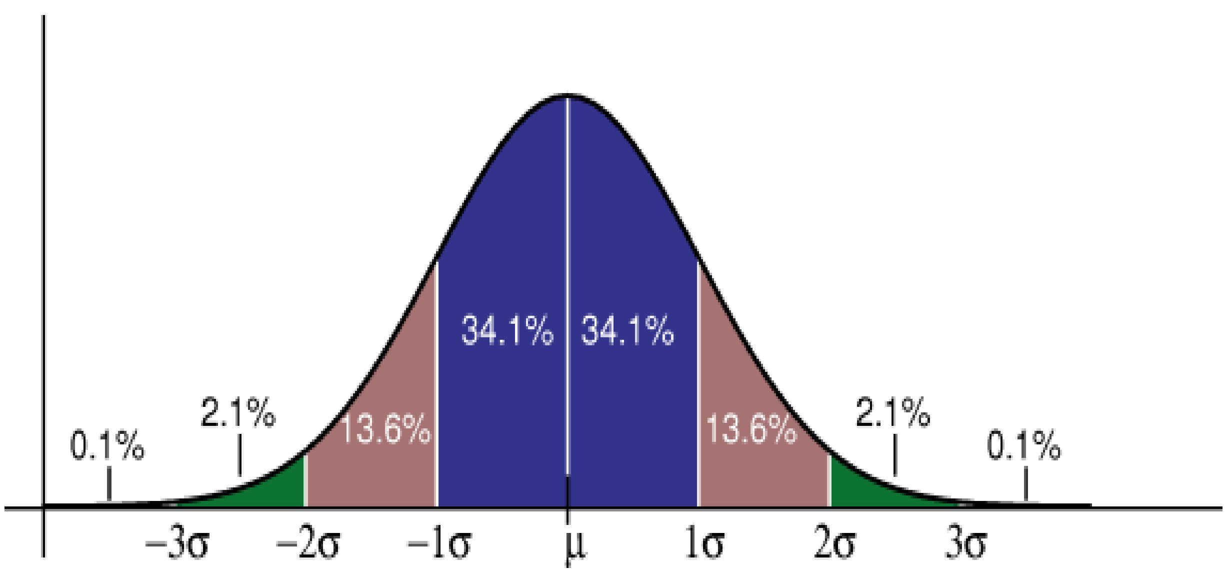
ომარ ფურთუხია

ალბათობის თეორია და მათემატიკური სტატისტიკა ეკონომისტებისათვის



თ ს უ – 2 0 1 2

სარჩევი

ალბათობის თეორია

მათემატიკური სტატისტიკა

§0. შესავალი

ადამიანის ყოველდღიურ ცხოვრებაში სიტყვა “ალბათობა” ხშირად გამოიყენება ამა თუ იმ ხდომილების მოხდენის ან არ მოხდენის დამაჯერებლობის ხარისხის ცვლილების გამოსახატავად, რაც გარკვეული აზრით დაკავშირებულია ჩვენს სუბიექტურ სურვილებთან. ასეთია მაგალითად, შემდეგი შინაარსის მტკიცებულები: “ხვალ ალბათ გამოიდარებს”, “თვის ბოლოსათვის ლარი ალბათ გამყარდება”, “ორ წელიწადში საქართველო ალბათ ნატოს წევრი გახდება” და ა. შ. იმ ხდომილებებს შორის, რომელთაც ჩვენ ვახასიათებთ როგორც ნაკლებად ალბათური, ან როგორც საკმარისად ალბათური, ან როგორც ძალიან ალბათური, შეიძლება განვასხვავოთ სამი კატეგორია: ისინი, რომლებიც შეეხება ჩვენს საკუთარ ქცევას; ისინი, რომლებიც შეეხება სხვა ადამიანების ქცევას, და ისინი, რომლებიც შეეხება ბუნების მოვლენებს. ამასთანავე, ამ სამ კატეგორიას შორის სხვაობა ზოგჯერ განუსაზღვრელია.

როცა მე ვამბობ: “სავსებით ალბათურია, რომ მე ხვალ დილით გავალ სახლიდან”, მე ზოგად ფორმაში ვაჯამებ მთელ რიგ მეტად თუ ნაკლებად რთულ მსჯელობებს, რომელთაგან ნაწილი შეეხება სხვადასხვა ადამიანის ქცევას, ხოლო დანარჩენი – ბუნების მოვლენებს. მაგალითად, მე გადავწყვიტე გავიდე სახლიდან, თუ კი ის ჩემი მეგობარი, რომელმაც მე გამაფრთხილა, რომ შესაძლოა მოვიდეს ჩემს სანახავად, არ მოვა ჩემთან; ან მე გადავწყვიტე გავიდე სახლიდან, თუ კი არ მოვა თოვლი (იმ შემთხვევაში, როცა დეკემბერია), -- ან თუ არ იქნება სეტყვა (იმ შემთხვევაში, როცა ივლისია). როგორც ერთ, ისე მეორე შემთხვევაში ჩემი სახლიდან გასვლის ალბათობა დამოკიდებულია იმ ალბატობაზე, რომელსაც მე მივაწერ მეორე პირის ქცევას ან ამა თუ იმ მეტეოროლოგიურ მოვლენას. გარდა ამისა, გასათვალისწინებელია ის ალბათობაც, რასაც კარგი ჯანმრთელობის მქონე ადამიანები უგულებელყოფენ, როდესაც ისინი გეგმავენ რაიმე მოკლევადიან ქმედებას: მე ვერ შევძლებ სახლიდან გავიდე, თუ მე ძალიან ავად გავხდები, და მით უმეტეს, თუ მოვკვდები.

მათემატიკაში კი სიტყვა “ალბათობა” გამოიყენება მკაცრად განსაზღვრული აზრით, რომელსაც არანაირი კავშირი არა აქვს დამაჯერებლობასთან, და მით უმეტეს სუბიექტურ სურვილებთან. ალბათობის თეორია წარმოადგენს მეცნიერებას შემთხვევითობის შესახებ. მისი საშუალებით აღიწერება სამყაროს მრავალი მოვლენა და სიტუაცია. ჯერ კიდევ შორეულ წარსულში პირველყოფილი ტომის ბელადმა იცოდა, რომ 10 მონადირეს გაცილებით მეტი “ალბათობით” შეუძლია ისარი მოარტყას ირემს, ვიდრე ერთ მონადირეს. ამიტომაც, ისინი კოლექტიურად ნადირობდნენ. არასწორი იქნებოდა გვეფიქრა, რომ მორიგი ომისათვის მზადების პროცესში, ალექსანდრე მაკედონელი ან სხვა რომელიმე გამოჩენილი მხედართმთავარი მხოლოდ მეომართა მამაცობაზე და საბრძოლო ხელოვნებაზე ამყარებდა იმედს. ეჭვგარეშეა, რომ დაკვირვებებისა და სამხედრო ხელმძღვანელობის გამოცდილების საფუძველზე მათ შეეძლოთ როგორღაც შეეფასებინათ თავიანთი გამარჯვების ან დამარცხების “ალბათობა”, იცოდნენ როდის უნდა ჩაბმულიყვნენ ომში და როდის უნდა აერიდებინათ მისთვის თავი. ცხადია, რომ ისინი არ იყვნენ შემთხვევითობის მონები, მაგრამ იმავდროულად ძალიან შორს იდგნენ ალბათობის თეორიისაგან.

მოგვიანებით, გამოცდილების დაგროვებასთან ერთად, ადამიანმა სულ უფრო ხშირად დაიწყო შემთხვევითი მოვლენების – დაკვირვებებისა და ცდების (ექსპერიმენტების) დაგეგმვა, მათი შედეგების კლასიფიცირება, როგორც შეუძლებელი, შესაძლებელი და აუცილებელი შედეგები. ადამიანმა შეამჩნია, რომ შემთხვევითობებს არც თუ ისე იშვიათად საფუძვლად უდევს (წარმართავს) ობიექტური კანონზომიერებები. განვიხილოთ უმარტივესი ცდა – მონეტის აგდება. გერბის ან საფასურის მოსვლა, ცხადია შემთხვევითი მოვლენაა. მაგრამ მონეტის მრავალჯერადი აგდებისას შესაძლებელია შევამჩნიოთ, რომ გერბის მოსვლა ხდება დაახლოებით ცდათა რიცხვის ნახევარჯერ (მე-18 საუკუნეში ბუნებისმეტყველმა ბიუფონმა მონეტა ააგდო 4040-ჯერ, საიდანაც გერბი მოვიდა 2048-ჯერ; მე-20 საუკუნის დასაწყისში მათემატიკოსმა პირსონმა მონეტა ააგდო 24000-ჯერ და გერბი მოვიდა 12012-ჯერ). მასასადამე, მიუხედავად იმისა, რომ მონეტის ცალკეული აგდების შედეგი შემთხვევითი ხდომილებაა, მონეტის მრავალჯერადი აგდების შედეგები ობიექტურ კანონს ემორჩილება.

განვიხილოთ მეორე მაგალითი – ექსპერიმენტი ე. წ. გალტონის დაფით. გვაქვს ვერტიკალურ დაფაზე სამკუთხედის ფორმით დამაგრებული რგოლები, ისე რომ წვეროში ერთი რგოლია, მეორე სტრიქონში წინასგან თანაბრ მანძილებზე ორი რგოლი, მესამე სტრიქონში ზედა ორი რგოლიდან თანაბარ მანძილებზე სამი რგოლი და ა.შ. პირველი რგოლის თავზე დგას რეზერვუარი, საიდანაც ვარდებიან ბურთულები ქვევით და გროვდებიან რგოლების ბოლოში მოთავსებულ მართკუთხედებში. თითოეულ ბურთულას მორიგ რგოლზე დაცემისას შეუძლია შემთხვევით გადავარდეს მარჯვნივ ან მარცხნივ და აღმოჩნდეს ბოლოში განთავსებულ ნებისმიერ მართკუთხედში. აღმოჩნდა, რომ ექსპერიმენტიდან ექსპერიმენტამდე მეორდება ბურთულების სიმეტრიული განლაგება მართკუთხედებში, რომლის დროსაც ცენტრალურ მართკუთხედებში ბურთულები ბევრია, ხოლო განაპირა მართკუთხედებში – ცოტა. ეს დამაჯერებლად მიუთითებს ბურთულების განაწილების ობიექტური კანონის არსებობის შესახებ. როცა ბურთულები ბევრია, მაშინ ამბობენ, რომ ისინი განაწილებულია ნორმალური კანონის მიხედვით.

ამრიგად, შეიძლება ითქვას, რომ შემთხვევითობა შეიძლება ემორჩილებოდეს შედარებით მარტივ და შედარებით რთულ კანონზომიერებას. მაგრამ, იბადება კითხვა, სად არის აქ მათემატიკა და მათემატიკური ამოცანები?

შეიძლება ითქვას, რომ ალბათობის თეორის განვითარება დაიწყო აზარტული თამაშების დროს წარმოშობილი ამოცანებიდან, თუმცა მისი საფუძვლების ფორმირებას ხელი შეუწყო დემოგრაფიულ მონაცემებში აღმოჩენილმა კანონზომიერებებმა (ახალშობილთა სტატისტიკის, სიკვდილიანობის სტატისტიკისა და უბედურ შემთხვევათა სტატისტიკის შესწავლამ), რაც თავის მხრივ, ეფექტურად გამოიყენებოდა სადაზღვევო კომპანიების საქმიანობაში. მოგვიანებით, ობიექტური კანონზომიერებები აღმოჩენულ იქნა შემთხვევითი მოვლენების შესწავლისას ადამიანის მოღვაწეობის ყველა სფეროში. ბუნებრივია დავიწყოთ მარტივი ამოცანების განხილვით.

შუა საუკუნეების ბოლომდე ძვლებით თამაში ყველაზე პოპულარული აზარტული თამაში იყო. თვითონ სიტყვა “აზარტი” ასევე დაკავშირებულია ძვლებით თამაშთან, რამდენადაც ის მოდის არაბული სიტყვიდან “alzar”, რომელიც ითარგმნება როგორც – “სათამაშო ძვალი”. სიტყვა “აზარ” არაბულად

აგრეთვე ნიშნავს რთულს. არაბები აზარტულ თამაშს უწოდებდნენ ქულების ისეთ კომბინაციას, რომელიც შეიძლება გამოჩნდეს ერთადერთი გზით, რამოდენიმე სათამაშო კამათლის გაგორებისას. მაგალითად, ორი კამათლის გაგორებისას რთულად ("აზარ") ითვლებოდა ჯამში ორი ან თორმეტი ქულის მოსვლა. აღსანიშნავია, რომ სიტყვა "hasard" ფრანგულად ნიშნავს შემთხვევითობას, ხოლო "jeu de hasard" კი – აზარტულ თამაშს.

მიუხედავად იმისა, რომ დღეს ალბათობის თეორიას იმდენივე საერთო აქვს აზარტულ თამაშებთან, რამდენიც გეომეტრიას ფართობების გაზომვასთან მიწის სამუშაოების დროს, ალბათობის თეორიის პირველი პარადოქსები დაკავშირებულია სწორედ პოპულარულ აზარტულ თამაშებთან. 1494 წელს იტალიელმა მათემატიკოსმა ლ. პაჩოლიმ (1445-1514) გამოაქვეყნა ნაშრომი, რომელშიც იხილავდა შემდეგ სიტუაციას: ორი ტოლძალოვანი მოთამაშე შეთანხმდა ეთამაშათ გარკვეული თამაში, მანამ სანამ ერთი მათგანი არ მოიგებდა n პარტიას. ამ შემთხვევაში ის იღებდა გარკვეულ თანხას (პრიზს). მაგრამ თამაში შეწყდა მას შემდეგ რაც პირველმა მოთამაშემ მოიგო k ($k<n$), ხოლო მეორემ -- m ($m<n$) პარტია. მოსაგები თანხის როგორი განაწილება იქნება სამართლიანი? ამ ამოცანას შემდგომში ეწოდა *პრიზის განაწილების პარადოქსი*. მიუხედავად იმისა, რომ სინამდვილეში ეს ამოცანა არ წარმოადგენს პარადოქსს, ზოგიერთი უდიდესი მეცნიერის მიერ ამ ამოცანის ამოხსნის წარუმატებელმა მცდელობამ, და არასწორმა ურთიერთსაწინააღმდეგო პასუხებმა წარმოქმნეს ლეგენდა პარადოქსის შესახებ.

თვითონ პაჩოლიმ სწორი ამოხსნა ვერ იპოვა. იგი ვერ ხედავდა ამ ამოცანის კავშირს ალბათობის თეორიასთან და იხილავდა მას როგორც ამოცანას პროპორციებზე. ამიტომ ის თვლიდა, რომ თანხა უნდა განაწილებულიყო პროპორციით $k:m$, არ ითვალისწინებდა რა პარტიათა იმ რაოდენობას, რომელიც უნდა მოიგოს ცალკეულმა მოთამაშემ, რათა მიიღოს მთლიანი თანხა. არასწორი ამოხსნა ეკუთვნის ნიკოლო ტარტალიასაც (1499-1557), მიუხედავად იმისა, რომ ის იყო საკმარისად გენიალური, რათა მათემატიკურ დუელში ერთი ღამის განმავლობაში ეპოვა კუბური განტოლების ამოხსნის ფორმულა. 50 წლის შემდეგ, მეორე იტალიელმა მათემატიკოსმა დ. კარდანომ (1501-1576), სამართლიანად გააკრიტიკა პაჩოლის მსჯელობა, მაგრამ სამწუხაროდ თვითონაც მოგვცა მცდარი ამოხსნა. გავიდა კიდევ 100-ზე მეტი წელი, და მხოლოდ 1654 წელს, აღნიშნული ამოცანა ამოხსნილ იქნა გამოჩენილი ფრანგი მათემატიკოსების ბ. პასკალისა (1623-1662) და პ. ფერმას (1601-1665) მიმოწერის პროცესში ერთმანეთისაგან დამოუკიდებლად. ეს აღმოჩენა იყო იმდენად მნიშვნელოვანი, რომ ბევრი თვლის ამ წელს ალბათობის თეორიის დაბადების წლად, ხოლო ყველა ადრინდელ შედეგს – წინაისტორიად. შევხედოთ როგორ ხსნიდა პასკალი ამოცანას, როცა $n=3$, $k=2$ და $m=1$. პასკალი და ფერმა განიხილავდნენ ამ პრობლემას როგორც ამოცანას ალბათობებზე. ამიტომ სამართლიანი იქნება ისეთი გაყოფა, რომელიც პროპორციულია თითოეული მოთამაშის მიერ პრიზის მოგების შანსის.

დავუშვათ, რომ თამაში შეწყდა, როცა პირველ მოთამაშეს მოგებული აქვს ორი პარტია, ხოლო მეორეს – ერთი. ჯერ-ჯერობით უცნობია როგორ გავანაწილოთ თანხა, მაგრამ ყველაფერი გამარტივდებოდა, თუ ისინი ითამაშებდნ-

ენ კიდევ ერთ პარტიას. სინამდვილეში, ამ შემთხვევაში შესაძლებელია ორი შედეგი:

I. თუ ამ პარტიას მოიგებს პირველი მოთამაშე, მაშინ მას დაუგროვდება მოგებათა შეთანხმებული რიცხვი და მიიღებს მთლიან თანხას;

II. თუ პარტიას მოიგებს მეორე მოთამაშე, მაშინ ორივე ექნება მოგებათა თანაბარი რაოდენობა და სამართლიანი იქნებოდა თანხის თანაბრად გაყოფა.

თითოეულ ამ შედეგს მოხდენის თანაბარი შესაძლებლობა აქვს.

ამრიგად, პირველ მოთამაშეს შეუძლია მოიგოს ან მთელი თანხა ან თანხის ნახევარი, ანუ საშუალოდ მას შეუძლია მოიგოს თანხის

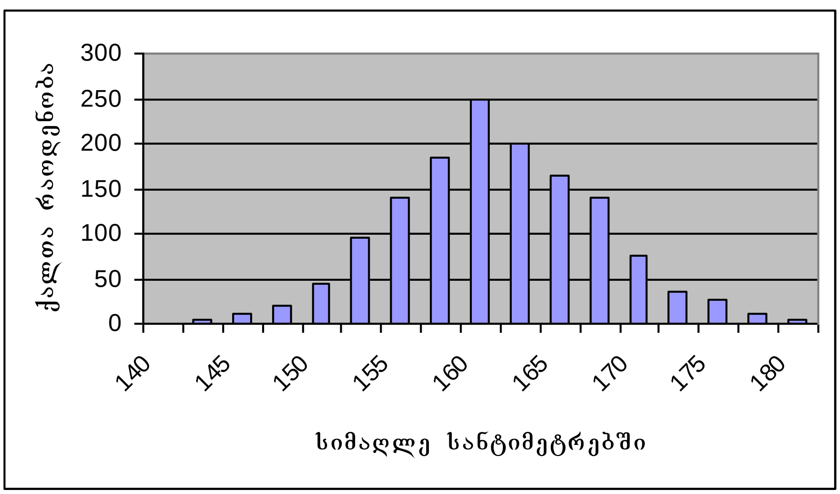
$$\frac{1+\frac{1}{2}}{2}=\frac{3}{4}$$

ნაწილი. მეორე მოთამაშის შესაძლებლობები უფრო მწირია: მან შეიძლება ან არაფერი მოიგოს ან თანხის ნახევარი მოიგოს, ანუ იგი საშუალოდ იგებს თა-ხის

$$\frac{0+\frac{1}{2}}{2}=\frac{1}{4}$$

ნაწილს. ამიტომ თანხა უნდა განაწილდეს პროპორციით 3:1 (და არა 2:1, როგორც თვლიდა პაჩოლი).

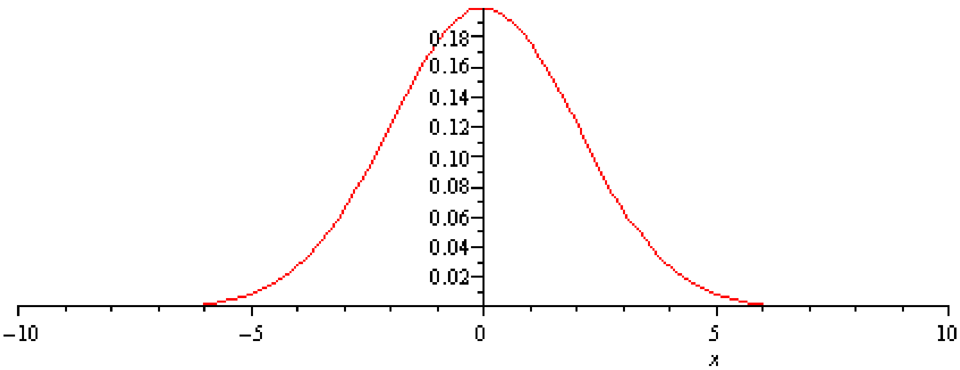
1718 წელს ლონდონში გამოვიდა ფრანგი მათემატიკოსის ა. მუავრის (1667-1754) წიგნი სახელწოდებით -- “სწავლება შემთხვევითობაზე”, რომლის მთავარი მიღწევაა იმ კანონზომიერების დადგენა, რომელიც ძალიან ხშირად შეიმჩნევა შემთხვევით მოვლენებში. მუავრმა პირველმა აღმოაჩინა და თეორიულად დაასაბუთა “ნორმალური” განაწილების როლი (გაიხსენეთ გალტონის დაფა). მუავრმა გაზომა 1375 შემთხვევით შერჩეული ქალის სიმაღლე. გაზომვის შედეგები მოყვანილია დიაგრამაზე:



ზარის მსგავსი წირი, რომელიც დაახლოებით “ედება” სიმაღლეთა განაწილების დიაგრამას, ახლოსაა

$$y=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$$

ფუნქციის გრაფიკთან, რომელსაც აქვს შემდეგი სახე:



ნორმალური განაწილების კანონს აქვს უდიდესი პრაქტიკული მნიშვნელობა. აღმოჩნდა, რომ ამ კანონითაა განაწილებული გაზის მოლეკულების სიჩქარე, ახალშობილების წონა, გაზომვის ცდომილებათა სიდიდე, და მრავალი სხვა ფიზიკური და ბიოლოგიური ბუნების მქონე შემთხვევითი სიდიდე. აღსანიშნავია, რომ ნორმალური განაწილების კანონი იძლევა ძალიან კარგ მიახლოებას ყოველთვის, როცა განსახილველი სიდიდე წარმოადგენს ბევრი დამოუკიდებელი კომპონენტის ერთობლივი მოქმედების შედეგს და ამასთანავე, ჯამურ ეფექტში თითოეული კომპონენტის წვლილი შედარებით მცირეა.

დავუბრუნდეთ ისევ პრიზის განაწილების პარადოქსს. იმ შემთხვევაში, როცა $n=6$, $k=5$ და $m=3$, პაჩოლის პასუხი იყო, რომ პრიზი უნდა განაწილებულიყო მოგებული პარტიების პროპორციულად, ანუ 5:3-ზე. ტარტალია თვლიდა, რომ განაწილება უნდა მომხდარიყო 2:1-თან პროპორციით (სავარაუდოდ ის მსჯელობდა შემდეგნაირად: ვინაიდან, პირველმა მოთამაშემ მოიგო მეორეზე ორი პარტიით მეტი, რაც შეადგენს მოგებისათვის აუცილებელი 6 პარტიის მესამედს, ამიტომ პირველმა მოთამაშემ უნდა მიიღოს პრიზის მესამედი, ხოლო დარჩენილი ნაწილი – 2/3 განაწილდეს თანაბრად. ე. ი. პირველმა უნდა მიიღოს $\frac{1}{3}+\frac{2}{3}:2=\frac{2}{3}$, ხოლო მეორემ -- $\frac{1}{3}$, ანუ გაყოფა უნდა მოხდეს 2:1-ზე). ვაჩვენოთ, რომ იმ დროს როდესაც მოგებისათვის პირველ მოთამაშეს ესაჭიროება მხოლოდ ერთი პარტიის მოგება, მეორეს კი – სამი პარტიისა, პრიზის სამართლიანი გაყოფაა 7:1-ზე.

ფერმას იდეის თანახმად, გავაგრძელოთ თამაში სამი ფიქტიური პარტიით, იმ შემთხვევაშიც კი როცა ზოგიერთი მათგანი აღმოჩნდება სრულიად ზედმეტი (ანუ, როცა ერთ-ერთ მოთამაშეს უკვე მოგებული აქვს თამაში). ასეთი გაგრძელებისას შესაძლებელია $2\cdot2\cdot2=8$ ერთნაირად მოსალოდნელი შედეგი: “მმმ”, “მმწ”, “მწმ”, “წმმ”, “მწწ”, “წმწ”, “წწმ” და “წწწ” (სადაც i-ურ ადგილზე მდგომი “მ”, შესაბამისად, “წ” აღნიშნავს, რომ i-ური პარტია მოიგო, შესაბამისად, წააგო პირველმა მოთამაშემ, $i=1,2,3$). ვინაიდან, მხოლოდ ერთ შემთხვევაში ღებულობს მეორე მოთამაშე პრიზს (როცა ის მოიგებს სამივე პარტიას), ხოლო დანარჩენ 7 შემთხვევაში იგებს პირველი მოთამაშე, ამიტომ სამართლიანია გაყოფა 7:1-ზე.

ამ ამოცანის ზოგად შემთხვევაში ამოხსნა აგრეთვე ეკუთვნით პასკალსა და ფერმას. 1654 წელს მთელი პარიზი ლაპარაკობდა ახალი მეცნიერების – ალბათობის თეორიის წარმოშობაზე. ფერმას ბრწყინვალე იდეა თამაშის გაგრ-

ძელების შესახებ 1977 წელს გამოიყენა ანდერსონმა. მან დაამტკიცა შემდეგი მნიშვნელოვანი თეორემა: მოთამაშეს, რომელიც არიგებს პირველი, აქვს ერთი და იგივე შანსი მოიგოს N პარტიაში თავის მოწინააღმდეგეზე უფრო ადრე მიუხედავად იმისა, მოთამაშეები არიგებენ მონაცვლეობით, თუ არიგებს ის ვინც მოიგო წინა პარტია.

ხდომილების ალბათობა განმარტებული იყო ლაპლასის მიერ შემდეგნაირად:

$$P(A) = \frac{m}{n},$$

სადაც n -- თანაბრადშესაძლებელ ხდომილებათა (შედეგთა) საერთო რაოდენობაა, ხოლო m -- იმ ხდომილებათა რაოდენობა, რომლის დროსაც ხდება A ხდომილება ("ხელშემწყობი შედეგების რაოდენობა"). მაგალითად, ვთქვათ, გამოსათვლელია ალბათობა ხდომილების A -- "ორი სათამაშო კამათლის გაგორებისას მოსულ ქულათა ჯამია 8". ორი კამათლის გაგორებისას შესაძლებელია მივიღოთ შემდეგი თანაბრადშესაძლებელი შედეგები:

(1,1) (1,2) (1,3) (1,4) (1,5) (1,6)   (2,1) (2,2) (2,3) (2,4) (2,5) (2,6)
(3,1) (3,2) (3,3) (3,4) (3,5) (3,6)   (4,1) (4,2) (4,3) (4,4) (4,5) (4,6)
(5,1) (5,2) (5,3) (5,4) (5,5) (5,6)   (6,1) (6,2) (6,3) (6,4) (6,5) (6,6).

როგორც ვხედავთ, სულ შესაძლებელი ვარიანტებია 36. ცალკე გამოყოფილია ის ვარიანტები, როცა მოხდა A ხდომილება. ასეთი შემთხვევებია 5, და ყველა ისინი თანაბრადშესაძლებელია. ამიტომ ლაპლასის განმარტების თანახმად

$$P(A) = 5/36.$$

ბუნებრივად იბადება კითხვა: როდის და რომელი შემთხვევითი ხდომილებები შეიძლება ჩაითვალოს თანაბრადშესაძლებლად? განვიხილოთ *დალამბერის ცნობილი შეცდომა.* ცნობილი ფრანგი მათემატიკოსი და ფილოსოფოსი ჟ. დალამბერი (1717-1783), მონეტის ორჯერადი აგდების ამოცანის განხილვისას თვლიდა, რომ გერბისა და საფასურის მოსვლის შანსი (ალბათობა) იყო 1/3. სინამდვილეში ეს არის 1/2. დალამბერის შეცდომამ თავის დროზე ბევრი კამათი გამოიწვია და ამიტომ გახდა ცნობილი. ეს შეცდომა ძალიან ჭკუის სასწავლებელია. დალამბერი მსჯელობდა შემდეგნაირად. არსებობს ამ ექსპერიმენტის სამი შესაძლო შედეგი: 1). მოვიდა ორი გერბი, 2). მოვიდა გერბი და საფასური, 3). მოვიდა ორი საფასური. თვლიდა რა დალამბერი ამ შედეგებს ტოლშესაძლებლად, ასკვნიდა, რომ გერბისა და საფასურის მოსვლის ალბათობაა 1/3.

ცნობილი მათემატიკოსის შეცდომა მდგომარეობდა იმის დაშვებაში, რომ აღნიშნული სამი შედეგი ერთნაირად შესაძლებელია. სინამდვილეში კი ეს ასე არ არის. გერბის მოსვლა ავღნიშნოთ ასოთი გ, ხოლო საფასურის – ასოთი ს. მაშინ მონეტის ორჯერადი აგდებისას ერთნაირად შესაძლებელი შედეგებია:

გგ, გს, სგ, სს.

ამ შედეგების რაოდენობა ოთხია (ე. ი. $n = 4$). აქედან "მოვიდა გერბი და საფასური" ხდომილებას ხელს უწყობს ორი შედეგი: გს და სგ, ანუ $m = 2$. შესაბამისად, გერბისა და საფასურის მოსვლის ალბათობაა: 2/4=1/2.

ალბათობის თეორიის შესწავლის დაწყებისას ბევრი, ისევე როგორც ოდესღაც დალამბერი, უშვებს შეცდომას, თვლის რა ტოლალბათურად ისეთ შედეგებს, რომლებიც სინამდვილეში ასეთები არ არიან. მაგალითად, მონეტის სამ-

ჯერ აგდებისას შეცდომაა იმის დაშვება, რომ ერთნაირად შესაძლებელია ოთხი შედეგი: 1). მოვიდა სამი გერბი, 2). მოვიდა ორი გერბი და ერთი საფასური, 3). მოვიდა ერთი გერბი და ორი საფასური, და 4). მოვიდა სამი საფასური. სინამდვილეში კი, ერთნაირადშესაძლებელია შემდეგი 8 შედეგი:

გგგ, გგს, გსგ, სგგ, გსს, სგს, სსგ, სსს.

შევადაროთ ჩამოთვლილი ოთხი ხდომილების სწორი და არასწორი ალბათობები:

	არასწორი ალბათობა	სწორი ალბათობა
მოვიდა 3 გერბი	1/4	1/8
მოვიდა 2 გერბი და 1 საფასური	1/4	3/8
მოვიდა 1 გერბი და 2 საფასური	1/4	3/8
მოვიდა 3 საფასური	1/4	1/8

სათამაშო კამათლის პარადოქსი. ორი სათამაშო კამათლის გაგორებისას მოსულ ქულათა ჯამი მოთავსებულია 2-სა და 12-ს შორის. როგორც 9, ისე 10 შესაძლებელია მივიღოთ ორი სხვადასხვა გზით: 9=3+6=4+5 და 10=4+6=5+5. სამი სათამაშო კამათლის გაგორებისას კი 9 და 10 მიიღება ექვსი სხვადასხვა გზით. მაშინ რითი აიხსნება ის გარემოება, რომ ორი კამათლის გაგორებისას უფრო ხშირად მოდის 9, ხოლო სამი კამათლის გაგორებისას კი – 10?

ამოცანა იმდენად მარტივია, რომ ძალიან გასაკვირია, რომ თავის დროზე ის ითვლებოდა ძალიან რთულად. როგორც კარდანო, ისე გალილეი აღნიშნავდნენ, რომ აუცილებელია ქულათა მოსვლის რიგის გათვალისწინება (წინააღმდეგ შემთხვევაში ყველა შედეგი არ იქნებოდა თანაბრადშესაძლებელი). ასეთ შემთხვევაში ორი კამათლის გაგორებისას 9 და 10 შესაბამისად მიიღებიან შემდეგნაირად: 9=3+6=6+3=4+5=5+4 და 10=4+6=6+4=5+5. ეს იმას ნიშნავს, რომ ორი კამათლის შემთხვევაში 9 შეიძლება გაგორდეს 4 გზით, ხოლო 10 – 3 გზით. შესაბამისად, 9-ის მიღების შანსები მეტია, ვიდრე 10-ის. სამი კამათლის შემთხვევაში სიტუაცია იცვლება საპირისპიროდ: 9 შეიძლება მიღებულ იქნეს 25 სხვადასხვა გზით, ხოლო 10 – 27 გზით. ასე, რომ ამ შემთხვევაში 10 უფრო ალბათურია, ვიდრე 9.

დე მერეს პარადოქსი. არსებობს ძველი ისტორია იმის შესახებ, რომ XVII საუკუნის ცნობილმა ფრანგმა მოთამაშემ შევალიე დე მერემ პასკალს დაუსვა აზარტულ თამაშებთან დაკავშირებული შემდეგი ამოცანა: ერთი სათამაშო კამათლის ოთხჯერ გაგორებისას ალბათობა იმისა, რომ ერთხელ მაინც მოვა 1, მეტია 1/2-ზე. მაშინ როცა ორი სათამაშო კამათლის 24-ჯერ გაგორებისას ალბათობა იმისა, რომ ერთხელ მაინც მოვა ერთდროულად ორი 1, ნაკლებია 1/2-ზე. ეს უცნაურად გამოიყურება ვინაიდან, ერთი 1-ის მოსვლის შანსები 6-ჯერ მეტია, ვიდრე ორი 1-ის მოსვლა, ხოლო 24 სწორედ 6-ჯერ მეტია 4-ზე.

პარადოქსის ახსნა: თუ წესიერ სათამაშო კამათელს ვაგორებთ k-ჯერ, მაშინ შესაძლებელ (და ტოლალბათურ) შედეგთა რაოდენობაა 6^k. აქედან 5^k შემთხვევაში არ მოვა 1, და შესაბამისად, ალბათობა იმისა, რომ კამათლის k-ჯერ გაგორებისას ერთხელ მაინც მოვა 1 ტოლია

$$(6^k - 5^k)/6^k = 1-(5/6)^k,$$

რაც მეტია 1/2-ზე თუ $k=4$. მეორეს მხრივ, ანალოგიურად, მიიღება, რომ ალბათობა იმისა, რომ კამათლის k-ჯერ გაგორებისას ერთხელ მაინც მოვა ერთდროულად ორი 1 ტოლია

$$(36^k - 35^k)/36^k = 1-(35/36)^k,$$

რაც $k=24$-სათვის ჯერ კიდევ ნაკლებია 1/2-ზე და მეტია 1/2-ზე $k=25$-დან დაწყებული.ამრიგად, ერთი კამათლისათვის "კრიტიკული მნიშვნელობაა" $k=4$, ხოლო ორი კამათლისათვის -- $k=25$.

კლასიკური ალბათობის თეორიის მიზანს წარმოადგენდა ხდომილების ფარდობითი სიხშირის გამოთვლის მეთოდების დადგენა მასთან დაკავშირებული სხვა ხდომილებების მოცემული ფარდობითი სიხშირეების საშუალებით. თანამედროვე გაგებით, ალბათობის თეორიის საგანს წარმოადგენს "შემთხვევითი მოვლენების" მათემატიკური ანალიზი.

მათემატიკური სტატისტიკა – ეს არის მათემატიკის დარგი, რომელიც შეისწავლის სტატისტიკური მონაცემების სისტემატიზაციის, ანალიზისა და გამოყენების მეთოდებს თეორიული და პრაქტიკული დასკვნების მიღების მიზნით. გარკვეული აზრით, მათემატიკური სტატისტიკის ამოცანები ალბათობის თეორიის ამოცანების საწინააღმდეგოა. კერძოდ, თუ ალბათობის თეორიის ძირითადი მიზანია რთული ხდომილებების ალბათობების გამოთვლა მოცემული ალბათური მოდელისათვის, მათემატიკური სტატისტიკა მიზნად ისახავს ხდომილებაზე დაკვირვების შედეგად გამოვავლინოთ ალბათურ-სტატისტიკური მოდელის სტრუქტურა.

ალბათობის თეორიის როგორც მეცნიერების აღმოცენება მიეკუთვნება მე-17 საუკუნის მეორე ნახევარს და დაკავშირებულია პასკალის (1623-1662), ფერმას (1601-1665) და ჰიუგენსის (1629-1695) სახელებთან. ამ მეცნიერების ჭეშმარიტი ისტორია კი იწყება ბერნულის (1654-1705), ლაპლასისა (1749-1827) და პუასონის (1781-1840) შრომებიდან, ხოლო თანამედროვე პერიოდი დაკავშირებულია ა. ნ. კოლმოგოროვის (1903-1987) სახელთან.

მათემატიკური სტატისტიკა როგორც მეცნიერება იწყება გაუსის (1777-1855) შრომებიდან. მის განვითარებაში დიდი წვლილი მიუძღვით პირსონს (1857-1936), ფიშერს (1860-1962), ნეიმანს (1894-1977), კოლმოგოროვს, სმირნოვს (1900-1966), ვალდს (1902-1950) და სხვებს.

საქართველოში ალბათობის თეორიისა და მათემატიკური სტატისტიკის სწავლების ტრადიციას საფუძველი დაუდო ანდრია რაზმაძემ (1889-1929), ხოლო მათემატიკის ამ დარგის განვითარებას საფუძველი ჩაუყარა გვანჯი მანიამ (1918-1985). მიმართულების შემდგომ განვითარებას ღრმა კვალი დააჩნია რევაზ ჩიტაშვილმა (1942-1995), მანვე დაუდო სათავე საქართველოში სტოქასტური ფინანსური ანალიზის განვითარებას.

ალბათობის თეორია და მათემატიკური სტატისტიკა ფართოდ გამოიყენება ტექნიკურ გამოკვლევებში, ეკონომიკაში (ეკონომეტრიკა), საბანკო სფეროში, მართვის (მენეჯმენტის) თეორიასა და პრაქტიკაში, სოციოლოგიაში, დემოგრაფიაში (დემოგრაფიული სტატისტიკა), ბილოგიასა და მედიცინაში (ბიომეტრია, ბიო-სამედიცინო სტატისტიკა), გეოლოგიაში, ლინგვისტიკაში, ისტორიაში და ა. შ. ალბათურ-სტატისტიკური მეთოდების გამოყენების გარეშე წარმოუდგენელია რისკის, დაზღვევის, ფინანსური საქმიანობის შეფასება და ანალიზი.

§1. ალბათობის თეორიის საგანი

ადამიანის ყოველდღიურ ცხოვრებაში სიტყვა “ალბათობა” ხშირად გამოიყენება ამა თუ იმ ხდომილების მოხდენის ან არ მოხდენის დამაჯერებლობის ხარისხის ცვლილების გამოსახატავად, რაც გარკვეული აზრით დაკავშირებულია ჩვენს სუბიექტურ სურვილებთან. ასეთია მაგალითად, შემდეგი შინაარსის მტკიცებულები: “ხვალ ალბათ გამოიდარებს”, “თვის ბოლოსათვის ლარი ალბათ გამყარდება”, “ორ წელიწადში საქართველო ალბათ ნატოს წევრი გახდება” და ა. შ. მათემატიკაში კი სიტყვა “ალბათობა” გამოიყენება მკაცრად განსაზღრული აზრით, რომელსაც არანაირი კავშირი არა აქვს დამაჯერებლობასთან და მით უმეტეს სუბიექტურ სურვილებთან.

როგორც აღნიშნული იყო ალბათობის თეორია წარმოადგენს მათემატიკის დარგს, რომელიც შეისწავლის ისეთი ექსპერიმენტების (მოვლენების) მათემატიკურ მოდელებს, რომელთა შედეგები ცალსახად არ განისაზღვრება ცდის პირობებით. ასეთ ექსპერიმენტებს ეწოდებათ *შემთხვევითი ექსპერიმენტები.* შემთხვევითი ექსპერიმენტებია: მონეტის აგდების შედეგი, მიზანში სროლისას მიზნის დაზიანება ან არდაზიანება, ხელსაწყოს მუშაობის ხანგრძლივობა, სატელეფონო სადგურში გამოძახებათა რიცხვი, ავტოსაგზაო შემთხვევათა რაოდენობა, არაბეჯითი სტუდენტის მიერ გამოცდის ჩაბარების შედეგი, ვალუტის კურსი და სხვა.

საჭიროა აღინიშნოს, რომ ალბათობის თეორია იკვლევს არა ნებისმიერ შემთხვევით ექსპერიმენტს, არამედ მხოლოდ ისეთ ექსპერიმენტებს, რომლებიც ხასიათდებიან *სტატისტიკური მდგრადობის ანუ სიხშირეთა მდგრადობის* თვისებით. ეს თვისება ხასიათდება შემდეგნაირად. განვიხილოთ შემთხვევითი ექსპერიმენტი და ვიგულისხმოთ, რომ ამ ექსპერიმენტის ჩატარების იდენტური პირობების შენარჩუნება და მისი განმეორებითი ჩატარება შესაძლებელია (თუ ფიზიკურად არა, აზრობრივად მაინც) ნებისმიერ სასურველ რაოდენობა რიცხვჯერ. აღვნიშნოთ A-თი ამ ექსპერიმენტის ერთერთი შესაძლო შედეგი. გავიმეოროთ ეს ექსპერიმენტი n-ჯერ და აღვნიშნოთ $\mu_n(A)$-თი A შედეგის (ხდომილების) მოხდენათა რიცხვი ამ n ექსპერიმენტში. მაშინ შეფარდებას $\mu_n(A)/n$ ეწოდება A ხდომილების *ფარდობითი სიხშირე,* ხოლო სიხშირეთა მდგრადობის თვისება მდგომარეობს შემდეგში:

დიდი n-ებისათვის A ხდომილების ფარდობითი სიხშირე მხოლოდ მცირედ ირხევა (n-ის ცვლილებისას) გარკვეული მუდმივი მნიშვნელობის ირგვლივ. რაც უნდა გავიგოთ შემდეგნაირად: თუ ჩავატარებთ ექსპერიმენტების რამოდენიმე სერიას, მაშინ *სტატისტიკური მდგრადობის თვისება* გულისხმობს, რომ: სიხშირეები $\mu_{n_i}(A)/n_i$ (სადაც n_i არის i-ურ სერიაში ექსპერიმენტების რიცხვია, ხოლო $\mu_{n_i}(A)$ -- კი A ხდომილების მოხდენათა რაოდენობა ამ სერიაში) ახლოს იქნებიან ერთმანეთთან (მცირედ განსხვავდებიან ერთმანეთისაგან ან რაიმე საშუალო სიდიდიდან) ყოველთვის როგორც კი n_i რიცხვები იქნებიან საკმაოდ დიდები. მაგალითად, ვ. ფელერის წიგნში მოყვანილია მონეტის აგდების 10 სერიაში ($i=1,\dots 10$), სადაც თით-

ეულ სერიაში $n_i = 1000$ ექსპერიმენტია, გერბის მოსვლათა μ_{n_i}("გ") რიცხვის შემდეგი მონაცემები: 501, 485, 509, 536, 485, 488, 500, 497, 494, 484. ცხადია, რომ აქ ფარდობითი სიხშირეები ახლოსაა ერთმანეთთან, და მაშასადამე, ექსპერიმენტს, რომელიც მდგომარეობს სიმეტრიული მონეტის აგდებაში, გააჩნია სიხშირის მდგრადობის თვისება.

შემდეგ ცხრილში მოგვყავს ის შედეგები, რომლებიც ექსპერიმენტალურად მიღებული იყო სხვადასხვა მკვლევარების მიერ მე-18 საუკუნიდან მოყოლებული სიმეტრიული მონეტის n-ჯერ აგდებისას გერბთა მოსვლის m/n ფარდობითი სიხშირისათვის:

მკვლევარი	აგდების რაოდენობა n	ფარდობითი სიხშირე m/n
ბიუფონი	4040	0.507
დე მორგანი	4092	0.5005
ჯევონსი	20480	0.5068
რომანოვსკი	80640	0.4923
კ. პირსონი	24000	0.5005
ფელერი	10000	0.4979

ეს ცხრილი გვიჩვენებს, რომ ფარდობითი სიხშირეები ახლოსაა 0.5-თთან.

მოვიყვანოთ კიდევ ერთი მაგალითი. გ. კრამერის მიერ მოყვანილი მონაცემები 1935 წელს შვეციაში დაბადებული ახალშობილების შესახებ (სადაც n -- ახალშობილთა რაოდენობაა, ხოლო m/n -- ვაჟების დაბადების ფარდობითი სიხშირე) ასე გამოიყურება:

თვეები	I	II	III	IV	V	VI	
n	7280	6957	7883	7884	7892	7609	
m/n	0.515	0.510	0.510	0.529	0.522	0.518	
თვეები	VII	VIII	IX	X	XI	XII	სულ
n	7585	7393	7203	6903	6552	7132	88273
m/n	0.523	0.514	0.515	0.509	0.518	0.527	0.517

მიუხედავად იმისა, რომ ახალშობილთა საერთო რაოდენობა იცვლება წლის განმავლობაში, ვაჟების დაბადების ფარდობითი სიხშირე საკმაოდ მდგრადად მერყეობს 0.517 – საშუალო მნიშვნელობის ირგვლივ.

ასეთი ტიპის სტატისტიკური კანონზომიერებები აღმოჩენილ იქნა უკვე მე-18 საუკუნეში დემოგრაფიულ მონაცემებში – ახალშობილთა სტატისტიკის შესწავლისას, სიკვდილიანობის სტატისტიკის შესწავლისას, უბედურ შემთხვევათა სტატისტიკის შესწავლისას და ა. შ. (რაც, თავის მხრივ, საკმაოდ ეფექტურად გამოიყენებოდა სადაზღვევო კომპანიების საქმიანობაში). მოგვიანებით, მე-19 საუკუნის ბოლოს და მე-20 საუკუნის დასაწყისში ახალი სტატისტიკური კანონზომიერებები აღმოჩენილ იქნა ფიზიკაში, ქიმიაში, ბიოლოგიაში, ეკონომიკაში და სხვა მეცნიერებებში.

ამ კანონზომიერებებს მივყავართ ალბათობის სტატისტიკური განმარტებისაკენ.

განმარტება. რიცხვს, რომლის ირგვლივაც ირხევა A ხდომილების ფარდობითი სიხშირე, ეწოდება A *ხდომილების ალბათობა* და აღინიშნება $P(A)$ სიმბოლოთი.

მათემატიკური დასაბუთება იმისა, რომ ფარდობითი სიხშირე ახლოსაა ალბათობასთან (ფარდობით სიხშირეთა მიმდევრობის ზღვარია ალბათობა და შესაბამისად, იგი ერთადერთია) მოყვანილია იაკობ ბერნულის ცნობილ თეორემაში, რომელიც ალბათობის თეორიაში ცნობილია აგრეთვე *დიდ რიცხვთა კანონის* სახელწოდებით და იგი წარმოადგენს ერთ-ერთ ფუნდამენტურ თეორემას.

აღსანიშნავია, რომ ყველა შესაძლო ექსპერიმენტი შეიძლება გაიყოს სამ კატეგორიად:

I. ექსპერიმენტები სრული მდგრადობით, სადაც საერთოდ არაა განუზღვრელობა;

II. ექსპერიმენტები სადაც არა გვაქვს სრული მდგრადობა, მაგრამ არის სტატისტიკური მდგრადობა;

III. ექსპერიმენტები სადაც სტატისტიკური მდგრადობაც კი არა გვაქვს.

პირველ ჯგუფს მიეკუთვნება უმრავლესობა ექსპერიმენტების, რომლებიც აღიწერებიან საბუნებისმეტყველო მეცნიერებების (ფიზიკა, ქიმია) კლასიკური კანონებით და მათი შესწავლა ხდება ალბათობის თეორიის გამოყენების გარეშე. მესამე კატეგორიისათვის ალბათობის თეორია გამოუსადეგარია. მეორე ჯგუფი წარმოადგენს სწორედ ალბათობის თეორიის გამოყენების არეს. თვითონ ალბათობის თეორიის მიერ გადასაწყვეტი ამოცანები იყოფა ორ დიდ ჯგუფად. პირველი ჯგუფის ამოცანებს შეიძლება ვუწოდოთ ამოცანები რთული ხდომილებების ალბათობების გამოთვლაზე, როცა ცნობილია მარტივი ხდომილებების ალბათობები. მაგალითად, თუ ცნობილია, რომ გერბის მოსვლის ალბათობა $P\{"გ"\}=1/2$, ვიპოვოთ ალბათობები იმისა, რომ ზემოთ მოყვანილ მაგალითში $\mu_{1000}("გ")$ ტოლია 501-ის, ან 485-ის, ან და ა. შ. 484-ის.

ამოცანების მეორე ჯგუფი გარკვეული აზრით პირველი ჯგუფის ამოცანების შებრუნებულია. აქ ექსპერიმენტების სერიის შედეგების საფუძველზე რაიმე გზით უნდა შევაფასოთ მარტივი ხდომილებების ალბათობები. ვინაიდან, მარტივი ხდომილებების ალბათობები უცნობია იმ ამოცანებში, რომლებიც პრაქტიკაში წარმოიშობა, ამიტომ ამოცანების ეს ჯგუფი განსაკუთრებით საინტერესოა გამოყენებების თვალსაზრისით. ალბათობის თეორიის იმ ნაწილს, რომელიც ახდენს მეორე გჯუფის ამოცანების ზუსტ დასმას და იკვლევს მათი ამოხსნის მეთოდებს, მათემატიკური სტატისტიკა ეწოდება.

§2. ელემენტარულ ხდომილებათა სივრცე

ალბათობის მათემატიკური თეორიის პრაქტიკული ღირებულება და მნიშვნელობა წარმოჩინდება ისეთ ნამდვილ თუ წარმოსახვით ცდებთან და მოვლენებთან დაკავშირებით, როგორიცაა მაგალითად, მონეტის ერთჯერადი აგდება, მონეტის აგდება 100-ჯერ, ორი სათამაშო კამათლის გაგორება, კარტის დარიგება, “რულეტკის” თამაში, ადამიანის ან რადიოაქტიური ატომის სიცოცხლის ხანგრძლივობაზე დაკვირვება, ადამიანების გარკვეული შემთხვევითი ჯგუფის შერჩევა და მათში ცაციების რაოდენობის დათვლა, ორი სახეობის მცენარის შეჯვარება და შედეგზე დაკვირვება, სატელეფონო სადგურის დაკავებული ხაზების ან სატელეფონო გამოძახებათა რიცხვის განსაზღვრა, ელექტრონული სისტემების შემთხვევითი ხმაურები, სამრეწველო პროდუქციის ხარისხის შერჩევითი კონტროლი, ავტოსაგზაო შემთხვევათა რაოდენობა, ახალშობილის სქესი, ორმაგი ვარსკვლავების რიცხვი ცის გარკვეულ უბანზე, სითხეში ჩავარდნილი მტვრის მცირე ნაწილაკის მდებარეობა, ვალუტის კურსი, ფასების ინდექსი. ჯერ-ჯერობით ყველა ამ მოვლენის აღწერა საკმაოდ ბუნდოვანია, და, იმისათვის, რომ თეორიას მიეცეს ზუსტი აზრი, ჩვენ უნდა შევთანხმდეთ იმაზე, თუ რა გვესმის ჩვენ განსახილველი ცდის ან დაკვირვების *შესაძლებელი შედეგების* ქვეშ.

მონეტის აგდების შედეგად არაა აუცილებელი მოვიდეს გერბი ან საფასური: მონეტა შეიძლება დადგეს წიბოზე ან გაგორდეს შორს (დაიკარგოს). მიუხედავად ამისა, ჩვენ ვთანხმდებით განვიხილოთ გერბი და საფასური როგორც მონეტის აგდების ორად-ორი შესაძლებელი შედეგი. ეს შეთანხმება ამარტივებს თეორიას და არ ახდენს გავლენას მის შესაძლო გამოყენებებზე. ხშირად ასეთი შეანხმებები აუცილებელია. შეუძლებელია უშეცდომოდ გაიზომოს რაიმე ატომის არსებობის ხანგრძლივობა ან რომელიმე პირის სიცოცხლის ხანგრძლივობა. მიუხედავად ამისა, მეცნიერული მიზნებისათვის სასარგებლოა ეს სიდიდეები ჩაითვალოს ზუსტ რიცხვებად. მაგრამ ამასთანავე, იბადება კითხვა: რომელი რიცხვი შეიძლება და რომელი -- არა წარმოადგენდეს ადამიანის სიცოცხლის ხანგრძლივობას? არსებობს თუ არა მაქსიმალური ასაკი, რომლის ზემოთაც სიცოცხლე შეუძლებელია, თუ ასაკი შეიძლება იყოს ნებისმიერი?

ჩვენ ცხადია არ შევუდგებით იმის მტკიცებას, რომ ადამიანს შეუძლია იცოცხლოს 1000 წელი, თუმცა ჩვეულებრივი სადაზღვევო პრაქტიკა არ აწესებს სიცოცხლის ხანგრძლივობის ასეთ საზღვარსაც. იმ ფორმულების თანახმად, რომელზეც დაფუძნებულია თანამედროვე სიკვდილიანობის ცხრილები იმ ადამიანების წილი, რომლებმაც იცოცხლეს 1000 წლამდე არის 10 ხარისხად (-10^{36}) რიგის. ბიოლოგიის თვალსაზრისით ეს მტკიცებულება აზრს მოკლებულია, მაგრამ ის არ ეწინააღმდეგება ცდას: საუკუნის განმავლობაში იბადება არაუმეტეს 10^{10} ადამიანი და იმისათვის, რომ ზემოთ მოყვანილი მტკიცებულობა სტატისტიკურად უარვყოთ საჭირო იქნებოდა 10 ხარისხად 10^{35} საუკუნე, რაც აჭარბებს დედამიწის ასაკს 10 ხარისხად 10^{34}-ჯერ. ცხადია, ასეთი მცირე შანსები თავსებადია შეუძლებელი შედეგის ჩვენს წარმოდგენასთან და შეიძლებოდა გვეფიქრა, რომ მისი გათვალისწინება აბსურდულია, თუმცა სინამდვილეში იგი ამარტივებს

ბევრ ფორმულებს. გარდა ამისა, თუ ჩვენ გამოვრიცხავდით, რომ შეუძლებელია 1000 წლამდე ცხოვრება, მაშინ წავაწყდებოდით უფრო დიდ სირთულეებს, ვინაიდან მაშინ ჩვენ უნდა დაგვეშვა რომ არსებობს მაქსიმალური ასაკი. მაგრამ დაშვება, რომ ადამიანმა შეიძლება იცოცხლოს x წლამდე, მაგრამ არ შეიძლება იცოცხლოს x წელი და 2 წამი, არაფრით არაა უკეთესი, ვიდრე ასაკის ზედა ზღვარის არარსებობა.

ნებისმიერი თეორია აუცილებლად გულისხმობს ზოგიერთ გამარტივებას. ჩვენი პირველი გამარტივება ეხება "ცდის" ან "დაკვირვების" შესაძლო შედეგებს. მათემატიკური თეორიის შესასწავლი ობიექტები შეიძლება იყვნენ მხოლოდ ეს შესაძლებელი შედეგები. თუ ჩვენ გვინდა ავაგოთ ცდის აბსტრაქტული მოდელი, ჩვენ თავიდან უნდა დავადგინოთ რას წარმოადგენს გამარტივებული (იდეალიზირებული) ცდის შესაძლო შედეგი. ტერმინოლოგიის ერთიანობისათვის ექსპერიმენტის (ცდის) ან დაკვირვების შედეგებს უწოდებენ *ხდომილებებს.*

განვიხილოთ ექსპერიმენტი, რომლის ყველა შესაძლო შედეგები ამოიწურება N სხვადასხვა მნიშვნელობით $\omega_1, \omega_2, ..., \omega_N$. ეს მნიშვნელობები არ არის აუცილებლად რიცხვითი და მათი ფიზიკური ბუნება არ არის არსებითი.

განმარტება 1. ექსპერიმენტის ცალკეულ შესაძლო შედეგებს *ელემენტარული ხდომილებები* ეწოდება, ხოლო მათ ერთობლიობას – *ელემენტარულ ხდომილებათა სივრცე* და აღინიშნება Ω ასოთი: $\Omega = \{\omega_1, \omega_2, ..., \omega_N\}$.

მოვიყვანოთ მაგალითები:

I. მონეტის ერთხელ აგდებისას -- Ω={გ, ს};

II. მონეტის ორჯერ აგდებისას, ან ორი მონეტის ერთდროულად აგდებისას -- Ω={გგ, გს, სგ, სს};

III. მონეტის სამჯერ აგდებისას, ან სამი მონეტის ერთდროულად აგდებისას -- Ω= {გგგ, გგს, გსგ, სგგ, გსს, სგს, სსგ, სსს};

IV. მონეტის n-ჯერ აგდებისას $\Omega = \{\omega : \omega = (a_1, ..., a_n), a_i =$ გ ან ს$\}$ და შედეგების საერთო რაოდენობა ტოლია 2^n-ის;

V. ერთი სათამაშო კამათლის გაგორებისას -- Ω={1, 2, 3, 4, 5, 6};

VI. ვთქვათ, თავიდან ვაგდებთ მონეტას. თუ მოვა გერბი, მაშინ ვაგორებთ სათამაშო კამათელს; ხოლო თუ მოვა საფასური, მაშინ კიდევ ერთხელ ვაგდებთ მონეტას. ამ შემთხვევაში $\Omega = \{$გ1, გ2, გ3, გ4, გ5, გ6, სგ, სს$\}$;

VII. ორი სათამაშო კამათლის გაგორებისას -- Ω={(1,1); (1,2); . . . ; (1,6); (2,1); (2,2); . . . ; (2,6); . . . ; (6,1); . . . ; (6,6)} ანუ

$$\Omega = \{(i, j) : i, j = 1, 2, ..., 6\};$$

VIII. პროდუქციის ვარგისიანობის დადგენისას -- Ω={"ვარგისი", "უვარგისი"};

IX. სატელეფონო სადგურში გამოძახებათა რაოდენობა -- $\Omega = \{0, 1, 2, \ldots\}$;

X. ძაბვა ქსელში -- $\Omega = \{[0, 220]\}$.

განმარტება 2. ელემენტარულ ხდომილებათა სივრცის ნებისმიერ ქვესიმრავლეს *ხდომილება* ეწოდება.

ცხადია, ელემენტარული ხდომილებები აგრეთვე ხდომილებებია, ისინი წარმოადგენენ ელემენტარულ ხდომილებათა სივრცის ერთელემენტიან ქვესიმრავლეებს. ყველა დანარჩენ ქვესიმრავლეს (მათ შორის ცარიელი სიმრავლისა და თვითონ სივრცის ჩათვლით) ხდომილებას უწოდებენ. ზოგჯერ (იმის აღსანიშნავად, რომ ქვესიმრავლეში ერთზე მეტი ელემენტია) ხმარობენ აგრეთვე *შედგენილი ან რთული* ხდომილების ცნებასაც. ჩვენ ვისარგებლებთ უბრალოდ ხდომილების ცნებით.

მონეტის ორჯერ აგდებისას (იხ. მაგალითი II) ხდომილების მაგალითებია: ა). ერთჯერ მაინც მოვიდა გერბი (ანუ მოვიდა ერთი ან მეტი, მაშასადამე, ორი, გერბი). იგი წარმოადგენს სიმრავლეს -- {გს,სგ,გგ}; ბ). გერბი მოვიდა არაუმეტეს ერთისა (ანუ მოვიდა ერთი ან ნაკლები, მაშასადამე, ნული – არცერთი, გერბი). იგი წარმოადგენს სიმრავლეს -- {გს,სგ,სს}; გ). გერბი მოვიდა ზუსტად ერთჯერ (ანუ პირველად მოვიდა გერბი და მეორედ კი საფასური ან პირიქით). ეს არის შემდეგი სიმრავლე -- {გს,სგ}; და ა. შ. აღსანიშნავია, რომ ამ შემთხვევაში სულ გვექნება 2^4=16 ხდომილება (როგორც ცნობილია n ელემენტიანი სიმრავლის ყველა შესაძლო ქვესიმრავლეთა რაოდენობაა 2^n ქვესიმრავლე).

განმარტება 3. თუ ექსპერიმენტის კონკრეტული შედეგი ეკუთვნის რაიმე ხდომილებას, მაშინ ამბობენ რომ ეს *ხდომილება მოხდა,* ხოლო რომელსაც არ ეკუთვნის – ის *ხდომილება არ მოხდა.*

ალბათობის თეორიაში ხდომილებები აღინიშნება დიდი ლათინური ასოებით: $A, B, C, D, \ldots$. ხდომილებას $A = \Omega$ უწოდებენ *აუცილებელ ხდომილებას,* ვინაიდან ის აუცილებლად ხდება (ის შეუძლებელია არ მოხდეს, რადგან ექსპერიმენტის ყველა შედეგი მას ეკუთვნის); ხოლო ხდომილებას, რომელიც არ შეიცავს არც ერთ ელემენტარულ ხდომილებას აღნიშნავენ Ø სიმბოლოთი და უწოდებენ *შეუძლებელ ხდომილებას,* ვინაიდან მისი მოხდენა შეუძლებელია (რადგან ექსპერიმენტის არც ერთი შედეგი მას არ ეკუთვნის).

შემოვიღოთ აღნიშვნები: A={მონეტის ორჯერ აგდებისას ერთჯერ მაინც მოვიდა გერბი}={გს,სგ,გგ}; B={ მონეტის ორჯერ აგდებისას გერბი მოვიდა არაუმეტეს ერთისა}={გს,სგ,სს}; C={მონეტის ორჯერ აგდებისას გერბი მოვიდა ზუსტად ერთჯერ}={ გს,სგ }: D={მონეტის ორჯერ აგდებისას ორივეჯერ მოვიდა საფასური}={სს}; E={მონეტის ორჯერ აგდებისას საფასური მოვიდა არაუმეტეს ერთისა}={გგ,სგ,გს}. ამ აღნიშვნებში, თუ მონეტის ორჯერ აგდებისას საფასური მოვიდა მხოლოდ მეორედ აგდებისას,

მაშინ შეგვიძლია ვთქვათ, რომ მოხდა A, B, C და E ხდომილებები, ხოლო D ხდომილება კი არ მოხდა.

თუ A ხდომილების მოხდენას მოსდევს B ხდომილების მოხდენა (სიმრავლეთა თეორიის ენაზე ეს ნიშნავს, რომ A ხდომილება ნაწილია, ქვესიმრავლეა B ხდომილების), მაშინ ჩვენ დავწერთ, რომ $A \subset B$ და ვიტყვით, რომ A ხდომილება *იწვევს* B ხდომილებას. გასაგებია, რომ ნებისმიერი A ხდომილება იწვევს აუცილებელ ხდომილებას -- $A \subset \Omega$. თუ A ხდომილება იწვევს B ხდომილებას და იმავდროულად B ხდომილება იწვევს A ხდომილებას, მაშინ ვიტყვით, რომ A და B ხდომილები ერთმანეთის *ტოლია* და დავწერთ $A = B$.

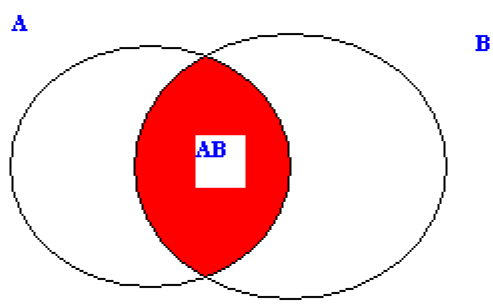
წინა აბზაცის აღნიშვნებში: C ხდომილება იწვევს A, B და E ხდომილებებს; D ხდომილება იწვევს B ხდომილებას; A და E ხდომილებები ერთმანეთის ტოლია.

§3. ოპერაციები ხდომილებებზე

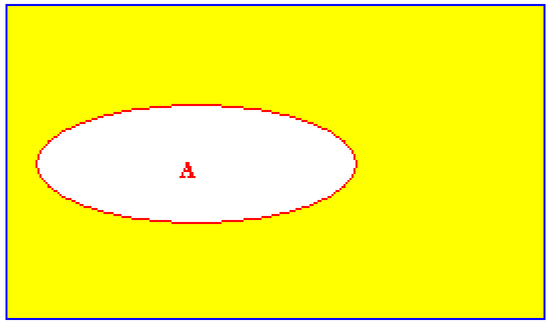
ხდომილებათა მოცემული სისტემის საშუალებით შესაძლებელია ახალი ხდომილებების აგება, ისევე როგორც სიმრავლეთა მოცემული სისტემის საშუალებით იგება ახალი სიმრავლეები მათი გაერთიანებებით, თანაკვეთებითა და დამატებებით.

ორი A და B ხდომილების *გაერთიანება (ან ჯამი)* ეწოდება ისეთ ხდომილებას, რომელიც ხდება მაშინ, როცა ამ ხდომილებებიდან ერთი მაინც ხდება და აღინიშნება სიმბოლოთი $A\cup B$ (ან $A+B$). სქემატურად ეს ასე გამოისახება:

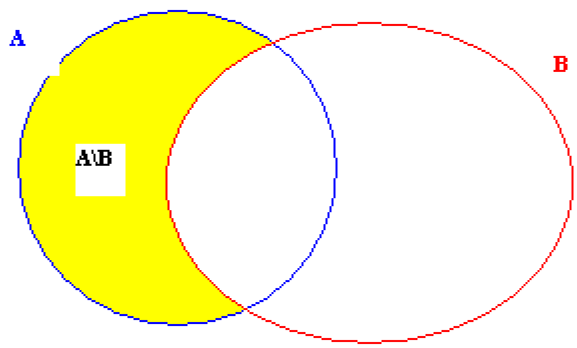
ორი A და B ხდომილების თანაკვეთა *(ან ნამრავლი)* ეწოდება ისეთ ხდომილებას, რომელიც ხდება მაშინ, როცა ეს ხდომილებები ერთდროულად ხდება და აღინიშნება სიმბოლოთი $A\cap B$ (ან AB). სქემატურად ეს ასე გამოისახება:



A ხდომილების *საწინააღმდეგო* ხდომილება ეწოდება ისეთ ხდომილებას, რომელიც ხდება მაშინ, როცა A არ ხდება და აღინიშნება სიმბოლოთი $\overline{A}$. სქემატურად, თუ წარმოვიდგენთ, რომ ელემენტარულ ხდომილებათა სივრცე მართკუთხედია, ხოლო A ხდომილება -- წრე, მაშინ საწინააღმდეგო ხდომილება იქნება ოთხკუთხედის გაფერადებული ნაწილი:

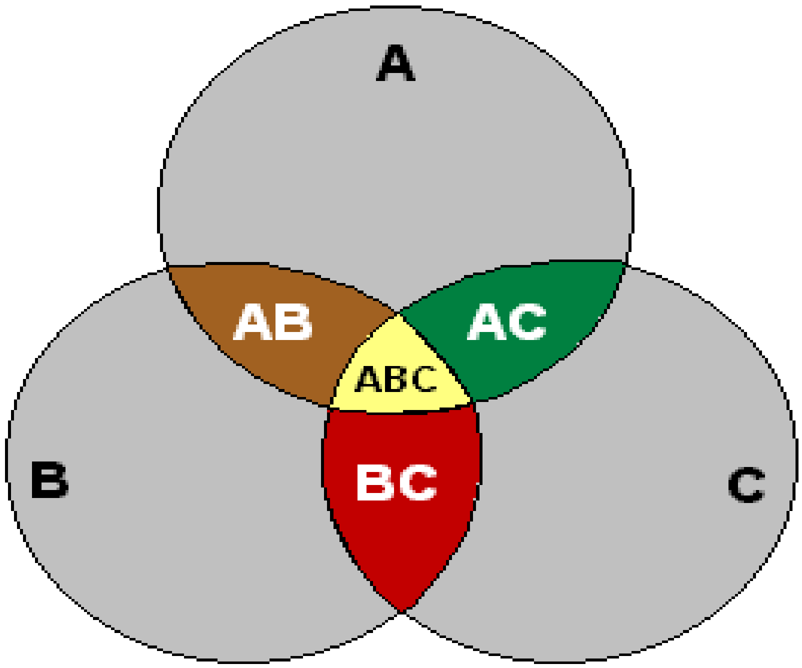


ორი A და B ხდომილების *სხვაობა* ეწოდება ისეთ ხდომილებას, რომელიც ხდება მაშინ, როცა ხდება A მაგრამ არ ხდება B და აღინიშნება სიმბოლოთი $A \setminus B$. სქემატურად ეს ასე გამოისახება:



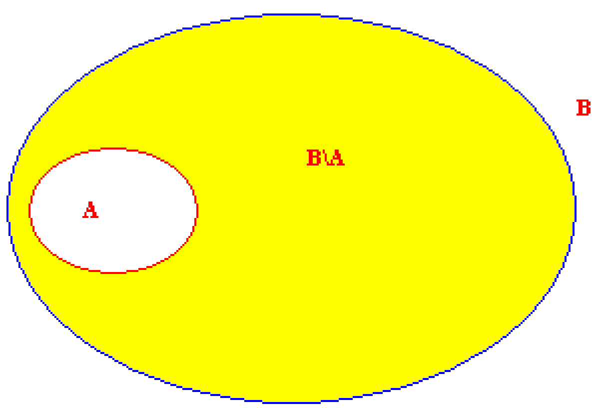
ცხადია, რომ ორი A და B ხდომილების სხვაობა აგრეთვე შეიძლება წარმოდგეს, როგორც A ხდომილებისა და B ხდომილების საწინააღმდეგო $\overline{B}$ ხდომილების თანაკვეთა: $A \setminus B = A \cap \overline{B}$.

გასაგებია, რომ მას შემდეგ რაც ჩვენ განვმარტეთ ორი ხდომილების გაერთიანება და თანაკვეთა, ბუნებრივად შესაძლებელია ხდომილებათა ნებისმიერი რაოდენობის გაერთიანებისა და თანაკვეთის განმარტება. ასე მაგალითად, სქემატურად სამი A, B და C ხდომილებისათვის თანაკვეთა ABC იქნება:



ორ A და B ხდომილებას ეწოდება *არათავსებადი (უთავსებადი, შეუთავსებელი),* თუ მათი ერთდროულად მოხდენა შეუძლებელია, ან რაც იგივეა მათი თანაკვეთა არის შეუძლებელი ხდომილება: $A\cap B=\varnothing$. სიმრავლეთა თეორიის ენაზე ეს ნიშნავს, რომ ეს ორი სიმრავლე თანაუკვეთია.

ცხადია, რომ რაიმე A ხდომილება და მისი საწინააღმდეგო $\overline{A}$ ხდომილება უთავსებადია -- $A\cap\overline{A}=\varnothing$. გარდა ამისა, $A\cup\overline{A}=\Omega$. თუ A ხდომილება იწვევს B ხდომილებას ($A\subset B$), მაშინ ცხადია, რომ $A\cap B=A$, $A\cup B=B$, $A\setminus B=\varnothing$, ხოლო სხვაობა $B\setminus A$ სქემატურად ასე გამ-ოისახება:



იმისათვის, რომ დავინახოთ რა განსხვავებაა და რა აქვთ საერთო სიმრავლეთა თეორიისა და ალბათობის თეორიის ტრადიციულ ტერმინებს, ქვემოთ ჩვენ მოვიყვანთ შესაბამის ცხრილს:

აღნიშვნები	სიმრავლეთა თეორიის ინტერპრეტაცია	ალბათობის თეორიის ინტერპრეტაცია
ω	ელემენტი, წერტილი	შედეგი, ელემენტარული ხდომილება
Ω	წერტილთა სიმრავლე	ელემენტარულ ხდომილებათა სივრცე, აუცილებელი ხდომილება
A	წერტილთა სიმრავლე	ხდომილება (თუ შედეგი $\omega \in A$, მაშინ ამბობენ, რომ მოხდა A ხდომილება)
$\overline{A} = \Omega \setminus A$	A სიმრავლის დამატება, ე.ი. იმ წერტილების სიმრავლე, რომლ-ებიც არ შედიან A-ში	ხდომილება, რომელიც მდგომარეობს A-ს არ მოხდენაში
$A \cup B$	A და B სიმრავლეების გაერ-თიანება, ე.ი. სიმრავლე იმ წერტილების, რომლებიც შედიან ან A-ში ან B-ში	ხდომილება, რომელიც მდგომარეობს A და B ხდომილებებიდან ერთის მაინც მოხდენაში
$A \cap B$	A და B სიმრავლეების თანაკ-ვეთა, ე.ი. სიმრავლე იმ წერტი-ლების, რომლებიც შედიან რო-გორც A, ისე B სიმრავლეში	ხდომილება, რომელიც მდგომარეობს A და B ხდომილებების ერთდროულ მოხდენაში
$\varnothing$	ცარიელი სიმრავლე	შეუძლებელი ხდომილება
$A \cap B = \varnothing$	A და B სიმრავლეები არ იკვეთებიან	A და B ხდომილებები არათავ-სებადია (მათი ერთდროულად მოხდენა შეუძლებელია)
$A + B$	სიმრავლეთა ჯამი, ე.ი. თანაუკვეთი სიმრავლეების გაერთიანება	ხდომილება, რომელიც მდგომარეობს ორი უთავსებადი ხდომილებიდან ერთის მოხდენაში
$A \setminus B$	A და B სიმრავლეების სხვა-ობა, ე. ი. სიმრავლე იმ წერტ-ილების, რომლებიც შედიან A-ში, მაგრამ არ შედიან B-ში	ხდომილება, რომელიც მდგომარეობს A ხდომილების მოხდენაში და B ხდომილების არ მოხდენაში
$A \Delta B$	სიმრავლეების სიმეტრიული სხვაობა, ე.ი. სიმრავლე $(A \setminus B) \cup (B \setminus A)$	ხდომილება, რომელიც მდგომარეობს A და B ხდომილებებიდან ერთის მოხდენაში, მაგრამ არა ორივეს ერთდროულად
$\bigcup_{n=1}^{\infty} A_n$	$A_1, A_2, \ldots$ სიმრავლეების გაერ-თიანება, ე.ი. იმ წერტილების სიმრავლე, რომლებიც შედიან ერთერთში მაინც	ხდომილება, რომელიც მდგომარეობს $A_1, A_2, \ldots$ ხდომილებებიდან ერთერთის მაინც მოხდენაში
$\sum_{n=1}^{\infty} A_n$	$A_1, A_2, \ldots$ სიმრავლეების ჯამი, ე.ი. გაერთიანება წყვილ-წყვილ-ად თანაუკვეთი სიმრავლეების	ხდომილება, რომელიც მდგომარეობს $A_1, A_2, \ldots$ უთავსებადი ხდომილებებ-იდან ერთის მოხდენაში
$\bigcap_{n=1}^{\infty} A_n$	$A_1, A_2, \ldots$ სიმრავლეების თანაკ-ვეთა, ე.ი. სიმრავლე იმ წერტი-ლების, რომლებიც შედიან ყვე-ლა სიმრავლეში	ხდომილება, რომელიც მდგომარეობს $A_1, A_2, \ldots$ ხდომილებების ერთდრო-ულ მოხდენაში

§4. ალბათობის განმარტება

განმარტება 1. თუ ელემენტარული ხდომილებათა სივრცის ნებისმიერ ელემენტარულ ხდომილებას ω_i შეესაბამება გარკვეული რიცხვები $p_i = P(\omega_i)$, რომლებიც აკმაყოფილებენ პირობებს: $0 \le p_i \le 1$ და $\sum_{i=1}^{N} p_i = 1$, მაშინ ამ რიცხვებს ეწოდებათ ω_i ელემენტარული ხდომილებების *ალბათობები* (P -- არის პირველი ასო ინგლისური სიტყვის "Probability", რომელიც ნიშნავს ალბათობას).

მონეტის ერთჯერ აგდებისას (Ω={გ, ს}) ცხადია უნდა დავუშვათ, რომ: $P\{\text{გ}\} = p$, $P\{\text{ს}\} = 1 - p$, სადაც $0 \le p \le 1$. შემთხვევას, როცა $p = 1/2$ ეწოდება *სიმეტრიული მონეტის აგდება.* მონეტის ორჯერ აგდებისას (Ω ={გგ, გს, სგ, სს}) -- $P\{\text{გგ}\} = p_1$, $P\{\text{გს}\} = p_2$, $P\{\text{სგ}\} = p_3$, $P\{\text{სს}\} = p_4$, სად-აც ყველა $0 \le p_i \le 1$ და $\sum_{i=1}^{4} p_i = 1$. შემთხვევას, როცა ყველა $p_i = 1/4$, ჩვენ ვუწოდებთ *სიმეტრიული მონეტის ორ დამოუკიდებელ აგდებას.*

განმარტება 2. A ხდომილების ალბათობა $P(A)$ ეწოდება მასში შემავალი ელემენტარული ხდომილებების ალბათობების ჯამს:

$$P(A) = \sum_{\omega_i \in A} P(\omega_i).$$

აქ და ყველგან შემდგომში სიმბოლოთი $\omega \in A$ აღინიშნება ის ფაქტი, რომ ω ელემენტარული ხდომილება *ეკუთვნის* A ხდომილებას.

ცხადია, რომ $P(\Omega) = 1$ და $P(\varnothing)=0$. გარდა ამისა, თუ $A \subset B$, მაშინ $P(A) \le P(B)$.

თუ ყველა ელემენტარული ხდომილება ტოლალბათურია ანუ ყველა $P(\omega_i) = 1/N$, $i = 1,2,...,N$, მაშინ ამბობენ, რომ გვაქვს *კლასიკური ალბათური მოდელი* (პრაქტიკულ სიტუაციაში ეს ნიშნავს, რომ შემთხვევითი ექსპერიმენტის ყველა შედეგი ერთნაირად მოსალოდნელია და ყველას განხორციელების ერთი და იგივე შანსი აქვს). ამ შემთხვევაში განმარტება 2 გვაძლევს, რომ კლასიკურ მოდელში ხდომილების ალბათობა დაემთხვევა *ალბათობის კლასიკურ განმარტებას,* რომელიც შემოღებული იყო 1812 წელს *ლაპლასის* მიერ:

ალბათობის კლასიკურ განმარტება. თუ შემთხვევითი ექსპერიმენტის შესაძლო შედეგთა რაოდენობა სასრულია და ცალკეულ შედეგს აქვს განხორციელების თანაბარი შანსი, მაშინ ხდომილების ალბათობა ტოლია მასში შემავალ ელემენტარულ ხდომილებათა რაოდენობა გაყოფილი ყველა შესაძლო ელემენტარულ ხდომილებათა რაოდენობაზე:

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

სადაც $|B|$ -- აღნიშნავს B ხდომილებაში შემავალ ელემენტარულ ხდომილებათა რაოდენობას.

მოცემულ ხდომილებაში შემავალ ელემენტარულ ხდომილებებს უწოდებენ *ხელშემწყობ* ელემენტარულ ხდომილებებს. შესაბამისად, *ალბათობის კლასიკური განმარტება* ასე გამოითქმის: კლასიკურ მოდელში ხდომილების ალბათობა ტოლია ხელშემწყობ ელემენტარულ ხდომილებათა რაოდენობა გაყოფილი ყველა შესაძლო ელემენტარულ ხდომილებათა რაოდენობაზე.

ამოცანა 1. სიმეტრიული მონეტის ორი დამოუკიდებელი აგდებისას ვიპოვოთ ალბათობა იმისა, რომ გერბი მოვა ერთჯერ მაინც.

ამოხსნა. ამ შემთხვევაში Ω={გგ, გს, სგ, სს}; $P\{\text{გგ}\}=P\{\text{გს}\}= =P\{\text{სგ}\}=P\{\text{სს}\}=1/4$; A={გგ, გს, სგ} და $P(A)=3/4$.

ამოცანა 2 ("ბედნიერ" ბილეთებზე). 25 საგამოცდო ბილეთიდან 5 "ბედნიერია", ხოლო დანარჩენი 20 – "არა ბედნიერი". რომელ სტუდენტს აქვს "ბედნიერი" ბილეთის აღების მეტი ალბათობა: ვინც პირველი იღებს ბილეთს, თუ ვინც მეორე იღებს ბილეთს?

ამოხსნა. ავღნიშნოთ პირველი სტუდენტის მიერ აღებული ბილეთის ნომერი i ასოთი, ხოლო მეორე სტუდენტის მიერ აღებული ბილეთის ნომერი j ასოთი. დავუშვათ, რომ "ბედნიერი" ბილეთების ნომრებია: 1, 2, 3, 4, 5. მაშინ ცხადია, რომ $\Omega=\{(i,j): i,j=1,2,...,25, i\neq j\}$, $|\Omega|=25\cdot 24=600$ და ბუნებრივია ჩავთვალოთ, რომ ყველა ელემენტარული ხდომილება ტოლალბათურია: $P(i,j)=1/600$.

შემოვიღოთ ხდომილებები:

A={პირველმა სტუდენტმა აიღო კარგი ბილეთი},

B={მეორე სტუდენტმა აიღო კარგი ბილეთი},

მაშინ ამ ხდომილებებს აქვთ შემდეგი სახე:

$A=\{(i.j): i=1,...,5;\ j=1,...,25; i\neq j\}$ და $B=\{(i,j): i=1,...,25;\ j=1,...,5; i\neq j\}$.

ადვილი დასანახია, რომ $|A|=|B|=5\cdot 24=120$. ამიტომ ალბათობის კლასიკური განმარტების თანახმად:

$$P(A)=P(B)=\frac{120}{600}=\frac{1}{5}.$$

ე. ი. ორივე სტუდენტს აქვს კარგი ბილეთის აღების ერთი და იგივე ალბათობა.

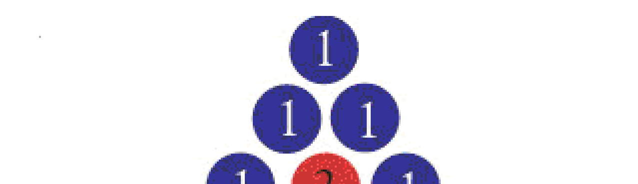
დავალება. წინა ამოცანაში ვიპოვოთ ალბათობა იმისა, რომ მესამე სტუდენტი ამოიღებს "ბედნიერ" ბილეთს.

ამოცანა 3 (გალტონის დაფა). გვაქვს დაფაზე სამკუთხედის ფორმით დალაგებული რგოლები, ისე რომ წვეროში ერთი რგოლია, მეორე რიგში წინასგან თანაბარ მანძილებზე ორი რგოლი, მესამე რიგში ზედა ორი რგოლიდან თანაბარ მანძილებზე სამი რგოლი და ა.შ. ბოლოში არის ექვსი რგოლი. მე-7 სტრიქონში კი არის ბოლო 6 რგოლიდან თანაბარ მანძილებზე 7 ღრმული. ზედა რგოლზე აგდებენ ბურთს და მას შეუძლია იგორაოს თანაბარი ალბათობით ან მარჯვნივ ან მარცხნივ რგოლიდან რგოლზე, რაც საბოლოოდ სრულდება რომელიმე ღრმულში ჩავარდნით. როგორია ალბათობა იმისა, რომ ბურთი ჩავარდება მეხუთე ღრმულში?

•
O
O O
O O O
O O O O
O O O O O
O O O O O O
U U U U U U U
1 2 3 4 5 6 7

ამოხსნა. როგორც ვხედავთ არსებობს პირველ და მე-7 ღრმულებში ბურთის ჩავარდნის ერთადერთი გზა (ტრაექტორია), მეორე და მე-6 ღრმულებში ბურთის ჩავარდნის -- ექვს-ექვსი გზა, მესამე და მეხუთე ღრმულებში – თხუთმეტ-თხუთმეტი გზა და ბოლოს, მეოთხე ღრმულში – ბურთის ჩავარდნის 20 გზა. გზების (შედეგების) სრული რაოდენობაა 1+6+15+20+15+6+1=64 და ყველა ეს შედეგი თანაბრად მოსალოდნელია, ვინაიდან თითოეული ტრაექტორიის გავლისას ბურთი განიცდის ექვს დაჯახებას რგოლებზე და ყოველი დაჯახებისას ის თანაბარი ალბათობებით გადაადგილდება ან მარჯვნივ ან მარცხნივ. თანაბარალბათური 64 შედეგიდან მეხუთე ღრმულში ჩავარდნას ხელს უწყობს 15 შედეგი და შესაბამისად, საძებნი ალბათობა იქნება 15/64.

ქვემოთ მოყვანილია გალტონის დაფაზე რგოლების 7 რიგის შემთხვევაში თითოეულ პოზიციაზე ბურთის მოხვედრის შესაძლო გზების რიცხვი.



ალბათობის კლასიკური განმარტება გამოსადეგია ამოცანათა მხოლოდ ძალიან ვიწრო კლასისათვის, სადაც ყველა შესაძლო შედეგთა სიმრავლე სასრულია და თითოეული შესაძლო შედეგი თანაბრად შესაძლებე-

ლია. უმრავლეს რეალურ ამოცანაში ეს პირობა ირღვევა, და შესაბამისად, კლასიკური განმარტების გამოყენება შეუძლებელია. ასეთ შემთხვევაში საჭიროა ხდომილების ალბათობის სხვა გზით განმარტება. ამ მიზნით, წინასწარ შემოვიღოთ A ხდომილების **ფარდობითი სიხშირის W(A)** ცნება, როგორც ცდათა იმ რიცხვის შეფარდება, რომელშიც დაფიქსირდა (განხორციელდა) A ხდომილება, ჩატარებული ექსპერიმენტების საერთო რაოდენობასთან:

$$W(A) = \frac{M}{N},$$

სადაც N – ცდათა საერთო რიცხვია, M -- A ხდომილების მოხდენათა რიცხვი.

ექსპერიმენტების დიდმა რაოდენობამ აჩვენა, რომ თუ ცდები ტარდება ერთი და იგივე პირობებში, მაშინ ცდათა (დაკვირვებათა) დიდი რიცხვისათვის, ფარდობითი სიხშირე უმნიშვნელოდ იცვლება, მერყეობს (ირხევა) რა გარკვეული მუდმივი რიცხვის ირგვლივ. ეს რიცხვი შეიძლება ჩაითვალოს განსახილეველი ხდომილების ალბათობად.

განმარტება. ხდომილების **სტატისტიკურ ალაბათობად** ითვლება ამ ხდომილების ფარდობითი სიხშირე ან მასთან ახლოს მყოფი რიცხვი (მათემატიკურად ზუსტი ფორმულირება ასეთია: $P(A) = \lim_{N \to \infty} W_N(A)$).

შენიშვნა. იმისათვის, რომ არსებობდეს A ხდომილების სტატისტიკური ალბათობა, საჭიროა:

ა). ექსპერიმენტების უსასრულოდ დიდი რიცხვის ჩატარების შესაძლებლობა;

ბ). სხვადასხვა სერიებში A ხდომილების მოხდენის ფარდობითი სიხშირეების მდგრადობა საკმარისად დიდი რაოდენობა ცდებისათვის.

§5. გეომეტრიული ალბათობა

ალბათობის კლასიკური განმარტების ერთ-ერთი ნაკლი ისაა, რომ მისი გამოყენება არ შეიძლება იმ ექსპერიმენტებისათვის, რომლებსაც გააჩნიათ შედეგების უსასრულო რაოდენობა. ასეთ შემთხვევაში შესაძლებელია ვისარგებლოთ **გეომეტრიული ალბათობის ცნებით**.

დავუშვათ რომ L მონაკვეთზე შემთხვევით აგდებენ წერტილს. რაც იმას ნიშნავს, რომ წერტილი აუცილებლად დაეცემა L მონაკვეთზე და ამასთანავე, თანაბარი შესაძლებლობებით ის შესაძლებელია დაემთხვეს ამ მონაკვეთის ნებისმიერ წერტილს. ამავე დროს, ალბათობა იმისა, რომ წერტილი დაეცემა L მონაკვეთის ნებისმიერ ნაწილზე არაა დამოკიდებული ამ ნაწილის განლაგებაზე მონაკვეთის შიგნით და პროპორციულია მისი სიგრძის. მაშინ ალბათობა იმისა, რომ აგდებული წერტილი დაეცემა l მონაკვეთზე, რომელიც არის ნაწილი L მონაკვეთის, გამოითვლება ფორმულით:

$$P = |l| / |L|, \qquad (1)$$

სადაც $|l|$ -- l მონაკვეთის სიგრძეა, ხოლო $|L|$ -- L მონაკვეთის სიგრძე.

ანალოგიურად, შესაძლებელია ამოცანის დასმა წერტილისათვის, რომელსაც ვაგდებთ ბრტყელ S არეზე და ალბათობა იმისა, რომ ის მოხვდება ამ არის s ნაწილში განიმარტება როგორც:

$$P = |s| / |S|, \qquad (2)$$

სადაც $|s|$ -- s არის ფართობია, ხოლო $|S|$ -- S არის ფართობი.

სამგანზომილებიან შემთხვევაში ალბათობა იმისა, რომ შემთხვევითი გზით V სხეულში ჩავარდნილი წერტილი აღმოჩნდება ამ სხეულის v ნაწილში, გამოითვლება ფორმულით:

$$P = |v| / |V|, \qquad (3)$$

სადაც $|v|$ -- v სხეულის მოცულობაა, ხოლო $|V|$ -- V სხეულის მოცულ-ობა.

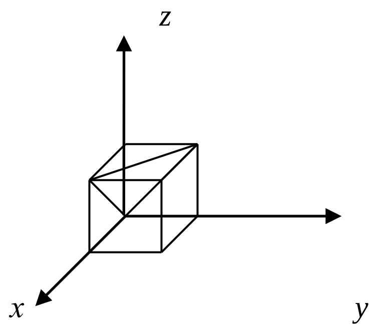
ამოცანა 1. ვიპოვოთ ალბათობა იმისა, რომ წრეში შემთხვევით ჩაგდებული წერტილი არ ჩავარდება ამ წრეში ჩახაზულ წესიერ ექვსკუთხედში.

ამოხსნა. დავუშვათ, რომ წრის რადიუსია R, მაშინ მასში ჩახაზული წესიერი ექვსკუთხედის გვერდიც იქნება R. ამასთანავე, წრის ფართობია $|S| = \pi R^2$, ხოლო ექვსკუთხედის ფართობია -- $|s| = \frac{3\sqrt{3}}{2}R^2$. ამიტომ საძიებელი ალბათობა იქნება:

$$P = \frac{|S|-|s|}{|S|} = \frac{\pi R^2 - \frac{3\sqrt{3}}{2}R^2}{\pi R^2} = \frac{2\pi - 3\sqrt{3}}{2\pi} \approx 0.174.$$

ამოცანა 2. AB მონაკვეთზე შემთხვევით აგდებენ სამ წერტილს C, D და M. ვიპოვოთ ალბათობა იმისა, რომ AC, AD და AM მონაკვეთებისაგან შეიძლება აიგოს სამკუთხედი?

ამოხსნა. ავღნიშნოთ AC, AD და AM მონაკვეთების სიგრძეები შესაბამისად x, y და z-ით და ელემენტარულ ხდომილებათა სივრცის როლში განვიხილოთ სამგანზომილებიანი სივრცის წერტილთა სიმრავლე კოორდინატებით (x, y, z). თუ ჩავთვლით, რომ AB მონაკვეთის სიგრძე ტოლია 1-ის, მაშინ ელემენტარულ ხდომილებათა სივრცე იქნება კუბი, რომლის წიბოა ერთი. ამავე დროს, ხელშემწყობ ელემენტარულ ხდომილებათა სიმრავლე (სამკუთხედის აქსიომის თანახმად)შედგება იმ წერტილებისაგან, რომელთა კოორდინატებისათვის სრულდება სამკუთხედის უტოლობები: : $x + +y > z, x + z > y, y + z > x$. ეს კი წარმოადგენს კუბის ნაწილს, რომელიც მოჭრილია მისგან სიბრტყეებით: $x + y = z, x + z = y, y + z = x$



(ერთ-ერთი ამ სიბრტყიდან, კერძოდ $x + y = z$, მოყვანილია ნახაზზე). ყოველი ასეთი სიბრტყე კუბიდან მოჭრის პირამიდას, რომლის მოცულობა ტოლია

$$\frac{1}{3} \cdot \frac{1}{2} \cdot 1 = \frac{1}{6}.$$

შესაბამისად, კუბის დარჩენილი ნაწილის მოცულობა იქნება

$$|v| = 1 - 3 \cdot \frac{1}{6} = \frac{1}{2}.$$

ამიტომ საძიებელი ალბათობა, განმარტების თანახმად, იქნება

$$|P| = \frac{|v|}{|V|} = \frac{1}{2} : 1 = \frac{1}{2}.$$

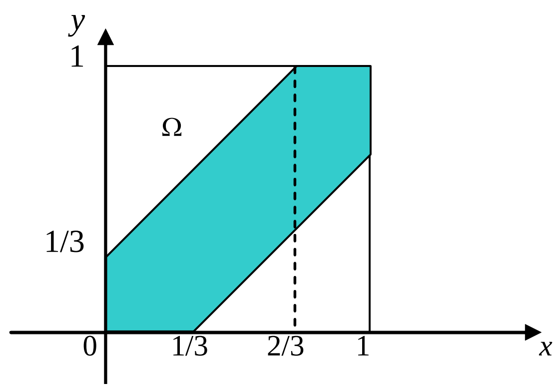
ამოცანა 3 (შეხვედრის ამოცანა). ორი პირი შეთანხმდა გარკვეულ ადგილზე შეხვდეს ერთმანეთს 6-დან 7 საათამდე. თითოეული მათგანი შემთხვევით მომენტში მიდის დათქმულ ადგილას და მეორეს ელოდება 20 წუთის მანძილზე (შემდეგ კი მიდის). ვიპოვოთ ალბათობა იმისა, რომ ეს პირები შეხვდებიან ერთმანეთს?

ამოხსნა. ავღნიშნოთ, ერთ-ერთი პირის დათქმულ ადგილზე მოსვლის დრო $6 + x$-ით, ხოლო მეორე პირის -- $6 + y$-ით (სადაც x და y გამოსახულია საათებში). ელემენტარულ ხდომილებათა სივრცის როლში შეგვიძლია ავიღოთ იმ (x, y) წერტილთა სიმრავლე, რომლებიც ეკუთვნიან ერთეულოვან კვადრატს. ამ შემთხვევაში ჩვენთვის საინტერესო წერტილთა (ხელშემწყობ ელემენტარულ ხდომილებათა) სიმრავლე იქნება ერთეულოვ-

ანი კვადრატის იმ წერტილთა სიმრავლე, რომელთა კოორდინატები ერთმანეთისაგან დაშორებულია არაუმეტეს $20/60=1/3$-ით:

$$A=\{(x,y);|x-y|\le 1/3\}\cap\Omega.$$

ვინაიდან, $|x-y|\le 1/3 \Leftrightarrow -1/3\le y-x\le 1/3 \Leftrightarrow x-1/3\le y\le x+1/3$. ამიტომ ადვილი გასაგებია, რომ A სიმრავლე იქნება ერთეულოვანი კვადრატის შიგნით $y=x-1/3$ და $y=x+1/3$ წრფეებს შორის მოქცეული გაფერადებული არე (როცა $0\le x\le 1/3$, მაშინ $y=x-1/3$ წრფის ნაცვლად ქვედა საზღვრის როლში იქნება x ღერძი: $0\le y\le x+1/3$, ხოლო როცა $2/3\le x\le 1$, მაშინ ზედა საზღვარი $y=x+1/3$ წრფის ნაცვლად იქნება $y=1$ წრფე: $x-1/3\le y\le 1$).



ამიტომ საძიებელი ალბათობა იქნება:

$$P(A)=\frac{|A|}{|\Omega|}=\frac{1-2/3\cdot 2/3}{1}=\frac{5}{9}.$$

დავალება 1. იპოვეთ ალბათობა იმისა, რომ ერთეულოვან გვერდიან კვადრატში შემთხვევით შერჩეული წერტილი ამ კვადრატის უახლოესი გვერდიდან დაშორებული იქნება არაუმეტეს 0,15-ით?

დავალება 2. ორი სიგნალი მიმღებ მოწყობილობაზე T დროის მანძილზე შემთხვევით მომენტში მიიღება. მოწყობილობა მათ განარჩევს, თუ ისინი t დროით მაინც არიან დაცილებული ერთმანეთს. იპოვეთ ალბათობა იმისა, რომ ორივე სიგნალი მიღებული იქნება?

§6. კომბინატორიკის ელემენტები

ძალიან ბევრ შემთხვევაში ელემენტარულ ხდომილებათა სივრცისა და ხდომილების ალბათობის განსაზღვრა პირდაპირი გადათვლით შეუძლებელია. ხდომილებათა ალბათობების გამოთვლის დროს ხშირად საჭირო ხდება კომბინატორიკის ფორმულების გამოყენება. **კომბინატორიკა** – არის მეცნიერება, რომელიც სწავლობს კომბინაციებს, რომელთა შედგენაც შესაძლებელია გარკვეული სასრული სიმრავლის ელემენტებიდან განსაზღრული წესების გამოყენებით. განვსაზღვროთ ძირითადი კომბინაციები.

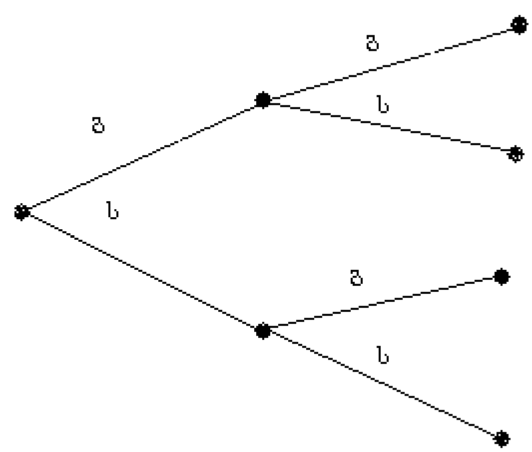
პირველ რიგში ჩამოვაყალიბოთ კომბინატორული ანალიზის ძირითადი პრინციპი, რომელსაც **გამრავლების პრინციპი** ეწოდება:

თუ ასარჩევია m ობიექტი და არსებობს პირველი ობიექტის არჩევის n_1 შესაძლებლობა, პირველი ობიექტის არჩევის შემდეგ არსებობს მეორე ობიექტის არჩევის n_2 შესაძლებლობა, და ა. შ., $m-1$ ობიექტის არჩევის შემდეგ არსებობს m-ური ობიექტის არჩევის n_m შესაძლებლობა, მაშინ არსებობს ამ m ობიექტის ამ მიმდევრობით არჩევის $n_1 \times n_2 \times \cdots \times n_m$ შესაძლებლობა.

მაგალითად, თუ მამაკაცს აქვს 4 პერანგი და 2 პიჯაკი, მაშინ ამ მამაკაცს აქვს პერანგისა და პიჯაკის შერჩევის $4 \times 2 = 8$ შესაძლებლობა (ვარიანტი).

ნამრავლის პრინციპთან დაკავშირებით ხშირად სასარგებლოა ხის მსგავსი (ხისებრი) დიაგრამის ანუ *დენდროგრამის* გამოყენება.

მაგალითად, დენდროგრამით გამოსახული მონეტის ორჯერ აგდების შესაბამისი შედეგების სიმრავლე იქნება:



განმარტება 1. გადანაცვლებები – ეს არის კომბინაციები, რომლებიც შედგენილია მოცემული n ელემენტიანი სიმრავლის ყველა n ელემენტისაგან და ერთამანეთისაგან განხსვავდება მხოლოდ ელემენტების განლაგების რიგით. n ელემენტიანი სიმრავლის ყველა შესაძლო გადანაცვლებათა რიცხვია:

$$P_n = n!.$$

მართლაც, თუ ამ ამოცანას შევხედავთ როგორც n ობიექტიდან n ობიექტის არჩევის ამოცანას, მაშინ არსებობს პირველი ობიექტის არჩევის n შესაძლებლობა, პირველი ობიექტის არჩევის შემდეგ არსებობს მეორე ობიექტის არჩევის $n-1$ შესაძლებლობა, და ა. შ., მე-n ობიექტის არჩევის ერთადერთი შესაძლებლობა. შესაბამისად, გამრავლების პრინციპის თანახმად:

$$P_n = n\cdot(n-1)\cdots 1 = n!$$

მაგალითი 1. რამდენი განსხვავებული სია შეიძლება შედგენილ იქნეს 7 სხვადასხვა გვარისაგან?

ამოხსნა. $P_7 = 7! = 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7 = 5040$.

განმარტება 2. წყობები – ეს არის m ელემენტიანი კომბინაციები n განხსვავებული ელემენტის მქონე სიმრავლიდან, რომლებიც ერთმანეთისაგან განსხვავდებიან ან ელემენტების შემადგენლობით ან ელემენტების განლაგების რიგით. სხვა სიტყვებით, წყობა არის n ელემენტიანი სიმრავლის m ელემენტიან, დალაგებულ ქვესიმრავლეთა რაოდენობა. ყველა შესაძლო წყობების რიცხვია:

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

ამ ამოცანას ჩვენ შეგვიძლია შევხედოთ, როგორც n ობიექტიდან m ობიექტის შერჩევის ამოცანას. რადგანაც, არსებობს პირველი ობიექტის არჩევის n შესაძლებლობა, პირველი ობიექტის არჩევის შემდეგ არსებობს მეორე ობიექტის არჩევის $n-1$ შესაძლებლობა, და ა. შ., მე-m ობიექტის არჩევის $n-m+1$ შესაძლებლობა, ამიტომ გამრავლების პრინციპის საფუძველზე გვაქვს:

$$A_n^m = n(n-1)(n-2)...(n-m+1) = n!/(n-m)!$$

მაგალითი 2. შეჯიბრებაში მონაწილე 10 სპორტსმენიდან პირველ სამ ადგილზე გასული სამი გამარჯვებული რამდენ სხვადასხვანაირად შეიძლება განთავსდეს დასაჯილდოებელ კვარცხლბეკზე?

ამოხსნა. $A_{10}^3 = 10\cdot 9\cdot 8 = 720$.

განმარტება 3. ჯუფდებები – ეს არის n ელემენტიანი სიმრავლის დაულაგებელი m ელემენტიანი ქვესიმრავლეები (ანუ ისეთი ერთობლიობები, რომლებიც ერთმანეთისაგან განსხვავდებიან მხოლოდ ელემენტების შემადგენლობით). ჯუფდებათა რიცხვი ტოლია:

$$C_n^m = \frac{n!}{m!(n-m)!}\,.$$

ვინაიდან, n ელემენტიანი სიმრავლის ყველა m ელემენტიანი, დალაგებული ქვესიმრავლის მისაღებად, ჩვენ შეგვიძლია ავიღოთ n ელემენტიანი სიმრავლის ყველა დაულაგებელი m ელემენტიანი ქვესიმრავლე (რომელთა რაოდენობაა C_n^m) და თითოეულში მოვახდინოთ ყველანაირი გადანაცვლებები (ამის შესაძლებლობაა $m!$), ამიტომ ადგილი აქვს თანაფარდობას:

$$C_n^m \cdot m! = A_n^m\,.$$

აქედან ცხადია, რომ

$$A_n^m = \frac{C_n^m}{m!} = \frac{n!}{m!(n-m)!}$$

მაგალითი 3. შესარჩევ შეჯიბრებაში მონაწილეობს 10 ადამიანი, რომელთაგან ფინალში გადის სამი. ფინალისტების რამდენი განსხვავებული სამეული შეიძლება გამოვლინდეს?

ამოხსნა. წინა მაგალითისაგან განსხვავებით, აქ ფინალისტების რიგს (დალაგებას) მნიშვნელობა არა აქვს. ამიტომ ვიყენებთ ჯუფდების ფორმულას. შესაბამისად, საძიებელი რიცხვია:

$$C_{10}^3 = \frac{10!}{3!\cdot 7!} = \frac{8\cdot 9\cdot 10}{6} = 120.$$

§7. ალბათობის გამოთვლა კომბინატორიკის გამოყენებით

განვიხილოთ მაგალითები, რომლებიც დაკავშირებულია ყუთიდან, რომელშიც M რაოდენობის ბურთია, სხვადასხვა გზებით n რაოდენობის ბურთის ამოღებასთან.

ამორჩევა დაბრუნებით. ამ შემთხვევაში ექსპერიმენტის ყოველ ნაბიჯზე ყუთიდან ამოღებული ბურთი უკან ბრუნდება. ვიგულისხმოთ, რომ ბურთები გადანომრილია რიცხვებით ერთიდან M-მდე. შესაბამისად, ელემენტარულ ხდომილებათა სივრცე შეიძლება წარმოდგენილ იქნეს იქნება როგორც რიცხვთა n-ეულები $a_1,...,a_n$, სადაც a_i აღნიშნავს i-ურ ნაბიჯზე ამოღებული ბურთის ნო-მერს. განასხვავებენ ორი ტიპის ამორჩევებს: **დალაგებული ამორჩევები** და **დაულაგებელი ამორჩევები.** პირველ შემთხვევაში ამორჩევები, რომლებიც შედგებიან ერთი და იგივე ელემენტებისგან, მაგრამ განსხვავდებიან ამ ელემენტების დალაგების რიგით, ცხადდება **განსხვავებულად**. მეორე შემთხვევაში ელემენტების დალაგების რიცხვი არ მიიღება მხედველობაში და ასეთი ამორჩევები ცხადდება **იდენტურად**. იმისათვის, რომ განვასხვავოთ ეს ამორჩევები დალაგებული ამორჩევები აღინიშნება სიმბოლოთი $(a_1,...,a_n)$, ხოლო დაულაგებელი ამორჩევები – $[a_1,...,a_n]$.

ამგვარად, **დაბრუნებით დალაგებული ამორჩევის** შემთხვევაში ელემენტარ-ულ ხდომილებათა სივრცეს აქვს შემდეგი სტრუქტურა

$$\Omega=\{\omega:\omega=(a_1,...,a_n):a_i=1,...,M;\,i=1,...,n\}$$

და განსხვავებულ შედეგთა (ამორჩევათა) რაოდენობა, რომელსაც კომბინატორი-კაში უწოდებენ *წყობებს* M-დან n *დაბრუნებებით,* ტოლია $|\Omega|=M^n$.

თუ კი ჩვენ ვიხილავთ **დაულაგებელ ამორჩევებს დაბრუნებით**, რომელსაც კომბინატორიკაში უწოდებენ *ჯუფდებას* M-დან n, მაშინ

$$\Omega=\{\omega:\omega=[a_1,...,a_n]:a_i=1,...,M;\,i=1,...,n\}.$$

გასაგებია, რომ განსხვავებულ დაულაგებელ ამორჩევათა რიცხვი ნაკლებია ვიდ-რე დალაგებულების რიცხვი. ვაჩვენოთ, რომ $|\Omega|=C^n_{M+n-1}$,

სადაც $C^l_k:=\dfrac{k!}{l!(k-l)!}$ -- არის *ჯუფდებათა რიცხვი* k*-დან* l.

დამტკიცებას ჩვენ ჩავატარებთ მათემატიკური ინდუქციის პრინციპით. აღვნიშნოთ $N(M,n)$-ით ჩვენთვის საინტერესო შედეგების რიცხვი. ცხადია, რომ თუ $k\le M$, მაშინ $N(k,1)=k=C^1_k$. დავუშვათ ახლა, რომ $N(k,n)=C^n_{k+n-1}$, $k\le M$ და ვაჩვენოთ, რომ ეს ფორმულა ძალაში დარჩება n-ის $n+1$-ით შეცვლისას. შევნიშნათ, რომ დაულაგებელი [$a_1,...,a_{n+1}$] შერჩევების განხილვისას შეიძლება ვიგულისხმოთ, რომ მისი ელემენტები დალაგებულია კლების მიხედვით: $a_1\le a_2\le\cdots\le a_{n+1}$.

ადვილი დასანახია, რომ იმ დაულაგებელი $a_1 \le a_2 \le \cdots \le a_{n+1}$ შერჩევების რაოდენობა, რომელთათვისაც $a_1 = 1$ ტოლია $N(M,n)$-ის, რომელთათვისაც $a_1 = 2$ -- $N(M-1,n)$-ის და ა. შ. ამიტომ

$$N(M,n+1) = N(M,n) + N(M-1,n) + \cdots + N(1,n) =$$
$$= C^n_{M+n-1} + C^n_{M-1+n-1} + \cdots C^n_n = (C^{n+1}_{M+n} - C^{n+1}_{M+n-1}) +$$
$$+ (C^{n+1}_{M-1+n} - C^{n+1}_{M-1+n-1}) + \cdots + (C^{n+1}_{n+1} - C^n_n) + C^n_n = C^{n+1}_{M+n},$$

სადაც ჩვენ ვისარგებლეთ ბინომის კოეფიციენტების შემდეგი თვისებით:

$$C^{l-1}_k + C^l_k = C^l_{k+1}.$$

ამორჩევა დაბრუნების გარეშე. ვიგულისხმოთ, რომ $n \le M$ და ამოღებული ბურთი უკან არ ბრუნდება. ამ შემთხვევაშიც აგრეთვე განიხილება ორი შესაძლებლობა დაკავშირებული ამორჩევის დალაგება -- დაულაგებლობასთან.

დაბრუნების გარეშე დალაგებული შერჩევის შემთხვევაში, რომელსაც კო-მბინატორიკაში უწოდებენ *წყობებს* M-დან n *დაბრუნებების გარეშე,* შედეგების სიმრავლეს აქვს სახე:

$$\Omega = \{\omega : \omega = (a_1,..., a_n) : a_k \ne a_l,\, k \ne l;\, a_i = 1,..., M;\, i = 1,..., n\},$$

ხოლო ამ სიმრავლის ელემენტთა რაოდენობა ტოლია $|\Omega| = M(M-1)\cdots(M-n+1)$. ეს რიცხვი აღინიშნება სიმბოლოთი $(M)_n$ ან A^n_M და ეწოდება *წყობათა რიცხვი* M-დან n.

დაულაგებელი ამორჩევისას დაბრუნების გარეშე, რომელსაც კომბინატორ-იკაში უწოდებენ *ჯუფდებას* M-დან n *დაბრუნებების გარეშე,* შედეგთა სიმრავ-ლეს აქვს სახე

$$\Omega = \{\omega : \omega = [a_1,..., a_n] : a_k \ne a_l,\, k \ne l;\, a_i = 1,..., M;\, i = 1,..., n\}$$

და მისი ელემენტების რაოდენობა ტოლია $|\Omega| = C^n_M$. მართლაც, დაულაგებელი $[a_1,..., a_n]$ ერთობლიობიდან, რომელიც შედგება განსხვავებული ელემენტებისაგან, შეიძლება მივიღოთ ზუსტად $n!$ რაოდენობა დალაგებული ერთობლიობები. ამიტ-ომ $|\Omega| \cdot n! = A^n_M$, და მაშასადამე, $|\Omega| = A^n_M / n! = C^n_M$.

საბოლოოდ ჩვენ გვაქვს ყუთიდან, რომელშიც M რაოდენობის ბურთია n რაოდენობის ბურთის ამოღების რიცხვის შემდეგი ცხრილი:

	დალაგებული	დაულაგებელი
დაბრუნებით	M^n	C^n_{M+n-1}
დაბრუნების გარეშე	A^n_M	C^n_M

მაგალითად, $M = 3$ და $n = 2$-ის შემთხვევაში შესაბამის ელემენტარულ ხდომილებათა სივრცეებს ექნებათ შემდეგი სახის სტრუქტურები:

			შერჩევა
	(1,1)(1,2)(1,3) (2,1)(2,2)(2,3) (3,1)(3,2)(3,3)	[1,1][2,2][3,3] [1,2][1,3][2,3]	დაბრუნებით
	(1,2)(1,3) (2,1)(2,3) (3,1)(3,2)	[1,2] [1,3] [2,3]	დაბრუნების გარეშე
ერთობლიობა	დალაგებული	დაულაგებელი	

ნაწილაკების განლაგება დანაყოფებში. განვიხილოთ საკითხი ელემენტარულ ხდომილებათა სივრცის სტრუქტურის შესახებ n ნაწილაკის (ბურთის) M დანაყოფში (ყუთში) განლაგების ამოცანაში. ასეთი ამოცანებთან საქმე გვაქვს სტატისტიკურ ფიზიკაში, როცა სწავლობენ n ნაწილაკის (ეს შეიძლება იყოს პროტონები, ნეიტრონები, და სხვა) M მდგომარეობაში (ეს შეიძლება იყოს ენერგეტიკული დონეები) განაწილების საკითხებს. ვიგულისხმოთ, რომ დანაყოფები გადანომრილია ნომრებით $1,2,...,M$ და თავიდან დავუშვათ, რომ ნაწილაკები გარჩევადია (განსხვავებულებია, აქვთ ნომრები $1,2,...,n$). მაშინ n ნაწილაკის M დანაყოფში განაწილების სრულიად აღიწერება დალაგებული ერთობლიობით $(a_1,...,a_n)$, სადაც a_i -- წარმოადგენს იმ დანაყოფის ნომერს, რომელშიც მოხვდა ნაწილაკი ნომრით i. თუ კი განვიხილავთ განურჩეველ ნაწილაკებს, მაშინ მათი განაწილება M დანაყოფში სრულიად აღიწერება დაულაგებელი ერთობლიობით $[a_1,...,a_n]$, სადაც a_i -- იმ დანაყოფის ნომერია, რომელშიც მოხვდა ნაწილაკი i-ურ ნაბიჯზე.

მაგალითები:

I. კურსზე, რომელზეც სამი ჯგუფია, ჯგუფხელების არჩევის ყველა შესაძლო ვარიანტების რიცხვია $n_1 \cdot n_2 \cdot n_3$, სადაც n_i -- i-ურ ჯგუფში სტუდენტების რაოდენობაა (ვიყენებთ ნამრავლის პრინციპს);

II. რაოდენობა ყველა შესაძლო კომბინაციების, რამდენნაირადაც შესაძლებელია m მგზავრი განვათავსოთ n ვაგონში ტოლია n^m (დალაგებული ამორჩევა დაბრუნებით, სადაც $M=n$ და $n=m$);

III. m ადამიანის დაბადების დღეების ყველა შესაძლო კომბინაცია (იმ პირობით, რომ თითეული დაბადების დღე არის რომელიმე 365 დღი-დან) ტოლია 365^m (საქმე გვაქვს ამორჩევასთან დაბრუნებით, ამასთანავე ამორჩევები ითვლება დალაგებულად, სადაც $M=365$ და $n=m$);

IV. რაოდენობა ყველა შესაძლო კომბინაციების, რამდენნაირადაც შესაძლებელია 5 ბურთი განვათავსოთ 5 ყუთში, ისე რომ ერთ ყუთში იყოს ერთი ბურთი, ტოლია 5! (ნამრავლის პრინციპის თანახმად);

V. პარტიების რაოდენობა, რომელიც უნდა შედგეს n მონაწი-ლისაგან შემდგარ წრიულ საჭადრაკო ტურნირში ტოლია $C_n^2 = n(n-1)/2$ (დაულაგებელი ამორჩევა დაბრუნების გარეშე).

ამოცანა 1 (დამთხვევებზე). ვიპოვოთ ალბათობა იმისა, რომ: ა). m შემთხ-ვევით არჩეული ადამიანის დაბადების დღეები არ დაემთხვევა

ერთმანეთს (იმ პირობით, რომ ყველა დღე ტოლალბათურია); ბ). m შემთხვევით არჩეულ ადამიანში მოიძებნება ორი მაინც ისეთი, რომელთა დაბადების დღეები დაემთხვევა ერთმანეთს.

ამოხსნა. ა). ელემენტარულ ხდომილებათა სივრცე შეესაბამება დალაგებ-ულ ამორჩევებს დაბრუნებით, სადაც $M=365$ და $n=m$, ანუ $|\Omega|=365^m$; ხოლო ხელშემწყობი ელემენტარული ხდომილებები კი დალაგებულ ამორჩევებს დაბრუნების გარეშე, სადაც აგრეთვე $M=365$ და $n=m$, ამიტომ მათი რაოდენობაა -- A_{365}^m. შესაბამისად, ალბათობის კლასიკური განმარტების თანახმად, გვაქვს

$$P(m)=\frac{A_{365}^m}{365^m}=(1-\frac{1}{365})(1-\frac{2}{365})\cdots(1-\frac{m-1}{365}).$$

ბ). საწინააღმდეგო ხდომილების ალბათობის გამოთვლის წესის თანახმად კი გვაქვს, რომ

$$Q(m)=1-\frac{A_{365}^m}{365^m}.$$

მოვიყვანოთ ამ ალბათობის მნიშვნელობების ცხრილი ზოგიერთი m -ის შემთხვევაში:

m	4	16	22	23	40	64	70
Q(m)	0.01636	0.28360	0.47569	0.50730	0.89123	0.99711	0.99916

საინტერესოა აღინიშნოს, რომ (მოლოდინის საწინააღმდეგოდ!) კლასის მოსწავლეთა რაოდენობა, რომელშიც 1/2-ის ტოლი ალბათობით მოიძებნება ორი მოსწავლე მაინც ერთი და იგივე დაბადების დღეებით, არც ისე დიდია: იგი ტოლია მხოლოდ 23-ის.

ამოცანა 2 (მოგება ლატარეაში). გვაქვს M ბილეთი გადანომრილი რიცხვებით ერთიდან M-მდე, რომელთაგან n ბილეთი ნომრებით ერთიდან n-მდე მომგებიანია ($M\ge 2n$). ვყიდულობთ n ცალ ბილეთს. რას უდრის ალბათობა იმისა, რომ ამ n ბილეთიდან ერთი მაინც იქნება მომგებიანი (ავღნიშნოთ ეს ხდომილება A-თი)?

ამოხსნა. ვინაიდან ბილეთების ამოღების (ყიდვის) თანმიმდევრობას არა აქვს მნიშვნელობა ნაყიდ ბილეთებში მომგებიანი ბილეთების არსებობის ან არ არსებობის თვალსაზრისით, ამიტომ ელემენტარულ ხდომილებათა სივრცეს ექნ-ება შემდეგი სტუქტურა:

$$\Omega=\{\omega:\omega=[a_1,...,a_n]:a_k\ne a_l,\, k\ne l;\, a_i=1,...,M\}.$$

შესაბამისად, ჩვენს მიერ ზემოთმოყვანილი ცხრილის თანახმად

$|\Omega|=C_M^n$.

ხდომილებას (ავღნიშნოთ იგი B_0-ით), რომ ნაყიდ ბილეთებში არ არის მო-მგებიანი ბილეთები, ექნება შემდეგი სტრუქტურა:

$$B_0=\{\omega:\omega=[a_1,...,a_n]:a_k\ne a_l,\, k\ne l;\, a_i=n+1,...,M\} \text{ და } |B_0|=C_{M-n}^n.$$

ამიტომ

$$P(B_0)=\frac{C_{M-n}^n}{C_M^n}=\frac{A_{M-n}^n}{A_M^n}=(1-\frac{n}{M})(1-\frac{n}{M-1})\cdots(1-\frac{n}{M-n+1})\,.$$

შესაბამისად, საძებნი ალბათობა იქნება

$$P(B)=1-P(B_0)=1-(1-\frac{n}{M})(1-\frac{n}{M-1})\cdots(1-\frac{n}{M-n+1})\,.$$

თუ მაგალითად, $M=n^2$ და $n\to\infty$, მაშინ $P(B_0)\to e^{-1}$ (აქ e ნეპერის რიცხვია) და $P(B)\to 1-e^{-1}\approx 0{,}632$, სადაც კრებადობის სიჩქარე საკმაოდ დიდია: უკვე როცა $n=10$ -- $P(B)=0{,}670$.

დავალება. წინა ამოცანის პირობებში ვიპოვოთ ალბათობა იმისა, რომ ნაყიდი n ბილეთიდან ზუსტად m ($m\le n$) იქნება მომგებიანი.

ამოცანა 3 (ურთიერთობის უპირატესობაზე). დავუშვათ კლასში, რომელშიც 10 მოსწავლეა ტარდება გამოკითხვა, სადაც თითოეულმა მოსწავლემ ანკეტაში უნდა მიუთითოს ის სამი ამხანაგი, რომელსაც აძლევს უპირატესობას ცხრა ამხანაგიდან. A იყოს ხდომილება, რომ ერთერთი მოსწავლე დასახელებულ იქნა ყველა შესაძლო ცხრა ანკეტაში. რას უდრის მისი ალბათობა, თუ ანკეტის შევსება იყო შემთხვევითი, ანუ ანკეტის შევსების ნებისმიერი კომბინაცია ტოლალბათურია.

ამოხსნა. ცალკეული მოსწავლისათვის ანკეტის შევსების სხვადასხვა კომბინაციათა რაოდენობა ტოლია C_9^3-ის, ხოლო 10 ანკეტის შევსების ვარიანტების რაოდენობა პირველი მაგალითის ანალოგიურად ტოლია $(C_9^3)^{10}$-ის. ვინაიდან ერთი ანკეტა შეიძლება შევსებულ იქნეს ნებისმიერად, ხოლო დანარჩენ ცხრა ანკეტაში ერთი პასუხი დაფიქსირებულია, ხოლო დანარჩენი ორი პასუხი კი შეიძლება ნებისმიერად ამოირჩეს რვა შესაძლებელი პასუხიდან, ამიტომ ელემენტარული ხდომილებების რაოდენობა A-ში ტოლია $|A|=C_9^3\cdot(C_8^2)^9$ (თუ წყვილის პირველ კომპონენტს შეუძლია მიიღოს m განსხვავებული მნიშვნელობა, ხოლო მეორე კომპონენტს კი პირველისგან დამოუკიდებლად -- n განსხვავებული მნიშვნელობა, მაშინ ასეთი წყვილების რაოდენობა ნამრავლის პრინციპის თანახმად იქნება -- $m\cdot n$). აქედან გვაქვს

$$P(A)=\frac{C_9^3\cdot(C_8^2)^9}{(C_9^3)^{10}}=(\frac{C_8^2}{C_9^3})^9=\frac{1}{3^9}\,.$$

დავალება. წინა ამოცანაში იპოვეთ ალბათობა იმისა, რომ ერთერთი სტუდენტი დასახელებული იქნება k-ჯერ ($k\le 9$).

ამოცანა 4 (ორ "ტუზზე"). განვიხილოთ პრეფერანსის თამაში, როდესაც კარტის შეკვრის მაღალი 32 კარტი შემთხვევით ნაწილდება (რიგდება) სამ მოთამაშეს შორის, ისე რომ თითეული ღებულობს 10 კარტს და ორი კარტი ინახება "საყიდლებში". როგორია ალბათობა იმისა, რომ "საყიდლებში" აღმოჩნდება ორი "ტუზი"?

ამოხსნა. ორი კარტის სხვადასხვა კომბინაციების რაოდენობა, რომელიც შეიძლება აღმოჩნდეს "საყიდლებში" ტოლია $C_{32}^2=496$. კარტის შეკვრაში ოთხი "ტუზია" და სხვადასხვა კომბინაციების რაოდენობა,

რომელიც მოგვცემდა ორ “ტუზს” ტოლია $C_4^2 = 6$. შესაბამისად, საძიებელი ალბათობა ტოლია

$$C_4^2 \big/ C_{32}^2 = 6 \big/ 496 = 0{,}012 .$$

დავალება. დავუშვათ წინა ამოცანაში ერთერთმა მოთამაშემ, ნახა რა თა-ვისი კარტები, იცის, რომ მას “ტუზი” არა აქვს. შეიცვლება თუ არა მაშინ ალბ-ათობა იმისა, რომ “საყიდლებში” ორი “ტუზია”? გამოთვალეთ ეს ალბათობა.

ამოცანა 5 (მხედველობით მოძებნაზე). დავუშვათ გვაქვს N ცალი შემთხვ-ევით დალაგებული გეომეტრიული ფიგურა, რომელთა შორის M ცალი მართკუ-თხედია ($M \le N$). მოითხოვება მოიძებნოს ყველა მართკუთხედი, თუ ძებნა წარმ-ოებს ელემენტების (ფიგურების) სათითაოდ სკანირებით ფიქსაციის მოცულობით ერთი ელემენტი, ამასთანავე ხდება დაკვირვებული ელემენტის პოზიციის დამახ-სოვრება და მას თავიდან აღარ ვუბრუნდებით. ვიპოვოთ ალბათობა იმისა, რომ დამკვირვებელი შეძლებს აღმოაჩინოს ყველა M მართკუთხედი არაუმეტეს n დაკვირვებისას ($n = M,\dots,N$)?

ამოხსნა. ყველა შესაძლო ელემენტარულ ხდომილებათა რაოდენობა იქნე-ბა $|\Omega| = C_N^n$. ხელშემწყობი ელემენტარული ხდომილებები ისეთი [$a_1,\dots,a_n$] ერთო-ბლიობებია, რომლებშიც M ადგილას განთავსებულია მართკუთხედები (ამის შესაძლებლობათა რაოდენობაა C_M^M), ხოლო დანარჩენ $N-M$ ადგილას კი არა მართკუთხედები (ამის შესაძლებლობათა რაოდენობაა C_{N-M}^{n-M}). ამიტომ პირველი მაგალითის ანალოგიურად ხელშემწყობი ელემენტარული ხდომილებების რაოდე-ნობა იქნება $C_M^M \cdot C_{N-M}^{n-M}$. შესაბამისად, საძიებელი ალბათობა ტოლია

$$P_M(n) = \frac{C_M^M \cdot C_{N-M}^{n-M}}{C_N^n} = \frac{C_{N-M}^{n-M}}{C_N^n} = \frac{C_n^M}{C_N^M} .$$

§8. ჯამისა და სხვაობის ალბათობის ფორმულები

გავიხსენოთ, რომ A ხდომილების ალბათობა ეწოდება მასში შემავალი ელემენტარული ხდომილებების ალბათობების ჯამს:

$$P(A)=\sum_{\omega_i\in A}P(\omega_i).$$

აქედან ჩვენ შეგვიძლია მივიღოთ ე. წ. **ალბათობათა შეკრების კანონი**: თუ A და B ხდომილებები უთავსებადია ($A\cap B=\varnothing$), მაშინ ხდომილებათა ჯამის ალბათობა ალბათობები ჯამის ტოლია

$$P(A+B)=P(A)+P(B). \qquad (1)$$

მართლაც, გვაქვს:

$$P(A+B)=\sum_{\omega_i\in(A+B)}P(\omega_i)=\sum_{\omega_i\in A}P(\omega_i)+\sum_{\omega_i\in B}P(\omega_i)=P(A)+P(B).$$

შედეგი 1. საწინააღმდეგო ხდომილების ალბათობა ტოლია ერთს გამოკლებული თვითონ ამ ხდომილების ალბათობა:

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

დამტკიცება. ვინაიდან $A\cap\overline{A}=\varnothing$ და $A\cup\overline{A}=\Omega$, ამიტომ ალბათობათა შეკრების კანონის თანახმად გვაქვს:

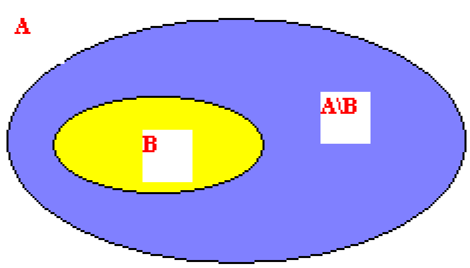
$$1=P(\Omega)=P(A\cup\overline{A})=P(A)+P(\overline{A}),$$

საიდანაც ვრწმუნდებით შედეგის სისწორეში.

შედეგი 2 (სხვაობის ალბათობის ფორმულა). თუ B ხდომილება იწვევს A ხდომილებას ($B\subset A$), მაშინ A და B ხდომილებების სხვაობის ალბათობა ალბათობათა სხვაობის ტოლია: $P(A\setminus B)=P(A)-P(B)$.

დამტკიცება. ამ შემთხვევაში A ხდომილება შეიძლება წარმოდგენილ იქნეს როგორც უთავსებადი B და $A\setminus B$ ხდომილებების ჯამი:

$$A=B+(A\setminus B).$$



ამიტომ ალბათობათა შეკრების კანონის თანახმად გვაქვს:

$$P(A)=P(B)+P(A\setminus B),$$

საიდანაც ჩანს შედეგის მართებულობა.

ცხადია, რომ ალბათობათა შეკრების კანონის განზოგადოება შესაძლებელია წყვილ-წყვილად უთავსებად ხდომილებათა ნებისმიერი რაოდენობისათვის:

თუ $A_1,A_2,...,A_n$ ხდომილებათა წყვილ-წყვილად უთავსებადი სისტემაა ($A_i\cap A_j =$ =Ø, როცა $i\neq j$ 0, მაშინ:

$$P(\sum_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) . \qquad (2)$$

ვისარგებლოთ მათემატიკური ინდუქციის მეთოდით. (2) ფორმულა $n = 2$-ის შემთხვევაში სამართლიანია (იხ. (1)). დავუშვათ, რომ იგი სამართლიანია n-სათვის და ვაჩვენოთ, რომ სამართლიანი იქნება $n+1$-ის შემთხვევაში. გვაქვს:

$$P(\sum_{i=1}^{n+1} A_i) = P(\sum_{i=1}^{n} A_i + A_{n+1}) \overset{(1)}{=} P(\sum_{i=1}^{n} A_i) + P(A_{n+1}) \overset{(2)}{=} \sum_{i=1}^{n} P(A_i) + P(A_{n+1}) = \sum_{i=1}^{n+1} P(A_i) .$$

ამით (2) თანაფარდობა დამტკიცებულია.

თეორემა 1. თუ A და B ნებისმიერი ხდომილებებია, მაშინ

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) . \qquad (3)$$

დამტკიცება. განმარტების თანახმად

$$P(A \cup B) = \sum_{\omega_i \in (A \cup B)} P(\omega_i) \quad \text{და} \quad P(A) + P(B) = \sum_{\omega_i \in A} P(\omega_i) + \sum_{\omega_i \in B} P(\omega_i) .$$

მაგრამ ჯამში $P(A) + P(B)$ ალბათობები $P(\omega_i)$, როცა $\omega_i \in A \cap B$, გათვალისწინებულია ორჯერ, ხოლო ამ $P(\omega_i)$-ების ჯამი არის $P(A \cap B)$. ამიტომ თუ $P(A) + P(B)$ ჯამს გამოვაკლებთ $P(A \cap B)$-ს, მივიღებთ სწორედ საძიებელ $P(A \cup B)$ ალბათობას.

დავალება. დაამტკიცეთ თეორემა 1 ალბათობათა შეკრების კანონისა და სხვაობის ალბათობის ფორმულის გამოყენებით.

სამი ხდომილების შემთხვევაში (თუ ვისარგებლებთ ე. წ. დე მორგანის კანონით $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$) გვექნება:

$$P(A \cup B \cup C) = P((A \cup B) \cup C) \overset{(3)}{=} P(A \cup B) + P(C) - P[(A \cup B) \cap C] \overset{(3)}{=}$$

$$\overset{(3)}{=} P(A) + P(B) - P(A \cap B) + P(C) - P[(A \cap C) \cup (B \cap C)] \overset{(3)}{=}$$

$$\overset{(3)}{=} P(A) + P(B) + P(C) - P(A \cap B) - \{P(A \cap C) + P(B \cap C) - P[(A \cap C) \cap (B \cap C)]\} =$$

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

ზოგად შემთხვევაში სამართლიანია შემდეგი თანაფარდობა:

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) - \sum_{i \neq j} P(A_i \cap A_j) + \sum_{i \neq j \neq k} P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n+1} P(\bigcap_{i=1}^{n} A_i) .$$

§9. პირობითი ალბათობის ფორმულა

ალბათობის თეორიაში ალბათობის ცნებასთან ერთად შემოდის ე. წ. პირობითი ალბათობის ცნება. იმისათვის, რომ დავინახოთ მისი საჭიროება მოვიყვანოთ შემდეგი მარტივი მაგალითი: დავუშვათ, რომ ადამიანს დაავიწყდა ტელეფონის ნომრის ორი ციფრი. ვიპოვოთ ალბათობა იმისა, რომ შემთხვევით აკრეფილი ორი ციფრით მოხერხდება სასურველ აბონენტთან დაკავშირება?

ცხადია, რომ ამ შემთხვევაში $\Omega=\{(i,j):i,j=0,1,...,9\}$ და ნამრავლის პრინციპის თანახმად $|\Omega|=10\cdot 10=100$. ვინაიდან სასურველ აბონენტთან შეერთებას ხელს უწყობს ციფრების ერთადერთი $(i.j)$ წყვილი, ამიტომ საძიებელი ალბათობა ტოლია 1/100-ის.

ვთქვათ, ადამიანს გაახსენდა რომ ეს ციფრები იყო სხვადასხვა. რა იქნება მაშინ სასურველ აბონენტთან დაკავშირების ალბათობა? ამ შემთხვევაში ელემენტარულ ხდომილებათა სივრცე იქნება $\Omega=\{(i,j):i=0,1,...,9;$ $j=0,1,...,9; i\neq j\}$. შესაბამისად, $|\Omega|=10\cdot 9=90$, ხოლო საძიებელი ალბათობა იქნება 1/90. როგორც ვხედავთ დამატებითი ინფორმაციის გაჩენამ შეცვალა (კერძოდ, გაზარდა) ალბათობა.

მოვიყვანოთ კიდევ ერთი მაგალითი. ვიპოვოთ ალბათობა იმისა, რომ ერთი კამათლის ერთჯერ გაგორებისას მოვა სამის ჯერადი ქულა, თუ ცნობილია, რომ მოვიდა ლუწი ქულა? ამ შემთხვევაში $\Omega=\{2,4,6\}$, $|\Omega|=3$, ხოლო ხელშემწყობი ელენმენტარული ხდომილება ერთადერთია (კერძოდ, 6-იანის მოსვლა) და საძიებელი ალბათობა იქნება 1/3.

შევხედოთ იგივე მაგალითს სხვანაირად. დავუშვათ, შემთხვევითი მოვლენა მდგომარეობს ერთი კამათლის ერთჯერ გაგორებაში. შემოვიღოთ ხდომილობები: A -- მოვიდა სამის ჯერადი ქულა და B -- მოვიდა ლუწი ქულა. ნათელია, რომ ამ დროს $A\cap B$ იქნება – მოვიდა სამის ჯერადი ლუწი ქულა. ცხადია, რომ ამ შემთხვევაში: $\Omega=\{1,2,3,4,5,6\}$; $A=\{3,6\}$; $B=\{2,4,6\}$; $A\cap B$={6}, ხოლო ალბათობის კლასიკური განმარტების თანახმად კი: $P(A)=2/6=1/3$; $P(B)=3/6=1/2$ და $P(A\cap B)=1/6$. ადვილი დასანახია, რომ ზემოთ გამოთვლილი ალბათობა 1/3 (რომელსაც ბუნებრივია დავარქვათ A ხდომილების ალბათობა პირობაში, რომ ადგილი ჰქონდა B ხდომილებას) ფორმალურად შეგვეძლო მიგვეღო შემდეგნაირად:

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

ამა თუ იმ მოვლენის ანალიზის დროს ხშირად იბადება კითხვა რა გავლენას ახდენს რაიმე A ხდომილების მოხდენის შესაძლებლობაზე რაიმე სხვა B ხდომილების მოხდენა. უმარტივესი მაგალითებია, როცა B ხდომილების მოხდენა აუცილებლად იწვევს A ხდომილების განხორციელებას, ე. ი. $B\subset A$, ან პირიქით, B ხდომილების მოხდენა გამორიცხავს A ხდომილების განხორციელების შესაძლებლობას, ე. ი. $A\cap B=\varnothing$. ალბათობის თეორიაში A და B ხდომილებებს შორის კავშირის დასახასიათებლად შემოდის A ხდომილების **პირობითი ალბათობის ცნება** B ხდომილების მიმართ.

განმარტება 1. *A ხდომილების პირობითი ალბათობა პირობაში, რომ ადგილი ჰქონდა B ხდომილებას* აღინიშნება $P(A|B)$ სიმბოლოთი და განიმარტება შემდეგნაირად

$$P(A|B) = P(A \cap B)/P(B), \tag{1}$$

თუ $P(B) \neq 0$.

ზემოთ მოყვანილი უკანასკნელი მაგალითი გარკვეული აზრით შეიძლება ჩაითვალოს პირობითი ალბათობის განმარტების მოტივაციად. მოვიყვანოთ ახლა მოსაზრება, რომელიც გაამართლებს A ხდომილების პირობითი ალბათობის (პირობაში, რომ მოხდა B ხდომილება) $P(A|B)$-ს განმარტებას (1) თანაფარდობით. დავუშვათ, რომ ექსპერიმენტის შედეგად შესაძლებელია მოხდეს A და B ხდომილები. A ხდომილების ***პირობითი სიხშირე,*** პირობაში რომ მოხდა B ხდომილება ვუწოდოთ A ხდომილების სიხშირეს გამოთვლილს არა ყველა ექსპერიმენტის მიმართ, არამედ იმ ექსპერიმენტების მიმართ, რომლებშიც ადგილი ჰქონდა B ხდომილებას. სხვა სიტყვებით, რომ ვთქვათ, თუ n -- ყველა ექსპერიმენტის რაოდენობაა, $\mu_n(B)$ -- B ხდომილების მოხდენათა რიცხვია, $\mu_n(A \cap B)$-- $A \cap B$ ხდომილების მოხდენათა რიცხვია (ანუ იმ ექსპერიმენტების რაოდენობა, სადაც ერთდროულად მოხდა A და B), მაშინ პირობითი სიხშირე არის

$$\frac{\mu_n(A \cap B)}{\mu_n(B)} = \frac{\mu_n(A \cap B)/n}{\mu_n(B)/n}.$$

n-ის დიდი მნიშვნელობებისათვის ამ გამოსახულების მარცხენა მხარეს შეიძლება შევხედოთ როგორც B ხდომილების მოხდენის პირობაში A ხდომილების მოხდენის პირობითი ალბათობის $P(A|B)$ მიახლოებით მნიშვნელობას, შეფარდება $\mu_n(A \cap B)/n$ -- არის $P(A \cap B)$ ალბათობის მიახლოებითი მნიშვნელობა, ხოლო შეფარდება $\mu_n(B)/n$ კი არის $P(B)$ ალბათობის მიახლოებითი მნიშვნელობა. ეს მსჯელობა უდევს სწორედ საფუძვლად პირობითი ალაბათობის (1) განმარტებას.

პირობით ალბათობებს გააჩნიათ ჩვეულებრივი ალბათობების ანალოგიური თვისებები:

I. ნებისმიერი A და B ხდომილებებისათვის ($P(B) \neq 0$):

$$0 \leq P(A|B) \leq 1;$$

II. A და B უთავსებადია ($A \cap B = \varnothing$), მაშინ $P(A|B) = 0$;

III. თუ B იწვევს A-ს ($B \subset A$), მაშინ $P(A|B) = 1$.

თეორემა 1. თუ $A_1, A_2, ..., A_n$ წყვილ-წყვილად უთავსებადი ხდომილებებია, მაშინ ნებისმიერი B ხდომილებისათვის ($P(B) \neq 0$) სამართლიანია ტოლობა:

$$P[(\sum_{i=1}^{n} A_i) | B] = \sum_{i=1}^{n} P(A_i | B).$$

დამტკიცება. ცხადია, რომ $(\sum_{i=1}^{n} A_i) \cap B = \sum_{i=1}^{n} (A_i \cap B)$. გარდა ამისა, ვინაიდან ხდომილებები $A_1, A_2, ..., A_n$ წყვილ-წყვილად უთავსებადი ხდომილებე-

ბია, წყვილ-წყვილად უთავსებადი იქნება ხდომილებები $A_1 \cap B, A_2 \cap, \ldots,$ $A_n \cap B$. ამიტომ პირობითი ალბათობის განმარტების თანახმად, ალბათობათა შეკრების კანონის გათვალისწინებით, ვწერთ:

$$P[(\sum_{i=1}^{n} A_i) \mid B] = \frac{P[(\sum_{i=1}^{n} A_i) \cap B]}{P(B)} = \frac{P[\sum_{i=1}^{n} (A_i \cap B)]}{P(B)} = \frac{\sum_{i=1}^{n} P(A_i \cap B)}{P(B)} =$$

$$= \sum_{i=1}^{n} \frac{P(A_i \cap B)}{P(B)} = \sum_{i=1}^{n} P(A_i \mid B).$$

ამოცანა 1. განვიხილოთ ოჯახები, სადაც ორ-ორი ბავშვია. როგორია ალბათობა იმისა, რომ ოჯახში ორივე ბავშვი ვაჟია პირობაში, რომ: ა). უფროსი ბავშვი – ვაჟია; ბ). ერთი ბავშვი მაინც – ვაჟია?

ამოხსნა. აქ ელემენტარულ ხდომილებათა სივრცე ასეთია

$$\Omega = \{ვვ, ვქ, ქვ, ქქ\},$$

სადაც "ვ" აღნიშნავს ვაჟს, ხოლო "ქ" – ქალს. ჩავთვალოთ, რომ ოთხივე შედეგი ტოლალბათურია. შემოვიღოთ ხდომილებები: A -- იყოს ხდომილება, რომ უფროსი ბავშვი -- ვაჟია, ხოლო B -- იყოს ხდომილება, რომ უმცროსი ბავშვი – ვაჟია. მაშინ $A \cap B$ -- იქნება ხდომილება, რომ ორივე ბავშვი ვაჟია, ხოლო $A \cup B$ -- კი იქნება ხდომილება, რომ ერთი ბავშვი მაინც ვაჟია. შესაბამისად, საძებნი ალბათობები იქნება: ა). $P(A \cap B \mid A)$ და ბ). $P(A \cap B \mid A \cup B)$. ადვილი დასანახია, რომ:

$$P(A \cap B \mid A) = \frac{P[(A \cap B) \cap A]}{P(A)} = \frac{P(A \cap B)}{P(A)} = \frac{1/4}{1/2} = \frac{1}{2},$$

$$P(A \cap B \mid A \cup B) = \frac{P[(A \cap B) \cap (A \cup B)]}{P(A \cup B)} = \frac{P(A \cap B)}{P(A \cup B)} = \frac{1/4}{3/4} = \frac{1}{3}.$$

ამოცანა 2 (საუკეთესოს შერჩევაზე). მოცემულია m ობიექტი გადანომრილი რიცხვებით $1,2,\ldots,m$, ამასთანავე ისე, რომ ვთქვათ, ობიექტი №1 კლასიფიცირდება როგორც "საუკეთესო", . . . , ობიექტი №m კი როგორც "ყველაზე უარესი". იგულისხმება რომ ობიექტები შემოდიან დროის მომენტებში $1,2,\ldots,m$ შემთხვევითი რიგით (ანუ ყველა შესაძლო $m!$ გადანაცვლება ტოლალბათურია). დამკვირვებელს შეუძლია ორი მათგანის შედარებით თქვას რომელია უკეთესი და რომელი უარესი. საჭიროა საუკეთესოს შერჩევა იმ პირობით რომ ობიექტები წარმოიდგინება სათითაოდ და უკუგდებულის დამახსოვრება ხდება დამკვირვებლის მიერ. არ შეიძლება საუკეთესოდ მიჩნეულ იქნეს ის ობიექტი, რომელიც დაკვირვებული ობიექტებიდან ერთზე მაინც უარესია ან რომელიც უკვე იქნა უკუგდებული. ვთქვათ, დამკვირვებელმა ობიექტი შეარჩია k-ურ ნაბიჯზე ($k \le m$), ანუ დათვალიერებული ობიექტებიდან უკანასკნელი აღმოჩნდა ყველა წინაზე უკეთესი და ამიტომ მოხდა მისი შერჩევა. როგორია ალბათობა იმისა, რომ ამორჩეული ობიექტი იქნება საუკეთესო მთელი ერთობლიობიდან როგორც უკვე განხილულ, ისე ჯერ კიდევ განუხილავ ობიექტებს შორის?

ამოხსნა. შემოვიღოთ ხდომილებები: A იყოს ხდომილება, რომ k-ური ობიექტი საუკეთეოა ყველა არსებულ m ობიექტს შორის და B იყოს

ხდომილება, რომ k-ური ობიექტი საუკეთეოა დაკვირვებულ k ობიექტს შორის. გასაგებია, რომ მოსაძებნია პირობითი ალბათობა $P(A|B)$.

ვინაიდან $A \subset B$, ამიტომ $A \cap B = A$ და $P(A \cap B) = P(A)$. შესაბამისად, პირობითი ალბათობის განმარტების თანახმად

$$P(A|B) = P(A)/P(B)$$

ვინაიდან ობიექტების ყველა შესაძლო გადანაცვლებები ტოლალბათურია, ამიტომ ალბათობის კლასიკური განმარტების თანახმად ადვილი დასანახია, რომ

$$P(B) = \frac{(k-1)!}{k!} = \frac{1}{k} \quad \text{და} \quad P(A) = \frac{(m-1)!}{m!} = \frac{1}{m}.$$

შესაბამისად, $P(A|B) = k/m$.

სტრატეგია. შეიძლება დამტკიცდეს, რომ საუკეთესო ობიექტის ამორჩევის ოპტიმალური სტრატეგია მოწყობილია შემდეგნაირად. ავღნიშნოთ სიმბოლოთი m^* ისეთი ნატურალური რიცხვი, რომლისთვისაც სამართლიანია უტოლობა

$$\frac{1}{m^*} + \cdots + \frac{1}{m-1} \le 1 < \frac{1}{m^*-1} + \cdots + \frac{1}{m-1}.$$

საუკეთესო ობიექტის არჩევის ოპტიმალური სტრატეგია მდგომარეობს იმაში, რომ დავაკვირდეთ და უკუვაგდოთ პირველი m^*-1 ობიექტი და შემდეგ გავაგრძელოთ დაკვირვება ისეთ τ^* მომენტამდე, როცა პირველად გამოჩნდება ყველა წინამორბედზე უკეთეს ობიექტი.

მაგალითად, თუ $m = 1,...,10$, მაშინ m^*-ის შესაბამისი მნიშვნელობებია:

m	1	2	3	4	5	6	7	8	9	10
m-ოპტიმალური	1	1	1	1	2	2	2	3	3	4

საკმაოდ დიდი m-ისათვის $m^* \approx m/e$ (სადაც e -- ნეპერის რიცხვია, $e \cong 2.718$) და საუკეთესო ობიექტის არჩევის ალბათობა დაახლოებით ტოლია $1/e \cong 0.368$ (თუმცა, ერთი შეხედვით ბუნებრივია მოგვჩვენებოდა, რომ განსახილველი ობიექტების m რაოდენობის ზრდასთან ერთად, საუკეთესო ობიექტის არჩევის ალბათობა უნდა წასულიყო ნულისაკენ). ე. ი. საუკეთესო ობიექტის არჩევის ოპტიმალური სტრატეგია მდგომარეობს იმაში, რომ უნდა უკუვაგდოთ ობიექტების საერთო რაოდენობის დაახლოებით მესამედი და შემდეგ ავირჩიოთ პირველი ისეთი ობიექტი, რომელიც ყველა წინაზე უკეთესია.

§10. ნამრავლის ალბათობის ფორმულა

ვიგულისხმოთ, რომ $P(B)\neq 0$, მაშინ პირობითი ალბათობის ფორმულიდან

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

შეგვიძლია დავწეროთ, რომ

$$P(A\cap B)=P(B)\cdot P(A\,|\,B)\,. \qquad (1)$$

ანალოგიურად, თუ ვიგულისხმებთ, რომ $P(A)\neq 0$, მაშინ $P(B\,|\,A)$ პირობითი ალბათობის ფორმულიდან მივიღებთ, რომ

$$P(A\cap B)=P(A)\cdot P(B\,|\,A)\,.$$

ე. ი. **ორი ხდომილების ნამრავლის ალბათობა** *ტოლია ერთ-ერთის ალბათობა გამრავლებული მეორის პირობით ალბათობაზე პირობაში, რომ მოხდა პირველი.*

ცხადია, რომ სამი ხდომილების შემთხვევაში (თუ $P(B)\neq 0$ და $P(C)\neq 0$) გვექნება:

$$P(A\cap B\cap C)=P[(A\cap B)\cap C]\overset{(1)}{=}P(A\cap B)\cdot P[C\,|\,(A\cap B)]\overset{(1)}{=}$$

$$\overset{(1)}{=}P(A)\cdot P(B\,|\,A)\cdot P[C\,|\,A\cap B]\,.$$

ანალოგიურად, $A_1, A_2, ..., A_n$ ხდომილებებისათვის გვექნება, რომ:

$$P(A_1\cap A_2\cap\cdots\cap A_n)=P(A_1)\cdot P(A_2\,|\,A_1)\cdots P(A_n\,|\,A_1\cap\cdots\cap A_{n-1})\,. \qquad (2)$$

დავალება 1. მიუთითეთ რა შემთხვევაშია სამართლიანი (2) თანაფარდობა და დაამტკიცეთ იგი.

ამოცანა 1 "ბედნიერ" ბილეთებზე). 25 საგამოცდო ბილეთიდან 5 "ბედნიერია", ხოლო დანარჩენი 20 – "არა ბედნიერი". რომელ სტუდენტს აქვს "ბედნიერი" ბილეთის აღების მეტი ალბათობა: ვინც პირველი იღებს ბილეთს, თუ ვინც მეორე იღებს ბილეთს?

ამოხსნა. ეს ამოცანა ჩვენ უკვე ამოვხსენით ალბათობის კლასიკური განმარტების გამოყენებით. ამოვხსნათ ახლა იგი ელემენტარულ ხდომილებათა სივრცის ახლებური შემოტანითა და ნამრავლის ალბათობის ფორმულის გამოყენებით. წინასწარ შემოვიღოთ შემდეგი ხდომილებები: A იყოს ხდომილება, რომ პირველმა სტუდენტმა აიღი ბედნიერი ბილეთი, ხოლო B იყოს ხდომილება, რომ მეორე სტუდენტმა აიღი ბედნიერი ბილეთი. მაშინ ცხადია, რომ ელემენტარულ ხდომილებათა სივრცე შედგება ოთხი ხდომილებისაგან

$$\Omega=\{A\cap B, A\cap\overline{B}, \overline{A}\cap B, \overline{A}\cap\overline{B}\}\,.$$

ალბათობის კლასიკური განმარტების თანახმად $P(A)=5/25=1/5$, ხოლო $P(\overline{A})=20/25=4/5$. მეორეს მხრივ, ამოცანის შინაარსიდან გამომდინარე, თუ ცნობილია, რომ პირველმა სტუდენტმა აიღო ბედნიერი ბილეთი, მაშინ ალბათობა იმისა რომ მეორე სტუდენტი აიღებს ბედნიერ ბილეთს ისევ შეიძლება გამოვითვალოთ ალბათობის კლასიკური განმარტებით: ამ შემთხვევაში ყველა შესაძლო შედეგთა რაოდენობაა 24, ხოლო ხე-

ლშემწყობ ელემენტარულ ხდომილებათა რაოდენობა კი მხოლოდ 4 (რადგან ერთი ბედნიერი ბილეთი უკვე აღებულია) და შესაბამისად,

$$P(B\,|\,A) = 4/24 = 1/6 .$$

ანალოგიურად, $P(\overline{B}\,|\,A) = 20/24 = 5/6$, $P(B\,|\,\overline{A}) = 5/24 = 5/24$ და $P(\overline{B}\,|\,\overline{A}) = 19/24$. ამიტომ ორი ხდომილების ნამრავლის ალბათობის ფორმულის თანახმად გვაქვს:

$$P(A \cap B) = P(A)\cdot P(B\,|\,A) = 1/5\cdot 1/6 = 1/30; \;\; P(A\cap \overline{B}) = 1/5\cdot 5/6 = 1/6;$$

$$P(\overline{A}\cap B) = 4/5\cdot 5/24 = 1/6 \;\; \text{და} \;\; P(\overline{A}\cap \overline{B}) = 4/5\cdot 19/24 = 19/30$$

(შევნიშნავთ, რომ ჩვენ აქ მხოლოდ სისრულისათვის გამოვთვალეთ ყველა შესაძლო პირობითი და ნამრავლის ალბათობები).

ცხადია, რომ $B = (A\cap B) + (\overline{A}\cap B)$. ამიტომ ალბათობათა შეკრების კანონის თანახმად

$$P(B) = P(A\cap B) + P(\overline{A}\cap B) = 1/30 + 1/6 = 1/5\; (= P(A)) .$$

ამოცანა 2. ყუთში m ბურთია, მათ შორის n თეთრია. ვიპოვოთ ალბათობა იმისა, რომ ყუთიდან ორი ბურთის მიმდევრობით დაბრუნების გარეშე ამოღებისას: ა). პირველი ბურთი თეთრია; ბ). მეორე ბურთი თეთრია; გ). ორივე ბურთი თეთრია.

ამოხსნა. A_i იყოს ხდომილება, რომ i-ური ბურთი თეთრია ($i = 1,2$). მაშინ ალბათობის კლასიკური განმარტების თანახმად:

ა). $P(A_1) = n/m$.

გარდა ამისა,

$$P(A_2\,|\,A_1) = (n-1)/(m-1) \;\; \text{და} \;\; P(A_2\,|\,\overline{A_1}) = n/(m-1) .$$

ამიტომ ნამრავლის ალბათობის ფორმულის თანახმად:

გ). $P(A_1\cap A_2) = P(A_1)\cdot P(A_2\,|\,A_1) = n(n-1)/m(m-1)$.

ანალოგიურად,

$$P(\overline{A_1}\cap A_2) = P(\overline{A_1})\cdot P(A_2\,|\,\overline{A_1}) = n(m-n)/m(m-1) .$$

ამიტომ:

ბ). $P(A_2) = P[(A_1\cap A_2) + (\overline{A_1}\cap A_2)] = P(A_1\cap A_2) + P(\overline{A_1}\cap A_2) = n/m$.

დავალება 2. დავუშვათ, რომ შესამოწმებელი ჯგუფის 1% ავადმყოფია, ხოლო დანარჩენი 99% კი ჯანმრთელია. ადამიანების შერჩევა ხდება შემთხვევით და ამიტომ

$$P(\text{ავადმყოფი}) = 1\% = 0.01 \;\; \text{და} \;\; P(\text{ჯანმრთელი}) = 99\% = 0.99 .$$

ვიგულისხმოთ, რომ იმ შემთხვევაში, როცა ტესტირება უტარდება ადამიანს, რომელსაც არა აქვს ავადმყოფობა, მაშინ 1%-ია ალბათობა იმისა, რომ მივიღოთ მცდარი დადებითი შედეგი, ე.ი.

$$P(\text{დადებითი}\,|\,\text{ჯანმრთელი}) = 1\% \;\; \text{და} \;\; P(\text{უარყოფითი}\,|\,\text{ჯანმრთელი}) = 99\% .$$

და ბოლოს, დავუშვათ, რომ იმ შემთხვევაში, როცა ტესტირება უტარდება ავადმყოფ ადამიანს, მაშინ 1%-ია ალბათობა იმისა, რომ მივიღოთ მცდარი უარყოფითი შედეგი, ე.ი.

$$P(\text{უარყოფითი}\,|\,\text{ავადმყოფი}) = 1\% \;\; \text{და} \;\; P(\text{დადებითი}\,|\,\text{ავადმყოფი}) = 99\% .$$

გამოთვალეთ ალბათობა იმისა, რომ: ა). ადამიანი ჯანმრთელია, ხოლო ტესტმა აჩვენა უარყოფითი შედეგი; ბ). ადამიანი ავადმყოფია, ხოლო ტესტმა აჩვენა დადებითი შედეგი; გ). ადამიანი ჯანმრთელია, ხოლო ტესტმა აჩვენა დადებითი შედეგი; დ). ადამიანი ავადმყოფია, ხოლო ტესტმა აჩვენა უარყოფითი შედეგი.

ამოცანა 3. ყუთს აქვს n უჯრა. ალბათობა იმისა რომ ბურთი არის ამ უჯრებიდან ერთ-ერთში ტოლია p-სი. იპოვეთ ალბათობა იმისა, რომ ბურთი არის i-ურ უჯრაში, თუ ცნობილია, რომ ბურთი თითეულ უჯრაში შესაძლებელია იყოს თანაბარი ალბათობებით?

ამოხსნა. A_i იყოს ხდომილება, რომ ბურთი არის i-ურ უჯრაში. A იყოს ამ ხდომილებების გაერთიანება $A = \bigcup_{i=1}^{n} A_i$. პირობის თანახმად

$$P(A) = p \quad \text{და} \quad P(A_i \mid A) = 1/n.$$

ამიტომ

$$P(A_i) = P(A_i \cap A) = P(A)P(A_i \mid A) = p \cdot 1/n = p/n.$$

§11. დამოკიდებული და დამოუკიდებელი ხდომილებები

ალბათობის თეორიაში ორ A და B ხდომილებას ეწოდება **დამოუკიდებელი**, თუ ერთ-ერთი მათგანის მოხდენა არ ცვლის მეორე მათგანის მოხდენის ალბათობას. წინააღმდეგ შემთხვევაში ამ ხდომილებებს ეწოდებათ **დამოკიდებული**. იმ შემთხვევაში, როცა ერთ-ერთი ხდომილების ალბათობა არანულოვანია, ვთქვათ, $P(B) \neq 0$, მაშინ გვაქვს შემდეგი

განმარტება 1. A ხდომილებას ეწოდება B ხდომილებისაგან *დამოუკიდებელი*, თუ

$$P(A \mid B) = P(A), \qquad (1)$$

ხოლო თუ $P(A \mid B) \neq P(A)$, მაშინ გვაქვს *დამოკიდებული* ხდომილებები.

თუ გავიხსენებთ პირობითი ალბათობის განმარტებას

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

მაშინ (1) თანაფარდობიდან მივიღებთ, რომ

$$P(A \cap B) = P(A) \cdot P(B). \qquad (2)$$

პირიქით, იმ შემთხვევაში, როცა $P(B) \neq 0$, (2) თანაფარდობიდან მიიღება (1) თანაფარდობა.

შენიშვნა. ზოგიერთ სახელმძღვანელოში ხდომილებათა დამოუკიდებლობა განიმარტება (2) თანაფარდობით ((1) და (2) ექვივალენტურია, თუ $P(B) \neq 0$). მას აქვს ის უპირატესობა, რომ მისი გამოყენება შესაძლებელია მაშინაც, როცა ხდომილებების ალბათობები ნულოვანია (მაგალითად, შეუძლებელი ხდომილება დამოუკიდებელია ნებისმიერი ხდომილებისაგან $P(A \cap \varnothing) = P(\Omega) = 0 = P(A) \cdot P(\varnothing)$) და გარდა ამისა, იგი სიმეტრიულია A და B-ს მიმართ (ასეთ შემთხვევაში, თუ A დამოუკიდებელია B-საგან, მაშინ B დამოუკიდებელია A-საგან), მაგრამ (1) თანაფარდობის უპირატესობა ის არის, რომ იქიდან ჩანს აღნიშნული განმარტების შინაარსი.

ცხადია, რომ აუცილებელი ხდომილება დამოუკიდებელია ნებისმიერი ხდომილებისაგან:

$$P(A \mid \Omega) = \frac{P(A \cap \Omega)}{P(\Omega)} = \frac{P(A)}{1} = P(A),$$

რაც ბუნებრივია იმის გამო, რომ ამ შემთხვევაში პირობა არ წარმოადგენს დამატებით ინფორმაციას (ჩვენ ისედაც ვიცოდით, რომ აუცილებელი ხდომილება ეს ის ხდომილებაა, რომელიც ყოველთვის ხდება) და ამიტომ ალბათობა არც უნდა შეცვლილიყო.

ვინაიდან A და B ხდომილებების დამოუკიდებლობა ნიშნავს, რომ ინფორმაცია A-ს მოხდენის შესახებ არ ცვლის B-ს ალბათობას, ბუნებრივია, რომ ინფორმაციამ A-ს არ მოხდენის შესახებ აგრეთვე არ უნდა შეცვალოს B-ს მოხდენის ალბათობა. მართლაც სამართლიანია შემდეგი

თეორემა 1. თუ A და B ხდომილებები დამოუკიდებელია, მაშინ ხდომილებები $\overline{A}$ და B აგრეთვე დამოუკიდებელია.

დამტკიცება. ადვილი დასანახია, რომ $\overline{A} \cap B = B \setminus (A \cap B)$ (ვინაიდან $\overline{A} \cap B = B \setminus (A \cap B)$), ამიტომ სხვაობის ალბათობის ფორმულის გამოყენებით თეორემის პირობებში ვღებულობთ, რომ

$$P(\overline{A}\cap B)=P(B)-P(A\cap B)\overset{(2)}{=}P(B)-P(A)P(B)=(1-P(A))P(B)=P(\overline{A})P(B).$$

შედეგი. თუ A და B ხდომილებები დამოუკიდებლია, მაშინ დამოუკიდებელია $\overline{A}$ და $\overline{B}$.

მაგალითი 1. კარტების ნაკრებიდან (რომელშიც 36 კარტია) შემთხვევით იღებენ ერთ კარტს. განვიხილოთ ხდომილებები: A იყოს ხდომილება, რომ ამოღებული კარტი "აგურისა", ხოლო B იყოს ხდომილება, რომ ამოღებული კარტი "მეფეა". არიან თუ არა ეს ხდომილებები დამოუკიდებელი?

ცხადია, რომ ამ შემთხვევაში

$$|\Omega|=36,\ P(A)=9/36=1/4,\ P(B)=4/36=1/9 \text{ და}$$

$$P(A\cap B)=1/36=1/4\cdot 1/9=P(A)\cdot P(B).$$

ე.ი. ეს ხდომილებები დამოუკიდებელია.

მაგალითი 2. დავუშვათ, ვაგორებთ ორ სათამაშო კამათელს. განვიხილოთ ხდომილებები: A -- პირველ კამათელზე მოვიდა კენტი ქულა, B -- მეორე კამათელზე მოვიდა კენტი ქულა, C -- ორივე კამათელზე მოსულ ქულათა ჯამი კენტია. გავარკვიოთ ამ ხდომილებების დამოუკიდებლობის საკითხი.

ცხადია, რომ $P(A)=P(B)=3/6=1/2$, ხოლო $P(A\cap B)=3\cdot 3/36=1/4$. ამიტომ A და B ხდომილებები დამოუკიდებლია.

დავალება 1. შეამოწმეთ, რომ $P(C)=1/2$.

შევნიშნოთ, რომ A და B ხდომილებიდან ერთ-ერთის მოხდენის პირობაში C ხდომილება ხდება მაშინ და მხოლოდ მაშინ, როცა შესაბამისად, ან პირველ ან მეორე კამათელზე მოვიდა ლუწი ქულა, ანუ გვაქვს თანაფარდობები:

$$A\cap C=A\cap\overline{B} \text{ და } B\cap C=\overline{A}\cap B.$$

ვინაიდან, თეორემა 1-ის ძალით, ხდომილებები A და $\overline{B}$ და $\overline{A}$ და B აგრეთვე დამოუკიდებლებია, ამიტომ

$$P(A\cap\overline{B})=P(A)P(\overline{B})=1/4 \text{ და } P(\overline{A}\cap B)=1/4.$$

შესაბამისად,

$$P(A\cap C)=P(A\cap\overline{B})=1/4 \text{ და } P(B\cap C)=P(\overline{A}\cap B)=1/4.$$

ეს თანაფარდობები კი, $P(C)=1/2$ ალბათობის გათვალისწინებით, ნიშნავს, რომ დამოუკიდებლებია A და C და B და C ხდომილებათა წყვილებიც.

განმარტება 2. ხდომილებათა ერთობლიობას $A_1,A_2,...,A_n$ ეწოდება **წყვილ-წყვილად დამოუკიდებელი** თუ ნებისმიერი ორი ხდომილება ამ ერთობლიობიდან დამოუკიდებელია, ანუ

$$P(A_i\cap A_j)=P(A_i)\cdot P(A_j),\quad \forall i\neq j.$$

წინა მაგალითში ჩვენ ვნახეთ, რომ ხდომილებები A, B და C წყვილ-წყვილად დამოუკიდებელია.

განმარტება 3. ხდომილებათა ერთობლიობას $A_1,A_2,...,A_n$ ეწოდება **ერთობლივად დამოუკიდებელი** თუ ნებისმიერი $k\leq n$ რაოდენობისათვის

და ერთმანეთისაგან განსხვავებული $i_1, i_2, ..., i_k$ ინდექსებისათვის, რომელთაგან თითოეული იცვლება ერთიდან n-მდე:

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

ცხადია, რომ თუ ხდომილებები ერთობლივად დამოუკიდებელია, მაშინ ისინი იქნებიან წყვილ-წყვილად დამოუკიდებელი. პირიქით, კი საზოგადოდ სწორი არ არის. ამის მაგალითად გამოდგება მაგალითი 2.

დავალება 2. შეამოწმეთ, რომ ხდომილებები მაგალითი 2-დან არ არიან ერთობლივად დამოუკიდებელი.

მაგალითი 3. დავუშვათ, ვაგდებთ სამ მონეტას. შემოვიღოთ ხდომილებები:

A_1 -- პირველი და მეორე მონეტა დაეცა ერთი და იგივე მხარეზე;

A_2 -- მეორე და მესამე მონეტა დაეცა ერთი და იგივე მხარეზე;

A_3 -- პირველი და მესამე მონეტა დაეცა ერთი და იგივე მხარეზე.

ადვილი შესამოწმებელია, რომ აქედან ნებისმიერი ორი ხდომილება დამოუკიდებელია, ხოლო სამივე ერთად დამოკიდებულია, ვინაიდან თუ ჩვენ გვეცოდინება რომ მაგალითად, A_1 და A_2 მოხდა, მაშინ ჩვენ ზუსტად ვიცით, რომ A_3 აგრეთვე მოხდა.

დავალება 3. შეამოწმეთ, რომ ხდომილებები A_1, A_2 და A_3 წყვილ-წყვილად დამოუკიდებელია.

თეორემა 2. თუ $A_1, A_2, ..., A_n$ ხდომილებები ერთობლივად დამოუკიდებელია, მაშინ ხდომილებები $\overline{A_1}, A_2, ..., A_n$ აგრეთვე ერთობლივად დამოუკიდებლებია.

დავალება 4. დაამტკიცეთ თეორემა 2.

ამოცანა 1 (დაბადების დღეებზე). ვიპოვოთ ალბათობა იმისა, კონკრეტული სკოლის რომ 150 მოსწავლიდან ერთი მაინც დაბადებულია მოცემულ ფიქსირებულ დღეს, მაგალითად პირველ სექტემბერს?

ამოხსნა. შემოვიღოთ ხდომილებები:

$$A_i = \{i-\text{ური სტუდენტი დაბადებულია } 1.09\}, \quad i = 1,2,...,150;$$

$$A = \{\text{ერთი მაინც 150 სტუდენტიდან დაბადებულია } 1.09\}.$$

ნათელია, რომ A_i ხდომილებები ერთობლივად დამოუკიდებელია და $P(A_i) = 1/365$. გარდა ამისა, $\boldsymbol{A = \bigcup_{i=1}^{150} A_i}$ და მაშასადამე, საპოვნელია დამოუკიდებელ ხდომილებათა გაერთიანების ალბათობა. გადავიდეთ საწინააღმდეგო ხდომილებაზე და ვისარგებლოთ დე-მორგანის კანონით. მაშინ გვაქვს:

$$P(A) = 1 - P(\overline{A}) = 1 - P(\overline{\bigcup_{i=1}^{150} A_i}) = 1 - P(\bigcap_{i=1}^{150} \overline{A_i}) = 1 - P(\overline{A_1}) \cdots P(\overline{A_n}).$$

რამდენადაც $P(\overline{A_i}) = 1 - 1/365$, თუ ვისარგებლებთ ნიუტონის ბინომის ფორმულით $(1+x)^n = \sum_{j=1}^{n} C_n^j x^j$, ვღებულობთ

$$P(A)=\frac{150}{365}-C_{150}^{2}(\frac{1}{365})^{2}+C_{150}^{3}(\frac{1}{365})^{3}-C_{150}^{4}(\frac{1}{365})^{4}+\cdots.$$

ვინაიდან, $150/365\cong 0{,}41$, ხოლო

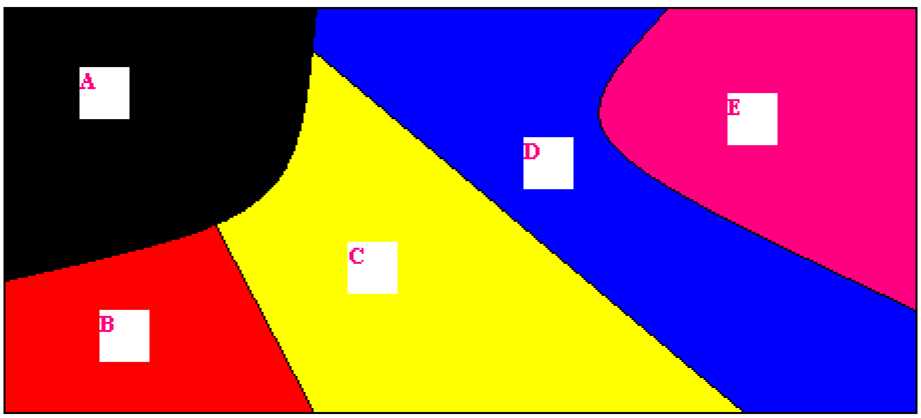
$$C_{150}^{4}(\frac{1}{365})^{4}<(\frac{150}{365})^{4}\frac{1}{24}\cong\frac{(0,41)^{4}}{24}<0.005,$$

ამიტომ მწკრივის ნიშანცვლადობის გამო, თუ გადავაგდებთ მწკრივის წევრებს დაწყებული მე-5 წევრიდან (მწკრივების ზოგადი თეორიიდან გამომდინარე), შესაძლებელია ვამტკიცოთ, რომ

$$P(A)\cong 0.41-\frac{(0.41)^{2}}{2}+\frac{(0.41)^{3}}{6}\cong 0.41-0.08+0.01=0.34.$$

§12. სრული ალბათობის ფორმულა

ხდომილებათა ერთობლიობას $A_1, A_2, ..., A_n$ ეწოდება **ხდომილებათა სრული სისტემა,** თუ ეს ხდომილებები წყვილ-წყვილად უთავსებადია და მათი გაერთიანება ემთხვევა აუცილებელ ხდომილებას: $A_i \cap A_j = \varnothing$, როცა $i \neq j$ და $\bigcup_{i=1}^{n} A_i = \Omega$. სხვა სიტყვებით, ელემენტარულ ხდომილებათა სივრცე დაყოფილია (დახლეჩილია) თანაუკვეთ ნაწილებად. ქვემოთ მოყვანილ ნახაზზე ელემენტარულ ხდომილებათა სივრცე წარმოდგენილა მართკუთხედის სახით და ხდომილებათა სრული სისტემა შედგება ხუთი თანაუკვეთი A, B, C, D, E ხდომილებისაგან.



ხდომილებათა სრული სისტემაა ნებისმიერი A ხდომილება და მისი საწინააღმდეგო $\bar{A}$ ხდომილება, ვინაიდან $A \cap \bar{A} = \varnothing$ და $A \cup \bar{A} = \Omega$.

ქვემოთ ჩვენ მოვიყვანთ ფორმულას, რომელსაც *სრული ალბათობის ფორმულა* ეწოდება და რომელიც წარმოადგენს ძირითად საშუალებას რთული ხდომილებების ალბათობების გამოსათვლელად პირობითი ალბათობების საშუალებით.

თეორემა 1. თუ $A_1, A_2, ..., A_n$ ხდომილებათა სრული სისტემაა ისეთი, რომ მისი თითოეული ხდომილების ალბათობა არანულოვანია ($P(A_i) > 0$, $i = 1, 2, ..., n$), მაშინ ნებისმიერი B ხდომილების ალბათობა გამოითვლება ფორმულით

$$P(B) = \sum_{i=1}^{n} P(A_i) P(B \mid A_i), \qquad (1)$$

რომელსაც **სრული ალბათობის ფორმულა** ეწოდება.

დამტკიცება. დე-მორგანის ფორმულის თანახმად

$$B = B \cap \Omega = B \cap (\bigcup_{i=1}^{n} A_i) = \bigcup_{i=1}^{n} (B \cap A_i).$$

თუ ახლა გავითვალისწინებთ, რომ ვინაიდან $A_1, A_2, ..., A_n$ ხდომილებები წყვილ-წყვილად უთავსებადია, მითუმეტეს წყვილ-წყვილად უთავსება-

დი იქნებიან ხდომილებები $B\cap A_1, B\cap A_2,..., B\cap A_n$, ამიტომ ხდომილებათა ჯამის ალბათობისა და ნამრავლის ალბათობის ფორმულების გამოყენებით გვექნება

$$P(B)=\sum_{i=1}^{n}P(B\cap A_i)=\sum_{i=1}^{n}P(A_i)P(B\,|\,A_i)\,.$$

ამოცანა 1 (ასოების ამოცნობა). გვაქვს ასოების ორი ერთობლიობა:

I={К, И, З, Н} და II={З, Ч, Н}

შემთხვევით ვირჩევთ ერთ ერთობლიობას და არჩეული ერთობლიობიდან ერთ ასოს. ამორჩეულ ასოს ძალიან მცირე დროის განმავლობაში ვუჩვენებთ დამკვირვებელს (ისე რომ მას არ შეუძლია მთლიანად აღიქვას ასო). როგორია ასოს სწორად გამოცნობის ალბათობა, თუ დამკვირვებელის პასუხია “Н”, როცა ის ასოს გამოსახულებაში დაინახავს ვერტიკალურ ხაზს და პასუხია “З”, როცა ასოს გამოსახულებაში ვერტიკალური ხაზი არ არის?

ამოხსნა. შემოვიღოთ ხდომილებები:

$$A_i=\{\text{ამორჩეულია } i-\text{ური ერთობლიობა}, \ i=1,2\};$$

“К”, “И”, “З”, “Н” და “Ч” – იყოს ხდომილება, რომ წარმოდგენილია შესაბამისად К, И, З, Н და Ч ასოები;

$$B=\{\text{დამკვირვებელმა სწორად უპასუხა}\}.$$

ამოცანის პირობებში ცხადია:

$$P(A_1)=P(A_2)=1/2\,;$$

$$P(\text{“К”}|\,A_1)=P(\text{“И”}|\,A_1)=P(\text{“З”}|\,A_1)=P(\text{“Н”}|\,A_1)=1/4,\quad P(\text{“Ч”}|\,A_1)=0;$$

$$P(\text{“З”}|\,A_2)=P(\text{“Ч”}|\,A_2)=P(\text{“Н”}|\,A_2)=1/3,\quad P(\text{“К”}|\,A_2)=P(\text{“И”}|\,A_2)=0.$$

ვინაიდან A_1 და A_2 ქმნიან ხდომილებათა სრულ სისტემას, ამიტომ თითოეული ასოს სწორად ამოცნობის ალბათობა შეგვიძლია გამოვთვალოთ სრული ალბათობის ფორმულით:

$$P(\text{“К”})=P(A_1)\,P(\text{“К”}|\,A_1)+P(A_2)\,P(\text{“К”}|\,A_2)=1/8;$$

$$P(\text{“И”})=P(A_1)\,P(\text{“И”}|\,A_1)+P(A_2)\,P(\text{“И”}|\,A_2)=1/8;$$

$$P(\text{“З”})=P(A_1)\,P(\text{“З”}|\,A_1)+P(A_2)\,P(\text{“З”}|\,A_2)=7/24;$$

$$P(\text{“Н”})=P(A_1)\,P(\text{“Н”}|\,A_1)+P(A_2)\,P(\text{“Н”}|\,A_2)=7/24;$$

$$P(\text{“Ч”})=P(A_1)\,P(\text{“Ч”}|\,A_1)+P(A_2)\,P(\text{“Ч”}|\,A_2)=1/6.$$

ამოცანის პირობებში ცხადია აგრეთვე, რომ სწორი პასუხის პირობითი ალბათობები სხვადასხვა ასოების წარმოდგენის შემთხვევაში შესაბამისად იქნება:

$P(B\,|\text{“К”})=P(B\,|\text{“И”})=P(B\,|\text{“Ч”})=0$ და $P(B\,|\text{“З”})=P(B\,|\text{“Н”})=1$.

ვინაიდან, “К”, “И”, “З”, “Н” და “Ч” აგრეთვე ხდომილებათა სრული სისტემაა, ამიტომ სრული ალბათობის ფორმულა გვაძლევს სწორი პასუხის ალბათობას:

$$P(B)=P(\text{“К”})\,P(B\,|\text{“К”})+P(\text{“Н”})\,P(B\,|\text{“Н”})+P(\text{“И”})\,P(B\,|\text{“И”})+$$
$$+P(\text{“З”})\,P(B\,|\text{“З”})+P(\text{“Ч”})\,P(B\,|\text{“Ч”})=P(\text{“Н”})+P(\text{“З”})=7/12.$$

ამოცანა 2 (მოთამაშის გაკოტრებაზე). განვიხილოთ ე. წ. “გერბი-საფასურის” თამაში: თუ მონეტის აგდებისას მოვა მოთამაშის მიერ წინასწარ დასახელებული მონეტის მხარე, მაშინ იგი იგებს 1 ლარს, წინააღმდეგ შემთხვევაში კი აგებს 1 ლარს. ვთქვათ, მოთამაშის საწყისი კაპიტალი შეადგენს x ლარს და მისი მიზანია მიიყვანოს ეს თანხა a ლარამდე. თამაში გრძელდება მანამ სანამ მოთამაშე არ მიიყვანს თავის თანხას წინასწარ განსაზღრულ a ლარამდე, ან იგი არ გაკოტრდება (ანუ წააგებს მის ხელთ არსებულ მთელ x ლარს). როგორია ალბათობა იმისა, რომ მოთამაშე გაკოტრდება?

ამოხსნა. ეს ალბათობა დამოკიდებული იქნება საწყის x კაპიტალზე. აღვნიშნოთ იგი $p(x)$ სიმბოლოთი. ცხადია, რომ იგი განმარტებულია ნებისმიერი $0 \le x \le a$ და ამასთანავე, $P(0)=1$ და $P(a)=0$. შემოვიღოთ ხდომილებები:

$$A_1 = \{\text{მოთამაშემ მოიგო პირველ ნაბიჯზე}\},$$

$$B = \{\text{მოთამაშე, რომელსაც გააჩნია საწყისი კაპიტალი } x, \text{გაკოტრდება}\}.$$

ამოცანის პირობებში გვაქვს:

$$P(A_1) = P(\overline{A_1}) = 1/2\,,\ \ P(B\,|\,A_1) = p(x+1) \ \ \text{და} \ \ P(B\,|\,\overline{A_1}) = p(x-1) \ \ (1 \le x \le a-1).$$

ვინაიდან, A_1 და $\overline{A_1}$ ხდომილებათა სრული სისტემაა, ამიტომ სრული ალბათობის ფორმულა $p(x)$ ალბათობისათვის გვაძლევს შემდეგ განტოლებას:

$$p(x) = \frac{1}{2}\,p(x+1) + \frac{1}{2}\,p(x-1)\,,\ \ 1 \le x \le a-1$$

(ამ ტიპის განტოლებებს მათემატიკაში *რეკურენტულ განტოლებებს* უწოდებენ). შეიძლება შემოწმდეს, რომ ამ განტოლების ამოხსნას აქვს სახე:

$$p(x) = bx + c\,,$$

სადაც b და c -- ნებისმიერი მუდმივებია. ამ კოეფიციენტების მოსაძებნად უნდა ვისარგებლოთ სასაზღვრო პირობებით $P(0)=1$ და $P(a)=0$. მაშინ მივიღებთ, რომ

$$c = 1 \ \ \text{და} \ \ ab + c = 0\,,$$

საიდანაც, $b = -1/a$ და საბოლოოდ $p(x) = 1 - x/a$, $0 \le x \le a$.

განვიხილოთ რეალური სიტუაცია, რომელიც გვიჩვენებს ერთი შეხედვით მოულოდნელ განსხვავებას $P(A\,|\,B)$ და $P(B\,|\,A)$ პირობით ალბათობებს შორის. იმისათვის, რომ გამოვავლინოთ სერიოზული ავადმყოფობის მქონე ადამიანები ადრეულ სტადიაზე, ხდება ადამიანების დიდი ჯგუფის ტესტირება. მიუხედავად წინასწარი შემოწმების სარგებლობისა, ამ მიდგომას გააჩნია უარყოფითი მხარე: თუ ადამიანს სინამდვილეში არ გააჩნია ავადმყოფობა და საწყისმა ტესტმა აჩვენა დადებითი შედეგი (დაუდგინა ავადმყოფობა), ის იქნება სტრესულ მდგომარეობაში (რაც თავის მხრივ უარყოფითად მოქმედებს მის ცხოვრებაზე) სანამ უფრო წარმატებული ტესტი არ აჩვენებს, რომ ის ჯანმრთელია. ამ პრობლემის მნიშვნელობა შესაძლებელია კარგად გავიგოთ პირობითი ალბათობების ტერმინებში.

ორი ხდომილების ნამრავლის ალბათობის ფორმულაში მოყვანილი დავალება 2-ის მონაცემებში გამოვთვალოთ ალბათობა იმისა, რომ ტესტი აჩვენებს დადებით შედეგს. სრული ალბათობის ფორმულის თანახმად:

$$P(\text{დადებითი}) = P(\text{ჯანმრთელი})P(\text{დადებითი} \mid \text{ჯანმრთელი}) +$$

$$+P(\text{ავადმყოფი})P(\text{დადებითი} \mid \text{ავადმყოფი}) = 0.99 \cdot 0.01 + 0.01 \cdot 0.99 = 0.0198\,.$$

როგორც ცნობილია, მაგალითის პირობებში

$$P(\text{დადებითი} \mid \text{ავადმყოფი}) = 99\%\,.$$

გამოვთვალოთ ახლა შებრუნებული პირობითი ალბათობა, რისთვის-აც ვისარგებლოთ პირობითი ალბათობის განმარტებითა და ნამრავლის ალბათობის ფორმულებით. მაშინ ზემოთ მიღებული

$$P(\text{დადებითი}) = 0.0198 = 1.98\%$$

შედეგის თანახმად:

$$P(\text{ავადმყოფი} \mid \text{დადებითი}) = \frac{P(\text{ავადმყოფი} \cap \text{დადებითი})}{P(\text{დადებითი})} =$$

$$= \frac{P(\text{ავადმყოფი})\ P(\text{დადებითი} \mid \text{ავადმყოფი})}{1.98\%} = \frac{1\% \cdot 99\%}{1.98\%} = 50\%\,.$$

როგორც ვხედავთ, პირობითი ალბათობა იმისა რომ ტესტი მოგვც-ემს დადებით შედეგს, პირობაში რომ ადამიანი ავადმყოფია ტოლია 99%-ის, მაშინ როდესაც პირობითი ალბათობა იმისა რომ ადამიანი ავადმყო-ფია, პირობაში რომ ტესტმა მოგვცა დადებითი შედეგი არის მხოლოდ 50%. აქ შერჩეული მონაცემების შემთხვევაში უკანასკნელი შედეგი შეიძლ-ება ჩაითვალოს მიუღებელად: ნახევარი ადამიანების, რომელთა ტესტირებ-ამ აჩვენა დადებითი შედეგი, ფაქტიურად არის ჯანმრთელი.

§13. ბაიესის ფორმულა

ვიგულისხმოთ, რომ A და B ხდომილებები ისეთია, რომ $P(A)>0$ და $P(B)>0$. მაშინ $P(A|B)$ და $P(B|A)$ პირობით ალბათობების განმარტებიდან:

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

საიდანაც მიიღება ე. წ. *ბაიესის ფორმულა:*

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

თუ $A_1,A_2,...,A_n$ ხდომილებათა სრული სისტემაა ისეთი, რომ $P(A_i)>0,\ i=1,2,...,n$, მაშინ ბაიესის ფორმულიდან სრული ალბათობის ფორმულის გამოყენებით ვღებულობთ ე. წ. **ბაიესის თეორემას:**

$$P(A_i\,|\,B)=\frac{P(A_i)P(B\,|\,A_i)}{\sum_{j=1}^{n}P(A_j)P(B\,|\,A_j)}.$$

შევნიშნოთ, რომ ორივე ამ ფორმულაში ერთი პირობითი ალბათობა იცვლება შებრუნებული პირობითი ალბათობებით, რომლებიც ხშირ შემთხვევაში შედარებით მარტივად გამოითვლება (ან პირდაპირ მოცემულია) და მათი კომბინაციით ითვლება პირდაპირი პირობითი ალბათობა.

ბაიესის ფორმულას შეიძლება მიეცეს შემდეგი ინტერპრეტაცია: დავუშვათ, სამეცნიერო გამოკვლევის დაწყებამდე ჩვენ გვაქვს n სხვადასხვა ვარაუდი (ჰიპოთეზა) $A_1,A_2,...,A_n$ შესასწავლი ობიექტის ბუნების შესახებ, ამასთანავე ჩვენ მათ მივაწერთ ალბათობებს $P(A_1),P(A_2),...,P(A_n)$ (ამ ალბათობებს უწოდებენ **აპრიორულ ალბათობებს**). შემდეგ ჩვენ ვატარებთ ექსპერიმენტს (ან დაკვირვებას), რომლის შედეგადაც შეიძლება მოხდეს ან არ მოხდეს B ხდომილება (ე. ი. მოხდეს $\overline{B}$ ხდომილება). თუ მოხდა B ხდომილება, ვახდენთ თითოეული ჰიპოთეზის სამართლიანობის შესახებ ჩვენი რწმენის გადაფასებას ვცვლით რა $P(A_i)$ ალბათობებს $P(A_i\,|\,B)$ ალბათობებით (ამ ალბათობებს ეწოდება **აპოსტერიორული ალბათობები**). ასე ჩვენ ვაგრძელებთ, სანამ რომელიმე $i=i_0$-სათვის A_{i_0} ხდომილების აპოსტერიორული ალბათობა არ გახდება თითქმის ერთის ტოლი. მაშინ A_{i_0} ჰიპოთეზა ფაქტიურად სამართლიანია. თუ კი გადაწყვეტილების მიღება საჭიროა N ექსპერიმენტის ჩატარების შემდეგ, ხოლო ამ მომენტისათვის აპოსტერიორული აბათობებიდან არც ერთი არ არის ერთთან საკმაოდ ახლოს, მაშინ გადაწყვეტილება მიიღება იმ ჰიპოთეზის სასარგებლოდ, რომლის აპოსტერიორული ალბათობაც მაქსიმალურია.

მოკლედ, რომ ვთქვათ: სტატისტიკურ გამოყენებებში $A_1,A_2,...,A_n$ ხდომილებებს, რომლებიც ქმნიან ხდომილებათა სრულ სისტემას, ხშირად "ჰიპოთეზებს" უწოდებენ, $P(A_i)$ -- ალბათობებს A_i ხდომილებების აპრიორულ (ცდამდე) ალბათობებს. პირობით ალბათობას $P(A_i\,|\,B)$ კი ეძლევა B

ხდომილების მოხდენის შემდეგ A_i ჰიპოთეზის აპოსტერიორული (ცდის შემდგომი) ალბათობის ინტერპრეტაცია.

ამოცანა 1. ყუთში მოთავსებულია ორი მონეტა: A_1 -- სიმეტრიული მონეტა გერბის მოსვლის ალბათობით 1/2, და A_2 -- არასიმეტრიული მონეტა გერბის მოსვლის ალბათობით 1/3. შემთხვევით ვიღებთ ერთ მონეტას და ვაგდებთ. დავუშვათ, რომ მოვიდა გერბი. როგორია ალბათობა იმისა, რომ ამოღებული მონეტა იყო სიმეტრიული?

ამოხსნა. ამ შემთხვევაში ელემენტარულ ხდომილებათა სივრცე იქნება:

$$\Omega = \{(A_1, \text{გ}), (A_1, \text{ს}), (A_2 . \text{გ}), (A_2, \text{ს})\},$$

სადაც მაგალითად, $(A_1, \text{გ})$ -- ნიშნავს, რომ ამოვიღეთ A_1 მონეტა და მისი აგდების შედეგად მოვიდა გერბი. ამოცანის პირობებში გვაქვს:

$$P(A_1) = P(A_2) = 1/2, \quad P(\text{გ} \mid A_1) = 1/2 \;\; \text{და} \;\; P(\text{გ} \mid A_2) = 1/3.$$

შესაბამისად, ნამრავლის ალბათობის ფორმულის გამოყენებით გამოვითვლით:

$$P\{(A_1, \text{გ})\} = 1/4, \;\; P\{(A_1, \text{ს})\} = 1/4, \;\; P\{(A_2, \text{გ})\} = 1/6 \;\; \text{და} \;\; P\{(A_2, \text{ს})\} = 1/3.$$

ამიტომ ბაიესის ფორმულის თანახმად

$$P(A_1 \mid \text{გ}) = \frac{P(A_1)P(\text{გ} \mid A_1)}{P(A_1)P(\text{გ} \mid A_1) + P(A_2)P(\text{გ} \mid A_2)} = \frac{3}{5}.$$

ცხადია, აგრეთვე რომ $P(A_2 \mid \text{გ}) = 2/5$.

ამოცანა 2 (კეთილ გამომცდელზე I). ვთქვათ, ჩვენ ჩასაბარებელი გვაქვს გამოცდა და შეგვიძლია ავირჩიოთ ნებისმიერი სამი გამომცდელიდან. დავუშვათ, ჩვენთვის ცნობილია, რომ ერთერთი სამი გამომცდელიდან (უცნობია რომელი) -- "კეთილია" და ალბათობა იმისა, რომ მასთან ჩააბარო გამოცდა ტოლია 0,4-ის, ხოლო დანარჩენი ორი გამომცდელი "ავია" და მათთან გამოცდის ჩაბარების ალბათობა ტოლია 0,1-ის. ჩვენ შემთხვევით ავირჩიეთ გამომცდელი და წარმატებით ჩავაბარეთ გამოცდა. როგორია ლაბათობა იმისა, რომ ჩვენ ავირჩიეთ "კეთილი" გამომცდელი?

ამოხსნა. შემოვიღოთ შემდეგი ხდომილებები: A -- ამორჩეული გამომცდელი "კეთილია" (მაშინ $\overline{A}$ -- იქნება ხდომილება, რომ ამორჩეული გამომცდელი "ავია") და B -- გამოცდა ჩაბარებულია (შესაბამისად, $\overline{B}$ -- გამოცდა არაა ჩაბარებული). ამოცანის პირობებში გვაქვს:

$$P(A) = 1/3, \qquad P(\overline{A}) = 1 - P(A) = 2/3;$$

$$P(B \mid A) = 0{,}4, \qquad P(\overline{B} \mid A) = 1 - P(B \mid A) = 0{,}6;$$

$$P(B \mid \overline{A}) = 0{,}1, \qquad P(\overline{B} \mid \overline{A}) = 1 - P(B \mid \overline{A}) = 0{,}9.$$

ცნობილია, რომ მოხდა B ხდომილება და გამოსათვლელია პირობითი ალბათობა $P(A|B)$. ვინაიდან, A და $\overline{A}$ ხდომილებები ქმნიან სრულ სისტემას, ბაიესის ფორმულის თანახმად საძიებელი ალბათობა იქნება:

$$P(A \mid B) = \frac{P(A)P(B \mid A)}{P(A)P(B \mid A) + P(\overline{A})P(B \mid \overline{A})} = \frac{2}{3}.$$

დავალება. ამოცანა 2-ის პირობებში გამოთვალეთ ალბათობა იმისა, რომ არჩეულ იქნა “ავი” გამომცდელი, თუ ცნობილია, რომ გამოცდა ჩაბარებულ იქნა წარმატებით?

ამოცანა 3 (კეთილ გამომცდელზე II). დავუშვათ, რომ გამომცდელთან, რომელთანაც წარმატებით ჩაიარა გამოცდამ (იხ. ამოცანა 2) გამოსაცდელად რიგ-რიგობით მივიდა კიდევ ორი მოსწავლე. ჯერ გამოცდა ვერ ჩააბარა მეორე მოსწავლემ, შემდეგ მივიდა მესამე და მანაც ვერ ჩააბარა გამოცდა. ამ ფაქტის შემდეგ რომელი ჰიპოთეზაა უფრო დასაჯერებელი: ეს გამომცდელი “კეთილია” თუ “ავი”?

ამოხსნა. ავღნიშნოთ $P_i(A)$ (შესაბამისად, $P_i(\overline{A})$) სიმბოლოთი ალბათობა (აპოსტერიორული) იმისა, რომ ეს გამომცდელი “კეთილია” (შესაბამისად, “ავია”) მას შემდეგ რაც გამოცდილ იქნა i-ური სტუდენტი, $i=1,2,3$. ჩვენ უკვე დავადგინეთ, რომ $P_1(A)=2/3$. შესაბამისად,

$$P_1(\overline{A})=1-P_1(A)=1/3.$$

მეორე მოსწავლის თვალსაზრისით ეს ალბათობები წარმოადგენენ ორი შესაძლო ჰიპოთეზის აპრიორულ ალბათობებს. ამიტომ, ბაიესის ფორმულის თანახმად, მეორე სტუდენტის ჩაჭრის შემდეგ აპოსტერიორული ალბათობები იქნება:

$$P_2(A)=\frac{P(\overline{B}\,|\,A)P_1(A)}{P(\overline{B}\,|\,A)P_1(A)+P(\overline{B}\,|\,\overline{A})P_1(\overline{A})}=\frac{4}{7} \quad \text{და} \quad P_2(\overline{A})=1-P_2(A)=\frac{3}{7}.$$

ანალოგიურად, ახლა მიღებული ალბათობები უკვე იქნება აპრიორული ალბათობები მესამე მოსწავლისათვის, და ამიტომ საძიებელი აპოსტერიორული ალბათობები, მას შემდეგ რაც მესამე მოსწავლემ ვერ ჩააბარა გამოცდა, გამოითვლება ისევ ბაიესის ფორმულით:

$$P_3(A)=\frac{P(\overline{B}\,|\,A)P_2(A)}{P(\overline{B}\,|\,A)P_2(A)+P(\overline{B}\,|\,\overline{A})P_2(\overline{A})}=\frac{8}{17} \quad \text{და} \quad P_3(\overline{A})=1-P_3(A)=\frac{9}{17}>P_3(A).$$

როგორც ვხედავთ, ექსპერიმენტის (გამოცდის) დაწყების წინ აპრიორული ალბათობა იმისა. რომ არჩეული გამომცდელი “კეთილია” ტოლი იყო 1/3-ის. ექსპერიმენტების შემდეგ ამ ხდომილების აპოსტერიორული ალბათობა გაიზარდა და გახდა 8/17. მიუხედავად ამისა, თუ სამი ექსპერიმენტის შემდეგ მისაღებია გადაწყვეტილება ამ გამომცდელის შესახებ, მაშინ უფრო სარწმუნოა ჩავთვალოთ იგი “ავად” (ვინაიდან, $P_3(\overline{A})>P_3(A)$).

§14. განმეორებითი ცდები. ბერნულის ფორმულა

განვიხილოთ ერთი და იგივე ექსპერიმენტების სერია, რომლებიც ტარდება ერთი და იგივე პირობებში ერთმანეთისაგან დამოუკიდებლად (ნებისმიერი ექსპერიმენტის შედეგი დამოუკიდებელია დანარჩენი ექსპერიმენტების შედეგებისაგან). ამასთანავე, ყოველ კონკრეტულ ექსპერიმენტში (ელემენტარული ხდომილებების როლში) ჩვენ განვასხვავებთ მხოლოდ ორ შედეგს: გარკვეული A ხდომილების მოხდენა (რომელსაც პირობითად *"წარმატებას"* უწოდებენ) და მისი არ მოხდენა -- $\overline{A}$ (ე. ი. A ხდომილების საწინააღმდეგო ხდომილების მოხდენა, რომელსაც *"მარცხს"* უწოდებენ), ასე რომ $A+\overline{A}=\Omega$. A ხდომილების მოხდენის ალბათობა ნებისმიერი ექსპერიმენტისათვის მუდმივია და ტოლია $P(A)=p$, სადაც $0<p<1$. შესაბამისად, $P(\overline{A})=1-P(A)=1-p:=q$ $(p+q=1)$.

დავუშვათ, ჩატარდა n დამოუკიდებელი ექსპერიმენტი, რომელსაც ჩვენ განვიხილავთ როგორც ერთ რთულ ექსპერიმენტს. ყოველი ექსპერიმენტის შედეგს ჩვენ წარმოვადგენთ n-ეულების სახით, სადაც თითოეულ ადგილზე დავწერთ ან A-ს ან $\overline{A}$-ს იმის მიხედვით მოხდა A თუ $\overline{A}$. მაგალითად, ორი ექსპერიმენტის შემთხვევაში შესაძლებელია $2^2=4$ შედეგი: $AA, A\overline{A}, \overline{A}A, \overline{A}\overline{A}$ (A ხდომილება მოხდა ორჯერ, A ხდომილება მოხდა პირველ და არ მოხდა მეორე ექსპერიმენტში, A ხდომილება არ მოხდა პირველ და მოხდა მეორე ექსპერიმენტში, A ხდომილება არ მოხდა ორჯერ). სამი ექსპერიმენტის შემთხვევაში მოსალოდნელია $2^3=8$ შედეგი:

$$AAA, AA\overline{A}, A\overline{A}A, \overline{A}AA, A\overline{A}\overline{A}, \overline{A}A\overline{A}, \overline{A}\overline{A}A, \overline{A}\overline{A}\overline{A}.$$

და ა. შ. n ექსპერიმენტის ყველა შესაძლო შედეგს (სულ იქნება 2^n შედეგი) შეესაბამება n ასოს მიმდევრობა A, $\overline{A}$ იმ რიგით რა მიმდევრობითაც შეგვხვდება ეს ხდომილებები n ექსპერიმენტში, მაგალითად, $A\overline{A}\overline{A}A\cdots\overline{A}$.

ვინაიდან ექსპერიმენტები დამოუკიდებელია, ამიტომ n ექსპერიმენტის თითეული შესაძლო შედეგის ალბათობა გამოითვლება შესაბამის ექსპერიმენტებში A და $\overline{A}$ ხდომილებების ალბათობების გადამრავლებით. ასე მაგალითად, ზემოთ დაწერილი შედეგისათვის (იმის გათვალისწინებით, რომ ყოველ ექსპერიმენტში $P(A)=p$ და $P(\overline{A})=q$) მივიღებთ ალბათობას:

$$P(A)P(\overline{A})P(\overline{A})P(A)\cdots P(\overline{A})=pqqp\cdots q.$$

ცხადია, რომ თუ დაწერილ შედეგში ასო A შეგვხვდა x, და შესაბამისად, ასო $\overline{A}$ გვხვდება $(n-x)$-ჯერ, მაშინ ასეთი შედეგის ალბათობა იქნება: $p^x q^{n-x}$, დამოუკიდებლად იმისგან რა თანმიმდევრობითაა განლაგებული n-ეულში x ასო A და $n-x$ ასო $\overline{A}$. სამი ექსპერიმენტის რვა შესაძლო შედეგისათვის ამ გზით დათვლილი ალბათობები იქნება:

$$P(AAA)=p^3,\ P(AA\overline{A})=P(A\overline{A}A)=P(\overline{A}AA)=p^2q,$$

$$P(A\overline{A}\overline{A})=P(\overline{A}A\overline{A})=P(\overline{A}\overline{A}A)=pq^2 \ \text{ და } \ P(\overline{A}\overline{A}\overline{A})=q^3.$$

ავღნიშნოთ $P_3(i)$ სიმბოლოთი ალბათობა იმისა, რომ სამ ექსპერიმენტში A ხდომილება შეგვხვდა (მოხდა) ზუსტად i-ჯერ. მაშინ შედეგს -- სამ ექსპერიმენტში A ხდომილება არც ერთხელ არ შეგვხვდა (A ხდომილება მოხდა 0-ჯერ) აქვს ალბათობა $P_3(0)=P(\overline{A}\overline{A}\overline{A})=q^3$. A ხდომილება მოხდა ზუსტად ერთჯერ განხორციელდება თუ მოხდა რომელიმე შემდეგი სამი ვარიანტიდან: $AA\overline{A}$ ან $A\overline{A}A$ ან $\overline{A}AA$, რომელთაგან თითეულის ალბათობაა pq^2. ამიტომ ალბათობათა ჯამის კანონის თანახმად:

$$P_3(1)=P(A\overline{A}\overline{A})+P(\overline{A}A\overline{A})+P(\overline{A}\overline{A}A)=3q^2p\,.$$

ანალოგიურად,

$$P_3(2)=P(AA\overline{A})+P(A\overline{A}A)+P(\overline{A}AA)=3qp^2\,.$$

და, ბოლოს, $P_3(3)=P(AAA)=p^3$.

ვნახოთ, რისი ტოლია ამ ალბათობების ჯამი. გვაქვს:

$$P_3(0)+P_3(1)+P_3(2)+P_3(3)=q^3+3q^2p+3qp^2+p^3=(q+p)^3=1^3=1,$$

რაც ბუნებრივია ასეც უნდა ყოფილიყო, ვინაიდან ჩვენ განვიხილეთ ალბათობების ჯამი იმ ხდომილებების, რომლებიც ქმნიან ხდომილებათა სრულ სისტემას: A ხდომილება სამ ექსპერიმენტში აუცილებლად მოხდება ან 0-ჯერ, ან 1-ჯერ, ან 2-ჯერ, ან 3-ჯერ.

თუ ახლა $P_n(x)$ სიმბოლოთი ავღნიშნავთ ალბათობას იმისა, რომ n ექსპერიმენტში A ხდომილება (წარმატება) შეგვხვდა (მოხდა) x-ჯერ, მაშინ ანალოგიური მსჯელობით, მივიღებთ ე. წ. **ბერნულის ფორმულას:**

$$P_n(x)=C_n^x p^x q^{n-x}=\frac{n!}{x!(n-x)!}p^x q^{n-x}=\frac{n(n-1)\cdots(n-x+1)}{x!}p^x q^{n-x}\,. \qquad (1)$$

მართლაც, n ექსპერიმენტის ისეთ შედეგთა რაოდენობა, რომლებიც ჩაიწერებიან x ასო A და $n-x$ ასო $\overline{A}$-ს სხვადასხვა კომბინაციით, ტოლი იქნება ჯუფდებათა რიცხვის n-დან x, ვინაიდან ნებისმიერი ასეთი n-ეული სავსებით განისაზღვრება, თუ n ადგილიდან ამოვარჩევთ ზუსტად x ადგილს ასო A-სათვის, ხოლო დანარჩენ $n-x$ ადგილს დავტოვებთ ასო $\overline{A}$-სათვის. მაგრამ x ნომერის ამორჩევა n ადგილიდან შესაძლებელია სწორედ C_n^x სხვადასხვა გზით, რადგანაც ჯგუფები შედგენილი ნომრებისაგან, რიგითობისაგან დამოუკიდებლად, უნდა განსხვავდებოდნენ ერთი მაინც ელემენტით.

მაგალითი 1. ყუთში 3 თეთრი და 5 შავი ბურთია. ყუთიდან შემთხვევით დაბრუნებით იღებენ 4 ბურთს. იპოვეთ ალბათობა იმისა, რომ ამოღებული ბურთებიდან თეთრი ბურთების რაოდენობა მეტი იქნება შავი ბურთების რაოდენობაზე?

ამოხსნა. თუ წარმატებად ჩავთვლით თეთრი ბურთის ამოღებას, მაშინ პირობის თანახმად ერთ ცდაში წარმატების ალბათობა იქნება $p=3/8$. საძიებელი ხდომილების ხელშემწყობ უთავსებად შემთხვევებს წარმოადგენს ოთხ ცდაში 3 ან 4 თეთრი ბურთის ამოღება, რომელთა ალბათობები ბერნულის ფორმულის თანახმად შესაბამისად ტოლია:

$$P_4(3) = C_4^3 \cdot (3/8)^3 \cdot (1-3/8)^{4-3} = 135/1024 \quad \text{და}$$
$$P_4(4) = C_4^4 \cdot (3/8)^4 \cdot (5/8)^0 = 81/4096.$$
შესაბამისად, საძიებელი ალბათობა ალბათობათა შეკრების კანონის თანახმად იქნება:

$$135/1024 + 81/4096 = 621/4096.$$

ალბათობების ერთობლიობას $P_n(x)$, როცა $x = 0,1,...,n$, ე. ი ($P_n(0)$, $P_n(1),...,P_n(n)$ ალბათობებს) ეწოდება **ალბათობების ბინომიალური განაწილება.** რადგანაც ეს ალბათობები შეესაბამება უთავსებად ხდომილებებს, რომლებიც ქმნიან სრულ სისტემას, ამიტომ გასაგებია, რომ:

$$\sum_{x=0}^{n} P_n(x) = 1,$$

რაც, მეორეს მხრივ, ადვილად მოწმდება ნიუტონის ბინომის ფორმულის გამოყენებითაც, რადგანაც ბერნულის ფორმულაში მონაწილეობენ სწორედ $(q+p)^n$ ნიუტონის ბინომის კოეფიციენტები (აქედან მოდის სახელწოდებაც: "ბინომიალური განაწილება"):

$$\sum_{x=0}^{n} P_n(x) = \sum_{x=0}^{n} C_n^x p^x q^{n-x} = (q+p)^n = 1^n = 1.$$

ხშირ შემთხვევაში საჭიროა გამოითვალოს ალბათობა იმისა, რომ A ხდომილება n ექსპერიმენტში შეგვხვდება არაუმეტეს x-ჯერ. ამ ალბათობას უწოდებენ ბინომიალური განაწილების **კუმულატიურ ანუ დაგროვილ ალბათობას.** აღნიშნოთ იგი $\overline{P}_n(x)$ სიმბოლოთი. მაშინ ალბათობათა შეკრების კანონის თანახმად კუმულატიური ალბათობა ასე გამოითვლება:

$$\overline{P}_n(x) = P_n(0) + P_n(1) + \cdots + P_n(x) = \sum_{i=0}^{x} P_n(i).$$

თუ ექსპერიმენტების n რიცხვი საკმაოდ დიდია, მაშინ $P_n(x)$ და $\overline{P}_n(x)$ ალბათობების გამოთვლა ხდება სპეციალური "ასიმპტოტური" ფორმულებით. თუ n მცირეა, მაშინ შეგვიძლია გამოვიყენოთ მარტივი თანაფარდობა, რომელიც აკავშირებს ბინომური განაწილების ორ მომდევნო $P_n(x)$ და $P_n(x+1)$ წევრს:

$$\frac{P_n(x+1)}{P_n(x)} = \frac{(n-x)p}{(x+1)q}. \tag{2}$$

თუ ნაპოვნია $P_n(x)$, მაშინ უკანასკნელი თანაფარდობიდან ადვილად გადავითვლით $P_n(x+1)$-ს.

დამოუკიდებელი ექსპერიმენტების სერიის ზემოთაღწერილი სქემა პირველად განხილული და შესწავლილი იყო შვეიცარიელი მათემატიკოსის იაკობ ბერნულის (1654-1705) მიერ და ამიტომ იგი ატარებს **ბერნულის სქემის** სახელს.

მაგალითი 2. დავუშვათ ვამოწმებთ დეფექტურობაზე საქონლის პარტიას, რომელიც შედგება 30 ნაწარმისაგან. ცნობილია, რომ დეფექტური პროდუქციის წილი შეადგენს 5%-ს. როგორია საქონლის ამ პარტიაში დეფექტური პროდუქციის ამა თუ იმ რიცხვის აღმოჩენის ალბათობები?

ამოხსნა. ამ შემთხვევაში ექსპერიმენტების რიცხვია $n=30$, ხოლო ალბათობის კლასიკური განმარტების თანახმად $p=5/100=0.05$ (შესაბამისად, $q=0.95$). ვისარგებლოთ (1) და (2) ფორმულებით. გვაქვს:

$$P_{30}(0)=C_{30}^{0}0.05^{0}0.95^{30-0}=0.95^{30}=0.2146,$$

გარდა ამისა,

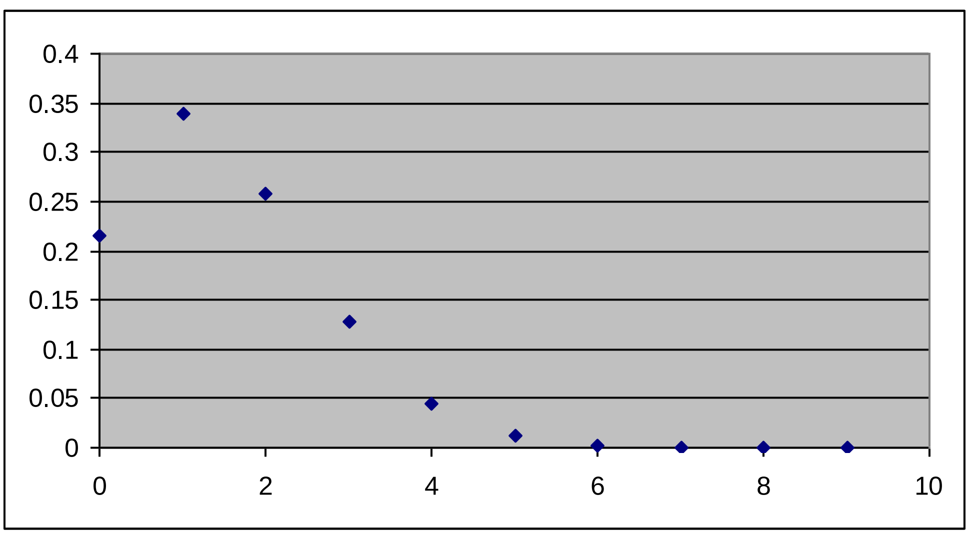
$$P_{30}(x+1)=\frac{30-x}{x+1}\frac{p}{q}P_{30}(x)=\frac{30-x}{x+1}\frac{0,05}{0,95}P_{30}(x)=\frac{30-x}{19(x+1)}P_{30}(x),$$

საიდანაც როცა $x=0$: $P_{30}(0+1)=\dfrac{30-0}{19\cdot(0+1)}\cdot 0.2146=0.3389$;

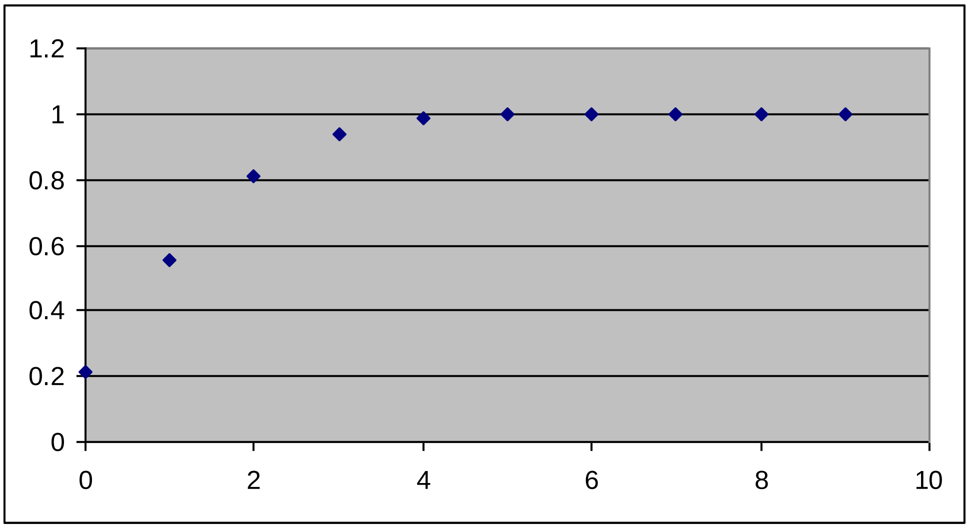
როცა $x=1$: $P_{30}(1+1)=\dfrac{30-1}{19\cdot(1+1)}\cdot 0.3389=0.2586$ და ა. შ. საბოლოოდ გვექნება შემდეგი ცხრილი:

დეფექტური ნაწარმის რიცხვი x	ალბათობა $P_n(x)$	კუმულატიური ალბათობა $\overline{P}_n(x)$
0	0.2146	0.2146
1	0.3389	0.5535
2	0.2586	0.8122
3	0.1270	0.9392
4	0.0451	0.9844
5	0.0124	0.9967
6	0.0027	0.9994
7	0.0005	0.9999
8	0.0001	0.999998
9	0.000001	0.999999

ამ ცხრილის შესაბამისი ალბათობების განაწილების გრაფიკი იქნება:



კუმულატიური ალბათობების შესაბამისი განაწილების გრაფიკი იქნება:



დავალება 1. ექსპერიმენტი მდგომარეობს სამი სათამაშო კამათლის გაგორებაში. იპოვეთ ალაბათობა იმისა, რომ ექსპერიმენტის 10-ჯერ გამეორებისას, ზუსტად 4 ექსპერიმენტში მოვა ზუსტად ორ-ორი “6”?

დავალება 2. რამდენი შემთხვევითი ციფრი უნდა ავიღოთ, რომ ციფრი “5” მოვიდეს ერთჯერ მაინც არანაკლებ 0.9-ის ტოლი ალბათობით?

განმარტება. ისეთ k_0 რიცხვს, რომლის შესაბამისი ალბათობა $P_n(k_0)$ უდიდესია $P_n(0)$, $P_n(1),...,P_n(n)$ ალბათობებს შორის **უალბათესი რიცხვი** ეწოდება.

უალბათესი რიცხვი გვიჩვენებს n დამოუკიდებელ ცდაში წარმატებათა რა რაოდენობაა ყველაზე უფრო მოსალოდნელი.

განვიხილოთ ფარდობა

$$a_k = P_n(k+1) \big/ P_n(k) = (n-k)p/[(k+1)q].$$

ადვილი დასანახია, რომ $a_k > 1$, როცა $k < np - q$; $a_k = 1$, როცა $k = np - q$ და $a_k < 1$, როცა $k > np - q$. საიდანაც ცხადია, რომ უალბათესი რიცხვი წარმოადგენს შემდეგი უტოლობის მთელ ამონახსნს:

$$np - q \le k_0 \le np + p.$$

უალბათესი რიცხვი შეიძლება იყოს ერთი ან ორი, იმის მიხედვით, ამ უტოლობის საზღვრები მთელი რიცხვებია, თუ ათწილადი. მაგალითად, მონეტის 11-ჯერ აგდებისას გერბის მოსვლის უალბათესი რიცხვი იქნება

$$11 \cdot 1/2 - 1/2 \le k_0 \le 11 \cdot 1/2 + 1/2 \Leftrightarrow 5 \le k_0 \le 6$$

უტოლობის მთელი ამონხსნი, ანუ 5 და 6. რაც იმას ნიშნავს, რომ მონეტის 11-ჯერ აგდებისას გერბის 5-ჯერ და 6-ჯერ მოსვლის ალბათობები ერთმანეთის ტოლია და ყველა დანარჩენ ალბათობებზე მეტი.

§15. პუასონის ფორმულა

გამოთვლების ჩატარება ბერნულის ფორმულის გამოყენებით, ცდათა დიდი რიცხვის შემთხვევაში, მოითხოვს ძალიან დიდ ძალისხმევას. მიახლოებითი გამოთვლების ჩასატარებლად შესაძლებელია უფრო მოხერხებული ფორმულის მიღება, თუ კი ცდათა დიდი რიცხვის შემთხვევაში ცალკეულ ცდაში A ხდომილების მოხდენის p ალბათობა მცირეა, ხოლო ნამრავლი $np=\lambda$ ინარჩუნებს მუდმივ მნიშვნელობას ექსპერიმენტების სხვადასხვა სერიაში (ანუ A ხდომილების მოხდენის საშუალო რიცხვი უცვლელი რჩება ექსპერიმენტების სხვადასხვა სერიაში). ბერნულის ფორმულა შეგვიძლია გადავწეროთ შემდეგი სახით:

$$p_n(k)=\frac{n(n-1)(n-2)...(n-k+1)}{k!}p^k(1-p)^{n-k}=\frac{n(n-1)...(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}.$$

გამოვთვალოთ მიღებული გამოსაულების ზღვარი, როცა $p\to 0$ და $n\to\infty$, ისე რომ $np\to\lambda$. ადვილი დასანახია, რომ:

$$P(k)=\frac{\lambda^k}{k!}\lim_{n\to\infty}(1\cdot(1-\frac{1}{n})\cdots(1-\frac{k-1}{n})(1-\frac{\lambda}{n})^{n-k})=\frac{\lambda^k}{k!}\lim_{n\to\infty}(1-\frac{\lambda}{n})^n(1-\frac{\lambda}{n})^{-k}=\frac{\lambda^k}{k!}\cdot e^{-\lambda}\cdot 1.$$

მიღებულ ფორმულას

$$\lim_{np\to\lambda}p_n(k)=\frac{\lambda^k e^{-\lambda}}{k!}$$

პუასონის ფორმულა ეწოდება. იგი საშუალებას იძლევა ვიპოვოთ n დამოუკიდებელ ცდაში A ხდომილების k-ჯერ მოხდენის ალბათობა (როცა n საკმაოდ დიდია, ხოლო p საკმაოდ მცირე, ამასთანავე $np=\lambda<15$) პუასონის მიახლოებითი ფორმულით:

$$p_n(k)\approx\frac{\lambda^k e^{-\lambda}}{k!}$$

აღსანიშნავია, რომ პუასონის ფორმულით სარგებლობისას, განსხვავებით ბერნულის ფორმულის შემთხვევისაგან, ჩვენ არ გვჭირდება მის გამოსახულებაში სიდიდეების (მონაცემების) შეტანა კონკრეტული ამოცანიდან, არამედ უბრალოდ ვსარგებლობთ პუასონის განაწილების ცხრილებით. ქვემოთ მოყვანილია ამ ცხრილის ერთი ფრაგმენტი:

	λ = 0.1	λ = 0.2	λ = 0.3	λ = 0.4	λ = 0.5	λ = 0.6	λ = 0.7	λ = 0.8	λ = 0.9
p(0)	**0.9048**	**0.8187**	**0.7408**	**0.6703**	**0.6065**	**0.5488**	**0.4966**	**0.4493**	**0.4066**
p(1)	0.0905	0.1637	0.2222	0.2681	0.3033	0.3293	0.3476	0.3595	0.3659
p(2)	0.0045	0.0164	0.0333	0.0536	0.0758	0.0988	0.1217	0.1438	0.1647
p(3)	0.0002	0.0011	0.0033	0.0072	0.0126	0.0198	0.0284	0.0383	0.0494
p(4)		0.0001	0.0003	0.0007	0.0016	0.0030	0.0050	0.0077	0.0111
p(5)				0.0001	0.0002	0.0004	0.0007	0.0012	0.0020
p(6)							0.0001	0.0002	0.0003

§16. შემთხვევითი სიდიდე. განაწილების კანონი

ალბათობის თეორიაში შემთხვევითი ხდომილების ცნებასთან ერთად გამოიყენება გარკვეული აზრით უფრო მოხერხებული *შემთხვევითი სიდიდის* ცნება. ცვლად სიდიდეს, რომლის მნიშვნელობები დამოკიდებულია შემთხვევითი ექსპერიმენტის ან მოვლენის შესაძლო შედეგებზე, შემთხვევით სიდიდეს უწოდებენ. შემთხვევითი სიდიდის მაგალითებია: სათამაშო კამათლის გაგორებისას მოსულ ქულათა რიცხვი; მონეტის განმეორებითი აგდებისას მონეტის რომელიმე მხარის გამოჩენათა რიცხვი; გასროლათა რაოდენობა მიზანში პირველად მოხვედრამდე; მანძილი სამიზნის ცენტრიდან დაზიანების წერტილამდე; სხვადასხვა დროს გარკვეულ პროდუქციაზე მოთხოვნათა რაოდენობა; სითხეში ჩაძირული მტვრის მცირე ნაწილაკის (რომელსაც ვაკვირდებით მიკროსკოპში) მდებარეობა და ა. შ.

განმარტება. შემთხვევითი ექსპერიმენტის ელემენტარულ ხდომილებათა სივრცეზე განსაზღრულ რიცხვით ფუნქციას **შემთხვევითი სიდიდე** ეწოდება. შემთხვევით სიდიდეს ეწოდება **დისკრეტული ტიპის** თუ ის ღებულობს ცალკეულ, იზოლირებულ შესაძლო მნიშვნელობებს. შემთხვევით სიდიდეს ეწოდება **უწყვეტი ტიპის** თუ მისი შესაძლო მნიშვნელობების სიმრავლე მთლიანად ავსებს რაიმე სასრულ ან უსასრულო რიცხვით შუალედს.

დისკრეტული ტიპის შემთხვევითი სიდიდე ღებულობს სასრულ ან თვლად რაოდენობა განსხვავებულ მნიშვნელობებს, ხოლო უწყვეტი ტიპის შემთხვევითი სიდიდის მნიშვნელობათა რაოდენობა კონტინიუმის სიმძლავრისაა.

შემთხვევით სიდიდეებს აღნიშნავენ დიდი ლათინური ასოებით: $X, Y, Z, \ldots$ (ან პატარა ბერძნული ასოებით $\xi, \eta, \varsigma, \ldots$), ხოლო შემთხვევითი სიდიდის შესაძლო მნიშვნელობებს აღნიშნავენ პატარა ლათინური ასოებით: $x_i, y_j, z_k, \ldots$.

მაგალითი 1. შემთხვევითი სიდიდე იყოს მონეტის სამჯერ აგდებისას მოსულ გერბთა რიცხვი. ამ შემთხვევაში ელემენტარულ ხდომილებათა სივრცე რვა ელემენტიანი სიმრავლეა:

$$\Omega = \{\text{გგგ}, \text{გგს}, \text{გსგ}, \text{სგგ}, \text{გსს}, \text{სგს}, \text{სსგ}, \text{სსს}\}$$

და, შესაბამისად, საძიებელი შემთხვევითი სიდიდე იქნება Ω-ზე განსაზღრული შემდეგი რიცხვითი ფუნქცია:

$$X(\text{გგგ}) = 3; \;\; X(\text{გგს}) = X(\text{გსგ}) = X(\text{სგგ}) = 2;$$

$$X(\text{გსს}) = X(\text{სგს}) = X(\text{სსგ}) = 1 \;\; \text{და} \;\; X(\text{სსს}) = 0.$$

ცხადია ეს შემთხვევითი სიდიდე დისკრეტული ტიპისაა, ის ღებულობს იზოლირებულ მნიშვნელობებს, მაგალითად, 1-სა და 2-ს შორის ის არ ღებულობს არცერთ მნიშვნელობას.

მაგალითი 2. შემთხვევითი სიდიდე იყოს ორი სათამაშო კამათლის გაგორებისას მოსულ ქულათა ჯამი. ამ შემთხვევაში ელემენტარულ ხდომილებათა სივრცე შედგება 36 ელემენტარული ხდომილებისაგან:

$$\Omega = \{(i, j) : i, j = 1, 2, \ldots, 6\},$$

ხოლო შემთხვევითი სიდიდე ცალკეულ ელემენტარულ ხდომილებას (i, j) (სადაც i -- პირველ კამათელზე მოსული ქულაა, ხოლო j -- მეორე კამათელზე მოსული ქულა) შეუსაბამებს: $X(i, j) = i + j$ (პირველ და მეორე კამათელზე მოსული ქულების ჯამი). მაგალითად, $X(1,3) = X(2,2) = X(3,1) = 4$. აღნიშნული შემთხვევითი სიდიდის შესაძლო მნიშვნელობებია: 2, 3, . . . , 12. ის ასევე დიკრეტული ტიპისაა.

ზემოთ ჩამოთვლილი მაგალითებიდან უწყვეტი ტიპის შემთხვევითი სიდიდეა მანძილი სამიზნის ცენტრიდან დაზიანების წერტილამდე და მტვრის ნაწილაკის მდებარეობა სითხეში. თითოეული მათგან ნებისმიერ ორ მიღებულ მნიშვნელობას შორის არ გამოტოვებს არცერთ მნიშვნელობას.

შემთხვევითი სიდიდე მოცემულია თუ ჩვენ ვიცით ექსპერიმენტის ამა თუ იმ შედეგს რა რიცხვი შეესაბამება. მაგრამ, იმისათვის რომ ალბათურად დავახასიათოთ შემთხვევითი სიდიდე, ჩვენ კიდევ უნდა ვიცოდეთ თუ რამდენად ხშირად ანუ რა ალბათობებით ღებულობს ეს შემთხვევითი სიდიდე თავის ამა თუ იმ მნიშვნელობას. შესაბამისობას, შემთხვევითი სიდიდის შესაძლო მნიშვნელობებსა და მათ შესაბამის ალბათობებს შორის, დისკრეტული ტიპის შემთხვევითი სიდიდის **განაწილების კანონი** ეწოდება. შემთხვევითი სიდიდის განაწილების კანონი შეიძლება მოცემული იყოს *ცხრილის, ფორმულის ან გრაფიკის* სახით.

ცხრილს, რომელშიც ჩამოთვლილია შემთხვევითი სიდიდის შესაძლო მნიშვნელობები და მათი შესაბამისი ალბათობები, დისკრეტული ტიპის შემთხვევითი სიდიდის **განაწილების მწკრივი** ეწოდება:

x_i	x_1	x_2	$\cdots$	x_n	$\cdots$
p_i	p_1	p_2	$\cdots$	p_n	$\cdots$

შევნიშნოთ, რომ ხდომილება, რომ შემთხვევითი სიდიდე მიიღებს ერთ-ერთ მნიშვნელობას თავისი შესაძლო მნიშვნელობებიდან, წარმოადგენს აუცილებელ ხდომილებას და ამიტომ: $\sum_i p_i = 1$ (ჩვენ არ ვუთითებთ შესაკრებთა რაოდენობას, ის შეიძლება იყოს როგორც სასრული, ისე უსასრულო).

ამოცანა 1. ორი მსროლელი თითოჯერ ესვრის სამიზნეს. მათ მიერ სამიზნის დაზიანების (მიზანში მოხვედრის) ალბათობებია შესაბმისად 0.6 და 0.7. შემთხვევითი სიდიდე X იყოს დაზიანებულ სამიზნეთა რაოდენობა. შევადგინოთ მისი განაწილების მწკრივი.

ამოხსნა. ცხადია, რომ X შემთხვევითმა სიდიდემ შეიძლება მიიღოს შემდეგი მნიშვნელობები: 0 (ვერც ერთმა მსროლელმა ვერ დააზიანა სამიზნე), 1 (მხოლოდ ერთმა მსროლელმა დააზიანა სამიზნე) და 2 (ორივე მსროლელმა დააზიანა სამიზნე). ვიპოვოთ შესაბამისი ალბათობები.

ბუნებრივია შეგვძლია ვიგულისხმოთ რომ პირველი და მეორე მსროლელის სროლის შედეგები ერთმანეთისაგან დამოუკიდებელია. შემოვიღოთ ხდომილებები: A -- პირველმა მსროლელმა დააზიანა სამიზნე და B -- მეორე მსროლელმა დააზიანა სამიზნე. მოცემულია, რომ $P(A) = 0.6$ და

$P(B)=0.7$. შესაბამისად, $P(\overline{A})=0.4$ და $P(\overline{B})=0.3$. გარდა ამისა, A და B დამოუკიდებელი ხდომილებებია. დამოუკიდებელი ხდომილებებია აგრეთვე: $\overline{A}$ და B, $\overline{A}$ დაB , $\overline{A}$ და $\overline{B}$.

ადვილი დასანახია, რომ ხდომილება – ვერც ერთმა მსროლელმა ვერ დააზიანა სამიზნე იქნება $\overline{A}\cap\overline{B}$, ხდომილება -- მხოლოდ ერთმა მსროლელმა დააზიანა სამიზნე იქნება $(A\cap\overline{B})\cup(\overline{A}\cap B)$ და ხდომილება -- ორივე მსროლელმა დააზიანა სამიზნე იქნება $A\cap B$. გასაგებია, რომ $(A\cap\overline{B})$ და $(\overline{A}\cap B)$ უთავსებადი ხდომილებებია $(A\cap\overline{B})\cap(\overline{A}\cap B)=\varnothing$.

ამიტომ, დამოუკიდებელ ხდომილებათა ნამრავლის ალბათობისა და უთავსებად ხდომილებათა ჯამის ალბათობის ფორმულების თანახმად გვექნება:

$$P(X=0)=P(\overline{A}\cap\overline{B})=P(\overline{A})\cdot P(\overline{B})=0,4\cdot 0,3=0,12;$$

$$P(X=1)=P\{(A\cap\overline{B})\cap(\overline{A}\cap B)\}=P(A\cap\overline{B})+P(\overline{A}\cap B)=$$

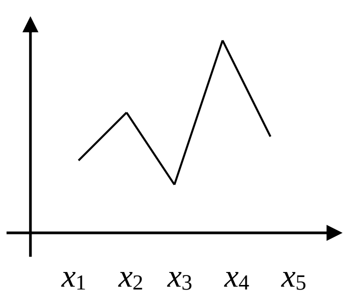
$$=P(A)P(\overline{B})+P(\overline{A})P(B)=0,6\cdot 0,3+0,4\cdot 0,7=0,46;$$

$$P(A\cap B)=P(A)\cdot P(B)=0,6\cdot 0,7=0,42$$

შესაბამისად, X შემთხვევითი სიდიდის განაწილების მწკრივი იქნება:

x_i	**0**	**1**	**2**
p_i	**0.12**	**0.46**	**0.42**

გრაფიკულად დისკრეტული შემთხვევითი სიდიდის განაწილების კანონი შეიძლება წარმოვადგინოთ *განაწილების მრავალკუთხედის სახით,* რომელიც წარმოადგენს ტეხილს სიბრტყეზე, რომელიც მიიღება საკოორდინატო სიბრტყეზე იმ წერტილების შეერთებით, რომელთა კოორდინატებია (x_i,p_i).



თუ მოცემულია დისკრეტული ტიპის შემთხვევითი სიდიდე X და რაიმე რიცხვითი g ფუნქცია, მაშინ $g(X)$ ისევ იქნება დისკრეტული ტიპის შემთხვევითი სიდიდე, რომლის განაწილების მწკრივის პირველ სტრიქონში იქნება $g(x_i)$ რიცხვები ($g(X)$ შემთხვევითი სიდიდის შესაძლო მნიშვნელობები), ხოლო მეორე სტრიქონში იგივე p_i ალბათობები, რაც გვქონდა X შემთხვევითი სიდიდის განაწილების მწკრივში, ვინაიდან:

$$P\{g(X)=g(x_i)\}=P(X=x_i)=p_i,$$

ანუ გვექნება განაწილების მწკრივი:

$g(x_i)$	$g(x_1)$	$g(x_2)$	$\cdots$	$g(x_n)$	$\cdots$
p_i	p_1	p_2	$\cdots$	p_n	$\cdots$

შევნიშნოთ, რომ შესაძლებელია X-ის რომელიმე ორი განსხვავებული $x_j \neq x_k$ მნიშვნელობისათვის $g(x_j) = g(x_k)$, მაშინ $g(X)$-ის განაწილების მწკრივში მხოლოდ ერთ ადგილას დავწერთ $g(x_j)$-ს და ქვეშ მივუწერთ შესაბამისი ალბათობის როლში $(p_j + p_k)$ სიდიდეს. მაგალითად, თუ X შემთხვევითი სიდიდის განაწილების მწკრივია:

x_i	-3	-1	0	1	2
p_i	0.15	0.12	0.2	0.18	0.35

მაშინ X^2-ის (ამ შემთხვევაში $g(x) = x^2$) განაწილების მწკრივი იქნება:

x_i^2	**0**	**1**	**4**	**9**
p_i	**0.2**	**0.3**	**0.35**	**0.15**

აქ $P(X^2 = 1) = P\{(X = -1) \cup (X = 1)\} = P(X = -1) + P(X = -1) = 0.12 + 0.18 = 0.3$.

ჰიპერგეომეტრიული განაწილება. დავუშვათ, რომ ყუთში N ბურთია და მათ შორის M თეთრია. შემთხვევით, დაბრუნების გარეშე ყუთიდან ვიღებთ n ბურთს. ვიპოვოთ ალბათობა იმისა, რომ ამოღებულ n ბურთს შორის ზუსტად k ცალი იქნება თეთრი?

ავღნიშნოთ μ_n-ით ამოღებულ n ბურთს შორის თეთრი ბურთების რაოდენობა. ჩვენ გვაინტერესებს $P(\mu_n = k)$ ალბათობა. ვისარგებლოთ ალბათობის კლასიკური განმარტებით. გასაგებია, რომ ყველა შესაძლო შედეგთა რაოდენობა დაემთხვევა N ელემენტიანი სიმრავლის n ელემენტიან ქვესიმრავლეთა რაოდენობას, ანუ $P(\Omega) = C_N^n$. ჩვენთვის საინტერესო n ელემენტიანი ქვესიმრავლეები უნდა შედგებოდნენ ზუსტად k ცალი თეთრი და $n-k$ ცალი შავი ბურთებისაგან. k ცალი თეთრი ბურთი შეიძლება შეირჩეს C_M^k სხვადასხვა გზით, ხოლო $n-k$ ცალი შავი ბურთი კი -- C_{N-M}^{n-k} სხვადასხვანაირად. ნამრავლის წესის თანახმად ხელშემწყობ ელემენტარულ ხდომილებთა რაოდენობა იქნება $C_M^k \cdot C_{N-M}^{n-k}$. შესაბამისად, კლასიკური განმარტების საფუძველზე გვაქვს:

$$P(N; M; n; k) := P(\mu_n = k) = \frac{C_M^k \cdot C_{N-M}^{n-k}}{C_N^n}, \quad k = 0, 1, \ldots, n.$$

რიცხვთა ამ მიმდევრობას **ჰიპერგეომეტრიული განაწილება** ეწოდება.

ამოცანა. აუდიტორიაში მყოფი 15 სტუდენტიდან 5 ვაჟია. ვიპოვოთ ალბათობა იმისა, რომ შემთხვევით შერჩეულ 6 სტუდენტს შორის 3 ვაჟია?

ამოხსნა. თუ მივუსადაგებთ ზემოთ განხილულ სქემას, გასაგებია, რომ: $N=15,\quad M=5,\quad n=6$ და $k=3$. ამიტომ საძიებელი ალბათობა იქნება:

$$P(15;5;6;3)=\frac{C_{10}^{3}C_{15-10}^{6-3}}{C_{15}^{6}}=\frac{C_{10}^{3}C_{5}^{3}}{C_{15}^{6}}=\frac{120\cdot 10}{5005}\approx 0.239$$

დავუშვათ, რომ ვატარებთ დამოუკიდებელი ორშედეგიანი ცდების სერიას ერთ-ერთი შედეგის (პირობითად მას ვუწოდოთ "წარმატება") პირველად მოხდენამდე. ცალკეულ ცდაში "წარმატების" ალბათობა იყოს p (მეორე შედეგის ალბათობა იქნება $1-p\equiv q$). შემთხვევითი სიდიდე იყოს ჩატარებული ცდების რაოდენობა. მაშინ ცხადია, რომ ეს შემთხვევითი სიდიდე მიიღებს მნიშვნელობას k ალბათობით pq^{k-1}, $k=1,2,\ldots$.

გეომეტრიული განაწილება. დისკრეტულ X შემთხვევით სიდიდეს, რომელიც ღებულობს ნატურალურ k მნიშვნელობებს ალბათობებით

$$P(X=k)=pq^{k-1},$$

სადაც $0<p<1$ ($q=1-p$), გეომეტრიული კანონით განაწილებული შემთხვევითი სიდიდე ეწოდება.

უსასრულოდ კლებადი გეომეტრიული პროგრესიის წევრთა ჯამის ფორმულის გამოყენებით ადვილი შესამოწმებელია, რომ ამ ალბათობების ჯამი 1-ის ტოლია:

$$\sum_{k=1}^{\infty}pq^{k-1}=p\cdot\sum_{k=1}^{\infty}q^{k-1}=p\cdot\frac{1}{1-q}=p\cdot\frac{1}{p}=1.$$

პუასონის განაწილება. განვიხილოთ დისკრეტული შემთხვევითი სიდიდე X, რომელიც ღებულობს მხოლოდ მთელ არაუარყოფით მნიშვნელობებს (0, 1, 2,…, m,…), რომელთა მიმდევრობა შემოუსაზღვრელია. ასეთ შემთხვევით სიდიდეს ეწოდება **პუასონის კანონით** განაწილებული, თუ ალბათობა იმისა, რომ ის მიიღებს მნიშვნელობას m, გამოისახება ფორმულით:

$$p(X=m)=\frac{a^m}{m!}e^{-a},$$

სადაც a -- გარკვეული დადებითი სიდიდეა, რომელსაც პუასონის კანონის (განაწილების) **პარამეტრი** ეწოდება. თუ ვისარგებლებთ e^x ფუნქციის გაშლით ხარისხოვან მწკრივად ($e^x=\sum_{m=0}^{\infty}x^m/m!$), ადვილად დავინახავთ, რომ ამ ალბათობების ჯამი 1-ის ტოლია. მართლაც,

$$\sum_{m=0}^{\infty}p(X=m)=e^{-a}\sum_{m=0}^{\infty}\frac{a^m}{m!}=e^{-a}\cdot e^{a}=1$$

განვიხილოთ ტიპიური ამოცანა, რომელსაც მივყავართ პუასონის განაწილებამდე. დავუშვათ, რომ აბსცისთა ღერძზე შემთხვევით განაწილდებიან წერტილები, ამასთანავე მათი განაწილება აკმაყოფილებს შემდეგ პირობებს:

1). ალბათობა იმისა, რომ გარკვეული რაოდენობის წერტილები მოხ-

ვდება l სიგრძის ინტერვალში დამოკიდებულია მხოლოდ ინტერვალის სიგრძეზე და არაა დამოკიდებული აბსცისთა ღერძზე მის მდებარეობაზე (ე. ი. წერტილები განაწილებულია ერთნაირი საშუალო სიმკვრივით);

2). წერტილები ნაწილდებიან ერთმანეთისაგან დამოუკიდებლად: ალბათობა იმისა, რომ წერტილთა რაიმე რაოდენობა მოხვდება მოცემულ ინტერვალში არ არის დამოკიდებული წერტილთა რაოდენობაზე, რომლებიც მოხვდნენ ნებისმიერ სხვა ინტერვალში;

3). პრაქტიკულად შეუძლებელია ორი ან მეტი წერტილის დამთხვევა.

მაშინ შემთხვევითი სიდიდე X -- l სიგრძის ინტერვალში მოხვედრილ წერტილთა რაოდენობა – განაწილებულია პუასონის კანონით, სადაც a -- არის l სიგრძის ინტერვალზე მოსულ წერტილთა საშუალო რიცხვი.

შენიშვნა. ვინაიდან პუასონის ფორმულა გამოსახავს ბინომიალურ განაწილებას ცდათა დიდი რიცხვისა და ხდომილების მცირე ალბათობის შემთხვევაში, ამიტომ პუასონის კანონს ხშირად უწოდებენ **იშვიათ მოვლენათა კანონს.**

პუასონის განაწილება წარმოადგენს კარგ მათემატიკურ მოდელს იშვიათ ხდომილებათა აღსაწერად: დროის ფიქსირებულ შუალედში მომხდარ ხდომილებათა რაოდენობა ხშირად ემორჩილება პუასონის განაწილებას. მაგალითად, შეიძლება გამოდგეს გეიგერის მთვლელის მიერ t დროში რეგისტრირებული რადიაქტიური დაშლის შედეგად α ნაწილაკების რაოდენობა, სატელეფონო სადგურში t დროის განმავლობაში რეგისტრირებულ გამოძახებათა რაოდენობა. როგორც ჩვენ უკვე ვნახეთ, წარმატებების მცირე ალბათობისა და ცდათა რიცხვის საკმაოდ დიდი რაოდენობის შემთხვევაში პუასონის განაწილება გვევლინება ბინომური განაწილების მიახლოებად.

§17. შემთხვევითი სიდიდის განაწილების ფუნქცია და განაწილების სიმკვრივე

შემთხვევითი სიდიდის განაწილება – ეს არის ფუნქცია, რომელიც ცალსახად განსაზღვრავს ალბათობას იმისა, რომ: შემთხვევითმა სიდიდემ მიიღო მოცემული მნიშვნელობა ან შემთხვევითი სიდიდე ეკუთვნის გარკვეულ მოცემულ ინტერვალს. თუ შემთხვევითი სიდიდე ღებულობს სასრულ რაოდენობა მნიშვნელობებს, მაშინ განაწილება მოიცემა ფუნქციით $P(X=x)$, რომელიც X შემთხვევითი სიდიდის ყველა შესაძლო x მნიშვნელობას შეუსაბამებს ალბათობას იმისა, რომ $X=x$ (ანუ განაწილების კანონით):

$$P(X\in\langle a,b\rangle)=\sum_{x\in\langle a,b\rangle}P(X=x),$$

სადაც a და b ($a<b$) ნებისმიერი ნამდვილი რიცხვებია, ხოლო $\langle a,b\rangle$ -- ნებისმიერი ტიპის ინტერვალია (როგორც ღია, ისე ნახევრად ღია და ჩაკეტილი).

თუ შემთხვევითი სიდიდე ღებულობს უსასრულოდ ბევრ მნიშვნელობას (რაც შესაძლებელია მხოლოდ მაშინ, როცა ელემენტარულ ხდომილებათა სივრცე, რომელზეც განმარტებულია შემთხვევითი სიდიდე შედგება უსასრულო რაოდენობა ელემენტარული ხდომილებებისაგან), მაშინ განაწილება მოიცემა $P(a<X\le b)$ ალბათობების ერთობლიობით რიცხვთა ნებისმიერი a,b წყვილისათვის, $a<b$. განაწილება შესაძლებელია მოცემულ იქნეს ე. წ. *განაწილების ფუნქციით:*

$$F(x):=P(X\le x),$$

რომელიც ნებისმიერი ნამდვილი x რიცხვისთვის განსაზღვრავს ალბათობას იმისა, რომ X შემთხვევითი სიდიდე მიიღებს ისეთ მნიშვნელობებს, რომელიც არ აღემატება x-ს. ადვილი დასანახია, რომ:

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

მართლაც, ხდომილებათა სხვაობის ალაბათობის ფორმულის თანახმად გვაქვს:

$$P(a<X\le b)=P\{(X\le b)\setminus(X\le a)\}=P(X\le b)-P(X\le a)=F(b)-F(a).$$

ეს თანაფარდობა გვიჩვენებს თუ როგორ შეიძლება განაწილების ფუნქციის საშუალებით გამოვთვალოთ განაწილება და პირიქით, როგორ გამოვთვალოთ განაწილების ფუნქცია განაწილების საშუალებით:

$$F(x)=P(X\le x)=P(\infty<X\le x).$$

განაწლების ფუნქციის თვისებები:

1). ნებისმიერი x-სათვის $0\le F(x)\le 1$ (როგორც $(X\le x)$ ხდომილების ალბათობა);

2). განაწილების ფუნქცია არაკლებადია: თუ $x^{'}<x^{''}$, მაშინ $F(x^{'})\le F(x^{''})$;

მართლაც, თუ $x^{'}<x^{''}$, მაშინ ხდომილება $(X\le x^{'})$ იწვევს ხდომილებას $(X\le x^{''})$, ამიტომ გვაქვს:

$$F(x^{'})=P(X\le x^{'})\le P(X\le x^{''})=F(x^{''});$$

3). განაწილების ფუნქცია უწყვეტია მარჯვნიდან (თუ განაწილების ფუნქციას განვმარტავთ როგორც: $F(x) := P(X < x)$, მაშინ ის იქნება მარცხნიდან უწყვეტი);

ამ თვისების შესამოწმებლად უნდა ვისარგებლოთ ე.წ. **ალბათობის უწყვეტობის თვისებით**: თუ $\{A_n\}_{n\geq 1}$ ხდომილებათა კლებადი (შესაბამისად, ზრდადი) მიმდევრობაა $A_n \supset A_{n+1}$ (შესაბამისად, $A_n \subset A_{n+1}$), მაშინ ადგილი აქვს თანაფარდობას

$$\lim_{n\to\infty} P(A_n) = P(\bigcap_{n=1}^{\infty} A_n) \text{ (შესაბამისად, } \lim_{n\to\infty} P(A_n) = P(\bigcup_{n=1}^{\infty} A_n)) .$$

აღვნიშნოთ $A_n := \{X \leq x + \frac{1}{n}\}$, მაშინ თუ გავითვალისწინებთ, რომ $\bigcap_{n=1}^{\infty}\{X \leq x + \frac{1}{n}\} = \{X \leq x\}$, ადვილად დავრწმუნდებით განაწილების ფუნქციის მარჯვნიდან უწყვეტობაში:

$$F(x+0) := \lim_{n\to\infty} F(x+\frac{1}{n}) = \lim_{n\to\infty} P(X \leq x + \frac{1}{n}) = P(\bigcap_{n=1}^{\infty}\{X \leq x + \frac{1}{n}\} = P\{X \leq x\} = F(x).$$

4). თუ მოცემულია დისკრეტული ტიპის შემთხვევითი სიდიდის განაწილების კანონი $p_k = P(X = x_k)$, მაშინ მისი განაწილების ფუნქცია იქნება:

$$F(x) = \sum_{x \leq x_k} P(X = x_k) \text{ (შესაბამისად, თუ } F(x) := P(X < x) \text{, მაშინ}$$

$$F(x) = \sum_{x < x_k} P(X = x_k));$$

მართლაც, ცხადია, რომ ხდომილება $(X \leq x)$ წარმოიდგინება უთავსებადი $(X = x_k)$ ხდომილებების გაერთიანების სახით $(X \leq x) = \bigcup_{x \leq x_k} (X = x_k)$.

ამიტომ უთავსებად ხდომილებათა ჯამის ალბათობის ფორმულის თანახმად გვაქვს:

$$P(X \leq x) = P\{\bigcup_{x \leq x_k} (X = x_k)\} = \sum_{x \leq x_k} P(X = x_k) = \sum_{x \leq x_k} p_k .$$

5). თუ მოცემულია დისკრეტული ტიპის შემთხვევითი სიდიდის განაწილების ფუნქცია $F(x) := P(X \leq x)$, მაშინ მისი განაწილების კანონი იქნება:

$$p_k = P(X = x_k) = F(x_k) - F(x_k - 0) \text{ (შესაბამისად, თუ } F(x) := P(X < x),$$

$$\text{მაშინ } p_k = P(X = x_k) = F(x_k + 0) - F(x_k)),$$

ანუ ალბათობა იმისა, რომ შემთხვევითი სიდიდე მიიღებს თავის კონკრეტულ მნიშვნელობას ტოლია ამ მნიშვნელობაზე განაწილების ფუნქციის ნახტომის სიდიდის;

აღვნიშნოთ $A_n := \{X \le x - \frac{1}{n}\}$. მაშინ თუ ვისარგებლებთ თანაფარდობით $\bigcup_{n=1}^{\infty}\{X \le x - \frac{1}{n}\} = \{X < x\}$, ალბათობის უწყვეტობის თვისების თანახმად დავასკვნით, რომ:

$$P\{X < x\} = P(\bigcup_{n=1}^{\infty}\{X \le x - \frac{1}{n}\}) = \lim_{n\to\infty} P(\{X \le x - \frac{1}{n}\}) = \lim_{n\to\infty} F(x - \frac{1}{n}) = F(x-0).$$

ამიტომ, ვინაიდან $(X = x_k) = (X \le x_k) \setminus (X < x_k)$, სხვაობის ალბათობის ფორმულის საშუალებით, ვრწმუნდებით, რომ:

$$p_k = P(X = x_k) = P(X \le x_k) - P(X < x_k) = F(x_k) - F(x_k - 0) := \Delta F(x_k).$$

განაწილების ფუნქცია შეიძლება იყოს ან დისკრეტული, ან უწყვეტი, ან მათი კომბინაცია. დისკრეტული განაწილების ფუნქცია შეესაბამება დისკრეტულ შემთხვევით სიდიდეს, რომელიც ღებულობს სასრულ რაოდენობა მნიშვნელობებს ან მნიშვნელობებს ისეთი სიმრავლიდან, რომლის ელემენტების გადანომვრაც შეიძლება ნატურალური რიცხვებით (ასეთ სიმრავლეებს, მათემატიკაში, *თვლად სიმრავლეებს* უწოდებენ). დისკრეტულ განაწილების ფუნქციას აქვს საფეხურა კიბის სახე. $1 < x \le 2$

მაგალითი 3. საქონლის პარტიაში დეფექტურ ნაწარმთა რიცხვი X ღებულობს მნიშვნელობა 0-ს ალბათობით – 0.3; მნიშვნელობა 1-ს ალბათობით –0.4; მნიშვნელობა 2-ს ალბათობით – 0.2 და მნიშვნელობა 3-ს ალბათობით – 0.1 ანუ X -ის განაწილების მწკრივს აქვს სახე:

x_i	0	1	2	3
p_i	0.3	0.4	0.2	0.1

გამოვთვალოთ X -ის $F(x) = P(X < x)$ განაწილების ფუნქცია და ავაგოთ მისი გრაფიკი.

თუ $x \le 0$, მაშინ $F(x) = P(X < x) = P(\varnothing) = 0$;

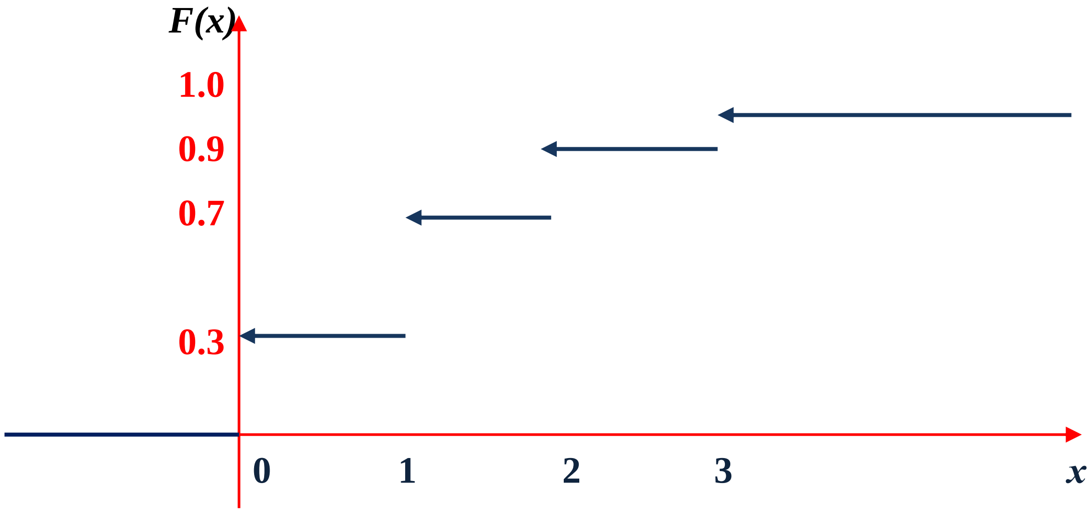
თუ $0 < x \le 1$, მაშინ $F(x) = P(X < x) = P(X = 0) = 0.3$;

თუ $1 < x \le 2$, მაშინ $F(x) = P(X < x) = P\{(X = 0) \cup (X = 1)\} =$
$= P(X = 0) + P(X = 1) = 0.3 + 0.4 = 0.7$;

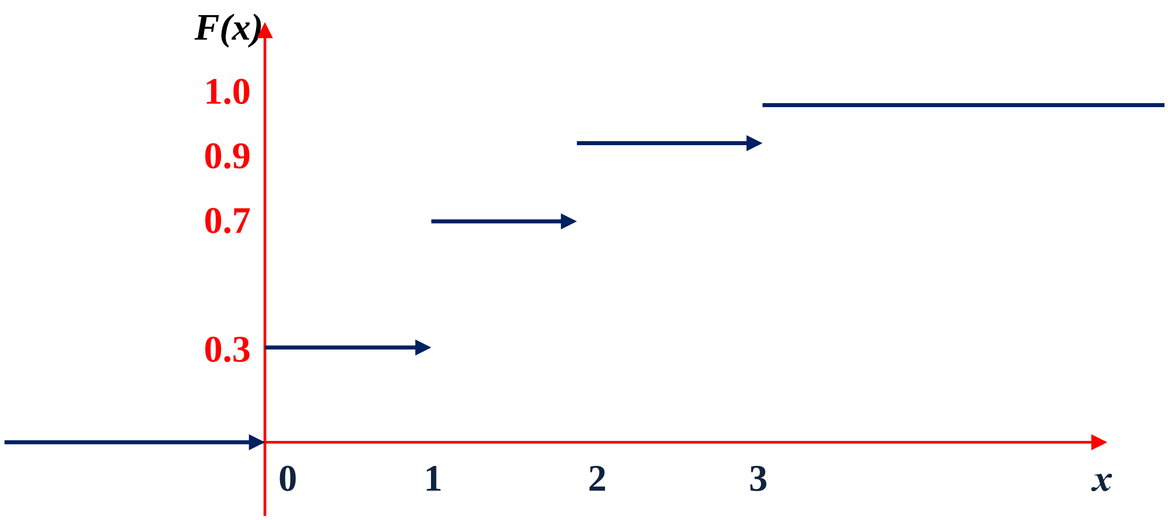
თუ $2 < x \le 3$, მაშინ $F(x) = P(X < x) = P\{(X = 0) \cup (X = 1) \cup (X = 2)\} =$
$= P(X = 0) + P(X = 1) + P(X = 2) = 0.3 + 0.4 + 0.2 = 0.9$;

და ბოლოს, თუ $x \ge 3$, მაშინ $F(x) = P(X < x) = P(\Omega) = 1$.

ამიტომ $F(x) = P(X < x)$ განაწილების ფუნქციის გრაფიკს ექნება შემდეგი სახე:



ანალოგიურად, ცხადია, რომ $F(x) := P(X \le x)$ განაწილების ფუნქციის გრაფიკი იქნება შემდეგი სახის:



უწყვეტ განაწილების ფუნქციას ნახტომები არა აქვს. ის მონოტონურად იზრდება არგუმენტის ზრდასთან ერთად 0-დან (როცა $x \to -\infty$) 1-მდე (როცა $x \to +\infty$). შემთხვევით სიდიდეს, რომელსაც აქვს უწყვეტი განაწილების ფუნქცია, უწოდებენ **უწყვეტ შემთხვევით** სიდიდეს.

თუ უწყვეტი შემთხვევითი სიდიდის განაწილების ფუნქცია $F(x)$ წარმოებადია, მაშინ მის წარმოებულს შემთხვევითი სიდიდის განაწილების სიმკვრივე ეწოდება და აღინიშნება $f(x)$-ით:

$$f(x) = \frac{dF(x)}{dx}.$$

შემთხვევითი სიდიდის განაწილების სიმკვრივიდან შეგვიძლია აღვადგინოთ განაწილების ფუნქცია:

$$F(x)=\int_{-\infty}^{x} f(y)dy.$$

ვინაიდან ნებისმიერი განაწილების ფუნქციისათვის:

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

ამიტომ, ნიუტონ-ლეიბნიცის ფორმულის გამოყენებით, გვაქვს:

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

მაგალითი 4. მოცემულია შემთხვევითი სიდიდის განაწილების ფუნქცია:

$$F(x)=\begin{cases} 0, & x<a, \\ \dfrac{x-a}{b-a}, & a\le x\le b, \\ 1, & x>b. \end{cases}$$

(სადაც a და b ($a<b$) ნებისმიერი ნამდვილი რიცხვებია). ვიპოვოთ შესაბამისი განაწილების სიმკვრივე. ცხადია, რომ განმარტების თანახმად:

$$f(x)=\begin{cases} 0, & x<a, \\ \dfrac{1}{b-a}, & a<x<b, \\ 0, & x>b \end{cases}$$

რაც შეეხება $x=a$ და $x=b$ წერტილებს, აქ $F(x)$ ფუნქციას წარმოებული არა აქვს და იქ შეგვიძლია $f(x)$ განვმარტოთ ნებისმიერად, ვთქვათ, $f(a)=f(b)=0$. შემთხვევით სიდიდეს, რომელსაც აქვს აღნიშნული განაწილების სიმკვრივე, ეწოდება **თანაბარად განაწილებული [*a,b*] მონაკვეთზე**.

შერეული ტიპის განაწილების ფუნქციები გვხვდება, როცა დაკვირვებები რომელიღაც მომენტში წყდება. მაგალითად, იმ სტატისტიკური მონაცემების ანალიზის დროს, რომლებიც მიღება ობიექტის საიმედობაზე გამოცდის (შემოწმების) ისეთი გეგმის გამოყენებისას, რომელიც გულისხმობს გამოცდის შეწყვეტას გარკვეული დროის ამოწურვის შემდეგ ან იმ ტექნიკური ნაწარმის მონაცემების ანალიზის დროს, რომლებსაც დასჭირდათ საგარანტიო შეკეთება.

მაგალითი 5. დავუშვათ, რომ ნათურის მუშაობის დრო არის შემთხვევითი სიდიდე განაწილების ფუნქციით $F(t)$, ხოლო ნათურის გამოცდა გრძელდება ნათურის მწყობრიდან გამოსვლამდე (გადაწვამდე), თუ ეს მოხდება გამოცდის დაწყებიდან არაუმეტეს 100 საათის განმავლობაში, ანუ $t_0=100$ სთ მომენტამდე. ვთქვათ, $G(t)$ -- არის ამ გამოცდის დროს ნათუ-

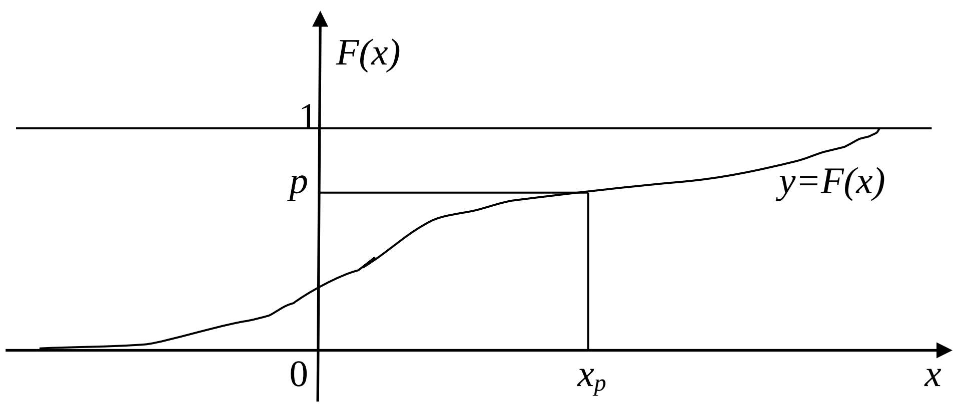
რის ნორმალურად მუშაობის დროის განაწილების ფუნქცია. მაშინ ცხადია, რომ:

$$G(t) = \begin{cases} F(t), & t \le 100 \\ 1, & t > 100. \end{cases}$$

$G(t)$ ფუნქციას აქვს ნახტომი t_0 წერტილში, ვინაიდან შესაბამისი შემთხვევითი სიდიდე ღებულობს t_0 მნიშვნელობას ალბათობით $1 - F(t_0) > 0$.

შემთხვევით სიდიდეთა მახასიათებლები. ალბათურ-სტატისტიკურ მეთოდებში გამოიყენება შემთხვევით სიდიდეთა სხვადასხვა მახასიათებლები, რომლებიც გამოისახება განაწილების ფუნქციითა და განაწილების სიმკვრივით.

კვანტილი. შემოსავლების დიფერენცირების აღწერისას, შემთხვევითი სიდიდის პარამეტრების შესაძლო (საიმედობის) საზღვრების დადგენისას და სხვა მრავალ შემთხვევაში გამოიყენება ე. წ. "***p* რიგის კვანტილის**" ცნება ($0 < p < 1$), რომელიც აღინიშნება x_p სიმბოლოთი. p რიგის კვანტილი ეწოდება შემთხვევითი სიდიდის იმ მნიშვნელობას, რომლისთვისაც განაწილების ფუნქცია ღებულობს მნიშვნელობას p ან ადგილი აქვს "ნახტომს" მნიშვნელობიდან, რომელიც ნაკლებია p-ზე მნიშვნელობისაკენ რომელიც მეტია p-ზე. შეიძლება მოხდეს, რომ ეს პირობა სრულდება x ყველა მნიშვნელობისათვის გარკვეული ინტერვალიდან (ე. ი. განაწილების ფუნქცია მუდმივია ამ ინტერვალზე და ტოლის p-სი), მაშინ x ყველა ასეთ მნიშვნელობას უწოდებენ p რიგის კვანტილს. უწყვეტი განაწილების ფუნქციების შემთხვევაში, როგორც წესი, არსებობს ერთადერთი p რიგის x_p კვანტილი, ამასთანავე $F(x_p) = p$.



მაგალითი 6. ვიპოვოთ p რიგის x_p კვანტილი თანაბარი განაწილების ფუნქციისათვის.

p ($0 < p < 1$) რიგის x_p კვანტილი უნდა ვეძებოთ როგორც $F(x) = p$ განტოლების ამონახსნი. თანაბარი განაწილების ფუნქციის შემთხვევაში ეს განტოლება მიიღებს სახეს:

$$\frac{x - a}{b - a} = p,$$

საიდანაც ცხადია, რომ:

$$x_p = a + p(b-a) = a(1-p) + bp .$$

როცა $p=0$, მაშინ ნებისმიერი $x \le a$ წარმოადგენს $p=0$ რიგის კვანტილს, ხოლო $p=1$ რიგის კვანტილი იქნება ნებისმიერი $x \ge b$ რიცხვი.

დისკრეტული განაწილების შემთხვევაში, როგორც წესი, არ არსებობს x_p, რომელიც აკმაყოფილებს $F(x_p) = p$ განტოლებას. უფრო ზუსტად, თუ დისკრეტულ შემთხვევით სიდიდეს აქვს n მნიშვნელობა $x_1 < x_2 < \cdots < x_n$ (ამასთანავე შესაბამისი ალბათობებია $P(X = x_i) = p_i$, $i = 1, 2, ..., n$), მაშინ $F(x_p) = p$ განტოლებას x_p-ს მიმართ გააჩნია ამონახსნი p-ს მხოლოდ n მნიშვნელობისათვის, კერძოდ, როცა:

$$p = p_1,$$

$$p = p_1 + p_2 ,$$

$$\cdots$$

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

p-ს ჩამოთვლილი n მნიშვნელობისათვის $F(x_p) = p$ განტოლების x_p ამონახსნი არაერთადერთია, კერძოდ,

$$F(x) = p_1 + p_2 + \cdots + p_i$$

ყველა ისეთი x-სათვის, რომლისთვისაც სრულდება უტოლობა $x_i < x \le x_{i+1}$. ე. ი. x_p -- ნებისმიერი რიცხვია $(x_i, x_{i+1}]$ ინტერვალიდან. ყველა დანარჩენი p-სათვის (0,1) ინტერვალიდან ადგილი აქვს "ნახტომს" p-ზე ნაკლები მნიშვნელობიდან p-ზე მეტი მნიშვნელობისაკენ. კერძოდ, თუ

$$p_1 + p_2 + \cdots + p_i < p \le p_1 + p_2 + \cdots + p_i + p_{i+1} ,$$

მაშინ $x_p = x_{i+1}$.

მდებარეობის მახასიათებლები მიუთითებენ განაწილების "ცენტრზე". სტატისტიკაში დიდი მნიშვნელობა აქვს $p = 1/2$ რიგის კვანტილს. მას X შემთხვევითი სიდიდის ან მისი განაწილების $F(x)$ ფუნქციის **მედიანა** ეწოდება და აღინიშნება x. ე. ი. $x = x_{1/2}$. ისევე როგორც გეომეტრიაში სამკუთხედის მედიანა წვეროს მოპირდაპირე გვერდს ყოფს ორ ტოლ ნაწილად, მათემატიკურ სტატისტიკაში მედიანა შუაზე ყოფს შემთხვევითი სიდიდის განაწილებას: ტოლობა $F(x_{1/2}) = 1/2$ ნიშნავს, რომ ალბათობა იმისა, რომ შემთხვევითი სიდიდე მიიღებს მნიშვნელობას $x_{1/2}$-ის მარცხნივ ან თვითონ $x_{1/2}$-ის ტოლ მნიშვნელობას და ალბათობა იმისა, რომ შემთხვევითი სიდიდე მიიღებს მნიშვნელობას $x_{1/2}$-ის მარჯვნივ ერთმანეთის ტოლია და ორივე 1/2-ია, ანუ

$$P(X \le x_{1/2}) = P(X > x_{1/2}) = 1/2$$

(მათემატიკურად უფრო ზუსტია განმარტება: $P(X \le x_{1/2}) \ge 1/2$ და $= P(X \ge x_{1/2}) \ge 1/2$).

თუ დისკრეტულ შემთხვევით სიდიდეს აქვს n მნიშვნელობა დალაგებული ზრდადობის მიხედვით $x_1 < x_2 < \cdots < x_n$ და თითოეული ეს მნიშვნელობა მიიღება ერთი და იგივე $1/n$-ის ტოლი ალბათობით, მაშინ გასაგებია, რომ როცა n კენტია მედიანა იქნება ზუსტად შუაში (ცენტრში) მდგომი მნიშვნელობა:

$$x = x_{(n+1)/2},$$

ხოლო როდესაც n ლუწია, მაშინ მედიანის როლში იღებენ ბოლოებიდან თანაბრად დაშორებული ორი (შუაში მდგომი ორი) მნიშვნელობის საშუალო არითმეტიკულს:

$$x = \frac{x_{(n/2)} + x_{(n/2)+1}}{2}.$$

მედიანა განეკუთვნება შემთხვევითი სიდიდის ცენტრალური ტენდენციის მდგრადი საზომების ჯგუფს. საზოგადოდ მონაცემთა რაიმე რიცხვითი მახასიათებლის **მდგრადობა** ნიშნავს, რომ დაკვირვებათა მცირე რაოდენობის ცვლილებას შეზღუდული გავლენა აქვს მასზე, მიუხედავად იმისა, თუ როგორია ამ ცვლილების სიდიდე.

შემთხვევითი სიდიდის ერთ-ერთი რეალური შინაარსის მქონე მახასიათებელია *მოდა*. **მოდა** ეწოდება შემთხვევითი სიდიდის იმ მნიშვნელობას (ან მნიშვნელობებს), რომელიც შეესაბამება განაწილების სიმკვრივის ლოკალურ მაქსიმუმს უწყვეტი შემთხვევითი სიდიდის შემთხვევაში ან ალბათობის ლოკალურ მაქსიმუმს დისკრეტული შემთხვევითი სიდიდის შემთხვევაში.

თუ x_0 უწყვეტი ტიპის შემთხვევითი სიდიდის (რომლის განაწილების სიმკვრივეა $f(x)$) მოდაა, მაშინ გასაგებია, რომ $\frac{df(x_0)}{dx} = 0$.

შემთხვევით სიდიდეს შეიძლება ჰქონდეს ბევრი მოდა. მაგალითად, თანაბარი განაწილების შემთხვევაში ნებისმიერი x წერტილი $a < x < b$ წარმოადგენს მოდას. თუმცა ეს გამონაკლისია. უმეტესობა შემთხვევით სიდიდეებს, რომლებიც გამოიყენებიან ალბათურ-სტატისტიკური მეთოდების საშუალებით გადაწყვეტილებების მიღებისას ან სხვა გამოყენებითი ხასიათის კვლევებში, გააჩნიათ ერთი მოდა. იმ შემთხვევით სიდიდეებს (შესაბამისად, იმ განაწილებებს ან განაწილებების სიმკვრივეებს), რომელთაც აქვთ ერთი მოდა – უწოდებენ **უნიმოდალურს.**

§18. ორგანზომილებიანი შემთხვევითი სიდიდე.

ერთგანზომილებიან შემთხვევით სიდიდეებთან ერთად, რომელთა შესაძლო მნიშვნელობები განისაზღვრება ერთი რიცხვით, ალბათობის თეორიაში განიხილება მრავალგანზომილებიანი შემთხვევითი სიდიდეებიც. ასეთი შემთხვევითი სიდიდის ნებისმიერი მნიშვნელობა წარმოადგენს რამოდენიმე რიცხვის დალაგებულ ერთობლიობას. ამ ცნების გეომეტრიულ ილუსტრაციას წარმოადგენს n –განზომილებიანი სივრცის წერტილები, რომელთა თითოეული კოორდინატა წარმოადგენს შემთხვევით სიდიდეს (დისკრეტულს ან უწყვეტს), ანუ n –განზომილებიან ვექტორს. ამიტომ, მრავალგანზომილებიან შემთხვევით სიდიდეებს ასევე შემთხვევით ვექტორებსაც უწოდებენ. სიმარტივისათვის ჩვენ განვიხილავთ ორგანზომილებიან შემთხვევით ვექტორებს. ჯერ განვიხილოთ დისკრეტული შემთხვევა.

დისკრეტული ორგანზომილებიანი (X, Y) შემთხვევითი სიდიდის **განაწილების კანონს** (ანუ X და Y შემთხვევითი სიდიდეების **ერთობლივ განაწილების კანონს**) აქვს ორგანზომილებიანი ცხრილის სახე, რომელიც გვაძლევს შესაძლო მნიშვნელობების ცალკეული კომპონენტების ჩამონათვალს და იმ $p(x_i, y_j)$ ალბათობებს, რა ალბათობებითაც მიიღება მნიშვნელობა (x_i, y_j):

Y	**X**					
	x_1	x_2	…	x_i	…	x_n
y_1	$p(x_1, y_1)$	$p(x_2, y_1)$	…	$p(x_i, y_1)$	…	$p(x_n, y_1)$
…	…	…	…	…	…	…
y_j	$p(x_1, y_j)$	$p(x_2, y_j)$	…	$p(x_i, y_j)$	…	$p(x_n, y_j)$
…	…	…	…	…	…	…
y_m	$p(x_1, y_m)$	$p(x_2, y_m)$	…	$p(x_i, y_m)$	…	$p(x_n, y_m)$

ამასთანავე, ცხრილის ყველა უჯრაში მდგომი ალბათობების ჯამი 1-ის ტოლია. თუ ცნობილია ორგანზომილებიანი შემთხვევითი სიდიდის განაწილების კანონი, ჩვენ შეგვიძლია ვიპოვოთ მისი შემადგენელი ცალკეული შემთხვევითი სიდიდის განაწილების კანონი. მართლაც, ხდომილება $X = x_1$ წარმოადგენს ჯამს არათავსებადი $(X = x_1, Y = y_1)$, $(X = x_1, Y = y_2)$,…, $(X = x_1, Y = y_m)$ ხდომილებების, ამიტომ $p(X = x_1) = p(x_1, y_1) + p(x_1, y_2) +\ldots+ +p(x_1, y_m)$ (მარჯვენა მხარეს წერია $X = x_1$ სვეტის შესაბამისი ალბათობების ჯამი). ანალოგიურად შეგვიძლია ვიპოვოთ X –ის დანარჩენი მნიშვნელობების ალბათობები. იმისათვის, რომ განვსაზღვროთ Y –ის შესაძლო მნიშვნ-ელობების ალბათობები, საჭიროა შეიკრიბოს $Y = y_j$ სვეტის შესაბამისი ალბათობები.

მაგალითი 7. მოცემულია ორგანზომილებიანი შემთხვევითი სიდიდის განაწილების კანონი:

Y	**X**		
	-2	**3**	**6**
-0.8	**0.1**	**0.3**	**0.1**
-0.5	**0.15**	**0.25**	**0.1**

ვიპოვოთ ცაკლკეული შემთხვევითი სიდიდის განაწილების კანონი.

ამოხსნა. ცხრილში მოყვანილი ალბათობების სვეტების მიხედვით შეკრებით მივიღებთ *X*–ის განაწილების მწკრივს:

X	-2	3	6
p	0.25	0.55	0.2

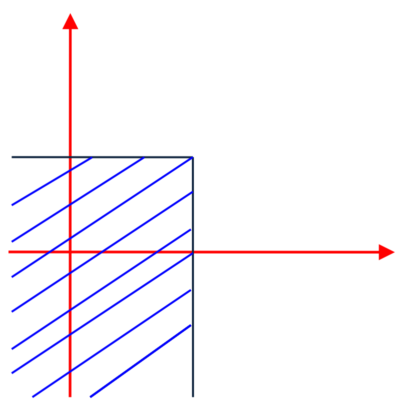
ალბათობების შეკრებით სტრიქონების მიხედვით, მივიღებთ *Y*–ის განაწილების მწკრივს:

Y	-0.8	-0.5
p	0.5	0.5

გადავიდეთ უწყვეტი ტიპის შემთხვევითი სიდიდეების განხილვაზე.

განმარტება 2. ორგანზომილებიანი (X, Y) შემთხვევითი სიდიდის **განაწილების ფუნქცია** (ან X და Y შემთხვევითი სიდიდეების **ერთობლივი განაწილების ფუნქცია**) ეწოდება $(X \le x, Y \le y)$ ხდომილების ალბათობას:

$$F(x, y) = P(X \le x, Y \le y).$$



ეს ნიშნავს, რომ (X, Y) წერტილი მოხვდება დაშტრიხულ არეში, თუ მართი კუთხის წვერო მოთავსებულია წერტილში (x, y).

შევნიშნავთ, რომ ერთობლივი განაწილების ფუნქცია განიმარტება როგორც დისკრეტული, ისე უწყვეტი შემთხვევითი სიდიდეებისათვის. ჩამოვაყალიბოთ მისი თვისებები:

1). $0 \le F(x, y) \le 1$ (ვინაიდან $F(x, y)$ ალბათობაა).

2). $F(x, y)$ არის თითოეული არგუმენტის მიმართ არაკლებადი, მარჯვნიდან უწყვეტი ფუნქცია.

3). ადგილი აქვს ზღვრულ თანაფარდობებს: $F(-\infty, y) = 0$; $F(x, -\infty) = 0$; $F(-\infty, -\infty) = 0$; $F(\infty, \infty) = 1$.

4). $F(x, \infty) = F_1(x)$; $F(\infty, y) = F_2(y)$.

შემთხვევით სიდიდეებს ეწოდება **დამოუკიდებელი**, თუ

$$F(x, y) = F_1(x)\, F_2(y).$$

დისკრეტული შემთხვევითი სიდიდები დამოუკიდებელია მაშინ და მხოლოდ მაშინ, როცა

$$p(x_i, y_j) = p(x_i)\, p(y_j), \ \forall i, j .$$

განმარტება 3. უწყვეტი ორგანზომილებიანი შემთხვევითი სიდიდის **ერთობლივი განაწილების სიმკვრივე** (ანუ **ორგანზომილებიანი სიმკვრივე**) ეწოდება ერთობლივი განაწილების ფუნქციის შერეულ მეორე რიგის კერძო წარმოებულს:

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} .$$

შენიშვნა. ორგანზომილებიანი განაწილების სიმკვრივე წარმოადგენს შემთხვევითი წერტილის Δx და Δy გვერდების მქონე მართკუთხედში მოხვედრის ალბათობის ამ მართკუთხედის ფართობთან შეფარდების ზღვარს, როცა $\Delta x \to 0, \ \Delta y \to 0.$

ორგანზომილებიანი განაწილების სიმკვრივის თვისებები:

1). $f(x, y) \geq 0$ (წერტილის მართკუთხედში მოხვედრის ალბათობა არაუარყოფითია, ამ მართკუთხედის ფართობი არაუარყოფითია, და, შესაბამისად, მათი შეფარდების ზღვარი არაუარყოფითია).

2). $F(x, y) = \int\limits_{-\infty}^{y} \int\limits_{-\infty}^{x} f(x, y) dx dy .$

3). $\int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} f(x, y) dx dy = 1$ (ვინაიდან ეს არის აუცილებელი ხდომილების, კერძოდ, წერტილის Oxy სიბრტყეზე მოხვედრის, ალბათობა).

4). წერტილის სიბრტყის   არეში მოხვედრის ალბათობა ტოლია:

$$p((X, Y) \subset D) = \iint\limits_{D} f(x, y) dx dy.$$

5). ორგანზომილებიანი განაწილების სიმკვრივიდან ერთ-ერთი შემთხვევითი სიდიდის განაწილების სიმკვრივის პოვნა:

$$f_1(x) = \frac{dF_1(x)}{dx} = \frac{dF(x, \infty)}{dx} = \frac{d\left(\int\limits_{-\infty}^{x} \int\limits_{-\infty}^{\infty} f(x, y) \right)}{dx} = \int\limits_{-\infty}^{\infty} f(x, y) dy ,$$

ანალოგიურად, $f_2(y) = \int\limits_{-\infty}^{\infty} f(x, y) dx.$

6). უწყვეტი შემთხვევითი სიდიდეები დამოუკიდებელია მაშინ და მხოლოდ მაშინ, როცა

$$f(x, y) = f_1(x)\, f_2(y).$$

მაგალითი 8. მოცემულია X და Y დამოუკიდებელი შემთხვევითი სიდიდეების განაწილების კანონი:

X	-2	3	6
p	0.2	0.5	0.3

Y	-0.8	-0.5
p	0.4	0.6

ვიპოვოთ $Z=\max\{X,Y\}$ შემთხვევითი სიდიდის განაწილების კანონი.

ამოხსნა. Z შემთხვევითი სიდიდის შესაძლო მნიშვნელობებია: -0.8; -0.5; 3 და 6. გამოვთვალოთ შესაბამისი ალბათობები. გვაქვს:

$$P(Z=-0.8)=P\{(X=-2)\cap(Y=-0.8)\}=P(X=-2)\cdot P(Y=-0.8)=0.2\cdot 0.4=0.08\ ;$$

$$P(Z=-0.5)=P\{(X=-2)\cap(Y=-0.5)\}=P(X=-2)\cdot P(Y=-0.5)=0.2\cdot 0.6=0.12\ ;$$

$$P(Z=3)=P\{[(X=3)\cap(Y=-0.8)]\cup[(X=3)\cap(Y=-0.5)]\}=$$

$$=P[(X=3)\cap(Y=-0.8)]+P[(X=3)\cap(Y=-0.5)]=0.5\cdot 0.4+0.5\cdot 0.6=0.5\ ;$$

$$P(Z=6)=P\{[(X=6)\cap(Y=-0.8)]\cup[(X=6)\cap(Y=-0.5)]\}=$$

$$=P[(X=6)\cap(Y=-0.8)]\cdot P[(X=6)\cap(Y=-0.5)]=0.3\cdot 0.4+0.3\cdot 0.6=0.3\,.$$

§19. შემთხვევითი სიდიდის მათემატიკური ლოდინი

ხშირად შემთხვევითი სიდიდის დასახასიათებლად უფრო მოხერხებულია რიცხვითი მახასიათებლები, ნაცვლად ფუნქციონალურისა (როგორიცაა განაწილების კანონი, განაწილების ფუნქცია ან განაწილების სიმკვრივე უწყვეტ შემთხვევაში). შემთხვევითი სიდიდის რიცხვით მახასიათებლებს შორის პირველ რიგში გამოყოფენ ისეთებს, რომელთა "ირგვლივ" ("გარშემოც") ლაგდება (ჯგუფდება) შემთხვევითი სიდიდის შესაძლო მნიშვნელობები. ერთერთ ასეთ რიცხვით მახასიათებლს წარმოადგენს შემთხვევითი სიდიდის *მათემატიკური ლოდინი,* რომელსაც მისი არსიდან გამომდინარე (რასაც ჩვენ ქვემოთ დავინახავთ) შემთხვევითი სიდიდის *საშუალო მნიშვნელობასაც* ეძახიან.

ალბათობის თეორიის ძალიან ბევრ საკითხში მოსახერხებელია შემოვიტანოთ მათემატიკური ლოდინის ცნება. როცა მოთამაშემ უნდა მიიღოს განსაზღრული თანხა, თუ მოხდება გარკვეული შემთხვევითი ხდომილება, რომლის ალბათობა ცნობილია, მაშინ მისი მათემატიკური ლოდინი არის ის თანხა, რომელიც სამართლიანად უნდა შემოუტანოს მას იმან, ვინც იყიდის მისგან მოგების შანსებს. მაგალითად, მოთამაშემ უნდა გააგოროს ერთხელ სათამაშო კამათელი და მიიღოს მოგება 6 ლარი, თუ მოვა ციფრი 4. ადვილი დასანახია, რომ მისი მათემატიკური ლოდინი ტოლია 1 ლარის, ე. ი. იმ თანხის (6 ლარის), რომლიც შეიძლება მიიღოს მოთამაშემ, ნამრავლი სასურველი შედეგის ალბათობაზე (1/6ზე).

მართლაც, დავუშვათ, რომ ბანკომატი გვთავაზობს გავაგოროთ კამათელი და ყოველ მსურველს აძლევს შესაძლებლობას დადოს სანაძლეო მის მიერ შერჩეულ წახნაგზე (ქულაზე), რათა მოგების შემთხვევაში მიიღოს 6 ლარი. თუ 6 სხვადასხვა მოთამაშე დადებს ფსონს შესაბამისად 6 სხვადასხვა წახნაგზე, მაშინ ბანკომატმა ნებისმიერ შემთხვევაში უნდა გადაიხადოს 6 ლარი, ვინაიდან იქნება ერთი და მხოლოდ ერთი მოგებული. იმისათვის რომ თამაში იყოს სამართლიანი, საჭიროა რომ 6 მოთამაშიდან თითოეულმა შეიტანოს ბანკომატში 1 ლარი, რადგანაც არ არსებობს არანაირი საფუძველი იმისათვის, რომ რომელიმე მათგანმა გადაიხადოს სხვაზე მეტი ან ნაკლები, ვინაიდან სათამაშო კამათლის ექვსივე წახნაგი ტოლალბათურია. აქედან ჩვენ ვასკვნით, რომ მათემატიკური ლოდინი თითოეული მოთამაშისათვის შეადგენს 1 ლარს.

განმარტება 1. $X:\Omega\to R^1$ დისკრეტული შემთხვევითი სიდიდის მათემატიკური ლოდინი აღინიშნება EX სიმბოლოთი (E არის პირველი ასო ინგლისური სიტყვისა *Expectation*, რომელიც ნიშნავს – ლოდინი, მოსალოდნელობა) და ეწოდება რიცხვს:

$$EX=\sum_{\omega\in\Omega}X(\omega)P(\omega)\,, \tag{1}$$

ე. ი. შემთხვევითი სიდიდის მათემატიკური ლოდინი წარმოადგენს შემთხვევითი სიდიდის მნიშვნელობების შეწონილ ჯამს წონებით, რომლებიც ტოლია შესაბამისი ელემენტარული ხდომილებების ალბათობების.

შევნიშნავთ, რომ მათემატიკური ლოდინის აღსანიშნავად ასევე გამოიყენება სიმბოლო MX (M არის პირველი ასო რუსული სიტყვისა Математическое ожидание).

მაგალითი 1. გამოვთვალოთ სათამაშო კამათელზე მოსული ქულათა რიცხვის მათემატიკური ლოდინი. (1) თანაფარდობიდან გამომდინარე გვაქვს:

$$EX = 1\times\frac{1}{6} + 2\times\frac{1}{6} + 3\times\frac{1}{6} + 4\times\frac{1}{6} + 5\times\frac{1}{6} + 6\times\frac{1}{6} = 3.5 .$$

თეორემა1. თუ შემთხვევითი სიდიდე ღებულობს მნიშვნელობებს $x_1, x_2, \ ..., x_n$, მაშინ სამართლიანია თანაფარდობა:

$$EX = \sum_{i=1}^{n} x_i P\{X = x_i\}, \qquad (2)$$

ე. ი. შემთხვევითი სიდიდის მათემატიკური ლოდინი წარმოადგენს შემთხვევითი სიდიდის მნიშვნელობების შეწონილ ჯამს წონებით, რომლებიც ტოლია ალბათობის იმისა, რომ შემთხვევითი სიდიდე ღებულობს გარკვეულ მნიშვნელობებს.

თუ შემოვიღებთ აღნიშვნებს $P\{X = x_i\} := p_i$, $i = 1, 2, ..., m$ მაშინ (2) თანაფარდობა ასე გადაიწერება

$$EX = \sum_{i=1}^{n} x_i p_i = x_1 p_1 + x_2 p_2 + \cdots + x_n p_n . \qquad (3)$$

განსხვავებით (1) თანაფარდობისაგან, სადაც აჯამვა ხდება უშუალოდ ელემენტარული ხდომილებების მიმართ, ხდომილება $\{X = x_i\} = \{\omega : X(\omega) = x_i\}$ შეიძლება შედგებოდეს რამოდენიმე ელემენტარული ხდომილებისაგან. ხშირ შემთხვევაში (2) თანაფარდობით განიმარტება მათემატიკური ლოდინი, თუმცა მათემატიკური ლოდინის თვისებების შესამოწმებლად უფრო მოხერხებულია (1) თანაფარდობა.

თეორემა 1ის დამტკიცება. დავაჯგუფოთ (1) თანაფარდობაში შემთხვევითი სიდიდის ერთი და იგივე მნიშვნელობიანი წევრები:

$$EX = \sum_{i=1}^{n} \Big(\sum_{\omega : X(\omega) = x_i} X(\omega) P(\omega) \Big).$$

ვინაიდან მუდმივი გადის ჯამის ნიშნის გარეთ, ამიტომ

$$\sum_{\omega : X(\omega) = x_i} X(\omega) P(\omega) = \sum_{\omega : X(\omega) = x_i} x_i P(\omega) = x_i \sum_{\omega : X(\omega) = x_i} P(\omega) .$$

მეორეს მხრივ, ალბათობის განმარტების თანახმად:

$$\sum_{\omega : X(\omega) = x_i} P(\omega) = P(X = x_i).$$

ორი უკანასკნელი თანაფარდობის გაერთიანება გვაძლევს:

$$EX = \sum_{i=1}^{n} \Big(x_i \sum_{\omega : X(\omega) = x_i} P(\omega) \Big) = \sum_{i=1}^{n} x_i P\{X = x_i\}. \qquad \blacksquare$$

მათემატიკური ლოდინის ცნება ალბათურსტატისტიკურ თეორიაში შეესაბამება სიმძიმის ცენტრის ცნებას მექანიკაში. რიცხვითი ღერძის $x_1, x_2, ..., \ x_n$ წერტილებში განვათავსოთ შესაბამისად

$$P\{X=x_1\}, P\{X=x_2\}, ..., P\{X=x_n\}$$

მასები. მაშინ (2) თანაფარდობა გვიჩვენებს, რომ მატერიალური წერტილე-ბის ამ სისტემის სიმძიმის ცენტრი ემთხვევა მათემატიკურ ლოდინს. ეს, თავის მხრივ, გვიჩვენებს განმარტება 1-ის ბუნებრიობას.

იმისათვის, რომ გასაგები გახდეს მათემატიკური ლოდინის შინაარ-სი, დავუშვათ, რომ ჩავატარეთ n დაკვირვება (ექსპერიმენტი) X შემთხვე-ვით სიდიდეზე და ვთქვათ, რომ მან n_1 ჯერ მიიღო მნიშვნელობა x_1, n_2 ჯერ – მნიშვნელობა x_2, და ა. შ. n_m ჯერ – მნიშვნელობა x_m. ცხადია $n_1+n_2+\cdots+n_m = \; =n$, ხოლო შემთხვევითი სიდიდის მიერ მიღებული მნიშვ-ნელობების საშუალო არითმეტიკული $\overline{x}$ გამოითვლება ფორმულით

$$\overline{x}=\frac{x_1n_1+x_2n_2+\cdots+x_mn_m}{n},$$

ანუ,

$$\overline{x}=x_1\frac{n_1}{n}+x_2\frac{n_2}{n}+\cdots+x_m\frac{n_m}{n}. \qquad (4)$$

აქ $\frac{n_1}{n}$ არის x_1 ის განხორციელების ფარდობითი სიხიშირე, $\frac{n_2}{n}$ არის x_2 ის განხორციელების ფარდობითი სიხიშირე და ა. შ. $\frac{n_m}{n}$ არის x_m ის განხორ-ციელების ფარდობითი სიხიშირე. თუ დავუშვებთ, რომ დაკვირვებათა რა-ოდენობა საკმარისად დიდია, მაშინ ფარდობითი სიხშირე ახლოსაა ხდო-მილების ალბათობასთან

$$\frac{n_1}{n}=p_1, \frac{n_2}{n}=p_2, \ldots, \frac{n_m}{n}=p_m.$$

თუ ახლა (4) თანაფარდობაში ფარდობით სიხშირეებს შევცვლით შესაბა-მისი ალბათობებით და გავითვალისწინებთ (3) თანაფარდობას, მივიღებთ, რომ

$$\overline{x}=x_1p_1+x_2p_2+\cdots+x_np_n=EX.$$

ე. ი. შემთხვევითი სიდიდის მათემატიკური ლოდინი დაახლოებით ტოლია ამ შემთხვევითი სიდიდის დაკვირვებული მნიშვნელობების საშუა-ლო არითმეტიკულის.

ცხადია, რომ არაა აუცილებელი შემთხვევითი სიდიდის მათემატი-კური ლოდინი ტოლი იყოს მისი რომელიმე შესაძლო მნიშვნელობის.

თუ დისკრეტული ტიპის შემთხვევითი სიდიდე თანაბარი კანონითაა განაწილებული ანუ ის ყველა თავის მნიშვნელობას $x_1, x_2, ..., x_n$ ღებულობს თანაბარი (ერთი და იგივე) ალბათობებით ($p_1=p_2=\cdots=p_n=1/n$), მაშინ მა-თემატიკური ლოდინი ზუსტად ემთხვევა მისი მნიშვნელობების საშუალო არითმეტიკულს:

$$EX=x_1\frac{1}{n}+x_2\frac{1}{n}+\cdots+x_n\frac{1}{n}=\frac{x_1+x_2+\cdots+x_n}{n}.$$

განმარტება 2. თუ დისკრეტული ტიპის შემთხვევითი სიდიდის შე-საძლო მნიშვნელობათა სიმრავლე თვლადია, მაშინ

$$EX = x_1 p_1 + x_2 p_2 + \cdots + x_n p_n + \cdots = \sum_{i=1}^{\infty} x_i p_i ,$$

თუ ცნობილია, რომ შესაბამისი მწკრივი აბსოლუტურად კრებადია –

$$\sum_{i=1}^{\infty} | x_i | p_i < \infty ,$$

სადაც $p_i := P\{X = x_i\}, i = 1, 2, ...$ და $p_1 + p_2 + \cdots = 1$.

განმარტება 3. უწყვეტი ტიპის X შემთხვევითი სიდიდის მათემატიკური ლოდინი ეწოდება რიცხვს

$$EX = \int_{-\infty}^{+\infty} x f(x) dx ,$$

სადაც $f(x)$ არის X შემთხვევითი სიდიდის განაწილების სიმკვრივე, თუ ცნობილია, რომ

$$\int_{-\infty}^{+\infty} | x | \, f(x) dx < \infty .$$

თეორემა 2. თუ X და Y ერთი და იგივე ელემენტარულ ხდომილებათა სივრცეზე განმარტებული შემთხვევითი სიდიდებია, ხოლო $c = const$ რაიმე მუდმივია, მაშინ:

ა). $Ec = c$;

ბ). $E(X+Y) = EX + EY$ და $E(X-Y) = EX - EY$;

გ). $E(cX) = cEX$; დ). $E(X - EX) = 0$ და

ე). $E(X-c)^2 = E(X-EX)^2 + (c - EX)^2$.

დამტკიცება. ა). ამ შემთხვევაში საქმე გვაქვს მუდმივ შემთხვევით სიდიდესთან $X(\omega) = c$, ანუ ფუნქცია $X(\omega)$ ასახავს ელემენტარულ ხდომილებათა სივრცეს Ω ერთადერთ c წერტილში. ვინაიდან მუდმივი მამრავლი შეგვიძლია გამოვიტანოთ ჯამის ნიშნის გარეთ, ამიტომ გვაქვს:

$$Ec = \sum_{\omega \in \Omega} cP(\omega) = c \sum_{\omega \in \Omega} P(\omega) = cP(\Omega) = c .$$

ბ). თუ ჯამის (შესაბამისად, სხვაობის) ყველა წევრი წარმოიდგინება ორ შესაკრებად (შესაბამისად, სხვაობად), მაშინ მთელი ჯამიც წარმოიდგინება ორი ჯამის ჯამად (შესაბამისად, სხვაობად), რომელთაგან პირველი შედგება თითოეული წევრის პირველი შესაკრებებისაგან (შესაბამისად, საკლებებისაგან), ხოლო მეორე – მეორე შესაკრებებისაგან (შესაბამისად, მაკლებებისაგან). ამიტომ

$$E(X+Y) = \sum_{\omega \in \Omega} [X(\omega) + Y(\omega)] P(\omega) = \sum_{\omega \in \Omega} X(\omega) P(\omega) + \sum_{\omega \in \Omega} Y(\omega) P(\omega) = EX + EY$$

და ანალოგიურად, $E(X-Y) = EX - EY$.

ე. ი. შემთხვევით სიდიდეთა ჯამის (შესაბამისად, სხვაობის) მათემატიკური ლოდინი მათი მათემატიკური ლოდინების ჯამის (შესაბამისად, სხვაობის) ტოლია.

გ). $E(cX) = \sum_{\omega \in \Omega} cX(\omega) P(\omega) = c \sum_{\omega \in \Omega} X(\omega) P(\omega) = cEX$.

დ). ბ) და ა) პუნქტების თანახმად გვაქვს

$$E(X-EX)\overset{\text{ბ).}}{=}EX-E(EX)\overset{\text{ა).}}{=}EX-EX=0.$$

ე). ვინაიდან

$$(X-c)^2=[(X-EX)+(EX-c)]^2=$$
$$=(X-EX)^2+2(X-EX)(EX-c)+(EX-c)^2.$$

ამიტომ

$$E(X-c)^2\overset{\text{ბ).}}{=}E(X-EX)^2+E[2(X-EX)(EX-c)]+E(EX-c)^2\overset{\text{გ).,ა).}}{=}$$

$$\overset{\text{გ).,ა).}}{=}E(X-EX)^2+2(EX-c)E(X-EX)+(EX-c)^2\overset{\text{დ).}}{=}E(X-EX)^2+(EX-c)^2. \quad \blacksquare$$

შედეგი 1. ბ) და გ) პუნქტების გაერთიანება გვაძლევს, რომ შემთხვევით სიდიდეთა წრფივი კომბინაციის მათემატიკური ლოდინი ტოლია მათი მათემატიკური ლოდინების წრფივი კომბინაციის:

$$E(aX+bY)=aEX+bEY,$$

სადაც a, b მუდმივებია.

დავალება. დაამტკიცეთ შედეგი 1 პირდაპირი გზით მათემატიკური ლოდინის განმარტების გამოყენებით.

შედეგი 2. ვინაიდან ე) პუნქტის თანაფარდობის მარჯვენა მხარეში მეორე შესაკრები ყოველთვის არაუარყოფითია და ნულია მხოლოდ მაშინ, როცა $c=EX$, ამიტომ გამოსახულება $E(X-c)^2$ თავის მინიმუმს c-ს მიმართ აღწევს როცა $c=EX$:

$$\min_{c\in(-\infty,+\infty)} E(X-c)^2=E(X-EX)^2.$$

შემთხვევითი სიდიდის ფუნქციის მათემატიკური ლოდინი. ხშირად მოცემულია რაიმე X შემთხვევითი სიდიდის განაწილების კანონი და გვაინტერესებს $Y=g(X)$ შემთხვევითი სიდიდის მათემატიკური ლოდინი, სადაც $g(x)$ ნამდვილი x ცვლადის რაიმე ფუნქციაა. ამისათვის ჯერ შეიძლება დავადგინოთ Y შემთხვევითი სიდიდის განაწილება და შემდეგ გამოვთვალოთ მისი მათემატიკური ლოდინი განმარტება 1ის გამოყენებით. მაგრამ უფრო მოხერხებულია $Y=g(X)$ შემთხვევითი სიდიდის მათემატიკური ლოდინი გამოვთვალოთ X შემთხვევითი სიდიდის განაწილების ტერმინებში.

ვთქვათ, X შემთხვევითი სიდიდის განაწილების კანონია:

x_i	x_1	x_2	$\cdots$	x_n
p_i	p_1	p_2	$\cdots$	p_n

მაშინ

$$Eg(X)=\sum_{i=1}^{m} g(x_i)P\{X=x_i\}=\sum_{i=1}^{m} g(x_i)p_i. \qquad (5)$$

ამ ფაქტის შესამოწმებლად ვისარგებლოთ მათემატიკური ლოდინის განმარტებით და დავაჯგუფოთ ის წევრები, რომლებიც შეესაბამებიან $X(\omega)$-ს ერთი და იგივე მნიშვნელობას:

$$Eg(X)=\sum_{\omega\in\Omega}g(X(\omega))P(\omega)=\sum_{i=1}^{m}(\sum_{\omega:X(\omega)=x_i}g(X(\omega))P(\omega)).$$

თუ კი აქ მუდმივ თანანმამრავლს გავიტანთ ჯამის ნიშნის გარეთ და ვისარგებლებთ ალბათობის განმარტებით, მივიღებთ შესამოწმებელ თანაფარდობას:

$$Eg(X)=\sum_{i=1}^{m}(\sum_{\omega:X(\omega)=x_i}g(x_i)P(\omega))=\sum_{i=1}^{m}(g(x_i)\sum_{\omega:X(\omega)=x_i}P(\omega))=\sum_{i=1}^{m}g(x_i)P\{X=x_i\}.$$

მაგალითი 2. დავუშვათ, $g(x)=x^3-4x$ და მოცემულია X შემთხვევითი სიდიდის განაწილების კანონი:

-2	-1	0	2
0.1	0.3	0.4	0.2

დავადგინოთ $Y=g(X)$ შემთხვევითი სიდიდის განაწილების კანონი და გამოვთვალოთ მისი მათემატიკური ლოდინი.

ცხადია, რომ $g(-2)=g(0)=g(2)=0$ და $g(-1)=3$. ამიტომ Y შემთხვევითი სიდიდის შესაძლო მნიშვნელობებია 0 და 3. დავადგინოთ მისი განაწილების კანონი. ამისათვის გამოვთვალოთ ალბათობები: $P\{Y=0\}$ და $P\{Y=3\}$. რადგან ხდომილებები $\{X=-2\},\{X=0\}$ და $\{X=2\}$ უთავსებადია, ამიტომ ალბათობათა შეკრების წესის თანახმად გვექნება:

$$P\{Y=0\}=P\{\{X=-2\}\cup\{X=0\}\cup\{X=2\}\}=$$
$$=P\{X=-2\}+P\{X=0\}+P\{X=2\}=0.1+0.4+0.2=0.7.$$

გარდა ამისა, $P\{Y=3\}=P\{X=-1\}=0.3$. ამიტომ $Y=g(X)$ შემთხვევითი სიდიდის განაწილების კანონს აქვს სახე:

0	3
0.7	0.3

ხოლო (3) თანაფარდობის ძალით $EY=0\cdot0.7+3\cdot0.3=0.9$.

ახლა გამოვთვალოთ $Y=g(X)$ შემთხვევითი სიდიდის მათემატიკური ლოდინი (5) თანაფარდობის საშუალებით. გვექნება:

$$EY=g(-2)\cdot0.1+g(-1)\cdot0.3+g(0)\cdot0.4+g(2)\cdot0.2=0\cdot0.1+3\cdot0.3+0\cdot0.4+0\cdot0.2=0.9.$$

მაგალითი 3. გამოვთვალოთ პუასონის განაწილების მათემატიკური ლოდინი (გავიხსენოთ, რომ პუასონის განაწილებას პარამეტრით λ აქვს სახე: $P\{X=k\}=\dfrac{\lambda^k}{k!}e^{-\lambda},\quad k=0,1,2,...$).

თუ ვისარგებლებთ წარმოდგენით $e^x=\sum_{m=0}^{\infty}\dfrac{x^m}{m!}$, მაშინ გვექნება:

$$EX=\sum_{k=0}^{\infty}k\frac{\lambda^k}{k!}e^{-\lambda}=e^{-\lambda}\sum_{k=1}^{\infty}k\frac{\lambda^k}{k!}=\lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}=\lambda e^{-\lambda}e\lambda=\lambda.$$

ე. ი. პუასონის განაწილების λ პარამეტრი წარმოადგენს ამ განაწილების მათემატიკურ ლოდინს (საშუალო მნიშვნელობას).

ჩვენ ზემოთ ვნახეთ როგორაა დამოკიდებული მათემატიკური ლოდინი ათვლის წერტილის შეცვლაზე და სხვა ზომის ერთეულზე გადასვლაზე (გადასვლა $Y = ax + b$), ასევე შემთხვევითი სიდიდის ფუნქციაზე. მიღებული შედეგები მუდმივად გამოიყენება ტექნიკურეკონომიკურ ანალიზში, ორგანიზაციების საფინანსოსამეურნეო მოქმედებების შეფასებისას, საგარეოეკონომიკურ გათვლებში ერთი ვალუტიდან მეორეზე გადასვლისას, ნორმატიულტექნიკურ დოკუმენტაციაში და ა. შ. განხილული შედეგები საშუალებას იძლევა გამოყენებულ იქნეს ერთი და იგივე გამოსათვლელი ფორმულები მასშტაბების სხვადასხვა პარამეტრების და გადახრების დროს.

განმარტება 4. თუ X შემთხვევითი სიდიდის შესაძლო მნიშვნელობებია $x_1, x_2, \ ..., x_n$, ხოლო B რაიმე ხდომილებაა ($P(B) > 0$), მაშინ X შემთხვევითი სიდიდის პირობითი მათემატიკური ლოდინი B ხდომილების მიმართ აღინიშნება სიმბოლოთი $E(X \mid B)$ და განიმარტება შემდეგნაირად:

$$E(X \mid B) = \sum_{i=1}^{n} x_i P(X = x_i \mid B) . \qquad (6)$$

ცხადია, რომ $E(X \mid \Omega) = EX$. გარდა ამისა, ადვილი შესამოწმებელია, რომ:

$$E(X \mid B) = \frac{1}{P(B)} E(I_B X) .$$

განმარტება 5. X შემთხვევითი სიდიდის Y შემთხვევით სიდიდეზე **რეგრესიის მრუდი (ფუნქცია)** ეწოდება ფუნქციას $R(y) = E(X \mid Y = y)$.

რეგრესიის ფუნქციის აზრი იმაში მდგომარეობს, რომ Y შემთხვევითი სიდიდის მნიშვნელობების პირობით განისაზღვრება X შემთხვევითი სიდიდის პირობითი საშუალო მნიშვნელობა.

განვიხილოთ შემთხვევითი სიდიდე $R(Y)$ (რეგრესიის ფუნქციაში ჩასმულია Y შემთხვევითი სიდიდე). ამ შემთხვევით სიდიდეს X შემთხვევითი სიდიდის **პირობით მათემატიკურ ლოდინს** უწოდებენ Y პირობით და $E(X \mid Y)$ სიმბოლოთი აღნიშნავენ.

§20. შემთხვევით სიდიდეთა დამოუკიდებლობა

ისევე როგორც ხდომილებათა დამოუკიდებლობა, შემთხვევით სიდიდეთა დამოუკიდებლობა წარმოადგენს ალბათობის თეორიის ერთერთ საბაზო ცნებას, რომელიც საფუძვლად უდევს გადაწყვეტილებების მიღების პრაქტიკულად ყველა ალბათურ-სტატისტიკურ მეთოდს.

განმარტება 1. ერთი და იგივე ელემენტარულ ხდომილებათა სივრცეზე განმარტებულ დისკრეტულ X და Y შემთხვევით სიდიდეებს ეწოდება **დამოუკიდებელი**, თუ ნებისმიერ ნამდვილი a და b რიცხვებისათვის დამოუკიდებელია ხდომილებები $\{\omega: X(\omega)=a\}$ და $\{\omega: Y(\omega)=b\}$.

ცხადია, რომ თუ X და Y დამოუკიდებელი შემთხვევით სიდიდებია, ხოლო a,b რაიმე რიცხვებია, მაშინ შემთხვევით სიდიდები $X+a$ და $Y+b$ აგრეთვე დამოუკიდებლებია.

მართლაც, ხდომილებები $\{X+a=c\}$ და $\{Y+b=d\}$ ემთხვევა შესაბამისად ხდომილებებს $\{X=c-a\}$ და $\{Y=d-b\}$, ამიტომ ისინი დამოუკიდებლებია.

დავალება. აჩვენეთ, რომ თუ X და Y დამოუკიდებელი შემთხვევით სიდიდებია, ხოლო a_1,b_1,a_2,b_2 რაიმე რიცხვებია, მაშინ შემთხვევითი სიდიდები a_1X+b_1 და a_2Y+b_2 აგრეთვე დამოუკიდებლებია.

განმარტება 2. ერთი და იგივე ელემენტარულ ხდომილებათა სივრცეზე განმარტებულ დისკრეტულ $X,Y,Z,...$ შემთხვევით სიდიდეებს ეწოდება **ერთობლივად დამოუკიდებელი,** თუ ერთობლივად დამოუკიდებელია ხდომილებები $\{X=a\},\{Y=b\}, \{Z=c\},...$.

მაგალითი 1. შემთხვევითი სიდიდეები, რომლებიც განიმარტებიან დამოუკიდებელ ცდათა სქემაში სხვადასხვა ცდის შედეგების მიხედვით, თვითონაც დამოუკიდებლებია. ეს გამოდის იქიდან, რომ ხდომილებები, რომელთა საშუალებითაც განიმარტება შემთხვევით სიდიდეთა დამოუკიდებლობა, განისაზღვრებიან სხვადასხვა ცდების შედეგების მიხედვით, და მაშასადამე, ისინი დამოუკიდებლები არიან თვითონ დამოუკიდებელ ცდათა სქემის განმარტების თანახმად.

გადაწყვეტილებების მიღების ალბათურსტატისტიკურ მეთოდებში მუდმივად გამოიყენება შემდეგი ფაქტი: თუ X და Y დამოუკიდებელი შემთხვევითი სიდიდებია, ხოლო $g(X)$ და $h(Y)$ შემთხვევითი სიდიდეებია, რომლებიც მიიღებიან X და Y შემთხვევით სიდიდების ჩასმით ნამდვილი ცვლადის g და h ფუნქციებში, მაშინ $g(X)$ და $h(Y)$ აგრეთვე დამოუკიდებელი შემთხვევითი სიდიდეებია. მაგალითად, თუ X და Y დამოუკიდებელია, მაშინ დამოუკიდებელია X^2 და $2Y+3$, $\ln|X|$ და 3^Y, და ა. შ.

როგორც უკვე ავღნიშნეთ ალბათურსტატისტიკური მეთოდების უმრავლესობა, რომლებიც გამოიყენება პრაქტიკაში, დაფუძნებულია შემთხვევითი სიდედეების დამოუკიდებლობის ცნებაზე, რამდენადაც, დაკვირვებების, გაზომვების, ექსპერიმენტების, ანალიზებისა და ცდების შედეგები ჩვეულებრივ მოდელირდება დამოუკიდებელი შემთხვევითი

სიდიდეებით. ხშირად ითვლება, რომ დაკვირვებები ხორციელდება დამოუკიდებებლი ცდების სქემის მიხედვით. მაგალითად, ორგანიზაციების საფინანსოსამეურნეო ქმედებების შედეგები, მუშახელის გამომუშავება, შესამოწმებელი ნაწარმის (რომელიც ამორჩეულია ტექნოლოგიური პროცესის სტატისტიკური რეგულირების დროს) საკონტროლო პარამეტრების გაზომვის შედეგები (მონაცემები), მარკეტინგული გამოკითხვების დროს გამოკითხული მომხმარებლების პასუხები და სხვა ტიპის მონაცემები, რომლებიც გამოიყენება გადაწყვეტილებების მიღების დროს, ჩვეულებრივ განიხილება როგორც დამოუკიდებელი შემთხვევითი სიდიდეები (ვექტორები ან ელემენტები). შემთხვევით სიდიდეთა ცნების ასეთი პოპულარობის მიზეზი მდგომარეობს იმაში, რომ ამ მომენტისათვის კვლევის შესაბამისი თეორია გაცილებით წინაა წასული დამოუკიდებელი შემთხვევითი სიდიდეებისათვის, ვიდრე დამოკიდებულებისათვის.

თეორემა 1. თუ X და Y დამოუკიდებელი შემთხვევით სიდიდებია, მაშინ მათი ნამრავლის მათემატიკური ლოდინი თითოეულის მათემატიკური ლოდინების ნამრავლის ტოლია, ე. ი.

$$E(XY) = EX \cdot EY .$$

დამტკიცება. დავუშვათ, რომ X შემთხვევითი სიდიდის მნიშვნელობებია $x_1, x_2, ..., x_n$, ხოლო Y შემთხვევითი სიდიდე კი ღებულობს მნიშვნელობებს y_1, $y_2, ..., y_n$. XY ნამრავლის მათემატიკური ლოდინის მომცემ ჯამში დავაჯგუფოთ ის წევრები, რომლებშიც X და Y ღებულობენ ფიქსირებულ მნიშვნელობებს, მუდმივი მამრავლები გავიტანოთ ჯამის ნიშნის გარეთ და გავიხსენოთ ალბათობის განმარტება. მივიღებთ:

$$E(XY) = \sum_{\omega \in \Omega} X(\omega) Y(\omega) P(\omega) = \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{\omega : X(\omega)=x_i, Y(\omega)=y_j} X(\omega) Y(\omega) P(\omega) =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{\omega : X(\omega)=x_i, Y(\omega)=y_j} x_i y_j P(\omega) = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j \sum_{\omega : X(\omega)=x_i, Y(\omega)=y_j} P(\omega) =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j P\{X = x_i, Y = y_j\} .$$

ვინაიდან X და Y დამოუკიდებელი შემთხვევით სიდიდებია, ამიტომ

$$P\{X = x_i, Y = y_j\} = P\{X = x_i\} P\{Y = y_j\} .$$

მეორეს მხრივ, თუ ვისარგებლებთ ჯამის სიმბოლოს შემდეგი თვისებით:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j = \sum_{i=1}^{n} a_i \cdot \sum_{j=1}^{m} b_j ,$$

საბოლოოდ, თეორემა 19.1ის ძალით, მივიღებთ, რომ

$$E(XY) = \sum_{i=1}^{n} \sum_{j=1}^{m} (x_i P\{X = x_i\})(y_j P\{Y = y_j\}) =$$

$$= \sum_{i=1}^{n} x_i P\{X = x_i\} \cdot \sum_{j=1}^{m} y_j P\{Y = y_j\} = EX \cdot EY . \qquad \blacksquare$$

აღსანიშნავია, რომ თეორემა 1ის შებრუნებული თეორემა არაა მართებული. მოვიყვანოთ შესაბამისი მაგალითი.

მაგალითი 2. დავუშვათ, რომ ელემენტარულ ხდომილებათა სივრცე შედგება სამი ტოლალბათური ელემენტარული ხდომილებისაგან $\Omega = \{\omega_1, \omega_2, \omega_3\}$, $P(\omega_1) = P(\omega_2) = P(\omega_3) = 1/3$. განვმარტოთ X და Y შემთხვევითი სიდიდეები შემდეგნაირად:

$$X(\omega_1) = 1, X(\omega_2) = 0, X(\omega_3) = -1;$$

$$Y(\omega_1) = 1, Y(\omega_2) = 0, Y(\omega_3) = 1.$$

მაშინ გასაგებია, რომ $XY = X$, $E(XY) = EX = 1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + (-1) \cdot \frac{1}{3} = 0$.

შესაბამისად, $E(XY) = EX \cdot EY$. მეორეს მხრივ,

$$P\{X = 0\} = P\{Y = 0\} = P\{X = 0, Y = 0\} = P(\omega_2) = 1/3,$$

მაშინ როდესაც X და Y შემთხვევითი სიდიდეები რომ იყვნენ დამოუკიდებლები $\{X = 0, Y = 0\}$ ხდომილების ალბათობა უნდა ყოფილიყო

$$\frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}.$$

მაგალითი 3. დამოკიდებული შემთხვევითი სიდიდეებისათვის საზოგადოდ:

$$E(\xi\eta) \neq E\xi \cdot E\eta.$$

მართლაც, დავუშვათ, რომ ξ ბერნულის შემთხვევითი სიდიდეა წარმატების ალბათობით 1/2: $P(\xi = 1) = P(\xi = 0) = 1/2$, ხოლო $\eta = 1 - \xi$. ცხადია, რომ η აგრეთვე ბერნულის შემთხვევითი სიდიდეა, მაგრამ ის დამოკიდებულია (ფუნქციონალურადაც კი) ξ-ზე. ამასთანავე,

$$\xi\eta \equiv 0, \; E\xi = E\eta = 1/2, \; E(\xi\eta) = E0 = 0 \neq E\xi \cdot E\eta = 1/4.$$

§21. შემთხვევითი სიდიდის დისპერსია

მათემატიკური ლოდინი გვიჩვენებს თუ რომელი წერტილის (მნიშვნელობის) ირგვლივ ჯგუფდება (ლაგდება) შემთხვევითი სიდიდის მნიშვნელობები. ხშირ შემთხვევაში საჭიროა შეგვეძლოს შემთხვევითი სიდიდის მნიშვნელობების ცვლილების გაზომვა მათემატიკური ლოდინის მიმართ. განვიხილოთ ორი დისკრეტული ტიპის შემთხვევითი სიდიდე განაწილების შემდეგი კანონებით:

X	-3	1
P	1/4	3/4

Y	-90	45
P	1/3	2/3

გამოვთვალოთ თითოეულის მათემატიკური ლოდინი:

$$EX=(-3)\cdot 1/4+1\cdot 3/4=0 \ \text{ და } \ EY=(-90)\cdot 1/3+45\cdot 2/3=0\,.$$

როგორც ვხედავთ ორივე შემთხვევით სიდიდეს აქვს ერთი და იგივე მათემატიკური ლოდინი, მაგრამ მათი განაწილებები განსხვავდებიან იმით, რომ X შემთხვევითი სიდიდის შესაძლო მნიშვნელობები გაცილებით ახლოსაა მათემატიკურ ლოდინთან (ამ შემთხვევაში ნულთან), ვიდრე Y შემთხვევითი სიდიდის მნიშვნელობები.

შემთხვევითი სიდიდის მნიშვნელობების მათემატიკური ლოდინის მიმართ გაფანტულობის ერთ-ერთ მნიშვნელოვან საზომს წარმოადგენს შემთხვევითი სიდიდის დისპერსია. ჩვენ უკვე ვნახეთ, რომ გამოსახულება $E(X-c)^2$ აღწევს მინიმუმს c-ს მიმართ როცა $c=EX$. ამიტომ შემთხვევითი სიდიდის მნიშვნელობების გაფანტულობის საზომად ბუნებრივია ავიღოთ $E(X-EX)^2$.

განმარტება 1. X შემთხვევითი სიდიდის დისპერსია (აღინიშნება DX-ით, D არის პირველი ასო ინგლისური სიტყვისა -- Dispersion) ეწოდება $(X-EX)^2$ შემთხვევითი სიდიდის მათემატიკურ ლოდინს

$$DX=E(X-EX)^2\,. \tag{1}$$

მათემატიკური ლოდინის თვისებების გამოყენებით დისპერსია შესაძლებელია გადაიწეროს სხვა ფორმით:

$$\begin{aligned} DX&=E(X-EX)^2=E[X^2-2X\cdot EX+(EX)^2]= \\ &=EX^2-2EX\cdot EX+(EX)^2=EX^2-(EX)^2\,, \end{aligned} \tag{2}$$

სადაც, EX^2-ს ეწოდება X შემთხვევითი სიდიდის **მეორე რიგის მომენტი** (თვითონ დისპერსიას უწოდებენ აგრეთვე – **მეორე რიგის ცენტრალურ მომენტს**).

თუ დისკრეტული ტიპის X შემთხვევითი სიდიდის განაწილების კანონია

x_i	x_1	x_2	$\cdots$	x_n
p_i	p_1	p_2	$\cdots$	p_n

მაშინ მათემატიკური ლოდინისა და შემთხვევითი სიდიდის ფუნქციებიდან მათემატიკური ლოდინის გამოსათვლელი ფორმულების თანახმად დისპე-

რსიის გამოსათვლელ ფორმულებს (1) და (2) ფორმულების მიხედვით ექნება შესაბამისად შემდეგი სახე:

$$DX=\sum_{i=1}^{n}(x_i-\sum_{j=1}^{n}x_j p_j)^2 p_i\ , \tag{3}$$

$$DX=\sum_{i=1}^{n}x_i^2 p_i-(\sum_{j=1}^{n}x_j p_j)^2\ . \tag{4}$$

მაგალითი 1. დისკრეტული ტიპის X შემთხვევითი სიდიდის განაწილების კანონია

x_i	-1	0	1	2	3
p_i	0.1	0.15	0.3	0.25	0.2

გამოვთვალოთ მისი დისპერსია.

ვინაიდან დისპერსიის გამოსათვლელად გვაქვს ორი (3) და (4) ფორმულები, შესაბამისად, გვექნება დისპერსიის გამოთვლის ორი ხერხი. მოხერხებულია ეს გამოთვლები ჩაიწეროს ცხრილების სახით.

დისპერსიის გამოთვლის პირველი ხერხი:

i	x_i	p_i	$x_i p_i$	$(x_i-EX)^2$	$(x_i-EX)^2 p_i$
1	-1	0.10	-0.1	5.29	0.5290
2	0	0.15	0	1.69	0.2535
3	1	0.30	0.3	0.09	0.0270
4	2	0.25	0.5	0.49	0.1225
5	3	0.20	0.6	2.89	0.5780
			$EX=\sum_{i=1}^{5}x_i p_i=1.3$		$DX=\sum_{i=1}^{5}(x_i-EX)^2 p_i=1.51$

დისპერსიის გამოთვლის მეორე ხერხი:

i	x_i	p_i	$x_i p_i$	x_i^2	$x_i^2 p_i$
1	-1	0.10	-0.1	1	0.1
2	0	0.15	0	0	0
3	1	0.30	0.3	1	0.3
4	2	0.25	0.5	4	1
5	3	0.20	0.6	9	1.8
			$EX=\sum_{i=1}^{5}x_i p_i=1.3$		$EX^2=\sum_{i=1}^{5}x_i^2 p_i=3.2$
			$DX=EX^2-(EX)^2=1.51$		

დავადგინოთ დისპერსიის თვისებები, რომლებიც მუდმივად გამოიყენება გადაწყვეტილებების მიღების ალბათურ-სტატისტიკურ მეთოდებში.

I. მუდმივის დისპერსია ნულის ტოლია -- $Dc=0$. მართლაც,

$$Dc=E(c-Ec)^2=E(c-c)^2=E0=0;$$

II. $D(aX+b)=a^2DX$. მართლაც, მათემატიკური ლოდინის ცნობილი თვისებების გამოყენებით გვაქვს:

$$D(aX+b)=E[(aX+b)-E(aX+b)]^2=E[aX+b-aEX-b)]^2=E[a(X-EX)]^2=$$
$$=E[a^2(X-EX)^2]=a^2E(X-EX)^2=a^2DX.$$

როგორც ვხედავთ, შემთხვევით სიდიდეზე მუდმივის დამატება მის დისპერსიას არ ცვლის (თუ ავიღებთ $a=1$, მაშინ მივიღებთ $D(X+b)=DX$), ხოლო მუდმივი მამრავლი დისპერსიის ნიშნის გარეთ გადის კვადარატში ახარისხებუბული (თუ ავიღებთ $b=0$, მივიღებთ $D(aX)=a^2DX$). კერძოდ, ეს ფორმულა გვიჩვენებს როგორ იცვლება დაკვირვების შედეგების დისპერსია ათვლის საწყისი წერტილისა და გაზომვის ერთეულის ცვლილებისას. ის გვაძლევს გამოსათვლელი ფორმულების გარდაქმნის წესს მასშტაბისა და გადატანის სხვა პარამეტრებზე გადასვლის დროს.

თეორემა 1. თუ X და Y დამოუკიდებელი შემთხვევითი სიდიდეებია, მაშინ მათი ჯამის დისპერსია თითოეულის დისპერსიების ჯამია

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

დამტკიცება. ვისარგებლოთ იგივეობით

$$[(X+Y)-(EX+EY)]^2=[(X-EX)+(Y-EY)]^2=$$
$$=(X-EX)^2+2(X-EX)(Y-EY)+(Y-EY)^2.$$

მაშინ მათემატიკური ლოდინის თვისებების ძალით, ვღებულობთ:

$$D(X+Y)=E[(X+Y)-E(X+Y)]^2=$$
$$=E(X-EX)^2+2E[(X-EX)(Y-EY)]+E(Y-EY)^2=DX+DY+2E[(X-EX)(Y-EY)]$$

როგორც ცნობილია, თუ X და Y დამოუკიდებელი შემთხვევითი სიდიდეებია, მაშინ აგრეთვე დამოუკიდებელია $X+a$ და $Y+b$ და $E(XY)=EX\cdot EY$. გარდა ამისა, $E(X-EX)=0$. ამიტომ საბოლოოდ გვაქვს:

$$D(X+Y)=DX+DY+2E(X-EX)\cdot E(Y-EY)]=DX+DY. \qquad \blacksquare$$

დავალება. შეამოწმეთ, რომ თუ X და Y დამოუკიდებელი შემთხვევითი სიდიდეებია, მაშინ მათი სხვაობის დისპერსია თითოეულის დისპერსიების ჯამია

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

თეორემა 2. თუ $X_1,X_2,...,X_n$ -- წყვილ-წყვილად დამოუკიდებელი შემთხვევითი სიდიდეებია (ე. ი. X_i და X_j დამოუკიდებელია, თუ $i\neq j$). მაშინ ჯამის დისპერსია ტოლია დისპერსიების ჯამის

$$D(X_1+X_2+\cdots+X_n)=DX_1+DX_2+\cdots+DX_n.$$

დამტკიცება. შემოვიღოთ აღნიშვნები $a_i=X_i-EX_i$ $(i=1,2,...,n)$. მაშინ თუ ვისარგებლებთ აჯამვის სიმბოლოს შემდეგი თვისებით:

$$(\sum_{i=1}^{n}a_i)^2=(\sum_{i=1}^{n}a_i)\cdot(\sum_{j=1}^{n}a_j)=\sum_{i=1}^{n}a_i^{\,2}+\sum_{i\neq j}a_ia_j.$$

მივიღებთ, რომ

$$(X_1+X_2+\cdots+X_n-EX_1-EX_2-\cdots-EX_n)^2=$$
$$=\sum_{i=1}^{n}(X_i-EX_i)^2+\sum_{i\neq j}(X_i-EX_i)(X_j-EX_j).$$

თუ ახლა გავითვალისწინებთ, რომ ჯამის მათემატიკური ლოდინი მათემატიკური ლოდინების ჯამის ტოლია, ხოლო დამოუკიდებელი შემთხვევითი სიდიდეების ნამრავლის მათემატიკური ლოდინი მათემატიკური ლოდინების ნამრავლის ტოლია, მივიღებთ შესამოწმებელ თანაფარდობას:

$$D(X_1+X_2+\cdots+X_n)=E[X_1+X_2+\cdots+X_n-E(X_1+X_2+\cdots+X_n)]^2=$$
$$=E\sum_{i=1}^{n}(X_i-EX_i)^2+E\sum_{i\neq j}(X_i-EX_i)(X_j-EX_j)=$$
$$=\sum_{i=1}^{n}E(X_i-EX_i)^2+\sum_{i\neq j}E(X_i-EX_i)\cdot E(X_j-EX_j)=\sum_{i=1}^{n}DX_i\,,$$

სადაც ბოლო ეტაპზე ჩვენ ვისარგებლეთ იგივეობით: $E(X-EX)=0$. ■

შენიშვნა. რაც შეეხება ჯამის მათემატიკურ ლოდინს ის ყოველთვის შესაკრებთა მათემატიკური ლოდინების ჯამის ტოლია, მიუხედავად იმისა დამოუკიდებელია თუ დამოკიდებული შემთხვევითი სიდიდეები. ორი შემთხვევითი სიდიდის შემთხვევაში ჩვენ ეს უკვე შევამოწმეთ. მისი განზოგადოება ადვილად შეიძლება მათემატიკური ინდუქციის მეთოდის გამოყენებით.

დავალება. დაამტკიცეთ, რომ ნებისმიერი $X_1,X_2,...,X_n$ შემთხვევითი სიდიდეებისათვის ჯამის მათემატიკური ლოდინი მათემატიკური ლოდინების ჯამის ტოლია

$$E(X_1+X_2+\cdots+X_n)=EX_1+EX_2+\cdots+EX_n\,.$$

ეს ორი თანაფარდობა არსებით როლს თამაშობს მათემატიკურ სტატისტიკაში მონაცემთა შერჩევითი მახასიათებლების შესწავლის დროს, ვინაიდან შერჩევაში მონაწილე დაკვირვებებისა და გაზომვების შედეგები, მათემატიკურ სტატისტიკაში, გადაწყვეტილებების მიღების თეორიაში და ეკონომეტრიკაში, როგორც წესი განიხილება როგორც დამოუკიდებელი შემთხვევითი სიდიდეების რეალიზაციები.

მაგალითი 2. განვიხილოთ რაიმე A ხდომილება და X შემთხვევითი სიდიდე, ისეთი, რომ $X(\omega)=1$, თუ $\omega\in A$ და $X(\omega)=0$, თუ $\omega\notin A$ (ასეთ შემთხვევით სიდიდეს A ხდომილების **მახასიათებელი ფუნქცია** ეწოდება). ვაჩვენოთ, რომ

$$EX=P(A),\ \ DX=P(A)\cdot(1-P(A)).$$

გასაგებია, რომ ამ X შემთხვევითი სიდიდის განაწილების კანონი იქნება

x_i	1	0
p_i	$P(A)$	$1-P(A)$

ასეთი კანონით განაწილებულ შემთხვევით სიდიდეს **ბერნულის შემთხვევით სიდიდეს** უწოდებენ.

მათემატიკური ლოდინის განმარტების თანახმად გვექნება

$$EX = 1 \cdot P(A) + 0 \cdot (1 - P(A)) = P(A).$$

ანალოგიურად, $Y = (X - EX)^2 = (X - P(A))^2$ შემთხვევითი სიდიდის განაწილების კანონი იქნება:

y_i	$(1-P(A))^2$	$(P(A))^2$
p_i	$P(A)$	$1-P(A)$

შესაბამისად,

$$DX = EY = E(X - EX)^2 = (1 - P(A))^2 \cdot P(A) + (P(A)) \cdot (1 - P(A)) =$$
$$= P(A) \cdot (1 - P(A)) \cdot [1 - P(A) + P(A)] = P(A) \cdot (1 - P(A)). \qquad \blacksquare$$

მაგალითი 3. განვიხილოთ n დამოუკიდებელი ექსპერიმენტი (ცდა), რომელთაგან თითოეულში გარკვეული A ხდომილება შეიძლება მოხდეს ან არ მოხდეს. შემოვიღოთ შემთხვევითი სიდიდეები $X_1, X_2, ..., X_n$ შემდეგნაირად: $X_i(\omega) = 1$, თუ i-ურ ექსპერიმენტში მოხდა A ხდომილება, და $X_i(\omega) = 0$ -- წინააღმდეგ შემთხვევაში. მაშინ როგორც უკვე ზემოთ ავღნიშნეთ $X_1, X_2, ..., X_n$ შემთხვევითი სიდიდეები წყვილ-წყვილად დამოუკიდებელია. ავღნიშნოთ $p = P(A)$ (ზოგჯერ p-ს უწოდებენ "*წარმატების ალბათობას*" – თუ A ხდომილების მოხდენა განიხილება როგორც "*წარმატება*"). მაშინ წინა მაგალითის თანახმად $EX_i = p$ და $DX_i = p(1-p)$.

ბინომიალური განაწილება. $B = X_1 + X_2 + \cdots + X_n$ შემთხვევით სიდიდეს, სადაც $X_1, X_2, ..., X_n$ წინა მაგალითიდანაა, **ბინომიალური** ეწოდება. ცხადია, რომ მისი მნიშვნელობები მერყეობს 0-დან n-მდე ცდების ნებისმიერი შედეგების დროს, $0 \le B \le n$. იმისათვის რომ ვიპოვოთ მისი განაწილება (ე. ი. $P\{B = k\}$ ალბათობები, როცა $k = 0, 1, ..., n$), საკმარისია ვიცოდეთ p -- ცალკეულ ცდაში განსახილველი A ხდომილების მოხდენის ალბათობა.

მართლაც, ხდომილება $\{B = k\}$ ხდება მაშინ და მხოლოდ მაშინ როცა A ხდომილება ხდება ზუსტად k ექსპერიმენტში (შესაბამისად, A ხდომილება არ ხდება ზუსტად $n-k$ ექსპერიმენტში, ამასთანავე მნიშვნელეობა არა აქვს რომელ k ექსპერიმენტში მოხდება A ხდომილება). ამდენად, $\{B = k\}$ ხდომილების ალბათობა ტოლია n დამოუკიდებელი ხდომილების ერთდროულად მოხდენის (ნამრავლის) ალბათობის, და რადაგანაც დამოუკიდებელ ხდომილებათა ნამრავლის ალბათობა ტოლია ალბათობების ნამრავლის, შესაბამისად, ალბათობა იმისა, რომ n დამოუკიდებელ ცდაში A ხდომილება მოხდა ზუსტად k-ჯერ, ხოლო მისი საწინააღმდეგო $\overline{A}$ -- კი $(n-k)$-ჯერ, ტოლია $\underbrace{p \cdot p \cdots p}_{k-\text{ჯერ}} \cdot \underbrace{(1-p) \cdot (1-p) \cdots (1-p)}_{(n-k)-\text{ჯერ}} = p^k (1-p)^{n-k}$.

რამდენ სხვადასხვანაირად შეიძლება k ადგილის შერჩევა n ადგილიდან? ეს შეიძლება განხორციელდეს C_n^k-ჯერ. ე. ი. ხდომილება $\{B = k\}$ წარმოიდგინება C_n^k ცალი უთავსებადი ხდომილების გაერთიანების სახით, რომელთაგან თითოეულის ალბათობაა $p^k (1-p)^{n-k}$. ამიტომ ალბათობათა შეკრების წესის თანახმად გვაქვს:

$$P\{B=k\}=C_n^k p^k(1-p)^{n-k}=\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}.$$

სახელწოდება “*ბინომიალური განაწილება*” მოდის იქიდან, რომ $P\{B=k\}$ ალბათობა წარმოადგენს $(k+1)$-ე წევრს ნიუტონის ბინომის გაშლაში:

$$(a+b)^n=\sum_{k=0}^{n} C_n^k a^k b^{n-k},$$

თუ დავუშვებთ, რომ $a=p$ და $b=1-p$.

ვინაიდან ბინომიალური შემთხვეითი სიდიდე წარმოიდგინება n დამოუკიდებელი ბერნულის შემთხვევითი სიდიდის ჯამის სახით, ამიტომ თეორემა 2-სა და მაგალითი 2-ის თანახმად გვექნება:

$$EB=np \text{ და } DB=np(1-p).$$

მაგალითი 4. გამოვთვალოთ პუასონის განაწილების

$$P\{X=k\}=\frac{\lambda^k}{k!}e^{-\lambda}, \quad k=0,1,2,...$$

დისპერსია.

როგორც ჩვენ უკვე ვნახეთ პუასონის განაწილების მათემატიკური ლოდინი ემთხვევა მის λ პარამეტრს. ვიპოვოთ ახლა პუასონის შემთხვევითი სიდიდის კვადრატის მათემატიკური ლოდინი. ამ მიზნით ვისარგებლოთ წარმოდგენით:

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

მაშინ

$$EX^2=E[X(X-1)]+EX=E[X(X-1)]+\lambda.$$

ამიტომ, თუ ვისარგებლებთ შემთხვევითი სიდიდის ფუნქციიდან მათემატიკური ლოდინის გამოთვლის წესითა და e^x ფუნხციის ხარისხოვან მწკრივად გაშლით, მივიღებთ, რომ:

$$E[X(X-1)]=\sum_{k=0}^{\infty}k(k-1)\frac{\lambda^k}{k!}e^{-\lambda}=e^{-\lambda}\sum_{k=2}^{\infty}k(k-1)\frac{\lambda^k}{k!}=$$

$$=\lambda^2e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}=\lambda^2e^{-\lambda}e^{\lambda}=\lambda^2.$$

საბოლოოდ, გვაქვს:

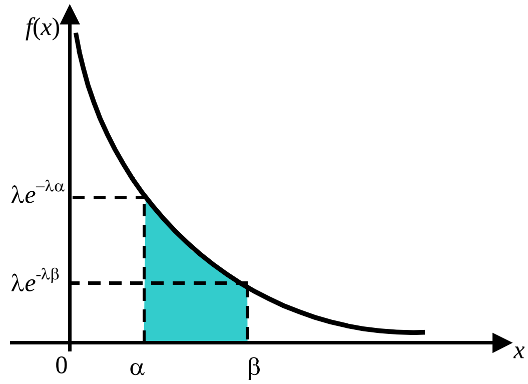
$$DX=EX^2-(EX)^2=\lambda^2+\lambda-\lambda^2=\lambda.$$

ე. ი. პუასონის განაწილების როგორც მათემატიკური ლოდინი, ისე დისპერსია ტოლია ამ განაწილების λ პარამეტრის.

ექსპონენციალური განაწილება. X შემთხვევით სიდიდეს ეწოდება **ექსპონენციალურად განაწილებული პარამეტრით** λ ($\lambda>0$), თუ მის განაწილების სიმკვრივეს აქვს სახე:

$$f(x)=\begin{cases}0, & x<0,\\ \lambda e^{-\lambda x}, & x\ge 0.\end{cases}$$

ექსპონენციელური განაწილების სიმკვრივის ფუნქციის გრაფიკს აქვს შემდეგი სახე:



აქ დაშტრიხული არის ფართობი ტოლია ექსპონენციალურად განაწილებული X შემთხვევით სიდიდის (α,β) ინტერვალში მოხვედრის ალბათობის.

ექსპონენციალური განაწილების სიმკვრივიდან შეგვიძლია აღვადგინოთ ექსპონენციალური განაწილების ფუნქცია. ადვილი დასანახია, რომ

$$F(x)=\begin{cases} 0, & \text{hj wf} \quad x<0, \\ 1-e^{-\lambda x}, & \text{hj wf} \quad x\geq 0. \end{cases}$$

ვნახოთ რა ალბათური შინაარსი გააჩნია λ პარამეტრს. ამ მიზნით, გამოვთვალოთ ექსპონენციალური განაწილების მათემატიკური ლოდინი. უწყვეტი შემთხვევითი სიდიდისათვის მათემატიკური ლოდინის განმარტების თანახმად გვაქვს:

$$EX=\int_{-\infty}^{+\infty} xf(x)dx=\int_{0}^{+\infty}\lambda xe^{-\lambda x}dx=\frac{1}{\lambda}.$$

ექსპონენციალური განაწილების λ პარამეტრი წარმოადგენს განაწილების მათემატიკური ლოდინის შებრუნებულ სიდიდეს. მეორეს მხრივ, თუ გამოვითვლით დისპერსიას, დავინახავთ, რომ:

$$DX={}^{1}\!/_{\lambda^2},$$

შესაბამისად, საშუალო კვადრატული გადახრა იხნება ${}^{1}\!/_{\lambda}$.

გარდა ამისა, ვინაიდან განაწილების სიმკვრივე კლებადია, ამიტომ ის თავის მაქსიმალურ მნიშვნელობას აღწევს $x=0$ წერტილში. შესაბამისად, ექსპონენციალური განაწილების მოდა იქნება $M_0=0$.

§22. სტანდარტული გადახრა. მომენტები

განმარტება 1. X შემთხვევითი სიდიდის **საშუალო კვადრატული გადახრა** ეწოდება არითმეტიკულ კვადრატულ ფესვს ამ შემთხვევითი სიდიდის დისპერსიიდან და აღინიშნება σ_x სიმბოლოთი:

$$\sigma_x = +\sqrt{DX} \ .$$

σ_x-ს ხშირად **სტანდარტულ გადახრასაც** უწოდებენ. გადახრის ამ მახასიათებლის შემოღება განპირობებულია იმით, რომ, განსხვავებით დისპერსიისაგან, იგი ზომის იგივე ერთეულებში გამოისახება, რაც X შემთხვევითი სიდიდე.

შემთხვევითი სიდიდის საშუალო კვადრატული გადახრა დაახლოებით მიუთითებს იმაზე, თუ რამდენად განსხვავდება შემთხვევითი სიდიდის დაკვირვებული მნიშვნელობა მათემატიკური ლოდინისაგან. კომერციული მოღვაწეობის ხშირ შემთხვევაში სტანდარტული გადახრა არის რისკის მახასიათებელი, მიუთიუთებს რა, თუ რამდენად განუსაზღვრელია სიტუაცია.

მაგალითი 1. მოცემულია ორი ტრესტის წლიური საპროცენტო შემოსავლის მონაცემები ბოლო 10 წლის განმავლობაში:

ტრესტი A	8.3	-6.2	20.9	-2.7	33.6	42.8	24.4	5.2	3.1	30.5
ტრესტი B	12.1	-2.8	6.4	12.2	27.8	25.3	18.2	10.7	-1.3	11.4

რომელი ტრესტი მიეკუთვნება რისკის უფრო მაღალ დონეს?

ამოხსნა. თითოეული ტრესტისათვის გამოვთვალოთ შესწორებული შერჩევითი დისპერსია შევადაროთ ისინი ერთმანეთს. გვაქვს:

$$\overline{x}_A = 16\,\%, \quad \overline{x}_B = 12\,\%, \quad s_A'^2 = 280.34\,(\%)^2 \text{ და } s_B'^2 = 99.38\,(\%)^2.$$

როგორც ვხედავთ, $s_A' > s_B'$. ეს იმას ნიშნავს, რომ A ტრესტი მიეკუთვნება რისკის უფრო მაღალ დონეს (უფრო რისკიანია), ვიდრე B ტრესტი, იმავდროულად, $\overline{x}_A > \overline{x}_B$, ანუ A ტრესტის ამონაგები საშუალოდ უფრო მაღალია, ვიდრე B ტრესტისა. ეს ფაქტი სავსებით ეთანხმება ჩვენს ინტუიციას: ინვესტიცია, რომელიც დაკავშირებულია რისკის უფრო მაღალ დონესთან, უნდა იძლეოდეს უფრო მაღალ საშუალო ამონაგებსაც.

მაგალითი 2. დავუშვათ, რომ 10 წლის წინ თქვენ მოახდინეთ ტოლი თანხების ინვესტირება A და B ტრესტებში, ე. ი. შექმენით საინვესტიციო პორტფელი, რომელშიც A და B ტრესტებში ინვესტიციათა წონები ტოლია და უდრის 1/2-ს. ახალი პორტფელი აღვნიშნოთ C-თი. შევადაროთ A, B და C პორტფელების რისკიანობა.

ამოხსნა. წინასწარ შემოვიღოთ *შემოსავლიანობის* ცნება: თუ საინვესტიციო თანხა წლის დასაწყისში უდრის P_0-ს და ამ თანხის ინვესტირებამ A ტრესტში წლის ბოლოს მოგვცა P_A თანხა, მაშინ A ტრესტის **ამონაგები** R_A მოიცემა ფორმულით (%-ში): $R_A = (P_A - P_0)\cdot 100\% / P_0$.

ვთქვათ, ახლა A და B ტრესტებიდან თითოეულში ინვესტირებულია $P_0/2$-ის ტოლი თანხა. მაშინ წლის ბოლოს C პორტფელისაგან მიღებული

თანხაა $(P_A+P_B)/2$. ამიტომ მისი ამონაგები იქნება: $R_C=(R_A+R_B)/2$. ამიტომ, C პორტფელის ამონაგებთა მწკრივი იქნება (გასაგებია, რომ წინა მაგალითში მოყვანილია ბოლო 10 წლის დაკვირვებები R_A-სა და R_B-ზე):

10.2 -4.5 13.65 4.75 30.7 34.1 21.3 7.95 0.9 20.95.

სტანდარტული გადახრის მეშვეობით შევადაროთ C პორტფელის რისკიანობა A და B ტრესტებში ინვესტიციათა რისკიანობას. ვინაიდან,

$$\bar{x}_A=(\bar{x}_A+\bar{x}_B)/2=14\%. \text{ ამიტომ } s_C'^2=159.45(\%)^2 \text{ და } s_C'=12.63\%$$

როგორც ვხედავთ, რისკიანობის მიხედვით A , B და C ინვესტიციები დალაგდა შემდეგი მიმდევრობით:

$$s_B'<s_C'<s_A',$$

ხოლო მათი ამონაგები კი შემდეგი მიმდევრობით:

$$R_A>R_C>R_B.$$

შემთხვევითი სიდიდის სტანდარტიზაცია. დავუშვათ, რომ X შემთხვევითი სიდიდის მათემატიკური ლოდინია EX, ხოლო საშუალო კვადარტული გადახრაა -- σ_x. განვიხილოთ ახალი შემთხვევითი სიდიდე

$$Y=\frac{X-EX}{\sigma_x} \qquad (1)$$

და გამოვთვალოთ მისი მათემატიკური ლოდინი და დისპერსია.

მათემატიკური ლოდინისა და დისპერსიის თვისებებიდან გამომდინარე ადვილი დასანახია, რომ $EY=0$ და $DY=1$. მართლაც, გვაქვს:

$$EY=E(\frac{X-EX}{\sigma_x})=E[\frac{1}{\sigma_x}\cdot(X-EX)]=\frac{1}{\sigma_x}\cdot E(X-EX)$$

და

$$DY=D(\frac{X-EX}{\sigma_x})=D[\frac{1}{\sigma_x}\cdot(X-EX)]=\frac{1}{{\sigma_x}^2}\cdot D(X-EX)=\frac{1}{DX}\cdot DX=1.$$

(1) გარდაქმნას ეწოდება X შემთხვევითი სიდიდის **ცენტრირება** (მათემატიკური ლოდინის გამოკლება) და **ნორმირება** (საშუალო კვავდრატულ გადახრაზე გაყოფა) ან უფრო მოკლედ -- X შემთხვევითი სიდიდის **სტანდარტიზაცია.**

სხვა სიტყვებით, რომ ვთქვათ სტანდარტიზაცია არის შემთხვევითი სიდიდის ისეთი წრფივი გარდაქმნა, რომელსაც გარკვეული მათემატიკური ლოდინისა და დისპერსიის მქონე შემთხვევითი სიდიდე დაყავს ნოლოვანი მათემატიკური ლოდინისა და ერთეულოვანი დისპერსიის მქონე (ანუ სტანდარტულ) შემთხვევით სიდიდეზე.

მომენტები, ასიმეტრია და ექსცესი. X შემთხვევითი სიდიდის n **რიგის საწყისი მომენტი** ეწოდება სიდიდეს $\mu_n:=EX^n$. შესაბამისად, პირველი რიგის მომენტი წარმოადგენს მათემატიკურ ლოდინს $\mu:=\mu_1=EX$.

X შემთხვევითი სიდიდის n **რიგის ცენტრალური მომენტი** ეწოდება სიდიდეს $\nu_n:=E(X-\mu)^n$. ამ აღნიშვნებში გასაგებია, რომ მეორე რიგის ცენტრალური მომენტი წარმოადგენს დისპერსიას $\sigma^2:=\nu_2=DX$. შესაბამისად,

$\sigma = \sigma_x$. ადვილი დასანახია, რომ საწყის და ცენტრალურ მომენტებს შორის არსებობს შემდეგი კავშირი:

$$\nu_2 = \mu_2 - \mu^2,$$

$$\nu_3 = \mu_3 - 3\mu_2 \cdot \mu + 2\mu^2,$$

$$\nu_4 = \mu_4 - 4\mu_3 \cdot \mu + 6\mu_2 \cdot \mu^2 - 3\mu^4, \text{ და ა. შ.}$$

ცენტრალური მომენტების საშუალებით განიმარტება შემთხვევითი სიდიდის შემდეგი მნიშვნელოვანი რიცხვითი მახასიათებლები, კერძოდ, ასიმეტრიისა და ექსცესის კოეფიციენტები.

ექსცესის კოეფიციენტი ეწოდება სიდიდეს

$$e = \frac{\nu_4}{\sigma^4} - 3 = \frac{E(X-EX)^4}{[+\sqrt{E(X-EX)^2}\,]^4} - 3.$$

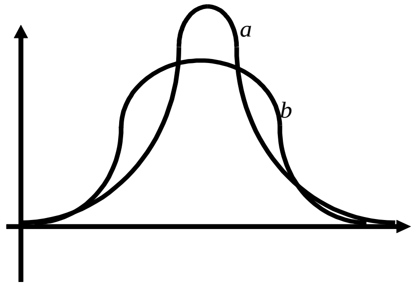
ექსცესის კოეფიციენტი ახასიათებს განაწილების კონცენტრაციის ხარისხს საშუალო მნიშვნელობის (μ-ს) ირგვლივ. რაც უფრო დიდია e, მით მეტადაა კონცენტრირებული განაწილება საშუალოს ირგვლივ, ანუ სიმკვრივეს μ წერტილში აქვს მაღალი პიკი, და პირიქით (იხ. ნახ. 1: შესაბამისად, a და b წირები).

ასიმეტრიის კოეფიციენტი ეწოდება სიდიდეს

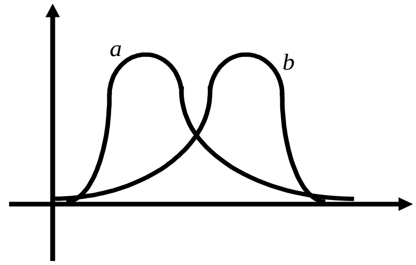
$$\alpha = \frac{\nu_3}{\sigma^3} = \frac{E(X-EX)^3}{[+\sqrt{E(X-EX)^2}\,]^3}.$$

ასიმეტრიის ზომას საფუძვლად უდევს საშუალო კუბური გადახრა, რომელიც საშუალებას იძლევა უფრო სრულად გავითვალისწინოთ შემთხვევითი სიდიდის დიდი გადახრები. განაწილების ასიმეტრიის შემთხვევაში განაწილების მრუდის ერთი მხარე იძლევა უფრო დიდ კუბურ გადახრას მეორე მხარესთან შედარებით და რადგან კუბური გადახრის დროს გადახრის ნიშანი ნარჩუნდება, ამიტომ კუბურ გადახრებს შორის განსხვავება აჩვენებს დადებით ან უარყოფით ასიმეტრიას.

თუ X შემთხვევითი სიდიდის განაწილება სიმეტრიულია თავისი საშუალო მნიშვნელობის ($\mu = EX$ მათემატიკური ლოდინის) მიმართ, მაშინ X შემთხვევითი სიდიდის ყველა კენტი რიგის ცენტრალური მომენტი ნულის ტოლია ($\nu_{2n-1} = 0$). შესაბამისად, ამ შემთხვევაში ასიმეტრიის კოეფიციენტიც ნულის ტოლი იქნება. თუ $\alpha < 0$, მაშინ განაწილება მარცხნივ ასიმეტრიულია, ხოლო თუ $\alpha > 0$, მაშინ განაწილება მარჯვნივ ასიმეტრიულია. (იხ. ნახ. 2: შესაბამისად, a და b წირები).



ნახ. 1



ნახ. 2

§23. კოვარიაცია. კორელაციის კოეფიციენტი

ორ ξ და η შემთხვევით სიდიდეს შორის კავშირის (დამოკიდებულების) მახასიათებელს წარმოადგენს ξ და η შემთხვევით სიდიდეების თავიანთი განაწილებების ცენტრებიდან (განაწილების ცენტრი ეწოდება შემთხვევითი სიდიდის მათემატიკურ ლოდინს) გადახრების ნამრავლის მათემატიკური ლოდინი, რომელსაც **კოვარიაციის კოეფიციენტი** ან უბრალოდ **კოვარიაცია** ეწოდება და აღინიშნება $\operatorname{cov}(\xi;\eta)$ სიმბოლოთი:

$$\operatorname{cov}(\xi;\,\eta) = E((\xi - E\xi)(\eta - E\eta)).$$

დავუშვათ, რომ ξ შემთხვევითი სიდიდის შესაძლო მნიშვნელობებია $\{x_1, x_2, x_3, \ldots, x_n\}$, ხოლო η შემთხვევითი სიდიდის შესაძლო მნიშვნელობებია $\{y_1, y_2, y_3, \ldots, y_k\}$. მაშინ მათემატიკური ლოდინის თვისებებიდან გამომდინარე გასაგებია, რომ:

$$\operatorname{cov}(\xi;\,\eta) = \sum_{i=1}^{n}\sum_{j=1}^{k}(x_i - E\xi)(y_j - E\eta)P((\xi = x_i)\bigcap(\eta = y_j)). \qquad (1)$$

(1) ფორმულას შეიძლება მიეცეს ასეთი ინტერპრეტაცია: თუ ξ-ს დიდი მნიშვნელობებისათვის უფრო ალბათურია η-ს დიდი მნიშვნელობები, ხოლო ξ-ს მცირე მნიშვნელობებისათვის უფრო ალბათურია η-ს მცირე მნიშვნელობები, მაშინ (1) ფორმულის მარჯვენა მხარეში დომინირებენ დადებითი შესაკრებები, და კოვარიაცია ღებულობს დადებით მნიშვნელობას.

თუ კი უფრო ალბათურია ისეთი ნამრავლები $(x_i - E\xi)(y_j - E\eta)$, რომლებიც შედგებიან სხვადასხვა ნიშნიანი თანამამრავლებისაგან, ანუ შემთხვევითი ექსპერიმენტის იმ შედეგებს, რომლებსაც მივყავართ ξ-ს დიდი მნიშვნელობებისაკენ, ძირითადად მივყავართ η-ს მცირე მნიშვნელობებთან, და პირიქით, მაშინ კოვარიაცია ღებულობს მოდულით დიდ უარყოფით მნიშვნელობებს.

პირველ შემთხვევაში მიღებულია ვილაპარაკოთ პირდაპირ კავშირზე: ξ-ს ზრდასთან ერთად η შემთხვევით სიდიდეს გააჩნია მატების (ზრდის) ტენდენცია. მეორე შემთხვევაში კი ლაპარაკობენ შებრუნებულ კავშირზე: ξ-ს ზრდასთან ერთად η შემთხვევით სიდიდეს გააჩნია შემცირების ანუ დაცემის ტენდენცია.

თუ კი ჯამში დაახლოებით ერთი და იგივე წვლილი შეაქვთ დადებით და უარყოფით ნამრავლებს $(x_i - E\xi)(y_j - E\eta)p_{ij}$, მაშინ შეიძლება ითქვას, რომ ჯამში ისინი "აქრობენ" ერთმანეთს და კოვარიაცია ახლოს იქნება ნულთან. ამ შემთხვევაში ერთი შემთხვევითი სიდიდის მეორეზე დამოკიდებულება არ იკვეთება.

ადვილი საჩვენებელია, რომ თუ

$$P[(\xi = x_i)\cap P(\eta = y_j)] = P(\xi = x_i)P(\eta = y_j), \quad i = 1,\ldots n;\ j = 1,\ldots,k \qquad (2)$$

მაშინ $\operatorname{cov}(\xi;\,\eta) = 0.$

მართლაც, (1) ფორმულის თანახმად, (2)-ის საფუძველზე, გვაქვს:

$$\operatorname{cov}(\xi,\eta)=\sum_{i=1}^{n}\sum_{j=1}^{k}(x_i-E\xi)(y_j-E\eta)P(\xi=x_i)P(\eta=y_j)=$$
$$=\sum_{i=1}^{n}(x_i-E\xi)P(\xi=x_i)\cdot\sum_{j=1}^{k}(y_j-E\eta)P(\eta=y_j)=$$
$$=E(\xi-E\xi)\cdot E(\eta-E\eta)=0\cdot 0=0.$$

აქ ჩვენ გამოვიყენეთ მათემატიკური ლოდინის შემდეგი მნიშვნელოვანი თვისება: *შემთხვევითი სიდიდის თავისი მათემატიკური ლოდინიდან გადახრის მათემატიკური ლოდინი ნულის ტოლია.*

მაგალითად, დისკრეტულ შემთხვევაში მართლაც გვაქვს:

$$E(\xi-E\xi)=\sum_{i=1}^{n}(x_i-E\xi)P(x_i)=\sum_{i=1}^{n}x_iP(x_i)-E\xi\sum_{i=1}^{n}P(x_i)=E\xi-E\xi=0.$$

კოვარიაცია შეიძლება გადაიწეროს შემდეგი მოხერხებული ფორმით:

$$\operatorname{cov}(\xi;\eta)=E(\xi\eta-\xi E\eta-\eta E\xi+E\xi E\eta)=E(\xi\eta)-E(\xi E\eta)-$$
$$-E(\eta E\xi)+E(E\xi E\eta)=$$
$$=E(\xi\eta)-E\eta E\xi-E\xi E\eta+E\xi E\eta=E(\xi\eta)-E\xi E\eta,$$

ე. ი. ორი შემთხვევითი სიდიდის კოვარიაცია ტოლია მათი ნამრავლის მათემატიკურ ლოდინს გამოკლებული მათემატიკური ლოდინების ნამრავლი.

ცნობილია, რომ თუ შემთხვევითი სიდიდეები დამოუკიდებელია (დისკრეტულ შემთხვევაში ეს იმას ნიშნავს, რომ შესრულებულია (2) თანაფარდობები), მაშინ ნამრავლის მათემატიკური ლოდინი მათემატიკური ლოდინების ნამრავლის ტოლია. ამიტომ, უკანასკნელი თანფარდობის თანახმად (ისევე როგორც ეს ზემოთ იყო აღნიშნული) დამოუკიდებელი ξ და η შემთხვევითი სიდიდეებისათვის $\operatorname{cov}(\xi;\eta)=0$. შევნიშნავთ, რომ საზოგადოდ შებრუნებული დებულება სწორი არ არის.

ამოცანა 1. მონეტას აგდებენ 5-ჯერ. ξ შემთხვევითი სიდიდე არის მოსულ გერბთა რიცხვი, ხოლო η შემთხვევითი სიდიდე კი ბოლო ორ აგდებაში მოსულ გერბთა რიცხვი. ავაგოთ ამ შემთხვევითი სიდიდეების ერთობლივი განაწილების კანონი და ვიპოვოთ კოვარიაცია.

ამოცანა 2. 32 კარტიდან შემთხვევით იღებენ 2-ს. ξ შემთხვევითი სიდიდე არის ამოღებული ტუზების რიცხვი, ხოლო η შემთხვევითი სიდიდე კი ამოღებული მეფეების რიცხვი. ავაგოთ ავაგოთ ამ შემთხვევითი სიდიდეების ერთობლივი განაწილების კანონი და ვიპოვოთ კოვარიაცია.

ადვილის დასანახია, რომ გარდა ზემოთმოყვანილი თვისებებისა, კოვარიაციის ფუნქციას გააჩნია შემდეგი თვისებები:

1). $\operatorname{cov}(\xi,\xi)=D\xi$;

2). $\operatorname{cov}(\xi,\eta)=\operatorname{cov}(\eta,\xi)$;

3). $\operatorname{cov}(c\cdot\xi,\eta)=c\cdot\operatorname{cov}(\xi,\eta)$;

4). $\operatorname{cov}(\xi\pm\eta,\varsigma)=\operatorname{cov}(\xi,\varsigma)\pm\operatorname{cov}(\eta,\varsigma)$;

5). $D(\xi \pm \eta) = D\xi \pm 2\operatorname{cov}(\xi, \eta) + D\eta$, კერძოდ, თუ ξ და η დამოუკიდებელი შემთხვევითი სიდიდეებია, მაშინ $D(\xi \pm \eta) = D\xi + D\eta$ (ვინაიდან ამ შემთხვევაში $\operatorname{cov}(\xi, \eta) = 0$);

6). $|\operatorname{cov}(\xi, \eta)| \leq \sqrt{D\xi} \cdot \sqrt{D\eta}$.

გასაგებია, რომ cov(ξ;η) დამოკიდებულია იმ ზომის ერთეულებზე, რომლებშიც გამოსახულია ξ და η შემთხვევითი სიდიდეები (მაგალითად, ვთქვათ, ξ და η -- გარკვეული დეტალის წრფივი სიგრძეებია. თუ ზომის ერთეულად ავიღებთ 1 სმ-ს, მაშინ cov(ξ;η) მიიღებს ერთ მნიშვნელობას, ხოლო თუ კი ზომის ერთეულად ავიღებთ 1მმ-ს, მაშინ cov(ξ;η) მიიღებს სხვა, უფრო მეტ მნიშვნელობას, თუ რა თქმა უნდა $\operatorname{cov}(\xi;\eta) \neq 0$). ამდენად, cov(ξ;η) -ს გამოყენება კავშირის მაჩვენებლად მოუხერხებელია.

იმისათვის, რომ საქმე გვქონდეს ზომის ერთეულისაგან დამოუკიდებელ მაჩვენებელთან, განვიხილოთ შემთხვევითი სიდიდეები:

$$\xi^* = \frac{\xi - M\xi}{\sqrt{D\xi}} = \frac{\xi - M\xi}{\sigma_\xi}; \quad \eta^* = \frac{\eta - M\eta}{\sqrt{D\eta}} = \frac{\eta - M\eta}{\sigma_\eta}.$$

ასეთ შემთხვევით სიდიდეებს ეწოდებათ ξ და η შემთხვევითი სიდიდეების **ნორმირებული გადახრები** ან **სტანდარტიზაცია.** თითოეულ მათგანს ცენტრად აქვს ნული, ხოლო დისპერსია კი ერთის ტოლია. მართლაც, გვაქვს:

$$E\xi^* = E\left(\frac{\xi - E\xi}{\sigma_\xi}\right) = \frac{1}{\sigma_\xi}\left(E\xi - EE\right) = \frac{1}{\sigma_\xi}\left(E\xi - E\xi\right) = 0;$$

$$D\xi^* = D\left(\frac{\eta - E\eta}{\sigma_\xi}\right) = \frac{1}{\sigma_\xi^{\ 2}} D(\eta - E\eta) = \frac{D\eta}{\sigma_\xi^{\ 2}} = 1.$$

ξ* და η* შემთხვევითი სიდიდეების კოვარიაციას ეწოდება ξ და η შემთხვევითი სიდიდეების **კორელაციის კოეფიციენტი** და აღინიშნება ρ(ξ;η) სიმბოლოთი:

$$\operatorname{cov}(\xi^*;\eta^*) = \rho(\xi;\eta) = E\left(\frac{\xi - E\xi}{\sigma_\xi} \cdot \frac{\eta - E\eta}{\sigma_\eta}\right) = \frac{E((\xi - E\xi)(\eta - E\eta))}{\sigma_\xi \sigma_\eta} =$$

$$= \frac{\operatorname{cov}(\xi;\eta)}{\sigma_\xi \sigma_\eta} = \frac{E(\xi\eta) - E\xi E\eta}{\sigma_\xi \sigma_\eta}; \quad \sigma_\xi = \sqrt{D\xi}; \ \sigma_\eta = \sqrt{D\eta}.$$

ξ და η დამოუკიდებელი შემთხვევითი სიდიდეებია, მაშინ ρ(ξ;η)=0, ვინაიდან ამ შემთხვევაში cov(ξ;η)=0. პირიქით, საზოგადოდ სწორი არ არის. შემთხვევითი სიდიდეები შეიძლება ფუნქციონალურადაც კი იყვნენ დამოკიდებულები (ერთი შემთხვევითი სიდიდის ნებისმიერ მნიშვნელობას შეესაბამება ერთადერთი მნიშვნელობა მეორე შემთხვევითი სიდიდის), მაგრამ მათი კორელაციის კოეფიციენტი ნულის ტოლი იყოს.

მაგალითი 1. დავუშვათ, რომ ξ შემთხვევითი სიდიდე სიმეტრიულადაა განაწილებული ნულის ირგვლივ. მაშინ $E\xi=0$. ვთქვათ, $\eta=\xi^2$. მაშინ $E(\xi\eta)=E(\xi^3)=0$, ვინაიდან ξ^3 აგრეთვე, სიამეტრიულადაა განაწილებული ნულის ირგვლივ. მეორეს მხრივ, $E\xi E\eta=0$, ვინაიდან $E\xi=0$. ამიტომ:

$$\rho(\xi,\eta)=\frac{E(\xi\eta)-E\xi E\eta}{\sigma_\xi\sigma_\eta}=0 .$$

მაგალითი 2. დავუშვათ, რომ ξ და η შემთხვევითი სიდიდეების ერთობლივი განაწილების კანონი მოცემულია შემდეგი ცხრილით:

ξ \ η	**1**	**2**	
1	1/5	0	1/5
2	0	3/5	3/5
3	1/5	0	1/5
	2/5	3/5	

ცხადია, რომ:

$$E\xi=1\cdot1/5+2\cdot3/5+3\cdot1/5=2;\ E\eta=1\cdot2/5+2\cdot3/5=8/5;$$
$$E(\xi\eta)=1\cdot1\cdot1/5+2\cdot2\cdot3/5+3\cdot1\cdot1/5=16/5;\ E(\xi\eta)-E\xi E\eta=0 .$$

აქედან გამომდინარე, კორელაციის კოეფიციენტი ნულია, მაშინ როდესაც ნათელია, რომ ადგილი აქვს η შემთხვევითი სიდიდის ფუნქციონალურ დამოკიდებულებას ξ შემთხვევით სიდიდეზე.

კორელაციის კოეფიციენტი არ შეიცვლება, თუ ξ შემთხვევითი სიდიდის ნაცვლად განვიხილავთ $\xi_1=\xi+a$ ან $\xi_2=k\xi$ შემთხვევით სიდიდეს (სადაც a და k—მუდმივებია, $k>0$), ვინაიდან კოორდინატთა სათავის შეცვლისას ან მასშტაბის ცვლილებისას ნორმირებული გადახრა არ იცვლება. იგივე შეიძლება ითქვას η შემთხვევით სიდიდეზეც.

კორელაციის კოეფიციენტის თვისებები:

1. $-1\le\rho(\xi;\eta)\le1$.
2. თუ $\rho(\xi;\eta)=1$, მაშინ $\eta=k\xi+b$, სადაც k და b— მუდმივებია, $k>0$.
3. თუ $\rho(\xi;\eta)=-1$, მაშინ $\eta=k\xi+b$, სადაც k და b— მუდმივებია, $k<0$.
4. თუ $\eta=k\xi+b$, $(k\neq0)$ ან $\xi=k_1\eta+b_1$ $(k_1\neq0)$, მაშინ
   $\rho(\xi;\eta)=1$ როცა $k_i>0$; $\rho(\xi;\eta)=-1$ როცა $k_i<0$ $(i=1,2)$.

კორელაციის კოეფიციენტი $\rho(\xi;\eta)$აღწევს თავის სასაზღვრო მნიშვნელობებს -1-სა და 1-ს მაშინ და მხოლოდ მაშინ, როცა ξ და η შემთხვევითი სიდიდეების ყველა მნიშვნელობა კონცენტრირებულია (თავმოყრილია) $\xi;\eta$ სიბრტყის გარკვეულ წრფეზე, ანუ ξ და η შემთხვევით სიდიდეებს შორის არსებობს წრფივი კავშირი.

თუ $|\rho(\xi;\eta| <1$, მაშინ ასეთი წრფივი კავშირი არ არსებობს. თუმცა, $|\rho(\xi;\eta)|$-ს ერთთან მიახლოებასთან ერთად ξ და η –ს ერთობლივ განაწილებას გააჩნია გარკვეული წრფის ირგვლივ კონცენტრირების ტენდენცია და $|\rho(\xi;\eta)|$ სიდიდე შეიძლება ჩაითვალოს ξ და η სიდიდეებს შორის სრული წრფივი დამოკიდებულების საზომად.

მაგალითი. ξ და η შემთხვევითი სიდიდეების ქვემოთ მოყვანილი ერთობლივი განაწილების კანონის მიხედვით გამოვთვალოთ კორელაციის კოეფიციენტი $\rho(\xi;\eta)$.

η / ξ	**1**	**2**	**3**	
10	1/36	0	0	1/36
20	2/36	1/36	0	3/36
30	2/36	2/36	2/36	6/36
40	1/36	9/36	16/36	26/36
	6/36	12/36	18/36	

$$E\xi = 10\cdot 1/36 + 20\cdot 3/36 + 30\cdot 6/36 + 40\cdot 26/36 \cong 35.83;$$
$$E\eta = 1\cdot 6/36 + 2\cdot 12/36 + 3\cdot 18/36 \cong 2.3;$$
$$D\xi = (10-35.83)^2\cdot 1/36 + (20-35.83)^2\cdot 3/36 + (30-35.83)^2\cdot 6/36 +$$
$$+(40-35.83)^2\cdot 26/36 \cong 57.64;$$
$$\sigma_\xi \cong 7.6;$$
$$D\eta = (1-2.3)^2\cdot 6/36 + (2-2.3)^2\cdot 12/36 + (3-2.3)^2\cdot 18/36 \cong 0.556;$$
$$\sigma_\eta \cong 0.746;$$
$$E(\xi\eta) = 10\cdot 1\cdot 1/36 + 20\cdot 1\cdot 2/36 + 20\cdot 2\cdot 1/36 + 30\cdot 1\cdot 2/36 + 30\cdot 2\cdot 2/36 +$$
$$+30\cdot 3\cdot 2/36 + 40\cdot 1\cdot 1/36 + 40\cdot 2\cdot 9/36 + 40\cdot 3\cdot 16/36 = 86.94;$$
$$\rho(\xi,\eta) = (6.94 - 2.3\cdot 35.83)/(7.6\cdot 0.746) \cong 0.8.$$

დავუშვათ, რომ მოცემულია ξ და η ორი შემთხვევითი სიდიდის ერთობლივი განაწილების კანონი, და ξ შემთხვევითი სიდიდის პირობითი მათემატიკური ლოდინი იცვლება η შემთხვევითი სიდიდის მნიშვნელობის მიხედვით. მაშინ ლაპარაკობენ ξ შემთხვევითი სიდიდის **კორელაციურ დამოკიდებულებაზე** η შემთხვევით სიდიდეზე. თუ ξ-ს პირობითი მათემატიკური ლოდინი არის η-ს წრფივი ფუნქცია, მაშინ ξ და η შემთხვევით სიდიდეებს შორის არსებობს წრფივი კორელაციური კავშირი ანუ დამოკიდებულება.

როგორც წესი, კორელაციურ დამოკიდებულებაზე საუბრისას, მხედველობაში აქვთ წრფივი კორელაციური დამოკიდებულება. არაწრფივი კორ-

ელაციური დამოკიდებულების არსებობისას, მას სპეციალურად აღნიშნავ-ენ.

ξ და η შემთხვევით სიდიდეებს შორის კორელაციური დამოკიდებულ-ება შეიძლება განიმარტოს როგორც კავშირი ξ და η-ს ზრდის ტენდენც-იებს შორის. მაგალითად, ξ და η-ს შორის არსებობს პირდაპირი კორელ-აციური დამოკიდებულება, თუ ξ-ს ზრდასთან ერთად η შემთხვევით სიდ-იდეს გააჩნია ზრდის ტენდენცია (ეს ნიშნავს, რომ ξ-ს დიდი მნიშვნელობ-ების შემთხვევაში დიდი ალბათობით შეგვხვდება η-ს დიდი მნიშვნელობე-ბიც). თუ ξ-ს დიდ მნიშვნელობებს დიდი ალბათობით შეესაბამება η-ს მცირე მნიშვნელობები, ანუ ξ-ს ზრდასთან ერთად η შემთხვევით სიდიდეს გააჩნია კლების ტენდენცია, მაშინ ამბობენ, რომ ξ და η-ს შორის არსებ-ობს შებრუნებული კორელაციური დამოკიდებულება.

კორელაციური დამოკიდებულების სიღრმე (ანუ სიმჭიდროვე) ხასიათ-დება კორელაციის კოეფიციენტით $\rho(\xi;\eta)$. რაც უფრო ახლოსაა $|\rho(\xi;\eta)|$ ერთთან. მით უფრო მჭიდროა კორელაციური დამოკიდებულება.

რაც უფრო ახლოსაა ξ შემთხვევითი სიდიდის პირობითი მათემატიკუ-რი ლოდინი η-ს მიმართ წრფივ დამოკიდებულებასთან და რაც უფრო მჭი-დროდ ლაგდებიან ξ-ს მნიშვნელობები პირობითი მათემატიკური ლოდინის ირგვლივ, მით უფრო ღრმაა (მჭიდროა) კორელაციური დამოკიდებულება.

ჩვენ შეგვიძლია ვისაუბროთ ორი უწყვეტი შემთხვევითი სიდიდის ერ-თობლივ განაწილების კანონზე. უმეტეს შემთხვევაში შესაძლებელია უწყვ-ეტი შემთხვევითი სიდიდეებიდან დისკრეტული შემთხვევითი სიდიდეების ერთობლივ განაწილებაზე გადასვლა შემდეგნაირად: ξ შემთხვევითი სიდი-დის ცვლილების შუალედი $[a; b]$ უნდა გაიყოს ტოლი სიგრძის მონაკვეთე-ბად $[c_0=a; c_1]; [c_1; c_2]; [c_2; c_3],\ldots,[c_{n-1}; c_n=b]$. ξ შემთხვევითი სიდიდის მნი-შვნელობებად მივიღოთ თითოეული მონაკვეთის შუაწერტილი. ანალოგიუ-რად, უნდა მოვიქცეთ η შემთხვევითი სიდიდის შემთხვევაში. მისი მნიშვნ-ელობათა არე $[a; b]$ უნდა გაიყოს ტოლი სიგრძის შუალედებად $[c_0=a; c_1]$; $[c_1; c_2]; [c_2; c_3],\ldots,[c_{n-1}; c_n=b]$, და η-ს მნიშვნელობებად გამოვაცხადოთ $[g_{i-1}; g_i]$ შუალედების შუაწერტილები. ასეთნაირად, ჩვენ მივიღებთ $\xi^*=\{x_1; x_2; \ldots x_n\}$ და $\eta^*=\{y_1; y_2; \ldots y_k\}$ დისკრეტულ შემთხვევით სიდიდეებს, ამასთ-ანავე ყოველ წყვილს უსაბამებენ ალბათობას

$$p_{ij} = P((\xi\in[c_{i-1}; c_i])\cap(\eta\in[g_{j-1}; g_j])).$$

§24. ჩებიშევის უტოლობა. დიდ რიცხვთა კანონი

ჩებიშევის უტოლობა აფასებს შემთხვევითი სიდიდის გადახრას თავისი მათემატიკური ლოდინიდან. თუ X რაიმე შემთხვევითი სიდიდეა, მაშინ ნებისმიერი $\varepsilon > 0$ რიცხვისათვის სამართლიანია უტოლობა:

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

ამ უტოლობას **ჩებიშევის უტოლობას** უწოდებენ. იგი სამართლიანია როგორც დისკრეტული, ისე უწყვეტი შემთხვევითი სიდიდეებისათვის.

შევამოწმოთ ჩებიშევის უტოლობა დისკრეტულ შემთხვევაში. დავუშვათ, რომ შემთხვევითი სიდიდე მოცემულია განაწილების კანონით:

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

ვინაიდან $|X - EX)| < \varepsilon$ და $|X - E(X)| \geq \varepsilon$ საწინააღმდეგო ხდომილებებია, ამიტომ $P\,(\,|X - E(X)| < \varepsilon\,) + P\,(\,|X - E(X)| \geq \varepsilon\,) = 1$, შესაბამისად,

$$P\,(\,|X - E(X)| < \varepsilon\,) = 1 - P\,(\,|X - E(X)| \geq \varepsilon\,).$$

ვიპოვოთ $P\,(\,|X - E(X)| \geq \varepsilon\,)$. დისპერსიის განმარტების თანახმად გვაქვს:

$$D(X) = (x_1 - E(X))^2 p_1 + (x_2 - E(X))^2 p_2 + \ldots + (x_n - E(X))^2 p_n\,.$$

გადავაგდოთ ამ ჯამიდან ის შესაკრებები, რომელთათვისაც $|X - E(X)| < \varepsilon$. ამის შედეგად ჯამი მხოლოდ შემცირდება, ვინაიდან ყველა მასში შემავალი შესაკრები არაუარყოფითია. გარკვეულობისათვის ვიგულისხმოთ, რომ გადაგდებულია პირველი k შესაკრები. მაშინ

$$D(X) \geq (x_{k+1} - E(X))^2 p_{k+1} + (x_{k+2} - E(X))^2 p_{k+2} + \ldots + (x_n - E(X))^2 p_n \geq$$

$$\geq \varepsilon^2\,(p_{k+1} + p_{k+2} + \ldots + p_n).$$

შევნიშნოთ, რომ $p_{k+1} + p_{k+2} + \ldots + p_n$ არის ალბათობა იმისა, რომ $|X - E(X)| \geq \varepsilon$, ვინაიდან ეს არის ჯამი X შემთხვევითი სიდიდის ყველა შესაძლო მნიშვნელობის, რომელთათვისაც აღნიშნული უტოლობა სამართლიანია. შესაბამისად, $D(X) \geq \varepsilon^2 P(|X - E(X)| \geq \varepsilon)$, ანუ $P\,(|X - E(X)| \geq \varepsilon) \leq D(X) / \varepsilon^2$. მაშინ საწინააღმდეგო ხდომილების ალბათობა იქნება

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

დიდ რიცხვთა კანონი. ჩებიშევის უტოლობა საშუალებას იძლევა დავამტკიცოთ ფუნდამენტური შედეგი, რომელიც საფუძვლად უდევს მათემატიკურ სტატისტიკას – ე. წ. *დიდ რიცხვთა კანონი.* ამ შედეგის თანახმად შერჩევითი მახასიათებლები ცდების (ექსპერიმენტების) რიცხვთა ზრდისას უახლოვდება თეორიულ მახასიათებლებს, რაც საშუალებას იძლევა ამა თუ იმ რეალური მოვლენის ალბათური მოდელების პარამეტრები შევაფასოთ ცდების მიერ მიღებული შედეგების გამოყენებით. დიდ რიცხვთა კანონის გარეშე არ გვექნებოდა გამოყენებითი მათემატიკური სტატისტიკის მნიშვნელოვანი ნაწილი.

სტატისტიკური კანონზომიერებების შესწავლა საშუალებას იძლევა დავადგინოთ, რომ გარკვეულ პირობებში შემთხვევით სიდიდეთა დიდი რაოდენობის ჯამური ქცევა (ეფექტი) თითქმის კარგავს შემთხვევით ხასიათს და ხდება კანონზომიერი (სხვა სიტყვებით რომ ვთქვათ, ურთიერთ ჩაიხშობა შემთხვევითი გადახრები გარკვეული საშუალო ქცევიდან). კერძოდ, თუ ცალკეული შესაკრებების გავლენა ჯამზე თანაბრად მცირეა, მაშინ ჯამის განაწილების კანონი უახლოვდება ნორმალურს. ამ მტკიცებულების მათემატიკური ფორმულირება ატარებს სწორედ დიდ რიცხვთა კანონის სახელს. მოვიყვანოთ ერთ-ერთი ამ ტიპის მტკიცებულება.

ჩებიშევის თეორემა. ვთქვათ, შემთხვევითი სიდიდეები $X_1, X_2, ..., X_n$ წყვილ-წყვილად დამოუკიდებელია და არსებობს ისეთი რიცხვი C, რომ $DX_i \le C$, $i = 1,2,...,n$. მაშინ ნებისმიერი დადებითი ε რიცხვისათვის სრულდება უტოლობა:

$$P\{|\frac{X_1 + X_2 + \cdots + X_n}{n} - \frac{EX_1 + EX_2 + \cdots + EX_n}{n}| \ge \varepsilon\} \le \frac{C}{n\varepsilon^2}. \qquad (1)$$

დამტკიცება. განვიხილოთ შემთხვევითი სიდიდეები $Y_n = X_1 + X_2 + \cdots + X_n$ და $Z_n = Y_n / n$. მათემატიკური ლოდინისა და დისპერსიის თვისებების ძალით გვექნება შემდეგი თანაფარდობები:

$$EY_n = EX_1 + EX_2 + \cdots + EX_n, \;\; DY_n = DX_1 + DX_2 + \cdots + DX_n.$$

გარდა ამისა,

$$EZ_n = EY_n / n \;\; \text{და} \;\; DZ_n = DY_n / n^2.$$

შესაბამისად,

$$EZ_n = [EX_1 + EX_2 + \cdots EX_n]/n \;\; \text{და} \;\; DZ_n = [DX_1 + DX_2 + \cdots DX_n]/n^2.$$

ამიტომ თეორემის პირობებში გვაქვს:

$$DZ_n \le Cn/n^2 = C/n.$$

თუ ახლა Z_n შემთხვევითი სიდიდისათვის გამოვიყენებთ ჩებიშევის უტოლობას, მაშინ (1) თანაფარდობის მარცხენა მხარისათვის მივიღებთ შეფასებას:

$$P\{|Z_n - EZ_n| \ge \varepsilon\} \le \frac{DZ_n}{\varepsilon^2} \le \frac{C}{n\varepsilon^2}. \qquad \blacksquare$$

ეს თეორემა, ისევე როგორც, საკუთრივ ჩებიშევის უტოლობები მიღებულ იქნა პ. ჩებიშევის მიერ 1867 წელს გამოქვეყნებულ ნაშრომში: "საშუალო მნიშვნელობების შესახებ".

მაგალითი 1. ვთქვათ, $C = 1$ და $\varepsilon = 0,1$. n-ის რომელი მნიშვნელობისათვის არ აღემატება (1) უტოლობის მარჯვენა მხარე 0.1-ს?, 0.05-ს?, 0.00001-ს?

განსახილველ შემთხვევაში (1) უტოლობის მარჯვენა მხარე ტოლია $100/n$ -ის. შესაბამისად, ის არ აღემატება 0.1-ს, თუ n არაა ნაკლები 1000-ზე, არ აღემატება 0.05-ს, თუ n არაა ნაკლები 2000-ზე და არ აღემატება 0.00001-ს, თუ n არაა ნაკლები 10 000 000-ზე.

(1) უტოლობის მარჯვენა მხარე, და მასთან ერთად მარცხენაც, n-ის ზრდასთან ერთად, ფიქსირებული C და ε-ის შემთხვევაში, უახლოვდება

ნულს. შესაბამისად, ალბათობა იმისა, რომ დამოუკიდებელი შემთხვევითი სიდიდეების საშუალო არითმეტიკული თავისი მათემატიკური ლოდინისაგან განსხვავდება ε-ზე ნაკლებით, უახლოვდება 1-ს შემთხვევით სიდიდეთა რიცხვის ზრდასთან ერთად, ნებისმიერი ε-ის შემთხვევაში. ამ მტკიცებულებას უწოდებენ **დიდ რიცხვთა კანონს.**

გადაწყვეტილებების მიღების ალბათურ-სტატისტიკური მეთოდებისათვის (და მთლიანად მათემატიკური სტატისტიკისათვის) განსაკუთრებით მნიშვნელოვანია ის შემთხვევა, როცა ყველა X_i შემთხვევით სიდიდეს ($i = 1, 2, \ldots$) გააჩნია ერთი და იგივე მათემატიკური ლოდინი EX_1 და ერთი და იგივე დისპერსია $\sigma^2 = DX_1$. მკვლევარისათვის უცნობი მათემატიკური ლოდინის ნაცვლად (შეფასებად) იყენებენ შერჩევით საშუალო არითმეტიკულს:

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}.$$

დიდ რიცხვთა კანონიდან გამომდინარეობს, რომ ცდების (ექსპერიმენტების, გაზომვების) რიცხვის ზრდასთან ერთად $\bar{X}$ რაგინდ ახლოს უახლოვდება EX_1-ს, რაც მოკლედ ასე ჩაიწერება:

$$\bar{X} \xrightarrow{P} EX_1, \text{ როცა } n \to \infty.$$

სიმბოლო $\xrightarrow{P}$ აღნიშნავს "ალბათობით კრებადობას". საჭიროა აღინიშნოს, რომ "ალბათობით კრებადობის" ცნება განსხვავდება მათემატიკურ ანალიზში მიღებული "ზღვარზე გადასვლის" ცნებისაგან. გავიხსენოთ, რომ a_n რიცხვით მიმდევრობას აქვს ზღვარი a, როცა $n \to \infty$, თუ ნებისმიერი, რაგინდ მცირე, $\delta > 0$ რიცხვისათვის, არსებობს ისეთი $n(\delta)$ რიცხვი, რომ ყოველი $n > n(\delta)$ ნომრისათვის სრულდება თანაფარდობა: $a_n \in (a - \delta, a + \delta)$. "ალბათობით კრებადობის" ცნების გამოყენებისას მიმდევრობის წევრები Y_n შემთხვევითი სიდიდეებია, ვიხილავთ რაგინდ მცირე $\varepsilon > 0$ რიცხვს და თანაფარდობა $Y_n \in (a - \varepsilon, a + \varepsilon)$ იგულისხმება რომ სრულდება არა გარანტირებულად, არამედ არანაკლებ $(1 - \varepsilon)$-ის ტოლი ალბათობით.

სიხშირეების კრებადობა ალბათობებისაკენ. როგორც აღნიშნული იყო რაიმე A ხდომილების ალბათობა – ეს არის ის რიცხვი, რომელსაც უახლოვდება A ხდომილების მოხდენათა რიცხვის შეფარდება ექსპერიმენტების საერთო რიცხვთან, როცა ექსპერიმნეტების რიცხვი უსასრულოდ იზრდება. ეს დებულება, მათემატიკური მოდელის ჩარჩოებში, მე-17 საუკუნის მიწურულს პირველად დაამტკიცა ცნობილმა მათემატიკოსმა იაკობ ბერნულმა, მაგრამ დამტკიცება გამოქვეყნებულ იქნა ი. ბერნულის სიკვდილის შემდეგ 1713 წელს (ი. ბერნული ცხოვრობდა შვეიცარიის ქალაქ ბაზელში 1654-1705 წლეებში). ბერნულის თეორემის თანამედროვე ფორმულირება შემდეგია:

ბერნულის თეორემა. დავუშვათ, m არის n დამოუკიდებელ ექსპერიმენტში A ხდომილების მოხდენათა რიცხვი, ხოლო p არის A ხდომილ-

ების მოხდენის ალბათობა ცალკეულ ექსპერიმენტში. მაშინ ნებისმიერი $\varepsilon > 0$ რიცხვისათვის სამართლიანია უტოლობა

$$P\{|\frac{m}{n} - p| \geq \varepsilon\} \leq \frac{p(1-p)}{n\varepsilon^2}. \qquad (2)$$

დამტკიცება. როგორც ჩვენ უკვე ვნახეთ m შემთხვევით სიდიდეს აქვს ბინომიალური განაწილება წარმატების ალბათობით p და იგი წარმოიდგინება n დამოუკიდებელი X_i, $i = 1, 2, ..., n$ შემთხვევითი სიდიდის ჯამის სახით, რომელთაგან თითოეული ბერნულის შემთხვევითი სიდიდეა: X_i ტოლია 1-ის ალბათობით p და ტოლია 0-ის ალბათობით $1-p$, ანუ $m = X_1 + X_2 + \cdots + X_n$. თუ ახლა გამოვიყენებთ ჩებიშევის თეორემას $X_1, X_2, ..., X_n$ შემთხვევითი სიდიდეებისათვის, სადაც $C = p(1-p)$, ადვილად დავრწმუნდებით (2) უტოლობის სამართლიანობაში. ■

შენიშვნა. ვინაიდან $1/4 - p(1-p) = (p - 1/2)^2 \geq 0$, შესაბამისად, $p(1-p) \leq 1/4$, ამიტომ ჩებიშევის თეორემაში ჩვენ შეგვეძლო აგვეღო $C = 1/4$. მაშინ ნებისმიერი p-სა და ფიქსირებული $\varepsilon > 0$ რიცხვისათვის (2) უტოლობის მარჯვენა მხარე n-ის ზრდასთან ერთად უახლოვდება ნულს. ეს კი თავის მხრივ გვიჩვენებს, რომ ალბათობის მათემატიკური განმარტება (მაგალითად, ა. ნ. კოლმოგოროვის მიხედვით) სრულ თანხვედრაშია ბუნებისმეტყველთა (მაგალითად, პ. მიზესის (1883-1953)) მიერ მოყვანილ განმარტებასთან, რომლის თანახმად ალბათობა არის სიხშირეების ზღვარი ექსპერიმენტების უსასრულო მიმდევრობაში.

რაც შეეხება პირდაპირ ექსპერიმენტალურ დადასტურებას იმისა, რომ გარკვეული ხდომილებების განხორციელების სიხშირეები ახლოსაა ალბათობებთან, რომლებიც განიმარტება თეორიული მოსაზრებებით, ეს ჩვენ ადრე უკვე მოვიყვანეთ შესავალ ნაწილში სხვადასხვა დროს სხვადასხვა მეცნიერის მიერ სიმეტრიული მონეტის აგდების მაგალითზე. ასე მაგალითად, მე-18 საუკუნეში ფრანგი მეცნიერის ბიუფონის მიერ მონეტის 4040-ჯერ აგდებისას გერბის მისვლის ფარდობითი სიხშირე იყო 0.507; ინგლისელი სტატისტიკოსის კ. პირსონის მიერ მონეტის 12000-ჯერ აგდებისას შესაბამისი სიხშირე აღმოჩნდა 0.5016, ხოლო მის მიერვე მონეტის 24000-ჯერ აგდებისას კი – 0.5005. ყველა შემთხვევაში სიხშირეები მხოლოდ უმნიშვნელოდ განსხვავდებოდნენ თეორიული ალბათობისაგან, რომელიც 0.5-ის ტოლია (ვინაიდან სიმეტრიული მონეტის შემთხვევაში გერბისა და საფასურის მოსვლა თანაბრადშესაძლებელია).

სტატისტიკური ჰიპოთეზების შემოწმების შესახებ. (2) უტოლობის გამოყენებით ჩვენ შეგვიძლია გარკვეული დასკვნები გამოვიტანოთ პროდუქციის ხარისხის წინასწარ მოცემულ მოთხოვნებთან შესაბამისობასთან დაკავშირებით.

დავუშვათ, რომ პროდუქციის 100 000 ერთეულიდან 30 000 აღმოჩნდა დეფექტური. ეთანხმება თუ არა ეს ჰიპოთეზას იმის შესახებ, რომ პროდუქციის დეფექტურობის ალბათობა ტოლია 0,23-ის? როგორი ალბათური მოდელის გამოყენებაა მიზანშეწონილი? ჩავთვალოთ, რომ ტარდება რთული ცდა, რომელიც შედგება 100 000 ექსპერიმენტისაგან, რომელთაგან თითოე-

ული გულისხმობს პროდუქციის 100 000 ერთეულიდან ცალკეულის შემოწმებას გამოსადეგიანობაზე. ითვლება, რომ ექსპერიმენტები წყვილ-წყვილად დამოუკიდებელია და ყოველ ექსპერიმენტში ალბათობა იმისა, რომ პროდუქციის ერთეული დეფექტურია ტოლია p-სი.

რეალურ ცდაში მიღებულია, რომ ხდომილება "პროდუქციის ერთეული დეფექტურია" განხორციელდა 30 000-ჯერ 100 000 ექსპერიმენტში. ეთანხმება თუ არა ეს ჰიპოთეზას იმის შესახებ, რომ პროდუქციის დეფექტურობის ალბათობაა 0,23?

ვისარგებლოთ (2) უტოლობით. განსახილველ შემთხვევაში $n=100000$, $m=30000$, $m/n=0.3$, $p=0.23$, $m/n-p=0.07$. ჰიპოთეზის შესამოწმებლად იქცევიან შემდეგნაირად. შევაფასოთ ალბათობა იმისა, რომ m/n განსხვავდება p-სა-გან ისევე როგორც განსახილველ შემთხვევაში, ან უფრო მეტით, ე. ი. შევაფასოთ $|m/n-p|\geq 0.07$ უტოლობის შესრულების ალბათობა. ვიგულისხმოთ, რომ (2) უტოლობაში $p=0.23$ და $\varepsilon=0.07$. მაშინ ბერნულის თეორემის თანახმად

$$P\{|\frac{m}{n}-0.23|\geq 0.07\}\leq\frac{0.23\cdot 0.77}{0.0049n}\approx\frac{36.11}{n}. \quad (3)$$

როცა $n=100000$ (3) უტოლობის მარჯვენა მხარე ნაკლებია 1/2500. ამიტომ ალბათობა იმისა, რომ გადახრა იქნება არანაკლები დაკვირვებულზე, ერთობ მცირეა. შესაბამისად, თუ საწყისი ჰიპოთეზა სამართლიანია, მაშინ განსახილველ ცდაში განხორციელდა ხდომილება, რომლის ალბათობა ნაკლებია 1/2500. ვინაიდან, 1/2500 – ძალიან პატარა რიცხვია, ამიტომ საწყისი ჰიპოთეზა უნდა უკუვაგდოთ (უარვყოთ).

§25. ნორმალური განაწილების კანონი

თუ ξ შემთხვევითი სიდიდის განაწილების სიმკვრივე განისაზღვრება ფორმულით

$$p(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad (1)$$

სადაც μ– ნებისმიერი რიცხვია, ხოლო σ-- დადებითი რიცხვია, მაშინ ამბობენ რომ ξ განაწილებულია ნორმალური კანონით ანუ ξ "ნორმალური" შემთხვევითი სიდიდეა.

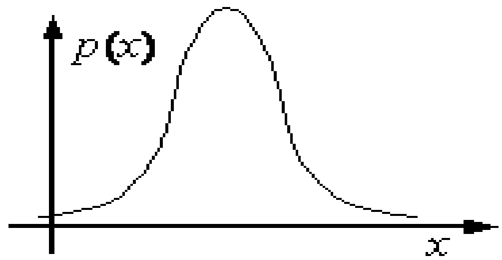
μ-სა და σ-ს მნიშვნელობები სრულად განსაზღვრავენ $p(x)$ ფუნქციას. ამ შემთხვევითი სიდიდისათვის სარგებლობენ აღნიშვნით: $\xi \cong N(\mu,\sigma^2)$. შესაბამისად, სიმკვრივე აღინიშნება სიმბოლოთი: $f_{N(\mu,\sigma^2)}(x)$

როცა $\mu=0$ და $\sigma=1$, გამოიყენება აღნიშვნა

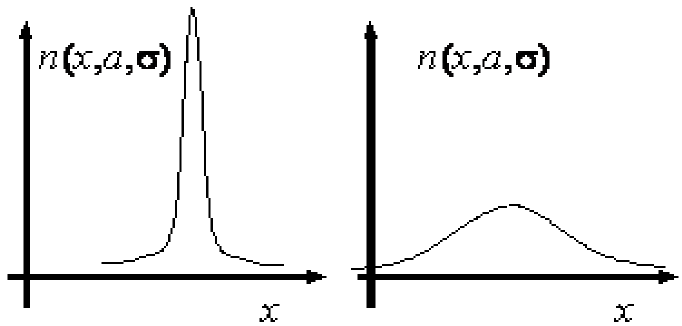
$$\varphi(x):=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

და მას ლაპლასის ფუნქციას უწოდებენ. ის ლუწი ფუნქციაა: $\varphi(-x)=\varphi(x)$.

ნორმალური შემთხვევითი სიდიდის განაწილების სიმკვრივის გრაფიკს μ-სა და σ-ს გარკვეული მნიშვნელობებისათვის აქვს შემდეგი სახე:



გრაფიკი სიმეტრიულია $x=\mu$ წრფის მიმართ, და სრულდება პირობა $p(x)\to 0$, როცა $x\to\pm\infty$. თუ დავიწყებთ μ-ს გაზრდას, ისე რომ σ-ს დავტოვებთ უცვლელად, მაშინ გრაფიკი დაიწყებს გადაადგილებას მარჯვნივ, ხოლო μ-ს შემცირებისას კი – მარცხნივ, ისე რომ არ შეიცვლის ფორმას. მეორეს მხრივ, თუ μ-ს მნიშვნელობა უცვლელია, მაშინ შედარებით მცირე σ-ს შეესაბამება $p(x)$-ის გრაფიკი აშკარად გამოხატული პიკით (როგორც ეს გამოსახულია ქვემოთ მოყვანილი ნახაზებიდან მარჯვენა ნახაზზე), ხოლო შედარებით დიდი σ-ს შემთხვევაში $p(x)$-ის გრაფიკი გაწოლილი წირია (როგორც ეს გამოსახულია ქვემოთ მოყვანილი ნახაზებიდან მარცხენა ნახაზზე).



ნორმალურად განაწილებული ξ შემთხვევითი სიდიდის განაწილების $F(x)$ ფუნქციის აღსანიშნავად ხმარობენ სიმბოლოს $F_{N(\mu,\sigma^2)}(x)$. ის მიიღება განაწილების სიმკვრივის ინტეგრირებით:

$$F_{N(\mu,\sigma^2)}(x) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}}\,dt\,.$$

როცა $\mu = 0$ და $\sigma = 1$, გამოიყენება აღნიშვნა

$$\Phi(x) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}\,dt$$

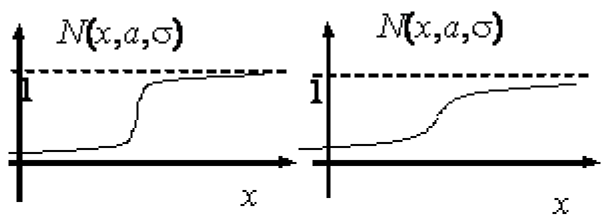
და მას ლაპლასის ინტეგრალურ ფუნქციას უწოდებენ.

გარდა ამისა, იხმარება აღნიშვნა:

$$\Phi_0(x) := \frac{1}{\sqrt{2\pi}}\int_{0}^{x} e^{-\frac{t^2}{2}}\,dt\,.$$

ცხადია, რომ $\Phi(-x) = 1 - \Phi(x)$ და $\Phi(x) = 0.5 + \Phi_0(x)$, $x > 0$.

ქვემოთ მოყვანილია ნორმალურად განაწილებული შემთხვევითი სიდიდეების განაწილების $F(x)$ ფუნქციების გრაფიკები σ-ს შესაბამისად შედარებით მცირე (მარცხენა გრაფიკი) და შედარებით დიდი (მარჯვენა გრაფიკი) მნიშვნელობბებისათვის:



ნორმალურად განაწილებული ξ შემთხვევითი სიდიდის განაწილებ-ის სიმკვრივის $p(x)$ ფუნქციის გრაფიკის $x=\mu$ წრფის მიმართ სიმეტრიულობიდან გამომდინარეობს, რომ $E\xi=\mu$.

თუ გამოვითვლით ნორმალურად განაწილებული შემთხვევითი სიდი-დის $D\xi$ დისპერსიას, აღმოჩნდება, რომ ის σ^2–ის ტოლია.

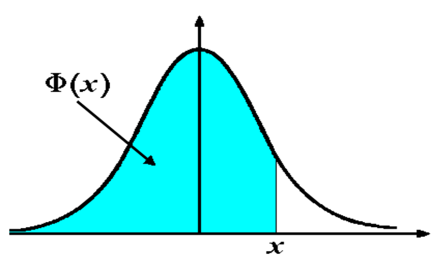
ამრიგად, μ და σ პარამეტრებს ნორმალურად განაწილებული შემთ-ხვევითი სიდიდის განაწილების სიმკვრივის ფორმულაში, გააჩნიათ შემდე-გი ალბათური შინაარსი: μ -- არის მათემატიკური ლოდინი, ხოლო σ^2 – დისპერსია.

ალბათობა იმისა, რომ ნორმალურად განაწილებული ξ შემთხვევითი სიდიდე მიიღებს მნიშვნელობას (x_1, x_2) შუალედიდან, გამოითვლება შემდ-ეგი ფორმულით:

$$P(x_1<\xi<x_2)=F(x_2)-F(x_1)=\Phi(\beta)-\Phi(\alpha).$$

აქ $\Phi(x)$– ლაპლასის ინტეგრალური ფუნქციაა, $\beta=\dfrac{x_2-\mu}{\sigma}$; $\alpha=\dfrac{x_1-\mu}{\sigma}$. სი-ნამდვილეში, შუალედი როგორი იქნება (ღია, ჩაკეტილი, თუ ნახევრადჩაკე-ტილი) მნიშვნელობა არაა აქვს (რადგანაც, უწყვეტი ტიპის განაწილება კონკრეტულ მნიშვნელობას ღებულობს ნულის ტოლი ალბათობით), ანუ ნებისმიერი სახის $\langle x_1, x_2\rangle$ შუალედისათვის:

$$P\{N(\mu,\sigma^2)\in\langle x_1,x_2\rangle\}=\Phi(\frac{x_2-\mu}{\sigma})-\Phi(\frac{x_1-\mu}{\sigma}).$$



თუ ξ შემთხვევითი სიდიდის განაწილების სიმკვრივეა $f_{N(0,1)}(x)=\varphi(x)$ მაშინ, ის არის ნორმალურად განაწილებული შემთხვევითი სიდიდე მათე-მატიკური ლოდინით ნული და დისპერსიით ერთი. მას **სტანდარტული ნო-რმალური** შემთხვევითი სიდიდე ეწოდება. ასეთი შემთხვევითი სიდიდის განაწილების სიმკვრივის გრაფიკი სიმეტრიულია ორდინატთა ღერძის მი-მართ, და მისთვის გვაქვს:

$$P\{N(\mu,\sigma^2)\in\langle x_1,x_2\rangle\}=\Phi(x_2)-\Phi(x_1).$$

ცხადია, რომ:

$$\frac{N(\mu,\sigma^2)-\mu}{\sigma}=N(0,1).$$

ნორმალური $N(\mu,\sigma^2)$ განაწილების p-კვანტილი აღვნიშნოთ სომბოლოთი x_p^{μ,σ^2} (ანუ, $P\{N(\mu,\sigma^2) \le x_p^{\mu,\sigma^2}\} = p$). შესაბამისად, $N(0,1)$-ის p-კვანტილი აღინიშნება სიმბოლოთი $x_p^{0,1}$. ამ კვანტილებს შორის ისეთივე კავშირია, რაც თვითონ $N(\mu,\sigma^2)$ და $N(0,1)$ შემთხვევით სიდიდეებს შორის, კერძოდ:

$$x_p^{0,1} = \frac{x_p^{\mu,\sigma^2} - \mu}{\sigma} \text{ და } x_p^{\mu,\sigma^2} = \sigma \cdot x_p^{0,1} + a.$$

დავუშვათ, რომ ξ და η – დამოუკიდებელი და ნორმალურად განაწილებული შემთხვევითი სიდიდეებია, ისეთი რომ $E\xi = \mu_1$, $D\xi = \sigma_1^2$, $E\eta = \mu_2$, $D\eta = \sigma_2^2$. მაშინ შემთხვევითი სიდიდე $\Psi = c_1\xi + c_2\eta$ (სადაც c_1 და c_2 -- ნებისმიერი მუდმივებია), აგრეთვე განაწილებულია ნორმალური კანონით. მისი მათემატიკური ლოდინი და დისპერსია გამოითვლება ფორმულებით:

$$E\Psi = c_1\mu_1 + c_2\mu_2, \; D\Psi = c_1^2\sigma_1^2 + c_2^2\sigma_2^2,$$

ანუ, თუ $N(\mu_1,\sigma_1^2)$ და $N(\mu_2,\sigma_2^2)$ დამოუკიდებლებია, მაშინ

$$c_1 N(\mu_1,\sigma_1^2) + c_2 N(\mu_2,\sigma_2^2) \cong N(c_1\mu_1 + c_2\mu_2, c_1^2\sigma_1^2 + c_2^2\sigma_2^2).$$

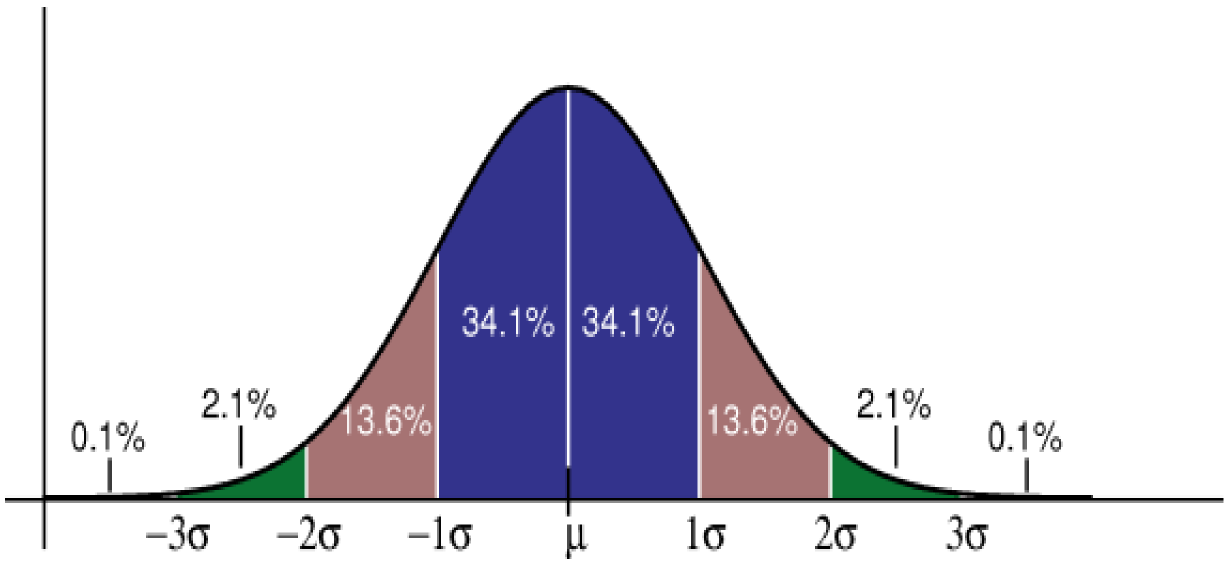
ამოცანა. 12 ბოთლიანი ყუთის წონა – ნორმალურად განაწილებული შემთხვევითი სიდიდეა მათემატიკური ლოდინით 2კგ და საშუალოკვადრატული გადახრით 0.01კგ. ბოთლის მასა ლიმონათით – აგრეთვე ნორმალურად განაწილებული შემთხვევითი სიდიდეა მათემატიკური ლოდინით 0.8კგ და საშუალოკვადრატული გადახრით 0.04კგ. ვიპოვოთ ალბათობა იმისა, რომ ყუთის წონა 12 ბოთლი ლიმონათით მოთავსებული იქნება საზღვრებში: 11 კგ-დან 11.5კგ-მდე.

სამი σ–ს ("სიგმას") წესი. დავუშვათ, მოცემულია ნორმალური კანონით განაწილებული ξ შემთხვევითი სიდიდე მათემატიკური ლოდინით μ და დისპერსიით σ^2. განვსაზღვროთ ξ შემთხვევითი სიდიდის $(\mu - 3\sigma;\ \mu + 3\sigma)$ ინტერვალში მოხვედრის ალბათობა, ანუ ალბათობა იმისა, რომ ξ მიიღებს მნიშვნელობას, რომელიც მათემატიკური ლოდინიდან განსხვავდება არაუმეტეს სამი საშუალოკვადრატული გადახრით. ცხადია, რომ

$$P(\mu - 3\sigma < \xi < \mu + 3\sigma) = \Phi(3) - \Phi(-3) = 2\Phi(3).$$

სტანდარტული ნორმალური განაწილების ფუნქციის (ლაპლასის ინტეგრალური ფუნქციის) ცხრილიდან ვპოულობთ, რომ $\Phi(3)=0,49865$, აქედან გამომდინარეობს, რომ $2\Phi(3)$ პრაქტიკულად ერთის ტოლია. ამრიგად, შეიძლება გაკეთდეს დასკვნა: ნორმალური შემთხვევითი სიდიდე ღებულობს მნიშვნელობებს, რომლებიც მისი მათემატიკური ლოდინიდან გადაიხრება არაუმეტეს 3σ-თი. ქვემოთ მოყვანილია ამ ფაქტის საილუსტარციო

ნახაზი, რომელზეც მითითებულია რომელ შუალედში რა ალბათობებით (პროცენტებში გამოსახული) ხვდება ნორმალური შემთხვევითი სიდიდე.



ხი კვადრატ განაწილება. დავუშვათ, რომ მოცემულია n დამოუკიდებელი, ნორმალური კანონით განაწილებული $\xi_1, \xi_2, \ldots, \xi_n$, შემთხვევითი სიდიდე მათემატიკური ლოდინით ნული და დისპერსიით ერთი. მაშინ შემთხვევითი სიდიდე

$$\chi^2(n) = \sum_{i=1}^{n} \xi_i^2$$

განაწილებულია კანონით, რომელსაც ეწოდება **"ხი კვადრატ განაწილება"** ანუ "**პირსონის განაწილება**" თავისუფლების n ხარისხით. თუ შესაკრებები დაკავშირებულია რაიმე თანაფარდობით (მაგალითად, $\sum_{i=1}^{n} \xi_i = n\bar{\xi}$), მაშინ თავისუფლების ხარისხია $k = n - 1$.

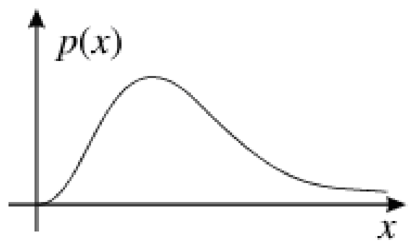
ამ განაწილების სიმკვრივეა

$$f(x) = \begin{cases} 0, & x \le 0; \\ \dfrac{1}{2^{\frac{k}{2}} \Gamma\left(\dfrac{k}{2}\right)} e^{-\frac{x}{2}} x^{\frac{k}{2}-1}, & x > 0. \end{cases}$$

фй სადაც $\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$ – გამა ფუნქციაა, $\Gamma(n+1) = n!$.ьф агтйсшфф6

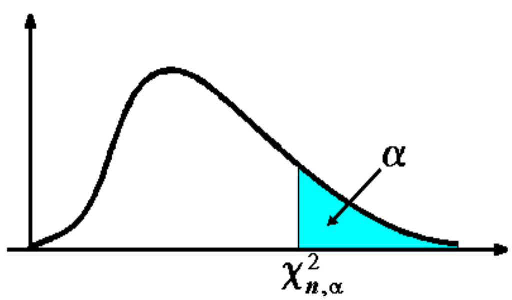
შესაბამისად, ხი კვადრატ განაწილება განისაზღვრება ერთი პარამეტრით, კერძოდ, თავისუფლების ხარისხის რიცხვით. აღსანიშნავია, რომ თავისუფლების ხარისხის რიცხვის ზრდასთან ერთად, ხი კვადრატ განაწილება თანდათანობით უახლოვდება ნორმალურ განაწილებას.

ცხადია, რომ $\chi^2(n)$ შემთხვევითი სიდიდე ღებულობს მხოლოდ არაუარყოფით მნიშვნელობებს. $n>1$ შემთხვევაში $\chi^2(n)$ შემთხვევითი სიდიდის განაწილების სიმკვრივის გრაფიკი გამოსახულია ქვემოთ მოყვანილ ნახაზზე:



იმისათვის, რომ განვსაზღვროთ $\chi^2(n)$ შემთხვევითი სიდიდის დადებით რიცხვთა სიმრავლიდან აღებულ რაიმე ინტერვალში მოხვედრის ალბათობა, გამოიყენება $\chi^2(n)$ განაწილების ცხრილი. ჩვეულებრივ, ასეთი ცხრილი საშუალებას იძლევა α ალბათობისა და თავისუფლების n ხარისხის მიხედვით განისაზღვროს ე. წ. **კვანტილი** $\chi^2_{n,\alpha}$, რომელიც განიმარტება შემდეგი თანაფარდობით:

$$P(\chi^2(n) > \chi^2_{n,\alpha}) = \alpha\ .$$



ეს ფორმულა ნიშნავს შემდეგს: ალბათობა იმისა, რომ $\chi^2(n)$ შემთხვევითი სიდიდე მიიღებს მნიშვნელობას, რომელიც მეტია ვიდრე გარკვეული $\chi^2_{n,\alpha}$ მნიშვნელობა, α-ს ტოლია.

ქვემოთ მოყვანილია $\chi^2(n)$ განაწილების ცხრილის ერთი ფრაგმენტი. ის გვიჩვენებს მაგალითად, რომ $\chi^2(n)$ შემთხვევითი სიდიდე თავისუფლების 10-ის ტოლი ხარისხით ალბათობით $\alpha = 0.95$ ღებულობს მნიშვნე-

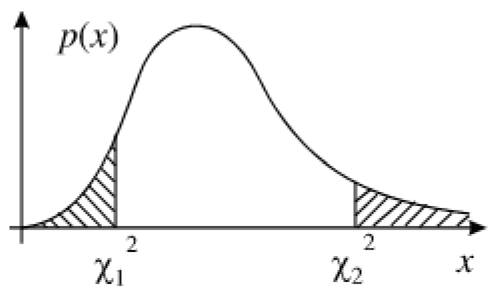
ლობას მეტს, ვიდრე 3.94, ხოლო იგივე შემთხვევითი სიდიდე 1-ის ტოლი თავისუფ-ლების ხარისხით ალბათობით $\alpha = 0.975$ არაღემატება 0.00098-ს.

α / n	0.99	0.975	0.95	...	0.1	0.05	0.01
1	$0.0^3 15$	$0.0^3 98$	$0.0^2 39$	...	2.71	3.84	6.63
...	...	...	...	...	...	...	...
10	2.56	3.25	3.94	...	16.0	18.3	23.2
...	...	...	...	...	...	...	...

(აქ $0.0^3 15$ აღნიშნავს 0.00015, $0.0^2 39 = 0.0039$).

ამოცანა. ვიპოვოთ ისეთი ინტერვალი (χ_1^2, χ_2^2), რომელშიც χ^2 შემთ-ხვევითი სიდიდე თავისუფლების 10-ის ტოლი ხარისხით მოხვდება ალბათ-ობით 0.9.

ამოხსნა. ქვემოთ სქემატურად მოყვანილია თავისუფლების 10-ის ტო-ლი ხარისხის მქონე χ^2 შემთხვევითი სიდიდის განაწილების სიმკვრივის გრაფიკი.



ჩავთვალოთ, რომ დაშტრიხული არეების ფართობები (აქ მარჯვენა არე არ არის შემოსაზღრული) ერთმანეთის ტოლია. თუ χ_1^2-სა და χ_2^2-ს შევარჩევთ პირობიდან

$$P(\chi^2 < \chi_1^2) = P(\chi^2 > \chi_2^2) = (1 - 0.9)/2 = 0.05, \qquad (1)$$

მაშინ შესრულდება პირობა $P(\chi_1^2 < \chi^2 < \chi_2^2) = 0.9$.

(1) ტოლებები საშუალებას გვაძლევს χ^2 განაწილების ცხრილიდან განვსაზღვროთ: $\chi_2^2 = 18.3$. საძებნი ინტერვალის მარცხენა საზღვრის დას-ადგენად ვისარგებლოთ ტოლობით $P(\chi^2 > \chi_1^2) = 1\text{-}0.05 = 0.95$. მაშინ ცხრი-ლიდან ვპოულობთ, რომ $\chi_1^2 = 3{,}94$, და ამიტომ ამოცანის პასუხი იქნება: χ^2 შემთხვევითი სიდიდის მნიშვნელობები ალბათობით 0.9 ეკუთვნის ინტ-ერვალს (3.94, 18.3).

სტიუდენტის განაწილება.

სტატისტიკის ბევრ ამოცანას მივყავართ შემდეგი სახის შემთხვევით სიდიდემდე

$$t(n) = \frac{\xi}{\sqrt{\eta / n}},$$

სადაც ξ და η დამოუკიდებელი შემთხვევითი სიდიდეებია, ამასთანავე ξ – ნორმალურად განაწილებული შემთხვევითი სიდიდეა პარამეტრებით $E\xi = 0$ და $D\xi = 1$, ხოლო η განაწილებულია χ^2 განაწილების კანონით თავისუფლე-ბის ხარისხით n, ანუ

$$t(n) = \frac{N(0,1)}{\sqrt{\chi^2(n)/n}},$$

სადაც $N(0,1)$ და $\chi^2(n)$ დამოუკიდებლებია.

$t(n)$ შემთხვევითი სიდიდის განაწილების კანონს ეწოდება **სტიუდენ-ტის განაწილება** თავისუფლების ხარისხით n.

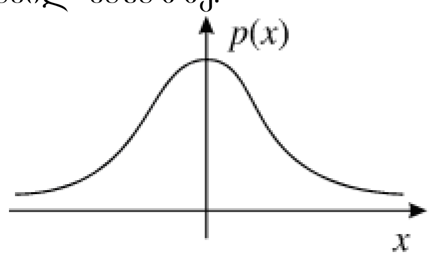
სტიუდენტის განაწილების სიმკვრივეს აქვს შემდეგი სახე:

$$s(t,n) = B_n \left(1 + \frac{t^2}{n-1}\right)^{-\frac{n}{2}},$$

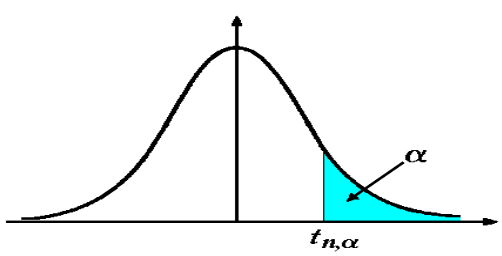
სადაც

$$B_n = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)}\Gamma\left(\frac{n-1}{2}\right)}.$$

სტიუდენტის განაწილების სიმკვრივის გრაფიკი სქემატურად გამოს-ახულია ქვემოთ მოყვანილ ნახაზზე:



განაწილების სიმკვრივის წირი მსგავსია ნორმალური განაწილების ანალ-ოგიური წირის.



სტიუდენტის განაწილების ცხრილები საშუალებას იძლევა თავისუ-ფლების ხარისხის მოცემული n რიცხვისათვის α ალბათობის მიხედვით განისაზღვროს ისეთი მნიშვნელობა t_α, რომლისთვისაც სრულდება

თანაფარდობა $P(|t|>t_\alpha)=\alpha$. ამ ცხრილის ერთი ფრაგმენტი მოყვანილია ქვემოთ:

α / n	0.1	0.05	...	0.01	0.005	...
1	6.314	12.71	...	63.57	318	...
...	...	...	...	...	...	...
12	1.782	2.179	...	3.055	3.428	...
...	...	...	...	...	...	...

ამოცანა 1. ვიპოვოთ სიმეტრიული ინტერვალი, რომელშიც სტიუდენტის კანონით გათვალისწინებული შემთხვევითი სიდიდე თავისუფლების ხარისხით 12, მოხვდება ალბათობით 0.9.

ამოხსნა. ცხადია, რომ

$$P(-x < t < x) = P(|t| < x) = 1 - P(|t| \ge x) = 0.9.$$

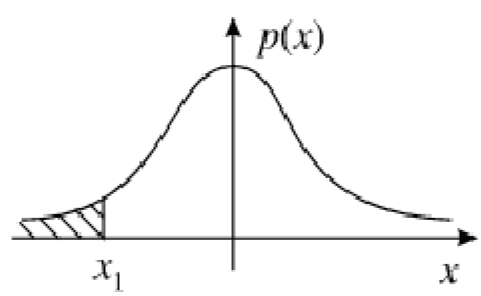
უკანასკნელი ტოლობიდან გამომდინარეობს, რომ:

$$P(|t| \ge x) = 0.1\ ,\ \ (n = 12).$$

არ უნდა გაგვიკვირდეს, რომ აქ გვაქვს არამკაცრი უტოლობა. რამდენადაც საქმე გვაქვს უწყვეტ შემთხვევით სიდიდესთან, ის კონკრეტულ მნიშვნელობას ღებულობს ნულოვანი ალბათობით. ამიტომ არამკაცრი უტოლობა იცვლება მისი ექვივალენტური მკაცრი უტოლობით. ცხრილიდან ვადგენთ, რომ: $x = 1.782$.

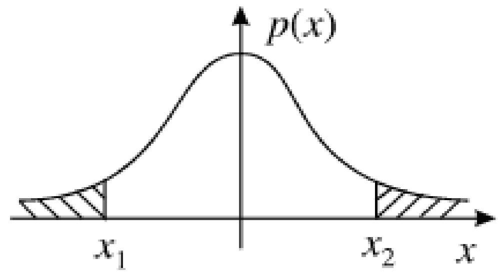
ამოცანა 2. ვიპოვოთ მნიშვნელობა x პირობიდან $P(t > x) = 0.995$, სადაც t -- სტიუდენტის კანონით განაწილებული შემთხვევითი სიდიდეა 12-ის ტოლი თავისუფლების ხარისხით.

ამოხსნა. ქვემოთ მოყვანილია თავისუფლების 12 ხარისხის მქონე სტიუდენტის განაწილების სიმკვრივის გრაფიკი:



ალბათობა იმისა, რომ შემთხვევითი სიდიდე მიიღებს მნიშვნელობას x_1 წერტილის მარჯვნივ მდებარე არიდან ტოლია 0.995-ის, შესაბამისად, ამ წერტილის მარცხნივ მდებარე არეში შემთხვევითი სიდიდე მოხვდება

0.005-ის ტოლი ალბათობით. იმისათვის, რომ ვიპოვოთ x_1, განვიხილოთ ორი სიმეტრიული არე, რომლებიც გამოსახულია ქვემოთ მოყვანილ ნახაზზე:



დავუშვათ, რომ თითოეულ ამ შუალედში, შემთხვევითი სიდიდის მნიშვნელობა აღმოჩნდება ალბათობით 0.005. მაშინ მივიღებთ: $x_1 = -x$, $x_2 = x$, ამასთანავე x განისაზღვრება პირობიდან $P(|t| > x) = 0.01$. ცხრილიდან გვაქვს, რომ: $x = 3.055$. ამიტომ ამოცანის პასუხი იქნება:

$$P(t > -3.055) = 0.995 .$$

ფიშერის განაწილება.

სტატისტიკაში მნიშვნელოვანი გამოყენებები გააჩნია შემთხვევით სიდიდეს

$$F(n,m) = \frac{\xi / n}{\eta / m} = \frac{m\xi}{n\eta},$$

სადაც ξ– შემთხვევითი სიდიდე განაწილებულია χ^2 განაწილების კანონით თავისუფლების ხარისხით n, ხოლო η– შემთხვევითი სიდიდე განაწილებულია χ^2 განაწილების ხარისხით თავისუფლების ხარისხით m, ამასთანავე ξ და η შემთხვევითი სიდიდეები დამოუკიდებელია, ანუ

$$F(n,m) = \frac{\chi^2(n)/n}{\chi^2(m)/m} = \frac{m\chi^2(n)}{n\chi^2(m)},$$

სადაც $\chi^2(n)$ და $\chi^2(m)$ დამოუკიდებელია.

F შემთხვევითი სიდიდე განაწილებულია კანონით, რომელსაც ეწოდება **ფიშერის განაწილება** თავისუფლების ხარისხებით k_1 და k_2. მისი განაწილების სიმკვრივეს აქვს შემდეგი სახე:

$$f(x) = \begin{cases} 0, \quad x \le 0; \\ C_0 \dfrac{x^{\frac{k_1-2}{2}}}{(k_2 + k_1 x)^{\frac{k_1+k_2}{2}}}, \quad x > 0, \end{cases}$$

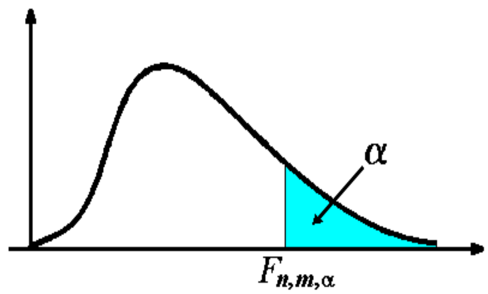
სადაც

$$C_0 = \frac{\Gamma\left(\frac{k_1+k_2}{2}\right) k_1^{\frac{k_1}{2}} k_2^{\frac{k_2}{2}}}{\Gamma\left(\frac{k_1}{2}\right)\Gamma\left(\frac{k_2}{2}\right)}.$$

ამრიგად, ფიშერის განაწილება განისაზღვრება ორი პარამეტრით, კერძოდ თავისუფლების ხარისხების რიცხვებით.

მოცემული n და m რიცხვებისათვის, და მოცემული α ალბათობ-ით ფიშერის განაწილების ცხრილიდან განისაზღვრება ისეთი მნიშვნელო-ბა $F_{n,m,\alpha}$, რომ

$$P(F(n,m) > F_{n,m,\alpha}) = \alpha .$$



როგორც წესი, ცხრილები დგება α-ს მინიშვნელობბებისათვის, რომელიც ტოლია 0.05-ის ან 0.01-ის, ხოლო ზოგჯერ ორივე მნიშვნელობისათვის. ამ ცხრილის ერთი ფრაგმენტი მოყვანილია ქვემოთ:

n / m	1	...	10	...	20	...
1	161.4 647.8	...	241.9 6056	...	248 6209	...
...	...	...	...	...	...	...
10	4.96 10.04	...	2.97 4.85	...	2.77 4.41	...
...	...	...	...	...	...	...

ამ ცხრილში ყოველი უჯრის ზედა ნაწილში მოცემულია $F_{n,m,\alpha}$-ს მნიშვნელობა, როცა $\alpha = 0.05$, ხოლო ქვედა ნაწილში კი როცა $\alpha = 0.01$.

დამოუკიდებელ, ნორმალურად განაწილებულ შემთხვევით სიდიდეთა საშუალო არითმეტიკულის განაწილების კანონი: თუ $\xi_1, \xi_2, ..., \xi_n$ დამოუკიდებელი, ნორმალურად განაწილებულ შემთხვევით სიდიდეებია საშუალოთი μ და სტანდარტული გადახრით σ ($\xi_i \cong N(\mu, \sigma^2)$), მაშინ

$$\bar{\xi} := \frac{\xi_1 + \xi_2 + \cdots + \xi_n}{n} \cong N(\mu, \frac{\sigma^2}{n}) \text{ (ანუ, } Z_n := \frac{\bar{\xi} - \mu}{\sigma / \sqrt{n}} \cong N(0,1) \text{)}.$$

ცნობილია, რომ დამოუკიდებელ, ნორმალურად განაწილებულ შემთხვევით სიდიდეთა საშუალო არითმეტიკული ისევ ნორმალური შემთხვევითი სიდიდეა. შესაბამისად, საკმარისია გამოითვალოს მისი რიცხვითი მახასიათებლები. მართლაც, მათემატიკური ლოდნისა და დისპერსიის ცნობილი თვისებების თანახმად, გვაქვს:

$$E\bar{\xi} = \frac{E\xi_1 + E\xi_2 + \cdots + E\xi_n}{n} = \frac{\mu + \mu + \cdots + \mu}{n} = \frac{n\mu}{n}\mu$$

და

$$D\bar{\xi} = \frac{D\xi_1 + D\xi_2 + \cdots + D\xi_n}{n^2} = \frac{\sigma^2 + \sigma^2 + \cdots + \sigma^2}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

დამოუკიდებელ, ნორმალურად განაწილებულ შემთხვევით სიდიდეთა საშუალო არითმეტიკულისაგან ჯამური კვადრატული გადახრის განაწილების კანონი: თუ $\xi_1, \xi_2, ..., \xi_n$ დამოუკიდებელი, ნორმალურად განაწილებულ შემთხვევით სიდიდეებია საშუალოთი μ და სტანდარტული გადახრით σ ($\xi_i \cong N(\mu, \sigma^2)$), მაშინ

$$\sum_{i=1}^{n} (\xi_i - \mu)^2 \cong \sigma^2 \chi^2(n).$$

ამ შემთხვევაში, ცხადია, რომ $(\xi_i - \mu)/\sigma \cong N(0,1)$. ამიტომ გასაგებია, რომ $(\xi_1 - \mu)/\sigma, (\xi_2 - \mu)/\sigma, \ldots, (\xi_n - \mu)/\sigma$ იქნებიან დამოუკიდებელი სტანდარტული ნორმალური შემთხვევითი სიდიდეები, როგორც დამოუკიდებელ, ნორმალურ შემთხვევით სიდიდეთა წრფივი გარდაქმნით მიღებული შემთხვევითი სიდიდეები. შესაბამისად, ხი კვადრატ განაწილების განმარტების თანახმად, ადვილად დავრწმუნდებით, რომ

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(\xi_i-\mu)^2=\sum_{i=1}^{n}(\frac{\xi_i-\mu}{\sigma})^2\cong\chi^2(n)\,.$$

თუ უკანასკნელ თანაფარდობაში μ-ს შევცვლით $\overline{\xi}$-ით, მივიღებთ ე. წ. განზოგადებულ ხი კვადრატ განაწილებას:

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(\xi_i-\overline{\xi})^2=\frac{n\cdot[\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\overline{\xi})^2]}{\sigma^2}:=\frac{nS_n^2}{\sigma^2}\cong\chi_G^2\,,$$

ანუ

$$\sum_{i=1}^{n}(\xi_i-\overline{\xi})^2=nS_n^2\cong\sigma^2\chi_G^2\,.$$

ადვილი დასანახია, რომ nS_n^2/σ^2 წარმოიდგინება შემდეგნაირად:

$$\frac{nS_n^2}{\sigma^2}=\sum_{i=1}^{n}(\frac{\xi_i-\overline{\xi}}{\sigma})^2=\sum_{i=1}^{n}(\frac{(n-1)\xi_i-\xi_1-\cdots-\xi_{i-1}-\xi_{i-1}-\cdots-\xi_n}{n\sigma})^2:=\sum_{i=1}^{n}(\frac{\eta_i}{n\sigma})^2\ .$$

ცხადია, რომ $\eta_1/(n\sigma)$, $\eta_1/(n\sigma)$, . . . , $\eta\ /(n\sigma)$ უკვე დამოკიდებული (ვინაიდან $\overline{\xi}$ ფუნქციაა $\xi_1,\xi_2,...,\xi_n$-ების), მაგრამ მაინც ნორმალური შემთხვევითი სიდიდეებია, როგორც დამოუკიდებელ, ნორმალურ შემთხვევით სიდიდეთა წრფივი კომბინაცია, რიცხვითი მახასიათებლებით:

$$E\frac{\eta_i}{n\sigma}=\frac{(n-1)E\xi_i-E\xi_1-\cdots-E\xi_{i-1}-E\xi_{i-1}-\cdots-E\xi_n}{n\sigma}=\frac{(n-1)\mu-(n-1)\mu}{n\sigma}=0\,,$$

$$D\frac{\eta_i}{n\sigma}=\frac{(n-1)^2D\xi_i+D\xi_1+\cdots+D\xi_{i-1}+D\xi_{i-1}+\cdots+D\xi_n}{n^2\sigma^2}=\frac{(n-1)^2\sigma^2+(n-1)\sigma^2}{n^2\sigma^2}=\frac{n-1}{n}.$$

მიუხედავად ამისა, nS_n^2/σ^2 შესაძლებელია გადაიწეროს როგორც $n-1$ ცალი სტანდარტული, დამოუკიდებელი ნორმალური შემთხვევითი სიდიდის კვადრატების ჯამი (ფიშერის თეორემა), და, შესაბამისად,

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(\xi_i-\overline{\xi})^2=\frac{nS_n^2}{\sigma^2}:=\frac{(n-1)S_n^{'2}}{\sigma^2}\cong\chi^2(n-1)\,,$$

ანუ

$$\sum_{i=1}^{n}(\xi_i-\overline{\xi})^2=nS_n^2=(n-1)S_n^{'2}\cong\sigma^2\chi^2(n-1)\,.$$

მტკიცდება, რომ (ფიშერის თეორემა) $\bar{\xi} := \dfrac{\xi_1 + \xi_2 + \cdots + \xi_n}{n}$ და $S_n^2 := \dfrac{1}{n}\sum_{i=1}^{n}(\xi_i - \bar{\xi})^2$ (და, შესაბამისად, $\bar{\xi}$ და $S_n^{'2} := \dfrac{n}{n-1}S_n^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(\xi_i - \bar{\xi})^2$) დამოუკიდებელი შემთხვევითი სიდიდეებია. აქედან გამომდინარე, დამოუკიდებლები იქნებიან $Z_n = \dfrac{\bar{\xi} - \mu}{\sigma/\sqrt{n}} \cong N(0,1)$ და $\dfrac{nS_n^2}{\sigma^2} = \dfrac{(n-1)S_n^{'2}}{\sigma^2} \cong \chi^2(n-1)$ შემთხვევითი სიდიდებიც (როგორც მათი წრფივი გარდაქმნები). ამიტომ, სტიუდენტის განაწილების განმარტების თანახმად, შემთხვევით სიდიდეს:

$$T = \frac{\bar{\xi} - \mu}{S_n/\sqrt{n-1}} = \frac{\bar{\xi} - \mu}{S_n^{'}/\sqrt{n}}$$

ექნება სტიუდენტის განაწილება თავისუფლების ხარისხით $n-1$, ვინაიდან ცხადია, რომ

$$T_n = \frac{\dfrac{\bar{\xi} - \mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{nS_n^2}{\sigma^2}/(n-1)}} = \frac{\dfrac{\bar{\xi} - \mu}{\sigma/\sqrt{n}}}{\sqrt{\dfrac{(n-1)S_n^{'2}}{\sigma^2}/(n-1)}} = \frac{N(0,1)}{\sqrt{\chi^2(n-1)/(n-1)}} = t(n-1)\,.$$

პარეტოს განაწილება.

ეკონომიკაში პარეტოს განაწილება გამოიყენება შემოსავლების განაწილების აღსაწერად. დავუშვათ, რომ ξ უწყვეტი ტიპის შემთხვევითი სიდიდეა მნიშვნელობებით $[0,\infty)$ ინტერვალიდან. ამბობენ, რომ ξ-ს გააჩნია პარეტოს განაწილება პარამეტრებით $\theta > 0$ და $x_{\min} \ge 0$, თუ მისი განაწილების სიმკვრივე მოიცემა ფორმულით:

$$f_{\xi(\theta, x_{\min})}(x) = \begin{cases} \dfrac{\theta (x_{\min})^{\theta}}{x^{\theta+1}}, & \text{თუ} \quad x \ge x_{\min}, \\ 0 \quad , & \text{სხვაგან}, \end{cases}$$

ხოლო რიცხვითი მახასიათებლებია: $E\xi(\theta, x_{\min}) = \dfrac{\theta x_{\min}}{\theta - 1}$, თუ $\theta > 1$ (როცა $\theta \le 1$ საშუალო არ არსებობს). $D\xi(\theta, x_{\min}) = (\dfrac{x_{\min}}{\theta - 1})^2 \cdot \dfrac{\theta}{\theta - 2}$, თუ $\theta > 2$ (როცა $\theta \le 2$ დისპერსია არ არსებობს).

ეკონომიკაში, შემოსავლების განაწილების განხილვისას, θ პარამეტრს პარეტოს ინდექსს უწოდებენ. ის წარმოადგენს შემოსავლების განაწი-

ლების სიგანის საზომს და გვეხმარება პარეტოს პრინციპის გაგებაში: ხშირ შემთხვევაში, პოპულაციის დაახლოებით 20%-ს აქვს შემოსავლების დაახლოებით 80%. უფრო ზუსტად, იმ შემოსავლების წილი, რომელიც აღემატება ნებისმიერ დადებით $x > x_{\min}$ რიცხვს შეადგენს $(\frac{x_{\min}}{x})^{\theta}$-ს.

თუ $x_{\min} = 1$, ცვლადის შეცვლით $x = 1 + y$, მივიღებთ სიმკვრივეს:

$$f_{\xi(\theta,1)}(x) = \begin{cases} \dfrac{\theta}{(1+x)^{\theta+1}}, & \text{თუ} \quad x \geq 0, \\ 0 \quad , & \text{სხვაგან.} \end{cases}$$

§26. ცენტრალური ზღვარითი თეორემა

დიდ რიცხვთა კანონი არ იკვლევს შემთხვევით სიდიდეთა ჯამის განაწილების კანონის სახეს. ეს საკითხი შეისწავლება თეორემების ჯგუფში, რომლებსაც **ცენტრალური ზღვარითი თეორემა** ეწოდება. ეს თეორემა ამტკიცებს, რომ შემთხვევით სიდიდეთა ჯამის განაწილების კანონი, რომელთაგან ცალკეულ შესაკრებს შეიძლება ჰქონდეს განსხვავებული განაწილება, უახლოვდება ნორმალურს შესაკრებთა საკმაოდ დიდი რიცხვის შემთხვევაში. ამით აიხსნება ნორმალური განაწილების კანონის უაღრესად დიდი მნიშვნელობა პრაქტიკულ გამოყენებებში.

ცენტრალური ზღვარითი თეორემა ერთნაირად განაწილებული შემთხვევითი სიდიდეებისათვის ასე ჩამოყალიბდება.

თეორემა 1. თუ $X_1, X_2, \ldots, X_n, \ldots$ – დამოუკიდებელი შემთხვევითი სიდიდეების მიმდევრობაა, ერთი და იგივე განაწილების კანონით, მათემატიკური ლოდინით m და დისპერსიით σ^2, მაშინ n –ის უსასრულოდ ზრდისას $Y_n = \sum_{k=1}^{n} X_k$ ჯამის განაწილების კანონი უახლოვდება ნორმალურ განაწილების კანონს:

$$\lim_{n\to\infty} P\{\frac{Y_n - nm}{\sigma\sqrt{n}} \le x\} = N(x;0;1).$$

ა. ლიაპუნოვმა დაამტკიცა ცენტრალური ზღვარითი თეორემა უფრო ზოგად შემთხვევაში.

თეორემა 2 (ლიაპუნოვის თეორემა). თუ X შემთხვევითი სიდიდე წარმოადგენს დამოუკიდებელ შემთხვევით სიდიდეთა ძალიან დიდი რიცხვის ჯამს, რომელთათვისაც შესრულებულია პირობა:

$$\lim_{n\to\infty}(\sum_{k=1}^{n} b_k)/[\left(\sum_{k=1}^{n} D_k\right)^{\frac{3}{2}}] = 0,$$

სადაც b_k – მესამე რიგის აბსოლუტური ცენტრალური მომენტია X_k შემთხვევითი სიდიდის , ხოლო D_k – მისი დისპერსია, მაშინ X შემთხვევით სიდიდეს გააჩნია განაწილება, რომელიც ახლოსაა ნორმალურ განაწილებასთან.

შევნიშნავთ, რომ ლიაპუნოვის თეორემის პირობა ნიშნავს იმას, რომ ცალკეული შესაკრების გავლენა ჯამზე მიზერულია.

აღსანიშნავია, რომ ცენტრალური ზღვარითი თეორემა პრაქტიკულად შესაძლებელია გამოყენებულ იქნეს შემთხვევით სიდიდეთა საკმაოდ არა დიდი რიცხვის შემთხვევაში. გამოცდილება აჩვენებს, რომ თუნდაც 10 ან უფრო ნაკლები შესაკრებების რაოდენობის შემთხვევაშიც ჯამის განაწილება შესაძლებელია შეცვლილ იქნეს ნორმალურით.

მუავრ-ლაპლასის თეორემა. ცენტრალური ზღვარითი თეორემის კერძო შემთხვევას დისკრეტული შემთხვევითი სიდიდეების შემთხვევაში წარმოადგენს მუავრ-ლაპლასის თეორემა.

თეორემა 3 (მუავრ-ლაპლასის თეორემა). თუ ტარდება n დამოუკიდებელი ცდა, რომელთაგან თითოეულში A ხდომილება ხდება ალბათობით p, $np > 15$, მაშინ სამართლიანია თანაფარდობა:

$$p\left(\alpha < \frac{Y-np}{\sqrt{npq}} < \beta\right) \approx \Phi(\beta) - \Phi(\alpha),$$

სადაც $Y-$ A ხდომილების მოხდენათა რიცხვია n ცდაში, $q = 1-p$, ხოლო

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

(ამ ფუნქციის მნიშვნელობები მოყვანილია სპეციალურ ცხრილებში, ამასთანავე $\Phi(-x) = 1 - \Phi(x)$).

დამტკიცება. შეგვიძლია ჩავთვალოთ, რომ $Y = \sum_{i=1}^{n} X_i$, სადაც X_i – A ხდომილების მოხდენათა რიცხვია i –ურ ცდაში (ანუ ბერნულის შემთხვევითი სიდიდეებია მნიშვნელობებით 0 ან 1). მაშინ თეორემა 1-ის თანახმად $Z = \frac{Y - m_y}{\sigma_y}$ შემთხვევითი სიდიდე შეიძლება ჩაითვალოს ნორმალური კანონით განაწილებულ, ნორმირებულ (სტანდარტიზებულ) შემთხვევით სიდიდედ. შესაბამისად, მისი (α, β) ინტერვალში მოხვედრის ალბათობა გამოითვლება ფორმულით

$$p(\alpha < Z < \beta) = \Phi(\beta) - \Phi(\alpha).$$

ვინაიდან Y შემთხვევით სიდიდეს აქვს ბინომიალური განაწილება,

$$m_y = np,\ D_y = npq,\ \sigma_y = \sqrt{npq}.$$

ამიტომ $Z = \frac{Y-np}{\sqrt{npq}}$. თუ ჩავსვამთ ამ გამოსახულებას წინა ფორმულაში, მივიღებთ დასამტკიცებელ თანაფარდობას.

ეს თეორემა ლიტერატურაში აგრეთვე ცნობილია მუავრ-ლაპლასის ინტეგრალური თეორემის სახელწოდებით.

შედეგი (მუავრ-ლაპლასის ლოკალური თეორემა). მუავრ-ლაპლასის თეორემის პირობებში $p_n(k)$ -- ალბათობა იმისა, რომ A ხდომილება n ცდაში მოხდება ზუსტად k –ჯერ, ცდათა დიდი რიცხვის შემთხვევაში, თუ $np > 15$, შეიძლება გამოითვალოს შემდეგი ფორმულით:

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

სადაც $x = \frac{k-np}{\sqrt{npq}}$, ხოლო $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ (ამ ფუნქციის მნიშვნელობები მოყვანილია სპეციალურ ცხრილებში, ამასთანავე $\varphi(-x) = \varphi(x)$).

შემოვიღოთ განაწილების ფუნქციებისათვის შემდეგი აღნიშვნები:

ჰიპერგეომეტრიული განაწილების ფუნქცია -- $H(x;t,s,n)=\sum_{k\le x}\frac{C_t^k\cdot C_{n-t}^{s-k}}{C_n^s}$;

ბინომური განაწილების ფუნქცია -- $Bi(x;p,n)=\sum_{k\le x}C_n^k p^k q^{n-k}$;

პუასონის განაწილების ფუნქცია -- $\Pi(x;\lambda)=\sum_{0\le k\le x}\frac{\lambda^k}{k!}e^{-\lambda}$;

სტანდარტული ნორმალური განაწილების ფუნქცია –

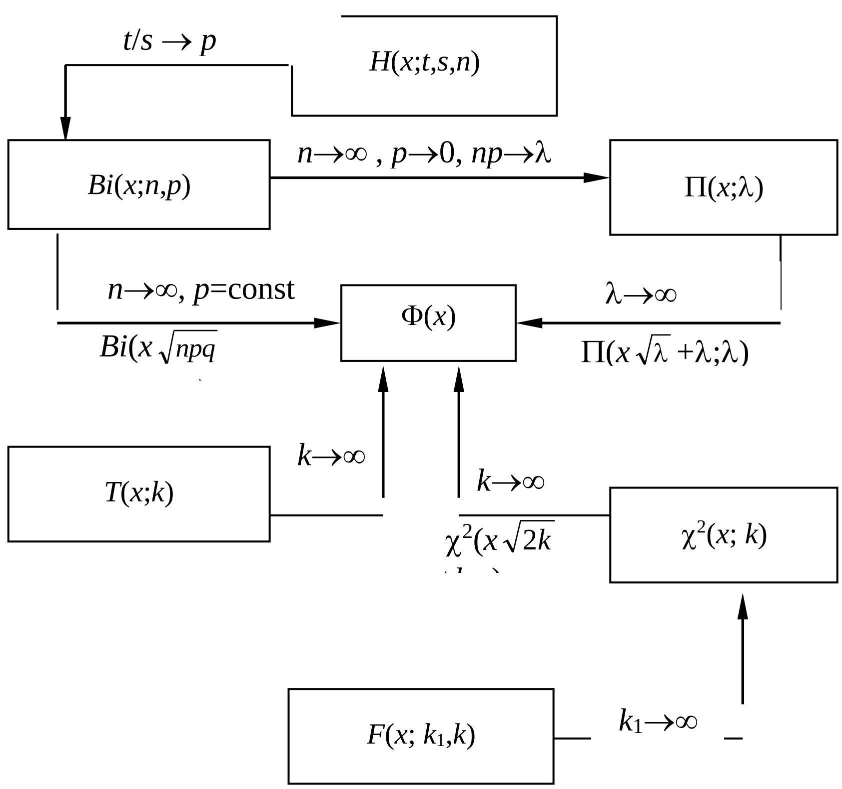
$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt;$$

ხი კვადრატ განაწილების ფუნქცია (იხ. §. 29) -- $\chi^2(x;k)$;

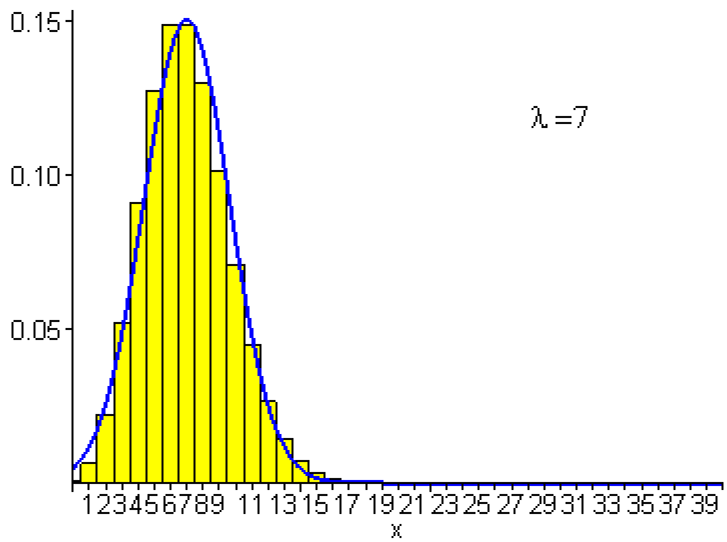
სტიუდენტის განაწილების ფუნქცია (იხ. §. 29) -- $T(x;k)$;

ფიშერის განაწილების ფუნქცია (იხ. §. 29) -- $F(x;k_1,k_2)$.

აღსანიშნავია, რომ ზემოთ მოყვანილი განაწილებების ფუნქციებს შორის ადგილი აქვს ქვემოთ მოყვანილ სქემაზე გამოსახულ ზღვრულ თანაფარდობებს:



ქვემოთ მოყვანილ ნახაზზე შედარებულია ნორმალური და პუასონის განაწილებები $\lambda=7$-ის შემთხვევაში:



მაგალითი 1. ვიპოვოთ ალბათობა იმისა, რომ მონეტის 100-ჯერ აგდებისას გერბთა მოსვლის რიცხვი აღმოჩნდება საზღვრებში 40-დან 60-მდე.

ვისარგებლოთ მუავრ-ლაპლასის ინტეგრალური თეორემით. ჩვენს შემთხვევაში $p = 0.5$, $np = 100 \cdot 0.5 = 50$. ამიტომ თუ $40 < Y < 60$, მაშინ

$$-2 < \frac{Y - 50}{5} < 2.$$

შესაბამისად, გვაქვს:

$$p(40 < Y < 60) = p\left(-2 < \frac{Y-50}{5} < 2\right) = \Phi(2) - \Phi(-2) = 0,9772 - 0,0228 = 0,9544.$$

მაგალითი 2. წინა მაგალითის პირობებში ვიპოვოთ ალბათობა იმისა, რომ გერბი მოვა 45-ჯერ.

ამ შემთხვევაში $x = \frac{k - np}{\sqrt{npq}} = \frac{45 - 50}{5} = -1$, ამიტომ

$$p_{100}(45) \approx \frac{1}{5} \cdot \varphi(-1) = \frac{1}{5} \cdot \varphi(1) = \frac{1}{5} \cdot 0.2420 = 0.0484.$$

§27. მათემატიკური სტატისტიკის ძირითადი ცნებები

მათემატიკური სტატისტიკის ძირითადი მიზანია ისეთი მეთოდების დამუშავება, რომლებიც დაკვირვებებისა და ექსპერიმენტების შედეგებზე დაყრდნობით, მასობრივ მოვლენებზე და პროცესებზე მეცნიერულად დასაბუთებული დასკვნების მიღების საშუალებას იძლევა. ეს დასკვნები შეეხება არა ცალკეულ ექსპერიმენტებს, რომელთა განმეორებითაც ყალიბდება მოცემული მასობრივი მოვლენა, არამედ წარმოადგენს მტკიცებულებებს მოცემული პროცესის ზოგადი ალბათური მახასიათებლების შესახებ (ანუ ალბათობებზე, განაწილებების კანონებზე, მათემატიკურ ლოდინებზე, დისპერსიებზე და ა. შ.).

დავუშვათ, რომ ჩვენ გაგვაჩნია მონაცემები, მაგალითად, გარკვეულ პირობებში დამზადებულ პროდუქციაში დეფექტური ნაწარმის რიცხვის შესახებ ან ნაწარმის გამძლეობაზე შემოწმების ექსპერიმენტის შედეგების შესახებ და ა. შ. ჩვენს მიერ შეგროვილი მონაცემები შესაძლებელია წარმოადგენდეს უშუალო ინტერესის საგანს პროდუქციის ამა თუ იმ პარტიის ხარისხზე ინფორმაციის თვალსაზრისით. სტატისტიკური ამოცანა კი იწყება მაშინ, როდესაც ჩვენ იმავე ინფორმაციაზე დაყრდნობით ვიწყებთ დასკვნების გაკეთებას მოვლენათა უფრო ფართო წრის შესახებ. ასე მაგალითად, ჩვენ შეიძლება გვაინტერესებდეს ტექნოლოგიური პროცესის ხარისხი, რისთვისაც ჩვენ ვაფასებთ ამ პროცესში დეფექტური ნაწარმის მიღების ალბათობას ან ნაწარმის საშუალო სიცოცხლის ხანგრძლივობას. ამ შემთხვევაში, შეგროვილ მასალას ჩვენ ვიხილავთ როგორც გარკვეულ საცდელ ჯგუფს ან შერჩევას, რომელიც წარმოადგენს მხოლოდ სერიას შესაძლო შედეგებიდან, რომლებიც შესაძლებელია შეგვხვდეს მოცემულ პირობებში მასობრივ პროცესზე დაკვირვებების გაგრძელების შემთხვევაში. დაკვირვებების შედეგების საფუძველზე გაკეთებული დასკვნები და შეფასებები ასახავენ საცდელი ჯგუფის შემთხვევით შემადგენლობას და ამიტომ ითვლება, რომ ისინი ალბათური ხასიათის მიახლოებითი შეფასებებია. თეორია გვიჩვენებს, თუ როგორ უნდა გამოვიყენოთ არსებული ინფორმაცია იმისათვის, რომ მივიღოთ რაც შეიძლება ზუსტი და საიმედო მახასიათებლები და ამასთანავე მივუთითოთ მონაცემთა მარაგის შეზღუდულობით გამოწვეული დასკვნების საიმედოობის ხარისხი.

მათემატიკური სტატისტიკა იხილავს ამოცანათა ორი ძირითად კატეგორიას: შეფასება და ჰიპოთეზათა სტატისტიკური შემოწმება. პირველი ამოცანა, თავის მხრივ, იყოფა განაწილების პარამეტრების წერტილოვან და ინტერვალურ შეფასებებად. მაგალითად, შესაძლებელია დაკვირვებების საფუძველზე წარმოიშვას მათემატიკური ლოდინისა და დისპერსიის წერტილოვანი შეფასების აუცილებლობა. თუ კი ჩვენ გვინდა მივიღოთ რაიმე ინტერვალი, რომელიც ამა თუ იმ საიმედოობის ხარისხით მოიცავს პარამეტრის ჭეშმარიტ მნიშვნელობას, მაშინ ეს არის ინტერვალური შეფასების ამოცანა.

მეორე ამოცანა – ჰიპოთეზათა შემოწმება – მდგომარეობს იმაში, რომ ჩვენ ვაკეთებთ დაშვებას შემთხვევითი სიდიდის ალბათური განაწილ-

ების შესახებ (მაგალითად, განაწილების ფუნქციის სახის შესახებ, ან განაწილების ფუნქციის ერთი ან რამოდენიმე პარამეტრის მნიშვნელობის შესახებ) და ვადგენთ არის თუ არა განაწილების სახე ან პარამეტრების მნიშვნელობები შესაბამისობაში (გარკვეული აზრით) დაკვირვებების მიღებულ შედეგებთან.

შერჩევითი მეთოდი. დავუშვათ, რომ ჩვენი მიზანია შევისწავლოთ საქონლის გარკვეული პარტიის რაოდენობრივი ნიშანი. პარტიის შემოწმება (კონტროლი) შესაძლებელია მოხდეს ორი გზით:

1. ჩავატაროთ მთელი პარტიის შემოწმება;
2. ჩავატაროთ პარტიის მხოლოდ ნაწილის შემოწმება.

პირველი გზა ყოველთვის არაა განხორციელებადი, მაგალითად, პარტიაში საქონლის დიდი რიცხვის გამო, შემოწმების ოპერაციის ჩატარების სიძვირის გამო ან იმის გამო, რომ შემოწმება იწვევს საქონლის განადგურებას (ელექტრო ნათურის შემოწმებისას მუშაობის ხანგრძლივობაზე).

მეორე შემთხვევაში, შემთხვევითი გზით შერჩეული ობიექტების სიმრავლეს **შერჩევითი ერთობლიობა ან შერჩევა** ეწოდება. ობიექტთა მთლიან ერთობლიობას, საიდანაც ხდება შერჩევა, **გენერალური ერთობლიობა** ეწოდება. შერჩევაში ელემენტთა რაოდენობას **შერჩევის მოცულობა** ეწოდება. როგორც წესი, ითვლება, რომ გენერალური ერთობლიობის მოცულობა უსასრულოა.

შერჩევა შეიძლება იყოს **განმეორებითი** (დაბრუნებით) და **განმეორების გარეშე** (დაბრუნების გარეშე).

ჩვეულებრივ, ხორციელდება განმეორების გარეშე შერჩევები, თუმცა გენერალური ერთობლიობის მოცულობის სიდიდის (უსასრულობის) გამო, მხოლოდ განმეორებითი შერჩევების დროს სამართლიანი გათვლები წარმოებს და დასკვნები კეთდება.

შერჩევა საკმარისად სრულად უნდა ასახავდეს გენერალური ერთობლიობის ყველა ობიექტის განსაკუთრებულობებს, ანუ სხვა სიტყვებით, რომ ვთქვათ, შერჩევა უნდა იყოს **რეპრეზენტატული** (წარმომადგენლობითი).

ობიექტების ამორჩევის ხერხების მიხედვით განასხვავებენ შემდეგი ტიპის შერჩევებს:

1. *მარტივი შემთხვევითი ამორჩევა.*

გენერალური ერთობლიობის ყველა ელემენტი გადაინომრება და შემთხვევით რიცხვთა ცხრილიდან იღებენ, მაგალითად, ნებისმიერ ერთმანეთის მომდევნო 50 რიცხვის მიმდევრობას და შერჩევაში შეყავთ ამოსული ნომრების მქონე ობიექტები.

2. *ტიპიური ამორჩევა.*

ასეთი ამორჩევა წარმოებს იმ შემთხვევაში, თუ გენერალური ერთობლიობა შესაძლებელია წარმოდგეს ისეთ ქვესიმრავლეთა გაერთიანებად, რომელთა ელემენტები ერთგვაროვანია რაიმე ნიშნის მიხედვით, თუმცა მთელ ერთობლიობას ასეთი ერთგვაროვნება არ გააჩნია (საქონლის პარტია შედგება რამოდენიმე ჯგუფისაგან, რომლებიც წარმოებულია სხვადასხვა საწარმოს მიერ). მაშინ, თითოეულ ქვესიმრავლეში ატარებენ მარტივ

შემთხვევით შერჩევას, და შერჩევაში აერთიანებენ ყველა მიღებულ ობიექტს.

3. *მექანიკური ამორჩევა.*

გენერალური ერთობლიობიდან იღებენ ყოველ მეცხრე (ორმოცდამეათე) ობიექტს.

4. *სერიული ამორჩევა.*

შერჩევაში აერთიანებენ იმ ობიექტებს, რომლებიც წარმოებულია რაიმე წარმოების სფეროში დროის გარკვეულ ინტერვალში.

შემდგომში, გენერალური ერთობლიობის ქვეშ, ჩვენ ვიგულისხმებთ არა თვითონ ობიექტთა სიმრავლეს, არამედ იმ შემთხვევითი სიდიდის მნიშვნელობათა სიმრავლეს, რომელიც ღებულობს რიცხვით მნიშვნელობებს თითოეულ ობიექტზე. სინამდვილეში, გენერალური ერთობლიობა, როგორც ობიექტთა სიმრავლე შეიძლება არც არსებობდეს. მაგალითად, აზრი აქვს ვილაპარაკოთ იმ დეტალების სიმრავლეზე, რომლებიც შეიძლება წარმოებულ იქნეს, თუ გამოვიყენებთ მოცემულ ტექნოლოგიურ პროცესს. გამოვიყენებთ რა ამ პროცესის ჩვენთვის ცნობილ მახასიათებლებს, ჩვენ შეგვიძლია შევაფასოთ დეტალების ამ არ არსებული სიმრავლის პარამეტრები. დეტალის ზომა – ეს შემთხვევითი სიდიდეა, რომლის მნიშვნელობა განისაზღვრება ტექნოლოგიური პროცესის შემადგენელი მრავალი ფაქტორის ზემოქმედებით. ჩვენ, მაგალითად, შეიძლება გვაინტერესებდეს თუ რა ალბათობით ღებულობს ეს შემთხვევითი სიდიდე მნიშვნელობას გარკვეული ინტერვალიდან. ამ კითხვაზე პასუხის გაცემა შესაძლებელია თუ ჩვენ გვეცოდინება ამ შემთხვევითი სიდიდის განაწილების კანონი და მისი ისეთი პარამეტრები, როგორიცაა ლოდინი და დისპერსია.

ამრიგად, გენერალური ერთობლიობის, როგორც ობიექტთა სიმრავლის ცნებიდან, რომლებიც ხასიათდებიან გარკვეული ნიშნით (თვისებით), ჩვენ გადავდივართ გენერალურ ერთობლიობაზე, როგორც შემთხვევით სიდიდეზე, რომლის განაწილების კანონი და პარამეტრები განისაზღვრება შერჩევითი მეთოდის საშუალებით.

განვიხილოთ n მოცულობის შერჩევა, რომელიც წარმოადგენს მოცემულ გენერალურ ერთობლიობას. პირველ შერჩევით მნიშვნელობას x_1-ს განვიხილავთ როგორც რეალიზაციას, როგორც ერთ-ერთ შესაძლებელ მნიშვნელობას ξ_1 შემთხვევითი სიდიდის, რომელსაც გააჩნია იგივე განაწილების კანონი რაც ξ შემთხვევით სიდიდეს. მეორე შერჩევითი მნიშვნელობა x_2-ს განვიხილავთ როგორც ერთ-ერთ შესაძლებელ მნიშვნელობას ξ_2 შემთხვევითი სიდიდის , რომელსაც გააჩნია იგივე განაწილების კანონი რაც ξ შემთხვევით სიდიდეს და ა. შ. იგივე შეიძლება ითქვას $x_3, x_4, ..., x_n$ მნიშვნელობებზე.

ამრიგად, შერჩევას ჩვენ ვუყურებთ როგორც ერთობლიობას დამოუკიდებელი $\xi_1, \xi_2, ..., \xi_n$ შემთხვევითი სიდიდეების, რომლებიც იმავე კანონით არიან განაწილებული როგორც ξ შემთხვევითი სიდიდე, რომელიც წარმოადგენს გენერალურ ერთობლიობას. შერჩევითი მნიშვნელობები $x_1, x_2,$

$..., x_n$ -- *ეს ის მნიშვნელობებია, რაც მიიღეს შემთხვევითმა სიდიდეებმა პირველი, მეორე, და ა. შ. მე-*n *ექსპერიმენტის შედეგად.*

ვარიაციული მწკრივი. დავუშვათ, რომ გენერალური ერთობლიობის ობიექტებისათვის განსაზღრულია გარკვეული ნიშანი ან რიცხვითი მახასიათებელი, რომლის გაზომვა შესაძლებელია (დეტალის სიგრძე, ნიტრატების ფარდობითი შემცველობა საზამთროში, ძრავის მუშაობის ხმაური). ეს მახასიათებელი არის შემთხვევითი სიდიდე ξ, რომელიც ყოველ ობიექტზე ღებულობს გარკვეულ რიცხვით მნიშვნელობას. n მოცულობის შერჩევიდან ვღებულობთ ამ შემთხვევითი სიდიდის მნიშვნელობებს n რიცხვისგან შედგენილი მწკრივის სახით:

$$x_1, x_2, ..., x_n . \qquad (1)$$

ამ რიცხვებს ნიშნის მნიშვნელობებს უწოდებენ.

(1) მწკრივის რიცხვებს შორის შესაძლებელია იყოს ერთი და იგივე რიცხვები. თუ ნიშნის მნიშვნელობებს დავალაგებთ, ანუ რიცხვებს განვალაგებთ ზრდადობის ან კლებადობის მიხედვით, ამასთანავე ყოველ მნიშვნელობას დავწერთ მხოლოდ ერთჯერ, ხოლო შემდეგ ყოველი x_i მნიშვნელობის ქვეშ დავწერთ m_i რიცხვს, რომელიც გვიჩვენებს თუ რამდენჯერ შეგვხვდა x_i მნიშვნელობა (1) მწკრივში, მივიღებთ ცხრილს, რომელსაც **დისკრეტული ვარიაციული მწკრივი** ეწოდება:

x_1	x_2	x_3	...	x_k
m_1	m_2	m_3	...	m_k

m_i რიცხვს ნიშნის i-ური მნიშვნელობის სიხშირე ეწოდება.

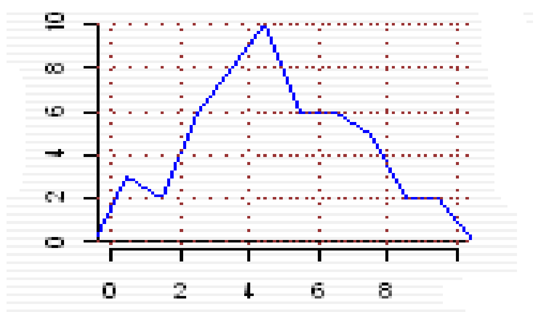
ცხადია, რომ (1) მწკრივის x_i შეიძლება არ ემთხვევოდეს x_i-ს ვარიაციული მწკრივიდან. ნათელია აგრეთვე, რომ

$$\sum_{i=1}^{k} m_i = n .$$

თუ შერჩევის მინიმალურ და მაქსიმალურ მნიშვნელობებს შორის ინტერვალს გავყოფთ ერთი და იგივე სიგრძის რამოდენიმე ინტერვალად, და ყოველ ინტერვალს შევუსაბამებთ ამ ინტერვალში მიხვედრილი ნიშნის შერჩევითი მნიშვნელობების რიცხვს, მაშინ მივიღებთ **ინტერვალურ ვარიაციულ მწკრივს.** თუ ნიშანს შეუძლია მიიღოს ნებისმიერი მნიშვნელობა გარკვეული ინტერვალიდან, ე. ი. წარმოადგენს უწყვეტ შემთხვევით სიდიდეს, მაშინ შერჩევა სწორედ ასეთი მწკრივის სახით უნდა წარმოვადგინოთ. თუ ინტერვალურ ვარიაციულ მწკრივში ყოველ $[\alpha_i, \alpha_{i+1})$ ინტერვალს შევცვლით ამ ინტერვალის შუაში მდებარე რიცხვით -- $(\alpha_i + \alpha_{i+1})/2$, მაშინ მივიღებთ დისკრეტულ ვარიაციულ მწკრივს. ასეთი შეცვლა სრულიად ბუნებრივია, ვინაიდან, მაგალითად, დეტალის სიგრძის გაზომვისას ერთი მილიმეტრის სიზუსტით, ყველა სიგრძეს [49.5,50.5) ინტერვალიდან შეესაბამება ერთი რიცხვი, რომელიც ტოლია 50-ის.

§28. ემპირიული განაწილების ფუნქცია

გამოსაკვლევი შემთხვევითი სიდიდის თვალსაჩინო წარმოსახვისათვის შერჩევის მიხედვით შესაძლებელია აიგოს სხვადასხვა გრაფიკები. ერთ-ერთიასეთი გრაფიკია – **სიხშირეთა პოლიგონი**: ტეხილი, რომლის მონაკვეთები აერთებენ წერტილებს კოორდინატებით $(x_1, n_1), (x_2, n_2),\ldots, (x_k, n_k)$, სადაც x_i გადაიზომება აბსცისთა ღერძზე, ხოლო n_i – ორდინატთა ღერძზე. თუ ორდინატთა ღერძზე გადავზომავთ არ აბსულუტურ (n_i), არამედ ფარდობით (w_i) სიხშირეებს, მაშინ მივიღებთ **გარდობით სიხშირეთა პოლიგონს:**



შემთხვევითი სიდიდის განაწილების ფუნქციის ანალოგიით, შეიძლება განსაზღრულ იქნეს გარკვეული ფუნქცია, კერძოდ, $X \le x$ ხდომილების ფარდობითი სიხშირე.

განმარტება. შერჩევით (ემპირიულ) განაწილების ფუნქციას უწოდებენ ფუნქციას $F^*(x)$, რომელიც x –ის ნებისმიერი მნიშვნელობისათვის განსაზღვრავს $X \le x$ ხდომილების ფარდობით სიხშირეს. ამრიგად,

$$F^*(x) = \frac{n_x}{n},$$

სადაც n_x – ვარიანტების რიცხვია, რომლებიც არ არემატება x –ს, ხოლო n – შერჩევის მოცულობა.

განსხვავებით ემპირიული განაწილების ფუნქციისაგან, რომელიც იგება შერჩევის მიხედვით, გენერალური ერთობლიობის $F(x)$ განაწილების ფუნქციას *თეორიული განაწილების ფუნქციას* უწოდებენ. იგი განსაზღვრავს $X \le x$ ხდომილების ალბათობას, ხოლო $F^*(x)$ -- მის ფარდობით სიხშირეს. საკმაოდ დიდი n–ებისათვის, როგორც ამას ამტკიცებს დიდ რიცხვთა კანონი, $F^*(x)$ ფუნქცია კრებადია ალბათობით $F(x)$ ფუნქციისაკენ.

ემპირიული განაწილების ფუნქციის განმარტებიდან ადვილი დასანახია, რომ მისი თვისებები ემთხვევა $F(x)$ ფუნქციის თვისებებს, კერძოდ:

1. $0 \le F^*(x) \le 1$.

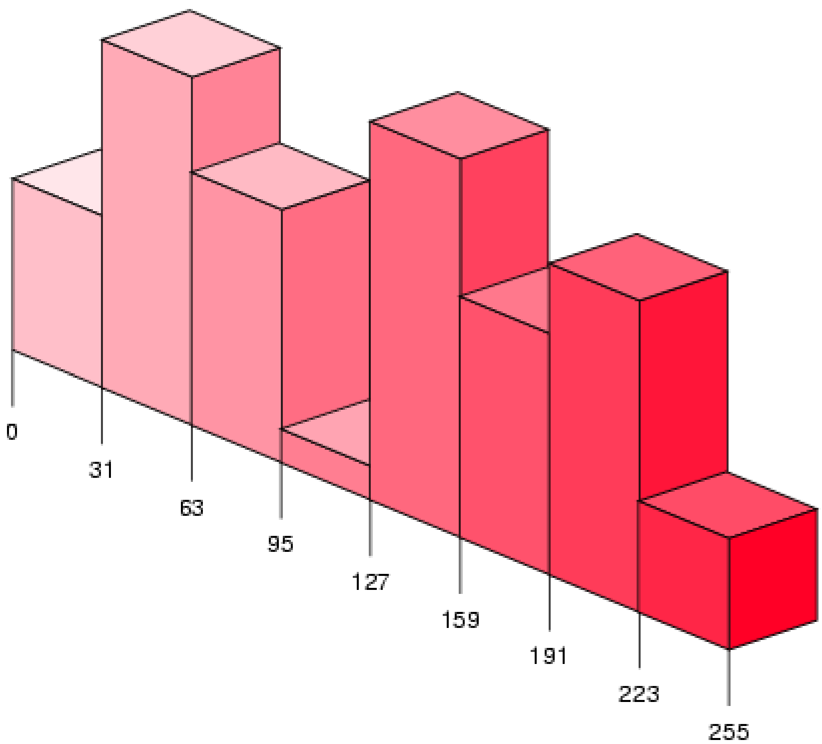
2. $F^*(x)$ – არაკლებადი ფუნქციაა.

3. თუ x_1– უმცირესი ვარიანტია, მაშინ $F^*(x) = 0$, როცა $x < x_1$; თუ x_k – უდიდესი ვარიანტია, მაშინ $F^*(x) = 1$, როცა $x \ge x_k$.

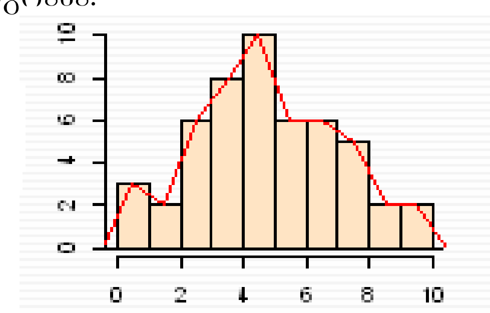
4. $F^*(x)$ – მარჯვნიდან უწყვეტი ფუნქციაა.

უწყვეტი მონაცემების შემთხვევაში გრაფიკულ ილუსტრაციას წარმოადგენს ე. წ. **ჰისტოგრამა,** ე. ი. საფეხურა ფიგურა, რომელიც შედგება მა-

რთკუთხედებისაგან, რომელთა ფუძეებია h სიგრძის ინტერვალები, ხოლო სიმაღლეები – მონაკვეთები სიგრძით n_i/h (სიხშირეების ჰისტოგრამა) ან w_i/h (ფარდობითი სიხშირეების ჰისტოგრამა). პირველ შემთხვევაში ჰისტოგრამის ფართობი ტოლია შერჩევის მოცულობის, ხოლო მეორე შემთხვევაში – ერთის.



ჰისტოგრამა წარმოდგენას გვაძლევს გენერალური ერთობლიობის განაწილების სიმკვრივეზე. შერჩევის დიდი მოცულობის შემთხვევაში ის ახლოსაა თეორიულ სიმკვრივესთან. ქვემოთ, ერთ ნახაზზე, მოყვანილია პოლიგონი და ჰისტოგრამა.



§29. გენერალური ერთობლიობის პარამეტრების წერტილოვანი შეფასებები

ძალიან ბევრ შემთხვევაში ჩვენ გაგვაჩნია ინფორმაცია შემთხვევითი სიდიდის განაწილების კანონის სახის შესახებ (ნორმალური, ბერნულის, თანაბარი და ა. შ.), მაგრამ არ ვიცით ამ განაწილების ისეთი პარამეტრები, როგორიცაა $E\xi$ და $D\xi$. ამ პარამეტრების განსაზღვრისათვის გამოიყენება შერჩევითი მეთოდი.

დავუშვათ, რომ n მოცულობის შერჩევა წარმოდგენილია ვარიაციული მწკრივის სახით. **შერჩევითი საშუალო** ეწოდება სიდიდეს:

$$\bar{x}=\frac{x_1m_1+x_2m_2+\cdots+x_km_k}{n}=x_1\frac{m_1}{n}+x_2\frac{m_2}{n}+\cdots+x_k\frac{m_k}{n}.$$

სიდიდეს $\omega_i=m_i/n$ ნიშნის x_i მნიშვნელობის **ფარდობითი სიხშირე** ეწოდება. თუ შერჩევიდან მიღებულ ნიშნის მნიშვნელობებს არ დავაჯგუფებთ და არ წარმოვადგენთ ვარიაციული მწკრივის სახით, მაშინ შერჩევითი საშუალოს გამოსათვლელად უნდა ვისარგებლოთ შემდეგი ფორმულით:

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i.$$

ბუნებრივია $\bar{x}$ სიდიდე ჩაითვალოს $E\xi$ პარამეტრის შერჩევით შეფასებად. პარამეტრის შერჩევით შეფასებას, რომელიც წარმოადგენს რიცხვს, **წერტილოვანი შეფასება** ეწოდება.

შერჩევითი დისპერსია ეწოდება სიდიდეს:

$$s^2=\sum_{i=1}^{k}(x_i-\bar{x})^2\omega_i=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2.$$

ის შეიძლება ჩაითვალოს გენერალური ერთობლიობის $D\xi$ დისპერსიის წერტილოვან შეფასებად.

დავუშვათ, რომ გენერალური ერთობლიობის ყოველი ობიექტი ხასიათდება ორი რაოდენობრივი x და y ნიშნით. მაგალითად, დეტალს შეიძლება ჰქონდეს ორი ზომა – სიგრძე და სიგანე, შეიძლება სხვადასხვა რეგიონში გაიზომოს მავნე ნივთიერებების კონცენტრაცია და დაფიქსირდეს თვის განმავლობაში მოსახლეობაში ფილტვების დაავადებების რაოდენობა, შეიძლება დროის ტოლ შუალედებში შევადაროთ მოცემული კორპორაციის აქციების შემოსავლიანობა რაიმე ინდექსს, რომელიც ახასიათებს აქციების მთელი ბაზრის საშუალო შემოსავლიანობას. ასეთ შემთხვევაში, გენერალური ერთობლიობა წარმოადგენს ორგანზომილებიან შემთხვევით სიდიდეს ξ,η. ეს შემთხვევითი სიდიდე გენერალური ერთობლიობის ობიექტების სიმრავლეზე ღებულობს მნიშვნელობებს x,y. თუ ჩვენ არ ვიცით ξ და η შემთხვევითი სიდიდეების ერთობლივი განაწილების კანონი, ჩვენ არ შეგვიძლია ვილაპარაკოთ მათ შორის კორელაციური კავშირის არსებ-

ობაზე ან სიძლიერეზე, მაგრამ, მიუხედავად ამისა, შერჩევითი მეთოდის გამოყენებით შესაძლებელია ზოგიერთი დასკვნის გაკეთება.

ასეთ შემთხვევაში, n მოცულობის შერჩევა წარმოიდგინება ცხრილის სახით, სადაც i-ური ამორჩეული ობიექტი ($i=1,2,...,n$) წარმოდგენილია რიცხვთა წყვილით x_i, y_i:

x_1	x_2	...	x_n
y_1	y_2	...	y_n

შერჩევითი კორელაციის კოეფიციენტი ეწოდება სიდიდეს:

$$r_{xy} = \frac{\overline{xy} - \overline{x}\,\overline{y}}{s_x s_y},$$

სადაც

$$\overline{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i, \quad s_x = \sqrt{s_x^{\ 2}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2},$$

$$s_y = \sqrt{s_y^{\ 2}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}.$$

შერჩევითი კორელაციის კოეფიციენტი შეიძლება განხილულ იქნეს როგორც წერტილოვანი შეფასება კორელაციის კოეფიციენტის $\rho_{\xi\eta}$, რომელიც ახასიათებს გენერალურ ერთობლიობას.

შერჩევითი პარამეტრები $\overline{x}$, s^2, r_{xy} ან ნებისმიერი სხვა დამოკიდებულია იმაზე, გენერალური ერთობლიობის რომელი ობიექტები მოხვდნენ შერჩევაში და განსხვავდებიან შერჩევიდან შერჩევამდე. ამიტომ ისინი თვითონ წარმოადგენენ შემთხვევით სიდიდეებს.

დავუშვათ, რომ შერჩევითი პარამეტრი δ განიხილება როგორც გენერალური ერთობლიობის Δ პარამეტრის შერჩევითი შეფასება. შერჩევით შეფასებას ეწოდება **გადაუადგილებადი** (ან **ჩაუნაცვლებელი**), თუ

$$E\delta = \Delta.$$

იმისათვის რომ დავამტკიცოთ ზოგიერთი წერტილოვანი შეფასების გადაუადგილებადობა, n მოცულობის შერჩევას განვიხილავთ როგორც სისტემას n დამოუკიდებელი $\xi_1, \xi_2, ..., \xi_n$ შემთხვევითი სიდიდის, რომელთაგან თითოეული გააჩნია იგივე განაწილების კანონი, იგივე პარამეტრებით, რაც ξ შემთხვევით სიდიდეს, რომელიც წარმოადგენს გენერალურ ერთობლიობას. ასეთი მიდგომის შემთხვევაში ცხადი ხდება თანაფარდობები:

$$Ex_i = E\xi_i = E\xi; \;\; Dx_i = D\xi_i = D\xi \;\; (\,k = 1,2,...n).$$

ახლა ვაჩვენოთ, რომ შერჩევითი საშუალო $\overline{x}$ წარმოადგენს გენერალური ერთობლიობის საშუალოს გადაუადგილებად შეფასებას, ან რაც იგივეა ξ შემთხვევითი სიდიდის მათემატიკური ლოდინის გადაუადგილებად შეფასებას. მართლაც, გვაქვს:

$$E\overline{x} = E\frac{x_1 + x_2 + ... + x_n}{n} = \frac{1}{n}\left(E\xi_1 + E\xi_2 + ... + E\xi_n\right) = \frac{1}{n}nE\xi = E\xi .$$

გამოვთვალოთ შერჩევითი საშუალოს დისპერსია. ცხადია, რომ:

$$D\overline{x} = D\frac{x_1 + x_2 + ... + x_n}{n} = \frac{1}{n^2}\left(D\xi_1 + D\xi_2 + ... + D\xi_{n1}\right) = \frac{1}{n^2}nD\xi = \frac{D\xi}{n} .$$

ვიპოვოთ ახლა რისი ტოლია შერჩევითი დისპერსიის მათემატიკური ლოდინი, რისთვისაც თავიდან σ^2 გარდავქმნათ შემდეგნაირად:

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - E\xi + E\xi - \overline{x}\right)^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(\left(x_i - E\xi\right)^2 - 2\left(x_i - E\xi\right)\left(\overline{x} - E\xi\right) + \left(\overline{x} - E\xi\right)^2\right) =$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(x_i - E\xi\right)^2 - \left(\overline{x} - E\xi\right)^2$$

(შევნიშნავთ, რომ ჩვენ აქ გამოვიყენეთ გარდაქმნა:

$$\sum_{i=1}^{n}2\left(x_i - E\xi\right)\left(\overline{x} - E\xi\right) = 2\left(\overline{x} - E\xi\right)\sum_{i=1}^{n}\left(x_i - E\xi\right) =$$

$$= 2\left(\overline{x} - E\xi\right)\left(\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}E\xi\right) = 2\left(\overline{x} - E\xi\right)\left(n\overline{x} - nE\xi\right) = 2n\left(\overline{x} - E\xi\right)^2).$$

ამიტომ შერჩევითი დისპერსიის მათემატიკური ლოდინი იქნება:

$$E\sigma^2 = E\left(\frac{1}{n}\sum_{i=1}^{n}\left(x_i - E\xi\right)^2 - \left(\overline{x} - E\xi\right)^2\right) =$$

$$= \frac{1}{n}\sum_{i=1}^{n}E\left(x_i - E\xi\right)^2 - E\left(\overline{x} - E\xi\right)^2 = \frac{1}{n}nD\xi - D\overline{x} =$$

$$= D\xi - \frac{D\xi}{n} = \frac{n-1}{n}D\xi .$$

როგორც ვხედავთ, $Es^2 \ne D\xi$. ამიტომ *შერჩევითი დისპერსია არ წარმოადგენს გენერალური ერთობლიობის დისპერსიის გადაუადგილებად შეფასებას.*

იმისათვის რომ მივიღოთ გენერალური ერთობლიობის დისპერსიის გადაუადგილებადი შეფასება, საჭიროა შერჩევითი დისპერსია გავამრავლოთ მამრავლზე $n/(n-1)$. მიღებული სიდიდე აღინიშნება s'^2-ით და მას **შესწორებული შერჩევითი დისპერსია** ეწოდება:

$$s'^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 .$$

თუ ჩვენ გვაქვს გენერალური ერთობლიობის ერთი და იგივე პარამეტრის რამოდენიმე გადაუადგილებადი შეფასება, მაშინ იმ შეფასებას, რომელსაც გააჩნია უმცირესი დისპერსია, **ეფექტური** ეწოდება.

n მოცულობის შერჩევიდან მიღებულ გენერალური ერთობლიობის Δ პარამეტრის წერტილოვან δ_n შეფასებას ეწოდება **ძალმოსილი**, თუ ის

ალბათობით კრებადია Δ-სკენ. ეს იმას ნიშნავს, რომ ნებისმიერი დადებითი ε და γ რიცხვებისათვის, მოიძებნება ისეთი რიცხვი $n_{\varepsilon\gamma}$, რომ ყველა n რიცხვისათვის, რომელიც აკმაყოფილებს უტოლობას $n > n_{\varepsilon\gamma}$, სრულდება პირობა

$$P(|\delta_n - \Delta| < \varepsilon) > 1 - \gamma .$$

შევნიშნავთ, რომ $\bar{x}$ და s'^2 შესაბამისად წარმოადგენენ გენერალური ერთობლიობის მათემატიკური ლოდინისა და დისპერსიის გადაუადგილებად, ძალმოსილ და ეფექტურ შეფასებებს.

§30. შერჩევითი პარამეტრების განაწილება ნორმალური პოპულაციისათვის

დავუშვათ, რომ $\xi_1, \xi_2, ..., \xi_n$ წარმოადგენს შერჩევას ნორმალური გენერალური ერთობლიობიდან, $\xi_i = N(x;\mu;\sigma^2)$, $i = 1, 2, ..., n$. რადაგანაც $\overline{x}$ წარმოადგენს დამოუკიდებელი ნორმალურად განაწილებული შემთხვევითი სიდიდეების წრფივ კომბინაციას, ამიტომ

$$\overline{x} = N(x;\mu;\sigma^2/n).$$

გავარკვიოთ ახლა შერჩევითი დიპერსიის განაწილების კანონი. ჯერ განვიხილოთ ის შემთხვევა, როცა გენერალური ერთობლიობის საშუალო (მათემატიკური ლოდინი) ცნობილია. რადგანაც, $\xi_i = N(x;\mu;\sigma^2)$, ამიტომ

$$(\xi_i - \mu)/\sigma = N(x;0;1).$$

თუ გავითვალისწინებთ, რომ შერჩევით დისპერსიას აქვს სახე

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(\xi_i - \mu)^2,$$

ადვილად დავრწმუნდებით, რომ სამართლიანია თანაფარდობა:

$$\frac{ns^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{\xi_i - \mu}{\sigma}\right)^2.$$

ამრიგად, ns^2/σ^2 წარმოიდგინება $N(x;0;1)$ კანონით განაწილებული დამოუკიდებელი შემთხვევითი სიდიდეების კვადრატების ჯამის სახით. ე. ი. მას აქვს χ^2 განაწილება თავისუფლების ხარისხით n:

$$\frac{ns^2}{\sigma^2} = \chi^2(n).$$

უცნობი საშუალოს შემთხვევაში (განსხვავებით განხილული შემთხვევისაგან, სადაც $\xi_i - \mu$ $(i = 1, 2, ..., n)$ დამოუკიდებელი შემთხვევითი სიდიდეებია), ns^2/σ^2 აღარ წარმოადგენს დამოუკიდებელი შემთხვევითი სიდიდეების კვადრატების ჯამს, $\xi_i - \overline{x}$ შემთხვევით სიდიდეებს გააჩნიათ ერთი „ბმა“. კერძოდ, ვინაიდან $\overline{x} = \frac{1}{n}\sum_{i=1}^{n}\xi_i$, ამიტომ $\sum_{i=1}^{n}(\xi_i - \overline{x}) = 0$. ამ შემთხვევაში ადგილი აქვს თანაფარდობას

$$\frac{ns^2}{\sigma^2} = \frac{(n-1)s'^2}{\sigma^2} = \chi^2(n-1).$$

გარდა ამისა, მტკიცდება რომ $\overline{x}$ და s^2 დამოუკიდებელი შემთხვევითი სიდიდეებია.

სტატისტიკაში ხშირად გამოიყენება ე. წ. Z სტატისტიკა, და T სტატისტიკა:

$$Z = \frac{\overline{x} - \mu}{\sigma/\sqrt{n}} \qquad \text{და вф} T = \frac{\overline{x} - \mu}{s/\sqrt{n-1}} = \frac{\overline{x} - \mu}{s'/\sqrt{n}}.$$

ცხადია, რომ $Z = N(x;0;1)$. რაც შეეხება T სტატისტიკას, ის გადავწეროთ შემდეგი სახით:

$$T = \frac{Z}{\sqrt{\frac{ns^2}{\sigma^2}/(n-1)}} = \frac{Z}{\sqrt{\frac{(n-1)s'^2}{\sigma^2}/(n-1)}}.$$

ვინაიდან, $\frac{(n-1)s'^2}{\sigma^2} = \chi^2(n-1)$ და Z და $\frac{(n-1)s'^2}{\sigma^2}$ შემთხვევითი სიდიდეები დამოუკიდებელია, ამიტომ T სტატისტიკას აქვს სტიუდენტის განაწილება თავისუფლების ხარისხით $n-1$.

§31.შეფასებათა აგების მეთოდები

მაქსიმალური დასაჯერობის მეთოდი.

დავუშვათ, რომ X – დისკრეტული შემთხვევითი სიდიდეა, რომელმაც ექსპერიმენტის შედეგად მიიღო მნიშვნელობები $x_1, x_2, \ldots, x_n$. დავუშვათ, რომ ჩვენთვის ცნობილია ამ შემთხვევითი სიდიდის განაწილების კანონი, რომელიც განისაზღვრება Θ პარამეტრით, მაგრამ უცნობია ამ პარამეტრის რიცხვითი მნიშვნელობა. ჩვენი მიზანია ვიპოვოთ ამ პარამეტრის წერტილოვანი შეფასება.

ვთქვათ, $p(x_i, \Theta)$ -- არის ალბათობა იმისა, რომ ექსპერიმენტის შედეგად X შემთხვევითი სიდიდე მიიღებს x_i მნიშვნელობას. დისკრეტული X შემთხვევითი სიდიდის **მაქსიმალური დასაჯერობის ფუნქცია** ეწოდება Θ არგუმენტის ფუნქციას, რომელიც განისაზღვრება ფორმულით:

$$L\,(x_1, x_2, \ldots, x_n; \Theta) = p(x_1,\Theta)p(x_2,\Theta)\ldots p(x_n,\Theta).$$

Θ პარამეტრის წერტილოვანი შეფასების როლში იღებენ მის ისეთ $\Theta^* = \Theta(x_1, x_2, \ldots, x_n)$ მნიშვნელობას, რომლის დროსაც მაქსიმალური დასაჯერობის ფუნქცია აღწევს თავის მაქსიმუმს. Θ^* შეფასებას **მაქსიმალური დასაჯერობის შეფასებას** უწოდებენ.

რადგანაც ფუნქციები L და $\ln L$ მაქსიმუმს აღწევენ Θ-ს ერთი და იგივე მნიშვნელობისათვის, უფრო მოხერხებულია მოვძებნოთ $\ln L$ ფუნქციის მაქსიმუმი (ვინაიდან ნამრავლის ლოგარითმი ლოგარითმების ჯამია და ამდენად კრიტიკული წერტილების პოვნისას ნამრავლის გაწარმოების ნაცვლად მოგვიწევს ჯამის გაწარმოება, რაც გაცილებით მარტივია). ამ ფუნქციას **მაქსიმალური დასაჯერობის ლოგარითმული ფუნქცია** ეწოდება.

$\ln L$ ფუნქციის მაქსიმუმის მიმნიჭებელი წერტილის მოსაძებნად საჭიროა შემდეგი პროცედურების ჩატარება:

1). ვიპოვოთ წარმოებული $\dfrac{d\ln L}{d\Theta}$;

2). გავუტოლოთ წარმოებული ნულს (მივიღებთ ე. წ. **მაქსიმალური დასაჯერობის განტოლებას**) და ვიპოვოთ კრიტიკული წერტილები’

3). ვიპოვოთ მეორე წარმოებული $\dfrac{d^2\ln L}{d\Theta^2}$; თუ ის უარყოფითია კრიტიკულ წერტილში, მაშინ ეს წერტილი – მაქსიმუმის წერტილია.

აღსანიშნავია, რომ მაქსიმალური დასაჯერობის მეთოდით მიღებული შეფასებები ძალმოსილია (თუმცა შესაძლებელია არ იყოს ჩაუნაცვლებელი), განაწილებული არიან ასიმპტოტურად ნორმალურად შერჩევის დიდი მოცულობის შემთხვევაში და გააჩნიათ უმცირესი დისპერსია სხვა ასიმპტოტურად ნორმალურ შეფასებებთან შედარებით. თუ შესაფასებელი Θ პარამეტრისათვის არსებობს ეფექტური Θ^* შეფასება, მაშინ მაქსიმალური დასაჯერობის განტოლებას გააჩნია ერთადერთი ამონახსნი Θ^*. ეს მეთოდი ყველაზე სრულად იყენებს შერჩევის მონაცემებს და ამიტომ განსაკუთრებით სასარგებლოა მცირე შერჩევების დროს.

მაქსიმალური დასაჯერობის მეთოდის ნაკლად შეიძლება ჩაითვალოს გამოთვლების სირთულე.

უწყვეტი შემთხვევითი სიდიდის შემთხვევაში, რომლის $f(x)$ განაწილების სიმკვრივის სახე ცნობილია, მაგრამ იგი შეიცავს უცნობ Θ პარამეტრს, მაქსიმალური დასაჯერობის ფუნქციას აქვს შემდეგი სახე:

$$L(x_1, x_2, ..., x_n; \Theta) = f(x_1, \Theta) \cdot f(x_2, \Theta) \cdots f(x_n, \Theta).$$

უცნობი პარამეტრის მაქსიმალური დასაჯერობის შეფასების საპოვნელად უნდა ჩავატაროთ იგივე პროცედურები, რაც დისკრეტულ შემთხვევაში.

მომენტთა მეთოდი.

მომენტთა მეთოდი დაფუძნებულია იმ გარემოებაზე, რომ საწყისი და ცენტრალური ემპირიული მომენტები წარმოადგენენ შესაბამისი საწყისი და ცენტრალური თეორიული მომენტების ძალმოსილ შეფასებებს. ამიტომ ჩვენ შეგვიძლია თეორიული მომენტები გავუტოლოთ იმავე რიგის შესაბამის ემპირიულ მომენტებს. თუ მოცემულია განაწილების $f(x, \Theta)$ სიმკვრივის სახე, რომელიც განისაზღვრება ერთი უცნობი Θ პარამეტრით (დამოკიდებულია ერთ უცნობ პარამეტრზე), მაშინ ამ პარამეტრის შესაფასებლად საკმარისია გვქონდეს ერთი განტოლება. მაგალითად, შეგვიძლია გავუტოლოთ ერთმანეთს პირველი რიგის საწყისი მომენტები:

$$\bar{x}_B = E(X) = \int_{-\infty}^{\infty} x f(x; \Theta) dx = \varphi(\Theta),$$

და მივიღებთ განტოლებას Θ პარამეტრის საპოვნელად. მისი ამონახსნი Θ* იქნება Θ პარამეტრის წერტილოვანი შეფასება, რომელიც წარმოადგენს შერჩევითი საშუალოს ფუნქციას და, შესაბამისად, შერჩევის ფუნქციას:

$$\Theta = \psi\,(x_1, x_2, \ldots, x_n).$$

თუ განაწილების სიმკვრივე განისაზღვრება ორი Θ_1 და Θ_2 პარამეტრით (დამოკიდებულია ორ პარამეტრზე), მაშინ მოითხოვება შევადგინოთ ორი განტოლება, მაგალითად, $\nu_1 = M_1$, $\mu_2 = m_2$.

აქედან ვღებულობთ ორი განტოლებისაგან შემდგარ სისტემას ორი Θ_1 და Θ_2 უცნობით: $\begin{cases} E(X) = \bar{x}_B \\ D(X) = D_B \end{cases}$.

მისი ამოხსნები Θ_1^* და Θ_2^* იქნებიან Θ_1 და Θ_2 პარამეტრების წერტილოვანი შეფასებები დამოკიდებული შერჩევაზე:

$$\Theta_1 = \psi_1\,(x_1, x_2, \ldots, x_n),$$
$$\Theta_2 = \psi_2(x_1, x_2, \ldots, x_n).$$

§32. ინტერვალური შეფასებები. ნდობის ინტერვალი მათემატიკური ლოდინისათვის

გენერალური ერთობლიობის პარამეტრების წერტილოვანი შეფასებები შეიძლება მიღებულ იქნეს შერჩევითი მონაცემების დამუშავების საორიენტაციო, პირველად შედეგებად. მათი ნაკლი იმაში მდგომარეობს, რომ უცნობია რა სიზუსტით ფასდება პარამეტრი. დიდი მოცულობის შერჩევებისათვის სიზუსტე როგორც წესი საკმარისია (შეფასებების გადაუადგილებადობის, ძალმოსილებისა და ეფექტურობის პირობებში), მაშინ როდესაც მცირე მოცულობის შერჩევებისათვის შეფასების სიზუსტის საკითხი ძალიან მნიშვნელოვანია.

შემოვიღოთ გენერალური ერთობლიობის (ან ξ შემთხვევითი სიდიდის, რომელიც განმარტებულია ამ გენერალური ერთობლიობის ობიექტების სიმრავლეზე) უცნობი პარამეტრის ინტერვალური შეფასების ცნება. ავღნიშნოთ ეს პარამეტრი Δ-თი. მოცემული შერჩევიდან გარკვეული წესით იძებნება ისეთი რიცხვები Δ_1 და Δ_2, რომ სრულდებოდეს პირობა:

$$P(\Delta_1 < \Delta < \Delta_2) = P(\Delta \in (\Delta_1; \Delta_2)) = \gamma.$$

Δ_1 და Δ_2 რიცხვებს უწოდებენ **ნდობის საზღვრებს,** ხოლო (Δ_1, Δ_2) ინტერვალს -- Δ პარამეტრის **ნდობის ინტერვალს.** γ რიცხვს ეწოდება **ნდობის ალბათობა** ან გაკეთებული შეფასების **საიმედოობა.**

თავიდან მოიცემა საიმედოობა. ჩვეულებრივ, მას ირჩევენ 0.95-ის, 0.99 -ის ან 0.999-ის ტოლს. მაშინ ალბათობა იმისა, რომ ჩვენთვის საინტერესო პარამეტრი მოხვდა (Δ_1, Δ_2) ინტერვალში საკმარისად მაღალია. რიცხვი ($\Delta_1 + \Delta_2$)/2 – ნდობის ინტერვალის შუაწერტილი – იძლევა Δ პარამეტრის მნიშვნელობას $(\Delta_2 - \Delta_1)/2$-ს ტოლი სიზუსტით, რომელიც წარმოადგენს ნდო-ბის ინტერვალის სიგრძის ნახევარს.

საზღვრები Δ_1 და Δ_2 განისაზღვრება შერჩევითი მონაცემებიდან და წარმოადგენენ $x_1, x_2, ..., x_n$ შემთხვევითი სიდიდეების ფუნქციებს. შესაბამისად, საზღვრები თვითონაც შემთხვევითი სიდიდეებია. აქედან გამომდინარე, ნდობის ინტერვალი (Δ_1, Δ_2) -- აგრეთვე შემთხვევითია. ის შეიძლება ფარავდეს ან არ ფარავდეს Δ პარამეტრს. სწორედ ასეთი აზრით უნდა გავიგოთ შემთხვევითი ხდომილება, რომელიც მდგომარეობს იმაში, რომ ნდობის ინტერვალი ფარავს Δ რიცხვს.

ნდობის ინტერვალი ნორმალური განაწილების მათემატიკური ლოდინისათვის ცნობილი დისპერსიის შემთხვევაში:

დავუშვათ, რომ შემთხვევითი სიდიდე ξ (შეიძლება ვილაპარაკოთ გენერალურ ერთობლიობაზე) განაწილებულია ნორმალური განაწილების კანონის მიხედვით, რომლის დისპერსია ცნობილია $D\xi = \sigma^2$ $(\sigma > 0)$. გენერალური ერთობლიობიდან (რომლის ობიექტების სიმრავლეზე განმარტებულია შემთხვევითი სიდიდე) კეთდება n მოცულობის შერჩევა. შერჩევა

$x_1, x_2, ..., x_n$ განიხილება როგორც ერთობლიობა n დამოუკიდებელი შემთხვევითი სიდიდის, რომლებიც იგივე კანონით არიან განაწილებული როგორც ξ. ამ შემთხვევაში, როგორც ჩვენ უკვე ვნახეთ:

$$Ex_1 = Ex_2 = \cdots = Ex_n = E\xi; \quad Dx_1 = Dx_2 = \cdots = Dx_n = D\xi;$$

$$E\bar{x} = E\xi; \quad D\bar{x} = D\xi / n.$$

ცნობილია, რომ მოცემულ შემთხვევაში შემთხვევითი სიდიდე $\bar{x}$ აგრეთვე განაწილებულია ნორმალური განაწილების კანონით. ავღნიშნოთ უცნობი მათემატიკური ლოდინი a-თი, $E\xi := a$ და მოცემული γ საიმედოობისათვის შევარჩიოთ $d > 0$ რიცხვი ისე, რომ შესრულდეს პირობა:

$$P(|\bar{x} - a| < d) = \gamma \qquad (1)$$

ვინაიდან შემთხვევითი სიდიდე $\bar{x}$ განაწილებულია ნორმალურად მათემატიკური ლოდინით $E\bar{x} = E\xi = a$ და დისპერსიით $D\bar{x} = D\xi / n = \sigma^2/n$, ამიტომ გვაქვს:

$$P(|\bar{x} - a| < d) = P(a - d < \bar{x} < a + d) =$$

$$= \Phi_0\big((a + d - a)\sqrt{n}/\sigma\big) - \Phi_0\big((a - d - a)\sqrt{n}/\sigma\big) = 2\Phi_0\big(d\sqrt{n}/\sigma\big).$$

ახლა შევარჩიოთ $d > 0$ ისე, რომ შესრულდეს ტოლობა

$$2\Phi_0\big(d\sqrt{n}/\sigma\big) = \gamma \text{ ანუ } \Phi_0\big(d\sqrt{n}/\sigma\big) = \gamma/2.$$

ნებისმიერი $\gamma \in [0;1]$ რიცხვისათვის ნორმალური განაწილების ფუნქციის ცხრილიდან შეგვიძლია ვიპოვოთ ისეთი t რიცხვი, რომ $\Phi_0(t) = \gamma/2$. ამ t რიცხვს $\gamma/2$-**კვანტილი** ეწოდება (მას აღნიშნავენ აგრეთვე $x_{\gamma/2}$ სიმბოლოთი).

ტოლობიდან $d\sqrt{n}/\sigma = t$ ვპოულობთ d-ს მნიშვნელობას: $d = \sigma t/\sqrt{n}$. საბოლოო შედეგს მივიღებთ, თუ (1) ფორმულას წარმოვადგენთ შემდეგი სახით:

$$P\big(\bar{x} - \sigma t/\sqrt{n} < a < \bar{x} + \sigma t/\sqrt{n}\big) = \gamma.$$

უკანასკნელი ფორმულის აზრი მდგომარეობს შემდეგში: საიმედოობით γ ნდობის ინტერვალი

$$\big(\bar{x} - \sigma t/\sqrt{n}; \bar{x} + \sigma t/\sqrt{n}\big)$$

ფარავს (მოიცავს) გენერალური ერთობლიობის უცნობ პარამეტრს $a = E\xi$-ს. შეიძლება ითქვას სხვანაირად: წერტილოვანი შეფასება $\bar{x}$ განსაზღვრავს $E\xi$ პარამეტრის მნიშვნელობას $d = \sigma t/\sqrt{n}$ სიზუსტითა და γ საიმედოობით.

ამოცანა. დავუშვათ გვაქვს გენერალური ერთობლიობა გარკვეული მახასიათებლით, რომელიც განაწილებულია ნორმალური კანონით, რომლის დისპერსია ტოლია 6.25-ის. ჩატარებულია $n = 27$ მოცულობის შერჩევა და მიღებულია მახასიათებლის საშუალო შერჩევითი მნიშვნელობა $\bar{x} = 12$. ვიპოვოთ ნდობის ინტერვალი, რომელიც ფარავს გენერალური ერთობლი-

ობის გამოსაკვლევი მახასიათებლის უცნობ მათემატიკურ ლოდინს საიმე-დოობით $\gamma = 0.99$.

ამოხსნა. პირველ რიგში, ლაპლასის ფუნქციის ცხრილებიდან ვიპოვ-ოთ t-ს მნიშვნელობა ტოლობიდან $\Phi_0(t) = \gamma / 2 = 0.495$. მიღებული $t = 2.58$ მნიშვნელობიდან განვსაზღვროთ შეფასების სიზუსტე (ანუ ნდობის ინტერვალის სიგრძის ნახევარი) d: $d = 2.5 \times 2.58 / \sqrt{27} \approx 1.24$. აქედან ვღებულობთ საძებნ ნდობის ინტერვალს: (10.76, 13.24).

ნდობის დონის სიზუსტე და შერჩევის მოცულობის მოძებნა:

მოცემული γ ნდობის ალბათობისათვის დავადგინოთ შერჩევის ის მინიმალური n^* მოცულობა, რომელიც უზრუნველყოფს შეფასების წინას-წარ ფიქსირებულ სიზუსტეს (შეფასება მით უფრო ზუსტია, რაც უფრო ნაკლებია ნდობის ინტერვალის სიგრძე). ცხადია, რომ რაც უფრო დიდია γ, მით უფრო დიდია $x_{\gamma/2}$, და შესაბამისად, განიერია ნდობის ინტერვალი და პირიქით. აქედან გამომდინარე, თუ შერჩევის მოცულობა ფიქსირებუ-ლია, ნდობის ინტერვალის სიგრძის (ანუ შეფასების სიზუსტის) შემცირება შესაძლებელია მხოლოდ ნდობის ალბათობის შემცირების ხარჯზე. ფიქსი-რებული ნდობის ალბათობის დროს ინტერვალის სიგრძე მით უფრო მცი-რეა, რაც უფრო დიდია შერჩევის მოცულობა. ყოველივე ზემოთ თქმული-დან ვასკვნით, რომ შეფასების ფიქსირებული სიზუსტე ნიშნავს ნდობის ინტერვალის ფიქსირებულ l სიგრძეს.

ვინაიდან, რომ γ ნდობის ინტერვალის სიგრძეა

$$2 \frac{\sigma}{\sqrt{n}} x_{\gamma/2},$$

ამიტომ შერჩევის n^* მოცულობა უნდა შეირჩეს, როგორც შემდეგი განტო-ლების ამონახსნი

$$2 \frac{\sigma}{\sqrt{n}} x_{\gamma/2} = l,$$

ე. ი.

$$n^* = (\frac{2\sigma}{l} x_{\gamma/2})^2.$$

ვინაიდან ასეთნაირად მოძებნილი n^* შეიძლება არ იყოს მთელი რი-ცხვი, ამიტომ n^*-ის როლში იღებენ მიღებული სიდიდის მთელ ნაწილს მი-მატებულ ერთს:

$$n^* = [(\frac{2\sigma}{l} x_{\gamma/2})^2] + 1.$$

ნდობის ინტერვალი ნორმალური განაწილების მათემატიკური ლოდინისათვის უცნობი დისპერსიის შემთხვევაში:

დავუშვათ, რომ ξ -- ნორმალური კანონით განაწილებული შემთხვევითი სიდიდეა უცნობი მათემატიკური ლოდინით $E\xi$, რომელიც ავღნიშნოთ a ასოთი. ჩავატაროთ n მოცულობის შერჩევა. განვსაზღვროთ შერჩევითი საშუალო $\bar{x}$ და შესწორებული შერჩევითი დისპერსია s^2 ზემოთ მოყვანილი ფორმულების მიხედვით.

ცნობილია, რომ შემთხვევითი სიდიდე

$$t=\left(\bar{x}-a\right)\sqrt{n}/s$$

განაწილებულია სტიუდენტის კანონის მიხედვით თავისუფლების $n-1$ ხარისხით. ამოცანა მდგომარეობს იმაში, რომ მოცემული γ საიმედოობისა და თავისუფლების $n-1$ ხარისხის მიხედვით, ვიპოვოთ ისეთი t_γ რიცხვი, რომ შესრულდეს ტოლობა:

$$P\left(\left|\left(\bar{x}-a\right)\sqrt{n}/s\right|<t_\gamma\right)=\gamma, \qquad (2)$$

ან მისი ექვივალენტური ტოლობა

$$P\left(\bar{x}-t_\gamma s/\sqrt{n}<a<\bar{x}+t_\gamma s/\sqrt{n}\right)=\gamma. \qquad (3)$$

აქ ფრჩხილებში წერია იმის პირობა, რომ უცნობი a პარამეტრის მნიშვნელობა ეკუთვნის გარკვეულ შუალედს, რომელიც არის სწორედ ნდობის ინტერვალი. მისი საზღვრები დამოკიდებულია γ საიმედოობაზე და აგრეთვე, შერჩევის $\bar{x}$ და s პარამეტრებზე.

იმისთვის, რომ γ სიდიდის მიხედვით ვიპოვოთ t_γ-ს მნიშვნელობა, (2) ტოლობა გადავწეროთ შემდეგი სახით:

$$P\left(\left|\left(\bar{x}-a\right)\sqrt{n}/s\right|\geq t_\gamma\right)=1-\gamma\,.$$

ახლა, t შემთხვევითი სიდიდის ცხრილის მიხედვით, რომელიც განაწილებულია სტიუდენტის კანონით, ალბათობით $1-\gamma$ და თავისუფლების $n-1$ ხარისხით, ვპოულობთ $t_\gamma=t_{n-1,(1-\gamma)/2}$-ს.

ამოცანა. 20 ელექტრონათურის საკონტროლო შემოწმებისას მათი მუშაობის საშუალო ხანგრძლივობა აღმოჩნდა 2000 საათის ტოლი, ხოლო საშუალო კვადრატული გადახრა (გამოთვლილი როგორც კვადრატული ფესვი შესწორებული შემთხვევითი დისპერსიიდან) კი 11 საათის ტოლი. ცნობილია, რომ ნათურის მუშაობის ხანგრძლივობა წარმოადგენს ნორმალური კანონით განაწილებულ შემთხვევით სიდიდეს. განვსაზღვროთ ამ შემთხვევითი სიდიდის მათემატიკური ლოდინის ნდობის ინტერვალი საიმედოობით 0.95.

ამოხსნა. ამ შემთხვევაში $1-\gamma=0.05$. სტიუდენტის განაწილების ცხრილიდან (თავისუფლების 19-ის ტოლი ხარისხით) ვპოულობთ, რომ $t_\gamma=2.093$. გამოვთვალოთ შეფასების სიზუსტე: $2.093\times 11/\sqrt{20}=5.2$. ამიტომ ნდობის ინტერვალი იქნება (1994.8, 2005.2).

§33. ნდობის ინტერვალი დისპერსიისათვის და სტანდარტული გადახრისათვის

ნდობის ინტერვალი ნორმალური განაწილების დისპერსიისათვის:

დავუშვათ, რომ ξ შემთხვევითი სიდიდე განაწილებულია ნორმალური განაწილების კანონით, რომლის დისპერსია $D\xi$ უცნობია. კეთდება n მოცულობის შერჩევა. მისი საშუალებით განისაზღვრება შესწორებული შერჩევითი დისპერსია s^2. ცნობილია რომ, შემთხვევითი სიდიდე

$$\chi^2 = (n-1)s^2 / D\xi$$

განაწილებულია χ^2 განაწილების კანონით თავისუფლების ხარისხით $n-1$. მოცემული γ საიმედოობისათვის შეიძლება ვიპოვოთ ინტერვალების ისეთი საზღვრები χ_1^2 და χ_2^2, რომ

$$P\left(\chi_1^2 < \chi^2 < \chi_2^2\right) = \gamma \qquad (1)$$

ვიპოვოთ χ_1^2 და χ_2^2 შემდეგი პირობებიდან:

$$P(\chi^2 \le \chi_1^2) = (1-\gamma)/2 \qquad (2)$$

$$P(\chi^2 \ge \chi_2^2) = (1-\gamma)/2 \qquad (3)$$

ნათელია, რომ ამ ორი უკანასკნელი პირობის შესრულებისას, სამართლიანი იქნება (1) ტოლობა.

χ^2 შემთხვევითი სიდიდის ცხრილებში, ჩვეულებრივ, მოიცემა $P(\chi^2 \ge \chi_q^2) = q$ განტოლების ამონახსნი. ასეთი ცხრილიდან, მოცემული q სიდიდისა თავისუფლების $n-1$ ხარისხის მიხედვით შეგვიძლია განვსაზღვროთ χ_q^2. ამრიგად, ჩვენ ვპოულობთ χ_q^2ის მნიშვნელობას (3) ფორმულაში.

χ_1^2ის საპოვნელად გადავწეროთ (2) შემდეგი ფორმით:

$$P(\chi^2 \ge \chi_1^2) = 1-(1-\gamma)/2 = (1+\gamma)/2.$$

მიღებული ტოლობა საშუალებას გვაძლევს ცხრილებიდან დავადგინოთ χ_1^2.

მას შემდეგ რაც ნაპოვნია χ_1^2 –ისა და χ_2^2 ის მნიშვნელობები, გადავწეროთ (1) ტოლება შემდეგი სახით:

$$P\left(\chi_1^2 < (n-1)s^2 / D\xi < \chi_2^2\right) = \gamma.$$

უკანასკნელი ტოლობა გადავწეროთ ისეთი ფორმით, რომ განსაზღრული იყოს უცნობი $D\xi$ პარამეტრის ნდობის ინტერვალის საზღვრები:

$$P\left((n-1)s^2/\chi_2^2 < D\xi < (n-1)s^2/\chi_1^2\right) = \gamma.$$

აქედან ადვილად მივიღებთ ფორმულას, რომლის მიხედვითაც განისაზღვრება სტანდარტული გადახრის ნდობის ინტერვალი:

$$P\left(\sqrt{(n-1)}s/\sqrt{\chi_2{}^2}<\sqrt{D\xi}<\sqrt{(n-1)}s/\sqrt{\chi_1{}^2}\right)=\gamma \qquad (4)$$

ამოცანა. ჩავთვალოთ, რომ ხმაური, ერთი და იგივე ტიპის ვერტმფრენის კაბინაში, გარკვეულ რეჟიმში მომუშავე ძრავის დროს, შემთხვევითი სიდიდეა, რომელიც განაწილებულია ნორმალური კანონით. შემთხვევით შერჩეულ იქნა 20 ვერტმფრენი და მოხდა მათში ხმის დონის გაზომვა (დეციბალებში). გაზომვების შესწორებული შერჩევითი დისპერსია აღმოჩნდა 22.5ის ტოლი. ვიპოვოთ ნდობის ინტერვალი, რომელიც ფარავს მოცემული ტიპის ვერტმფრენების კაბინაში ხმაურის სიდიდის უცნობ სტანდარტულ გადახრას 98%იანი საიმედოობით.

ამოხსნა. თავისუფლების 19ის ტოლი ხარისხითა და (1 – 0.98)/2 = 0,01 ალბათობის საშუალებით χ^2 –ის განაწილების ცხრილიდან ვპოულობთ სიდიდეს: $\chi_2{}^2 = 36.2$. ანალოგიურად, (1 + 0.98)/2 = 0.99 ალბათობის საშუალებით ვპოულობთ: $\chi_1{}^2 = 7.63$. შესაბამისად, (4) ფორმულის გამოყენებით ვღებულობთ, რომ საძიებელი ნდობის ინტერვალია: (3.44, 7.49).

ნდობის ინტერვალი საშუალო კვადრატული გადახრისათვის:

ავაგოთ $(s-\delta, s+\delta)$ სახის ნდობის ინტერვალი ნორმალურად განაწილებული შემთხვევითი სიდიდის საშუალო კვადრატული გადახრისათვის, სადაც s – შესწორებული შერჩევითი საშუალო კვადრატული გადახრაა, ხოლო δ –სათვის სრულდება პირობა: $P(\ |\sigma - s| < \delta\ ) = \gamma$. გადავწეროთ ეს უტოლობა შემეგი სახით

$$s\left(1-\frac{\delta}{s}\right)<\sigma<s\left(1+\frac{\delta}{s}\right),$$

ან თუ ავღნიშნავთ $q=\delta/s$, მაშინ გვექნება:

$$s(1-q)<\sigma<s(1+q) \qquad (1)$$

განვიხილოთ შემთხვევითი სიდიდე χ, რომელიც განისაზღვრება ფორმულით

$$\chi=\frac{s}{\sigma}\sqrt{n-1}\,.$$

როგორც ცნობილია, მას აქვს ხი კვადრატ განაწილება თავისუფლების ხარისხით $n1$. მისი განაწილების სიმკვრივე:

$$R(\chi,n)=\frac{\chi^{n-2}e^{-\frac{\chi^2}{2}}}{2^{\frac{n-3}{2}}\Gamma\left(\frac{n-1}{2}\right)}$$

არაა დამოკიდებული შესაფასებელ σ პარამეტრზე, და შესაბამისად, დამოკიდებულია მხოლოდ შერჩევის n მოცულობაზე. გარდავქმნათ (1) უტოლობა ისე, რომ მან მიიღოს სახე $\chi_1 < \chi < \chi_2$. ამ უტოლობის შესრულების ალბათობა ტოლია ნდობის γ ალბათობის, შესაბამისად,

$$\int_{\chi_1}^{\chi_2} R(\chi,n)d\chi = \gamma.$$

დავუშვათ, რომ $q < 1$, მაშინ (1) უტოლობა შეიძლება ჩაიწეროს შემდეგნაირად:

$$\frac{1}{s(1+q)} < \frac{1}{\sigma} < \frac{1}{s(1-q)},$$

ან, $s\sqrt{n-1}$ მამრავლზე გამრავლების შემდეგ, გვექნება:

$$\frac{\sqrt{n-1}}{1+q} < \frac{s\sqrt{n-1}}{\sigma} < \frac{\sqrt{n-1}}{1-q}.$$

შესაბამისად,

$$\frac{\sqrt{n-1}}{1+q} < \chi < \frac{\sqrt{n-1}}{1-q}.$$

საბოლოოდ, ვღებულობთ თანაფარდობას:

$$\int_{\sqrt{n-1}/(1+q)}^{\sqrt{n-1}/(1-q)} R(\chi,n)d\chi = \gamma.$$

არსებობს ხი კვადრატ განაწილების ცხრილები, რომელიც საშუალ ებას იძლევა უკანასკნელი განტოლების ამოხსნის გარეშე, მოცემული n და γ–სათვის ვიპოვოთ q. ამრიგად, თუ გამოვითვლით შერჩევის მიხედვით s–ის მნიშვნელობას და ცხრილიდან ვიპოვით q–ს, ჩვენ ავაგებთ (1) ნდობის ინტერვალს, რომელშიც σ–ს მნიშვნელობა მოხვდება γ–ს ტოლი ალბათობით.

შენიშვნა. თუ $q > 1$, მაშინ $\sigma > 0$ პირობის გათვალისწინებით, ნდობის ინტერვალს σ–სათვის ექნება სახე:

$$0 < \sigma < s(1+q).$$

მაგალითი. ვთქვათ, $n = 20$, $s = 1.3$. ვიპოვოთ ნდობის ინტერვალი σ–სათვის მოცემული $\gamma = 0.95$–ის ტოლი ნდობის ალბათობისათვის.

შესაბამისი ცხრილიდან ვპოულობთ, რომ q (n = 20, γ = 0.95) = 0.37. შესაბამისად, ნდობის ინტერვალის საზღვრები იქნება: 1.3(10.37) = 0.819 და 1.3(1+0.37) = =1.781. მასასადამე, $0.819 < \sigma < 1.781$ ალბათობით 0.95.

§34. ნდობის ინტერვალი ბერნულის სქემაში

ბერნულის სქემაში (დამუკიდებელ ცდათა სქემაში) უცნობი p ალბათობის წერტილოვანი შეფასებაა ფარდობითი სიხშირე:

$$w_n = \frac{S_n}{n},$$

სადაც $S_n = \sum_{i=1}^{n} X_i$ ($X_i = 1$, თუ i-ურ ცდაში მოხდა წარმატება, და $X_i = 0$, თუ i-ურ ცდაში მოხდა მარცხი) – წარმატებათა რაოდენობაა n დამოუკიდებელ ცდაში, ამასთან

$$Ew_n = p \quad \text{და} \quad Dw_n = \frac{p(1-p)}{n}.$$

უცნობი p ალბათობისათვის ნდობის ინტერვალის ასაგებად იყენებენ ფარდობითი სიხშირის სტანდარტიზაციის შედეგად მიღებულ სტატისტიკას:

$$T_n = \frac{w_n - p}{\sqrt{\frac{p(1-p)}{n}}},$$

რომელიც, ცენტრალური ზღვარითი თეორემის თანახმად, დაახლოებით ნორმალურადაა განაწილებული ნულოვანი საშუალოთო და ერთეულოვანი დისპერსიით, თუ შერჩევის მოცულობა n საკმაოდ დიდია. მაგრამ, სამწუხაროდ, ამ სტატისტიკის გამოსახულების მნიშვნელშიც შედის შესაფასებელი p პარამეტრი, რაც საშუაულებას არ იძლევა სტანდარტული გზით მივიღოთ ნდობის ინტერვალი.

არსებობს ასეთი გამოსავალი. შეიძლება გამოვიყენოთ გამარტივებული მიდგომა, რომლის თანახმადაც მნიშვნელში მდგომი უცნობი p ალბათობა უნდა შევცვალოთ მისი w_n შეფასებით და შესაბამისად, T_n სტატისტიკის ნაცვლად გამოვიყენოთ შემდეგი სტატისტიკა:

$$\hat{T}_n = \frac{w_n - p}{\sqrt{w_n \cdot (1-w_n)}} \cdot \sqrt{n}.$$

გასაგებია, რომ ამ სტატისტიკას ასიმპტოტურად ექნება იგივე ყოფაქცევა რაც T_n სტატისტიკას. ამის შემდეგ ნდობის ინტერვალი იგება სტანდარტული გზით, რის შედეგადაც ვღებულობთ, რომ შერჩევის დიდი მოცულობის შემთხვევაში $(1-\alpha)$ ნდობის ალბათობის მქონე ნდობის ინტერვალს უცნობი p ალბათობისათვის აქვს შემდეგი სახე:

$$\left(w_n - z_{\alpha/2}\sqrt{\frac{w_n(1-w_n)}{n}}; w_n + z_{\alpha/2}\sqrt{\frac{w_n(1-w_n)}{n}}\right), \tag{1}$$

სადაც $z_{\alpha/2}$ -- სტანდარტული ნორმალური განაწილების ზედა $\alpha/2$ კრიტიკული წერტილია (სტანდარტული ნორმალური განაწილების ზედა α კრიტიკული წერტილი ეწოდება ისეთ z_α რიცხვს, რომლისთვისაც

$$P\{N(x;0;1) > z_\alpha\} = \alpha \quad \text{ანუ} \quad \Phi(z_\alpha) = 1-\alpha).$$

მეორე მიდგომა ეყრდნობა აგრეთვე ნორმალურ აპროქსიმაციას: მოვძებნოთ ისეთი p რიცხვი, რომ სრულდებოდეს უტოლობა

$$P\{-z_{\alpha/2} \le \frac{w_n - p}{\sqrt{\frac{p(1-p)}{n}}} \le -z_{\alpha/2}\} \approx 1-\alpha .$$

რაც ტოლფასია იმისა, რომ ამოვხსნათ p ცვლადის მიმართ განტოლება

$$\frac{w_n - p}{\sqrt{\frac{p(1-p)}{n}}} = z_{\alpha/2} .$$

საბოლოოდ, ამ გზით მიიღებული დაზუსტებული $(1-\alpha)$ ნდობის ალბათობის მქონე ნდობის ინტერვალი უცნობი p ალბათობისათვის იქნება:

$$\left(\frac{w_n + \frac{z_{\alpha/2}^2}{2n} - z_{\alpha/2}\sqrt{\frac{w_n(1-w_n)}{n} + \frac{z_{\alpha/2}^2}{4n^2}}}{1+\frac{z_{\alpha/2}^2}{n}}, \frac{w_n + \frac{z_{\alpha/2}^2}{2n} + z_{\alpha/2}\sqrt{\frac{w_n(1-w_n)}{n} + \frac{z_{\alpha/2}^2}{4n^2}}}{1+\frac{z_{\alpha/2}^2}{n}}\right) .$$

ცხადია, რომ თუ n იმდენად დიდია, რომ $\frac{z_{\alpha/2}^2}{n}$-ისა და $\frac{z_{\alpha/2}^2}{n^2}$-ის უგულებელყოფა (ნულთან გატოლება) შეიძლება, მაშინ უკანასკნელი ინტერვალი დაემთხვევა (1) ინტერვალს.

§35. ჰიპოთეზათა სტატისტიკური შემოწმების ამოცანები

ჰიპოთეზების სტატისტიკური შემოწმება წარმოადგენს მათემატიკური სტატისტიკის უმნიშვნელოვანესს ნაწილს. მათემატიკური სტატისტიკის მეთოდები საშუალებას იძლევა შევამოწმოთ დაშვებები გარკვეული შემთხვევითი სიდიდის (გენერალური ერთობლიობის) განაწილების კანონის შესახებ, ამ კანონის პარამეტრების (მაგალითად, $E\xi$, $D\xi$) მნიშვნელობების შესახებ, ერთი და იგივე გენერალური ერთობლიობის ობიექტების სიმრავლეზე განმარტებულ შემთხვევით სიდიდეებს შორის კორელაციური კავშირის არსებობის შესახებ.

დავუშვათ, რომ გარკვეული მონაცემების მიხედვით, გვაქვს საფუძველი წამოვაყენოთ წინადადება განაწილების კანონის შესახებ ან შემთხვევითი სიდიდის (ან გენერალური ერთობლიობის, რომელთა ობიექტების სიმრავლეზე განმარტებულია მოცემული შემთხვევითი სიდიდე) განაწილების კანონის პარამეტრის შესახებ. ამოცანა მდგომარეობს იმაში, რომ დავადასტუროთ ან უარვყოთ ეს წინადადება შერჩევითი (ექსპერიმენტალური) მონაცენების გამოყენების საფუძველზე.

განაწილების პარამეტრების მნიშვნელობების შესახებ ან ორი განაწილების პარამეტრების სიდიდეების შედარების ჰიპოთეზებს, **პარამეტრული ჰიპოთეზები** ეწოდება. ჰიპოთეზებს განაწილების სახის შესახებ კი **არაპარამეტრული ჰიპოთეზები** ეწოდება.

სტატისტიკური ჰიპოთეზის შემოწმება ნიშნავს, რომ შევამოწმოთ შერჩევიდან მიღებული მონაცემები არის თუ არა შესაბამისობაში მოცემულ ჰიპოთეზასთან (მონაცემები ეთანხმება თუ არა მოცემულ ჰიპოთეზას). შემოწმება ხორციელდება **სტატისტიკური კრიტერიუმის** საშუალებით. *სტატისტიკური კრიტერიუმი – ეს არის შემთხვევითი სიდიდე, რომლის განაწილების კანონი (პარამეტრების მნიშვნელობებთან ერთად) ცნობილია იმ შემთხვევაში, თუ მიღებული ჰიპოთეზა სამართლიანია (ზოგჯერ სტატისტიკურ კრიტერიუმს უბრალოდ* ***სტატისტიკას*** *უწოდებენ).* ამ კრიტერიუმს უწოდებენ აგრეთვე **თანხმობის კრიტერიუმს** (მხედველობაში აქვთ რა მიღებული ჰიპოთეზის თანხმობა შერჩევიდან მიღებულ შედეგებთან).

ჰიპოთეზას, რომელიც წამოყენებულია შერჩევით მონაცემებთან მისი თანხმობის შესამოწმებლად, **ნულოვანი ჰიპოთეზა** ეწოდება და აღინიშნება H_0-ით. H_0 ჰიპოთეზასთან ერთად იხილავენ (წამოაყენებენ) **ალტერნატიულ** ანუ **საწინააღმდეგო** ჰიპოთეზასაც, რომელსაც H_1-ით აღნიშნავენ. მაგალითად:

1) H_0: $E\xi = 0$, H_1: $E\xi \neq 0$ 2) H_0: $E\xi = 0$, H_1: $E\xi > 0$ 3) H_0: $E\xi = 0$, H_1: $E\xi = 2$.

დავუშვათ, რომ შემთხვევითი სიდიდე K -- არის გარკვეული H_0 ჰიპოთეზის შემოწმების სტატისტიკური კრიტერიუმი. H_0 ჰიპოთეზის სამართლიანობის შემთხვევაში K შემთხვევითი სიდიდის განაწილების კანონი

ხასიათდება გარკვეული, ჩვენთვის ცნობილი განაწილების სიმკვრივით $p_K(x)$. ამოვირჩიოთ გარკვეული მცირე ალბათობა α, რომელიც ტოლია 0.05-ის, 0.01-ის ან კიდევ უფრო მცირეა. განვმარტოთ **კრიტერიუმის კრიტიკული მნიშვნელობა** $K_{კრ}$ როგორც შემდეგი სამი განტოლებიდან ერთ-ერთის ამონახსნი, იმის მიხედვით თუ რა სახისაა ნულოვანი და ალტერნატიული ჰიპოთეზები:

$$P(K > K_{კრ}) = \alpha, \qquad (1)$$

$$P(K < K_{კრ}) = \alpha, \qquad (2)$$

$$P((K < K_{კრ1}) \cup (K > K_{კრ2})) = \alpha. \qquad (3)$$

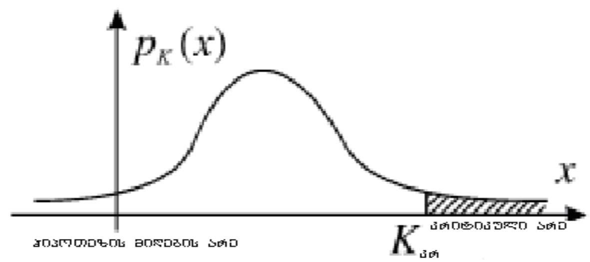
შესაძლებელია სხვა სახის განტოლებებიც, მაგრამ ყველაზე ხშირად გვხვდება სწორედ ასეთები.

(1) განტოლების ამოხსნა (ისევე როგორც (2) და (3) განტოლებების) მდგომარეობს შემდეგში: მოცემული α ალბათობით, ვიცით რა $p_K(x)$ ფუნქცია, რომელიც როგორც წესი მოცემულია ცხრილით, საჭიროა განისაზღვროს $K_{კრ}$.

რას ნიშნავს (1) პირობა?

თუ სამართლიანია H_0 ჰიპოთეზა, მაშინ ალბათობა იმისა, რომ K კრიტერიუმი გადააჭარბებს გარკვეულ $K_{კრ}$ მნიშვნელობას ძალიან მცირეა – 0.05, 0.01 ან კიდევ უფრო მცირე, იმის მიხედვით თუ ჩვენ რას ამოვირჩევთ. თუ $K_{შ}$ -- შერჩევითი მონაცემებით გამოთვლილი K კრიტერიუმის სიმძლავრე მეტია ვიდრე $K_{კრ}$, ეს იმას ნიშნავს, რომ შერჩევითი მონაცემები არ იძლევიან საფუძველს ნულოვანი H_0 ჰიპოთეზის მისაღებად (მაგალითად, თუ $\alpha = 0.01$, მაშინ შეიძლება ითქვას, რომ მოხდა ისეთი ხდომილება, რომელიც H_0 ჰიპოთეზის სამართლიანობის შემთხვევაში საშუალოდ გვხვდება არა უმეტეს ვიდრე ერთჯერ 100 შერჩევიდან). ამ შემთხვევაში ამბობენ, რომ H_0 **ჰიპოთეზა არ ეთანხმება შერჩევით მონაცემებს და ის უნდა იქნეს უკუგდებული.** თუ $K_{შ}$ არ აღემატება $K_{კრ}$-ს, მაშინ ამბობენ, რომ **შერჩევითი მონაცემები არ ეწინააღმდეგებიან** H_0 **ჰიპოთეზას,** და არა გვაქვს საფუძველი ამ ჰიპოთეზის უკუგდების.

(1) განტოლების შემთხვევში არეს -- $K > K_{კრ}$ ეწოდება **კრიტიკული არე. თუ $K_{შ}$-ს მნიშვნელობა მოხვდება კრიტიკულ არეში, მაშინ H_0 ჰიპოთეზა უკუგდებულ იქნება.** ასეთ კრიტიკულ არეს მარჯვენა კრიტიკული არე ეწოდება. ქვემოთ მოყვანილია (1) განტოლების საილუსტრაციო ნახაზი:

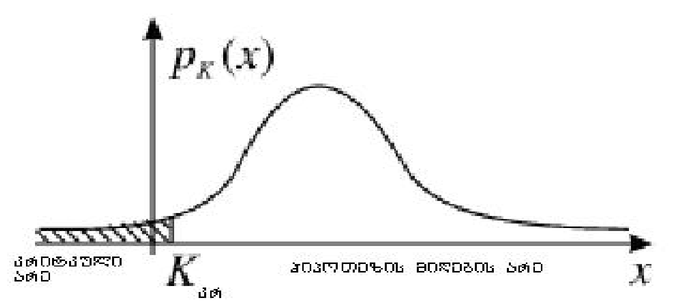


აქ $p_K(x)$ -- K შემთხვევითი სიდიდის ცნობილი განაწილების სიმკვრივეა H_0 ჰიპოთეზის სამართლიანობის შემთხვევაში. შევნიშნავთ, რომ დაშტრიხული ფიგურის ფართობი აქ α-ს ტოლია.

დავუშვათ, რომ შერჩეულია გარკვეული მცირე მნიშვნელობა α ალბათობის, ამ მნიშვნელობის მიხედვით განსაზღრულია $K_{კრ}$ და შერჩევითი მონაცემების მიხედვით განსაზღრულია $K_{დ}$-ს მნიშვნელობა, რომელიც მოხვდა კრიტიკულ არეში. ამ შემთხვევაში H_0 ჰიპოთეზა უკუგდებულ იქნება, მაგრამ ის შეიძლება აღმოჩნდეს სამართლიანი. უბრალოდ, შემთხვევით მოდხა ხდომილება, რომელსაც გააჩნია ძალიან მცირე ალბათობა α. ამ აზრით α არის ალბათობა იმისა, რომ უკუგდებულ იქნება სამართლიანი H_0 ჰიპოთეზა.

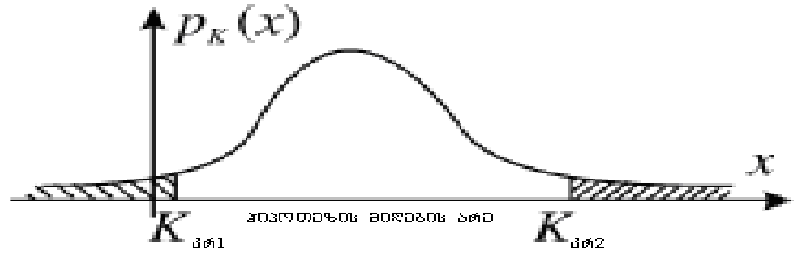
სამართლიანი ჰიპოთეზის უკუგდებას **პირველი გვარის შეცდომა** ეწოდება. α ალბათობას **მნიშვნელოვნების დონე** ეწოდება. ამრიგად, *მნიშვნელობნების დონე – ეს არის პირველი გვარის შეცდომის დაშვების ალბათობა.*

(2) განტოლება განსაზღვრავს **მარცხენა კრიტიკულ არეს.** მის გამოსახულებას აქვს შემდეგი სახე:



კრიტიკული არის (დაშტრიხული ფიგურის ფართობი) აქაც α-ს ტოლია.

და ბოლოს, (3) განტოლება განსაზღვრავს **ორმხრივ კრიტიკულ არეს.** ასეთი არე გამოსახულია ქვემოთ მოყვანილ ნახაზზე:



აქ კრიტიკული არე შედგება ორი ნაწილისაგან. მისი საზღვრები განისაზღვრება ისე, რომ სრულდებოდეს პირობა:

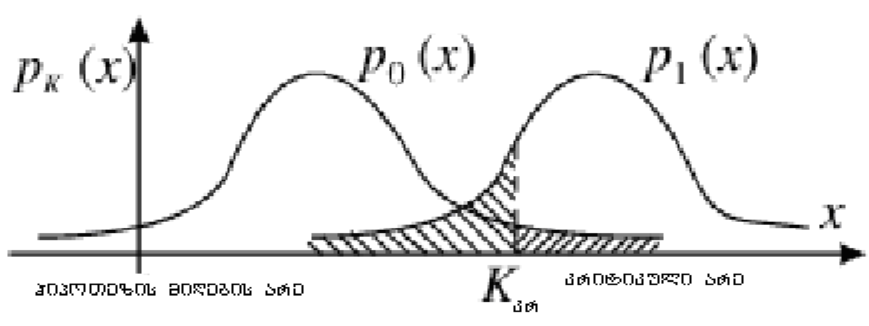
$$P(K \le K_{კრ1}) = P(K \ge K_{კრ2} = \alpha / 2 .$$

ამ შემთხვევაში თითოეული დაშტრიხული ფიგურის ფართობი ტოლია $\alpha / 2$-ის.

კრიტიკული არის სახე დამოკიდებულია იმაზე, თუ როგორია ალტერნატიული ჰიპოთეზაა.

რაც უფრო პატარაა მნიშვნელოვნების დონე, მით უფრო მცირეა ალბათობა იმისა, რომ უკუვაგდოთ შესამოწმებელი H_0 ჰიპოთეზა, როცა ის სამართლიანია, ანუ დავუშვათ პირველი გვარის შეცდომა. მაგრამ, მნიშვნელოვნების დონის შემცირებასთან ერთად ფართოვდება H_0 ჰიპოთეზის მიღების არე და შესაბამისად, იზრდება ალბათობა იმისა, რომ მივიღოთ შესამოწმებელი ჰიპოთეზა, როცა ის არაა სამართლიანი, ანუ მაშინ როცა უპირატესობა უნდა მინიჭოს ალტერნატიულ ჰიპოთეზას.

დავუშვათ რომ H_0 ჰიპოთეზის სამართლიანობისას K სტატისტიკურ კრიტერიუმს გააჩნია სიმკვრივე $p_0(x)$, ხოლო ალტერნატიული H_1 ჰიპოთეზის სამართლიანობისას კი -- $p_1(x)$ განაწილების სიმკვრივე. ამ ფუნქციების გრაფიკები გამოსახულია ნახაზზე:



მნიშვნელოვნების გარკვეული დონისათვის ვპოულობთ კრიტიკულ მნიშვნელობას $K_{კრ}$ და მარჯვენა კრიტიკული არეს. თუ შერჩევითი მონაცემების საშუალებით განსაზღრული $K_{შ}$-ს მნიშვნელობა აღმოჩნდება უფრო ნაკლები, ვიდრე $K_{კრ}$, მაშინ H_0 ჰიპოთეზა მიიღება. დავუშვათ, რომ სინამდვილეში სამართლიანია H_1 ჰიპოთეზა. მაშინ ალბათობა იმისა, რომ კრიტერიუმი მოხვდება H_0 ჰიპოთეზის მიღების არეში, არის გარკვეული რიცხვი β, რომელიც ტოლია იმ ფიგურის ფართობის, რომელიც შემოსაზღრულია $p_1(x)$ ფუნქციის გრაფიკითა და ჰორიზონტალური საკოორდინატო ღერძის ნახევრადუსასრულო ნაწილით, რომელიც ძევს $K_{კრ}$ წერტილის მარცხნივ. ცხადია, რომ β -- არის ალბათობა იმისა, რომ მიღებული იქნება არასამართლიანი H_0 ჰიპოთეზა.

არასამართლიანი ჰიპოთეზის მიღებას **მეორე გვარის შეცდომა** ეწოდება. განსახილველ შემთხვევაში რიცხვი β არის მეორე გვარის შეცდომის ალბათობა. **რიცხვს** $1-\beta$**,** რომელიც ტოლია ალბათობის იმისა, რომ არ იქნება დაშვებული მეორე გვარის შეცდომა, **კრიტერიუმის სიმძლავრე** ეწოდება. ზემოთ მოყვანილ ნახაზზე, კრიტერიუმის სიმძლავრე ტოლია იმ ფიგურის ფართობის, რომელიც რომელიც შემოსაზღრულია $p_1(x)$ ფუნქციის გრაფიკითა და ჰორიზონტალური საკოორდინატო ღერძის ნახევრადუსასრულო ნაწილით, რომელიც ძევს $K_{კრ}$ წერტილის მარჯვნივ.

სტატისტიკური კრიტერიუმისა და კრიტიკული არის სახის შერჩევა ხდება ისე, რომ კრიტერიუმის სიმძლავრე იყოს მაქსიმალური.

§36. ჰიპოთეზის შემოწმება ლოდინის შესახებ

სტატისტიკური ჰიპოთეზის შემოწმება ნორმალური განაწილების მათემატიკური ლოდინის შესახებ ცნობილი დისპერსიის შემთხვევაში:

დავუშვათ, რომ მოცემულია ნორმალური კანონით განაწილებული შემთხვევითი სიდიდე ξ, რომელიც განმარტებულია გარკვეული გენერალური ერთობლიობის ობიექტების სიმრავლეზე. ცნობილია, რომ $D\xi = \sigma^2$. მათემატიკური ლოდინი $E\xi$ უცნობია. დავუშვათ, რომ ჩვენ გაგვაჩნია საფუძველი იმისა, რომ დავუშვათ: $E\xi = a$, სადაც a -- გარკვეული რიცხვია (ასეთი საფუძველი შეიძლება იყოს ინფორმაცია გენერალური ერთობლიობის ობიექტების შესახებ, მსგავსი ერთობლიობების კვლევის გამოცდილება და სხვა). ვიგულისხმოთ, რომ გვაქვს აგრეთვე მეორე ინფორმაცია, რომელიც გვიცვენებს, რომ $E\xi = a_1$, სადაც $a_1 > a$.

I. ვაყენებთ ნულოვან ჰიპოთეზას -- $H_0 : E\xi = a$ ალტერნატიული ჰიპოთეზის წინააღმდეგ -- $H_1 : E\xi = a_1$.

ვაკეთებთ n მოცულობის შერჩევას $x_1, x_2, ..., x_n$. შემოწმებას საფუძვლად უდევს ის ფაქტი, რომ შემთხვევითი სიდიდე $\bar{x}$ (შერჩევითი საშუალო) განაწილებულია ნორმალური განაწილების კანონით σ^2/n-ის ტოლი დისპერსიითა და a-ს (შესაბამისად, a_1-ის) ტოლი მათემატიკური ლოდინით H_0 (შესაბამისად, H_1) ჰიპოთეზის სამართლიანობის შემთხვევაში.

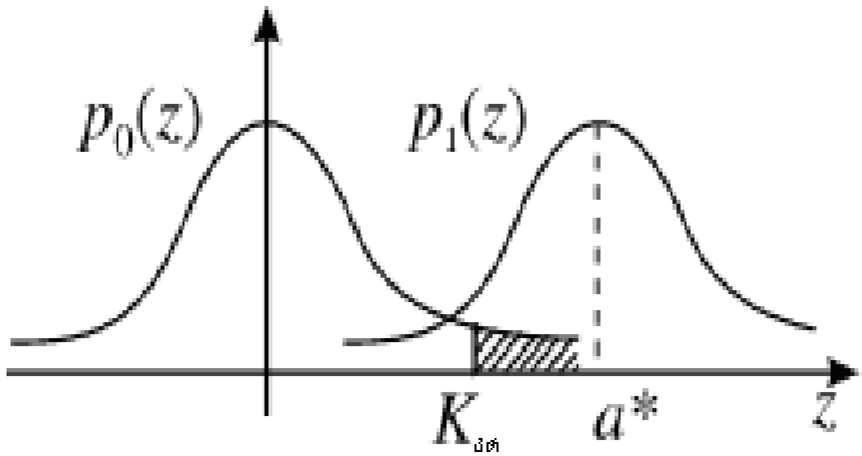
ცხადია, რომ თუ სიდიდე $\bar{x}$ აღმოჩნდება საკმარისად მცირე, მაშინ ეს გვაძლევს საფუძველს H_0 ჰიპოთეზას მივანიჭოთ უპირატესობა H_1 ჰიპოთეზასთან შედარებით. მეორეს მხრივ, $\bar{x}$-ის საკმარისად დიდი მნიშვნელობის შემთხვევაში უფრო ალბათურია H_1 ჰიპოთეზის სამართლიანობა. ამოცანა შეიძლება ასე დაისვას: საჭიროა მიოძებნოს გარკვეული კრიტიკული რიცხვი, რომელიც შერჩევითი საშუალოს ყველა შესაძლო მნიშვნელობებს (ამ ამოცანის შემთხვევაში, ეს მთლიანად ნამდვილ რიცხვთა სიმრავლეა) გაყოფს ნახევრადუსასრულო შუალედად. $\bar{x}$ შერჩევითი საშუალოს მარცხენა ინტერვალში მოხვედრისას უნდა მივიღოთ H_0 ჰიპოთეზა, ხოლო $\bar{x}$-ის მარჯვენა ინტერვალში მოხვედრისას უპირატესობა უნდა მიენიჭოს H_1 ჰიპოთეზას. თუმცა სინამდვილეში იქცევიან რამდენადმე სხვანაირად.

სტატისტიკური კრიტერიუმის როლში ირჩევენ შემთხვევით სიდიდეს

$$z = (\bar{x} - a)\sqrt{n}/\sigma,$$

რომელიც განაწილებულია ნორმალური განაწილების კანონით, პარამეტრრებით: $Ez = 0$ და $Dz = 1$ H_0 ჰიპოთეზის სამართლიანობის შემთხვევაში. თუ კი სამართლიანია H_1 ჰიპოთეზა, მაშინ $Ez = a^* = (a_1 - a)\sqrt{n}/\sigma$ და $Dz =$

=1. ქვემოთ მოყვანილია $p_0(z)$ და $p_1(z)$ ფუნქციების გრაფიკები, რომლებიც წარმოადგენენ z შემთხვევითი სიდიდის განაწილების სიმკვრივის ფუნქციებს შესაბამისად H_0 და H_1 ჰიპოთეზების სამართლიანობისას.



თუ შერჩევითი მონაცემებიდან მიღებული $\bar{x}$-ის მნიშვნელობა შედარებით დიდია, მაშინ z სიდიდეც დიდი იქნება, რაც წარმოადგენს მტკიცებულებას H_1 ჰიპოთეზის სასარგებლოდ. $\bar{x}$-ის შედარებით მცირე მნიშვნელობებს მივყავართ z-ის მცირე მნიშვნელობებამდე, რაც მეტყველებს H_0 ჰიპოთეზის სასარგებლოდ. აქედან გამომდინარეობს, რომ უნდა შეირჩეს მარჯვენა კრიტიკული არე. არჩეული მნიშვნელოვნების α დონისათვის (მაგალითად, $\alpha = 0.05$), ვისარგებლებთ რა იმ გარემოებით, რომ შემთხვევითი სიდიდე z განაწილებულია ნორმალური განაწილების კანონით, განვსაზღვრავთ $K_{კრ}$-ს შემდეგი თანაფარდობიდან:

$$\alpha = P(K_{კრ} < z < \infty) = \Phi_0(\infty) - \Phi_0(K_{კრ}) = 0.5 - \Phi_0(K_{კრ}).$$

აქედან $\Phi_0(K_{კრ}) = (1-2\alpha)/2$, და $K_{კრ}$-ს საპოვნელად საჭიროა ვისარგებლოთ სტანდარტული ნორმალური განაწილების ცხრილით. ცხადია, რომ

$$K_{კრ} = x_{1-\alpha} = z_\alpha .$$

თუ z-ის მნიშვნელობა, გამოთვლილი $\bar{x}$ შერჩევითი საშუალოს მიხედვით, მოხვდება ჰიპოთეზის მიღების არეში ($z < K_{კრ}$), მაშინ H_0 ჰიპოთეზა მიიღება (კეთდება დასკვნა, რომ შერჩევითი მონაცემები არ ეწინააღმდეგება H_0 ჰიპოთეზას). თუ z სიდიდე ხვდება კრიტიკულ არეში, მაშინ H_0 ჰიპოთეზას უკუაგდებენ.

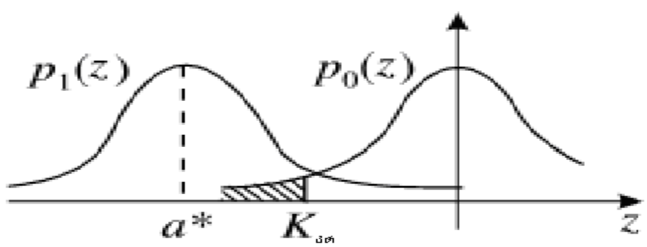
გამოვთვალოთ ამ ამოცანაში კრიტერიუმის სიმძლავრე. გვაქვს:

$$1-\beta = \phi(\infty) - \phi[K_{კრ} - (a_1 - a)\sqrt{n}/\sigma].$$

აქედან ჩანს, რომ კრიტერიუმის სიმძლავრე მით უფრო დიდია, რაც უფრო დიდია სხვაობა $a_1 - a$.

II. თუ წინა ამოცანაში დავსვამთ სხვა პირობას, კერძოდ, $a_1 < a$, ანუ

განვიხილავთ ნულოვან ჰიპოთეზას -- $H_0 : E\xi = a$ ალტერნატიული ჰიპოთეზის წინააღმდეგ -- $H_1 : E\xi = a_1$, $a_1 < a$, მაშინ ზემოთ მოყვანილი მსჯელობის ანალოგიით გასაგებია, რომ უნდა განვიხილოთ მარცხენა კრიტიკული არე. ნახაზი იქნება შემდეგი სახის:



აქ, ისევე როგორც წინა შემთხვევაში, $a^* = (a_1 - a)\sqrt{n}/\sigma$, ხოლო კრიტიკული რიცხვი $K_{კრ}$ განისაზღვრება შემდეგი თანაფარდობიდან:

$$\alpha = P(-\infty < z < K_{კრ}) = \Phi_0(K_{კრ}) - \Phi_0(-\infty) = \Phi_0(K_{კრ}) + 0.5 .$$

თუ ვისარგებლებთ თანაფარდობით $-\Phi_0(K_{კრ}) = \Phi_0(-K_{კრ})$, მივიღებთ:

$$\Phi_0(-K_{კრ}) = (1 - 2\alpha)/2 ,$$

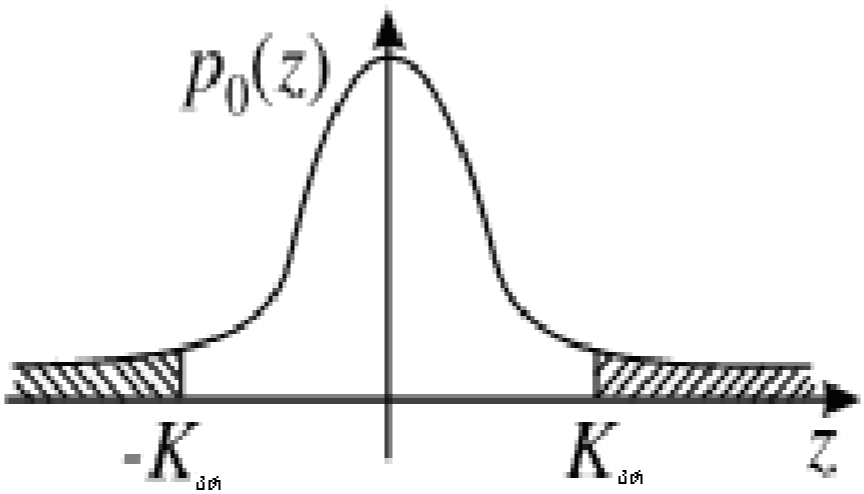
ე. ი. $-K_{კრ} = x_{1-\alpha} = z_\alpha$, ანუ $K_{კრ} = -x_{1-\alpha} = -z_\alpha$ (შევნიშნავთ, რომ ამოცანის შინაარსიდან გამომდინარე, აქ $K_{კრ}$ უარყოფითი რიცხვია).

შერჩევითი მონაცემებით გამოთვლილი z-ის ის მნიშვნელობები, რომლებიც მეტია $K_{კრ}$-ზე ეთანხმება H_0 ჰიპოთეზას. თუ z სიდიდე ხვდება კრიტიკულ არეში ($z < K_{კრ}$), მაშინ უპირატესობა ენიჭება H_1 ჰიპოთეზას და უკუაგდებენ H_0 ჰიპოთეზას.

III. განვიხილოთ ახლა ასეთი ამოცანა:

$$H_0 : E\xi = a ;$$
$$H_1 : E\xi \neq a .$$

ამ შემთხვევაში z სიდიდის დიდი გადახრები ნულისაგან როგორც დადებით, ისე უარყოფით მხარეს, მეტყველებს H_0 ჰიპოთეზის საწინააღმდეგოდ, ანუ აქ საჭიროა განხილულ იქნეს ორმხრივი კრიტიკული არე, ისე როგორც ეს ნაჩვენებია ნახაზზე:



ამ შემთხვევაში კრიტიკული მნიშვნელობა $K_{კრ}$ განისაზღვრება თანაფარდობიდან:

$$P(-K_{კრ} < z < K_{კრ}) = 1-\alpha = \Phi_0(K_{კრ}) - \Phi_0(-K_{კრ}) = 2\Phi_0(K_{კრ}),$$

საიდანაც ვღებულობთ, რომ:

$$\Phi_0(K_{კრ}) = (1-\alpha)/2,$$

და შესაბამისად,

$$K_{კრ} = x_{1-\alpha/2} = z_{\alpha/2}.$$

p-მნიშვნელობის მეთოდი.

ჰიპოთეზის შემოწმების მეორე ექვივალენტური მეთოდი მდგომარეობს შემდეგში: $x_1,\dots,x_n$ შერჩევის საფუძველზე გამოვთვალოთ კრიტერიუმის $Z=\left(\overline{X}-a\right)\sqrt{n}/\sigma$ სტატისტიკის $z=(\bar{x}-a)\sqrt{n}/\sigma$ მნიშვნელობა და შემდეგი ალბათობა:

$$p = P\{Z \ge z \mid H_0\} = P\{\frac{\overline{X}-a}{\sigma/\sqrt{n}} \ge \frac{\bar{x}-a}{\sigma/\sqrt{n}} \mid H_0\} = 1-\Phi(\frac{\bar{x}-a}{\sigma/\sqrt{n}}).$$

ცხადია, რომ თუ $p \le \alpha$, მაშინ $z \ge z_\alpha$, რაც ჰიპოთეზის გარჩევის მარჯვენა ცალმხრივი კრიტერიუმის შესაბამისად იწვევს H_0 ჰიპოთეზის უარყოფას. აქედან შეიძლება დავასკვნათ, რომ: $x_1,\dots,x_n$ შერჩევის საფუძველზე ნებისმიერი α, $\alpha \ge p$ მნიშვნელოვნების დონით ხდება H_0 ჰიპოთეზის უარყოფა. შესაბამისად, p-მნიშვნელობა არის ის მინიმალური მნიშვნ-

ელოვნების დონე, რომლითაც $x_1,...,x_n$ მონაცემებით H_0 ჰიპოთეზას უარვყოფთ. ამრიგად, ჰიპოთეზათა გარჩევის p-მნიშვნელობის მეთოდი ასე ყალიბდება:

მოცემული α მნიშვნელოვნების დონისათვის, $x_1,...,x_n$ მონაცემებით H_0 ჰიპოთეზას უარვყოფთ, თუ p-მნიშვნელობა ნაკლებია ან ტოლი α-ზე ($p \le \alpha$).

სტატისტიკური ჰიპოთეზის შემოწმება ნორმალური განაწილების მათემატიკური ლოდინის შესახებ უცნობი დისპერსიის შემთხვევაში:

ამ შემთხვევაში სტატისტიკური კრიტერიუმის (სტატისტიკის) როლში იღებენ შემდეგ შემთხვევით სიდიდეს:

$$T = \frac{(\bar{X} - a_0)\sqrt{n}}{S^{'}},$$

სადაც $S^{'}$ – შესწორებული საშუალო კვადრატული გადახრაა. ცნობილია, რომ ამ შემთხვევით სიდიდეს გააჩნია სტიუდენტის განაწილება თავისუფლების ხარისხით $k = n - 1$. განვიხილოთ იგივე ალტერნატიული ჰიპოთეზები და შესაბამისი კრიტიკული არეები, რაც გვქონდა ცნობილი დისპერსიის შემთხვევაში. წინასწარ გამოვთვალოთ კრიტერიუმის დაკვირვებული (შერჩევითი) მნიშვნელობა

$$T_{დ} = \frac{(\bar{x}_{შ} - a_0)\sqrt{n}}{s^{'}}.$$

დავუშვათ, რომ გვაქვს ნულოვანი ჰიპოთეზა -- $H_0 : E\xi = a_0$ ალტერნატიული ჰიპოთეზის წინააღმდეგ -- $H_1 : E\xi \ne a_0$. მოვცემული α-სა და $k = n-1$-სათვის სტიუდენტის განაწილების ცხრილიდან ვპოულობთ სტიუდენტის ორმხრივ კრიტიკულ წერტილს $t_{კრ} = t_{n-1,\alpha/2}$. თუ აღმოჩნდა, რომ $|T_{დ}| < t_{კრ}$, მაშინ ნულოვანი ჰიპოთეზა მიიღება. თუ $|T_{დ}| > t_{კრ}$, მაშინ ნულოვანი ჰიპოთეზა უკუგდებულ იქნება.

თუ იგივე ნულოვანი ჰიპოთეზის წინააღმდეგ განვიხილავთ ალტერნატიულ $H_1 : E\xi > a_0$ ჰიპოთეზას, მაშინ შესაბამისი ცხრილიდან ვპოულობთ მარჯვენა კრიტიკული არის კრიტიკულ წერტილს $t_{კრ}(\alpha,k) = t_{n-1,\alpha}$, და მივიღებთ ნულოვან ჰიპოთეზას, თუ $T_{დ} < t_{კრ}(\alpha,k)$ (წინააღმდეგ შემთხვევაში მიიღება ალტერნატიული ჰიპოთეზა).

ბოლოს, თუ ალტერნატიული ჰიპოთეზაა $H_1 : E\xi < a_0$, გვექნება მარცხენა კრიტიკული არე და ნულოვანი ჰიპოთეზა მიიღება იმ შემთხვევაში, როცა $T_{დ} > -t_{n-1,\alpha}$. თუ კი $T_{დ} < -t_{n-1,\alpha}$, მაშინ ნულოვან ჰიპოთეზას უარყოფენ.

ჰიპოთეზათა შემოწმების ნდობის ინტერვალის მეთოდი

თუ ჰიპოთეზის შემოწმების ამოცანაში ხდება ნულოვანი ჰიპოთეზის უარყოფა, მაშინ იმავე მნიშვნელოვნების დონის მქონე ნდობის ინტერვალი პოპულაციის პარამეტრისათვის *არ მოიცავს* პარამეტრის ჰიპოთეტურ მნიშვნელობას.

თუ ჰიპოთეზის შემოწმების ამოცანაში არ ხდება ნულოვანი ჰიპოთეზის უარყოფა, მაშინ იმავე მნიშვნელოვნების დონის მქონე ნდობის ინტერვალი პოპულაციის პარამეტრისათვის *მოიცავს* პარამეტრის ჰიპოთეტურ მნიშვნელობას.

მაგალითი. შაქარი დაფასოებულია 5 ფუნტიან (1 ფუნტი = 453.6 გრ) ფუთებში. კონტროლიორს ეჭვი აქვს, რომ ფუთაში არ არის 5 ფუნტი შაქარი. 50 შაქრის ფუთისაგან შემდგარი შერჩევის საშუალო აღმოჩნდა 4.6 ფუნტი, ხოლო შესწორებული სტანდარტული გადახრა კი 0.7 ფუნტი. $\alpha = 0.05$ მნიშვნელოვნების დონით გვაქვს თუ არა საფუძველი დავასკვნათ, რომ ფუთაში საშუალოდ არ არის 5 ფუნტი შაქარი? ააგეთ 95%-იანი ნდობის ინტერვალი ჭეშმარიტი საშუალოსათვის.

ამოხსნა. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0 : E\xi = 5$, $H_1 : E\xi \neq 5$. ამ ორმხრივი კრიტერიუმის შესაბამისი კრიტიკული მნიშვნელობებია: $-z_{\alpha/2} = -z_{0.025} = -1.96$ და $z_{\alpha/2} = z_{0.025} = 1.96$. კრიტერიუმის დაკვირვებული მნიშვნელობა იქნება:

$$z = \frac{\bar{x} - E\xi}{s' / \sqrt{n}} = \frac{4.5 - 5}{0.7 / \sqrt{50}} = \frac{-0.4}{0.099} = -4.04 .$$

რადგანაც $-4.04 < -1.96$, ამიტომ ნულოვანი ჰიპოთეზა უნდა უკუვაგდოთ, ანუ ჩვენ გვაქვს საკმარისი საფუძველი, რათა დავასკვნათ, რომ ფუთაში საშუალოდ არ არის 5 ფუნტი შაქარი.

ავაგოთ 95%-იანი ნდობის ინტერვალი საშუალოსათვის:

$$\bar{x} - z_{\alpha/2} \cdot \frac{s'}{\sqrt{n}} < E\xi < \bar{x} + z_{\alpha/2} \cdot \frac{s'}{\sqrt{n}} ,$$

$$4.6 - 1.96 \cdot \frac{0.7}{\sqrt{50}} < E\xi < 4.6 + 1.96 \cdot \frac{0.7}{\sqrt{50}} ,$$

$$4.4 < E\xi < 4.8 .$$

როგორც ვხედავთ, 95%-იანი ნდობის ინტერვალი (ანუ ნდობის ინტერვალი მნიშვნელოვნების დონით $\alpha = 0.05$) საშუალოსათვის $E\xi$ არ მოიცავს საშუალოს ჰიპოთეტურ მნიშვნელობას $E\xi = 5$. შესაბამისად, ნდობის ინტერვალის მეთოდის თანახმად, ხდება ნულოვანი ჰიპოთეზის უარყოფა. როგორც ვხედავთ, გამოტანილი დასკვნები, როგორც ტრადიციული, ისე ნდობის ინტერვალის მეთოდით, იდენტურია.

§37. ჰიპოთეზათა შემოწმება ბერნულის სქემაში

დავუშვათ, რომ ჩატარებულია n დამოუკიდებელი ექსპერიმენტი (n – საკმაოდ დიდი რიცხვია), რომელთაგან თითოეულში გარკვეული A ხდომილება ხდება ერთი და იგივე, მაგრამ უცნობი p ალბათობით; ნაპოვნია ექსპერიმენტების ამ სერიაში A ხდომილების მოხდენის ფარდობითი სიხშირე $\frac{m}{n}$. მნიშვნელოვნების მოცემული α დონისათვის შევამოწმოთ H_0 ჰიპოთეზა, რომელიც მდგომარეობს იმაში, რომ p ალბათობა ტოლია გარკვეული p_0 რიცხვის.

სტატისტიკური კრიტერიუმის როლში ავიღოთ შემთხვევითი სიდიდე

$$U=\frac{\left(\frac{m}{n}-p_0\right)\sqrt{n}}{\sqrt{p_0q_0}},$$

რომელსაც გააჩნია ნორმალური განაწილება პარამეტრებით $E(U)=0$, $\sigma(U)=1$. აქ $q_0=1-p_0$. დასკვნა კრიტერიუმის ნორმალურად განაწილებულობის შესახებ გამოდის ლაპლასის თეორემიდან (საკმაოდ დიდი n–ებისათვის ფარდობითი სიხშირე დაახლოებით შეიძლება ჩაითვალოს ნორმალურად განაწილებულად მათემატიკური ლოდინით p და საშუალო კვადრატული გადახრით $\sqrt{\frac{pq}{n}}$). კრიტიკული არე იგება ალტერნატიული ჰიპოთეზის სახის მიხედვით.

1). თუ $H_0: p=p_0$, ხოლო $H_1: p\neq p_0$, მაშინ კრიტიკული არე უნდა ავაგოთ ისე, რომ კრიტერიუმის ამ არეში მოხვედრის ალბათობა ტოლი იყოს მოცემული α მნიშვნელოვნების დონის. ამასთანავე კრიტერიუმის უდიდესი სიმძლავრე მიიღწევა მაშინ, როცა კრიტიკული არე შედგება ორი ინტერვალისაგან, რომელთაგან თითოეულში მოხვედრის ალბათობაა $\frac{\alpha}{2}$. ვინაიდან U სიმეტრიულია ორდინატთა ღერძის მიმართ, ამიტომ მისი $(-\infty;0)$ და $(0;+\infty)$ ინტერვალებში მოხვედრის ალბათობებია 0.5. შესაბამისად, კრიტიკული არე აგრეთვე უნდა იყოს სიმეტრიული ორდინატთა ღერძის მიმართ. ამიტომ, $u_{კრ}$ განისაზღვრება ნორმალური განაწილების ცხრილიდან, ისე რომ შესრულდეს პირობა $\Phi_0(u_{კრ})=\frac{1-\alpha}{2}$, ხოლო კრიტიკულ არეს აქვს სახე: $(-\infty;-u_{კრ})\cup(u_{კრ};+\infty)$ (აქ $\Phi_0(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-t^2/2}dt$).

შემდგომ უნდა გამოვთვალოთ კრიტერიუმის დაკვირვებული (შერჩევითი) მნიშვნელობა

$$U_{\text{დ}}=\frac{\left(\frac{m}{n}-p_0\right)\sqrt{n}}{\sqrt{p_0q_0}}.$$

თუ აღმოჩნდა, რომ $|U_{\text{ვ}}| < u_{\text{კრ}}$, მაშინ მიიღება ნულოვანი ჰიპოთეზა, ხოლო თუ $|U_{\text{ვ}}| > u_{\text{კრ}}$, მაშინ ნულოვან ჰიპოთეზას უკუვაგდებთ.

2). თუ ალტერნატიული ჰიპოთეზა H_1: $p > p_0$ სახისაა**,** მაშინ კრიტიკული არე განისაზღვრება უტოლობით $U > u_{\text{კრ}}$, ე. ი. გვაქვს მარჯვენა კრიტი-
კული არე, ამასთან $P(U > u_{\text{კრ}}) = \alpha$. შესაბამისად,

$$P(0 < U < u_{\text{კრ}}) = \frac{1}{2} - \alpha = \frac{1-2\alpha}{2}.$$

ამიტომ ნორმალური განაწილების ცხრილიდან ვიპოვით $u_{\text{კრ}}$-ს ისე, რომ

$$\varPhi_0(u_{\text{კრ}}) = \frac{1-2\alpha}{2}.$$

შემდეგ ვითვლით კრიტერიუმის შერჩევით მნიშვნელობას

$$U_{\text{ვ}} = \frac{\left(\frac{m}{n} - p_0\right)\sqrt{n}}{\sqrt{p_0 q_0}}.$$

და, თუ აღმოჩნდა, რომ $U_{\text{ვ}} < u_{\text{კრ}}$, მაშინ მიიღება ნულოვანი ჰიპოთეზა. თუ კი $U_{\text{ვ}} > u_{\text{კრ}}$, მაშინ მიიღება ალტერნატიული ჰიპოთეზა.

3). ალტერნატიული ჰიპოთეზისათვის H_1: $p < p_0$, კრიტიკული არა მარცხენა ცალმხრივია და მოიცემა უტოლობით $U < -u_{\text{კრ}}$, სადაც $u_{\text{კრ}}$ გამოითვლება ისე, როგორც წინა შემთხვევაში.

თუ $U_{\text{ვ}} > -u_{\text{კრ}}$, მაშინ მიიღება ნულოვანი ჰიპოთეზა.

თუ $U_{\text{ვ}} < -u_{\text{კრ}}$, მაშინ მიიღება ალტერნატიული ჰიპოთეზა.

მაგალითი. დავუშვათ ჩატარებულია 50 დამოუკიდებელი ექსპერიმენტი, A ხდომილების მოხდენის ფარდობითი სიხშირე აღმოჩნდა 0,12. მნიშვნელოვნების $\alpha = 0.01$ დონისათვის შევამოწმოთ ნულოვანი H_0: $p = 0.1$ ჰიპოთეზა ალტერნატიული H_1: $p > 0.1$ ჰიპოთეზის შემთხვევაში.

ამოხსნა. ვიპოვოთ კრიტერიუმის შერჩევითი მნიშვნელობა

$$U_{\text{ვ}} = \frac{(0.12 - 0.1)\sqrt{50}}{\sqrt{0.1 \cdot 0.9}} = 0.471.$$

კრიტიკული არე იქნება მარჯვენა ცალმხრივი, ხოლო კრიტიკული წერტილი უნდა ვიპოვოთ პირობიდან

$$\Phi_0(u_{\text{კრ}}) = \frac{1 - 2 \cdot 0.01}{2} = 0.49.$$

ნორმალური განაწილების ფუნქციის ცხრილიდან ვპოულობთ, რომ $u_{\text{კრ}} = 2.33$. ვინაიდან, $U_{\text{ვ}} < u_{\text{კრ}}$, ამიტომ მიიღება ჰიპოთეზა იმის შესახებ, რომ $p = 0.1$.

§ 38. ჰიპოთეზის შემოწმება ნორმალური პოპულაციის დისპერსიისათვის.

ნორმალური გენერალური ერთობლიობიდან მიღებული შერჩევისათვის შევამოწმოთ შემდეგი სახის ჰიპოთეზები:

ა). $H_0 : \sigma^2 = \sigma_0^2$, ბ). $H_0 : \sigma^2 = \sigma_0^2$, გ). $H_0 : \sigma^2 = \sigma_0^2$,

$H_1 : \sigma^2 > \sigma_0^2$, $H_1 : \sigma^2 < \sigma_0^2$, $H_1 : \sigma^2 \neq \sigma_0^2$,

სადაც ყველგან σ_0^2 წარმოადგენს რაიმე ცნობილ რიცხვს. შესაბამისად, გვაქვს მარჯვენა ცალმხრივი ალტერნატივა, მარცხენა ცალმხრივი ალტერნატივა და ორმხრივი ალტერნატივა. ამასთანავე, იმ შემთხვევაში, როცა პოპულაციის საშუალო μ ცნობილია, ძირითადი ჰიპოთეზა მარტივია, ხოლო ალტერნატივა რთულია. როცა μ უცნობია, როგორც ნულოვანი, ისე ალტერნატიული ჰიპოთეზები რთულია.

ვინაიდან, უცნობი დისპერსიის შეფასებას წარმოადგენს შერჩევითი დისპერსია, ამიტომ ბუნებრივია ცნობილი μ-ს შემთხვევაში კრიტერიუმის სტატისტიკად ავიღოთ ეს უკანასკნელი:

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2 , \qquad (1)$$

ხოლო უცნობი μ-ს შემთხვევაში კი შესწორებული შერჩევითი დისპერსია:

$$S_n^{'2} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}^2) . \qquad (2)$$

როგორც ცნობილია, ნულოვანი ჰიპოთეზის სამართლიანობის შემთხვევაში

$$\chi_n^2 := \frac{nS_n^2}{\sigma^2} \sim \chi^2(n)$$

და

$$\chi_{n-1}^2 := \frac{(n-1)S_n^{'2}}{\sigma^2} \sim \chi^2(n-1) .$$

ვინაიდან (1) და (2) დისპერსიის ჩაუნაცვლებელი შეფასებებია, ამიტომ შერჩევითი (შესაბამისად, შესწორებული შერჩევითი) დისპერსიის მნიშვნელოვანი გადახრა σ_0^2-საგან მეტყველებს ალტერნატიული ჰიპოთეზის სასარგებლოდ. ამიტომ, ფიქსირებული α მნიშვნელოოვნების დონისათვის კრიტიკული არეები შემდეგნაირად აიგება:

ა). ცალმხრივი მარჯვენა ალტერნატივა. კრიტიკულია არეა χ^2 განაწილების α ზომის ზედა არე, ანუ

$U_1 = [\chi_{n,\alpha}^2, +\infty)$, როცა μ ცნობილია,

$U_1 = [\chi_{n-1,\alpha}^2, +\infty)$, როცა μ უცნობია,

სადაც $\chi_{n,\alpha}^2$ არის $\chi^2(n)$ განაწილების ზედა α-კრიტიკული წერტილი.

ბ). ცალმხრივი მარცხენა ალტერნატივა. კრიტიკულია არეა χ^2 განაწილების α ზომის ქვედა არე, ანუ

$$U_1 = (0, \chi^2_{n,1-\alpha}]\text{, როცა } \mu \text{ ცნობილია,}$$

$$U_1 = (0, \chi^2_{n-1,1-\alpha}]\text{, როცა } \mu \text{ უცნობია.}$$

გ). ორმხრივი ალტერნატივა. კრიტიკული არე შედგება ორი ნაწილისაგან: χ^2 განაწილების $\alpha/2$ ზომის ზედა და χ^2 განაწილების $\alpha/2$ ზომის ქვედა არეებისაგან:

$$U_1 = (0, \chi^2_{n,1-\alpha/2}] \cup [\chi^2_{n,\alpha/2}, +\infty)\text{, როცა } \mu \text{ ცნობილია,}$$

$$U_1 = (0, \chi^2_{n-1,1-\alpha/2}] \cup [\chi^2_{n-1,\alpha/2}, +\infty)\text{, როცა } \mu \text{ უცნობია.}$$

შერჩევის მოცემული $x_1,...,x_n$ მნიშვნელობებით ვითვლით შერჩევით დისპერსიასა s^2 (შესაბამისად, შესწორებულ შერჩევით დისპერსიას s'^2) და იმ შემთხვევაში, როცა μ ცნობილია, თუ $\frac{ns^2}{\sigma_0^2} \in U_1$ უარვყოფთ H_0 ჰიპოთეზას (შესაბამისად, როცა μ უცნობია, თუ $\frac{(n-1)s'^2}{\sigma_0^2} \in U_1$ უარვყოფთ H_0 ჰიპოთეზას), წინააღმდეგ შემთხვევაში H_0 ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს.

ჰიპოთეზის შემოწმება პუასონის პოპულაციის μ პარამეტრის შესახებ (მცირე მოცულობის შერჩევებისათვის)

პუასონის პოპულაციის μ პარამეტრის შესახებ ჰიპოთეზის შემოწმებისას მცირე მოცულობის შერჩევის შემთხვევაში, იყენებენ p-მნიშვნელობის მეთოდს.

დავუშვათ, რომ ვამოწმებთ ნულოვან ჰიპოთეზას $H_0: \mu = \mu_0$ ალტერნატივის $H_1: \mu \neq \mu_0$ წინააღმდეგ.

აქაც, თუ $x \leq \mu_0$, სადაც x არის პუასონის შემთხვევითი სიდიდის დაკვირვებული მნიშვნელობა, მაშინ

$$p/2 = P\{X \leq x \mid \mu = \mu_0\} = \sum_{k=0}^{x} \frac{\mu_0^k}{k!} \cdot e^{-\mu_0}\ , \qquad (1)$$

ხოლო თუ $x > \mu_0$, მაშინ

$$p/2 = P\{X \geq x \mid \mu = \mu_0\} = 1 - \sum_{k=0}^{x-1} \frac{\mu_0^k}{k!} \cdot e^{-\mu_0}\ . \qquad (2)$$

თუ p-ს გამოთვლილი მნიშვნელობა აღმოჩნდება 0.05-ზე პატარა, მაშინ ამბობენ, რომ შედეგი (ლაპარაკია *b*-ზე) ***სტატისტიკურად მნიშვნელოვანია***, წინააღმდეგ შემთხვევაში კი – ***ის სტატისტიკურად უმნიშვნელოა*** და ნულოვან ჰიპოთეზას არ უარყოფენ.

დიდი მოცულობის შემთხვევაში აქაც იყენებენ ნორმალურ აპროქსიმაციას, კერძოდ, იმ ფაქტს, რომ

$$(X - \mu_0)^2 / \mu_0 \cong N^2(0,1) \cong \chi^2(1)\ .$$

§39. ჰიპოთეზათა შემოწმება ორამოკრეფიანი ამოცანებში

სტატისტიკურ გადაწყვეტილებათა თეორიის ერთ-ერთ უმნიშვნელოვანეს ამოცანას წარმოადგენს ორი პოპულაციის შედარება, რათა დადგინდეს, იძლევა თუ არა მოქმედების (ქცევის) ალტერნატიული გზა უკეთეს შედეგს. აქმდე ჩვენ ვამოწმებდით ჰიპოთეზებს ერთი პოპულაციის პარამეტრების კონკრეტული მნიშვნელობის შესახებ, რომელიც შეიძლებოდა ყოფილიყო გარკვეული მოსაზრებებით დასახელებული რიცხვი, ან ცნობილი ყოფილიყო წინა გამოკვლევებიდან ან მოტანილი ყოფილიყო უფრო დიდი (ან მცირე) მოცულობის პოპულაციაზე დაკვირვებებიდან.

ახლა ჩვენ შევუდგებით ე.წ. *ორამოკრეფიანი ამოცანების* განხილვას, რაც გულისხმობს ორი პოპულაციის პარამეტრების შედარებას, ისე, რომ არცერთის კონკრეტული მნიშვნელობა ცნობილი არ არის. მაშასადამე, ამოცანა მდგომარეობს შემდეგში: მოცემულია ორი შერჩევა. ერთი შერჩევა აღებულია A, ხოლო მეორე B პოპულაციებიდან, რომელთა მახასიათებლები (პარამეტრები) ჩვენთვის უცნობია და გვაინტერესებს, არის თუ არა ამ პოპულაციების საშუალოები (ან დისპერსიები, ან პროპორციები) ტოლი. ეს იქნება ძირითადი (ნულოვანი) ჰიპოთეზა, ხოლო ალტერნატივები აქაც შეიძლება იყოს ცალმხრივი (მარჯვენა და მარცხენა) ან ორმხრივი. ჩვენ ხშირად ჩავთვლით, რომ შესადარებელი პოპულაციები ერთმანეთისაგან დამოუკიდებელია, რაც იმას ნიშნავს, რომ ის შემთხვევითი სიდიდეები, რომლებიც ორივე შერჩევისათვის რიცხვითი მონაცემების უკან დგას, არიან დამოუკიდებლები.

ორი პოპულაციის მახასიათებლების შედარების ამოცანა ჩნდება ე. წ. კონტროლირებად ექსპერიმენტებში, როცა შეისწავლება რაიმე ახალი ტექნოლოგიის, მედიკამენტის, რეკლამის თუ სხვა ფაქტორების ზემოქმედების ეფექტურობის საკითხი. ამ ექსპერიმენტს უკავშირდება ორი შერჩევა, რომლებიდანაც ერთი მიღებულია კონსერვატიული (დამკვიდრებული) ქმედების, ხოლო მეორე – შეცვლილი (ახალი) ქმედების შედეგად. მაგალითად, ბანკმა შეცვალა მოლარეები სპეციალური აპარატებით და აინტერესებს თუ რა გავლენას მოახდენს ეს კლიენტების მომსახურებაზე. აქ ბუნებრივად ჩნდება ორი შერჩევა: ერთი შედგება იმ კლიენტების მომსახურების დროებისაგან, რომლებსაც ემსახურებოდნენ მოლარეები, ხოლო მეორე – იმ კლიენტების მომსახურების დროებისაგან, რომლებიც სარგებლობენ სპეციალური აპარატებით. შესაბამისად, კონტროლირებად ექსპერიმენტებში ერთი შერჩევა არის საკონტროლო (მიღებულია ადრინდელ პირობებში), ხოლო მეორე – ექსპერიმენტული. სწორედ ასეთ ამოცანებს უწოდებენ ორამოკრეფიან ამოცანებს.

ამით არ ამოიწურება ორამოკრეფიანი ამოცანების ჩამონათვალი. ხშირად შესადარებელია განსხვავებული სოციალური სტატუსის, სქესის, რეგიონის ან სხვა ფაქტორების მიხედვით განსხვავებული ორი პოპულაცია. განასხვავებენ ორ შემთხვევას: პირველი – გვაქვს ორი დამოუკიდებელი შერჩევა ორი პოპულაციიდან და მეორე – გვაქვს წყვილების შერჩევა, სადაც თვითონ წყვილები ერთმანეთისაგან დამოუკიდებელია, მაგრამ მათი

კომპონენტები შეიძლება არ იყვნენ დამოუკიდებლები (ასეთ მონაცემებს დაწყვილებული მომაცემები ეწოდება).

დაწყვილებული მონაცემები

განვიხილოთ ასეთი მაგალითი: დავუშვათ, ჩვენ გვაინტერესებს მოქმედებს თუ არა გარკვეული ტიპის კონტრაცეპტივის მიღება ქალის სისხლის წნევაზე. ამ მიზნით მოპოვებული მონაცემები წარმოდგენილია შემდეგი ცხრილის სახით:

i	წნევა კონტრაცეპტივის გამოყენებამდე (x_{i1})	წნევა კონტრაცეპტივის გამოყენების შემდეგ (x_{i2})	სხვაობა $d_i = x_{i2} - x_{i1}$
1	115	128	13
2	112	115	3
3	107	106	-1
4	119	128	9
5	115	122	7
6	138	145	7
7	126	132	6
8	105	109	4
9	104	102	-2
10	115	117	2

ვთქვათ, i-ური ქალის სისხლის წნევა კონტრაცეპტივის მიღებამდე განაწილებულია ნორმალურად საშუალოთი μ_i და დისპერსიით σ^2, ხოლო კონტრაცეპტივის მიღების შემდეგ განაწილებულია ისევ ნორმალურად საშუალოთი $\mu_i + \Delta$ და იმავე σ^2 დისპერსიით.

ნულოვანი ჰიპოთეზა ყალიბდება შემდეგნაირად: $H_0 : \Delta = 0$, რაც იმას ნიშნავს, რომ კონტრაცეპტივის მიღებას წნევა არ შეუცვლია, ხოლო ალტერნატივაა $H_1 : \Delta \neq 0$, რაც იმას ნიშნავს, რომ კონტრაცეპტივის მიღება ცვლის წნევის სიდიდეს ანუ მოქმედებს სისხლის წნევაზე. ცხადია, რომ კრიტერიუმის სტატისტიკა ორივე შერჩევას უნდა ეყრდნობოდეს, რადგან ცალ-ცალკე პოპულაციების პარამეტრები ჩვენთვის უცნობია. სამაგიეროდ, ჰიპოთეზის სამართლიანობის შემთხვევაში, ჩვენთვის ცნობილია, რომ შემთხვევითი სიდიდეები $Y_i := X_{i2} - X_{i1}$ $(i = 1, 2, ..., n)$ განაწილებულია ნორმალურად საშუალთი 0 და დისპერსიით, σ_Y^2, რომელიც მართალია არ ვიცით, მაგრამ რომლის შეფასებაც შეგვიძლია გაერთიანებული შერჩევიდან. ამრიგად, კრიტერიუმის სტატისტიკად შეგვიძლია ავიღოთ:

$$T_n \equiv \frac{\sqrt{n} \cdot \overline{Y}}{S_{n,Y}} \qquad (1)$$

შემთხვევითი სიდიდე, სადაც $\overline{Y} = \frac{1}{n} \cdot \sum_{i=1}^{n} Y_i$, ხოლო $S_{n,Y}$ აღნიშნავს სხვაობე-ბის შერჩევით სტანდარტულ გადახრას, ანუ

$$S_{n,Y} = \sqrt{\frac{1}{n-1} \cdot \left(\sum_{i=1}^{n} Y_i^2 - (1/n) \cdot \left(\sum_{i=1}^{n} Y_i \right)^2 \right)} \quad . \qquad (2)$$

ისევე როგორც ერთამოკრეფიან ამოცანაში (ნორმალური პოპულაცი-ის შესახებ ჰიპოთეზის შემოწმებისას უცნობი დისპერსიის შემთხვევაში), T_n შემთხვევით სიდიდეს ექნება $t(n-1)$ განაწილება (სტიუდენტის განაწი-ლება თავისუფლების ხარისხით $n-1$) და ამიტომ სტატისტიკური კრიტერ-იუმი შემდეგნაირად ჩამოყალიბდება:

თუ დასახელებული α-სათვის T_n შემთხვევით სიდიდის დაკვირვებუ-ლი მნიშვნელობა t_n (ანუ ის მნიშვნელობა, რომელიც მიიღება T_n შემთხვ-ევითი სიდიდის გამოსახულებაში Y_i-ების შეცვლისას d_i-ებით) აკმაყოფი-ლებს პირობას $-t_{n-1,\alpha/2} \le t_n \le t_{n-1,\alpha/2}$, მაშინ H_0 ჰიპოთეზის უარყოფის საფუძვ-ელი არა გვაქვს, წინააღმდეგ შემთხვევაში კი – მას უარვყოფთ.

დავუბრუნდეთ ჩვენს მაგალითს და გამოვთვალოთ T_n შემთხვევითი სიდიდის დაკვირვებული $t_n = t_{10}$ მნიშვნელობა. ცხადია, რომ

$$\overline{d}_{10} = (13+3-1+9+7+7+6+4+2-2)/10 = 4.8,$$

$$s_{n,Y} = \sqrt{\frac{13^2+3^2+(-1)^2+9^2+7^2+7^2+6^2+2^2+(-2)^2-10\cdot(4.8)^2}{9}} =$$

$$= \sqrt{20.844} = 4.566$$

და ამიტომ $t_{10} = 3.17 \cdot 4.8 / 4.566 = 3.32$. $t(9)$-განაწილების ზედა კრიტიკული წერტილების ცხრილებიდან კი, ვპოულობთ, რომ $t_{9,0.025}$ = 2.262. ვინაიდან $t_{10} = 3.32$ მნიშვნელობა არ ეკუთვნის [-2.262, 2.262] რიცხვით ინტერვალს, H_0 ჰიპოთეზას უარვყოფთ. მაშასადამე, შეგვიძლია გავაკეთოთ დასკვნა, რომ კონტრაცეპტივი მოქმედებს ქალის წნევაზე.

დავსვათ ასეთი კითხვა: როგორ ინტერვალშია მოთავსებული სხვაო-ბის ჭეშმარიტი საშუალო μ_Y? ნდობის ინტერვალის ასაგებად ისევ ვიყენ-ებთ იმავე ფაქტს, რომ $T_n \equiv \frac{\sqrt{n} \cdot \overline{Y}}{S_{n,Y}} \cong t(n-1)$ და ამ შემთხვევაში ნდობის ინტ-ერვალს აქვს შემდეგი სახე:

$$\left(\overline{d}_n - t_{n-1,\alpha/2} \cdot s_{n,Y} / \sqrt{n}\, ; \overline{d}_n - t_{n-1,\alpha/2} \cdot s_{n,Y} / \sqrt{n} \right). \qquad (3)$$

ამიტომ ჩვენს მაგალითში მივიღებთ, რომ:

(4.8-2.262·4.566/3.17, 4.8+2.262·4.566/3.17) = (1.53, 8.07).

მაშასადამე, წნევათა სხვაობის ჭეშმარიტი საშუალო μ_Y მნიშვნელ-ობა 0.95-ის ტოლი ალბათობით მოთავსებულია ინტერვალში 1.53 მმ ვწყ. სვ.-დან 8.07 მმ ვწყ.სვ.-მდე.

ორამოკრეფიანი t –კრიტერიუმი ტოლი, უცნობი დისპერსიების შემთხვევაში

ვთქვათ, მოცემულია ორი დამოუკიდებელი ნორმალურად განაწილებული პოპულაცია საშუალოებით μ_1 და μ_2 და ცნობილია, რომ მათ აქვთ ტოლი დისპერსიები $\sigma_1^2 = \sigma_2^2 = \sigma^2$, რომელთა საერთო σ^2 მნიშვნელობა უცნობია. ჩვენი ამოცანაა შევამოწმოთ $H_0 : \mu_1 = \mu_2$ ჰიპოთეზა, $H_1 : \mu_1 \neq \mu_2$ ალტერნატივის წინააღმდეგ. ცხადია, რომ $\bar{X}_{1n_1} - \bar{X}_{2n_2} \cong N(\mu_1 - \mu_2, \sigma^2 \cdot (1/n_1 + 1/n_2))$, სადაც n_1 და n_2 შესაბამისად, პირველი და მეორე შერჩევის მოცულობებია. ამიტომ H_0 ჰიპოთეზის სამართლიანობისას

$$\frac{\bar{X}_{1n_1} - \bar{X}_{2n_2}}{\sigma \cdot \sqrt{1/n_1 + 1/n_2}} \cong N(0,1) . \qquad (1)$$

მაგრამ, უკანასკნელი სიდიდე კრიტერიუმის სტატისტიკად არ გამოდგება, ვინაიდან ის შეიცავს უცნობ σ პარამეტრს. ამიტომ, როგორც წესი, იქცევიან შემდეგნაირად: (1) გამოსახულებაში σ-ს მაგივრად სვამენ მის შეფასებას და კრიტერიუმის სტატისტიკად იღებენ ასეთნაირად მიღებულ სიდიდეს. ამ შემთხვევაში ავტომატურად იბადება კითხვა: რომელი შერჩევიდან შევაფასოთ უცნობი σ პარამეტრი? ამ კითხვაზე პასუხის გასაცემად, გავიხსენოთ, რომ შერჩევის მოცულობის ზრდა იწვევს საშუალოს სტანდარტული შეცდომის შემცირებას. ამიტომ იქცევიან ასე: ვინაიდან, ძირითადი ჰიპოთეზის სამართლიანობის შემთხვევაში, ორივე პოპულაცია ერთნაირადაა განაწილებული, აერთიანებენ ამ შერჩევებს (რითაც იზრდება შერჩევის საერთო მოცულობა) და ისე აფასებენ σ-ს მნიშვნელობას, ანუ σ^2-ის შეფასების როლში აიღება შემდეგი სიდიდე:

$$S_{n_1,n_2}^2 = \frac{(n_1 - 1)S_{1n_1}^2 + (n_2 - 1)S_{1n_2}^2}{n_1 + n_2 - 2} . \qquad (2)$$

მტკიცდება, რომ ასე აგებული შეფასება წარმოადგენს σ^2–ის ჩაუნაცვლებელ შეფასებას (გავიხსენოთ, რომ ერთი შერჩევის შემთხვევაში შესწორებული შერჩევითი დისპერსია იყო დისპერსიის ჩაუნაცვლებელი შეფასება).

საბოლოოდ, კრიტერიუმის სტატისტიკას ექნება შემდეგი სახე

$$T_{n_1,n_2} \equiv \frac{\bar{X}_{1n_1} - \bar{X}_{2n_2}}{S_{n_1,n_2} \cdot \sqrt{1/n_1 + 1/n_2}} , \qquad (3)$$

რომლისთვისაც მტკიცდება, რომ

$$T_{n_1,n_2} := \frac{\bar{X}_{1n_1} - \bar{X}_{2n_2}}{S_{n_1,n_2} \cdot \sqrt{1/n_1 + 1/n_2}} \cong t(n_1 + n_2 - 2) . \qquad (4)$$

ამიტომ სტატისტიკური კრიტერიუმი შემდეგნაირად ყალიბდება:

თუ დასახელებული α-სათვის T_{n_1,n_2} შემთხვევით სიდიდის დაკვირვებული მნიშვნელობა t_{n_1,n_2} აკმაყოფილებს პირობას

$$-t_{n_1+n_2-2,\alpha/2} \le t_{n_1,n_2} \le t_{n_1+n_2-2,\alpha/2}, \qquad (5)$$

მაშინ H_0 ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს, წინააღმდეგ შემთხვევაში კი – მას უარვყოფთ.

რადგან კრიტერიუმის სტატისტიკის განაწილება ცნობილია, ჩვენ შეგვიძლია გამოვთვალოთ $1-\alpha$ ალბათობის მქონე ნდობის ინტერვალი $\mu_1-\mu_2$ სხვაობისათვის, რომელსაც შემდეგი სახე აქვს:

$$\left(\bar{x}_{1n_1} - \bar{x}_{1n_2} - t_{n_1+n_2-2,\alpha/2} \cdot s_{n_1,n_2} \cdot \sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \,;\, \bar{x}_{1n_1} - \bar{x}_{1n_2} + t_{n_1+n_2-2,\alpha/2} \cdot s_{n_1,n_2} \cdot \sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \right).$$

ორამოკრეფიანი $t-$კრიტერიუმი არატოლი დისპერსიების შემთხვევაში

ვთქვათ, მოცემულია ორი დამოუკიდებელი ნორმალურად განაწილებული პოპულაცია საშუალოებით μ_1 და μ_2 და ცნობილია, რომ მათ აქვთ არატოლი დისპერსიები $\sigma_1^2 \ne \sigma_2^2$. ჩვენი ამოცანაა შევამოწმოთ $H_0:\mu_1=\mu_2$ ჰიპოთეზა, $H_1:\mu_1\ne\mu_2$ ალტერნატივის წინააღმდეგ. ცნობილია, რომ

$$\bar{X}_{1n_1} - \bar{X}_{2n_2} \cong N(\mu_1-\mu_2, \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}),$$

სადაც n_1 და n_2 შესაბამისად, პირველი და მეორე შერჩევის მოცულობებია. აქედან გამომდინარე, თუ ცნობილია σ_1^2 და σ_2^2, მაშინ H_0 ჰიპოთეზის სამართლიანობის შემთხვევაში

$$\frac{\bar{X}_{1n_1} - \bar{X}_{2n_2}}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}} \cong N(0,1). \qquad (1)$$

ამიტომ, თუ ცნობილია σ_1^2 და σ_2^2, მაშინ (1) შემთხვევითი სიდიდე გამოდგება კრიტერიუმის სტატისტიკად. მაგრამ ეს სიდიდე კრიტერიუმის სტატისტიკად არ გამოგვადგება, როცა ის შეიცავს უცნობ σ_1^2 და σ_2^2 პარამეტრებს. ამ შემთხვევაში მათ მაგივრად (1) გამოსახულებაში ჩავსვათ მათი შეფასებები (შესაბამისი შერჩევითი სტანდარტული გადახრები) და კრიტერიუმის სტატისტიკად ავიღოთ შემდეგი შემთხვევითი სიდიდე:

$$T_{n_1,n_2} \equiv \frac{\bar{X}_{1n_1} - \bar{X}_{2n_2}}{\sqrt{S_{1n_1}^2/n_1+S_{2n_2}^2/n_2}}. \qquad (2)$$

როგორც ვიცით, კრიტერიუმის სტატისტიკა მხოლოდ გამოთვლადი სიდიდე კი არ უნდა იყოს, არამედ კრიტერიუმის ასაგებად, არსებითია მისი ზუსტი ან ასიმპტოტური განაწილების ცოდნაც. როგორაა განაწილებული T_{n_1,n_2} შემთხვევითი სიდიდე? მათემატიკურ სტატისტიკაში ეს პრობლემა ცნობილია ***ბერენს-ფიშერის პრობლემის*** სახელით. მისი გადაჭრის ერთ-ერთი, შედარებით მარტივი გზა ცნობილია ***სატერტვაიტის*** **(Satterthwaite)** ***მეთოდის*** სახელით, რომელიც შემდეგნაირად ყალიბდება:

თუ მოცემული α მნიშვნელოვნების დონისათვის T_{n_1,n_2} სტატისტიკის დაკვირვებული მნიშვნელობა t_{n_1,n_2} აკმაყოფილებს შემდეგ უტოლობას

$$-t_{[c],\alpha/2} \le t_{n_1,n_2} \le t_{[c],\alpha/2}, \qquad (3)$$

სადაც $[c]$ აღნიშნავს c რიცხვის მთელ ნაწილს (ანუ c-ს უახლოეს მთელ რიცხვს მარცხნიდან), ხოლო c გამოითვლება შემდგი წესით

$$c = \frac{(s_{1n_1}^2/n_1 + s_{2n_2}^2/n_2)^2}{(s_{1n_1}^2/n_1)^2/(n_1-1) + (s_{2n_2}^2/n_2)^2/(n_2-1)}, \qquad (4)$$

მაშინ α მნიშვნელოვნების დონით $H_0: \mu_1 = \mu_2$ ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს, წინააღმდეგ შემთხვევაში, მას უარვყოფთ.

აქვე აღვნიშნავთ, რომ $1-\alpha$ ალბათობის მქონე ნდობის ინტერვალს ექნება შემდეგი სახე:

$$\left(\bar{x}_{1n_1} - \bar{x}_{1n_2} - t_{[c],\alpha/2} \cdot \sqrt{\frac{s_{1n_1}^2}{n_1} + \frac{s_{2n_2}^2}{n_2}}\,;\, \bar{x}_{1n_1} - \bar{x}_{1n_2} + t_{[c],\alpha/2} \cdot \sqrt{\frac{s_{1n_1}^2}{n_1} + \frac{s_{2n_2}^2}{n_2}}\right).$$

შერჩევათა მოცულობების განსაზღვრა. ორი პოპულაციის საშუალოების შედარების კრიტერიუმის სიმძლავრე

შერჩევათა მოცულობების განსაზღვრა საჭიროა ექსპერიმენტის წინასწარ დაგეგმვისათვის. ქვემოთ ჩვენ მოვიყვანთ ორამკრეფიანი ამოცანებისათვის შერჩევათა მოცულობების გამოსათვლელ ფორმულებს, რომლებიც საჭიროა პოპულაციათა საშუალოების შედარების ამოცანაში (მოცემული α მნიშვნელოვნების დონით) დასახელებული $1-\beta$ სიმძლავრის მისაღწევად ორმხრივი ალტერნატივების დროს.

დავუშვათ, რომ მოცემულია ორი ნორმალური პოპულაცია საშუალოებით μ_1 და μ_2 და ცნობილი დისპერსიებით σ_1^2 და σ_2^2. შესამოწმებელია $H_0: \mu_1 = \mu_2$ ჰიპოთეზა, $H_1: \mu_1 \ne \mu_2$ ალტერნატივის წინააღმდეგ. ჩვენი ამოცანაა, განვსაზღვროთ პოპულაციათა ის n_1 და $n_2 = k \cdot n_1$ მოცულობები, რომელთათვისაც α მნიშვნელოვნების დონით მიიღწევა მოცემული $1-\beta$ სიმძლავრე. მოცულობების გამოსათვლელ ფორმულებს აქვთ შემდეგი სახე:

$$n_1 \ge (\sigma_1^2 + \sigma_2^2/k) \cdot (z_{1-\alpha/2} + z_{1-\beta})^2/(\mu_2 - \mu_1)^2 \qquad (1)$$

და

$$n_2 \ge (k \cdot \sigma_1^2 + \sigma_2^2) \cdot (z_{1-\alpha/2} + z_{1-\beta})^2/(\mu_2 - \mu_1)^2. \qquad (2)$$

შევნიშნოთ, რომ $k=1$-ის შემთხვევაში $n_2 = n_1$ და (1) და (2) შეფასებებიც იძლევა ერთსა და იმავე შედეგს.

კრიტერიუმის სიმძლავრის გამოსათვლელად $H_0: \mu_1 = \mu_2$ ჰიპოთეზის წინააღმდეგ განვიხილოთ სპეციფიკური ალტერნატივა: კერძოდ, ალტერნატივა $H_1: \mu_2 = \mu_1 + \Delta$. მაშინ ორმხრივი ალტერნატივისათვის სიმძლავრე გამოითვლება შემდეგი თანაფარდობიდან:

$$1-\beta=\Phi\left(-z_{\alpha/2}+\frac{\Delta\cdot\sqrt{n_1}}{\sqrt{\sigma_1^2+\sigma_2^2/k}}\right), \qquad (3)$$

სადაც $k=n_2/n_1$.

§40. ორამოკრეფიანი ამოცანები ბერნულის სქემაში

ჰიპოთეზათა შემოწმება წარმატებათა ალბათობებისათვის ბერნულის ცდათა ორი დამოუკიდებელი მიმდევრობისათვის

$X_1,\dots,X_n$ და $Y_1,\dots,Y_m$ ორი დამოუკიდებელი შერჩევაა ბერნულის კანონით განაწილებული პოპულაციიდან შესაბამისად წარმატების უცნობი p_1 და p_2 ალბათობებით; $q_1=1-p_1$, $q_2=1-p_2$; $S_1=\sum_{i=1}^{n}X_i$, $S_2=\sum_{i=1}^{m}Y_i$; $\overline{P}_1=\frac{S_1}{n}$, $\overline{P}_2=\frac{S_2}{m}$; $\overline{p}_1=\frac{s_1}{n}=\frac{1}{n}\cdot\sum_{i=1}^{n}x_i$, $\overline{p}_2=\frac{s_2}{m}=\frac{1}{m}\cdot\sum_{i=1}^{m}y_i$; $\overline{P}=\frac{n}{n+m}\overline{P}_1+\frac{m}{n+m}\overline{P}_2$; $\overline{p}=\frac{n}{n+m}\overline{p}_1+\frac{m}{n+m}\overline{p}_2$; $\overline{Q}=1-\overline{P}$; $\overline{q}=1-\overline{p}$.

კრიტერიუმი:

ორმხრივი	მარჯვენა ცალმხრივი	მარცხენა ცალმხრივი
$H_0: p_1=p_2$	$H_0: p_1=p_2$ ან $H_0: p_1\le p_2$	$H_0: p_1=p_2$ ან $H_0: p_1\ge p_2$
$H_1: p_1\ne p_2$	$H_1: p_1>p_2$	$H_1: p_1<p_2$

ჰიპოთეზა: $H_0: p_1=p_2$

მნიშვნელოვნების დონე: α

კრიტერიუმის სტატისტიკა: $Z=\dfrac{(\overline{P}_1-\overline{P}_2)-(p_1-p_2)}{\sqrt{\overline{P}\,\overline{Q}(1/n+1/m)}}\overset{as}{\cong}N(0,1)$

კრიტერიუმის მნიშვნელობა T.V.: $z=\dfrac{(\overline{p}_1-\overline{p}_2)-(p_1-p_2)}{\sqrt{\overline{p}\,\overline{q}(1/n+1/m)}}$

ალტერნატივა	კრიტიკული არე C.R. (H_0-ის უარყოფის არე)
$H_1: p_1>p_2$	$z\ge z_\alpha$,
$H_1: p_1<p_2$	$z\le -z_\alpha$,
$H_1: p_1\ne p_2$	$z\le -z_{\alpha/2}$ ან $z\ge z_{\alpha/2}$

(სადაც z_α არის $N(0,1)$-ის ზედა α-კრიტიკული წერტილი ანუ **კრიტიკული მნიშვნელობა C.V.**).

გადაწყვეტილება: თუ $z\in$ **C.R.**, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

$(1-\alpha)$ **საიმედოობის ნდობის ინტერვალი $p_1 - p_2$ სხვაობისათვის:**

$$((\overline{p}_1 - \overline{p}_2) - z_{\alpha/2}\sqrt{\frac{\overline{p}_1\overline{q}_1}{n} + \frac{\overline{p}_2\overline{q}_2}{m}}, (\overline{p}_1 - \overline{p}_2) + z_{\alpha/2}\sqrt{\frac{\overline{p}_1\overline{q}_1}{n} + \frac{\overline{p}_2\overline{q}_2}{m}})$$

(სადაც $\overline{q}_1 = 1 - \overline{p}_1$, $\overline{q}_2 = 1 - \overline{p}_2$).

შეზღუდვები: $np_1, nq_1, mp_2, mq_2 \geq 5$.

მაგალითი. მკვლევარმა დაადგინა, რომ 34 მცირე კლინიკიდან 12-ში, ისევე როგორც 24 დიდი კლინიკიდან 17-ში, დაკავებული საწოლების წილი ნაკლებია 80%-ზე. $\alpha = 0.05$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ პროპორციები იმ მცირე და დიდ კლინიკების რომლებშიც დაკავებული საწოლების წილი ნაკლებია 80%-ზე ერთიდაიგივეა. ააგეთ 95%-იანი ნდობის ინტერვალი პროპორციათა სხვაობისათვის.

ამოხსნა. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0 : p_1 = p_2$, $H_1 : p_1 \neq p_2$. აღვნიშნოთ $\overline{p}_1$ (შესაბამისად, $\overline{p}_2$) სიმბოლოთი პროპორცია იმ მცირე (შესაბამისად, იმ დიდი) კლინიკების რომლებშიც დაკავებული საწოლების წილი ნაკლებია 80%-ზე. გვაქვს:

$$\overline{p}_1 = \frac{s_1}{n} = \frac{12}{34} = 0.35 \quad \text{და} \quad \overline{p}_2 = \frac{s_2}{m} = \frac{17}{24} = 0.71.$$

ამიტომ $\overline{p} = \dfrac{n}{n+m}\overline{p}_1 + \dfrac{m}{n+m}\overline{p}_2 = \dfrac{34}{34+24}\cdot 0.35 + \dfrac{24}{34+24}\cdot 0.71 = 0.5$, $\overline{q} = 1 - \overline{p} = 0.5$.

ვიპოვოთ კრიტიკული მნიშვნელობა. რადაგანაც კრიტერიუმი ორმხრივია და $\alpha = 0.05$, კრიტიკული მნიშვნელობები იქნება: $C.V. = \pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96$.

გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$T.V. = z = \frac{(\overline{p}_1 - \overline{p}_2) - (p_1 - p_2)}{\sqrt{\overline{p}\overline{q}(1/n + 1/m)}} = \frac{(0.35 - 0.71) - 0}{\sqrt{0.5 \cdot 0.5 \cdot (1/34 + 1/24)}} = -2.7 .$$

ვინაიდან -2.7 < -1.96 (ე.ი. $T.V. \in C.R.$), ამიტომ ჩვენ უნდა უკუვაგდოთ ნულოვანი ჰიპოთეზა, ანუ ჩვენ გვაქვს საკმარისი საფუძველი უარვყოთ ჰიპოთეზა იმის შესახებ, რომ პროპორციები არ განსხვავდება.

ავაგოთ ნდობის ინტერვალი. აქ $\overline{q}_1 = 1 - \overline{p}_1 = 0.65$, $\overline{q}_2 = 1 - \overline{p}_2 = 0.29$. ამიტომ საძიებელი ნდობის ინტერვალი იქნება:

$$((0.35 - 0.71) - 1.96 \cdot \sqrt{\frac{0.35\cdot 0.65}{34} + \frac{0.71 \cdot 0.29}{24}}, (0.35 - 0.71) + 1.96 \cdot \sqrt{\frac{0.35\cdot 0.65}{34} + \frac{0.71 \cdot 0.29}{24}}),$$

$$(-0.36 - 0.242, -0.36 + 0.242),$$

$$(-0.602, -0.118).$$

ვინაიდან ნდობის ინტერვალი მოიცავს 0-ს, გადაწყვეტილება ისევ იქნება: უკუვაგდოთ ნულოვანი ჰიპოთეზა.

§41. ჰიპოთეზის შემოწმება დისპერსიების ტოლობის შესახებ

დისპერსიების შესახებ ჰიპოთეზები ძალიან მნიშვნელოვან როლს თამაშობენ ეკონომიკურ-მათემატიკური მოდელირებისას, ვინაიდან ექსპერიმენტული შერჩევითი მონაცემების გაბნევის სიდიდე შესაბამისი პარამეტრების გათვლილი მნიშვნელობებიდან, რომელიც ხასიათდება დისპერსიით, შესაძლებლობას გვაძლევს გადავწყვიტოთ იმ მოდელის გამოსადეგობა (ადექვატურობა), რომლის საფუძველზეც იგება თეორია.

დავუშვათ, რომ ნორმალური კანონით განაწილებული ξ შემთხვევითი სიდიდე განსაზღრულია გარკვეულ სიმრავლეზე, რომელიც ქმნის გენერალურ ერთობლიობას, ხოლო ნორმალური კანონით განაწილებული η შემთხვევითი სიდიდე განმარტებულია სხვა სიმრავლეზე, რომელიც აგრეთვე შეადგენს გენერალურ ერთობლიობას. ორივე ერთობლიობიდან კეთდება შერჩევა: პირველიდან -- n_1 მოცულობის მქონე, ხოლო მეორედან -- n_2 მოცულობის მქონე (შევნიშნავთ, რომ შერჩევის მოცულობა ყოველთვის არ შეიძლება თავიდანვე განსაზღრული იყოს, მაგალითად, იმ შემთხვევაში, თუ შერჩევა არის ბადეში მოხვედრილი თევზები). თითოეული შერჩევისათვის გამოითვლება შესწორებული შერჩევითი დისპერსია: s_1^2 -- შერჩევი-სათვის პირველი ერთობლიობიდან და s_2^2 -- შერჩევისათვის მეორე ერთობლიობიდან.

ამოცანა მდგომარეობს შემდეგში: შერჩევითი მონაცემების საშუალებით შევამოწმოთ სტატისტიკური ჰიპოთეზა $H_0 : D\xi = D\eta$. ალტერნატიული ჰიპოთეზის როლში განვიხილავთ იდეას, რომელიც მდგომარეობს იმაში, რომ იმ ერთობლიობის დისპერსია, რომლის შესწორებული შერჩევითი დისპერსია აღმოჩნდა უდიდესი, მეტია ვიდრე მეორე ერთობლიობის დისპერსია. განიხილება შემდეგი სახის კრიტერიუმი:

$$F = S^{**} / S^{*},$$

სადაც S^{**}-- უდიდესია s_1^2 და s_2^2 შეფასებებს შორის, ხოლო S^{*}-- კი მათ შორის უმცირესი.

ცნობილია, რომ F კრიტერიუმი განაწილებულია ფიშერის განაწილების კანონით თავისუფლების k_1 და k_2 ხარისხებით, სადაც:

$$k_1 = n_1 - 1, \quad k_2 = n_2 - 1, \text{ თუ } S^{**} = s_1^2;$$

$$k_1 = n_2 - 1, \quad k_2 = n_1 - 1, \text{ თუ } S^{**} = s_2^2.$$

ამ ამოცანაში ბუნებრივია განვიხილოთ მარჯვენა კრიტიკული არე, ვინაიდან F კრიტერიუმის საკმარისად დიდი შერჩევითი მნიშვნელობა მეტყველებს ალტერნატიული ჰიპოთეზის სასარგებლოდ.

მოცემული მნიშვნელოვნების დონისათვის q (ჩვეულებრივ, $q = 0.05$ ან $q = 0.01$) კრიტიკული მნიშვნელობა $F_{კრ}$ განისაზღვრება ფიშერის განაწილების ცხრილიდან. *იმ შემთხვევაში, როცა $F > F_{კრ}$ ხდება H_0 ჰიპოთე-ზის უარყოფა, ხოლო როცა $F < F_{კრ}$ -- H_0 ჰიპოთეზა მიიღება.*

დავუშვათ, რომ გარკვეული ობიექტების ორი სიმრავლე, რომელთაც გააჩნიათ რაოდენობრივი ნიშანი, ექვემდებარება შერჩევით კონტროლს. რაოდენობრი-

ვი ნიშნის მნიშვნელობები არიან ნორმალური განაწილების კანონით განაწილ-ებული შემთხვევითი სიდიდეები, რომელთაც ჩვენ ავღნიშნავთ ξ_1-ით და ξ_2-ით შესაბამისად, პირველი და მეორე სიმრავლეებისათვის. პირველი სიმრავლიდან გაკეთებულია $n_1 = 21$ მოცულობის შერჩევა და ნაპოვნია შესწორებული შერჩე-ვითი დისპერსია, რომელიც აღმოჩნდა 0.75-ის ტოლი. მეორე სიმრავლიდან გაკ-ეთებულია $n_2 = 11$ მოცულობის შე-რჩევა. მისი შესწორებული შერჩევითი დისპერ-სიაა 0.25. ვაყენებთ ჰიპოთეზას: $H_0 : D\xi_1 = D\xi_2$. ალტერნატიული ჰიპოთეზა მდგო-მარეობს იმასი, რომ $H_1 : D\xi_1 > D\xi_2$. ამ შემთხვევაში ფიშერის კრიტერიუმის შერ-ჩევითი მნიშვნე-ლობა $F_{ფ} = 3$. არჩეული $q = 0.05$ მნიშვნელოვნების დონისათვის, თავისუფ-ლების $k_1 = 20$ და $k_2 = 10$ ხარისხით, ფიშერის განაწილების ცხრილიდან ვპოულობთ, რომ $F_{კრ} = 2.77$. ვინაიდან, $F_{ფ} > F_{კრ}$, ჰიპოთეზა დისპერსიების ტოლო-ბის შესახებ უნდა უკუგდებულ იქნეს.

$(1-\alpha)$ **საიმედოობის ნდობის ინტერვალი დისპერსიათა σ_1^2 / σ_2^2 ფარდო-ბისათვის**

$$\overline{F}_{ფ} \cdot F_{k_2, k_1 .1-\alpha/2} = (\overline{F}_{ფ} / F_{k_1, k_2, \alpha/2}) < \sigma_1^2 / \sigma_2^2 < \overline{F}_{ფ} \cdot F_{k_1, k_2, \alpha/2} .$$

მაგალითი. მკვლევარს აინტერესებს არის თუ არა განსხვავება მწეველი და არამწეველი ადამიანების პულსის რიცხვთა დისპერსიებს შორის. შემთხვევით შერჩეული 26 მწეველისაგან და 18 არამწეველისაგან შემდგარი ორი შერჩევის შესაბამისი შესწორებული შერჩევითი დისპერსიებია: $s_1^{'2} = 36$, $s_2^{'2} = 10$. $\alpha = 0.05$ მნიშვნელოვნების დონით გვაქვს თუ არა საკმარისი საფუძველი დავასკვნათ, რომ დისპერსიები განსხვავებულია? ჩათვალეთ, რომ სიდიდე განაწილებულია ნორმალურად.

ამოხსნა. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0 : \sigma_1^2 = \sigma_2^2$, $H_1 : \sigma_1^2 \neq \sigma_2^2$. ვისარგებლოთ გამარტივებული პროცედურით. ვინაიდან კრიტერიუმი ორმხრივია, $s_1^{'2} > s_2^{'2}$ და თავისუფლების ხარისხებია 26 – 1 = 25 , 18 – 1 =17, ამიტომ კრიტიკული მნიშვნელობა იქნება: $C.V. = F_{k,l,\alpha/2} = F_{25,17,0.025} = 2.56$. გამოვთვალოთ კრიტერიუმის მნიშვნელობა: $T.V. = s_1^{'2} / s_2^{'2} = 36/10 = 3.6$. რადგანაც 3.6 > 2.56, ამიტომ ძირითადი ჰიპოთეზა უნდა უკუვაგდოთ, ანუ ჩვენ გვაქვს საკმ-არისი საფუძველი დავასკვნათ, რომ დისპერსიები განსხვავებულია.

§42. ჰიპოთეზის შემოწმება შერჩევითი კორელაციის კოეფიციენტის სტატისტიკური მნიშვნელოვნების შესახებ

გენერალური ერთობლიობის Δ პარამეტრის შერჩევითი δ შეფასების სტატისტიკური მნიშვნელოვნების შემოწმება ეწოდება $H_0 : \Delta = 0$ სტატისტიკური ჰიპოთეზის შემოწმებას ალტერნატიული $H_1 : \Delta \neq 0$ ჰიპოთეზის წინააღმდეგ. თუ მოხდება H_0 ჰიპოთეზის უარყოფა, მაშინ δ შეფასება ითვლება სტატისტიკურად მნიშვნელოვნად.

დავუშვათ, რომ მოცემულია ერთი და იგივე გენერალური ერთობლიობის ობიექტთა სიმრავლეზე განსაზღრული ნორმალური კანონით განაწილებული ორი შემთხვევითი სიდიდე ξ და η. ჩვენი მიზანია შევამოწმოთ სტატისტიკური ჰიპოთეზა ξ და η შემთხვევით სიდიდეებს შორის კორელაციური კავშირის არ არსებობის შესახებ:

$H_0 : \rho(\xi,\eta) = 0$; $H_1 : \rho(\xi,\eta) \neq 0$.

ვატარებთ n მოცულობის შერჩევას და გამოითვლება შერჩევითი კორელაციის კოეფიციენტი r. სტატისტიკური კრიტერიუმის როლში განიხილება შემთხვევითი სიდიდე

$$t = r\sqrt{n-2}\,/\sqrt{1-r^2}\ ,$$

რომელიც განაწილებულია სტიუდენტის განაწილების კანონით თავისუფლების ხარისხით $n-2$.

შევნიშნავთ, რომ შერჩევითი კორელაციის r კოეფიციენტის ყველა შესაძლო მნიშვნელობა მოთავსებულია [-1,1] ინტერვალში. გასაგებია, რომ t სიდიდის შედარებით დიდი გადახრები ნულიდან ნებისმიერ მხარეს მიიღება შედარებით დიდი, ანუ მოდულით 1-თან ახლოს მდგომი r-ის მნიშვნელობებისათვის. ვინაიდან, მოდულით 1-თან ახლოს მდგომი r-ის მნიშვნელობები ეწინააღმდეგებიან H_0 ჰიპოთეზას, ამიტომ ბუნებრივია, რომ აქ განვიხილოთ ორმხრივი კრიტიკული არე t კრიტერიუმისათვის.

მნიშვნელოვნების α დონისა და თავისუფლების ხარისხის $n-2$ რიცხვის მიხედვით სტიუდენტის განაწილების ცხრილიდან ვპოულობთ კრიტიკულ მნიშვნელობას $t_{კრ}$. თუ კრიტერიუმის შერჩევითი მნიშვნელობის $t_{უ}$ მოდული აღემატება $t_{კრ}$-ს, მაშინ ხდება H_0 ჰიპოთეზის უარყოფა და კორელაციის შერჩევითი კოეფიციენტი ითვლება სტატისტიკურად მნიშვნელოვნად. წინააღმდეგ შემთხვევაში, ანუ თუ $|t_{უ}| < t_{კრ}$, მიიღება H_0 ჰიპოთეზა და კორელაციის შერჩევითი კოეფიციენტი ითვლება სტატისტიკურად არა მნიშვნელოვნად.

§43. თანხმობის კრიტერიუმები. ხი კვადრატ კრიტერიუმი

აქამდე ჩვენ ვიხილავდით ჰიპოთეზებს, რომლებშიც გენერალური ერთობლიობის განაწილების კანონი ითვლებოდა, რომ იყო ცნობილი. ახლა ჩვენ შევუდგებით ჰიოპოთეზების შემოწმებას უცნობი განაწილების კანონის სავარაუდო სახის შესახებ, ე. ი. შევამოწმებთ ნულოვან ჰიპოთეზას იმის შესახებ, რომ გენერალური ერთობლიობა განაწილებულია გარკვეული ცნობილი კანონის მიხედვით. ასეთი ჰიპოთეზების შემოწმების სტატისტიკურ კრიტერიუმებს, ჩვეულებრივ, **თანხმობის კრიტერიუმებს** უწოდებენ.

პირსონის კრიტერიუმი (ხი კვადრატ კრიტერიუმი). პირსონის კრიტერიუმის საშუალებით შესაძლებელია სხვადასხვა განაწილების კანონის შესახებ ჰიპოთეზების შემოწმება.

I. ჰიპოთეზის შემოწმება განაწილების ნორმალურობის შესახებ.

ვიგულისხმოთ, რომ მიღებულია საკმარისად დიდი n მოცულობის შერჩევა განსხვავებული ვარიანტების დიდი რიცხვით. მისი დამუშავების მოხერხებულობის მიზნით ვარიანტების უმცირესი მნიშვნელობიდან უდიდეს მნიშვნელობამდე ინტერვალი დავყოთ s ტოლ ნაწილად და ჩავთვალოთ, რომ ვარიანტების მნიშვნელობები, რომლებიც მოხვდენენ ცალკეულ ინტერვალში დაახლოებით ტოლია ამ ინტერვალის შუაწერტილის მომცემი რიცხვის. დავთვალოთ თოთოეულ ინტერვალში მოხვედრილი ვარიანტების რაოდენობა და შევადგინოთ ე. წ. დაჯგუფებული შერჩევა

ვარიანტები	x_1	x_2	$\cdots$	x_n
სიხშირე	n_1	n_2	$\cdots$	n_k

სადაც x_i – ინტერვალის შუაწერტილის მნიშვნელობაა, ხოლო n_i – ვარიანტების რიცხვია, რომლებიც მოხვდნენ i –ურ ინტერვალში (ემპირიული სიხშირეები).

მიღებული მონაცემებით გამოვთვალოთ შერჩევითი საშუალო $\bar{x}_შ$ და შერჩევითი საშუალო კვადრატული გადახრა $\sigma_შ$. შევამოწმოთ წინადადება, რომ გენერალური ერთობლიობა განაწილებულია ნორმალური კანონით პარამატრებით $E\xi = \bar{x}_შ$ და $D\xi = \sigma_შ$. მაშინ ჩვენ შეგვიძლია დავითვალოთ რიცხვების რაოდენობა n მოცულობის შერჩევიდან, რამდენიც უნდა აღმოჩნდეს თითოეულ ინტერვალში ამ დაშვების დროს (ე. ი. თეორიული სიხშირეები). ამ მიზნით, ნორმალური განაწილების ფუნქციის ცხრილიდან ვპოულობთ i –ურ ინტერვალში მოხვედრის ალბათობას:

$$p_i = \Phi\left(\frac{b_i - \bar{x}_შ}{\sigma_შ}\right) - \Phi\left(\frac{a_i - \bar{x}_შ}{\sigma_შ}\right)$$

სადაც a_i და b_i -- i -ური ინტერვალის საზღვრებია. მიღებული ალბათობების შერჩევის მოცულობაზე გამრავლებით ვპოულობთ თეორიულ სიხშირე-

ებს: $n_i' = n \cdot p_i$. ჩვენი მიზანია – შევადაროთ ემპირიული და თეორიული სიხშირეები, რომლებიც, რა თქმა უნდა, განსხსხვავდებიან ერთმანეთისაგან, და გავარკვიოთ, არიან თუ არა ეს განსხვავებები არაარსებითი, რომლებიც არ უარყოფენ ჰიპოთეზას გამოსაკვლევი შემთხვევითი სიდიდის ნორმალური განაწილების შესახებ, ან ეს განსხვავებები იმდენად დიდია, რომ ეწინააღმდეგებიან ამ ჰიპოთეზას. ამ მიზნით გამოიყენება კრიტერიუმი შემდეგი შემთხვევითი სიდიდის სახით

$$\chi^2 = \sum_{i=1}^{s} \frac{(n_i - n_i')^2}{n_i'}. \quad (1)$$

შეზღუდვები: ყველა $n_i' \geq 5$.

ამ კრიტერიუმის აღების აზრი შემდეგში მდგომარეობს: იკრიბება ის წილები, რასაც შეადგენს ემპირიული სიხშირეების თეორიული სიხშირეებისაგან გადახრის კვადრატები, შესაბამისი თეორიული სიხშირეებისაგან. შეიძლება დამტკიცდეს, რომ გენერალური ერთობლიობის რეალური განაწილების კანონისაგან დამოუკიდებლად (1) შემთხვევითი სიდიდის განაწილების კანონი უახლოვდება (მიისწრაფის) χ^2 განაწილებისაკენ თავისუფლების ხარისხით $k = s - 1 - r$, (როცა $n \to \infty$), სადაც r – შერჩევის მონაცემებით შესაფასებული სავარაუდო განაწილების პარამეტრების რაოდენობაა.

ნორმალური განაწილება ხასიათდება ორი პარამეტრით, ამიტომ $k = s - 3$. არჩეული კრიტერიუმისათვის იგება მარჯვენა ცალმხრივი კრიტიკული არე, რომელიც განისაზღვრება უტოლობით

$$p(\chi^2 > \chi^2_{კრ}(\alpha, k)) = \alpha, \quad (2)$$

სადაც α –მნიშვნელოვნების დონეა. შესაბამისად, კრიტიკული არე მოიცემა უტოლობით $\chi^2 > \chi^2_{კრ}(\alpha, k)$, ხოლო ჰიპოთეზის მიღების არეა -- $\chi^2 < \chi^2_{კრ}(\alpha, k)$.

ამრიგად, იმისათვის, რომ შევამოწმოთ ნულოვანი ჰიპოთეზა H_0: გენერალური ერთობლიობა განაწილებულია ნორმალურად – უნდა გამოვთვალოთ შერჩევის მიხედვით კრიტერიუმიდს დაკვირვებული მნიშვნელობა:

$$\chi^2_{ფ} = \sum_{i=1}^{s} \frac{(n_i - n_i')^2}{n_i'}, \quad (3)$$

ხოლო χ^2 განაწილების კრიტიკული წერტილების ცხრილიდან ვიპოვოთ კრიტიკული წერტილი $\chi^2_{კრ}(\alpha, k)$ ცნობილი α და $k = s - 3$ მნიშვნელობებისათვის. თუ აღმოჩნდა, რომ $\chi^2_{ფ} < \chi^2_{კრ}(\alpha, k)$ -- ვღებულობთ ნულოვან ჰიპოთეზას, თუ $\chi^2_{ფ} > \chi^2_{კრ}(\alpha, k)$, მაშინ – უკუვაგდებთ.

II. ჰიპოთეზის შემოწმება თანაბარი განაწილების შესახებ.

პირსონის კრიტერიუმის გამოყენებისას გენერალური ერთობლიობის თანაბარი განაწილების შესახებ ჰიპოთეზის შემოწმებისას სავარაუდო განაწილების სიმკვრივით

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & x\in(a,b)\\ 0, & x\notin(a,b)\end{cases}$$

აუცილებელია არსებული შერჩევის მიხედვით გამოვთვალოთ შერჩევითი საშუალო $\bar{x}_{შ}$ და შევაფასოთ a და b პარამეტრები ფორმულებით:

$$a^*=\bar{x}_B-\sqrt{3}\sigma_B,\ \ b^*=\bar{x}_B+\sqrt{3}\sigma_B, \qquad (4)$$

სადაც a^* და b^* -- a-სა და b-ს შეფასებებია. მართლაც, ვინაიდან თანაბარი განაწილებისათვის:

$$E\xi=\frac{a+b}{2},\quad \sigma(x)=\sqrt{D(X)}=\sqrt{\frac{(a-b)^2}{12}}=\frac{a-b}{2\sqrt{3}},$$

აქედან შეგვიძლია მივიღოთ განტოლებათა სისტემა a^*-სა და b^*-სათვის:

$$\begin{cases}\dfrac{b^*+a^*}{2}=\bar{x}_B\\ \dfrac{b^*-a^*}{2\sqrt{3}}=\sigma_B\end{cases}$$

რომლის ამოხსნასაც წარმოადგენს სწორედ (4) გამოსახულებები.

შემდეგ, ვუშვებთ, რომ $f(x)=\dfrac{1}{b^*-a^*}$ და ვპოულობთ თეორიულ სიხშირეებს ფორმულებიდან:

$$n_1'=np_1=nf(x)(x_1-a^*)=n\cdot\frac{1}{b^*-a^*}(x_1-a^*);$$

$$n_2'=n_3'=...=n_{s-1}'=n\cdot\frac{1}{b^*-a^*}(x_i-x_{i-1}),\ \ i=1,2,...,s-1;$$

$$n_s'=n\cdot\frac{1}{b^*-a^*}(b^*-x_{s-1}).$$

აქ s – იმ ინტერვალების რიცხვია, რამდენ ინტერვალადაც გაიყო შერჩევა.

პირსონის კრიტერიუმის დაკვირვებული მნიშვნელობა გამოითვლება (3) ფორმულიდან, ხოლო კრიტიკული წერტილი $\chi^2_{კრ}(\alpha,k)$ -- χ^2 განაწილების კრიტიკული წერტილების ცხრილიდან თავისუფლების ხარისხის $k=$ $=s-3$ რიცხვის გათვალისწინებით. ამის შემდეგ ვიქცევით ისე. როგორც წინა შემთხვევაში. კერძოდ, თუ აღმოჩნდა, რომ $\chi^2_{შ}<\chi^2_{კრ}(\alpha,k)$ -- ვღებულობთ ნულოვან ჰიპოთეზას, თუ $\chi^2_{შ}>\chi^2_{კრ}(\alpha,k)$, მაშინ – უკუვაგდებთ.

III. ჰიპოთეზის შემოწმება მაჩვენებლიანი (ექსპონენციალური) განაწილების შესახებ.

ამ შემთხვევაში მოცემულ შერჩევას ვყოფთ თანაბარი სიგრძის ინტერვალებად და ვიხილავთ ვარიანტების მიმდევრობას $x_i^*=\dfrac{x_i+x_{i+1}}{2}$, რომლებიც თანაბრად დაშორებული არიან ერთმანეთისაგან (ითვლება, რომ ყვე-

ლა ვარიანტი, რომელიც მოხვდა i –ურ ინტერვალში ღებულობს მნიშვნელობას, რომელიც ემთხვევა ამ ინტერვალის შუაწერტილს), და შესაბამისი n_i სიხშირეების მიმდევრობას (i –ურ ინტერვალში მიხვედრილი ვარიანტების რიცხვი). ამ მონაცემებით გამოვთვალოთ შერჩევითი საშუალო $\bar{x}_ვ$ და მივიღოთ λ პარამეტრის შეფასებად $\lambda^* = \frac{1}{\bar{x}_B}$. მაშინ თეორიული სიხშირეები გამოითვლება ფორმულით

$$n_i' = n_i p_i = n_i p(x_i < X < x_{i+1}) = n_i (e^{-\lambda x_i} - e^{-\lambda x_{i+1}}).$$

შემდეგ პირსონის კრიტერიუმის დაკვირვებული მნიშვნელობა გამოითვლება (3) ფორმულიდან, ხოლო კრიტიკული წერტილი $\chi^2_{კრ}(\alpha,k)$ -- χ^2 განაწილების კრიტიკული წერტილების ცხრილიდან თავისუფლების ხარისხის $k = s-2$ რიცხვის გათვალისწინებით.

თუ აღმოჩნდა, რომ $\chi^2_ვ < \chi^2_{კრ}(\alpha,k)$ -- ვღებულობთ ნულოვან ჰიპოთეზას, თუ $\chi^2_ვ > \chi^2_{კრ}(\alpha,k)$, მაშინ – უკუვაგდებთ.

IV. ჰიპოთეზის შემოწმება ბინომიალური განაწილების შესახებ.

ჩვენი მიზანია შევამოწმოთ ჰიპოთეზა იმის შესახებ, რომ გენერალური ერთობლიობა განაწილებულია ბინომიალური კანონით $Bi(N,p)$. ავღნიშნოთ ν_i-თი იმ x-ების რაოდენობა $x_1,\dots,x_n$ შერჩევიდან, რომელთათვისაც $x=i,\quad i=0,1,\dots,N$. ჰიპოთეზის სამართლიანობის შემთხვევაში:

$$p_i = P(X_j = i) = C_N^i p^i (1-p)^{N-i}, i=0,1,\dots,N;\ j=1,\dots,n\,.$$

ბინომიალური განაწილების ცხრილებიდან მოცემული p-სათვის მიღებული p_i ალბათობებით გამოვიანგარიშოთ პირსონის სტატისტიკის რიცხვითი მნიშვნელობა:

$$\chi_n^2(k) = \frac{1}{n}\sum_{j=1}^{k}\frac{\nu_j^2}{p_j} - n\,.$$

თუ აღმოჩნდა, რომ $\chi_n^2(k) < \chi^2_{კრ}(\alpha,k-1)$, მაშინ ჰიპოთეზას ვღებულობთ, თუ $\chi_n^2(k) > \chi^2_{კრ}(\alpha,k-1)$, მაშინ – უკუვაგდებთ.

იმ შემთხვევაში, როცა წარმატების p ალბათობა უცნობია, p_i ალბათობების როლში უნდა ავიღოთ მათი შეფასებები:

$$\overline{p}_i = P(X_j = i) = C_N^i \overline{p}^i (1-\overline{p})^{N-i},\ \text{სადაც}\ \overline{p} = \frac{1}{Nn}\sum_{j=1}^{n} x_j\,,$$

და, თუ აღმოჩნდა, რომ $\chi_n^2(k) < \chi^2_{კრ}(\alpha,k-2)$ -- მივიღოთ ჰიპოთეზა, ხოლო წინააღმდეგ შემთხვევაში კი – უკავაგდოთ.

§44. კოლმოგოროვ-სმირნოვის კრიტერიუმი

მცირე შერჩევის დროს მიზანშეწონილია ისეთი კრიტერიუმის გამოყენება, რომელიც (განსხვავებით χ^2 კრიტერიუმისაგან) დაეყრდნობა ინდივიდუალურ და არა დაჯგუფებულ მონაცემებს. ერთ-ერთი ასეთი უმნიშვნელოვანესი კრიტერიუმია კოლმოგოროვის კრიტერიუმი. იგი გამოიყენება H_0 ჰიპოთეზის შესამოწმებლად იმის შესახებ, რომ დამოუკიდებელ და ერთნაირად განაწილებულ $X_1, X_2, \ldots, X_n$ შემთხვევით სიდიდეებს გააჩნიათ მოცემული უწყვეტი $F(x)$ განაწილების ფუნქცია. განვიხილოთ ჰიპოთეზა

$H_0 : F(x) = F_0(x)$ ორმხრივი ალტერნატივის წინააღმდეგ

$H_1 : \max\limits_{|x|<\infty} |F(x) - F_0(x)| > 0$.

განვიხილოთ აგრეთვე ცალვმხრივი ალტერნატიული ჰიპოთეზები

$H_1^+ : \max\limits_{|x|<\infty}(F(x) - F_0(x)) > 0$ და $H_1^- : \max\limits_{|x|<\infty}(F(x) - F_0(x)) < 0$.

H_0 ჰიპოთეზის შესამოწმებლად H_1, H_1^+ და H_1^- ალტერნატივების წინააღმდეგ გამოიყენება კოლმოგოროვისა და სმირნოვის კრიტერიუმები, რომელთა შესაბამისი სტატისტიკებია:

$$D_n = \max_{|x|<\infty} |F_n(x) - F(x)|,\ \ D_n^+ = \max_{|x|<\infty}(F_n(x) - F(x)) \ \text{ და } \ D_n^- = -\min_{|x|<\infty}(F_n(x) - F(x)).$$

ვიპოვოთ ემპირიული განაწილების ფუნქცია $F_n(x)$ და ორმხრივი კრიტიკული არის საზღვრები მოვძებნოთ პირობიდან:

$$D_n = \sup_{|x|<\infty} |F_n(x) - F(x)| > \lambda_n. \quad (1)$$

ა. კოლმოგოროვმა დაამტკიცა, რომ H_0 ჰიპოთეზის სამართლიანობის შემთხვევაში D_n სტატისტიკის განაწილება არ არის დამოკიდებული $F(x)$ ფუნქციაზე, და როცა $n \to \infty$, ადგილი აქვს კრებადობას:

$$p(\sqrt{n}D_n < \lambda) \to K(\lambda),\ \ \lambda > 0, \quad (2)$$

სადაც

$$K(\lambda) = \sum_{m=-\infty}^{\infty} (-1)^m e^{-2m^2\lambda^2} \text{ --}$$

არის კოლმოგოროვის კრიტერიუმი, რომლის მნიშვნელობების პოვნა შესაძლებელია შესაბამისი ცხრილებიდან. კრიტერიუმის კრიტიკული მნიშვნელობა $\lambda_n(\alpha)$ გამოითვლება მოცემული მნიშვნელოვნების α დონის მიხედვით, როგორც $p(D_n \geq \lambda) = \alpha$ განტოლების ამონახსნი.

მტკიცდება, რომ კრიტერიუმის კრიტიკული მნიშვნელობა $\lambda_n(\alpha)$ გამოითვლება შემდეგი მიახლოებითი ფორმულით:

$$\lambda_n(\alpha) \approx \sqrt{\frac{z}{2n}} - \frac{1}{6n},$$

სადაც $z -$ არის $1 - K\left(\sqrt{\frac{\lambda}{2}}\right) = \alpha$ განტოლების ამონახსნი.

პრაქტიკულ ამოცანებში D_n სტატისტიკის გამოსათვლელად გამოიყენება თანაფარდობა:

$$D_n = \max(D_n^+, D_n^-),$$

სადაც

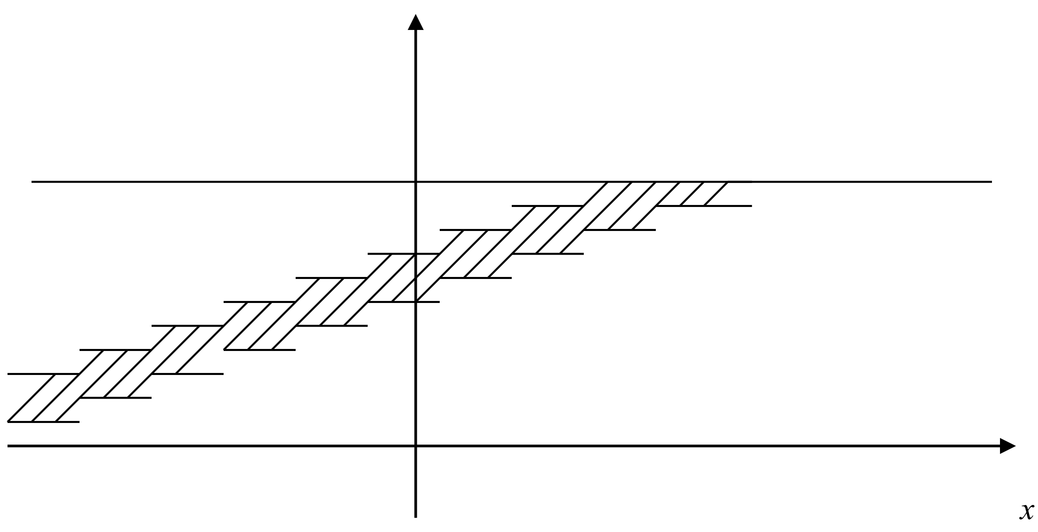
$$D_n^+ = \max_{1\le m\le n}\left(\frac{m}{n} - F(X_{(m)})\right),\ \ D_n^- = \max_{1\le m\le n}\left(F(X_{(m)}) - \frac{m-1}{n}\right),$$

ხოლო $X_{(1)} \le X_{(2)} \le ... \le X_{(n)}$ – ვარიაციული მწკრივია, აგებული $X_1, X_2, \ldots, X_n$ შერჩევის მიხედვით.

თუ H_0 ჰიპოთეზა სამართლიანია, მაშინ D_n^+ და D_n^- სტატისტიკები ერთნაირად არიან განაწილებული. ცნობილია, რომ თუ $\alpha < 0.2$, მაშინ დიდი სიზუსტით $\lambda_n^+(\alpha) \approx \lambda_n(2\alpha)$, სადაც $\lambda_n^+(\alpha)$ არის D_n^+ კრიტერიუმის კრიტიკული მნიშვნელობა.

ჰიპოთეზების შემოწმების წესი შემდეგში მდგომარეობს: ა). H_0 ჰიპოთეზის შემოწმებისას H_1 ალტერნატივის წინააღმდეგ ვიწუნებთ H_0 ჰიპოთეზას, როცა $D_n > \lambda_n(\alpha)$: ბ). თუ $D_n^+ > \lambda_n^+(\alpha)$, მაშინ H_0 ჰიპოთეზას უარვყოფთ H_1^+ ალტერნატივის სასარგებლოდ.

კოლმოგოროვის კრიტერიუმს შეიძლება მიეცეს შემდეგი გეომეტრიული ინტერპრეტაცია: თუ საკოორდინატო სიბრტყეზე გამოვსახავთ $F(x)$ და $F_n(x) \pm\lambda_n(\alpha)$ ფუნქციების გრაფიკებს, მაშინ H_0 ჰიპოთეზა სამართლიანია, თუ $F(x)$ ფუნქციის გრაფიკი არ გამოდის $F_n(x) -\lambda_n(\alpha)$ და $F_n(x) +\lambda_n(\alpha)$ ფუნქციების გრაფიკებს შორის მოთავსებული არიდან:



შენიშვნა. აღსანიშნავია, რომ კოლმოგოროვ-სმირნოვის ტიპის სტატისტიკების კვლევაში დიდი წვლილი მიუძღვით ქართველ მეცნიერებს. 1949-1951 წლებში პროფ. გ. მანიამ დაადგინა აღნიშნული სტატისტიკების ზღვარითი განაწილება და გამოთვალა კრიტიკული მნიშვნელობები. 1964-1965 წლებში პროფ. ე. ნადარაიამ აჩვენა უცნობი განაწილების სიმკვრივის გულოვანი შეფასების კრებადობა თეორიული სიმკვრივისაკენ და დაადგინა

შეფასების სიზუსტე. ე. ნადარაიას მიერ შემოთავაზებული იყო აგრეთვე უცნობი რეგრესიის ფუნქციისათვის გულოვანი შეფასებები, რომელიც ლიტერატურაში ნადარაია-ვატსონის შეფასების სახელითაა ცნობილი.

განაწილების ნორმალურობის შემოწმების მიახლოებითი მეთოდი

განვიხილოთ განაწილების ნორმალურობის შემოწმების მიახლოებითი მეთოდი, რომელიც დაკავშირებულია ასიმეტრიისა და ექსცესის კოეფიციენტების შეფასებებთან. განვმარტოთ ემპირიული განაწილებისათვის ასიმეტრიისა და ექსცესის კოეფიციენტები თეორიული განაწილების შესაბამისი ცნებების ანალოგიურად.

ემპირიული (შერჩევითი) განაწილების ასიმეტრია განიმარტება თანაფარდობით:

$$a_{ე} = \frac{m_3}{\sigma_{ე}^3},$$

სადაც m_3 -- მესამე რიგის ცენტრალური ემპირიული მომენტია.

ემპირიული (შერჩევითი) განაწილების ექსცესი განიმარტება ტოლობით:

$$e_{ე} = \frac{m_4}{\sigma_{ე}^4} - 3,$$

სადაც m_4 -- მეოთხე რიგის ცენტრალური ემპირიული მომენტია.

ცნობილია, რომ ნორმალურად განაწილებული შემთხვევითი სიდიდისათვის ასიმეტრია და ექსცესი ნულია. ამიტომ, თუ შესაბამისი ემპირიული სიდიდეები საკმარისად მცირეა, შეიძლება დავუშვათ, რომ გენერალური ერთობლიობა განაწილებულია ნორმალური განაწილების კანონით.

§45. დამოუკიდებლობის ჰიპოთეზის შემოწმება

განვიხილოთ ერთი პოპულაციის ორი სხვადასხვა ნიშნის (ან ფაქტორის) ერთმანეთთან დამოკიდებულების საკითხი. დავუშვათ, რომ პოპულაციიდან აღებულია n მოცულობის შერჩევა და ამ შერჩევის ელემენტები კლასიფიცირებულია ორი A და B ნიშნის მიხედვით. დავუშვათ, რომ დაკვირვებათა ყველა შესაძლო შედეგი დაყოფილია A ნიშნით $A_1,...,A_k$, ხოლო B ნიშნით $B_1,...,B_r$ კატეგორიებად. პოპულაციის ყოველი ელემენტი ეკუთვნის ზუსტად ერთ კატეგორიას A ნიშნის შესაბამისი რომელიმე კლასიდან და ასევე ზუსტად ერთ რომელიმე კატეგორიას B ნიშნის მიხედვით. ამიტომ დაკვირვებული მონაცემები იყოფა $r\times k$ რაოდენობის A_iB_j არათავსებად ჯგუფად. თუ n_{ij}-ით ავღნიშნავთ იმ მონაცემთა რაოდენობას, რომლებიც ერთდროულად ეკუთვნიან A ნიშნის i-ურ და B ნიშნის j-ურ კატეგორიას და ჩავწერთ ამ სიდიდეს ცხრილის i-ური სვეტისა და j-ური სტრიქონის გადაკვეთაზე, მივიღებთ ორგანზომილებიან ნიშანთა შეუღლების ქვემოთ მოყვანილ ცხრილს, რომელსაც იყენებენ A და B ნიშნების დამოუკიდებლობის ჰიპოთეზის შესამოწმებლად.

$A \backslash B$	B_1	B_2	…	B_j	…	B_r	Σ
A_1	n_{11}	n_{12}	…	n_{1j}	…	n_{1r}	$n_{1\bullet}$
A_2	n_{21}	n_{22}	…	n_{2j}	…	n_{2r}	$n_{2\bullet}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
A_i	n_{i1}	n_{i2}	…	n_{ij}	…	n_{ir}	$n_{i\bullet}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
A_k	n_{k1}	n_{k2}	…	n_{kj}	…	n_{kr}	$n_{k\bullet}$
Σ	$n_{\bullet 1}$	$n_{\bullet 2}$	…	$n_{\bullet j}$	…	$n_{\bullet r}$	n

ამ ცხრილში $(n_{i*},\ 1\le i\le k)$ და $(n_{*j},1\le j\le r)$ სიდიდეები აღნიშნავს A და B ნიშნების შესაბამის მარგინალურ სიხშირეებს. n_{i*} წარმოადგენს შერჩევის იმ ელემენტთა სიხშირეს, რომლებიც მოხვდნენ i-ურ კლასში A ნიშნის მიხედვით, ხოლო n_{*j} არის შერჩევის იმ ელემენტთა რაოდენობა, რომლებიც მოხვდნენ j-ურ კლასში B ნიშნით. ამასთანავე

$$n_{i\bullet}=\sum_{j=1}^{r}n_{ij}\ ,\quad n_{\bullet j}=\sum_{i=1}^{k}n_{ij}\ ,\quad \sum_{i=1}^{k}n_{i\bullet}=\sum_{j=1}^{r}n_{\bullet j}=\sum_{i=1}^{k}\sum_{j=1}^{r}n_{ij}=n.$$

ავღნიშნოთ P_{ij} სიმბოლოთი ალბათობა იმისა, რომ პოპულაციიდან შემთხვევით ამორჩეული ელემენტი აღმოჩნდება ერთდროულად A ნიშნის

i-ურ და B ნიშნის j-ურ კატეგორიაში. მაშინ $P_{i*}=\sum_{j=1}^{r}P_{ij}$ იქნება ალბათობა იმისა, რომ პოპულაციის ელემენტი აღმოჩნდება i-ურ კატეგორიაში A ნიშნის მიხედვით, ხოლო $P_{*j}=\sum_{i=1}^{k}P_{ij}$ -- ალბათობა იმისა, რომ პოპულაციის ელემენტი მოხვდება j-ურ კატეგორიაში B ნიშნით.

ჩვენ შეგვიძლია ჩავთვალოთ, რომ ზემოთ მოყვანილი ცხრილი წარმოადგენს n დამოუკიდებელი დაკვირვების შედეგს ალბათურ მოდელზე, რომლის ერთობლივი განაწილების კანონია:

$X \backslash Y$	B_1	B_2	…	B_j	…	B_r
A_1	P_{11}	P_{12}	…	P_{1j}	…	P_{1r}
A_2	P_{21}	P_{22}	…	P_{2j}	…	P_{2r}
⋮	⋮	⋮	⋮	⋮	⋮	⋮
A_i	P_{i1}	P_{i2}	…	P_{ij}	…	P_{ir}
⋮	⋮	⋮	⋮	⋮	⋮	⋮
A_k	P_{k1}	P_{k2}	…	P_{kj}	…	P_{kr}

სადაც $A_1,...,A_k$ -- არის A ნიშნის შესაძლო შედეგი, ხოლო $B_1,...,B_r$ -- არის B ნიშნის შესაძლო შედეგი.

როდესაც P_{ij} ალბათობები მოცემულია, მაშინ მარტივად გამოითვლება n დამოუკიდებელ ცდაში შესაძლო შედეგთა სავარაუდო სიხშირეები:

$$n_{ij}^{*}=n\cdot P_{ij},\quad i=1,2,...,k;\quad j=1,2,...,r\ .$$

განვიხილოთ დამოუკიდებლობის შემდეგი ჰიპოთეზის შემოწმების ამოცანა.

$H_0: P_{ij}=P_{i*}\cdot P_{*j},\quad i=1,2,...,k;\quad j=1,2,...,r;$

$H_1: H_0$ არ არის მართებული.

როცა A და B ნიშნები დამოუკიდებელია, მაშინ $\forall i,j$-სათვის უნდა შესრულდეს ტოლობა:

$$P_{ij}=P_{i*}\cdot P_{*j},\quad i=1,2,...,k;\quad j=1,2,...,r\ ,$$

სადაც

$$\sum_{i=1}^{k}P_{i*}=1 \ \text{და}\ \sum_{j=1}^{r}P_{*j}=1\,.$$

ჩვენ განვიხილავთ დამოუკიდებლობის ჰიპოთეზის შემოწმების ამოცანას, როცა P_{i*} და P_{*j} მარგინალური განაწილებები უცნობია. ამ შემთხვევაში H_0 ჰიპოთეზა არ აზუსტებს უცნობ პარამეტრთა მნიშვნელობას და საჭიროა მათი შეფასება შერჩევის საშუალებით. შეფასების როლში ავიღოთ ფარდობითი სიხშირე:

$$\overline{P}_{i*} = \frac{n_{i*}}{n}, \quad \overline{P}_{*j} = \frac{n_{*j}}{n}, \; i = 1,2,...,k; \quad j = 1,2,...,r,$$

მაშინ ჰიპოთეტური სიხშირეები გამოითვლება ფორმულებით

$$n_{ij}^* = \frac{n_{i*} \cdot n_{j*}}{n}, \; i = 1,2,...,k; \quad j = 1,2,...,r. \tag{1}$$

ნიშანთა დამოუკიდებლობის ჰიპოთეზის შესამოწმებლად გამოიყენება

$$\hat{\chi}^2 = \sum_{i=1}^{k} \sum_{j=1}^{r} \frac{\left(n_{ij} - n_{ij}^*\right)^2}{n_{ij}^*}$$

სტატისტიკა, რომელიც მიახლოებით χ^2 კანონით არის განაწილებული თავისუფლების ხარისხით $(k-1)(r-1)$. მოცემული მნიშვნელოვნების α დონისათვის χ^2 განაწილების ცხრილიდან ვპოულობთ $\chi^2_{\alpha,(k-1)(r-1)}$ კრიტიკულ წერტილს. თუ აღმოჩნდა, რომ $\hat{\chi}^2 \geq \chi^2_{\alpha,(k-1)(r-1)}$, მაშინ ნულოვან ჰიპოთეზას უარვყოფთ. წინააღმდეგ შემთხვევაში ვასკვნით, რომ A და B ნიშნები დამოუკიდებელია.

მაგალითი. სოციოლოგს სურს 395 ოჯახზე დაკვირვებით მიღებული შერჩევის საფუძველზე შეამოწმოს ჰიპოთეზა იმის შესახებ, რომ ოჯახში ბავშვების რაოდენობა არ არის დამოკიდებული ოჯახის შემოსავალზე:

ბავშვების რაოდენობა	0–6 A ჯგუფი	6–12 B ჯგუფი	12–18 C ჯგუფი	18-ზე მეტი D ჯგუფი
0	10	9	18	24
1	8	12	25	31
2	24	28	23	28
3	26	24	20	6
4 ან მეტი	32	22	18	7

ამოხსნა. ავიღოთ 0.01-ის ტოლი ნდობის ალბათობა. (1) ფორმულების თანახმად გვექნება ჰიპოთეტური სიხშირეების შემდეგი ცხრილი:

ბავშვების რაოდენობა	0–6 A ჯგუფი	6–12 B ჯგუფი	12–18 C ჯგუფი	18-ზე მეტი D ჯგუფი
0	15.44	14.67	16.06	14.83
1	19.24	18.28	20.01	18.47
2	26.08	24.77	27.12	25.03
3	19.24	18.28	20.01	18.47
4 ან მეტი	20.00	19.00	20.80	19.20

გამოთვლების მოხერხებულობის მიზნით χ^2 სტატისტიკის დაკვირვე-ბული მნიშვნელობების გამოსათვლელად ვისარგებლოთ შემდეგი ცხრილ-ით:

	n_{ij}^0	n_{ij}^*	$n_{ij}^0 - n_{ij}^*$	$\left(n_{ij}^0 - n_{ij}^*\right)^2$	$\left(n_{ij}^0 - n_{ij}^*\right)^2 / n_{ij}^*$
A0	10	15.44	-5.44	29.63	1.92
A1	8	19.24	-11.24	126.35	6.57
A2	24	26.08	-2.08	4.31	0.17
A3	26	19.24	6.76	45.69	2.37
A4	32	20.00	12.00	144.00	7.20
B0	9	14.67	-5.67	32.16	2.19
B1	12	18.28	-6.28	39.42	2.16
B2	28	24.77	3.23	10.42	0.42
B3	24	18.28	5.72	32.74	1.79
B4	22	19.00	3.00	9.00	0.47
C0	18	16.06	1.94	3.76	0.23
C1	25	20.01	4.99	24.90	1.24
C2	23	27.12	-4.12	16.97	0.63
C3	20	20.01	-0.01	0.00	0.00
C4	18	20.80	-2.80	7.84	0.38
D0	24	14.83	9.17	84.17	5.68
D1	31	18.47	12.53	156.98	8.50
D2	28	25.03	2.97	8.80	0.35
D3	6	18.47	-12.47	155.52	8.42
D4	7	19.20	-12.20	148.84	7.75
					$\sum = 58.44$

როგორც ვხედავთ სტატისტიკის დაკვირვებული მნიშვნელობაა $\hat{\chi}^2 =$ $=58.44$. მეორეს მხრივ, რადგან თავისუფლების ხარისხია $(r-1)(k-1)=12$, ამიტომ (0.01 მნიშვნელოვნების დონისათვის) კრიტიკული მნიშვნელობაა $\chi^2_{12,0.01} = 26.217$. ვინაიდან, $\hat{\chi}^2 > \chi^2_{12,0.01}$, სოციოლოგი დაასკვნის, რომ ოჯახში ბავშვების რაოდენობა და ოჯახის შემოსავალი დამოკიდებულია ერთმანე-თზე.

§46. ერთგვაროვნების შემოწმების ჰიპოთეზა

დავუშვათ, რომ მოცემულია k რაოდენობის სხვადასხვა პოპულაცია და ყოველი პოპულაციიდან, ერთმანეთისაგან დამოუკიდებლად, აღებულია $n_1,\dots,n_k$ მოცულობის შერჩევები. ვიგულისხმოთ, რომ ყველა პოპულაცია კლასიფიცირებულია ერთი და იგივე A ნიშნის $A_1,\dots A_r$ კატეგორიის მიხედვით. i-ური შერჩევის იმ ელემენტთა სიხშირე, რომლებსაც აღმოაჩნდათ j-ური კატეგორია ავღნიშნოთ n_{ij} სიმბოლოთი. მაშინ მონაცემები განლაგდება ნიშანთა შეუღლების შემდეგ ცხრილში:

	კატეგორიები						$\sum$
	A_1	A_2	...	A_j	...	A_r	
შერჩევა I პოპულაციიდან	n_{11}	n_{12}	...	n_{1j}	...	n_{1r}	n_1
შერჩევა II პოპულაციიდან	n_{21}	n_{22}	...	n_{2j}	...	n_{2r}	n_2
	...	...	...	...	...	...	...
შერჩევა j-ური პოპულაციიდან	n_{i1}	n_{i2}	...	n_{ij}	...	n_{ir}	n_i
	...	...	...	...	...	...	...
შერჩევა k-ური პოპულაციიდან	n_{k1}	n_{k2}	...	n_{kj}	...	n_{kr}	n_k
$\sum$	$n_{\bullet 1}$	$n_{\bullet 2}$	...	$n_{\bullet j}$	...	$n_{\bullet r}$	n

ასეთ შემთხვევაში ხშირად ჩნდება პოპულაციათა ერთგვაროვნების (შერჩევები, რომ აღებულია ერთი და იგივე გენერალური ერთობლიობიდან) ჰიპოთეზის შემოწმების აუცილებლობა. ასეთი ჰიპოთეზა ექვივალენტურია ჰიპოთეზისა, რომ პოპულაციიდან შემთხვევით არჩეული ელემენტის ყოველ A_j კლასში მოხვედრის P_j ალბათობა ერთი და იგივეა ყველა პოპულაციისათვის.

ცხრილში n_i არის i-ური პოპულაციიდან აღებული შერჩევის მოცულობა

$$n_i = \sum_{j=1}^{r} n_{ij}, \quad i = 1,2,\dots,k .$$

n_{*j} სიმბოლოთი აღნიშნულია ყველა შერჩევის იმ ელემენტთა რაოდენობა, რომლებსაც აღმოაჩნდათ A ნიშნის j-ური კატეგორია. ცხადია, რომ n_i სიდიდეებისგან განსხვავებით, n_{*j} სიდიდეები (შერჩევის აღებამდე) შემთხვევით სიდიდეებს წარმოადგენენ და

$$n_{*j} = \sum_{i=1}^{k} n_{ij}, \quad j = 1,2,\dots,r .$$

ალბათობა იმისა, რომ i-ური პოპულაციიდან შემთხვევით არჩეულ ელემენტს აღმოაჩნდება j-ური კატეგორია (ან i-ური პოპულაციიდან j-ური კატეგორიის ელემენტთა პროპორცია) ავღნიშნოთ P_{ij} სიმბოლოთი.

$$\sum_{j=1}^{r} P_{ij} = 1 .$$

განვიხილოთ შემდეგი ჰიპოთეზები.

$H_0 : P_j \equiv P_{1j} = P_{2j} = \cdots = P_{kj}, \quad j = 1, 2, ..., r$;

$H_1 : H_0$ არ არის მართებული.

H_0 ჰიპოთეზის დროს მოსალოდნელი რაოდენობა i-ური შერჩევის ელემენტებისა, რომლებიც j-ური კატეგორიის აღმოჩნდნენ ტოლია:

$$n_{ij}^{*} = n_i \cdot P_j .$$

P_j პარამეტრის შეფასების როლში ავიღოთ

$$\overline{P}_j = \frac{n_{\bullet j}}{n}, \quad j = 1,2,..., r .$$

მაშინ, H_0 ჰიპოთეზის დროს i-ური შერჩევის j-ური კატეგორიის ელემენტების მოსალოდნელი რაოდენობა იქნება:

$$n_{ij}^{*} = \frac{n_i \cdot n_{*j}}{n} .$$

დაკვირვებულ n_{ij} სიდიდეებსა და H_0 ჰიპოთეზის დროს მათ მოსალოდნელ n_{ij}^{*} მნიშვნელობებს შორის გადახრის საზომად აიღება

$$\hat{\chi}^2 = \sum_{i=1}^{k} \sum_{j=1}^{r} \frac{\left(n_{ij} - n_{ij}^{*}\right)^2}{n_{ij}^{*}}$$

სტატისტიკა.

ნულოვანი ჰიპოთეზის უარყოფის არეა $\chi^2 \geq \chi^2_{\alpha,(k-1)(r-1)}$. წინააღმდეგ შემთხვევაში ვასკვნით, რომ პოპულაცია ერთგვაროვანია.

მაგალითი. ქვემოთ მოყვანილია სამი სხვადასხვა საწარმოს მიერ წარმოებულ ერთი და იგივე ტიპის პროდუქციაში ვარგის და უვარგის ნაწარმთა რაოდენობები:

	ვარგისი	უვარგისი	სულ
I საწარმო	240	10	250
II საწარმო	191	9	200
III საწარმო	139	11	150
სულ	570	30	600

არის თუ არა განსხვავება ამ საწარმოთა მიერ გამოშვებული პროდუქციის ხარისხში?

ამოხსნა. განვიხილოთ ნულოვანი ჰიპოთეზა: შერჩევები ერთგვაროვა-

ნია. ამ ჰიპოთეზის დროს სავარაუდო სიხშირეებია

237.5	12.5
190	10
142.5	7.5

ცხრილის საშუალებით გამოვთვალოთ ხი კვადრატ სტატისტიკის დაკვირვებული მნიშვნელობები:

	n_{ij}	n^*_{ij}	$n_{ij} - n^*_{ij}$	$(n_{ij} - n^*_{ij})^2$	$(n_{ij} - n^*_{ij})^2 / n^*_{ij}$
I საწარმო/ვარგისი	240	237.5	2.5	6.25	0.026
II საწარმო/ვარგისი	191	190	1	1	0.005
III საწარმო/ვარგისი	139	142.5	-3.5	12.25	0.086
I საწარმო/უვარგისი	10	12.5	-2.5	6.25	0.500
II საწარმო/უვარგისი	9	10	-1	1	0.100
III საწარმო/უვარგისი	11	7.5	3.5	12.25	1.633
					$\hat{\chi}^2=2.35$

ვიპოვოთ, მნიშვნელოვნების 0.1 დონისათვის თავისუფლების ხარისხით (3-1)(2-1)=2, კრიტერიუმის კრიტიკული მნიშვნელობა: $\chi^2_{2,0.1}=4.60517$.

რადგანაც $\hat{\chi}^2=2.35<4.60517$, ამიტომ აღნიშნული მონაცემები არ იძლევა ერთგვაროვნების ჰიპოთეზის უარყოფის საფუძველს.

§47. შემთხვევით სიდიდეთა მოდელირება. მონტე-კარლოს მეთოდი

მონტე-კარლოს მეთოდი გამოიყენება შემდეგი ამოცანის ამოსახსნელად: საჭიროა მოიძებნოს შესასწავლი შემთხვევითი სიდიდის მნიშვნელობა a. მისი განსაზღვრისათვის ირჩევენ X შემთხვევით სიდიდეს, რომლის მათემატიკური ლოდინი ტოლია a–სი, და X შემთხვევითი სიდიდის n ცალი მნიშვნელობის შერჩევიდან, რომელიც მიიღება n ექსპერიმენტში, გამოითვლება შერჩევითი საშუალო:

$$\bar{x} = \frac{\sum x_i}{n},$$

რომელიც მიიღება საძიებელი a რიცხვის შეფასებად:

$$a \approx a^* = \bar{x}.$$

ეს მეთოდი მოითხოვს ექსპერიმენტების დიდი რიცხვის ჩატარებას, ამიტომ მას სხვანაირად **სტატისტიკური ექსპერიმენტების მეთოდი** ეწოდება. მონტე-კარლოს მეთოდის თეორია იკვლევს: როგორ უფრო მიზანშეწონილია აირჩეს X შემთხვევითი სიდიდე, როგორ უნდა ვიპოვოთ მისი შესაძლო მნიშვნელობები, როგორ შევამციროთ გამოყენებული შემთხვევითი სიდიდეების დისპერსია, რათა ცდომილება a–ს a^* –თი შეცვლისას იყოს რაც შეიძლება მცირე.

X შემთხვევითი სიდიდის შესაძლო მნიშვნელობების მოძებნას უწოდებენ **შემთხვევითი სიდიდის გათამაშებას (მოდელირებას).** ქვემოთ ჩვენ განვიხილავთ შემთხვევითი სიდიდის მოდელირების ზოგიერთ მეთოდს და გავარკვევთ თუ როგორ შევაფასოთ ამ დროს დაშვებული შეცდომა.

თუ ჩვენ გვინდა განვსაზღვროთ დაშვებული შეცდომის ზედა საზღვარი მოცემული საიმედოობის γ ალბათობით, ანუ მოვძებნოთ δ რიცხვი, რომლისთვისაც $p(|\overline{X} - a| \le \delta) = \gamma$, ჩვენ ვღებულობთ გენერალური ერთობლიობის მათემატიკური ლოდინისათვის ნდობის ინტერვალის მოძებნის ცნობილ ამოცანას. ამიტომ ჩვენ ამ ამოცანაზე ცალკე არ შევჩერდებით.

განმარტება 1. [0; 1] ინტერვალზე თანაბრად განაწილებული R შემთხვევითი სიდიდის შესაძლო r მნიშვნელობებს **შემთხვევითი რიცხვები** ეწოდება.

დისკრეტული შემთხვევითი სიდიდის მოდელირება

დავუშვათ, რომ გასათამაშებელია დისკრეტული X შემთხვევითი სიდიდე, ე. ი. X შემთხვევითი სიდიდის ცნობილი განაწილების კანონის მიხედვით მივიღოთ მისი შესაძლო მნიშვნელობების მიმდევრობა:

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

განვიხილოთ [0; 1] ინტერვალზე თანაბრად განაწილებული R შემთხვევითი სიდიდე და დავყოთ [0; 1] ინტერვალი $p_1,$ $p_1 + p_2$, …, $p_1 + p_2 + \ldots + p_{n-1}$

კოორდინატების მქონე წერტილებით n ქვეინტერვალად: $\Delta_1, \Delta_2, ..., \Delta_n$, რომელთა სიგრძეები ტოლია შესაბამისი ინდექსის მქონე ალბათობების.

თეორემა 1. თუ ნებისმიერ შემთხვევით რიცხვს $r_j (0 \le r_j < 1)$, რომელიც მოხვდა Δ_i ინტერვალში, შევუსაბამებთ x_i შესაძლო მნიშვნელობას, მაშინ გასათამაშებელ სიდიდეს ექნება მოცემული განაწილების კანონი:

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

დამტკიცება. შემთხვევითი სიდიდის შესაძლო მნიშვნელობები ემთხვევა $\{x_1, x_2, \ldots, x_n\}$ სიმრავლეს, რადგანაც ინტერვალების რაოდენობა ტოლია n–ის, და r_j–ს Δ_i ინტერვალში მოხვედრისას შემთხვევით სიდიდეს შეუძლია მიიღოს მხოლოდ ერთი $x_1, x_2, \ldots, x_n$ მნიშვნელობებიდან. ვინაიდან R განაწილებულია თანაბრად, ამიტომ მისი თითოეულ ინტერვალში მოხვედრის ალბათობა ტოლია ამ ინტერვალის სიგრძის, საიდანაც გამოდის, რომ ნებისმიერ x_i მნიშვნელობას შეესაბამება ალბათობა p_i. ამრიგად, გასათამაშებელი შემთხვევითი სიდიდეს გააჩნია მოცემული განაწილების კანონი.

მაგალითი. გავათამაშოთ 10 მნიშვნელობა დისკრეტული X შემთხვევითი სიდიდის, რომლის განაწილების კანონია:

$$\begin{array}{lllll} X & 2 & 3 & 6 & 8 \\ P & 0.1 & 0.3 & 0.5 & 0.1. \end{array}$$

ამოხსნა. დავყოთ [0; 1] ინტერვალი ქვეინტერვალებად: Δ_1- [0; 0.1), Δ_2 – [0.1; 0.4), Δ_3 - [0.4; 0.9), Δ_4 – [0.9; 1]. შემთხვევითი რიცხვების ცხრილიდან ამოვწეროთ 10 რიცხვი: 0.09; 0.73; 0.25; 0.33; 0.76; 0.52; 0.01; 0.35; 0.86; 0.34. პირველი და მეშვიდე რიცხვი ძევს Δ_1 ინტერვალში, შესაბამისად, ამ ორ შემთხვევაში გასათამაშებელი შემთხვევითი სიდიდე მიიღებს მნიშვნელობას $x_1 = 2$; მე-3, მე-4, მე-8 და მე-10 რიცხვები ჩავარდნენ Δ_2 ინტერვალში, რასაც შეესაბამება $x_2 = 3$; მე-2, მე-5, მე-6 და მე-9 რიცხვები აღმოჩნდნენ Δ_3 ინტერვალში, ამასთანავე $X = x_3 = 6$; და ბოლოს, უკანასკნელ ინტერვალში არ ჩავარდა არც ერთი რიცხვი. ამრიგად, X შემთხვევითი სიდიდის გათამაშებული მნიშვნელობებია: 2, 6, 3, 3, 6, 6, 2, 3, 6, 3.

საწინააღმდეგო ხდომილებების მოდელირება

დავუშვათ, რომ უნდა გავითამაშოთ ექსპერიმენტები, რომელთაგან თითოეულში A ხდომილება ჩნდება (ხდება) ცნობილი p ალბათობით. განვიხილოთ დისკრეტული X შემთხვევითი სიდიდე, რომელიც ღებულობს მნიშვნელობას 1 (იმ შემთხვევაში, როცა ხდება A) ალბათობით p და მნიშვნელობას 0 (თუ არ მოხდა A) ალბათობით $q = 1 - p$. შემდეგ ვათამაშებთ ამ შემთხვევით სიდიდეს, ისე როგორც ეს იყო წინა პუნქტში.

მაგალითი. გავათამაშოთ 10 ექსპერიმენტი, რომელთაგან თითოეულში A ხდომილება ხდება ალბათობით 0.3.

ამოხსნა. X შემთხვევითი სიდიდისათვის განაწილების კანონით

$$\begin{array}{lll} X & 1 & 0 \\ P & 0.3 & 0.7 \end{array}$$

მივიღებთ ინტერვალებს Δ_1 – [0; 0,3) и Δ_2 – [0,3; 1]. გამოვიყენოთ შემთხვევითი რიცხვების იგივე შერჩევა, რაც გვქონდა წინა მაგალითში: 0.09; 0.73; 0.25; 0.33; 0.76; 0.52; 0.01; 0.35; 0.86; 0.34. Δ_1 ინტერვალში მოხვდება პირველი, მე-3 და მე-7 რიცხვი, ხოლო დანარჩენი კი -- Δ_2 ინტერვალში. შესაბამისად, შეგვიძლია ჩავთვალოთ, რომ A ხდომილება მოხდა პირველ, მე-3 და მე-7 ექსპერიმენტში, ხოლო დანარჩენებში კი – არ მოხდა.

ხდომილებათა სრული სისტემის მოდელირება

თუ ხდომილებები $A_1, A_2, \ldots, A_n$, რომელთა ალბათობებია შესაბამისად p_1 , p_2 ,… p_n, ქმნიან ხდომილებათა სრულ ჯგუფს, მაშინ მათი მოდელირებისათვის (ე. ი. ექსპერიმენტების სერიაში მათი გამოჩენის მიმდევრობის მოდელირება) უნდა გავათამაშოთ დისკრეტული X შემთხვევითი სიდიდე განაწილების კანონით:

$$\begin{array}{cccc} X & 1 & 2 \ldots & n \\ P & p_1 & p_2 \ldots & p_n \end{array}$$

ამასთანავე ითვლება, რომ თუ X მიიღებს მნიშვნელობას $x_i = i$, მაშინ ამ ექსპერიმენტში მოხდა A_i ხდომილება.

უწყვეტი შემთხვევითი სიდიდის მოდელირება

ა). ***შებრუნებული ფუნქციების მეთოდი.*** დავუშვათ, რომ უნდა გავათამაშოთ უწყვეტი X შემთხვევითი სიდიდე, ე. ი. უნდა მივიღოთ მისი შესაძლო მნიშვნელობების მიმდევრობა x_i (i = 1, 2, …, n), როცა ცნობილია მისი განაწილების ფუნქცია $F(x)$.

თეორემა 2. თუ r_i – შემთხვევითი რიცხვია, მაშინ მოცემული მკაცრად ზრდადი $F(x)$ განაწილების ფუნქციის მქონე გასათამაშებელი უწყვეტი X შემთხვევითი სიდიდის შესაძლო x_i მნიშვნელობა, რომელიც შეესაბამება r_i –ს, წარმოადგენს შემდეგი განტოლების ამონახსნს

$$F(x_i) = r_i\,. \qquad (1)$$

დამტკიცება. ვინაიდან $F(x)$ მკაცრად იზრდება ინტერვალში 0-დან 1-მდე, ამიტომ მოიძებნება (ამასთანავე ერთადერთი) არგუმენტის ისეთი მნიშვნელობა x_i, რომლის დროსაც განაწილების ფუნქცია მიიღებს მნიშვნელობას r_i, ანუ (1) განტოლებას გააჩნია ერთადერთი ამონახსნი: $x_i = F^{-1}(r_i)$, სადაც F^{-1}– არის F ფუნქციის შექცეული ფუნქციაა. ვაჩვენოთ, რომ (1) განტოლების ამონახსნი წარმოადგენს განსახილველი X შემთხვევითი სიდიდის შესაძლო მნიშვნელობას.

წინასწარ ვაჩვენოთ, რომ თუ x_i – შესაძლო მნიშვნელობაა გარკვეული ξ შემთხვევითი სიდიდის, მაშინ ξ შემთხვევითი სიდიდის (c, d) ინტერვალში მოხვედრის ალბათობაა $F(d) - F(c)$. მართლაც, $F(x)$ ფუნქციის მონოტონურობის გამო, $F(x_i) = r_i$ ტოლობის გათვალისწინებით გვაქვს:

$$c < x_i < d \Leftrightarrow F(c) < r_i < F(d)\,.$$

ამიტომ

$$c < x_i < d \Leftrightarrow F(c) < r_i < F(d),$$

შესაბამისად,

$$p(c < \xi < d) = p(F(c) < R < F(d)) = F(d) - F(c).$$

ე. ი. ξ შემთხვევითი სიდიდის (c, d) ინტერვალში მოხვედრის ალბათობა ტოლია ამ ინტერვალზე $F(x)$ განაწილების ფუნქციის ნაზრდის, შესაბამისად, $\xi = X$.

მაგალითი. გავათამაშოთ [5; 8] ინტერვალზე თანაბრად განაწილებული უწყვეტი X შემთხვევითი სიდიდის 3 შესაძლო მნიშვნელობა.

ამოხსნა. გასაგებია, რომ

$$F(x) = \frac{x-5}{3}.$$

ამიტომ უნდა ამოვხსნათ განტოლება $\frac{x_i - 5}{3} = r_i$, საიდანაც $x_i = 3r_i + 5$.

ავირჩიოთ 3 შემთხვევითი რიცხვი: 0.23; 0.09; 0.56 და ჩავსვათ ისინი ამ განტოლებაში. მივიღებთ X შემთხვევითი სიდიდის შესაბამის შესაძლო მნიშვნელობებს: $x_1 = 5.69$; $x_2 = 5.27$; $x_3 = 6.68$.

ბ). ***სუპერპოზიციის მეთოდი.*** თუ გასათამაშებელი შემთხვევითი სიდიდის განაწილების ფუნქცია შეიძლება წარმოდგეს ორი განაწილების ფუნქციის წრფივი კომბინაციის სახით:

$$F(x) = C_1F_1(x) + C_2F_2(x) \quad (C_1, C_2 > 0),$$

მაშინ $C_1 + C_2 = 1$, ვინაიდან, $F(x) \to 1$, როცა $x \to \infty$.

შემოვიღოთ დამხმარე დისკრეტული შემთხვევითი სიდიდე Z განაწილების ფუნქციით:

Z	1	2
P	C_1	C_2

ავირჩიოთ 2 დამოუკიდებელი შემთხვევითი რიცხვი r_1 და r_2 გავათამაშოთ Z შემთხვევითი სიდიდე r_1 რიცხვის მიხედვით. თუ $Z = 1$, მაშინ X –ის შესაძლო მნიშვნელობას ვეძებთ განტოლებიდან $F_1(x) = r_2$, ხოლო თუ $Z =$ $= 2$, მაშინ ვხსნით განტოლებას $F_2(x) = r_2$. შეიძლება დამტკიცდეს, რომ ამ შემთხვევაში გასათამაშებელი შემთხვევითი სიდიდის განაწილების ფუნქცია ტოლია მოცემული განაწილების ფუნქციის.

გ). ***ნორმალური შემთხვევითი სიდიდის მიახლოებითი გათამაშება.*** ვინაიდან (0, 1) ინტერვალში თანაბრად განაწილებული R შემთხვევითი სიდიდისათვის: $E(R) = \frac{1}{2}$, $D(R) = \frac{1}{12}$, ამიტომ [0, 1] ინტერვალზე თანაბრად განაწილებული დამოუკიდებელი R_j $(j = 1, 2, ..., n)$ შემთხვევითი სიდიდეების ჯამისათვის $\sum_{j=1}^{n} R_j$:

$$E(\sum_{j=1}^{n} R_j) = \frac{n}{2}, \quad D(\sum_{j=1}^{n} R_j) = \frac{n}{12}, \quad \sigma = \sqrt{\frac{n}{12}}.$$

ამიტომ ცენტრალური ზღვარითი თეორემის თანახმად ნორმირებულ შემთხვევით სიდიდეს $(\sum_{j=1}^{n} R_j - \frac{n}{2})/\sqrt{n/12}$, როცა $n \to \infty$ ექნება ნორმალურთან ახლოს მყოფი განაწილება, პარამეტრებით $a = 0$ და $\sigma = 1$. კერძოდ, საკმაოდ კარგი მიახლოება მიიღება, როცა $n = 12$:

$$\sum_{j=1}^{12} R_j - 6 .$$

ამრიგად, იმისათვის, რომ გავათამაშოთ ნორმირებული ნორმალური შემთხვევითი სიდიდის შესაძლო მნიშვნელობა, უნდა შევკრიბოთ 12 დამოუკიდებელი შემთხვევითი რიცხვი და ჯამს გამოვაკლოთ 6.

ინტეგრალის გამოთვლა მონტე-კარლოს მეთოდით

ვნახოთ, თუ როგორ შეიძლება

$$\int_0^1 f(x)dx \qquad (2)$$

ინტეგრალის მიახლოებითი გამოთვლა, სადაც $f:[0,1] \to [0,1]$ უწყვეტი ფუნქციაა. განვიხილოთ [0,1] შუალედში თანაბრად განაწილებულ დამოუკიდებელ შემთხვევით სიდიდეთა მიმდევრობა $X_1, X_2, \ldots$ და ავაგოთ ახალი მიმდევრობა:

$$Z_i = f(X_i), \; i \geq 1 . \qquad (3)$$

მტკიცდება, რომ $Z_i, i \geq 1$ აგრეთვე დამოუკიდებელ ერთნაირად განაწილებულ შემთხვევით სიდიდეთა მიმდევრობაა და $\forall i$:

$$EZ_i = \int_0^1 f(x)dx .$$

ამიტომ დიდ რიცხვთა კანონის თანახმად ადგილი აქვს კრებადობას:

$$\frac{1}{n}\sum_{i=1}^{n} Z_i \to \int_0^1 f(x)dx \;\; (\text{ალბათობით } 1).$$

შესაბამისად, (2) ინტეგრალის მიახლოებითი გამოთვლისათვის უნდა დამოდელირდეს შემთხვევით სიდიდეთა $(X_i, Z_i), i \geq 1$ მიმდევრობა და გამოითვალოს (3) წესით შედგენილ სიდიდეთა საშუალო არითმეტიკული.

დანართი (სტატისტიკური ცხრილები)

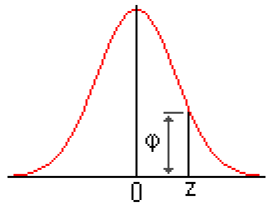
პუასონის განაწილების ცხრილები ($P(k)=\dfrac{\lambda^k}{k!}e^{-\lambda}$)

	λ = 0.1	λ = 0.2	λ = 0.3	λ = 0.4	λ = 0.5	λ = 0.6	λ = 0.7	λ = 0.8	λ = 0.9
p(0)	**0.9048**	**0.8187**	**0.7408**	**0.6703**	**0.6065**	**0.5488**	**0.4966**	**0.4493**	**0.4066**
p(1)	0.0905	0.1637	0.2222	0.2681	0.3033	0.3293	0.3476	0.3595	0.3659
p(2)	0.0045	0.0164	0.0333	0.0536	0.0758	0.0988	0.1217	0.1438	0.1647
p(3)	0.0002	0.0011	0.0033	0.0072	0.0126	0.0198	0.0284	0.0383	0.0494
p(4)		0.0001	0.0003	0.0007	0.0016	0.0030	0.0050	0.0077	0.0111
p(5)				0.0001	0.0002	0.0004	0.0007	0.0012	0.0020
p(6)							0.0001	0.0002	0.0003

	λ = 1.0	λ = 1.5	λ = 2.0	λ = 2.5	λ = 3.0	λ = 3.5	λ = 4.0	λ = 4.5	λ = 5.0
p(0)	**0.3679**	0.2231	0.1353	0.0821	0.0498	0.0302	0.0183	0.0111	0.0067
p(1)	**0.3679**	**0.3347**	**0.2707**	0.2052	0.1494	0.1057	0.0733	0.0500	0.0337
p(2)	0.1839	0.2510	**0.2707**	**0.2565**	**0.2240**	0.1850	0.1465	0.1125	0.0842
p(3)	0.0613	0.1255	0.1804	0.2138	**0.2240**	**0.2158**	**0.1954**	0.1687	0.1404
p(4)	0.0153	0.0471	0.0902	0.1336	0.1680	0.1888	**0.1954**	**0.1898**	**0.1755**
p(5)	0.0031	0.0141	0.0361	0.0668	0.1008	0.1322	0.1563	0.1708	**0.1755**
p(6)	0.0005	0.0035	0.0120	0.0278	0.0504	0.0771	0.1042	0.1281	0.1462
p(7)	0.0001	0.0008	0.0034	0.0099	0.0216	0.0385	0.0595	0.0824	0.1044
p(8)		0.0001	0.0009	0.0031	0.0081	0.0169	0.0298	0.0463	0.0653
p(9)			0.0002	0.0009	0.0027	0.0066	0.0132	0.0232	0.0363
p(10)				0.0002	0.0008	0.0023	0.0053	0.0104	0.0181
p(11)					0.0002	0.0007	0.0019	0.0043	0.0082
p(12)					0.0001	0.0002	0.0006	0.0016	0.0034
p(13)						0.0001	0.0002	0.0006	0.0013
p(14)							0.0001	0.0002	0.0005
p(15)								0.0001	0.0002

სტანდარტული ნორმალური განაწილების სიმკვრივის

($\varphi(z)=\frac{1}{2\pi}e^{-z^2/2}$) მნიშვნელობები

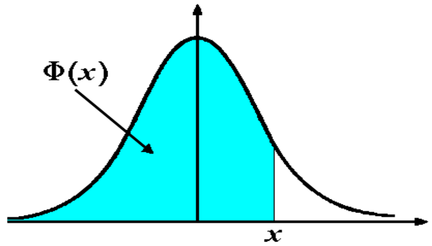


Z	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0.0	.398942	.398922	.398862	.398763	.398623	.398444	.398225	.397966	.397668	.397330
0.1	.396953	.396536	.396080	.395585	.395052	.394479	.393868	.393219	.392531	.391806
0.2	.391043	.390242	.389404	.388529	.387617	.386668	.385683	.384663	.383606	.382515
0.3	.381388	.380226	.379031	.377801	.376537	.375240	.373911	.372548	.371154	.369728
0.4	.368270	.366782	.365263	.363714	.362135	.360527	.358890	.357225	.355533	.353812
0.5	.352065	.350292	.348493	.346668	.344818	.342944	.341046	.339124	.337180	.335213
0.6	.333225	.331215	.329184	.327133	.325062	.322972	.320864	.318737	.316593	.314432
0.7	.312254	.310060	.307851	.305627	.303389	.301137	.298872	.296595	.294305	.292004
0.8	.289692	.287369	.285036	.282694	.280344	.277985	.275618	.273244	.270864	.268477
0.9	.266085	.263688	.261286	.258881	.256471	.254059	.251644	.249228	.246809	.244390
Z	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1.0	.241971	.239551	.237132	.234714	.232297	.229882	.227470	.225060	.222653	.220251
1.1	.217852	.215458	.213069	.210686	.208308	.205936	.203571	.201214	.198863	.196520
1.2	.194186	.191860	.189543	.187235	.184937	.182649	.180371	.178104	.175847	.173602
1.3	.171369	.169147	.166937	.164740	.162555	.160383	.158225	.156080	.153948	.151831
1.4	.149727	.147639	.145564	.143505	.141460	.139431	.137417	.135418	.133435	.131468
1.5	.129518	.127583	.125665	.123763	.121878	.120009	.118157	.116323	.114505	.112704
1.6	.110921	.109155	.107406	.105675	.103961	.102265	.100586	.098925	.097282	.095657
1.7	.094049	.092459	.090887	.089333	.087796	.086277	.084776	.083293	.081828	.080380
1.8	.078950	.077538	.076143	.074766	.073407	.072065	.070740	.069433	.068144	.066871
1.9	.065616	.064378	.063157	.061952	.060765	.059595	.058441	.057304	.056183	.055079
Z	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
2.0	.053991	.052919	.051864	.050824	.049800	.048792	.047800	.046823	.045861	.044915
2.1	.043984	.043067	.042166	.041280	.040408	.039550	.038707	.037878	.037063	.036262
2.2	.035475	.034701	.033941	.033194	.032460	.031740	.031032	.030337	.029655	.028985
2.3	.028327	.027682	.027048	.026426	.025817	.025218	.024631	.024056	.023491	.022937
2.4	.022395	.021862	.021341	.020829	.020328	.019837	.019356	.018885	.018423	.017971
2.5	.017528	.017095	.016670	.016254	.015848	.015449	.015060	.014678	.014305	.013940
2.6	.013583	.013234	.012892	.012558	.012232	.011912	.011600	.011295	.010997	.010706
2.7	.010421	.010143	3z98712	3z96058	3z93466	3z90936	3z88465	3z86052	3z83697	3z81398
2.8	3z79155	3z76965	3z74829	3z72744	3z70711	3z68728	3z66793	3z64907	3z63067	3z61274
2.9	3z59525	3z57821	3z56160	3z54541	3z52963	3z51426	3z49929	3z48470	3z47050	3z45666

$\varphi(z)$-ის მნიშვნელობები (გაგრძელება)

Z	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
3.0	3z44318	3z43007	3z41729	3z40486	3z39276	3z38098	3z36951	3z35836	3z34751	3z33695
3.1	3z32668	3z31669	3z30698	3z29754	3z28835	3z27943	3z27075	3z26231	3z25412	3z24615
3.2	3z23841	3z23089	3z22358	3z21649	3z20960	3z20290	3z19641	3z19010	3z18397	3z17803
3.3	3z17226	3z16666	3z16122	3z15595	3z15084	3z14587	3z14106	3z13639	3z13187	3z12748
3.4	3z12322	3z11910	3z11510	3z11122	3z10747	3z10383	3z10030	4z96886	4z93577	4z90372
3.5	4z87268	4z84263	4z81352	4z78534	4z75807	4z73166	4z70611	4z68138	4z65745	4z63430
3.6	4z61190	4z59024	4z56928	4z54901	4z52941	4z51046	4z49214	4z47443	4z45731	4z44077
3.7	4z42478	4z40933	4z39440	4z37998	4z36605	4z35260	4z33960	4z32705	4z31494	4z30324
3.8	4z29195	4z28105	4z27053	4z26037	4z25058	4z24113	4z23201	4z22321	4z21473	4z20655
3.9	4z19866	4z19105	4z18371	4z17664	4z16983	4z16326	4z15693	4z15083	4z14495	4z13928
Z	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
4.0	4z13383	4z12858	4z12352	4z11864	4z11395	4z10943	4z10509	4z10090	5z96870	5z92993
4.1	5z89262	5z85672	5z82218	5z78895	5z75700	5z72626	5z69670	5z66828	5z64095	5z61468
4.2	5z58943	5z56516	5z54183	5z51942	5z49788	5z47719	5z45731	5z43821	5z41988	5z40226
4.3	5z38535	5z36911	5z35353	5z33856	5z32420	5z31041	5z29719	5z28449	5z27231	5z26063
4.4	5z24942	5z23868	5z22837	5z21848	5z20900	5z19992	5z19121	5z18286	5z17486	5z16719
4.5	5z15984	5z15280	5z14605	5z13959	5z13340	5z12747	5z12180	5z11636	5z11116	5z10618
4.6	5z10141	6z96845	6z92477	6z88297	6z84298	6z80472	6z76812	6z73311	6z69962	6z66760
4.7	6z63698	6z60771	6z57972	6z55296	6z52739	6z50295	6z47960	6z45728	6z43596	6z41559
4.8	6z39613	6z37755	6z35980	6z34285	6z32667	6z31122	6z29647	6z28239	6z26895	6z25613
4.9	6z24390	6z23222	6z22108	6z21046	6z20033	6z19066	6z18144	6z17265	6z16428	6z15629
5.0	6z14867	6z14141	6z13450	6z12791	6z12162	6z11564	6z10994	6z10451	7z99339	7z94414

სტანდარტული ნორმალური განაწილების ფუნქციის ($\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$) მნიშვნელობები

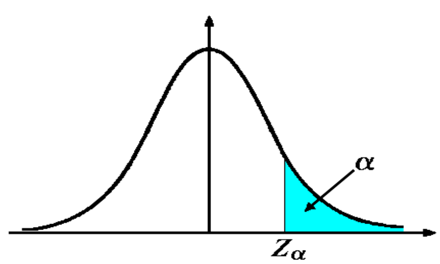


x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.00	0.500	**0.33**	0.629	**0.66**	0.745	**0.99**	0.838	**1.32**	0.906	**1.65**	0.950
0.01	0.503	**0.34**	0.633	**0.67**	0.748	**1.00**	0.841	**1.33**	0.908	**1.66**	0.951
0.02	0.507	**0.35**	0.636	**0.68**	0.751	**1.01**	0.843	**1.34**	0.909	**1.67**	0.952
0.03	0.511	**0.36**	0.640	**0.69**	0.754	**1.02**	0.846	**1.35**	0.911	**1.68**	0.953
0.04	0.515	**0.37**	0.644	**0.70**	0.758	**1.03**	0.848	**1.36**	0.913	**1.69**	0.954
0.05	0.519	**0.38**	0.648	**0.71**	0.761	**1.04**	0.850	**1.37**	0.914	**1.70**	0.955
0.06	0.523	**0.39**	0.651	**0.72**	0.764	**1.05**	0.853	**1.38**	0.916	**1.71**	0.956
0.07	0.527	**0.40**	0.655	**0.73**	0.767	**1.06**	0.855	**1.39**	0.917	**1.72**	0.957
0.08	0.531	**0.41**	0.659	**0.74**	0.770	**1.07**	0.857	**1.40**	0.919	**1.73**	0.958
0.09	0.535	**0.42**	0.662	**0.75**	0.773	**1.08**	0.859	**1.41**	0.920	**1.74**	0.959
0.10	0.539	**0.43**	0.666	**0.76**	0.776	**1.09**	0.862	**1.42**	0.922	**1.75**	0.959
0.11	0.543	**0.44**	0.670	**0.77**	0.779	**1.10**	0.864	**1.43**	0.923	**1.76**	0.960
0.12	0.547	**0.45**	0.673	**0.78**	0.782	**1.11**	0.866	**1.44**	0.925	**1.77**	0.961
0.13	0.551	**0.46**	0.677	**0.79**	0.785	**1.12**	0.868	**1.45**	0.926	**1.78**	0.962
0.14	0.555	**0.47**	0.680	**0.80**	0.788	**1.13**	0.870	**1.46**	0.927	**1.79**	0.963
0.15	0.559	**0.48**	0.684	**0.81**	0.791	**1.14**	0.872	**1.47**	0.929	**1.80**	0.964
0.16	0.563	**0.49**	0.687	**0.82**	0.793	**1.15**	0.874	**1.48**	0.930	**1.81**	0.964
0.17	0.567	**0.50**	0.691	**0.83**	0.796	**1.16**	0.876	**1.49**	0.931	**1.82**	0.965
0.18	0.571	**0.51**	0.694	**0.84**	0.799	**1.17**	0.879	**1.50**	0.933	**1.83**	0.966
0.19	0.575	**0.52**	0.698	**0.85**	0.802	**1.18**	0.881	**1.51**	0.934	**1.84**	0.967
0.20	0.579	**0.53**	0.701	**0.86**	0.805	**1.19**	0.882	**1.52**	0.935	**1.85**	0.967
0.21	0.583	**0.54**	0.705	**0.87**	0.807	**1.20**	0.884	**1.53**	0.936	**1.86**	0.968
0.22	0.587	**0.55**	0.708	**0.88**	0.810	**1.21**	0.886	**1.54**	0.938	**1.87**	0.969
0.23	0.590	**0.56**	0.712	**0.89**	0.813	**1.22**	0.888	**1.55**	0.939	**1.88**	0.969
0.24	0.594	**0.57**	0.715	**0.90**	0.815	**1.23**	0.890	**1.56**	0.940	**1.89**	0.970
0.25	0.598	**0.58**	0.719	**0.91**	0.818	**1.24**	0.892	**1.57**	0.941	**1.90**	0.971
0.26	0.602	**0.59**	0.722	**0.92**	0.821	**1.25**	0.894	**1.58**	0.942	**1.91**	0.971
0.27	0.606	**0.60**	0.725	**0.93**	0.823	**1.26**	0.896	**1.59**	0.944	**1.92**	0.972
0.28	0.610	**0.61**	0.729	**0.94**	0.826	**1.27**	0.897	**1.60**	0.945	**1.93**	0.973
0.29	0.614	**0.62**	0.732	**0.95**	0.828	**1.28**	0.899	**1.61**	0.946	**1.94**	0.973
0.30	0.617	**0.63**	0.735	**0.96**	0.831	**1.29**	0.901	**1.62**	0.947	**1.95**	0.974
0.31	0.621	**0.64**	0.738	**0.97**	0.833	**1.30**	0.903	**1.63**	0.948	**1.96**	0.975
0.32	0.625	**0.65**	0.742	**0.98**	0.836	**1.31**	0.904	**1.64**	0.949	**1.97**	0.975

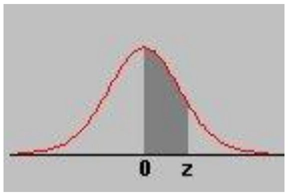
$\Phi(x)$-ის მნიშვნელობები (გაგრძელება)

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
1.98	0.976	**2.26**	0.988	**2.54**	0.994	**2.82**	0.997	**3.10**	0.999
1.99	0.976	**2.27**	0.988	**2.55**	0.994	**2.83**	0.997	**3.11**	0.999
2.00	0.977	**2.28**	0.988	**2.56**	0.994	**2.84**	0.997	**3.12**	0.999
2.01	0.977	**2.29**	0.988	**2.57**	0.994	**2.85**	0.997	**3.13**	0.999
2.02	0.978	**2.30**	0.989	**2.58**	0.995	**2.86**	0.997	**3.14**	0.999
2.03	0.978	**2.31**	0.989	**2.59**	0.995	**2.87**	0.997	**3.15**	0.999
2.04	0.979	**2.32**	0.989	**2.60**	0.995	**2.88**	0.998	**3.16**	0.999
2.05	0.979	**2.33**	0.990	**2.61**	0.995	**2.89**	0.998	**3.17**	0.999
2.06	0.980	**2.34**	0.990	**2.62**	0.995	**2.90**	0.998	**3.18**	0.999
2.07	0.980	**2.35**	0.990	**2.63**	0.995	**2.91**	0.998	**3.19**	0.999
2.08	0.981	**2.36**	0.990	**2.64**	0.995	**2.92**	0.998	**3.20**	0.999
2.09	0.981	**2.37**	0.991	**2.65**	0.995	**2.93**	0.998	**3.21**	0.999
2.10	0.982	**2.38**	0.991	**2.66**	0.996	**2.94**	0.998	**3.22**	0.999
2.11	0.982	**2.39**	0.991	**2.67**	0.996	**2.95**	0.998	**3.23**	0.999
2.12	0.983	**2.40**	0.991	**2.68**	0.996	**2.96**	0.998	**3.24**	0.999
2.13	0.983	**2.41**	0.992	**2.69**	0.996	**2.97**	0.998	**3.25**	0.999
2.14	0.983	**2.42**	0.992	**2.70**	0.996	**2.98**	0.998	**3.26**	0.999
2.15	0.984	**2.43**	0.992	**2.71**	0.996	**2.99**	0.998	**3.27**	0.999
2.16	0.984	**2.44**	0.992	**2.72**	0.996	**3.00**	0.998	**3.28**	0.999
2.17	0.985	**2.45**	0.992	**2.73**	0.996	**3.01**	0.998	**3.29**	0.999
2.18	0.985	**2.46**	0.993	**2.74**	0.996	**3.02**	0.998	**3.30**	0.999
2.19	0.985	**2.47**	0.993	**2.75**	0.997	**3.03**	0.998	**3.31**	0.999
2.20	0.986	**2.48**	0.993	**2.76**	0.997	**3.04**	0.998	**3.32**	0.999
2.21	0.986	**2.49**	0.993	**2.77**	0.997	**3.05**	0.998	**3.33**	0.999
2.22	0.986	**2.50**	0.993	**2.78**	0.997	**3.06**	0.998	**3.34**	0.999
2.23	0.987	**2.51**	0.993	**2.79**	0.997	**3.07**	0.998	**3.35**	0.999
2.24	0.987	**2.52**	0.994	**2.80**	0.997	**3.08**	0.998	**3.36**	0.999
2.25	0.987	**2.53**	0.994	**2.81**	0.997	**3.09**	0.999	**3.37**	0.999

სტანდარტული ნორმალური განაწილების ზედა α-კრიტიკული წერტილები (z_α)



α	0.1	0.05	0.025	0.125	0.01	0.005	0.0025	0.001
z_α	1.28	1.64	1.96	2.24	2.33	2.57	2.81	3.08

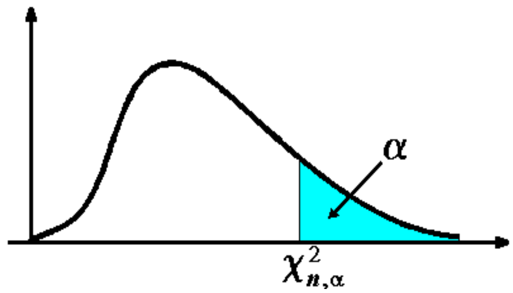


$$\Phi_0(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt$$

ფუნქციის ცხრილები

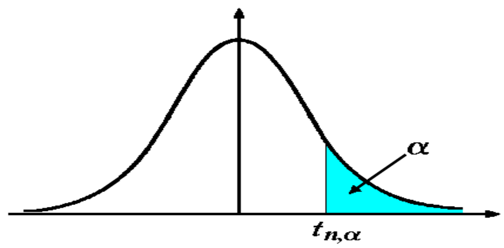
	0.00	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

χ^2 (ხი კვადრატ) განაწილების ზედა α-კრიტიკული წერტილები ($\chi^2_{\alpha,n}$)



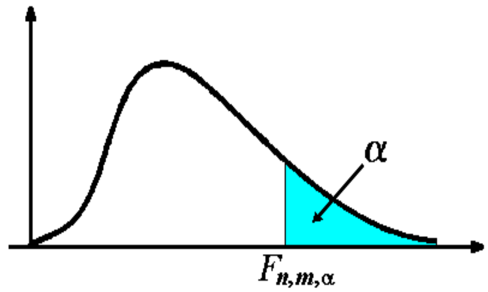
	α							
n	**0.99**	**0.975**	**0.95**	**0.9**	**0.1**	**0.05**	**0.025**	**0.01**
1	0.0002	0.0010	0.0039	0.0158	2.7055	3.8415	5.0239	6.6349
2	0.0201	0.0506	0.1026	0.2107	4.6052	5.9915	7.3778	9.2104
3	0.1148	0.2158	0.3518	0.5844	6.2514	7.8147	9.3484	11.3449
4	0.2971	0.4844	0.7107	1.0636	7.7794	9.4877	11.1433	13.2767
5	0.5543	0.8312	1.1455	1.6103	9.2363	11.0705	12.8325	15.0863
6	0.8721	1.2373	1.6354	2.2041	10.6446	12.5916	14.4494	16.8119
7	1.2390	1.6899	2.1673	2.8331	12.0170	14.0671	16.0128	18.4753
8	1.6465	2.1797	2.7326	3.4895	13.3616	15.5073	17.5345	20.0902
9	2.0879	2.7004	3.3251	4.1682	14.6837	16.9190	19.0228	21.6660
10	2.5582	3.2470	3.9403	4.8652	15.9872	18.3070	20.4832	23.2093
11	3.0535	3.8157	4.5748	5.5778	17.2750	19.6752	21.9200	24.7250
12	3.5706	4.4038	5.2260	6.3038	18.5493	21.0261	23.3367	26.2170
13	4.1069	5.0087	5.8919	7.0415	19.8119	22.3620	24.7356	27.6882
14	4.6604	5.6287	6.5706	7.7895	21.0641	23.6848	26.1189	29.1412
15	5.2294	6.2621	7.2609	8.5468	22.3071	24.9958	27.4884	30.5780
16	5.8122	6.9077	7.9616	9.3122	23.5418	26.2962	28.8453	31.9999
17	6.4077	7.5642	8.6718	10.0852	24.7690	27.5871	30.1910	33.4087
18	7.0149	8.2307	9.3904	10.8649	25.9894	28.8693	31.5264	34.8052
19	7.6327	8.9065	10.1170	11.6509	27.2036	30.1435	32.8523	36.1908
20	8.2604	9.5908	10.8508	12.4426	28.4120	31.4104	34.1696	37.5663
21	8.8972	10.2829	11.5913	13.2396	29.6151	32.6706	35.4789	38.9322
22	9.5425	10.9823	12.3380	14.0415	30.8133	33.9245	36.7807	40.2894
23	10.1957	11.6885	13.0905	14.8480	32.0069	35.1725	38.0756	41.6383
24	10.8563	12.4011	13.8484	15.6587	33.1962	36.4150	39.3641	42.9798
25	11.5240	13.1197	14.6114	16.4734	34.3816	37.6525	40.6465	44.3140
26	12.1982	13.8439	15.3792	17.2919	35.5632	38.8851	41.9231	45.6416
27	12.8785	14.5734	16.1514	18.1139	36.7412	40.1133	43.1945	46.9628
28	13.5647	15.3079	16.9279	18.9392	37.9159	41.3372	44.4608	48.2782
29	14.2564	16.0471	17.7084	19.7677	39.0875	42.5569	45.7223	49.5878
30	14.9535	16.7908	18.4927	20.5992	40.2560	43.7730	46.9792	50.8922

t (სტიუდენტის) განაწილების ზედა α-კრიტიკული წერტილები ($t_{\alpha,n}$)



	α						
n	**0.1**	**0.05**	**0.025**	**0.01**	**0.005**	**0.0025**	**0.001**
1	3.078	6.314	12.706	31.821	63.656	127.321	318.289
2	1.886	2.920	4.303	6.965	9.925	14.089	22.328
3	1.638	2.353	3.182	4.541	5.841	7.453	10.214
4	1.533	2.132	2.776	3.747	4.604	5.598	7.173
5	1.476	2.015	2.571	3.365	4.032	4.773	5.894
6	1.440	1.943	2.447	3.143	3.707	4.317	5.208
7	1.415	1.895	2.365	2.998	3.499	4.029	4.785
8	1.397	1.860	2.306	2.896	3.355	3.833	4.501
9	1.383	1.833	2.262	2.821	3.250	3.690	4.297
10	1.372	1.812	2.228	2.764	3.169	3.581	4.144
11	1.363	1.796	2.201	2.718	3.106	3.497	4.025
12	1.356	1.782	2.179	2.681	3.055	3.428	3.930
13	1.350	1.771	2.160	2.650	3.012	3.372	3.852
14	1.345	1.761	2.145	2.624	2.977	3.326	3.787
15	1.341	1.753	2.131	2.602	2.947	3.286	3.733
16	1.337	1.746	2.120	2.583	2.921	3.252	3.686
17	1.333	1.740	2.110	2.567	2.898	3.222	3.646
18	1.330	1.734	2.101	2.552	2.878	3.197	3.610
19	1.328	1.729	2.093	2.539	2.861	3.174	3.579
20	1.325	1.725	2.086	2.528	2.845	3.153	3.552
21	1.323	1.721	2.080	2.518	2.831	3.135	3.527
22	1.321	1.717	2.074	2.508	2.819	3.119	3.505
23	1.319	1.714	2.069	2.500	2.807	3.104	3.485
24	1.318	1.711	2.064	2.492	2.797	3.091	3.467
25	1.316	1.708	2.060	2.485	2.787	3.078	3.450
26	1.315	1.706	2.056	2.479	2.779	3.067	3.435
27	1.314	1.703	2.052	2.473	2.771	3.057	3.421
28	1.313	1.701	2.048	2.467	2.763	3.047	3.408
29	1.311	1.699	2.045	2.462	2.756	3.038	3.396
30	1.310	1.697	2.042	2.457	2.750	3.030	3.385

$F(n,m)$ (ფიშერის) განაწილების ზედა $\alpha = 0.05$ კრიტიკული წერტილები ($F_{n,m,\alpha}$)



m \ n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	50	100
1	161	199	215	224	230	234	236	238	240	241	243	243	244	245	245	246	246	247	247	248	251	253
2	18.5	19.0	19.1	19.2	19.3	19.3	19.3	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.4
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74	8.73	8.71	8.70	8.69	8.68	8.67	8.67	8.66	8.58	8.55
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.89	5.87	5.86	5.84	5.83	5.82	5.81	5.80	5.70	5.66
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.66	4.64	4.62	4.60	4.59	4.58	4.57	4.56	4.44	4.41
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.98	3.96	3.94	3.92	3.91	3.90	3.88	3.87	3.75	3.71
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.55	3.53	3.51	3.49	3.48	3.47	3.46	3.44	3.32	3.27
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.26	3.24	3.22	3.20	3.19	3.17	3.16	3.15	3.02	2.97
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.05	3.03	3.01	2.99	2.97	2.96	2.95	2.94	2.80	2.76
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.89	2.86	2.85	2.83	2.81	2.80	2.79	2.77	2.64	2.59
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.76	2.74	2.72	2.70	2.69	2.67	2.66	2.65	2.51	2.46
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.66	2.64	2.62	2.60	2.58	2.57	2.56	2.54	2.40	2.35
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60	2.58	2.55	2.53	2.51	2.50	2.48	2.47	2.46	2.31	2.26
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53	2.51	2.48	2.46	2.44	2.43	2.41	2.40	2.39	2.24	2.19
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48	2.45	2.42	2.40	2.38	2.37	2.35	2.34	2.33	2.18	2.12
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42	2.40	2.37	2.35	2.33	2.32	2.30	2.29	2.28	2.12	2.07
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38	2.35	2.33	2.31	2.29	2.27	2.26	2.24	2.23	2.08	2.02
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.31	2.29	2.27	2.25	2.23	2.22	2.20	2.19	2.04	1.98
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31	2.28	2.26	2.23	2.21	2.20	2.18	2.17	2.16	2.00	1.94
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	2.25	2.22	2.20	2.18	2.17	2.15	2.14	2.12	1.97	1.91
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.99	1.95	1.92	1.89	1.87	1.85	1.83	1.81	1.80	1.78	1.60	1.52
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97	1.93	1.89	1.85	1.82	1.79	1.77	1.75	1.73	1.71	1.69	1.68	1.48	1.39

კოლმოგოროვის განაწილების ($K(\lambda) = \sum_{m=-\infty}^{\infty} (-1)^m e^{-2m^2\lambda^2}$) კვანტილები

α	0.20	0.10	0.05	0.02	0.01
$K_{1-\alpha}$	1.073	1.224	1.358	1.517	1.628

D_n სტატისტიკის კრიტიკული $k_{n,\alpha}$ მნიშვნელობები

n	$\alpha = 0.05$	$\alpha = 0.02$	$\alpha = 0.01$
1	0.975	0.990	0.995
5	0.563	0.627	0.668
10	0.409	0.456	0.489
15	0.338	0.377	0.404
20	0.294	0.329	0.352
25	0.264	0.295	0.326
30	0.243	0.270	0.290
35	0,224	0.250	0.268
40	0.210	0.234	0.252
45	0.198	0.221	0.232
50	0.188	0.210	0.226
55	0.180	0.201	0.216
60	0.172	0.193	0.207
70	0.160	0.179	0.192
80	0.150	0.167	0.179
90	0.141	0.158	0.169
100	0.134	0.150	0.161

შემთხვევითი რიცხვების ცხრილი
39634 62349 74088 65564 16379 19713 39153 69459 17986 24537
14595 35050 40469 27478 44526 67331 93365 54526 22356 93208
30734 71571 83722 79712 25775 65178 07763 82928 31131 30196
64628 89126 91254 24090 25752 03091 39411 73146 06089 15630
42831 95113 43511 42082 15140 34733 68076 18292 69486 80468
80583 70361 41047 26792 78466 03395 17635 09697 82447 31405
00209 90404 99457 72570 42194 49043 24330 14939 09865 45906
05409 20830 01911 60767 55248 79253 12317 84120 77772 50103
95836 22530 91785 80210 34361 52228 33869 94332 83868 61672
65358 70469 87149 89509 72176 18103 55169 79954 72002 20582
72249 04037 36192 40221 14918 53437 60571 40995 55006 10694
41692 40581 93050 48734 34652 41577 04631 49184 39295 81776
61885 50796 96822 82002 07973 52925 75467 86013 98072 91942
48917 48129 48624 48248 91465 54898 61220 18721 67387 66575
88378 84299 12193 03785 49314 39761 99132 28775 45276 91816
77800 25734 09801 92087 02955 12872 89848 48579 06028 13827
24028 03405 01178 06316 81916 40170 53665 87202 88638 47121
86558 84750 43994 01760 96205 27937 45416 71964 52261 30781
78545 49201 05329 14182 10971 90472 44682 39304 19819 55799
14969 64623 82780 35686 30941 14622 04126 25498 95452 63937
58697 31973 06303 94202 62287 56164 79157 98375 24558 99241
38449 46438 91579 01907 72146 05764 22400 94490 49833 09258
62134 87244 73348 80114 78490 64735 31010 66975 28652 36166
72749 13347 65030 26128 49067 27904 49953 74674 94617 13317
81638 36566 42709 33717 59943 12027 46547 61303 46699 76243

შემთხვევითი რიცხვების ცხრილის გაგრძელება
46574 79670 10342 89543 75030 23428 29541 32501 89422 87474
11873 57196 32209 67663 07990 12288 59245 83638 23642 61715
13862 72778 09949 23096 01791 19472 14634 31690 36602 62943
08312 27886 82321 28666 72998 22514 51054 22940 31842 54245
11071 44430 94664 91294 35163 05494 32882 23904 41340 61185
82509 11842 86963 50307 07510 32545 90717 46856 86079 13769
07426 67341 80314 58910 93948 85738 69444 09370 58194 28207
57696 25592 91221 95386 15857 84645 89659 80535 93233 82798
08074 89810 48521 90740 02687 83117 74920 25954 99629 78978
20128 53721 01518 40699 20849 04710 38989 91322 56057 58573
00190 27157 83208 79446 92987 61357 38752 55424 94518 45205
23798 55425 32454 34611 39605 39981 74691 40836 30812 38563
85306 57995 68222 39055 43890 36956 84861 63624 04961 55439
99719 36036 74274 53901 34643 06157 89500 57514 93977 42403
95970 81452 48873 00784 58347 40269 11880 43395 28249 38743
56651 91460 92462 98566 72062 18556 55052 47614 80044 60015
71499 80220 35750 67337 47556 55272 55249 79100 34014 17037
66660 78443 47545 70736 65419 77489 70831 73237 14970 23129
35483 84563 79956 88618 54619 24853 59783 47537 88822 47227
09262 25041 57862 19203 86103 02800 23198 70639 43757 52064

მილიონი შემთხვევითი რიცხვის სიხშირეები

No.	0	1	2	3	4	5	6	7	8	9	χ^2
1	4923	5013	4916	4951	5109	4993	5055	5080	4986	4974	7.556
2	4870	4956	5080	5097	5066	5034	4902	4974	5012	5009	10.132
3	5065	5014	5034	5057	4902	5061	4942	4946	4960	5019	6.078
4	5009	5053	4966	4891	5031	4895	5037	5062	5170	4886	15.004
5	5033	4982	5180	5074	4892	4992	5011	5005	4959	4872	13.846
6	4976	4993	4932	5039	4965	5034	4943	4932	5116	5070	7.076
7	5011	5152	4990	5047	4974	5107	4869	4925	5023	4902	14.116
8	5003	5092	5163	4936	5020	5069	4914	4943	4914	4946	13.051
9	4860	4899	5138	4959	5089	5047	5030	5039	5002	4937	13.410
10	4998	4957	4964	5124	4909	4995	5053	4946	4995	5059	7.212
11	4948	5048	5041	5077	5051	5004	5024	4886	4917	5004	7.142
12	4958	4993	5064	4987	5041	4984	4991	4987	5113	4882	6.992
13	4968	4961	5029	5038	5022	5023	5010	4988	4936	5025	2.162
14	5110	4923	5025	4975	5095	5051	5035	4962	4942	4882	10.172
15	5094	4962	4945	4891	5014	5002	5038	5023	5179	4852	16.261
16	4957	5035	5051	5021	5036	4927	5022	4988	4910	5053	4.856
17	5088	4989	5042	4948	4999	5028	5037	4893	5004	4972	5.347
18	4970	5034	4996	5008	5049	5016	4954	4989	4970	5014	1.625
19	4998	4981	4984	5107	4874	4980	5057	5020	4978	5021	6.584
20	4963	5013	5101	5084	4956	4972	5018	4971	5021	4901	6.584
$\sum$	99802	100050	100641	100311	100094	100214	99942	99559	100107	99280	13.316

ლიტერატურა

1. გ. მანია. ალბათობის თეორია და მათემატიკური სტატისტიკა. სახელმძღვანელო ეკონომიკური ფაკულტეტის სტუდენტებისათვის. თსუ, თბილისი, 1976
2. ნ. ანთელავა, ა. ედიბერიძე, გ. მანია. ალბათობის თეორიისა და მათემატიკური სტატისტიკის ამოცანათა კრებული. თსუ, თბილისი, 1980.
3. მარი გ., მოსიძე ა., ციგროშვილი ზ., სტატისტიკა. დამხმარე სახელმძღვანელო ESM-თბილისის სტუდენტებისათვის, ESM-თბილისი, 1996.
4. ნ. ლაზრიევა, მ. მანია, გ. მარი, ა. მოსიძე, ა. ტორონჯაძე, თ. ტორონჯაძე, თ. შერვაშიძე. ალბათობის თეორია და მათემატიკური სტატისტიკა ეკონომისტებისათვის. ფონდი «ევრაზია», თბილისი, 2000.
5. ო. ფურთუხია. ალბათობის თეორია და მათემატიკური სტატისტიკა. თბილისი, 2007.
6. ო.ფურთუხია. აღწერითი სტატისტიკა, ალბათობა, სტატისტიკური დასკვნების თეორია. თბილისი, 2008.
7. ო. ფურთუხია. ალბათობა და სტატისტიკა მაგალითებსა და ამოცანებში. თბილისი, 2009.
8. ო. ფურთუხია, ზ. ციგროშვილი, ქ. მანჯგალაძე. ალბათობა და მათემატიკური სტატისტიკა. თსუ, თბილისი, 2009.
9. ო. ფურთუხია. ალბათურ-სტატისტიკური ამოცანები. თსუ გამომცემლობა, თბილისი 2012.
10. ე. ნადარაია, ბ. დოჭვირი, პ. ბაბილუა, მ. ბერაძე, გ. ლომინაშვილი, მ. მნაცაკანიანი. ალბათობის თეორის ამოცანათა კრებული. ქუთაისი 2008.
11. ე. ნადარაია, ბ. დოჭვირი, თ. ბოკელავაძე, გ. ლომინაშვილი, მ. მნაცაკანიანი, მ. ფაცაცია. ალბათობის თეორია და მათემატიკური სტატისტიკა (ამოცანათა კრებული). ქუთაისი 2008.
12. Allan G. Bluman. Ementary Statistics: a brief version, second edition. Published by McGraw-Hill, New York, 2003.
13. P. Newbold, W. L. Carlson, B. M. Thorne. Statistics for Business and Economics, sixth edition. Prentice Hall, Upper Saddle River, New Jersey, 2007.
14. В. Феллер. ВВедение в теорию вероятностей и ее приложения. Москва, 1967.
15. УА. Г. Дьячков. Теория вероятностей. Москва, 1980.
16. А. Н. Колмогоров, И. Г. Журбенко, А. В. Прохоров. ВВедение в теорию вероятностей. Москва, 1982.
17. В. К. Захаров, Б. А. Севастьянов, В. П. Чистяков. Теория вероятностей. Москва, 1988.
18. В. Е. Гмурман. Руководство к решению задач по теории вероятностей и математической статистике. Москва, 1988.