

ზაქარია გიუნაშვილი

სტატისტიკის ელემენტები, რომლებიც პრაქტიკაში გამოგვადგება

შესავალი

ტრადიციულად, სასკოლო მათემატიკის კურსში ძირითადი ყურადღება ექცევა *აღწერითი სტატისტიკის* მიმართულებას, რომლის ძირითადი მიზანია ამა თუ იმ ვითარებასთან დაკავშირებული მონაცემების რიცხვითი მახასიათებლების დადგენა. მაგალითად: მოსახლეობის საშუალო სიმჭიდროვე, საშუალო ხელფასი, ნალექების საშუალო წლიური მაჩვენებელი, ყველაზე მოთხოვნადი პროდუქტი (რაოდენობრივი მონაცემების მოდა).

აღწერითი სტატისტიკის დანიშნულება, პირველ რიგში, არის ის, რომ იგი იძლევა შემჭიდროებულ და კონცენტრირებულ ინფორმაციას შესასწავლი მოვლენის შესახებ. მაგალითად, ქვეყნის მოსახლეობის აღწერის მონაცემები შეიცავს სრულ ინფორმაციას მოსახლეობის შესახებ, მაგრამ ამ შემთხვევაში მონაცემების რაოდენობა იმდენად დიდია, რომ თითქმის უსარგებლოა მხოლოდ მონაცემთა შეგროვების ეტაპით შემოფარგვლა. მონაცემთა შეგროვების შემდეგ მნიშვნელოვანია ისეთი აღწერითი შინაარსის მონაცემები, თუ როგორია მოსახლეობის საშუალო სიმჭიდროვე სხვადასხვა რეგიონში, განაწილება ასაკობრივი ჯგუფების მიხედვით და სხვა. ამ სახის მახასიათებლების დადგენა თავისთავად მოითხოვს იმ პროცედურების შესრულებას, რომლებიც აქტუალურია სასკოლო მათემატიკის კურსში. ამდენად, შეიძლება ჩავთვალოთ, რომ აღწერითი სტატისტიკის სწავლების ერთ-ერთი ძირითადი დანიშნულება ამ პროცედურული ცოდნის განმტკიცებაა. მიუხედავად ამისა, არანაკლებ მნიშვნელოვანია, მეტი ყურადღება მიექცეს სტატისტიკის მეორე მიმართულებას, რომელსაც ეწოდება *ინფერენციალური (დასკვნებზე ორიენტირებული) სტატისტიკა*. მისი ძირითადი მიზანია იმ მეთოდების შემუშავება, რომლებიც საშუალებას მოგვცემს, გამოვიტანოთ დასკვნები სტატისტიკური მონაცემების საფუძველზე, განვაზოგადოთ შერჩევითი მონაცემების ანალიზის შედეგები. ყოველივე ამას საფუძვლად უდევს ის ფაქტი, რომ შერჩევითი დაკვირვების შედეგების მიხედვით მიღებული დასკვნები შეიძლება განვავრცოთ მთლიან ერთობლიობაზე. ამასთან, ალბათობის თეორიის მეთოდების გამოყენება საშუალებას გვაძლევს, შევაფასოთ ამ დასკვნების სანდოობა. აქედან გამომდინარე, ინფერენციალური სტატისტიკის ამოცანებს შორისაა ისეთი შერჩევითი დაკვირვებების ხერხების შემუშავება, რომლებიც უზრუნველყოფს მათ საფუძველზე მიღებული დასკვნების სანდოობას.

შემთხვევითი სიდიდეები

მეცნიერულ კვლევაში, ყოველდღიურ ცხოვრებაში, წარმოებაში და ადამიანის სხვა საქმიანობაში მუდმივად ვაწყდებით ისეთ ვითარებებს, როდესაც ჩვენთვის საინტერესო სიდიდე იღებს სხვადასხვა მნიშვნელობებს, რომლებიც შემთხვევით გარემოებებზეა დამოკიდებული. რამდენი წუნდებული საქონელი აღმოჩნდება საქონლის პარტიის შერჩევითი შემოწმების შემდეგ? რისი ტოლი იქნება ექპერიმენტის დროს რომელიმე

სიდიდის გაზომვის შედეგი? რისი ტოლი იქნება მრიცხველის მიერ რეგისტრირებული კოსმოსური ნაწილაკების რაოდენობა? ასეთი სიდიდეების წინასწარი ცოდნა შეუძლებელია. მართალია, თითოეული დაკვირვების შედეგად მიიღება სრულიად გარკვეული, კონკრეტული რიცხვი, მაგრამ ეს რიცხვი შესაძლოა შეიცვალოს ექსპერიმენტის გამეორების შედეგად. მაგალითად, თუ საქონლის პარტიიდან, რომელიც შეიცავს 100 ერთეულს, შემთხვევით ვირჩევთ 10 ერთეულს და მათ შორის წუნდებული აღმოჩნდა 2 ერთეული, 10 ერთეულის შემთხვევითი შერჩევის გამეორების შედეგად წუნდებული შეიძლება აღმოჩნდეს 1 ცალი. რიცხვით მნიშვნელობას, რომელიც მიიღება შემთხვევითი ექსპერიმენტის შედეგად, უწოდებენ **შემთხვევით სიდიდეს**. იმისათვის, რომ ნათელი წარმოდგენა შეგვექმნას ამ ცნების შესახებ, განვიხილოთ ასეთი მაგალითი: მონეტის აგდება შემთხვევითი ექსპერიმენტია; ამ ექსპერიმენტის შედეგია ორი შესაძლო ხდომილება (მნიშვნელობა) - გერბი ან ცხაური; ეს ორი მნიშვნელობა არ არის რიცხვითი, ამიტომ შეიძლება ითქვას, რომ ამ შემთხვევაში საქმე არ გვაქვს შემთხვევით სიდიდესთან. მაგრამ თუ ამ ხდომილებებს რიცხვებს შევუსაბამებთ, მაგალითად: გერბი - 0, ცხაური - 1, მივიღებთ შემთხვევით სიდიდეს, რომლის მნიშვნელობათა სიმრავლეა ელემენტარული რიცხვითი სიმრავლე - $\{0, 1\}$. შემთხვევითი სიდიდის კიდევ ერთი გავრცელებული მაგალითია კამათლების წყვილის გაგორება. ამ შემთხვევითი ექსპერიმენტის ხდომილებათა სივრცეა (m, n) სახის წყვილების სიმრავლე, სადაც m და n ღებულობს მნიშვნელობებს 1 - დან 6-ის ჩათვლით. ეს თავისთავად არ არის შემთხვევითი სიდიდე, მაგრამ ზოგჯერ სინამდვილეში მნიშვნელობა აქვს არა კამათლებზე "მოსული" რიცხვების წყვილს, არამედ მათ ჯამს, $m + n$ - ს. ეს უკანასკნელი კი შემთხვევითი სიდიდეა: **კამათლების წყვილის გაგორების შედეგად მიღებული რიცხვების ჯამი**. რომ შევაჯამოთ, ჩამოვაყალიბოთ შემთხვევითი სიდიდის მათემატიკური განმარტება. როგორც ცნობილია, შემთხვევითი ექსპერიმენტის ალბათური მოდელის საყრდენს წარმოადგენს **ელემენტარული ხდომილებების სივრცე**, რომელიც, მარტივად რომ ვთქვათ, წარმოადგენს ექსპერიმენტის (დაკვირვების) შესაძლო შედეგების სიმრავლეს. შემთხვევითი სიდიდე განიმარტება, როგორც **რიცხვითი ფუნქცია შემთხვევითი ექსპერიმენტის ელემენტარული ხდომილებების სივრცეზე**. მონაცემთა ანალიზის ერთ-ერთი მნიშვნელოვანი ასპექტია შემთხვევითი სიდიდის სხვადასხვა მახასიათებლის გამოთვლა და მათ საფუძველზე შემთხვევითი მოვლენების შესახებ კანონზომიერებების დადგენა.

შემთხვევითი სიდიდე შეიძლება იყოს **დისკრეტული** ან **უწყვეტი**. დისკრეტულია ისეთი შემთხვევითი სიდიდე, რომლის შესაძლო მნიშვნელობების ჩამოთვლა შესაძლებელია, ხოლო უწყვეტია ისეთი შემთხვევითი სიდიდე, რომელმაც შესაძლოა მიიღოს ნებისმიერი მნიშვნელობა რომელიმე რიცხვით ინტერვალში. ამ განსხვავების უკეთ გასააზრებლად განვიხილოთ ასეთი მაგალითი:

ტბაზე ბადით თევზაობისას შემთხვევითი ექსპერიმენტის ხდომილება შეიძლება განვიხილოთ, როგორც ბადის ყოველი ამოღების შემდეგ მასში მოხვედრილი თევზების ერთობლიობა; ამ ხდომილებას შეიძლება დავუკავშიროთ ორი სიდიდე:

I - ბადეში მოხვედრილი თევზების რაოდენობა;

II - ბადეში მოხვედრილი თევზების მთლიანი მასა.

ამ სიდიდეებიდან პირველი დისკრეტული სიდიდეა, რადგან მან შეიძლება მიიღოს მხოლოდ მთელი რიცხოვანი მნიშვნელობები: 0, 1, 2, 3..., ხოლო მეორე შეიძლება ჩავთვალოთ უწყვეტად, რადგან მასამ შეიძლება (თეორიულად) მიიღოს არა მხოლოდ მთელი რიცხოვანი მნიშვნელობები, არამედ ნებისმიერი მნიშვნელობა 0-დან თევზების იმ მაქსიმალურ მასამდე, რომლის დაჭერაც შესაძლებელია.

რა უნდა ვიცოდეთ შემთხვევითი სიდიდის შესახებ მის შესასწავლად? პირველი მახასიათებელია მისი შესაძლო მნიშვნელობების სიმრავლე. თუმცა უნდა ითქვას, რომ ეს საკმარისი არ არის. მაგალითად, კლასში მოსწავლეთა შემაჯამებელი შეფასების ქულა შეიძლება განვიხილოთ, როგორც შემთხვევითი სიდიდე, რომლის შესაძლო მნიშვნელობებია რიცხვები 0-დან 10-მდე, მაგრამ ეს ინფორმაცია არაფერს გვეუბნება მოცემული კლასის მიღწევის დონის შესახებ. მნიშვნელოვანია იმის ცოდნა, თუ რა სიხშირით ღებულობს შემთხვევითი სიდიდე სხვადასხვა მნიშვნელობას მისი ცვლილების დიაპაზონში. შესაბამისობას, რომელიც შემთხვევითი სიდიდის თითოეულ მნიშვნელობას შეუსაბამებს ამ მნიშვნელობის სიხშირეს, ეწოდება შემთხვევითი სიდიდის განაწილების კანონი. როდესაც შემთხვევითი ექსპერიმენტის ხდომილებათა სივრცე სასრული სიმრავლეა, შემთხვევითი სიდიდის განაწილების კანონის აღსაწერად შესაძლებელია ცხრილის გამოყენება. მაგალითად, შემდეგ ცხრილში წარმოდგენილია კლასში მოსწავლეთა მიერ მიღებული ქულების განაწილება:

ქულა	0	1	2	3	4	5	6	7	8	9	10
სიხშირე	0	2	0	2	5	2	3	5	2	3	1

ამ ცხრილიდან შეუძლებელია იმის დადგენა, თუ კლასის რომელმა მოსწავლემ რამდენი ქულა მიიღო. განაწილების კანონი ინფორმაციას გვაძლევს იმის შესახებ, თუ რა სიხშირით ღებულობს შემთხვევითი სიდიდე ამა თუ იმ მნიშვნელობას. მაგალითად, ამ შემთხვევაში, შეიძლება ითქვას, რომ 9-ქულიანი შეფასება მიიღო სამმა მოსწავლემ. როგორც წესი, შემთხვევითი სიდიდის განაწილების კანონის მოცემისას იყენებენ არა სიხშირეების აბსოლუტურ მნიშვნელობას, არამედ **ფარდობით სიხშირეებს**. ფარდობითი სიხშირე არის შემთხვევითი სიდიდის მოცემული მნიშვნელობის შესაბამისი აბსოლუტური სიხშირის შეფარდება ხდომილებათა სივრცის ელემენტების რაოდენობასთან. მაგალითად, ზემოაღნიშნული ცხრილის საფუძველზე შესაძლებელია დავასკვნათ, რომ ტესტირებაში მონაწილეობდა 25 მოსწავლე (აბსოლუტური სიხშირეების ჯამი), ხოლო (ფარდობითი) სიხშირეების ცხრილს ექნება ასეთი სახე:

ქულა	0	1	2	3	4	5	6	7	8	9	10
სიხშირე	0	0.08	0	0.08	0.2	0.08	0.12	0.2	0.08	0.12	0.04

შეიძლება ითქვას, რომ ეს ცხრილი გვაძლევს ინფორმაციას იმის შესახებ, თუ რა ალბათობით ღებულობს შემთხვევითი სიდიდე ამა თუ იმ მნიშვნელობას: ალბათობა იმისა, რომ ამ კლასში შემთხვევით შერჩეულმა მოსწავლემ ტესტირებაში მიიღო 6 ქულა, ტოლია $\frac{3}{25} = 0.12$ -ის. ეს ფაქტი შეიძლება ასე ჩაიწეროს: $P(X = 6) = 0.12$. სადაც X არის შემთხვევითი სიდიდე, რომელიც სიტყვიერად ასე აღიწერება: *"კლასში შემთხვევით შერჩეული მოსწავლის მიერ მიღებული ქულა"*. ე.ი. შეიძლება ითქვას, რომ დისკრეტული შემთხვევითი სიდიდის განაწილების კანონი აღიწერება ასეთი ცხრილით:

X შემთხვევითი სიდიდის მნიშვნელობა	x_1	x_2	$\dots$	x_n
ალბათობა P	p_1	p_2	$\dots$	p_n

ამ ცხრილში $x_1, x_2, \dots, x_n$ არის X შემთხვევითი სიდიდის ყველა შესაძლო განსხვავებული მნიშვნელობები, ხოლო $p_1 = P(X = x_1)$, $p_2 = P(X = x_2)$, $\dots$, $p_n = P(X = x_n)$ არის ის ალბათობები, რომლებითაც ეს შემთხვევითი სიდიდე ღებულობს შესაბამის მნიშვნელობას. რადგან შემთხვევითი სიდიდის განაწილების კანონის შესაბამისი ცხრილის მეორე სტრიქონში წარმოდგენილია ალბათობები, ადვილი მისახვედრია, რომ მათი ჯამი უნდა იყოს 1-ის ტოლი. შემთხვევითი სიდიდის განაწილების კანონის განსაზღვრა შესაძლებელია ექსპერიმენტის შედეგებზე დაკვირვებით, თუმცა როდესაც ზოგიერთი ხდომილების ალბათობა წინასწარ ცნობილია, განაწილების კანონის დასადგენად გამოიყენება ალბათობის გამოთვლის ხერხები.

მაგალითი

შევადგინოთ იმ შემთხვევითი სიდიდის განაწილების კანონი, რომელიც შეესაბამება ბურთის ორჯერ სროლის შედეგად კალათში ჩავარდნის რაოდენობას, თუ თითოეული სროლისას ჩავარდნის ალბათობა 0.7-ის ტოლია.

ამ ექსპერიმენტისას X შემთხვევითი სიდიდის შესაძლო მნიშვნელობების სიმრავლეა $\{0, 1, 2\}$ - *ორჯერ სროლის შედეგად ბურთი კალათში საერთოდ არ ჩავარდა, ჩავარდა მხოლოდ ერთხელ ან ჩავარდა ორივეჯერ*. იმისათვის, რომ შევადგინოთ ამ შემთხვევითი სიდიდის განაწილების კანონის ცხრილი, საჭიროა გამოვთვალოთ თითოეული მნიშვნელობის შესაბამისი ალბათობა. თუ მივყევით ალბათური მოდელის აგების სტანდარტულ სქემას, ჯერ უნდა დავადგინოთ ელემენტარული ხდომილებების სივრცე. ეს სივრცე შედგება 4 ელემენტისაგან: A - ბურთი საერთოდ არ ჩავარდნილა კალათში, B - პირველი ბურთი ჩავარდა, ხოლო მეორე - არა, C - პირველი არ ჩავარდა, მეორე ჩავარდა, C - ორივე ბურთი ჩავარდა კალათში. გამოვთვალოთ თითოეული ამ ხდომილების ალბათობა. როგორც ავღნიშნეთ, ამისათვის გამოვიყენოთ ალბათობის გამოთვლის ხერხები. თუ ბურთის ჩავარდნის ალბათობა 0.7-ის ტოლია, მაშინ არჩავარდნის ალბათობა იქნება $1 - 0.7 = 0.3$.

რადგან ბურთის სროლები დამოუკიდებელი ხდომილებებია, ამიტომ წყვილის როლის შედეგის ალბათობა ტოლია თითოეული სროლის შედეგების ალბათობების ნამრავლის:

$$P(A) = 0.3 \cdot 0.3 = 0.09,$$

$$P(B) = 0.3 \cdot 0.7 = 0.21,$$

$$P(C) = 0.7 \cdot 0.3 = 0.21,$$

$$P(D) = 0.7 \cdot 0.7 = 0.49.$$

შემთხვევითი სიდიდე, რომელიც შეესაბამება ბურთების ჩავარდნის რაოდენობას, არის ფუნქცია ელემენტარული ხდომილებების სივრცეზე: $X(A) = 0$, $X(B) = X(C) = 1$, $X(D) = 2$. ამ შედეგებზე დაყრდნობით მიიღება შემთხვევითი სიდიდის განაწილების შემდეგი კანონი:

კალათში ჩავარდნის რაოდენობა	0	1	2
ალბათობა	0.09	0.21+0.21=0.42	0.49

მათემატიკური ლოდინი

რა თქმა უნდა, შემთხვევითი სიდიდის განაწილების კანონი იძლევა სრულ ინფორმაციას ამ შემთხვევითი სიდიდის შესახებ, მაგრამ ხშირად შემთხვევითი სიდიდის თვისებების შესახებ წარმოდგენას გვაძლევს მისი ზოგიერთი ცალკეული მახასიათებელი. ასეთ მახასიათებლებს შორისაა შემთხვევითი სიდიდის *მათემატიკური ლოდინი*. განვიხილოთ ასეთი მაგალითი: ორ პარალელურ კლასში ჩატარდა საკონტროლო წერა, რომელთა შედეგების (ქულების) განაწილების კანონები მოცემულია ცხრილებით.

A კლასი

ქულა	0	1	2	3	4	5	6	7	8	9	10
სიხშირე	0	0.04	0.04	0.12	0.16	0.08	0.12	0.12	0.2	0.08	0.04

B კლასი

ქულა	0	1	2	3	4	5	6	7	8	9	10
სიხშირე	0	0.1	0.2	0.05	0.1	0.2	0	0.05	0.1	0.15	0.05

ისმის შეკითხვა: როგორ შევადაროთ ამ ორი კლასის შედეგები ერთმანეთს? რომელ მათგანშია მთლიანობაში უკეთესი შედეგი?

როგორც წარმოდგენილი მონაცემებიდან ჩანს, B კლასში 9 და 10-ქულიანი შედეგების სიხშირეები უფრო მაღალია, თუმცა A კლასში უფრო მაღალია 7 და 8-ქულიანი შედეგების

სიხშირები. ასეთი განსხვავებები სხვა შედეგებისთვისაც შეიძლება შევამჩნიოთ. გარდა ამისა, არ ვიცით კლასებში მოსწავლეთა რაოდენობები. შესაძლოა ეს რაოდენობები განსხვავებული იყოს. ასეთ შემთხვევებში მონაცემების შესადარებლად იყენებენ სიდიდეს, რომელსაც ეწოდება შემთხვევითი სიდიდის მათემატიკური ლოდინი. იგი ტოლია შემთხვევითი სიდიდის მნიშვნელობებისა და შესაბამისი სიხშირეების (ალბათობების) ნამრავლის ჯამის: $MX = p_1x_1 + p_2x_2 + \dots + p_nx_n$. ზემოაღნიშნულ მაგალითში მათემატიკური ლოდინი A კლასისათვის იქნება $MA = 5.8$ ხოლო B კლასისათვის იქნება $MB = 5.05$. თუ გავიხსენებთ (ფარდობითი) სიხშირის განსაზღვრებას, ადვილი მისახვედრია, რომ მათემატიკური ლოდინი, ფაქტობრივად, შემთხვევითი სიდიდის *საშუალო მნიშვნელობის* ტოლია. ამ მოსაზრებიდან გამომდინარე, შეიძლება ითქვას, რომ A კლასში საშუალო ქულა უფრო მაღალია, ვიდრე B კლასში. მათემატიკურ ლოდინს ხშირად **შემთხვევითი სიდიდის საშუალო მნიშვნელობას** ან **მოსალოდნელ მნიშვნელობას** უწოდებენ. ტერმინი *მათემატიკური ლოდინი* მომდინარეობს აზარტული თამაშების სფეროდან (რომელმაც ასევე დასაბამი დაუდო ალბათობის თეორიას) და იგი გამოხატავს იმას, რომ ამ დროს მნიშვნელობა აქვს არა მხოლოდ იმას, თუ როგორია მოგების ალბათობა, არამედ იმას, თუ როგორია მოსალოდნელი სიდიდე მოგების შემთხვევაში. განვიხილოთ ასეთი მაგალითი: დავუშვათ გამოშვებულია 1000000 ცალი ლატარიის ბილეთი. მათ შორის მომგებიანი ბილეთების განაწილება ასეთია: 200 ბილეთი იგებს 100 ლარს, 100 ბილეთი - 1000 ლარს, ხოლო 2 ბილეთი - 1000000 ლარს. ერთი ბილეთის ღირებულებაა 5 ლარი.

ამ ვითარების გასაანალიზებლად შეიძლება გამოვიყენოთ ასეთი მსჯელობა: თუ ჩვენ ვიყიდით 1 ბილეთს, 1000000 ლარის მოგების ალბათობა 0.000002-ის ტოლია; ზუსტად 1000 ლარის მოგების ალბათობა 0.0001-ის ტოლია და ა.შ. მაგრამ არსებობს მეორე მიდგომა, რომელიც იყენებს შემთხვევითი სიდიდის მოსალოდნელი მნიშვნელობის (მათემატიკური ლოდინის) ცნებას. ამ შემთხვევაში შეკითხვა შეიძლება დაისვას ასე: *როგორია მოგების მოსალოდნელი სიდიდე 1 ბილეთის შეძენისას?* ამ კითხვაზე პასუხის გასაცემად დავადგინოთ შემთხვევითი სიდიდის განაწილების კანონი. აქ შემთხვევითი სიდიდე ასე განისაზღვრება: X არის შესაძლო მოგებული თანხა 1 ბილეთის შეძენისას. მისი განაწილების კანონი ასეთია:

$$1000000 - 5 = 999995 \text{ ლარი} - 2 \text{ ბილეთი}$$

$$1000 - 5 = 995 \text{ ლარი} - 100 \text{ ბილეთი}$$

$$100 - 5 = 95 \text{ ლარი} - 200 \text{ ბილეთი}$$

$$-5 \text{ ლარი} - 999698 \text{ (დანარჩენი) ბილეთი}$$

აქედან გამომდინარე, განაწილების კანონი, რომელიც აღწერს სიდიდეებს და მათ შესაბამის ფარდობით სიხშირეებს, გამოისახება ასეთი ცხრილით:

შესაძლო მოგება	-5	95	995	999995
ალბათობა	0.999698	0.0002	0.0001	0.000002

უშუალო გამოთვლებით ადვილი შესამოწმებელია, რომ ამ შემთხვევითი სიდიდის მოსალოდნელი მნიშვნელობა არის

$$MX = -5 \cdot 0.999698 + 95 \cdot 0.0002 + 995 \cdot 0.0001 + 999995 \cdot 0.000002 = -2.88$$

ე.ი. ერთი ბილეთის შეძენის შემთხვევაში მოგების მოსალოდნელი მნიშვნელობა უარყოფითი რიცხვია (წაგება). თუ გავაგრძელებთ ლატარიის ბილეთებთან დაკავშირებული მაგალითების შესწავლას და განვიხილავთ ერთზე მეტი ბილეთის ყიდვის შემთხვევებს, ვნახავთ, რომ სხვადასხვა რაოდენობის ბილეთების შემთხვევაში მოგების მოსალოდნელი მნიშვნელობა სხვადასხვაა. თუ აღმოჩნდა, რომ შეძენილი ბილეთების რომელიმე რაოდენობისათვის მოსალოდნელი მნიშვნელობა დადებითი რიცხვია, მაშინ შეიძლება ვიფიქროთ იმაზე, რომ შევიძინოთ ბილეთების ეს რაოდენობა. მაგალითად, რა მოხდება, თუ ყველა ბილეთს შევიძენთ? შევნიშნავთ, რომ ლატარიის კომპანიები წინასწარ თვითონვე ატარებენ ამ სახის ანალიზს (სპეციალური პროგრამული უზრუნველყოფის გამოყენებით) და ისე არჩევენ მოგებების განაწილებას, რომ ბილეთების ნებისმიერი რაოდენობისათვის მოგების მათემატიკური ლოდინი უარყოფითი რიცხვია.

გადაწყვეტილების მიღება მონაცემების საფუძველზე

როგორც აღვნიშნეთ, სტატისტიკის მიზანი არის არა მხოლოდ მონაცემების მახასიათებლების დადგენა, არამედ ამ მონაცემების საფუძველზე სხვადასხვა სახის დასკვნების ჩამოყალიბება. შემთხვევითი სიდიდის მათემატიკური ლოდინი პრაქტიკაში გადაწყვეტილების მიღებისას ხშირად გამოიყენება. იმის სადემონსტრაციოდ, თუ როგორ შეიძლება გადაწყვეტილების მიღება შემთხვევითი სიდიდის მოსალოდნელი მნიშვნელობის გამოყენებით, განვიხილოთ მარტივი მაგალითი.

დავუშვათ მაღაზიაში იყიდება ორი სახეობის ნაყინი – შოკოლადიანი და ხილის. გამყიდველს სურს მიიღოს გადაწყვეტილება, ღირს თუ არა რომელიმე მათგანზე გამოაცხადოს ფასდაკლება იმისათვის, რომ გაზარდოს ნაყინის გაყიდვით მიღებული მოგება. ამისათვის მან, პირველ რიგში, შეადგინა ცხრილი, რომელიც გამოსახავს ნაყინის თითოეული სახეობის ერთი ერთეულის გაყიდვით მიღებულ შემოსავალს ფასდაკლების გარეშე და ფასდაკლებით.

I. ნაყინის ერთი ერთეულის გაყიდვით მიღებული შემოსავალი სხვადასხვა შემთხვევაში

ნაყინის სახეობები	ფასდაკლების გარეშე	ფასდაკლება შოკოლადისაზე	ფასდაკლება ხილისაზე
შოკოლადის	3 ლარი	2 ლარი	3 ლარი

ნაყინი			
ხილის ნაყინი	2 ლარი	2 ლარი	1 ლარი
არც ერთი	0 ლარი	0 ლარი	0 ლარი

ამასთან გამყიდველი რაღაც დროის განმავლობაში აგროვებდა მონაცემებს იმის შესახებ, თუ როგორი სიხშირით იყიდება თითოეული ნაყინი სხვადასხვა ფასდაკლების შემთხვევაში და შეადგინა ცხრილი, რომელიც გამოსახავს განაწილების შესაბამის კანონებს.

II. სიხშირეთა ცხრილი, რომელიც გამოსახავს მყიდველის არჩევანს სხვადასხვა შემთხვევაში

მყიდველის არჩევანი	ფასდაკლების გარეშე	ფასდაკლება შოკოლადისაზე	ფასდაკლება ხილისაზე
შოკოლადის	0.2	0.4	0.2
ხილის	0.3	0.3	0.5
არც ერთი	0.5	0.3	0.3

ამ მონაცემების საფუძველზე შესაძლებელია გამოვითვალოთ საშუალო მნიშვნელობები, რომლებიც არის ნაყინის ერთი ერთეულით მიღებული საშუალო შემოსავალი თითოეულ შემთხვევაში.

	ფასდაკლების გარეშე	ფასდაკლება შოკოლადისაზე	ფასდაკლება ხილისაზე
საშუალო მნიშვნელობა	1.2	1.4	1.1

რომ შევაჯამოთ, გვაქვს ასეთი სურათი: შემთხვევითია მაღაზიაში შემოსული მყიდველის არჩევანი; შესაბამის ხდომილებებს მიეწერება ალბათობები (სიხშირეები) სხვადასხვა შემთხვევაში (იხ. ცხრილი II); სხვადასხვა შემთხვევაში შემთხვევითი სიდიდეები რომელიმე სახის ნაყინის გაყიდვით მიღებული თანხაა (იხ. ცხრილი I); ამ შემთხვევითი სიდიდეების მოსალოდნელი მნიშვნელობებია შესაბამისი სიხშირეებისა და შემოსავლების ნამრავლების ჯამი. მათი შედარების საფუძველზე შესაძლებელია დავასკვნათ, რომ ყველაზე ხელსაყრელია ფასდაკლება შოკოლადის ნაყინზე, რადგან ამ დროს შემოსავლის საშუალო მნიშვნელობა ყველაზე მეტია.

როგორც ეს მაგალითი გვიჩვენებს, სხვადასხვა საქმიანობისას ძალზე მნიშვნელოვანია მონაცემების შეგროვება და მათი შენახვა სტრუქტურირებული სახით. მონაცემთა სტრუქტურირებული შენახვა გულისხმობს როგორც მათ შენახვას ისეთი სახით, რომ მაქსიმალურად გამარტივდეს დამუშავება და ანალიზი, ასევე იმას, რომ თითოეული ჩანაწერი ინახავდეს რაც შეიძლება მეტ ინფორმაციას, რომელიც დაკავშირებულია კონკრეტულ საქმიანობასთან.