# *<u>სალექციო მასალა საგანში:</u>*

## <u>სტატისტიკა ეკონომიკისა და ბიზნესისთვის-1</u>

## თავი IV. ვარიაციული ანალიზი ეკონომიკასა და ბიზნესში

### 1. ვარიაცია და მისი შესწავლის აუცილებლობა

ეკონომიკაში, ბიზნესსა და მენეჯმენტში მიმდინარე მოვლენები განიცდიან მუდმივ ცვალებადობას დროსა და სივრცეში. ეს ცვალებადობანი ხასიათდებიან გარკვეული წესებითა და კანონებით, რომელთა გათვალისწინება აუცილებელია ეკონომიკურ პოლიტიკაში, აგრეთვე ბისნესმენურ და მენეჯმენტურ გადაწყვეტილებებში. ამიტომ ასეთი საკითხების შესწავლისათვის არასაკმარისია ჩვენთვის აქამდე განხილული სტატისტიკური მაჩვენებლები. მაგალითად, საშუალო სიდიდე არის ვარიაციული ნიშნის განზოგადებული მახასიათებელი, მაჩვენებელი. მაგრამ ის ვერ ახასიათებს ვარიაციული მწკრივის ვარიანტების განლაგებას, ნიშნის ვარიაციის ხარისხს და მის რხევადობას. მაგალითად:

ფირმის ორი საწარმოო ბრიგადის მუშათა საშუალო დღიური ხელფასის მონაცემები

ცხრილი №9

| მაჩვენებლები / საწარმოო ბრიგადები | მუშათა რიცხვი | ცალკეული მუშების საშუალო დღიური ხელფასი | | | | | ერთი მუშის საშუალო დღიური გამომუშავება (ლარ) $(\overline{x})$ |
|---|---|---|---|---|---|---|---|
| ბრიგადა №1 | 5 | 2.5 | 3.0 | 2.6 | 2.8 | 3.1 | 2.8 |
| ბრიგადა №2 | 5 | 1.5 | 6.4 | 1.6 | 1.0 | 3.5 | 2.8 |

როგორც ცხრილიდან ჩანს, ორივე ბრიგადის მუშათა საშუალოდღიური ხელფასი ( $\overline{x}$ ) ერთი და იგივეა და შეადგენს

2.8 ლარს. მაგრამ თუ გადავხედავთ ცალკეული მუშების საშუალო დღიურ ხელფასს, ე.ი. ვარიაციული მწკრივის ვარიანტები როგორაა განლაგებული საშუალოს ირგვლივ, დავინახავთ, რომ მეორე ბრიგადის მუშათა ხელფასის განსხვავება უფრო მკვეთრია საშუალო ხელფასისაგან, ვიდრე I ბრიგადაში. I ბრიგადაში საშუალოსაგან მაქსიმალური გადახრა შეადგენს 3.1-2.8=0.3 ლარს, ხოლო მეორე ბრიგადაში 6.4-2.8=3.6 ლარს. ეკონომიკურად ეს იმას ნიშნავს, რომ I ბრიგადაში დაახლოებით თანაბარი კვალიფიკაციის მუშებია, ვიდრე მეორეში. მასასადამე ირკვევა, რომ განაწილების ცენტრის მაჩვენებელთან ერთად (საშუალო არითმეტიკული, მედიანა, მოდა) საჭიროა რაღაც სხვა, დამატებითი მაჩვენებლები, რომლებიც დაახასიათებენ ვარიაციის ხარისხს, მის რხევადობას და ა. შ. სწორედ ასეთია ვარიაციის მაჩვენებლები.

## 2. ვარიაციის გაქანება (დიაპაზონი)

ვარიაციის ყველაზე მარტივი მახასიათებელი ანუ მაჩვენებელია ვარიაციის დიაპაზონი (გაქანება), რომელიც წარმოადგენს ვარიაციული მწკრივის ვარიანტების მაქსიმალურ ($x_{\min}$) და მინიმალურ ($x_{\max}$) მნიშვნელობათა შორის სხვაობას და აღინიშნება R-ით.

$$R = X_{\max} - X_{\min} \qquad (7.1)$$

ზემოთ მოყვანილ მაგალითზე I ბრიგადის საშუალოდღიური ხელფასის ვარიაციის დიაპაზონი იქნება $R$=3.1-2.5=0.6, ხოლო მეორე ბრიგადაში $R$=6.4-1.0=5.4.

საზოგადოდ თუ გვაქვს $X$ შემთხვევით სიდიდეზე დაკვირვებით მიღებული მნიშვნელობანი $X_1$, $X_2$,... $X_n$ და ზოგადობის შეუზღუდავად ვიგულისხმებთ, რომ ისინი დალაგებულია ზრდადი მიმდევრობით, ე.ი.

$$X_1 \le X_2 \le, \ldots, \le X_n \qquad (7.2)$$

ისე რომ უმცირესია $X_1$, უდიდესი $X_n$, მიმდევრობის უდიდეს და უმცირეს წევრს შორის სხვაობას ამ მიმდევრობის დიაპაზონი ანუ ვარიაციის გაქანება ეწოდება:

$$R = X_n - X_1, \qquad (7.3)$$

ვარიაციის გაქანება (დიაპაზონი) იმ ერთეულებში გამოისახება, რა ერთეულებშიც გამოისახება ვარიანტები.

ვარიაციის გაქანების თავისებურება ისაა, რომ ის დამოკიდებულია მხოლოდ ორ წევრზე და ვერ ასახავს მინიმალურ და მაქსიმალურ ვარიანტებს შორის არსებულ წევრებს, რაც მისი ნაკლია. ამიტომ ეს მაჩვენებელი გამოიყენება მაშინ, როცა თითოეული ვარიანტი გვხვდება მხოლოდ ერთხელ, ან როცა მინიმალურ ან მაქსიმალურ მნიშვნელობას განსაკუთრებული როლი აქვს.

ვარიაციის გაქანების (დიაპაზონის) ნაკლოვანებას ავსებს საშუალო წრფივი და საშუალო კვადრატული გადახრა.

## 3. საშუალო წრფივი გადახრა

საშუალო წრფივი გადახრა ახასიათებს ვარიაციული მწკრივის საშუალოსაგან დანარჩენი ვარიანტების საშუალო გადახრას. ეს მაჩვენებელი გვიჩვენებს თუ რამდენად სწორად შეუძლია დაახასიათოს მწკრივის ცენტრის მახასიათებელმა ანუ მაჩვენებელმა შესასწავლი ერთობლიობა.

**საშუალო წრფივი გადახრა წარმოადგენს** ერთობლიობის საშუალო არითმეტიკულისაგან ცალკეული ვარიანტების გადახრის აბსოლუტური სიდიდეების საშუალო არითმეტიკულს და აღინიშნება $\bar{d}$ -ით.

ვინაიდან შესასწავლი ნიშნის ცალკეული ინდივიდუალური მნიშვნელობების (ვარიანტების) მათი საშუალო არითმეტიკულისაგან გადახრების ჯამი ნორმალური განაწილებისათვის ყოველთვის უდრის ნულს, საშუალო წრფივი გადახრის გაანგარიშებისათვის აიღება გადახრების აბსოლუტური მნიშვნელობების ჯამი. (გვაქვს თუ არა ამის

უფლება? თუ განვიხილავთ ამ გადახრებს კოორდინატთა სისტემაზე, დავინახავთ, რომ ისინი წარმოადგენს საშუალო არითმეტიკულისაგან ცალკეული ვარიანტების დაცილების მანძილებს, რაც შეიძლება მივიჩნიოთ დადებითად).

საშუალო წრფივი გადახრის
გასაანგარიშებელი ფორმულა:

$$\bar{d} = \frac{\sum|x-\bar{x}|}{n}. \qquad (7.4)$$

ამას ეწოდება მარტივი საშუალო წრფივი გადახრის გასაანგარიშებელი ფორმულა, ხოლო შეწონილს აქვს სახე:

$$\bar{d} = \frac{\sum|x-\bar{x}|f}{\sum f}. \qquad (7.5)$$

(გადახრების პირდაპირი ფრჩხილები უჩვენებს, რომ ეს გადახრები აიღება ნიშნის გარეშე),

სადაც $x$ – ვარიანტის ინდივიდუალური მნიშვნელობა,

$\bar{x}$ – საშუალო არითმეტიკული,

$f$ -თითოეული ვარიანტის სიხშირე, წონა,

$n$ –ვარიანტების რაოდენობა მწკრივში.

საშუალო წრფივი გადახრა ჩვენ მიერ ზემოთ მოყვანილი მაგალითისათვის იქნება:

I ბრიგადისათვის

$$\bar{d} = \sum_{i=1}^{5}\frac{|x-\bar{x}|}{5} = \frac{|2\times5-2\times|8 + 3\times0-|2}{5} + \frac{|2\times6-2\times|8 + 2\times8-|2| + 3\times1-2|\times8}{5} =$$

$$= \frac{0\times3+0\times2+0\times2+0\times0+0\times3}{5} = \frac{1+0}{5} = 0\times2,$$

II ბრიგადისათვის

$$\bar{d} = \frac{1\times 3+3\times 6+1\times 2+1\times 8+0\times 7}{5} = 1\times 72.$$

ე. ი. მეორე ბრიგადის მუშებში საშუალოდღიური ხელფასის საშუალო წრფივი გადახრა 8-ჯერ მეტია, ვიდრე I ბრიგადის მუშების საშუალოდღიურ ხელფასებში. ეს იმას ნიშნავს, რომ II ბრიგადაში უფრო განსხვავებული კვალიფიკაციის მუშებია, ვიდრე I-ში.

I ბრიგადაში საშუალო გადახრა საშუალო არითმეტიკულისაგან შეადგენს 20 თეთრს, ხოლო მეორეში 1 ლარსა და 72 თეთრს. ე. ი. საშუალო წრფივი გადახრა იმავე ერთეულებში გაიზომება, რა ერთეულებშიც გაზომილია თვით ვარიანტები.

## 4. დისპერსია და საშუალო-კვადრატული გადახრა.

საშუალო წრფივ გადახრასთან შედარებით მეტ გავრცელებას პოულობს დისპერსია და საშუალო კვადრატული გადახრა. მათი ფართო გამოყენება დაკავშირებულია მათ მათემატიკურ თვისებებთან. ისინი ფართოდ გამოიყენება კორელაციური კავშირის სიმჭიდროვის და აგრეთვე შერჩევითი დაკვირვების შეცდომის გასაზომად, ბიზნესში ხარისხის სტატიკური კონტროლის დროს და ა.შ.

აქვე უნდა შევნიშნოთ, რომ საშუალო სიდიდეს სხვაგვარად მათემატიკურ ლოდინს უწოდებენ. თუ მოცემულია X შემთხვევითი სიდიდის განაწილება.

$$X\begin{cases} x_{1,} x_{2,\dots,} x_n \\ p_1 \times p_1, \dots, p_n \end{cases} \qquad (7.6)$$

მაშინ შემთხვევითი სიდიდის მათემატიკური ლოდინი ვუწოდოთ ამ შემთხვევითი სიდიდის შესაძლო რიცხვით მნიშვნელობათა შესაბამის ალბათობაზე ნამრავლთა ჯამს.

თუ X შემთხვევითი სიდიდის მათემატიკურ ლოდინს $E(x)$ ით აღვნიშნავთ,

$$E(x) = x_1 p_1 + x_3 p_3 + ... + x_n p_n = \sum_{i-1}^{n} x_i p_i \ . \quad (7.7)$$

$p_i$ -არის შემთხვევითი სიდიდის ალბათობა, ე.ი.

$$\frac{f_i}{\sum_{i=1}^{n} f_t} \qquad f_i \text{- წონაა, სიხშირეა.} \qquad (7.8)$$

აქედან შეიძლება დავასკვნათ, რომ საშუალო წრფივი გადახრა წრფივი გადახრების მათემატიკური ლოდინია, ხოლო დისპერსია კვადრატული გადახრის მათემატიკური ლოდინია და აღინიშნება ბერძნული ასო მცირე სიგმა კვადრატით ($\sigma^2$)

**გადახრის კვადრატის საშუალო მნიშვნელობას შემთხვევითი სიდიდის ანუ ვარიაციული მწკრივის დისპერსია ეწოდება.**

$$\text{მარტივი} \quad \sigma^2 = \frac{\sum (x - \bar{x})^2}{n} . \qquad (7.9)$$

$$\text{შეწონილი} \qquad \sigma^2 = \frac{\sum (x - \bar{x})^2 f}{\sum f} . \qquad (7.10)$$

საშუალო კვადრატული გადახრა ეწოდება კვადრატულ ფესვს დისპერსიიდან

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} , \qquad \text{მარტივი} \quad (7.11)$$

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2 f}{\sum f}} , \qquad \text{შეწონილი} \quad (7.12)$$

## 5. ალტერნატიული ნიშნის დისპერსია

ხშირ შემთვევაში ჩვენ გვაინტერესებს არა ნიშნის საშუალო მნიშვნელობა, არამედ ამ ნიშნის მქონე ერთეულთა ხვედრითი წილი. მაგ., მამაკაცებისა და ქალების ხვედრითი წილი მოსახლეობაში, სტიპენდიანტებისა და არასტიპენდიანტების ხვედრითი წილი სტუდენტებში, წუნდებული და ვარგისი პროდუქციის ხვედრითი წილი და ა. შ. ასეთ შემთხვევას ეკონომიკაში ეწოდება ნიშნის ხარისხობრივი ვარიაცია. თუ არსებობს მხოლოდ ორი ურთიერთგამომრიცხავი ვარიანტი, მაშინ ნიშნის ვარიაციას **ეწოდება ალტერნატიული**.

ნიშნის არსებობა აღინიშნება 1-ით, არარსებობა 0-ით. ერთეულთა ხვედრითი წილი, რომლებიც ხასიათდება ამ ნიშნით, p-ით აღვნიშნოთ, ხოლო, რომლებიც არ ხასიათდებიან ამ ნიშნით, აღვნიშნოთ q-ით, მაშინ ალტერნატიული ნიშნის საშუალო არითმეტიკული და საშუალო კვადრატული გადახრა იქნება:

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{1\times p + 0\times q}{p+q} = p; \qquad (p+q=1) \qquad (7.13)$$

$$\sigma = \sqrt{\frac{\sum(x-\bar{x})^2 f}{\sum f}} = \sqrt{\sum(x-\bar{x})^2 f} = \sqrt{(1-p)^2 p + (0-p)^2 q} = \sqrt{(1-1+q)^2 p + p^2 q} = \sqrt{q^2 p + p^2 q} = \sqrt{pq(q+p)} = \sqrt{pq} = \sqrt{p(1-p)} \qquad (7.14)$$

რადგან q+p=1.

მაგ., 10000 მცხოვრებიდან 4000 მამაკაცია და 6000 ქალი

$$p = \frac{4000}{10000} = 0{,}4$$

$$q = \frac{6000}{10000} = 0{,}6$$

$$\sigma = \sqrt{0{,}4\times 0{,}6} = \sqrt{0{,}24} = 0{,}49$$ **ე.ი.** 49%

## 6. დისპერსიის თვისებები

ჩვენ საშუალო სიდიდეების თემაში ვაჩვენეთ საშუალო არითმეტიკულის ზოგიერთი თვისება, რომელიც შემდგომ გამოვიყენეთ გაანგარიშებათა გამარტივების მიზნებისათვის. აქაც შეიძლება მოვიტანოთ დისპერსიის ზოგიერთი თვისება, რაც შემდგომ გამოყენებულ იქნება დისპერსიების გაანგარიშების გამარტივებისათვის.

1. თუ ვარიანტების მნიშვნელობას ( $x$ ) გავადიდებთ რაიმე მუდმივი რიცხვით ( $a$ ), ამით დისპერსიის მნიშვნელობა არ შეიცვლება.

$$\sigma^2_{x+a} = \frac{\sum\left(x+a-\overline{x}\right)^2 f}{\sum f} = \sigma^2 . \qquad (7.15)$$

მართლაც თუ გავითვალისწინებთ, რომ ვარიანტების მნიშვნელობათა გადიდება ან შემცირება ერთიდაიგივე მუდმივი რიცხვით იწვევს საშუალო არითმეტიკულის გადიდებას ან შემცირებას იმავე მუდმივი რიცხვით, გვექნება:

$$\sigma^2_{x+a} = \frac{\sum\left[(x+a)-(\overline{x}+a)\right]^2 f}{\sum f} = \frac{\sum(x+a-\overline{x}-a)^2 f}{\sum f} = \sigma^2$$

ამ თვისებას მხოლოდ თეორიული მნიშვნელობა აქვს, ვინაიდან გაანგარიშებებს ვერ ამარტივებს და არც გამოიყენება პრაქტიკულად.

2. თუ ვარიანტების მნიშვნელობებს (x) შევამცირებთ რაიმე მუდმივი რიცხვით ( $a$ ), ამით დისპერსიის მნიშვნელობა არ შეიცვლება

$$\sigma^2_{x-a} = \frac{\sum\left(x-a-\overline{x}\right)}{\sum f} = \sigma^2 \qquad (7.16)$$

(დამტკიცება იქნება წინა თვისების მსგავსი).

ეს თვისება კი შეიძლება ფართოდ გამოვიყენოთ დისპერსიის გაანგარიშებათა გამარტივებისათვის, ვინაიდან როგორი

 147

სიდიდითაც არ უნდა შევამციროთ ვარიანტების მნიშვნელობანი, ამით დისპერსიის მნიშვნელობა არ შეიცვლება. ეს კი გაამარტივებს გაანგარიშებებს.

3. თუ ვარიანტების მნიშვნელობებს შევამცირებთ ან გავადიდებთ ერთი და იგივე მუდმივ რიცხვჯერ (h), მაშინ დისპერსიის მნიშვნელობა გადიდდება ან შემცირდება ამ მუდმივი რიცხვის კვადრატჯერ ($h^2$) და საშუალო კვადრატული გადახრა h-ჯერ. ეს თვისებაც ფართოდ გამოიყენება დისპერსიისა და საშუალო კვადრატული გადახრის გაანგარიშების გამარტივებისათვის. მუდმივი რიცხვები, რომლებითაც შეიძლება შევამციროთ ვარიანტების მნიშვნელობანი უმჯობესია ავიღოთ თანაბარინტერვალიანი ვარიაციული მწკრივისათვის ცენტრალური ვარიანტი და ინტერვალის მნიშვნელობა, რაც ყველაზე მეტად ამარტივებს სათანადო გაანგარიშებებს.

4. საშუალო კვადრატული გადახრა ყოველთვის მეტია საშუალო წრფივ გადახრაზე $\sigma > \bar{d}$, კერძოდ ნორმალური განაწილებისათვის $\sigma = 1{,}25\bar{d}$. ამავე დროს $\bar{x} \pm 1\sigma$ ფარგლებში თავსდება დაკვირვებათა 68,3%, $\bar{x} \pm 2\sigma$ ფარგლებში–დაკვირვებათა რიცხვის 95,4%, ხოლო $\bar{x} \pm 3\sigma$ ფარგლებში–დაკვირვებათა რიცხვის 99.7%. მაშასადამე პრაქტიკულად ნორმალური განაწილებისათვის საშუალოდ ყველა სახის გადახრა არ აღემატება 3σ–ს, რასაც **სტატისტიკაში „სამი სიგმას კანონს“** უწოდებენ.

5. მუდმივი რიცხვის (C) დისპერსია ნულის ტოლია. ადვილი მისახვედრია, რომ მუდმივი რიცხვის საშუალო მნიშვნელობა იგივე მუდმივი რიცხვია და ამიტომ ამ შემთხვევაში დისპერსია, რომელიც საშუალოდან ვარიანტების მნიშვნელობათა გადახრების კვადრატების საშუალოს წარმოადგენს, ნულის ტოლი იქნება. ეს თვისებაც ზოგჯერ შეიძლება გამოყენებულ იქნას დისპერსიისა და საშუალო კვადრატული გადახრის გაანგარიშების გამარტივების მიზნებისათვის.

 148

## 7. დისპერსიისა და საშუალო კვადრატული გადახრის გაანგარიშების მარტივი წესები

დისპერსიისა და საშუალო-კვადრატული გადახრის გაანგარშების ზემოთ მოყვანილი ფორმულები მეტად შრომატევადია. მათი გამარტივების ორი წესი არსებობს:

I წესი: გამოსახულება დისპერსიისა და საშუალო კვადრატული გადახრის გაანგარიშებისას გამოყენებული

$$\frac{\sum(x-\overline{x})^2 f}{\sum f} \qquad (7.17)$$

გარდავქმნათ: ავიყვანოთ კვადრატში, გადავამრავლოთ f-ზე და შემდეგ წევრ-წევრად გავყოთ $\sum f$ -ზე:

$$\frac{\sum(x-\overline{x})^2 f}{\sum f} = \frac{\sum\left(x^2 f - 2x\overline{x}f + \overline{x}^2 f\right)}{\sum f} = \frac{\sum x^2 f}{\sum f} - 2\overline{x}\frac{\sum xf}{\sum f} + \frac{\sum(\overline{x})^2 f}{\sum f} = \overline{x^2} - (\overline{x})^2$$

მაშასადამე $\sigma^2 = \overline{x^2} - (\overline{x})^2 \qquad (7.18)$

$$\sigma = \sqrt{\overline{x^2} - (\overline{x})^2} \qquad (7.19)$$

საშუალო კვადრატული გადახრა უდრის კვადრატული ფესვი, ვარიანტების კვადრატების საშუალოდან მინუს ვარიანტების საშუალოს კვადრატი.

II წესი – მომენტების წესი ანუ პირობითი ნულიდან ათვლის წესია.

სანამ ამ წესს განვიხილავდეთ, საჭიროა გავარჩიოთ განაწილების მომენტების ცნება[1].

მათემატიკურ სტატისტიკაში განაწილების მომენტს უწოდებენ ამა თუ იმ რიცხვიდან ვარიანტების ამა თუ იმ ხარისხის გადახრის საშუალო სიდიდეებს. თუ ეს რიცხვი

[1]განაწილების მომენტები შემოთავაზებულია რუსი მათემატიკოსის, პ.ლ. ჩებიშევის მიერ.

 149

საშუალო არითმეტიკულია, როგორც ჩვენ ამას აქამდე ვიხილავდით მოყვანილ მაჩვენებლებში, მაშინ განაწილების მომენტს უწოდებენ **ცენტრალურს**. თუ გადახრები იანგარიშება რომელიღაც ნებისმიერი რიცხვიდან, განაწილების მომენტს ეწოდება **პირობითი**, ხოლო თუ ნულიდან ხდება გადახრების ათვლა, მაშინ განაწილების მომენტს ეწოდება **საწყისი**.

სტატისტიკაში ჩვეულებრივად გვაქვს I, II, III, და IV რიგის მომენტები. ფორმულები ასეთია:

| განაწილების მომენტი | ცენტრალური | პირობითი | საწყისი |
|---|---|---|---|
| პირველი რიგის | $\frac{\sum(x-\overline{x})f}{\sum f}$ | $\frac{\sum(x-A)f}{\sum f}$ | $\frac{\sum(x-O)f}{\sum f}=\frac{\sum xf}{\sum f}$ |
| II რიგის | $\frac{\sum(x-\overline{x})^2 f}{\sum f}$ | $\frac{\sum(x-A)^2 f}{\sum f}$ | $\frac{\sum(x-O)^2 f}{\sum f}=\frac{\sum x^2 f}{\sum f}$ |
| III რიგის | $\frac{\sum(x-x)^3 f}{\sum f}$ | $\frac{\sum(x-A)^3 f}{\sum f}$ | $\frac{\sum(x-O)^3 f}{\sum f}=\frac{\sum x^3 f}{\sum f}$ |
| IV რიგის | $\frac{\sum(x-\overline{x})^4 f}{\sum f}$ | $\frac{\sum(x-A)^4 f}{\sum f}$ | $\frac{\sum(x-O)^4 f}{\sum f}=\frac{\sum x^4 f}{\sum f}$ |

გავიხსენოთ საშუალო არითმეტიკულის ზოგიერთი თვისება: თუ ნიშნის ყველა ვარაიანტს შევამცირებთ ან გავადიდებთ გარკვეული ერთი და იგივე რიცხვით ან რიცხვჯერ, მაშინ სათანადოდ შეიცვლება მათი საშუალოც. მაგ., ვარიანტები ჯერ შევამციროთ $x_0$-ით და შემდეგ $h$-ჯერ. ახალი ვარიანტებია:

$$x^1 = \frac{x - x_0}{h}. \qquad (7.20)$$

და მათი საშუალო არითმეტიკული იქნება:

$$\overline{x}' = \frac{\sum\left(\frac{x - x_0}{h}\right) f}{\sum f} \qquad (7.21)$$

ამასთან საშუალო არითმეტიკულის თვისების თანახმად საშუალო ახალი ვარიანტებიდან $\left(\frac{x-x_0}{h}\right)$ იქნება ანუ უდრის თავდაპირველი ვარიანტების საშუალო შემცირებული ჯერ $x_0$–ით და შემდეგ $h$–ჯერ:

$\overline{x}^1 = \frac{\overline{x}-x_0}{h}$, საიდანაც $\overline{x} = x_0 + \overline{x}^1 h$ (მივიღებთ საშუალოს გამარტივებული გაანგარიშების წესი).

$x_0$ და $h$ ნებისმიერი რიცხვებია, მაგრამ გაანგარიშების გამარტივების მიზნით უნდა ავიღოთ მწკრივის რომელიმე ცენტრალური ვარიანტი, ხოლო $h$ – $(x-x_0)$ სიდიდის საერთო უდიდესი გამყოფი, მაგ., თანაბარი ინტერვალების შემთხვევაში ინტერვალი უნდა მივიჩნიოთ $h$-ად.

მომენტების წესით დისპერსიის გასაანგარიშებელი ფორმულა ასეთი სახისაა:

$$\sigma^2 = h^2\left[\overline{x'^2} - \left(\overline{x'}\right)^2\right], \qquad (7.22)$$

სადაც $h$- ინტერვალის სიდიდეა;

$$\overline{x'^2} = \frac{\sum x'^2 f}{\sum f}, \qquad \left(\overline{x'}\right)^2 = \left(\frac{\sum x' f}{\sum f}\right)^2.$$

$f$ - თითოეული ვარიანტის წონა.

ამ ფორმულის (7.22) გამოსაყვანად გავიხსენოთ დისპერსიის თვისებები:

1. თუ ვარიანტების მნიშვნელობებს შევამცირებთ რაიმე მუდმივი რიცხვით, დისპერსია ანუ გადახრების კვადრატების საშუალო არ შეიცვლება;
2. თუ ვარიანტებს შევამცირებთ რაიმე მუდმივ რიცხვჯერ, მაშინ დისპერსია შემცირდება ამ რიცხვის კვადრატჯერ.

 151

ამ საფუძველზე შეიძლება გავამარტივოთ დისპერსიის გაანგარიშების პროცესი. $x$ ვარიატებს ჯერ გამოვაკლოთ რაღაც მუდმივი $x_0$ და შემდეგ გავყოთ $h$ ინტერვალის სიდიდეზე.

მაშასადამე ახალი ვარიანტები და ამ ვარაიანტების საშუალო არითმეტიკული, როგორც ზემოთ დავინახეთ

$$(7.20,\ 7.21)\ \text{იქნება:}\ x' = \frac{x - x_0}{h}, \overline{x}' = \frac{\sum \left( \frac{x - x_0}{h} \right) f}{\sum f}$$

დისპერსიების თვისებების თანახმად ახალი ვარიანტების დისპერსია $\frac{\sum (x - x_0)^2 f}{\sum f}$ არ შეიცვლება $x_0$-ით, არამედ შემცირდა $h^2$-ჯერ თავიდან მოცემული $x$ ვარიანტების დისპერსიასთან ($\sigma^2$) შედარებით. მაშასადამე სრული უფლება გვაქვს დავწეროთ $x'$ ვარიანტების დისპერსია (7.22) უდრის $x$ ვარიანტების დისპერსია $\sigma^2$ გაყოფილი $h^2$-ზე:

$$\frac{\sum (x' - \overline{x}')^2 f}{\sum f} = \frac{\sigma^2}{h^2} \qquad (7.23)$$

თუ 7.23 ტოლობის მარცხენა ნაწილის მრიცხველს ავიყვანთ კვადრატში, ფრჩხილების შიგნით შევიტანთ $\sum$ და $f$ სიმბოლოებს და შემდეგ მრიცხველს წევრწევრად გავყოფთ მნიშვნელზე $\left(\sum f\right)$, მივიღებთ:

$$\frac{\sum \left[ x'^2 - 2x'\overline{x}' + (\overline{x}')^2 \right] f}{\sum f} = \frac{\sigma^2}{h^2} \qquad (7.24)$$

(7.24) ფორმულიდან მივიღებთ:



152

153

$$\frac{\sum x'^2 f}{\sum f} - \frac{2\sum x'\overline{x'} f}{\sum f} + \frac{\sum \left(\overline{x'}\right)^2 f}{\sum f} = \frac{\sigma^2}{h^2} \qquad (7.25)$$

ამ გამოსახულებიდან:

$$\frac{\sum x'^2 f}{\sum f} = \overline{x'^2}, \qquad (7.26)$$

$$\frac{\sum \left(\overline{x'}\right)^2 f}{\sum f} = \left(\overline{x'}\right)^2, \qquad (7.27)$$

ხოლო $\frac{2\sum x'\overline{x'} f}{\sum f}$ გარდავქმნათ შემდეგნაირად:

$$2\overline{x}' \frac{\sum x' f}{\sum f} = 2\overline{x}' \times \overline{x}' = 2(\overline{x}')^2 \qquad (7.28)$$

აქედან მივიღებთ: $\overline{x'^2} - \left(\overline{x'}\right)^2 = \frac{\sigma^2}{h^2}$. (7.29)

საბოლოოდ გვექნება: $\sigma^2 = h^2\left[\overline{x'^2} - \left(\overline{x'}\right)^2\right]$ (7.30)

რისი დამტკიცებაც გვინდოდა.

(7.30) ფორმულიდან საშუალო კვადრატული გადახრა იქნება:

$$\sigma = h\sqrt{\overline{x'^2} - \left(\overline{x'}\right)^2}, \qquad (7.31)$$

მე-6 თავის მე-4 პარაგრაფში მოტანილი კონკრეტული მაგალითის მიხედვით გავიანგარიშოთ დისპერსიები როგორც ჩვეულებრივი, ისე სამომენტო წესებით (შედეგები ერთმანეთს შევადაროთ). აღნიშნულ პარაგრაფში მოტანილი მასალის შესაბამისი ცხრილი ასეთი სახისაა:

| $x$ | $\frac{x-a}{h}=\frac{x-140}{10}=x'$ | $f$ | $x'f$ |
|---|---|---|---|
| 120 | -2 | 5 | -10 |
| 130 | -1 | 6 | -6 |
| 140 | 0 | 7 | 0 |
| 150 | +1 | 8 | 8 |
| 160 | +2 | 9 | 10 |

ამ ცხრილის მიხედვით, $\overline{x}'=\frac{2}{31}$, ხოლო $x=140,6$, ჩვეულებრივი წესით დისპერსია ($\sigma^2$) გავიანგარიშოთ 7.17 ფორმულით: $\sigma^2=173,2$.

$$\sigma^2=\frac{\sum\left(x-\overline{x}\right)^2}{\sum f}=\frac{(140,6-120)^2 5+(140,6-130)^2 6}{5+6+7+8+5}$$
$$+\frac{(140,6-140)^2 7+(140,6-150)^2 8+(140,6-160)^2 5}{5+6+7+8+5}=173,2.$$

$\sigma^2=173,2.$

სამომენტო წესით:

$$\sigma^2=h^2\left[\overline{x'^2}-(\overline{x'})^2\right]=$$
$$=10^2\left[\frac{(-2)^2 5+(-1)^2 6+0+1^2+8+2^2+5}{31}-\left|\frac{2}{31}\right|^2\right]=173,2$$

როგორც ვხედავთ, ჩვენს მაგალითზე ჩვეულებრივი და სამომენტო წესებით გაანგარიშებული დისპერსიის მნიშვნელობანი ერთმანეთს დაემთხვა, რაც იმაზე მიანიშნებს, რომ ზოგჯერ უმჯობესია სამომენტო წესის გამოყენება, რომელიც ძალიან ამარტივებს გაანგა-რიშებებს.


155

156

## 8. ვარიაციის კოეფიციენტები

ვარიაციის გაქანება (დიაპაზონი), საშუალო წრფივი გადახრა და საშუალო კვადრატული გადახრები წარმოადგენს ვარიაციის აბსოლუტურ მაჩვენებლებს და აისახებიან იმ ერთეულებში, რა ერთეულებშიც ასახულია ვარიანტები. ეს არ იძლევა საშუალებას შევადაროთ სხვადასხვა მაჩვენებლების (ნიშნების) ვარიაცია. ამისათვის იყენებენ ვარიაციის შეფასების ხარისხობრივ მაჩვენებელს, ვარიაციის კოეფიციენტს.

ვარიაციის გაქანების (დიაპაზონი), საშუალო წრფივი და საშუალო კვადრატული გადახრის პროცენტული შეფარდება საშუალო არითმეტიკულთან გვაძლევს ვარიაციის კოეფიციენტებს. აქედან ვარიაციის გაქანების (დიაპაზონის) პროცენტული შეფარდებით საშუალო არითმეტიკულთან მიიღება **ოსცილაციის** კოეფიციენტი:

$$V = \frac{R}{\bar{x}}100. \qquad (7.32)$$

საშუალო წრფივი გადახრის ($\bar{d}$) საშუალო არითმეტიკულთან ($\bar{x}$) შეფარდება გვაძლევს ვარიაციის კოეფიციენტს, რომელიც ასე გამოისახება:

$$V = \frac{\bar{d}}{\bar{x}}100. \qquad (7.33)$$

საშუალო კვადრატული გადახრის ($\sigma$) საფუძველზე გაანგარიშებული ვარიაციის კოეფიციენტი გამოისახება ფორმულით:

$$V = \frac{\sigma}{\bar{x}}100. \qquad (7.34)$$

## 9. დისპერსიის შეკრების კანონი და მისი გამოყენება კორელაციურ ანალიზში

თუ რომელიმე ერთობლიობა დაყოფილია გარკვეულ ჯგუფებად, მაშინ საერთო დისპერსიასთან ($\sigma^2$) ერთად შეიძლება გავიანგარიშოთ ჯგუფური დისპერსიების საშუალო, საშუალო ჯგუფური დისპერსია და ჯგუფთაშორისი დისპერსია.

ჯგუფური დისპერსიების საშუალო ახასიათებს ნიშნის შიგაჯგუფურ ვარიაციას:

$$\sigma_i{}^2 = \frac{\sum_i (x_i - \bar{x}_i)^2 f_i}{\sum_i f_i}, \qquad (7.35)$$

სადაც $\bar{x}_i -$ $i$-ური ჯგუფის საშუალო არითმეტიკულია,

$x_i -$ჯგუფის -ური ვარიანტის მნიშვნელობა.

მათი საშუალო იქნება:

$$\overline{\sigma^2} = \frac{\sum_i \sigma_i{}^2 f_i}{\sum_i f_i}. \qquad (7.36)$$

ჯგუფური საშუალოების ვარიაციას ანუ რხევადობას საერთო საშუალოს გარშემო ახასიათებს ჯგუფთაშორისი დისპერსია, რომელიც აღინიშნება დელტა კვადრატით ($\delta^2$-ით)

$$\delta^2 = \frac{\sum_i \left(\overline{x_i} - \overline{x}\right)^2 f_i}{\sum f_i} \qquad (7.37)$$

საერთო დისპერსიას, კერძო, ჯგუფური დისპერსიების საშუალოსა და ჯგუფთაშორის დისპერსიას შორის არსებობს ასეთი კავშირი:

$$\sigma^2 = \overline{\sigma_i^{\,2}} + \delta^2 \qquad (7.38)$$

ე. ი. კერძო, ჯგუფური დისპერსიების საშუალოს მიმატებული ჯგუფთაშორისი დისპერსია უდრის საერთო დისპერსიას, რასაც ეწოდება დისპერსიების შეკრების წესი.

დამტკიცება: დისპერსიების შეკრების წესის დასამტკიცებლად შიგაჯგუფური დისპერსია და ჯგუფთაშორისი დისპერსია შეიძლება განვიხილოთ როგორც ორი ურთიერთდამოუკიდებელი შემთხვევითი სიდიდეების, $(x)$ და $(y)$, დისპერსიები. დავამტკიცოთ, რომ $x+y$ ჯამის დისპერსია უდრის $x$ და $y$ შემთხვევითი ურთიერთდამოუკიდებელი სიდიდეების დისპერდიების ჯამს ($x$ და $y$ ურთიერთდამოუკიდებელი სიდიდეებია იმიტომ, რომ ჯგუფის ანუ ინტერვალის ფარგლებში ვარიანტების ცვალებადობა არ იწვევს სხვა ჯგუფში ცვალებადობას).

დავამტკიცოთ, რომ

$$D(x+y) = D(x) + D(y) \qquad (7.39)$$

დისპერსიის განმარტებით

$$D(x) = E\left(x - \overline{x}\right)^2 \qquad (7.40)$$

$E$ მათემატიკური ლოდინია და უდრის შემთხვევითი სიდიდეების ყველა მნიშვნელობათა მათ ალბათობაზე ნამრავლთა ჯამს

ე. ი. $$D(x) = \Sigma\left(x - \overline{x}\right)^2 P_i \qquad (7.41)$$

მაშასადამე

$$D(x+y) = E\left[(x+y) - (\overline{x} + \overline{y})\right]^2 \qquad (7.42)$$

დავაჯგუფოთ

$$(x - \overline{x}) + (y - \overline{y})$$

ანუ რაც იგივეა

$$D(x+y) = E\left[(x+y) - E(x) + E(y)\right]^2 \qquad (7.43)$$

გვექნება

$$D(x+y)=E\left[(x-\bar{x})+(y-\bar{y})\right]^2 =$$
$$=E\left[(x-\bar{x})^2+2(x-\bar{x})(y-\bar{y})+(y-\bar{y})\right]^2 \quad (7.44)$$

ჯამის მათემატიკური ლოდინი უდრის შესაკრებთა მათემატიკური ლოდინის ჯამს. გვექნება

$$D(x+y)=E(x-\bar{x})^2+E\left[2(x-\bar{x})(y-\bar{y})\right]+E(y-\bar{y})^2 \quad (8.45)$$

$$E(x-\bar{x})^2=D(x) \quad (7.46)$$

$$E(y-\bar{y})^2=D(y) \quad (7.47)$$

ხოლო მეორე შესაკრები ნამრავლის ლოდინია და უდრის მათ ლოდინთა ნამრავლს და რადგან გადახრების ალგებრული ჯამი ნულს უდრის, მივიღებთ

$$D(x+y)=D(x)+D(y) \quad (7.48)$$

თუ საშუალო ჯგუფურ დისპერსიას გავყოფთ საერთო დისპერსიაზე, მივიღებთ დეტერმინაციის კოეფიციენტს ($\eta^2$ - ეტა კვადრატი), რომელიც გვიჩვენებს მთლიანი ვარიაციის რა ნაწილია განსაზღვრული დაჯგუფების ნიშნით. ფესვი დეტერმინაციის კოეფიციენტიდან არის კორელაციური დამოკიდებულება, რაც გვიჩვენებს დაჯგუფებისა და საშედეგო ნიშნებს შორის კავშირის სიმჭიდროვის ხარისხს.

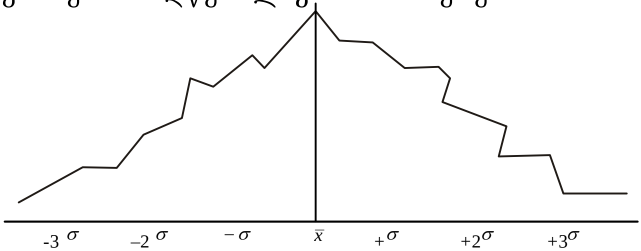
## 10. ნორმალური განაწილების კანონი და თანადობის კრიტერიუმები

ეკონომიკაში, ბიზნესსა და მენეჯმენტში ძალიან ხშირად გვხვდება განაწილების ვარიაციული მწკრივები (დისკრეტული ან ინტერვალური), რომლებშიც კავშირი ვარიანტების მნიშვნელობებსა ($x$) და წონებს (სიხშირეებს $f$) შორის ემორჩილება გარკვეულ კანონზომიერებებს. ასეთია, მაგალითად, სასოფლო-სამეურნეო მიწის ნაკვეთებზე მოსავლიანობისა და სასუქების, მოსახლეობის რიცხოვნობისა და ფეხსაცმელების ზომების, ხელფასისა


160

161

და მუშათა რიცხოვნობის და სხვა განაწილების ვარიაციული მწკრივები. ასეთი სახის მწკრივებში ვარიანტის მნიშვნელობათა ზრდის შესაბამისად სიხშირეები (წონები) ჯერ დიდდება, მწკრივის რომელიღაც შუა ნაწილში აღწევს მაქსიმალურ მნიშვნელობას და შემდეგ იწყებს კლებას. მაშასადამე, წონების მნიშვნელობანი სავარიაციო ნიშნის ცვალებადობის მიხედვით იცვლებიან კანონზომიერად. მაგალითად, თუ ავიღებთ ქალების პროცენტულ განაწილებას ფეხსაცმელების ზომების მიხედვით აღმოჩნდება, რომ დაბალი ზომის ფეხსაცმელს ქალების ნაკლები პროცენტული რაოდენობა ატარებს. შემდეგ ფეხსაცმელების ზომების ( $x$ ) გადიდებასთან ერთად ქალების პროცენტული რაოდენობა ( $f$ ) თანდათან იზრდება და შესაძლებელია 36 ან 37 ზომის ფეხსაცმელს ქალების ყველაზე მეტი რაოდენობა ატარებდეს. შემდეგ ისევ იკლებს ფეხსაცმლის ზომის გადიდებასთან ერთად ქალების პროცენტული შემადგენლობა. ისე, რომ თუ კოორდინატთა სისტემაზე გადავიტანთ ამ ურთიერთ-დამოკიდებულებას, მივიღებთ გარკვეულ მრუდს, რომელიც ასახავს ზემოთ აღწერილ კანონზომიერებას.




ნახ. 16. ფაქტობრივ მონაცემთა განაწილების მრუდი

ეს მრუდი ფაქტობრივი (ემპირიული) მონაცემების საფუძველზეა აგებული და ამიტომ ტეხილი ხაზით გამოისახება. კანონზომიერების გამოვლენის მიზნით საჭიროა მისი მოსწორება, რასაც ახდენენ ნორმალური განაწილების მრუდის დახმარებით.


162

ნორმალური განაწილება ემორჩილება ნორმალური განაწილების კანონს, რომელიც გრაფიკულად კარგად ჩანს მე-17 და მე-16 გრაფიკულ გამოსახულებაზე.

ალბათობის თეორიიდან ცნობილია, რომ თუ გვაქვს $x$ შემთხვევითი სიდიდის ისეთი განაწილება, რომელიც ემორჩილება სიდიდის ნორმალურ განაწილებას შეიძლება განისაზღვროს $x$-ის წონების (სიხშირეების $f$) მიხედვით $x_1$-დან $x_2$ ინტერვალში მოხვედრის ალბათობა. ამისათვის უნდა ვიპოვოთ $x$ ფუნქციის განსაზღვრული ინტეგრალი, ანუ

$$P[x_1 \leq x \leq x_2] = \frac{1}{\sqrt{2\Pi}} \int_{x_1}^{x_2} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} dx \quad (7.49)$$ [1]

ამ გამოსახულებას ამარტივებენ ნორმირებული გადახრის $t = \frac{x-\bar{x}}{\sigma}$ სიდიდის გამოყენებით. აქედან გამოსახულების მარჯვენა მხარე მიიღებს სახეს:

$$\frac{1}{\sqrt{2\Pi}} \int_{t_1}^{t_2} e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\Pi}} \int_{-t_2}^{t_2} e^{-\frac{t^2}{2}} dt - \frac{1}{\sqrt{2\Pi}} \int_{-t_1}^{t_1} e^{-\frac{t^2}{2}} dt = F(t_2) - F(t_1)$$

სადაც $t$, $\left(t_1 = \frac{x_1 - \bar{x}}{\sigma}, t_2 = \frac{x_2 - \bar{x}}{\sigma}\right)$ როგორც ვიცით, არის

---

[1](7.49) გამოსახულების ინტეგრალი ლაპლასის ინტეგრალის სახელწოდებათაა ცნობილი სტატისტიკაში. ის ნორმირებული ($t = \frac{x-\bar{x}}{\sigma}$) შემთხვევითი სიდიდისათვის ასახავს ნორმალური მრუდის ქვეშ მოთავსებულ ფართობს $t=1$-ისათვის ეს ფართობია 0.683, $t=2$-ისათვის 0.954, $t=3$-ისათვის 0.997)

[2]ამ ფუნქციის მნიშვნელობანი $t$-ს სიდიდის მიხედვით ტაბულირებულია და მოცემულია ცხრილებში.

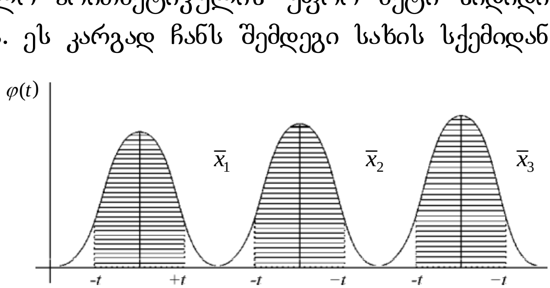
ნორმირებული, სტანდარტიზებული გადახრა, $F(t)$–განაწილების ინტეგრალური ნორმირებული სტანდარტიზებული ფუნქციაა[2]. ასეთი განაწილების ინტეგრალური ფუნქცია, ანუ ნორმალური განაწილების სიმჭიდროვის (ანუ მრუდზე სიხშირეების განაწილების ხარისხი, მრუდის თითოეულ წერტილში სიხშირეების სიდიდე) მაჩვენებელია ინტეგრალის ქვეშა გამოსახულება:

$$\varphi(t) = \frac{1}{\sqrt{2\Pi}} \times e^{-\frac{t^2}{2}} \qquad (7.50)$$

(ეს გამოსახულებაც ტაბულირებულია და მოიძებნება ცხრილებში).

სადაც $t = \dfrac{x - \bar{x}}{\sigma}$, $\Pi$–როგორც წრეხაზის სიგრძის შეფარდება დიამეტრთან არის მუდმივი სიდიდე და უდრის 3.1415-ს, $e$ ნატურალური ლოგარითმის ფუძეა და უდრის დაახლოებით 2.7182-ს, $\sigma$ - საშუალო კვადრატული გადახრაა.

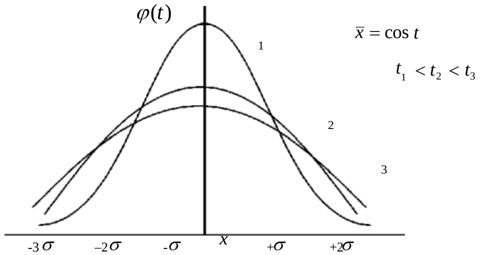
როგორც ამ გამოსახულებიდან ჩანს, ნორმალური განაწილების მრუდეები შეიძლება ავაგოთ ორი პარამეტრის ($\bar{x}$ და $\sigma$) მიხედვით. ამასთან თუ საშუალო კვადრატული გადახრა ($\sigma$) არ იცვლება და იცვლება მხოლოდ საშუალო არიტმეტიკული, ვთქვათ $\bar{x}_1 < \bar{x}_2 < \bar{x}_3$,მაშინ ნორმალური განაწილების მრუდული ფორმა არ იცვლება. მათ შორის მხოლოდ ის განსხვავებაა რომ მაქსიმალური მდგომარეობა შედარებით მაღალი აქვს საშუალო არითმეტიკულის უფრო მეტი სიდიდის შესაბამის მრუდს. ეს კარგად ჩანს შემდეგი სახის სქემიდან.




ნახ. 17. ნორმალური განაწილების მრუდეები $t$-ს უცვლელობისას

ამ სქემიდან ჩანს, რომ $t = \cos t, \bar{x}_1 < \bar{x}_2 < \bar{x}_3$.

იმ შემთხვევაში, თუ საშულო არითმეტიკული არ იცვლება ($\bar{x} = \cos t$) და იცვლება მხოლოდ ნორმირებული გადახრა ($t$), მაშინ ნორმალური განაწილების მრუდეები ასეთ სახეს მიიღებს:




ნახ. 18. ნორმალური განაწილების მრუდეები $\bar{x}$ -ის უცვლელობის პირობებში.

ამ სქემიდან ჩანს ნორმალური განაწილების შემდეგი ძირითადი თავისებურებანი:

ა. ნორმალური განაწილების მრუდი სიმეტრიულია ორდინატთა ღერძის მიმართ და მაქსიმალური გადაღუნვა გააჩნია ორ $\bar{x} = x_{მედ.} = x_{მოდ.}$ წერტილში;

ბ. მრუდე ასიპტონურად უახლოვდება აბცისთა ღერძს ორთავე, როგორც პლიუსი ისე მინუსი მიმართულებით უსასრულობამდე;

გ. ფართობი (დაშტრიხული მე-17 ნახაზზე), რომელიც მოთავსებულია $\bar{x} \pm \sigma$ მანძილზე გავლებულ ორდინატებს შორის შეადგენს 0.683-ს ანუ გამოსაკვლევი ერთეულების (ჩვენს შემთხვევაში $x$ -ის წონების, სიხშირეების, ხშირადობის) 68.3%-ს, იხრება საშუალო არითმეტიკულისაგან არაუმეტეს $1\sigma$ სიდიდით, ე. ი. იმყოფება $\bar{x} \pm \sigma$ საზღვრებში. $\bar{x} \pm 2\sigma$ საზღვრებში მოთავსებულია გამოსაკვლევი ერთეულების 95.4%, ხოლო $\bar{x} \pm 3\sigma$ საზღვრებში–99.7%.

თეორიული ნორმალური მრუდის (იხ.ნახ.17, ნახ.18), რომელიც განსხვავებულია ემპირიული მონაცემებით

ტეხილხაზოვანი მრუდისაგან (ნახ.16) ასაგებად საჭიროა ვარიაციული მწკრივის მოსწორება.

ვარიაციული მწკრივის მოსწორება ეწოდება წონების ემპირიული მახასიათებლების ( $f$ ) შეცვლას მათთან ახლომდგომი თეორიული მახასიათებლებით. ნორმალური განაწილების მრუდის დახმარებით ვარიაციული მწკრივის მოსწორების დროს თეორიული სიხშირეები ( $f'$ ) განისაზღრება ფორმულით:

$$f' = \frac{w \times i}{\sigma} \times \frac{1}{\sqrt{2\Pi}} \times e^{-\frac{t^2}{2}} = \frac{w \times i}{\sigma}\varphi(t) \qquad (7.51)$$

სადაც $W = \Sigma f$ – ვარიაციული მწკრივის სიხშირეთა ჯამია; i-ინტერვალის მნიშვნელობა ვარიაციულ მწკრივში.

$$\varphi(t) = \frac{1}{\sqrt{2\Pi}} \times e^{-\frac{t^2}{2}} \qquad (7.52)$$

$\varphi(t)$ მნიშვნელობანი $t$-ს შესაბამისად ტაბულირებულია და ადვილად შეიძლება ვიპოვოთ (იხ. დანართი №1).

ვარიაციული მწკრივის მოსწორების შემდეგ საჭიროა გაიზომოს ემპირიულ და თეორიულ სიხშირეებს შორის განსხვავების არსებითობა ან არარსებითობა. თუ განსხვავება არარსებითია, მაშინ მოსაწორებლად ჩვენს მიერ შერჩეული მრუდი სწორად ასახავს ემპირიულ განაწილებას და მაშასადამე ემორჩილება ნორმალური განაწილების კანონის მოთხოვნებს.

ემპირიულ ( $f$ ) და თეორიულ ( $f'$ ) სიხშირეებს შორის განსხვავების გასაზომად ყველაზე გავრცელებულია პირსონის ( $\chi^2$ ) – 'ხი კვადრატის'', რომანოვსკის და კოლმოგოროვის ( $\lambda$ ) - ,,ლამბდა'' კრიტერიუმების გამოყენება. **მათ უწოდებენ თანხმობის ანუ თანადობის კრიტერიუმებს.**

პირსონის (ინგლისელი მათემატიკოსი (1857-1936) თანხმობის კრიტერიუმი ( $\chi^2$ ) გაიზომება ფორმულით:

$$\chi^2 = \Sigma \frac{(f - f')^2}{f'} \qquad (7.53)$$

არსებობს ცხრილები (იხ. დანართი 11), რომლებშიც $\chi^2$–ისა და თავისუფლების ხარისხის ($K$)[1] მიხედვით მოიძებნება $\chi^2$–ის შემთხვევითი დადგომის ალბათობა $p(\chi^2)$. ამის გარდა $\chi^2$-ის ფაქტობრივი მაჩვენებელი უნდა შევუდაროთ ცხრილურ მონაცემს (იხ. დანართი №4). ემპირიულ (ფაქტობრივ) და თეორიულ სიხშირეთა სრული თანდამთხვევის პირობებში $\chi^2 = 0$-ს. წინააღმდეგ შემთხვევაში $\chi^2 > 0$. ამასთან თუ $\chi^2_{ფაქტ.} < \chi^2_{ცხრ.}$, მაშინ ჩვენს მიერ წინასწარ დაშვებული **ჰიპოთეზა** თეორიულ და ემპირიულ სიხშირეთა შორის განსხვავების არაარსებითობის (შემთხვევითობის) შესახებ მიღებულია და მოცემული ემპირიული განაწილება ემორჩილება ნორმალური განაწილების კანონს.

**რომანოვსკის** თანხმობის კრიტერიუმი გაიანგარიშება ცხრილების გამოყენების გარეშე შემდეგი ფორმულით:

$$\frac{\left|\chi^2 - K\right|}{\sqrt{2k}} \qquad (7.54).$$

თუ ეს მაჩვენებელი 3-ზე მეტია, მაშინ განსხვავება ითვლება არსებათად, და პირიქით, თუ 3-ზე ნაკლებია–არაარსებითად.

კიდევ უფრო მარტივია კოლმოგოროვის თანხმობის კრიტერიუმი

$$\lambda = \frac{D}{\sqrt{W}} \qquad (7.55)$$

სადაც $D$–ემპირიული და თეორიული სიხშირეების მზარდ

---

[1]თავისუფლების ხარისხი ($K$) ეწოდება ვარიაციულ მწკრივში ჯგუფების რაოდენობას გამოკლებული ემპირიული მწკრივის იმ მაჩვენებლების რიცხვი,

–



164

რომლებიც გამოიყენება თეორიულ სიხშირეთა გაანგარიშებისათვის (ასეთია
$X$ და $\sigma$ )


165

ჯამებს შორის განსხვავების აბსოლუტური მნიშვნელობით მაქსიმალური სიდიდეა: $W$ – სიხშირეთა ჯამია. არსებობს ცხრილები (იხ. დანართი №6), სადაც $\lambda$-ს შესაბამისად ვიპოვით ალბათობას, რომლის მიხედვით ვმსჯელობთ თეორიულ და ფაქტიურ სიხშირეთა შორის განსხვავების არსებითობაზე. მოვიტანოთ კონკრეტული მაგალითი (ციფრები პირობითია):

ფირმის მუშების მიერ გამომუშავების
ნორმების შესრულების მაჩვენებლები

| შესრულების % | 78-82 | 82-86 | 86-90 | 90-94 | 94-98 | 98-102 | 102-106 | 106-110 |
|---|---|---|---|---|---|---|---|---|
| მუშების რიცხვი | 4 | 8 | 21 | 27 | 36 | 28 | 19 | 5 |

მუშების რიცხვი არის ემპირიული სიხშირეები ( $f$ ), რომელთა მოსწორებისათვის ანუ თეორიული სიხშირეების ( $f'$ ) გაანგარიშებისათვის გამოვიყენოთ ნორმალური განაწილება და შემდგომ შევაფასოთ, რამდენად ემორჩილება მოცემული თანაბარინტერვალიანი ვარიაციული მწკრივი შერჩეული განაწილების კანონს. ჩვენთვის უკვე ცნობილი მეთოდების გამოყენებით გავიანგარიშოთ საშუალო არითმეტიკული ( $\bar{x}$ ) და საშუალოკვადრატული გადახრა ( $\sigma$ ). გაანგარიშების შედეგად საშუალო არითმეტიკული შეადგენს 95%–ს, ხოლო საშუალოკვადრატული გადახრა 6.5%-ს. როგორც ზამოთ იყო ნაჩვენები, ( $f'$ ) თეორიული სიხშირეების გაიანგარიშება განტოლებით:

$$f' = \frac{W \times i}{\sigma} \varphi(t), \qquad (7.56)$$

სადაც $W$ –მუშების რიცხვია,
ხოლო $i$ –ინტერვალის სიდიდე.
მაშასადამე, მუდმივი თანამამრავლი


166

$$\frac{W \times i}{\sigma} = \frac{148 \times 4}{6.5} = 91.$$

მოვიტანოთ ცხრილი, რომელშიც გვექნება თეორიული სიხშირეების, აგრეთვე თანადობის კრიტერიუმების გაანგარიშებისათვის საჭირო მონაცემები.

ცხრილი №10

| ინტერვალის შუალედი, $x$ | ემპირიული სიხშირეები $f$ | $t = \frac{x - \bar{x}}{\sigma}$ | $\varphi(t)$ [1] | თეორიული სიხშირეები $f' = 91\varphi(t)$ | $\frac{(f - f')^2}{f'}$ |
|---|---|---|---|---|---|
| 80 | 4 | -2.3 | 0.0283 | 2.6 | 0.754 |
| 84 | 8 | -1.6 | 0.1109 | 10.2 | 0.474 |
| 88 | 21 | -1.07 | 0.2251 | 20.6 | 0.007 |
| 92 | 27 | -0.46 | 0.3589 | 32.7 | 0.995 |
| 96 | 36 | 0.15 | 0.3945 | 36.0 | 0 |
| 100 | 28 | 0.78 | 0.2943 | 26.8 | 0.053 |
| 104 | 19 | 1.38 | 0.1539 | 14.1 | 1.702 |
| 108 | 5 | 2.0 | 0.0540 | 5.0 | 0 |
| | 148 | | | 148 | 3.985 |

თავისუფლების ხარისხის (8-2=6) და $\alpha$-ს 0.05 მნიშვნელობისათვის $\chi^2_{ცხრ.} = 12.59$-ს. რადგან $\chi^2_{ფაქტ.} < \chi^2_{ცხრ.}$ (3.985<12.59), ამიტომ 1-0.05=0.95 ალბათობით შეიძლება ვამტკიცოთ რომ მოცემული ემპირიული განაწილება ემორჩილება ნორმალური განაწილების კანონს.

იგივე შედეგს ვღებულობთ რომანოვსკის კრიტერიუმით,

$$\frac{|\chi^2 - k|}{\sqrt{2k}} = \frac{3.985 - 6}{\sqrt{2 \times}} = \frac{2.015}{3.46} = 0.582,$$

6

[1]t-ს – მნიშვნელობათა შესაბამისად სპეციალური ცხრილიდან (იხ.

დანართი 1) ნაპოვნია $\varphi(t)$ მნიშვნელობანი



168

რაც ნაკლებია 3-ზე, ე.ი. ემპირიულ და თეორიულ სიხშირეებს შორის განსხვავება არაარსებითია. მაშასადამე, მოცემული ვარიაციული მწკრივი ემორჩილება ნორმალური განაწილების კანონს.

კოლმოგოროვის $\lambda$ (“ლამბდა‘) თანადობის კრიტერიუმის გასაანგარიშებლად მოვიყვანოთ ცხრილი:

ცხრილი №11

| $f$ | $f'$ | ნაზარდი ჯამი | | $\|F-F'\|$ |
|---|---|---|---|---|
| | | ემპირიული $F$ | თეორიული $F'$ | |
| 4 | 2.6 | 4 | 2.6 | 1.4 |
| 8 | 10.2 | 12 | 12.8 | 0.8 |
| 21 | 20.6 | 33 | 33.4 | 0.4 |
| 27 | 32.7 | 60 | 66.1 | 6.1 |
| 36 | 36.0 | 96 | 102.1 | 6.0 |
| 28 | 26.8 | 124 | 128.9 | 4.9 |
| 19 | 14.1 | 143 | 143 | 0 |
| 5 | 5.0 | 148 | 148 | 0 |

$$\lambda = \frac{D}{\sqrt{W}} = \frac{6.1}{\sqrt{148}} = 0.5$$

ახლა $P(\lambda) = P(0.5)$ ცხრილში (იხ. დანართი 6) ვიპოვოთ ალბათობა, რაც შეადგენს: $P(0.5) = 0.9639$

მაშასადამე, თამამად შეიძლება ვამტკიცოთ, რომ განსხვავება თეორიულ და ემპირიულ სიხშირეებს შორის არაარსებითია ანუ შემთხვევითი.

## 11. ვარიაციული მწკრივის განაწილების ფორმის სტატისტიკური შესწავლა

ვარიაციული მწკრივის განაწილებას ორი ფორმა გააჩნია: სიმეტრიული და ასიმეტრიული. სიმეტრიული განაწილება ეწოდება X შემთხვევითი სიდიდის ისეთ განაწილებას,


169

რომელშიაც განაწილების ცენტრიდან თანაბრად დაშორებული ნებისმიერი ორი ვარიანტის წონები ერთმანეთის ტოლია.

ზემოთ ფეხსაცმელების ზომებისა და ამ ზომების მიხედვით ქალების პროცენტული ურთიერთდამოკიდებულება განაწილების მრუდის სახით გამოვსახეთ, რომელსაც სიმეტრიული ფორმა აქვს. **ზოგადად განაწილების მრუდი ეწოდება ვარიაციული მწკრივის ვარიანტებსა და მათ წონებს შორის ურთიერთდამოკიდებულების უწყვეტი ხაზის გრაფიკულ გამოსახულებას.** განაწილების მრუდს, რომელიც აგებულია ემპირიული (ფაქტობრივი) მასალების მიხედვით, შეიძლება ჰქონდეს პოლიგონის ან ჰისტოგრამის სახე, რაც ტეხილი ხაზითაა წარმოდგენილი. თუ ემპირიულ მონაცემებს გავანთავისუფლებთ შევთხვევითი ფექტორების გავლენისაგან და მათ საფუძველზე ავაგებთ განაწილების მრუდს, მაშინ მივიღებთ **განაწილების თეორიულ მრუდს,** რომელიც ასახავს ვარიანტების მნიშვნელობებსა და მათ წონებს შორის ურთიერთდამოკიდებულების კანონზომიერებას.

სიმეტრიული განაწილება თავისთავად შეიძლება იყოს ნორმალური და ექსცესური ფორმის. როგორც მე-17 და მე-18 ნახაზებიდან ჩანს, ნორმალური სიმეტრიული განაწილების მრუდს წვერი ნორმალური კუთხითაა წარმოდგენილი. თუ მახვილი კუთხითაა წარმოდგენილი, მაშინ მას ეწოდება **ექსცესი ზემო ნაწილში,** ხოლო თუ ბლაგვი კუთხითაა წარმოდგენილი – **ექსცესი ქვემო ნაწილში**. ექსცესის ხარისხს მათემატიკურ სტატისტიკაში ზომავენ მეოთხე რიგის ცენტრალური ნორმირებული მომენტით (სხვადასხვა რიგის ცენტრალური მომენტები იხ. ამ თავის მე-7 პარაგრაფში). ცენტრალური მომენტების საფუძველზე მათემატიკურ სტატისტიკაში გაიანგარიშება შესაბამისი რიგის ცენტრალური ნორმირებული მომენტები (საერთოდ განაწილების მომენტები-საწყისი, პირობითი, ცენტრალური


170

როგორც I, ისე II და III რიგის დამუშავებულია რუსი მათემატიკოსის პ. ლ. ჩებიშევის და შემდგომ ნორმალური

168



171

განაწილებისათვის გამოყენებულ იქნა ა. ა. მარკოვის მიერ). მათემატიკურ სტატისტიკაში კვადრატული ფესვი მეორე რიგის ცენტრალური მომენტიდან $\left(\sqrt{\dfrac{\sum(x-\bar{x})^2 f}{\sum f}}\right)$, რომელსაც სხვაგვარად საშუალო კვადრატული გადახრა ($\sigma$) ეწოდება მიღებულია სტანდარტად. მის მიმართ თითოეული რიგის ცენტრალური მომენტის შეფარდებით და შესაბამის ხარისხში აყვანით ვღებულობთ შესაბამისი რიგის ნორმირებულ ცენრტალურ მომენტებს. თუ მათ აღვნიშნავთ $z$ - სიმბოლოთი, მაშინ მივიღებთ:

ა) საწყისი ნორმირებული ცენტრალური მომენტი:

$$z_1 = \frac{\sum(x-\bar{x})f}{\sum f} : \sigma = 0 : \sigma = 0$$

ბ) მეორე რიგის ნორმირებული ცენტრალური მომენტი:

$$z_2 = \frac{\sum(x-\bar{x})^2 f}{\sum f} : \sigma^2 = 1$$

გ) მესამე რიგის ნორმირებული ცენრალური მომენტი:

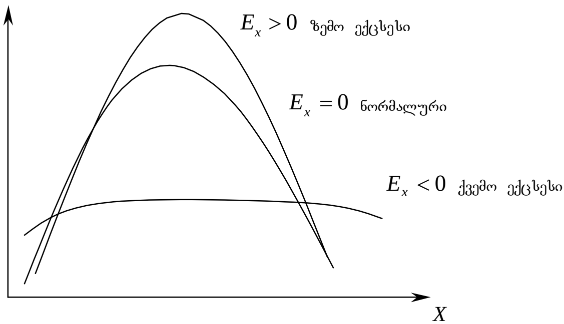
$$z_3 = \frac{\sum(x-\bar{x})^3 f}{\sum f} : \sigma^3 = \frac{M^3}{\sigma^3}$$

დ) მეოთხე რიგის ნორმირებული ცენტრალური მომენტი: $z_4 = \dfrac{\sum(x-\bar{x})^4 f}{\sum f} : \sigma^4 = \dfrac{M^4}{\sigma^4}$

ნორმალური განაწილებისათვის $z_4 = \dfrac{M^4}{\sigma^4} = 3$. ამიტომ ნორმალურ განაწილებასთან შედარებით სხვა სახის

განაწილების ექცესს ანგარიშობენ ფორმულით: $E_x = \frac{M^4}{\sigma^4} - 3$, რომელსაც ეწოდება ექცსესური განაწილების მაჩვენებელი. თუ $E_x > 0$, მაშინ საქმე გვაქვს ზემოდან ექცცსესთან, თუ $E_x = 0$ -ნორმალურ განაწილებასთან, ხოლო თუ $E_x < 0$, მაშინ ქვემო ექცსესთან გვაქვს საქმე.

გრაფიკულად ეს შემთხვევები ასე გამოისახება:



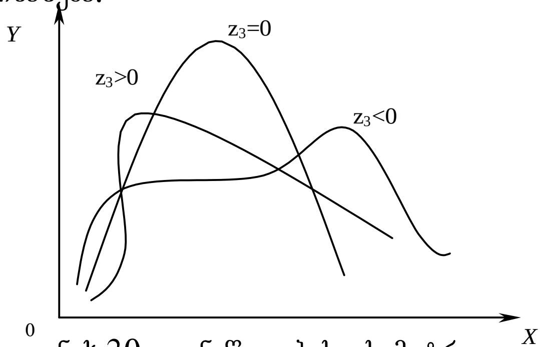

ნახ.19 განაწილების ექცსესი

**ასიმეტრიული** განაწილება ეწოდება ისეთ განაწილებას, რომლის დროსაც ვარიაციული მწკრივის ცენტრიდან მარჯვნივ ან მარცხნივ არსებული ვარიანტები ერთმანეთისაგან განსხვავებული წონებით გამოირჩევიან. ამიტომ ასიმეტრიული განაწილებაც ორი სახისაა: მარჯვნივ ასიმეტრიული და მარცხნივ ასიმეტრიული. ასიმეტრიის ყველაზე ზოგადი მაჩვენებელია $z_3 = \frac{M^3}{\sigma^3} = 3$, რომელიც ნორმალური განაწილებისათვის უდრის ნულს $z_3 = \frac{M^3}{\sigma^3} = \frac{0}{\sigma^3} = 0$. თუ



173

$z_3 > 0$ , მაშინ იმ ვარიანტებს აქვს მეტი წონა, რომლებიც საშუალო არიტმეტიკულის მარჯვნივ მდებარეობენ და გრაფიკულადაც მარჯვნივ აქვს გრაფიკს უფრო გრძელი ფრთა, თუ $z_3 < 0$, მაშინ ასეთივე მიზეზებით მარცხნივ ასიმეტრიასთან გვაქვს საქმე. გრაფიკულად ეს მაჩვენებლები ასე გამოისახება:




ნახ.20 განაწილების ასიმეტრია

ასიმეტრიის კოეფიციენტი $z_3 = \dfrac{M^3}{\sigma^3}$ არა მარტო ასიმეტრიის არსებობის, არამედ მისი არსებითობის ან არაარსებითობის კარგი მაჩვენებელია. საერთოდ მიღებულია სტატისტიკაში, რომ თუ $z_3$ აჭარბებს 0,05-ს, მაშინ ასიმეტრია მნიშვნელოვანია ანუ არსებითია და შესასწავლი ნიშნის განაწილება გენერალურ ერთობლიობაში არასიმეტრიულია.

სტატისტიკაში არსებობს, აგრეთვე, უფრო მარტივი ფორმულები, რომელთა გამოყენებით გაიზომება ასიმეტრია განაწილებაში. ასეთია, მაგალითად, **ასიმეტრიის პირსონის** კოეფიციენტი:

$$K_x = \frac{\bar{x} - \bar{x}_{\text{მოდ}}}{\sigma} \quad (7.61) \qquad \text{ან} \quad K_x = \frac{\bar{x} - \bar{x}_{\text{მედ}}}{\sigma} \quad (7.62)$$


175

სადაც -$K_x$ ასიმეტრიის კოეფიციენტია;

$\bar{x}, \bar{x}_{\text{მოდ}}$ და $\bar{x}_{\text{მედ}}$ -შესაბამისად საშუალო არითმეტიკული, მოდა და მედიანა,

$\sigma$ -საშუალოკვადრატული გადახრა.

თუ $K_x = 0$- ვარიაციული მწკრივი სიმეტრიულია;

$K_x < 0$ გვაქვს მარცხენამხრივი ასიმეტრია;

$K_x > 0$ გვაქვს მარჯვენამხრივი ასიმეტრია.

არსებობს აგრეთვე **ლინდბერგის** ასიმეტრიის კოეფიციენტი:

$$K_x = \lambda - 50, \qquad (7.63)$$

სადაც $\lambda$ -ვარიანტების იმ რაოდენობის ხვედრითი წილია პროცენტობით ვარიანტების საერთო რაოდენობაში, რომლებიც მეტია თავიანთი სიდიდით საშუალო არითმეტიკულზე. 50 იგივე მაჩვენებელია სიმეტრიული განაწილებისათვის. აქაც თუ $K_z = 0$, განაწილება სიმეტრიულია, თუ $K_x < 0$ ან $K_x > 0$, გვაქვს შესაბამისად მარჯვენამხრივი და მარცხენამხრივი ასიმეტრია.

ასიმეტრიისა და ექცესის მაჩვენებელთა გასაანგარიშებლად მოვიტანოთ მარტივი მაგალითი. ქ. თბილისის ერთერთმა სუპერმარკეტმა წინასაახალწლო დღეს გაჰყიდა 1 ლარის ფასის მქონე საქონელი 200 ერთეული, 2 ლარის –300 ერთეული, 3 ლარის – 220 ერთეული, 4 ლარის – 180 ერთეული, 5 ლარის 100 ერთეული.

ამ მონაცემებით შევადგინოთ ცხრილი:

ცხრილი №12

| $x$ | $f$ | $xf$ | $x-x$ | $(x-x)^2 f$ | $(x-x)^3 f$ | $(x-x)^4 f$ |
|---|---|---|---|---|---|---|
| 1 | 200 | 200 | -1.68 | 564.4 | -948.3 | 1593.1 |
| 2 | 300 | 600 | -0.68 | 138.7 | -94.3 | 64.1 |
| 3 | 220 | 660 | +0.32 | 22.5 | 7.20 | 2.30 |
| 4 | 180 | 720 | +1.32 | 313.6 | 413.9 | 546.3 |
| 5 | 100 | 500 | +2.32 | 538.2 | 1248.7 | 2896.9 |
| | $\sum f = 1000$ | $\sum xf = 2680$ | $\sum(x-x) = 160$ | $\sum(x-x)^2 f = 1577,4$ | $\sum(x-x)^3 f = 627$ | $\sum(x-x)^4 f = 5102,7$ |


176

 177

ამ ცხრილის საფუძველზე საშუალო არითმეტიკული $\bar{x} = \frac{\sum xf}{\sum f} = \frac{2680}{1000} = 2.68$ -ს, საშუალო კვადრატული გადახრა ანუ სტანდარტი $\sigma = \sqrt{\frac{\sum (x - x)^2 f}{\sum f}} = \sqrt{\frac{1577.4}{1000}} = 1.25$, მოდა,

$\bar{x}_{მოდ} = 2$ (რადგან მას ყველაზე მეტი წონა (300) აქვს ვარიაციულ მწკრივში. მესამე რიგის ცენტრალური მომენტი

$$M_3 = \frac{\sum (x - \bar{x})^3 f}{\sum f} = \frac{627}{1000} = 0.627$$

და $z_3 = \frac{M^3}{\sigma^3} = \frac{0.627}{1.25^3} = \frac{0.627}{1.95} = 0.32$

მესამე რიგის ნორმირებული მომენტია.

$$M_4 = \frac{\sum (x - \bar{x})^4 f}{\sum f} = \frac{5102.7}{1000} = 5.10$$

მეოთხე რიგის ცენტრალური მომენტი, ხოლო

$$z_4 = \frac{M^4}{\sigma^4} = \frac{5.10}{1.25^4} = \frac{5.10}{2.44} = 2.1$$

მეოთხე რიგის ნორმირებული მომენტია.

ამ მონაცემების მიხედვით ასიმეტრიის პირსონის კოეფიციენტი $K_x = \frac{\bar{x} - \bar{x}_{მოდ.}}{\sigma} = \frac{2.68 - 2.0}{1.25} = 0.54$. ეს ნიშნავს, რომ ასიმეტრია არის მარჯვენამხრივი (რადგან $K_x > 0$) და არა დიდი. ამავე შედეგს იძლევა ლინდბერგის ასიმეტრიის კოეფიციენტი: $K_x = \lambda - 50 = 60 - 50 = 10 > 0$

$(\lambda = \frac{3+4+5}{5} \times 100 = 60\%)$, რაც ნიშნავს, რომ ასიმეტრია დადებითია, ე. ი. მარჯვენამხრივი.

განაწილების ექსცესი $E_x = Z_4 = 2.1 - 3 = -0.90$ . რადგან



179

$E_x < 0$, ამიტომ ექცსესი დაბალსიმაღლიანია.

## 13. ვარიაციული მწკრივის კონცენტრაციის მაჩვენებლები

**ვარიაციული მწკრივის კონცენტრაცია ეწოდება მის მაჩვენებელთა, ვარიანტებისა ($x$) და მათი წონების ($f$) თავმოყრას,** განაწილებას, ამა თუ იმ წერტილში. საილუსტრაციოდ მოვიტანოთ კონკრეტული მაგალითი საბაზრო ეკონომიკაზე გარდამავალი პერიოდისათვის დამახასიათებელი სიტუაციიდან:

პრივატიზებული ქონება მესაკუთრეთა მიხედვით (ციფრები პირობითია)

ცხრილი №15

| პრივატიზებული ქონების ინტერვალები (ათას ლარებში) | მესაკუთრეთა რაოდენობა (%%-ობით ჯამის მიმართ) |
|---|---|
| 10.0 | 30.0 |
| 10-20 | 32.0 |
| 20-30 | 20.0 |
| 30-40 | ს0.0 |
| 40-50 | 5.0 |
| 50-ზე მეტი | 3.0 |
| სულ | 100.0 |

როგორც ცხრილიდნ ჩანს, ქონების დიდი ნაწილი (50.0 ათას მეტი ღირებულების) კონცენტრირებულია, თავმოყრილია მესაკუთრეთა მცირე ჯგუფის (3%) ხელში. მაშასადამე განაწილება მეტისმეტად უთანაბროა. მოცემული უთანაბრობა ეხება პრივატიზებული ქონების მეპატრონეებს. მით უმეტეს უთანაბრო იქნება ეს მაჩვენებელი მთელი მოსახლეობის მიმართ, რომელთა დიდი და მნიშვნელოვანი ნაწილი საქართველოში საერთოდ ქონების გარეშე, ღარიბ-ღატაკად დარჩა.

მესაკუთრეთა მიხედვით პრივატიზებული ქონების თანაბარი განაწილების შემთხვევაში მესაკუთრეთა 3%–ის ხელში ქონების 3% უნდა მოხვედრილიყო, 5%-ის ხელში ქონების 5%, 10%-ის ხელში – 10% და ა.შ.

სოციალურ-ეკონომიკური მოვლენებისა და პროცესების მაჩვენებელთა არათანაბარი განაწილების მრავალი მაგალითია მსოფლიოს როგორც განვითარებული ისე განვითარებადი საბაზრო ეკონომიკის ქვეყნებში. ასეთია, მაგალითად, მოსახლეობის შემოსავლები, დანახარჯები, პრივატიზებული ქონება და ა.შ.

მსოფლიო ეკონომიკურ-სტატისტიკურ თეორიასა და პრაქტიკაში ასეთ ეკონომიკურ მაჩვენებელთა გრაფიკული გამოსახვისათვის ცნობილია ლორენცის მრუდი. ამ მრუდის ასაგებად საჭიროა გვქონდეს ქონებისა და მესაკუთრეთა

ხშირადობანი[1] კოეფიციენტების სახით და მათ საფუძველზე გაანგარიშებული ხშირადობათა კუმულიატური ანუ ნაზრდი ჯამები. როგორც ცხრილიდან ჩანს, მესაკუთრეთა პროცენტებია 30%, 32%, 20%, 10%, 5%, 3%, მათი შესაბამისი ხშირადობანი: 0.30, 032, 0.20, 0.10, 0.05, 0.03. ამ ხშირადობათა კუმულიატური ნაზრდი ჯამებია 0.30, 0.62, 0.82, 0.92, 0.97, 1.0. პრივატიზებული ქონების ხშირადობათა გასაანგარიშებლად ჯერ მოცემული ინტერვალური ვარიაციული მწკრივი უნდა დავიყვანოთ დისკრეტულ ანუ წყვეტილ ვარიაციულ მწკრივზე, რისთვისაც ინტერვალის ქვედა და ზედა მნიშვნელობათა ჯამს ვყოფთ ორზე. (ამასთან რადგან მოცემული ინტერვალური ვარიაციული მწკრივი ღიაა როგორც ქვემოდან, ისე ზემოდან, საჭიროა ქვედა ინტერვალად მივიჩნიოთ 1-10 ათასი ლარი, ხოლო ზედა ინტერვალად 50-60 ათასი ლარი). გვექნება დისკრეტული ვარიაციული მწკრივი: 5,15,25,35,45,55. შესაბამისი ხშირადობანი იქნება:

$$\frac{5}{5+15+25+35+45+55}=\frac{5}{180}=0.03;$$

$$\frac{15}{180}=0.08,\frac{25}{180}=0.14,\frac{35}{180}=0.13,\frac{45}{180}=0.25,\frac{55}{180}=0.30.$$

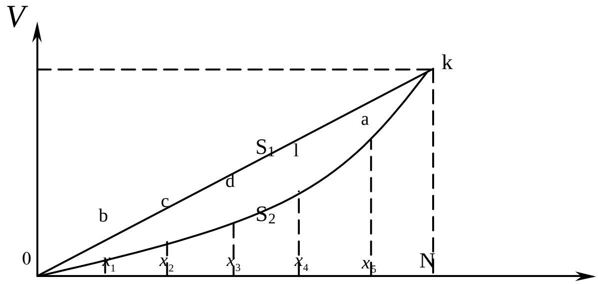
მათი შესაბამისი კუმულიატური ნაზრდი ჯამებია 0.03,0.11,0.25,0.44,0.70, 1.0. მაშასადამე მივიღეთ: პრივატიზებული ქონების კუმულიატური ხშირადობანი ($x$) 0.03, 0.11, 0.25,0.44, 0.70, 1.0. მესაკუთრეთა შესაბამისი ხშირადობანი ($f$) 0.30,0.62, 0.82,0.92,0.97,1.0. თუ მესაკუთრეთა კუმულიატურ ხსირადობებს გადავზომავთ დეკარტეს მართკუთხა კოორდინატთა სისტემის აბსცისთა

---

[1]სტატისტიკაში ხშირადობანი ეწოდება სტატისტიკური ერთობლიობის თითოეული ნაწილის უშუალო შეფარდებას მთლიან ერთობლიობასთან. ჩვენს მიერ ზემოთ მოტანილ მაგალითზე 3%-ის შესაბამისი ხშირადობა იქნება 0.03, 5%-ის 0.05, 8%-ის 0.08 და ა.შ.


180

ღერძზე, ხოლო პრივატიზებული ქონების შესაბამის მაჩვენებლებს ორდინატთა ღერძზე, მივიღებთ ლორენცის ცნობილ მრუდს:




ნახ. 22 ლორენცის მრუდი

თანაბარი განაწილების შემთხვევაში აბსცისთა და ორდინატთა ღერძებიდან აღმართული პერპენდიკულარების გადაკვეთის წერტილები განლაგდებოდა OK ღერძზე (დიაგონალზე). რაც უფრო სცილდება obcdea ლორენცის მრუდი ok დიაგონალს, მით უფრო ღრმავდება განაწილების უთანაბრობა. ამიტომ უთანაბრობის მაჩვენებლად მიჩნეულია მრუდის მიერ დიაგონალის ქვემოდ მოჭრილი $S_1$ ფართობის და მთლიანი OKN სამკუთხედის ფართობთა ურთიერთშეფარდება $\left(\frac{S_1}{S_1+S_2}\div\right)$. ამ შეფარდების საფუძველზეა გაანგარიშებული **ჯინის კოეფიციენტი,** რომელსაც ეკონომიკაში ფართოდ იყენებენ განაწილების უთანაბრობის ანუ კონცენტრაციის გასაზომავად. თუ შეფარდება ანუ კონცე-ნტრაციის (დიფერენციაციის) კოეფიციენტი $K_{\text{კონც}} = \frac{S_1}{S_1+S_2}$ ნულის ტოლია, ე.ი. OK დიაგონალიდან ქვემოთ მოჭრილი ფართობი არ არსებობს, მაშინ შესასწავლი ნიშნის განაწილება ერთობლიობაში თანაბარია. რაც უფრო იზრდება, OK დიაგონალიდან ქვემოთ მოჭრილი ფართობი, მით უფრო


181

კონცენტრაციის კოეფიციენტი ($K_{\text{კონც}} = \frac{S_1}{S_1 + S_2}$) უახლოვდება ერთს. ეს იმას ნიშნავს, რომ იზრდება ვარიაცლიული ნიშნის კონცენტარციის ხარისხი და ჩვენს მაგალითზე, ქონების სულ უფრო და უფრო მეტი ნაწილი კონცენტრირდება ერთ მესაკუთრეთა ხელში.

კონცენტრაციის ჯინის კოეფიციენტის გასაანგარიშებელი საბოლოო ფორმულის მისაღებად ვნახოთ რას წარმოადგენს ONK სამკუთხედის მთლიანი $S_1 + S_2$ ფართობი. ის OVKN ოთკუთხედის ფართობის ნახევარია. იმდენად, რამდენადაც აბსცისთა და ორდინატთა ღერძზე გადაზომილი მაჩვენებლების მაქსიმალური მნიშვნელობანი ერთის ტოლია, ე. ი. ამ ოთკუთხედის ფართობიც ერთის ტოლი იქნება და OKN სამკუთხედის ფართობი 1/2ის ტოლია. თუ ზემოთ მოტანილ ფორმულაში ჩავსვამთ ამ მნიშვნელობას, მივიღებთ:

$$K_{\text{კონც}} = \frac{S_1}{S_1 + S_2} = \frac{S_1}{\frac{1}{2}} = 2S_1. \qquad (7.66)$$

თუ $S_1$-ს გამოვსახავთ $S_2$ ფართობითა და OKN სამკუთხედის მთლიანი ფართობის (1/2) საფუძველზე გვექნება

$$S_1 = \frac{1}{2} - S_2 \qquad (7.67)$$

შევიტანოთ ეს მნიშვნელობა კონცენტრაციის კოეფიციენტის ($K_{\text{კონც}} = 2S_1$) მნიშვნელობაში, გვექნება:

$$K_{\text{კონც}} = 2S_1 = 2\left(\frac{1}{2} - S_2\right) = 1 - 2S_2$$

$$K_{\text{კონც}} = 1 - 2S_2$$


182

ახლა ვნახოთ რა არის $S_2$ ფართობი? თუ დავაკვირდებით ნახაზს, ადვილად შევამჩნევთ, რომ $S_2$ ფართობი დაფარულია გეომეტრიული სხეულებით. აქედან პირველი (კოორდინატთა სისტემის ცენტრთან) მართკუთხა სამკუთხედია, ხოლო დანარჩენი მართკუთხა ტრაპეციებია, რომელთა ფუძეებია $x_1b$ (პირველი ტრაპეცია), $x_2c$ (მეორე ტრაპეცია), $x_3d$ (მესამე ტრაპეცია), $x_4l$ (მეოთხე ტრაპეცია), $x_5a$ (მეხუთე ტრაპეცია), და $Nk$ (მეექვსე ტრაპეცია). ამ ტრაპეციების სიმაღლეებია შესაბამისად $x_1x_2, x_2x_3, x_3x_4, x_4x_5, x_5x_n$. მაშასადამე, თუ გამოვთვლით აღნიშნული გეომეტრიული სხეულების (გაანგარიშების გამარტივებისათვის სამკუთხედი $ox_1b$ ჩათვალოთ ტრაპეციად, რომლის ერთ-ერთი ფუძე ნულის ტოლია, რაც შეგვიძლია, რადგან გაანგარიშების ტექნოლოგიას და შედეგებს არ ცვლის) ფართობებს, მათი ჯამი $S_2$ ფართობი იქნება. ამის გაანგარიშება კი თავისუფლად შეგვიძლია, ვინაიდან ზემოთ ჩამოთვლილი ტრაპეციათა სიმაღლეები და ფუძეები ჩვენთვის უკვე ცნობილია. კერძოდ, აბსცისთა ღერძზე ჩვენ გადავზომავთ მესაკუთრეთა პროცენტებს ხშირადობათა სახით. ამიტომ $ox_1$, მონაკვეთი (სიმაღლე) უდრის 0.30-ს, $x_1x_2 = 0.32$, $x_2x_3 = 0.20$ , $x_3x_4 = 0.10$ , $x_4x_5 = 0.05$ , $x_5x_N = 0.03$ ორდინატთა ღერძზე გადაზომილია პრივატიზებებული ქონების კუმულიატიური ხშირადობანი. ამიტომ, $x_1b$ მონაკვეთი უდრის 0.03-ს, $x_2c = 0.11$, $x_3d = 0.25$ და ა.შ. ელემენტალური მათემატიკიდან ცნობილია, რომ ტრაპეციის ფართობი ფუძეების ნახევარჯამისა და სიმაღლის ნამრავლის ტოლია. აქედან გამომდინარე, ზოგადად $S_2$ ფართობი ასე შეიძლება გამოვსახოთ:


183

$$S_2 = \frac{1}{2}h_1(l_0 + l_1) + \frac{1}{2}h_2(l_1 + l_2) + ... + \frac{1}{2}h_n(l_{n-1} + l_n) = \frac{1}{2}\sum_{i=1}^{n} h_i(l_{i-1} + l_i)$$

სადაც $h_1, h_2$ და $h_n$ ტრაპეციების სიმაღლეებია, ხოლო $l_0 (l_0 = 0)$ და ა.შ. $J_i$– ტრაპეციათა ფუძეებია.

თუ მოცემულ გამოსახულებას შევიტანთ კონცენტრაციის კოეფიციენტის $(1 - 2S_2)$ ფორმულაში, მივიღებთ:

$$K_{კონც} = 1 - 2S_2 = 1 - 2\frac{1}{2}\sum_{i=1}^{n} h_i(l_{i-1} + l_i) = 1 - \sum_{i=1}^{n} h_i(l_{i-1} + l_i)$$

$$K_{კონც} = 1 - \sum_{i=1}^{n} h_i(l_{i-1} + l_i)$$

ეს ფორმულა ვარიაციული მწკრივის ათწილადებით გამოსახულ მაჩვენებლებში (ვარიანტების მნიშვნელობანი– $x$ და წონები– $f$ ) ასე გამოისახება:

$$K_{კონც} = 1 - \sum h_i\left(x_{i-1} + x_i\right) \quad (7.69)$$

ესაა სწორედ კონცენტრაციის ჯინის კოეფიციენტი.

თუ ამ ფორმულაში ჩვენს მონაცემებს ჩავსვამთ, მივიღებთ:

$$K_{კონც} = 1 - [0.30(0 + 0.03) + 0.32(0.03 + 0.11) +$$
$$+ 0.20(0.11 + 0.25) + 0.10(0.25 + 0.44) +$$
$$+ 0.05(0.44 + 0.70) + 0.03(0.70 + 1.0)] = 1 - 0.306 = 0.694$$

$K_{კონც.} = 0.694$ ანუ $69.4\%$.

ძალიან ხშირად ჯინის კოეფიციენტი გამოიყენება მოსახლეობის შემოსავლების, დანახარჯების კონცენტრაციის და დიფერნციაციის გასაზომავად. ამ მიზნებისათვის ჩვენი რეკომენდაციით[1] უმჯობესია გამოვიყენოთ შედარებით მარტივი

[1]ჩვენს მიერ (ავტორი) ჯინის კოეფიციენტის მოდიფიცირებული ფორმულები მიღებულია მოსახლეობის შემოსავლების მიხედვით კვარტილური, დეცილური და პერცენტილური განაწილებისათვის.

185

ფორმულები. კერძოდ მოსახლეობის შემოსავლების ან დანახარჯების კვარტილური ჯგუფების მიხედვით განაწილებისათვის $K_{კონც} = 1 - 0.25\sum_{i=1}^{4}(l_{i-1} + l_i)$, დეცილური ჯგუფების მიხედვით განაწილებისათვის $K_{კონც} = 1 - 0.10\sum_{i=1}^{10}(l_{i-1} + l_i)$, ხოლო პერცენტილური ჯგუფების მიხედვით განაწილებისათვის $K = 1 - 0.01\sum_{i=1}^{100}(l_{i-1} + l_i)$.

კონცენტრაციის ჯინის კოეფიციენტი შეიძლება გამოყენებულ იქნას მრავალი ვარიაციული მწკრივის მაჩვენებელთა განაწილების უთანაბრობის გასაზომავად. მაგრამ უმთავრესად ის საბაზრო ეკონომიკის პირობებში მოსახლეობის შემოსავლებისა და დანახარჯების დიფერენციაციის (კონცენტრაციის) გასაზომავად გამოიყენება პრაქტიკულად. რაც შეეხება პროდუქციის წარმოებისა და რეალიზაციის, აგრეთვე საბანკო ბიზნესს, აქ მსგავსი ამოცანების გადაწყვეტისას სტატისტიკოსები უპირატესობას შედარებით მარტივ მეთოდებს ანიჭებენ. მათ შორის გამოიყოფა **ჰერფინდალისა და როზენბლიუტის** კოეფიციენტები. გერფინდალის კონცენტრაციის გაანგარიშების ფორმულა შემდეგი სახისაა:

$$K_{კონც} = \sum\left(\frac{x}{\sum x}\right)^2 \quad (7.70)$$

სადაც $x$ –ვარიაციული მწკრივის ვარიანტის მნიშვნელობაა.

როზენ ბლიუტის მიერ კი შემოთავაზებულია შემდეგი ფორმულა: $K_{კონც.} = \dfrac{1}{2\sum id - 1}$ (7.71)

სადაც $i$ – ობიექტების ნომერი (ფირმის, საწარმოს, ქარხნის,



186

ფაბრიკის, მეწარმის და სხვ.)

$d$ –ობიექტის ხვედრითი წილი ერთობლიობაში. გავიანგარიშოთ ეს კოეფიციენტები კომერციული ბანკების აქტივების საფუძველზე.

**კომერციული ბანკების აქტივები (მლნ აშშ დოლარი)**

| კომერციული ბანკების რიცხვი | აქტივები (მლნ. დოლარი) |
|---|---|
| 20 | 20.5 |
| 30 | 100.5 |
| 40 | 1200.0 |
| 60 | 2000.0 |
| სულ 150 | 3321.0 |

ჰერფინდალის კონცენტრაციის კოეფიციენტი:

$$K_{კონც} = \left(\frac{20.5}{3321}\right)^2 + \left(\frac{100.5}{3321.0}\right)^2 + \left(\frac{1200.0}{3321.0}\right)^2 + \left(\frac{2000.0}{3321.0}\right)^2 = 0.401.$$

როზენბლიუტის კოეფიციენტი:

$$K_{კონც} = \frac{1}{2\left(1\cdot\frac{2000}{3321} + 2\frac{1200}{3321} + 3\frac{100.5}{3321} + 4\frac{20.5}{3321}\right) - 1} =$$

$$= \frac{1}{2\left(\frac{2000}{3321} + \frac{2400}{3321} + \frac{301.5}{3321} + \frac{82.0}{3321}\right) - 1} =$$

$$= \frac{1}{2(0.602 + 0.722 + 0.09 + 0.02) - 1} = \frac{1}{3488} = 0.286$$

როგორც ჩანს, ამ ორ კოეფიციენტს შორის მნიშვნელოვანი განსხვავებაა, რაც მათ ერთერთ ნაკლზე მიანიშნებს. ამიტომ თავისი პრაქტიკული მნიშვნელობით კონცენტრაციისა და დიფერენციაციის მაჩვენებლებს შორის მაინც ჯინის კოეფიციენტი სარგებლობს შედარებითი უპირატესობით.



187

# თემა V. კორელაციურ-რეგრესული[1] ანალიზი ეკონომიკასა და ბიზნესში

## 1. მოვლენათა შორის ურთიერთკავშირის ფორმები და სახეები

ნებისმიერი მოვლენის ადა პროცესის ღრმა ეკონომიკური ანალიზი, აგრეთვე ოპტიმალურ ბიზნესმენურ და მენეჯმენტურ გადაწყვეტილებათა მიღება მოითხოვს ისეთი მძლავრი მეთოდების გამოყენებას, როგორიცაა კორელაციური, რეგრესიული და მათთან დაკავშირებული მეცნიერული აპარატი. ამიტომ სტატისტიკა საზოგადოებრივ-ეკონომიკურ მოვლენებსა და პროცესებს განიხილავს არა იზოლირებულად, არამედ მათ მჭიდრო ურთიერთკავშირში. თითოეული მოვლენის განვითარება გავლენას ახდენს სხვა რომელიმე, მასთან დაკავშირებული მოვლენებისა და პროცესების განვითარებაზე, ან კიდევ პირიქით. ამიტომ მოვლენათა შორის ურთიერთკავშირის განხილვისას გამოყოფენ ფაქტორულ (მოვლენები, რომლებიც მოქმედებენ სხვა მოვლენების განვითარებაზე) და საშედეგო (მოვლენები, რომელთა განვითარება მოცემულ შემთხვევაში განპირობებულია სხვა, მათზედ მოქმედი მოვლენებისა და პროცესების განვითარებით) ნიშნებს.

სტატისტიკა ერთიმეორისაგან განასხვავებს მოვლენებს შორის ურთიერთკავშირის ორ ფორმას: **ფუნქციურსა და კორელაციურს, ანუ სტატისტიკურს.**

ფუნქციურ კავშირს სხვაგვარად სრულ კავშირს უწოდებენ. ამ ფორმის შემთხვევაში ფაქტორული ნიშნის თითოეულ მნიშვნელობას შეესაბამება საშედეგო ნიშნის მხოლოდ ერთი მნიშვნელობა. ის ერთნაირი ძალით გამოვლინდება ყველგან

---

[1]სიტყვა „კორელაცია“ ლათინური სიტყვა Correlatio –ისგანაა წარმოშობილი და ნიშნავს ურთიერთდამოკიდებულებას, შეფარდებას, ხოლო სიტყვა – რეგრესი, ლათინური სიტყვა – Regressus-ისგანაა წარმოშობილი და ნიშნავს უკუსვლას, დაქვეითებას, პროგრესის საპირისპიროს.

და ყველა შემთხვევაში. ასეთია, მაგალითად, კავშირი წრის ფართობსა და რადიუსს შორის, ან წრეხაზის სიგრძესა და რადიუს შორის, ბიზნესში სანარდო ანაზღაურებაზე მყოფი მუშის ხელფასსა და გამოშვებულ პროდუქციას შორის კავშირი და სხვ. რადიუსის ცვალებადობა (მოზეზობრივი მოვლენა ანუ ფაქორული ნიშანი), ყოველთვის, ყველგან და ერთი და იმავე ძალით იწვევს წრეხაზის სიგრძის ან წრეხაზის ფართობის (საშედეგო მოვლენები) ცვალებადობას. ფუნქციურ კავშირს უმთავრესად ადგილი აქვს საბუნებისმეტყველო მოვლენებსა და პროცესებში. საზოგადოებრივ-ეკონომიკურ მოვლენებსა და პროცესებში კი სხვა ფორმის კავშირი შეიმჩნევა. აქ ერთი მოვლენის ცვალებადობა მეორეზე მოქმედებს არა ერთობლიობის ყველა კონკრეტული ერთეულისათვის, შემთხვევისათვის, არამედ საერთოდ, საშუალოდ, დაკვირვების საკმარისი რიცხვის პირობებისათვის. ასეთია, მაგალითისათვის, კავშირი შრომის ნაყოფიერებასა და მუშების კვალიფიკაციას შორის, წარმოების მოცულობასა და პროდუქციის ერთეულის თვითღირებულებას შორის და ა.შ. მუშების კვალიფიკაციის ამაღლება იწვევს შრომის ნაყოფიერების გადიდებას არა ყველა კონკრეტული შემთხვევისათვის, არამედ საერთოდ, საშუალოდ მთელს ქარხანაში ან დარგში. რატომ ხდება ასე? საშედეგო მოვლენის ცვლილება დამოკიდებულია მრავალ, მასზედ (ზოგჯერ ურთიერთსაწინააღმდეგო) მოქმედი ფაქტორების განვითარებაზე. ხდება ხოლმე ისე, რომ ზოგჯერ ერთი ფაქტორის გავლენას გადაფარავს სხვა ფაქორების ზემოქმედება და შედეგიც ვერ გამოავლენს მიზეზ-შედეგობრივ კავშირ-ურთიერთობას. მაგრამ თუ განვიხილავთ რამდენიმე შემთხვევას, მაშინ საშუალოდ, საერთოდ ყველა შემთხვევისათვის მივიღებთ, რომ მიზეზობრივი მოვლენის განვითარებამ გავლენა მოახდინა საშედეგო მოვლენის ცვალებადობაზე. მაგალითად, სასუქების შეტანამ მიწის ნაკვეთზე ყოველთვის შეიძლება არ გაადიდოს მოსავლიანობა

(ვინაიდან შესაძლებელია მოსავლიანობა შეამციროს ცუდმა კლიმატურმა პირობებმა). მაგრამ საერთოდ, თუ ავიღებთ დაკვირვების დიდ (საკმარის) რიცხვს, მაშინ გამოვლინდება კანონზომიერება: სასუქების რაოდენობის გადიდება (გარკვეულ საზღვრამდე) იწვევს მოსავლიანობის გადიდებას.

არსებობს კავშირების **სხვადასხვა სახეობანი.** მათ შორის აღსანიშნავია **პირდაპირი და უკუ, წრფივი და არაწრფივი, ერთფაქტორიანი და მრავალფაქტორიანი კავშირები.** პირდაპირი კავშირის დროს მიზეზობრივი ფაქტორის ცვლილების მიმართულება (შემცირება ან გადიდება) ემთხვევა საშედეგო მოვლენის ცვლილების მიმართულებას. მაგალითად, მუშების კვალიფიკაციის ამაღლება ან შემცირება გამოიწვევს შრომის ნაყოფიერების ამაღლებას ან შემცირებას ან პირიქით, შემცირება-გადიდებას. უკუკავშირების შემთხვევაში კი პირიქითაა. მიზეზობრივი მოვლენის გადიდება იწვევს საშედეგო მოვლენის შემცირებას და შემცირება-გადიდებას. მაგალითად, წარმოების მოცულობის გადიდება იწვევს პროდუქციის ერთეულის თვითღირებულების შემცირებას და პირიქით.

მიზეზობრივი ფაქტორის გადიდებასთან ერთად ზოჯერ მიმდინარეობს საშედეგო მოვლენის უწყვეტი გადიდება ან შემცირება, რასაც ეწოდება წრფივი კავშირის სახეობა და მათემატიკურად გამოისახება წრფივი განტოლებით $y = a_0 + a_1x,$ სადაც $y$ –საშედეგო მოვლენის რაოდენობრივი გამოსახულება, $a_0$ და $a_1$ კავშირის გამომსახველი პარამეტრებია და $x$ – მიზეზობრივი მოვლენის რაოდენობრივი გამოსახულებაა.

არაწრფივი კავშირის შემთხვევაში მიზეზობრივი ფაქტორის გადიდებისას საშედეგო ფაქტორი დიდდება ან მცირდება უთანაბროდ, ზოგჯერ ცვლილების მიმართულება საწინააღმდეგოა. ასეთი სახის კავშირი მათემატიკურად გამოისახება არაწრფივი (პარაბოლა, ჰიპერბოლა, მაჩვენებლიანი ფუნქცია და ა.შ.) განტოლებით.


190

ზოგჯერ საჭიროა მოვახდინოთ საშედეგო მოვლენის განვითარებაზე მოქმედი ფაქტორების აბსტრაჰირება (გარდა ერთისა) და განვიხილოთ მხოლოდ ჩვენთვის საინტერესო ფაქტორის გავლენა. ამ შემთხვევაში გვაქვს ერთფაქტორიანი სახეობის კავშირი. სხვაგვარად ასეთ კავშირს უწოდებენ **წყვილად კავშირს, წყვილად კორელაციას.** იმ შემთხვევაში, თუ მრავალი ფაქტორის გავლენას ვიხილავთ საშედეგო მოვლენის განვითარებაზე, მაშინ საქმე გვაქვს **მრავალფაქტორულ კავშირთან, მრავლობით კორელაციასთან.**

სტატისტიკის ერთერთი მთავარი და მნიშვნელოვანი ამოცანაა სოციალურ-ეკონომიკურ მოვლენებსა და პროცესებს შორის კავშირის კორელაციურ-რეგრესიული ანალიზი. თვით კორელაციურ-რეგრესიული ანალიზის შემქნელებად ითვლებიან (ფ. გალტონი (1822-1911) და ლ. პირსონი (1857-1936წ). ფ. გალტონი დაინტერესებული იყო მამებისა და შვილების სიმაღლეთა შორის კავშირით. მან გამოიკვლია 200-ზე მეტი ოჯახი. კვლევის შედეგებმა აჩვენა, რომ მაღალი მამების შვილები მამებთან შედარებით ნაკლები სიმაღლის ვითარდებოდნენ, ანუ ხასიათდებოდნენ სიმაღლის რეგრესით, ხოლო დაბალი მამების შვილები, პირიქით უფრო მაღალი სიმაღლით ვითარდებოდნენ მამებთან შედარებით (აქედან წარმოდგა რეგრესიული ანალიზი). საშუალოდ კი საზოგადოებაში შვილების სიმაღლე უფრო მეტია, ვიდრე მამებისა (ეს კი გამოწვეულია ცოლების საპირისპირო (მაღალი ან დაბალი) სიმაღლითა და სხვა ფაქტორებით).

## 2. კავშირის შესწავლის სტატისტიკური მეთოდები

მოვლენათა შორის კავშირის ფორმისა და სახეობის, რაოდენობრივი თანაფარდობის დადგენისათვის გამოიყენება სხვადასხვა ხერხები და მეთოდები. მათ შორის აღსანიშნავია

**პარალელურ მწკრივთა, საბალანსო, ანალიზური დაჯგუფების, კორელაციურ-რეგრესული** და სხვ. მეთოდები. **პარალელურ მწკრივთა** შედარების მეთოდი გულისხმობს, რომ მიზეზობრივი და საშედეგო მოვლენების რაოდენობრივი გამოსახულებანი ჩაიწერება ერთმანეთის პარალელურად. ასე შეიძლება ჩაიწეროს, მაგალითად, ფირმაში მუშების საშუალო თანრიგი და შრომის ნაყოფიერების მაჩვენებლები, მოსავლიანობა და პროდუქციის ერთეულის თვითღირებულება და ა.შ. ჩაწერილი პარალელური მწკრივების ურთიერთშედარებით შეიძლება დავადგინოთ კავშირის ფორმა, სახეობა, აგრეთვე კავშირის სიმჭიდროვის ხარისხის მაჩვენებლები სხვადასხვა კოეფიციენტებით და ა.შ.

**საბალანსო მეთოდი** გულისხმობს რესურსების მოძრაობის აბსოლუტურ მაჩვენებლთა ურთიერთდაკავშირებული სისტემის შედგენას. ასეთია, მაგალითად, მატერიალური ბალანსები, რომლებიც სქემატურად შეიძლება წარმოვადგინოთ რესუსებისა და მათი გამოყენების წყაროების ტოლობით:

**ნაშთი პერიოდის დასაწყისში+შემოსულობანი= =დანახარჯები+ნაშთი პერიოდის ბოლოსათვის.** ეს ბალანსი გვიჩვენებს მატერიალური რესურსების მოძრაობის ერთიან პროცესს და ახასიათებს ამ პროცესის ცალკეულ ელემენტებს შორის კავშირსა და პროპორციებს.

კავშირების შესწავლის სტატისტიკურ მეთოდებს შორის ერთ-ერთი მნიშვნელოვანია **ანალიზური დაჯგუფების მეთოდი.** ეს მეთოდი გულისხმობს არა მარტო დაჯგუფების გამოყენებას, არამედ განზოგადებული მაჩვენებლების გაანგარიშებას ანალიზისათვის. ამ მეთოდის გამოყენებით მოვლენათა შორის ურთიერთკავშირის გამოვლენისათვის საჭიროა ჯერ მიზეზობრივი მოვლენის მიხედვით დავაჯგუფოთ საშედეგო მოვლენის მაჩვენებლები. შემდეგ კი თითოეული ჯგუფისათვის გავიანგარიშოთ საშედეგო მოვლენის საშუალო ან შეაფარდებითი მაჩენებლები. ამის შემდეგ შეგვიძლია საშედეგო მოვლენისა და მიზეზობრივი ფაქტორის


192

ცვლილებანი ერთმანეთს შევადაროთ და გამოვალინოთ კავშირის ხასიათი, სახეობა, ფორმა და რაოდენობრივი თანაფარდობანი. მაგალითად, 24 ფირმის ძირითადი ფონდებისა და გამოშვებული პროდუქციის შესახებ საანგარიშო პერიოდისათვის გვაქვს შემდეგი მონაცემები:

ცხრილი №16

| ფირმის ნომერი რიგზე | ფირმის ძირიტადი კაპიტალის საშუალო წლიური ღირებულება (მლნ. ლარობით) | პროდუქცია საანგარიშო პერიოდში, შესადარის ფასებში (მლნ. ლარობით) |
|---|---|---|
| 1 | 2 | 3 |
| 1 | 2.0 | 1.5 |
| 2 | 3.9 | 4.2 |
| 3 | 3.3 | 6.4 |
| 4 | 3.3 | 4.3 |
| 5 | 3.0 | 1.4 |
| 6 | 3.1 | 3.0 |
| 7 | 3.1 | 2.5 |
| 8 | 4.4 | 7.9 |
| 9 | 3.1 | 3.6 |
| 10 | 5.6 | 8.9 |
| 11 | 3.5 | 2.5 |
| 12 | 4.0 | 2.8 |
| 13 | 1.0 | 1.6 |
| 14 | 7.0 | 2.9 |
| 15 | 3.5 | 5.6 |
| 16 | 4.9 | 4.4 |
| 17 | 2.8 | 2.8 |
| 18 | 5.5 | 9.4 |
| 19 | 6.6 | 1.9 |
| 20 | 2.0 | 2.5 |
| 21 | 4.7 | 3.5 |
| 22 | 2.7 | 2.3 |
| 23 | 3.0 | 3.2 |
| 24 | 6.1 | 9.3 |

ძირითადი კაპიტალისა და გამოშვებული პროდუქციის ღირებულებათა შორის არსებული კავშირის გამოვლენისათვის საჭიროა ფირმები დავაჯგუფოთ ძირითადი კაპიტალის ღირებულების მიხედვით (ვთქვათ, 4 ჯგუფად თანაბარი ინტერვალებით). თითოეული ჯგუფისათვის გავიანგარიშოთ ფირმების რიცხვი, ძირითადი კაპიტალის ღირებულება სულ და საშუალოდ ერთ ფირმაზე, პროდუქციის ღირებულება სულ და საშუალოდ ერთ ფირმაზე, ცალკეული ჯგუფების

ხვედრითი წილი როგორც ძირითადი კაპიტალის, ასევე პროდუქციის ღირებულების მიხედვით.

წინა მასალიდან გავიხსენოთ თანაბა-რინტერვალიანი დაჯგუფების ინტერვალის სიდიდის (h) განმსაზღვრელი ფორმულა:

$$h = \frac{x_{\max} - x_{\min}}{n}$$

ჩვენს მაგალითზე:

$$x_{\max} = 7.0\,,\; x_{\min} = 1.0\,,\; n = 4\,,\; h = \frac{7.0 - 1.0}{4} = 1.5$$

I ჯგუფში მოხვდება ფირმები, რომელთა ძირითადი კაპიტალის ღირებულება არის 1.0 მლნ ლარიდან 2.5 მლნ ლარამდე, II ჯგუფში 2.5 მლნ ლარიდან 2.5+1.5=4 მლნ ლარამდე, III ჯგუფში 4.0 მლნ ლარიდან 5.5 მლნ ლარამდე, IV ჯგუფში 5.5 მლნ ლარიდან 7.0 მლნ ლარამდე. შევადგინოთ ცხრილი, სადაც მოთავსდება ყველა საძიებელი სიდიდე.

ცხრილი №17

| ჯგუფები | ფირმების რიცხვი | ძირითადი კაპიტალის ღირ. (მლნ. ლარი) | | საერთო პროდუქცია (მლნ. ლარი) | | ხვედრითი წილი (%) | |
|---|---|---|---|---|---|---|---|
| | | სულ | ერთ ფირმაზე | სულ | საშუალოდ ერთ ფირმაზე | ფირმების მიხედვით | საერთო პროდუქციის მიხედვით |
| I 1.0-2.5 | 3 | 5.0 | 1.6 | 5.4 | 1.8 | 5.3 | 4.5 |
| II 2.5-4.0 | 12 | 38.8 | 3.2 | 39.0 | 3.2 | 41.6 | 32.8 |
| III 4.0-5.5 | 5 | 24.0 | 4.8 | 31.0 | 6.2 | 25.8 | 26.2 |
| IV 5.7-7.0 | 4 | 25.3 | 6.3 | 43.3 | 10.8 | 27.3 | 36.5 |
| სულ | 24 | 93.1 | 3.8 | 118.7 | 4.9 | 100.0 | 100.0 |

როგორც ჩანს, ძირითადი კაპიტალის ღირებულების ზრდასთან ერთად იზრდება გამოშვებული პროდუქციის რაოდენობა საშუალოდ ერთ ფირმაზე. ეს იმას ნიშნავს, რომ ძირითადი კაპიტალისა და გამოშვებული პროდუქციის

ღირებულებათა შორის არსებობს პირდაპირი კორელაციური კავშირი. ამ მაჩვენებლების საფუძველზე შეგვიძლია გავიანგარიშოთ ძირითადი კაპიტალის გამოყენების დონეები ფირმების ჯგუფების მიხედვით. ეს დონეები გამოისახება კაპიტალუკუგებით და გაიანგარიშება პროდუქციის გაყოფით ძირითადი კაპიტალის ღირებულებაზე. ეს მაჩვენებლები გვიჩვენებს ძირითადი კაპიტალის ერთ ლარზე შექმნილი პროდუქციის რაოდენობას. ფირმების I ჯგუფში კაპიტალუკუგება შეადგენს 1.12 ლარს, II ჯგუფში – 1.0 ლარს, III-ში – 1.55 ლარს, IV-ში 1.71 ლარს, მთელი დარგის მიხედვით – 1.27 ლარს. მაშასადამე, წარმოების მოცულობის გადიდებასთან დაკავშირებით იზრდება ძირითადი კაპიტალის გამოყენების დონეც, რაც მეტყველებს მსხვილი წარმოების ეკონომიკურ უპირატესობაზე.

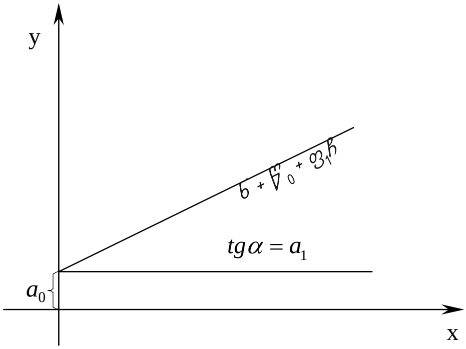
## 3. კორელაციურ-რეგრესული ანალიზის მეთოდები

**კორელაციურ-რეგრესული ანალიზის** მეთოდი გულისხმობს კავშირის ადეკვატური ამსახველი მოდელის აგებას და მისი მეშვეობით მიზეზობრივ-შედეგობრივი კავშირის რაოდენობრივი თანაფარდობის გაანგარიშებას. თუ მოვლენებს შორის კავშირი წრფივი ფორმისაა, მაშინ ის გამოისახება წრფივი განტოლებით $y = a_0 + a_1 x$

$a_0$ და $a_1$ პარამეტრების გასაანგარიშებლად ვიყენებთ ნორმალურ განტოლებათა სისტემას:

$$\begin{cases} na_0 + a_1\Sigma x = \Sigma y \\ a_0\Sigma x + a_1\Sigma x^2 = \Sigma xy \end{cases} \qquad (8.1)$$

განტოლებათა სისტემის ამოხსნით მივიღებთ $a_0$ და $a_1$ პარამეტრების მნიშვნელობებს, რომელთა დახმარებით ვადგენთ ემპირიულ განტოლებას. როგორია პარამეტრების გეომეტრიული და ეკონომიკური შინაარსი? თუ ავაგებთ დეკარტეს მართკუთხა კოორდინატთა სისტემაზე x-ისა და y-ის მნიშვნელობებს, მივიღებთ:




ნახ.23. წრფივი განტოლების გრაფიკი

$a_0$ გეომეტრიულად არის მანძილი კოორდინატთა ცენტრიდან გრაფიკული გამოსახულების ორდინატთა ღერძის გადაკვეთამდე. ხოლო $a_1$ იმ კუთხის ტანგენსია, რომელსაც ქმნის გრაფიკი აბსცისთა ღერძთან.

ეკონომიკურად $a_0$ არის საშედეგო მოვლენის რაღაც საწყისი მნიშვნელობა, ხოლო $a_1$ გვიჩვენებს მიზეზობრივი მოვლენის ერთი ერთეულით ცვლილება რამდენი ერთეულით შეცვლის საშედეგო მოვლენას.

ზემოთმოყვანილი განტოლებით გამოისახება, მაგალითად, კავშირი შრომის ნაყოფიერებასა და მუშების კვალიფიკაციას, მოსავლიანობასა და სასუქების რაოდენობას შორის და ა. შ.

ზოგჯერ მოვლენებს შორის კავშირი არაწრფივია, მაშინ მას ასახავს ჰიპერბოლა

$$y = a_0 + a_1 x + a_2 \frac{1}{x} \qquad (8.2),$$

პარაბოლა

$$y = a_0 + a_1 x + a_2 x^2 \qquad (8.3),$$

ან მაჩვენებლიანი ფუნქცია

$$y = a_0 a_1^x \qquad (8.4).$$

ჰიპერბოლარული კავშირის პარამეტრების განსაზღვრისათვის ამოიხსნება შემდეგი ნორმალურ განტოლებათა სისტემა:

$$\begin{cases} na_0 + a_1 \Sigma \frac{1}{x} = \Sigma y \\ a_0 \Sigma \frac{1}{x} + a_1 \Sigma \frac{1}{x^2} = \Sigma \frac{y}{x} \end{cases} \qquad (8.5),$$

პარაბოლის შემთხვევაში:

$$\begin{cases} na_0 + a_1 \Sigma x + a_2 \Sigma x^2 = \Sigma y \\ a_0 \Sigma x + a_1 \Sigma x^2 + a_2 \Sigma x^3 = \Sigma xy \\ a_0 \Sigma x^2 + a_1 \Sigma x^3 + a_2 \Sigma x^4 = \Sigma x^2 y \end{cases} \qquad (8.6).$$

მაჩვენებლიანი ფუნქციის შემთხვევაში ჯერ საჭიროა გავაწრფივოთ გალოგარითმების წესით:

$$\log a_0 + \log a_1 x = \log y \qquad (8.7).$$

ახლა შეგვიძლია გამოვიყენოთ წრფივ განტოლებათა სისტემა და ვიპოვოთ $a_0$ და $a_1$ ლოგარითმები:

$$\begin{cases} n \log a_0 + \log a_1 \sum x = \log y \\ \log a_0 \sum x + \log a_1 \sum x^2 = \log yx \end{cases} \qquad (8.8).$$



ანტილოგარითმების დახმარებით ვიპოვით $a_0$ და $a_1$

პარამეტრების მნიშვნელობებს.

ზემომოყვანილი განტოლებანი ასახავს წყვილად კორელაციას. მაგრამ თითოეულ მოვლენაზე მოქმედებს არა მხოლოდ ერთი ფაქტორი, არამედ მრავალი. ამ უკანასკნელ შემთხვევაში საქმე გვაქვს მრავლობით კორელაციასთან, რომლის განტოლება წრფივი დამოკიდებულების შემთხვევაში ასეთია:

$$y = a_0 + ax_1 + a_2x_2 + ... + a_nx_n \qquad (8.10),$$

სადაც $x_1, x_2, ..., x_n$ არის $y$ საშედეგო მოვლენის განვითრებაზე მოქმედი ფაქრორები. პარამეტრების გასაანგარიშებლად გვაქვს განტოლებათა სისტემა:

$$\begin{cases} na_0 + a_1\Sigma x_1 + a_2\Sigma x_2 + ... + a_n\Sigma x_n = \Sigma y \\ a_0\Sigma x_1 + a_1\Sigma x_1^{\ 2} + a_2\Sigma x_1x_2 + ... + a_n\Sigma x_1x_n = \sum x_1y \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_0\Sigma x_n + a_1\Sigma x_1x_n + a_2\Sigma x_2x_n + ... + a_n\sum x_n^2 = \Sigma yx_n \end{cases} \qquad (8.11),$$

$x_i$-ს თითოეული კოეფიციენტი გვიჩვენებს, რამდენი ერთეულით შეიცვლება $y$, $x_i$-ის ერთი ერთეულით ცვალებადობისას და სხვა ფაქტორების ფიქსირებული მნიშვნელობის შემთხვევაში.

## 4. წყვილადი კორელაცია და მისი სტატისტიკური შესწავლა

ეკონომიკაში, ბიზნესსა და მენეჯმენტში სოცილიალურ-ეკონომიკური მოვლენებისა და პროცესების სტატისტიკური ანალიზის შემთხვევაში უმეტესწილად განიხილავენ ორ მოვლენას შორის ურთიერთკავშირს. ასეთი მოვლენები და პროცესები მრავლად გვხვდება სოციალურ-ეკონომიკურ სფეროში. მაგალითად, კავშირი შრომის ნაყოფიერებასა და მუშათა კვალიფიკაციას შორის (მუშათა კვალიფიკაციის ამაღლება იწვევს შრომის ნაყოფიერების გადიდებას), კავშირი წარმოების მოცულობასა და პროდუქციის ერთეულის თვითღირებულების შემცირებას (ფირმაში, ქარხანასა და ფაბრიკებში პროდუქციის წარმოების მოცულობის გადიდება იწვევს პროდუქციის ერთეულის თვითღირებულების შემცირებას), კავშირი ქვეყანაში დანაშაულობათა ზრდასა და სიღარიბეს შორის (სიღარიბის დონის, სიღრმისა და სიმწვავის გადიდება ადიდებს სხვადასხვა სახის დანაშაულობათა რაოდენობას ქვეყანაში) და ა.შ.

**ორ მოვლენას შორის კავშირს უწოდებენ წყვილად კორელაციას, ხოლო ასეთი კავშირების სტატისტიკურ ანალიზს–კორელაციურ ანალიზს.**

**ხშირად წყვილად კორელაციას წყვილად რეგრესიას უწოდებენ, რაც სავსებით დასაშვებია, ვინაიდან კორელაცია კავშირის ფორმას ასახავს, ხოლო რეგრესია კავშირის ფორმის გამომსახველი განტოლებაა.**

**წყვილადი კორელაცია ანუ რეგრესია შეიძლება იყოს ორი სახის: წრფივი ანუ სწორხაზოვანი და არაწრფივი ანუ მრუდხაზოვანი.** როგორც ზემოთ დავინახეთ, პირველ შემთხვევაში კავშირის ანალიზური გამომსახველი განტოლებაა $y = a_0 + a_1 x$, ხოლო მეორე შემთხვევაში–ჰიპერბოლა $y = a_0 + a_1 \dfrac{1}{x}$ ან პარაბოლა

$y = a_0 + a_1x + a_2x^2$. ზოგადად არაწრფივ კავშირს ხშირად გამოსახავენ ნახევარლოგა-რითმული ფუნქციით:

$$y = a_0 + a_1 \log x \qquad (8.18).$$

წყვილადი კორელაციის წრფივი და არაწრფივი ფორმების გამოსავლენად მრავალი მეთოდი არსებობს. მათგან გავრცელბულია ორი: კავშირის ფორმის გამოვლენის გრაფიკული მეთოდი და ვიზუალური მეთოდი. გრაფიკული მეთოდი გულისხმობს ემპირიული მონაცემების საფუძველზე შესაბამისი გრაფიკის აგებას. თუ გრაფიკი სწორხაზოვანია, ამბობენ, რომ მოცემული წვილადი კორელაცია წრფივი სახეობისაა, ხოლო თუ მრუდხაზოვანია – არაწრფივი სახეობის.

**ვიზუალური მეთოდი** ყველაზე მარტივია და გულისხმობს მიზეზობრივი და საშედეგო მოვლენების განვითარების შესწავლას ვიზუალურად, განვითარების დათვალიერებას. თუ ეს მოვლენები იცვლება (იზრდება ან მცირდება) არითმეტიკული პროგრესიით, მაშინ განვითარება წრფივი სახეობისაა, ხოლო თუ გეომეტრიული პროგრესიით იცვლება, მაშინ არაწრფივი სახეობისაა: პარაბოლური, ჰიპერბოლური ან მაჩვენებლიან-ხარისხობრივი.

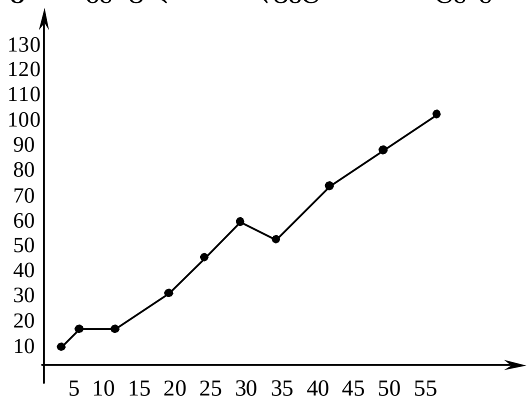
ფირმის საწარმოთა ძირითადი კაპიტალი და გამოშვევული პროდუქცია (მლნ. ლარი)

ცხრილი №18

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ძირიტადი კაპიტალი x | 6 | 8 | 12.5 | 19 | 22.5 | 27.5 | 30 | 40 | 45.5 | 50 |
| | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ | $x_9$ | $x_{10}$ |
| გამოშვებული როდუქციის მოცულობა y | 14 | 20 | 19 | 32.5 | 40 | 50.5 | 47.5 | 62.5 | 91.5 | 122.5 |
| | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | $y_8$ | $y_9$ | $y_{10}$ |

როგორც ჩანს ძირითადი კაპიტალის ღირებულება და გამოშვებული პროდუქციის მოცულობა დაახლოებით არითმეტიკული პროგრესიით დიდდება. ამიტომ მოცემულ შემთხვევაში წყვილადი კორელაცია წრფივი ფორმისაა. ჩვენი მოსაზრების სისწორის დამტკიცების მიზნით გრაფიკული

მეთოდიც გამოვიყენოთ. ამისათვის დეკარტეს მართკუთხა კოორდინატთა სისტემის აბსცისთა ღერძზე გადავზომოთ ძირითადი კაპიტალის ღირებულებანი, ხოლო ორდინატთა ღერძზე – გამოშვებული პროდუქციის მონაცემები.




ნახ. 24. ძირითადი კაპტალისა და გამოშვებული პროდუქციის ურთიერთდამოკიდებულების გრაფიკი

ნახაზი ნათლად გვიჩვენებს, რომ მაჩვენებელთა შორის ურთიერთდამოკიდებულების გრაფიკი სწორხაზოვანია, რადგან მათი წრფეზე ურთიერთდამკვეთი წერტილები ფაქტობრივად სწორ ხაზზეა განლაგებული ან მასთან მიახლოებულია. ამიტომ მოცემულ შემთხვევაში ძირითადი კაპტალის მოცულობასა და გამოშვებული პროდუქციის ურთიერთდამოკიდებულების ანალიზური ფორმა შეიძლება წრფივი განტოლებით ($y = a_0 + a_1 x$) გამოისახოს.

ახლა ისმის კითხვა? კი მაგრამ რამდენად მოქმედებს მიზეზობრივი მოვლენის განვითარების ცვალებადობა საშედეგო მოვლენის განვითარების ცვალებადობაზე? როგორია მათ შორის კავშირის რაოდენობრივი თანაფარდობანი? ამისათვის ჯერ უნდა შევარჩიოთ მოცემული ემპირიული მონაცემების განვითარების ამსახველი განტოლება. როგორც ზემოთ

დავინახეთ, ჩვენს მაგალითზე, ძირითადი კაპტალისა და გამოშვებული პროდუქციის ურთიერთკავშირი წრფივი ფორმისაა და გამოისახება წრფივი განტოლებით: $\hat{y} = a_0 + a_1 x$, სადაც $\hat{y}$ - გამოშვებული პროდუქციის მოცულობის (საშედეგო მოვლენის) მოსწორებული დონეებია;

$x$ – ძირითადი კაპიტალის ღირებულება (მიზეზობრივი მოვლენა). $a_0$ და $a_1$ პარამეტრებია, რომლებიც ამ ორ მოვლენას შორის კავშირის რაოდენობრივ თანაფარდობაზე მიანიშნებენ. მაშასადამე, ჩვენი შემთხვევა წყვილადი კოლეციის წრფივი სახეობის შემთხვევაა. $a_0$ და $a_1$ პარამეტრების გასაანგარიშებლად ზემოთმოტანილი განტოლებათა სისტემები მოცემულია სტატისტიკაში მზამზარეული ფორმით, რაც ზოგჯერ გაუგებრობას იწვევს. ამიტომ მიზანშეწონილია ვაჩვენოთ რა მათემატიკური აპარატის გამოყენებით მიიღება ასეთი სისტემები. ამოსავალი ამ სისტემების მისაღებად არის უმცირეს კვადრატთა მეთოდი, რომელიც შეიმუშავა კ.ფ. გაუსმა (1777-1855) და რომლის კრიტერიუმია შემდეგი სახის გამოსახულება: $\Sigma(y - \hat{y})^2 = \min$

სადაც $y$ - საშედეგო მოვლენის ემპირიული, ფაქტობრივი დონეებია;

$\hat{y}$ – მოსწორებული, თეორიული დონეები.

მოსწორებული ანუ თერიული დონეები ($\hat{y}$) ისეთი მაჩვენებლებია, რომლებიც აღმოფხვრის ემპირიული დონეების ნახტომისებური განვითარების ჭრელ სურათს და წარმოადგენს მას მზარდი ან კლებადი ტენდენციის სურათის სახით. ჩვენს მაგალითზე, ფირმის საწარმოთა მიერ გამოშვებული პროდუქციის მოცულობა ჯერ იზრდება (პირველი და მეორე დონეები), შემდეგ მცირდება (მესამე დონე), შემდეგ ისევ იზრდება და ა.შ. სწორედ ასეთი ზიგზაგისებური განვითარების სურათის

ალმოსაფხვრელადაა საჭირო მოსწორებული ანუ თეორიული დონეების გაანგარიშება. უმცირეს კვადრატთა მეთოდის კრიტერიუმი გულისხმობს ისეთი თეორიული დონეების გაანგარიშებას, რომლისგანაც ემპირიული დონეების გადახრების კვადრატების ჯამი მინიმალური იქნება. ასეთი დონეების გასაანგარიშებელი განტოლებათა სისტემის მოსაძებნად უნდა მოიძებნოს $S=\Sigma\left(y-\hat{y}\right)^2$ მინიმუმი. თუ $\hat{y}$-ის ნაცვლად ჩავსვამთ მის მნიშვნელობას ($a_0+a_1x$), გვექნება:

$$S=\Sigma\left(y-a_0-a_1x\right)^2=\min$$

ასეთი ფუნქციის მინიმალური მნიშვნელობის საპოვნელად არსებობს ორი გზა. პირველი გზა გულისხმობს აღნიშნული განტოლების ისეთ გარდაქმნაში, რომელიც მას მისცემს კვადრატული სამწევრის სახეს[1]. ამისათვის საჭიროა მოცემული ფუნქციის პირველ წევრად ჩათვალოდ $y$, ხოლო მეორე წევრად $a_0-a_1x$ და სახვაობა ავიყვანოთ კვადრატში. გვექნება:

$$\begin{aligned}
S&=\Sigma\left(y-a_0-a_1x\right)^2=\Sigma\left[y^2-2y(a_0-a_1x)+(a_0-a_1\ x)^2\right]=\\
&=\Sigma\left(y^2-2a_0y+2a_1xy+a_0{}^2-2a_0a_1x+a_1^2x^2\right)=\\
&=\Sigma\left(y^2+a_0{}^2+a_1{}^2x^2-2a_0y-2a_0a_1x+2a_1xy\right)=\\
&=\Sigma y^2+\Sigma a_0{}^2+a_1{}^2\Sigma x^2-2a_0\Sigma y-2a_0a_1\Sigma x+2a_1\Sigma xy
\end{aligned}$$

ტოლობის მარჯვენა მხარე შეგვიძლია წარმოვადგინოთ ორი კვადრატული სამწევრის სახით: 1) კვადრატული სამწევრა $a_0$-ის, $f(a_0)$ მიმართ და 2) კვადრატული სამწევრა $a_1$-ის მიმართ. გვექნება გამოსახულება ($\Sigma a_0$ შეცვლილია მისი ექვივალენტური სიდიდთ $na_0$):

---

[1]იხ. Ï. Ëàíãå À. Áàíàñèíñêèé, Òåîðèÿ ñòàòèñòèêè (ïåðåâîä) Ñòàòèñòèêà, Ìîñêâà, 1971, ñòð 256-257

$$S = f(a_0) = na_0^{\ 2} - 2a_0\Sigma y + 2a_0a_1\Sigma x + c\ ,$$

სადაც $c$ როგორც $ax^2+bx+c$ სამწევრის თავისუფალი წევრი ამ შემთხვევაში მოიცავს იმ წევრების ჯამს, რომლებიც არ შეიცავენ საძიებელი ანუ უცნობი პარამეტრის ($a_0$) რაიმე მნიშვნელობას $\sum y^2 + a_1^2 \sum x^2 + 2a_1 \sum xy$. გამოსახულებას $na_0^{\ 2} - 2a_0\Sigma y + 2a_0a_1\Sigma x + c$ თუ გარდავქმნით, გვექნება $S = f(a_0) = na_0^{\ 2} + 2a_0(a_1\Sigma x - \Sigma y) + c$. მიღებული ტოლობის მარჯვენა მხარე $f(x) = ax^2 + bx + c$ კვადრატული სამწევრას ტიპიური გამოსახულებაა, სადაც $a = n, b = 2(a_1\Sigma x - \Sigma y)$. მათემატიკიდან ცნობილია, რომ კვადრატული სამწევრის $f(x) = ax^2 + bx + c$ წარმოებული $f'(x) = 2ax + b$ არსებობს ყველა $x \in b$- ისათვის და ერთადერთ $x = -\dfrac{b}{2a}$ წერტილში. ამ წერტილში ფუნქციის მნიშვნელობა

$$f\left(-\frac{b}{2a}\right) = a\frac{b^2}{4a^2} + b\left(-\frac{b}{2a}\right) + c = \frac{D}{4a},$$

სადაც $D = b^2 - 4ac$, რაც $f(x) = ax^2 + bx + c$ ფუნქციის დისკრიმინანტია. ამასთან $-\dfrac{b}{2a}$ ექსტრემალურ წერტილში ფუნქცია არის მაქსიმალური მნიშვნელობის, თუ $a < 0$, ხოლო თუ $a > 0$, მაშინ – მინიმალური მნიშვნელობის.

ჩვენს შემთხვევაში $a > 0$ ($n < 0$) და რადგან საძიებელი სიდიდეა $a_0$, ამიტომ შეგვიძლია დავწეროთ:

$$a_0 = -\frac{2(a_1\sum x - \sum y)}{2n} = -\frac{a_1\sum x - \sum y}{n}\ ; \quad na_0 = \sum y - a_1\sum x$$

აქედან $na_0 + a_1\Sigma x = \Sigma y$. ეს არის პირველი განტოლება, რომელიც მივიღეთ იმ შემთხვევისათვის, როცა სამწევრის მინიმალური მნიშვნელობა უმცირეს კვადრატთა მეთოდით გამოყენებისათვის ვიპოვეთ $a_0$–ის მიმართ.

ეხლა $a_1$–ის მიმართ შევადგინოთ სამწევრას გამოსახულება.

ზემოთ მოტანილი ტოლობის მარჯვენა მხარე

$$S = \sum y^2 + \sum a_0{}^2 + a_1^2 \sum x^2 - 2a_0 \sum y - 2a_0 a_1 \sum x + 2a_1 \sum xy \quad (8.13)$$

შეგვიძლია $a_1$-ის მიმართ წარმოვადგინოთ შემდეგნაირად:

$$S = f(a_1) = a_1{}^2\Sigma x^2 - 2a_0 a_1 \Sigma x + 2a_1 \Sigma xy + c, \qquad (8.14)$$

სადაც თავისუფალი წევრი $c$ იმ შესაკრებელთა ჯამია, რომლებიც არ შეიცავენ $a_1$-ს, როგორც საძიებელ სიდიდეს. თუ $2a_1$-ს გავიტანთ ფრჩხილებს გარეთ, გვექნება:

$$S = f(a_1) = a_1{}^2\Sigma x^2 + 2a_1\left(a_0\Sigma x - \Sigma xy\right) + c.$$

ამ კვადრატულ სამწევრში $a = \Sigma x^2$-ს. აქედან $a_1$-ის მინიმალური მნიშვნელობა იქნება $-\dfrac{b}{2a}$,

$$\text{ანუ} \quad a_1 = -\frac{2(a_0 \sum x - \sum xy)}{2\sum x^2} = -\frac{a_0 \sum x - \sum xy}{\sum x^2}$$

$$a_1\Sigma x^2 = \Sigma xy - a_0\Sigma x \qquad\qquad a_0\Sigma x + a_1\Sigma x^2 = \Sigma xy.$$

ეს არის მეორე განტოლება $a_1$-ის მიმართ. ახლა შეგვიძლია დავწეროთ $a_0$ და $a_1$ პარამეტრების გასაანგარიშებელ განტოლებათა სისტემა:

$$\begin{cases} na_0 + a_1\Sigma x = \Sigma y \\ a_0\Sigma x + a_1\Sigma x^2 = \Sigma xy \end{cases}$$

მეორე გზა ასეთი განტოლებათა სისტემის მისაღებად $S = \sum(y - a_0 - a_1 x)^2$ გაწარმოების წესია.

უმაღლესი მათემატიკიდან ცნობილია, რომ ორი უცნობის $S = f(a_0, a_1) = \sum(y - a_0 - a_1x)^2$ ფუნქციის მნიშვნელობამ ექსტრემუმს შეიძლება მიაღწიოს მხოლოდ იმ შემთხვევაში, როცა მოცემული ფუნქციის პირველი რიგის კერძო წარმოებულები ნულის ტოლია, ე. ი. როცა

$$\frac{\partial S}{\partial a_0} = 0 \quad \text{და} \quad \frac{\partial S}{\partial a_1} = 0 \qquad (8.15)$$

გავიხსენოთ გაწარმოების ზოგიერთი წესი უმაღლესი მათემატიკიდან.

პირველ რიგში უნდა გავიხსენოთ, რომ ფუნქცია $S = f(a_0, a_1) = \Sigma\left[y - \left(a_0 + a_1x\right)\right]^2 = \min$, რომელიც საჭიროა $a_0$ და $a_1$ პრამეტრების გასაანგარიშებელ განტოლებათა სისტემის მისაღებად ჯერ უნდა გავაწარმოოთ $a_0$ მიმართ ანუ მოვძებნოთ პირველი რიგის კერძო წარმოებული $a_0$–ის მიმართ, შემდეგ ასეთნაირად უნდა მოვიქცეთ $a_1$–ის მიმართ, არის რთული ფუნქცია. გრ. ხიდაშელს[1] მოჰყავს ორი მაგალითი:

$$y = \lg x \qquad (8.16)$$

$$y = \lg(x - 2x) \qquad (8.17)$$

და მიანიშნებს, რომ (1) ფუნნქციაში არგუმენტი არის $x$,

ხოლო (2) ფუნქციაში–გამოსახულება $x - 2x$, რომელიც დამოუკიდებელი $x$ ცვლადის მიმართ თავისთავად წარმოადგენს ფუნქციას. **ფუნქციას, რომლის არგუმენტი თავის მხრივ ფუნქციას წარმოადგენს, ფუნქციის ფუნქცია, ანუ რთული ფუნქცია ეწოდება. თუ $x - 2x$ გამოსახულებას აღვნიშნავთ $U$-თი:**

---

[1]იხ. გრ. ხიდეშელი, უმაღლესი მატემატიკის ელემენტები, სახელმძღვანელო, თბილისის უნივერსიტეტის გამომცემლობა, თბ., 1973, გვ. 368

$$U = x - 2x, \qquad (8.18)$$

მაშინ რთული ფუნქცია 8.17შეიძლება ასე ჩავწეროთ:

$y = \lg U$ , სადაც$U = x - 2x$.

საზოგადოდ, მათემატიკაში რთულ ფუნქციას ასე ჩაწერენ:

$$y = f\left[\varphi(x)\right]$$

როგორც ჩანს „ ...$y$ ცვლადი საბოლოოდ $x$-ის ფუნქციაა, მაგრამ $x$-ზე დამოკიდებულია არა უშუალოდ, არამედ დამხმარე ცვლადის $U$ მეშვეობით".

**რთული ფუნქციის გაწარმოების წესი შემდეგნაირია: ჯერ რთული ფუნქცია უნდა გავაწარმოოთ დამხმარე ცვლადის მიმართ, ხოლო დამხმარე ცვლადი, დამოუკიდებელი ცვლადის მიმართ და მიღებული შედეგები ერთმანეთზე გადავამრავლოთ.**

დამხმარე ცვლადი ჩვენს შემთხვევაში არის ხარისხოვანი ფუნქცია $\left[\left(y-(a_0+a_1x)\right]^2$, რომელიც საჭიროებს ხარისხოვანი ფუნქციის გაწარმოების წესის გამოყენებას. ხარისხოვანი ფუნქციის წარმოებული უდრის ხარისხის მაჩვენებელი გამრავლებული იმავე არგუმენტზე 1-ით ნაკლებ ხარისხში. მაგალითად, $y'(x) = x^3 = 3x^2$ $y'(x) = x^2 = 2x$ და ა შ. ჩვენს მაგალითზე რთული ფუნქცია გაწარმოებული დამხმარე ცვლადით იქნება:

$$\frac{\partial S}{\partial a_0} = 2\sum(y - a_0 - a_1x) = 0 \qquad (8.19)$$

$$\frac{\partial S}{\partial a_1} = 2\sum(y - a_0 - a_1x) = 0 \qquad (8.20)$$

მაგრამ თუ არგუმენტს $(y - a_0 - a_1x)$ განვიხილავთ თავისთავად ფუნქციის სახით დამოუკიდებელი ჯერ $a_0$ და შემდეგ $a_1$ ცვლადების მიმართ, გვექნება:

$$\frac{\partial S}{\partial a_0} = 2\sum(0-1-0) = -1 \qquad (8.21),$$

$$\frac{\partial S}{\partial a_1} = 2\sum(0-0-x) = -x \qquad (8.22).$$

რადგანაც დამოუკიდებელი ცვლადის წარმოებული უდრის ერთს ($a_0$ პირველ შემთხვევაში და $a_1$–მეორე შემთხვევაში ერთის ტოლია, ისე როგორც $x' = 1$), ხოლო მუდმივების წარმოებული ნულის ტოლია, ისე როგორც $c' = 0$. პირველ შემთხვევაში მუდმივებია $y$ და $a_0$, ხოლო მეორე შემთხვევაში $y$ და $a_1$. საბოლოოდ როგორც რთული ფუნქციის გაწარმოების წესი მოითხოვს თუ გაწარმოების შედეგებს ერთმანეთზე გადავამრავლებთ და თითოეულ მათგანს გავუტოლებთ ნულს, მივიღებთ:

$$\frac{\partial S}{\partial a_1} = 2\sum(y-a_0-a_1x)(-1) = -2\sum(y-a_0-a_1x) = 0$$

$$\frac{\partial S}{\partial a_1} = 2\sum(y-a_0-a_1x)(-x) = -2\sum(xy-a_0\sum x-a_1\sum x^2) = 0$$

თუ ორთავე ტოლობის მარჯვენა მხარეს – 2-ზე შევკვეცავთ, მივიღებთ:

$$\Sigma y - na_0 - a_1\Sigma x = 0 \qquad na_0 + a_1\Sigma x = \Sigma y$$

$$\Sigma xy - a_0\Sigma x - a_1\Sigma x^2 = 0 \qquad a_0\Sigma x + a_1\Sigma x^2 = \Sigma xy$$

მივიღეთ განტოლებათა სისტემა, რომელთა ამოსახსნელად და $a_0$ და $a_1$ პარამეტრების გასაანგარიშებლად მრავალი ჩვენთვის ცნობილი ხერხი არსებობს მათემატიკაში. მაგრამ აქედან ყველაზე მოსახერხებელი, ჩვენი აზრით, კრამერის

ფორმულების გამოყენებაა.

კრამერის ფორმულები და საერთოდ მატრიცული უმაღლესი ალგებრის ცემენტების გამოყენება ძალიან დიდ ეფექტს იძლევა ეკონომიკურ გაანგარიშებებში საბაზრო ეკონომიკაზე გარდამავალ პერიოდში. ასეთია მაგალითად, მატრიცული ალგებრის ელემენტების გამოყენება საანგარიშო დარგთაშორისი ბალანსის ოპტიმალურ გაანგარიშებებში და სხვა. ჩვენი ნორმალურ განტოლებათა სისტემა მატრიცული ფორმითა და მატრიცების ერთმანეთზე გადამრავლების წესის გათვალიწინებით შემდეგნაირად წარმოვადგინოთ:

$$\begin{pmatrix} n & \Sigma x \\ \Sigma x & \Sigma x^2 \end{pmatrix} \begin{vmatrix} a_0 \\ a_1 \end{vmatrix} = \begin{vmatrix} \Sigma y \\ \Sigma xy \end{vmatrix} \qquad (8.23).$$

ეს ჩანაწერი მსგავსია წრფივ ალგებრაში ცნობილი ჩანაწერისა $\overline{A}x = \overline{B}$, სადაც $A$, $a$ სახის კოეფიციენტებით შედგენილი მატრიცაა, $\overline{x} - x$-ის ვექტორი, ანუ იგივე ერთსვეტოვანი მატრიცა და $\overline{B} - B$ სახის ვექტორი. როგორც ვიცით აქ საძიებელი სიდიდეებია $a_0$ და $a_1$. კრამერის ფორმულების თანახმად თითოეული უცნობი უდრის წილადს, რომლის მნიშვნელია ამ სისტემის დეტერმინანტი $\begin{pmatrix} n & \Sigma x \\ \Sigma x & \Sigma x^2 \end{pmatrix}$,

ხოლო მრიცხველი იგივე დეტერმინანტია, მაგრამ იმ განსხვავებით, რომ მასში საძიებელი უცნობის შესაბამისი კოეფიციენტებით შედგენილი ვექტორი შეცვლილია თავისუფალი წევრების $\begin{vmatrix} \Sigma y \\ \Sigma xy \end{vmatrix}$ ვექტორით.

აქედან გამომდინარე შეგვიძლია დავწეროთ:

$$a_0 = \frac{\begin{vmatrix} \Sigma y & \Sigma x \\ \Sigma xy & \Sigma x^2 \end{vmatrix}}{\begin{vmatrix} n & \Sigma x \\ \Sigma x & \Sigma x^2 \end{vmatrix}} = \frac{\Sigma y \Sigma x^2 - \Sigma xy \Sigma x}{n\Sigma x^2 - (\Sigma x)^2} \qquad (8.24),$$

$$a_1 = \frac{\begin{vmatrix} n & \Sigma y \\ \Sigma x & \Sigma xy \end{vmatrix}}{\begin{vmatrix} n & \Sigma x \\ \Sigma x & \Sigma x^2 \end{vmatrix}} = \frac{n\Sigma yx - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} \qquad (8.25).$$

გავიანგარიშოთ პარამეტრები ჩვენს მიერ მოტანილი მაგალითის საფუძველზე.

ძირითად კაპიტალისა და გამოშვებულ პროდუქციას შორის წრფიული კავშირის ინფორმაცია.

ცხრილი №19

| ძირითადი კაპიტალის ღირებულება (x) (მლნ. ლარი) | გამოშვებული პროდუქციის ღირებულება (y) (მლნ. ლარი) | $x^2$ | $y^2$ | ?=7.2+2.26x |
|---|---|---|---|---|
| 6 | 14 | 36 | 84 | |
| 8 | 20 | 64 | 160 | |
| 12.5 | 19 | 156.3 | 237.5 | |
| 19 | 32.5 | 361 | 617.5 | |
| 22.5 | 40 | 506.3 | 900 | |
| 27.5 | 50.5 | 756.3 | 1388.8 | |
| 30 | 47.5 | 900 | 1425 | |
| 40 | 62.5 | 1600 | 2500 | |
| 45.5 | 91.5 | 2070.3 | 4163.3 | |
| 50 | 122.5 | 2500 | 5125 | |
| Σ 261.0 | 500.3 | 8950.2 | 17901 | |

ამ მონაცემების საფუძველზე შეგვიძლია დავწეროთ ნორმალურ განტოლებათა სისტემა:

$$10a_0 + 261a_1 = 500.3$$

$$261a_0 + 8950.2a_1 = 17901$$

კრამერის ფორმულების გამოყენებით ამ სისტემის $a_0$ და $a_1$ პარამეტრების გასაანგარიშებლად გვექნება:

$$a_0 = \frac{\Sigma x^2 \Sigma y - \Sigma xy \Sigma x}{n\Sigma x^2 - (\Sigma x)^2} = \frac{192376}{21382} = -7.2$$

$$a_1 = \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} = \frac{48432}{21381} = 2.2$$

განტოლება მიიღებს სახეს: $\hat{y} = -7.2 + 2.26x$ (მოსწორებული დონეები იხილეთ ცხრილის ბოლო სვეტში). მოსწორებული დონეების ჯამი უდრის 500-ს, ხოლო ემპირიული დონეების ჯამი–500,3-ს, $\sum(\text{y - ŷ})^2 = (500.3 - 500)^2 = 0.3^2 = 0.09$, რაც აკმაყოფილებს ამოცანის მინიმიზაციის პირობას. ეს იმას ნიშნავს, რომ შერჩეულმა წრფივმა განტოლებამ ადექვატურად ასახა ემპირიული დონეების განვითარება.

შევნიშვავთ, რომ ზოგჯერ განტოლების თავისუფალი წევრი $a_0$ ეკონომიკური შინაარსის მიხედვით უნდა იყოს დადებითი, რადგანაც ის საშედეგო მოვლენის რაღაც საწყისი დონეა, ხოლო რეგრესიის კოეფიციენტი $a_1$ გვიჩვეებს მიზეზობრივი ფაქტორის ვარიაციისა და საშედეგო მოვლენის ვარიაციას შორის კავშირის სიმჭიდროვის ძალას. რაოდენობრივად ის გვიჩვენებს შესაბამის ზომის ერთეულებში მიზეზობრივი მოვლენის ინდივიდუალური მნიშვნელობის ამავე მოვლენის (ფაქტორის) საშუალო მნიშვნელობიდან ერთი ერთეულით გადახრა (გადიდება ან შემცირება), რამდენი ერთეულით გამოიწვევს საშედეგო ფაქტორის ( $y$ ) ინდივიდუალური მნიშვნელობის გადახრას ამავე საშედეგო ფაქტორის საშუალო მნიშვნელობიდან.

ჩვენს მაგალითზე შეიძლება ითქვას, რომ ძირითადი კაპიტალის ერთი მილიონი ლარით გადიდება ძირითადი კაპიტალის საშუალო წლიურ ღირებულებასთან შედარებით 2,2 მლნ ლარით გაადიდებს გამოშვებული პრიდუქციის წლიურ მოცულობას ამავე მაჩვენებლის საშუალო წლიურ მაჩვენებელთან შედარებით. რით აიხსნება $a_0$ - უარყოფითი მნიშვნელობა?

ზოგიერთი ავტორის[1] მოსაზრებით ეს იმით აიხსნება, რომ საშედეგო ნიშნის ($y$) არსებობის არეალი არ მოიცავს მიზეზობრივი ფაქტორის ($x$) ნულოვან ან მასთან ახლო მდგომ მნიშვნელობებს. ამისათვის ავტორის აზრით, შეიძლება გავიანგარიშოთ $x$ ფაქტორის მინიმალური შესაძლებელი მნიშვნელობა, რომლისთვისაც უზრუნველყოფილი იქნება $y$ საშედეგო ფაქტორის მინიმალური მნიშვნელობა (ცხადია დადებითი).

ჩვენს მაგალითზე $x_{\min} = a_0 : a_1 = 7.2 : 2.2 = 3.3$ მლნ. ლარი.

ეს არის ძირითადი ფონდების მინიმალური მოცულობა, რომლითაც მიიღწევა გამოშვებული პროდუქციის მინიმალური წლიური მოცულობა.

ავტორის (მ. იუზბაშევი) აზრით თუ $y$-ის არსებობის არეალი მოიცავს $x$-ის ნულოვან მნიშვნელობას, მაშინ თავისუფალი წევრი ($a_0$) დადებითია და აღნიშნავს საშედეგო მოვლენის საშუალო მნიშვნელობას.

## 5. პარაბოლური წყვილადი კორელაცია

წყვილადი კორელაციის ერთ-ერთი მნიშვნელოვანი სახეობაა პარაბოლური, არაწრფივი კორელაცია, რომელიც შეიძლება


211

[1]იხილეთ, მაგალითად, Åëèñåååâà È. È. Þçáàøåâ Ì. Ì. Îáùàÿ òåîðèÿ ñòàòèñòèêè , Ó÷åáíèê, ïîä ðåäàêöèåé È. È. Åëèñååâîé – Ì. Ôèíàíñû è ñòàòèñòèêà , 1995 ñ 208.

 212

იყოს მე-2 რიგის ( $\hat{y} = a_0 + a_1x + a_2x^2$ ), მე-3 რიგის ( $\hat{y} = a_0 + a_1x^2 + a_2x^3$ ) და ა. შ.

პარაბოლური სახის წყვილად კორელაციასთან მაშინ გვექნება საქმე, როცა მიზეზობრივი მოვლენის ( $x$ ) თანაბარი (არითმეტიკული პროგრესიის) ცვლილებასთან დაკავშირებით საშედეგო მოვლენა იცვლება (იზრდება ან მცირდება) შედარებით სწრაფად. ამ ცვალებადობის სისწრაფის ხარისხის მიხედვითაა სწორედ განსხვავებული პარაბოლის სახეობანი (მე-2 რიგის ანუ კვადრატული, მე–3 რიგის ანუ კუბური პარაბოლა და ა. შ.).

პარაბოლური წყვილადი კორელაციის შემთხვევაში რაგრესიის განტოლება ასეთი სახისაა:

$$\hat{y} = a_0 + a_1x + a_2x^2 \qquad (8.26).$$

აქაც პარამეტრების $a_0, a_1, a_2$ -ის საპოვნელად ვიყენებთ უმცირეს კვადრატთა მეთოდს, რომლის მიხედვით $\Sigma(y - \hat{y})^2 = \min$ .

თუ ამ გამოსახულებაში $\hat{y}$ - ის ნაცვლად ჩავსვამთ $\hat{y} = a_0 + a_1x + a_2x^2$ გამოსახულებას, გვექნება $\Sigma\left(y - a_0 - a_1x - a_2x^2\right)^2 = \min$ . როგოც ვიცით ეს ფუნქცია $S = f(a_0, a_1, a_2)$ მინიმალურ მნიშვნელობას ღებულობს მისი პირველი რიგის წარმოებულის ნულთან გატოლების შემთხვევაში.

ამიტომ საჭიროა ამ ფუნქციის, როგორც რთული ფუნქციის პირველი რიგის, კერძო წარმოებულის მოძებნა ჯერ $a_0$-ის მიმართ და ნულთან გატოლება, ასეთნაირი ოპერაციის ჩატარებაა საჭირო $a_1$-ის და გვექნება:

$a_2$ -ის მიმართ.



214

$$\frac{\partial s}{\partial a_0} = 2\Sigma\left(y - a_0 - a_1 x - a_2 x^2\right) = 0$$

$$\frac{\partial s}{\partial a_1} = 2\Sigma\left(y - a_0 - a_1 x - a_2 x^2\right) = 0 \qquad (8.27).$$

$$\frac{\partial s}{\partial a_2} = 2\Sigma\left(y - a_0 - a_1 x - a_2 x^2\right) = 0$$

როგორც უკვე ჩვენთვის ცნობილია, რთული ფუნქციის წარმოებულის მოსაძებნად ტოლობის მარჯვენა მხარეზე არსებული გამოსახულება $a_0 - a_1 x - a_2 x^2$, რომელიც თავიდან განხილული იყო როგორც არგუმენტი, ამჯერად უნდა განხვიხილოთ, როგორც $x$ არგუმენტის ფუნქცია და გავაწარმოოთ ჩვეულებრივი ხერხებით, გვექნება:

$$\frac{\partial s}{\partial a_0} = -2\Sigma\left(y - a_0 - a_1 x - a_2 x^2\right) = 0$$

$$\frac{\partial s}{\partial a_1} = -2x\Sigma\left(y - a_0 - a_1 x - a_2 x^2\right) = 0 \qquad (8.28).$$

$$\frac{\partial s}{\partial a_2} = -2x^2\Sigma\left(y - a_0 - a_1 x - a_2 x^2\right) = 0$$

თუ სამივე ტოლობას შევკვეცავთ – 2-ზე და მოვახდენთ მარტივ ალგებრულ გარდაქმნებს, მივიღებთ $a_0, a_1$ და $a_2$ პარამეტრების ანუ რეგრესიის კოეფიციენტების გასაანგარიშებელ ნორმალურ განტოლებათა სისტემას:

$$\begin{cases} na_0 + a_1\Sigma x + a_1\Sigma x^2 = \Sigma y \\ a_0\Sigma x + a_1\Sigma x^2 + a_2\Sigma x^3 = \Sigma xy \\ a_0\Sigma x^2 + a_1\Sigma x^3 + a_2\Sigma x^4 = \Sigma x^2 y \end{cases} \qquad (8.29).$$


215

ამ სისტემის ამოხსნაც მოსახერხებელია იგივე კრამერის დეტერმინანტების გამოყენებით. სტატისტიკოსები ამჯობინებენ სისტემის გამარტივებას $x$-ის ნაცვლად $\left(x-\overline{x}\right)$-ს გამოყენებით. ამ შემთხვევაში სისტემა მიიღებს სახეს:

$$\begin{cases} na_0 + a_1\Sigma\left(x-\overline{x}\right) + a_2\Sigma\left(x-\overline{x}\right)^2 = \Sigma y \\ a_0\Sigma\left(x-\overline{x}\right) + a_1\Sigma\left(x-\overline{x}\right)^2 + a_2\Sigma\left(x-\overline{x}\right)^3 = \Sigma\left(x-\overline{x}\right)y \\ a_0\Sigma\left(x-\overline{x}\right)^2 + a_1\Sigma\left(x-\overline{x}\right)^3 + a_2\Sigma\left(x-\overline{x}\right)^4 = \Sigma\left(x-\overline{x}\right)^2 y \end{cases} \qquad (8.30).$$

ვინაიდან $\Sigma\left(x-\overline{x}\right)=0$, ამიტომ სისტემა მიიღებს სახეს:

$$\begin{cases} na_0 + a_2\Sigma\left(x-\overline{x}\right)^2 = \Sigma y \\ a_1\Sigma\left(x-\overline{x}\right)^2 = \Sigma\left(x-\overline{x}\right)y \\ a_0\Sigma\left(x-\overline{x}\right)^2 + a_2\Sigma\left(x-\overline{x}\right)^4 = \Sigma\left(x-\overline{x}\right)^2 y \end{cases} \qquad (8.31).$$

მეორე განტოლებიდან $a_1 = \dfrac{\Sigma\left(x-\overline{x}\right)y}{\Sigma\left(x-\overline{x}\right)^2}$, ხოლო $a_0$ და $a_2$ მიიღება ორუცნობიანი განტოლებათა შემდეგი სისტემის ამოხსნით:

$$\begin{cases} na_0 + a_2\Sigma\left(x-\overline{x}\right)^2 = \Sigma y \\ a_0\Sigma\left(x-\overline{x}\right)^2 + a_2\Sigma\left(x-\overline{x}\right)^4 = \Sigma\left(x-\overline{x}\right)^2 y \end{cases} \qquad (8.32).$$

ასეთი გზით ჩვენ ვღებულობთ $y$ და $\left(x-\overline{x}\right)^2$-ს შორის ურთიერთკავშირის რეგრესულ განტოლებას

$$\hat{y} = a_0 + a_1\left(x-\overline{x}\right) + a_2\left(x-\overline{x}\right)^2 \qquad (8.33).$$


216

თუ ასეთ განტოლებაში ბოლოს, $a_0$, $a_1$ და $a_2$ პარამეტრების განსაზღვრის შემდეგ, ჩავსვამთ $\overline{x}$, მნიშვნელობას და მიღებულ გამოსახულებას გარდავქმნით, მივიღებთ $y$-სა და $x$ -ს შორის ურთიერთკავშირის პარაბოლურ განტოლებას.

მოვიტანოთ პრაქტიკული მაგალითი აგრობიზნესში სავარგულებზე სასუქების შეტანის რაოდენობასა და მოსავლიანობას შორის ურთიერთკავშირის შესახებ. როგორც ცნობილია სასუქების სასოფლო-სამეურნეო სავარგულებზე შეტანის რაოდენობის გადიდება გარკვეულ საზღვრამდე სხვა თანაბარ პირობებში იწვევს სასოფლო-სამეურნეო შესაბამისი კულტურის (სიმინდის, ბრინჯის, სოიას, ჩაის მწვანე ფოთლის და სხვა) მოსავლიანობის გადიდებას. მაგრამ ეს ხდება არა ყოველთვის, ყველა კონკრეტულ შემთხვევაში, არამედ ზოგადად, საბოლოოდ დაკვირვების საკმარისი რიცხვის პირობებში. ამ შემთხვევაში ამ ორ მოვლენას შორის არსებობს კორელაციური ანუ სტატისტიკური კავშირი, რომელიც ცალკეულ კონკრეტულ შემთხვევაში მოსავლიანობაზე მოქმედი სხვა ფაქტორების (მიწის ნაყოფიერება, ნიადაგის დამუშავების აგროტექნიკური ვადები, ნალექების მოსვლის რეჟიმი წლის მანძილზე და სხვ.) ზეგავლენით შეიძლება გადაიფაროს სასუქების შეტანის ზემოქმედების ფაქტორი და მივიღოთ საწინააღმდეგო სურათი. ვთქვათ გვაქვს ასეთი სურათი:

ცხრილი №20

| **მინერალური სასუქების შენატანის რაოდენობა** ($x$) | 2 | 4 | 6 | 8 | 10 |
|---|---|---|---|---|---|
| მოსავლიანობა, ც/ჰა ($y$) | 32 | 38 | 40 | 44 | 46 |


217

შევადგინოთ, $a_0$, $a_1$ და $a_2$ პარამეტრების გასაანგარიშებელი



218

მონაცემების ცხრილი $x$-ის $x-\bar{x}$-სხვაობით შეცვლის შემთხვევისათვის.

საანგარიშო ცხრილი

ცხრილი. №21

| $(x)$ | $(y)$ | $(x-\bar{x})$ | $(x-\bar{x})^2$ | $(x-\bar{x})^4$ | $y(x-\bar{x})$ | $y(x-\bar{x})^2$ | $\hat{y}$ |
|---|---|---|---|---|---|---|---|
| 2<br>4<br>6<br>8<br>10 | 32<br>38<br>40<br>44<br>46 | -4<br>-2<br>0<br>+2<br>+4 | 16<br>4<br>0<br>4<br>16 | 256<br>16<br>0<br>16<br>256 | -128<br>-76<br>0<br>88<br>184 | 512<br>152<br>0<br>176<br>736 | 32<br>38<br>41<br>44<br>45 |
| $\bar{x}=6$ | 200 | **0** | 40 | 544 | 68 | 1576 | 200 |

დავწეროთ ნორმალურ განტოლებათა სისტემა $(x)$-ის $(x-\bar{x})$ სხვაობებით შეცვლის პირობებისათვის $n=5$, გვექნება:

$$5a_0 + 40a_2 = 200$$
$$40a_1 = 68$$
$$40a_0 + 544a_2 = 1576$$

აქედან მეორე განტოლების მიხედვით $a_1 = \dfrac{68}{40} = 1.7$, ხოლო $a_0$, და $a_2$ პარამეტრების საპოვნელად გვაქვს შემდეგი სახის ორუცნობიანი განტოლებათა სისტემა:

$$\begin{cases} 5a_0 + 40a_2 = 200 \\ 40a_0 + 544a_2 = 1576 \end{cases}$$

ამ სისტემის ამოსახსნელად თუ კრამერის ფორმულებს გამოვიყენებთ, მივიღებთ:

$$a_0 = 40.85, \quad a_1 = 1.7 \quad \text{და} \quad a_2 = -0.107$$

პარაბოლური განტოლება მიიღებს სახეს:


219

$$\hat{y} = 40.85 + 1.7(x - \bar{x}) - 0.107(x - \bar{x})^2$$

 220

თუ ამ განტოლებაში ჩავსვამთ $\bar{x}$ -ს მნიშვნელობას $(\bar{x}=6)$ და მოვახდენთ ელემენტარულ ალგებრულ გარდაქმნებს, მივიღებთ:

$$\hat{y} = 40.85 + 1.7(x-6) - 0.107(x-6)^2 = 40.85 + 1.7x + 10.2 - 0.107(x^2 - 12x + 36) =$$
$$= 40.85 + 1.7x - 10.2 - 0.107x^2 + 1.284x - 3.852 = 26.7 + 2.984x - 0.107x^2$$

საბოლოოდ ჩვენისაძიებელი პარაბოლური განტოლება:

$$\hat{y} = 26.7 + 2.984x - 0.107x^2 \quad (8.34)$$

მოსწორებული დონეები ცხრილის ბოლო სვეტში და მათი ჯამი მეტყველებს, რომ შერჩეული პარაბოლური ფორმულა ადექვატურად ასახავს ემპირიული დონეების განვითარების სურათს.

## 6. ჰიპერბოლური არაწრფივი წყვილადი კორელაცია

ჰიპერბოლური წყვილადი, არაწრფივი ურთიერთკავშირები ეკონომიკურ ანალიზში, ბიზნესსა და მენეჯმენტში გამოიყენება ისეთ შემთხვევებში, როცა ერთი მოვლენის (მიზეზობრივი მოვლენა – $x$) გადიდება იწვევს მეორე მოვლენის (საშედეგო მოვლენა – $y$) შემცირებას ან პირიიქით, შემცირება იწვევს გადიდებას. მოვლენებისა და პროცესების ამ სახის ურთიერთკავშირის შემთხვევები ძალიან ხშირია ეკონომიკაში. ასეთია, მაგალითად, ურთი-ერთდამოკიდებულება წარმოების მოცულობასა და პროდუქციის ერთეულის თვითღირებულებას და სხვა მოვლენებს სორის.

მაგალითად, ცნობილია, რომ ამა თუ იმ ფირმის მიერ წარმოებული პროდუქციის მოცულობის გადიდება იწვევს პროდუქციის ერთეულის თვითრირებულების[1] შემცირებას და

[1]თვითღირებულება ეწოდება პროდუქციის წარმოებისა და რეალიზაციის დანახარჯების ფულად გამოხატულებას. განსხვავება სააბაზრო ფასსა და თვითღირებულებას შორის წარმოქმნის ფირმის მოგებას. ამიტომ საქმიანი

ბიზნესმენები კონკურენტულ ბრძოლაში ცდილობენ ნაკლები დანახარჯებით აწარმოონ მეტი რაოდენობის პროდუქცია და მიიღონ მაღალი მოგება.



222

ამ საფუძველზე მოგების გადიდებას, ან პირიქით, წარმოების მოცულობის შემცირება იწვევს პროდუქციის ერთეულის თვითღირებულების გადიდებას. პროდუქციის ერთეულის თვითღირრებულებასა $(c)$ და წარმოების მოცულობას $(x)$ შორის ურთიერთკავშირი გამოისახება შემდეგნაირად:

$$c = a + \frac{b}{x} \qquad (8.35)$$

სადაც $a$ – პირობით–ცვალებადი ხარჯებია პროდუქციის ერთეულზე;

$b$ – პირობით-უცვლელი ხარჯები მოცემულ პერიოდში (თვე, კვარტალი, წელი);

$x$ – მოცემულ პერიოდში პროდუქციის გამოშვების მოცულობა;

$c$ – პროდუქციის ერთეულის თვითღირებულება.

პირობით-ცვალებადი ხარჯები ის ხაჯებია, რომელთა საერთო მოცულობა წარმოების მოცულობის ცვალებადობასთან ერთად იცვლება, მაგრამ უცვლელი რჩება პროდუქციის ერთეულზე. ასეთია, მაგალითად, ძირითადი ნედლეულისა და მასალების, აგრეთვე ტექნოლოგიური სათბობის, ტექნოლოგიური ელექტროენერგიის, ძირითადი მუშების ხელფასის და სხვა დანახარჯები. ისე, რომ ეს ხარჯები პირობით ცვალებადია წარმოების მოცულობის მიმართ, ხოლო პირობით-უცვლელია პროდუქციის ერთეულის მიმართ ანუ წარმოების მოცულობის ერთეულზე. პირობით-უცვლელი ხარჯები არის, მაგალითად, ფირმის დროით ანაზღაურებაზე მყოფი მუშაკების ხელფასი, შენობების ამორტიზაცია ან შენობის ქირის და სხვა დანახარჯები, რომლებიც მოცემულ პერიოდში (თვე, კვარტალი, წელი) არ იცვლება და ამიტომ წარმოების მოცულობის გადიდებისას პროდუქციის ერთეულზე მცირდება (ეს კარგად ჩანს ზემოთმოტანილი ფორმულიდან, სადაც $x$-ის გადიდებასთან ერთად მცირდება $c$). ასეთივე

 223

ურთიერთდამოკიდებულებაა მეცხოველეობის ბიზნესში ცხოველის გამოკვების დანახარჯებსა და ასაკს შორის. თავიდან


224

(მეღორეობა, მეძროხეობა, მეფრინველეობას და ა.შ.) ცხოველის გარკვეულ ასაკამდე რაციონალურ გამოკვებასთან ერთად ცხოველის წონითი ნამატის ერთეულზე (მაგალითად, 1 კგ-ზე) ნაკლები დანახარჯებია პირუტყვის გამოკვებაზე საჭირო იმდენად, რამდენადაც პირუტყვის წონა უფრო მეტად მატულობს, ვიდრე გამოკვებაზე საჭირო დანახარჯები. ზრდადასრულებული პირუტყვის გარკვეულ ასაკში შეიძლება ეს ტენდენცია შეიცვალოს. ამიტომ უნარიანი ბიზნესმენები ითვალისწინებენ ამ ფაქტორის გავლენას წარმოების მოგების გადიდების საქმეში და ცდილობენ იმ ასაკში გაუშვან დაკლული პირუტყვი ან ფრინველი და ღორი რეალიზაციაში, რომლის შემდეგ მათ გამოკვებაზე დანახარჯები წონითი ნამატის ერთეულზე გადიდებას იწყებს. ამ თავის მე-3 საკითხში მოტანილია მოვლენათა შორის ურთიერთშებრუნებული კავშირის ჰიპერბოლური განტოლება: $\hat{y} = a_0 + a_1\frac{1}{x}$. აქაც უმცირეს კვადრატითა მეთოდის გამოყენებით $\Sigma(y-\bar{y})^2 = \min$, ან $\Sigma\left(y - a_0 - a_1\frac{1}{x}\right)^2 = \min$ განტოლებაში $\bar{y}$-ის ნაცვლად მისი ტოლი სიდიდის $a_0 + a_1\frac{1}{x}$ ჩასმით, ჯერ $a_0$-ის, შემდეგ $a_1$-ის მიმართ პირველი რიგის კერძო წარმოებულების პოვნისა და ნულთან გატოლებით, ანუ $S$ ფუნქციის $f(a_0, a_0)$ მინიმუმის მოძებნით ვღებულობთ ნორმალურ განტოლებათა სისტემას:

$$\begin{cases} na_0 + a_1\sum\frac{1}{x} = \sum y \\ a_0\sum\frac{1}{x} + a_1\sum\frac{1}{x^2} = \sum\frac{y}{x} \end{cases} \qquad (8.36)$$


225

0 $x$ 1 $x^2$ $x$



226

მოვიტანოთ კონკრეტული მაგალითი:

ფირმას გააჩნია ხუთი სახის წარმოება, რომელთა მიხედვით ეკონომიკური მაჩვენებლები შემდეგ სურათს იძლევა:

ცხრილი №22

| წარმოების ნომრები | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| სასაქონლო პროდუქცია (მლნ. ლარი) $(x)$ | 5.0 | 6.5 | 10.8 | 11.2 | 15.0 |
| დანახარჯები სასაქონლო პროდუქციის 1 ლარზე (ლარი) $(y)$ | 0.98 | 0.94 | 0.91 | 0.85 | 0.80 |

შევადგიოთ ნორმალურ განტოლებათა სისტემის ამოსახსნელი ინფორმაციის ცხრილი:

ცხრილი №23

| წარმოების ნომერი | $(x)$ | $(y)$ | $\frac{1}{x}$ | $\frac{1}{x^2}$ | $y\frac{1}{x^2}$ | $\hat{y} = 0.78 + 1.2\frac{1}{x}$ |
|---|---|---|---|---|---|---|
| 1 | 5.0 | 0.98 | 0.2 | 0.040 | 0.196 | 0.98 |
| 2 | 6.5 | 0.94 | 0.15 | 0.023 | 0.141 | 0.93 |
| 3 | 10.8 | 0.91 | 0.09 | 0.0086 | 0.082 | 0.86 |
| 4 | 11.2 | 0.85 | 0.08 | 0.0079 | 0.068 | 0.86 |
| 5 | 15 | 0.80 | 0.07 | 0.0044 | 0.056 | 0.85 |
| $\Sigma$ | 48.5 | 4.48 | 0.59 | 0.084 | 0.543 | 4.48 |

ჰიპერბოლური განტოლების პარამეტრების ამოსახსნელი ნორმალურ განტოლებათა ზემოთმოტანილი სისტემა განსხვავდება წრფივი განტოლების პარამეტრების ამოსახსნელი სისტემისაგან მხოლოდ იმით, რომ მასში $x$-ის ნაცვლად ჩასმულია $\frac{1}{x}$. ამიტომ კრამერის ფორმულების გამოყენებით წრფივი განტოლების მსგავსად პირდაპირ შეგვიძლია დავწეროთ $a_0$ და $a_1$ პარამეტრების გასაანგარიშებელი ფორმულები:

$$a_0 = \frac{\Sigma y \times \Sigma\frac{1}{x^2} - \Sigma\frac{y}{x} \times \Sigma\frac{1}{x}}{n\Sigma\frac{1}{\ } - \Sigma\frac{1}{\ } \times \Sigma\frac{1}{\ }} \qquad (8.37)$$


227

$x^2$ $x$ $x$


228

$$a_1 = \frac{n\Sigma \frac{y}{x} - \Sigma \frac{1}{x} \times \Sigma y}{n\Sigma \frac{1}{x^2} - \Sigma \frac{1}{x} \times \Sigma \frac{1}{x}} \qquad (8.38)$$

ჩვენი მონაცემების ამ ფორმულაში ჩასმის შედეგად მივიღებთ:

$$a_0 = \frac{4.48 \times 0.084 - 0.543 \times 0.59}{5 \times 0.084 - 0.59 \times 0.59} = 0.78$$

$$a_1 = \frac{5 \times 0.543 - 0.59 \times 4.48}{5 \times 0.084 - 0.59 \times 0.59} = 1.0$$

ამრიგად მივიღეთ განტოლება: $\hat{y} = 0.78 + 1.0\frac{1}{x}$ (8.39),


229

რომელშიც ემპირიული ანუ ფაქტობრივი მონაცემების ( $x$ -ის მნიშვნელობების) შეტანით მივიღებთ $\hat{y}$ -ის შესაბამის მოსწორებულ დონეებს (ეს დონეები ნაჩვენებია ცხრილის ბოლო სვეტში). როგორც ჩანს განსხვავებათა ჯამი ემპირიულ და თეორიულ (მოსწორებულ) დონეთა შორის ნულის ტოლია, რაც მიანიშნებს რეგრესიული ჰიპერბოლური განტოლების შერჩევის სისწორეზე.

## 8. მრავლობითი კორელაცია

წყვილადი კორელაციის განხილვისას საქმე გვქონდა ორ ურთიერთდამოკიდებულ მოვლენასთან, რომელთაგან ერთი იყო მიზეზობრივი ფაქტორი, ხოლო მეორე – საშედეგო. მაგრამ ეკონომიკაში, ბიზნესსა და მენეჯმენტში ამა თუ იმ სახის მოვლენის განვითარებაზე მოქმედებს არა მხოლოდ ერთი, არამედ მრავალი ფაქტორი. ავიღოთ, მაგალითად, ნიადაგებში სასუქების შეტანის რაოდენობასა და ამ საფუძველზე სასოფლო-სამეურნეო კულტურის მოსავლიანობის ცვალებადობა, ან კიდევ მუშათა კვალიფიკაცია და მასთან დაკავშირებული შრომის ნაყოფიერება. მოსავლიანობაზე ზემოქმედებს არა მარტო სასუქების შეტანის რაოდენობა ნიადაგებში, არამედ ნიადაგების დამუშავების აგროტექნიკური ვადები, ნიადაგის ნაყოფიერება, წლის მანძილზე ნალექების მოსვლის რეჟიმი და სხვ. შრომის ნაყოფიერებაზე მოქმედებს არა მარტო მუშების კვალიფიკაციის დონე, არამედ მოწყობილობის წარმადობა, ნედლეულით, სათბობით, ელექტროენერგიით და სხვა საჭირო საბრუნავი სახსრებით ფირმის მომარაგება, ბიზნესის ორგანიზაცია, ხელმძღვანელობის უნარჩვევები, შრომისა და წარმოების ორგანიზაცია და სხვა ფაქტორები. ამიტომ რთული სოციალურ-ეკონომიკური მოვლენებისა და პროცესების ანალიზის დროს საჭიროა ამ მოვლენებისა და პროცესების განვითარებაზე მოქმედი მრავალი


230

ფაქტორის განხილვა, რომელიც წარმოშობს მრავალფაქტორულ ანუ მრავლობით კორელაციას.

**მრავალფაქტორული ანუ მრავლობითი კორელაცია** მოვლენებს შორის სტატიტიკური, სტოქასტური კავშირებია, ხოლო რეგრესია – კავშირის გამომსახველი განტოლებებია. ამიტომ ხშირად მრავლობით კორეალაციას მრავლობით რეგრესიასაც უწოდებენ სტატისტიკურ მეცნიერებაში.

მრავლობითი კორელაციურ-რეგრესული ანალიზის საწყის ეტაპზე აუცილებელია შეირჩეს საანალიზო მოვლენაზე მოქმედი უმნიშვნელოვანესი ფაქორები და მრავალ ფაქტორსა და



231

საშედეგო მოვლენას შორის ურთიერთკავშირის ადექვატურად ამსახველი შესაბამისი მათემატიკური ფუნქცია. პირევლივე ამოცანა ეკონომიკური ამოცანაა და მას ყველაზე კარგად რთულ ეკონომიკურ მოვლენებსა და პროცესებში გათვითცნობიერებული მაღალკვალიფიციური ეკონომისტი გადაწყვეტს, ხოლო მეორე – მეტად რთული მათემატიკური სახის ამოცანაა და მას ამჟამად კომპიუტერული ტექნიკის გამოყენებით წყვეტენ. ამ ეტაპზე საშედეგო მოვლენასა და მასზედ მოქმედ მრავალ ფაქტორს შორის ურთიერთკავშირის გამომსახველი მრავალი მოდელიდან ხდება ისეთის შერჩევა, რომელიც როგორც წინა მასალაში იყო ნაჩვენები, უზრუნველყოფს უმცირეს კვადრატთა მეთოდის მიხედვით $\Sigma(y-\hat{y})^2 = \min$ გამოსახულების მინიმიზაციის კრიტერიუმებით ამოცანის გადაწყვეტას. ასეთი თეორიული მოდელები კი მრავლობითი კორელაციის შემთხვევაში არის როგორც წრფივი ისე არაწრფივი სახის.

წრფივი მოდელია $\hat{y} = a_0 + a_1x_1 + a_2x_2 + ... + a_nx_n$ (8.43)

ხოლო არაწრფივია :

1. პარაბოლური: $\hat{y} = a_0 + a_1x_1^2 + a_2x_2^2 + ... + a_nx_n^2$ (8.44)

2. ჰიპერბოლური $\hat{y} = a_0 + \frac{a_1}{x_1} + \frac{a_2}{x_2} + ... + \frac{a_n}{x_n}$ (8.45)

3. ხარისხოვანი $\hat{y} = a_0x_1^{a_1}x_2^{a_2}...x_n^{a_n}$ (8.46)

4. მაჩვენებლიანი $\hat{y} = e^{a_0 + a_1x_1 + a_2x_2 + ... + a_nx_n}$ (8.47)

ამ მოდელებში $\hat{y}$ საშედეგო მოვლენის გამომსახველი სიმბოლოა, $x(x_1, x_2, ..., x_n)$– მიზეზობრივი მოვლენის ანუ საშედეგო მოვლენაზე მოქმედი ფაქტორის გამომსახველი სიმბოლოა.

$a_0$– თავისუფალი წევრი,

$a_1, a_2, ..., a_n$ – რეგრსიის კოეფიციენტებია ანუ საძიებელი

პარამეტრებია, რომელთა მნიშვნელობანი მეტყველებს თუ როგორ მოქმედებს თითოეული მათგანი საშედეგო მოვლენის განვითარებაზე.

$a_0, a_1, a_2, ..., a_n$ პარამეტრების გასაანგარიშებელ განტოლებათა სისტემა მიიღება მოდელის გაწარმოებით ჯერ $a_0$, შემდეგ $a_1, a_2, ..., a_n$ -ის მიმართ ცალცალკე და პირველი რიგის კერძო წარმოებულის ნულთან გატოლებით.

თუ წრფივი განტოლების (დანარჩები არაწრფივი განტოლებანი გალოგარითმების წესით ჯერ უნდა დავიყვანოთ წრფივ ფორმაზე და შემდეგ ვაწარმოოთ წრფივი განტოლებით მსგავსი მოქმედებანი) მიმართ ვაწარმოებთ მოქმედებებს, მაშინ უმცირეს კვადრატთა მეთოდით მოიძებნება

$$S = f(a_0, a_1, a_2, ..., a_n) = \Sigma(y - a_0 - a_1x_1 - a_2x_2 ... - a_nx_n) = \min$$

ფუნქციის მინიმალური მნიშვნელობანი. ამ მნიშვნელობებს ეს ფუნქცია, როგორც ვიცით ღებულობს მხოლოდ მაშინ, როცა მისი პირველი რიგის კერძო წარმოებულები გაუტოლდება ნულს. მაშასადამე, უნდა ვვიპოვოთ $S$ ფუნქციის კერძო წარმოებულები ცალცალკე $a_0, a_1, a_2, ..., a_n$ -ის მიმართ და გავუტოლოთ ნულს. ე.ი.

$$\frac{\partial S}{\partial a_0} = 0, \frac{\partial S}{\partial a_1} = 0, \frac{\partial S}{\partial a_2} = 0, ..., \frac{\partial S}{\partial a_n} = 0. \qquad (8.48)$$

მაგალითად, $a_2$ პარამეტრის მიმართ გვექნება:

$$\frac{\partial S}{\partial a_2} = \Sigma 2(y - a_0 - a_1x_1 - a_2x_2 ... - a_nx_n) \times (-x_2) = 0$$

მარტივი ალგებრული გარდაქმნის შემდეგ მივიღებთ:

$$a_0 \Sigma x_1 + a_2 \Sigma x_1 x_2 + a_2 \Sigma x_2{}^2 + ... + a_n \Sigma x_2 x_n = \Sigma y x_2$$

თუ ყველა პარამეტრის მიხედვით ასეთ მოქმედებებს ჩავატარებთ მივიღებთ განტოლებათა სისტემას:

$$\begin{cases} na_0 + a_1\Sigma x_1 + a_2\Sigma x_2 + \ldots + a_n\Sigma x_n = \Sigma y \\ a_0\Sigma x_1 + a_1\Sigma x_1^{\,2} + a_2\Sigma x_1 x_2 + .. + a_n\Sigma x_n x_1 = \Sigma x_1 y \\ a_0\Sigma x_2 + a_1\Sigma x_1 x_2 + a_2\Sigma x_2^{\,2} + \ldots + a_n\Sigma x_2 x_n = \Sigma y x_2 \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ a_n\Sigma x_n + a_1\Sigma x_1 x_n + a_2\Sigma x_2 x_n + \ldots + a_n\Sigma x_n^{\,2} = \Sigma y x_n \end{cases} \qquad (8.49)$$

განტოლებათა სისტემის ამოხსნა $a_0, a_1, a_2, \ldots, a_n$ პარამეტრების მიმართ ადვილად შეიძლება დეტერმინანტთა თეორიისა და კრამერის ფორმულების გამოყენებით.

მოვიყვანოთ კონკრეტული მაგალითი:

როგორც ცნობილია მაღალი ხარისხის ჩაის ხვედრითი წილი ($y$) ჩაის მზა პროდუქციის საერთო რაოდენობაში დამოკიდებულია პირველ რიგში, უმაღლესი ხარისხის ნედლეულის ხვედრით წილზე დამზადებული ჩაის მწვანე მასის საერთო რაოდენობაში და აგრეთვე ჩაის მწვანე ფოთლის მოკრეფიდან მის გადამუშავებამდე დაყოვნების დროზე, იმდენად, რამდენადაც ნედლეულში დროის დაყოვნებასთან დაკავშირებით იკარგება მშრალი ნივთიერება, რაც იწვევს მზა პროდუქციის ხარისხის გაუარესებას. ვთქვათ გვაქვს შემდეგი მონაცემები (ციფრები პირობითია):

უმაღლესი და I სორტის ჩაის მზა პროდუქციის დამოკიდებულება ნედლეულის ხარისხსა და გადამუშავების დაყოვნების დროზე.

ცხრილი №24

| ჩაის მწვანე ფოთლის დამზადების ზონები | უმაღლესი და I ხარისხის მზა პროდუქციის ხვედრითი წილი (%) $y$ | I სორტის ნედლეულის ხვედრითი წილი ნედლეულის საერთო მასაში (%) $x_1$ | ნედლეულის გადამუშავების დაყოვნების საშუალო დრო (საათებში) $x_2$ |
| --- | --- | --- | --- |


226

| | | | |
|---|---|---|---|
| I ზონა | 42.5 | 55.6 | 25 |
| II ზონა | 43.8 | 57.7 | 20 |
| III ზონა | 45.6 | 60.8 | 26 |
| IV ზონა | 44.8 | 61.5 | 17 |
| V ზონა | 46.7 | 62.8 | 17 |


227

მრავლობითი რეგრესიის ზოგადი განტოლება ამ შემთხვევაში იქნება:

$$\hat{y} = a_0 + a_1 x_1 + a_2 x_2 \quad (8.50)$$

შევადგინოთ რეგრესიის პარამეტრების გასაანგარიშებელი ცხრილი:

ცხრილი №25

| $y$ | $x_1$ | $x_2$ | $x_1^2$ | $x_2^2$ | $x_1 x_2$ | $x_1 y$ | $x_2 y$ | $\hat{y}$ |
|---|---|---|---|---|---|---|---|---|
| 42.5 | 55.6 | 25 | 3091.4 | 625 | 1390.0 | 2363.0 | 1062.5 | 42.8 |
| 43.8 | 57.7 | 20 | 3329.3 | 400 | 1154.0 | 2527.3 | 876.0 | 43.7 |
| 45.6 | 60.8 | 26 | 3696.6 | 676 | 1580.8 | 2772.5 | 1185.6 | 45.2 |
| 44.8 | 61.5 | 19 | 3782.3 | 361 | 1168.5 | 2755.2 | 851.2 | 45.4 |
| 46.7 | 62.8 | 17 | 3943.8 | 289 | 1067.5 | 2932.8 | 793.9 | 46.0 |
| 223.4 | 298.4 | 107 | 17843.4 | 2351 | 6360.9 | 13350.8 | 4769.2 | 223.11 |

ზოგადად ამ ამოცანის გადაწყვეტისათვის საჭირო ნორმალურ განტოლებათა სისტემა ასეთი სახისაა:

$$\begin{cases} na_0 + a_1\Sigma x_1 + a_2\Sigma x_2 = \Sigma y \\ a_0\Sigma x_1 + a_1\Sigma x_1{}^2 + a_2\Sigma x_1 x_2 = \Sigma x_1 y \\ a_0\Sigma x_2 + a_1\Sigma x_1 x_2 + a_2\Sigma x_2{}^2 = \Sigma x_2 y \end{cases}$$

ჩვენი მონაცემების საფუძველზე ეს განტოლებათა სისტემა მიიღებს სახეს:

$$\begin{cases} 5a_0 + 298.4a_1 + 107a_2 = 223.4 \\ 298.4a_0 + 17843.4a_1 + 6360.9a_2 = 13350.8 \\ 107a_0 + 6360.9a_1 + 2351a_2 = 4769.2 \end{cases}$$

გავყოთ თითოეული განტოლება $a_0$-ის კოეფიციენტებზე (პირველი განტოლება 5-ზე, მეორე 298.4-ზე, ხოლო მესამე – 107-ზე), მივიღებთ:

$$\begin{cases} a_0 + 59.68a_1 + 21.40a_2 = 44.68 \\ a_0 + 59.80a_1 + 21.32a_2 = 44.74 \\ a_0 + 59.45a_1 + 21.97a_2 = 44.57 \end{cases}$$


228

თუ მეორე და მესამე განტოლებებს ცალცალკე გამოვაკლებთ პირველ განტოლებას, გვექნება:

$$\begin{cases} 0.21a_1 + 0.08a_2 = 0.06 \\ -0.23a_1 + 0.57a_2 = 0.11 \end{cases}$$

თუ იმავე პროცედურას გავიმეორებთ მიღებულ განტოლებათა სისტემის მიმართ, მივიღებთ პარამეტრების მნიშვნელობებს:

$$a_0 = 17.59$$
$$a_1 = 0.45$$
$$a_2 = 0.011$$

მრავლობითი რეგრესიის განტოლება ასეთი სახის იქნება:

$$\hat{y} = 17.59 + 0.45x_1 + 0.011x_2 \quad (8.51)$$

ამ განტოლებით მოსწორებული დონეები, რომლებიც, მოთავსებულია ცხრილის ბოლო სვეტში, ჯამში იძლევა 223.11-ს, რაც მცირედითაა განსხავავებული ეპპირიული დონეების ჯამისაგან (223.4). განსხვავება გამოწვეულია ციფრების დამრგვალებით. ეს კი იმას ნიშნავს, რომ ჩვენს მიერ შერჩეულმა წრფივმა მრავაფაქტორულმა მოდელმა ადექვატურად ასახა ემპირიული მონაცემების, ჩვენს შემთხვევაში ჩაის ხარისხის ამაღლების დამოკიდებულება ნედლეულის ხარისხსა და გადამუშავების დაყოვნების დროზე.

მრავლობითი კორელაციურ-რეგრეციული ანალიზის დროს უნდა გავითვალისწინოთ შერჩეულ ფაქტორთა არაერთგვაროვნება, მათი სხვადასხვა ზომის ერთეულებში გამოსახვა და სხვა მოვლენები და პროცესები, რომლებიც ზოგჯერ ამახინჯებს მიზეზ-შედეგობრივი ურთიერთკავშირის სურათს. ამიტომ სტატისტიკაში შემოღებულია ფაქტორთა გადაყვანის პრაქტიკა ერთგვაროვან, ფართობით სიდიდეებში. ასეთ ქმედებას უწოდებენ ცვლადების **სტანდარტიზაციას** ანუ მათ წარმოდგენას **სტანდარტულ** მასშტაბებში.


229

ცვლადების $y, x_1, x_2, x_n$ სტანდარტულ მასშტაბებში გადაყვანა



230

წარმოებს ფორმულებით:

$$t_{iy}=\frac{y_i-\bar{y}}{\sigma_y},\quad t_{ix}=\frac{x_i-\bar{x}}{\sigma_x} \qquad (8.52),$$

სადაც $t_{iy}$ და $t_{ix}$ შესაბამისად $y$ და $x$ ნატურალურ მნიშვნელობათა სტანდარტული შეფარდებითი სიდიდეებია, $\bar{y}$ და $\bar{x}$ მონაცემთა საშუალო არითმეტიკულია, $\sigma_y$ და $\sigma_x$ - შესაბამისად $y$ და $x$ მონაცემთა საშუალოკვადრატული გადახრაა.

სტანდარტულ მასშტაბში გადაყვანილი ცვლადების საშუალო მნიშვნელობა ნულის ტოლია ($\bar{t}_{iy}=0$, $\bar{t}_{ix}=0$), ხოლო საშუალოკვადრატული გადახრა უდრის ერთს.

სტანდარტულ მასშტაბებში რეგრესის წრფივი განტოლება მიიღებს სახეს:

$$t_y=B_1t_1+B_2t_2+\ldots+B_nt_n \qquad (8.53)$$

($a_0$ განტოლებაში არა გვაქვს, ვინაიდან მისი მიღება შეიძლება განტოლებით: $a_0=\bar{y}-a_1\bar{x}_1-a_2\bar{x}_2-..-a_n\bar{x}_n$)

$B$ კოეფიციენტები განტოლებაში არის რეგრესიის სტანდარტიზებული კოეფიციენტები, რომლებიც გვიჩვენებენ ამა თუ იმ ფაქტორის საშუალოკვადრატული გადახრით ცვალებადობისას, ამ საშუალოკვადრატული გადახრის რა ნაწილით შეიცვლება საშედეგო მოვლენა სხვა ფაქტორების უცვლელობის პირობებში.

$B$ კოეფიციენტების გაანგარიშება შეიძლება უმცირეს კვადრატთა მეთოდით $\Sigma(t-\hat{t})^2=\min$ .

თუ $\hat{t}$ - ნაცვლად ამ გამოსახულებაში ჩავსვამთ მის შესაბამის მნიშვნელობას ($B_1t_1+B_2t_2+\ldots+B_nt_n$), გვექნება:



231

$$\sum\left[t-(B_1t_1+B_2t_2+..+B_nt_n)\right]^2=\min \qquad (8.54).$$


232

როგორც წინა მასალიდანაა ცნობილი ასეთი რთული ფუნქცია მინიმუმს აღწევს მხოლოდ ამ ფუნქციის პირველი რიგის კერძო წარმოებულის ნულთან გატოლების წერტილში.

მაშასადამე, საჭიროა მოვძებნოთ მოცემული ფუნქციის პირველი რიგის კერძო წარმოებული და თითოეული მათგანი ჯერ $B_1$-ის, შემდეგ $B_2$-ის და ა.შ. $B_n$-ის მიმართ გავუტოლოთ ნულს.

გვექნება:

$$\begin{aligned}
&2\sum\left[t-(B_1t_1+B_2t_2+..+B_nt_n)\right](-t)=0\\
&2\sum\left[t-(B_1t_1+B_2t_2+..+B_nt_n)\right](-t_2)=0\\
&. \quad . \quad . \quad . \quad . \quad .\\
&2\sum\left[t-(B_1t_1+B_2t_2+..+B_nt_n)\right](-t_n)=0
\end{aligned} \qquad (8.55)$$

მარტივი ალგებრული გარდაქმნების შედეგად მივიღებთ $B_1, B_2 ... B_n$ რეგრესიული განტოლების სტანდარტული მაჩვენებლების (კოეფიციენტების) გასაანგარიშებელ განტოლებათა სიტემას:

$$\begin{cases}
B_1\Sigma t_1{}^2+B_2\Sigma t_1t_2+...+B_n\Sigma t_1t_n=\Sigma tt_1\\
B_1\Sigma t_1t_2+B_2\Sigma t_2{}^2+...+B_n\Sigma t_2t_n=\Sigma tt_2\\
. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad .\\
B_1\Sigma t_1t_n+B_2\Sigma t_2t_n+...+B_n\Sigma t_n{}^2=\Sigma tt_n
\end{cases} \qquad (8.56)$$

აქაც ვფიქრობთ ყველაზე მოსახერხებელია სისტემის ამოხსნისათვის გამოვიყენოთ კრამერის ფორმულები.

$B$ სტანდარტული კოეფიციენტები შეიძლება გადავიყვნოთ ნატურალურ $a$ კოეფიციენტებში შემდეგი ფორმულის გამოყენებით:


233

$$a_i = \frac{\sigma_y}{\sigma_{xi}} \beta_i \qquad (i = 1,2,...,n) \qquad (8.57)$$

$a_0$ კი, როგორც ზემოთ დავინახეთ, გაიანგარიშება ფორმულით:

$$a_0 = y - a_1 x_1 - a_2 x_2 - .. - a_n x_n \qquad (8.58)$$

გაანგარიშებათა შედეგად შეგვიძლია ჩავწეროთ $x$ და $y$ ნიშნებს შორის ნატურალურ კოეფიციენტებში გამოსახული რეგრესიული განტოლება.

## 12. წყვილადი კორელაციის კავშირის სიმჭიდროვის ხარისხის მაჩვენებლები

საერთოდ კორელაციური კავშირის სიმჭიდროვის ხარისხის დადგენა გულისხმობს გაიზომოს თუ როგორ მოქმედებს საშედეგო მოვლენის განვითარებაზე ანუ ვარიაციის ცვლილებაზე მიზეზობრივი ფაქტორის (წყვილადი კორელაცია) ან ფაქტორების (მრავლობითი კორელაცია) ვარიაციული ცვალებადობანი. წყვილადი კორელაციური კავშირის


234

შესწავლისას გამოყოფენ მოვლენათა ორ ჯგუფს: პარამეტრული და **არაპარამეტრული** მოვლენები და პროცესები. **პარამეტრული** სოციალურ-ეკონომიკური მოვლენები და პროცესები ეწოდება რაოდენობრივი შინაარსის მქონე მოვლენებსა და პროცესებს, რომელთა განვითარების სურათი რაოდენობრივი პარამეტრებით გაიზომება, ხოლო არაპარამეტრული მოვლენები და პროცესები ისეთი მოვლენები და პროცესებია, რომლებიც გამოსახულია არა რაოდენობრივი, არამედ არაპარამეტრული მაჩვენებლებით. ასეთს მიეკუთვნება, მაგალითად, რანგების მიხედვით გამოსახული მოვლენების ურთიერთკავშირი.

**რანგი** ეწოდება შესასწავლი ნიშნის მზარდი და კლებადი ტენდეციით დალაგებულ მწკრივში რიგით ნომერს, რომელიც მოვლენის მნიშვნელობის ადგილს მიანიშნებს. **წყვილადი კორელაციური კავშირების პარამეტრული სიმჭიდროვის ხარისხის მაჩვენებლებს განეკუთვნება კორელაციის წრფივი კოეფიციენტი, დეტერმინაციის ემპირიული კოეფიციენტი, დეტერმინაციის თეორიული კოეფიციენტი, ემპირიული კორელაციური დამოკიდებულება, თეორიული კორელაციის დამოკიდებულება, კორელაციის კერძო კოეფიციენტი და სხვა, ხოლო არაპარამეტრულს განეკუთვნება კორელაციის რანგების კოეფიციენტი, ასოციაციისა და კონტიგენციის კოეფიციენტი და ა. შ.**

განვიხილოთ ისინი ცალცალკე.

## 13. კორელაციის წრფივი კოეფიციენტი

მისი ცვალებადობის ინტერვალები და სანდოობის კრიტერიუმები მენეჯმენტში წრფივი კავშირის შემთხვევაში კორელაციის სიმჭიდროვის ხარისხის ყველაზე გავრცელებული მაჩვენებელია კორელაციის წრფივი კოეფიციენტი, რომელიც


235

შემუშავებული იქნა მე-19 **საუკუნის 90-იან წლებში ინგლისელი სტატისტიკოსისა და ფილოსოფოსის კარლ**



236

**პირსონის (1857-1936), აგრეთვე ეჯვორტისა და ველდონის მიერ.** ზოგადად მისი გასაანგარიშებელი ფორმულები და გამოყენება ამ თემის წინა მასალაშია ნაჩვენები. აქ მხოლოდ ვაჩვენებთ მისი გასაანგარიშებელი ფორმულების მოდიფიკაციას, აგრეთვე მნიშვნელობათა ცვალებადობის ინტერვალებსა და სანდოობის კრიტერიუმებს. **კ. ფიშერის** ინტერპრეტაციით კორელაციის წრფივი კოეფიციენტი ($R_{xy}$) სხვა არაფერია თუ არა რეგრესიის სტანდარტიზებული კოეფიციენტი ($a_1$), რომელიც გამოსახულია არა აბსოლუტურ ერთეულებში, როგორც ეს რეგრესიის კოეფიციენტის შემთხვევაში გვაქვს, არამედ საშუალო კვადრატული გადახრის ნაწილებში. ამიტომ შეგვიძლია დავწეროთ:

$$R_{xy} = a_1 \frac{\sigma_x}{\sigma_y}$$

$a_1$-ის სიდიდე შეიძლება განვსაზღვროთ ნორმალურ განტოლებათა სისტემიდან:

$$\begin{cases} na_0 + a_1\Sigma x = \Sigma y \\ a_1\Sigma x + a_1\Sigma x^2 = \Sigma xy \end{cases}$$

თუ სისტემის ორივე განტოლებას ცალცალკე გავყოფთ $n$-ზე მივიღებთ:

$$\begin{cases} a_0 + a_1 \dfrac{\Sigma x}{n} = \dfrac{\Sigma y}{n} \\ a_0 \dfrac{\Sigma x}{n} + a_1 \dfrac{\Sigma x^2}{n} = \dfrac{\Sigma xy}{n} \end{cases}$$

ანუ $$\begin{cases} a_0 + a_1\bar{x} = \bar{y} \\ a_0\bar{x} + a_1\overline{x^2} = \overline{xy} \end{cases}$$


237

თუ პირველი განტოლებიდან $a_0$- ის მნიშვნელობას განვსაზღვრავთ ($a_0 = \bar{y} - a_1\bar{x}$) და ჩავსვამთ მეორე განტოლებაში მივიღებთ $a_1$- ის მნიშვნელობას:

$$\bar{x}(\bar{y} - a_1\bar{x}) + a_1\overline{x^2} = \overline{xy}$$

$$\bar{x}\,\bar{y} - a_1\bar{x}^2 + a_1\overline{x^2} = \overline{xy}$$

$$a_1\overline{x^2} - a_1\bar{x}^2 = \overline{xy} - \bar{x}\,\bar{y}$$

აქედან $a_1 = \dfrac{\overline{xy} - \bar{x}\,\bar{y}}{x^2 - (x)^2}$ და ვინაიდან დისპერსიების გაანგარიშების წესებიდან გვახსოვს, რომ $\overline{x^2} - (\bar{x})^2 = \sigma_x^2$, გვექნება: $a_1 = \dfrac{\overline{xy} - \bar{x}\,\bar{y}}{\sigma_x^2}$. თუ მიღებულ გამოსახულებას ჩავსვამთ კორელაციის კოეფიციენტის ($R_{xy} = a_1\dfrac{\sigma_x}{\sigma_y}$) ფორმულაში, მივიღებთ:

$$R_{xy} = \frac{\overline{xy} - \bar{x}\,\bar{y}}{\sigma_x^2} \times \frac{\sigma_x}{\sigma_y} = \frac{\overline{xy} - \bar{x}\,\bar{y}}{\sigma_x \times \sigma_y}$$

მივიღეთ ჩვენთვის უკვე ცნობილი კორელაციის წრფივი კოეფიციენტის გასაანგარიშებელი ერთი ფორმულა.

მეორეს მხრივ კორელაციის წრფივი კოეფიციენტის ფიშერისეული ინტერპრეტაცია მასში მდგომარეობს, რომ ის წარმოადგენს $x$ და $y$ ნიშნების მიხედვით ნორმირებული გადახრების საშუალო სიდიდეს. $x$ ნიშნის მიხედვით ნორმირებული გადახრა, როგორც უკვე ჩვენთვის ცნობილია, უდრის– $t_x = \dfrac{x - \bar{x}}{\sigma_x}$


238

, ხოლო $t = \frac{y - y}{\sigma_y}$ $y$ —

$y$ ნიშნის მიხედვით .



239

მაშასადამე შეგვიძლია დავწეროთ კორელაციის წრფივი კოეფიციენტის ფორმულა:

$$R_{xy} = \frac{\sum\left(\frac{x-\bar{x}}{\sigma_x}\right)\left(\frac{y-\bar{y}}{\sigma_y}\right)}{n} \qquad (8.61)$$

როგორც მუდმივი რიცხვები $\sigma_x$ და $\sigma_y$ შეგვიძლია გავიტანოთ ფრჩხილებს გარეთ, გვექნება:

$$R_{xy} = \frac{\frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sigma_x\sigma_y}}{n} = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{n\sigma_x\sigma_y} =$$

$$= \frac{\Sigma(x-\bar{x})(y-\bar{y})}{n\sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}\times\sqrt{\frac{\Sigma(y-\bar{y})^2}{n}}} = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{n\sqrt{\frac{\Sigma(x-\bar{x})^2\times\Sigma(y-\bar{y})^2}{n^2}}}$$

თუ $n^2$-დან ამოვიღებთ ფესვს და $\frac{1}{n}$ გავიტანთ რადიკალის გარეთ, გვექნება:

$$R_{xy} = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\underset{2}{\Sigma}(x-\bar{x})^2\times\Sigma(y-\bar{y})}} \qquad (8.62)$$

მივიღეთ კორელაციის წრფივი კოეფიციენტის გასაანგარიშებელი მეორე ფორმულა.

მაშასადამე თუ გვაქვს რეგრესიის კოეფიცინტი $a_1$, აგრეთვე $x$ და $y$ ნიშნების მიხედვით საშუალო კვადრატული გადახრები ($\sigma_x$ და $\sigma_y$), შეგვიძლია გავიანგარიშოთ კორელაციის წრფივი კოეფიციენტი ფორმულით: $R_{xy} = a_1\frac{\sigma_x}{\sigma_y}$,

 240

აქედან ცხადია რეგრესიის კოეფიციენტი თავის მხრივ ამ ფორმულიდან უდრის: $a_1 = \dfrac{R_{xy}\sigma_y}{\sigma_x}$ და ა.შ.

დანარჩენ შემთხვევაში მოსახერხებელია კრელაციის წრფივი კოეფიციენტი გავიანგარიშოთ მოტანილი რომელიმე ფორმულით. კორელაციის წრფივი კოეფიციენტი იცვლება 1-დან – 1-მდე ან პირიქით. თუ ის უდრის ნულს, მოვლენებს შორის კავშირი არ არსებობს, თუ მეტია ნულზე და ნაკლებია 1-ზე ($o < R_{xy} < 1$) პირდაპირი კავშირია მოვლენებს შორის, თუ მეტია –1-ზე და ნაკლებია ნულზე ($-1 < R_{xy} < 0$) , მაშინ შებრუნებული კავშირი გვაქვს, თუ ერთს უდრის ($R_{xy} = 1$), მაშინ საქმე გვაქვს არა კორელაციურ, არამედ ფუნქციონალურ კავშირთან.

ჩვენ ზემოთ ვაჩვენეთ, რომ კორელაციის კოეფიციენტი ($R_{xy}$) იცვლება 1 დან +1 მდე. მაგრამ თუ ნულის ტოლია კავშირი აღნიშნულს მოვლენებს შორის არ არსებობს, ხოლო თუ 1-ის ტოლია, არსებობს არა კორელაციური, არამედ ფუნქციონალური ანუ სრული კავშირი. მაგრამ –1 დან +1 ინტერვალში რა სიდიდის კორელაციის კოეფიციენტი ჩაითვლება არსებითად ანუ მნიშვნელოვან სიდიდედ, რის გამო უნდა განხორციელდეს მენეჯმენტური ღონისძიებანი ასეთი ფაქტორის მართვის მიმართულებით? სტატისტიკაში შემოღებულია მოვლენებს შორის კავშირის რეგრესიისა და კორელაციის კოეფიციენტების კრიტერიუმები, რომელთა დახმარებით დადგინდება მათი არსებითობა ან არარსებითობა.

სტატისტიკოსები[1] კორელაციის კოეფიციენტის არსებითობის საკითხის გარკვევის ამოსავალ პუნქტად კორელაციის

 241

[1]იხ. Òåîðèÿ ñòàòèñòèêè. Ó÷åáíèê ïîä ðåäàêöèåé ïðîô Ð. Ë. Ãðîìûêî Ì.: ÈÍÔÐÀ-.,Ì 2002, ñòð. 211, 212. Òåîðèÿ ñòàòèñòèêè ïîä ðåäàêöèåé Ð. À.Øìîéëîâîé Ì. Ôèíàíñû è ñòàòèñòèêà, 2002, ñòð. 304-302 და სხვ



242

კოეფიციენტის მის საშუალოკვადრატულ გადახრასთან შედარებას მიიჩნევენ. ამის საფუძველზე ამ ორი მაჩვენებლის შეფარდებით გაინგარიშებენ კოეფიციენტს $\left(\frac{|R_{xy}|}{\sigma_R}\right)$ ანუ $R_{xy}$-ს აბსოლუტური მნიშვნელობის მისივე საშუალოკვადრატულ გადახრასთან შეფარდებით. არსებობს კორელაციის კოეფიციენტის საშუალო კვადრატულ გადახრასთან შეფარდებით გაანგარიშების წესები დაკვირვების მცირე რიცხვის ($n < 50$) პირობებისათვის ანუ მცირე შერჩევისათვის და დაკვირვების დიდი რიცხვისათვის ($n > 50$).

დაკვირვების მცირე რიცხვისათვის: $$\sigma_R = \frac{\sqrt{1-R^2}}{\sqrt{n}} \qquad (8.63)$$

ხოლო დაკვირვების დიდი რიცხვისათვის: $$\sigma_R = \frac{1-R^2}{\sqrt{n}} \qquad (8.64)$$

ამიტომ შედარებითი კოეფიცეინტი $\left(t = \frac{|R|}{\sigma_R}\right)$, რომელსაც სტატისტიკოსები $t$-სტატისტიკას უწოდებენ, მიიღებს სახეს:

მცირე შერჩევისათვის:

$$t = \frac{|R|}{\frac{\sqrt{1-R^2}}{\sqrt{n}}} = \frac{|R|}{\sqrt{1-R^2}} \times \sqrt{n} \qquad (8.64)$$

ხოლო დაკვირვების დიდი რიცხვის პირობებისათვის:

$$t = \frac{|R|}{\frac{1-R^2}{\sqrt{n}}} = \frac{|R|}{1-R^2} \times \sqrt{n} \qquad (8.65)$$



243

სადაც $n$ – დაკვირვების რიცხვია.

ბოლო ფორმულიდან შეიძლება განვსაზღვროთ[1]

$|R| = t\frac{1-R^2}{\sqrt{n}} = t\sigma_R$. გავიხსენოთ. რომ ნორმალური განაწილების დროს კონკრეტული ვარიაციული მაჩვენებელი თავსდება $\pm 3\sigma$-ს ფარგლებში. აქედან შეიძლება დავასკვნათ, რომ თუ დაკვირვების დიდი რიცხვის ($n$) პირობებში კორელაციის კოეფიციენტი ($R$) 3-ჯერ მეტია თავის საშუალოკვადრატულ გადახრაზე ($R > \pm 3\sigma$ ანუ $\frac{|R|}{\sigma} > 3$), მაშინ ის ითვლება მნიშვნელოვნად (არსებითად) და კავშირი რეალურად. ნდომის ინტერვალი ამ შემთხვევაში არის $\pm 3\sigma$. ცხადია თუ ნდომის ინტერვალზე დაბალია კორელაციის კოეფიციენტი, მაშინ ის არარსებითად ითვლება და კავშირი-მეტისმეტად სუსტად, რომელიც შეიძლება არ გავითვალისწინოთ მენეჯმენტურ მართვაში.

მცირე შერჩევისათვის გაიანგარიშება $t$-სტატისტიკა (8.65) ფორმულის მიხედვით და შეუდარდება სტიუდენტის კრიტერიუმს. რადგან სტიუდენტის $t$ კრიტერიუმი (იხილეთ დანართი 9) გაანგარიშებულია გარკვეული ალბათობისა ($\alpha$-ს მიიჩნევენ 0.05-ის ანუ 5%-ის ფარგლებში) და $n-2$ თავისუფლების ხარისხისათვის, ამიტომ $t$ ფაქტიური ანუ საანგარიშო მიიღებს სახეს:

---

[1]მოტანილ ფორმულებში დაკვირვების რიცხვი ყოველი კონკრეტული ამოცანის გადაწყვეტისას შეიძლება შეიცვალოს თავისუფლების ხარისხით. სტატისტიკაში თავისუფლების ხარისხს უწოდებენ დაკვირვების რიცხვსა და არათავისუფალ პარამეტრებს შორის სხვაობით მიღებულ რიცხვს. მაგალითად, წრფივი კავშირის შემთხვევაში თავისუფლება არა აქვს ორ პარამეტრს ($a_0$, $a_1$), რომლებსაც ჩვენ ვანგარიშობთ და ვაფიქსირებთ გარკვეულ დონეზე. მასასადამე, ამ შემთხვევაში თავისუფლების ხარისხი იქნება n-2.

$$t_{\text{გაანგ.}} = \frac{|R|}{\sigma} = \frac{|R|\sqrt{n-2}}{\sqrt{1-R^2}} \qquad (8.67)$$

ამ ფორმულით გაანგარიშებული ნდობის ინტერვალი ($t_{\text{გაანგ.}}$) უნდა შევუდაროთ სტიუდენტის ცხრილურ მაჩვენებელს. ამასთან წინასწარ მიიღება ნულოვანი ჰიპოთეზა იმის შესახებ, რომ $R = 0$, ე. ი. ამ მოვლენებს შორის კავშირი არ არსებობს.

თუ გაანგარიშებული მაჩვენებელი მეტია ცხრილურ მაჩვენებელზე ($t_{\text{გაანგ.}} > t_{\text{ცხრ.}}$), მაშინ ნულოვანი ჰიპოთეზა უარიყოფა და კორელაციის კოეფიციენტი ითვლება არსებითად, ხოლო კავშირი მნიშვნელოვნად (სტიუდენტის ცხრილებში მოცემულია $t$-სტატისტიკური მაჩვენებლის კრიტიკული შეფასებანი, რომლებიც სამართლიანია მხოლოდ ნულოვანი ჰიპოთეზის პირობებისათვის. ცხადია გაანგარიშებითი მაჩვენებელი თუ ნაკლებია ცხრილურ კრიტიკულ ზღვარზე, მაშინ მიიღება ნულოვანი ჰიპოთეზა იმის შესახებ, რომ კავშირი ამ მოვლენებს შორის არ არსებობს და კორელაციის კოეფიციენტის ფაქტობრივი მნიშვნელობა არარსებითია).

ზემოთ მოტანილი ფორმულებიდან შეიძლება განისაზღვროს კორელაციის კოეფიციენტის მიმართულების საზღვრები გენერალურ ერთობლიობაში.

$$\tilde{R} = R \pm t\sigma \qquad (8.68)$$

სადაც $\tilde{R}$ - გენერალური ერთობლიობის კორელაციის კოეფიციენტია,

$R$ შერჩევითი ერთობლიობის კორელაციის კოეფიციენტი,

$t$ – სტიუდენტის კრიტერიუმი (0.95 ალბათობისათვის. ცხრილის მიხედვით ის უდრის 1.96, 0.71 ალბათობისათვის –1.06-ს, 0.99 ალბათობისათვის 3-ს და ა.შ.)

$\sigma$ – კორელაციის კოეფიციენტის საშუალო კვადრატული გადახრა.

ჩვენს მიერ ამავე თავში ძირითად კაპიტალსა და გამოშვებულ



245

პროდუქციას შორის კორელაციური ურთიერთკავშირის შემთხვევაში კორელაციის კოეფიციენტი შეადგენს 0.95-ს, $n = 8$; აქედან საშუალო კვადრატული გადახრაა:

$$\sigma_R = \frac{\sqrt{1-R^2}}{\sqrt{n-2}} = \frac{\sqrt{1-0.95^2}}{\sqrt{8-2}} = 0.13 \qquad t_{\text{გაანგ.}} = \frac{|R|}{\sigma_R} = \frac{0.95}{0.13} = 7.3$$

ცხრილური მაჩვენებლები (თავისუფლების ხარისხის $n-2=6$ და $\alpha = 0.05$ მნიშვნელობისათვის იხ. მე-9 დანართი) შეადგენს 2.4469-ს $t_{\text{ცხრ.}} = 2.4469$.

მაშასადამე, რადგან $t$-სტატისტიკის გაანგარიშებითი მაჩვენებელი (7.3) მეტია ცხრილურ მაჩვენებელზე (2.4469), ამიტომ შეიძლება დავასკვნათ, რომ გენერალურ ერთობლიობაში $x$-სა და $y$-ს შორის ნულოვანი ჰიპოთეზა უარიყოფა. აქედან ცხადია, რომ ამ მოვლენებს შორის კავშირის კორელაციის კოეფიციენტი არსებითია, მნიშვნელოვნად განსხვავდება ნულისაგან და ამიტომ მოცემულ მოვლენებს შორის კორელაციური კავშირი არსებითია.

## 14. კორელაციის ფეხნერის კოეფიციენტი

კორელაციის ფეხნერის კოეფიციენტი სტატისტიკაში ცნობილია **გერმანელი ფსიქოლოგის (1801-1887) გ. ტ. ფეხნერის ავტორობით.** სწორედ მან შემოგვთავაზა მოვლენებს შორის კავშირის სიმჭიდროვის ხარისხი გავზომოთ არა თვით ნიშნის რაოდენობრივი მახასიათებლებით, არამედ თითოეული ნიშნის რაოდენობრივი ინდივიდუალური მნიშვნელობის მათი საშუალოსაგან გადახრების ნიშნების მიხედვით. ფეხნერის კორელაციის კოეფიციენტის გასაანგარიშებლად საჭიროა ჯერ გავიანგარიშოთ ცალცალკე $x$ მიზეზობრივ და $y$ საშედეგო მოვლენების ინდივიდუალურ



246

მნიშვნელობათა გადახრები ანუ სხვაობანი მათი საშუალო არითმეტიკულისაგან. ამის შემდეგ უნდა დავითვალოთ $\left(x - x\right)$

 247

და $\left(y-\bar{y}\right)$ გადახრების ნიშანთა თანამთხვევისა და არათანამთხვევის რაოდენობა. ნიშანთა თანდამთხვევათა და არათანდამთხვევათა რიცხვებს შორის სხვაობას თუ გავყოფთ **ორთავეს ჯამზე, მივიღებთ კორელაციის ფეხნერის** კოეფიციენტს. ზოგადად კოეფიციენტი გამოისახება ფორმულით:

$$K_{\text{ფეხნ.}}=\frac{a-b}{a+b} \qquad (8.69)$$

სადაც $K$ - ფეხნერის კორელაციის კოეფიციენტია;

$a-$ ნიშნების მნიშვნელობათა მათი საშუალოსაგან გადახრების ნიშანთა თანდამთვევის რიცხვი,

$b-$ არათანამთხვევის რიცხვი.

მოვიტანოთ კონკრეტული მაგალითი:

წყვილადი კორელაციის კოეფიციენტების
გასაანგარიშებელი საანაგარიშო

ცხრილი №30

| № რიგზე | $y$ | $x$ | $(x-x)$ | $(y-y)$ | $xy$ | $y^2$ | $x^2$ |
|---|---|---|---|---|---|---|---|
| 1 | 110 | 45 | + | - | 4950 | 12100 | 2025 |
| 2 | 500 | 35 | - | + | 17500 | 250000 | 1225 |
| 3 | 400 | 35 | - | + | 14000 | 160000 | 1225 |
| 4 | 300 | 46 | + | - | 13800 | 90000 | 2116 |
| 5 | 350 | 42 | + | - | 14700 | 122500 | 1764 |
| 6 | 450 | 40 | - | + | 18000 | 202500 | 1600 |
| ჯამი Σ | 2110 | 243 | - | - | 82950 | 937100 | 9955 |
| საშაულო | 351.7 | 40.5 | - | - | 13825 | 139516 | 1659 |

როგორც ცხრილიან ჩანს კორელაციის ფეხნერის კოეფიციენტი შეადგენს:

$$K_{\text{ფეხნ.}}=\frac{a-b}{a+b}=\frac{0-6}{0+6}=-1$$

ამავე ცხრილის მონაცემებით შეგვიძლია გავიანგარიშოთ კორელაციის წრფივი კოეფიციენტი ($R_{xy}$):

$$R_{xy}=\frac{\overline{xy}-\bar{x}\ \bar{y}}{\sigma_x\times\sigma_y}=\frac{13825-351.7\times 40.5}{\sigma_x\times\sigma_y}$$

გავიხსენოთ საშუალოკვადრატული გადახრის გაანგარიშების გამარტივებული წესი:

$$\sigma_x = \sqrt{\overline{x^2} - (\overline{x})^2}, \sigma_y = \sqrt{\overline{y^2} - (\overline{y})^2},$$

გვექნება:

$$\sigma_x = \sqrt{1659 - 40.5 \times 40.5} = 4.4$$

$$\sigma_y = \sqrt{139515 - 351.7 \times 351.7} = 125$$

$$R_{xy} = \frac{-418.85}{4.4 \times 125} = -\frac{418.85}{550} = -0.76$$

როგორც ჩანს, კორელაციის ფეხნერის კოეფიციენტი შეადგენს $-1$-ს, რაც ნიშნავს ამ მოვლენებს შორის არა კორელაციურ, არამედ ფუნქციონალურ დამოკიდებულებას, მაშინ როდესაც ჩვეულებრივი კორელაციის წრფივი კოეფიციენტი $B_{xy} = -0.76$, რაც მნიშვნელოვნად განსხვავებულია ფეხნერის კორელაციის კოეფიციენტისაგან. აქედან დასკვნა იმის შესახებ, რომ რაოდენობრივად კორელაციის ფეხნერის კოეფიციენტი ყოველთვის არ იძლევა ზუსტ შედეგს. ამიტომ იგი საიმედოა მხოლოდ კავშირის მიმართულების განსაზღვრისათვის. ამ შემთხვევაში მიმართულება შებრუნებულ კავშირს გვიჩვენებს, რაც დასტურდება კორელაციის წრფივი კოეფიციენტის მიხედვითაც.

## 16. ასოციაციის, კონტინგენციისა და ბისერიალური კორელაციის კოეფიციენტები.

მე-15 პარაგრაფში მოტანილი ‘ოთხ მინდვრიანი“ ცხრილის მონაცემების მიხედვით, სადაც სიხშირეები აღნიშნულია $a, b, c$

 258

და $d$ სიმბოლოებით, ასოციაციის[1] კოეფიციენტი ზოგადად ასე ჩაიწერება:

$$K_{\text{ასოც.}} = \frac{ad - bc}{ad + bc} \quad (8.71)$$

კონკრეტული მონაცემებით (იხ, ცხრილი 38) ის შეადგენს:

$$K_{\text{ასოც.}} = \frac{ad - bc}{ad + bc} = \frac{(60 \times 140) - (15 \times 40)}{(60 \times 140) + (15 \times 40)} = 0.87$$

მაშასადამე ასოციაციის კოეფიციენტი გვიჩვენებს, რომ კავშირის სიმჭიდროვის ხარისხი პირთა მუშაობისა და ამ პირთა უმაღლესის დამთავრების ნიშანს შორის მაღალია.

სტატისტიკოსები აღნიშნავენ ამ კოეფიციენტის უარყოფით მხარეს. კერძოდ თუ რომელიმე უჯრის სიხშირე არა გვაქვს, ანუ უდრის ნულს, მაშინ როგორც ფორმულიდან ჩანს კოეფიციენტი უდრის +1-ს ან –1, რომელიც წარმოადგენს ფუნქციონალურ ანუ სრულ კავშირს და ამახინჯებს კორელაციური კავშირის სურათს. ამიტომ გვთავაზობენ მეორე ფორმულას, რომელსაც **ჰქვია კონტინგენციის კოეფიციენტი. ის ზოგადად გაიანგარიშება ფორმულით:**

$$K_{\text{ასოც.}} = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}} \qquad (8.72)$$

კონკრეტული ცხრილის მონაცემებით:

$$K_{\text{ასოც.}} = \frac{(60 \times 140) - (15 \times 40)}{\sqrt{(60 + 40)(15 + 40)(60 + 15)(40 + 140)}} = 0.545$$

მივიღეთ კორელაციის კონტინგენციის კოეფიციენტი, რომელიც გაცილებით ნაკლებია ასოციაცის კოეფიციენტზე. სტატისტიკოსები მიიჩნევენ[2], რომ კონტიგენციის

---

[1]სიტყვა „ასოციაცია“ წარმომდგარია ლათინური სიტყვა associo-საგან, რაც ნიშნავს გაერთიანებას, კავშირს.

[2]იხ. Òåîðèÿ ñòàòèñòèêè. Ó÷åáíèê, ïîä ðåäàêöèåé ïðîô. Ð. Ë. Ãðîìûêî, Ì.: ÈÍÔÐÀ-Ì, 2002 ñòð. 203

კოეფიციენტი ყოველთვის ნაკლებია აცოციაციის კოეფიციენტზე და ამიტომ თუ კონტიგენციის კოეფიციენტი $\left(|K|_{\text{კონტ.}} > 0.3\right)$ აბსოლუტური სიდიდით მეტია 0.3-ზე ხოლო ასოციაციის კოეფიციენტი $\left(|K|_{\text{ასოც.}} > 0.5\right)$, კავშირი აღნიშნული მოვლენებს შორის მნიშვნელოვანი და არსებითია. ასოციაციისა და კონტინგენციის კოეფიცენტები გამოიყენება ეკონომიკის, ბიზნესისა და მენეჯმენტის მოვლენებისა და პროცესებში ალტერნატიულ ნიშანთა ორი ჯგუფის წარმოქმნის პირობებში. მაგრამ თუ ჯგუფების რაოდენობა ორზე მეტია, მაშინ სტატისტიკაში არსებობს ურთიერთშეუღლებულ მაჩვენებელთა **პირსონის და ჩუპროვის კორელაციის** კოეფიციენტები (მათი გამოყენება შეიძლება, აგრეთვე, ორი ჯგუფის წარმოქმნის შემთხვევაშიაც).

პირსონის კოეფიციენტი ($K_{\text{პირს.}}$) გამოიანგარიშება ფორმულით:

$$K_{\text{პირს.}} = \sqrt{\frac{\chi^2}{\chi^2 + n}} \quad (8.73)$$

რუსმა სტატისტიკოსმა ა. ა. ჩუპროვმა (1874-1926)

შეიმუშავა კავშირის განსხვავებული მაჩვენებელი, რომელიც გამოისახება ფორმულით:

$$K_{\text{ჩუპრ.}} = \sqrt{\frac{\chi^2}{n(K_1 - 1)(K_2 - 1)}} \quad (8.74)$$

სადაც $\chi^2$- პირსონის ხი-კვადრატია, რომელიც ცნობილია წინა მასალიდან.

$n$ - დაკვირვების ერთეულთა რიცხვი

$K_1$ და $K_2$ –შესაბამისად სტრიქონებისა და სვეტების რიცხვი ცხრილში (ჯამის გარდა).



260

ჩვენი ცხრილი №29-ის მონაცემებით დაკვირვების საერთო რიცხვი (უმაღლესდამთავრებულები და უმაღლე-დაუმთავრებლები) შეადგენს 255-ს. $\chi^2$ –წინა მასალაში გაანგარიშების მიხედვით შეადგენს 72,48-ს, $K_1 = 2$, $K_2 = 2$, აქედან გამომდინარე პირსონის კოეფიციენტი შეადგენს:

$$K_{\text{პირს.}} = \sqrt{\frac{\chi^2}{\chi^2 + n}} = \sqrt{\frac{72.48}{72.48 + 255}} = 0.47$$

ჩუპროვის კოეფიციენტი:

$$K_{\text{ჩუპრ.}} = \sqrt{\frac{\chi^2}{n(K_1 - 1)(K_2 - 1)}} = \sqrt{\frac{72.48}{72.48(2-1)(2-1)}} = 0.53$$

როგორც ჩანს ალტერნატიული ნიშნის ორი ჯგუფის შემთხვევისთვისაც შეიძლება გამოვიყენოთ პირსონისა და ჩუპროვის კოეფიციენტები, მაგრამ ამ შემთხვევაში შედეგები ნაკლებსაიმედოა, ვინაიდან ჩუპროვის კოეფიციენტი მეტია, ვიდრე პირსონის კოეფიციენტი. 2-ზე მეტი რაოდენობის ჯგუფების წარმოქმნისას კი პირიქითაა, ჩუპროვის კოეფიციენტი ყოველთვის ნაკლებია პირსონის კოეფიციენტზე. ამიტომ სტატისტიკოსები 3-ზე ნაკლები ჯგუფების შემთხვევაში არ იძლევიან პირსონისა და ჩუპროვის კოეფიციენტების გამოყენების რეკომენდაციას.

ზოგჯერ სოციალურ-ეკონომიკურ მოვლენებსა და პროცესებში საჭიროა დავადგინოთ კორელაციური კავშირის სიმჭიდროვის ხარისხი ალტერნატიული (თვისებრივ, ხარისხობრივ) მაჩვენებლისა და რაოდენობრივ მაჩვენებლებს შორის. მაგალითად, იგივე უმაღლესდამთავრებულები და უმაღლესდაუმთავრებლები განსხვავდებიან ერთმანეთისაგან არა მარტო სამუშაოზე დასაქმებით, არამედ შემოსავლების ოდენობითაც. მოვიტანოთ პირობითი მაგალითი:


261

ფირმის მუშაკთა შემოსავლის თვიური დონე მათი განათლებისაგან დამოკიდებულებით.

ცხრილი №33

| | საშუალო თვიური შემოსავლები (ლარებში) | | | | სულ თანამშრომლები (კაცი) |
|---|---|---|---|---|---|
| | 80 | 120 | 150 | 200 | |
| უმაღლეს დამთავრებულები | 10 | 15 | 25 | 30 | 80 |
| უმაღლეს დაუმთავრებლები | 20 | 20 | 15 | 10 | 65 |
| სულ | 30 | 30 | 40 | 40 | 145 |

როგორც ჩანს უმაღლესდამთავრებულთა რაოდენობა იზრდება ხელფასის გადიდებასთან ერთად (მაგალითად, 80 ლარი აქვს 10 კაცს, ხოლო 200 ლარი 30 კაცს), ხოლო უმაღლესდაუმთავრებელთა რიცხვი ხელფასის ზრდასთან ერთად კლებულობს. მაშასადამე, ვიზუალურად ჩანს, რომ მომუშავეთა რიცხოვნობის განაწილება ხელფასის დონის მიხედვით არაა შემთხვევითი, არამედ დამოკიდებულია მათი განათლების დონეზე.

ასეთ შემთხვევაში ანუ ისეთ შემთხვევებში, როდესაც ეკონომიკაშიბიზნესსა და მენეჯმენტში მიზეზობრივი მოვლენა (ფაქტორი) ხარისხობრივია (თვისებრივი), ხოლო საშედეგო მოვლენა რაოდენობრივი (ან პირიქით), სტატისტიკაში კავშირის სიმჭიდროვის ხარისხის გასაზომავად გამოიყენება ბისერიალური კორელაციის კოეფიციენტი, რომლის ფორმულაა:

$$B = \frac{|\bar{y}_2 - \bar{y}_1|}{\sigma_y} \times \frac{pq}{z} \qquad (8.75)$$

სადაც $B$ - ბისერიალური კოეფიციენტია,

$\bar{y}_2 - \bar{y}$ – შესაბამისად I და II ჯგუფების საშუალო მაჩვენებელია,

$\sigma_y$ – საშუალო კვადრატული გადახრა.

$p$ და $q$ – შესაბამისად I და II ჯგუფების ხვედრითი წილი საერთო რაოდენობაში.

$z$ – ფიშერის ცხრილური – განაწილების მაჩვენებელი, რომელიც მოიძებნება ცხრილში (იხ. დანართი 10) შესაბამისი ალბათობის მიხედვით. ალტერნატიულ ნიშნებში ალბათობად შეიძლება ჩავთვალოთ თითოეული ნიშნის წილი საერთო მოცულობაში (ხელშემწყობ რიცხვთა შეფარდება ყველა შესაძლო რიცხვთან).

ჩვენს მიერ მოტანილი ცხრილის მონაცემებით:

$$y_1 = \frac{(80\times 20)+(120\times 15)+(150\times 25)+(200\times 30)}{10+15+25+30} = \frac{12350}{80} = 154;$$

$$y_2 = \frac{(80\times 20)+(120\times 20)+(150\times 15)+(200\times 10)}{20+20+15+10} = \frac{8255}{65} = 125$$

$$y_{\text{საერთ.}} = \frac{(80\times 30)+(120\times 35)+(150\times 40)+(200\times 40)}{145} = \frac{20605}{145} = 145$$

$$\sigma_y = \sqrt{\overline{y^2} - (\bar{y})^2} = 43.3$$

$$\sigma_y = 43.3$$

$$p = 0.55$$

$$q = 0.45$$

$Z =$ (რადგან $p = 0.55$, რასაც ჩვენ მივიჩნევთ ალბათობად, ცხრილში – დანართი 10, ის შეადგენს 0.6184) $= 0.6184$

ამ მონაცემების საფუძველზე კორელაციის ბისერიალური კოეფიციენტი შეადგენს:

$$B_{\text{ბისერ.}} = \frac{|\bar{y}_2 - \bar{y}_1|}{\sigma_y} \times \frac{pq}{z} = \frac{127-154}{43.3} \times \frac{0.55\times 0.45}{0.6184} = 0.25$$

$$B_{\text{ბისერ.}} = 0.25$$

კოეფიციენტის მნიშვნელობა გვიჩვენებს, რომ ამ შემთხვევაში შემოსავლებსა და უმაღლესის დამთავრებას შორის კავშირი არსებობს, მაგრამ ძალიან სუსტი.


263

## 17. რანგების კორელაციის კოეფიციენტი

ეკონომიკაში, ბიზნესსა და მენეჯმენტში წყვილადი კორელაციური კავშირების სიმჭიდროვის ხარისხის გასაზომავად ხშირად იყენებენ რანგების კორელაციის კოეფიციენტებს, რომლებიც გამოირჩევიან გაანგარიშებათა სიმარტივით და მოხერხებულობით. ასეთია **ინგლისელი სტატისტიკოსების სპირმენისა და კენდელის** რანგების კორელაციის კოეფიციენტები. ორთავე მეცნიერის მიერ შემოთავაზებული კოეფიციენტის გაანგარიშების მეთოდოლოგია ეყრდნობა $x$ და $y$ ნიშნების მიხედვით მოცემული ვარიაციული მწკრივების რანჟირებას და შესაბამისი რანგების ვარიაციული მწკრივების წარმოქმნას.

სოციალურ-ეკონომიკური მოვლენების ვარიაციული მწკრივების რანჟირება ეწოდება ამ მწკრივების ინდივიდუალურ მნიშვნელობათა ადგილების (პირველი 1, მეორე 2 და ა. შ.) განსაზღვრას, ანუ მათდამი რანგების მიკუთვნებას, სიდიდის მიხედვით ზრდის ან კლების ტენდენციის გათვალისწინებით.

აქედან, ცხადია, **რომ რანგი მაჩვენებელთა რიგითი ნომერია ზრდის ან კლების ტენდენციით დაალაგებულ ვარიაციულ მწკრივში.** მოვიტანოთ პირობითი მაგალითი, რომლის მიხედვით ადვილი გასაგები გახდება ვარიაციული მწკრივების რანჟირების პროცესი.

ბიზნესში კომერციული ბანკების აქტივები და წლიური მოგება (მლნ ლარობით)

ცხრილი №34

| ბანკების ნომერი | აქტივები (მლნ. ლარი) $x$ | წლიური მოგება (მლნ. ლარი) $y$ | რანჟირებული ვარიაციული მწკრივები | | რანგებს შორის სხვაობა $(N_x - N_y)$ | $d^2$ |
|---|---|---|---|---|---|---|
| | | | $x$ ნიშნით $N_x$ | $y$ ნიშნით $N_y$ | | |
| 1 | 10.0 | 0.5 | 10 | 10 | 10-10=0 | 0 |
| 2 | 15.0 | 0.7 | 9 | 8 | 9-8=1 | 1 |
| 3 | 20.0 | 0.6 | 8 | 9 | 8-9=-1 | 1 |
| 4 | 25.0 | 0.8 | 7 | 7 | 7-7=0 | 0 |
| 5 | 30.0 | 1.2 | 6 | 5 | 6-5=1 | 1 |
| 6 | 40.0 | 1.3 | 5 | 4 | 5-4=1 | 1 |
| 7 | 50.0 | 1.0 | 4 | 6 | 4-6=-2 | 4 |
| 8 | 60.0 | 1.5 | 3 | 3 | 3-3=0 | 0 |
| 9 | 70.0 | 1.8 | 2 | 2 | 2-2=0 | 0 |
| 10 | 80.0 | 2.0 | 1 | 1 | 1-1=0 | 0 |
| $n=10$ | | | | | | $\Sigma d^2 = 8$ |

 264

ავიღოთ, მაგალითად, 10 მსხვილი კომერციული ბანკი, რომელთა მიხედვით გვაქვს აქტივები და წლიური მოგების მაჩვენებლები:

მოვლენებს შორის პირდაპირი კავშირების დროს რანჟირება იწყება უდიდესიდან უმცირესის მიმართულებით. უდიდესი მნიშვნელობის მქონე ვარიანტს ენიჭება უპირატესობა და ესმის რანგი 1, ანუ თავსდება პირველ ადგილზე, ხოლო თუ შებრუნებულ ანუ უკუკავშირების შემთხვევაში რომელიმე ერთერთი ნიშნის მნიშვნელობათა რანჟირება წარმოებს, მაშინ–უმცირესიდან უდიდესამდე. მაგალითად, ბიზესში წარმოების მოცულობისა და პროდუქციის ერთეულის თვითღირებულების ურთიერთკავშირის შესწავლისას წარმოების მოცულობის ვარიაციული მწკრივის რანჟირება მოხდება უდიდესიდან უმცირესამდე, ხოლო პროდუქციის ერთეულის თვითღირებულების რანჟირება–უმცირესიდან უდიდესამდე (ბუნებრივია, რომ პირველ ადგილზე მოხვდება უმცირესი დანახარჯების მქონე წარმოებანი).

(ცხრილში მიზეზობრივი ფაქტორი აღნიშნულია $x$-ით, საშედეგო ფაქტორი $y$-ით, ნიშნის მიხედვით რანგი $N_x$-ით, $y$ ნიშნით მიხედვით რანგი $N_y$-ით). ცხრილში რანგები წარმოიქმნება მაჩვენებელთა რაიმე ნიშნის უპირატესობის მიხედვით. როგორც საწესდებო კაპიტალის ასევე წლიური მოგების მიხედვით უპირატესობა ენიჭება მათ სიდიდეს. ცხადია პირველ ადგილზე უნდა დავსვათ ის ბანკები, რომლებსაც ყველაზე დიდი საწესდებო კაპიტალი და წლიური მოგება აქვს, შემდეგ ადგილებზე განაწილდება ამ მაჩვენებელთა უფრო ნაკლები სიდიდის ბანკები და ა. შ.

იმ შემთხვევაში თუ რამდენიმე ბანკს აქვს სიდიდით ერთი და იგივე მაჩევენებელი, მაშინ თითოეულ მათგანს რანგი მიეკუთვნება ვარიაციულ მწკრივში მათი რიგითი ადგილების საშუალო არითმეტიკულის მიხედვით.

 265

მაგალითად, თუ მე-4 რანგის შემდეგ ორი მაჩვენებელია ერთნაირი სიდიდის, მაშინ თითოეულ მათგანს მიეკუთვნება 4.5 რანგი $\left(\frac{4+5}{2}=4.5\right)$, ხოლო თუ სამია ასეთი, მაშინ თითოეულ მათგანს მიენიჭება მე-5 რანგი $\left(\frac{4+5+6}{3}=5\right)$ და ა.შ.

რანჟირებული ვარიაციული მწკრივების რანგების საფუძველზე **სპირმენმა** შეიმუშავა კორელაციის რანგების კოეფიციენტის გასაანგარიშებლად შემდეგი ფორმულა:

$$R_{xy}=1-\frac{6\sum d_i^2}{n(n^2-1)} \qquad (8.76)$$

სადაც $R_{xy}$ - რანგების კორელაციის კოეფიციენტია,

$d$, $i$–ური რიგის $x$ და $y$ ნიშნების რანგებს შორის სხვაობაა $\left(N_x-N_y\right)$

$n$–დაკვირვების რიცხვი.

ჩვენი ცხრილის მონაცემებით $n=10$, $\Sigma d^2=8$. ამ საფუძველზე **სპირმენის** რანგების კორელაციის კოეფიციენტი შეადგენს:

$$R_{xy}=1-\frac{6\sum d_i^2}{n(n^2-1)}=1-\frac{48}{990}=0.95$$

როგორც ჩანს კორელაციის რანგების კოეფიციენტმა აჩვენა მოცემულ მოვლენებს შორის პირდაპირი და ამავე დროს მაღალი სიმჭიდროვის ხარისხის კავშირი. სტატისტიკოსები


266

იტყობინებიან, რომ კორელაციის რანგების კოეფიციენტის მაჩვენებელი ნაკლებ საიმედოა, ვიდრე წრფივი კორელაციის

 267

კოეფიციენტის მაჩვენებელი. მაგრამ გაანგარიშებათა სიმარტივის გამო მას შედარებით ფართო გამოყენება აქვს სოციალურ-ეკონომიკური მოვლენებისა და პროცესების ანალიზის დროს[1].

ინგლისელმა სტატისტიკოსმა გ. ჯ. კენდელმა რანგების კორელაციის კოეფიციენტის გაანგარიშება შემოგვთავაზა $x$ და $y$ ნიშნების რანგების ზრდისა და კლების შესადარისობის საფუძველზე. საქმე ისაა, რომ $x$ ნიშნის რანგები აიგება მკაცრად განსაზღვრულ ზრდის ტენდენციის მიხედვით. მაგრამ მის შესაბამისად $y$ ნიშნის რანგების ($N_y$) ზრდა ყოველთვის არ შეესაბამება $x$-ის ზრდას ე. ი. არაა დაცული მათ შორის ზრდის თანმიმდევრობა. მაგალითად, ჩვენს მიერ ზემოთმოტანილ ცხრილში რანგები ხასიათდება მუდმივი ზრდის ტენდენციით, ხოლო ზემოდან მე-2 და მე-6 რიგის რანგები ნაკლებია შესაბამისად მათ წინამდებარე რანგებზე. ამიტომ განსაზღვრავენ იმ რანგების რაოდენობას, რომლებიც თავისი მნიშვნელობით მეტია მის წინამდებარე რანგების მნიშვნელობაზე, ე.ი. სწორი „თანმიმდევრობის“ რაოდენობას.

---

[1]შევნიშნავთ, რომ კორელაციის რანგების კოეფიცენტი სხვა არაფერია თუ არა წრფივი კორელაციის კოეფიციენტი მხოლოდ იმ განსხვავებით, რომ ამ უკანასკნელის გასაანგარიშებელ ფორმულაში x და y–ის ნაცვლად რანგები ($N_x$ და $N_y$) გამოიყენება. ნატურალური რიცხვებისათვის საშუალო სიდიდე განისაზღვრება ფორმულით $\frac{n^2+n}{2n}=\frac{n+1}{2}$, ხოლო საშუალოდან ნატურალური რიცხვების გადახრის კვადრატების საშუალო $\frac{n^3-n}{12}$ ფორმულით. ამ მაჩვენებლების ფიშერის კორელაციის წრფივი კოეფიციენტის ფორმულაში შეტანითა და მარტივი გარდაქმნებით ვღებულობთ სპირმენის კორელაცის რანგების კოეფიციენტს. (დამტკიცება იხილეთ: Äæ, Ýäæ, Þë, Ì. Äæ. Êåíäåë, Òåîðèÿ Ñòàòèñòèêè, Ãîñòàòèçäàò ÖÑÓ, Ì., 1960, ñ. 304).

 268

მათ რაოდენობას „+“ ნიშნით გამოსახავენ და აღინიშნება $p$ სიმბოლოთი. პარალელურად დაითვლიან ამ რანგების რაოდენობას, რომლებიც თავისი მნიშვნელობით ნაკლებია მის წინამდებარე რანგების მნიშვნელობაზე. ასეთ შემთხვევებს გამოსახავენ „–“ ნიშნით და აღნიშნავენ $Q$ სიმბოლოთი. ორთავეს ჯამს გამოსახავენ $S$ სიმბოლოთი ($S = p + Q$).

ადვილი მისახვედრია, რომ $p$ მაქსიმალურ მნიშვნელობას მიაღწევს მაშინ, როდესაც $N_x$ და $N_y$ რანგები სრულ შესაბამისობაშია, ე.ი. როგორც $N_x$, ისე $N_y$ იზრდებიან ნატურალური მთელი რიცხვების შესაბამისად, 1-დან $n$-მდე. მაშინ პირველი წყვილ ($N_x = 1$ და $N_y = 1$) რანგების შემდეგ რანგების მნიშვნელობათა წინა რანგების მნიშვნელობებზე გადაჭარბების რიცხვი იქნება $n-1$, რანგების მეორე წყვილის ($N_x = 2$ და $N_y = 2$) შემდეგ $n-2$ და ა.შ. მაშასადამე სულ გვექნება:

$$P_{\max} = (n-1) + (n-2) + \ldots + 3 + 2 + 1 = \frac{n(n-1)}{2}\ .$$

იმ შემთხვევაში, როცა $y$-ის რანგებს აქვს $x$-ის რანგების ტენდენციის საწინააღმდეგო ტენდენცია, უნდა ავიღოთ აბსოლუტური მნიშვნელობით:

$$\left|P_{\max}\right| = \frac{n(n-1)}{2}\ .$$

**მ. ჯ. კენდალმა** კორელაციის რანგების კოეფიციენტი შეიმუშავა $S = P + Q$–ს $P$ ან $Q$-ს მაქსიმალურ შესაძლებელ რიცხვთან შეფარდებით:



269

$$R_{xy} = \frac{S}{\frac{n(n-1)}{2}} = \frac{2S}{n(n-1)} \qquad (8.77).$$

ჩვენს მიერ ზემოთ მოტანილი ცხრილის მონაცემებით ვაჩვენოთ კენდელის კორელაციის რანგების კოეფიციენტის გაანგარიშების წესი.

ცხრილი №35

| რ ა ნ გ ე ბ ი | | ბალების გაანგარიშება | |
|---|---|---|---|
| $N_x$ | $N_y$ | „+“ | „-“ |
| 1. | 1 | 9 | 0 |
| 2. | 2 | 8 | 0 |
| 3. | 3 | 7 | 0 |
| 4. | 6 | 4 | 1 |
| 5. | 4 | 4 | 2 |
| 6. | 5 | 5 | 0 |
| 7. | 7 | 3 | 0 |
| 8. | 9 | 1 | 1 |
| 9. | 8 | 1 | 0 |
| 10. | 10 | - | - |
| $n=10$ | | $P=42$ | $Q=-4$ |

ბალების გაანგარიშება „+“ და „–“ მინუს ნიშნების მიხედვით შემდეგნაირად ხდება: პირველ სტრიქონში ($N_x = 1, N_y = 1$) აღმოჩნდა, რომ 1-ზე გადაჭარბების 9 შემთხვევაა და არცერთი ერთზე ნაკლები შემთხვევა სვეტის რანგებიდან. ამიტომ „+“ ნიშანია 9 ანუ 9 ბალი, ხოლო მინუსი ნიშანი 0. ასეთივე მდგომარეობაა მეორე და მესამე სტრიქონებში. მეოთხე სტრიქონის შემდეგ 6-ზე მეტია 4 რანგი, ხოლო ნაკლები 1. ამიტომ მეოთხე

სტრიქონში გვაქვს „+“ 4, ხოლო მინუსი 1 და ა.შ.

მაშასადამე სულ გვაქვს $P = 42$, $Q = -4$ აქედან $S = P + (-4) = 38$

კორელაციის კენდელის რანგების კოეფიციენტი უდრის:

$$R_{xy} = \frac{2S}{n(n-1)} = \frac{2 \times 38}{10(10-1)} = \frac{76}{90} = 0.84.$$

როგორც ჩანს კენდელის რანგების კოეფიციენტი აღმოჩნდა უფრო ნაკლები თავისი სიდიდით, ვიდრე სპირმენის კოეფიციენტია.

კენდელის აღნიშნული კოეფიციენტი გამოიყენება იმ პირობებისათვის, როცა $x$ და $y$ ნიშნების ვარიაციულ მწკრივებში ვარიანტები არ მეორდება და აქედან გამომდინარე არ ხდება რანგების გაერთიანება საშუალო მაჩვენებლებში (რანგებში). თუ ვარიანტების ერთი და იგივე მნიშვნელობანი მეორდება და მაშასადამე წარმოიქმნება ერთი და იგივე რანგები, მაშინ კორელაციის რანგების კოეფიციენტის გასაანგარიშებლად გამოიყენება **კენდელის მეორე ფორმულა:**

$$B_{xy} = \frac{S}{\sqrt{\left[\frac{n(n-1)}{2} - K_x\right]\left[\frac{n(n-1)}{2} - K_y\right]}} \qquad (8.78)$$

სადაც $K_x = K_y = \dfrac{\sum t(t-1)}{2}$ - ბალების რიცხვი, რომლითაც კორექტირდება თითოეულ მწკრივში განმეორებული $t$ რანგების ხარჯზე (უნდა გვახსოვდეს, რომ ერთნაირი ბალების თანმიმდევრობის შემთხვევები ნებიესმიერ მწკრივში ფასდება ნულოვანი ბალით (0), რის გამო ისინი კოეფიციენტის გაანგარიშებისას არც პლიუსი და არც მინუსი ნიშნებით მხედველობაში არ

მიიღება). აღნიშნული შემთხვევისათვის ვაჩვენოთ კენდელის რანგების კორელაციის კოეფიციენტის გაანგარიშების წესი პირობით მაგალითზე.

რანგების კორელაციის კენდელის კოეფიციენტის გაანგარიშების ცხრილი.

ცხრილი №36

| $x$ | $y$ | $N_x$ | $N_y$ | ბალების დათვლის შედეგები | |
|---|---|---|---|---|---|
| | | | | „+“ ნიშნით | „-“ ნიშნით |
| 26 | 62 | 1 | 2 | 8 | 1 |
| 30 | 60 | 2.5 | 1 | 7 | 0 |
| 30 | 64 | 2.5 | 3 | 7 | 0 |
| 32 | 66 | 4 | 4.5 | 5 | 0 |
| 36 | 66 | 6 | 4.5 | 3 | 0 |
| 36 | 68 | 6 | 6 | 3 | 0 |
| 36 | 70 | 6 | 7.5 | 2 | 0 |
| 38 | 70 | 8 | 7.5 | 2 | 0 |
| 40 | 76 | 9 | 9.5 | 0 | 0 |
| 44 | 76 | 10 | 9.5 | - | - |
| $n = 10$ | | | | $P = 37$ | $Q = -1$ |

მოქმედი წესის მიხედვით[1] თავიდანვე უნდა განვსაზღვროთ $x$ ნიშნისათვის ერთნაირ მნიშვნელობათა რანგები. მინიმალურ მნიშვნელობას ($x = 26$) მიენიჭება რანგი 1. მის მომდევნო ორ ერთნაირ მნიშვნელობებს, რომლებსაც უჭირავს მეორე და მესამე ადგილი 2.5 ($\frac{3+2}{2} = 2.5$), $x = 32$ მნიშვნელობას მიენიჭება რანგი 4. მის მომდევნო თითოეულ ერთნაირ მნიშვნელობებს ($x = 36$), რომლებსაც უჭირავს შესაბამისად მე-5, მე-6 და მე-7 ადგილები, ბალი 6 ($\frac{5+6+7}{3} = 6$). რადგანაც

[1] იხ. Òåîðèÿ ñòàòèñòèêè. Ó÷åáíèê ïîä ðåäàêöèåé ïðîô. Ð. Ë. Ãðîìûêî Ì.: ÈÍÔÐÀ–Ì, 2002, ñòð. 219.

ამის შემდგომ მწკრივის მაჩვენებლები არაა ერთნაირი, ამიტომ თითოეულის დანარჩენ ვარიანტების ბალები მიენიჭება შესაბამისად 8,9,10.

ამის ანალოგიურად წარმოებს $y$ ნიშნის მიხედვით რანგების დათვლის პროცესი მხოლოდ იმ განსხვავევით, რომ რანგების მეორე წყვილის ($N_x = 2.5$, $N_y = 1$) შემდეგ მესამე წყვილს მხედველობაში არ ვღებულობთ არც პლუსი და არც მინუსი ნიშნებით, ვინაიდან 2.5 რანგის მნიშვნელობა იმეორებს მეორე წყვილის მნიშვნელობას. ამავე მიზეზით, მაგალითად, მეხუთე წყვილის ($N_x = 6, N_y = 4.5$) განხილვისას, მხედველობაში არ ვღებულობთ მე-6 და მე-7 წყვილებს, რომელთათვისაც $N_x = 6$. მე-7 წყვილის განხილვისას ($N_x = 6, N_y = 7.5$) მხედველობაში არ ვღებულობთ მე-8 წყვილს, რომლისთვისაც $N_y = 7.5$ იმეორებს მე-7 წყვილის მნიშვნელობას და ა.შ.

დავითვალოთ $K_x$ და $K_y$. $N_x$ რანგებში განმეორებით ორი შემთხვევა გვაქვს: ერთ შემთხვევაში რანგი 2.5 მეორდება 2-ჯერ, მეორე შემთხვევაში რანგი 6 მეორდება 3-ჯერ. მაშასადამე

$$K_x = \frac{2(2-1)+3(3-1)}{2} = 4 .$$

$N_y$ -ის მწკრივში გვაქვს რანგების განმეორების სამი შემთხვევა. პირველ შემთხვევაში რანგი 2.5 მეორდება 2-ჯერ, რანგი 7.5 აგრეთვე ორჯერ და რანგი 2.5-ორჯერ. მაშასადამე $N_y = \frac{2(2-1)+2(2-1)+2(2-1)}{2} = 3$ ($t$ როგორც აღვნიშნეთ ერთზე მეტად განმეორებადი რანგის განმეორების სიხშირე).

$$S = P + Q = 37 + (-1) = 36$$


276

მაქსიმალური ბალების რაოდენობა

$$\frac{n(n-1)}{2}=\frac{10\times 9}{2}=45$$

კორელაციის რანგების კენდელის კოეფიციენტი შეადგენს:

$$B_{xy}=\frac{S}{\sqrt{\left[\frac{n(n-1)}{2}-K_x\right]\left[\frac{n(n-1)}{2}-K_y\right]}}=\frac{36}{41}=0.867$$

როგორც ჩანს კორელაციის რანგების კოეფიციენტი ძალიან მაღალია და აჩვენებს $x$ და $y$ მოვლენებს შორის მჭიდრო ურთიერთკავშირს.

ამიტომ სტატისტიკოსები  კორელაციის სპირმენისა და კენდელის რანგების კოეფიციენტებს ანიჭებს უპირატესობას იმის გამო, რომ მათი გამოთვლა მარტივია. მათი მეშვეობით შეიძლება კავშირების სიმჭიდროვის ხარისხი გაიზომოს არა მარტო რაოდენობრივ, არამედ თვისებრივ (ხარისხობრივ) მოვლენებსა და პროცესებში და ა.შ.

## 18. პირსონის თეორიული კორელაციური შეფარდება

სოციალურ-ეკონომიკურ მოვლენებსა და პროცესებში ურთიერთკავშირის სომჭიდროვის ხარისხის განხილული კორელაციის კოეფიციენტები უმთავრესად გამოიყენება წრფივი კავშირების შემთხვევებისათვის.

**პირსონის თეორიული კორელაციური შეფარდება** უნივერსალური კორელაციის კოეფიციენტია, რომლითაც შეიძლება გაიზომოს მოლენებს შორის არსებული კავშირების სიმჭიდროვის ხარისხი არა მარტო წრფივი, არამედ არაწრფივი კავშირების (პარაბოლური, ჰიპერბოლური, ხარისხოვანი და

 277

ა. შ.) შემთხვევისათვის.

წინა მასალაში, დაჯგუფებულ მონაცემებს შორის კორელაციური ურთიერთკავშირების შესწავლისას, დისპერსიების შეკრების კანონის ძალით ვაჩვენეთ ემპირიული კორელაციური შეფარდების გაანგარიშება დეტერმინაციის ემპირიული კოეფიციენტიდან ფესვის ამოღების გზით:

$$\left(\eta_{ემპ.} = \sqrt{\frac{\delta^2}{\sigma_y^2}}\right).$$

პირსონის თეორიული კორელაციური შეფარდება გაინგარიშება დეტერმინაციის თეორიული კოეფიციენტიდან კვადრატული ფესვის ამოღებით. თავისთავად დეტერმინაციის თეორიული კოეფიციენტი დგინდება ნებისმიერი ფუნქციის (წრფივი და არაწრფივი) საშედეგო მოვლენების დონეთა მოსწორებით, ანუ ისეთი თეორიული დონეების გაანგარიშებით, რომლებიც უმცირეს კვადრატთა მეთოდით საერთო ჯამში მინიმალური მნიშვნელობით იქნება განსხვავებული ემპირიული დონეებისაგან. კოეფიციენტის სიდიდე განისაზღვრება საშედეგო მოვლენის თეორიული დონეების დისპერსიის $\dfrac{\Sigma(\hat{y}-\hat{\bar{y}})^2}{n}$ შეფარდებით ემპირიული დონეების დისპერსიასთან $\left(\dfrac{\Sigma(y-\bar{y})^2}{n}\right)$. ამის შედეგად თეორიული დეტერმინაციის კოეფიციენტი შეადგენს:

$$K_{თეორ.} = \frac{\Sigma(\hat{y}-\bar{\hat{y}})^2}{n} : \frac{\Sigma(y-\bar{y})}{n} = \frac{\Sigma(\hat{y}-\bar{\hat{y}})^2}{\Sigma(y-\bar{y})^2},$$

სადაც $K_{თეორ.}$ თეორიული დეტერმინაციის კოეფიციენტია,

$n$ - დაკვირების რიცხვი,

$y$ - ემპირიული დონეები,


278

$\hat{y}$ - თეორიული ანუ მოსწორებული დონეები

$\bar{\hat{y}}$ -თეორიული დონეების საშუალო. თეორიული და ემპირიული დონეების საშუალო (რამდენადაც $\Sigma y = \Sigma \hat{y}$ ერთმანეთის ტოლია. ამიტომ იქ, სადაც ფორმულებში გვაქვს $\bar{\hat{y}}$ თავისუფლად ვიხმაროთ სიდიდით მისი შემცვლელი $\bar{y}$ ).

როგორია თეორიული ანუ მოსწორებული და ემპირიული დონეების დისპერსიების ეკონომიკური შინაარსი? რის ცვალებადობას ასახავენ ისინი? ემპირიული დონეების $\left(\sigma^2 = \dfrac{\Sigma(y-\bar{y})^2}{n}\right)$ ანუ საერთო დისპერსია ასახავს მოცე-მული ეკონომიკური მაჩვენებლის ვარიაციის (ცვალება-დობის) ხარისხს უამრავი სოციალურ-ეკონომიკური ფაქტორის ზეგავლენით. ასეთი ფაქტორები იწვევენ საშედეგო მოვლენის ზრდას, კლებას ან სტაბილიზაციას. ძნელია სრულყოფილად დავასახელოთ საზოგადოებრივ-ეკონომიკური მოვლენებისა და პროცესების განვითარებაზე მოქმედი მზეზობრივი ფაქტორების სია ეკონომიკაში, ბიზნესსა და მენეჯმენტში. ყოველ კონკრეტულ შემთხვევაში მათი ჩამონათვალი მხოლოდ თვალშისაცემ ფაქტორებს ეხება და არა ყველას, რომელთა ნაწილი არა აშკარა, არამედ ფარული გზითაც, ზოგჯერ პოლიტიკურ ზეგავლენასაც კი ახდენენ ამა თუ იმ მოვლენის განვითარებაზე.

რაც შეეხება თეორიული, მოსწორებული დონეების დისპერსიას $\left(\sigma^2 = \dfrac{\Sigma(\hat{y}-\bar{y})^2}{n}\right)$ ის იმდენად რამდენადაც თვით თეორიული, მოსწორებული დონეები $x$ (მიზეზობრივი)

ფაქტორის ზეგავლენით ჩამოყალიბდა, ასახავს საშედეგო მოვლენის $x$ ფაქტორის ზემოქმედებით გამოწვეულ


279

ცვალებადობის დისპერსიას. ამიტომ დეტერმინაციის თეორიული კოეფიციენტი $\left(\frac{\delta^2}{\sigma^2} = K_{\text{დეტ.}}\right)$ გვიჩვენებს საშედეგო მოვლის ($y$) საერთო ცვალებადობის (დისპერსიის) რა ნაწილია გამოწვეული მოდელში გათვალისწინებული ფაქტორის (ან ფაქტორების) ცვალებადობით.

ამიტომ დეტერმინაციის თეორიული კოეფიციენტის ეკონომიკური შინაარსი ერთნაირია როგორც წყვილადი, ისე მრავლობითი კორელაციური კავშირურთიერთობის განხილვის პირობებისათვის.

**პირსონის თეორიული შეფარდება** ანუ კვადრატული ფესვი დეტერმინაციის თეორიული კოეფიციენტიდან $\left(\eta_{\text{თეორ.}} = \sqrt{\frac{\sum(\hat{y} - \bar{\bar{y}})^2}{\sum(y - \bar{y})^2}}\right)$ წოდებულია კორელაციური **კავშირის სიმჭიდროვის ხარისხის უნივერსალურ** მაჩვენებლად, რომელიც შეიძლება თავისუფლად გამოვიყენოთ როგორც წყვილადი ისე, მრავლობითი, აგრეთვე წრფივი და არაწრფივი კორელაციური კავშირურთიერთობის კორელაციურ-რეგრესიულ ანალიზსა და პროგნოზირებაში.

პირსონის თეორიულ კორელაციურ შეფარდებას ზოგჯერ **უწოდებენ კორელაციის ინდექსს** და გამოიყენება როგორც წყვილადი ასევე მულტიკორელაციური ანუ მრავლობითი კორელაციის დროს. მისი მნიშვნელობა იცვლება 0-დან 1-მდე. სტატისტიკოსები ამბობენ, რომ თუ მისი მნიშვნელობა ნაკლებია 0.3-ზე $(\eta < 0.3)$ კორელაციური კავშირი სუსტია, თუ 0.3-0.6 ფარგლებში – საშუალო, ხოლო თუ მეტია 0.6-ზე $(\eta > 0.6)$- კავშირი ძლიერია. თუ კორელაციის ინდექსი 0-ის ტოლია, მაშინ მოდელში შეყვანილი ფაქტორი არავითარ


280

ზეგავლენას არ ახდენს საშედეგო მოვლენის განვითარებაზე, თუ კი ერთის ტოლია, მაშინ შეიძლება ითქვას, რომ საშედეგო მოვლენის ცვალებადობა (ვარიაცია) მთლიანად განისაზღვრება მიზეზობრივი ფაქტორით.

მოვიტანოთ კონკრეტული მაგალითი საბანკო ბიზნესიდან:

კორელაციური შეფარდების საანგარიშო ცხრილი

ცხრილი №37

| წლიური საბანკო პროცენტი | ბანკებში დეპოზიტებზე თანხები (მლნ. ლარი) | | $y-\bar{y}$ | $(y-\bar{y})^2$ | $\hat{y}_x-\bar{y}$ | $(\hat{y}_x-\bar{y})^2$ | $y-\bar{y}_x$ | $(y-\bar{y}_x)^2$ |
|---|---|---|---|---|---|---|---|---|
| | ფაქტობრივი $y$ | თეორიული, [illegible] $\hat{y}$ | | | | | | |
| 1 | 32 | 32.4 | -8 | 64 | -7.6 | 57.76 | -0.4 | 0.16 |
| 2 | 38 | 37.0 | -2 | 4 | -0.3 | 9.0 | 1 | 1 |
| 3 | 40 | 40.8 | 0 | 0 | 0.8 | 0.64 | -0.8 | 0.64 |
| 4 | 44 | 43.8 | 24 | 16 | +3.8 | 14.44 | 0.2 | 0.4 |
| 5 | 46 | 46.0 | 36 | 36 | +6 | 36 | 0 | 0 |
| $\Sigma x=15$ | $\Sigma y=200$ | $\Sigma\hat{y}_x=200$ | 0 | 120 | 0 | 117.84 | 0 | 1.84 |

მოცემულ შემთხვევაში ბანკებში დეპოზიტებზე არსებული თანხების საშუალო შეადგენს:

$$\bar{y}=\frac{\Sigma y}{n}=\frac{200}{5}=40 \text{ მლნ. ლარს.}$$

ასეთივე მაჩვენებლი იქნება მოსწორებული დონეების საშუალო:

$$\bar{\hat{y}}=\frac{\Sigma\hat{y}}{n}=\frac{200}{5}=40 \text{ მლნ. ლარი.}$$

დეტერმინაციის თეორიული კოეფიციენტი იქნება:

$$\eta^2=\frac{\delta^2}{\sigma^2}=\frac{\sum(\hat{y}-\bar{y})^2}{\sum(y-y)^2}=\frac{117.84}{120}=0.982$$

პირსონის თეორიული შეფარდება:

$$\eta=\sqrt{\frac{\Sigma(\hat{y}-\bar{y})^2}{\Sigma(y-y)^2}}=0.990$$

ეს იმას ნიშნავს, რომ დეპოზიტების ზრდის პროცესი

 281

ბანკებში 98%-ით გამოწვეულია საბანკო წლიური პროცენტის მატებით. ამ ორ მოვლენას შორის, ე.ი. საბანკო პროცენტის მატებასა და ბანკებში დეპოზიტების ზრდას შორის კორელაციის კოეფიციენტი შეადგენს 0.99-ს. ე. ი. კავშირის სიმჭიდრვის ხარისხი ძალიან მაღალია.

## 19. მრავლობითი კორელაციური კავშირის სიმჭიდროვის ხარისხის მაჩვენებლები

მრავლობითი კორელაციური კავშირების სიმჭიდროვის ხარისხის გასაზომავად სტატისტიკაში ცნობილია:

მრავლობითი კორელაციის საერთო კოეფიციენტი,

მრავლობითი კორელაციის დეტერმინაციის საერთო კოეფიციენტი,

დეტერმინაციის კერძო კოეფიციენტები,

კორელაციის კერძო კოეფიციენტები,

ელასტიურობის კერძო კოეფიციენტები,

კონკორდაციის კორელაციის კოეფიციენტი.

წინა პარაგრაფში ვაჩვენეთ დეტერმინაციის ემპირიული და თეორიული კოეფიციენტების გაანგარიშების წესები. ისიც აღვნიშნეთ, რომ დეტერმინაციის კოეფიციენტებიდან ამოღებული კვადრატული ფესვი გვაძლევს შესაბამისი სახის (ემპირიული და თეორიული) კორელაციის კოეფიციენტებს და ის წოდებულია მოვლენებს შორის კავშირის სიმჭიდროვის უნივერსალური მაჩვენებლის სახელწოდებით. ამიტომ ეს მაჩვენებელი თანაბარი ძალით გამოიყენება როგორც წყვილადი, ისე მრავლობითი კორელაციის ანალიზის მიზნებისათვის.

დეტერმინაციის საერთო კოეფიციენტი გაიანგარიშება ფორმულით:

$$R_{y/x_1,x_2,...,x_n} = \frac{\delta^2}{\sigma^2} \quad (8.79)$$



282

სადაც $\delta^2$- მრავლობითი კორელაციის რეგრესიული განტოლების საფუძველზე გაანგარიშებული თეორიული დონეების დისპერსია,

$\sigma^2$–საერთო დისპერსიაა, რომელიც გაიანგარიშება ემპირიული (ფაქტობრივი) დონეების საფუძვლზე.

$R_{y/x_1,x_2,..,x_n}$–დეტერმინაციის კოეფიციენტია, რომელიც ასახავს საშედეგო $y$ მოვლენის დამოკიდებულებას $x_1,x_2,..,x_n$ ფაქტორებისაგან.

აქედან ამოღებული კვადრატული ფესვი გვიჩვენებს მრავლობითი კორელაციის კოეფიციენტის მნიშვნელობას:

$$R_{y/x_1,x_2,..,x_n} = \sqrt{\frac{\delta^2}{\sigma^2}} \qquad (8.80).$$

წრფივი კორელაციური კავშირის შემთხვევაში თუ გამოვიყენებთ წყილად კორელაციურ კოეფიციენტებს, დეტერმინაციის საერთო კოეფიციენტი შეგვიძლია შემდეგნაირად ჩავწეროთ:

$$R^2_{y/x_1,x_2,..,x_n} = \frac{a_1 z_{yx_1}\sigma_{x_1} + a_2 z_{yx_2}\sigma_{x_2} + .. + a_n z_{yx_n}\sigma_{x_n}}{\sigma_y} \qquad (8.81)$$

სადაც $a_1,a_2,..,a_n$–მრავლობითი რეგრესიული განტოლების პარამეტრებია ნატურალურ გამოსახულებაში.

$z_{yx_1}, z_{yx_2},.., z_{yx_n}$–წყვილადი კორელაციის კოეფიციენტებია, რომლებიც გვიჩვენებენ $y$ საშედეგო მოვლენის კორელაციურ ურთიერთკავშირს შესაბამისად $x_1,x_2,..,x_n$ ფაქტორებთან ცალცალკე, $\sigma_{x_1},\sigma_{x_2},..,\sigma_{x_n}$ - შესაბამისად $x_1,x_2,..,x_n$ – მიზეზობრივი ფაქტორების საშუალო კვადრატული გადახრაა,

$\sigma_y$– საშედეგო მოვლენის საშუალო კვადრატული გადახრაა.

აქედან გამომდინარე, მრავლობითი კორელაციის მთლიანი (საერთო) კოეფიციენტი გაინგარიშება ფორმულით:

$$R_{y/x_1,x_2,..,x_n} = \sqrt{\frac{a_1 z_{yx^1} \sigma_{x^1} + a_2 z_{yx^2} \sigma_{x^2} + .. + a_n z_{yx^n} \sigma_{x^n}}{\sigma_y}}\ .$$

სტატისტიკოსები გვთავაზობენ[1], აგრეთვე, დეტერმინაციის საერთო კოეფიციენტი გავიანგარიშოთ სტანდარტიზებულ მასშტაბში გაზომილი მრავლობითი რეგრესიული განტოლებით:

$$R^2{}_{y/x_1,x_2,...,x_n} = \beta_1 r_{yx_1} + \beta_2 r_{yx_2} + ... + \beta_n r_{yx_n} \qquad (8.82),$$

სადაც წინა მასალაში მოტანილი სტანდარტიზებულ მასშტაბებში წრფივი რეგრესიული განტოლების $t_1, t_2, .., t_n$ ცვლადები შეცვლილია წყვილადი კორელაციის კოეფიციენტებით. ამ ფორმულის საფუძველზე გამოიანგარიშება მრავლობითი კორელაციის კოეფიციენტი:

$$R^2{}_{y/x_1,x_2,...,x_n} = \sqrt{\beta_1 r_{yx_1} + \beta_2 r_{yx_2} + ... + \beta_n r_{yx_n}} \qquad (8.83)$$

თუ გვაქვს $y$ საშედეგო მოვლენასა და ორ მიზეზობრივ ფაქტორთან წყვილადი კავშირის კორელაციის კოეფიციენტები, აგრეთვე ამ ორ მიზეზობრივ ფაქტორს შორის კორელაციური კავშირის წყვილადი კოეფიციენტი ($z_{yx_1}, z_{yx_2}$ და $z_{x_1x_2}$), მაშინ მრავლობითი კორელაციის კოეფიციენტი გამოიანგარიშება ფორმულით[2]:

---

[1]იხ., მაგალითად, Òåîðèÿ ñòàòèñòèêè, ïîä ðåäàêöèåé Ð. À. Øìîéëîâîé, Ì.: Ôèíàíñû è ñòàòèñòèêà 2002, ñòð. 307. Òåîðèÿ ñòàòèñòèêè. Ó÷åáíèê ïîä ðåäàêöèåé ïðîô. Ð. Ë. Ãðîìûêî Ì.: ÈÍÔÐÀ–Ì 2002, ñòð. 252 და სხვ

[2]შევნიშნავთ, რომ მრავლობითი კორელაციის კოეფიციენტის სიდიდე დამოკიდებულია არა მარტო საშედეგო მოვლენასა ( $y$ ) და მიზეზობრივ ფაქტორებს, არამედ თვით მიზეზობრივ ფაქტორებს შორის კორელაციურ ურთიერთკავშირზე. ამასთან თუ წყვილადი კორელაციის წრფივი კოეფიციეტი ( $R_{xy}$ ) ძალიან მაღალია და აჭარბებს 0.8-ს, მაშინ სტატიტიკაში ორ ფაქტორს შორის ასეთი კავშირი წოდებულია **კოლინეარობის,** ხოლო მრავალ ფაქტორს შორის – **მულტიკოლინეარობის** სახელწოდებით



284

$$R_{y/x_1,x_2,...,x_n} = \sqrt{\frac{z^2{}_{yx_1} + z^2{}_{yx_2} - 2z_{yx_1}z_{yx_2}z_{x_1x_2}}{1 - z^2{}_{x_1x_2}}} \qquad (8.84).$$

გავიანგარიშოთ მრავლობითი კორელაციის საერთო კოეფიციენტი 24-ე ცხრილის მონაცემებით. ცხრილის პირველი და ბოლო სვეტის მიხედვით გავიანგარიშოთ დისპერსიები. აქედან პირველი სვეტის დისპერსია ($\sigma^2$) გვაძლევს საერთო დისპერსიას ანუ საშედეგო ნიშნის საერთო დისპერსიას.

ამ სვეტის მიხედვით მონაცემთა საშუალო არითმეტიკული შეადგენს 44,6-ს $\left(\frac{223.4}{5}\right)$. ამავე სიდიდეს მივიღებთ ბოლო სვეტის მიხედვით. ბოლო სვეტის მონაცემთა დისპერსია არის მოსწორებული დონეების ($\hat{y}$) დისპერსია, რომელიც შეადგენს:

$$\delta^2 = \frac{\Sigma(\hat{y} - \hat{\underline{y}})}{n} = \frac{(42.8-44.6)^2 + (43.7-44.6)^2 + (45.2-44.6)^2}{5} +$$

$$\frac{+(45.4-44.6)^2 + (46.0-44.6)^2}{5} = 1.402$$

პირველი სვეტის მონაცემების ($y$) მიხედვით საერთო დისპერსია უდრის:

$$\sigma^2 = \frac{\Sigma(y - \bar{y})}{n} = \frac{(42.5-44.6)^2 + (43.8-44.6)^2 + (45.6-44.6)^2}{5} +$$

$$+\frac{+(44.8-44.6)^2 + (46.7-44.6)^2}{5} = 2.1$$

ამ მონაცემებით დეტერმინაციის საერთო კოეფიციენტი ($R_{y/x_1,x_2}$) შეადგენს:

$$R_{y/x_1,x_2} = \frac{\delta^2}{\sigma^2} = \frac{1.402}{2.1} = 0.667$$


285

ეს გვიჩვენებს, რომ საშედეგო ნიშნის ანუ უმაღლსი და 1 ხარისხის ჩაის ხვედრითი წილის ვარიაციის (ცვალებადობის) 66,7% განისაზღვრება ჩვენს მიერ აღებული ფაქტორების (პირველი სორტის ნედლეულის ხვედრითი წილის ($x_1$), ნედლეულის გადამუშავების დაყოვნების დროის ($x_2$) ცვალებადობით, ვარიაციით.

მრავლობითი კორელაციის საერთო კოეფიციენტი, რომელიც განსაზღვრავს ჩვენს მიერ აღებულ საშედეგო მოვლენასა და მასზედ მოქმედ ფაქტორებს ($x_1, x_2$) შორის კორელაციური ურთიერთკავშირის სიმჭიდროვის ხარისხს, გაიანგარიშება **დეტერმინაციის** საერთო კოეფიციენტიდან კვადრატული ფესვის ამოღებით:

$$R_{y/x_1,x_2} = \sqrt{\frac{\delta_{\ 2}}{\sigma^{\ 2}}} = \sqrt{0.667} = 0.817$$

ეს იმას ნიშნავს, რომ საშედეგო ნიშნის (ჩაის მზა პროდუქციის ხარისხი) კავშირის სიმჭიდროვის ხარისხი შეადგენს 81,7%-ს, რაც ძალიან მაღალია და მეტყველებს ჩვენს მიერ შერჩეული ფაქტორების არსებითობაზე.

ახლა იმავე ცხრილის მონაცემების მოშველიებით ვაჩვენოდ მრავლობითი კორელაციის საერთო კოეფიციენტის გაანგარიშების სხვა წესებიც. ამისათვის ცხრილის მონაცემებით გავიანგარიშოთ საშუალო მაჩვენებლები საშუალო არითმეტიკული ფორმულის გამოყენებით $\left(\bar{x} = \frac{\Sigma x}{n}\right)$, საშუალო კვადრატული გადახრები დისპერსიისა და საშუალო კვადრატული გადახრის გაანგარიშების სამომენტო ანუ გამარტივებული ფორმულის $\left(\sigma = \sqrt{\overline{x^2} + (\bar{x})^2} \div\right)$ დახმარებით და წყვილადი წრფივი კორელაციის კოეფიციენტები ჩვენთვის


286

უკვე ცნობილი ფორმულის $\left( r_{xy} = \frac{\overline{xy} - \bar{x} \times \bar{y}}{\sigma_x \sigma_y} \right)$ გამოყენებით.

მათშორის: ა) საშუალოები:

$$\bar{y} = \frac{\Sigma y}{n} = \frac{223.4}{5} = 44.68\% \qquad \overline{x_1 x_2} = \frac{6360.9}{5} = 1272.18$$

$$\bar{x}_1 = \frac{\Sigma x_1}{n} = \frac{298.4}{5} = 59.68\% \qquad \overline{x_1 y} = \frac{13350.8}{5} = 2670.18$$

$$\bar{x}_2 = \frac{107}{5} = 21.4 \text{ საათი} \qquad \overline{x_2 y} = \frac{4769.2}{5} = 953.84$$

$$\overline{x_1^2} = \frac{17843.4}{5} = 3568.68 \qquad \overline{y^2} = \frac{9991.99}{5} = 1998.398$$

$$\overline{x_2^2} = \frac{2351}{5} = 470.2$$

ბ) საშუალო კვადრატული გადახრები:

$$\sigma_y = \sqrt{\overline{y^2} - (\bar{y})^2} = \sqrt{1998.398 - 1996.302} = 1.447$$

$$\sigma_{x_1} = \sqrt{\overline{x_1^2} - (\bar{x}_1)^2} = \sqrt{3568.68 - 3561.70} = 2.642$$

$$\sigma_{x_2} = \sqrt{\overline{x_1^2} - (\bar{x}_2)^2} = \sqrt{470.2 - 457.96} = 3.498$$

გ) კორელაციის წყვილადი წრფივი კოეფიციენტები:

$$r_{yx_1} = \frac{\overline{yx_1} - \bar{y} \times \bar{x}_1}{\sigma_y \times \sigma_x} = \frac{2670.18 - 44.68 \times 59.68}{1.447 \times 2.642} = 0.962$$

$$r_{yx_2} = \frac{\overline{yx_2} - \bar{y} \times \bar{x}_2}{\sigma_y \times \sigma_x} = \frac{953.84 - 21.4 \times 44.68}{1.447 \times 3.498} = -0.456$$

$$r_{x_1x_2} = \frac{\bar{x}_1\bar{x}_2 - \bar{x}_1 \times \bar{x}_2}{\sigma_{x_1} \times \sigma_x} = \frac{1272.18 - 1277.15}{2.642 \times 3.498} = -0.545$$

წყვილადი კორელაციის კოეფიციენტებიდან ორი მათგანის კავშირის სიმჭიდროვის ხრისხი უარყოფითი მაჩვენებელია. ეს ბუნებრივიცაა, ვინაიდან ნედლეულის გადამუშავების დაყოვნების დროის შემცირებასთან დაკავშირებით, როგორც ცხრილიდან ჩანს, იზრდება როგორც უმაღლესი და I სორტის მზა პროდუქციის, ასევე ნედლეულის საერთო მოცულობაში მაღალი ხარისხის ნედლეულის ხვედრითი წილი.

ამ კოეფიციენტების საფუძველზე შეგვიძლია დავწეროთ:

$$\begin{aligned} r_{y_1} &= \beta_1 + r_{12}\beta_2 + \ldots + r_{1k}\beta_k \\ r_{y_2} &= r_{21}\beta_1 + \beta_2 + \ldots + r_{2k}\beta_k \\ &. \quad . \quad . \quad . \quad . \quad . \quad . \\ r_{y_k} &= r_{k1}\beta_1 + r_{k2}\beta_2 + \ldots + \beta_k \end{aligned}$$

სადაც $r_{y_i}$ - საშედეგო ნიშედეგო ნიშნის თითოეული $i$-ურ ფაქტორულ ნიშანთან კავშირის წყვილადი, წრფივი კორელაციის კოეფიციენტებია,

 288

$r_{ij}$ –თითოეული $j$-ური ფაქტორის სხვა $i$-ურ ფაქტორებთან კავშირის წყვილადი, წრფივი კორელაციის კოეფიციენტებია.

მოცემული სისტემის წყვილადი კოეფიციენტები


289

წინასწარაა გაანგარიშებული და ნაცნობ სიდიდეებად ითვლებიან. სისტემის ამოხსნით მიიღება $\beta_1, \beta_2, \ldots, \beta_k$ კოეფიციენტები.

ეს ნორმალურ განტოლებათა ზოგადი სისტემა ჩვენი კონკრეტული ორფაქტორიანი მოდელისათვის $\left(R^2_{y/x_1,x_2} = \beta_1 r_{yx_1} + \beta_2 r_{yx_2}\right)$ ჩაიწერება შემდეგნაირად:

$$\begin{aligned} r_{yx_1} &= \beta_1 + r_{12}\beta_2 \\ r_{yx_2} &= r_{21}\beta_1 + \beta_2 \end{aligned} \qquad (8.85)$$

ჩვენს მაგალითზე:

$$\begin{aligned} r_{yx_1} &= 0.962 \\ r_{yx_2} &= -0.456 \\ r_{x_1x_2} &= -0.545 \end{aligned}$$

გავიანგარიშოთ მრავლობითი კორელაციის საერთო კოეფიციენტი ზემოთმოტანილი შემდეგი ფორმულის გამოყენებით:

$$R^2_{y/x_1,x_2} = \sqrt{\frac{r^2_{yx_1} + r^2_{yx_2} - 2r_{yx_1} \times r_{yx_2} \times r_{x_1x_2}}{1 - r^2_{x_1x_2}}}$$

ჩავსვათ ჩვენს მიერ გასაანგარიშებელი სიდიდეები:

$$R^2_{y/x_1,x_2=} \sqrt{\frac{0.962^2 + (-0.456)^2 - 2\times(-0.456)(-0.545)\times 0.962}{1-(-0.545)^2}} =$$

$$= \sqrt{\frac{0.962 + 0.207 - 2\times 0.962\times 0.248\times 0.962}{1-0.297}} = 0.965$$

იგივე მრავლობითი კორელაციის საერთო კოეფიციენტი შეგვიძლია გავიანგარიშოთ სტანდარტულ მასშტაბებში რეგრესიული განტოლების ფორმის საფუძველზე. ამისათვის საჭიროა წინა მასალაში მოტანილი დეტერმინაციის კოეფიციენტის შესაბამისი რეგრესიული

$R^2{}_{y/x_1,x_2} = \beta_1 r_{x_1} + \beta_2 r_{x_2} + ... + \beta_n r_{x_n}$ განტოლების $\beta_1, \beta_2, ..., \beta_n$ პარამეტრების გასაანგარიშებელი განტოლებათა სისტემა ჩავწეროთ შემდეგი ფორმით[1], გვექნება:

$$0.962 = \beta_1 - 0.545\beta_2$$

$$-0.456 = -0.545\beta_1 + \beta_2$$

ამოვხსნათ მოცემული სისტემა $\beta_1$ და $\beta_2$- ის მიმართ. ამისათვის მოსახერხებელია პირველი განტოლებიდან განვსაზღვროთ $\beta_1$.

$$\beta_1 = 0.962 + 0.545\beta_2$$

შევიტანოთ $\beta_1$-ის მნიშვნელობა მეორე განტოლებაში. გვექნება:

$$-0.456 = -0.545(0.962 + 0.545\beta_2) + \beta_2 = -0.524 - 0.297\beta_2 + \beta_2$$

$$0.703\beta_2 = 0.068,$$

$$\beta_2 = 0.096,$$

$$\beta_1 = 0.962 + 0.096 \times 0.545 = 1.014.$$

არიგად, სტანდარტიზებული სახით ჩვენი კონკრეტული შემთხვევისათვის მრავლობითი რეგრესიული განტოლება შეიძლება შემდეგი სახით ჩავწეროთ:

$$t_y = 1.014t_1 + 0.096t_2.$$

$t$-ს კოეფიციენტები მიუთითებს აშკარა ჭეშმარიტებაზე იმის შესახებ, რომ ჩაის მზა პროდუქციის ხარისხის ამაღლებაზე ყველაზე მეტ გავლენას ახდენს ნედლეულის ხარისხის

[1]ეს განტოლებათა სისტემა მიიღება უმცირეს კვადრატთა მეთოდის გამოყენებით და გულისხმობს, რომ მრავლობითი კორელაციის დეტერმინაციისა და კორელაციის საერთო კოეფიციენტები წყვილადი წრფივი კორელაციის კოეფიციენტების ფუნქციაა და მათ საფუძველზედაც შეიძლება ისინი გავიანგარიშოთ

გაუმჯობესება ($\beta_1 > \beta_2$). ეკონომიკურად ეს კოეფიციენტები იმაზე მიუთითებენ, რომ ნედლეულის ხარისხის ამაღლება საშუალოკვადრატული გადახრის ($\sigma$) ოდენობით, ანუ ჩვენს შემთხვევაში, რადგან $\sigma_{x_2} = 3.498$, ნედლეულის საერთო მასაში I სორტის ნედლეულის ხვედრითი წილის 3.498%-ით ამაღლება, ნედლეულის გადამუშავების საშუალო დროის უცვლელობის პირობებში, ჩაის მზა პროდუქციის ხარისხს აამაღლებს 1.014 საშუალოკვადრატული გადახრის ოდენობით ანუ ჩვენს შემთხვევაში უმაღლესი და პირველი ხარისხის მზა პროდუქციის ხვედრითი წილი გადიდდება 2.642X1.014=2.678 პროცენტით (რადგან $\sigma_y = 2.642$).

სტანდარტიზებული კოეფიციენტებიდან ნატურალურ კოეფიციენტებში გადასვლისათვის, როგორც წინა მასალაში იყო ნაჩვენები, ვიყენებთ ამ მაჩვენებლებს შორის შემდეგ თანაფარდობებს:

$$a_i = \frac{\sigma_y}{\sigma_{x_1}} \beta_1, \; a_0 = \bar{y} - a_1 x_1 - a_2 x_2 .$$

ჩვენს კონკრეტულ მაგალითზე:

$$a_1 = \frac{\sigma_y}{\sigma_{x_1}} \beta_1 = \frac{1.447}{2.642} \times 1.014 = 0.555 ,$$

$$a_2 = \frac{\sigma_y}{\sigma_{x_1}} \beta_2 = \frac{1.447}{3.498} \times 0.096 = 0.039 .$$

$$a_0 = \bar{y} - a_1 \bar{x}_1 - a_2 \bar{x}_2 = 44.68 - 0.555 \times 59.68 - 0.039 \times 21.4 =$$

10.72. მრავლობითი კორელაციის რეგრესიის განტოლება მიიღებს
სახეს:


291

$$\hat{y}_x = 10.72 + 0.555x_1 + 0.039x_2 .$$

მოცემული კონკრეტული ორფაქტორიანი მოდელის



292

შემთხვევისათვის მრავლობითი დეტერმინაციის კოეფიციენტი განისაზღვრება ფომულით:

$$R^2{}_{y/x_1,x_2} = \beta_1 r_{yx_1} + \beta_2 r_{yx_2}$$

თუ შევიტანთ ჩვენს მონაცემებს განტოლებაში, გვექნება:

$$R^2{}_{y/x_1,x_2} = 1.014 \times 0.962 + 0.096 \times (-0.456) = 0.929$$

აქედან ამოღებული კვადრატული ფესვი გვაძლევს მრავლობითი კორელაციის კოეფიციენტს, რაც შეადგენს 0.963-ს. ეს კი ემთხვევა წყვილადი კორელაციის კოეფიციენტების მიხედვით გაანგარიშებულ მაჩვენებელს. ეს მაჩვენებლი ე.ი. $y$ საშედეგო მოვლენასა და $x_1, x_2$ ფაქტორებს შორის მრავლობითი კორელაციის კოეფიციენტის მნიშვნელობა ემთხვევა $y$-ის ერთ-ერთ ფაქტორთან ($x$) კორელაციის წყვილადი კოეფიციენტის სიდიდეს $\left(r_{yx_1} = 0.962\right)$. ეს გამოწვეულია იმით, რომ ფაქტორებს შორის კორელაციური კავშირი მნიშენელოვანია $\left(r_{x_1x_2} = -0.545\right)$. რაც უფრო მაღალია ფაქტორებს შორის მულტიკორელაციის ხარისხი მით უფრო მრავლობითი კორელაციის საერთო კოეფიციენტის მნიშვნელობა უახლოვდება ერთერთი წყვილადი კორელაციური კავშირის სიმჭიდროვის მაჩვენებელს.

მრავლობითი კორელაციის საერთო კოეფიციენტი განსაზღვრავს შერჩეული ფაქტორების საშედეგო ნიშნის განვითარებაზე ზემოქმედების სიმჭიდროვის ხარისხს. ამასთან ერთად მნიშვნელოვანია, აგრეთვე, საშედეგო ნიშანზე თითოეული ფაქტორის ზემოქმედების დადგენა სხვა ფაქტორების უცვლელობის, ანუ თითოეული ფაქტორის სხვა ფაქტორებთან კავშირის ელიმინირების პირობებში. **ამისათვის სტატისტიკაში გამოიყენება დეტერმინაციისა და კორელაციის კერძო კოეფიციენტები.**

დეტერმინაციის კერძო კოეფიციენტი $\left(r^2{}_{yk(x_1,x_2,\dots,x_{n-1})}\right)$ გამოიანგარიშება დეტერმინაციის საერთო კოეფიციენტების

გამოყენებით:

$$r^2_{yk(x_1,x_2,...,x_{n-1})} = \frac{R^2_{y/x_1,x_2,...,x_n} - R^2_{y/x_1,x_2,...,x_{n-1}}}{1-R^2_{y/x_1,x_2,...,x_{n-1}}}$$

სადაც $r^2_{yk}$ - $k$-ური ფაქტორის საშედეგო მოვლენაზე ზემოქმედების კოეფიციენტია ამ ფაქტორის სხვა ფაქტორებთან ურთიერთკავშირის ელიმინირების პირობებში;

$R^2_{y/x_1,x_2,...,x_n}$ – დეტერმინაციის საერთო კოეფიციენტია, რომელიც ასახავს საშედეგო ნიშნის განვითარებაზე ანუ ვარიაციაზე მოდელში ჩართული ყველა ფაქტორის ერთდროული ზემოქმედების ხარისხს;

$R^2_{y/x_1,x_2,..,x_{n-1}}$ – დეტერმინაციის საერთო კოეფიციენტია, რომელიც ასახავს ყველა ფაქტორის გავლენას საშედეგო მოვლენის განვითარებაზე გარდა ერთი, $K$-ური ფაქტორისა, რომლის ზემოქმედებასაც ჩვენ ვზომავთ.

ჩვენი შემთხვევისათვის კონკრეტული მაგალითის მიხედვით:

$$R^2_{y/x_1,x_2} = 0.929,$$

$$R^2_{y/x_1} = 1.014 \times 0.962 = 0.975,$$

$$R^2_{y/x_2} = -0.043.$$

აქედან პირველი ფაქტორის ანუ ჩაის მზა პროდუქციის ამაღლებაზე ნედლეულის ხარისხის გაუმჯობესება იმოქმედებს შემდეგი ოდენობით:

$$r^2_{y/x_1} = \frac{R^2_{y/x_1,x_2} - R^2_{y/x_2}}{1-R^2_{y/x_2}} = \frac{0.929+0.043}{1+0.043} = \frac{0.972}{1.043} = 0.932,$$

ხოლო მეორე ფაქტორის ანუ ნედლეულის გადამუშავების დაყოვნების დრო:

$$r^2_{y/x_2} = \frac{R^2_{y/x_1,x_2} - R^2_{y/x_1}}{1 - R^2_{y/x_1}} = \frac{0.929 - 0.975}{1 - 0.975} = \frac{-0.046}{0.025} = -1.84$$

მიღებული კოეფიციენტები მიუთითებენ იმ ბუნებრივ გარემოებაზე, რომ ჩაის მწვანე ფოთლის ნედლეულის ხარისხის გაუმჯობესება ყველაზე დიდ გავლენას ახდენს ჩაის მზა პროდუქციის ხარისხის ამაღლებაზე, ხოლო ნედლეულის გადამუშავების დაყოვნების დროის გადიდება, პირიქით, დიდ უარყოფით ზემოქმედებას ახდენს ჩაის მზა პროდუქციის ხარისხზე, ვინაიდან კერძო კოეფიციენტი შეადგენს $-1.84$-ს.

დეტერმინაციის კერძო კოეფიციენტებიდან ამოღებული კვადრატული ფესვი გვაძლევს კორელაციის კერძო კოეფიციენტებს. პირველი ფაქტორის მიხედვით კორელაციის კერძო კოეფიციენტი იქნება: $r_{y/x_1} = \sqrt{0.932} = 0.965$, ხოლო მეორე ფაქტორის მიხედვით $r_{y/x_2} = \sqrt{-1.84} = -1.356$. ამის გარდა, თითოეული ფაქტორის გავლენის ხარისხს ზომავენ ელასტიურობის კერძო კოეფიციენტების დახმარებით. ელასტიურობის კერძო კოეფიციენტი გვიჩვენებს მიზეზობრივი ფაქტორის 1%-ით ცვალებადობისას საშედეგო ნიშნის რამდენი პროცენტით ცვალებადობას უნდა ველოდოთ. მის გასაანგარიშებლად სტატისტიკოსების მიერ შემოთავაზებულია ფორმულა:

$$Э_i = a_i \frac{\bar{x}_i}{y} \qquad (8.86),$$

სადაც $Э_i$ - ფაქტორის ელასტიურობის კოეფიციენტია,

$a_i$ – ფაქტორის მიხედვით რეგრესიის კოეფიციენტია,

$\bar{x}_i$ და $\bar{y}$ –შესაბამისად, $x_i$ ფაქტორისა და $y$ საშედეგო მოვლენის საშუალო არითმეტიკული სიდიდეებია.

ჩვენს კონკრეტულ მაგალითზე ელასტიურობის

კოეფიციენტების გასაანგარიშებელი


296

სიდიდეებია:

$$a_1 = 0.45,\ a_2 = -0.011$$
$$\bar{y} = 44.68,\ \bar{x}_1 = 59.68$$
$$\bar{x}_2 = 21.4$$

ამ მონაცემებით ელასტიურობის კოეფიციენტები იქნება:

$$Э_1 = a_1 \frac{\bar{x}_1}{\bar{y}} = 0.45 \frac{59.68}{44.68} = 0.60$$

$$Э_2 = a_2 \frac{\bar{x}_2}{\bar{y}} = \frac{21.4}{44.68} \times (-0.011) = -0.005$$

აქედან ჩანს, რომ ნედლეულის ხარისხის 1%-ით გაუმჯობესება ჩაის მზა პროდუქციის ხარისხს აუმჯობესებს 0,6%-ით, ხოლო ნედლეულის გადამუშავების დაყოვნების დროის გადიდება 0,005%-ით აუარესებს მზა პროდუქციის ხარისხს.

## 20. მრავლობითი კორელაციის რანგების კონკორდაციის კოეფიციენტი

წინა მასალაში განხილული **რანგების სპირმენისა და კენდელის** კოეფიციენტები გამოიყენება მხოლოდ და მხოლოდ წყვილადი კორელაციური კავშირების შემთხვევებისთვის. ორზე მეტი ნიშნის რანჟირებული მწკრივების განხილვისას საქმე გვაქვს არა წყვილად, არამედ მრავლობით კორელაციასთან. ამ შემთხვევისათვის ინგლისელი **სტატისტიკოსების მ. კენდალისა და ბ. სმიტის** მიერ მოვლენებს შორის ურთიერთკორელაციური კავშირის სიმჭიდროვის ხარისხის გასაზომავად შემოთავაზებულია **რანგების კონკორდაციის** კოეფიციენტები. მათი გაზომვისათვის გამოიყენება ფორმულები:

 297

$$K_{\text{კონკ.}} = \frac{12S}{m^2(n^3 - n)} \qquad (8.87).$$


298

ეს ფორმულა გამოსაყენებელია მხოლოდ იმ შემთხვევაში, როცა თითოეული ნიშნის რანგები არ მეორდება რანჟირებულ მწკრივში. თუ ეს ასე არ არის და რანგები რანჟირებულ მწკრივებში მეორდება ორჯერ და მეტად, მაშინ გამოიყენება მეორე ფორმულა:

$$K_{\text{კონკ.}} = \frac{12S}{m^2(n^3 - n) - m\sum_{1}^{m}(t^3 - t)} \qquad (8.88).$$

სადაც $S$ – ნიშანთა მიხედვით რანგების ჯამის მისი საშუალო მნიშვნელობიდან ($T$) გადახრების კვადრატების ჯამია,

$m$ – რანჟირებულ ნიშანთა რიცხვია,

$n$ – დაკვირვების რიცხვი,

$t$ – თითოეული ნიშნის მიხედვით ერთნაირი რანგების რიცხვი.

მოვიტანოთ კონკრეტული მაგალითი, რომლითაც გასაგები გახდება რანგების კონკორდაციის კოეფიციენტის გაანგარიშება როგორც პირველი ისე მეორე ფორმულის გამოყენებით. მაგალითად, საქართველო-საპარლამენტო საარჩევნო პროგრამებში მეტად პრიორიტულია ისეთი საკვანძო, ქვეყნისათვის სასიცოცხლო მნიშვნელობის საკითხები, როგორიცაა ეკონომიკის აღორძინება, ქვეყნის ტერიტორიულ-ეკონომიკური მთლიანობის აღდგენა, მოსახლეობის სოციალური მდგომარეობის გაუმჯობესება, ოპტიმალური, დამოუკიდებელი ქვეყნისათვის შესაბამისი საკადრო პოლიტიკის გატარება, დაპირებათა შესრულების საიმედოობა და ა.შ.

თუ პირველი ოთხი ფაქტორის შეფასებას რანგების მიხედვით მივანდობთ სამ ექსპერტს, შესაძლებელია ასეთი პირობითი ციფრები მივიღოთ.

მენეჯერმა ბაზრობებზე ფირმის პროდუქციის რეალიზაციაზე მომქმედი ოთხი ფაქტორი გამოჰყო და ექპერტებს მოუხმო, ამ ფაქტორტა ინდივიდუალური შეფასებისათვის.


299

ფაქტორთა საექსპერტო შეფასებანი ასეთია (რანგებში)

ცხრილი №38

| ფაქორული ნიშნები $x_i$ | ექსპერტთა მიერ მიკუთვნებული რანგები $R_{ij}$ | | | რანგების ჯამი $\sum_1^m R_{ij}$ | რანგების ჯამის კვადრატი $\left(\sum_1^m R_{ij}\right)^2$ | რანგების ჯამის მისი საშუალოდან $(T)$ გადახრების კვადრატი $\left(\sum R_{ij} - T\right)^2$ |
|---|---|---|---|---|---|---|
| | I ექსპერტი $R_{1j}$ | II ექსპერტი $R_{2j}$ | III ექსპერტი $R_{3j}$ | | | |
| 1. | 2. | 3. | 4. | 5. | 6. | 7. |
| $x_1$ | 1 | 2 | 1 | 4 | 16 | 12.25 |
| $x_2$ | 2 | 1 | 2 | 5 | 25 | 6.25 |
| $x_3$ | 3 | 3 | 4 | 10 | 200 | 6.25 |
| $x_4$ | 4 | 4 | 3 | 11 | 121 | 12.25 |
| $\Sigma$ | 10 | 10 | 10 | 30 | 262 | 37.0 |

ჩვენი ცხრილის მონაცემებით $m = 3$, $n = 4$, $S$ -ის მნიშვნელობის გასაანგარიშებლად სტატისტიკოსები ორ მეთოდს გვთავაზობენ[1]:

I მეთოდით

$$S = \sum_1^n \left( \sum_1^m R_{ij} \right)^2 - \frac{\left( \sum_1^n \sum_1^m R_{ij} \right)^2}{n} = 262 - \frac{30}{4} = 37,$$

II მეთოდით

$$S = \sum_1^n \left( \sum_1^m R_{ij} - T \right) = (4 - 7.5)^2 + (5 - 7.5)^2 + (10 - 7.5)^2 + (11 - 7.5)^2 = 37.$$

$$T = \frac{30}{4} = 7.5.$$



300

თუ ჩვენს მონაცემებს შევიტანთ რანგების კონკორდაციის გასაანგარიშებელ პირველ ფორმულაში, გვექნება:

$$K_{\text{კოეფ.}} = \frac{12\times 37}{m^2(n^3-n)} = -\frac{12\times 37}{3^2(4^3-4)} = 0.822 \; .$$



301

როგორც ჩანს, კონკორდაციის კოეფიციენტი ძალიან მაღალია. ეს იმას ნიშნავს, რომ ექსპერტთა შეფასებებს შორის კავშირი ძალიან მჭიდროა, რაც ამაღლებს ამ შეფასებათა საიმედოობის ხარისხს.

ამავე წესებით ადვილია, აგრეთვე, კონკორდაციის კოეფიციენტის გაანგარიშება ცალკეული ექსპერტის შეფასებათა შორის განმეორებათა შემთხვევაში.


302