონია (ზოგადად პარამეტრებით განსხვავებული). ნორმალურ კანონს გაჩნია მდგრადობის თვისება: ნორმალური კანონების კომპოზიციას ასევე გააჩნია ნორმალური განაწილება (ამ კოპოზიციის მათემატიკური ლოდინი და დისპერსია ტოლია შესაკრებთა შესაბამის მათემატიკური ლოდინებისა და დისპერსიების ჯამების). მაგალითად, თუ X და Y დამოუკიდებელი ნორმალურად განაწილებული შემთხვევითი სიდიდეებია, რომელთა მათემატიკური ლოდინი და დისპერსია შესაბამისად ტოლია $a_1 = 3, a_2 = 4, D_1 = 1,\ D_2 = 0{,}5$, მაშინ ამ სიდიდეების კომპოზიცია (ე.ი. $Z = X + Y$ ჯამის ალბათობის სიმკვრივე) ასევე განაწილებულია ნორმალურად, ამასთან კომპოზიციის მათემატიკური ლოდინი და დისპერსია შესაბამისად ტოლია $a = 3 + 4 = 7; d = 1 + 0{,}5 = 1{,}5$.

13.13.განაწილება „ხი კვადრატი“

ვთქვათ, $X_i\ (i = 1,2, \dots, n)$ - ნორმალურად განაწილებული დამოუკიდებლი შემთხვევითი სიდიდეებია, ამასთან თითოეული მათგანის მათემატიკური ლოდინი ნულის ტოლია, ხოლო საშუალო კვადრატული გადახრა - ერთის. მაშინ ამ სიდიდეების

კვადრატების ჯამი $\chi^2 = \sum_{i=1}^{n} X_i^2$, χ^2 -ის კანონითაა განაწილებული $k = n$ თავისუფლების ხარისხებით; თუ კი ეს სიდიდეები დაკავშირებულნი არიან ერთი წრფივი თანაფრადობით, მაგალითად $\sum X_i = n\bar{X}$, მაშინ თავისუფლების ხარისხის რიცხვი $k = n - 1$.

ამ განაწილების სიმკვრივეა:

$$f(x) = \begin{cases} 0 & \text{როცა } x \leq 0, \\ \dfrac{1}{2^{k/2}\Gamma(k/2)} & \text{როცა } x > 0, \end{cases}$$

სადაც, $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\, dt$- გამა ფუნქციაა.

$$\Gamma(n+1) = n!$$

აქედან ჩანს, რომ „ხი კვადრატი“ განაწილება განისაზღვრება ერთი პარამეტრით - k თავისუფლების ხარისხის რიცხვით.

თავისუფლების ხარისხის რიცხვის გაზრდით განაწილება თანდათან უახლოვდება ნორმალურს.

13.14. სტიუდენტის განაწილება

ვთქვათ, Z ნორმალურად განაწილებული შემთხვევითი სიდიდეა, ამასთან $M(Z) = 0$, $\sigma(Z) = 1$, ხოლო V არის Zსგან დამოუკიდებელი სიდიდე, რომელიც განაწილებულია χ^2 კანონით k თავისუფლების ხარისხებით. მაშინ სიდიდეს

$$T = \frac{Z}{\sqrt{V/k}} \tag{13.14.1}$$

გააჩნია განაწილება, რომელსაც უწოდებენ t-განაწილებას, ან სტიუდენტის განაწილებას k თავისუფლების ხარისხებით.

ამგვარად, ნორმირებული ნორმალური სიდიდის ფარდობა დამოუკიდებელი შემთხვევითი სიდიდიდან ფესვთან, რომელიც განაწილებულია „ხი კვადრატით“ k

თავისუფლების ხარისხებით და გაყოფილია k -ზე, განაწილებულია სტიუდენტის კანონით k თავისუფლების ხარისხით.

თავისუფლების ხარისხის რიცხვის გაზრდით სტიუდენტის განაწილება თანდათან უახლოვდება ნორმალურს.

თავი 14. მაჩვენებლიანი განაწილება

14.1 მაჩვენებლიანი განაწილების განმარტება

მაჩვენებლიანი (ექსპონენციალური) ეწოდება X უწყვეტი შემთხვევითი სიდიდის ალბათობების განაწილებებს, რომლებიც შემდეგი სიმკვრივით აღიწერებიან

$$f(x) = \begin{cases} 0\,, & \textit{როცა} \quad x < 0 \\ \lambda e^{-\lambda x} & \textit{როცა} \ x \geq 0 \end{cases}$$

სადაც, λ - მუდმივი დადებითი სიდიდეა.

ვხედავთ, რომ მაჩვენებლიანი განაწილება განისაზღვრება ერთი λ პარამეტრით. მაჩვენებლიანი განაწილების ეს თავისებურება მიუთითებს მის უპირატესობაზე სხვა განაწილებებთან შედარებით, რომლებიც მეტი რაოდენობის პარამეტრებზე არიან დამოკიდებულნი. ჩვეულებრივ პარამეტრები ცნობილი არაა და გვიწევს მათი შეფასებების მოძებნა (უახლოესი მნიშვნელობის); რა თქმა უნდა ერთი პარამეტრის შეფასება უფრო მარტივია, ვიდრე ორის ან სამის და ა.შ. უწყვეტი შემთხვევითი სიდიდის მაგალითად, რომელიც მაჩვენებლიანი კანონითაა განაწილებული, შეიძლება გამოგვადგეს უმარტივესი ნაკადის ორი მიმდევრობითი ხდომილების მოხდენის დრო.

ვიპოვოთ მაჩვენებლიანი კანონის განაწილების ფუნქცია (იხ. პარაგრაფი 12.3):

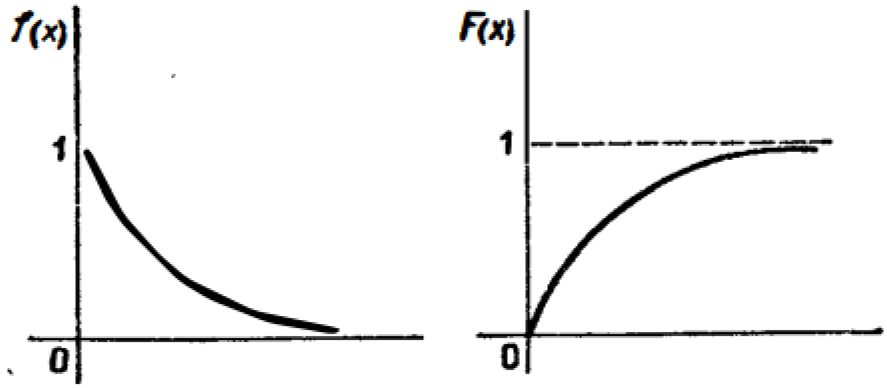
$$F(x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{0} 0dx + \lambda \int_{0}^{x} e^{-\lambda x}dx = 1 - e^{-\lambda x}.$$

ამრიგად,

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & x \geq 0. \end{cases}$$

მაჩვენებლიანი კანონი განვსაზღვრეთ განაწილების სიმკვრივის მეშვეობით; ცხადია, რომ ის შეიძლება განისაზღვროს განაწილების ფუნქციის გამოყენებით.

სიმკვრივის და მაჩვენებლიანი განაწილების ფუნქციის გრაფიკები გამოსახულია ნახ. 14-ზე.



ნახ.14.

14.2. მაჩვენებლიანად განაწილებული შემთხვევითი სიდიდის მოცემულ ინტერვალში მოხვედრის ალბათობა

ვიპოვოთ შემდეგი განაწილების ფუნქციის მიხედვით, მაჩვენებლიანი კანონით განაწილებული X შემთხვევითი სიდიდის (a, b) ინტერვალში მოხვედრის ალბათობა

$$F(x) = 1 - e^{-\lambda x}(x \geq 0).$$

გამოვიყენოთ შემდეგი ფორმულა (იხ. 10.2.2)

$$P(a \leq X < b) = F(b) - F(a).$$

თუ გავითვალისწინებთ, რომ $F(a) = 1 - e^{-\lambda a}$, $F(b) = 1 - e^{-\lambda b}$, მივიღებთ

$$P(a \leq X < b) = e^{-\lambda a} - e^{-\lambda b}. \qquad (14.2.1)$$

e^{-x}ფუნქციის მნიშვნელობები მოცემულია ცხრილებში.

14.3 მაჩვენებლიანი განაწილების რიცხვითი მახასიათებლები

ვთქვათ X უწყვეტი შემთხვევითი სიდიდე განაწილებულია შემდეგი მაჩვენებლიანი კანონით

$$f(x) = \begin{cases} 0\,, & \textit{როცა} \quad x < 0 \\ \lambda e^{-\lambda x} & \textit{როცა} \ x \geq 0 \end{cases}$$

ვიპოვოთ მათემატიკური ლოდინი

$$M(X) = \int_0^\infty x f(x) dx = \lambda \int_0^\infty x e^{-\lambda x} dx.$$

ნაწილობრივი ინტეგრებით მივიღებთ

$$M(X) = 1/\lambda. \tag{14.3.1}$$

ამრიგად, *მაჩვენებლიანი განაწილების მათემატიკური ლოდინი პარამეტრ λ -ს შებრუნებული მნიშვნელობის ტოლია.*

ვიპოვოთ დისპერსია

$$D(X) = \int_0^\infty x^2 f(x) dx - [M(X)]^2 = \lambda \int_0^\infty x^2 e^{-\lambda x} dx - 1/\lambda^2.$$

ნაწილობრივი ინტეგრირებით მივიღებთ

$$\lambda \int_0^\infty x^2 e^{-\lambda x} dx = 2/\lambda^2.$$

აქედან

$$D(X) = 1/\lambda^2.$$

ვიპოვოთ საშუალო კვადრატული გადახრა, ამისათვის ამოვიღოთ კვადრატული ფესვი დისპერსიიდან:

$$\sigma(X) = 1/\lambda. \tag{13.3,2}$$

თუ შევადარებთ (13.3.1)-ს და (13.3.2)-ს, დავასკვნით, რომ

$$M(X) = \sigma(X) = 1/\lambda.$$

ე.ი. *მაჩვენებლიანი განაწილების მათემატიკური ლოდინი და საშუალო კვადრატული გადახრა ერთმანეთის ტოლია.*

მაჩვენებლიანი განაწილება ფართოდ გამოიყენება საიმედოობის თეორიაში, რომლის ძირითად ცნებას წარმოადგენს საიმედოობის ფუნქცია.

14.4 საიმედოობის ფუნქცია

ელემენტს ვუწოდებთ რაიმე მოწყობილობას იმისდა მიუხედავად „რთულია“ ის თუ „მარტივი“.

ვთქვათ ელემენტი იწყებს მუშაობას დროის $t_0 = 0$ მომენტში, ხოლო t დროის გასვლის შემდეგ დგება მტყუნება. T-თი აღვნიშნოთ უწყვეტი შემთხვევითი სიდიდე - ელემენტის უმტყუნო მუშაობის დროის ხანგრძლივობა. თუ ელემენტმა უმტყუნოდ იმუშავა (მტყუნობის დადგომამდე) t -ზე ნაკლები დროის მანძილზე, აქედან გამომდინარე t დროის მიმდინარეობის დროს დადგება მტყუნება.

ამრიგად, განაწილების $F(t) = P(T < t)$ ფუნქცია განსაზღვრავს t დროის ხანგრძლივობის დროს *მტყუნების ალბათობას.* აქედან გამიმდინარე, უმტყუნო მუშაობის ალბათობა დროის იმავე t ხანგრძლივობისას, ე.ი. მოპირდაპირე ხდომილების ალბათობა $T > t$, ტოლია

$$R(t) = P(T > t) = 1 - F(t). \tag{13.4.1}$$

$R(t)$ საიმედოობის ფუნქციას უწოდებენ ფუნქციას, რომელიც განსაზღვრავს ელემენტის უმტყუნო მუშაობის ალბათობას t დროის ხანგრძლივობის დროს:

$$R(t) = P(T > t).$$

მათემატიკური სტატისტიკის ელემენტები

თავი 15. შერჩევის მეთოდი

15.1 მათემატიკური სტატისტიკის ამოცანები

კანონზომიერებების ჩამოყალიბება, რომლებიც მართავენ მასობრივ შემთხვევით მოვლენებს, ეფუძნება სტატისტიკური მონაცემების - დაკვირვების შედეგების - შესწავლას ალბათობის თეორიის მეთოდების გამოყენებით.

მათემატიკური სტატისტიკის უპირველესი ამოცანაა სტატისტიკური ინფორმაციის შეგროვებისა და დაჯგუფების გზების მითითება, რომელიც მიღებულია დაკვირვების ან სპეციალურად ჩატარებული ექსპერიმენტების შედეგად.

მათემატიკური სტატისტიკის მეორე ამოცანაა კვლევის მიზნებიდან გამომდინარე სტატისტიკური მონაცემების ანალიზის მეთოდების შემუშავება. მასში შედის:

ა) ხდომილების უცნობი ალბათობის შეფასება; უცნობი განაწილების ფუნქციის შეფასება; ცნობილი განაწილების უცნობი პარამეტრების შეფასება; შემთხვევითი სიდიდის დამოკიდებულების შეფასება ერთ ან რამდენიმე შემთხვევით სიდიდეებზე და ა.შ.

ბ) უცნობი განაწილების ფორმის ან განაწილების პარამეტრების სიდიდის შესახებ სტატისტიკური ჰიპოთეზების შემოწმება.

თანამედროვე მათემატიკური სტატისტიკა ამუშავებს ხერხებს იმისათვის, რომ დადგინდეს საჭირო ცდების რაოდენობა კვლევის დაწყებამდე (ექსპერიმენტის დაგეგმვა), ასევე კვლევის დროს (მიმდევრობითი ანალიზი) და სხვა მრავალი ამოცანის გადასაწყვეტად. თანამედროვე მათემატიკური სტატისტიკა განისაზღვრება, როგორც განუსაზღვრელობის პირობებში გადაწყვეტილების მიღების მეცნიერება.

ასე რომ, მათემატიკური სტატისტიკის ამოცანაა შექმნას სტატისტიკური მონაცემების შეგროვებისა და დამუშავების მეთოდები მეცნიერული და პრაქტიკული დასკვნების მისაღებად.

15.2. გენერალური და შერჩევითი ერთობლიობები

ვთქვათ, საჭიროა ერთგვაროვანი ობიექტების სიმრავლის შესწავლა ამ ობიექტების დამახასიათებელი **თვისებრივი** ან **რაოდენობრივი** ნიშნის მიმართ. მაგალითად, თუ არსებობს დეტალების ჯგუფი, მაშინ დეტალის სტანდარტი შეიძლება იყოს ხარისხობრივი მაჩვენებელი, ხოლო დეტალის კონტროლირებადი ზომა შეიძლება იყოს რაოდენობრივი მაჩვენებელი.

ზოგჯერ ტარდება სრული გამოკვლევა, ანუ ერთობლიობის თითოეული ობიექტის გამოკვლევა ხდება იმ ნიშნის მიმართ, რომლითაც ინტერესდებიან. თუმცა პრაქტიკაში სრული გამოკვლევა შედარებით იშვიათად გამოიყენება. მაგალითად, თუ ერთობლიობა შეიცავს ობიექტთა ძალიან დიდ რაოდენობას, მაშინ ფიზიკურად შეუძლებელია სრული გამოკვლევის ჩატარება. თუ ობიექტის კვლევა დაკავშირებულია მის განადგურებასთან, ან მოითხოვს დიდ მატერიალურ ხარჯებს, სრული გამოკვლევის ჩატარებას აზრი არ აქვს. ასეთ დროს შემთხვევით ირჩევენ მთელი ნაკრებიდან შეზღუდული რაოდენობის ობიექტებს და ხდება მათი შესწავლა.

შერჩევის ერთობლიობა ან უბრალოდ *შერჩევა* არის შემთხვევით შერჩეული ობიექტების ერთობლიობა.

გენერალური ერთობლიობა, ანუ პოპულაცია არის ობიექტების ერთობლიობა, საიდანაც კეთდება შერჩევა.

ერთობლიობის (შერჩევის ან პოპულაციის) მოცულობას უწოდებენ ამ ერთობლიობის ობიექტების რაოდენობას. მაგალითად, თუ 1000 დეტალიდან შეირჩევა 100 დეტალი კვლევისთვის, მაშინ საერთო პოპულაციის მოცულობა არის $N = 1000$, ხოლო შერჩევის მოცულობა არის $n = 100$.

შენიშვნა. ხშირად პოპულაცია შეიცავს ობიექტთა სასრულ რაოდენობას. თუმცა, თუ ეს რიცხვი საკმარისად დიდია, მაშინ ზოგჯერ, გამოთვლების გასამარტივებლად ან თეორიული დასკვნების გასაადვილებლად ვარაუდობენ, რომ პოპულაცია შედგება ობიექტების უსასრულო რაოდენობისგან. ეს ვარაუდი გამართლებულია იმით, რომ

პოპულაციის მოცულობის (საკმარისად დიდი მოცულობის) ზრდა პრაქტიკულად არ ახდენს გავლენას შერჩევის მონაცემების დამუშავების შედეგებზე.

15.3.განმეორებადი შერჩევები და შერჩევები განმეორებების გარეშე. რეპრეზენტატული შერჩევა

შერჩევის მიღებისას შეგიძლიათ გააკეთოთ ორი რამ: მას შემდეგ რაც ობიექტი შეირჩევა და მას დააკვირდით, ის შეიძლება დაუბრუნდეს ან არ დაუბრუნდეს პოპულაციას. ზემოაღნიშნულის შესაბამისად შერჩევები იყოფა განმეორებით და განმეორებების გარეშე შერჩევებად.

განმეორებითი შერჩევა ეწოდება შერჩევას, რომელშიც შერჩეული ობიექტი (შემდეგის არჩევამდე) უბრუნდება პოპულაციას.

შერჩევა განმეორების გარეშე ეწოდება შერჩევას, რომელშიც შერჩეული ობიექტი (შემდეგის არჩევამდე) არ უბრუნდება პოპულაციას.

პრაქტიკაში, ჩვეულებრივ იყენებენ შემთხვევით შერჩევას განმეორების გარეშე.

იმისათვის, რომ ნიმუშის მონაცემები იყოს საკმარისად დამაჯერებელი, რათა ვიმსჯელოთ პოპულაციის ჩვენთვის საინტერესო თვისებაზე აუცილებელია, რომ შერჩევის ობიექტები სწორად წარმოადგენდნენ მას. სხვა სიტყვებით რომ ვთქვათ, შერჩევა სწორად უნდა წარმოადგენდეს ერთობლიობის პროპორციებს. ეს მოთხოვნა მოკლედ ჩამოყალიბებულია შემდეგნაირად: ნიმუში უნდა იყოს *რეპრეზენტატული* (წარმომადგენლობითი).

დიდ რიცხვთა კანონის მიხედვით შეიძლება ითქვას, რომ შერჩევა იქნება წარმომადგენლობითი, თუ ის განხორციელდება შემთხვევით: შერჩევის თითოეული ობიექტი შემთხვევითად შეირჩევა პოპულაციიდან, თუ ყველა ობიექტს აქვს შერჩევაში მოხვედრის ერთნაირი ალბათობა.

თუ პოპულაციის მოცულობა საკმარისად დიდია და შერჩევა ამ ერთობლიობის მხოლოდ უმნიშვნელო ნაწილია, მაშინ წაშლილია განსხვავება განმეორებით და

განმეორების გარეშე შერჩევებს შორის; უკიდურეს შემთხვევაში, როდესაც განიხილება უსასრულო პოპულაცია და შერჩევას აქვს სასრული მოცულობა, ეს განსხვავება ქრება.

15.4. შერჩევის მეთოდები

პრაქტიკაში გამოიყენება შერჩევის სხვადასხვა მეთოდი. ეს მეთოდები შეიძლება დაიყოს ორ ტიპად:

1.შერჩევა, რომელიც არ საჭიროებს პოპულაციის ნაწილებად დაშლას. მათ შორისაა: ა) მარტივი შემთხვევითი შერჩევა განმეორების გარეშე; ბ) მარტივი შემთხვევითი შერჩევა განმეორებით.

2. შერჩევა, რომელშიც პოპულაცია იყოფა ნაწილებად. მას მიეკუთვნებიან: ა) ტიპიური შერჩევა; ბ) მექანიკური შერჩევა; გ) სერიული შერჩევა.

მარტივი შემთხვევითი შერჩევა ეწოდება შერჩევას, რომლის დროსაც ობიექტები ამოიღება თითო-თითოდ მთელი პოპულაციიდან. მარტივი შერჩევა შეიძლება გაკეთდეს სხვადასხვა გზით. მაგალითად, N მოცულობის პოპულაციიდან n ობიექტის ამოსაღებად აკეთებენ შემდეგს: ჩაწერენ რიცხვებს 1-დან N-მდე ბარათებზე, რომელთაც საფუძვლიანად აურევენ და ამოიღებენ ერთ ბარათს შემთხვევით; განიხილება ამოღებული ბარათის იგივე ნომრის მქონე ობიექტი, შემდეგ ბარათი ბრუნდება დასტაში და პროცესი მეორდება; ანუ ბარათებს ურევენ, შემდეგ ერთ-ერთს იღებენ შემთხვევით და ა.შ. ეს კეთდება n-ჯერ; შედეგად, მიიღება n მოცულობის მარტივი შემთხვევითი შერჩევა განმეორებებით.

თუ ამოღებული ბარათები არ დაბრუნდება დასტაში, მაშინ შერჩევა წარმოადგენს მარტივ შემთხვევით შერჩევას განმეორებების გარეშე. პოპულაციის დიდი მოცულობის დროს აღწერილი პროცესი ძალიან შრომატევადია. ამ შემთხვევაში გამოიყენება "შემთხვევითი რიცხვების" მზა ცხრილები, რომლებშიც ნომრები დალაგებულია შემთხვევითი თანმიმდევრობით. იმისათვის, რომ შეარჩიოთ, მაგალითად, 50 ობიექტი

დანომრილი პოპულაციიდან, გახსენით შემთხვევითი რიცხვების ცხრილის ნებისმიერი გვერდი და ჩაწერეთ ზედიზედ 50 რიცხვი; შერჩევა მოიცავს იმ ობიექტებს, რომელთა ნომრებიც ემთხვევა ჩამოწერილ შემთხვევით რიცხვებს. თუ აღმოჩნდება, რომ ცხრილის შემთხვევითი რიცხვი აღემატება N რიცხვს, მაშინ ასეთი შემთხვევითი რიცხვი გამოიტოვება. განმეორებების გარეშე შერჩევის შესრულებისას ასევე უნდა გამოტოვოთ შემთხვევითი რიცხვების ცხრილიდან ის რიცხვები, რომლებიც უკვე შეგხვდათ.

შერჩევას უწოდებენ *ტიპურს,* რომლის დროსაც ობიექტები შეირჩევა არა პოპულაციიდან, არამედ მისი თითოეული "ტიპური" ნაწილიდან. მაგალითად, თუ დეტალები მზადდება რამდენიმე დაზგაზე, შემდეგ შერჩევა ხდება არა ყველა დაზგაზე წარმოებული დეტალების მთელი ნაკრებიდან, არამედ თითოეული დაზგის პროდუქტებისგან ცალკ-ცალკე. ტიპური შერჩევა გამოიყენება, როდესაც გამოკვლეული თვისება მკვეთრად იცვლება პოპულაციის სხვადასხვა ნაწილში. მაგალითად, თუ პროდუქტი იწარმოება რამდენიმე მანქანაზე, რომელთაგან ზოგიერთი მეტ-ნაკლებად გაცვეთილია, მაშინ აქ მიზანშეწონილია ტიპური შერჩევა.

შერჩევას უწოდებენ *მექანიკურს,* რომლის დროსაც პოპულაცია "მექანიკურად" იყოფა იმდენ ჯგუფად, რამდენი ობიექტიც უნდა იყოს შეყვანილი შერჩევაში და თითოეული ჯგუფიდან შეირჩევა ერთი ობიექტი. მაგალითად, თუ თქვენ გჭირდებათ ჩარხზე დამზადებული დეტალების 20%-ის შერჩევა, მაშინ შეირჩევა ყოველი მეხუთე დეტალი; თუ საჭიროა დეტალების 5%-ის შერჩევა, მაშინ შეირჩევა ყოველი მეოცე დეტალი და ა.შ. უნდა აღინიშნოს, რომ ზოგჯერ მექანიკურმა შერჩევამ შეიძლება არ უზრუნველყოს რეპრეზენტატული შერჩევა. მაგალითად, თუ შეირჩევა ყოველი მეოცე გამოჩარხული ლილვაკი და შერჩევისთანავე საჭრისი შეიცვლება, მაშინ შერჩევაში მოხვედრილი ყველა ლილვაკი, რომლებიც ბლაგვი საჭრისით იქნება გამოჩარხული. ამ შემთხვევაში აუცილებელია შერჩევის რიტმის დამთხვევის აღმოფხვრა საჭრისის გამოცვლის რიტმთან, რისთვისაც საჭიროა, ვთქვათ, ოცი გამოჩარხულიდან ყოველი მეათე ლილვაკის შერჩევა.

ხაზს ვუსვამთ იმას, რომ პრაქტიკაში ხშირად გამოიყენება კომბინირებული შერჩევა, რომელშიც გაერთიანებულია ზემოთ ჩამოთვლილი მეთოდები. მაგალითად, ზოგჯერ

ზოგადი ერთობლიობა იყოფა ერთნაირი მოცულობის მწკრივებად, შემდეგ რამდენიმე მწკრივი შეირჩევა მარტივი შემთხვევითი შერჩევით და, ბოლოს, ცალკეული ობიექტები ამოიღება თითოეული მწკრივიდან მარტივი შემთხვევითი შერჩევით.

15.5 შერჩევის სტატისტიკური განაწილება

ვთქვათ ზოგადი ერთობლიობიდან ამოღებულია შერჩევა, ამასთან დაიკვირვება $x_1 n_1$-ჯერ, $x_2\, n_2$-ჯერ, $x_k n_k$-ჯერ და შერჩევის მოცულობაა$\sum n_i = n$. x_i-ის დაკვირვებად მნიშვნელობებს უწოდებენ *ვარიანტებს*, ხოლო ვარიანტების თანმიმდევრობას, რომლებიც ჩაწერილია ზრდადი მიმდევრობით, ეწოდება *ვარიაციული მწკრივი.* დაკვირვების რაოდენობას *სიხშირეებს* უწოდებენ, ხოლო მათ შეფარდებას შერჩევის მოცულობასთან $n_i/n = W_i$ უწოდებენ ფარდობით სიხშირეებს.

შერჩევის სტატისტიკურ განაწილებას უწოდებენ ვარიანტების ჩამონათვალს და მათ შესაბამის სიხშირეებს ან ფარდობით სიხშირეებს. სტატისტიკური განაწილება ასევე შეიძლება მოიცეს ინტერვალების თანმიმდევრობით და მათი შესაბამისი სიხშირეებით (ინტერვალის შესაბამის სიხშირედ მიიღება სიხშირეების ჯამი ამ ინტერვალში).

შევნიშნოთ, რომ ალბათობის თეორიაში განაწილება გაგებულია, როგორც შემთხვევითი ცვლადის შესაძლო მნიშვნელობებისა და მათი ალბათობების შესაბამისობა, ხოლო მათემატიკურ სტატისტიკაში კი დაკვირვებად ვარიანტებსა და მათ სიხშირეებს, ან ფარდობით სიხშირეებს შორის შესაბამისობა.

მაგალითი. მოცემულია n = 20 მოცულობის შერჩევის სიხშირეების განაწილება:

x_i2 6 12

n_i3 10 7

დაწერეთ ფარდობითი სიხშირეების განაწილება.

ამოხსნა. ვიპოვოთ ფარდობითი სიხშირეები, რისთვისაც გავყოთ სიხშირეები შერჩევის მოცულობაზე:

W_1= 3/20= 0,15 ,W_2= 10/20 = 0,50, W_3 = 7/20= 0,35.

დავწეროთ ფარდობითი სიხშირეების განაწილება:

x_i	2	,6	1
W_i	0,15	0,50	0,35

კონტროლი: 0,15+0,50+0,35= 1.

15.6 განაწილების ემპირიული ფუნქცია

ვთქვათ, ცნობილია X რაოდენობრივი თვისების სიხშირეების სტატისტიკური განაწილება. შემოვიღოთ აღნიშვნა: n_x- იყოს იმ დაკვირვებების რაოდენობა, რომლებშიც მახასიათებლის მნიშვნელობა არის ნაკლები x-ზე; n არის დაკვირვებების საერთო რაოდენობა (შერჩევის მოცულობა). ცხადია, რომ მოვლენის ფარდობითი სიხშირე X <x უდრის $n_x\ /\ n$. თუ x იცვლება, მაშინ, ზოგადად რომ ვთქვათ, ფარდობითი სიხშირე ასევე იცვლება, ანუ ფარდობითი სიხშირე $n_{\ x}/\ n$ არის x-ის ფუნქცია. ვინაიდან ეს ფუნქცია იძებნება ემპირიული გზით (ცდის მიხედვით), მას ემპირიულს უწოდებენ. განაწილების ემპირიულ ფუნქციას (შერჩევის განაწილების ფუნქცია) ეწოდება ფუნქცია *F*(x)*-ს, რომელიც x-ის თითოეული მნიშვნელობისთვის განსაზღვრავს X <x მოვლენის ფარდობით სიხშირეს.

ამრიგად, განსაზღვრის თანახმად,

$$F^*(x) = n_{\ x}/\ n,$$

სადაც $n_{\ x}$ - x-ზე ნაკლები ვარიანტების რიცხვია, n- შერჩევის მოცულობა.

ამრიგად, იმისათვის, რომ ვიპოვოთ, მაგალითად, F*(x_2), საჭიროა გავყოთ x_2-ზე ნაკლები ვარიანტების რაოდენობა შერჩევის მოცულობაზე:

$$F^*(\mathrm{x}_2) = n_{\ x_2}/\ n.$$

ნიმუშის ემპირიული განაწილების ფუნქციისგან განსხვავებით, პოპულაციის განაწილების ფუნქციას F(x) უწოდებენ *თეორიული განაწილების ფუნქციას*. განსხვავება ემპირიულ და თეორიულ ფუნქციებს შორის არის ის, რომ თეორიული ფუნქცია F(x)

განსაზღვრავს X<x ხდომილების ალბათობას, ხოლო ემპირიული ფუნქცია F*(x) განსაზღვრავს იმავე ხდომილების ფარდობით სიხშირეს. ბერნულის თეორემიდან გამომდინარეობს, რომ X <x ხდომილების ფარდობითი სიხშირე, ანუ F*(x) მიისწრაფის ალბათობით ამ მოვლენის F(x) ალბათობისაკენ. სხვა სიტყვებით რომ ვთქვათ, დიდი n-სთვის რიცხვები F*(x) და F(x) მცირედით განსხვავდება ერთმანეთისგან იმ გაგებით, რომ $\lim_{n\to\infty} P\ [|F(x) - F^*(x)| < \varepsilon] = 1\ (\varepsilon > \ 0)$. უკვე აქედან გამომდინარეობს შერჩევის ემპირიული განაწილების ფუნქციის გამოყენების მიზანშეწონილობა პოპულაციის თეორიული (ინტეგრალური) განაწილების ფუნქციის მიახლოებითი წარმოდგენისთვის.

ამ დასკვნას ადასტურებს ის ფაქტი, რომ F*(x) აქვს F (x)-ის ყველა თვისება. მართლაც, F*(x) ფუნქციის განმარტებიდან გამომდინარეობს მისი შემდეგი თვისებები:

1) ემპირიული ფუნქციის მნიშვნელობები ეკუთვნის [0, 1] მონაკვეთს;

2) F*(x) -არაკლებადი ფუნქციაა;

3) თუ x_1 არის ვარიანტებიდან ყველაზე მცირე, მაშინ F*(x) = 0 x ≤ x_1-სთვის; თუ x_k ყველაზე დიდია ვარიანტებიდან, მაშინ F*(x) = 1 ყოველი x >x_k-სთვის.

ამრიგად, შერჩევის ემპირიული განაწილების ფუნქცია წარმოადგენს ზოგადი ერთობლიობის თეორიული განაწილების ფუნქციის შეფასებას.

მაგალითი. ააგეთ ემპირიული ფუნქცია მოცემული შერჩევის განაწილებისთვის:

ვარიანტები x_i2 6 10

სიხშირეები n_i1218 30

ამოხსნა. ვიპოვოთ შერჩევის მოცულობა: 12 + 18 + 30 = 60. უმცირესი ვარიანტი ტოლია 2-ის, შესაბამისად

F*(x) = 0 როცა x ≤ 2.

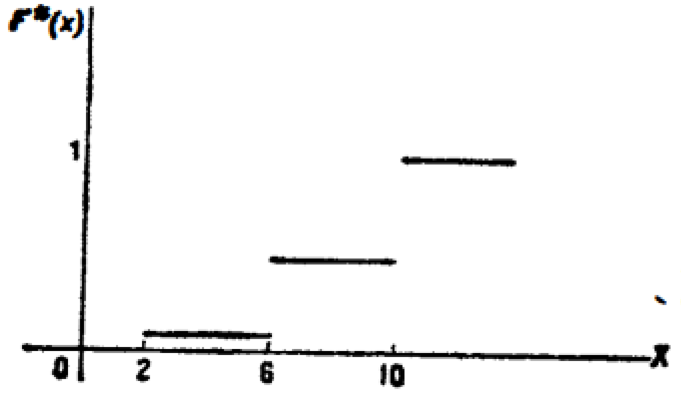
მნიშვნელობა X<6, კერძოდ $x_1 = 2$, დაიკვირვება 12-ჯერ, შესაბამისად F*(x) = 12/60=0,2, როცა 2<x≤6. მნიშვნელობა X<10, კერძოდ $x_1 = 2$,$x_2 = 6$ დაიკვირვება 12 + 18 = 30-ჯერ, შესაბამისად F*(x) = 30/60=0,5, როცა 6<x≤10. რამდენადაც x = 10 - ვარიანტებიდან უდიდესია, ამდენად

F*(x) = 1 როცა x > 10.

საძიებელ ემპირიულ ფუნქციას შემდეგი სახე აქვს:

$$F^*(x) = \begin{cases} 0 \quad \text{როცა}\, x \leq 2, \\ 0{,}2\ \text{როცა}\, 2 < x \leq 6 \\ 0{,}5\ \text{როცა} 6 < x \leq 10 \\ 1\ \text{როცა}\, x > 10 \end{cases}$$

ამ ფუნქციის გრაფიკი გამოსახულია ნახ.15-ზე



ნახ. 15

15.7 პოლიგონი და ჰისტოგრამა

თვალსაჩინოებისათვის აგებენ სტატისტიკური განაწილების სხვადასხვა გრაფიკებს. კერძოდ, პოლიგონს და ჰისტოგრამას.

სიხშირეების პოლიგონს უწოდებენ ტეხილს, რომლის მონაკვეთები ერთმანეთთან აერთებს $(x_1; n_1)$, $(x_2; n_2)$, ⋯, $(x_k; n_k)$ წერტილებს. პოლიგონის ასაგებად აბსცისათა ღერძზე გადაიზომება x_i ვარიანტები, ხოლო ორდინატაზე - მათი შესაბამისი n_i სიხშირეები. წერტილები (x_i ; n_i) ერთდება წრფის მონაკვეთებით და მიიღება სიხშირეების პოლიგონი.

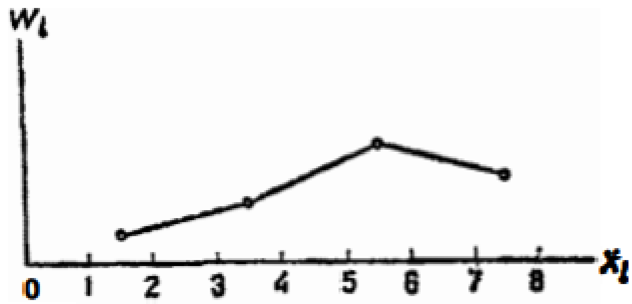
ფარდობითი სიხშირეების პოლიგონს უწოდებენ ტეხილს, რომლის მონაკვეთები ერთმანეთთან აერთებს $(x_1; W_1)$, $(x_2; W_2)$, ⋯, $(x_k; W_k)$ წერტილებს. პოლიგონის ასაგებად აბსცისათა ღერძზე გადაიზომება x_i ვარიანტები, ხოლო ორდინატაზე - მათი შესაბამისი

W_i ფარდობითი სიხშირეები. წერტილები (x_i; W_i) ერთდება წრფის მონაკვეთებით და მიიღება ფარდობითი სიხშირეების პოლიგონი.

ნახ. 16-ზე გამოსახულია ფარდობითი სიხშირეების შემდეგი განაწილების პოლიგონი:

X 1,53,5 5,5 7,5

W 0,10,2 0,4 0,3



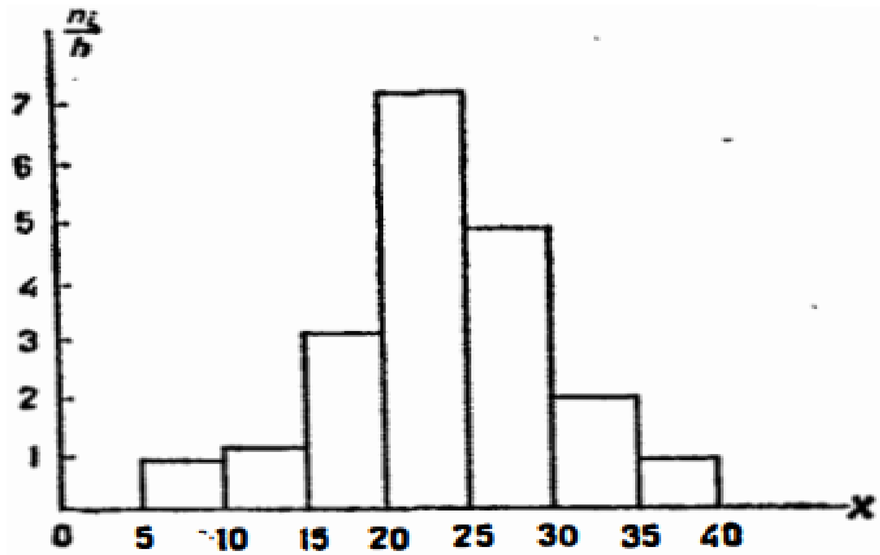
ნახ. 16

იმ შემთხვევაში, როცა დაკვირვების შედეგები უწყვეტია, მიზანშეწონილია ჰისტოგრამის აგება, რომლისთვისაც ყველა დაკვირვებული მნიშვნელობის შემცველ ინტერვალს ყოფენ h სიგრძის რამდენიმე ინტერვალად და თითოეული ინტერვალისთვის პოულობენ იმ ვარიანტების სიხშირეების ჯამს n_i , რომლებიც მოხვდნენ i-ურ ინტერვალში.

სიხშირის ჰისტოგრამა არის საფეხურებიანი ფიგურა, რომელიც შედგება მართკუთხედებისგან, რომელთა ფუძეები არის h სიგრძის ინტერვალები, ხოლო სიმაღლეები უდრის ფარდობას n_i/ h (სიხშირის სიმკვრივე).

სიხშირეების ჰისტოგრამის ასაგებად, ინტერვალები გადაზომილია აბსცისათა ღერძზე, ხოლო მათ ზემოთ გავლებულია აბსცისათა ღერძის პარალელური მონაკვეთები n_i/ h მანძილზე.

i-ური მართკუთხედის ფართობი უდრის $\frac{hn_i}{h} = n_i$ - i-ურ ინტერვალში შემავალი ვარიანტების სიხშირეების ჯამს; აქედან გამომდინარე, სიხშირის ჰისტოგრამის ფართობი უდრის ყველა სიხშირის ჯამს, ანუ შერჩევის მოცულობას.



ნახ. 17

ნახ. 17 გვიჩვენებს n = 100 მოცულობის სიხშირეების განაწილების ჰისტოგრამას, რომელიც მოცემულია ცხრილში. 6.

ფარდობითი სიხშირეების ჰისტოგრამას უწოდებენ საფეხურებიან ფიგურას, რომელიც შედგება მართკუთხედებისგან, რომელთა ფუძეები არის h სიგრძის ინტერვალები, ხოლო სიმაღლეები უდრის W_i/h თანაფარდობას (ფარდობითი სიხშირის სიმკვრივე).

ფარდობითი სიხშირეების ჰისტოგრამის ასაგებად აბსცისათა ღერძზე გადაზომავენ ინტერვალებს და მათ ზემოთ ავლებენ აბსცისათა ღერძის პარალელურად მონაკვეთებს W_i/h ტოლ მანძილზე. i-ური მართკუთხედის ფართობი უდრის $\frac{hW_i}{h} = W_i$- იმ ვარიანტის ფარდობითი სიხშირეს, რომელიც მოხვდა i-ურ ინტერვალში. შესაბამისად, *ფარდობითი სიხშირეების ჰისტოგრამის ფართობი უდრის ყველა ფარდობითი სიხშირის ჯამს, ანუ ერთს.*

ცხრილი6

h =5 სიგრძის ნაწილობრივი ინტერვალი	n_i ნაწილობრივი ინტერვალის სიხშირეების ჯამი	n_i / h სიხშირის სიმკვრივე
5 – 10	4	0,8
10 -15	6	1,2
15 – 20	16	3,2
20 – 25	36	7,2
25 – 30	24	4,8
30 – 35	10	2,0
35 - 40	4	0,8

15.8 შეფასების სიზუსტე სანდოობა და ნდობის ინტერვალი

შეფასებას უწოდებენ წერტილოვანს, თუ ის განისაზღვრება ერთი რიცხვით. შერჩევის მცირე მოცულობის დროს, წერტილოვანი შეფასება შეიძლება მნიშვნელოვნად განსხვავდებოდეს შესაფასებელი პარამეტრისგან, ანუ მიგვიყვანოს უხეშ შეცდომებამდე. ამ მიზეზით, როდესაც შერჩევის მოცულობა მცირეა, გამოყენებული უნდა იქნას ინტერვალური შეფასებები.

ინტერვალური ეწოდება შეფასებას, რომელიც განისაზღვრება ორი რიცხვით - ინტერვალის ბოლოებით. ინტერვალური შეფასებები საშუალებას იძლევა დადგინდეს შეფასებების სიზუსტე და სანდოობა.

ვთქვათ, შერჩევის მონაცემებით ნაპოვნი სტატისტიკური მახასიათებელი θ^* წარმოადგენს უცნობი θ პარამეტრის შეფასებას. ჩვენ განვიხილავთ θ -ს როგორც მუდმივ რიცხვს (θ შეიძლება იყოს შემთხვევითი ცვლადიც). ნათელია, რომ θ^* რაც უფრო ზუსტად განსაზღვრავს θ პარამეტრს, მით უფრო მცირეა სხვაობის $|\ \theta - \theta^*\ |$ აბსოლუტური მნიშვნელობა. სხვა სიტყვებით რომ ვთქვათ, თუ $\delta > 0$ და $|\ \theta - \theta^*| < \delta$, მაშინ რაც უფრო ნაკლებია δ, შეფასება მით უფრო ზუსტია. ამრიგად დადებითი რიცხვი δ ახასიათებს *შეფასების სიზუსტეს.*

თუმცა, სტატისტიკური მეთოდები არ იძლევა კატეგორიულად მტკიცების საშუალებას, რომ შეფასება θ^* აკმაყოფილებს უტოლობას $|\theta - \theta^*| < \delta$; აქ მხოლოდ იმ γ ალბათობაზე საუბარი, რომლითაც ეს უტოლობა სრულდება.

θ პარამეტრის θ^* შეფასების სანდოობას (ნდობის ალბათობას) უწოდებენ γ ალბათობას, რომლითაც სრულდება $|\theta - \theta^*| < \delta$ უტოლობა. ჩვეულებრივ, შეფასების სანდოობა მოიცემა წინასწარ, ამასთან γ-ს მნიშვნელობად იღებენ ერთთან ახლოს მდებარე რიცხვს. ყველაზე ხშირად აღებული სანდოობა 0,95-ის; 0.99-ის და 0.999-ის ტოლია.

ვთქვათ, იმის ალბათობა, რომ $|\theta - \theta^*| < \delta$ არის γ-ს ტოლი

$$P[|\theta - \theta^*| < \delta] = \gamma.$$

თუ $|\theta - \theta^*| < \delta$ უტოლობას შევცვლით მისი ეკვივალენტური ორმაგი უტოლობით $-\delta < \theta - \theta^* < \delta$ ანუ $\theta^* - \delta < \theta < \theta^* + \delta$, გვექნება

$$P[\theta^* - \delta < \theta < \theta^* + \delta] = \gamma$$

ეს თანაფარდობა შემდეგნაირად უნდა გავიგოთ: ალბათობა იმისა, რომ ინტერვალი $(\theta^* - \delta, \theta^* + \delta)$ შეიცავს (ფარავს) უცნობ θ პარამეტრს, უდრის γ-ს.

ნდობის ინტერვალს უწოდებენ $(\theta^* - \delta, \theta^* + \delta)$ ინტერვალს რომელიც მოიცავს უცნობ პარამეტრს მოცემული სანდოობით.

ნდობის ინტერვალის მეთოდი შეიმუშავა ამერიკელმა სტატისტიკოსმა ჯ. ნეიმანმა, ინგლისელი სტატისტიკოსის რ. ფიშერის იდეებზე დაყრდნობით.

15.9 სტატისტიკური ჰიპოთეზა. ნულოვანი და ალტერნატული, მარტივი და რთული ჰიპოთეზები

ხშირად საჭიროა პოპულაციის განაწილების კანონის ცოდნა. თუ განაწილების კანონი უცნობია, მაგრამ არსებობს საფუძველი ვივარაუდოთ, რომ მას აქვს გარკვეული ფორმა (აღვნიშნოთ ის A-თი) და შეგვიძლია წამოვაყენოთ ჰიპოთეზა: პოპულაცია განაწილებულია A კანონის მიხედვით. ამრიგად, ამ ჰიპოთეზაში ჩვენ ვსაუბრობთ სავარაუდო განაწილების ტიპზე.

შესაძლებელია შემთხვევა, როდესაც ცნობილია განაწილების კანონი, ხოლო მისი პარამეტრები კი უცნობია. თუ არსებობს საფუძველი ვივარაუდოთ, რომ უცნობი პარამეტრი θ უდრის განსაზღვრულ მნიშვნელობას θ_0-ს, წამოაყენებენ ჰიპოთეზას: $\theta = \theta_0$. ამრიგად, ეს ჰიპოთეზა ეხება ცნობილი განაწილების უცნობი პარამეტრის სავარაუდო მნიშვნელობას.

შესაძლებელია სხვა ჰიპოთეზებიც: ორი ან მეტი განაწილების პარამეტრების ტოლობის შესახებ, შერჩევების დამოუკიდებლობის შესახებ და მრავალი სხვა.

სტატისტიკური ეწოდება ჰიპოთეზას უცნობი განაწილების ფორმის, ან ცნობილი განაწილების პარამეტრების შესახებ.

მაგალითად, სტატისტიკურ ჰიპოთეზებს წარმოადგენენ:

1) პოპულაცია განაწილებულია პუასონის კანონის მიხედვით;

2) ორი ნორმალური შერჩევის დისპერსიები ერთმანეთის ტოლია.

პირველ ჰიპოთეზაში კეთდება ვარაუდი უცნობი განაწილების ფორმის შესახებ, მეორეში კი ორი ცნობილი განაწილების პარამეტრებზე.

ჰიპოთეზა „მარსზე სიცოცხლე არსებობს“, არ არის სტატისტიკური, რადგან მასში არ არის საუბარი არც განაწილების ფორმაზე და არც განაწილების პარამეტრებზე.

წამოყენებულ ჰიპოთეზასთან ერთად განიხილება ჰიპოთეზაც, რომელიც ეწინააღმდეგება მას. თუ წამოყენებული ჰიპოთეზა უარყოფილია, მაშინ ადგილი აქვს საწინააღმდეგო ჰიპოთეზას. ამ მიზეზით, მიზანშეწონილია განასხვავოთ ეს ჰიპოთეზები.

ნულოვანი (ძირითადი) ჰიპოთეზა ეწოდება შემოთავაზებულ H_0ჰიპოთეზას.

ალტერნატიული (*კონკურენტული*) ჰიპოთეზა ეწოდება H_1 ჰიპოთეზას, რომელიც ეწინააღმდეგება ნულოვანს.

მაგალითად, თუ ნულოვანი ჰიპოთეზა შედგება იმ დაშვებისგან, რომ ნორმალური განაწილების მათემატიკური ლოდინი a არის 10, მაშინ კონკურენტული ჰიპოთეზა, კერძოდ, შეიძლება მდგომარეობდეს იმის დაშვებაში, რომ a ≠10. მოკლედ ეს ასე ჩაიწერება: $H_0: a = 10; H_1: a \neq 10$.

ანსხვავებენ ჰიპოთეზებს, რომელებიც შეიცავენ მხოლოდ ერთ ან ერთზე მეტ ვარაუდებს.

მარტივი ეწოდება ჰიპოთეზას, რომელიც შეიცავს მხოლოდ ერთ ვარაუდს. მაგალითად, თუ λ არის ექსპონენციალური განაწილების პარამეტრი, მაშინ ჰიპოთეზა $H_0: \lambda = 5$ - მარტივია. ჰიპოთეზა H_0: ნორმალური განაწილების მათემატიკური ლოდინი უდრის 3-ს (σ ცნობილია) - მარტივია.

რთული ეწოდება ჰიპოთეზას, რომელიც შედგება მარტივი ჰიპოთეზების სასრული ან უსასრულო რაოდენობისგან. მაგალითად, რთული ჰიპოთეზა $H_0: \lambda > 5$ შედგება უამრავი მარტივი ფორმის $H_0: \lambda = b_i$ ჰიპოთეზების სიმრავლისგან, სადაც b_i არის ნებისმიერი 5-ზე მეტი რიცხვი.

15.10 პირველი და მეორე გვარის შეცდომები

წამოყენებული ჰიპოთეზა შეიძლება იყოს ჭეშმარიტი ან მცდარი, ამიტომ საჭირო ხდება მისი შემოწმება, რომელსაც ატარებენ სტატისტიკური მეთოდების გამოყენებით. ჰიპოთეზის სტატისტიკური შემოწმების შედეგად ორ შემთხვევაში შეიძლება მივიღოთ არასწორი გადაწყვეტილება, ანუ შეიძლება ორგვარი შეცდომის დაშვება.

პირველი გვარის შეცდომა მდგომარეობს სწორი ჰიპოთეზის უარყოფაში.

მეორე გვარის შეცდომა მდგომარეობს არასწორი ჰიპოთეზის მიღებაში.

ხაზგასმით აღვნიშნავთ, რომ ამ შეცდომების შედეგები შეიძლება ძალიან განსხვავებული იყოს. მაგალითად, თუ უარყოფილია სწორი ჰიპოთეზა „საცხოვრებელი შენობის მშენებლობის გაგრძელება“, მაშინ ეს პირველი გვარის შეცდომა გამოიწვევს მატერიალურ ზიანს; თუ არასწორი გადაწყვეტილება მიიღება „მშენებლობის გაგრძელების შესახებ“, მიუხედავად სამშენებლო მოედნის დანგრევის საფრთხისა, მაშინ ამ მეორე გვარის შეცდომამ შეიძლება გამოიწვიოს ადამიანების სიკვდილი. შეიძლება მოვიყვანოთ ისეთი მაგალითებიც, როდესაც პირველი გვარის შეცდომა იწვევს უფრო სერიოზულ შედეგებს, ვიდრე მეორე გვარის შეცდომა.

გამოყენებული ლიტერატურა:

1. Berenson, Mark, David Levine, and Timothy Krehbiel. *Basic Business Statistics*, 10th ed. Upper Saddle River, NJ: Prentice Hall, 2006.
2. Campbell, S. *Flaws and Fallacies in Statistical Thinking*. Upper Saddle River, NJ: Prentice Hall, 1974.
3. Campbell, S. *Flaws and Fallacies in Statistical Thinking*. Upper Saddle River, NJ: Prentice Hall, 1974.
4. Groebner, David, Patrick Shannon, Phillip Fry, and Kent Smith. *Business Statistics*, 8th ed. Upper Saddle River, NJ: Prentice Hall, 2011.
5. Nedosekin A. Fuzzy Financial Management. AFA Library, Moscow, 2003.
6. Peck R, devore J., Statistics: The Exploration and Analysis of Data, Thomson Learning, Inc., 2008.
7. Render B., Stair R.M., Hanna M.E., Quantitative Analysis For management, Pearson, 2012.
8. Гмурман В.Е. теория вероятностей и математическая статистика. М. "Высшая школа", 1999.