ივანე ჯავახიშვილის სახელობის თბილისის
სახელმწიფო უნივერსიტეტი

ომარ ფურთუხია

ამოცანათა კრებული
სტატისტიკური დასკვნების თეორიაში
(შემოკლებული ვარიანტი)

თსუ - 2020

სარჩევი

სტატისტიკური დასკვნების თეორია

თავი I

წერტილოვანი და ინტერვალური შეფასებები. ნდობის ინტერვალი საშუალოსათვის

შეფასების ამოცანა: n მოცულობის $X=(X_1,...,X_n)$ შერჩევის საფუძველზე გავაკეთოთ დასკვნები გენერალური ერთობლიობის უცნობი θ პარამეტრის შესახებ. შერჩევის ნებისმიერ $T_n(X)=T_n(X_1,...,X_n)$ ფუნქციას **სტატისტიკა (შეფასება)** ეწოდება. **წერტილოვანი შეფასების ამოცანაა** მოიძებნოს ისეთი სტატისტიკა $T_n(X_1,...,X_n)$, რომლის შერჩევითი მნიშვნელობა $T_n(x_1,...,x_n)$, გარკვეული აზრით, შეიძლება ჩაითვალოს უცნობი θ პარამეტრის ჭეშმარიტი (რეალური) მნიშვნელობის მიახლოებად (შეფასებად) და გამოყენებულ იქნეს მის ნაცვლად. ასეთ სტატისტიკას (შეფასებას) **წერტილოვანი სტატისტიკა (შეფასება)** ეწოდება.

$T_n(X)=T_n(X_1,...,X_n)$ შეფასებას ეწოდება **ჩაუნაცვლებელი (ანუ გადაუადგილებადი),** თუ $E_\theta T(X)=\theta$.

$T_n(X)=T_n(X_1,...,X_n)$ შეფასებას ეწოდება **ძალმოსილი (ანუ ძალდებული),** თუ $T_n \xrightarrow{P_\theta} \theta$ (T_n ალბათობით კრებადია θ-სკენ), როცა $n\to\infty$.

ჩაუნაცვლებელ შეფასებას ეწოდება **ოპტიმალური (ანუ ეფექტური),** თუ მას სხვა ჩაუნაცვლებელ შეფასებებს შორის გააჩნია უმცირესი დისპერსია.

თუ n მოცულობის შერჩევა წარმოდგენილია ვარიაციული მწკრივის სახით, მაშინ **შერჩევითი საშუალო** ეწოდება სიდიდეს:

$$\bar{x}=\frac{x_1n_1+x_2n_2+...+x_kn_k}{n}=x_1\frac{n_1}{n}+x_2\frac{n_2}{n}+...+x_k\frac{n_k}{n}.$$

თუ შერჩევიდან მიღებული მნიშვნელობები არაა დაჯგუფებული, მაშინ შერჩევითი საშუალო

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i \ .$$

შერჩევითი საშუალო წარმოადგენს უცნობი მათემატიკური ლოდინის ჩაუნაცვლებელ და ძალმოსილ შეფასებას.

შერჩევითი დისპერსია ეწოდება სიდიდეს:

$$s^2=\sum_{i=1}^{k}\left(x_i-\bar{x}\right)^2\frac{n_i}{n}=\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2 \ .$$

შესწორებული შერჩევითი დისპერსია ეწოდება სიდიდეს:

$$s'^2=\frac{n}{n-1}s^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2 \ .$$

შესწორებული შერჩევითი დისპერსია წარმოადგენს უცნობი დისპერსიის ჩაუნაცვლებელ და ძალმოსილ შეფასებას.

შერჩევითი სტანდარტული გადახრა (შესაბამისად, **შესწორებული შერჩევითი სტანდარტული გადახრა)** ეწოდება არითმეტიკულ კვადრატულ ფესვს შერჩევითი დისპერსიიდან $s=\sqrt{s^2}$ (შესაბამისად, შესწორებული შერჩევითი დისპერსიიდან $s'=\sqrt{s'^2}$).

შერჩევითი ასიმეტრიის კოეფიციენტი

$$a_{შერ}=\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^3}{\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^3}$$

შერჩევითი ექსცესის კოეფიციენტი

$$e_{შერ}=\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^4}{\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^4}-3$$

შერჩევითი კორელაციის კოეფიციენტი ეწოდება სიდიდეს:

$$r_{xy}=\frac{\overline{xy}-\bar{x}\,\bar{y}}{s_x s_y}\text{, სადაც } \overline{xy}=\frac{1}{n}\sum_{i=1}^{n}x_i y_i \ ;$$

$$s_x=\sqrt{s_x{}^2}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}\ ; \qquad s_y=\sqrt{s_y{}^2}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2}\ .$$

შერჩევითი პარამეტრების განაწილება ნორმალური პოპულაციისათვის.

დავუშვათ, რომ $X_1, X_2, ..., X_n$ წარმოადგენს შერჩევას ნორმალური გენერალური ერთობლიობიდან, $X_i \cong N(a,\sigma^2)$ $(i=1,2,...,n)$, მაშინ: $\overline{X}$ და S^2 ($S^{'2}$) დამოუკიდებელი შემთხვევითი სიდიდეებია; $\overline{X} \cong N(a,\sigma^2/n)$;

$$\sum_{i=1}^{n}(X_i - a)^2/\sigma^2 \cong \chi^2(n)\,;\ \frac{(n-1)S^{'2}}{\sigma^2} \cong \chi^2(n-1)\,;$$

$$Z = \frac{\overline{X}-a}{\sigma/\sqrt{n}} \cong N(0,1)\,;\ T = \frac{\overline{X}-a}{S/\sqrt{n-1}} = \frac{\overline{X}-a}{S^{'}/\sqrt{n}} \cong t(n-1)\,.$$

მაქსიმალური დასაჯერობის მეთოდი: ვთქვათ, $p(x_i,\theta)$ არის ალბათობა იმისა, რომ ექსპერიმენტის შედეგად დისკრეტული ξ შემთხვევითი სიდიდე მიიღებს x_i მნიშვნელობას. **მაქსიმალური დასაჯერობის ფუნქცია** ეწოდება θ არგუმენტის ფუნქციას: $L(x_1, x_2, \dots, x_n;\ \theta) = p(x_1,\ \theta)p(x_2,\ \theta)\dots p(x_n,\ \theta)$ ($\ln L$ ფუნქციას **მაქსიმალური დასაჯერობის ლოგარითმული ფუნქცია** ეწოდება). **მაქსიმალური დასაჯერობის შეფასებას** უწოდებენ θ-ს იმ მნიშვნელობას, სადაც მაქსიმალური დასაჯერობის ფუნქცია (ან რაც იგივეა $\ln L$) აღწევს თავის მაქსიმუმს. მის მოსაძებნად საჭიროა: 1. ვიპოვოთ წარმოებული $\partial \ln L/\partial\theta$; 2. გავუტოლოთ წარმოებული ნულს (მივიღებთ ე. წ. **მაქსიმალური დასაჯერობის ლოგარითმულ განტოლებას**) და ვიპოვოთ კრიტიკული წერტილები; 3. ვიპოვოთ მეორე წარმოებული $\partial^2 \ln L/\partial\theta^2$; თუ ის უარყოფითია კრიტიკულ წერტილში, მაშინ ეს წერტილი – მაქსიმუმის წერტილია.

უწყვეტი შემთხვევითი სიდიდის შემთხვევაში, რომლის $f(x,\ \theta)$ განაწილების სიმკვრივის სახე ცნობილია, მაგრამ იგი შეიცავს უცნობ θ პარამეტრს, მაქსიმალური დასაჯერობის ფუნქციას აქვს შემდეგი სახე:

$$L(x_1, x_2, ..., x_n;\theta) = f(x_1,\theta)\cdot f(x_2,\theta)\cdots f(x_n,\theta)\,.$$

უცნობი პარამეტრის მაქსიმალური დასაჯერობის შეფასების საპოვნელად უნდა ჩავატაროთ იგივე პროცედურები, რაც დისკრეტულ შემთხვევაში.

მომენტთა მეთოდი: მომენტთა მეთოდი დაფუძნებულია იმ გარემოებაზე, რომ საწყისი და ცენტრალური ემპირიული მომენტები წარმოადგენენ შესაბამისი საწყისი და ცენტრალური

თეორიული მომენტების ძალმოსილ შეფასებებს. ამიტომ ამა თუ იმ შეფასების მისაღებად თეორიული მომენტები უნდა გავუტოლოთ იმავე რიგის შესაბამის ემპირიულ მომენტებს და ამოვხსნათ მიღებული განტოლება ან განტოლებათა სისტემა.

θ უცნობი პარამეტრის γ (ან $1-\alpha$) **საიმედოობის მქონე ანუ** $100\gamma\%$**-იანი (ან** $100(1-\alpha)\%$**-იანი) ნდობის ინტერვალი** ეწოდება. ინტერვალს (T_1, T_2), რომლისთვისაც: $P\{T_1 < \theta < T_2\} = \gamma$ (ან $1-\alpha$), სადაც T_1 და T_2 θ პარამეტრის გარკვეული წერტილოვანი შეფასებებია, $\gamma \in (0,1)$ ($\alpha \in (0,1)$); α-ს **მნიშვნელოვნების დონე** ეწოდება; ნდობის ინტერვალის სიგრძის ნახევარს **შეფასების სიზუსტე** ეწოდება.

α **მნიშვნელოვნების დონის მქონე ნდობის ინტერვალი ნორმალური პოპულაციის მათემატიკური ლოდინისათვის ცნობილი** σ^2 **დისპერსიის შემთხვევაში** არის:

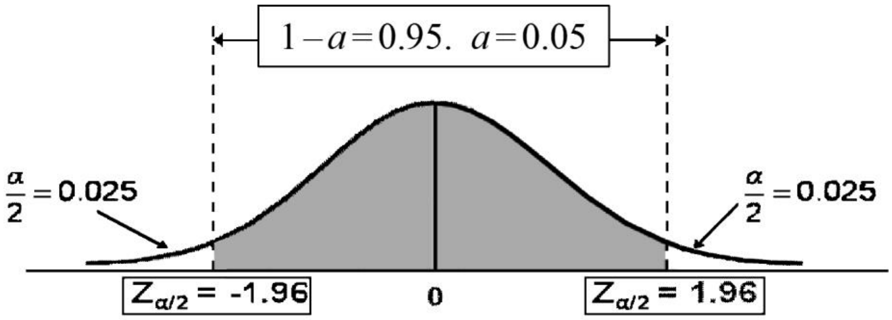
$$\left(\overline{X} - \sigma \cdot \frac{z_{\alpha/2}}{\sqrt{n}}, \overline{X} + \sigma \cdot \frac{z_{\alpha/2}}{\sqrt{n}}\right),$$

სადაც $z_{\alpha/2}$ სტანდარტული ნორმალური განაწილების ზედა $\alpha/2$-კრიტიკული წერტილია. აქ შეფასების სიზუსტეა $l = \sigma \cdot \frac{z_{\alpha/2}}{\sqrt{n}}$.

შეფასების წინასწარ დაფიქსირებული l სიზუსტის უზრუნველსაყოფად საჭირო შერჩევის მინიმალური მოცულობაა:

$$n^* = [(\frac{\sigma}{l} z_{\alpha/2})^2] + 1.$$

95%-იანი ნდობის ინტერვალი



α მნიშვნელოვნების დონის მქონე ნდობის ინტერვალს ნორმალური პოპულაციის მათემატიკური ლოდინისათვის უცნობი დისპერსიის შემთხვევაში აქვს შემდეგი სახე:

$$(\overline{X}-t_{n-1,\alpha/2}\cdot\frac{S'}{\sqrt{n}},\overline{X}+t_{n-1,\alpha/2}\cdot\frac{S'}{\sqrt{n}}),$$

სადაც $t_{n-1,\alpha/2}$ არის თავისუფლების $n-1$ ხარისხის მქონე სტიუდენტის განაწილების ზედა $\alpha/2$-კრიტიკული წერტილი.

α მნიშვნელოვნების დონის მქონე ასიმპტოტური ნდობის ინტერვალი (არანორმალური) პოპულაციის საშუალოსათვის შერჩევის დიდი მოცულობის შემთხვევაში ($n\geq 30$), როცა დისპერსია ცნობილია, არის:

$$(\overline{X}_n-z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}},\overline{X}_n+z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}}),$$

ხოლო უცნობი დისპერსიის შემთხვევაში კი:

$$(\overline{X}_n-z_{\alpha/2}\cdot\frac{S'}{\sqrt{n}},\overline{X}_n+z_{\alpha/2}\cdot\frac{S'}{\sqrt{n}}).$$

α მნიშვნელოვნების დონის მქონე ნდობის ინტერვალი (არანორმალური) პოპულაციის საშუალოსათვის, როცა σ უცნობია და $n<30$:

$$(\overline{X}-t_{n-1,\alpha/2}\cdot\frac{S'}{\sqrt{n}},\overline{X}+t_{n-1,\alpha/2}\cdot\frac{S'}{\sqrt{n}}).$$

როდის გამოიყენება Z ან T განაწილება

ცნობილია თუ არა σ? — დიახ → გამოიყენე $z_{\alpha/2}$ შერჩევის მოცულობის მიუხედავად*

↓ არა

არის თუ არა $n\geq 30$? — დიახ → გამოიყენე $z_{\alpha/2}$ და ფორმულაში σ შეცვალე S'-ით

↓ არა

გამოიყენე $t_{n-1,\alpha/2}$ და S' ფორმულაში**

* სიდიდე უნდა იყოს ნორმალურად განაწილებული როცა $n<30$.

** სიდიდე უნდა იყოს მიახლოებით ნორმალური.

მაგალითი 1. უნივერსიტეტის რექტორს სურს შეაფასოს წელს ჩარიცხული სტუდენტების საშუალო ასაკი. წარსული გამოკვლევებიდან ცნობილია, რომ სტანდარტული გადახრა არის 2 წელი. აღებულია 50 სტუდენტისაგან შემდგარი შერჩევა და მისთვის გამოთვლილი საშუალო ტოლია 23.2 წლის. იპოვეთ 95%-იანი ნდობის ინტერვალი პოპულაციის საშუალოსათვის. იგულისხმება, რომ ასაკი დაახლოებით ნორმალურადაა განაწილებული.

ამოხსნა. 95%-იანი ნდობის ინტერვალი ნიშნავს, რომ ნდობის ინტერვალის საიმედოობა $1-\alpha = 95\% / 100\% = 0.95$, ანუ $\alpha = 0.05$. ამიტომ ნორმალური განაწილების ცხრილიდან ვპოულობთ $z_{\alpha/2} = z_{0.025} = 1.96$. შესაბამისად, საძიებელი ნდობის ინტერვალი იქნება:

$$(23.2 - 1.96 \cdot \frac{2}{\sqrt{50}}, 23.2 + 1.96 \cdot \frac{2}{\sqrt{50}}),\ (22.6, 23.8),$$

ანუ 23.2 ± 0.6. ამრიგად, 50 სტუდენტის ასაკიდან გამომდინარე, 95%-იანი საიმედოობით, უნივერსიტეტის რექტორს შეუძლია თქვას, რომ სტუდენტების საშუალო ასაკი მოთავსებულია 22.6 წელსა და 23.8 წელს შორის.

მაგალითი 3. კოლეჯის პრეზიდენტმა დაავალა სტატისტიკის მასწავლებელს შეაფასოს კოლეჯის სტუდენტების საშუალო ასაკი. რა მოცულობის შერჩევაა აუცილებელი? სტატისტიკის მასწავლებელი თვლის, რომ საიმედოობა უნდა იყოს 99%, რათა შეფასება იყოს სწორი ერთი წლის სიზუსტით. წინა გამოკვლევებიდან ცნობილია, რომ ასაკთა სტანდარტული გადახრა არის 3 წელი. იგულისხმება, რომ ასაკი დაახლოებით ნორმალურდაა განაწილებული

ამოხსნა. ამ შემთხვევაში $l = 1$; $\sigma = 3$; $\alpha = 1 - 0.99 = 0.01$. ამიტომ $z_{\alpha/2} = z_{0.005} = 2.58$. შესაბამისად, შერჩევის მინიმალური მოცულობა იქნება:

$$n^* = [(\frac{2.58}{1} \cdot 3)^2] + 1 = [59.9] + 1 = 60.$$

ამრიგად, იმისათვის, რომ მასწავლებელმა შეაფასოს სტუდენტთა ასაკის საშუალოს ჭეშმარიტი მნიშვნელობა ერთი წლის სიზუსტით 99%-იანი საიმედოობით, მას სჭირდება სულ ცოტა 60 სტუდენტისაგან შემდგარი შერჩევა.

მაგალითი 5. მოცემულია გენერალური ერთობლიობა გარკვეული მახასიათებლით, რომელიც განაწილებულია ნორმალურად 6.25-ის ტოლი დისპერსიით. ჩატარებულია $n = 27$ მოცულობის შერჩევა და მიღებულია მახასიათებლის საშუალო შერჩევითი მნიშვნელობა $\bar{x} = 12$. ვიპოვოთ ნდობის ინტერვალი, რომელიც ფარავს გენერალური ერთობლიობის გამოსაკვლევი მახასიათებლის უცნობ მათემატიკურ ლოდინს საიმედოობით $\gamma = 0.99$.

ამოხსნა. გვაქვს: $\sigma = \sqrt{6.25} = 2.5$; $n = 27$; $\bar{x} = 12$; $\alpha = 1 - 0.99 = 0.01$; $z_{\alpha/2} = z_{0.005} = 2.57$. აქედან ვღებულობთ საძებნ ნდობის ინტერვალს: (10.76, 13.24).

ამოცანები (*σ ცნობილია ან $n \geq 30$*)

1. იპოვეთ $z_{\alpha/2}$ (სტანდარტული ნორმალური განაწილების $\alpha/2$-ზედა კრიტიკული წერტილი): ა) 99%-იანი; ბ) 98%-იანი; გ) 95%-იანი; დ) 90%-იანი და ე) 94%-იანი ნდობის ინტერვალებისათვის.

3. ქვემოთ მოყვანილი შერჩევის მონაცემების მიხედვით გამოთვალეთ შერჩევითი საშუალო, შესწორებული სტანდარტული გადახრა და ააგეთ 95%-იანი ნდობის ინტერვალი პოპულაციის საშუალოსათვის:

47.596 68.751 5.838 69.831 28.843 53.107 31.391 48.829 50.706 32.785
62.892 55.105 63.974 56.674 38.362 51.549 31.938 31.851 56.088 48.321
34.906 38.359 72.086 34.009 50.850 43.801 46.127 49.926 54.960 49.671

5. გამოთვალეთ შერჩევითი საშუალო, შესწორებული სტანდარტული გადახრა და ააგეთ 90%-იანი ნდობის ინტერვალი ევროპელი მოქალაქის წლიური შემოსავლის საშუალოსათვის, თუ ქვემოთ მოყვანილია შემთხვევით შერჩეული 50 ევროპელი მოქალაქის შემოსავლები გაზომილი ათასობით ევროებში:

84	14	31	72	26	49	252	104	31	8
3	18	72	23	55	133	16	29	225	138
85	24	391	72	158	4340	346	19	5	846
461	254	125	61	123	60	29	10	366	47
28	254	6	77	21	97	6	17	8	82

7. შემთხვევით შერჩეული 40 თავდამსხმელის საშუალო ქულა არის 186, ხოლო თავდამსხმელთა პოპულაციის სტანდარტული გადახრაა 6. ა) ააგეთ 95%-იანი ნდობის ინტერვალი თავდამსხმელთა პოპულაციის საშუალო ქულისათვის; ბ) აა-

გეთ 95%-იანი ნდობის ინტერვალი თავდამსხმელთა პოპულაციის საშუალო ქულისათვის იმ შემთხვევაში, როცა 40 თავდამსხმელის ნაცვლად განიხილება 100 თამდამსხმელისაგან შედგენილი შერჩევა იმავე საშუალო ქულით; გ) რომელი ინტერვალია უფრო პატარა (ზუსტი) და რატომ?

9. გამოკვლევებმა აჩვენა, რომ აშშ-ს მოქალაქეს საშუალოდ ჭირდება 5.9 თვე ახალი სამუშაოს მოსაძებნად. ააგეთ 95%-იანი ნდობის ინტერვალი სამუშაოს მოძებნის საშუალოსათვის, თუ გამოკითხული 36 სამუშაოს მძებნელისათვის შესწორებული სტანდარტული გადახრა აღმოჩნდა 0.8 თვე.
11. 48 შემთხვევით შერჩეული დღის მონაცემების მიხედვით დიდ ჰოსპიტალში დღის განმავლობაში შემოსული პაციენტების საშუალოდ 38% საჭიროებს გადაუდებელ სასწრაფო დახმარებას. პოპულაციის სტანდარტული გადახრაა 4. ა) ააგეთ 99%-იანი ნდობის ინტერვალი იმ პაციენტების რაოდენობის საშუალოსათვის, ვინც საჭიროებს გადაუდებელ სასწრაფო დახმარებას; ბ) ააგეთ 99%-იანი ნდობის ინტერვალი იმ პაციენტების რაოდენობის საშუალოსათვის, ვინც საჭიროებს გადაუდებელ სასწრაფო დახმარებას, თუ პოპულაციის სტანდარტული გადახრაა 8 ნაცვლად 4-სა; გ) რატომაა მეორე შემთხვევაში ნდობის ინტერვალი უფრო განიერი?
13. მკვლევარს სურს შეაფასოს დედაქალაქის პოლიციის ოფიცრის საშუალო ხელფასი. მისი მიზანია, რომ შეფასების საიმედოობა იყოს 95%. ცნობილია, რომ პოლიციის ოფიცრის ხელფასის სტანდარტული გადახრაა 1050 ლარი. რა მოცულობის შერჩევა დასჭირდება მკვლევარს იმისათვის, რომ შეფასების სიზუსტე იყოს 200 ლარი?
15. ცენტრალური სამხედრო ჰოსპიტალის 117 სხვადასხვა ოთახში ხმაურის საშუალო დონე იყო 58 დეციბალი, ხოლო შესწორებული სტანდარტული გადახრა 4.8 დეციბალი. ააგეთ 90%-იანი ნდობის ინტერვალი ჰოსპიტალში ხმაურის დონის ჭეშმარიტი საშუალოსათვის.
17. რესტორნის მეპატრონეს სურს დაადგინოს 99%-იანი ნდობის ინტერვალი შამპანიურის ფასის რეალური საშუალოსათვის. რა მოცულობის უნდა იყოს შერჩევა, რომ შეფასების სიზუსტე იყოს 0.1 ლარი, თუ წინა გამოკვლევის თანახმად ფასის სტანდარტული გადახრა იყო 0.12 ლარი.

ამოცანები *(σ უცნობია და $n<30$)*

19. იპოვეთ $t_{n-1,\alpha/2}$ (თავისუფლების $(n-1)$ ხარისხის მქონე სტუდენტის განაწილების $\alpha/2$-ზედა კრიტიკული წერტილი): ა) 99%-იანი ნდობის ინტერვალისათვის, როცა $n=18$; ბ) 95%-იანი ნდობის ინტერვალისათვის, როცა $n=23$; გ) 98%-იანი ნდობის ინტერვალისათვის, როცა $n=15$; დ) 90%-იანი ნდობის ინტერვალისათვის, როცა $n=10$; ე) 95%-იანი ნდობის ინტერვალისათვის, როცა $n=20$.

ქვემოთ მოყვანილ ამოცანებში იგულისხმეთ, რომ ყველა სიდიდე განაწილებულია დაახლოებით ნორმალურად:

21. ქვემოთ მოყვანილი მონაცემების მიხედვით გამოთვალეთ შერჩევითი საშუალო, შესწორებული სტანდარტული გადახრა და ააგეთ 95%-იანი ნდობის ინტერვალი პოპულაციის საშუალოსათვის:

    62 81 86 79 73 88 90 98 78 93 87 82 78 59 63 97 93 84.

23. მეტეოროლოგმა 15 ცივი ჰაერის მასივზე დაკვირვების შედეგად დაადგინა, რომ მათი გავრცელების საშუალო სიჩქარეა 18 მილი/სთ, ხოლო სტანდარტული გადახრა კი 2 მილი/სთ. ააგეთ 95%-იანი ნდობის ინტერვალი რეალური საშუალო სიჩქარისათვის.
25. 20 თინუსზე დაკვირვებამ აჩვენა, რომ ისინი საშუალოდ ცურავენ 8.6 მილს საათში. შერჩევის შესწორებული სტანდარტული გადახრაა 1.6. ააგეთ 90%-იანი ნდობის ინტერვალი ჭეშმარიტი საშუალოსათვის.
27. რეგიონის 8 სკოლის სათადარიგო მასწავლებლის ყოველდღიური ხელფასებია: 60ლ, 56ლ, 60ლ, 55ლ, 70ლ, 55ლ, 60ლ, 55ლ. იპოვეთ წერტილოვანი შეფასებები და ააგეთ 90%-იანი ნდობის ინტერვალი რეგიონში სათადარიგო მასწავლებლის ყოველდღიური ხელფასების ჭეშმარიტი საშუალოსათვის.
29. ავტომაღაზიის მენეჯერმა დაადგინა, რომ მის 6 თანამშრომელს ავტომობილის წყლის ტუმბოს შეცვლა შეუძლია საშუალოდ 18 წუთში, ხოლო შერჩევის შესწორებული სტანდარტული გადახრა არის 3 წუთი. ააგეთ 99%-იანი ნდობის ინტერვალი ჭეშმარიტი საშუალოსათვის.
31. სტრესულ სიტუაციაში მყოფი 10 კაციანი ჯგუფის გულისცემის საშუალო რიცხვი წუთში შეადგენს 126-ს, ხოლო შესწორებული სტანდარტული გადახრა არის 4. ააგეთ 95%-იანი

ნდობის ინტერვალი სტრესულ სიტუაციაში მყოფი კაცების გულისცემის რეალური საშუალოსათვის.

33. ჰოსპიტლის 24 საოპერაციო ოთახში ხმაურის დონის საშუალო იყო 41.6 დეციბალი, ხოლო შერჩევის შესწორებული სტანდარტული გადახრა კი 7.5 დეციბალი. ააგეთ 95%-იანი ნდობის ინტერვალი საოპერაციო ოთახებში ხმაურის დონის რეალური საშუალოსათვის.

35. ეროვნულ გამოცდაზე მათემატიკაში 20 აბიტურიენტის გულისცემის საშუალო იყო 96 დარტყმა წუთში, ხოლო შერჩევის შესწორებული სტანდარტული გადახრა კი 5. ააგეთ 95%-იანი ნდობის ინტერვალი გულისცემის რეალური საშუალოსათვის.

თავი II

ნდობის ინტერვალი ბერნულის სქემაში

ნდობის ინტერვალი და შერჩევის მოცულობა ბერნულის სქემისათვის (პროპორციისათვის): ბერნულის სქემაში უცნობი p **ალბათობის (პოპულაციის პროპორციის) წერტილოვანი შეფასებაა** ფარდობითი სიხშირე $w_n = S_n / n$, სადაც $S_n = \sum_{i=1}^{n} X_i$ ($X_i = 1$, თუ i-ურ ცდაში მოხდა წარმატება, და $X_i = 0$, თუ i-ურ ცდაში მოხდა მარცხი) წარმატებათა რაოდენობაა n დამოუკიდებელ ცდაში.

α **მნიშვნელოვნების დონის მქონე ასიმპტოტური ნდობის ინტერვალს უცნობი** p **ალბათობისათვის (პროპორციისათვის) აქვს შემდეგი სახე:**

$$(w_n - z_{\alpha/2} \cdot \sqrt{\frac{w_n(1-w_n)}{n}},\ w_n + z_{\alpha/2} \cdot \sqrt{\frac{w_n(1-w_n)}{n}}) .$$

პოპულაციის პროპორციის l **სიზუსტის ინტერვალური შეფასებისათვის საჭირო შერჩევის მინიმალური მოცულობაა:**

$$n^* = [w_n(1-w_n)(\frac{z_{\alpha/2}}{l})^2] + 1$$

(ამასთანავე, როცა w_n-ს აპროქსიმაცია უცნობია, მაშინ უნდა გამოვიყენოთ $w_n = 0.5$, ვინაიდან ნამრავლის $w_n(1-w_n)$ მაქსიმალურია, როცა $w_n = 0.5$).

დაზუსტებული α **მნიშვნელოვნების დონის მქონე ასიმპტოტური ნდობის ინტერვალი უცნობი** p **ალბათობისათვის არის:**

$$(\frac{w_n + \frac{z_{\alpha/2}^2}{2n} - z_{\alpha/2}\sqrt{\frac{w_n(1-w_n)}{n} + \frac{z_{\alpha/2}^2}{4n^2}}}{1 + \frac{z_{\alpha/2}^2}{n}},\ \frac{w_n + \frac{z_{\alpha/2}^2}{2n} + z_{\alpha/2}\sqrt{\frac{w_n(1-w_n)}{n} + \frac{z_{\alpha/2}^2}{4n^2}}}{1 + \frac{z_{\alpha/2}^2}{n}}) .$$

α მნიშვნელოვნების დონის მქონე ასიმპტოტური ნდობის ინტერვალი პუასონის პოპულაციის უცნობი λ პარამეტრისათვის არის:

$$(\overline{X}_n - z_{\alpha/2} \cdot \sqrt{\frac{\overline{X}_n}{n}},\ \overline{X}_n + z_{\alpha/2} \cdot \sqrt{\frac{\overline{X}_n}{n}}).$$

მაგალითი 1. მოხუცების 500 მომვლელისაგან შემდგარ შერჩევაში 60 მამაკაცია. იპოვეთ 90%-იანი საიმედოობის ნდობის ინტერვალი მოხუცების მოვლის პროგრამაში მონაწილე მამაკაცთა ჭეშმარიტი პროპორციისათვის.

ამოხსნა. ვინაიდან $\alpha = 1 - 0.90 = 0.10$, ამიტომ $z_{\alpha/2} = z_{0.05} = 1.65$. გარდა ამისა, $S_n = 60$, $n = 500$. შესაბამისად, $w_n = 60/500 = 0.12$ და $1 - w_n = 1 - 0.12 = 0.88$.

ამიტომ ნდობის ინტერვალი იქნება:

$$(0.12 - 1.65 \cdot \sqrt{\frac{0.12 \times 0.88}{500}}, 0.12 + 1.65 \cdot \sqrt{\frac{0.12 \times 0.88}{500}}),\ (0.096, 0.144)$$

ანუ

$$9.6\% < p < 14.4\%.$$

ამრიგად, მოხუცების მოვლის პროგრამაში მონაწილე მამაკაცების რეალური პროპორცია 90%-იანი საიმედოობით მოთავსებულია 9.6%-სა და 14.4%-ს შორის.

მაგალითი 3. მკვლევარს სურს 90%-იანი საიმედოობითა და 5%-იანი სიზუსტით შეაფასოს აღმასრულებელი ხელისუფლების იმ წარმომადგენელთა პროპორცია, რომლებსაც აქვთ მანქანის ტელეფონი.

ამოხსნა. ვინაიდან უცნობია w_n-ის შეფასება, სტატისტიკოსი ირჩევს $w_n = 0.5$ და $1 - w_n = 0.5$ მნიშვნელობებს.

გარდა ამისა, ამ შემთხვევაში $l = 0.05$, $\alpha = 1 - 0.90 = 0.1$, $z_{\alpha/2} = z_{0.05} = 1.65$. შესაბამისად,

$$n = [0.5 \times 0.5 \times (\frac{1.65}{0.05})^2] + 1 = [272.25] + 1 = 273.$$

ე. ი. უნდა გამოიკითხოს აღმასრულებელი ხელისუფლების 273 წარმომადგენელი.

ამოცანები

1. 1500 გამოკითხული ახალგაზრდიდან 39%-მა თქვა, რომ ისინი აპირებენ მომავალ წელს აიღონ უფრო ხანგრძლივი შვებულება. ააგეთ 95%-იანი ნდობის ინტერვალი იმ ახალგაზრდების რეალური პროპორციისათვის, რომლებიც აპირებენ მომავალ წელს უფრო ხანგრძლივი შვებულების აღებას.
3. 100 გამოკითხულ უმუშევარს შორის 65% არაა დაინტერესებული დაბრუნდეს ძველ სამსახურში. ააგეთ 95%-იანი ნდობის ინტერვალი იმ უმუშევრების რეალური პროპორციისათვის, რომლებსაც არ სურთ ძველ სამსახურში დაბრუნება.
5. 200 გამოკითხული მუშიდან 168 მუშამ განაცხადა, რომ 1 საათის განმავლობაში სატელეფონო ზარმა არანაკლებ 3-ჯერ გააწყვეტინა მუშაობა. ააგეთ 95%-იანი ნდობის ინტერვალი იმ მუშების რეალური პროპორციისათვის, რომლებსაც სატელეფონო ზარმა არანაკლებ 3-ჯერ გააწყვეტინა მუშაობა.
7. ჯანდაცვის სამინისტროს მონაცემებით 13-14 წლის მოზარდებიდან ყოველი მეხუთე ზოგჯერ ეწევა სიგარეტს. მთელი ქვეყნის მასშტაბით 13-14 წლის მოზარდებში სიგარეტის მწეველების პროპორციასთან შესადარებლად განათლების სამინისტრომ გამოკითხა 200 13-14 წლის სკოლის მოსწავლე და დაადგინა, რომ მათი 23% ზოგჯერ ეწევა სიგარეტს. ააგეთ 99%-იანი ნდობის ინტერვალი სკოლის 13-14 წლის მოსწავლეებში სიგარეტის მწეველთა რეალური პროპორციისა და შეადარეთ ის ჯანდაცვის სამინისტროს მონაცემებს.
9. 90 გამოკითხული ოჯახიდან 40 ფლობს ერთ იარაღს მაინც. ააგეთ 95%-იანი ნდობის ინტერვალი იმ ოჯახების რეალური პროპორციისათვის, რომლებიც ფლობენ ერთ იარაღს მაინც.
11. გარკვეული ასაკობრივი ჯგუფის 100 გარდაცვლილი ადამიანის კვლევამ აჩვენა, რომ მათგან 25% გარდაიცვალა კიბო-

თი. ააგეთ 98%-იანი ნდობის ინტერვალი ამ ასაკობრივი ჯგუფის გარდაცვლილ ადამიანებში იმ ადამიანების რეალური პროპორციისათვის, რომლებიც დაიღუპნენ კიბოთი.

13. მედიკოსს სურს 99%-იანი საიმედოობითა და 2%-ის სიზუსტით შეაფასოს იმ ქალების რეალური პროპორცია, რომლებიც ღებულობენ ვიტამინს. ადრე ჩატარებული კვლევის თანახმად 180 ქალიდან 25% ღებულობდა ვიტამინს. ა) რა მოცულობის უნდა იყოს საჭირო შერჩევა? ბ) თუ არ იქნებოდა ხელმისაწვდომი შერჩევითი პროპორცია, მაშინ რა მოცულობის შერჩევის აღება მოგვიწევდა?

15. მკვლევარს სურს 90%-იანი საიმედოობითა და 5%-ის სიზუსტით შეაფასოს იმ მამაკაცების რელური პროპორცია, რომელთა სიმაღლე 5 ფუტსა (1 ფუტი = 30.48 სმ) და 5 დიუმზე (1 დიუმი = 2.54 სმ) ნაკლებია. ა) რა მოცულობის შერჩევაა ამისათვის აუცილებელი თუ ცნობილია, რომ 300 მამაკაციდან 30-ის სიმაღლე 5 ფუტსა და 5 დიუმზე ნაკლებია; ბ) რა მოცულობის შერჩევა უნდა ავიღოთ იმ შემთხვევაში, როცა არ იქნება ხელმისაწვდომი შერჩევითი პროპორცია?

17. იპოვეთ საიმედოობის ხარისხი, თუ შერჩევის მოცულობაა 600 და მკვლევარს სურს, რომ რეალური პროპორციის შეფასების მაქსიმალური ცდომილება იყოს 4%. ამასთანავე, ცნობილია, რომ ადრინდელი კვლევისას პროპორცია იყო 50%.

თავი III
ნდობის ინტერვალი დისპერსიისათვის

α მნიშვნელოვნების დონის მქონე ნდობის ინტერვალი ნორმალური პოპულაციის დისპერსიისათვის ცნობილი საშუალოს შემთხვევაში არის:

$$\left(\frac{\sum_{i=1}^{n}(x_i-a)^2}{\chi^2_{n,\alpha/2}}, \frac{\sum_{i=1}^{n}(x_i-a)^2}{\chi^2_{n,1-\alpha/2}} \right),$$

ხოლო უცნობი საშუალოს შემთხვევაში კი:

$$\left(\frac{(n-1)s'^2}{\chi^2_{n-1,\alpha/2}}, \frac{(n-1)s'^2}{\chi^2_{n-1,1-\alpha/2}} \right).$$

ზოგჯერ გამოიყენება აღნიშვნები: $\chi^2_{\text{მარჯვ.}} := \chi^2_{n-1,\alpha/2}$ და $\chi^2_{\text{მარცხ.}} := \chi^2_{n-1,1-\alpha/2}$.

α მნიშვნელოვნების დონის მქონე ნდობის ინტერვალი ნორმალური პოპულაციის სტანდარტული გადახრისათვის შერჩევის დიდი მოცულობის შემთხვევაში დაახლოებით არის:

$$(s' - z_{\alpha/2} \cdot \frac{s'}{\sqrt{2n}}, s' + z_{\alpha/2} \cdot \frac{s'}{\sqrt{2n}}).$$

შერჩევითი დისპერსია ეწოდება სიდიდეს:

$$s^2 = \sum_{i=1}^{k}\left(x_i-\overline{x}\right)^2 \frac{n_i}{n} = \frac{1}{n}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2 \ \ (\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \ ).$$

შესწორებული შერჩევითი დისპერსია ეწოდება სიდიდეს:

$$s'^2 = \frac{n}{n-1}s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 .$$

შერჩევითი სტანდარტული გადახრა (შესაბამისად, **შესწორებული შერჩევითი სტანდარტული გადახრა**) ეწოდება არითმეტიკულ კვადრატულ ფესვს შერჩევითი დისპერსიიდან $s = \sqrt{s^2}$ (შესაბამისად, შესწორებული შერჩევითი დისპერსიიდან $s' = \sqrt{s'^2}$).

მაგალითი 1. იპოვეთ თამბაქოში ნიკოტინის შემცველლობის დისპერსიისა და სტანდარტული გადახრისათვის 95%-იანი ნდობის ინტერვალი თუ ცნობილია, რომ 20 სიგარეტისაგან შედგენილ შერჩევაში ნიკოტინის შემცველლობის შესწორებული სტანდარტული გადახრა 1.6 მილიგრამის ტოლია.

ამოხსნა. ვინაიდან ამ შემთხვევაში $\alpha = 1 - 0.95 = 0.05$, ამიტომ ხი-კვადრატ განაწილების ორი კრიტიკული მნიშვნელობა, რომელიც შეესაბამება $\alpha/2 = 0.025$-სა და $1-\alpha/2 = 0.975$-ს და თავისუფლების ხარისხის რიცხვს, რომელიც ტოლია $n-1 = 20-1 = 19$, შესაბამისად იქნება: $\chi^2_{19,0.025} = 32.852$ და $\chi^2_{19,0.975} = 8.907$. ამიტომ საძიებელი ნდობის ინტერვალი იქნება:

$$(\frac{(20-1)\cdot(1.6)^2}{32.852}, \frac{(20-1)\cdot(1.6)^2}{8.907}), \text{ ანუ } (1.5, 5.5).$$

ამრიგად, ჩვენ შეგვიძლია 95%-იანი საიმედოობით ვიგულისხმოთ, რომ სიგარეტში ნიკოტინის შემცველლობის რეალური დისპერსია მოთავსებულია 1.5 მილიგრამ^2-სა და 5.5 მილიგრამ^2-ს შორის.

ნდობის ინტერვალი სტანდარტული გადახრისათვის იქნება:

$$(\sqrt{1.5}, \sqrt{5.5}), \text{ ანუ } (1.2, 2.3).$$

მაშასადამე, სიგარეტში ნიკოტინის შემცველლობის რეალური სტანდარტული გადახრის 95%-იანი ნდობის ინტერვალია (1.2 მგ, 2.3 მგ).

ამოცანები

1. ცხრილის საშუალებით იპოვეთ $\chi^2_{\text{მარცხ.}}$ და $\chi^2_{\text{მარჯვ.}}$, თუ: ა) $\alpha = 0.05$, $n = 16$; ბ) $\alpha = 0.1, n = 5$; გ) $\alpha = 0.01, n = 23$; დ) $\alpha = 0.05$, $n = 29$; ე) $\alpha = 0.1, n = 14$.

ქვემოთ მოყვანილ ამოცანებში იგულისხმეთ, რომ ყველა სიდიდე განილებულია ნორმალურად:

3. ააგეთ 90%-იანი ნდობის ინტერვალი ავტობუსის უსაფრთხოების შესამოწმებლად საჭირო დროის დისპერსიისა და სტანდარტული გადახრისათვის თუ ცნობილია, რომ 20 ავტობუსის უსაფრთხოების შესამოწმებლად საჭირო დროის შესწორებული სტანდარტული გადახრაა 6.8 წთ.
5. ქვემოთ მოყვანილია ვაშლის წვენში შაქრის შემცველობა გრამებში:

18.6	19.5	20.2	20.4	19.3
21	20.3	19.6	20.7	18.9
22.1	19.7	20.8	18.9	20.7
21.6	19.5	20.1	20.3	19.9

ააგეთ 99%-იანი ნდობის ინტერვალი პოპულაციის დისპერსიისა და სტანდარტული გადახრისათვის.

7. ააგეთ 98%-იანი ნდობის ინტერვალი დროის იმ შუალედის დისპერსიისა და სტანდარტული გადახრისთვის, რომელიც სჭირდება სატელეფონო კომპანიას ზარის სწორი მიმართულებით გადართვისათვის თუ ცნობილია, რომ 15 ზარის შემთხვევაში ამ დროის შესწორებული სტანდარტული გადახრა ტოლია 1.6 წთ.
9. ავტოტექმომსახურების ცენტრი აკეთებდა რეკლამას, რომ ზეთის შეცვლისას კლიენტს არ დასჭირდებოდა 30 წუთზე მეტი ლოდინი. ზეთის შეცვლის 28 შემთხვევაში დროის შესწორებული სტანდარტული გადხრა იყო 5.2 წუთი. ააგეთ 95%-იანი ინტერვალი პოპულაციის სტანდარტული გადახრისათვის.
11. ააგეთ 95%-იანი ნდობის ინტერვალი ელემენტების მუშაობის ხანგრძლივობის სტანდარტული გადახრისათვის თუ

ცნობილია, რომ 200 ელემენტისათვის შესწორებული სტან-დარტული გადახრაა 18 თვე.

ამოცანები გამეორებისათვის

13. პარკში მოსეირნე 36 ადამიანზე დაკვირვებამ აჩვენა, რომ ისინი საშუალოდ გადაადგილდებიან 2.6 კმ/სთ სიჩქარით. შერჩევის შესწორებული სტანდარტული გადახრაა 0.4 კმ/სთ. ააგეთ 95%-იანი ნდობის ინტერვალი რეალური საშუალოსათვის.
15. 1500 გამოკითხული ადამიანიდან შვებულების განმავლობაში საშუალოდ 7.5 ღამე გაატარა სასტუმროში, შესაბამისი შესწორებული სტანდარტული გადახრა იყო 0.8 ღამე. ააგეთ 95%-იანი ნდობის ინტერვალი რეალური საშუალოსათვის.
17. ქალაქის გარკვეულ ნაწილში 5 თვიანი დაკვირვების შედეგად თვეში საშუალოდ 28 ფოსტალიონი იყო დაკბენილი მაწანწალა ძაღლების მიერ. შერჩევის შესწორებული სტანდარტული გადახრაა 3. ჩათვალეთ, რომ დაკბენილი ფოსტალიონების რიცხვი ნორმალურადაა განაწილებული და ააგეთ 90%-იანი ნდობის ინტერვალი თვეში დაკბენილი ფოსტალიონების რეალური საშუალოსათვის.
19. მკვლევარს სურს შეაფასოს უნივერსიტეტის ყოველწლიური საშუალო დანახარჯები საფოსტო მომსახურებაზე 90%-იანი საიმედოობითა და 25 ლარის სიზუსტით. რა მინიმალური მოცულობის შერჩევა უნდა იყოს განხილული თუ ცნობილია, რომ საფოსტო დანახარჯების სტანდარტული გადახრაა 80 ლარი?
21. 200 უბედური შემთხვევიდან, რომელიც საჭიროებს გადაუდებელ სამედიცინო დახმარებას, 40% მოხდა სახლში. ააგეთ 90%-იანი ნდობის ინტერვალი სახლში მომხდარი უბედური შემთხვევების რეალური პროპორციისათვის.
23. დიეტოლოგს სურს 95%-იანი საიმედოობითა და 2%-იანი სიზუსტით შეაფასოს იმ მოზარდების რეალური პროპორცია, რომლებიც ჭამენ ძილის წინ. რა მოცულობის შერჩევა იქნება ამისათვის აუცილებელი, თუ ადრინდელი კვლევის თანახმად 100 გამოკითხულიდან 18% ჭამდა ძილის წინ?

25. 28 ფორთოხლის შესწორებული სტანდარტული გადახრაა 0.34 სმ. ააგეთ 99%-იანი ნდობის ინტერვალი ფორთოხლის დიამეტრის ჭეშმარიტი სტანდარტული გადახრისათვის.
27. შემთხვევით შერჩეული 15 თოვლმავალი მანქანის ბატარეების მუშაობის ხანგრძლივობის (თვეებში) შესწორებული დისპერსია იყო 8.6. ააგეთ 90%-იანი ნდობის ინტერვალი რეალური დისპერსიისათვის.

ამოცანები გამოცდისათვის

29. 49 შემთხვევით შერჩეული წიგნის მოყვარული კვირის განმავლობაში წიგნის შეძენაზე საშუალოდ ხარჯავს 23.45 ლარს. შესაბამისი შესწორებული სტანდარტული გადახრაა 2.8 ლარი. ააგეთ 90%-იანი ნდობის ინტერვალი რეალური საშუალოსათვის.
31. 40 შემთხვევით შერჩეული სასკოლო ავტობუსის საშუალო წონაა 4150 ფუნტი, ხოლო შესწორებული სტანდარტული გადახრა კი 480 ფუნტი. ააგეთ 99%-იანი ნდობის ინტერვალი ავტობუსის წონის რეალური საშუალოსათვის.
33. რესპუბლიკურ საავადმყოფოში, შემთხვევით შერჩეული 8 კვირაზე დაკვირვებისას, კვირაში საშუალოდ 438 პაციენტი მიმართავდა სასწრაფო დახმარების განყოფილებას. შესწორებული სტანდარტული გადახრა იყო 16. ჩათვალეთ, რომ აღნიშნული კატეგორიის პაციენტთა რიცხვი განაწილებულია ნორმალურად და ააგეთ 99%-იანი ნდობის ინტერვალი პაციენტთა რეალური საშუალოსათვის.
35. ფაკულტეტის დეკანს სურს შეაფასოს საათების რაოდენობა კვირის მანძილზე, რომელსაც პირველკურსელები უთმობენ სწავლას. წინა კვლევის სტანდარტული გადახრა იყო 2.6 სთ. რა მინიმალური მოცულობის შერჩევა უნდა იყოს განხილული, რათა 99%-იანი საიმედოობით რეალური საშუალოს განსხვავება შერჩევითი საშუალოსაგან არ არემატებოდეს 0.5 საათს.
37. 75 გამოკითხული მუშიდან 53 სამსახურში ყოველდღიურად დადის ავტობუსით. ააგეთ 95%-იანი ნდობის ინტერვალი იმ მუშების რეალური პროპორციისათვის, ვინც სამსახურში დადის ავტობუსით.

39. 90 გამოკითხული ოჯახიდან 40 ფლობს ერთ ტელევიზორს მაინც. ააგეთ 95%-იანი ნდობის ინტერვალი იმ ოჯახების რეალური პროპორციისათვის, რომლებიც ფლობენ ერთ ტელევიზორს მაინც.
41. 25 შემთხვევით შერჩეული ნოველის გვერდების რიცხვის შესწორებული სტანდარტული გადახრაა 9 გვერდი. ააგეთ 95%-იანი ნდობის ინტერვალი პოპულაციის სტანდარტული გადახრისათვის.
43. 20 შემთხვევით შერჩეული ავტომობილის მიერ 1 გალონი (1 გალონი = 3.38 ლიტრი)ბენზინის გამოყენების შედეგად გამოყოფილი მავნე ნივთიერებების რაოდენობის შესწორებული სტანდარტული გადახრაა 2.3 უნცია. ააგეთ 90%-იანი ნდობის ინტერვალი ავტომობილების მიერ გამოყოფილი მავნე ნივთიერებების სტანდარტული გადახრისათვის.

თავი IV

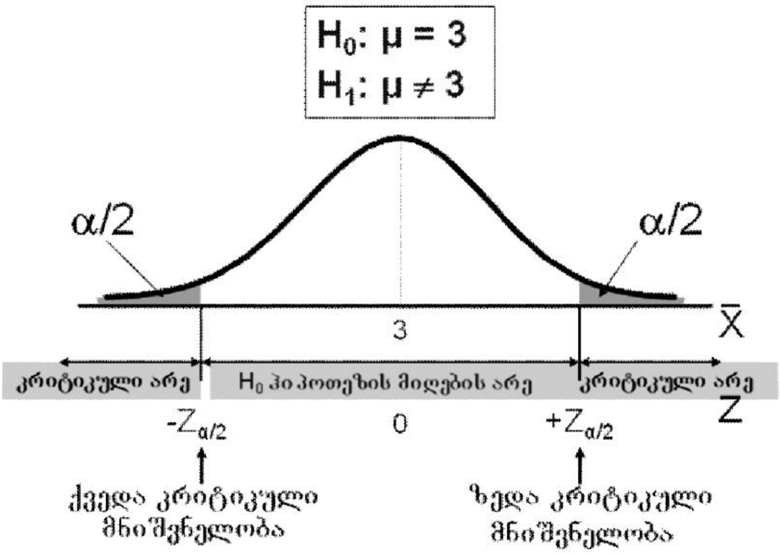
ჰიპოთეზათა შემოწმება საშუალოსათვის (σ^2 ცნობილია)

განაწილების პარამეტრების მნიშვნელობების შესახებ ან ორი განაწილების პარამეტრების სიდიდეების შედარების ჰიპოთეზებს **პარამეტრული ჰიპოთეზები** ეწოდება. ჰიპოთეზებს განაწილების სახის შესახებ კი **არაპარამეტრული ჰიპოთეზები** ეწოდება. ჰიპოთეზას, რომელიც წამოყენებულია შერჩევით მონაცემებთან მისი თანხმობის შესამოწმებლად, **ნულოვანი ჰიპოთეზა** ეწოდება და აღინიშნება H_0-ით (H_0 ამტკიცებს, რომ არ არსებობს განსხვავება პარამეტრსა და მის კონკრეტულ მნიშვნელობას შორის ან არ არსებობს განსხვავება ორ პარამეტრს შორის). H_0 ჰიპოთეზასთან ერთად იხილავენ (წამოაყენებენ) **ალტერნატიულ** ანუ **საწინააღმდეგო** ჰიპოთეზასაც, რომელსაც H_1-ით აღნიშნავენ (H_1 ამტკიცებს, რომ არსებობს განსხვავება პარამეტრსა და მის კონკრეტულ მნიშვნელობას შორის ან არსებობს განსხვავება ორ პარამეტრს შორის). ნულოვანი ჰიპოთეზა მოიცავს ტოლობის ნიშანს:

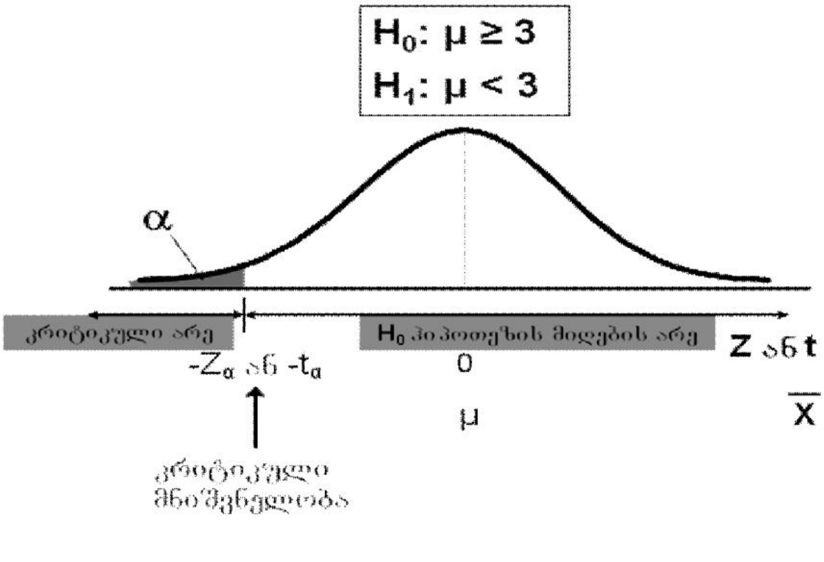
კრიტერიუმი:

ორმხრივი	მარჯვენა ცალმხრივი	მარცხენა ცალმხრივი
$H_0: \theta = \theta_0$	$H_0: \theta \le \theta_0$	$H_0: \theta \ge \theta_0$
$H_1: \theta \ne \theta_0$	$H_1: \theta > \theta_0$	$H_1: \theta < \theta_0$

ორმხრივი ჰიპოთეზა



მარცხენა ცალმხრივი ჰიპოთეზა



სამართლიანი ჰიპოთეზის უკუგდებას **პირველი გვარის შეცდომა** ეწოდება. პირველი გვარის შეცდომის დაშვების ალბათობას **მნიშვნელოვნების დონე** ეწოდება და α-თი აღინიშნება. არასამართლიანი ჰიპოთეზის მიღებას **მეორე გვარის შეცდომა** ეწოდება. მისი ალბათობა აღინიშნება β ასოთი. რიცხვს $1-\beta$,

რომელიც ტოლია ალბათობის იმისა, რომ არ იქნება დაშვებული მეორე გვარის შეცდომა **კრიტერიუმის სიმძლავრე** ეწოდება.

სტატისტიკური ჰიპოთეზის შემოწმება ნორმალური განაწილების მათემატიკური ლოდინის შესახებ ცნობილი დისპერსიის შემთხვევაში:

$\xi \cong N(\cdot,\sigma^2)$; $D\xi=\sigma^2$ ცნობილია; $E\xi$ უცნობია.

ჰიპოთეზა: $H_0: E\xi = a_0$

მნიშვნელოვნების დონე: α

კრიტერიუმის სტატისტიკა: $Z=\dfrac{\overline{X}-a_0}{\sigma/\sqrt{n}} \cong N(0,1)$

კრიტერიუმის მნიშვნელობა T.V.: $z=\dfrac{\overline{x}-a_0}{\sigma/\sqrt{n}}$

ალტერნატივა	**კრიტიკული არე C.R.** **(H_0-ის უარყოფის არე)**
$H_1: E\xi = a_1 > a_0$	$z \ge z_\alpha$,
$H_1: E\xi = a_1 < a_0$	$z \le -z_\alpha$,
$H_1: E\xi \ne a_0$	$z \le -z_{\alpha/2}$ ან $z \ge z_{\alpha/2}$

(სადაც z_α არის $N(0,1)$-ის ზედა α-კრიტიკული წერტილი ანუ **კრიტიკული მნიშვნელობა C.V.**).

P **- მნიშვნელობა:**

$$P=\begin{cases} P\{Z>z\,|\,H_0\},\ \text{თუ}\ H_1: E\xi=a_1>a_0; \\ P\{Z<z\,|\,H_0\},\ \text{თუ}\ H_1: E\xi=a_1<a_0; \\ P\{|Z|>|z|\,|\,H_0\},\ \text{თუ}\ H_1: E\xi\ne a_1. \end{cases} = \begin{cases} 1-\Phi(z),\ \text{თუ}\ H_1: E\xi=a_1>a_0; \\ \Phi(z),\ \text{თუ}\ H_1: E\xi=a_1<a_0; \\ 2\cdot[1-\Phi(|z|)],\ \text{თუ}\ H_1: E\xi\ne a_1. \end{cases}$$

გადაწყვეტილება: თუ $z\in$ **C.R.**, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

P **-მნიშვნელობის მეთოდი:** თუ $P\le\alpha$, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ α მნიშვნელოვნების დონით, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

შერჩევის მინიმალური რაოდენობა $n^{\bullet}$, რომლისთვისაც I გვარის შეცდომის ალბათობაა α, ხოლო II გვარის შეცდომის ალბათობა ნაკლებია β-ზე:

$$n^{\bullet}=[\sigma^2(z_\alpha+z_\beta)^2/(a_1-a_0)^2]+1$$

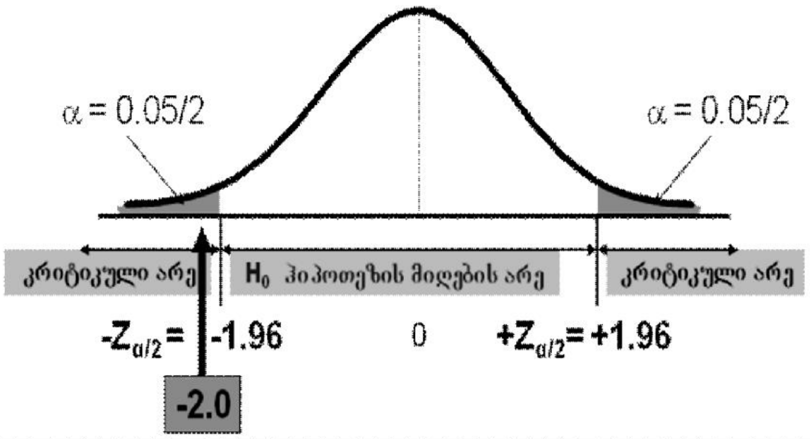
შენიშვნა: თუ შერჩევის მოცულობა მეტია ან ტოლი 30-ის და სტანდარტული გადახრა σ უცნობია, **კრიტერიუმის სტატისტიკად** განიხილება:

$$Z=\frac{\overline{X}-E\xi}{s^{'}/\sqrt{n}}\ .$$

ორმხრივი ჰიპოთეზის შემოწმება (σ ცნობილია)

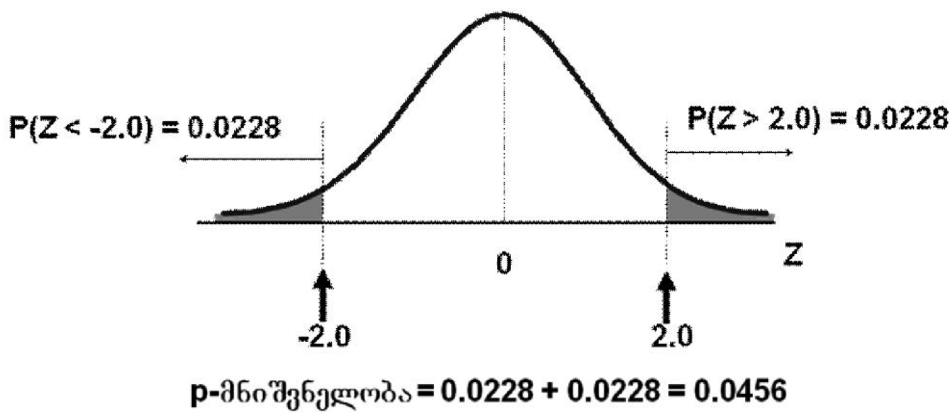
$a=0.05$ მნიშვნელოვნების დონით შევამოწმოთ ჰიპოთეზა იმის შესახებ, რომ ტელევიზორების საშუალო რიცხვი აშშ-ს მოსახლეობის ოჯახებში არის 3-ის ტოლი, თუ ცნობილია, რომ $\sigma=0.8$ და შემთხვევით შერჩეულ 100 ოჯახში ტელევიზორების საშუალო რიცხვი აღმოჩნდა 2.84.

გვაქვს: $a=0.05;\ \sigma=0.8\ n=100;\ \bar{x}=2.84$. ამიტომ $\pm z_{a/2}=\pm 1.96$ და $T.V.=\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}=\dfrac{2.84-3}{0.8/\sqrt{100}}=\dfrac{-0.16}{0.08}=-2.0$. შესაბამისად, გვექნება შემდეგი სურათი:



ვინაიდან T.V. = -2.0 < -1.96, უკუვაგდებთ ძირითად ჰიპოთეზას, ანუ საკმარისი საფუძველი გვაქვს დავასკვნათ, რომ ტელევიზორების საშუალო რიცხვი არ არის 3-ის ტოლი.

ორმხრივი ჰიპოთეზის შემოწმება *p*-მნიშვნელობის მეთოდით



ვინაიდან *p*-მნიშვნელობა = 0.0456 < a = 0.05, ამიტომ უკუვაგდებთ ძირითად ჰიპოთეზას

მაგალითი 1. სამეცნიერო ანგარიშის თანახმად სრული პროფესორის საშუალო წლიური შემოსავალი აღემატება 42000 ლარს. შემთხვევით შერჩეული 30 სრული პროფესორის წლიური შემოსავლის საშუალო აღმოჩნდა 43260 ლარი. 0.05 მნიშვნელოვნების დონისათვის შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ სრული პროფესორის საშუალო ხელფასი მეტია 42000 ლარზე, თუ პოპულაციის სტანდარტული გადახრაა 5230 ლარი.

ამოხსნა. ვინაიდან შერჩევის მოცულობა ≥ 30, ამიტომ შეგვიძლია ვიგულისხმოთ, რომ საქმე გვაქვს დაახლოებით ნორმალურ პოპულაციასთან $N(a, 5230^2)$. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0 : a \leq 42000$ და $H_1 : a > 42000$. ვინაიდან, $\alpha = 0.05$ და კრიტერიუმი მარჯვენა ცალმხრივია, კრიტიკული მნიშვნელობა იქნება $z_\alpha = 1.65$. გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$z = \frac{\bar{x} - E\xi}{\sigma / \sqrt{n}} = \frac{43260 - 42000}{5230 / \sqrt{30}} = 1.32 .$$

რადგანაც 1.32 < 1.65 (ანუ კრიტერიუმის მნიშვნელობა არ ვარდება კრიტიკულ არეში), ამიტომ არა გვაქვს საფუძველი ნულოვანი ჰიპოთეზის უარყოფის.

მაგალითი 3. ჯანდაცვის სამინისტროს ანგარიშის მიხედვით ინსულტის მკურნალობის საშუალო ღირებულება შეადგენს 24672 ლარს. იმის გასარკვევად, კონკრეტულ საავადმყოფოში ინსულტის მკურნალობის საშუალო ღირებულება არის თუ არ განსხვავებული, მკვლევარმა შემთხვევით შეარჩია ამ საავადმყოფოში ინსულტის მკურნალობის 35 შემთხვევა და მკურნალობის საშუალო ღირებულება გამოვიდა 25226 ლარი. პოპულაციის სტანდარტული გადახრაა 3251 ლარი. შეუძლია თუ არა მკვლევარს 0.01 მნიშვნელოვნების დონით ამტკიცოს, რომ ამ ჰოსპიტალში ინსულტის მკურნალობის ღირებულება განსხვავებულია 24672 ლარისაგან?

ამოხსნა. $H_0 : E\xi = 24672$, $H_1 : E\xi \neq 24672$. კრიტიკული მნიშვნელობებია: 2.58 და -2.58. კრიტერიუმის მნიშვნელობა იქნება:

$$z = \frac{\bar{x} - E\xi}{\sigma / \sqrt{n}} = \frac{25226 - 24672}{3251 / \sqrt{35}} = 1.01.$$

რამდენადაც 1.01 < 2.58 (ანუ კრიტერიუმის მნიშვნელობა არ ვარდება კრიტიკულ არეში), ამიტომ არ არსებობს საკმარისი საფუძველი იმისა, რომ კონკრეტულ საავადმყოფოში მკურნალობის ღირებულება განსხვავებულია.

მაგალითი 5. მეტეოროლოგის აზრით ქალაქში ქარის საშუალო სიჩქარეა 8კმ/სთ. შემთხვევით შერჩეული 32 დღის მონაცემებით ქარის საშუალო სიჩქარე აღმოჩნდა 8.2 კმ/სთ, ხოლო შესწორებული სტანდარტული გადახრა კი 0.6კმ/სთ. $\alpha = 0.05$ მნიშვნელოვნების დონით გვაქვს თუ არა საფუძველი არ დავეთანხმოთ მეტეოროლოგს? გამოიყენეთ P -მნიშვნელობის მეთოდი.

ამოხსნა. $H_0 : E\xi = 8$, $H_1 : E\xi \neq 8$. გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$z = \frac{8.2 - 8}{0.6 / \sqrt{32}} = 1.89.$$

ვიპოვოთ P -მნიშვნელობა: $P = 2 \cdot [1 - \Phi(|z|)] = 2 \cdot (1 - 0.9706) = = 0.0588$. ვინაიდან $P > \alpha$, ამიტომ H_0 არ უნდა უარვყოთ, ანუ

არა გვაქვს საკმარისი საფუძველი, რომ არ გავიზიაროთ მეტეოროლოგის აზრი.

ამოცანები

1. ისარგებლეთ სტანდარტული ნორმალური განაწილების ცხრილით და იპოვეთ კრიტიკული მნიშვნელობა (ან მნიშვნელობები) როცა: ა) $\alpha = 0.01$, კრიტ. ორმხრივია; ბ) $\alpha = 0.05$, კრიტ. მარჯ. ცალმხრივია; გ) $\alpha = 0.05$, კრიტ. მარც. ცალმხრივია; დ) $\alpha = 0.1$, კრიტ. მარც. ცალმხრივია; ე) $\alpha = 0.05$, კრიტ. ორმხრივია; ვ) $\alpha = 0.04$, კრიტ. მარჯ. ცალმხრივია; ზ) $\alpha = 0.01$, კრიტ. მარც. ცალმხრივია; თ) $\alpha = 0.1$, კრიტ. ორმხრივია; ი) $\alpha = 0.02$, კრიტ. მარჯ. ცალმხრივია; კ) $\alpha = 0.02$, კრიტ. ორმხრივია.
3. მკვლევარი ფიქრობს, რომ სოფლის საშუალო ბიუჯეტი შეადგენს 25000 ლარს. $\alpha = 0.05$ მნიშვნელოვნების დონით გვაქვს თუ არა საკმარისი საფუძველი უარვყოთ მკვლევრის მოსაზრება, თუ შემთხვევით შერჩეული 40 სოფლის ბიუჯეტი ათასობით ლარებში შემდეგია:

16.7	17.6	26.5	6.3	16.5	11.9	23.7	14.3	94	4.7
11.6	26.5	5.6	58.6	3.2	14.2	3.5	10.9	11.8	15.2
30.1	19.7	11.7	38.8	36.3	4.8	7.9	14.2	18	24.5
69.2	8.5	19.2	5	15.3	41	27.1	10.3	3.7	13.6

5. ქარხნის მენეჯერის აზრით მუშების საშუალო საათობრივი ანაზღაურება 9.78 ლარზე ნაკლებია. შემთხვევით შერჩეული 18 მუშის საშუალო საათობრივი ხელფასი აღმოჩნდა 9.6 ლარი, ხოლო შესწორებული სტანდარტული გადახრა კი 1.42 ლარი. ჩათვალეთ, რომ ხელფასი ნორმალურადაა განაწილებული. $\alpha = 0.1$ მნიშვნელოვნების დონით გვაქვს თუ არა საკმარისი საფუძველი დავადასტუროთ მენეჯერის მოსაზრება?

7. გაყინული კერძის მწარმოებელი ფირმის დირექტორი აცხადებს, რომ კერძის საშუალო კალორიულობა არის 800, ხოლო სტანდარტული გადახრა კი 25. მკვკლევარმა შეამოწმა 12 კერძი და დაადგინა, რომ მათი საშუალო კალორიულობა იყო 873. გვაქვს თუ არა საკმარისი საფუძველი $\alpha = 0.02$ მნიშვნე-

ლოვნების დონით უარვყოთ დირექტორის მტკიცებულობა? ჩავთვალოთ, რომ კალორიულობა კერძში განაწილებულია ნორმალურად.

9. დიეტოლოგის განცხადებით მისი დიეტით პაციენტები 20 კვირის მანძილზე საშუალოდ იკლებენ 24 ფუნტს. შესაბამისი სტანდარტული გადახრაა 5 ფუნტი. დიეტოლოგს სურს მიიღოს უკეთესი შედეგი და ამცირებს მარილის მოხმარებას. ახალი მეთოდის გამოყენებით 40 შემთხვევით შერჩეული პაციენტი 20 კვირაში საშუალოდ იკლებს 16.3 ფუნტს. $\alpha = 0.05$ მნიშვნელოვნების დონით შეიძლება თუ არა ითქვას, რომ დიეტა შეიცვალა?
11. გამოკითხვის თანახმად 55 წელზე მეტი ასაკის ქალები დღეში საშუალოდ ხარჯავენ 1660 კალორიას. იმის შესამოწმებლად, ჯანდაცვის სფეროში მომუშავე ქალები ხარჯავენ თუ არა იმავე რაოდენობის კალორიას, შემთხვევით შერჩეულ იქნა ჯანდაცვის სფეროში მომუშავე 55 წ. მეტი ასაკის 43 ქალი და აღმოჩნდა, რომ მათ მიერ დღეში დახარჯული კალორიების საშუალოა 1446, ხოლო სტანდარტული გადახრა კი 56 კალ. $\alpha = 0.05$ მნიშვნელოვნების დონით შეიძლება თუ არა იმის მტკიცება, რომ ჯანდაცვის სფეროში დასაქმებული ქალების მიერ დახარჯული კალორიების საშუალო განსხვავდება პოპულაციის საშუალოსაგან?
13. უძრავი ქონების აგენტის მტკიცებით დედაქალაქში სახლების საშუალო გასაყიდი ფასია 60000 ლარი. $\alpha = 0.05$ მნიშვნელოვნების დონით გვაქვს თუ არა საკმარისი საფუძველი უარვყოთ აგენტის მტკიცება, თუ დედაქალაქში შემთხვევით შერჩეული 36 გაყიდული სახლის ფასებია (ათასობით ლარებში):

9.5	54	99	94	80	29	121.5	184.75	15
164.45	6	13	188.4	121	308	42	7.5	32.9
126.9	25.225	95	92	38	60	211	15	28
53.5	27	21	76	85	25.225	40	97	284

15. მიუთითეთ უნდა უკუვაგდოთ თუ არა ნულოვანი ჰიპოთეზა მოცემული P-მნიშვნელობის მიხედვით, თუ: ა) P-მნიშვ. = 0.258, $\alpha = 0.05$, კრიტ. ცალმხრივია; ბ) P-მნიშვ. = 0.0684, $\alpha = 0.1$, კრიტ. ორმხრივია; გ) P-მნიშვ. = 0.0153, $\alpha = 0.01$,

კრიტ. ცალმხრივია; დ) P -მნიშვ. = 0.0232, $\alpha = 0.05$, კრიტ. ორმხრივია; ე) P -მნიშვ. = 0.002, $\alpha = 0.01$, კრიტ. ცალმხრივია.

17. ავტობუსების ფირმის მენეჯერის მტკიცებით 50 მილი/სთ სიჩქარით მოძრაობისას სასკოლო ავტობუსის საშუალო სამუხრუჭე მანძილი შეადგენს 264 ფუტს (1 ფუტი = 30.48სმ). მკვლევარის მიერ შემთხვევით შერჩეული 20 სასკოლო ავტობუსის საშუალო სამუხრუჭე მანძილი აღმოჩნდა 262.3 ფუტი, ხოლო შესწორებული სტანდარტული გადახრა კი 3 ფუტი. მკვლევარის აზრით საშუალო სამუხრუჭე მანძილი ნაკლებია 264 ფუტზე. ჩათვალეთ, რომ სამუხრუჭე მანძილი განაწილებულია ნორმალურად. იპოვეთ P -მნიშვნელობა. $\alpha = 0.01$ მნიშვნელოვნების დონით უნდა უკუვაგდოთ თუ არა ნულოვანი ჰიპოთეზა?
19. ტელევიზორების ფირმის მენეჯერის მტკიცებით მათი ტელევიზორების მუშაობის საშუალო ხანგრძლივობა მეტია 84 თვეზე. პოპულაციის სტანდარტული გადახრა არის 10 თვე. შემთხვევით შეარჩიეს 100 ტელევიზორი და შეამოწმეს. შერჩევის საშუალო აღმოჩნდა 85.1 თვე. შეამოწმეთ ჰიპოთეზა, რომ ტელევიზორების მუშაობის საშუალო ხანგრძლივობა მეტია 84 თვეზე და გამოთვალეთ P -მნიშვნელობა. მიღებული P -მნიშვნელობის მიხედვით $\alpha = 0.01$ მნიშვნელოვნების დონით უნდა უკუვაგდოთ თუ არა ნულოვანი ჰიპოთეზა?
21. გასულ წელს დედაქალაქის სამედიცინო მუშაკების საშუალო საათობრივი ანაზღაურება იყო 6.32 ლარი. პოპულაციის სტანდარტული გადახრაა 0.54 ლარი. მიმდინარე წელს შემთხვევით შერჩეული 50 სამედიცინო მუშაკის საშუალო საათობრივი ანაზრაურება შეადგენს 6.51 ლარს. P -მნიშვნელობის მიხედვით, $\alpha = 0.05$ მნიშვნელოვნების დონით, შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ საშუალო არ შეიცვალა.
23. ავტომობილისტის მტკიცებით ის დღეში საშუალოდ გადის 60 კმ-ს. ქვემოთ მოყვანილია ავტომობილისტის მიერ თვის განმავლობაში ყოველდღიურად გავლილი მანძილები. დავუშვათ, რომ პოპულაციის სტანდარტული გადახრაა 13.42 კმ. გვაქვს თუ არა საკმარისი საფუძველი უარვყოთ ავტომობილისტის მტკიცებულება $\alpha = 0.05$ მნიშვნელოვნების დონით. ისარგებლეთ P -მნიშვნელობის მეთოდით.

72	45	36	68	69	71	57	60	83	26
60	72	58	87	48	59	60	56	64	68
42	57	57	58	63	49	73	75	42	63

თავი V

ჰიპოთეზათა შემოწმება საშუალოსათვის (σ^2 უცნობია)

სტატისტიკური ჰიპოთეზის შემოწმება ნორმალური განაწილების მათემატიკური ლოდინის შესახებ უცნობი დისპერსიის შემთხვევაში:

$\xi \cong N(\cdot,\cdot)$; $D\xi$ უცნობია; $E\xi$ უცნობია.

ჰიპოთეზა: $H_0 : E\xi = a_0$

მნიშვნელოვნების დონე: α

კრიტერიუმის სტატისტიკა: $T = \dfrac{\overline{X} - a_0}{S^{'} / \sqrt{n}} \cong T(n-1)$

კრიტერიუმის მნიშვნელობა **T.V.**: $t = \dfrac{\overline{x} - a_0}{s^{'} / \sqrt{n}}$

ალტერნატივა კრიტიკული არე **C.R.** (H_0-ის უარყოფის არე)

$H_1 : E\xi = a_1 > a_0$ $\qquad t \geq t_{n-1,\alpha}$,

$H_1 : E\xi = a_1 < a_0$ $\qquad t \leq -t_{n-1,\alpha}$,

$H_1 : E\xi \neq a_0$ $\qquad t \leq -t_{n-1,\alpha/2}$ ან $t \geq t_{n-1,\alpha/2}$

(სადაც $t_{n-1,\alpha}$ არის თავისუფლების $n-1$ ხარისხის მქონე სტიუდენტის განაწილების ზედა α-კრიტიკული წერტილი ანუ **კრიტიკული მნიშვნელობა C.V.**).

P - მნიშვნელობა:

$$P = \begin{cases} P\{T > t \mid H_0\}, \text{ თუ } H_1 : E\xi = a_1 > a_0; \\ P\{T < t \mid H_0\}, \text{ თუ } H_1 : E\xi = a_1 < a_0; \\ P\{|T| > |t| \mid H_0\}, \text{ თუ } H_1 : E\xi \neq a_1. \end{cases} = \begin{cases} 1 - F_T(t), \text{ თუ } H_1 : E\xi = a_1 > a_0; \\ F_T(t), \text{ თუ } H_1 : E\xi = a_1 < a_0; \\ 2 \cdot [1 - F_T(|t|)], \text{ თუ } H_1 : E\xi \neq a_1. \end{cases}$$

გადაწყვეტილება: თუ $t \in$ **C.R.**, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

როდის გამოიყენება Z ან T განაწილება

ცნობილია თუ არა σ? — **დიახ** → გამოიყენე $Z = \dfrac{\overline{X} - EX}{\sigma / \sqrt{n}}$

↓ **არა**

არის თუ არა $n \geq 30$? — **დიახ** → გამოიყენე $Z = \dfrac{\overline{X} - EX}{S' / \sqrt{n}}$

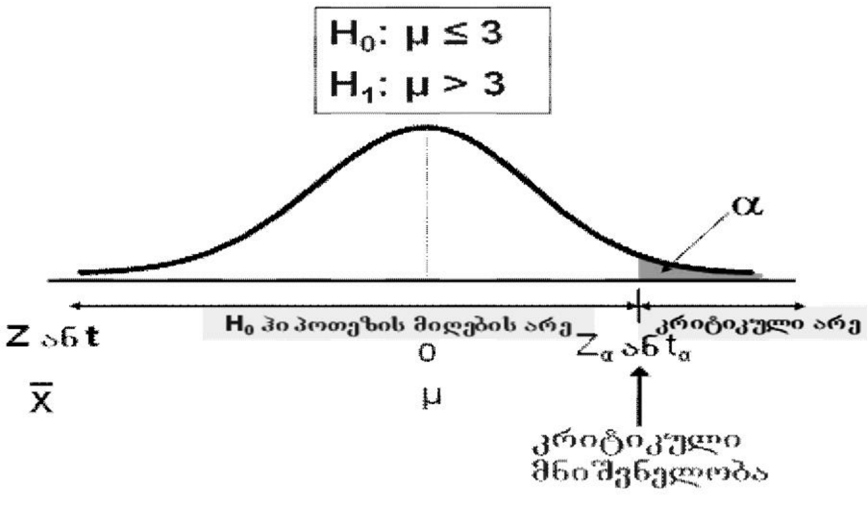
↓ **არა**

გამოიყენე $T = \dfrac{\overline{X} - EX}{S' / \sqrt{n}}$, **თ.ხ.** $= n - 1$

[*] **პოპულაცია უნდა იყოს დაახლოებით ნორმალურად განაწილებული.**

P **-მნიშვნელობის მეთოდი:** თუ $P \leq \alpha$, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

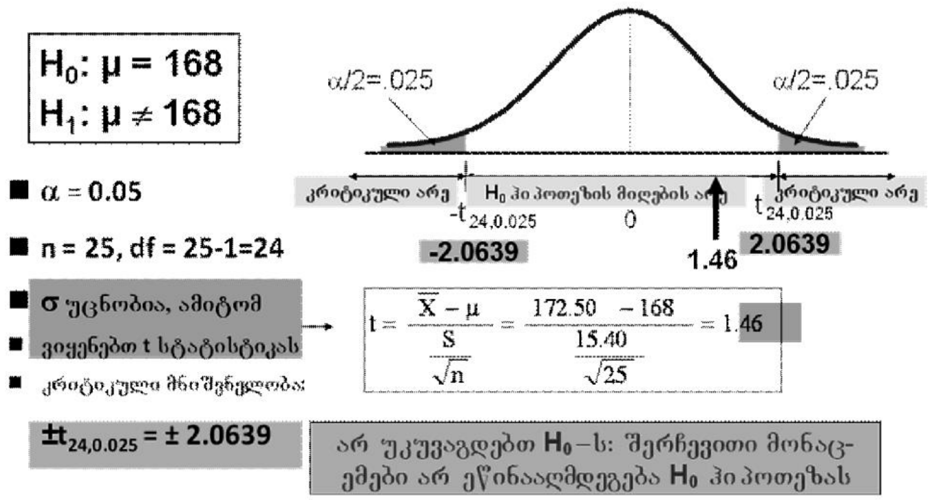
მარჯვენა ცალმხრივი ჰიპოთეზა



ორმხრივი ჰიპოთეზის შემოწმება (σ უცნობია)

$a = 0.05$ მნიშვნელოვნების დონით შევამოწმოთ ჰიპოთეზა იმის შესახებ, რომ პარიზის სასტუმროებში ოთახების საშუალო ფასია 168 ევრო, თუ ცნობილია, რომ შემთხვევით შერჩეულ 25 სასტუმროში ოთახების საშუალო ფასი აღმოჩნდა 172.50 ევრო, ხოლო შესწორებული სტანდარტული გადახრა კი 15.40 ევრო.

გვაქვს: $a = 0.05$; $n = 25$; $df = 24$; $\bar{x} = 172.50$; $s' = 15.40$. ამიტომ $\pm t_{24,0.025} = \pm 2.0639$ და $T.V. = \dfrac{\bar{x} - \mu}{s' / \sqrt{n}} = \dfrac{172.50 - 168}{15.40 / \sqrt{25}} = \dfrac{4.5}{3.08} = 1.46$. შესაბამისად, გვექნება შემდეგი სურათი:



მარჯვენა ცალმხრივი *t* კრიტერიუმი საშუალოსათვის (σ უცნობია)

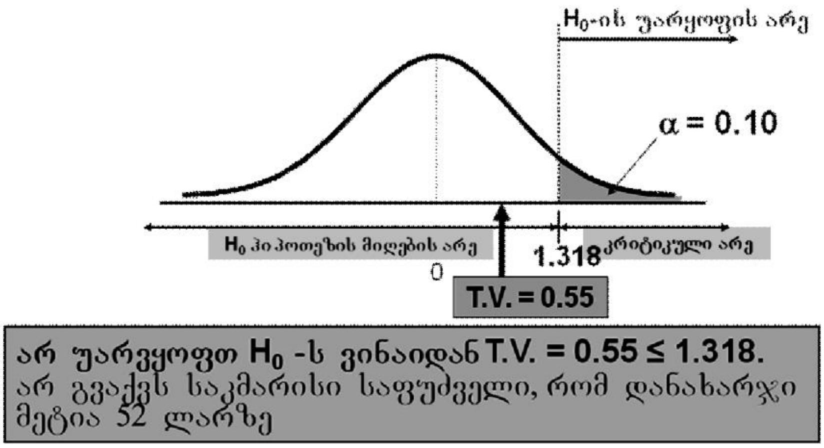
სატელეფონო კომპანიის მენეჯერის აზრით საქმიანი ადამიანების მობილური ტელეფონის დანახარჯები გაიზარდა და ახლა საშუალოდ მეტია ვიდრე 52 ლარი თვეში. კომპანიამ შეარჩია 25 მომხმარებელი, რომელთა საშუალო დანახარჯმა შეადგინა 53.1 ლარი, ხოლო შესწორებული სტანდარტული გადახრა იყო 10 ლარი. 0.1 მნიშვნელოვნების დონით შეამოწმეთ მენეჯერის მტკიცებულების სისწორე (იგულისხმება, რომ პოპულაცია ნორმალურია).

ამ შემთხვევაში გვაქვს: $n = 25$; $\bar{x} = 53.1$; $s' = 10$; $a = 0.1$. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები:

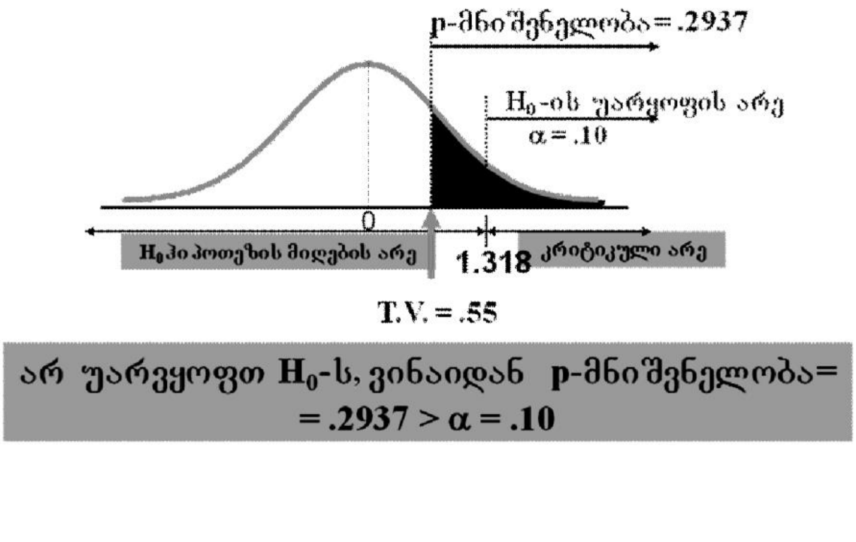
$H_0: \mu \le 52$ საშუალო დანახარჯი არ აღემატება 52 ლარს;

$H_1: \mu > 52$ საშუალო დანახარჯი მეტია 52 ლარზე.

შესაბამისად, კრიტერიუმი იქნება მარჯვენა ცალმხრივი: $(t_{a,n-1}, +\infty) = (t_{0.1,24}, +\infty) = (1.318, +\infty)$. კრიტერიუმის სტატისტიკად უნდა ავიღოთ t სტატისტიკა, $t = \dfrac{\bar{x} - \mu}{s' / \sqrt{n}}$. გამოვთვალოთ კრიტერიუმის მნიშვნელობა $T.V. = \dfrac{53.1 - 52}{10 / \sqrt{25}} = 0.55$.



იგივე ამოცანა ამოვხსნათ *p*-მნიშვნელობის მეთოდით, გვაქვს:



მაგალითი 1. ჯანდაცვის მინისტრის მტკიცებით ექიმის საშუალო წლიური შემოსავალი შეადგენს 24000 ლარს. შემთხვევით შერჩეული 10 ექიმის წლიური შემოსავლის საშუალო აღმოჩნდა 23450 ლარი, ხოლო შესწორებული სტანდარტული გადახრა კი 400 ლარი. გვაქვს თუ არა საკმარისი საფუძველი მინისტრის მტკიცების უარსაყოფად $\alpha = 0.05$ მნიშვნელოვნების დონით?

ამოხსნა. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0 : E\xi = 24000$, $H_1 : E\xi \neq 24000$. პირობის თანახმად, გვაქვს: $\bar{x} = 23450$, $s' = 400$, $n = 10$. სტიუდენტის განაწილების ცხრილიდან, $\alpha = 0.05$-ისა და თავისუფლების $n-1=9$ ხარისხისათვის, ვპოულობთ საჭირო კრიტიკულ მნიშვნელობებს: $-t_{n-1,\alpha/2} = -2.262$ და $t_{n-1,\alpha/2} = 2.262$. კრიტერიუმის მნიშვნელობა კი იქნება:

$$t = \frac{\bar{x} - E\xi}{s' / \sqrt{n}} = \frac{23450 - 24000}{400 / \sqrt{10}} = -4.35 .$$

ვინაიდან, $-4.35 < -2.262$, ამიტომ ნულოვანი ჰიპოთეზა უნდა უკუვაგდოთ, ანუ ჩვენ გვაქვს საკმარისი საფუძველი მინისტრის მტკიცების უარსაყოფად.

მაგალითი 3. ვიპოვოთ P-მნიშვნელობა, თუ კრიტერიუმის მნიშვნელობა სტიუდენტის განაწილებისათვის (T კრიტერიუმის სტატისტიკის t კრიტერიუმის მნიშვნელობა) არის 2.056, შერჩევის მოცულობაა 11 და კრიტერიუმი მარჯვენა ცალმხრივია.

ამოხსნა. სტიუდენტის განაწილების ზედა კრიტიკული წერტილების ცხრილში თავისუფლების $11-1=10$-ის ტოლი ხარისხის გასწვრივ მოვძებნოთ ისეთი ორი მნიშვნელობა, რომელთა შორის ვარდება (მოექცევა) 2.056. ეს მნიშვნელობებია: 1.812 და 2.228. ვინაიდან კრიტერიუმი მარჯვენა ცალმხრივია, შევხედოთ სტრიქონს წარწერით „One tail, α“ და ვიპოვოთ α-ს ორი მნიშვნელობა, რომელიც შეესაბამება 1.812-სა და 2.228-ს. ეს მნიშვნელობებია: 0.05 და 0.025. შესაბამისად, P-მნიშვნელობა სწორედ მათ შორისაა მოთავსებული: 0.025 < P-მნიშვნელობა < 0.05. მაგალითად, თუ $\alpha = 0.05$, მაშინ ჩვენ უნდა უკუვაგდოთ H_0, რადგანაც P-მნიშვნელობა < 0.05; მაგრამ, თუ $\alpha = 0.01$, მა-

შინ H_0-ის უარყოფის საფუძველი არა გვაქვს, რადგანაც P-მნიშვნელობა > 0.025 > 0.01.

შენიშვნა: P-მნიშვნელობის პოვნა შეიძლება სპეციალური კალკულატორით ან კომპიუტერული პროგრამით. ამ მაგალითში კალკულატორით გამოთვლილი P-მნიშვნელობა = 0.033.

მაგალითი 5. ექიმების აზრით მორბენალი ადამიანი უფრო მეტ ჟანგბადს მოიხმარს ვიდრე საშუალოდ ყველა ადამიანი. შემთხვევით შერჩეული 15 მორბენალისათვის ჟანგბადის მოხმარების საშუალო იყო 40.6 მლ/კგ, ხოლო შესწორებული სტანდარტული გადახრა კი 6 მლ/კგ. თუ ყველა ადამიანის საშუალო მოხმარება შეადგენს 36.7 მლ/კგ, გვაქვს თუ არა საკმაო საფუძველი დავუჯეროთ ექიმებს $\alpha = 0.05$ მნიშვნელოვნების დონით?

ამოხსნა. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0 : E\xi \le 36.7$, $H_1 : E\xi > 36.7$. გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$t = \frac{\bar{x} - E\xi}{s' / \sqrt{n}} = \frac{40.6 - 36.7}{6 / \sqrt{15}} = 2.517.$$

გამოვთვალოთ P-მნიშვნელობა: სტიუდენტის განაწილების ცხრილში $n-1=14$-ის გასწვრივ 2.517 ვარდება 2.145-სა და 2.624-ს შორის, რომლებიც შეესაბამება $\alpha = 0.025$-სა და $\alpha = 0.01$-ს (კრიტერიუმი მარჯვენა ცალმხრივია). ამიტომ 0.01 < P-მნიშვნელობა < 0.025, ე. ი. P-მნიშვნელობა < α (ვინაიდან 0.025 < 0.05) და ნულოვანი ჰიპოთეზა უნდა უკუვაგდოთ, ანუ ჩვენ გვაქვს საკმარისი საფუძველი გავიზიაროთ ექიმების მტკიცებულება.

ამოცანები

1. ისარგებლეთ სტიუდენტის განაწილების ცხრილით და იპოვეთ კრიტიკული მნიშვნელობა (ან მნიშვნელობები), როცა: ა) $n=10$, $\alpha = 0.05$, კრიტერიუმი მარჯვენა ცალმხრივია; ბ) $n=18$, $\alpha = 0.1$, კრიტერიუმი ორმხრივია; გ) $n=6$, $\alpha = 0.01$,

კრიტერიუმი მარცხენა ცალმხრივია; დ) $n=9$, $\alpha=0.025$, კრიტერიუმი მარჯვენა ცალმხრივია; ე) $n=15$, $\alpha=0.05$, კრიტერიუმი ორმხრივია; ვ) $n=23$, $\alpha=0.005$, კრიტერიუმი მარცხენა ცალმხრივია; ზ) $n=28$, $\alpha=0.01$, კრიტერიუმი ორმხრივია; თ) $n=17$, $\alpha=0.02$, კრიტერიუმი ორმხრივია.

ქვემოთ მოყვანილ ამოცანებში იგულისხმეთ, რომ პოპულაცია დაახლოებით ნორმალურადაა განაწილებული.

3. ეკვატორის სამხრეთით ზაფხულის თვეებში მოსული ნალექების საშუალო შეადგენს 11.52 დიუმს (1 დიუმი = 2.54 სმ). 2000 წელს მკვლევარმა შემთხვევით შეარჩია ეკვატორის სამხრეთით მდებარე 10 ქალაქი და დაადგინა, რომ მოსული ნალექების საშუალო არის 7.42 დიუმი. შერჩევის შესწორებული სტანდარტული გადახრა შეადგენს 1.3 დიუმს. $\alpha=0.05$ მნიშვნელოვნების დონით შეუძლია თუ არა მკვლევარს დაასკვნას, რომ 2000 წელს მოსული ნალექების საშუალო ნაკლებია 11.52 დიუმზე?
5. უძრავი ქონების აგენტის მტკიცებით დედაქალაქში მცირე ბიზნესის ოფისის ხარჯების საშუალოა 800 ლარი. შემთხვევით შერჩეული 10 მცირე ბიზნესისათვის ოფისის ხარჯების საშუალომ შეადგინა 863 ლარი, ხოლო შესწორებული სტანდარტული გადახრა იყო 20 ლარი. $\alpha=0.05$ მნიშვნელოვნების დონით გვაქვს თუ არა საკმარისი საფუძველი უკუვაგდოთ უძრავი ქონების აგენტის მტკიცებულება?
7. მშენებლობის სამინისტროს მტკიცებულებით დედაქალაქში შენობების საშუალო სიმაღლე სულ ცოტა 700 ფუტია. $\alpha=0.025$ მნიშვნელოვნების დონით გვაქვს თუ არა საკმარისი საფუძველი უკუვაგდოთ მშენებლობის სამინისტროს მტკიცებულება, თუ შემთხვევით შერჩეული 10 შენობის სიმაღლეებია:

485	511	841	725	615
582	616	635	535	520

9. უძრავი ქონების აგენტის მტკიცებით სამოთახიანი ბინის საშუალო საიჯარო გადასახადი დედაქალაქში შეადგენს 750 ლარს. მკვლევარმა შემთხვევით შეარჩია 12 სამოთახიანი ბი-

ნა და დაადგინა, რომ მათი საიჯარო გადასახადის საშუალოა 732 ლარი, ხოლო შესწორებული სტანდარტული გადახრა კი 17 ლარი. $\alpha = 0.01$ მნიშვნელოვნების დონით გვაქვს თუ არა საკმარისი საფუძველი უკუვაგდოთ აგენტის მტკიცებულება?

11. თბილგაზის მტკიცებით მცირე კომპანიების გაზის საშუალო დანახარჯი თვეში არ აღემატება 350 ლარს. კომპანიების მფლობელები კი ეჭვობენ, რომ გაზის დანახარჯი უფრო მეტია. შემთხვევით შეირჩა 12 მცირე კომპანია და აღმოჩნდა, რომ მათი საშუალო დანახარჯი იყო 358 ლარი, ხოლო შესწორებული სტანდარტული გადახრა კი 16 ლარი. შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ $\alpha = 0.05$ მნიშვნელოვნების დონით გაზის საშუალო დანახარჯი მეტია 350 ლარზე.
13. გამოკითხვის თანახმად, მარტოხელა ადამიანების სახლში ტელეფონი თვეში საშუალოდ 37-ჯერ რეკავს. ამ ჰიპოთეზის შესამოწმებლად მკვლევარმა გამოკითხა 29 მარტოხელა ადამიანი და დაადგინა, რომ ტელეფონის ზარების საშუალო იყო 34.9, ხოლო შესწორებული სტანდარტული გადახრა კი 6. $\alpha = 0.05$ მნიშვნელოვნების დონით გვაქვს თუ არა საფუძველი უკუვაგდოთ აღნიშნული ჰიპოთეზა? გამოიყენეთ P-მნიშვნელობის მეთოდით.
15. გასული წლების გამოცდილებიდან გამომდინარე მასწავლებელს სჯერა, რომ გამოცდაზე სტუდენტების საშუალო ქულა არის 75. გამოიყენეთ P-მნიშვნელობის მეთოდი და $\alpha = 0.01$ მნიშვნელოვნების დონით შეამოწმეთ, რომ სტუდენტების საშუალო ქულა ისევ არის 75, თუ ცნობილია, რომ მიმდინარე წელს 20 შემთხვევით შერჩეული სტუდენტის ქულებია:

80	68	72	73	76	81	71	71	50	65
63	71	70	76	75	69	70	72	70	74

თავი VI

ჰიპოთეზათა შემოწმება დისპერსიისათვის

სტატისტიკური ჰიპოთეზის შემოწმება ნორმალური განაწილების დისპერსიის შესახებ:

ჰიპოთეზა: $H_0 : D\xi = \sigma_0^2$

მნიშვნელოვნების დონე: α

კრიტერიუმის სტატისტიკა: $$\begin{cases} \dfrac{\sum\limits_{i=1}^{n}(X_i - E\xi)^2}{\sigma_0^2} \cong \chi^2(n), \text{ თუ } E\xi \text{ ცნობილია,} \\ \dfrac{(n-1)S'^2}{\sigma_0^2} \cong \chi^2(n-1), \text{ თუ } E\xi \text{ უცნობია.} \end{cases}$$

კრიტერიუმის მნიშვნელობა **T.V.=** $$\begin{cases} \dfrac{\sum\limits_{i=1}^{n}(x_i - E\xi)^2}{\sigma_0^2}, \text{ თუ } E\xi \text{ ცნობილია,} \\ \dfrac{(n-1)s'^2}{\sigma_0^2}, \text{ თუ } E\xi \text{ უცნობია.} \end{cases}$$

ალტერნატივა	კრიტიკული არე **C.R.** **(**H_0-ის უარყოფის არე**)**
$H_1 : D\xi > \sigma_0^2$	$\begin{cases} [\chi^2_{n,\alpha}, +\infty), \text{ თუ } E\xi \text{ ცნობილია,} \\ [\chi^2_{n-1,\alpha}, +\infty), \text{ თუ } E\xi \text{ უცნობია.} \end{cases}$
$H_1 : D\xi < \sigma_0^2$	$\begin{cases} (0, \chi^2_{n,1-\alpha}], \text{ თუ } E\xi \text{ ცნობილია,} \\ (0, \chi^2_{n-1,1-\alpha}], \text{ თუ } E\xi \text{ უცნობია.} \end{cases}$
$H_1 : D\xi \neq \sigma_0^2$	$\begin{cases} (0, \chi^2_{n,1-\alpha/2}] \cup [\chi^2_{n,\alpha/2}, +\infty), \text{ თუ } E\xi \text{ ცნობილია,} \\ (0, \chi^2_{n-1,1-\alpha/2}] \cup [\chi^2_{n-1,\alpha/2}, +\infty), \text{ თუ } E\xi \text{ უცნობია.} \end{cases}$

(სადაც $\chi^2_{n,\alpha}$ არის $\chi^2(n)$ განაწილების ზედა α-კრიტიკული წერტილი ანუ **კრიტიკული მნიშვნელობა C.V.**).

ალტერნატივა $\qquad$ P **- მნიშვნელობა:**

$H_1: D\xi > \sigma_0^2$ $\qquad$ $P = \begin{cases} P\{\chi^2(n) > T.V.\}, \text{ თუ } E\xi \text{ ცნობილია,} \\ P\{\chi^2(n-1) > T.V.\}, \text{ თუ } E\xi \text{ უცნობია.} \end{cases}$

$H_1: D\xi < \sigma_0^2$ $\qquad$ $P = \begin{cases} P\{\chi^2(n) < T.V.\}, \text{ თუ } E\xi \text{ ცნობილია,} \\ P\{\chi^2(n-1) < T.V.\}, \text{ თუ } E\xi \text{ უცნობია.} \end{cases}$

$H_1: D\xi \neq \sigma_0^2$

$$P = \begin{cases} 2\cdot\min(P\{\chi^2(n) > T.V.\},\ P\{\chi^2(n) < T.V.\}), \quad \text{თუ } E\xi \text{ ცნობილია,} \\ 2\cdot\min(P\{\chi^2(n-1) > T.V.\},\ P\{\chi^2(n-1) < T.V.\}), \text{თუ } E\xi \text{ უცნობია.} \end{cases}$$

გადაწყვეტილება: თუ $\mathbf{T.V.} \in \mathbf{C.R.}$, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

P **- მნიშვნელობის მეთოდი:** თუ $P \le \alpha$, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

მაგალითი 1. პროფესორს სურს თავის ჯგუფში, სადაც 23 სტუდენტია, გაარკვიოს ქულების ცვალებადობა არის თუ არა ნაკლები, ვიდრე პოპულაციის დისპერსია 225. პროფესორის ჯგუფში ქულების შესწორებული დისპერსიაა 198. $\alpha = 0.05$ მნიშვნელოვნების დონით გვაქვს თუ არა საკმარისი საფუძველი ვამტკიცოთ, რომ პროფესორის ჯგუფში ქულების ცვალებადობა ნაკლებია პოპულაციის დისპერსიაზე? ჩათვალეთ, რომ ქულები ნორმალურადაა განაწილებული.

ამოხსნა. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები:

$H_0: D\xi \ge 225$, $\quad H_1: D\xi < 225$. ვიპოვოთ კრიტიკული მნიშვნელობა: ვინაიდან კრიტერიუმი მარცხენა ცალმხრივია, $1-\alpha = 0.95$ და თავისუფლების ხარისხია $n-1 = 22$, ამიტომ ხი კვადრატ განაწილების ზედა კრიტიკული წერტილების ცხრილიდან ვპოულობთ, რომ $\chi^2_{n-1,1-\alpha} = \chi^2_{22,0.95} = 12.338$. მაშასადამე,

კრიტიკული არეა $C.R. = (0,12.338]$. გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$T.V. = \frac{(n-1)s'^2}{\sigma_0^2} = \frac{(23-1)\cdot 198}{225} = 19.36\,.$$

ვინაიდან, კრიტერიუმის მნიშვნელობა არ ვარდება კრიტიკულ არეში, ამიტომ ნულოვან ჰიპოთეზას ვღებულობთ ანუ არ არსებობს საკმარისი საფუძველი იმისათვის, რომ პროფესორის ჯგუფში ქულების ცვალებადობა ნაკლებია პოპულაციის დისპერსიაზე.

მაგალითი 3. სიგარეტის კომპანიას სურს შეამოწმოს ჰიპოთეზა, რომ მის სიგარეტში ნიკოტინის შემცველობის დიპერსია არის 0.644. ნიკოტინის შემცველობა იზომება მილიგრამებში და იგულისხმება, რომ ის ნორმალურად განაწილებულია. 20 სიგარეტისგან აღებული შერჩევის შესწორებული სტანდარტული გადახრაა 1 მილიგრამი. $\alpha = 0.05$ მნიშვნელოვნების დონით გვაქვს თუ არა საკმარისი საფუძველი უარვყოთ კომპანიის ჰიპოთეზა?

ამოხსნა. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0 : D\xi = 0.644$, $H_1 : D\xi \neq 0.644$. ამ შემთხვევაში საქმე გვაქვს ორმხრივ კრიტერიუმთან. $\alpha = 0.05$-ისა და თავისუფლების $n-1=19$ ხარისხისათვის ხი კვადრატ განაწილების ზედა კრიტიკული წერტილების ცხრილიდან ვპოულობთ, რომ კრიტიკული მნიშვნელობებია: $\chi^2_{n-1,\alpha/2} = \chi^2_{19,0.025} = 32.852$ და $\chi^2_{n-1,1-\alpha/2} = \chi^2_{19.0.975} = 8.907$. კრიტერიუმის მნიშვნელობა კი იქნება:

$$T.V. = \frac{(n-1)s'^2}{\sigma_0^2} = \frac{(20-1)\cdot(1)^2}{0.644} = 29.5\,.$$

კრიტერიუმის მნიშვნელობა მოთავსებულია ორ კრიტიკულ მნიშვნელობას შორის (8. 907 < 29.5 < 32.852) ანუ ის არ ვარდება კრიტიკულ არეში. ამიტომ H_0 არ უნდა უკუვაგდოთ. შესაბამისად, ჩვენ არა გვაქვს საკმარისი საფუძველი უარვყოთ კომპანიის მოსაზრება.

მაგალითი 5. ვიპოვოთ P-მნიშვნელობა, თუ კრიტერიუმის მნიშვნელობა ხი კვადრატ განაწილებისათვის არის 3.823, შერჩევის მოცულობაა 13 და კრიტერიუმი მარცხენა ცალმხრივია.

ამოხსნა. ხი კვადრატ განაწილების ზედა კრიტიკული წერტილების ცხრილში თავისუფლების $13-1=12$-ის ტოლი ხარისხის გასწვრივ მოვძებნოთ ისეთი ორი მნიშვნელობა, რომელთა შორის ვარდება (მოექცევა) 3.823. ეს მნიშვნელობებია: 3.571 და 4.404. შევხედოთ პირველ სტრიქონს და ვიპოვოთ α-ს ორი მნიშვნელობა, რომელიც შეესაბამება 3.571 -სა და 4.404-ს. ეს მნიშვნელობებია: 0.99 და 0.975. როდესაც მიღებული მნიშვნელობები განაწილების მარცხენა მხარესაა (> 0.5), 1-ს უნდა გამოვაკლოთ α-ს ეს მნიშვნელობები და მივიღებთ P-მნიშვნელობის საზღვრებს. ეს საზღვრებია: 1 – 0.99 =0.01 და 1 – 0.975 = 0.025. შესაბამისად, 0.01 < P-მნიშვნელობა < 0.025.

შევნიშნავთ, რომ აქ კალკულატორით გამოთვლილი P-მნიშვნელობა არის 0.014.

შენიშვნა: როდესაც ხი კვადრატ კრიტერიუმი ორმხრივია, ცალმხრივი კრიტერიუმის შესაბამისი ორივე საზღვარი უნდა გაორმაგდეს. შესაბამისად, წინა მაგალითში: $2\cdot 0.01 <$ P-მნიშვნელობა $< 2\cdot 0.025$, ანუ 0.02 < P-მნიშვნელობა < 0.05.

ამოცანები

1. ჩამოაყალიბეთ შესაძლო ძირითადი და ალტერნატიული ჰიპოთეზები და გამოთვალეთ კრიტიკული მნიშვნელობა, როცა $D\xi = 225$, თუ: ა) $\alpha = 0.05$, $n = 18$, კრიტერიუმი მარჯვენა ცალმხრივია; ბ) $\alpha = 0.1$, $n = 23$, კრიტერიუმი მარცხენა ცალმხრივია; გ) $\alpha = 0.05$, $n = 15$, კრიტერიუმი ორმხრივია; დ) $\alpha = 0.1$, $n = 8$, კრიტერიუმი ორმხრივია; ე) $\alpha = 0.01$, $n = 17$,

კრიტერიუმი მარჯვენა ცალმხრივია; ვ) $\alpha = 0.025$, $n = 20$, კრიტერიუმი მარცხენა ცალმხრივია; ზ) $\alpha = 0.01$, $n = 13$, კრიტერიუმი ორმხრივია; თ) $\alpha = 0.025$, $n = 29$, კრიტერიუმი მარცხენა ცალმხრივია.

ქვემოთ მოყვანილ ამოცანებში იგულისხმეთ, რომ სიდიდეები ნორმალურია ან დაახლოებით ნორმალურადაა განაწილებული.

3. დიეტოლოგის მტკიცებით სხვადასხვა სახის ერთ სუფრის კოვზ სიროფში კალორიების რიცხვის სტანდარტული გადახრა არის 60. $\alpha = 0.1$ მნიშვნელოვნების დონით შეიძლება თუ არ ამ მტკიცებულების უარყოფა, თუ შემთხვევით შერჩეული 18 სხვადასხვა სახის სიროფის ერთ კოვზში კალორიების რიცხვია:

53	210	100	200	100	220
210	100	240	200	100	210
100	210	100	210	100	60

5. კომპანიის მენეჯერის მტკიცებით მათ მიერ გამოშვებულ იოგურტში შაქრის შემცველობის დისპერსია არ აღემატება 25 გრამს. შემთხვევით შერჩეულ 20 იოგურტში გაზომეს შაქრის შემცველობა და აღმოჩნდა, რომ შესწორებული შერჩევითი დისპერსია ტოლია 36-ის. $\alpha = 0.1$ მნიშვნელოვნების დონით შეიძლება თუ არ ამ მტკიცებულების უარყოფა?
7. კოპანიის მენეჯერის მტკიცებით იმ სატელეფონო საუბრების ხანგრძლივობების სტანდარტული გადახრა, რომელიც მას სჭირდება ფირმის საქმეების მოსაწესრიგებლად, არ აღემატება 1.2 წუთს. მენეჯერის შემთხვევით შერჩეული 15 სატელეფონო საუბრის ხანგრძლივობის შესწორებული სტანდარტული გადახრა აღმოჩნდა 1.8 წუთი. $\alpha = 0.01$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ სტანდარტული გადახრა არ აღემატება 1.2-ს. ისარგებლეთ P -მნიშვნელობის მეთოდით.

9. ქვემოთ მოყვანილია სამშენებლო კომპანიის მიერ შესყიდული 12 ფოლადის მავთულის წინააღმდეგობების მაჩვენებლები. მავთულის პარტიის დაწუნება ხდება იმ შემთხვევაში,

როცა შესწორებული სტანდარტული გადახრა მეტია 2-ზე. $\alpha = 0.01$ მნიშვნელოვნების დონით გვაქვს თუ არა საფუძველი დავიწუნოთ მავთულების პარტია?

2001	1998	2002	2000	1998	1999
1997	2005	2003	2001	1999	2006

თავი VII

ჰიპოთეზათა შემოწმება ბერნულის სქემაში

სტატისტიკური ჰიპოთეზის შემოწმება ბერნულის სქემაში უცნობი p ალბათობის შესახებ (დიდი მოცულობის შემთხვევაში):

ჰიპოთეზა: $H_0 : p = p_0$

მნიშვნელოვნების დონე: α

კრიტერიუმის სტატისტიკა: $Z = \dfrac{\overline{X} - p_0}{\sqrt{p_0(1-p_0)/n}}$ $\left(\overline{X} \equiv \dfrac{S_n}{n} \equiv w_n\right)$;

კრიტერიუმის მნიშვნელობა **T.V.:** $z = \dfrac{\overline{x} - p_0}{\sqrt{p_0(1-p_0)/n}}$;

ალტერნატივა	კრიტიკული არე **C.R.** (H_0-ის უარყოფის არე)
$H_1 : p > p_0$	$z \geq z_\alpha$,
$H_1 : p < p_0$	$z \leq -z_\alpha$,
$H_1 : p \neq p_0$	$z \leq -z_{\alpha/2}$ ან $z \geq z_{\alpha/2}$,

სადაც z_α არის $N(0,1)$-ის ზედა α-კრიტიკული წერტილი ანუ **კრიტიკული მნიშვნელობა C.V.** (აქ იგულისხმება, რომ: $np \geq 5$ და $nq \geq 5$. თუკი ამ პირობებიდან ერთი მაინც არ სრულდება, მაშინ კრიტიკული მნიშვნელობა უნდა ვიპოვოთ ბინომიალური განაწილების ცხრილიდან).

$$P \text{ - მნიშვნელობა: } P = \begin{cases} 1-\Phi(z), \text{ თუ } H_1 : p > p_0; \\ \Phi(z), \text{ თუ } H_1 : p < p_0; \\ 2\cdot[1-\Phi(|z|)], \text{ თუ } H_1 : p \neq p_0. \end{cases}$$

გადაწყვეტილება: თუ $z \in$ **C.R.**, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

P **-მნიშვნელობის მეთოდი:** თუ $P \leq \alpha$, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

შენიშვნა. მცირე მოცულობის შერჩევისათვის (კერძოდ, თუ $n < 30$) კრიტერიუმის სტატისტიკად გამოიყენება $S_n \cong Bi(n, p_0)$. მაგალითად, ცალმხრივი მარჯვენა ალტერნატივის დროს **C.R.** $= \{S_n \geq c_\alpha\}$, სადაც c_α მთელი დადებითი რიცხვი უნდა შეირჩეს პირობიდან, რომ:

$$P\{S_n \geq c_\alpha \mid H_0\} = \sum_{k=c_\alpha}^{n} C_n^k p_0^k (1-p_0)^{n-k} = \alpha .$$

თუ ასეთი c_α არ არსებობს, მაშინ $c = c_\alpha$ -ს არჩევენ პირობიდან

$$P\{S_n \geq c \mid H_0\} \leq \alpha < P\{S_n \geq c-1 \mid H_0\},$$

ან რაც იგივეა:

$$\sum_{k=c}^{n} C_n^k p_0^k (1-p_0)^{n-k} \leq \alpha < \sum_{k=c-1}^{n} C_n^k p_0^k (1-p_0)^{n-k} .$$

ჰიპოთეზათა შემოწმება პროპორციისათვის

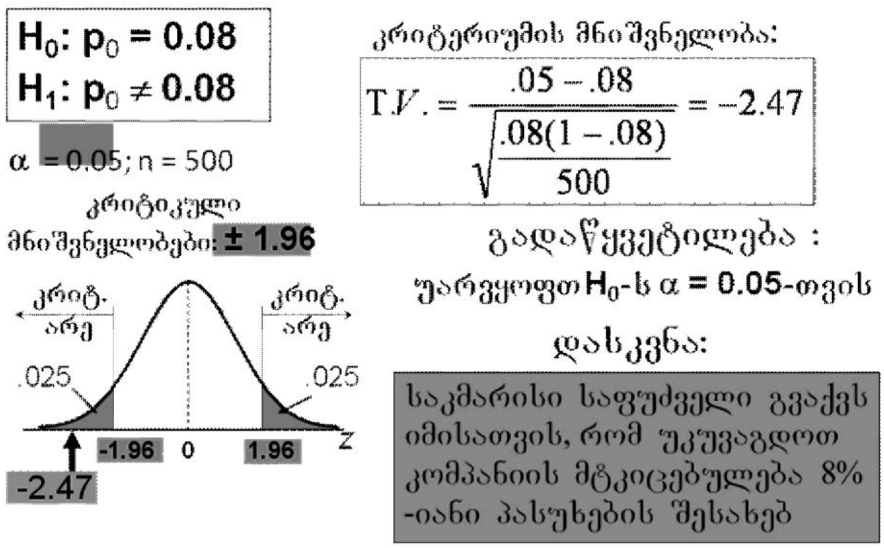
მარკეტინგული კომპანიის მტკიცებით კომპანიის მიერ გაგზავნილი ელექტრონული წერილებიდან 8%-ზე ღებულობს პასუხებს. ამ მტკიცებულების შესამოწმებლად შემთხვევით შერჩეულ 500 ადრესატს დაეგზავნა ელექტრონული წერილი და მათგან 25-მა გამოაგზავნა პასუხი. შეამოწმეთ შესაბამისი ჰიპოთეზა 0.05-ის ტოლი მნიშვნელოვნების დონით.

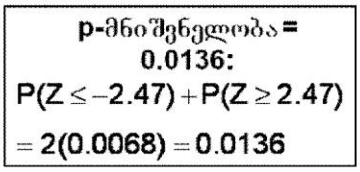
ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0: p = p_0$ და $H_1: p \neq p_0$. ამ შემთხვევაში:

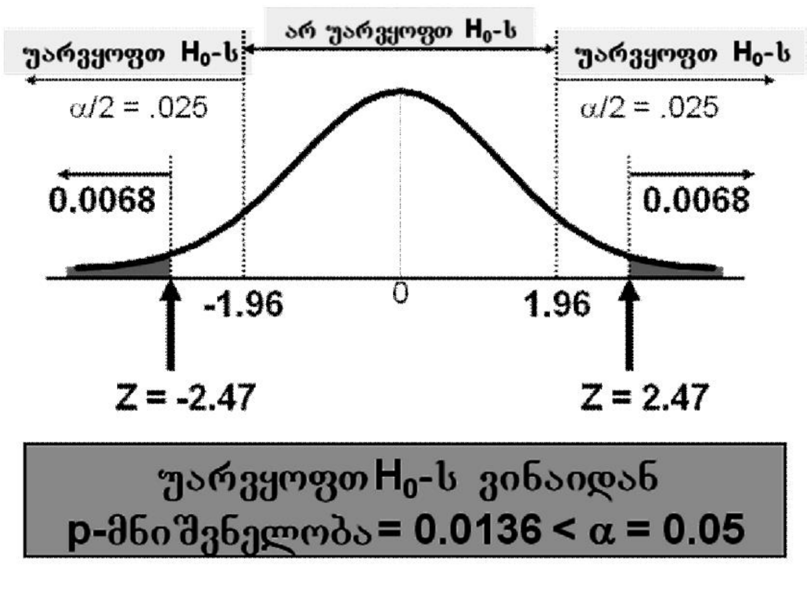
$$p_0 = \frac{8}{100} = 0.08;\ n = 500;\ S_n = 25;\ \bar{x} = \frac{25}{500} = 0.05;\ a = 0.05;\ \pm z_{a/2} = \pm 1.96 .$$

პირველ რიგში ვამოწმებთ პირობებს: $np_0 \geq 5$ და $n(1-p_0) \geq 5$. მართლაც, გვაქვს: $500 \cdot 0.08 = 40 > 5$ და $500 \cdot (1-0.08) = 460 > 5$.

კრიტერიუმის სტატისტიკაა: $Z = \dfrac{\overline{x} - p_0}{\sqrt{p_0(1-p_0)/n}}$ და, საბოლოოდ, გვექნება შემდეგი სურათი:







ჰიპოთეზის შემოწმება პუასონის პოპულაციის λ პარამეტრის შესახებ (მცირე მოცულობის შერჩევებისათვის):

პუასონის პოპულაციის λ პარამეტრის შესახებ ჰიპოთეზის შემოწმებისას, მცირე მოცულობის შერჩევის შემთხვევაში, იყენებენ P -მნიშვნელობის მეთოდს.

ჰიპოთეზა: $H_0 : \lambda = \lambda_0$;

ალტერნატივა: $H_1 : \lambda \neq \lambda_0$;

მნიშვნელოვნების დონე: α;

კრიტერიუმის სტატისტიკა: $\Pi \cong Po(\lambda_0)$ **;**

კრიტერიუმის მნიშვნელობა T.V.: π (პუასონის შემთხვევითი სიდიდის დაკვირვებული მნიშვნელობა).

P **- მნიშვნელობა:**

$$P = \begin{cases} 2 \cdot P\{\Pi \le \pi \mid H_0\} = 2 \cdot \sum_{k=0}^{\pi} \frac{\lambda_0^k}{k!} \cdot e^{-\lambda_0},\ \text{თუ}\ \pi \le \lambda_0; \\ 2 \cdot P\{\Pi \ge \pi \mid H_0\} = 2 \cdot (1 - \sum_{k=0}^{\pi-1} \frac{\lambda_0^k}{k!} \cdot e^{-\lambda_0}),\ \text{თუ}\ \pi > \lambda_0. \end{cases}$$

თუ $P \le \alpha$, მაშინ ამბობენ, რომ შედეგი (ლაპარაკია π-ზე) ***სტატისტიკურად მნიშვნელოვანია*** და H_0 ჰიპოთეზას უარყოფენ, წინააღმდეგ შემთხვევაში კი ***ის სტატისტიკურად უმნიშვნელოა*** და ნულოვან ჰიპოთეზას არ უარყოფენ.

დიდი მოცულობის შემთხვევაში აქაც იყენებენ ნორმალურ აპროქსიმაციას, კერძოდ იმ ფაქტს, რომ

$$(\Pi - \lambda_0)^2 / \lambda_0 \cong [N(0,1)]^2 \cong \chi^2(1).$$

მაგალითი 1. დავუშვათ ჩატარებულია 50 დამოუკიდებელი ექსპერიმენტი, A ხდომილების მოხდენის ფარდობითი სიხშირე აღმოჩნდა 0.12. მნიშვნელოვნების α = 0.01 დონისათვის შევამოწმოთ ნულოვანი H_0: $p \le 0.1$ ჰიპოთეზა ალტერნატიული H_1: $p > 0.1$ ჰიპოთეზის წინააღმდეგ.

ამოხსნა. ვიპოვოთ კრიტერიუმის მნიშვნელობა

$$z = \frac{(0.12 - 0.1)\sqrt{50}}{\sqrt{0.1 \cdot 0.9}} = 0.471.$$

კრიტიკული არე იქნება მარჯვენა ცალმხრივი, ხოლო კრიტიკული მნიშვნელობა უნდა ვიპოვოთ პირობიდან:

$$\Phi(C.V.) = 1 - 0.01 = 0.99.$$

ნორმალური განაწილების ფუნქციის ცხრილიდან ვპოულობთ, რომ $z_\alpha = 2.33$. ვინაიდან $z < z_\alpha$, ამიტომ მიიღება ჰიპოთეზა იმის შესახებ, რომ $p \le 0.1$.

მაგალითი 3. პროკურორის მტკიცებით ადვოკატების 25%-ზე მეტი იყენებს რეკლამას. 200 შემთხვევით შერჩეულ ადვოკატზე დაკვირვებამ აჩვენა, რომ 63 მათგანი ამა თუ იმ ფორმით იყენებდა რეკლამას. $\alpha = 0.05$ მნიშვნელოვნების დონით გვაქვს თუ არა საკმარისი საფუძველი გავიზიაროთ პროკურორის მტკიცებულება?

ამოხსნა. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები:

$$H_0 : p \le 0.25, \; H_1 : p > 0.25.$$

გვაქვს: $p_0 = 0.25$, $n = 200$, $\overline{x} = w_n = \frac{63}{200} = 0.315$. გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$z = \frac{\overline{x} - p_0}{\sqrt{p_0(1 - p_0)/n}} = \frac{0.315 - 0.25}{\sqrt{0.25 \cdot 0.75/200}} = 2.12.$$

გამოვთვალოთ P -მნიშვნელობა:

$$P = 1 - \Phi(z) = 1 - \Phi(2.12) = 1 - 9830 = 0.0170.$$

რადგანაც P -მნიშვნელობა $< \alpha$ (0.017 < 0.05), ამიტომ H_0 უნდა უკუვაგდოთ. მაშასადამე, ჩვენ გვაქვს საკმარისი საფუძველი გავიზიაროთ პროკურორის მოსაზრება იმის შესახებ, რომ ადვოკატების 25%-ზე მეტი იყენებს რეკლამას.

ამოცანები

1. სატელეფონო კომპანიის შეფასებით მათი მომხმარებლების 40%-ს გააჩნია ლოდინის რეჟიმის მქონე მოწყობილობა. ამ ჰიპოთეზის შესამოწმებლად შეირჩა 100 მომხმარებელი და გაირკვა, რომ მათ 37%-ს აქვს ლოდინის რეჟიმის მქონე მოწყობილობა. $\alpha = 0.01$ მნიშვნელოვნების დონით გვაქვს თუ არა საკმარისი საფუძველი უკუვაგდოთ კომპანიის მოსაზრება?
3. სამშობიარო სახლის ჩანაწერების მიხედვით არადღენაკლული ბავშვების 37%-ის საშუალო წონა მეტია ვიდრე 7 ფუნტი და 2 უნცია (1 ფუნტი = 453.6 გრ; 1 უნცია = 28.3 გრ). მიმდინარე წელს 100 დაბადებული ბავშვიდან 23-ის წონა მეტი იყო ვიდრე 7 ფუნტი და 2 უნცია. $\alpha = 0.01$ მნიშვნელოვნების დონით გვაქვს თუ არა საკმარისი საფუძველი ვამტკიცოთ რომ პროპორცია შეიცვალა?
5. უკანასკნელი კვლევის მიხედვით ავიაკატასტროფაში მოყოლილი ადამიანების არა უმეტეს 32% იღუპება. 100 ადამიანისაგან შემდგარ შერჩევაში, რომლებიც მოყვნენ ავიაკატასტროფაში, 38 დაიღუპა. $\alpha = 0.05$ მნიშვნელოვნების დონით უნდა უარვყოთ თუ არა კვლევის შედეგი?
7. სტატისტიკური ანგარიშის მიხედვით სრულწლოვანი მოსახლეობის 17% გასულ წელს დაესწრო წარმოდგენას ოპერაში. ამ ჰიპიოთეზის შესამოწმებლად მკვლევარმა გამოკითხა 90 ადამიანი და დაადგინა, რომ მათგან გასულ წელს ოპერას დაესწრო 22. $\alpha = 0.05$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა, რომ რეალური პროპორცია შეადგენს 17%-ს.
9. კვლევის მიხედვით უნივერსიტეტის სტუდენტების არა უმეტეს 25%-ის უნივერსიტეტამდე მისასვლელად გადის 10 კმ-ზე მეტს. შეთხვევით შერჩეული 100 სტუდენტიდან აღმოჩნდა, რომ 30 უნივერსიტეტამდე მისასვლელად გადის 10 კმ-ზე მეტს. $\alpha = 0.1$ მნიშვნელოვნების დონით შეამოწმეთ კვლევის შედეგების სისწორე. ისარგებლეთ P -მნიშვნელობის მეთოდით.
11. კრიმინალისტების მტკიცებით მკვლელობების 10% ჩადენილია ქალების მიერ. გვაქვს თუ არა საკმარისი საფუძველი

$\alpha = 0.01$ მნიშვნელოვნების დონით უკუვაგდოთ ეს მტკიცებულება, თუ 67 შემთხვევით შერჩეული მკვლელობიდან 10 აღმოჩნდა ქალების მიერ ჩადენილი? გამოიყენეთ P-მნიშვნელობის მეთოდი.

13. გასულ წელს თვითმფრინავის მგზავრების 20% მგზავრობდა პირველი კლასით. მიმდინარე წელს შემთხვევით შერჩეული 15 მგზავრიდან 5-მა იმგზავრა პირველი კლასით. $\alpha = 0.1$ მნიშვნელოვნების დონით შეგვიძლია თუ არა დავასკვნათ, რომ პროპორცია შეიცვალა?

თავი VIII

ნდობის ინტერვალი და ჰიპოთეზათა შემოწმება

თუ ჰიპოთეზის შემოწმების ამოცანაში ხდება ნულოვანი ჰიპოთეზის უარყოფა, მაშინ იმავე მნიშვნელოვნების დონის მქონე ნდობის ინტერვალი პოპულაციის პარამეტრისათვის არ მოიცავს პარამეტრის ჰიპოთეტურ მნიშვნელობას.

თუ ჰიპოთეზის შემოწმების ამოცანაში არ ხდება ნულოვანი ჰიპოთეზის უარყოფა, მაშინ იმავე მნიშვნელოვნების დონის მქონე ნდობის ინტერვალი პოპულაციის პარამეტრისათვის მოიცავს პარამეტრის ჰიპოთეტურ მნიშვნელობას.

მაგალითი 1. შაქარი დაფასოებულია 5 ფუნტიან (1 ფუნტი = 453.6 გრ) ფუთებში. კონტროლიორს ეჭვი აქვს, რომ ფუთაში არ არის 5 ფუნტი შაქარი. 50 შაქრის ფუთისაგან შემდგარი შერჩევის საშუალო აღმოჩნდა 4.6 ფუნტი, ხოლო შესწორებული სტანდარტული გადახრა კი 0.7 ფუნტი. $\alpha = 0.05$ მნიშვნელოვნების დონით გვაქვს თუ არა საფუძველი დავასკვნათ, რომ ფუთაში საშუალოდ არ არის 5 ფუნტი შაქარი? ააგეთ 95%-იანი ნდობის ინტერვალი ჭეშმარიტი საშუალოსათვის.

ამოხსნა. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0: E\xi = 5$, $H_1: E\xi \neq 5$. ამ ორმხრივი კრიტერიუმის შესაბამისი კრიტიკული მნიშვნელობებია: $-z_{\alpha/2} = -z_{0.025} = -1.96$ და $z_{\alpha/2} = z_{0.025} = 1.96$. კრიტერიუმის მნიშვნელობა იქნება:

$$z = \frac{\bar{x} - E\xi}{s' / \sqrt{n}} = \frac{4.5 - 5}{0.7 / \sqrt{50}} = \frac{-0.5}{0.099} = -5.05.$$

რადგანაც $-5.05 < -1.96$, ამიტომ ნულოვანი ჰიპოთეზა უნდა უკუვაგდოთ, ანუ ჩვენ გვაქვს საკმარისი საფუძველი, რათა დავასკვნათ, რომ ფუთაში საშუალოდ არ არის 5 ფუნტი შაქარი.

ავაგოთ 95%-იანი ნდობის ინტერვალი საშუალოსათვის:

$$\bar{x} - z_{\alpha/2} \cdot \frac{s^{'}}{\sqrt{n}} < E\xi < \bar{x} + z_{\alpha/2} \cdot \frac{s^{'}}{\sqrt{n}},$$

$$4.6 - 1.96 \cdot \frac{0.7}{\sqrt{50}} < E\xi < 4.6 + 1.96 \cdot \frac{0.7}{\sqrt{50}},$$

$$4.4 < E\xi < 4.8.$$

როგორც ვხედავთ, 95%-იანი ნდობის ინტერვალი (ანუ ნდობის ინტერვალი მნიშვნელოვნების დონით $\alpha = 0.05$) საშუალოსათვის $E\xi$ არ მოიცავს საშუალოს ჰიპოთეტურ მნიშვნელობას $E\xi = 5$.

ამოცანები

1. თხილამურების მაღაზიის მენეჯერის მტკიცებით ზამთრის თვეების განმავლობაში მისი მაღაზიის საშუალო დღიური ბრუნვა შეადგენს 1800 ლარს. შემთხვევით შერჩეული ზამთრის 10 დღის საშუალო დღიური ბრუნვა აღმოჩნდა 1830 ლარი. პოპულაციის სტანდარტული გადახრაა 200 ლარი. $\alpha = 0.05$ მნიშვნელოვნების დონით გვაქვს თუ არა საფუძველი უარვყოთ მენეჯერის მტკიცებულება? ააგეთ 95%-იანი ნდობის ინტერვალი საშუალოსათვის. არის თუ არა ნდობის ინტერვალის ინტერპრეტაცია თანხვედრაში ჰიპოთეზის შემოწმების შედეგთან? იგულისხმეთ, რომ სიდიდე განაწილებულია ნორმალურად.
3. უძრავი ქონების აგენტის მტკიცებით საკუთარი სახლების კომუნალური გადასახადების საშუალო თვეში შეადგენს 86 ლარს. წარსული კვლევის თანახმად პოპულაციის სტანდარტული გადახრაა 6 ლარი. შემთხვევით შერჩეული 15 სახლის მფლობელის საშუალო გადასახადი აღმოჩნდა 84 ლარი. $\alpha = 0.01$ მნიშვნელოვნების დონით შეამოწმეთ აგენტის მტკიცებულება. ააგეთ 99%-იანი ნდობის ინტერვალი საშუალოსათვის. არის თუ არა ნდობის ინტერვალის ინტერპრეტაცია თანხვედრაში ჰიპოთეზის შემოწმების შედეგთან? იგულისხმეთ, რომ სიდიდე განაწილებულია ნორმალურად.

5. წინა კვლევების თანახმად პირველკურსელი სტუდენტები კვირაში საშუალოდ 22 საათს ხარჯავენ სწავლაზე. სტანდარტული გადახრა შეადგენს 4 საათს. მიმდინარე წელს გამოკითხეს 60 სტუდენტი და აღმოჩნდა, რომ მათ მიერ კვირის განმავლობაში სწავლაზე დახარჯული დროის საშუალო შეადგენდა 20.8 საათს. $\alpha = 0.01$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა, რომ სწავლაზე დახარჯული დროის საშუალო შეიცვალა. ააგეთ 99%-იანი ნდობის ინტერვალი საშუალოსათვის. არის თუ არა ნდობის ინტერვალის ინტერპრეტაცია თანხვედრაში ჰიპოთეზის შემოწმების შედეგთან?

ამოცანები გამოცდისათვის

7. კვლევის თანახმად მწეველი ადამიანი საშუალოდ დღეში ეწევა 14 ცალ სიგარეტს. ამ ჰიპოთეზის შესამოწმებლად შემთხვევით შეირჩა 40 მწეველი და აღმოჩნდა, რომ ისინი დღეში საშუალოდ 18 ცალ სიგარეტს ეწეოდნენ. შერჩევის შესწორებული სტანდარტული გადახრა იყო 6. $\alpha = 0.05$ მნიშვნელოვნების დონით გვაქვს თუ არა საფუძველი ჩავთვალოთ, რომ მწეველების მიერ დღეში მოწეული სიგარეტის რიცხვი სინამდვილეში განსხვავებულია 14-საგან?
9. მკვლევარს სურს შეამოწმოს არის თუა არა დედაქალაქის მოსახლეობის საშუალო ასაკი 61.2 წელი. 22 შემთხვევით შერჩეული მოქალაქის საშუალო ასაკი აღმოჩნდა 59.8 წელი, ხოლო შესწორებული შერჩევითი სტანდარტული გადახრა კი 1.5 წელი. $\alpha = 0.01$ მნიშვნელოვნების დონით განსხვავდება თუ არა რეალურად საშუალო ასაკი 61.2 წლისაგან? იგულისხმეთ, რომ სიდიდე დაახლოებით ნორმალურადაა განაწილებული.
11. წინანდელი კვლევის მიხედვით მკვლელობის მსხვერპლთა საშუალო ასაკი არ აღემატება 23.2 წელს. მიმდინარე წელს შემთხვევით შერჩეული 18 მკვლელობის მსხვერპლთა საშუალო ასაკი აღმოჩნდა 22.6 წელი, ხოლო შესწორებული სტანდარტული გადახრა კი 2 წელი. $\alpha = 0.05$ მნიშვნელოვნების დონით შეგვიძლია თუ არა დავასკვნათ, რომ მიმდინარე წელს მსხვერპლთა საშუალო ასაკი გაიზარდა? იგულის-

ხმეთ, რომ სიდიდე დაახლოებით ნორმალურადაა განაწილებული. ისარგებლეთ P -მნიშვნელობის მეთოდით.

13. უნივერსიტეტის ფინანსური დახმარების დეპარტამენტი იმედოვნებს, რომ სტუდენტთა სულ ცოტა 30% მიიღებს ამა თუ იმ სახის ფინანსურ დახმარებას. იმის გასარკვევად, სწორია თუ არა ეს მოსაზრება, შემთხვევით შერჩეულ იქნა 60 სტუდენტი და აღმოჩნდა, რომ მათგან 15 მიიღო ფინანსური დახმარება. $\alpha = 0.05$ მნიშვნელოვნების დონით შევამოწმოთ ჰიპოთეზა, რომ სტუდენტთა სულ ცოტა 30% მიიღებს ფინანსურ დახმარებას.
15. სამშენებლო კომპანიის მტკიცებით ბინის მყიდველთა 80%-ს სურს ბინაში ქონდეს ბუხარი. ამ მოსაზრების შესამოწმებლად მკვლევარმა შემთხვევით შეარჩია 30 ბინის მყიდველი და დაადგინა, რომ მათგან 20-ს სურდა ბინაში ქონოდა ბუხარი. $\alpha = 0.02$ მნიშვნელოვნების დონით შეამოწმეთ კომპანიის მტკიცებულების სისწორე.
17. ფეხბურთის ფედერაციის პრეზიდენტის მტკიცებით ფეხბურთელების საშუალო წონა შეადგენს 225 ფუნტს (1 ფუნტი = 453.6 გრ). ამ მოსაზრების შესამოწმებლად, მკვლევარმა შემთხვევით შეარჩია 50 ფეხბურთელი და დაადგინა, რომ მათი საშუალო წონა იყო 230 ფუნტი, ხოლო შესწორებული სტანდარტული გადახრა კი 15 ფუნტი. $\alpha = 0.01$ მნიშვნელოვნების დონით შეამოწმეთ პრეზიდენტის მტკიცებულება. ისარგებლეთ P -მნიშვნელობის მეთოდით.
19. „ჟიგულის“ მიერ მოხმარებული საწვავის სტანდარტული გადახრა შეადგენს 4.3 ლიტრს. შემთხვევით შერჩეული 20 „ჟიგულის“ მიერ მოხმარებული საწვავის შესწორებული სტანდარტული გადახრა აღმოჩნდა 2.6 ლიტრი. $\alpha = 0.05$ მნიშვნელოვნების დონით, შეგვიძლია თუ არა დავასკვნათ, რომ შერჩევის სტანდარტული გადახრა ნაკლებია 4.3 ლიტრზე? ისარგებლეთ P -მნიშვნელობის მეთოდით.

21. საღებავების ფირმის მენეჯერის მტკიცებით გარკვეული საღებავის გაშრობის დროის სტანდარტული გადახრა შეადგენს 18 წუთს. ხუთმა სხვადასხვა ტესტირებამ აჩვენა რომ

საღებავის გაშრობის დროის შესწორებული სტანდარტული გადახრა შეადგენს 21 წუთს. $\alpha = 0.05$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა სტანდარტული გადახრის შესახებ?

23. უძრავი ქონების აგენტის მტკიცებით ახალი სახლის ყიდვისას დამატებითი ხარჯების საშუალო შეადგენს 6500 ლარს. შემთხვევით შერჩეული 40 გაყიდული სახლის მყიდველის დამატებითი ხარჯების საშუალომ შეადგინა 6600 ლარი. პოპულაციის სტანდარტული გადახრაა 120 ლარი. $\alpha = 0.05$ მნიშვნელოვნების დონით შეამოწმეთ აგენტის მტკიცებულება.

25. ტაქსების ფირმის მენეჯერის მტკიცებით მათი მძღოლების მუშაობის სტაჟის საშუალო შეადგენს სულ ცოტა 12.4 წელს. შემთხვევით შერჩეული 15 მძღოლის მუშაობის სტაჟის საშუალო აღმოჩნდა 11.2 წელი, ხოლო შესწორებული სტანდარტული გადახრა კი 2 წელი. $\alpha = 0.1$ მნიშვნელოვნების დონით არის თუ არა სინამდვილეში სტაჟის საშუალო უფრო ნაკლები, ვიდრე ამას ამტკიცებს მენეჯერი?

თავი IX

ორამოკრეფიანი ამოცანები პოპულაციათა საშუალოებისათვის

ორი დამოუკიდებელი პოპულაციის საშუალოს შორის განსხვავების ჰიპოთეზათა შემოწმება I

$\xi \cong N(a_1, \sigma_1^2)$ და $\eta \cong N(a_2, \sigma_2^2)$ ან ორივე შერჩევის მოცულობა მეტია ან ტოლი 30-ზე; ξ და η დამოუკიდებელია; σ_1^2 და σ_2^2 ცნობილია. $X_1, ..., X_n$ და $Y_1, ..., Y_m$ ორი შერჩევაა შესაბამისად ξ და η პოპულაციებიდან.

კრიტერიუმი:

ორმხრივი	მარჯვენა ცალმხრივი	მარცხენა ცალმხრივი
$H_0 : a_1 - a_2 = 0$	$H_0 : a_1 - a_2 = 0$	$H_0 : a_1 - a_2 = 0$
	ან $H_0 : a_1 - a_2 \leq 0$	ან $H_0 : a_1 - a_2 \geq 0$
$H_1 : a_1 - a_2 \neq 0$	$H_1 : a_1 - a_2 > 0$	$H_1 : a_1 - a_2 < 0$

ჰიპოთეზა: $H_0 : a_1 - a_2 = 0$

მნიშვნელოვნების დონე: α

კრიტერიუმის სტატისტიკა: $Z = \dfrac{(\overline{X}_n - \overline{Y}_m) - (a_1 - a_2)}{\sqrt{\sigma_1^2 / n + \sigma_2^2 / m}} \cong N(0,1)$

კრიტერიუმის მნიშვნელობა **T.V.:** $z = \dfrac{(\overline{x}_n - \overline{y}_m) - (a_1 - a_2)}{\sqrt{\sigma_1^2 / n + \sigma_2^2 / m}}$

ალტერნატივა	კრიტიკული არე **C.R.** (H_0-ის უარყოფის არე)
$H_1 : a_1 - a_2 > 0$	$z \geq z_{\alpha}$,
$H_1 : a_1 - a_2 < 0$	$z \leq -z_{\alpha}$,
$H_1 : a_1 - a_2 \neq 0$	$z \leq -z_{\alpha/2}$ ან $z \geq z_{\alpha/2}$

(სადაც z_α არის $N(0,1)$-ის ზედა α-კრიტიკული წერტილი ანუ **კრიტიკული მნიშვნელობა C.V.**).

P **-მნიშვნელობა:**

$$P=\begin{cases}P\{Z>z\mid H_0\}, \text{ თუ } H_1: a_1-a_2>0;\\ P\{Z<z\mid H_0\}, \text{ თუ } H_1: a_1-a_2<0;\\ P\{|Z|>|z|\mid H_0\}, \text{ თუ } H_1: a_1-a_2\neq 0.\end{cases}=\begin{cases}1-\Phi(z), \text{ თუ } H_1: a_1-a_2>0;\\ \Phi(z), \text{ თუ } H_1: a_1-a_2<0;\\ 2\cdot[1-\Phi(|z|)], \text{ თუ } H_1: a_1-a_2\neq 0.\end{cases}$$

გადაწყვეტილება: თუ $z\in$ **C.R.**, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

P **-მნიშვნელობის მეთოდი:** თუ $P\le\alpha$, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ α მნიშვნელოვნების დონით, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

$(1-\alpha)$ **საიმედოობის ნდობის ინტერვალი საშუალოთა სხვაობისათვის:**

$$(\bar{x}_n-\bar{y}_m)-z_{\alpha/2}\sqrt{\sigma_1^2/n+\sigma_2^2/m}<a_1-a_2<(\bar{x}_n-\bar{y}_m)+z_{\alpha/2}\sqrt{\sigma_1^2/n+\sigma_2^2/m}\,.$$

მაგალითი 1. ერთი ნორმალური პოპულაციიდან, რომლის სტანდარტული გადახრაა 5, აღებულია 40 მოცულობის მქონე შერჩევა. შერჩევითი საშუალო შეადგენს 102-ს. პირველისაგან დამოუკიდებელი მეორე ნორმალური პოპულაციიდან, რომლის სტანდარტული გადახრაა 6, აღებულია 50 მოცულობის მქონე შერჩევა. ამ უკანასკნელის შერჩევითი საშუალოა 99. $\alpha=0.04$ მნიშვნელოვნების დონით შეამოწმეთ არის თუ არა განსხვავება საშუალოებს შორის.

ამოხსნა. გვაქვს: $\xi\cong N(\cdot,25)$, $\eta\cong N(\cdot,36)$, $n=40$, $m=50$, $\bar{x}_n=102$, $\bar{y}_m=99$.

ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0: a_1-a_2=0$, $H_1: a_1-a_2\neq 0$. ალტერნატივა ორმხრივია, ამიტომ კრიტიკული არე $C.R.=(-\infty,-z_{0.02}]\bigcup[z_{0.02},+\infty)=(-\infty,-2.055]\bigcup[2.055,+\infty)$. გამოვთვალოთ კრიტერიუმის მნიშვნელობა. გვაქვს:

$$T.V.\equiv z=\frac{(\bar{x}_n-\bar{y}_m)-(a_1-a_2)}{\sqrt{\sigma_1^2/n+\sigma_2^2/m}}=\frac{102-99}{\sqrt{25/40+36/50}}=2.59\,.$$

ვინაიდან, 2.59 > 2.055 (ანუ $T.V. \in C.R.$), ამიტომ ძირითად ჰიპოთეზას უკუვაგდებთ $\alpha = 0.04$ მნიშვნელოვნების დონით.

გამოვთვალოთ P -მნიშვნელობა. P -მნიშვნელობა$=2[1-\Phi(2.59)]=$ $=0.0016$. ვინაიდან, P -მნიშვნელობა < $\alpha = 0.04$, ამიტომ P -მნიშვნელობის მეთოდითაც იგივე დასკვნა გამოგვაქვს: $\alpha = 0.04$ მნიშვნელოვნების დონით უკუვაგდებთ H_0 ჰიპოთეზას.

ორი დამოუკიდებელი პოპულაციის საშუალოს შორის განსხვავების ჰიპოთეზათა შემოწმება II

$\xi \cong N(a_1, \sigma_1^2)$ **და** $\eta \cong N(a_2, \sigma_2^2)$**;** ξ **და** η **დამოუკიდებელია;** $\sigma_1^2 = \sigma_2^2 = \sigma^2$ **უცნობია.** $X_1, ..., X_n$ **და** $Y_1, ..., Y_m$ **ორი შერჩევაა შესაბამისად** ξ **და** η **პოპულაციებიდან.**

ჰიპოთეზა: $H_0 : a_1 - a_2 = 0$

მნიშვნელოვნების დონე: α

კრიტერიუმის სტატისტიკა:

$$T = \frac{(\overline{X}_n - \overline{Y}_m) - (a_1 - a_2)}{S'_{n,m}\sqrt{1/n + 1/m}} \cong T(n+m-2),$$

აქ

$$S'^2_{n,m} = \frac{1}{n+m-2}[(n-1)S'^2_1 + (m-1)S'^2_2] = \frac{1}{n+m-2}[\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 + \sum_{i=1}^{m}(Y_i - \overline{Y}_n)^2],$$

$$(n+m-2)\frac{S^2_{n,m}}{\sigma^2} \cong \chi^2(n+m-2).$$

კრიტერიუმის მნიშვნელობა T.V.: $t = \dfrac{(\overline{x}_n - \overline{y}_m) - (a_1 - a_2)}{s'_{n,m}\sqrt{1/n + 1/m}}$

ალტერნატივა	კრიტიკული არე C.R. (H_0 -ის უარყოფის არე)
$H_1 : a_1 - a_2 > 0$	$t \geq t_{n+m-2,\alpha}$,
$H_1 : a_1 - a_2 < 0$	$t \leq -t_{n+m-2,\alpha}$,
$H_1 : a_1 - a_2 \neq 0$	$t \leq -t_{n+m-2,\alpha/2}$ ან $t \geq t_{n+m-2,\alpha/2}$

(სადაც $t_{n+m-2,\alpha}$ არის თავისუფლების $n+m-2$ ხარისხის მქონე სტიუდენტის განაწილების ზედა α-კრიტიკული წერტილი ანუ **კრიტიკული მნიშვნელობა C.V.**).

P **- მნიშვნელობა:**

$$P=\begin{cases} P\{T>t\,|\,H_0\},\ \text{თუ}\ H_1:a_1-a_2>0; \\ P\{T<t\,|\,H_0\},\ \text{თუ}\ H_1:a_1-a_2<0; \\ P\{|T|>|t|\,|\,H_0\},\ \text{თუ}\ H_1:a_1-a_2\neq 0. \end{cases}=\begin{cases} 1-F_T(t),\ \text{თუ}\ H_1:a_1-a_2>0; \\ F_T(t),\ \text{თუ}\ H_1:a_1-a_2<0; \\ 2\cdot[1-F_T(|t|)],\ \text{თუ}\ H_1:a_1-a_2\neq 0. \end{cases}$$

გადაწყვეტილება: თუ $t\in$ **C.R.**, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

P **- მნიშვნელობის მეთოდი:** თუ $P\le\alpha$, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ α მნიშვნელოვნების დონით, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

$(1-\alpha)$ **საიმედოობის ნდობის ინტერვალი საშუალოთა სხვაობისათვის:**

$$(\overline{x}_n-\overline{y}_m)-t_{n+m-2,\alpha/2}s'_{n,m}\sqrt{1/n+1/m}<a_1-a_2<$$
$$<(\overline{x}_n-\overline{y}_m)+t_{n+m-2,\alpha/2}s'_{n,m}\sqrt{1/n+1/m}.$$

მაგალითი 1. ერთი ნორმალური პოპულაციიდან აღებულია $n=10$ მოცულობის შერჩევა, რომლისთვისაც $\overline{x}_n=23$, $s'_1=4$ და მისგან დამოუკიდებელი მეორე ნორმალური პოპულაციიდან აღებულია $m=8$ მოცულობის შერჩევა, რომლისთვისაც $\overline{y}_m=26$, $s'_2=5$. ორივე პოპულაციის დისპერსიები ტოლია. $\alpha=0.05$ მნიშვნელოვნების დონით შეამოწმეთ არის თუ არა განსხვავება საშუალოებს შორის.

ამოხსნა. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0:a_1-a_2=0$, $H_1:a_1-a_2\neq 0$. კრიტერიუმი ორმხრივია. კრიტიკული მნიშვნელობებია: $C.V.=\pm t_{n+m-2,\alpha/2}=\pm t_{16,0.025}=\pm 2.12$. ვიპოვოთ საერთო დისპერსიის შეფასება:

$$s'^2_{n,m}=\frac{1}{n+m-2}[(n-1)s'^2_1+(m-1)s'^2_2]=\frac{1}{10+8-2}\cdot(9\cdot 16+7\cdot 25)=19.94\,.$$

გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$T.V. \equiv t = \frac{(\overline{x}_n - \overline{y}_m) - (a_1 - a_2)}{s'_{n,m}\sqrt{1/n + 1/m}} = \frac{(23-26)-0}{\sqrt{19.94}\cdot\sqrt{1/10+1/8}} = -1.42 \,.$$

ვინაიდან $T.V. \notin C.R.$, ამიტომ ძირითადი ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს, ანუ საშუალოებს შორის განსხვავება არ არის არსებითი.

ორი დამოუკიდებელი პოპულაციის საშუალოს შორის განსხვავების ჰიპოთეზათა შემოწმება III

პოპულაციები არაა ნორმალური, მაგრამ ორივე შერჩევის მოცულობა მეტია ან ტოლი 30-ზე ($n \geq 30$, $m \geq 30$) და დისპერსიებიც უცნობია. $X_1,\ldots,X_n$ და $Y_1,\ldots,Y_m$ ორი შერჩევაა შესაბამისად ξ და η პოპულაციებიდან.

ჰიპოთეზა: $H_0 : a_1 - a_2 = 0$

მნიშვნელოვნების დონე: α

კრიტერიუმის სტატისტიკა: $Z = \dfrac{(\overline{X}_n - \overline{Y}_m) - (a_1 - a_2)}{\sqrt{S_1'^2/n + S_2'^2/m}} \overset{as}{\cong} N(0,1)$

(სიმბოლო $\overset{as}{\cong}$ ნიშნავს, რომ სიდიდე მიახლოებით ნორმალურადაა განაწილებული; $S_1'^2 := \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$, $S_2'^2 := \dfrac{1}{m-1}\sum_{i=1}^{m}(Y_i - \overline{Y}_m)^2$)

კრიტერიუმის მნიშვნელობა **T.V.:** $z = \dfrac{(\overline{x}_n - \overline{y}_m) - (a_1 - a_2)}{\sqrt{s_1'^2/n + s_2'^2/m}}$

ალტერნატივა	კრიტიკული არე **C.R.** (H_0-ის უარყოფის არე)
$H_1 : a_1 - a_2 > 0$	$z \geq z_\alpha$,
$H_1 : a_1 - a_2 < 0$	$z \leq -z_\alpha$,
$H_1 : a_1 - a_2 \neq 0$	$z \leq -z_{\alpha/2}$ ან $z \geq z_{\alpha/2}$

(სადაც z_α არის $N(0,1)$-ის ზედა α-კრიტიკული წერტილი ანუ კრიტიკული მნიშვნელობა **C.V.**).

P - მნიშვნელობა: $P = \begin{cases} 1-\Phi(z), \text{ თუ } H_1: a_1 - a_2 > 0; \\ \Phi(z), \text{ თუ } H_1: a_1 - a_2 < 0; \\ 2\cdot[1-\Phi(|z|)], \text{ თუ } H_1: a_1 - a_2 \neq 0. \end{cases}$

გადაწყვეტილება: თუ $z \in$ **C.R.**, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

P **-მნიშვნელობის მეთოდი:** თუ $P \le \alpha$, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ α მნიშვნელოვნების დონით, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

$(1-\alpha)$ **საიმედოობის ნდობის ინტერვალი საშუალოთა სხვაობისათვის:**

$$(\bar{x}_n - \bar{y}_m) - z_{\alpha/2}\sqrt{s_1^{'2}/n + s_2^{'2}/m} < a_1 - a_2 < (\bar{x}_n - \bar{y}_m) + z_{\alpha/2}\sqrt{s_1^{'2}/n + s_2^{'2}/m}\,.$$

შენიშვნა: იმ შემთხვევაშიც, როცა ჩვენ ვამოწმებთ ჰიპოთეზას იმის შესახებ, რომ საშუალოთა შორის განსხვავება არის არა ნული, არამედ რაიმე კონკრეტული რიცხვი ($a_1 - a_2 = c$), კრიტერიუმის სტატისტიკად ისევ განიხილება სტატისტიკა:

$$Z = \frac{(\overline{X}_n - \overline{Y}_m) - (a_1 - a_2)}{\sqrt{S_1^{'2}/n + S_2^{'2}/m}} \overset{as}{\cong} N(0,1)\,.$$

მაგალითი 1. საერთაშორისო სტატისტიკური ასოციაციის მონაცემებით ლონდონში სასტუმროს ოთახის საშუალო ღირებულება არის \$88.42, ხოლო პარიზში კი \$80.61. დავუშვათ, რომ თითოეულ შემთხვევაში აღებულია 50 სხვადასხვა სასტუმროს მონაცემი და შესწორებული შერჩევითი სტანდარტული გადახრებია შესაბამისად \$5.62 და \$4.83. $\alpha = 0.05$ მნიშვნელოვნების დონით არის თუ არა არსებითი განსხვავება საშუალოებს შორის? ააგეთ 95%-იანი ნდობის ინტერვალი საშუალოთა სხვაობისათვის.

ამოხსნა. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0: a_1 - a_2 = 0$, $H_1: a_1 - a_2 \neq 0$. ვიპოვოთ კრიტიკული მნიშვნელობები:

$$C.V. = \pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96\,.$$

გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$z = \frac{(\bar{x}_n - \bar{y}_m) - (a_1 - a_2)}{\sqrt{s_1^{'2}/n + s_2^{'2}/m}} = \frac{(88.42 - 80.61) - 0}{\sqrt{5.62^2/50 + 4.83^2/50}} = 7.45\,.$$

ვინაიდან, 7.45 > 1.96, ამიტომ ნულოვანი ჰიპოთეზა უნდა უკუვაგდოთ, ანუ საშუალოებს შორის განსხვავება არსებითია.

ავაგოთ ნდობის ინტერვალი. ამ შემთხვევაში $95\% = (1-\alpha)\cdot 100\%$, საიდანაც $\frac{\alpha}{2} = 0.025$. შესაბამისად, $z_{\alpha/2} = 1.96$. ამიტომ ნდობის ინტერვალი იქნება:

$$(88.42 - 80.61) - 1.96\sqrt{5.62^2/50 + 4.83^2/50} < a_1 - a_2 <$$

$$< (88.42 - 80.61) + 1.96\sqrt{5.62^2/50 + 4.83^2/50}\,,$$

$$5.76 < a_1 - a_2 < 9.86.$$

რამდენადაც ნდობის ინტერვალი არ მოიცავს ნულს, ჩვენ უნდა მივიღოთ გადაწყვეტილება ნულოვანი ჰიპოთეზის უკუგდების შესახებ, რაც ეთანხმება ჰიპოთეზის შემოწმებისას მიღებულ შედეგს.

ამოცანები

1. მკვლევარს სურს შეამოწმოს არის თუ არა ამერიკის მთავარი მდინარეების საშუალო სიგრძე იგივე, რაც ევროპის მთავარი მდინარეების საშუალო სიგრძე. $\alpha = 0.01$ მნიშვნელოვნების დონით გვაქვს თუ არა საკმარისი საფუძველი უარვყოთ ეს მოსაზრება თუ ცნობილია, რომ შემთხვევით შერჩეული მთავარი მდინარეების სიგრძეთა მონაცემებია:

ამერიკა				ევროპა			
729	329	450	330	481	532	1776	1224
329	600	1243	525	877	447	824	634
850	532	710	300	565	675	724	357

560	332	2315	410	1122	634	326	580
800	1310	605	926	567	932	1124	405
310	375	545	470	454	820	505	496
434	360	865	1036	230	626	210	252
447	652	360	722	600	1575	2290	
430	1979	259	425				

3. მკვლევარს სურს გაარკვიოს არის თუ არა მწეველი ადამიანის პულსი უფრო მაღალი ვიდრე არამწეველის. შემთხვევით შერჩეული 100 მწეველისა და 100 არამწეველის გამოკვლევის შედეგებია შესაბამისად: $\overline{x}_{100}=90$, $s_1'=5$; $\overline{y}_{100}=88$, $s_2'=6$. შეუძლია თუ არა მკვლევარს $\alpha=0.05$ მნიშვნელოვნების დონით დაასკვნას, რომ მწეველი ადამიანის პულსი უფრო მაღალია ვიდრე არამწეველის?

5. $\alpha=0.02$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ საავადმყოფოების დერეფნებში ხმაურის დონე (გაზომილი დეციბალებში) უფრო მაღალია, ვიდრე ოთახებში, თუ შესაბამისი შერჩევითი მახასიათებლებია: $\overline{x}_{84}=61.2$, $s_1'=7.9$; $\overline{y}_{34}=59.4$, $s_2'=7.9$.

7. შემოწმებულ იქნა ქალთა ორი ჯგუფის ცოდნის დონე. პირველ ჯგუფში შედიოდნენ ისინი ვინც ინსტიტუტის დამთავრებიდან რამდენიმე თვეში თავი დაანებეს სამუშაოს თავიანთი პროფესიის მიხედვით, ხოლო მეორე ჯგუფში კი ისინი ვინც დარჩნენ სამუშაოზე თავიანთი პროფესიის მიხედვით. კვლევის შედეგად მიღებული შერჩევითი მახასიათებლებია: $\overline{x}_{103}=3.16$, $s_1'=0.52$; $\overline{y}_{225}=3.28$, $s_2'=0.46$. $\alpha=0.05$ მნიშვნელოვნების დონით შეგვიძლია თუ არა დავასკვნათ, რომ ცოდნის დონე პირველ ჯგუფში უფრო მაღალია?

9. კოლეჯის ადმინისტრატორის მტკიცებით კოლეჯის წლიური დანახარჯი ვაჟების სპორტზე მეტია ვიდრე ქალების სპორტზე. ქვემოთ მოყვანილი მონაცემების მიხედვით, გვაქვს თუ არა საკმარისი საფუძველი $\alpha=0.01$ მნიშვნელოვნების დონით დავადასტუროთ ადმინისტრატორის მტკიცებულება?

ვაჟები

13351	22220	11456	12244	8383	8796	7551	5254	6576	3377
10128	14029	10160	16175	623	5544	5811	7550	1664	7656
10652	15048	11480	10126	8811	7120	9119	5472	9505	8033
11015	12371	11267	12703	9732	8605	2063	6670	6797	7040
12919	13763	12403	5286	9571	9544	8725	9463	7723	9626

ქალები

10333	11248	14698	7654	7054	6869	933	9883	6959	6030
12745	16249	11597	9331	7300	7874	6989	5232	8478	7058
12016	10248	13371	5468	6407	6909	6502	6815	9959	9907
10082	11041	10353	7055	8324	9110	9277	5933	3991	5832
11324	10127	7435	8917	8411	7235	8903	6925	5922	6502

11. ქალთა ორ ჯგუფს შესთავაზეს კითხვარი, რათა გაერკვიათ საკუთარი თავისადმი პატივისცემის ხარისხი. პირველ ჯგუფში შედიოდნენ ისინი ვინც ინსტიტუტის დამთავრებიდან რამდენიმე თვეში თავი დაანებეს სამუშაოს თავიანთი პროფესიის მიხედვით, ხოლო მეორე ჯგუფში კი ისინი ვინც დარჩნენ სამუშაოზე თავიანთი პროფესიის მიხედვით. კვლევის შედეგად მიღებული შერჩევითი მახასიათებლებია: $\overline{x}_{103} = 3.05$, $s_1^{'} = 0.75$; $\overline{y}_{225} = 2.96$, $s_2^{'} = 0.75$. $\alpha = 0.05$ მნიშვნელოვნების დონით შეგვიძლია თუ არა დავასკვნათ, რომ ამ ჯგუფებში თვითპატივისცემის ხარისხი სხვადასხვაა? ისარგებლეთ P -მნიშვნელობის მეთოდით.

13. ორი ჯგუფის ტესტირებისას მიღებული ქულების მიხედვით გამოთვლილი შესაბამისი შერჩევითი მახასიათებლებია: $\overline{x}_{36} = 83.6$, $s_1^{'} = 4.3$; $\overline{y}_{36} = 79.2$, $s_2^{'} = 3.8$. ააგეთ 90%-იანი ნდობის ინტერვალი საშუალოთა სხვაობისათვის.
15. ორი სახის ელემენტის შემთხვევით შერჩეული პარტიების შემოწმების მიხედვით გამოთვლილი შესაბამისი შერჩევითი მახასიათებლებია: $\overline{x}_{27} = 9.2$, $s_1^{'} = 0.3$; $\overline{y}_{30} = 8.8$, $s_2^{'} = 0.1$. ჩათვალეთ, რომ სიდიდეები ნორმალურადაა განაწილებული და ააგეთ 95%-იანი ნდობის ინტერვალი საშუალოთა სხვაობისათვის.

თავი X

ორამოკრეფიანი ამოცანები პოპულაციათა საშუალოებისათვის მცირე შერჩევების შემთხვევაში (გამარტივებული პროცედურა). სატერტვაიტის მეთოდი

ორი დამოუკიდებელი ნორმალური პოპულაციის საშუალოს შორის განსხვავების ჰიპოთეზათა შემოწმება (გამარტივებული პროცედურა) IV

$X_1,...,X_n$ და $Y_1,...,Y_m$ ორი შერჩევაა დამოუკიდებელი ξ და η ნორმალური პოპულაციებიდან, რომელთა დისპერსიები განსხვავებულია; ერთ-ერთი ან ორივე შერჩევის მოცულობა ნაკლებია 30-ზე.

ჰიპოთეზა: $H_0: a_1 - a_2 = 0$

მნიშვნელოვნების დონე: α

კრიტერიუმის სტატისტიკა: $T = \dfrac{(\overline{X}_n - \overline{Y}_m) - (a_1 - a_2)}{\sqrt{S_1^{'2}/n + S_2^{'2}/m}} \cong T(k)$,

სადაც თავისუფლების ხარისხი $k = \min(n-1, m-1)$.

კრიტერიუმის მნიშვნელობა **T.V.:** $t = \dfrac{(\overline{x}_n - \overline{y}_m) - (a_1 - a_2)}{\sqrt{s_1^{'2}/n + s_2^{'2}/m}}$

ალტერნატივა   კრიტიკული არე **C.R.** (H_0-ის უარყოფის არე)

$H_1: a_1 - a_2 > 0$ $t \ge t_{k,\alpha}$,

$H_1: a_1 - a_2 < 0$ $t \le -t_{k,\alpha}$,

$H_1: a_1 - a_2 \ne 0$ $t \le -t_{k,\alpha/2}$ ან $t \ge t_{k,\alpha/2}$

(სადაც $t_{k,\alpha}$ არის თავისუფლების k ხარისხის მქონე სტიუდენტის განაწილების ზედა α-კრიტიკული წერტილი ანუ **კრიტიკული მნიშვნელობა C.V.**).

P - **მნიშვნელობა:**

$$P=\begin{cases}P\{T>t\,|\,H_0\},\ \text{თუ}\ H_1:a_1-a_2>0;\\ P\{T<t\,|\,H_0\},\ \text{თუ}\ H_1:a_1-a_2<0;\\ P\{|T|>|t|\,|\,H_0\},\ \text{თუ}\ H_1:a_1-a_2\neq 0.\end{cases}=\begin{cases}1-F_T(t),\ \text{თუ}\ H_1:a_1-a_2>0;\\ F_T(t),\ \text{თუ}\ H_1:a_1-a_2<0;\\ 2\cdot[1-F_T(|t|)],\ \text{თუ}\ H_1:a_1-a_2\neq 0.\end{cases}$$

გადაწყვეტილება: თუ $t\in$ **C.R.**, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

P - **მნიშვნელობის მეთოდი:** თუ $P\le\alpha$, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ α მნიშვნელოვნების დონით, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

$(1-\alpha)$ **საიმედოობის ნდობის ინტერვალი საშუალოთა სხვაობისათვის:**

$$(\overline{x}_n-\overline{y}_n)-t_{k,\alpha/2}\sqrt{s_1^{'2}/n+s_2^{'2}/m}<a_1-a_2<$$
$$<(\overline{x}_n-\overline{y}_n)+t_{k,\alpha/2}\sqrt{s_1^{'2}/n+s_2^{'2}/m}.$$

შენიშვნა: *საშუალოთა შესახებ ტოლობის შესამოწმებლად T კრიტერიუმის გამოყენებამდე წინასწარ უნდა გამოვიყენოთ F კრიტერიუმი, რათა გავარკვიოთ არის თუ არა დისპერსიები ტოლი. ამ უკანასკნელის გამოსაყენებლად კი აუცილებელია პოპულაციები იყოს ნორმალური ან დაახლოებით ნორმალური.*

მაგალითი 1. ფერმერული მეურნეობების საშუალო ფართობი ინდიანას შტატში შეადგენს 191 აკრს (1 აკრი = 0.4 ჰა), ხოლო აიოვას შტატში კი 199 აკრს. დავუშვათ, რომ ეს მონაცემები მიღებულია შესაბამისად 8-ისა და 10-ის ტოლი შერჩევებიდან და შესაბამისი შესწორებული სტანდარტული გადახრებია 38 აკრი და 12 აკრი. $\alpha=0.05$ მნიშვნელოვნების დონით შეიძლება თუ არა დავასკვნათ, რომ ფერმერული მეურნეობების საშუალო ფართობი ამ ორ შტატში განსხვავებულია? იგულისხმება, რომ პოპულაციები ნორმალურია. ააგეთ 95%-იანი ნდობის ინტერვალი საშუალოთა სხვაობისათვის.

ამოხსნა. წინასწარ შევამოწმოთ ჰიპოთეზა იმის შესახებ არის თუ არა პოპულაციათა დისპერსიები ტოლი. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0:\sigma_1^2=\sigma_2^2$,

$H_1 : \sigma_1^2 \neq \sigma_2^2$. ამ შემთხვევაში $38^2 = s_1^{'2} > s_2^{'2} = 12^2$. შესაბამისად, თავისუფლების ხარისხებია: 8 – 1 = 7 და 10 – 1 = 9. ამიტომ კრიტიკული მნიშვნელობა იქნება: $C.V. = F_{k,l,\alpha/2} = F_{7,9,0.025} = 4.2$. გამოვთვალოთ კრიტერიუმის მნიშვნელობა: $T.V. \equiv \overline{f} = s_1^{'2} / s_2^{'2} = = 38^2 / 12^2 = 10.03$. რადგანაც 10.03 > 4.2, ამიტომ ნულოვანი ჰიპოთეზა უნდა უკუვაგდოთ, ანუ პოპულაციათა დისპერსიები განსხვავებულია. შესაბამისად, საშუალოთა შესახებ ტოლობის შესამოწმებლად უნდა ვისარგებლოთ T კრიტერიუმით.

ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0 : a_1 = a_2$, $H_1 : a_1 \neq a_2$. რადგანაც კრიტერიუმი ორმხრივია და დისპერსიები განსხვავებული, ამიტომ თავისუფლების ხარისხი ტოლია $k = \min(8-1, 10-1) = 7$. შესაბამისად, კრიტიკული მნიშვნელობებია $C.V. = \pm t_{k,\alpha/2} = \pm t_{7,0.025} = \pm 2.365$. გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$T.V. \equiv t = \frac{(\overline{x}_n - \overline{y}_m) - (a_1 - a_2)}{\sqrt{s_1^{'2}/n + s_2^{'2}/m}} = \frac{(191-100) - 0}{\sqrt{38^2/8 + 12^2/10}} = -0.57 .$$

როგორც ვხედავთ კრიტერიუმის მნიშვნელობა არ შედის კრიტიკულ არეში, ამიტომ ნულოვან ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს, ანუ ჩვენ არა გვაქვს საკმარისი საფუძველი რათა დავასკვნათ, რომ საშუალოები განსხვავებულია.

ავაგოთ ნდობის ინტერვალი. ამისათვის საჭირო სიდიდეები შევიტანოთ შესაბამის ფორმულაში. გვაქვს:

$$(191-199) - 2.365 \cdot \sqrt{\frac{38^2}{8} + \frac{12^2}{10}} < a_1 - a_2 < (191-199) + 2.365 \cdot \sqrt{\frac{38^2}{8} + \frac{12^2}{10}} ,$$

$$-41.02 < a_1 - a_2 < 25.02 .$$

ორამოკრეფიანი t-კრიტერიუმი არატოლი დისპერსიების შემთხვევაში, სატერტვაიტის მეთოდი:

ვთქვათ, მოცემულია ორი დამოუკიდებელი ნორმალურად განაწილებული პოპულაცია საშუალოებით μ_1 და μ_2 და ცნობი-

ლია, რომ მათ აქვთ არატოლი დისპერსიები $\sigma_1^2 \neq \sigma_2^2$. ჩვენი ამოცანაა შევამოწმოთ $H_0 : \mu_1 = \mu_2$ ჰიპოთეზა, $H_1 : \mu_1 \neq \mu_2$ ალტერნატივის წინააღმდეგ. ცნობილია, რომ

$$\bar{X}_{1n_1} - \bar{X}_{2n_2} \cong N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}),$$

სადაც n_1 და n_2, შესაბამისად, პირველი და მეორე შერჩევის მოცულობებია. აქედან გამომდინარე, თუ ცნობილია σ_1^2 და σ_2^2, მაშინ H_0 ჰიპოთეზის სამართლიანობის შემთხვევაში

$$\frac{\bar{X}_{1n_1} - \bar{X}_{2n_2}}{\sqrt{\sigma_1^2 / n_1 + \sigma_2^2 / n_2}} \cong N(0,1). \qquad (1)$$

ამიტომ, თუ ცნობილია σ_1^2 და σ_2^2, მაშინ (1) შემთხვევითი სიდიდე გამოდგება კრიტერიუმის სტატისტიკად. მაგრამ ეს სიდიდე კრიტერიუმის სტატისტიკად არ გამოგვადგება, როცა ის შეიცავს უცნობ σ_1^2 და σ_2^2 პარამეტრებს. ამ შემთხვევაში მათ მაგივრად (1) გამოსახულებაში ჩავსვათ მათი შეფასებები (შესაბამისი შესწორებული შერჩევითი დისპერსიები $S_{1n_1}^2$ და $S_{2n_2}^2$) და კრიტერიუმის სტატისტიკად ავიღოთ შემდეგი შემთხვევითი სიდიდე:

$$T_{n_1,n_2} \equiv \frac{\bar{X}_{1n_1} - \bar{X}_{2n_2}}{\sqrt{S_{1n_1}^2 / n_1 + S_{2n_2}^2 / n_2}}. \qquad (2)$$

როგორც ვიცით, კრიტერიუმის სტატისტიკა მხოლოდ გამოთვლადი სიდიდე კი არ უნდა იყოს, არამედ კრიტერიუმის ასაგებად, არსებითია მისი ზუსტი ან ასიმპტოტური განაწილების ცოდნაც. როგორაა განაწილებული T_{n_1,n_2} შემთხვევითი სიდიდე? მათემატიკურ სტატისტიკაში ეს პრობლემა ცნობილია ***ბერენს-ფიშერის პრობლემის*** სახელით. მისი გადაჭრის ერთ-ერთი, შედარებით მარტივი გზა ცნობილია ***სატერტვაიტის*** **(Satterthwaite)** ***მეთოდის*** სახელით, რომელიც შემდეგნაირად ყალიბდება:

თუ მოცემული α მნიშვნელოვნების დონისათვის T_{n_1,n_2} სტა-ტისტიკის დაკვირვებული მნიშვნელობა t_{n_1,n_2} აკმაყოფილებს შემდეგ უტოლობას

$$-t_{[c],\alpha/2} \le t_{n_1,n_2} \le t_{[c],\alpha/2}, \qquad (3)$$

სადაც $[c]$ აღნიშნავს c რიცხვის მთელ ნაწილს (ანუ c -ს უახლო-ეს მთელ რიცხვს მარცხნიდან), ხოლო c გამოითვლება შემდგი წესით

$$c = \frac{(s_{1n_1}^2 / n_1 + s_{2n_2}^2 / n_2)^2}{(s_{1n_1}^2 / n_1)^2 / (n_1 - 1) + (s_{2n_2}^2 / n_2)^2 / (n_2 - 1)}, \qquad (4)$$

მაშინ α მნიშვნელოვნების დონით $H_0 : \mu_1 = \mu_2$ ჰიპოთეზის უარ-ყოფის საფუძველი არა გვაქვს, წინააღმდეგ შემთხვევაში, მას უარვყოფთ.

მაგალითი. დავუშვათ, რომ შეისწავლება ქოლესტერინის შემცველობა გულის დაავადებით გარდაცვალებული ადამიანე-ბის მემკვიდრეებში. დავუშვათ, რომ 100 ასეთ ბავშვზე დაკვირ-ვებამ მოგვცა ქოლესტერინის საშუალო დონე 207.3 მგ/დლ, ხო-ლო შესწორებული სტანდარტული გადახრაა 35.6 მგ/დლ. გარდა ამისა, დააკვირდნენ ქოლესტერინის დონეებს იმ ბავშვებში, რო-მელთა მშობლებიც ცოცხლებია და არა აქვთ გულის დაავადება. ასეთნაირად შერჩეული 74 ბავშვის მონაცემების მიხედვით ქო-ლესტერინის საშუალო დონე აღმოჩნდა 193.4 მგ/დლ, ხოლო შესწორებული სტანდარტული გადახრა –17.3 მგ/დლ. ჩვენი ამოცანაა $\alpha = 0.05$ მნიშვნელოვნების დონისათვის შევამოწმოთ ჰიპოთეზა პოპულაციების საშუალოების ტოლობის შესახებ.

ამოხსნა. დავუშვათ, რომ პოპულაციები განაწილებულია ნორმალურად, მაგრამ მათი პარამეტრები უცნობია. ეს ამოცანა გავყოთ ორ ეტაპად: პირველ ეტაპზე შევამოწმოთ ჰიპოთეზა დისპერსიების ტოლობის შესახებ. თუ ჰიპოთეზას დისპერსიე-ბის ტოლობის შესახებ ვერ უარვყოფთ, შემდეგ შევამოწმოთ ჰი-პოთეზა პოპულაციების საშუალოების ტოლობის შესახებ ტოლი დისპერსიების შემთხვევაში. იმ შემთხვევაში თუ არ დადასტურ-და დისპერსიათა ტოლობის ჰიპოთეზა, საშუალოთა ტოლობის ჰიპოთეზა შევამოწმოთ სატერტვაიტის მეთოდით.

განვიხილოთ შერჩევით დისპერსიათა შეფარდება და შევადაროთ ის ფიშერის განაწილების ზედა კრიტიკული წერტილის მნიშვნელობას $F_{99,73,0.025}$ (იხ. თავი XI). $F(100,74)$ სტატისტიკის დაკვირვებული მნიშვნელობა ტოლია

$$f = 35.6^2 / 17.3^2 = 4.23 .$$

ვინაიდან ცხრილებში არ არის ფიშერის განაწილების ზედა კრიტიკული წერტილის მნიშვნელობა – $F_{99,73,0.025}$, ამიტომ გამოვთვალოთ p-მნიშვნელობა, რომელიც ალტერნატივის ორმხრივობის გამო ტოლია

$$p = 2 \cdot P\{F(100,74) \geq 4.23\} \leq 0.0002 ,$$

რაც იმის მაჩვენებელია, რომ შედეგი სტატისტიკურად მნიშვნელოვანია და, მაშასადამე, ჰიპოთეზა დისპერსიათა ტოლობის შესახებ უნდა უარვყოთ.

გადავიდეთ მეორე ეტაპზე. რადგან დავადგინეთ, რომ პოპულაციათა დისპერსიები განსხვავებულია, ვიყენებთ სატერტვაიტის მეთოდს საშუალოთა ტოლობის შესახებ ჰიპოთეზის შესამოწმებლად. ჯერ გამოვთვალოთ (2) თანაფარდობით განსაზღვრული სტატისტიკის დაკვირვებული მნიშვნელობა:

$$t = \frac{207.3 - 193.4}{\sqrt{35.6^2 / 100 + 17.3^2 / 74}} = \frac{13.9}{4.089} = 3.4 .$$

ახლა გამოვთვალოთ თავისუფლების ხარისხის მიახლოებითი მნიშვნელობა c. (4) ტოლობის მიხედვით მოვიღებთ, რომ:

$$c = \frac{(35.6^2 / 100 + 17.3^2 / 74)^2}{(35.6^2 / 100)^2 / 99 + (17.3^2 / 74)^2 / 73} = \frac{16.718^2}{1.8465} = 151.4 ,$$

და, მაშასადამე, $[c] = 151$.

ვინაიდან $t = 3.4 > 1.98 = t_{151,0.025}$, ამიტომ $\alpha = 0.05$ მნიშვნელოვნების დონით პოპულაციათა საშუალოების ტოლობის ჰიპოთეზას უარვყოფთ.

აქვე აღვნიშნავთ, რომ $1-\alpha$ ალბათობის მქონე ნდობის ინტერვალს ექნება შემდეგი სახე:

$$\left(\overline{x}_{1n_1} - \overline{x}_{1n_2} - t_{[c],\alpha/2} \cdot \sqrt{\frac{s_{1n_1}^2}{n_1} + \frac{s_{2n_2}^2}{n_2}} \,;\, \overline{x}_{1n_1} - \overline{x}_{1n_2} + t_{[c],\alpha/2} \cdot \sqrt{\frac{s_{1n_1}^2}{n_1} + \frac{s_{2n_2}^2}{n_2}} \right).$$

ამოცანები

ქვემოთ მოყვანილ ამოცანებში ყველა სიდიდე ნორმალურია ან დაახლოებით ნორმალური

1. უძრავი ქონების აგენტის მტკიცებით საგადასახადო ინსპექტორის მიერ შეფასებული სახლის ღირებულება შეესაბამება სახლის გაყიდვის ფასს. შემთხვევით შერჩეული 10 - 10 სახლისათვის საგადასახადო ინსპექტორის შეფასებისა და გაყიდვის ფასის მიხედვით გამოთვლილი შესაბამისი შერჩევითი მახასიათებლები აღმოჩნდა: $\overline{x}_{10} = 83256$, $s_1^{'} = 3256$; $\overline{y}_{10} = 88354$, $s_2^{'} = 2341$. $\alpha = 0.05$ მნიშვნელოვნების დონით არის თუ არა საშუალოებს შორის განსხვავება მნიშვნელოვანი? ააგეთ 95%-იანი ნდობის ინტერვალი საშუალოთა სხვაობისათვის.
3. პროგრამირების პედაგოგის აზრით მათემატიკის მიმართულების სტუდენტს კომპიუტერული პროგრამის დაწერა და შესწორება შეუძლია უფრო სწრაფად ვიდრე ბიზნესის მიმართულების სტუდენტს. შემთხვევით შერჩეული 12 მათემატიკის (შესაბამისად, 18 ბიზნესის) მიმართულების სტუდენტის მიერ კონკრეტული პროგრამის დაწერაზე და შესწორებაზე დახარჯული დროის საშუალო აღმოჩნდა 36 წუთი (შესაბამისად, 39 წუთი), ხოლო შესაბამისი შესწორებული სტანდარტული გადახრებია: 4 წუთი და 9 წუთი. $\alpha = 0.1$ მნიშვნელოვნების დონით შეიძლება თუ არა დავადასტუროთ პედაგოგის მოსაზრება?
5. კლინიკის ადმინისტრატორს აინტერესებს არის თუ არა ბინაზე ექიმის ყოველდღიური გამოძახებების საშუალო უფრო მეტი ვიდრე კლინიკაში მოსული პაციენტების საშუალო. ქვემოთ მოყვანილი შერჩევების მონაცემების მიხედვით, $\alpha = 0.01$ მნიშვნელოვნების დონით, შეიძლება თუ არა დავადასტუროთ, რომ ექიმის ბინაზე გამოძახებათა საშუალო მეტია კლინიკაში მოსულ პაციენტთა საშუალოზე?

ბინაზე გამოძახება			კლინიკაში მოსვლა		
25	42	57	5	28	37
21	34	44	16	16	48

7. ფინანსური დეკლარაციის მიმღები სახელმწიფო ფირმა (შესაბამისად, კერძო ფირმა) საშუალოდ 10 ადამიანის მომსახურებას ანდომებს 21 წუთს (შესაბამისად, 14 ადამიანის მომსახურებას ანდომებს 27 წუთს), ხოლო შესწორებული შერჩევითი სტანდარტული გადახრაა 5.6 წუთი (შესაბამისად, 4.3 წუთი). $\alpha = 0.02$ მნიშვნელოვნების დონით არის თუ არა განსხვავება ფირმების მომსახურების საშუალო დროებს შორის? ააგეთ 98%-იანი ნდობის ინტერვალი საშუალოთა სხვაობისათვის.
9. 12 ექიმის (შესაბამისად, 27 მედდის) დაზღვევის საშუალო თვიური პრემია შეადგენს 56 ლარს (შესაბამისად, 63 ლარს), ხოლო შესწორებული შერჩევითი სტანდარტული გადახრა ექიმისა და მედდისათვის ტოლია 3 ლარისა და 5.75 ლარის. $\alpha = 0.05$ მნიშვნელოვნების დონით შეიძლება თუ არა დავასკვნათ, რომ მედდის სადაზღვევო პრემია უფრო მეტია ვიდრე ექიმის? ისარგებლეთ P -მნიშვნელობის მეთოდით.
11. ქვემოთ მოყვანილია დრო, რომელიც სჭირდება 6 – 6 თეთრ და ყავისფერ თაგვს, რათა ისწავლოს მარტივ ლაბირინთში გარბენა. $\alpha = 0.05$ მნიშვნელოვნების დონით გაარკვიეთ იწვევს თუ არა თაგვის ფერი ლაბირინთში გარბენის შესასწავლად საჭირო დროში განსხვავებას? ააგეთ 95%-იანი ნდობის ინტერვალი საშუალოთა სხვაობისათვის.

თეთრი თაგვი	18	24	20	13	15	12
ყავისფერი თაგვი	25	16	19	14	16	10

თავი XI

ორამოკრეფიანი ამოცანები პოპულაცია-თა დისპერსიებისათვის

ჰიპოთეზათა შემოწმება ორი დამოუკიდებელი ნორმალური პოპულაციის დისპერსიების შესახებ

$\xi \cong N(a_1, \sigma_1^2)$ და $\eta \cong N(a_2, \sigma_2^2)$ დამოუკიდებელია; ორივე პოპულაციის პარამეტრები უცნობია.

კრიტერიუმი:

ორმხრივი	მარჯვენა ცალმხრივი	მარცხენა ცალმხრივი
$H_0 : \sigma_1^2 / \sigma_2^2 = 1$	$H_0 : \sigma_1^2 / \sigma_2^2 = (\text{ან} \le)1$	$H_0 : \sigma_1^2 / \sigma_2^2 = (\text{ან} \ge)1$
$H_1 : \sigma_1^2 / \sigma_2^2 \ne 1$	$H_1 : \sigma_1^2 / \sigma_2^2 > 1$	$H_1 : \sigma_1^2 / \sigma_2^2 < 1$

ჰიპოთეზა: $H_0 : \sigma_1^2 / \sigma_2^2 = 1$;

მნიშვნელოვნების დონე: α.

კრიტერიუმის სტატისტიკა: $F = \dfrac{S_1^{'2} / \sigma_1^2}{S_2^{'2} / \sigma_2^2} \cong F(n-1, m-1)$ – ფიშერის განაწილება $n-1$ და $m-1$ თავისუფლების ხარისხებით (სადაც n და m შერჩევების მოცულობებია).

კრიტერიუმის მნიშვნელობა **T.V.**: $f = \dfrac{s_1^{'2}}{s_2^{'2}}$.

ალტერნატივა	კრიტიკული არე **C.R.** (H_0-ის უარყოფის არე):
$H_1 : \sigma_1^2 / \sigma_2^2 > 1$	$f \ge F_{n-1,m-1,\alpha}$,
$H_1 : \sigma_1^2 / \sigma_2^2 < 1$	$0 \le f \le F_{n-1,m-1,1-\alpha}$,
$H_1 : \sigma_1^2 / \sigma_2^2 \ne 1$	$0 \le f \le F_{n-1,m-1,1-\alpha/2}$ ან $f \ge F_{n-1,m-1,\alpha/2}$

(სადაც $F_{k,l,\alpha}$ არის k და l თავისუფლების ხარისხის მქონე ფიშერის განაწილების ზედა α-კრიტიკული წერტილია ანუ **კრიტიკული მნიშვნელობა C.V.**).

$(1-\alpha)$ **საიმედოობის ნდობის ინტერვალი დისპერსიათა σ_1^2/σ_2^2 ფარდობისათვის**

$$F_{m-1,n-1,1-\alpha/2} \cdot s_1^{'2}/s_2^{'2} < \sigma_1^2/\sigma_2^2 < F_{m-1,n-1,\alpha/2} \cdot s_1^{'2}/s_2^{'2}.$$

დისპერსიათა ტოლობის შესახებ ჰიპოთეზათა შემოწმების გამარტივებული პროცედურა:

ჰიპოთეზა: $H_0 : \sigma_1^2/\sigma_2^2 = 1$;

მნიშვნელოვნების დონე: α.

კრიტერიუმის სტატისტიკის მნიშვნელობა T.V.:

$$\overline{f} = \frac{\max\{s_1^{'2}, s_2^{'2}\}}{\min\{s_1^{'2}, s_2^{'2}\}}.$$

ალტერნატივა	**კრიტიკული არე C.R. (H_0-ის უარყოფის არე)**
$H_1 : \sigma_1^2/\sigma_2^2 > 1$	$\overline{f} \geq F_{k,l,\alpha}$,
$H_1 : \sigma_1^2/\sigma_2^2 < 1$	$0 \leq \overline{f} \leq F_{k,l,1-\alpha}$
$H_1 : \sigma_1^2/\sigma_2^2 \neq 1$	$\overline{f} \geq F_{k,l,\alpha/2}$

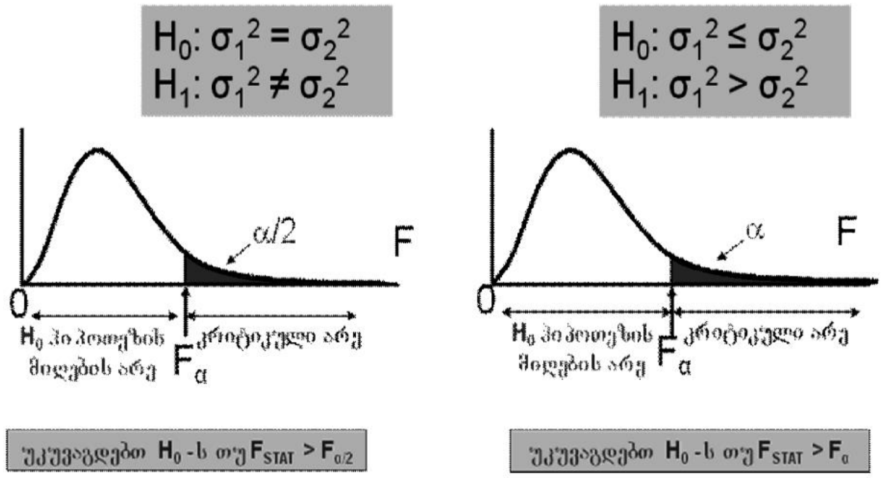
(სადაც $F_{k,l,\alpha}$ არის k და l თავისუფლების ხარისხის მქონე ფიშერის განაწილების ზედა α-კრიტიკული წერტილია ანუ **კრიტიკული მნიშვნელობა C.V.**; $k = n-1$ და $l = m-1$, თუ $s_1^{'2} > s_2^{'2}$ და, პირიქით, $k = m-1$ და $l = n-1$, თუ $s_1^{'2} < s_2^{'2}$).

$(1-\alpha)$ **საიმედოობის ნდობის ინტერვალი დისპერსიათა σ_1^2/σ_2^2 ფარდობისათვის**

$(\overline{f}/F_{n-1,m-1,\alpha/2} =)\ \overline{f} \cdot F_{m-1,n-1,1-\alpha/2} < \sigma_1^2/\sigma_2^2 < \overline{f} \cdot F_{m-1,n-1,\alpha/2}$, თუ $s_1^{'2} > s_2^{'2}$;

$(1/[\overline{f} \cdot F_{n-1,m-1,\alpha/2}] =)\ F_{m-1,n-1,1-\alpha/2}/\overline{f} < \sigma_1^2/\sigma_2^2 < F_{m-1,n-1,\alpha/2}/\overline{f}$, თუ $s_1^{'2} < s_2^{'2}$.

ორამოკრეფიანი ამოცანა (*F* კრიტერიუმი)



ჰიპოთეზები	სტატისტიკა
$H_0: \sigma_1^2 = \sigma_2^2$ $H_1: \sigma_1^2 \neq \sigma_2^2$	$F_{STAT} = S_1^2 / S_2^2$
$H_0: \sigma_1^2 \leq \sigma_2^2$ $H_1: \sigma_1^2 > \sigma_2^2$	

სადაც:

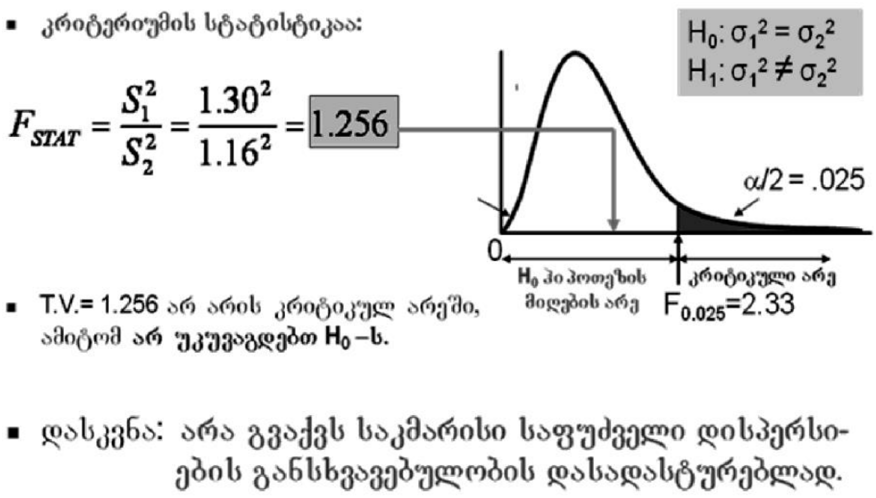
- $s_1^2 = \mathrm{I}$ შერჩევის შესწორებული შერჩევითი დისპერსია (უდიდესი შესწორებული შერჩევითი დისპერსია).
- $n_1 = \mathrm{I}$ შერჩევის მოცულობა.
- $s_2^2 = \mathrm{II}$ შერჩევის შესწორებული შერჩევითი დისპერსია (უმცირესი შესწორებული შერჩევითი დისპერსია).
- $n_2 = \mathrm{II}$ შერჩევის მოცულობა.
- $n_1 - 1 =$ მრიცხველის თავისუფლების ხარისხი.
- $n_2 - 1 =$ მნიშვნელის თავისუფლების ხარისხი.

თქვენ გსურთ შეადაროთ NYSE-ისა და NASDAQ-ის აქციებიდან მოსალოდნელი დივიდენდები შემდეგი მონაცემების მიხედვით:

	NYSE	NASDAQ
რაოდენობა	21	25
საშუალო	3.27	2.53
სტანდარტული გადახრა	1.30	1.16

არის თუ არა $\alpha = 0.05$ მნიშვნელოვნების დონით განსხვავება NYSE - ისა და NASDAQ-ის აქციების დისპერსიებს შორის.

გვაქვს: $a = 0.05$; $n = 21$; $m = 25$; $S_1 = 1.30$; $S_2 = 1.16$; $\bar{x}_1 = 3.27$; $\bar{x}_2 = 2.53$; $df1 = 21 - 1 = 20$ და $df2 = 25 - 1 = 24$ (ვინაიდან $S_1^2 > S_2^2$). ამიტომ $F_{n-1,m-1,\alpha/2} = F_{20,24,0.025} = 2.33$ და $T.V. = \frac{S_1^2}{S_2^2} = \frac{1.30}{1.16} = 1.256$. შესაბამისად, გვექნება შემდეგი სურათი:



მაგალითი 1. მკვლევარს აინტერესებს არის თუ არა განსხვავება მწეველი და არამწეველი ადამიანების პულსის რიცხვთა დისპერსიებს შორის. შემთხვევით შერჩეული 26 მწეველისაგან და 18 არამწეველისაგან შემდგარი ორი შერჩევის შესაბამისი შესწორებული შერჩევითი დისპერსიებია: $s_1^{'2} = 36$, $s_2^{'2} = 10$. $\alpha = 0.05$ მნიშვნელოვნების დონით გვაქვს თუ არა საკმარისი სა-

ფუძველი დავასკვნათ, რომ დისპერსიები განსხვავებულია? ჩათვალეთ, რომ სიდიდე განაწილებულია ნორმალურად.

ამოხსნა. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0:\sigma_1^2=\sigma_2^2$, $H_1:\sigma_1^2\neq\sigma_2^2$. ვისარგებლოთ გამარტივებული პროცედურით. ვინაიდან კრიტერიუმი ორმხრივია, $s_1^{'2}>s_2^{'2}$ და თავისუფლების ხარისხებია 26 – 1 = 25 , 18 – 1 =17, ამიტომ კრიტიკული მნიშვნელობა იქნება: $C.V.=F_{k,l,\alpha/2}=F_{25,17,0.025}=2.56$. გამოვთვალოთ კრიტერიუმის მნიშვნელობა: $T.V.=s_1^{'2}/s_2^{'2}=36/10=3.6$. რადგანაც 3.6 > 2.56, ამიტომ ძირითადი ჰიპოთეზა უნდა უკუვაგდოთ ანუ ჩვენ გვაქვს საკმარისი საფუძველი დავასკვნათ, რომ დისპერსიები განსხვავებულია.

მაგალითი 3. ორი დამოუკიდებელი ნორმალური პოპულაციიდან აღებულია 15-ის ტოლი მოცულობის შერჩევები, რომელთათვისაც: $\overline{x}_n=132.45$, $s_1^{'2}=123$; $\overline{y}_m=128.06$, $s_2^{'2}=95$. შეამოწმეთ ჰიპოთეზა დისპერსიათა ტოლობის შესახებ $\alpha=0.05$ მნიშვნელოვნების დონით და ააგეთ 95%-იანი ნდობის ინტერვალი დისპერსიათა ფარდობისათვის.

ამოხსნა. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0:\sigma_1^2/\sigma_2^2=1$, $H_1:\sigma_1^2/\sigma_2^2>1$. ვისარგებლოთ გამარტივებული პროცედურით. ვინაიდან $s_1^{'2}>s_2^{'2}$, ამიტომ $k=n-1=14$ და $l=m-1=14$. გამოვთვალოთ კრიტერიუმის მნიშვნელობა: $\overline{f}=123/95\approx1.3$. რადგანაც საქმე გვაქვს მარჯვენა ცალმხრივ ალტერნატივასთან, ამიტომ კრიტიკული არეა $[F_{k,l,\alpha},+\infty)=[F_{14,14,0.05},+\infty)=[2.48,+\infty)$. რამდენადაც 1.3 < 2.48, ამდენად ნულოვანი ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს ანუ $\alpha=0.05$ მნიშვნელოვნების დონით შეგვიძლია დავასკვნათ, რომ პოპულაციათა დისპერსიები ტოლია.

ავაგოთ ნდობის ინტერვალი. თანაფარდობიდან $(1-\alpha)\cdot100\%=95\%$ გამომდინარე $\alpha=0.05$, $\alpha/2=0.025$. ამიტომ საძიებელი ინტერვალი იქნება:

$$(\overline{f}/F_{14,14,0.025},\overline{f}\cdot F_{14,14,0.025}),$$

$$(1.3/2.9,1.3\cdot2.9) \text{ ანუ } (0.45,3.77).$$

ამოცანები

1. ისარგებლეთ ფიშერის განაწილების ცხრილებით და იპოვეთ კრიტიკული მნიშვნელობები თუ მოცემულია ორი დამოუკიდებელი ნორმალური შერჩევის შესწორებული შერჩევითი დისპერსიები: ა) $s_1^{'2}=128$, $n=23$; $s_2^{'2}=162$, $m=16$; კრიტერიუმი ორმხრივია და $\alpha=0.01$. ბ) $s_1^{'2}=37$, $n=14$; $s_2^{'2}=89$, $m=25$; კრიტერიუმი მარჯვენა ცალმხრივია და $\alpha=0.01$. გ) $s_1^{'2}=232$, $n=30$; $s_2^{'2}=387$, $m=46$; კრიტერიუმი ორმხრივია და $\alpha=0.05$. დ) $s_1^{'2}=164$, $n=21$; $s_2^{'2}=53$, $m=17$; კრიტერიუმი ორმხრივია და $\alpha=0.1$. ე) $s_1^{'2}=92.8$, $n=11$; $s_2^{'2}=43.6$, $m=11$; კრიტერიუმი მარჯვენა ცალმხრივია და $\alpha=0.05$.

ქვემოთ მოყვანილ ამოცანებში იგულისხმება, რომ ყველა სიდიდე განაწილებულია ნორმალურად:

3. პედაგოგის მტკიცებით, იმ შემთხვევაში, როცა ლექციების კურსი შერწყმულია კომპიუტერული დამუშავების კომპონენტთან, სტუდენტთა საგამოცდო ქულების დისპერსია უფრო დიდია, ვიდრე კომპიუტერული დამუშავების გარეშე. შემთხვევით შეირჩა სტუდენტთა ორი ჯგუფი. საგამოცდო ქულების შესწორებული შერჩევითი დისპერსია სტუდენტებისათვის, რომელთა სალექციო კურსი შერწყმული იყო კომპიუტერული დამუშავების კომპონენტთან აღმოჩნდა 103, ხოლო მეორე ჯგუფისათვის კი 73. თითოეული შერჩევა შედგება 20 – 20 სტუდენტისაგან. $\alpha=0.05$ მნიშვნელოვნების დონით შეიძლება თუ არა დავადასტუროთ პედაგოგის მტკიცებულება?
5. საგადასახადო ინსპექტორის მტკიცებით ორ A და B ქალაქში დაბეგვრისაგან თავისუფალი შემოსავლების დისპერსიები განსხვავებულია. ქვემოთ მოყვანილია ამ ქალაქებში დაბეგვრისაგან თავისუფალი შემოსავლების შერჩევები (შემოსავლები გაზომილია მილიონ დოლარებში). $\alpha=0.05$ მნიშვნელოვნების დონით შეიძლება თუ არა დავადასტუროთ ინსპექტორის მტკიცებულება?

ქალაქი A	**ქალაქი B**
113 25 44 31	82 295 12 20

22	23	11	19		11	50	68	16
14	23	19	5		5	12	81	4
8	30	7	2		15	9	2	5

7. პედიატრიის მტკიცებით ახალშობილი ვაჟების სიმაღლეთა ცვალებადობა განსხვავდება ახალშობილი გოგონების სიმაღლეთა ცვალებადობისაგან. შემთხვევით შერჩეული 15 ახალშობილი ვაჟის სიმაღლეთა შესწორებული სტანდარტული გადახრა აღმოჩნდა 1.3 დიუმი (1 დიუმი = 2.54 სმ), ხოლო შემთხვევით შერჩეული 15 ახალშობილი გოგონასათვის კი 0.9 დიუმი. $\alpha = 0.1$ მნიშვნელოვნების დონით შეიძლება თუ არა დავადასტუროთ პედიატრიის მტკიცებულება?

9. მკვლევარის მტკიცებით არტერიული წნევის ცვალებადობა ჭარბწონიან ადამიანებში უფრო დიდია ვიდრე ნორმალურწონიან ადამიანებში. შემთხვევით შერჩეულ 28 ჭარბწონიან ადამიანში არტერიული წნევის შესწორებული სტანდარტული გადახრა აღმოჩნდა 6.2 მმ. ვერცხ. სვ., ხოლო 25 ნორმალურწონიან ადამიანში კი 2.7 მმ. ვერცხ. სვ.. $\alpha = 0.01$ მნიშვნელოვნების დონით შეამოწმეთ მკვლევრის მტკიცებულება.

11. მკვლევარს სურს შეაფასოს ორი პოპულარული A და B დიეტის შედეგად პაციენტთა მიერ დაკარგული წონის დისპერსია. A დიეტის მიმდევარი შემთხვევით შერჩეული 10 ადამიანის მიერ ერთ თვეში დაკარგული წონების შესწორებული სტანდარტული გადახრა აღმოჩნდა 6.3 ფუნტი (1 ფუნტი = 453.6 გრ), ხოლო B დიეტის მიმდევარი შემთხვევით შერჩეული 12 ადამიანის კი 4.8 ფუნტი. $\alpha = 0.05$ მნიშვნელოვნების დონით, შეგვიძლია თუ არა დავასკვნათ, რომ A დიეტის შემთხვევაში დაკარგული წონის დისპერსია უფრო დიდია ვიდრე B დიეტის შემთხვევაში?

13. $\alpha = 0.1$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ დენვერში მაღალი შენობების სიმაღლეების დისპერსია ტოლია დეტროიტში მაღალი შენობების სიმაღლეების დისპერსიის, თუ სახლების შესაბამისი შერჩევების ფუტებში გაზომილი (1 ფუტი = 30.48 სმ) სიმაღლეებია:

დენვერი				დეტროიტი		
714	504	404		620	562	534
698	438	544		472	448	436
408				430	420	

15. ქვემოთ მოყვანილია A და B ტიპის მტვერსასრუტების შერჩევების წონები (გაზომილი ფუნტებში). $\alpha = 0.05$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა დისპერსიების განსხვავებულობის შესახებ.

A ტიპი						**B ტიპი**	
21	16	23	13	18		24	12
17	17	16	15	17		15	13
15	17	16	20	20		11	
17	18						

თავი XII

სტატისტიკური დასკვნები დაწყვილებული მონაცემებისათვის

სტატისტიკური დასკვნები დაწყვილებულ მონაცემთა საშუალოების სხვაობებისათვის (დამოკიდებული შერჩევები)

მონაცემები შედგება n დამოუკიდებლად შერჩეული წყვილისაგან (X_1,Y_1), . . . , (X_n,Y_n) $D_i = X_i - Y_i, i=1,...,n$. $D_1,...,D_n$ წარმოადგენს დამოუკიდებელ ნორმალურად განაწილებულ შემთხვევით სიდიდეთა მიმდევრობას უცნობი პარამეტრებით: $a_D = a_1 - a_2$ და σ_D^2 (აქ $X_1,...,X_n$ და $Y_1,...,Y_m$ – დამოკიდებული შერჩევებია)

$$\overline{D}_n := \frac{1}{n}\sum_{i=1}^{n} D_i\,,\; S_D^{'2} = \frac{1}{n-1}\sum_{i=1}^{n}(D_i - \overline{D}_n)^2 = \frac{1}{n-1}[\sum_{i=1}^{n} D_i^2 - \frac{1}{n}(\sum_{i=1}^{n} D_i)^2]\,;$$

$$\overline{d}_n := \frac{1}{n}\sum_{i=1}^{n} d_i = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i),$$

$$s_D^{'2} = \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \overline{d}_n)^2 = \frac{1}{n-1}[\sum_{i=1}^{n} d_i^2 - \frac{1}{n}(\sum_{i=1}^{n} d_i)^2]\;.$$

ჰიპოთეზა: $H_0 : a_D = 0$.

მნიშვნელოვნების დონე: α.

კრიტერიუმის სტატისტიკა: $T = \dfrac{\overline{D}_n - a_D}{S_D^{'} / \sqrt{n}} \cong T(n-1)$.

კრიტერიუმის მნიშვნელობა **T.V.:** $t = \dfrac{\overline{d}_n - a_D}{s_D^{'} / \sqrt{n}}$, სადაც $\overline{d}_n$-ით აღნიშნულია $\overline{D}_n$ სტატისტიკის დაკვირვებული მნიშვნელობა.

ალტერნატივა კრიტიკული არე C.R. (H_0-ის უარყოფის არე):

$H_1: a_D > 0$ $t \geq t_{n-1,\alpha}$,

$H_1: a_D < 0$ $t \leq -t_{n-1,\alpha}$,

$H_1: a_D \neq 0$ $t \leq -t_{n-1,\alpha/2}$ ან $t \geq t_{n-1,\alpha/2}$

(სადაც $t_{n-1,\alpha}$ არის თავისუფლების $n-1$ ხარისხის მქონე სტიუდენტის განაწილების ზედა α-კრიტიკული წერტილი ანუ **კრიტიკული მნიშვნელობა C.V.**).

P **- მნიშვნელობა:** $P = \begin{cases} 1 - F_T(t), \text{ თუ } H_1: a_1 - a_2 > 0; \\ F_T(t), \text{ თუ } H_1: a_1 - a_2 < 0; \\ 2 \cdot [1 - F_T(|t|)], \text{ თუ } H_1: a_1 - a_2 \neq 0. \end{cases}$

გადაწყვეტილება: თუ $t \in$ **C.R.**, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

P **- მნიშვნელობის მეთოდი:** თუ $P \leq \alpha$, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ α მნიშვნელოვნების დონით, წინააღმდეგ შემთხვევაში, ამის საფუძველი არა გვაქვს.

$(1-\alpha)$ საიმედოობის ნდობის ინტერვალი დაწყვილებული მონაცემების საშუალოთა $a_D = a_1 - a_2$ სხვაობისათვის:

$$\overline{d}_n - \frac{s_D'}{\sqrt{n}} \cdot t_{n-1,\alpha/2} < a_D < \overline{d}_n + \frac{s_D'}{\sqrt{n}} \cdot t_{n-1,\alpha/2}.$$

მაგალითი 1. ფარმაცევტის მტკიცებით კონკრეტული ვიტამინი ზრდის ათლეტის მიერ სიმძიმის აწევის შესაძლებლობას. შემთხვევით შეარჩიეს 8 ათლეტი და შეამოწმეს რა მაქსიმალური სიმძიმის აწევა შეუძლია თითოეულ მათგანს ვიტამინის მიღებამდე და ვიტამინის ერთი თვის განმავლობაში მიღების შემდეგ. იგულისხმეთ, რომ სიდიდე დაახლოებით ნორმალურია და $\alpha = 0.05$ მნიშვნელოვნების დონით შეამოწმეთ ვიტამინის ეფექტურობა.

ათლეტის №	1	2	3	4	5	6	7	8
ვიტ. მიღებამდე x_i	210	230	182	205	262	253	219	216
ვიტ. მიღების შემდეგ y_i	219	236	179	204	270	250	222	216

ამოხსნა. ვიტამინის ეფექტურობა ნიშნავს, რომ ვიტამინის მიღებამდე აწეული სიმძიმე უნდა იყოს არსებითად (მნიშვნელოვნად) ნაკლები, ვიდრე ვიტამინის მიღების შემდეგ. შესაბამისად ძირითადი და ალტერნატიული ჰიპოთეზები იქნება: $H_0 : a_D \geq 0$, $H_1 : a_D < 0$.

გამოვთვალოთ კრიტიკული მნიშვნელობა. აქ თავისუფლების ხარისხი იქნება $d.f. = n - 1 = 8 - 1 = 7$. ამიტომ $C.V. = -t_{n-1,\alpha} = -t_{7,0.05} = -1.895$.

გამოვთვალოთ კრიტერიუმის მნიშვნელობა.

ა) შევავსოთ შემდეგი ცხრილი:

ათლეტის №	ვიტ. მიღებამდე, x_i	ვიტ. მიღების შემდეგ, y_i	$d_i = x_i - y_i$	$d_i^2 = (x_i - y_i)^2$
1	210	219	-9	81
2	230	236	-6	36
3	182	179	3	9
4	205	204	1	1
5	262	270	-8	64
6	253	250	3	9
7	219	222	-3	9
8	216	216	0	0
			$\sum d_i = -19$	$\sum d_i^2 = 209$

ბ) ვიპოვოთ სხვაობათა საშუალო: $\overline{d}_8 := \frac{1}{8} \cdot \sum_{i=1}^{8} d_i = \frac{-19}{8} = -2.375$;

გ) ვიპოვოთ სხვაობათა შესწორებული შერჩევითი სტანდარტული გადახრა:

$$s_D^{'}=\sqrt{\frac{1}{n-1}[\sum_{i=1}^{n}d_i^2-\frac{1}{n}(\sum_{i=1}^{n}d_i)^2]}=\sqrt{\frac{1}{8-1}\cdot[209-\frac{1}{8}\cdot(-19)^2]}=4.84;$$

დ) ვიპოვოთ კრიტერიუმის მნიშვნელობა:

$$t=\frac{\overline{d}_n-a_D}{s_D^{'}/\sqrt{n}}=\frac{-2.375-0}{4.84/\sqrt{8}}=-1.388.$$

ვინაიდან კრიტერიუმის მნიშვნელობა არ ეკუთვნის კრიტიკულ არეს – $C.R.=(-\infty,-t_{7,0.05}]=(-\infty,-1.895]$ (-1.388 > -1.895), ამიტომ ჩვენ არა გვაქვს ნულოვანი ჰიპოთეზის უარყოფის საფუძველი, ანუ არ არსებობს საკმარისი საფუძველი, რათა დავასკვნათ, რომ ვიტამინი აძლიერებს ათლეტის შესაძლებლობებს.

მაგალითი 3. $\alpha=0.05$ მნიშვნელოვნების დონით გაარკვიეთ თუ რა გავლენას ახდენს სქესი იმ კურსდამთავრებულთათვის შეთავაზებული ხელფასის სიდიდეზე, რომელთაც გააჩნიათ ერთნაირი განათლება, ასაკი, სამუშაო გამოცდილება და ააგეთ 90%-იანი ნდობის ინტერვალი საშუალოთა სხვაობისათვის, თუ მათთვის შეთავაზებული დღიური ხელფასების მონაცემებია:

წყვილის ნომერი	1	2	3	4	5	6	7	8	9	10
ქალი X_i	22	17	21	19	26	23	21	31	25	18
ვაჟი Y_i	25	18	27	17	29	25	19	27	36	23
სხვაობები D_i	-3	-1	-6	2	-3	-2	2	4	-11	-5

ამოხსნა. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0:a_D=0$, $H_1:a_D<0$. გამოვთვალოთ სხვაობების შერჩევითი საშუალო და შესწორებული შერჩევითი დისპერსია. გვაქვს:

$$\overline{d}_{10} = \frac{1}{10} \cdot (-3-1-\cdots-5) = -2.3,$$

$$s_D^{'2} = \frac{1}{9} \cdot [(-3+2.3)^2 + \cdots + (-5+2.3)^2] = 19.78 .$$

შესაბამისად, კრიტერიუმის მნიშვნელობა იქნება:

$$T.V. \equiv t = \frac{\overline{d}_n - a_D}{s_D^{'} / \sqrt{n}} = \frac{-2.3-0}{\sqrt{19.78} / \sqrt{10}} = -1.635.$$

რამდენადაც კრიტერიუმი მარცხენა ცალმხრივია, ამიტომ კრიტიკული მნიშვნელობა იქნება: $C.V. = -t_{n-1,\alpha} = -t_{9,0.05} = -1.833$, ხოლო კრიტიკული არეა: $C.R. = (-\infty, -1.833]$. როგროც ვხედავთ, $T.V. \notin C.R.$, ამიტომ ნულოვანი ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს, ანუ არ არსებობს საკმარისი საფუძველი, რათა დავასკვნათ, რომ ხელფასზე გავლენას ახდენს სქესი.

საძიებელი ნდობის ინტერვალი იქნება:

$$(-2.3 - \frac{\sqrt{19.78}}{\sqrt{10}} \cdot t_{9,0.05}, -2.3 + \frac{\sqrt{19.78}}{\sqrt{10}} \cdot t_{9,0.05}),$$

$$(-2.3 - 1.4064 \cdot 1.833, -2.3 + 1.4064 \cdot 1.833),$$

$$(-4.278, 0.278).$$

ამოცანები

ქვემოთ მოყვანილ ამოცანებში იგულისხმება, რომ ყველა სიდიდე განაწილებულია ნორმალურად ან დაახლოებით ნორმალურად:

1. რესტორანთა ქსელის მენეჯერმა თანამშრომლებს შესთავაზა გამაჯანსაღებელი პროგრამა, რათა შეემცირებინა ავადმყოფობის გამო სამუშაოს გაცდენათა რიცხვი. ქვემოთ ნაჩვენებია შემთხვევით შერჩეული 10 მუშაკის მიერ თვის განმავლობაში სამუშაოს გაცდენათა რიცხვი გამაჯანსაღებელი პროგრამის გავლამდე და გავლის შემდეგ. $\alpha = 0.05$ მნიშვნე-

ლოვნების დონით შეგვიძლია თუ არა დავასკვნათ, რომ პროგრამამ გაამართლა?

პროგრამის გავლამდე	2	3	6	7	4	5	3	1	0	0
პროგრამის გავლის შემდეგ	1	4	3	8	3	3	1	0	1	0

3. პროფესორს აინტერესებს რა გავლენას ახდენს სავარჯიშოების შესახებ ფილმი სტუდენტის სავარჯიშოს მიმართ დამოკიდებულების ხარისხზე. შემთხვევით შერჩეული 10 სტუდენტის გამოკითხვის შედეგების მიხედვით (უფრო დიდი რიცხვი გვიჩვენებს დამოკიდებულების უფრო მაღალ ხარისხს), $\alpha = 0.05$ მნიშვნელოვნების დონით, გვაქვს თუ არა საკმარისი საფუძველი დავასკვნათ, რომ ფილმის ნახვის შემდეგ დამოკიდებულება შეიცვალა? ააგეთ 95%-იანი ნდობის ინტერვალი საშუალოთა სხვაობისათვის.

ფილმის ნახვამდე	12	11	14	9	8	6	8	5	4	7
ფილმის ნახვის შემდეგ	13	12	10	9	8	8	7	6	5	5

5. ოფისის მენეჯერს სურს გაარკვიოს გაზრდის თუ არა მდივნების მუშაობის სისწრაფეს საბეჭდი მანქანების შეცვლა კომპიუტერებით. ქვემოთ ნაჩვენებია 10 მდივნის მიერ წუთში დაბეჭდილი სიტყვების რაოდენობა საბეჭდ მანქანაზე და კომპიუტერზე. $\alpha = 0.1$ მნიშვნელოვნების დონით, გვაქვს თუ არა საკმარისი საფუძველი დავასკვნათ, რომ კომპიუტერზე გადასვლა ზრდის დაბეჭდილი სიტყვების რაოდენობას?

მდივანი	1	2	3	4	5	6	7	8	9	10
საბეჭდი მანქანა	63	72	85	97	82	101	73	62	58	75

7. ფეხსაცმლის მწარმოებლის მტკიცებით ის სპორტსმენები, რომლებსაც აცვიათ მისი ფირმის მიერ გამოშვებული ფეხსაცმელი დარბიან უფრო სწრაფად ვიდრე სხვა სპორტსმენები. ქვემოთ მოყვანილი შემთხვევით შერჩეული 8 მორბე-

ნალის სისწარაფის მონაცემების მიხედვით, $\alpha = 0.025$ მნიშვნელოვნების დონით, შეგვიძლია თუ არა დავადასტუროთ მწარმოებლის მტკიცებულება?

მორბენალი	1	2	3	4	5	6	7	8
ფირმის ფეხსაცმელით	8.2	6.3	9.2	8.6	6.8	8.7	8	6.9
სხვა ფეხსაცმელით	7.1	6.8	9.8	8	5.8	8	7.4	8

9. უძრავი ქონების აგენტის მტკიცებით დედაქალაქში სამოთახიანი ბინის ფასები შეიცვალა. ქვემოთ მოყვანილია შემთხვევით შერჩეული 16 ბინის საექსპერტო ფასები (გაზომილი ათასობით ლარებში) შესაბამისად 2003 და 2008 წლებში. $\alpha = 0.05$ მნიშვნელოვნების დონით შეგვიძლია თუ არა დავასკვნათ, რომ ბინების საექსპერტო ფასები შეიცვალა? ისარგებლეთ P -მნიშვნელობის მეთოდით.

2003	**2008**	**2003**	**2008**	**2003**	**2008**	**2003**	**2008**
184	161	116	120	282	297	12	20
414	382	49	52	25	40	37	38
22	22	24	28	141	148	9	9
99	109	50	50	45	56	17	19

თავი XIII

ორამოკრეფიანი ამოცანები ბერნულის სქემაში

ჰიპოთეზათა შემოწმება წარმატებათა ალბათობებისათვის ბერნულის ცდათა ორი დამოუკიდებელი მიმდევრობისათვის

$X_1,...,X_n$ და $Y_1,...,Y_m$ ორი დამოუკიდებელი შერჩევაა ბერნულის კანონით განაწილებული პოპულაციიდან შესაბამისად წარმატების უცნობი p_1 და p_2 ალბათობებით; $q_1=1-p_1$, $q_2=1-p_2$;
$S_1=\sum_{i=1}^{n}X_i$, $S_2=\sum_{i=1}^{m}Y_i$; $\overline{P}_1=\frac{S_1}{n}$, $\overline{P}_2=\frac{S_2}{m}$; $\overline{p}_1=\frac{s_1}{n}=\frac{1}{n}\cdot\sum_{i=1}^{n}x_i$, $\overline{p}_2=\frac{s_2}{m}=\frac{1}{m}\cdot\sum_{i=1}^{m}y_i$; $\overline{P}=\frac{n}{n+m}\overline{P}_1+\frac{m}{n+m}\overline{P}_2$; $\overline{p}=\frac{n}{n+m}\overline{p}_1+\frac{m}{n+m}\overline{p}_2$; $\overline{Q}=1-\overline{P}$; $\overline{q}=1-\overline{p}$.

კრიტერიუმი:

ორმხრივი	მარჯვენა ცალმხრივი	მარცხენა ცალმხრივი
$H_0: p_1=p_2$	$H_0: p_1=p_2$	$H_0: p_1=p_2$
	ან $H_0: p_1\le p_2$	ან $H_0: p_1\ge p_2$
$H_1: p_1\ne p_2$	$H_1: p_1>p_2$	$H_1: p_1<p_2$

ჰიპოთეზა: $H_0: p_1=p_2$;

მნიშვნელოვნების დონე: α;

კრიტერიუმის სტატისტიკა: $Z=\frac{(\overline{P}_1-\overline{P}_2)-(p_1-p_2)}{\sqrt{\overline{P}\,\overline{Q}(1/n+1/m)}}\overset{as}{\cong}N(0,1)$;

კრიტერიუმის მნიშვნელობა **T.V.:** $z=\frac{(\overline{p}_1-\overline{p}_2)-(p_1-p_2)}{\sqrt{\overline{p}\,\overline{q}(1/n+1/m)}}$;

ალტერნატივა კრიტიკული არე C.R. (H_0-ის უარყოფის არე):

$H_1: p_1 > p_2$ $\quad z \ge z_\alpha$,

$H_1: p_1 < p_2$ $\quad z \le -z_\alpha$,

$H_1: p_1 \ne p_2$ $\quad z \le -z_{\alpha/2}$ ან $z \ge z_{\alpha/2}$

(სადაც z_α არის $N(0,1)$-ის ზედა α-კრიტიკული წერტილი ანუ **კრიტიკული მნიშვნელობა C.V.**).

P **-მნიშვნელობა:** $P = \begin{cases} 1-\Phi(z), \text{ თუ } H_1: a_1 - a_2 > 0; \\ \Phi(z), \text{ თუ } H_1: a_1 - a_2 < 0; \\ 2\cdot[1-\Phi(|z|)], \text{ თუ } H_1: a_1 - a_2 \ne 0. \end{cases}$

გადაწყვეტილება: თუ $z \in$ **C.R.**, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

P **-მნიშვნელობის მეთოდი:** თუ $P \le \alpha$, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ α მნიშვნელოვნების დონით, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

$(1-\alpha)$ **საიმედოობის ნდობის ინტერვალი** $p_1 - p_2$ **სხვაობისათვის:**

$$((\overline{p}_1 - \overline{p}_2) - z_{\alpha/2}\sqrt{\frac{\overline{p}_1\overline{q}_1}{n} + \frac{\overline{p}_2\overline{q}_2}{m}}, (\overline{p}_1 - \overline{p}_2) + z_{\alpha/2}\sqrt{\frac{\overline{p}_1\overline{q}_1}{n} + \frac{\overline{p}_2\overline{q}_2}{m}})$$

(სადაც $\overline{q}_1 = 1 - \overline{p}_1$, $\overline{q}_2 = 1 - \overline{p}_2$).

შეზღუდვები: $n\overline{p}_1, n\overline{q}_1, m\overline{p}_2, m\overline{q}_2 \ge 5$.

მაგალითი 1. მკვლევარმა დაადგინა, რომ 34 მცირე კლინიკიდან 12-ში, ისევე როგორც 24 დიდი კლინიკიდან 17-ში, დაკავებული საწოლების წილი ნაკლებია 80%-ზე. $\alpha = 0.05$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ პროპორციები იმ მცირე და დიდ კლინიკებში, რომლებშიც დაკავებული საწოლების წილი ნაკლებია 80%-ზე – ერთი და იგივეა. ააგეთ 95%-იანი ნდობის ინტერვალი პროპორციათა სხვაობისათვის.

ამოხსნა. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0: p_1 = p_2$, $H_1: p_1 \ne p_2$. აღვნიშნოთ $\overline{p}_1$ (შესაბამი-

სად, $\overline{p}_2$) სიმბოლოთი პროპორცია იმ მცირე (შესაბამისად, იმ დიდი) კლინიკების, რომლებშიც დაკავებული საწოლების წილი ნაკლებია 80%-ზე. გვაქვს:

$$\overline{p}_1 = \frac{s_1}{n} = \frac{12}{34} = 0.35 \quad \text{და} \quad \overline{p}_2 = \frac{s_2}{m} = \frac{17}{24} = 0.71 .$$

ამიტომ $\overline{p} = \frac{n}{n+m}\overline{p}_1 + \frac{m}{n+m}\overline{p}_2 = \frac{34}{34+24}\cdot 0.35 + \frac{24}{34+24}\cdot 0.71 = 0.5$, $\overline{q} = 1 - \overline{p} = 0.5$.

ვიპოვოთ კრიტიკული მნიშვნელობა. რადაგანაც კრიტერიუმი ორმხრივია და $\alpha = 0.05$, კრიტიკული მნიშვნელობები იქნება: $C.V. = \pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96$.

გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$T.V. = z = \frac{(\overline{p}_1 - \overline{p}_2) - (p_1 - p_2)}{\sqrt{\overline{p}\overline{q}(1/n + 1/m)}} = \frac{(0.35 - 0.71) - 0}{\sqrt{0.5\cdot 0.5\cdot (1/34 + 1/24)}} = -2.7 .$$

ვინაიდან -2.7<-1.96 (ე.ი. $T.V. \in C.R.$), ამიტომ ჩვენ უნდა უკუვაგდოთ ნულოვანი ჰიპოთეზა ანუ ჩვენ გვაქვს საკმარისი საფუძველი უარვყოთ ჰიპოთეზა იმის შესახებ, რომ პროპორციები არ განსხვავდება.

ავაგოთ ნდობის ინტერვალი. აქ $\overline{q}_1 = 1 - \overline{p}_1 = 0.65$, $\overline{q}_2 = 1 - \overline{p}_2 = 0.29$. ამიტომ საძიებელი ნდობის ინტერვალი იქნება:

$$((0.35 - 0.71) - 1.96\cdot\sqrt{\frac{0.35\cdot 0.65}{34} + \frac{0.71\cdot 0.29}{24}} ,$$

$$(0.35 - 0.71) + 1.96\cdot\sqrt{\frac{0.35\cdot 0.65}{34} + \frac{0.71\cdot 0.29}{24}}) ;$$

$$(-0.36 - 0.242, -0.36 + 0.242), \ (-0.602, -0.118) .$$

ვინაიდან ნდობის ინტერვალი არ მოიცავს 0-ს, გადაწყვეტილება ისევ იქნება: უკუვაგდოთ ნულოვანი ჰიპოთეზა.

ამოცანები

1. მუშათა დასახლებიდან 150 შემთხვევით შერჩეულ ადამიანს შორის 80 ფილტვებითაა დაავადებული, ხოლო 100 შემთხვევით შერჩეული სოფლის მოსახლიდან კი 30 დაავადებული. $\alpha = 0.05$ მნიშვნელოვნების დონით არის თუ არა გნსხვავება ამ ორ დასახლებაში ფილტვებით დაავადებულ ადამიანთა პროპორციებს შორის?
3. შემთხვევით შერჩეული 100 მომხმარებლიდან 43 ანგარიშსწორებისათვის იყენებს „მასტერქარდს", ხოლო შემთხვევით შერჩეული სხვა 100 მომხმარებლიდან 58 ანგარიშსწორებისათვის იყენებს „ვიზაქარდს". $\alpha = 0.05$ მნიშვნელოვნების დონით არის თუ არა გნსხვავება იმ მომხმარებელთა პროპორციებს შორის, რომლებიც ანგარიშსწორებისათვის იყენებენ სხვადასხვა ტიპის საკრედიტო ბარათებს?
5. გამოკითხვამ აჩვენა, რომ გამოკითხული მამაკაცების 83% ლექციასთან შედარებით უპირატესობას ანიჭებს კომპიუტერულ სწავლებას, ხოლო იგივე მაჩვენებელი ქალებისათვის შეადგენს 75%-ს. თითოეული პოპულაციისათვის შეირჩა 100 – 100 ადამიანი. $\alpha = 0.05$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ მამაკაცებსა და ქალებში არ არსებობს განსხვავება იმ ადამიანების პროპორციებს შორის, რომლებიც ლექციასთან შედარებით უპირატესობას ანიჭებენ კომპიუტერულ სწავლებას. ააგეთ 95%-იანი ნდობის ინტერვალი პროპორციათა სხვაობისათვის.
7. 80 გამოკითხული ამერიკელიდან 55% თვლის, რომ ის მდიდარია, ხოლო 90 გამოკითხული ევროპელიდან 45%. $\alpha = 0.01$ მნიშვნელოვნების დონით არის თუ არა განსხვავება პროპორციებს შორის? ააგეთ 99%-იანი ნდობის ინტერვალი პროპორციათა სხვაობისათვის.
9. 80 გამოკითხული თბილისელიდან 45-ს აქვს კონდიციონერი, ხოლო 120 გამოკითხული ქუთაისელიდან 63-ს აქვს კონდიციონერი. $\alpha = 0.05$ მნიშვნელოვნების დონით არის თუ არა განსხვავება პროპორციებს შორის? ააგეთ 95%-იანი ნდობის ინტერვალი პროპორციათა სხვაობისათვის.

11. 200 გამოკითხული მოზარდიდან 50-ს სჯერა, რომ ომი შეუქცევადია, მაშინ როცა 300 ასაკოვანი ადამიანიდან 93-ს სჯერა, რომ ომი შეუქცევადია. $\alpha = 0.01$ მნიშვნელოვნების დონით არის თუ არა განსხვავება პროპორციებს შორის? ააგეთ 99%-იანი ნდობის ინტერვალი პროპორციათა სხვაობისათვის.

ამოცანები გამოცდისათვის

იგულისხმეთ, რომ სიდიდეები განაწილებულია ნორმალურად ან დაახლოებით ნორმალურად.

13. შემთხვევით შერჩეული 100 – 100 ქირურგისა და სტომატოლოგის საშუალო წლიური შემოსავალი შეადგენს, შესაბამისად, 54107-სა და 58417 ლარს. ორივე შერჩევის შესწორებული სტანდარტული გადახრაა 81 ლარი. $\alpha = 0.1$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ ამ პოპულაციათა საშუალოები არ განსხვავდება. ააგეთ 90%-იანი ნდობის ინტერვალი საშუალოთა სხვაობისათვის.

15. განათლების ინსპექტორს აინტერესებს შეადაროს მოსწავლეებზე დახარჯული თანხების დისპერსიები ქალაქად და სოფლად. შესაბამისი შემთხვევითი შერჩევებიდან მიღებული შესწორებული შერჩევითი დისპერსიებია: $s_1'^2 = 585$, $s_2'^2 = 261$ (შერჩევათა მოცულობებია: $n = 18$, $m = 16$). $\alpha = 0.05$ მნიშვნელოვნების დონით არის თუ არა მნიშვნელოვანი განსხვავება პოპულაციათა დისპერსიებს შორის?

17. კლინიკის ინტენსიური თერაპიის შემთხვევით შერჩეულ 11 ოთახში ხმაურის დონის შესწორებული სტანდარტული გადახრა აღმოჩნდა 4.2 დეციბალი, ხოლო შემთხვევით შერჩეულ არასამკურნალო დანიშნულების 24 ოთახში – 13.2 დეციბალი. $\alpha = 0.1$ მნიშვნელოვნების დონით არის თუ არა მნიშვნელოვანი განსხვავება კლინიკის სამკურნალო და არასამკურნალო დანიშნულების ოთახებში ხმაურის დონის სტანდარტულ გადახრებს შორის?

19. მკვლევარის ვარაუდით ქარხნის მუშაკთა მიერ ავადმყოფობის გამო წელიწადში გაცდენილი დღეების რაოდენობის ცვალებადობა უფრო მეტია ვიდრე კლინიკის მუშაკთათვის. დიდი

კლინიკის შემთხვევით შერჩეული 42 მუშაკისათვის შესწორებული სტანდარტული გადახრა აღმოჩნდა 2.1 დღე, ხოლო დიდი ქარხნის შემთხვევით შერჩეული 65 მუშაკისათვის – 3.2 დღე. შეამოწმეთ ჰიპოთეზა $\alpha = 0.1$ მნიშვნელოვნების დონით.

21. შემთხვევით შერჩეულ 25 დღიან პერიოდში A და B საქონელზე არსებული ფასების მონაცემების მიხედვით, $\alpha = 0.1$ მნიშვნელოვნების დონით, შეიძლება თუ არა დავასკვნათ, რომ A საქონელი უფრო ძვირია, ვიდრე B საქონელი?

A საქონელი

78	82	68	67	68
75	73	75	64	68
62	73	77	78	79
74	72	73	78	68
73	79	82	71	66

B საქონელი

70	74	73	60	77
71	72	71	74	76
71	80	65	70	83
67	76	75	62	65
66	65	77	66	64

23. დედაქალაქის შემთხვევით შერჩეული 16 ოჯახის საშუალო წლიური შემოსავალი არის 54356 ლარი, ხოლო შესწორებული სტანდარტული გადახრა კი 8256 ლარი. გარეუბნის შემთხვევით შერჩეული 12 ოჯახის საშუალო წლიური შემოსავალი შეადგენს 46512 ლარს, ხოლო შესწორებული სტანდარტული გადახრა ტოლია 1311 ლარის. $\alpha = 0.05$ მნიშვნელოვნების დონით შეიძლება თუ არა დავასკვნათ, რომ დედაქალაქის მცხოვრებთა წლიური შემოსავალი მეტია ვიდრე გარეუბანში მცხოვრებლების? ისარგებლეთ P -მნიშვნელობის მეთოდით?

25. შრომის ნაყოფიერების გაზრდის მიზნით ავტონაწილების საამქროს მენეჯერმა გადაწყვიტა სამუშაო ადგილას სამუშაო დროის განმავლობაში ჩაერთო მსუბუქი მუსიკა. შემთხვევით შეარჩიეს რვა მუშა და დათვალეს მათ მიერ გამოშვებული დეტალების რაოდენობა მუსიკის ჩართვამდე და მუსიკის ჩართვიდან ერთი კვირის შემდეგ. ქვემოთ მოყვანილი შესაბამისი მონაცემების მიხედვით, $\alpha = 0.05$ მნიშვნელოვნების დონით, შეუძლია თუ არა მენეჯერს დაასკვნას, რომ მუსიკამ გაზარდა შრომის ნაყოფიერება?

მუშის ნომერი	1	2	3	4	5	6	7	8
მუსიკის ჩართვამდე	6	8	10	9	5	12	9	7
მუსიკის შემდეგ	10	12	9	12	8	13	8	10

27. დედაქალაქის 50 გამოკითხული მცხოვრებიდან 32 აქვს მიკროტალღოვანი ქურა, ხოლო გარეუბანში 60 გამოკითხული მცხოვრებიდან 24 აქვს მიკროტალღოვანი ქურა. $\alpha = 0.05$ მნიშვნელოვნების დონით შეგვიძლია თუ არა ვთქვათ, რომ პროპორციები ერთი და იგივეა? ააგეთ 95%-იანი ნდობის ინტერვალი პროპორციათა სხვაობისათვის.

თავი XIV

თანხმობის კრიტერიუმები

ხი-კვადრატ თანხმობის კრიტერიუმის ფორმულა:

$$\chi^2 = \sum \frac{(O-E)^2}{E},$$

სადაც O – **დაკვირვებული სიხშირეა** (observed frequency), E – **მოსალოდნელი სიხშირე** (expected frequency), ხოლო თავისუფლების ხარისხი ტოლია კატეგორიათა (კლასთა) რაოდენობას (n) გამოკლებული 1 და გამოკლებული პოპულაციის უცნობი პარამეტრების რაოდენობა r ($d.f. = n-1-r$).

შეზღუდვები: 1. მონაცემები მიღებულია შემთხვევითი შერჩევიდან; 2. თითოეული კატეგორიის მოსალოდნელი სიხშირე უნდა იყოს მეტი ან ტოლი 5-ის. თუ უკანასკნელი პირობა არ სრულდება, უნდა მოხდეს კატეგორიის გაერთიანება სხვა კატეგორიასთან ისე, რომ გაერთიანებული კატეგორიის მოსალოდნელი სიხშირე გახდეს მეტი ან ტოლი 5-ის (თუ $n = 2$, მაშინ ყველა $e_1 \geq 10$).

კრიტერიუმის მნიშვნელობა T.V. $\equiv \chi^2 = \sum_i \frac{(o_i - e_i)^2}{e_i}$.

კრიტერიუმი მარჯვენა ცალმხრივია.

კრიტიკული მნიშვნელობა $C.V. = \chi^2_{k-1,\alpha}$.

კრიტიკული არე C.R. (H_0**-ის უარყოფის არე) =** $[\chi^2_{k-1,\alpha}, +\infty)$

P **- მნიშვნელობა =** $P\{\chi^2(k-1) > T.V.\}$

გადაწყვეტილება: თუ კრიტერიუმის მნიშვნელობა მეტია ან ტოლი კრიტიკულ მნიშვნელობაზე (T.V. $\geq$C.V.), მაშინ ნულოვან

ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

P **- მნიშვნელობის მეთოდი:** თუ $P \le \alpha$, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

მაგალითი 1. მაღაზიის მენეჯერს აინტერესებს ანიჭებს თუ არა მომხმარებელი უპირატესობას ლიმონათის ხუთი განსხვავებული გემოდან რომელიმეს. 100 შემთხვევით შერჩეული მომხმარებლიდან 32-მა აარჩია ალუბლის, 28-მ მარწყვის, 16-მა ფორთოხლის, 14-მა ფეიხოას და 10-მა ყურძნის გემოს მქონე ლიმონათი. $\alpha = 0.05$ მნიშვნელოვნების დონით გვაქვს თუ არა საკმარისი საფუძველი უკუვაგდოთ ჰიპოთეზა იმის შესახებ, რომ მომხმარებელი არ ანიჭებს უპირატესობას არც ერთ გემოს?

ამოხსნა. იმ შემთხვევაში, თუ მომხმარებელი არ ანიჭებს უპირატესობას არც ერთ გემოს, მაშინ უნდა ველოდოთ, რომ ყველა გემოს სიხშირე ტოლია, ანუ ჩვენს შემთხვევაში მოსალოდნელი სიხშირეებია: 100/5 = 20. შესაბამისად, ჩვენ გვაქვს სიხშირეთა განაწილების შემდეგი ცხრილი:

სიხშირე	**ალუბალი**	**მარწყვი**	**ფორთოხალი**	**ფეიხოა**	**ყურძენი**
დაკვირვებული, o_i	32	28	16	14	10
მოსალოდნელი, e_i	20	20	20	20	20

ნაბიჯი 1. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები:

H_0 : მომხმარებლებისათვის მნიშვნელობა არა აქვს გემოს.

H_1 : მომხმარებლებისათვის მნიშვნელობა აქვს გემოს.

ნაბიჯი 2. ვიპოვოთ კრიტიკული მნიშვნელობა. აქ თავისუფლების ხარისხია 5−1=4, $\alpha = 0.05$. შესაბამისად, ხი-კვადრატ განაწილების ცხრილიდან ვპოულობთ, რომ $C.V. = \chi^2_{k-1,\alpha} = \chi^2_{4,0.05} = 9.488$.

ნაბიჯი 3. გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$\chi^2 = \sum_i \frac{(o_i - e_i)^2}{e_i} = \frac{(32-20)^2}{20} + \frac{(28-20)^2}{20} + \frac{(16-20)^2}{20} +$$

$$+ \frac{(14-20)^2}{20} + \frac{(10-20)^2}{20} = 18.$$

ნაბიჯი 4. გადაწყვეტილების მიღება: ვინაიდან 18 > 9.488, ამიტომ ნულოვანი ჰიპოთეზა უნდა უკუვაგდოთ, ანუ ჩვენ გვაქვს საკმარისი საფუძველი რათა ჩავთვალოთ, რომ მომხმარებლებისათვის მნიშვნელობა აქვს ლიმონათის გემოს.

მაგალითი 3. უნივერსიტეტის ეკოლოგიური კლუბის მენეჯერის მტკიცებით კლუბის 10%-ს შეადგენს პირველკურსელები, 20% – მეორეკურსელია, 40% – მესამეკურსელი და 30% კი მეოთხეკურსელი. მიმდინარე წელს კლუბში ირიცხება 14 პირველკურსელი, 19 მეორეკურსელი, 51 მესამეკურსელი და 16 მეოთხეკურსელი. $\alpha = 0.1$ მნიშვნელოვნების დონით შეამოწმეთ მენეჯერის ჰიპოთეზა.

ამოხსნა.

ნაბიჯი 1. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები:

H_0: ეკოლოგიური კლუბის შემადგენლობის 10%, 20%, 40% და 30% შესაბამისად არის I, II, III და IV კურსელი.

H_1: განაწილება არ არის ისეთი როგორც ნულოვან ჰიპოთეზაშია.

ნაბიჯი 2. ვიპოვოთ კრიტიკული მნიშვნელობა: $C.V. = \chi^2_{k-1,\alpha} =$ $= \chi^2_{3,0.1} = 6.251.$

ნაბიჯი 3. გამოვთვალოთ კრიტერიუმის მნიშვნელობა. მოსალოდნელ სიხშირეთა მნიშვნელობებია:

$$e_1 = 100 \cdot 0.1 = 10,\ e_2 = 100 \cdot 0.2 = 20,\ e_3 = 100 \cdot 0.4 = 40,\ e_4 = 100 \cdot 0.3 = 30.$$

ამიტომ

$$\chi^2 = \sum_i \frac{(o_i - e_i)^2}{e_i} = \frac{(14-10)^2}{10} + \frac{(19-20)^2}{20} + \frac{(51-40)^2}{40} + \frac{(16-30)^2}{30} = 11.208.$$

ნაბიჯი 4. გადაწყვეტილების მიღება: რადგანაც 11.208 > 6.251, ამიტომ ძირითად ჰიპოთეზას უკუვაგდებთ, ანუ ჩვენ გვაქვს საკმარისი საფუძველი უარვყოთ მენეჯერის მტკიცებულება კლუბის შემადგენლობის შესახებ.

ნორმალურობის შემოწმების ხი-კვადრატ კრიტერიუმი

მაგალითი 4. ხი-კვადრატ კრიტერიუმის გამოყენებით შეამოწმეთ ξ სიდიდე, რომლის სიხშირული განაწილება მოყვანილია ქვემოთ. $\alpha = 0.05$ მნიშვნელოვნების დონით არის თუ არა ნორმალურად განაწილებული.

კლასის საზღვრები	სიხშირე
$89.5 - 104.5$	24
$104.5 - 119.5$	62
$119.5 - 134.5$	72
$134.5 - 149.5$	26
$149.5 - 164.5$	12
$164.5 - 179.5$	4
$\sum$	200

ამოხსნა.

ნაბიჯი 1. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები:

H_0 : სიდიდე განაწილებულია ნორმალურად.

H_1 : სიდიდე არ არის განაწილებული ნორმალურად.

ნაბიჯი 2. ვიპოვოთ სიდიდის საშუალო და შესწორებული სტანდარტული გადახრა. ვისარგებლოთ შემდეგი ცხრილით:

საზღვრები	სიხშირე, f_m	შუაწერტილები, x_m	$f_m \cdot x_m$	$f_m \cdot x_m^2$
$89.5 - 104.5$	24	97	2328	225816
$104.5 - 119.5$	62	112	6944	777728
$119.5 - 134.5$	72	127	9144	1161288

134.5 – 149.5	26	142	3692	524264
149.5 – 164.5	12	157	1884	295788
164.5 – 179.5	4	172	688	118336
$\sum$	200		24680	3103220

შესაბამისად გვაქვს: $\bar{x} = 24680/200 = 123.4$ და

$$s^{'} = \sqrt{\frac{1}{n-1}\left[\sum_{m=1}^{k} f_m \cdot x_m^2 - (\sum_{m=1}^{k} f_m \cdot x_m)^2 / n\right]} = \sqrt{\frac{1}{199} \cdot (3103220 - 24680^2 / 200)} = 17.03.$$

ნაბიჯი 3. ვიპოვოთ კლასის საზღვრებში ξ სიდიდის მოხვედრის ალბათობები მისი ნორმალურად განაწილებულობის დაშვების შემთხვევაში. ამისათვის მოვახდინოთ ξ სიდიდის და, შესაბამისად, კლასის საზღვრების სტანდარტიზაცია, ანუ თითოეული საზღვარი გადავიყვანოთ ე. წ. z საზღვარში $z = \frac{x - \bar{x}}{s^{'}}$ ფორმულის მიხედვით, მაშინ საზღვრები 104.5, 119.5, 134.5, 149.5, 164.5 შესაბამისად შეიცვლება საზღვრებით: -1.11, -0.23, 0.65, 1.53, 2.41 და ნორმალური განაწილების ცხრილის გამოყენებით გვექნება:

$$P\{\xi \le 104.5\} = P\{z \le -1.11\} = 0.1335;$$
$$P\{104.5 \le \xi \le 119.5\} = P\{-1.11 \le z \le -0.23\} = 0.2755;$$
$$P\{119.5 \le \xi \le 134.5\} = P\{-0.23 \le z \le 0.65\} = 0.3332;$$
$$P\{134.5 \le \xi \le 149.5\} = P\{0.65 \le z \le 1.53\} = 0.1948;$$
$$P\{149.5 \le \xi \le 164.5\} = P\{1.53 \le z \le 2.41\} = 0.055;$$
$$P\{\xi \ge 164.5\} = P\{z \ge 2.41\} = 0.008.$$

ნაბიჯი 4. ვიპოვოთ მოსალოდნელ სიხშირეთა მნიშვნელობები, რისთვისაც ზემოთ მიღებული ალბათობები გავამრავლოთ 200-ზე. მოსალოდნელ სიხშირეთა მნიშვნელობებია:

$$e_1 = 200 \cdot 0.1335 = 26.7,\ e_2 = 200 \cdot 0.2755 = 55.1,$$
$$e_3 = 200 \cdot 0.3332 = 66.64,\ e_4 = 200 \cdot 0.1948 = 38.96;$$
$$e_5 = 200 \cdot 0.055 = 11;\ e_6 = 200 \cdot 0.008 = 1.6.$$

შენიშვნა: ვინაიდან უკანასკნელი კატეგორიის მოსალოდნელი სიხშირე ნაკლებია 5-ზე, ის უნდა გაერთიანდეს წინა კატეგორიასთან და საბოლოოდ გვექნება სიხშირეთა განაწილების შემდეგი ცხრილი:

o_i	24	62	72	26	16
e_i	26.7	55.1	66.64	38.96	12.6

(აქ $16=12+4$ და $12.6=11+1.6$)

ნაბიჯი 5. გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$\chi^2=\sum_i\frac{(o_i-e_i)^2}{e_i}=\frac{(24-26.7)^2}{26.7}+\frac{(62-55.1)^2}{55.1}+\frac{(72-66.64)^2}{66.64}+$$
$$+\frac{(26-38.96)^2}{38.96}+\frac{(16-12.6)^2}{12.6}=6.797.$$

ნაბიჯი 6. ვიპოვოთ კრიტიკული მნიშვნელობა. ვინაიდან კატეგორიათა რაოდენობა გახდა 5 ამიტომ თავისუფლების ხარისხი იქნება $d.f.=5-1-2=2$. შესაბამისად,
$C.V.=\chi^2_{k-1-r,\alpha}=\chi^2_{2,0.05}=5.991$.

ნაბიჯი 7. გადაწყვეტილების მიღება: რადგანაც 6.797>5.991, ამიტომ ნულოვანი ჰიპოთეზა უნდა უკუვაგდოთ ანუ ξ სიდიდის განაწილება არ შეიძლება ჩაითვალოს დაახლოებით ნორმალურად.

ბოვმან-შელტონის ნორმალურობის კრიტერიუმი: თუ შერჩევის მოცულობა საკმაოდ დიდია, მაშინ, პოპულაციის ნორმალურობის შემთხვევაში, ბოვმან-შელტონის სტატისტიკას $B=n[(a_{შერ})^2/6+(e_{შერ}-3)^2/24]\cong\chi^2(2)$ გააჩნია ხი-კვადრატ განაწილება თავისუფლების ხარისხით 2. შესაბამისად, თუ $T.V.\geq C.V.=\chi^2_{2,\alpha}$, მაშინ პოპულაცია არაა ნორმალური, ხოლო წინააღმდეგ შემთხვევაში ის ნორმალურია.

ამოცანები

1. გადაუდებელი სამედიცინო სამსახურის ხელმძღვანელს სურს დაადგინოს არის თუ არა თანაბრად განაწილებული კვირის განმავლობაში გამოძახებათა რიცხვი. ქვემოთ მოყვანილია შემთხვევით არჩეული კვირის განმავლობაში სამსახურში შემოსულ გამოძახებათა რაოდენობები. $\alpha = 0.05$ მნიშვნელოვნების დონით გვაქვს თუ არა საკმარისი საფუძველი უარვყოთ ჰიპოთეზა იმის შესახებ, რომ კვირის განმავლობაში გამოძახებათა რიცხვი განაწილებულია თანაბრად?

დღე	ორშ.	სამშ.	ოთხშ.	ხუთშ.	პარას.	შაბ.	კვირა
გამოძ. რაოდ.	28	32	15	14	38	43	19

3. ბარის მენეჯერს სურს გაარკვიოს ანიჭებს თუ არა მომხმარებელი უპირატესობას შარბათში რომელიმე ხილის გემოს. ქვემოთ მოყვანილი მონაცემების მიხედვით, $\alpha = 0.01$ მნიშვნელოვნების დონით, შეგვიძლია თუ არა დავასკვნათ, რომ მომხმარებელთა მიერ შეკვეთილ შარბათში ხილის გემო განაწილებულია თანაბრად?

გემო	ლიმონი	ფორთოხალი	მარწყვი	ფეიხოა
შეკვ. რაოდ.	12	24	19	9

5. აშშ-ის წითელი ჯვრის საზოგადოების მონაცემებით ამერიკელთა 42%-ს აქვს O ჯგუფის სისხლი, 44%-ს – A ჯგუფის სისხლი, 10%-ს – B ჯგუფის სისხლი და 4%-ს – AB ჯგუფის სისხლი. გარკვეული საგრაფოს სამედიცინო ექსპერტის მოსაზრებით მის საგრაფოში სისხლის ჯგუფის განაწილება იგივეა რაც ზოგადნაციონალური. ამ საგრაფოდან შემთხვევით შერჩეული 200 ადამიანის სისხლის ჯგუფის მონაცემების მიხედვით, $\alpha = 0.1$ მნიშვნელოვნების დონით, შეამოწმეთ ექსპერტის მოსაზრება.

ჯგუფი	A	O	B	AB
სიხშირე	58	65	55	22

7. კვლევის თანახმად მომხმარებელთა 53% არჩევს ნაღდი ფულის გადახდას სალაროში, 30% სარგებლობს ჩეკებით, 16%

იყენებს საკრედიტო ბარათებს და 1%-სათვის არა აქვს მნიშვნელობა რა ფორმით გადაიხდის. დიდი მაღაზიის მეპატრონემ შემთხვევით შეარჩია 800 მყიდველი და შეეკითხა მათ თუ რა ფორმით არჩევდნენ გადახდას. აღმოჩნდა, რომ მათ შორის 400 იხდის ნაღდი ფულით, 210 იხდის ჩეკებით, 170 სარგებლობს საკრედიტი ბარათებით და 20-სათვის მნიშვნელობა არა აქვს გადახდის ფორმას. $\alpha = 0.01$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ ამ მაღაზიის მომხმარებლების გადახდის ფორმებთან დამოკიდებულება ემთხვევა კვლევის შედეგებს.

9. პროგრამული უზრუნველყოფის დეპარტამენტის მენეჯერის მოსაზრებით მისი მომხმარებლების 50% ყიდულობს ტექსტის დასამუშავებელ პროგრამებს, 25% – ელექტრონული გათვლის პროგრამებს და 25% – მონაცემთა ბაზების პროგრამებს. შემთხვევით შერჩეულ მყიდველთა ქვემოთ - მოყვანილი მონაცემების მიხედვით, $\alpha = 0.05$ მნიშვნელოვნების დონით, სწორია თუ არა მენეჯერის მოსაზრება? ისარგებლეთ P -მნიშვნელობის მეთოდით.

პროგრამა	**ტექსტის დამუშ.**	**ელექტრ. გათვ.**	**მონაც. ბაზები**
გაყიდვ. რაოდ.	38	23	19

თავი XV

დამოუკიდებლობის შემოწმების ხი-კვადრატ კრიტერიუმი

მოწმდება ერთი პოპულაციის ორი სხვადასხვა A და B ნიშნის ერთმანეთთან დამოკიდებულების საკითხი. A ნიშანი იყოფა R კატეგორიად, ხოლო B ნიშანი – C კატეგორიად.

ნულოვანი ჰიპოთეზა: H_0: A და B ნიშნები დამოუკიდებელია.

ალტერნატივა: H_1: A და B ნიშნები დამოკიდებელია.

მნიშვნელოვნების დონე: α.

კრიტერიუმის სტატისტიკა: ხი-კვადრატი

კრიტერიუმის მნიშვნელობა $\mathbf{T.V.} \equiv \chi^2 = \sum_{i,j} \frac{(o_{i,j} - e_{i,j})^2}{e_{i,j}}$, სადაც $o_{i,j}$ – დაკვირვებული სიხშირეებია, ხოლო $e_{i,j}$ კი მოსალოდნელი სიხშირეები:

$$e_{i,j} = \frac{\sum_i o_{i,j} \times \sum_j o_{i,j}}{\sum_{i,j} o_{i,j}}.$$

შეზღუდვა: ყველა $e_{i,j} \geq 5$ (წინააღმდეგ შემთხვევაში ვახდენთ დაჯგუფებას), თუ $R = C = 2$, მაშინ ყველა $e_i \geq 10$.

კრიტერიუმი მარჯვენა ცალმხრივია.

კრიტიკული მნიშვნელობა $C.V. = \chi^2_{k,\alpha}$, სადაც თავისუფლების ხარისხი $k = (R-1)(C-1)$.

კრიტიკული არე C.R. (H_0**-ის უარყოფის არე) =** $[\chi^2_{k,\alpha}, +\infty)$.

P **- მნიშვნელობა =** $P\{\chi^2(k) > T.V.\}$, სადაც $k = (R-1)(C-1)$.

გადაწყვეტილება: თუ კრიტერიუმის მნიშვნელობა მეტია ან ტოლი კრიტიკულ მნიშვნელობაზე (T.V. $\geq$C.V.), მაშინ ნულოვან

ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

P **- მნიშვნელობის მეთოდი:** თუ $P \le \alpha$, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

მაგალითი 1. 200 გამოკითხული ექთანიდან (შესაბამისად, ექიმიდან) 100 (შესაბამისად, 50) უპირატესობას ანიჭებს ახალ პროცედურას, 80 (შესაბამისად, 120) – ძველ პროცედურას, ხოლო 20-სათვის (შესაბამისად, 30-სათვის) მნიშვნელობა არა აქვს რომელ პროცედურას გამოიყენებს. $\alpha = 0.05$ მნიშვნელოვნების დონით არის თუ არა ექთანისა და ექიმის შეხედულებებში განსხვავება?

ამოხსნა.

ნაბიჯი 1. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები:

H_0: პროცედურაზე შეხედულება *დამოუკიდებელია* პროფესიისაგან.

H_1: პროცედურაზე შეხედულება *დამოკიდებულია* პროფესიაზე.

ნაბიჯი 2. შევადგინოთ ნიშანთა შეუღლების $R \times C$ ცხრილი:

ჯგუფი	**ამჯობინებს ახალ პროცედ.**	**ამჯობინებს ძველ პროცედ.**	**ერთი და იგივეა**
ექთნები	100	80	20
ექიმები	50	120	30

ამ შემთხვევაში გვაქვს 2 სტრიქონი და 3 სვეტი ($R = 2, C = 3$), ანუ გვაქვს 2×3 შეუღლების ცხრილი. ამ ცხრილის i-ური სტრიქონისა და j-ური სვეტის გადაკვეთაზე მდგომი ელემენტი ((i, j) ბლოკში მდგომი ელემენტი) აღვნიშნოთ $o_{i,j}$-თი. მაგალითად, $o_{1,2} = 80$. შესაბამისად, ნიშანთა შეუღლების ცხრილი სქემატურად ასე გამოისახება:

	სვეტი 1	**სვეტი 2**	**სვეტი 3**
სტრიქონი 1	$o_{1,1}$	$o_{1,2}$	$o_{1,3}$
სტრიქონი 2	$o_{2,1}$	$o_{2,2}$	$o_{2,3}$

ნაბიჯი 3. ვიპოვოთ კრიტიკული მნიშვნელობა. აქ თავისუფლების ხარისხი: $d.f.=(R-1)(C-1)=(2-1)\cdot(3-1)=2$. შესაბამისად, ხი-კვადრატ განაწილების ცხრილიდან ვპოულობთ, რომ: $C.V.=\chi^2_{2,0.05}=5.991$.

ნაბიჯი 4. გამოვთვალოთ სტრიქონებში და სვეტებში მდგომი სიდიდეებისა და ყველა სიდიდის ჯამები.

ჯგუფი	**ამჯობინებს ახალ პროცედ.**	**ამჯობინებს ძველ პროცედ.**	**ერთი და - იგივეა**	
ექთნები	100	80	20	$\sum=200$
ექიმები	50	120	30	$\sum=200$
	$\sum=150$	$\sum=200$	$\sum=50$	$\sum=400$

ნაბიჯი 5. თითოეული $(i,\ j)$ ბლოკისათვის გამოვთვალოთ მოსალოდნელ სიხშირეთა მნიშვნელობები $e_{i,j}$ შემდეგი ფორმულის მიხედვით:

$$\text{მოსალოდნელი მნიშვნელობა} \equiv e_{i,j}=\frac{\text{სტრიქონის ჯამი}\times\text{სვეტის ჯამი}}{\text{მთლიანი ჯამი}}.$$

მაგალითად, $(1,1)$ ბლოკისათვის მოსალოდნელი სიხშირის მნიშვნელობა $e_{1,1}$ იქნება:

$$e_{1,1}=\frac{(\text{I სტრიქონის ჯამი})\times(\text{I სვეტის ჯამი})}{\text{მთლიანი ჯამი}}=\frac{200\cdot150}{400}=75.$$

დანარჩენი მოსალოდნელი სიხშირეები ანალოგიურად იქნება:

$$e_{1,2}=100,\quad e_{1,3}=25,\quad e_{2,1}=75,\quad e_{2,2}=100,\quad e_{2,3}=25.$$

თვალსაჩინოებისათვის მოსალოდნელი სიხშირეები შეიძლება ჩავწეროთ ფრჩხილებში ნიშანთა შეუღლების ცხრილის შესაბამის ბლოკში:

ჯგუფი	**ამჯობინებს ახალ პროცედ.**	**ამჯობინებს ძველ პროცედ.**	**ერთი და იგივეა**	
ექთნები	100 (75)	80 (100)	20 (25)	$\sum=200$
ექიმები	50 (75)	120 (100)	30 (25)	$\sum=200$
	$\sum=150$	$\sum=200$	$\sum=50$	$\sum=400$

ნაბიჯი 6. გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$\chi^2 = \sum_{i,j} \frac{(o_{i,j} - e_{i,j})^2}{e_{i,j}} = \frac{(100-75)^2}{75} + \frac{(80-100)^2}{100} + \frac{(20-25)^2}{25} +$$

$$+ \frac{(50-75)^2}{75} + \frac{(120-100)^2}{100} + \frac{(30-25)^2}{25} = 26.67 .$$

ნაბიჯი 7. გადაწყვეტილების მიღება: ვინაიდან $26.67 > 5.991$, ნულოვანი ჰიპოთეზა უნდა უკუვაგდოთ ანუ ჩვენ გვაქვს საკმარისი საფუძველი რათა დავასკვნათ, რომ პროცედურაზე შეხედულება დამოკიდებულია პროფესიაზე – ე. ი. ექთანისა და ექიმის შეხედულება პროცედურაზე განსხვავებულია ერთმანეთისაგან.

მაგალითი 3. მკვლევარს აინტერესებს არის თუ არა კავშირი ადამიანის სქესსა და მის მიერ მიღებული ალკოჰოლის რაოდენობას შორის. შემთხვევით შერჩეული 68 ადამიანისათვის მიღებული მონაცემების მიხედვით, $\alpha = 0.1$ მნიშვნელოვნების დონით, შეუძლია თუ არა მკვლევარს დაასკვნას, რომ მიღებული ალკოჰოლის რაოდენობა დაკავშირებულია სქესთან?

ალკოჰოლის მოხმარება

სქესი	**დაბალი**	**საშუალო**	**მაღალი**	**ჯამი**
მამრობ.	10	9	8	27
მდედრობ.	13	16	12	41
ჯამი	23	25	20	68

ამოხსნა.

ნაბიჯი 1. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები:

H_0: მიღებული ალკოჰოლის რაოდენობა *დამოუკიდებელია* ადამიანის სქესისაგან.

H_1: ალკოჰოლის რაოდენობა *დამოკიდებელია* ადამიანის სქესზე.

ნაბიჯი 2. ვიპოვოთ კრიტიკული მნიშვნელობა. აქ თავისუფლების ხარისხია $d.f. = (R-1)(C-1) = (2-1)\cdot(3-1) = 2$ და $\alpha = 0.1$. შესაბამისად, კრიტიკული მნიშვნელობა იქნება: $C.V. = \chi^2_{2,0.1} = 4.605$.

ნაბიჯი 3. გამოვთვალოთ მოსალოდნელი სიხშირეები. გვაქვს:

$$e_{1,1}=\frac{27\cdot 23}{68}=9.13,\quad e_{1,2}=\frac{27\cdot 25}{68}=9.93,\quad e_{1,3}=\frac{27\cdot 20}{68}=7.94,$$

$$e_{2,1}=\frac{41\cdot 23}{68}=13.87,\quad e_{2,2}=\frac{41\cdot 25}{68}=15.07,\quad e_{2,3}=\frac{41\cdot 20}{68}=12.06.$$

შესაბამისად, დასრულებულ შეუღლების ცხრილს ექნება შემდეგი სახე:

ალკოჰოლის მოხმარება

სქესი	დაბალი	საშუალო	მაღალი	ჯამი
მამრობ.	10 (9.13)	9 (9.93)	8 (7.94)	27
მდედრობ.	13 (13.87)	16 (15.07)	12 (12.06)	41
ჯამი	23	24	20	68

ნაბიჯი 4. გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$\chi^2=\sum_{i,j}\frac{(o_{i,j}-e_{i,j})^2}{e_{i,j}}=\frac{(10-9.13)^2}{9.13}+\frac{(9-9.93)^2}{9.93}+\frac{(8-7.94)^2}{7.94}+$$

$$+\frac{(13-13.87)^2}{13.87}+\frac{(16-15.07)^2}{15.07}+\frac{(12-12.06)^2}{12.06}=0.283$$

ნაბიჯი 5. გადაწყვეტილების მიღება: რამდენადაც 0.283<4.605, ამიტომ ნულოვანი ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს ანუ არ არსებობს საკმარისი საფუძველი ვამტკიცოთ, რომ ალკოჰოლის მოხმარება დამოკიდებულია სქესზე.

ამოცანები

1. სწავლობენ არის თუ არა კავშირი სპორტსა და სისხლის წნევას შორის. შემთხვევით შერჩეული 210 ადამიანის ქვემოთ - მოყვანილი მონაცემების მიხედვით, $\alpha=0.05$ მნიშვნელოვნების დონით, შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ სპორტი და წნევა არ არის დამოკიდებული.

სისხლის წნევა

	დაბალი	საშუალო	მაღალი
სპორტსმენი	34	57	21
არა სპორტსმენი	15	63	20

3. მკვლევარს აინტერესებს არის თუ არა დამოკიდებული ერთმანეთზე მყიდველის ასაკი და შეძენილი ავტომობილის ფასი. შემთხვევით შერჩეული 222 ავტომფლობელის ქვემოთ - მოყვანილი მონაცემების მიხედვით, $\alpha = 0.05$ მნიშვნელოვნების დონით, არის თუ არა ავტომობილის ფასი დამოუკიდებელი ავტომფლობელის ასაკისაგან?

გაყიდვის ფასი

ასაკი	არა უმეტეს 20000 ლ	20000 ლ-დან 30000 ლ-მდე	30000 ლ-დან 40000 ლ-მდე
21–30	16	25	3
31–40	44	23	15
41–50	31	15	18
≥ 51	9	11	12

5. უნივერსიტეტის რექტორს აინტერესებს ახდენს თუ არა გავლენას ლექტორის სამეცნიერო ხარისხი სტუდენტის შეხედულებაზე ამ ლექტორის მუშაობის ხარისხის შესახებ. გამოკითხული სტუდენტების ქვემოთ მოყვანილი მონაცემების მიხედვით, $\alpha = 0.1$ მნიშვნელოვნების დონით, შეუძლია თუ არა რექტორს დაასკვნას, რომ სტუდენტის შეხედულება ლექტორის მუშაობის ეფექტურობის შესახებ დამოკიდებულია ლექტორის სამეცნიერო ხარისხზე?

ხარისხი

შეხედულება	ბაკალავრი	მაგისტრი	დოქტორი
უმაღლესი	14	9	4
საშუალო	16	5	7
დაბალი	3	12	16

7. იკვლევენ კავშირს ფილმის ტიპსა და მაყურებლის ასაკს შორის. ქვემოთ მოყვანილი მონაცემების მიხედვით, $\alpha = 0.1$ მნიშვნელოვნების დონით, არის თუ არა კავშირი ფილმის ტიპსა და მაყურებლის ასაკს შორის?

ასაკი	დოკუმენტური	კომედია	მისტერია
12–20	14	9	8
21–29	15	14	9
30–38	9	21	39

39–47	7	22	17
≥ 48	6	38	12

9. იკვლევდნენ ფეხბურთის თამაშის დროს მაყურებლების მიერ არჩეულ საკვების სახეობას. ქვემოთ მოყვანილი მონაცემების მიხედვით, $\alpha = 0.1$ მნიშვნელოვნების დონით, არის თუ არა მაყურებლის მიერ არჩეული საკვები დამოუკიდებელი მაყურებლის სქესისაგან?

სქესი	„ჰოთ დოგი“	არახისი	ღვეზელი
მამრობითი	12	21	19
მდედრობითი	13	8	25

11. წიგნის გამომცემელს აინტერესებს არის თუ არა განსხვავება მამაკაცებისა და ქალების მიერ წასაკითხად არჩეულ წიგნებს შორის. ქვემოთ მოყვანილი მონაცემების მიხედვით, $\alpha = 0.05$ მნიშვნელოვნების დონით, შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ არჩეული წიგნის ტიპი დამოუკიდებელია მკითხველის სქესისაგან. ისარგებლეთ P -მნიშვნელობის მეთოდით.

სქესი	მისტიკა	რომანი	დეტექტივი
მამრობითი	243	201	191
მდედრობითი	135	149	202

თავი XVI

ერთგვაროვნების შემოწმების ხი-კვადრატ კრიტერიუმი

მოწმდება პოპულაციათა ერთგვაროვნების ჰიპოთეზა, რომელიც ეკვივალენტურია ჰიპოთეზისა იმის შესახებ, რომ ამა თუ იმ ნიშნის პროპორციები სხვადასხვა პოპულაციებში ერთი და იგივეა.

ნულოვანი ჰიპოთეზა: $H_0 : p_1 = p_2 = \cdots = p_R$.

ალტერნატივა: H_1: **ერთი პროპორცია მაინც განსხვავდება დანარჩენებისაგან.**

მნიშვნელოვნების დონე: α ,

კრიტერიუმის სტატისტიკა: ხი-კვადრატი.

კრიტერიუმის მნიშვნელობა $\mathbf{T.V.} \equiv \chi^2 = \sum_{i,j} \frac{(o_{i,j} - e_{i,j})^2}{e_{i,j}}$, სადაც $o_{i,j}$ – დაკვირვებული სიხშირეებია, ხოლო $e_{i,j}$ კი მოსალოდნელი სიხშირეები:

$$e_{i,j} = \frac{\sum_i o_{i,j} \times \sum_j o_{i,j}}{\sum_{i,j} o_{i,j}}.$$

შეზღუდვა: ყველა $e_{i,j} \geq 5$ (წინააღმდეგ შემთხვევაში ხდება დაჯგუფება), თუ $R = C = 2$, მაშინ ყველა $e_i \geq 10$.

კრიტერიუმი მარჯვენა ცალმხრივია.

კრიტიკული მნიშვნელობა $C.V. = \chi^2_{(R-1)(C-1),\alpha}$, სადაც R წარმოადგენს შერჩევათა რაოდენობას, ხოლო C კი კლასების რაოდენობაა.

როდესაც შეუღლების ცხრილის თავისუფლების ხარისხი ერთის ტოლია $(d.f. = 1)$, გამოიყენება კრიტერიუმის მნიშვნელობა:

$$T.V. \equiv \chi^2 = \sum_{i,j} \frac{(|o_{i,j} - e_{i,j}| - 0.5)^2}{e_{i,j}}.$$

კრიტიკული არე C.R. (H_0 **-ის უარყოფის არე) =** $[\chi^2_{k,\alpha}, +\infty)$ **.**

P **- მნიშვნელობა =** $P\{\chi^2(k) > T.V.\}$, სადაც $k = (R-1)(C-1)$ **.**

გადაწყვეტილება: თუ კრიტერიუმის მნიშვნელობა მეტია ან ტოლი კრიტიკულ მნიშვნელობაზე (T.V. $\geq$C.V.), მაშინ ნულოვან ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

P **- მნიშვნელობის მეთოდი:** თუ $P \leq \alpha$, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

მაგალითი 1. მკვლევარმა შემთხვევით შეარჩია 150 დამამთავრებელი კურსის სტუდენტი სამი უმაღლესი სასწავლებლიდან (ინსტიტუტიდან) და თითოეულს დაუსვა შეკითხვა: “დადის თუ არა იგი ინსტიტუტში თავისი ან მშობლების მანქანით“. გამოკითხვის შედეგები მოყვანილია ცხრილის სახით. $\alpha = 0.05$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ იმ სტუდენტების პროპორცია, რომლებიც ინსტიტუტში დადიან საკუთარი ან მშობლების მანქანებით ერთი და იგივეა სამივე ინსტიტუტისათვის.

	ინსტიტუტი I	**ინსტიტუტი II**	**ინსტიტუტი III**	**ჯამი**
კი	18	22	16	56
არა	32	28	34	94
ჯამი	50	50	50	150

ამოხსნა.

ნაბიჯი 1. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები:

H_0 : $p_1 = p_2 = p_3$ (პროპორციები ერთი და იგივეა).

H_1 : სულ ცოტა ერთი პროპორცია მაინც განსხვავდება დანარჩენებისაგან.

ნაბიჯი 2. ვიპოვოთ კრიტიკული მნიშვნელობა. თავისუფლების ხარისხი ტოლია $d.f. = (R-1)(C-1) = (2-1)\cdot(3-1) = 2$. ამიტომ

ხი-კვადრატ განაწილების ცხრილიდან ვპოულობთ, რომ $\alpha = 0.05$-სათვის კრიტიკული მნიშვნელობა იქნება: $C.V. = \chi^2_{2,0.05} = 5.991$.

ნაბიჯი 3. გამოვთვალოთ მოსალოდნელი სიხშირეების მნიშვნელობები. ვისარგებლოთ ფორმულით:

$$\text{მოსალოდნელი სიხშირე} \equiv e_{i,j} = \frac{\text{სტრიქონის ჯამი} \times \text{სვეტის ჯამი}}{\text{მთლიანი ჯამი}}.$$

$$e_{1,1} = \frac{56 \cdot 50}{150} = 18.67,\ e_{1,2} = \frac{56 \cdot 50}{150} = 18.67,\ e_{1,3} = \frac{56 \cdot 50}{150} = 18.67,$$

$$e_{2,1} = \frac{94 \cdot 50}{150} = 31.33,\ e_{2,2} = \frac{94 \cdot 50}{150} = 31.33,\ e_{2,3} = \frac{94 \cdot 50}{150} = 31.33.$$

დასრულებული შეუღლების ცხრილი იქნება:

	ინსტიტუტი I	**ინსტიტუტი II**	**ინსტიტუტი III**	**ჯამი**
კი	18 (18.67)	22 (18.67)	16 (18.67)	56
არა	32 (31.33)	28 (31.33)	34 (31.33)	94
ჯამი	50	50	50	150

ნაბიჯი 4. გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$\chi^2 = \sum_{i,j} \frac{(o_{i,j} - e_{i,j})^2}{e_{i,j}} = \frac{(18-18.67)^2}{18.67} + \frac{(22-18.67)^2}{18.67} + \frac{(16-18.67)^2}{18.67} +$$

$$+ \frac{(32-31.33)^2}{31.33} + \frac{(28-31.33)^2}{31.33} + \frac{(34-31.33)^2}{31.33} = 1.596.$$

ნაბიჯი 5. გადაწყვეტილების მიღება: რადგანაც 1.596 < 5.991, ამიტომ ნულოვანი ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს ანუ არ არსებობს საკმარისი საფუძველი ვამტკიცოთ, რომ პროპორციები განსხვავებულია.

ამოცანები

1. კვლევის თანახმად 6-დან 17-წლამდე მოზარდების 64%-ს არ შეუძლია გადალახოს საბაზო შესაბამისობის ტესტი. მკვლევარს აინტერესებს არის თუ არა ამ კატეგორიის მოსწავლეე-

ბის პროპორცია ერთი და იგივე სხვადასხვა სკოლებში. ტესტირება ჩაუტარდა შემთხვევით შერჩეულ 120-120 მოსწავლეს ოთხი სხვადასხვა სკოლიდან. $\alpha = 0.05$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ იმ მოსწავლეების პროპორცია, რომლებმაც გადალახეს შესაბამისობის ტესტი, ერთი და იგივეა.

	I სკოლა	II სკოლა	III სკოლა	IV სკოლა
გადალახა	49	38	46	34
ვერ გადალახა	71	82	74	86
ჯამი	120	120	120	120

3. სადაზღვევო კომპანიას აინტერესებს იცვლება თუ არა ასაკის მიხედვით პროპორცია იმ მძღოლების, ვინც დალევის შემდეგ ჯდება საჭესთან. გამოკითხულ იქნა 4 ასაკობრივი ჯგუფის 86-86 მძღოლი. $\alpha = 0.05$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ პროპორცია იმ მძღოლების, ვინც კითხვას – „ჯდებით თუ არა ნასვამი საჭესთან“ – უპასუხა „დიახ“ ერთი და იგივეა ყველა ასაკობრივი ჯგუფისათვის.

ასაკი	21-29	30-39	40-49	50-ზე მეტი წლის
დიახ	32	28	26	21
არა	54	58	60	65
ჯამი	86	86	86	86

5. გამოკითხულ იქნა 100-100 ადამიანი ქვეყნის ოთხივე კუთხეში. მათ დაუსვეს კითხვა: „აქვთ თუ არა მუდმივი სამუშაო“. $\alpha = 0.1$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ პროპორციები იმ ადამიანების, რომელთაც აქვთ მუდმივი სამუშაო ტოლია ქვეყნის ოთხივე კუთხეში.

	აღმოსავლეთი	დასავლეთი	ჩრდილოეთი	სამხრეთი
დიახ	43	39	22	28
არა	57	61	78	72
ჯამი	100	100	100	100

7. მსოფლიო ჯანდაცვის ორგანიზაციის მონაცემებით ზოოპარკებში ბავშვების მიერ მიღებული ტრამვების 55% მოდის

მაიმუნის ნაკბენზე. გამოკვლეულ იქნა 4 სხვადასხვა ზოოპარკში ბავშვთა ტრამვების 30-30 შემთხვევა. $\alpha = 0.05$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა პროპორციების ტოლობის შესახებ. ისარგებლეთ P -მნიშვნელობის მეთოდით.

ტრამვა	**ზოოპარკი A**	**ზოოპარკი B**	**ზოოპარკი C**	**ზოოპარკი D**
მაიმუნის კბენა	15	18	13	16
სხვა ტრამვა	15	12	17	14
ჯამი	30	30	30	30

თავი XVII

დაწყვილებული მონაცემები, კორელაცია

ამა თუ იმ ორ მახასიათებელს (ან ორ მაჩვენებელს) შორის კავშირის შესწავლის მიზნით ამ მახასიათებლების დაკვირვებული მნიშვნელობების სიმრავლეს წარმოადგენენ წყვილების სიმრავლის სახით: (x_1, y_1), (x_2, y_2), . . . , (x_n, y_n). ამ მონაცემების მიხედვით აგებენ ე. წ. **გაბნევის დიაგრამას** შემდეგი წესით: სიბრტყეზე, მართკუთხა კოორდინატთა სისტემაში, მონიშნავენ წერტილებს, რომელთა კოორდინატებია (x_i, y_i), $i = 1, 2, ..., n$.

მაგალითი 1. მკვლევარს აინტერესებს დაადგინოს არსებობს თუ არა კავშირი ველოსიპედით წვიმიან ამინდში მომხდარ მსუბუქ ავარიათა რიცხვსა და სიკვდილით (ფატალური შედეგით) დამთავრებულ ავარიათა რიცხვს შორის. ქვემოთ მოყვანილია 10 წლის განმავლობაში წვიმიან ამინდში ველოსიპედით მომხდარი ავარიების მონაცემები:

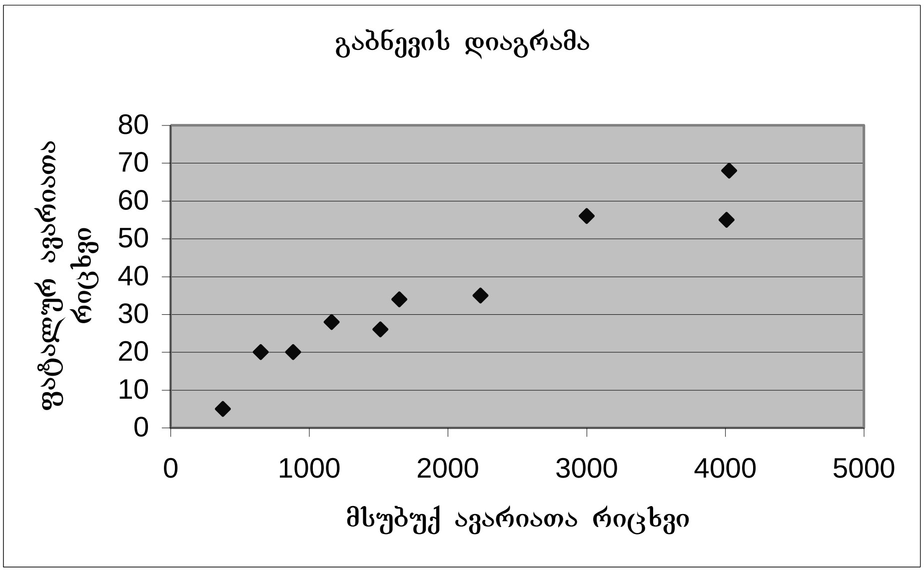
მსუბ. ავარ. რიცხვი, x	376	650	884	1162	1513	1650	2236	3002	4028	4010
ფატ. ავარ. რიცხვი, y	5	20	20	28	26	34	35	56	68	55

ავაგოთ გაბნევის დიაგრამა ამ მონაცემებისათვის.

ამოხსნა.

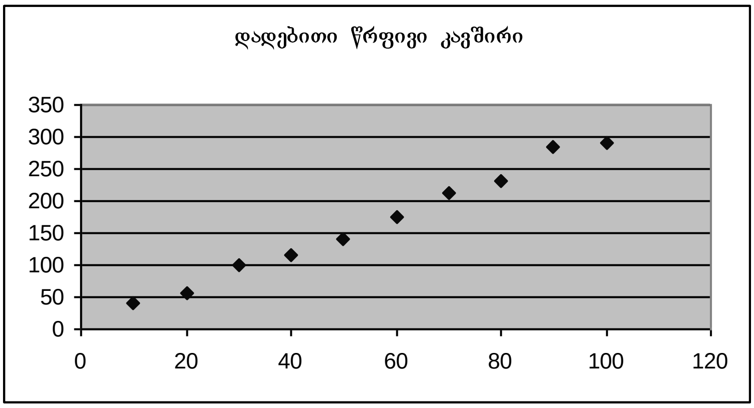
ნაბიჯი 1. სიბრტყეზე ავაგოთ მართკუთხა კოორდინატთა სისტემა და მოვნიშნოთ x და y ღერძები.

ნაბიჯი 2. ავაგოთ წერტილთა წყვილები საკოორდინატო სიბრტყეზე ისე როგორც ეს ქვემოთაა ნაჩვენები:

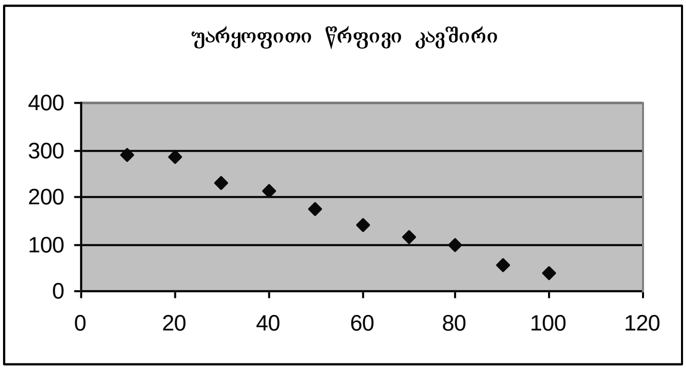


x და y სიდიდეებს შორის შესაძლებელია არსებობდეს რამდენიმე განსხვავებული ტიპის დამოკიდებულება:

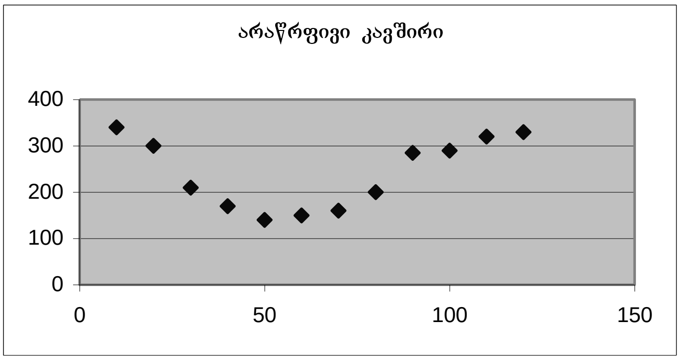
1. **დადებითი წრფივი კავშირი** არსებობს (გვაქვს) იმ შემთხვევაში, როდესაც წერტილები დაახლოებით კონცენტრირებულია აღმავალი სწორი ხაზის ირგვლივ და ერთდროულად ორივე x და y სიდიდის მნიშვნელობები ზრდადია. ასეთი კავშირი გვაქვს მაშინ, როდესაც x სიდიდის ზრდასთან ერთად იზრდება y სიდიდეც.



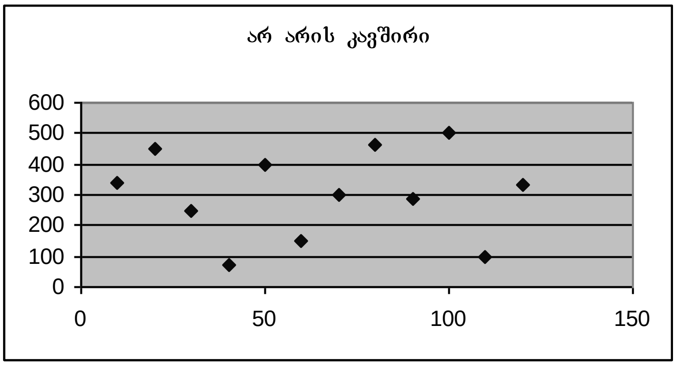
2. უარყოფითი წრფივი კავშირი არსებობს იმ შემთხვევაში, როდესაც წერტილები დაახლოებით კონცენტრირებულია დაღმავალი სწორი ხაზის ირგვლივ. ასეთი კავშირი გვაქვს მაშინ, როდესაც x სიდიდის ზრდასთან ერთად კლებულობს y სიდიდე და პირიქით.



3. არაწრფივი კავშირი არსებობს იმ შემთხვევაში, როდესაც წერტილები დაახლოებით კონცენტრირებულია არაწრფივი (მრუდი) ხაზის (წირის) ირგვლივ. ეს დამოკიდებულება აღიწერება შესაბამისი წირის ბუნებით.



4. არ არის კავშირი იმ შემთხვევაში, როდესაც წერტილები უწესრიგოდაა მიმოფანტული, ანუ არ ჩანს რომ ისინი რაიმე წირის ირგვლივაა კონცენტრირებული.



ზემოთ მოყვანილ მაგალითში x (მსუბუქ ავარიათა რიცხვი) და y (ფატალურ ავარიათა რიცხვი) სიდიდეებს შორის ადგილი აქვს წრფივ დადებით კავშირს. წირს, რომელიც აღწერს კავშირს (დამოკიდებულებას) x და y სიდიდეების მნიშვნელობებს შორის **მისადაგების წირი** ეწოდება. ამ წირის პოვნის საკითხებს სწავლობს მათემატიკური სტატისტიკის ნაწილი, რომელსაც **რეგრესიული ანალიზი** ეწოდება. წირის მოსაძებნად იყენებენ ე. წ. **უმცირეს კვადრატთა მეთოდს.**

თუ სიდიდეებს შორის გვაქვს წრფივი კავშირი $y(x) = ax + b$, მაშინ ამბობენ რომ x და y სიდიდეებს შორის **წრფივი რეგრესიული კავშირია** *და* $y(x) = ax + b$ წრფეს უწოდებენ **რეგრესიის წრფის განტოლებას** (ხოლო ამ ფორმულით გამოთვლილ $y_i^{'} = y(x_i) = ax_i + b$ მნიშვნელობას **პროგნოზი** ეწოდება). უმცირეს კვადრატთა მეთოდი გვაძლევს, რომ ამ შემთხვევაში:

$$a = \frac{n(\sum_{i=1}^{n} x_i y_i) - (\sum_{i=1}^{n} x_i)\cdot(\sum_{i=1}^{n} y_i)}{n(\sum_{i=1}^{n} x_i^2) - (\sum_{i=1}^{n} x_i)^2} \text{ და } b = \frac{(\sum_{i=1}^{n} y_i)\cdot(\sum_{i=1}^{n} x_i^2) - (\sum_{i=1}^{n} x_i)\cdot(\sum_{i=1}^{n} x_i y_i)}{n(\sum_{i=1}^{n} x_i^2) - (\sum_{i=1}^{n} x_i)^2}.$$

შეფასების სტანდარტული შეცდომა ეწოდება სიდიდეს:

$$s_{\text{შეფ}} := \sqrt{\frac{\sum_{i=1}^{n}(y_i - y_i^{'}))^2}{n-2}}.$$

იგი წარმოადგენს დამოკიდებული ცვლადის დაკვირვებული მნიშვნელობების სტანდარტულ გადახრას საპროგნოზო მნიშვნელობებისაგან.

$y(x)$ მნიშვნელობის α მნიშვნელოვნების დონის მქონე **საპროგნოზო ინტერვალი** (ნდობის ინტერვალის ანალოგი) მოიცემა თანაფარდობით:

$$\left(y(x)-t_{n-2,\alpha/2}s_{\text{შეფ}}\sqrt{1+\frac{1}{n}+\frac{n(x-\bar{x})}{n(\sum_{i=1}^{n}x_i^2)-(\sum_{i=1}^{n}x_i)^2}},\; y(x)+t_{n-2,\alpha/2}s_{\text{შეფ}}\sqrt{1+\frac{1}{n}+\frac{n(x-\bar{x})}{n(\sum_{i=1}^{n}x_i^2)-(\sum_{i=1}^{n}x_i)^2}}\right).$$

მაგალითი 3. მენეჯერმა შეაგროვა მონაცემები და დაადგინა, რომ გასამრავლებელი მანქანის ასაკსა და თვეში ტექნიკური მომსახურების ხარჯებს შორის მნიშვნელოვანი კავშირია. რეგრესიის წრფის განტოლებაა $y(x)=8.13x+55.57$. იპოვეთ შეფასების სტანდარტული შეცდომა, თუ მონაცემებია:

მანქანა	A	B	C	D	E	F
ასაკი, x	1	2	3	4	5	6
ხარჯი, y	62	78	70	90	93	103

ამოხსნა.

ნაბიჯი 1. გავაკეთოთ შემდეგი ცხრილი:

x_i	y_i	y_i'	y_i-y_i'	$(y_i-y_i')^2$
1	62			
2	78			
3	70			
4	90			
4	93			
6	103			

ნაბიჯი 2. ვისარგებლოთ რეგრესიის წრფის განტოლებით და გამოვთვალოთ საპროგნოზო მნიშვნელობები თითოეული x_i-სათვის: $y_i' := y(x_i)=8.13x_i+55.57$. გვაქვს: $y_1'=63.7$, $y_2'=71.83$,

$y_3^{'} = 79.96$, $y_4^{'} = y_5^{'} = 88.09$ და $y_6^{'} = 104.35$. შევიტანოთ ეს მონაცემები ცხრილის შესაბამის სვეტში.

ნაბიჯი 3. თითოეულ y_i-ს გამოვაკლოთ შესაბამისი $y_i^{'}$ და შედეგები შევიტანოთ შესაბამის სვეტში:

$$y_1 - y_1^{'} = -1.7,\quad y_2 - y_2^{'} = 6.17,\quad y_3 - y_3^{'} = -9.96,$$
$$y_4 - y_4^{'} = 1.91,\quad y_5 - y_5^{'} - 1.35.$$

ნაბიჯი 4. მიღებული სხვაობები ავიყვანოთ კვადრატში და შევიტანოთ შესაბამის სვეტში.

ნაბიჯი 5. შევკრიბოთ ბოლო სვეტის მონაცემები:

x_i	y_i	$y_i^{'}$	$y_i - y_i^{'}$	$(y_i - y_i^{'})^2$
1	62	63.7	-1.7	2.89
2	78	71.83	6.17	38.0689
3	70	79.96	-9.96	99.2016
4	90	88.09	1.91	3.6481
4	93	88.09	4.91	24.1081
6	103	104.35	-1.35	1.8225
ჯამი				169.7392

ნაბიჯი 6. გამოვთვალოთ შეფასების სტანდარტული შეცდომა:

$$s_{შეფ} := \sqrt{\frac{\sum_{i=1}^{n}(y_i - y_i^{'}))^2}{n-2}} = \sqrt{\frac{169.7392}{6-2}} = 6.51.$$

შეფასების სტანდარტული შეცდომა შეიძლება აგრეთვე გამოითვალოს ფომულით:

$$s_{შეფ} := \sqrt{\frac{\sum_{i=1}^{n} y_i^2 - b\sum_{i=1}^{n} y_i - a\sum_{i=1}^{n} x_i y_i}{n-2}}.$$

ორ სიდიდეს შორის დამოკიდებულების ხარისხის გასაზომად იყენებენ ე. წ. **კორელაციის კოეფიციენტს.** არსებობს სხვადასხვა სახის კორელაციის კოეფიციენტი. ჩვენ აქ მოვიყვანთ ე. წ. **პირსონის მომენტთა ნამრავლის კორელაციის კოეფიციენტს (Pearson product moment correlation coefficient – PPMCC),** რომე-

ლიც დაკავშირებულია ცნობილი სტატისტიკოსის კარლ პირსონის სახელთან.

კორელაციის კოეფიციენტი, გამოთვლილი მონაცემთა შერჩევის მიხედვით, წარმოადგენს ორ სიდიდეს შორის წრფივი დამოკიდებულების ხარისხისა და მიმართულების საზომს. შერჩევითი კორელაციის კოეფიციენტი აღინიშნება r ასოთი, ხოლო პოპულაციის მიხედვით გამოთვლილი კორელაციის კოეფიციენტი აღინიშნება ბერძნული ρ (რო) ასოთი.

ორ x და y სიდიდეს შორის **კოვარიაციის კოეფიციენტი** ($\mathrm{cov}(x,y)$) განიმარტება შემდეგი თანაფარდობით: $\mathrm{cov}(x,y)=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})=\frac{1}{n}\sum_{i=1}^{n}x_iy_i-\bar{x}\cdot\bar{y}$, სადაც $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$, $\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$.

შერჩევითი კორელაციის კოეფიციენტი მოიცემა თანაფარდობით:

$$r=\frac{\mathrm{cov}(x,y)}{s_xs_y}=\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2}}=$$

$$=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}=\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{s_x}\right)\left(\frac{y_i-\bar{y}}{s_y}\right).$$

კორელაციის კოეფიციენტის თვისებები:

1. $-1\le r\le 1$. r-ის დადებითი მნიშვნელობა მიუთითებს სიდიდეებს შორის პოზიტიური დამოკიდებულების არსებობაზე, ხოლო უარყოფითი მნიშვნელობა კი – უარყოფითზე.

2. $|r|=1$ მაშინ და მხოლოდ მაშინ, როცა x და y სიდიდეებს შორის არსებობს წრფივი კავშირი $y_i=ax_i+b$. ამასთანავე, თუ $r=1$, მაშინ $a>0$ და თუ $r=-1$, მაშინ $a<0$.

3. კორელაციის კოეფიციენტი უცვლელი რჩება სიდიდეების წრფივი გარდაქმნისას.

4. კორელაციის კოეფიციენტი ზომავს სიდიდეებს შორის წრფივი დამოკიდებულების სიძლიერეს. სხვა, არაწრფივი, თუნდაც დეტერმინისტული კავშირი კორელაციაში არ აისახება.

5. რაც უფრო კონცენტრირებულია მონაცემები რაიმე წრფის მიდამოში, მით უფრო დიდია კორელაციის კოეფიციენტი.

6. ორ სიდიდეს შორის კავშირი არსებობს, თუ ადგილი აქვს თანაფარდობას: $|r|\geq 2/\sqrt{n}$, სადაც n შერჩევის მოცულობაა.

7. თუ მოცემულია მონაცემთა ორი სიმრავლე $x_1,...,x_n$ და $y_1,...,y_n$, მაშინ მონაცემთა ახალი $z_i = x_i \pm y_i$ $(i=1,...,n)$ სიმრავლისათვის შერჩევითი დისპერსია ასე გამოითვლება: $s_z^2 = s_x^2 + s_y^2 \pm 2\operatorname{cov}(x,y)$ (თუ $\operatorname{cov}(x,y)=0$, მაშინ $s_z^2 = s_x^2 + s_y^2$). ანალოგიურად, $s_z^{'2} = s_x^{'2} + s_y^{'2} \pm 2\operatorname{cov}'(x,y)$, სადაც

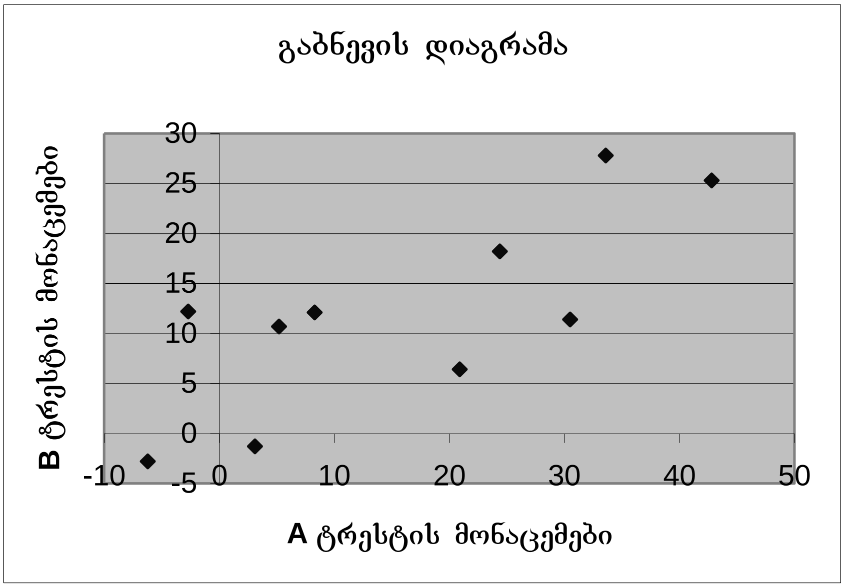
$$\operatorname{cov}'(x,y) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) = \frac{n}{n-1}\operatorname{cov}(x,y) .$$

მაგალითი 5. მოცემულია ორი ტრესტის წლიური საპროცენტო შემოსავლის მონაცემები ბოლო 10 წლის განმავლობაში:

ტრესტი A	8.3	-6.2	20.9	-2.7	33.6	42.8	24.4	5.2	3.1	30.5
ტრესტი B	12.1	-2.8	6.4	12.2	27.8	25.3	18.2	10.7	-1.3	11.4

ავაგოთ გაბნევის დიაგრამა და გამოვთვალოთ კორელაციის კოეფიციენტი.

ამოხსნა. A ტრესტის მონაცემები გადავზომოთ x ღერძზე, B ტრესტისა კი y ღერძზე.



ამ შემთხვევაში $\overline{x}=16$, $\overline{y}=12$, $s_x=17.65$ და $s_y=10.51$. ამიტომ გვაქვს:

$$\operatorname{cov}(x,y)=\frac{1}{n}\sum_{i=1}^{n}x_i y_i-\overline{x}\cdot\overline{y}=$$

$$=\frac{1}{10}\cdot[8.3\cdot 12.1+(-6.2)\cdot(-2.8)+\cdots+30.5\cdot 11.4]-16\cdot 12=71.48,$$

შესაბამისად,

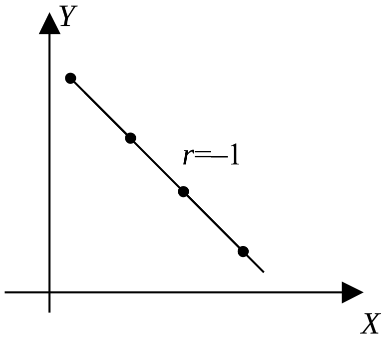
$$r=\frac{\operatorname{cov}(x,y)}{s_x s_y}=\frac{71.48}{17.65\cdot 10.51}\approx 0.39.$$

კორელაციის კოეფიციენტის ეს მნიშვნელობა გვარწმუნებს, რომ არსებობს სუსტად გამოხატული პოზიტიური წრფივი კავშირი A და B ტრესტების ამონაგებს შორის, რასაც გაბნევის დიაგრამაც ადასტურებს.

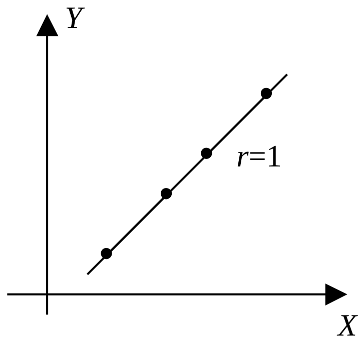
ამრიგად, *კორელაციის კოეფიციენტის* მნიშვნელობა მოთავსებულია -1-სა და 1-ს შორის. თუ ორ სიდიდეს შორის ადგილი აქვს *მკაცრად დადებით წრფივ დამოკიდებულებას,* მაშინ r -

ის მნიშვნელობა ახლოსაა +1-თან. თუ ორ სიდიდეს შორის ადგილი აქვს *მკაცრად უარყოფით წრფივ დამოკიდებულებას,* მაშინ r-ის მნიშვნელობა ახლოსაა -1-თან. როდესაც ორ სიდიდეს შორის არა გვაქვს წრფივი დამოკიდებულება ან ამ სიდიდეებს შორის სუსტი კავშირია, მაშინ r-ის მნიშვნელობა ახლოსაა 0-თან. ქვემოთ მოყვანილია კორელაციის ხარისხის სიტყვიერი დახასიათებისათვის მიღებული ტერმინოლოგია და შესაბამისი დიაგრამები:

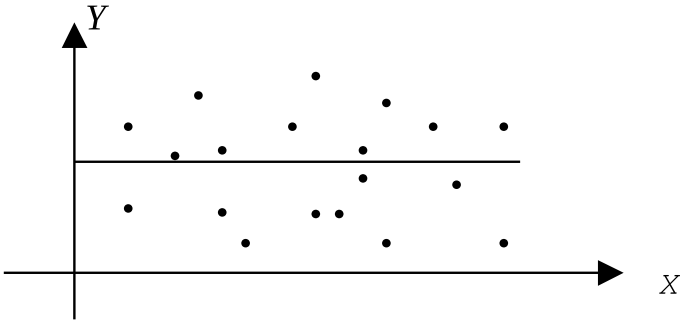
სრულყოფილი უარყოფითი კორელაცია		საშუალო უარყოფითი კორელაცია		არ არის კორელაცია		საშუალო დადებითი კორელაცია		სრულყოფილი დადებითი კორელაცია
	ძლიერი უარყოფითი კორელაცია		სუსტი უარყოფითი კორელაცია		სუსტი დადებითი კორელაცია		ძლიერი დადებითი კორელაცია	
-1		-0.5		0		0.5		1
უარყოფითი კორელაცია					დადებითი კორელაცია			



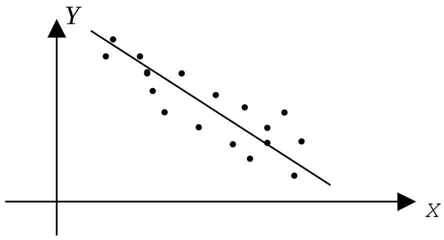
სრულყოფილი უარყოფითი კორელაცია



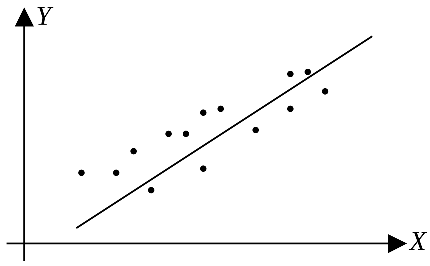
სრულყოფილი დადებითი კორელაცია



არ არის კორელაცია



ძლიერი უარყოფითი
კორელაცია



ძლიერი დადებითი
კორელაცია

კორელაციის კოეფიციენტის **პირდაპირი გამოთვლისათვის** იყენებენ ფორმულას:

$$r=\frac{n(\sum_{i=1}^{n}x_iy_i)-(\sum_{i=1}^{n}x_i)\cdot(\sum_{i=1}^{n}y_i)}{\sqrt{[n(\sum_{i=1}^{n}x_i^2)-(\sum_{i=1}^{n}x_i)^2]\cdot[n(\sum_{i=1}^{n}y_i^2)-(\sum_{i=1}^{n}y_i)^2]}}.$$

მაგალითი 7. გამოვთვალოთ კორელაციის კოეფიციენტის მნიშვნელობა 6 შემთხვევით შერჩეული პიროვნების ასაკისა და არტერიული წნევის მონაცემების მიხედვით:

პერსონა	**A**	**B**	**C**	**D**	**E**	**F**
ასაკი	43	48	56	61	67	70
წნევა	128	120	135	143	141	152

ამოხსნა.

ნაბიჯი 1. გავაკეთოთ ცხრილი ისე როგორც ეს ქვემოთაა ნაჩვენები:

პერსონა	ასაკი, x	წნევა, y	xy	x^2	y^2
A	43	128			
B	48	120			
C	56	135			
D	61	143			
E	67	141			
F	70	152			

ნაბიჯი 2. გამოვთვალოთ xy, x^2 და y^2 გამოსახულებათა მნიშვნელობები და შევიტანოთ ცხრილის შესაბამის სვეტებში. შევკრიბოთ სვეტებში მოთავებლი მნიშვნელობები. დასრულებულ ცხრილს ექნება შემდეგი სახე:

პერსონა	ასაკი, x	წნევა, y	xy	x^2	y^2
A	43	128	5.504	1.849	16.384
B	48	120	5.760	2.304	14.400
C	56	135	7.560	3.136	18.225
D	61	143	8.723	3.721	20.449
E	67	141	9.447	4.489	19.881
F	70	152	10.640	4.900	23.104
	$\sum x = 345$	$\sum y = 819$	$\sum xy = 47.634$	$\sum x^2 = 20.399$	$\sum y^2 = 112.443$

ნაბიჯი 3. მიღებული მნიშვნელობები შევიტანოთ კორელაციის კოეფიციენტის ფორმულაში და გამოვთვალოთ r:

$$r = \frac{6 \cdot 47.634 - 345 \cdot 819}{\sqrt{(6 \cdot 20.399 - 345^2) \cdot (6 \cdot 112.443 - 819^2)}} = 0.897 .$$

დასკვნა. კორელაციის კოეფიციენტის მიღებული მნიშვნელობა გვეუბნება, რომ ადამიანის ასაკსა და არტერიულ წნევას შორის არსებობს მკაცრად დადებითი წრფივი დამოკიდებულება.

დასკვნები კორელაციის კოეფიციენტის შესახებ

შერჩევითი კორელაციის კოეფიციენტი შეიძლება ავიღოთ პოპულაციის კორელაციის კოეფიციენტის შეფასებად, თუ x და y შემთხვევითი სიდიდეები წრფივადაა დაკავშირებული და მათი ერთობლივი განაწილება ნორმალურია.

ჰიპოთეზა: $H_0:\rho=0$ (სიდიდეებს შორის არ არის კორელაცია პოპულაციაში);

მნიშვნელოვნების დონე: α ;

კრიტერიუმის სტატისტიკა: $T=r\sqrt{\dfrac{n-2}{1-r^2}}\cong T(n-2)$.

კრიტერიუმის მნიშვნელობა T.V.: $t=r\sqrt{\dfrac{n-2}{1-r^2}}$.

ალტერნატივა კრიტიკული არე C.R. (H_0**-ის უარყოფის არე):**

$H_1:\rho>0$ $\qquad t\ge t_{n-2,\alpha}$,

$H_1:\rho<0$ $\qquad t\le -t_{n-2,\alpha}$,

$H_1:\rho\ne 0$ $\qquad t\le -t_{n-2,\alpha/2}$ ან $t\ge t_{n-2,\alpha/2}$

(სადაც $t_{n-2,\alpha}$ არის თავისუფლების $n-2$ ხარისხის მქონე სტიუდენტის განაწილების ზედა α-კრიტიკული წერტილი ანუ **კრიტიკული მნიშვნელობა C.V.**).

P **- მნიშვნელობა:**

$$P=\begin{cases}P\{T>t\mid H_0\},\ \text{თუ}\ H_1:\rho>0;\\ P\{T<t\mid H_0\},\ \text{თუ}\ H_1:\rho<0;\\ P\{|T|>|t|\mid H_0\},\ \text{თუ}\ H_1:\rho\ne 0.\end{cases}=\begin{cases}1-F_T(t),\ \text{თუ}\ H_1:\rho>0;\\ F_T(t),\ \text{თუ}\ H_1:\rho<0;\\ 2\cdot[1-F_T(|t|)],\ \text{თუ}\ H_1:\rho\ne 0.\end{cases}$$

გადაწყვეტილება: თუ $t\in$ **C.R.**, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

P **- მნიშვნელობის მეთოდი:** თუ $P\le\alpha$, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

მაგალითი 8. $\alpha = 0.05$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა წინა მაგალითში გამოთვლილი კორელაციის კოეფიციენტის მნიშვნელოვნების შესახებ.

ამოხსნა (სტანდარტული მეთოდი).

ნაბიჯი 1. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0 : \rho = 0$, $H_1 : \rho \neq 0$.

ნაბიჯი 2. ვიპოვოთ კრიტიკული მნიშვნელობა. აქ $n = 6$, ე. ი. თავისუფლების ხარისხია $d.f. = 6 - 2 = 4$ და რადგანაც $\alpha = 0.05$, ამიტომ სტიუდენტის განაწილების ცხრილიდან ვპოულობთ, რომ $C.V. = \pm t_{n-2,\alpha/2} = \pm t_{4,0.025} = \pm 2.776$.

ნაბიჯი 3. გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$t = r\sqrt{\frac{n-2}{1-r^2}} = 0.897 \cdot \sqrt{\frac{6-2}{1-0.897^2}} = 4.059.$$

ნაბიჯი 4. გადაწყვეტილების მიღება: ვინაიდან კრიტერიუმის მნიშვნელობა ვარდება კრიტიკულ არეში, ამიტომ ნულოვან ჰიპოთეზას უკუვაგდებთ, ანუ ასაკსა და წნევას შორის არის მნიშვნელოვანი წრფივი კავშირი.

ამოხსნა (P-მნიშვნელობის მეთოდი).

ნაბიჯი 1. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები: $H_0 : \rho = 0$, $H_1 : \rho \neq 0$.

ნაბიჯი 2. ვიპოვოთ თავისუფლების ხარისხი. ვინაიდან $n = 6$, ამიტომ თავისუფლების ხარისხია $d.f. = 6 - 2 = 4$.

ნაბიჯი 3. გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$t = r\sqrt{\frac{n-2}{1-r^2}} = 0.897 \cdot \sqrt{\frac{6-2}{1-0.897^2}} = 4.059.$$

ნაბიჯი 4. სტიუდენტის განაწილების ცხრილში თავისუფლების 4-ის ტოლი ხარისხისათვის მოვძებნოთ ორი მნიშვნელობა, რომელთა შორის ვარდება კრიტერიუმის მნიშვნელობა 4.059. ასეთი მნიშვნელობებია 3.747 და 4.604. მათი შესაბამისის α-ების მოძებნით ვპოულობთ, რომ, შესაბამისად, 0.01 < P-მნიშვნელობა < 0.02.

ნაბიჯი 5. გადაწყვეტილების მიღება: ვინაიდან P-მნიშვნელობა < α, ამიტომ ნულოვან ჰიპოთეზას უკუვაგდებთ.

მაგალითი 9. დაწერეთ რეგრესიის წრფის განტოლება მე-5 მაგალითში და გააკეთეთ 50 წლის ასაკის ადამიანის არტერიული წნევის პროგნოზი.

ამოხსნა. გვაქვს: $n=6$, $\sum_{i=1}^{6} x_i = 345$, $\sum_{i=1}^{6} y_i = 819$, $\sum_{i=1}^{6} x_i y_i = 47.634$, $\sum_{i=1}^{6} x_i^2 = 20.399$. გამოვთვალოთ რეგრესიის წრფის კოეფიციენტები:

$$a = \frac{n(\sum_{i=1}^{n} x_i y_i) - (\sum_{i=1}^{n} x_i)\cdot(\sum_{i=1}^{n} y_i)}{n(\sum_{i=1}^{n} x_i^2) - (\sum_{i=1}^{n} x_i)^2} = 0.964,$$

$$b = \frac{(\sum_{i=1}^{n} y_i)\cdot(\sum_{i=1}^{n} x_i^2) - (\sum_{i=1}^{n} x_i)\cdot(\sum_{i=1}^{n} x_i y_i)}{n(\sum_{i=1}^{n} x_i^2) - (\sum_{i=1}^{n} x_i)^2} = 81.048.$$

ამიტომ რეგრესიის წრფის განტოლება იქნება:

$$y = 0.964x + 81.048 .$$

გავაკეთოთ პროგნოზი:

$$y = 0.964 \cdot 50 + 81.048 = 129.248 \approx 129 .$$

მაშასადამე, 50 წლის ადამიანის სავარაუდო არტერიული წნევა იქნება 129.

შენიშვნა: *იმ შემთხვევაში, როცა კორელაციის კოეფიციენტი მნიშვნელოვნად არ განსხვავდება ნულისაგან, y -ის საუკეთესო პროგნოზია y -ის მნიშვნელობების საშუალო.*

ამოცანები

1. გამოთვალეთ კორელაციის კოეფიციენტი და გააკეთეთ შესაბამისი დასკვნები x და y სიდიდეებს შორის კავშირის შესახებ:

x	99	150	130	105	219	167	180
y	1200	1000	1050	1150	700	850	900

3. სადაზღვევო კომპანია სწავლობს კავშირს სიცოცხლის დაზღვევაზე გამოყოფილ სახსრებსა (y) და ოჯახის შემოსავლებს (x) შორის. ამ მიზნით შეისწავლეს 6 კლიენტის შესაბამისი მაჩვენებლები:

x	45	20	40	40	30	55
y	70	50	60	50	55	105

გამოთვალეთ კორელაციის კოეფიციენტი.

ქვემოთ მოყვანილ ამოცანებში $\alpha = 0.05$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა კორელაციის კოეფიციენტის მნიშვნელოვნების შესახებ.

5. მკვლევარს აინტერესებს არის თუ არა კავშირი ადამიანის ასაკსა და მის მიერ კვირის განმავლობაში გამაჯანსაღებელ ვარჯიშზე დახარჯულ საათებს შორის:

ასაკი, x	18	26	32	38	52	59
საათები, y	10	5	2	3	1.5	1

7. მაღაზიის მენეჯერს აინტერესებს არის თუ არა კავშირი გამყიდველის ასაკსა და მის მიერ წლის განმავლობაში ავადმყოფობის გამო გაცდენილ დღეების რაოდენობას შორის:

ასაკი, x	18	26	39	48	53	58
დღეები, y	16	12	9	5	6	2

9. სადაზღვევო კომპანიას აინტერესებს რამდენად ძლიერი კავშირია ადამიანის კვირის განმავლობაში მუშაობის საათების რაოდენობასა და მის მიერ ამავე პერიოდში მიღებული ტრამვების რაოდენობას შორის:

საათები, x	40	32	36	44	41
ტრამვები, y	1	0	3	8	5

ქვემოთ მოყვანილ ამოცანებში $\alpha = 0.05$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა კორელაციის კოეფიციენტის მნიშვნელოვნების შესახებ და დაწერეთ რეგრესიის წრფის განტოლება.

11. მკვლევარს აინტერესებს რა დამოკიდებულებაა ადამიანის ასაკსა და კვირის განმავლობაში ფიზიკურ ვარჯიშემზე დახარჯული საათების რაოდენობას შორის:

ასაკი, x	34	22	48	56	62
საათები, y	5.5	7	3.5	3	1

13. პედაგოდს აინტერესებს როგორაა დამოკიდებული სტუდენტის მიერ ლექციების გაცდენათა რიცხვი მის მიერ მიღებულ საბოლოო შეფასებაზე:

გაცდენა, x	10	12	2	0	8	5
შეფასება, y	70	65	96	94	75	82

ქვემოთ მოყვანილ **ამოცანებში ისარგებლეთ** P **-მნიშვნელობის მეთოდით**

15. ავტომობილების დილერს აინტერესებს არის თუ არა კავშირი გამყიდველის მუშაობის სტაჟსა და მის მიერ თვეში გაყიდული ავტომობილების რაოდენობას შორის:

სტაჟი, x	3	9	2	5	1
რაოდენობა, y	5	14	12	21	8

ამოცანები გამოცდისათვის

ქვემოთ მოყვანილ 4 ამოცანაში ააგეთ გაბნევის დიაგრამა, იპოვეთ კორელაციის კოეფიციენტი, შეამოწმეთ კორელაციის კოეფიციენტის *მნიშვნელოვნება* $\alpha = 0.05$-სათვის, დაწერეთ რეგრესიის წრფის განტოლება, გააკეთეთ პროგნოზი.

17. მკვლევარს აინტერესებს კავშირი მძღოლის ასაკსა და წელიწადში მომხდარ ავტო-საგზაო შემთხვევათა რაოდენობას შორის. გააკეთეთ 64 წლის მძღოლის ავტოსაგზაო შემთხვევათა რაოდენობის პროგნოზი, თუ:

ასაკი, x	63	65	60	62	66	67	59
რაოდენობა, y	2	3	1	0	3	1	4

19. იკვლევენ ყოველდღიურად მიღებული ცხიმის რაოდენობის დამოკიდებულებას პაციენტის ქოლესტერინის დონეზე. გააკეთეთ ქოლესტერინის დონის პროგნოზი 8.5 გრამი ცხიმის მიღებისას, თუ:

ცხიმის რაოდენობა, x	6.8	5.5	8.2	10	8.6	9.1	8.6	10.4
ქოლესტერ. დონე, y	183	201	193	283	222	250	190	218

21. მე-19 ამოცანაში იპოვეთ შეფასების სტანდარტული შეცდომა.
23. მე-19 ამოცანაში ააგეთ 95%-იანი საპროგნოზო ინტერვალი იმ პაციენტის ქოლესტერინის დონისათვის, ვინც მოიხმარს 10 გრამ ცხიმს.

თავი XVIII

რეგრესია და კორელაცია

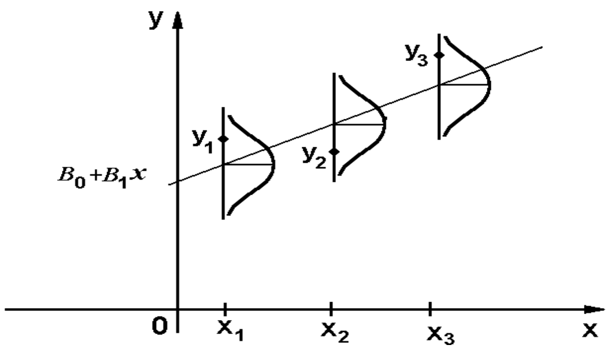
1. მარტივი წრფივი რეგრესია.

ზოგადად ამოცანა ასე აღიწერება: მოცემულია (x_i, Y_i), $i = 1, 2, ..., n$ წყვილები, რომელთა შორის შემდეგი დამოკიდებულებაა

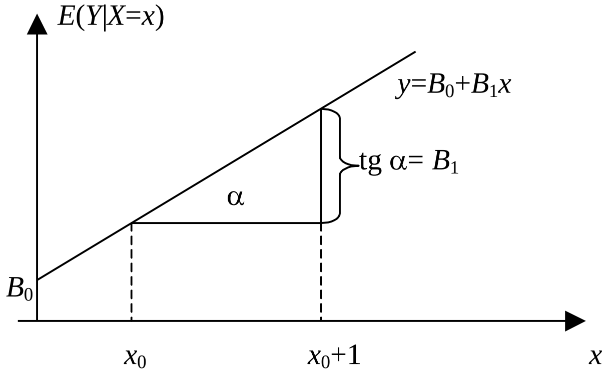
$$Y_i = B_0 + B_1 x_i + \varepsilon_i \, , \; i = 1, 2, ..., n \, ,$$

სადაც B_0 და B_1 უცნობი მუდმივებია, $\varepsilon_i \cong N(0, \sigma^2)$ კი ნორმალურად განაწილებული შემთხვევითი სიდიდეებია საშუალოთი ნული და უცნობი σ^2 დისპერსიით.

$y = B_0 + B_1 x$ წრფეს **რეგრესიის წრფე** ეწოდება. B_0 და B_1 კოეფიციენტებს – **რეგრესიის კოეფიციენტები,** ხოლო ε_i შემთხვევით სიდიდეს – **შემთხვევითი გადახრა** ეწოდება. ε_i - შემთხვევითი სიდიდეა, რომელიც განაპირობებს $y = B_0 + B_1 x$ წრფიდან გაბნევას. x ცვლადს უწოდებენ ***ამხსნელ*** **(დამოუკიდებელ) ცვლადს** ან **პრედიქტორს,** ხოლო Y -ს კი – ***მოპასუხე*** **(დამოკიდებულ) ცვლადს**.



B_1 კოეფიციენტს შეიძლება მივცეთ შემდეგი ინტერპრეტაცია: ამხსნელი x ცვლადის ერთი ერთეულით ცვლილება იწვევს დამოკიდებული Y ცვლადის ცვლილებას საშუალოდ B_1 სიდიდით. B_0 კოეფიციენტი წარმოადგენს რეგრესიის წრფის y ღერძთან გადაკვეთის წერტილს.



საჭიროა $(x_1,Y_1),(x_2,Y_2),...,(x_n,Y_n)$ დამოუკიდებელი დაკვირვებების საფუძველზე აიგოს B_0 და B_1 კოეფიციენტებისა და σ^2 დისპერსიის შეფასებები.

2. უმცირეს კვადრატთა მეთოდი.

B_0 და B_1 კოეფიციენტების b_0 და b_1 შეფასებებს ეძებენ ე. წ. უმცირეს კვადრატთა მეთოდით: $\sum_{i=1}^{n}(Y_i-b_0-b_1x_i)^2\to\min$.

$$b_1=\frac{n\sum_{i=1}^{n}x_iY_i-\sum_{i=1}^{n}x_i\sum_{i=1}^{n}Y_i}{n\sum_{i=1}^{n}x^2-\left(\sum_{i=1}^{n}x_i\right)^2}:=\frac{SS_{XY}}{SS_X},\ b_0=\frac{1}{n}\sum_{i=1}^{n}Y_i-b_1\frac{1}{n}\sum_{i=1}^{n}x_i=\overline{Y}-b_1\overline{x}\ ,$$

$$SS_x=\sum_{i=1}^{n}(x_i-\overline{x})^2=\sum_{i=1}^{n}x_i^2-n\overline{x}^2\ ,$$

$$SS_{XY}=\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})=\sum_{i=1}^{n}x_iy_i-n\overline{x}\cdot\overline{y}\ .$$

$\hat{Y}=b_0+b_1x$ წრფეს ეწოდება **რეგრესიის წრფის შეფასება** ან უმცირეს კვადრატთა მეთოდით აგებული წრფე.

შერჩევითი კორელაციის რიცხვითი მნიშვნელობა: $r=\dfrac{SS_{XY}}{\sqrt{SS_X\cdot SS_Y}}$.

შეფასებული რეგრესიის წრფის განტოლება შემდეგნაირად გადაიწერება: $y=\bar{Y}+b_1(x-\bar{x})$. ეს წრფე გადის $(\bar{x},\bar{Y})$ წერტილზე და $\sum_{i=1}^{n}(Y_i-\hat{Y}_i)=0$.

მოდელი რომელიც $Y=B_0+B_1x+\varepsilon$ ფორმულითაა მოცემული გულისხმობს, რომ გაბნევის დიაგრამაზე წერტილების გაბნევა განპირობებულია Y ცვლადის მნიშვნელობათა გაბნევით რეგრესიის მრუდიდან, ხოლო x ცვლადის მნიშვნელობები განიხილება ზუსტ მნიშვნელობებად. თუ რეგრესიის წრფის განტოლებაში x-ის ნაცვლად ჩავსვამთ დაკვირვებულ $x_1,...,x_n$ მნიშვნელობებს, მაშინ შესაბამის $\hat{Y}_1=b_0+b_1x_1,...,\hat{Y}_n=b_0+b_1x_n$ მნიშვნელობებს ეწოდება Y ცვლადის პროგნოზირებული მნიშვნელობები, ხოლო $e_1=Y_1-\hat{Y}_1,...,e_n==Y_n-\hat{Y}_n$ სიდიდეებს ნაშთები ეწოდება. სიდიდეს

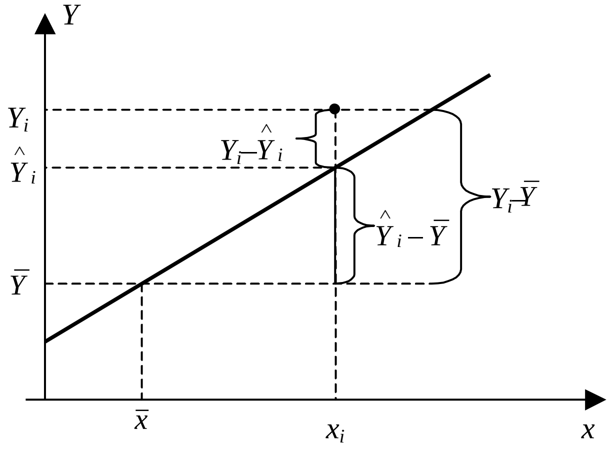
$$SSE:=\sum_{i=1}^{n}e_i^2=\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2=\sum_{i=1}^{n}\left[Y_i-(b_0+b_1x_i)\right]^2=SS_Y-\frac{SS_{XY}^2}{SS_X}$$

ეწოდება ***კვადრატების ნარჩენი ჯამი*** (ნაშთთა კვადრატების ჯამი, *sum of square for eror*).

$$SST:=\sum_{i=1}^{n}(Y_i-\bar{Y})^2=\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2+\sum_{i=1}^{n}(\hat{Y}_i-\bar{Y})^2:=SSE+SSR.$$

SST (*total sum of squares)* სიდიდეს ***სრულ ვარიაციას*** უწოდებენ, სიდიდე $SSR:=\sum_{i=1}^{n}(\hat{Y}_i-\bar{Y})^2$ არის გაბნევის ის ნაწილი, რომელიც ახსნილია რეგრესიის წრფით (*regression sum of squares).* *SSE* არის სრული გაბნევის ის ნაწილი, რომელიც არ აიხსნება

რეგრესიის საშუალებით. $\frac{SST}{\sigma^2} \sim \chi^2(n-1)$; $\frac{SSR}{\sigma^2} \sim \chi^2(1)$; $\frac{SSE}{\sigma^2} \sim \chi^2(n-2)$. $\underset{n-1}{SST} = \underset{n-2}{SSE} + \underset{1}{SSR}$.



დეტერმინაციის კოეფიციენტი. რეგრესიის ხარისხის დასახასიათებლად შემოჰყავთ ე. წ. შერჩევითი დეტერმინაციის კოეფიციენტი:

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2} = \frac{SSR}{SST}, \left(\frac{SSR}{SST} = 1 - \frac{SSE}{SST} \right)$$

ე. ი. R^2 წარმოადგენს რეგრესიით ახსნილ გაბნევის წილს სრულ გაბნევაში. მტკიცდება, რომ R^2 – ახალს არაფერს იძლევა, ვინაიდან $R^2 = r^2$, სადაც r შერჩევითი კორელაციის კოეფიციენტია.

თუ r^2 მცირეა, ე.ი. რეგრესიით ახსნილია სრული გაბნევის მცირე წილი, მაშინ მოსაძებნია სხვა ალტერნატიული მოდელი (მაგალითად, არაწრფივი ან მრავლობითი რეგრესიის მოდელი), რომელიც უფრო ეფექტურად ახსნიდა Y ცვლადის დაკვირვებული მნიშვნელობების სრულ გაბნევას თავისი საშუალო მნიშვნელობების მიმართ.

$$\sum_{i=1}^{n}(\hat{Y}_i-\bar{Y})^2 = R^2\cdot\sum_{i=1}^{n}(Y_i-\bar{Y})^2 = r^2\cdot\sum_{i=1}^{n}(Y_i-\bar{Y})^2 \ ,$$

$$\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2 = (1-r^2)\cdot\sum_{i=1}^{n}(Y_i-\bar{Y})^2 \ .$$

3. სტატისტიკური დასკვნები რეგრესიის კოეფიციენტების შესახებ.

მტკიცდება, რომ b_0 და b_1 შესაბამისად B_0 და B_1 პარამეტრების ჩაუნაცვლებელი შეფასებებია ($Eb_0 = B_0, \quad Eb_1 = B_1$) და

$$\sigma_{b_0}^2 := Db_0 = \sigma^2\cdot\frac{\sum_{i=1}^{n}x_i^2}{n\sum_{i=1}^{n}x_i^2-\left(\sum_{i=1}^{n}x_i\right)^2}, \qquad \sigma_{b_1}^2 := Db_1 = \sigma^2\cdot\frac{n}{n\sum_{i=1}^{n}x_i^2-\left(\sum_{i=1}^{n}x_i\right)^2}$$

უცნობი σ^2-ის ჩაუნაცვლებელი შეფასება მოიცემა თანაფარდობით

$$S^2 = (\sum_{i=1}^{n}e_i^2)/(n-2) = (\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2)/(n-2) = SSE/(n-2)\ .$$

სიდიდეს $S=\sqrt{SSE/(n-2)}$ **სტანდარტული შეცდომა** ეწოდება.

b_0 და b_1 სტატისტიკების დისპერსიების შეფასებებია შესაბამისად

$$S_{b_0}^2 = S^2\cdot\frac{\sum x_i^2}{n\cdot SS_X}\ ,\ S_{b_1}^2 = S^2\cdot\frac{1}{SS_X}\ .$$

$T_{b_j} = \dfrac{b_j-B_j}{S_{b_j}} \cong t(n-2), \quad j=0,1$. აქედან გამომდინარე იგება ნდობის ინტერვალი B_j-სათვის. α მნიშვნელოვნების დონის მქონე ნდობის ინტერვალი B_j-სათვის შემდეგი სახისაა: $(b_j - s_{b_j}\cdot t_{n-2,\alpha/2}, b_j + s_{b_j}\cdot t_{n-2,\alpha/2})$.

განვიხილოთ ძირითადი $H_0: B_J = B_j^0$ ჰიპოთეზა $H_1: B_J \neq B_j^0$ ალტერნატივის წინააღმდეგ. ამ შემთხვევაში კრიტერიუმის სტატისტიკა იქნება

$$T_{b_j} = (b_j - B_j^0)/S_{b_j} \cong t(n-2)\,.$$

თუ T_{b_j} სტატისტიკის დაკვირვებული მნიშვნელობისათვის t_{b_j} და α მნიშვნელოვნების დონისათვის სრულდება პირობა: $\left| t_{b_j} \right| > t_{n-k-1,\alpha/2}$, მაშინ უარვყოფთ $H_0 : B_J = B_j^0$ ჰიპოთეზას (შესაბამისად, ვღებულობთ $H_1 : B_J \neq B_j^0$ ჰიპოთეზას), წინააღმდეგ შემთხვევაში H_0-ის უარყოფის საფუძველი არ გაგვაჩნია.

შევნიშნავთ, რომ თუ $B_1 = 0$, მაშინ რეგრესიის წრფე აბსცისთა ღერძის პარალელურია, ე. ი. X-ის ცვლილებით Y არ იცვლება. თუ ჩვენ გვინდა, α მნიშვნელოვნების დონით, დავრწმუნდეთ, რომ არ არსებობს რაიმე (დადებითი ან უარყოფითი) წრფივი კავშირი X და Y სიდიდეებს შორის, ნულოვანი ჰიპოთეზა $H_0 : B_1 = 0$ უნდა შემოწმდეს ორმხრივი $H_1 : B_1 \neq 0$ ალტერნატივის წინააღმდეგ. თუ გვსურს დავრწმუნდეთ დადებითი (შესაბამისად, უარყოფითი) წრფივი კავშირის არსებობაში, მაშინ ვიხილავთ მარჯვენა ცალმხრივ $H_1 : B_1 > 0$ (შესაბამისად, მარცხენა ცალმხრივ $H_1 : B_1 < 0$) ალტერნატივას. ყველა შემთხვევაში ჰიპოთეზის შემოწმება ხდება $T_1 := T_{b_1} = b_1 / S_{b_1} \cong t(n-2)$ სტატისტიკის გამოყენებით. *გადაწყვეტილების მიღების წესი ასე ყალიბდება:*

ა) ორმხრივი ალტერნატივის $H_1 : B_1 \neq 0$ შემთხვევაში $H_0 : B_1 = 0$ ჰიპოთეზას უარვყოფთ, თუ $|t_1| > t_{n-2,\alpha/2}$;

ბ) მარჯვენა ცალმხრივი ალტერნატივის $H_1 : B_1 > 0$ შემთხვევაში $H_0 : B_1 = 0$ ჰიპოთეზას უარვყოფთ, თუ $t_1 > t_{n-2,\alpha}$;

გ) მარცხენა ცალმხრივი ალტერნატივის $H_1 : B_1 < 0$ შემთხვევაში $H_0 : B_1 = 0$ ჰიპოთეზას უარვყოფთ, თუ $t_1 < -t_{n-2,\alpha}$, სადაც t_1 არის T_1 სტატისტიკის დაკვირვებული მნიშვნელობა, ხოლო $t_{n-2,\alpha}$ – სტიუდენტის $t(n-2)$ განაწილების ზედა α-კრიტიკული წერტილი.

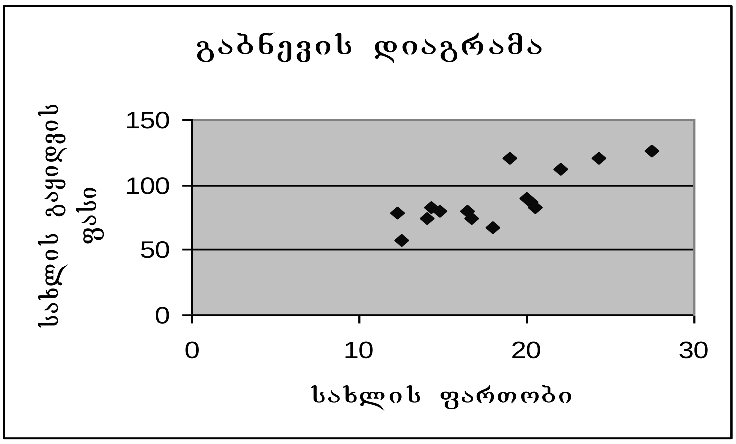
მაგალითი 1. უძრავი ქონების კომპანიის აგენტს სურს განჭვრიტოს სახლის გასაყიდი ფასი. მისი აზრით სახლის ფასი ყველაზე მჭიდრო კავშირშია მის ფართობთან. მან შემთხვევით შე-

არჩია ადრე გაყიდული 15 სახლის მონაცემები, რომლებიც მოყვანილია ცხრილის სახით (სადაც x – სახლის ფართობია (× 10 კვ. მ.), ხოლო Y – სახლის გაყიდვის ფასი (×1000$)). აგენტის მიზანია სახლის ფასის (დამოკიდებელი ცვლადი) პროგნოზირება მისი ფართობის (პრედიქტორის) მიხედვით.

x_i	Y_i	x_i	Y_i	x_i	Y_i
20.0	89.5	14.3	82.5	22.0	112.6
14.8	79.9	27.5	126.3	19.0	120.8
20.5	83.1	16.5	79.3	12.3	78.5
12.5	56.9	24.3	119.9	14.0	74.3
18.0	66.6	20.2	87.6	16.7	74.8

ამოხსნა.

ნაბიჯი 1. შევადგინოთ გაბნევის დიაგრამა



აქედან ჩანს, რომ Y ცვლადის X -ისაგან დამოკიდებულებაში იკვეთება წრფივობის ტენდენცია და შესაძლებელია, რომ მარტივმა წრფივმა რეგრესიამ კარგად აღწეროს იგი. ამიტომ მათემატიკურ მოდელად შეგვიძლია ავირჩიოთ

$$Y_i = B_0 + B_1 x_i + \varepsilon_i .$$

ნაბიჯი 2. გამოვთვალოთ სიდიდეები:

$$\sum_{i=1}^{n} x_i\ ,\quad \sum_{i=1}^{n} Y_i\ ,\quad \sum_{i=1}^{n} x_i^2\ ,\quad \sum_{i=1}^{n} Y_i^2\ ,\quad \sum_{i=1}^{n} x_i Y_i\ .$$

$$\sum_{i=1}^{n} x_i = 172.6\,,\quad \sum_{i=1}^{n} Y_i = 1332.6\,,\quad \sum_{i=1}^{n} x_i^2 = 5222.24\,,$$

$$\sum_{i=1}^{n} Y_i^2 = 124618.42\,,\quad \sum_{i=1}^{n} x_i Y_i = 25257.97\,.$$

ნაბიჯი 3. გამოვთვალოთ b_0-ისა და b_1-ის გამოსათვლელად საჭირო სიდიდეები:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{272.6}{15} = 18.17\,,\quad \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1332.6}{15} = 88.84\,.$$

$$SS_x = \sum_{i=1}^{n} x_i^2 - n\overline{x}^2 = 268.19\,,\quad SS_Y = \sum_{i=1}^{n} Y_i^2 - n\overline{Y}^2 = 6230.24\,,$$

$$SS_{xY} = \sum_{i=1}^{n} x_i Y_i - n\overline{x}\,\overline{Y} = 1040.18\,.$$

ნაბიჯი 4. გამოვთვალოთ b_0 და b_1 შეფასებები:

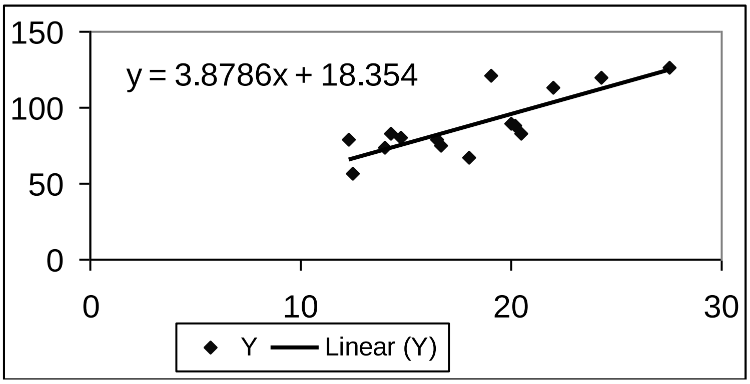
$$b_1 = \frac{SS_{xY}}{SS_x} = \frac{1040.18}{268.19} = 3.8786\,,$$

$$b_0 = \overline{Y} - b_1\overline{x} = 88.84 - 3.8786 \cdot 18.17 = 18.354\,.$$

ნაბიჯი 5. დავწეროთ რეგრესიის წრფის შეფასება:

$$\hat{Y} = 18.354 + 3.8786x\,.$$

მიღებული მონაცემები გვაძლევს შემდეგ გრაფიკულ სურათს:



დასკვნა. b_1-ის მიღებული მნიშვნელობა $b_1 = 3.8786$ ნიშნავს, რომ სახლის ფართობის ყოველი ერთი ერთეულით გაზრდა იწვევს სახლის ფასის საშუალოდ $3.8786 \cdot 1000\$ = 3878.6\$$-ით გაზრდას.

ნაბიჯი 6. გამოვთვალოთ ESS . გვაქვს:

$$SSE = SS_Y - \frac{SS_{xY}^2}{SS_x} = 6230.24 - \frac{(1040.18)^2}{268.19} = 21995.88,$$

საიდანაც მიიღება უცნობი σ^2 დისპერსიის შეფასება

$$S^2 = \frac{SSE}{n-2} = \frac{21995.88}{13} = 168.91.$$

ნაბიჯი 7. გამოვთვალოთ $S_{b_0}^2$ და $S_{b_1}^2$:

$$S_{b_0}^2 = S^2 \cdot \frac{\sum x_i^2}{n \cdot SS_X} = 168.91 \cdot \frac{5222.24}{15 \cdot 268.19} = 219.26,$$

$$S_{b_1}^2 = S^2 \cdot \frac{1}{SS_X} = 168.91 \cdot \frac{1}{268.19} = 0.6298,$$

შესაბამისად, $S_{b_1} = 0.7936$.

ნაბიჯი 8. გამოვთვალოთ დეტერმინაციის კოეფიციენტი:

$$R^2 = 1 - \frac{SSE}{SST} = 0.648.$$

დეტერმინაციის კოეფიციენტის გამოთვლილი მნიშვნელობა გვიჩვენებს, რომ მარტივი წრფივი რეგრესიით აიხსნება სახლის გასაყიდი ფასის ვარიაციის მხოლოდ 64.8%, დანარჩენი 35.2%

აუხსნელი რჩება. საჭიროა შეირჩეს სხვა მოდელი (არაწრფივი რეგრესია ან მრავლობითი რეგრესია).

მაგალითი 2. მაგალით 1-ში:

ა) ავაგოთ $\alpha = 0.05$ მნიშვნელოვნების დონის მქონე ნდობის ინტერვალი B_1-სათვის;

ბ) შევამოწმოთ $H_0 : B_1 = 0$ ჰიპოთეზა $H_1 : B_1 \neq 0$ ალტერნატივის წინააღმდეგ.

ამოხსნა. ა) ცნობილია, რომ $T_1 := \dfrac{b_1 - B_1}{S_{b_1}} \cong t(n-2)$. სტიუდენტის ზედა კრიტიკული წერტილების ცხრილში ვპოულობთ $t_{n-2,\alpha/2} = t_{13,0.025} = 2.16$. ამიტომ საძიებელი ნდობის ინტერვალი იქნება:

$$3.88 - 2.16 \cdot 0.7936 < B_1 < 3.88 + 2.16 \cdot 0.7936 ,$$

ანუ $2.1658 < B_1 < 5.5942$.

ბ) T_1 სტატისტიკის გამოთვლილი მნიშვნელობა t_1 იქნება:

$$t_1 = \frac{b_1 - B_1}{S_{b_1}} = \frac{3.88 - 0}{0.7936} = 4.8891 > t_{13,0.025} = 2.16 .$$

ამიტომ $H_0 : B_1 = 0$ ჰიპოთეზას უარვყოფთ (ვინაიდან $|t_1| > t_{n-2,\alpha/2}$).

ამრიგად, $\alpha = 0.05$ მნიშვნელოვნების დონით სარწმუნოა მტკიცება, რომ სახლის ფასსა და მის ფართობს შორის არსებობს წრფივი კავშირი.

4. მრავლობითი რეგრესია.

განვიხილოთ წრფივი მრავლობითი რეგრესიის მოდელი

$$Y = B_0 + B_1 X_1 + \cdots + B_k X_k + \varepsilon ,$$

სადაც Y – დამოკიდებული ცვლადია, $X_1, \ldots, X_k$ – დამოუკიდებელი ცვლადებია (მათ **ამხსნელ ცვლადებს, პრედიქტორებს ან რეგრესორებს** უწოდებენ), ε – შეშფოთებაა, რომელსაც ხშირად **ჭეშმარიტ გადახრას ან შეცდომას** უწოდებენ. ε შემთხვევითი სიდიდეა, რომლისთვისაც $E\varepsilon = 0$, ხოლო $D\varepsilon = \sigma^2 > 0$. იგულისხმება, რომ დამოუკიდებელი $X_1, \ldots, X_k$ ცვლადების მნიშვნე-

ლობები ზუსტადაა ცნობილი და გადახრები მიეწერება მხოლოდ Y ცვლადს.

ცხადია, რომ Y შემთხვევითი სიდიდის მათემატიკური ლოდინი პირობაში, რომ $X_1 = x_1, ..., X_k = x_k$ მოიცემა ფორმულით:

$$E(Y \mid X_1 = x_1, ..., X_k = x_k) = B_0 + B_1 x_1 + \cdots + B_k x_k ,$$

ხოლო დისპერსია იმავე პირობებში იქნება:

$$D(Y \mid X_1 = x_1, ..., X_k = x_k) = \sigma^2 .$$

k ცვლადის $y = B_0 + B_1 x_1 + \cdots + B_k x_k$ ფუნქციას **ჭეშმარიტი რეგრესიის ფუნქცია** ეწოდება, ხოლო $B_0, B_1, ..., B_k$ სიდიდეებს კი – **რეგრესიის კოეფიციენტები.**

საბოლოოდ, მრავლობითი რეგრესიის ამოცანა შემდეგნაირად აღიწერება: მოცემულია n მოცულობის შერჩევა $(Y_i, x_{1i}, ..., x_{ki})$, $i = 1, 2, ..., n, \quad n > k$, სადაც

$$Y_i = B_0 + B_1 x_{1i} + \cdots + B_k x_{ki} + \varepsilon_i, \quad i = 1, 2, ..., n , \qquad (3)$$

ხოლო ε_i შემთხვევითი სიდიდეა, რომელიც განაპირობებს Y_i ცვლადის გადახრას ჭეშმარიტი რეგრესიის ფუნქციის $B_0 + B_1 x_{1i} + \cdots + B_k x_{ki}$ მნიშვნელობიდან: $\varepsilon_i \cong N(0, \sigma^2)$, $E(\varepsilon_i \cdot \varepsilon_j) = 0$, $i \neq j$.

ისევე როგორც მარტივი რეგრესიის შემთხვევაში, ასევე პირველ ეტაპზე ხდება რეგრესიის თეორიული $B_0, B_1, ..., B_k$ კოეფიციენტებისა და უცნობი σ^2 დისპერსიის შეფასება. რეგრესიის კოეფიციენტების შეფასება ისევ წარმოებს უმცირეს კვადრატთა მეთოდის გამოყენებით. მას შემდეგ რაც მიღებულია თეორიული რეგრესიის კოეფიციენტების $b_0, b_1, ..., b_k$ შეფასებები, $\hat{Y} = b_0 + b_1 x_1 + \cdots + b_k x_k$ ფუნქცია განსაზღვრავს რეგრესიის ფუნქციის შეფასებას. სიდიდეს

$$e_i = Y_i - \hat{Y}_i = Y_i - b_0 - b_1 x_{1i} - \cdots - b_k x_{ki}$$

ნაშთი ეწოდება.

ისევე როგორც მარტივი რეგრესიის მოდელში, უცნობი $\sigma^2 = D\varepsilon_i$ დისპერსიის შეფასება ეყრდნობა ნაშთების (შესწორე-ბების) კვადრატების ჯამს (*sum of square errors (residuals) – SSE):*

$$SSE := \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)_i^2 .$$

ვინაიდან ამ გამოსახულებაში შედის $(k+1)$ შეფასებული პა-რამეტრი, მისი თავისუფლების ხარისხია $n-(k+1)$ და ამიტომ დისპერსიის შეფასებად იღებენ სიდიდეს:

$$S^2 = \frac{SSE}{n-(k+1)} = MSE .$$

რადგანაც, დაშვების თანახმად $\varepsilon_i \cong N(0,\sigma^2)$, ამიტომ Y_i ნორმა-ლურებია და როგორც მათი წრფივი კომბინაცია, ასევე ნორმალუ-რი იქნება b_j შეფასებებიც. მტკიცდება, რომ: $T_j := (b_j - B_j)/S_{b_j} \cong$ $\cong t(n-(k+1))$.

დეტერმინაციის კოეფიციენტი. დავწეროთ დისპერსიული თანაფარდობა

$$\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 + \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2 .$$

შემოვიღოთ აღნიშვნები:

$$SST = \sum_{i=1}^{n}(Y_i - \bar{Y})^2 , \quad SSE = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \quad \text{და} \quad SSR = \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2 .$$

მაშინ გვაქვს: $\underset{n-1}{SST} = \underset{n-k-1}{SSE} + \underset{k}{SSR}$, სადაც SSR – არის სრული გაბნევის რეგრესიით ახსნილი ნაწილი, SSE – სრული გაბნევის ის ნაწილი, რომელიც არ აიხსნება რეგრესიით. თითოეულ წევრს ქვევით მიწერილი აქვს თავისუფლების ხარისხი.

ე. წ. **მრავლობითი დეტერმინაციის კოეფიციენტი** წარმოად-გენს სრული ვარიაციის იმ ნაწილს, რომელიც ახსნილია მრავ-ლობითი რეგრესიის მოდელით:

$$R^2 = 1 - \frac{SSE}{SST} \left(= \frac{SSR}{SST}\right).$$

გარდა ამისა, შემოაქვთ ე. წ. **თავისუფლების ხარისხთან შე-თანხმებული (adjusted) დეტერმინაციის კოეფიციენტი:**

$$adjusted\ R^2 = 1 - \frac{SSE/(n-k-1)}{SST/(n-1)},$$

რაც აიხსნება იმ გარემოებით, რომ თუ დამოუკიდებელ ცვლად-თა რაოდენობა k შედარებადია n-თან, R^2-ის მნიშვნელობა შე-იძლება არარეალურად დიდი გამოვიდეს და მიგვიყვანოს მცდარ დასკვნამდე. იმ შემთხვევაში, როცა n მნიშვნელოვნად დიდია k-სთან შედარებით, მაშინ ეს კოეფიციენტები დაახლოე-ბით ტოლია, საზოგადოგ კი $adjusted\ R^2 < R^2$.

მტკიცდება, რომ R^2 სტატისტიკა ე. წ. შერჩევითი მრავლო-ბითი კორელაციის კოეფიციენტის კვადრატის ტოლია $R^2 = r^2$, რომელიც წარმოადგენს შერჩევითი კორელაციის კოეფიციენტს Y და $\hat{Y}$ შემთხვევით სიდიდეებს შორის:

$$r = \frac{\sum_{i=1}^{n}(Y_i - \overline{Y})(\hat{Y}_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2} \cdot \sqrt{\sum_{i=1}^{n}(\hat{Y}_i - \overline{Y})^2}}.$$

მოდელის ვარგისიანობის შემოწმება. მოდელის ვარგისიანო-ბის შემოწმება მრავლობითი რეგრესიის შემთხვევაში მდგომა-რეობს $H_0 : B_1 = B_2 = \cdots = B_k = 0$ ჰიპოთეზის შემოწმებაში $H_1 :$ *ერთი მაინც* $B_i \neq 0, i = 1, 2, ..., k$ ალტერნატივის წინააღმდეგ.

გასაგებია, რომ თუ H_0 ჰიპოთეზა დაწუნებული არ იქნება, არც ერთი დამოუკიდებელი ცვლადი (რეგრესორი) არ არის წრფივად დაკავშირებული Y-თან და ამიტომ მოდელი უვარგი-სია პროგნოზირების მიზნებისათვის. მაშინ როცა, თუ ერთი მა-ინც $B_i \neq 0$, მოდელი რაიმე თვალსაზრისით მაინც გამოსადეგია.

მტკიცდება, რომ ნულოვანი ჰიპოთეზის სამართიანობისას

$$F = \frac{SSR/k}{SSE/(n-(k+1))} = \frac{MSR}{MSE} \cong F(k, n-(k+1)).$$

F -სტატისტიკა შეიძლება ჩაიწეროს მრავლობითი დეტერმინაციის კოეფიციენტის საშუალებით:

$$F = \frac{R^2/k}{(1-R^2)/(n-(k+1))} = \frac{n-(k+1)}{k} \cdot \frac{R^2}{1-R^2},$$

ე.ი. F -სტატისტიკა $[n-(k+1)]/k$ თანამამრავლის სიზუსტით ემთხვევა სტატისტიკას: $R^2/(1-R^2)$, რომელიც წარმოადგენს ახსნილი ვარიაციის ფარდობას აუხსნელთან და თუ ეს ფარდობა საკმარისად დიდია, ბუნებრივია უარვყოთ H_0 ჰიპოთეზა H_1 ალტერნატივის სასარგებლოდ. შესაბამისად, კრიტიკული არე იქნება – $F > F_{კრიტ}$ სახის.

გადაწყვეტილების მიღების წესი: თუ F სტატისტიკის დაკვირვებული მნიშვნელობისათვის f და α მნიშვნელოვნების დონისათვის სრულდება პირობა

$$f = \frac{n-(k+1)}{k} \cdot \frac{R^2}{1-R^2} > F_{k,n-k-1,\alpha},$$

მაშინ უარვყოფთ H_0 ჰიპოთეზას, წინააღმდეგ შემთხვევაში H_0-ის უარყოფის საფუძველი არ გაგვაჩნია.

5. სტატისტიკური დასკვნები მრავლობითი მოდელის პარამეტრების შესახებ.

თუ H_0 ჰიპოთეზა დაწუნებულია, მაშინ ვამოწმებთ ჰიპოთეზას ცალკეული რეგრესიის კოეფიციენტის ნულთან ტოლობის შესახებ. ამრიგად, ყოველი j -სათვის ($j = 1, 2, ..., k$) მოწმდება $H_0 : B_j = 0$ ჰიპოთეზა $H_1 : B_j \neq 0$ ალტერნატივის წინააღმდეგ. ამ შემთხვევაში ტესტის სტატისტიკაა $T_{b_j} = b_j / S_{b_j}$, რომელსაც ნულოვანი ჰიპოთეზის სამართლიანობისას აქვს სტი-

უდენტის განაწილება თავისუფლების ხარისხით $n-(k+1)$: $T_{b_j}=b_j/S_{b_j}\cong t(n-(k+1))$.

გადაწყვეტილების მიღების წესი: თუ T_{b_j} სტატისტიკის დაკვირვებული მნიშვნელობისათვის t_{b_j} და α მნიშვნელოვნების დონისათვის სრულდება პირობა

$$|t_{b_j}|>t_{n-k-1,\alpha/2},$$

მაშინ H_0 ჰიპოთეზას უარვყოფთ, წინააღმდეგ შემთხვევაში H_0-ის უარყოფის საფუძველი არ გაგვაჩნია.

α მნიშვნელოვნების დონის ნდობის ინტერვალი B_j კოეფიციენტისათვის შემდეგია: $b_j-s_{b_j}\cdot t_{n-k-1,\alpha/2}<B_j<b_j+s_{b_j}\cdot t_{n-k-1,\alpha/2}$, სადაც $t_{n-k-1,\alpha/2}$ თავისუფლების $n-(k+1)$ ხარისხის მქონე სტიუდენტის განაწილების ზედა $\alpha/2$-კრიტიკული წერტილია.

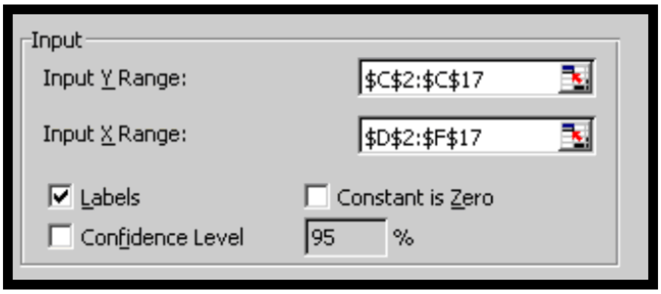
დავუბრუნდეთ უძრავი ქონების აგენტის ამოცანას, რომელსაც ჩვენ ვიხილავდით მარტივი წრფივი რეგრესიის კონტექსტში. როგორც ვნახეთ, სახლის ფასის ცვალებადობის ახსნა მხოლოდ ფართობის ცვალებადობის ხარჯზე არ იძლევა დამაკმაყოფილებელ შედეგს (დეტერმინაციის კოეფიციენტი იყო $R^2=0.648$, ანუ ახსნილია ცვალებადობის მხოლოდ 64.8%). აგენტმა გადაწყვიტა განეხილა მრავლობითი წრფივი რეგრესიის მოდელი:

$$Y=B_0+B_1X_1+B_2X_2+B_3X_3+\varepsilon,$$

სადაც Y – სახლის გაყიდვის ფასია ($\times 1000\$$-ში), X_1 – სახლის ფართობია (× 10 კვ. მ.-ში), X_2 – სახლის ექსპლოატაციის ხანგრძლივობაა (წლებში), ხოლო X_3 – სახლის კუთვნილი მიწის ნაკვეთის ფართობია (× 100 კვ. მ.-ში). შეგროვებული მონაცემები ჩაწერილია Excel-ის დავთრის ფურცელში (ქვეპროგრამა: **Tools/ Data Analysis/Regression):**

	A	B	C	D	E
1	სახლის ნომერი	**Y**	X_1	X_2	X_3
2	1	89.5	20	5	4.1
3	2	79.9	14.8	10	6.8
4	3	83.1	20.5	8	6.3
5	4	56.9	12.5	7	5.1
6	5	66.6	18	8	4.2
7	6	82.5	14.3	12	8.6
8	7	126.3	27.5	1	4.9
9	8	79.3	16.5	10	6.2
10	9	119.9	24.3	2	7.5
11	10	87.6	20.2	8	5.1
12	11	112.6	22	7	6.3
13	12	120.8	19	11	12.9
14	13	78.5	12.3	16	9.6
15	14	74.3	14	12	5.7
16	15	74.8	16.7	13	4.8

დეტერმინაციის კოეფიციენტი. მოცემული მონაცემები შევიყვანოთ კომპიუტერში შემდეგი წესით: საწყისი (შემავალი) მონაცემების Y ცვლადის არეში ჩავწეროთ დამოკიდებული ცვლადის რიცხვითი მნიშვნელობების მისამართი $\$C\$2:\ \$C\$17$, X ცვლადის არეში – $\$D\$2:\$F\$17$:



ქვეპროგრამა **Regression** გვაძლევს შემდეგ ამონაბეჭდს:

SUMMARY OUTPUT

Regression Statistics	
Multiple R / შერჩევითი კორელაციის კოეფიციენტი	0.957

R Square / დეტერმინაციის კოეფიციენტი	0.916
Adjusted R Square / შეთანხმებული დეტერმინაციის კოეფიციენტი	0.893
Standard Error / სტანდარტული შეცდომა	6.894
Observations / დაკვირვებები	15.000

როგორც ვხედავთ, ამ შემთხვევაში დეტერმინაციის კოეფიციენტი (R-square) გამოვიდა $R^2 = 0.916$ ანუ 91.6%, ხოლო $adjusted\ R^2 = 0.893$ ანუ 89.3%. გავიხსენოთ, რომ მარტივი წრფივი რეგრესიის შემთხვევაში $R^2 = 0.648$. ამრიგად, ორი დამატებითი ცვლადის შემოყვანამ საგრძნობლად გაზარდა დეტერმინაციის კოეფიციენტი, სახლის ფასის ვარიაციის 91.6% აიხსნა სამი დამოუკიდებელი ცვლადით და აუხსნელი დარჩა მხოლოდ 8.4%.

მოდელის ვარგისიანობის შემოწმება. კომპიუტრული ამონაბეჭდის მეორე ბლოკი (ANOVA) წარმოადგენს დისპერსიული ანალიზისათვის საჭირო სიდიდეების რიცხვით მნიშვნელობებს:

ANOVA დისპერსიული ანალიზი

	df თავისუფლების ხარისხი	SS *კვადრ. ჯამი*	*MS საშ. კვადრ. გადახრა*	*F* *F სტატისტიკა*	*Significance F მნიშვნელოვნების დონე*
Regression	3	*SSR*= =5707.43	*MSR*= $\frac{SSR}{k}$ = =1902.479	$f = \frac{MSR}{MSE} =$ =40.029	0.000
Residual	11	*SSE*= =522.79	MSE= $\frac{MSE}{n-k-1}$ = =47.527		
Total	14	*SST*= =6230.23			

აქ ბოლო სვეტში მითითებულია $\alpha = P\{F > f \mid H_0\}$-ის რიცხვითი მნიშვნელობა, ანუ მნიშვნელოვნების ის მინიმალური დონე, რომლითაც არსებული მონაცემების საფუძველზე ხდება H_0 ჰიპოთეზის უარყოფა.

თუ ავიღებთ, მაგალითად, $\alpha = 0.05$, მაშინ $F_{k,n-k-1,\alpha} = F_{3,11,0.05} = 8.76$ და ვინაიდან $40.029 = f > F_{3,11,0.05} = 8.76$, ამიტომ ძირითად ჰიპოთეზას უკუვაგდებთ $\alpha = 0.05$ მნიშვნელოვნების დონით და ვასკვნით, რომ ერთი მაინც რეგრესიის კოეფიციენტი განსხვავებულია ნულისაგან.

სტატისტიკური დასკვნები მრავლობითი მოდელის პარამეტრების შესახებ. ამონაბეჭდის მესამე ბლოკიდან ვიღებთ ინფორმაციას უცნობი რეგრესიის კოეფიციენტების შეფასების შესახებ:

	Coefficients კოეფიციენტი	*Standard Error სტანდარტული შეცდომა*	*t Stat t ტესტი*	*P-value P მნიშვნელობა*	*Lower 95%*	*Upper 95%*
Intercept (b_0)	-16.058	19.071	-0.842	0.418	-58.033	25.917
x1 (b_1)	4.146	0.751	5.520	0.000	2.493	5.800
x2 (b_2)	-0.236	0.881	-0.268	0.794	-2.176	1.703
x3 (b_3)	4.831	0.901	5.361	0.000	2.848	6.814

პირველ სვეტში მოცემულია b_0, b_1, b_2, b_3 შეფასებები; მეორეში – მათი სტანდარტული გადახრები, s_{b_j}; მესამეში – $T_{b_j} = b_j / S_{b_j}$ სტატისტიკის დაკვირვებული მნიშვნელობები, $t_{b_j} = b_j / s_{b_j}$; მეოთხე სვეტში – T_{b_j} სტატისტიკებისათვის p-მნიშვნელოებები შესაბამისი $H_0 : B_j = 0$ ჰიპოთეზის შესამოწმებლად $H_1 : B_j \neq 0$ ალტერნატივის წინააღმდეგ, ხოლო ბოლო ორ სვეტში კი მითითებულია 95%-იანი ნდობის ინტერვალები B_j-სათვის.

$b_0 = -16.058$ მნიშვნელობა მიუთითებს საშუალო გასაყიდ ფასს, როცა $X_1 = X_2 = X_3 = 0$ და მისი ინტერპრეტაცია აზრს მოკლებულია; $b_1 = 4.146$ მნიშვნელობა მიუთითებს იმაზე, რომ ფართობის ყოველი დამატებითი ერთეულისათვის სახლის გაყიდვის ფასი იზრდება საშუალოდ 4146 დოლარით; $b_2 = -0.236$ მნიშვნელობა მიუთითებს იმაზე, რომ სახლის ექსპლოატაციის ყოველი დამატებითი წლისათვის ფასი იკლებს 236 დოლარით; $b_3 = 4.831$

მნიშვნელობა კი მიუთითებს იმაზე, რომ მიწის ნაკვეთის ყოველი დამატებითი 100 კვ. მეტრისათვის სახლის ფასი იზრდება 4831 დოლარით.

ტესტის სტატისტიკის $T_{b_2}=b_2/S_{b_2}$ დაკვირვებული მნიშვნელობაა $t_2=-0.268$, ხოლო p −მნიშვნელობაა 0.794, რაც იმის მაჩვენებელია, რომ ნულოვანი ჰიპოთეზის უარყოფა ხდება მნიშვნელოვნების მხოლოდ ისეთი α დონით, რომელიც არანაკლებია 0.794-ზე. ამიტომ მნიშვნელოვნების $\alpha=0.05$ დონით (0.05<0.794) H_0 ჰიპოთეზას არ უარვყოფთ.

B_3-ის შემთხვევაში p −მნიშვნელობა ნულია. ე. ი. ნებისმიერი $\alpha>0$ დონით (კერძოდ, როცა $\alpha=0.05$) $H_0:B_3=0$ ჰიპოთეზას უარვყოფთ $H_1:B_3\neq 0$ ალტერნატივის წინააღმდეგ.

ცხრილის თანახმად B_3 კოეფიციენტის 95%-იანი ნდობის ინტერვალია (2.848, 6.814). ვნახოთ, რომ იმავე შედეგს მივიღებდით უშუალო გამოთვლებითაც. მართლაც, 95%-იანი ნდობის ინტერვალი B_3 კოეფიციენტისათვის იქნება:

$$b_j-s_{b_j}\cdot t_{n-k-1,\alpha/2}<B_j<b_j+s_{b_j}\cdot t_{n-k-1,\alpha/2}\,,\ \alpha/2=(1-0.95)/2=0.025\,;$$

$$b_3-s_{b_3}\cdot t_{11,0.025}<B_3<b_3+s_{b_3}\cdot t_{11,0.025}\,,\ t_{11,0.025}=2.201\,;$$

$$4.831-0.901\cdot 2.201<B_3<4.831+0.901\cdot 2.201\,.$$

ამიტომ საბოლოოდ მივიღებთ, რომ B_3 კოეფიციენტის 95%-იანი ნდობის ინტერვალია (2.848, 6.814).

6. რანგობრივი კორელაცია.

განვიხილოთ ინდივიდთა ან ობიექტთა ერთობლიობა, რომლის ყველა წევრი ხასითდება რაიმე ორი ნიშნით. ვიგულისხმოთ, რომ ორი ნიშნით დახასიათების შედეგები ჩაწერილია შემდეგი სახით $\begin{pmatrix}x_1\\y_1\end{pmatrix},\begin{pmatrix}x_2\\y_2\end{pmatrix},\ldots,\begin{pmatrix}x_n\\y_n\end{pmatrix}$,სადაც x_i და y_i არის i-ური ობიექტის დამახასიათებელი ნიშნები. ამ ნიშნებს შეიძლება არც

ჰქონდეთ ზუსტი რიცხობრივი მნიშვნელობები, მაგრამ შესაძლებელი იყოს მათი ისე დალაგება, რომ თითოეული ნიშნის მიხედვით ყოველ ობიექტს მიეცეს რიგითი ნომერი – რანგი.

მაგალითი 1. ქვემოთ მოყვანილია ერთი და იმავე სახის პროდუქციის რანჟირების შედეგები ჩატარებული ორი A და B ექსპერტის მიერ. I სვეტში მოცემულია შესამოწმებელი პროდუქციის პირობითი ნომერი, ხოლო II და III სვეტში A და B ექსპერტების შეფასებები. მაგალითად, A ექსპერტი საუკეთესოდ მიიჩნევს პროდუქციის 3 ნიმუშს, მაშინ როცა B ექსპერტი მას მეოთხე ადგილს მიაწერს. რამდენიმე ნიმუშის ერთნაირი შეფასებისას, თითოეულს მიეწერება შესაბამისი რანგების საშუალო არითმეტიკული. მაგალითად, A ექსპერტი ერთნაირად აფასებს პროდუქციის 5 და 7 ნიმუშებს, ისინი იყოფენ II-III ადგილს, ამიტომ ორივეს მიეწერება $(2+3)/2=2.5$-ის ტოლი რანგი. IV სვეტში A ექსპერტის რანგები დალაგებულია ზრდის მიხედვით, ხოლო მომდევნო V სვეტში მას მიწერილი აქვს იმავე ნიმუშისათვის B ექსპერტის მიერ მიკუთვნებული რანგი.

პროდუქციის პირობითი ნომერი	A ექსპერტის შეფასება	B ექსპერტის შეფასება	R_A	R_B	R_A-R_B	$(R_A-R_B)^2$
1	4	8	1	4	-3	9
2	9	7	2.5	1	1.5	2.25
3	1	4	2.5	5	-2.5	6.25
4	7	6	4	8	-4	16
5	2.5	5	5	3	2	4
6	8	10	6	2	4	16
7	2.5	1	7	6	1	1
8	6	2	8	10	-2	4
9	5	3	9	7	2	4
10	10	9	10	9	1	1
						$\sum = 63.5$

იმისათვის, რომ შევისწავლოთ A და B ექსპერტების კრიტერიუმებს შორის ურთიერთდამოკიდებულება, გამოვიყენოთ **სპირმენის ე. წ. რანგობრივი კორელაციის კოეფიციენტი,** რომელიც შემდეგი ფორმულით გამოითვლება:

$$\rho = 1 - \frac{6}{n(n^2-1)}\sum_{i=1}^{n}\left(R_A - R_B\right)^2 .$$

თუ ექსპერტების მიერ გამოყენებულ კრიტერიუმებს შორის სრული თანხვედრაა, მაშინ $\sum_{i=1}^{n}\left(R_A - R_B\right)^2 = 0$ და $\rho = 1$; თუ უარყოფითი კავშირია, მაშინ $\rho = -1$ და თუ კავშირი არ არსებობს, მაშინ $\rho = 0$.

განხილულ მაგალითში

$$\rho = 1 - \frac{6}{10\cdot(10^2-1)}\cdot 63.5 = 1 - 0.385 = 0.615 .$$

სპირმენის რანგობრივი კორელაციის კოეფიციენტი გამოიყენება შემდეგი ჰიპოთეზების შესამოწმებლად:

H_0 : ნიშნები დამოუკიდებელია, ე. ი $\rho = 0$,

H_1 : ნიშნებს შორის დადებითი კავშირია, ე. ი. $\rho > 0$.

ცნობილია, რომ დიდი მოცულობის შემთხვევაში

$$t = \sqrt{n-2}\,/\sqrt{1-\rho^2} \cong t(n-2) .$$

ამიტომ კრიტერიუმი ასე ყალიბდება: თუ t სტატისტიკის დაკვირვებული $t_{ჟ}$ მნიშვნელობისათვის და მოცემული α მნიშვნელოვნების დონისათვის სრულდება პირობა: $t_{ჟ} > t_{n-2,\alpha}$, მაშინ H_0 ჰიპოთეზას უარვყოფთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს. გასაგებია, რომ ორმხრივი ალტერნატივის შემთხვევაში – $H_1 : \rho \neq 0$, კრიტიკული არე იქნება: $U_1 = \{|t| > t_{n-2,\alpha/2}\}$.

მაგალითი 2. დღისა და საღამოს სწავლების ეკონომისტ სტუდენტებს ეკითხებოდნენ თუ როგორ აფასებენ ისინი სხვადასხვა პროფესიის პრესტიჟულობას 8 ბალიანი სისტემით. იპოვეთ სპირმენის რანგობრივი კორელაციის კოეფიციენტი, თუ გამოკითხვის მონაცემებია:

პროფესია	დღის	საღამოს
ბუღალტერი	6	3
პროგრამისტი	7	2
ბანკის მენეჯერი	2	6
ადმინისტრატორი	5	4

სტატისტიკოსი	1	7
მარკეტინგის მენეჯერი	4	8
ფინანსისტ-ანალიტიკოსი	3	5
წარმოების მენეჯერი	8	1

ამოხსნა. ამ შემთხვევაში გვაქვს პროფესიებისაგან შედგენილი ერთობლიობა, ხოლო ნიშნების როლში განიხილება დღისა და საღამოს სწავლების სტუდენტთა მიერ ამ პროფესიების შეფასებები. გავაკეთოთ შემდეგი ცხრილი:

პროფესია	დღის	საღამოს	$R_A - R_B$	$(R_A - R_B)^2$
ბუღალტერი	6	3	3	9
პროგრამისტი	7	2	5	25
ბანკის მენეჯერი	2	6	–4	16
ადმინისტრატორი	5	4	1	1
სტატისტიკოსი	1	7	–6	36
მარკეტინგის მენეჯერი	4	8	–4	16
ფინანსისტ-ანალიტიკოსი	3	5	–2	4
წარმოების მენეჯერი	8	1	7	49
$\sum$				156

$$\rho = 1 - \frac{6S_\rho}{n(n^2-1)} = 1 - \frac{6\cdot 156}{8\cdot(8^2-1)} = -0.857 .$$

ამოცანები

1. A და B ექსპერტებმა შეაფასეს 7 სახის ერთი და იგივე პროდუქცია:

A	19	17	15	14	13	10	7
B	16	18	10	14	15	11	5

გამოთვალეთ სპირმენის რანგობრივი კორელაციის კოეფიციენტი.

3. იპოვეთ დეტერმინაციის კოეფიციენტი, თუ $SSR = 30$ და $SSE = 10$.

5. რაც უფრო დაბალია საპროცენტო განაკვეთი (x), მით უფრო დაბალია სესხების დაუბრუნებლობის ანუ დეფოლტების

(y) განაკვეთი. ააგეთ 95%-იანი საპროგნოზო ინტერვალი დეფოლტების განაკვეთისათვის, როდესაც საპროცენტო განაკვეთი 8%-ია, თუ შემთხვევით შერჩეული 9 ბანკი მონაცემებია:

x	7	6.6	6	8.5	8	7.5	6.5	7	8
y	38	40	35	46	48	39	36	37	44

7. არსებობს თუ არა კავშირი მარტოობის გრძნობასა და დეპრესიას შორის? ცხრილში მოცემულია 10 პირის ქულები მარტოობისა და დეპრესიის სკალებზე.

მარტოობის ქულა	4	27	18	7	30	12	18	23	19	12
დეპრესიის ქულა	16	37	33	23	34	32	24	29	26	26

ა) ააგეთ გაბნევის დიაგრამა;
ბ) გამოთვალეთ კორელაციის კოეფიცენტი;
გ) როგორია კავშირი ამ ორ ცვლადს შორის?

9. ახსენით რატომ არის ჯვარედინა ნამრავლები მონაცემთა ორ სიმრავლეს შორის კორელაციის საზომი.

11. დაკავშირებულია თუ არა იუმორის გრძნობა ფიზიკურ ჯანმრთელობასთან?

იუმორის კითხვარი	36	68	61	47	35	48	42	30	56	39	51	54	60	29	65	47
ფიზიკური ჯანმრთელობა	9	3	12	12	15	17	10	17	11	17	14	8	3	13	7	8

გამოიყენეთ მნიშნელოვნობის 0.05 დონე.

13. ცხრილში მოცემულია აბიტურენტების მიერ ზოგადი უნარების ვერბალურ და მათემატიკურ ნაწილებში მიღებული ქულები

ვერბალური	15	60	89	22	32	50	75	75
მათემატიკა	20	58	80	12	28	46	60	89

რა დასკვნის გაკეთება შეიძლება ვერბალურ და მათემატიკურ ნაწილებში მიღებულ ქულებს შორის კავშირის შესახებ.

15. დაალაგეთ 5 კორელაციის კოეფიციენტი ისეთნაირად, რომ პირველი აღნიშნავდეს უმცირეს წრფივ კავშირს, უკანასკნელი კი უდიდეს წრფივ კავშირს ცვლადებს შორის.

17. ექიმს აინტერესებდა არის თუ არა დაკავშირებული ეპილეფსიით დაავადებული ბავშვების დოპამინის დონე დღე-ღამეში

Mal-ის ეპიზოდებთან. ორი კვირის დაკვირვების შედეგები მოყვანილია ცხრილში:

Petit mal Episode	1	4	3	4	3	2	2	1
დოპამინის დონე	6.5	8.3	9.7	8.1	7.5	7.0	6.1	6.3

19. სადაზღვევო კომპანია იკვლევს რამდენად ძლიერია კავშირი სამუშაო დღის ხანგრძლივობასა და უბედურ შემთხვევათა სიხშირეს შორის.

საათების რაოდენობა	40	32	36	44	41
შემთხვევათა სიხშირე	1	1	3	8	5

რისი ტოლი იქნება უბედურ შემთხვევათა სიხშირე, თუ კვირაში 42 სამუშაო საათია?

თავი XIX

დისპერსიული ანალიზი (ANOVA)

მოწმდება ჰიპოთეზა სამი ან მეტი პოპულაციის საშუალოს ტოლობის შესახებ. აღებულია R შერჩევა $(x_{1,1},...,x_{1,n_1})$, ..., $(x_{R,1},...,x_{R,n_R})$ მოცულობებით $n_1,n_2,...,n_R$; $N=n_1+n_2+\cdots+n_R$; i-ური შერჩევის შერჩევითი საშუალო და შესწორებული შერჩევითი დისპერსია აღნიშნულია შესაბამისად $\overline{x}_i$ და $s_i^{'2}$ სიმბოლოებით; $\overline{x}_{\text{ერთ.}}=\dfrac{\sum_{i,j}x_{i,j}}{N}$ – არის ერთობლივი საშუალო; $SBB:=\sum_i n_i(\overline{X}_i-\overline{X}_{\text{ერთ.}})^2$, $SBB/\sigma^2\cong\chi^2(R-1)$; $SSW:=\sum_i(n_i-1)S_i^{'2}$, $SSW/\sigma^2\cong\chi^2(N-R)$ ($N-R=\sum_i(n_i-1)$).

ნულოვანი ჰიპოთეზა: H_0 **:** $a_1=a_2=\cdots=a_R$ **.**

ალტერნატივა: H_1 **:** ერთი საშუალო მაინც განსხვავდება დანარჩენებისაგან.

მნიშვნელოვნების დონე: α .

კრიტერიუმის სტატისტიკა: $F=\dfrac{S_B^2}{S_W^2}\cong F(R-1,N-R)$ **.**

კრიტერიუმის მნიშვნელობა T.V.$\equiv f=\dfrac{s_B^2}{s_W^2}$ **, სადაც** s_B^2 **(შესაბამისად,** s_W^2**) არის ჯგუფთა შორის ვარიაცია (შესაბამისად, ჯგუფებში ვარიაცია),**

$$s_B^2:=\frac{\sum_i n_i(\overline{x}_i-\overline{x}_{\text{ერთ.}})^2}{R-1},\quad s_W^2:=\frac{\sum_i(n_i-1)s_i^{'2}}{\sum_i(n_i-1)}.$$

კრიტერიუმი მარჯვენა ცალმხრივია.

კრიტიკული მნიშვნელობა $C.V. = F_{R-1,N-R,\alpha}$.

კრიტიკული არე C.R. (H_0 **-ის უარყოფის არე) =** $[F_{R-1,N-R,\alpha}, +\infty)$.

P **- მნიშვნელობა =** $P\{F(R-1, N-R) > T.V.\}$.

შეზღუდვები: 1. პოპულაციები უნდა იყოს ნორმალური ან დაახლოებით ნორმალური; 2. შერჩევები უნდა იყოს ერთმანეთისაგან დამოუკიდებელი; 3. პოპულაციათა დისპერსიები უნდა იყოს ტოლი.

გადაწყვეტილება: თუ კრიტერიუმის მნიშვნელობა მეტია ან ტოლი კრიტიკულ მნიშვნელობაზე (T.V. ≥C.V.), მაშინ ნულოვან ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

P **-მნიშვნელობის მეთოდი:** თუ $P \le \alpha$, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

მაგალითი 1. მკვლევარს სურს გამოსცადოს მაღალი არტერიული წნევის დაწევის სამი განსხვავებული მეთოდი. შემთხვევით შეირჩა პაციენტთა სამი ჯგუფი. I ჯგუფს აძლევდნენ გარკვეულ პრეპარატს, II ჯგუფს უტარდებოდა სპეციალური ვარჯიშები, ხოლო III ჯგუფი იცავდა სპეციალურ დიეტას. ოთხი კვირის შემდეგ თითოეულ პაციენტს გაუზომეს არტერიული წნევა (მონაცემები მოყვანილია ქვემოთ). $\alpha = 0.05$ მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა იმის შესახებ, რომ პოპულაციათა საშუალოებს შორის არ არსებობს განსხვავება.

პრეპარატი	ვარჯიში	დიეტა
10	6	5
12	8	9
9	3	12
15	0	8
13	2	4

ამოხსნა.

ნაბიჯი 1. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები:

H_0 : $a_1 = a_2 = a_3$, H_1 : ერთი საშუალო მაინც განსხვავდება დანარჩენებისაგან.

ნაბიჯი 2. ვიპოვოთ კრიტიკული მნიშვნელობა. რადგანაც თავისუფლების ხარისხებია: $R-1=3-1$ და $N-R=15-3=12$, ამიტომ ფიშერის განაწილების ცხრილიდან ვპოულობთ, რომ $C.V.=F_{R-1,N-R,\alpha}=F_{2,12,0.05}=3.89$.

ნაბიჯი 3. გამოვთვალოთ კრიტერიუმის მნიშვნელობა.

ა) თითოეული შერჩევისათვის ვიპოვოთ შერჩევითი საშუალო და შესწორებული შერჩევითი დისპესია. გვაქვს:

$$\bar{x}_1=11.8,\ s_1^{'2}=5.7;\ \bar{x}_2=3.8,\ s_2^{'2}=10.2;\ \bar{x}_3=7.6,\ s_3^{'2}=10.3.$$

ბ) გამოვთვალოთ ერთობლივი საშუალო:

$$\bar{x}_{\text{ერთ.}}=\frac{\sum_{i,j}x_{i,j}}{N}=\frac{10+12+\cdots+4}{15}=7.73.$$

გ) გამოვთვალოთ ჯგუფთა შორის ვარიაცია:

$$s_B^2:=\frac{\sum_i n_i(\bar{x}_i-\bar{x}_{\text{ერთ.}})^2}{R-1}=$$
$$=\frac{5\cdot(11.8-7.73)^2+5\cdot(3.8-7.73)^2+5\cdot(7.6-7.73)^2}{3-1}=80.07.$$

დ) გამოვთვალოთ ჯგუფებში ვარიაცია:

$$s_W^2:=\frac{\sum_i(n_i-1)s_i^{'2}}{\sum_i(n_i-1)}=\frac{(5-1)\cdot5.7+(5-1)\cdot10.2+(5-1)\cdot10.3}{3\cdot(5-1)}=8.73.$$

ე) გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$f=\frac{s_B^2}{s_W^2}=\frac{80.07}{8.73}=9.17.$$

ნაბიჯი 4. გადაწყვეტილების მიღება: ვინაიდან 9.17 > 3.89, ამიტომ ნულოვანი ჰიპოთეზა უნდა უკუვაგდოთ ანუ $\alpha=0.05$ მნიშვნელოვნების დონით ჩვენ გვაქვს საკმარისი საფუძველი დავასკვნათ, რომ სულ ცოტა ერთი საშუალო მაინც განსხვავდება დანარჩენებისაგან.

მაგალითი 2. მარკეტინგის სპეციალისტს სურს გაარკვიოს არის თუ არა განსხვავება მომხმარებლების მიერ სამი სხვადასხვა სუპერმარკეტის სალაროსთან რიგში დგომის დროების საშუალოებს შორის (მონაცემები მოყვანილია ქვემოთ). $\alpha = 0.05$ მნიშვნელოვნების დონით არის თუ არა არსებითი განსხვავება რიგში დგომის დროების საშუალოებს შორის?

მაღაზია A	**მაღაზია B**	**მაღაზია C**
3	5	1
2	8	3
5	9	4
6	6	2
3	2	7
1	5	3

ამოხსნა.

ნაბიჯი 1. ჩამოვაყალიბოთ ძირითადი და ალტერნატიული ჰიპოთეზები:

$H_0: a_1 = a_2 = a_3$, H_1 : ერთი საშუალო მაინც განსხვავდება დანარჩენებისაგან.

ნაბიჯი 2. ვიპოვოთ კრიტიკული მნიშვნელობა. აქ თავისუფლების ხარისხებია: $R-1=3-1$ და $N-R=18-3=15$, ამიტომ ფიშერის განაწილების ცხრილიდან ვპოულობთ, რომ $C.V. = F_{R-1,N-R,\alpha} = F_{2,15,0.05} = 3.68$.

ნაბიჯი 3. გამოვთვალოთ კრიტერიუმის მნიშვნელობა.

ა) თითოეული შერჩევისათვის ვიპოვოთ შერჩევითი საშუალო და შესწორებული შერჩევითი დისპესია. გვაქვს:

$$\bar{x}_1 = 3.33,\ s_1^{'2} = 3.47;\ \bar{x}_2 = 5.83,$$

$$s_2^{'2} = 6.17;\ \bar{x}_3 = 3.33,\ s_3^{'2} = 4.27.$$

ბ) გამოვთვალოთ ერთობლივი საშუალო:

$$\bar{x}_{\text{ერთ.}} = \frac{\sum_{i,j} x_{i,j}}{N} = \frac{3+2+\cdots+3}{18} = 4.17.$$

გ) გამოვთვალოთ ჯგუფთა შორის ვარიაცია:

$$s_B^2 := \frac{\sum_i n_i(\overline{x}_i - \overline{x}_{\text{ერთ.}})^2}{R-1} =$$

$$= \frac{6\cdot(3.33-4.17)^2 + 6\cdot(5.83-4.17)^2 + 6\cdot(3.33-4.17)^2}{3-1} = 12.5.$$

დ) გამოვთვალოთ ჯგუფებში ვარიაცია:

$$s_W^2 := \frac{\sum_i (n_i - 1)s_i'^2}{\sum_i (n_i - 1)} = \frac{(6-1)\cdot 3.47 + (6-1)\cdot 6.17 + (6-1)\cdot 4.27}{3\cdot(6-1)} = 4.64\,.$$

ე) გამოვთვალოთ კრიტერიუმის მნიშვნელობა:

$$f = \frac{s_B^2}{s_W^2} = \frac{12.5}{4.64} = 2.69\,.$$

ნაბიჯი 4. გადაწყვეტილების მიღება: ვინაიდან 2.69 < 3.68, ამიტომ ნულოვანი ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს ანუ $\alpha = 0.05$ მნიშვნელოვნების დონით ჩვენ გვაქვს საკმარისი საფუძველი დავასკვნათ, რომ საშუალოები არსებითად არ განსხვავდება.

ამოცანები

იგულისხმეთ, რომ: პოპულაციები ნორმალური ან დაახლოებით ნორმალურია; შერჩევები ერთმანეთისაგან დამოუკიდებელია და პოპულაციათა დისპერსიები ტოლია.

1. მენეჯერს აინტერესებს არის თუ არა რესპუბლიკური საავადმყოფოს ექიმების, ექთანებისა და ტექნიკური პერსონალის საშუალო ასაკი განსხვავებული. საავადმყოფოს შემთხვევით შერჩეული თანამშრომლების ქვემოთ მოყვანილი მონაცემების მიხედვით, $\alpha = 0.05$ მნიშვნელოვნების დონით, შეუძლია თუ არა მენეჯერს დაასკვნას, რომ ამ სამი ჯგუფის საშუალო ასაკი განსხვავებულია?

ექიმები	ექთნები	ტექნიკური პერსონალი
60	23	33
36	25	28
29	26	35
56	35	29
32	42	23
54	22	41
58		

3. კოლეჯის სპორტულ პროგრამებში მონაწილე სტუდენტების მიერ მოპოვებული ქულების ქვემოთ მოყვანილი მონაცემების მიხედვით, $\alpha = 0.1$ მნიშვნელოვნების დონით, შეგვიძლია თუ არა დაასკვნათ, რომ ამ სამ ჯგუფში ქულების საშუალო განსხვავებულია?

ფეხბურთი	კალათბურთი	ხელბურთი
3.2	3.8	2.6
2.6	3.1	1.9
2.4	2.6	1.7
2.4	3.9	2.5
1.8	3.3	1.9

5. შემთხვევით შერჩეული ათლეტები დაყვეს სამ ჯგუფად და დაუნიშნეს სამი სახის დიეტა ერთი თვის განმავლობაში. ერთი თვის გასვლის შემდეგ თითოეული ათლეტის მიერ დიეტების მიხედვით დაკლებული კილოგრამები მოყვანილია ქვემოთ. $\alpha = 0.05$ მნიშვნელოვნების დონით შეგვიძლია თუ არა დაასკვნათ, რომ დიეტები განსხვავებულია?

დიეტა A	დიეტა B	დიეტა C
3	10	8
6	12	3
7	11	2
4	14	5
	8	
	6	

7. მანქანათმშენებელი ქარხნის მუშები შემთხვევით მიამაგრეს ოთხ ასაწყობ ხაზს ("კონვეირს"). თითოეული მუშის მიერ გა-

კეთებული დეფექტური ნაწილების რაოდენობის ქვემოთ - მოყვანილი რაოდენობების მიხედვით, $\alpha = 0.05$ მნიშვნელოვნების დონით, შეგვიძლია თუ არა დაასკვნათ, რომ დეფექტური ნაწილების რაოდენობის საშუალო ამ ხაზებზე ერთი - და იგივეა?

I ხაზი	II ხაზი	III ხაზი	IV ხაზი
3	8	10	9
2	6	9	15
0	2	8	3
6	0	11	0
4	1	12	2
3	9	15	0
5	7	17	1

9. მკვლევარს აინტერესებს არის თუ არა განსხვავება ოთხი სხვადასხვა ტიპის ბალახის საკრეჭი მანქანის წონას შორის. $\alpha = 0.1$ მნიშვნელოვნების დონით შეგვიძლია თუ არა დავასკვნათ, რომ მანქანების საშუალო წონა განსხვავებულია?

გაზზე	ბენზინზე	ელექტრონული	მექანიკური
95	73	55	37
101	69	52	24
108	72	51	25
97	71	37	29
101	67	57	22
	62	54	17
	68	34	17
	71	45	22
		41	20
		53	18
			21

ამოცანები გამოცდისათვის

11. კომპანიის მფლობელს აინტერესებს არის თუ არა გაყიდვათა რაოდენობა ერთნაირად განაწილებული რეგიონების მიხედ-

ვით. მან შემთხვევით შეარჩია თვეები და დათვალა 5 სხვადასხვა რეგიონში გაყიდვათა რაოდენობა. $\alpha = 0.05$ მნიშვნელოვნების დონით შეუძლია თუ არა კომპანიის მფლობელს დაასკვნას, რომ გაყიდვათა რაოდენობა თითოეულ რეგიონში ერთი და იგივეა?

რეგიონი	**სოხუმი**	**ზუგდიდი**	**გორი**	**თელავი**	**ქუთაისი**
გაყიდვები	324	236	182	221	365

13. იუსტიციის სამინისტროს აინტერესებს სრულწლოვანი მოსახლეობის შეხედულება ნაფიც მსაჯულთა ინსტიტუტის შემოღებასთან დაკავშირებით. ჩატარებული გამოკითხვის ქვემოთ მოყვანილი შედეგების მიხედვით, $\alpha = 0.1$ მნიშვნელოვნების დონით, შეგვიძლია თუ არა დაასკვნათ, რომ შეხედულება დაკავშირებულია ადამიანის სქესთან?

სქესი	**ეთანხმება**	**არ ეთანხმება**	**გაურკვეველია**
მდედრობითი	136	16	8
მამრობითი	114	30	6

15. გამოკითხულ იქნა 40 და 50 წლის ინვესტორები ფულის განთავსების გზების შესახებ. მიღებული მანაცემების მიხედვით, $\alpha = 0.05$ მნიშვნელოვნების დონით, არის თუ არა კავშირი ინვესტორის ასაკსა და ფულის განთავსების გზებს შორის?

	დიდი სააქციო ფონდი	**მცირე სააქციო ფონდი**	**საერთაშორისო სააქციო ფონდი**	**ფულადი ბაზრის ფონდი**	**ობლიგაცია**
40 წლის	20	10	10	15	45
50 წლის	42	24	24	6	24

17. გამოკითხულ იქნა 3 მაღაზიის 200 – 200 მომხმარებელი, რომლებიც სარგებლობენ საკრედიტო ბარათებით. მათ დაუსვეს კითხვა „ქონდათ თუ არა საკრედიტო ბარათით სარგებლობისას პრობლემები?“ $\alpha = 0.01$ მნიშვნელოვნების დონით შეიძლება თუ არა დავასკვნათ, რომ პროპორციები ერთი და იგივეა?

	მაღაზია I	**მაღაზია II**	**მაღაზია III**
დიახ	87	56	43

არა	113	144	157
ჯამი	200	200	200

19. მამაკაცის სამი ტიპის ფეხსაცმელის ქვემოთ მოყვანილი ფასების მიხედვით, $\alpha = 0.05$ მნიშვნელოვნების დონით, შეგვიძლია თუ არა დავასკვნათ, რომ ფასების საშუალოებს შორის მნიშვნელოვანი განსხვავებაა?

ოფიციალური	**ყოველდღიური**	**სპორტული**
110	80	64
95	100	66
95	135	70
265	90	92
59	80	
70		
50		

21. ქვემოთ მოყვანილია ნაწლავის ჩხირების რაოდენობა სამი სხვადასხვა ტბის გარკვეულ ფართობზე ხუთდღიანი პერიოდის განმავლობაში. $\alpha = 0.05$ მნიშვნელოვნების დონით არის თუ არა განსხვავება ჩხირების რაოდენობის საშუალოებში?

კუს ტბა	**ლისის ტბა**	**ფარავნის ტბა**
45	97	33
53	82	35
41	99	31
38	84	28
55	79	26

დანართი 1

EXCEL-ის სტატისტიკური ფუნქციების აღწერა

ალბათობა

ფაქტორიალი – **FACT(*n*): = *n*!;**

წყობა – **PERMUT(*n*,*k*): = *n*!/(*n* – *k*)!;**

ჯუფდება – **COMBIN(n,k): = *n*!/[*k*!(*n* – *k*)!];**

EXCEL-ში დისკრეტული განაწილებების ჩაწერისას გამოიყენება შემდეგი აღნიშვნები:

აღნიშვნა	შინაარსი
x	განაწილების სიმკვრივის ან განაწილების ფუნქციის არგუმენტის რიცხვითი მნიშვნელობა
X range	არე, სადაც მითითებულია შემთხვევითი სიდიდის შესაძლო მნიშვნელობები
Prob_range	არე, სადაც მითითებულია შემთხვევითი სიდიდის შესაძლო მნიშვნელობების ალბათობები
Lower_limit	ქვედა საზღვარი
Upper_limit	ზედა საზღვარი
Mean	განაწილების მათემატიკური ლოდინი
Numbers	რაოდენობა (წარმატებათა)
Trials	ცდათა სერიის სიგრძე
Probability_s	ალბათობა
Sample_s	I სახის ობიექტთა რაოდენობა შერჩევაში
Number_sample	შემთხვევით (დაბრუნების გარეშე) ამორჩეულ ობიექტთა რაოდენობა – შერჩევის მოცულობა
Population_s	I სახის ობიექტთა რაოდენობა პოპულაციაში
Number_pop.	ობიექტების რაოდენობა პოპულაციაში
Cumullative	თუ **Cumullative** პარამეტრის რიცხვითი მნიშვნელობა

	ნულის ტოლია, მაშინ გამოიანგარიშება $P\{X=k\}$ ალბათობა; თუ **Cumullative** პარამეტრის რიცხვითი მნიშვნელობა ერთის ტოლია, მაშინ გამოიანგარიშება $P\{X \le k\}$ ალბათობა

PROB(x_range,prob_range,lower_limit,upper_limit) – ფუნქციით გამოითვლება ალბათობა: $P\{Lowr\ limit \le X \le Upper\ limit\}$; მაგალითად, თუ X შემთხვევითი სიდიდის კანონი EXCEL-ში არის:

	A	B	C	D	E	F
1	-1	0	2	4	6	9
2	0.05	0.15	0.18	0.22	0.3	0.1

მაშინ **PROB(A1:F1,A2:F2,0,7)** – გამოითვლის $P\{0 \le X \le 5\}=0.45$-ს.

BINOMDIST(number_s,trials,probabilitys,cumullative) – ფუნქციით გამოითვლება ბინომურად განაწილებული შემთხვევითი სიდიდისათვის $P\{X=k\}$ ალბათობა, თუ **cumullative** = 0-ს და $P\{X \le k\}$ ალბათობა, თუ **cumullative** =1-ს.

CRITBINOM(trials,probability_s,alpha) – ფუნქციით, მოცემული a-სათვის, გამოითვლება ის მინიმალური k, რომლისთვისაც: $\sum_{m=0}^{k} P_n(m) \ge a$.

HYPGEOMDIST(sample_s,number_sample,population_s,number_pop.) – ფუნქციით გამოითვლება ჰიპერგეომეტრიული განაწილებისათვის $P\{X=k\}$ ალბათობა. მაგალითად, თუ ყუთში 100 ბურთია, რომელთა შორის 30 თეთრია და შემთხვევით, დაბრუნების გარეშე, ვიღებთ 45 ბურთს, მაშინ ალბათობა იმისა, რომ მათ შორის აღმოჩნდება 10 თეთრი ბურთი გამოითვლება შემდეგნაირად:

HYPGEOMDIST(10,40,30,100) = $C_{30}^{10} \cdot C_{70}^{35} / C_{100}^{45}$.

POISSON(x,mean,cumulative) – ფუნქციით გამოითვლება პუასონის კანონით განაწილებული შემთხვევითი სიდიდისათვის $P\{X=k\}$ ალბათობა, თუ **cumulative** = 0-ს და $P\{X \le k\}$ ალბათობა, თუ **cumulative** = 1-ს.

EXCEL-ში უწყვეტი განაწილებების ჩაწერისას გამოიყენება შემდეგი აღნიშვნები:

აღნიშვნა	შინაარსი
x	განაწილების სიმკვრივის ან განაწილების ფუნქციის არგუმენტის რიცხვითი მნიშვნელობა
Mean	განაწილების მათემატიკური ლოდინი
Standard_dev	განაწილების საშუალო კვადრატული გადახრა
Cumullative	თუ **Cumullative** პარამეტრის რიცხვითი მნიშვნელობა ნულის ტოლია, მაშინ გამოითვლება განაწილების სიმკვრივის რიცხვითი მნიშვნელობა; თუ **Cumullative** პარამეტრის რიცხვითი მნიშვნელობა ერთის ტოლია, მაშინ გამოითვლება განაწილების ფუნქციის რიცხვითი მნიშვნელობა
Lambda	λ პარამეტრი
Probability	ალბათობის P დონე, რომლისთვისაც გამოითვლება კვანტილი
Deg_freedom	განაწილების თავისუფლების ხარისხი
Probability	$P\{X > k\}$ ალბათობის მნიშვნელობა
Tails	თუ **Tails** პარამეტრის რიცხვითი მნიშვნელობა ერთის ტოლია, მაშინ **TDIST** ფუნქცია გვაძლევს ცალმხრივ $T(n)>x$ არეში მოხვედრის ალბათობას, $P\{T(n) > x\}$, სტიუდენტის განაწილებისათვის; თუ **Tails** პარამეტრის რიცხვითი მნიშვნელობა ორის ტოლია, მაშინ **TDIST** ფუნქცია გვაძლევს ორმხრივ $\lvert T(n)\rvert > x$ არეში მოხვედრის ალბათობას, $P\{\lvert T(n)\rvert > x\}$-ს, სტიუდენტის განაწილებისათვის

NORMSDIST(x) – ფუნქციით გამოითვლება სტანდარტული ნორმალური $N(0,1)$ შემთხვევითი სიდიდისათვის $P\{N(0,1) \le x\}$ ლბათობა.

NORMDIST(x,mean,standard_dev,cumullative) – ფუნქციით, თუ **cumullative** პარამეტრის რიცხვითი მნიშვნელობა ნულის ტოლია, გამოითვლება $N(\mu,\sigma^2)$-ის (არასტანდარტული ნორმალური) განაწილების სიმკვრივის მნიშვნელობა; ხოლო თუ **cumullative** პარამეტრის რიცხვითი მნიშვნელობა ერთის ტოლია, გამოითვლება $N(\mu,\sigma^2)$-ის განაწილების ფუნქციის მნიშვნელობა.

NORMSINV(probability) – ფუნქციით გამოითვლება სტანდარტული ნორმალური განაწილების p დონის კვანტილი. მაგალითად, **NORMSINV(0.65)** $= x_{0.65} = 0.39$.

NORMINV(probability,mean,standard_dev) – ფუნქციით გამოითვლება არასტანდარტული ნორმალური განაწილების p დონის კვანტილი.

EXPONDIST(x,lambda,cumullative) – ითვლის $\lambda e^{-\lambda x}$ ფუნქციის მნიშვნელობას, როცა $\text{Cumullative} = 0$-ს და $1 - e^{-\lambda x}$ ფუნქციის მნიშვნელობას, როცა Cumullative = 1-ს.

CHIDIST(x,deg_freedom) – ითვლის $P\{\aleph^2(k) > x\}$ ალბათობას ხი-კვადრატ განაწილებისათვის.

CHIINV(probability,deg_freedom) – ითვლის ხი-კვადრატ განაწილების ზედა კრიტიკულ წერტილს, ანუ ისეთ x რიცხვს, რომლისთვისაც $P\{\aleph^2(Deg\ freedom) > x\} = Probability$.

TDIST(x,deg_freedom,tails): TDIST(x,k,1) – ითვლის $P\{T(k) > x\}$ ალბათობას სტიუდენტის $T(k)$ შემთხვევითი სიდიდისათვის, ხოლო **TDIST(x,k,2)** – კი ითვლის $P\{|T(k)| > x\}$ ალბათობას სტიუდენტის $T(k)$ შემთხვევითი სიდიდისათვის.

TINV(probability,deg_freedom): TINV(α,*k*) – პოულობს ისეთ ***x***-ს, რომლისთვისაც $P\{|T(k)| > x\} = a$ (ფაქტიურად ***x*** იქნება სტიუდენტის $T(k)$ შემთხვევითი სიდიდის ზედა $a/2$-კრიტიკული წერტილი).

FDIST(x,deg_freedom1,deg_freedom2): FDIST(x,n,m) – ითვლის $P\{F(n,m) > x\}$ ალბათობას ფიშერის $F(n,m)$ შემთხვევითი სიდიდისათვის.

FINV(probability,deg_freedom1,deg_freedom2): FDIST(*a*,*n*,*m*) – პოულობს ფიშერის ***F*(*n*,*m*)** განაწილების ზედა ***a***-კრიტიკულ წერტილს, ანუ ისეთ ***x***-ს, რომლისთვისაც $P\{F(n,m) > x\} = a$.

სტატისტიკა

EXCEL-ის ქვეპროგრამა **Tools/Data Analysis/Random Number Generation** საშუალებას იძლევა მივიღოთ ფსევდოშემთხვევითი რიცხვების რეალიზაციები. ამ დროს გამოიყენება შემდეგი აღნიშვნები:

აღნიშვნა	შინაარსი
Number of variables	დაკვირვებულ სერიათა რაოდენობა
Number of Random Numbers	მონაცემთა რაოდენობა სერიაში
Distribution	განაწილების კანონი
Discrete	დისკრეტული
Value and Probability Input Range	ეთითება ვერტიკალურად ჩაწერილი დისკრეტული განაწილების კანონი, რომლის მიხედვითაც უნდა გათამაშდეს ფსევდოშემთხვევითი სიდიდეები
Uniform	თანაბარი
Parameters	ეთითება თანაბარი განაწილების a და b პარამეტრები
Normal	ნორმალური
Mean	საშუალო
Standard Deviation	სტანდარტული გადახრა
Bernouli	ბერნული
Binomial	ბინომური
***p*-Value**	p პარამეტრის რიცხვითი მნიშვნელობა
Number of Trials	n პარამეტრის რიცხვითი მნიშვნელობა ბერნულის სქემაში
Poisson	პუასონი
Lambda	λ პარამეტრი

CONFIDENCE(alpha,standard_dev,size): CONFIDENCE(a,σ,n) – ნორმალური პოპულაციის საშუალოსათვის ითვლის $(1-\alpha)$ საიმედოობის ნდობის ინტერვალის სიგრძის ნახევარს – $z_{\alpha/2}\cdot\sigma/\sqrt{n}$ -ს.

p-მნიშვნელობა

ZTEST ფუნქცია გამოიყენება Z სტატისტიკის *p*-მნიშვნელობის გამოსათვლელად. ამ დროს გამოიყენება შემდეგი აღნიშვნები:

აღნიშვნა	შინაარსი
Array	დაკვირვებულ მონაცემთა არე
x	პოპულაციის საშუალო
Sigma	პოპულაციის სტანდარტული გადახრა

ZTEST(array,x,sigma) – Z სტატისტიკის დაკვირვებული $z=\dfrac{\overline{x}-a_0}{\sigma/\sqrt{n}}$ მნიშვნელობისათვის (მარჯვენა ცალმხრივი კრიტერიუმის შემთხვევაში) ითვლის *p*-მნიშვნელობას, ანუ $P\left\{Z\geq\dfrac{\overline{x}-a_0}{\sigma/\sqrt{n}}\right\}$ ალბათობას.

CHITEST ფუნქცია გამოიყენება ხი-კვადრატ განაწილების *p*-მნიშვნელობის გამოსათვლელად. ამ დროს გამოიყენება შემდეგი აღნიშვნები:

აღნიშვნა	შინაარსი
Actual_range	დაკვირვებული სიხშირეები
Expected_range	მოსალოდნელი სიხშირეები

CHITEST(actual_range,expected_range) – ითვლის $(R-1)(C-1)$ თავისუფლების ხარისხის მქონე ხი-კვადრატ განაწილების *p*-მნიშვნელობას.

ორამოკრეფიანი ამოცანები

TTEST(array1,array2,tails,type)-ისა და **FTEST(array1,array2)**-ით სარგებლობის დროს გამოიყენება აღნიშვნები:

აღნიშვნა	შინაარსი
Array	არე
Tails	კრიტიკული არე
Type	ტიპი

და ქვეპროგრამები:

- **z-Test: Two Sample for Means** – Z-ტესტი (ორამოკრეფიან ამოცანაში) საშუალოების შესახებ;
- **t-Test: Paired Two Sample for Mean** – T-ტესტი დაწყვილე-ბული მონაცემთა საშუალოებისათვის;
- **t-Test: Two Sample Assuming Equal Variances** – T-ტესტი უცნობი ტოლი დისპერსიების შემთხვევაში;
- **t-Test: Two Sample Assuming Unequal Variances** – T-ტესტი უცნობი არატოლი დისპერსიების შემთხვევაში;
- **F-Test: Two Sample for Variances** – F-ტესტი დისპერსიებისა-თვის.

ამ ქვეპროგრამებით სარგებლობისას გამოიყენება შემდეგი აღ-ნიშვნები:

აღნიშვნა	შინაარსი
Input	შემავალი მონაცემები
Variable 1(2) Range	დაკვირვებათა არე
Variance (known)	პოპულაციათა დისპერსიების მნიშვნელობები
Labels	მონიშვნა
Hypothesized Mean Difference	ძირითადი ჰიპოთეზით განსაზღრული პოპულაციათა საშუალოების სხვაობა (ანუ ნული)
Alpha	მნიშვნელოვნების დონე (0.05)

z-Test-ის გამოყენების მაგალითი. ორი დამოუკიდებელი ნორმალური პოპულაციიდან დისპერსიებით შესაბამისად, 5 და 9 მიღებულია შემდეგი შერჩევები: I) 2.4, -2.3, 5.2, 10.3, 9.9, 12.6, -6.9, 2.8, 9.4, -1.4 და II) -8.2, 17.4, 16.4, 4, 21.7, -3.2, -2.9. 0.05 მნიშვნელოვნების დონით შევამოწმოთ ჰიპოთეზა საშუალოთა ტოლობის შესახებ.

ამოხსნა. მოწმდება ძირითადი ჰიპოთეზა $H_0: a_1 - a_2 = 0$ ცა-ლმხრივი $H_1: a_1 - a_2 < 0$ და ორმხრივი $H_1: a_1 - a_2 \neq 0$ ალტერნატი-ვების წინააღმდეგ. ქვეპროგრამა **z-Test: Two Sample for Means** გვაძლევს:

	I	II
Mean / საშუალო	4.2	6.5
known Variance / ცნობილი დისპერსიები	5	9
Observations / დაკვირვებები	10	7
Hypothesized Mean Difference / ძირითადი ჰიპოთეზით განსაზღრული საშუალოთა სხვაობა	0	
z Stat /Z-სტატისტიკის დაკვირვებული მნიშვნელობა	-1.689	
$P\{Z \le z\}$ **one-tail** / p-მნიშვნელობა ცალმხრივი კრიტიკული არისათვის	0.046	
z **Critical one-tail** / კრიტიკული წერტილი ცალმხრივი კრი-ტიკული არისათვის	1.645	
$P\{Z \le z\}$ **two-tail** / p-მნიშვნელობა ორმხრივი კრიტიკული არისათვის	0.091	
z **Critical two-tail** / კრიტიკული წერტილი ორმხრივი კრიტი-კული არისათვის	1.96	

დასკვნა:

- **ცალმხრივი $H_1: a_1 - a_2 < 0$ ალტერნატივისათვის *p*-მნიშვნელობა $= 0.046 < \alpha = 0.05$. შესაბამისად, 0.05 მნიშვნელოვნების დონით ძირითად ჰიპოთეზას უარვყოფთ;**
- **ორმხრივი $H_1: a_1 - a_2 \neq 0$ ალტერნატივისათვის *p*-მნიშვნელობა $= 091 < \alpha = 0.05$. შესაბამისად, 0.05 მნიშვნელოვნების დონით ძირითადი ჰიპოთეზის უარვყოფის საფუძველი არ გვაქვს.**

t-Test-ის გამოყენების მაგალითი. კომპანიის მენეჯერს სურს შეადაროს ორი ***A*** და ***B*** შეთანხმებიდან შემოსული თვიური შემოსავლები. ***A*** შეთანხმება გულისხმობს დაბალ საარენდო გადასახადს, მაგრამ ავალდებულებს არენდატორს დაზიანებული საარენდო ქონების აღდგენის ხარჯების დაფარვას. ***B*** შეთანხმება ადგენს მაღალ საარენდო გადასახადს, მაგრამ კომპანია თავის თავზე იღებს დაზიანებული საარენდო ქონების აღდგენის ხარჯებს. მენეჯერმა შეაგროვა თითოეული შეთანხმების თვიური შემოსავლები და $\alpha = 0.05$ მნიშვნელოვნების დონით სურს შეამოწმოს ჰიპოთეზა, რომ ამ შეთანხმებების საშუალო შემოსავლები არ განსხვავდება ერთმანეთისაგან. ეს მონაცემებია:

A შეთანხმება	84.5	102.6	79.1	92.5	101.7	78.7	88.5	93	72.2	96.8	92.6	93.7
B შეთანხმება	106	102.1	105.1	91.5	82.3	88.4	101.1	97.5	75.7	93.6		

ამოხსნა 1. ვიგულისხმოთ, რომ ორივე პოპულაცია ნორმალურადაა განაწილებული და უცნობი დისპერსიები ტოლია. უნდა შემოწმდეს ძირითადი ჰიპოთეზა $H_0: a_1 - a_2 = 0$ ცალმხრივი $H_1: a_1 - a_2 < 0$ და ორმხრივი $H_1: a_1 - a_2 \neq 0$ ალტერ-ნატივების წინააღმდეგ. ქვეპროგრამა **t-Test: Two Assuming Equal Variances** გვაძლევს:

	A	***B***
Mean / საშუალო	89.658	94.33
Variance / დისპერსია	88.366	100.438
Observations / დაკვირვებები	12	10
Pooled Variance / დაწყვილებული მონაცემების დისპერსია	93.799	
Hypothesized Mean Difference / ძირითადი ჰიპოთეზით განსაზღრული საშუალოთა სხვაობა	0	
df / თავისუფლების ხარისხი	20	
t Stat /T-სტატისტიკის დაკვირვებული მნიშვნელობა	-1.1266	
$P\{T \le t\}$ **one-tail** / ***p***-მნიშვნელობა ცალმხრივი კრიტიკუ-ლი არისათვის	0.1366	
t **Critical one-tail** / კრიტიკული წერტილი ცალმხრივი კრიტიკული არისათვის	1.7247	
$P\{T \le t\}$ **two-tail** / ***p***-მნიშვნელობა ორმხრივი კრიტიკული არისათვის	0.2733	
t **Critical two-tail** / კრიტიკული წერტილი ორმხრივი კრ-იტიკული არისათვის	2.086	

დასკვნა:

- **ცალმხრივი $H_1: a_1 - a_2 < 0$ ალტერნატივისათვის p-მნიშვნე-ლობა = 0.1366 > $\alpha = 0.05$. შესაბამისად, 0.05 მნიშვნელოვნების დონით, ძირითადი ჰიპოთეზის უარვყოფის საფუძველი არ გა-გვაჩნია;**

- ❖ **ორმხრივი $H_1: a_1 - a_2 \neq 0$ ალტერნატივისათვის p-მნიშვნელობა = 0.2733 > α = 0.05. შესაბამისად, 0.05 მნიშვნელოვნების დონით, ძირითადი ჰიპოთეზის უარვყოფის საფუძველი არ გვაქვს.**

შენიშვნა. სტატისტიკური ფუნქციით **TTEST(array1,array2,1,2)** – გამოითვლება სტიუდენტის განაწილების p-მნიშვნელობა ცალმხრივი ალტერნატივისათვის, ხოლო **TTEST(array1,array2,2,2)**-ით კი p-მნიშვნელობა ორმხრივი ალტერნატივისათვის.

ამოხსნა 2. ვიგულისხმოთ, რომ ორივე პოპულაცია ნორმალურადაა განაწილებული და უცნობი დისპერსიები არაა ტოლი. ვამოწმებთ ძირითად ჰიპოთეზას H_0:a_1–a_2=0 ცალმხრივი H_1:a_1–a_2<0 და ორმხრივი $H_1: a_1 - a_2 \neq 0$ ალტერნატივების წინააღმდეგ. ქვეპროგრამა **t-Test: Two Assuming Unequal Variances** გვაძლევს:

	A	***B***
Mean / საშუალო	89.658	94.33
Variance / დისპერსია	88.366	100.438
Observations / დაკვირვებები	12	10
Hypothesized Mean Difference / ძირითადი ჰიპოთეზით განსაზღრული საშუალოთა სხვაობა	0	
df / თავისუფლების ხარისხი	19	
t Stat /T-სტატისტიკის დაკვირვებული მნიშვნელობა	-1.1197	
***P*{*T* ≤ *t*} one-tail** / *p*-მნიშვნელობა ცალმხრივი კრიტიკული არისათვის	0.1384	
t **Critical one-tail** / კრიტიკული წერტილი ცალმხრივი კრიტიკული არისათვის	1.7291	
***P*{*T* ≤ *t*} two-tail** / *p*-მნიშვნელობა ორმხრივი კრიტიკული არისათვის	0.2768	
t **Critical two-tail** / კრიტიკული წერტილი ორმხრივი კრიტიკული არისათვის	2.093	

დასკვნა:

- ❖ **ცალმხრივი $H_1: a_1 - a_2 < 0$ ალტერნატივისათვის p-მნიშვნელობა= 0.1384 > a = 0.05. შესაბამისად, 0.05 მნიშვნელოვნების დონით, ძირითადი ჰიპოთეზის უარვყოფის საფუძველი არ გაგვაჩნია;**

- **ორმხრივი $H_1: a_1 - a_2 \neq 0$ ალტერნატივისათვის p-მნიშვნელობა=0.2768 > a = 0.05. შესაბამისად, 0.05 მნიშვნელოვნების დონით, ძირითადი ჰიპოთეზის უარყოფის საფუძველი არ გვაქვს.**

შენიშვნა 1. როგორც ვხედავთ, ამოხსნა 1-საგან განსხვავებით აქ ამოიბეჭდა სხვა თავისუფლების ხარისხი და T სტატისტიკის სხვა დაკვირვებული მნიშვნელობა. ეს აიხსნება იმით, რომ უკანასკნელ შემთხვევაში სარგებლობენ ე. წ. **სატერტვაიტის** მეთოდით.

შენიშვნა 2. სტატისტიკური ფუნქციით **TTEST(array1,array2,1,3)** – გამოითვლება სტიუდენტის განაწილების p-მნიშვნელობა ცალმხრივი ალტერნატივისათვის, ხოლო **TTEST(array1,array2,2,3)**-ით კი p-მნიშვნელობა ორმხრივი ალტერნატივისათვის.

t-Test-ი დაწყვილებული მონაცემებისათვის. ორჰექტარიანი 10 ნაკვეთი დაიყო ორ-ორ ტოლ ნაწილად, ერთ ნაწილზე გამოიყენეს I ტიპის სასუქი, ხოლო მეორეზე კი II ტიპის სასუქი. მოსავლიანობამ შესაბამისად შეადგინა:

	მოსავლიანობა									
ნაკვეთის №	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
I სასუქი	**1.2**	**1.4**	**1.6**	**1.1**	**1.4**	**1.1**	**0.7**	**1.5**	**1.3**	**1**
II სასუქი	**1.4**	**1.2**	**1.5**	**1.1**	**1.6**	**1.2**	**1**	**1.4**	**1.3**	**1.5**

0.05 მნიშვნელოვნების დონით შეამოწმეთ ზრდის თუ არა II ტიპის სასუქის გამოყენება მოსავლიანობას I ტიპის სასუქის გამოყენებასთან შედარებით.

ამოხსნა. ვამოწმებთ ძირითად ჰიპოთეზას $H_0: a_D = a_1 - a_2 = 0$ ცალმხრივი $H_1: a_D = a_1 - a_2 < 0$ და ორმხრივი $H_1: a_D = a_1 - a_2 \neq 0$ ალტერნატივების წინააღმდეგ. ქვეპროგრამა **t-Test: Paired Two Sample for Means** გვაძლევს:

	Variable 1	*Variable 2*
Mean / საშუალო	1.23	1.32
Variance / დისპერსია	0.0712	0.0373
Observations / დაკვირვებები	10	10

Pearson Correlation / პირსონის კორელაცია	0.612
Hypothesized Mean Difference / ძირითადი ჰიპოთეზით განსაზღრული საშუალოთა სხვაობა	0
df / თავისუფლების ხარისხი	9
t Stat /T-სტატისტიკის დაკვირვებული მნიშვნელობა	-1.3351
$P\{T\le t\}$ **one-tail** / p-მნიშვნელობა ცალმხრივი კრიტიკული არისათვის	0.1073
t **Critical one-tail** / კრიტიკული წერტილი ცალმხრივი კრიტიკული არისათვის	1.8331
$P\{T\le t\}$ **two-tail** / p-მნიშვნელობა ორმხრივი კრიტიკული არისათვის	0.2146
t **Critical two-tail** / კრიტიკული წერტილი ორმხრივი კრიტიკული არისათვის	2.2622

დასკვნა:

- **ცალმხრივი $H_1: a_D = a_1 - a_2 < 0$ ალტერნატივისათვის p-მნიშვნელობა = 0.1073 > a = 0.05. შესაბამისად, 0.05 მნიშვნელოვნების დონით ძირითადი ჰიპოთეზის უარვყოფის საფუძველი არ გაგვაჩნია;**
- **ორმხრივი $H_1: a_D = a_1 - a_2 \neq 0$ ალტერნატივისათვის p-მნიშვნელობა = 0.2146 > a = 0.05. შესაბამისად, 0.05 მნიშვნელოვნების დონით ძირითადი ჰიპოთეზის უარვყოფის საფუძველი არ გვაქვს.**

შენიშვნა. სტატისტიკური ფუნქციით **TTEST(array1,array2,1,1)** – გამოითვლება სტიუდენტის განაწილების p-მნიშვნელობა ცალმხრივი ალტერნატივისათვის, ხოლო **TTEST(array1,array2, 2,1)**-ით კი p-მნიშვნელობა ორმხრივი ალტერნატივისათვის.

F-Test-ის გამოყენების მაგალითი. 0.05 მნიშვნელოვნების დონით შეამოწმეთ ჰიპოთეზა ორი დამოუკიდებელი ნორმალური პოპულაციის დისპერსიების ტოლობის შესახებ, თუ მათგან მიღებული შერჩევებია:

I შერჩევა	12.1	4.1	4.6	15.2	-0.9	3.2	2.5	10.1	7.2	13.5
II შერჩევა	7	11.6	11.6	9.8	7.3	9	6.3	13.7		

ამოხსნა. მოწმდება ძირითადი ჰიპოთეზა $H_0:\sigma_1^2=\sigma_2^2$ ცალმ-ხრივი $H_1:\sigma_1^2>\sigma_2^2$ ალტერნატივის წინააღმდეგ. ქვეპროგრამა **F-Test: Two Sample for Variances** გვაძლევს:

	I	II
Mean / საშუალო	7.16	9.54
Variance / დისპერსია	28.463	6.874
Observations / დაკვირვებები	10	8
df / თავისუფლების ხარისხი	9	7
F Stat /F--სტატისტიკის დაკვირვებული მნიშვნელობა	4.141	
***P*{*F* ≤ *t*} one-tail** / *p*-მნიშვნელობა ცალმხრივი კრიტიკული არისათვის	0.037	
***F* Critical one-tail** / კრიტიკული წერტილი ცალმხრივი კრ-იტიკული არისათვის	3.677	

დასკვნა: ცალმხრივი $H_1:\sigma_1^2>\sigma_2^2$ ალტერნატივისათვის *p*- მნიშვ-ნელობა=0.037<*a* = 0.05. შესაბამისად, 0.05 მნიშვნელოვნების დონით ძირითად ჰიპოთეზას უარვყოფთ.

შენიშვნა. სტატისტიკური ფუნქციით **FTEST(array1,array2)** – გამოითვლება ფიშერის განაწილების *p*-მნიშვნელობა ცალმხრივი ალტერნატივისათვის.

დისპერსიული ანალიზი

ქვეპროგრამა **Anova: Single Factor** სარგებლობისას გამოი-ყენება შემდეგი აღნიშვნები:

აღნიშვნა	შინაარსი
Input range	შემავალი მონაცემები
Grouped by	დაჯგუფებული
Labels	მონიშვნა

მაგალითი. მენეჯერის მიზანია 0.05 მნიშვნელოვნების დო-ნით შეამოწმოს შემთხვევით შერჩეული 3 ჩარხის ერთგვაროვ-ნება მათზე გამოჩარხული ერთი და იმავე დეტალების ზომების მონაცემების ცვალებადობის მიხედვით:

I ჩარხი	6.783	6.781	6.776	6.788	6.785	6.789
II ჩარხი	6.775	6.77	6.772	6.771	6.779	6.773
III ჩარხი	6.778	6.776	6.769	6.772	6.78	6.777

Anova: Single Factor ქვეპროგრამის გამოყენება გვაძლევს:
SUMMARX

Groups	**Caunt**	**Sum**	**Average**	**Variance**
Colomn 1	**6**	**502**	**83.67**	**23.07**
Colomn 2	**6**	**440**	**73.33**	**10.67**
Colomn 3	**6**	**452**	**75.33**	**16.67**

ANOVA

Source of Variation SS df MS F P-value F crit
ვარიაციის წყარო

Between Groups 360.44 2 180.22 10.73 0.00 3.68

Within Groups 252 15 16.8

Total 612.44 17

დასკვნა: ვინაიდან $f = 10.73 > F_{2,15,0.05} = 3.68$ ამიტომ უარვყოფთ ჰიპოთეზას ჩარხების ერთგვაროვნების შესახებ – 0.05 მნიშვნელოვნების დონით ჩარხები არაერთგვაროვანია.

კორელაცია

ფუნქცია **COVAR(array1,array2)** – გამოიყენება კოვარიაციის გამოსათვლელად.

ფუნქცია **CORREL(array1,array2)** – გამოიყენება კორელაციის კოეფიციენტის გამოსათვლელად (**CORREL**-ის იდენტურია **PEARSON** ფუნქცია)

Covariance ქვეპროგრამით გამოითვლება ერთნაირი მოცულობის მონაცემთა რამდენიმე მასივის კოვარიაციული მატრიცა.

Correlation ქვეპროგრამით გამოითვლება ერთნაირი მოცულობის მონაცემთა რამდენიმე მასივის კორელაციური მატრიცა.

რეგრესია

წრფივი რეგრესიული ანალიზის ჩასატარებლად გამოიყენება ფუნქცია **Tools/Data Analysis/Regression**. თუ აღნიშნული ფუნქცია მიუწვდომელია, მაშინ უნდა შესრულდეს შემდეგი მოქმედებები: **Tools/ add-ins**-ში უნდა გააქტიურდეს **Analysis ToolPak, Analysis ToolPak-VBA.**

მარტივ წრფივ რეგრესიულ მოდელთან დაკავშირებით გამოიყენება შემდეგი აღნიშვნები:

აღნიშვნა	შინაარსი
Known_y's	ცნობილი დამოკიდებული ცვლადის მნიშვნელობები
Known_x's	ცნობილი დამოუკიდებული ცვლადის მნიშვნელობები
x	დამოუკიდებული ცვლადის მნიშვნელობა
Const	თუ **Const = 0**, მაშინ რეგრესიის წრფე გადის სათავეზე

ფუნქცია **DEVSQ(number1,2umber2,…)** – ითვლის დაკვირვებული მონაცემების საშუალოდან გადახრების კვადრატების ჯამს.

ფუნქცია **INTERCEPT(known_y's,known_x's)** – ითვლის B_0 კოეფიციენტის b_0 შეფასებას.

ფუნქცია **LINEST(known_y's,known_x's,const,stats)** – ითვლის B_1 კოეფიციენტის b_1 შეფასებას (თუ Const = 0, მაშინ გამოითვლება B_1 კოეფიციენტის შეფასება $Y = B_1 x + \varepsilon$ მოდელისათვის, ხოლო თუ Const = 1, მაშინ გამოითვლება B_1-ის შეფასება $Y = B_0 + B_1 x + \varepsilon$ მოდელისათვის). **LINEST**-ის იდენტურია **SLOPE** ფუნქცია.

ფუნქცია **FORECAST(x,known_y's,known_x's)** – ითვლის საპროგნოზო მნიშვნელობას.

ფუნქცია **RSQ(known_y's,known_x's)** – ითვლის პირსონის კორელაციის კოეფიციენტის კვადრატს (ანუ ე. წ. დეტერმინაციის კოეფიციენტს).

ფუნქცია **TREND(known_y's,known_x's,new_x's,const)** – ითვლის საპროგნოზო მნიშვნელობას (თუ Const = 0, მაშინ გამოითვ-

ლება B_1 კოეფიციენტის შეფასება $Y = b_1 x$ მოდელისათვის, ხოლო თუ Const = 1, მაშინ გამოითვლება B_1-ის შეფასება $Y = b_0 + b_1 x$ მოდელისათვის).

ფუნქცია **GROWTH(known_y's,known_x's,new_x's,const)** – ითვლის საპროგნოზო მნიშვნელობას (თუ Const = 0, მაშინ გამოითვლება Y-ის შეფასება $Y = (\beta_1)^x$ მოდელისათვის, ხოლო თუ Const = 1, მაშინ კი გამოითვლება Y-ის შეფასება $Y = \beta_0 \cdot (\beta_1)^x$ მოდელისათვის).

მაგალითი XVIII. 1-ის მონაცემების ანალიზი. Regression ფუნქციის გამოყენება გვაძლევს:

SUMMARY OUTPUT

Regression Statistics	
Multiple	**R 0.8047**
R Square	**0.6476**
Adjusted R Square	**0.6204**
Standard Error	**12.9965**
Observations	**15**

შესაბამისად, ამონაბეჭდი გვაძლევს კორელაციისა (**Multiple R**) და დეტერმინაციის (**R Square**) კოეფიციენტების, შეთანხმებული დეტერმინაციის კოეფიციენტის (**Adjusted R Square**) და სტანდარტული შეცდომის (**Standard Error**) რიცხვით მნიშვნელობებს. დეტერმინაციის კოეფიციენტი 0.6476 გვიჩვენებს, რომ მარტივი რეგრესიით აიხსნება სახლის გასაყიდი ფასის ვარიაციის მხოლოდ 64.76%, დანარჩენი 35.24% კი აუხსნელი რჩება. საჭიროა შეირჩეს სხვა მოდელი.

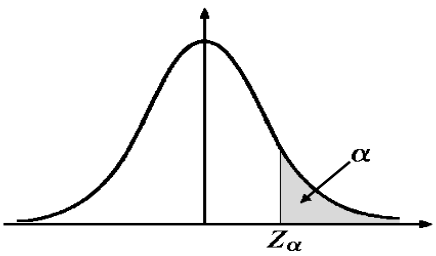
დანართი 2

(სტატისტიკური ცხრილები)

პუასონის განაწილების ცხრილები ($P(k)=\frac{\lambda^k}{k!}e^{-\lambda}$)

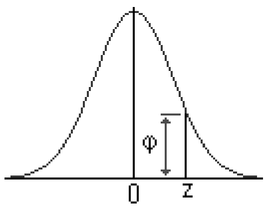
	λ=1.0	λ=1.5	λ=2.0	λ=2.5	λ=3.0	λ=3.5	λ=4.0	λ=4.5	λ=5.0
p(0)	**0.3679**	0.2231	0.1353	0.0821	0.0498	0.0302	0.0183	0.0111	0.0067
p(1)	**0.3679**	**0.3347**	**0.2707**	0.2052	0.1494	0.1057	0.0733	0.0500	0.0337
p(2)	0.1839	0.2510	**0.2707**	**0.2565**	**0.2240**	0.1850	0.1465	0.1125	0.0842
p(3)	0.0613	0.1255	0.1804	0.2138	**0.2240**	**0.2158**	**0.1954**	0.1687	0.1404
p(4)	0.0153	0.0471	0.0902	0.1336	0.1680	0.1888	**0.1954**	**0.1898**	**0.1755**
p(5)	0.0031	0.0141	0.0361	0.0668	0.1008	0.1322	0.1563	0.1708	**0.1755**
p(6)	0.0005	0.0035	0.0120	0.0278	0.0504	0.0771	0.1042	0.1281	0.1462
p(7)	0.0001	0.0008	0.0034	0.0099	0.0216	0.0385	0.0595	0.0824	0.1044
p(8)		0.0001	0.0009	0.0031	0.0081	0.0169	0.0298	0.0463	0.0653
p(9)			0.0002	0.0009	0.0027	0.0066	0.0132	0.0232	0.0363
p(10)				0.0002	0.0008	0.0023	0.0053	0.0104	0.0181
p(11)					0.0002	0.0007	0.0019	0.0043	0.0082
p(12)					0.0001	0.0002	0.0006	0.0016	0.0034
p(13)						0.0001	0.0002	0.0006	0.0013
p(14)							0.0001	0.0002	0.0005
p(15)								0.0001	0.0002

სტანდარტული ნორმალური განაწილების ზედა α -კრიტიკული წერტილები (z_α)



α	0.1	0.05	0.025	0.125	0.01	0.005	0.0025	0.001
z_α	1.28	1.64	1.96	2.24	2.33	2.57	2.81	3.08

$N(0.1)$ -ის სიმკვრივის ($\varphi(z) = \dfrac{1}{2\pi} e^{-z^2/2}$) მნიშვნელობები

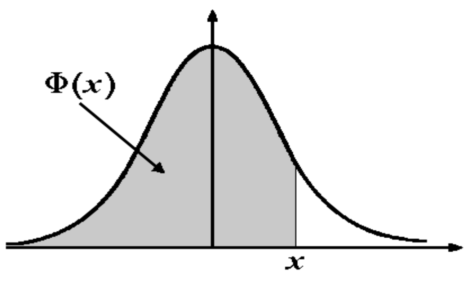


Z	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0.0	.398942	.398922	.398862	.398763	.398623	.398444	.398225	.397966	.397668	.397330
0.1	.396953	.396536	.396080	.395585	.395052	.394479	.393868	.393219	.392531	.391806
0.2	.391043	.390242	.389404	.388529	.387617	.386668	.385683	.384663	.383606	.382515
0.3	.381388	.380226	.379031	.377801	.376537	.375240	.373911	.372548	.371154	.369728
0.4	.368270	.366782	.365263	.363714	.362135	.360527	.358890	.357225	.355533	.353812
0.5	.352065	.350292	.348493	.346668	.344818	.342944	.341046	.339124	.337180	.335213
0.6	.333225	.331215	.329184	.327133	.325062	.322972	.320864	.318737	.316593	.314432
0.7	.312254	.310060	.307851	.305627	.303389	.301137	.298872	.296595	.294305	.292004
0.8	.289692	.287369	.285036	.282694	.280344	.277985	.275618	.273244	.270864	.268477
0.9	.266085	.263688	.261286	.258881	.256471	.254059	.251644	.249228	.246809	.244390
Z	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**

1.0	.241971	.239551	.237132	.234714	.232297	.229882	.227470	.225060	.222653	.220251
1.1	.217852	.215458	.213069	.210686	.208308	.205936	.203571	.201214	.198863	.196520
1.2	.194186	.191860	.189543	.187235	.184937	.182649	.180371	.178104	.175847	.173602
1.3	.171369	.169147	.166937	.164740	.162555	.160383	.158225	.156080	.153948	.151831
1.4	.149727	.147639	.145564	.143505	.141460	.139431	.137417	.135418	.133435	.131468
1.5	.129518	.127583	.125665	.123763	.121878	.120009	.118157	.116323	.114505	.112704
1.6	.110921	.109155	.107406	.105675	.103961	.102265	.100586	.098925	.097282	.095657
1.7	.094049	.092459	.090887	.089333	.087796	.086277	.084776	.083293	.081828	.080380
1.8	.078950	.077538	.076143	.074766	.073407	.072065	.070740	.069433	.068144	.066871
1.9	.065616	.064378	.063157	.061952	.060765	.059595	.058441	.057304	.056183	.055079
Z	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
2.0	.053991	.052919	.051864	.050824	.049800	.048792	.047800	.046823	.045861	.044915
2.1	.043984	.043067	.042166	.041280	.040408	.039550	.038707	.037878	.037063	.036262
2.2	.035475	.034701	.033941	.033194	.032460	.031740	.031032	.030337	.029655	.028985
2.3	.028327	.027682	.027048	.026426	.025817	.025218	.024631	.024056	.023491	.022937
2.4	.022395	.021862	.021341	.020829	.020328	.019837	.019356	.018885	.018423	.017971
2.5	.017528	.017095	.016670	.016254	.015848	.015449	.015060	.014678	.014305	.013940
2.6	.013583	.013234	.012892	.012558	.012232	.011912	.011600	.011295	.010997	.010706
2.7	.010421	.010143	3z98712	3z96058	3z93466	3z90936	3z88465	3z86052	3z83697	3z81398
2.8	3z79155	3z76965	3z74829	3z72744	3z70711	3z68728	3z66793	3z64907	3z63067	3z61274
2.9	3z59525	3z57821	3z56160	3z54541	3z52963	3z51426	3z49929	3z48470	3z47050	3z45666

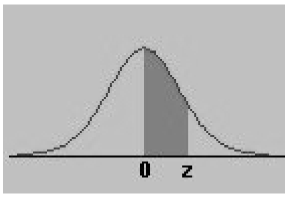
$N(0.1)$ -ის განაწილების ფუნქციის ($\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt$)

მნიშვნელობები



x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.00	0.500	**0.33**	0.629	**0.66**	0.745	**0.99**	0.838	**1.32**	0.906	**1.65**	0.950
0.01	0.503	**0.34**	0.633	**0.67**	0.748	**1.00**	0.841	**1.33**	0.908	**1.66**	0.951
0.02	0.507	**0.35**	0.636	**0.68**	0.751	**1.01**	0.843	**1.34**	0.909	**1.67**	0.952
0.03	0.511	**0.36**	0.640	**0.69**	0.754	**1.02**	0.846	**1.35**	0.911	**1.68**	0.953
0.04	0.515	**0.37**	0.644	**0.70**	0.758	**1.03**	0.848	**1.36**	0.913	**1.69**	0.954

0.05	0.519	**0.38**	0.648	**0.71**	0.761	**1.04**	0.850	**1.37**	0.914	**1.70**	0.955
0.06	0.523	**0.39**	0.651	**0.72**	0.764	**1.05**	0.853	**1.38**	0.916	**1.71**	0.956
0.07	0.527	**0.40**	0.655	**0.73**	0.767	**1.06**	0.855	**1.39**	0.917	**1.72**	0.957
0.08	0.531	**0.41**	0.659	**0.74**	0.770	**1.07**	0.857	**1.40**	0.919	**1.73**	0.958
0.09	0.535	**0.42**	0.662	**0.75**	0.773	**1.08**	0.859	**1.41**	0.920	**1.74**	0.959
0.10	0.539	**0.43**	0.666	**0.76**	0.776	**1.09**	0.862	**1.42**	0.922	**1.75**	0.959
0.11	0.543	**0.44**	0.670	**0.77**	0.779	**1.10**	0.864	**1.43**	0.923	**1.76**	0.960
0.12	0.547	**0.45**	0.673	**0.78**	0.782	**1.11**	0.866	**1.44**	0.925	**1.77**	0.961
0.13	0.551	**0.46**	0.677	**0.79**	0.785	**1.12**	0.868	**1.45**	0.926	**1.78**	0.962
0.14	0.555	**0.47**	0.680	**0.80**	0.788	**1.13**	0.870	**1.46**	0.927	**1.79**	0.963
0.15	0.559	**0.48**	0.684	**0.81**	0.791	**1.14**	0.872	**1.47**	0.929	**1.80**	0.964
0.16	0.563	**0.49**	0.687	**0.82**	0.793	**1.15**	0.874	**1.48**	0.930	**1.81**	0.964
0.17	0.567	**0.50**	0.691	**0.83**	0.796	**1.16**	0.876	**1.49**	0.931	**1.82**	0.965
0.18	0.571	**0.51**	0.694	**0.84**	0.799	**1.17**	0.879	**1.50**	0.933	**1.83**	0.966
0.19	0.575	**0.52**	0.698	**0.85**	0.802	**1.18**	0.881	**1.51**	0.934	**1.84**	0.967
0.20	0.579	**0.53**	0.701	**0.86**	0.805	**1.19**	0.882	**1.52**	0.935	**1.85**	0.967
0.21	0.583	**0.54**	0.705	**0.87**	0.807	**1.20**	0.884	**1.53**	0.936	**1.86**	0.968
0.22	0.587	**0.55**	0.708	**0.88**	0.810	**1.21**	0.886	**1.54**	0.938	**1.87**	0.969
0.23	0.590	**0.56**	0.712	**0.89**	0.813	**1.22**	0.888	**1.55**	0.939	**1.88**	0.969
0.24	0.594	**0.57**	0.715	**0.90**	0.815	**1.23**	0.890	**1.56**	0.940	**1.89**	0.970
0.25	0.598	**0.58**	0.719	**0.91**	0.818	**1.24**	0.892	**1.57**	0.941	**1.90**	0.971
0.26	0.602	**0.59**	0.722	**0.92**	0.821	**1.25**	0.894	**1.58**	0.942	**1.91**	0.971
0.27	0.606	**0.60**	0.725	**0.93**	0.823	**1.26**	0.896	**1.59**	0.944	**1.92**	0.972
0.28	0.610	**0.61**	0.729	**0.94**	0.826	**1.27**	0.897	**1.60**	0.945	**1.93**	0.973
0.29	0.614	**0.62**	0.732	**0.95**	0.828	**1.28**	0.899	**1.61**	0.946	**1.94**	0.973
0.30	0.617	**0.63**	0.735	**0.96**	0.831	**1.29**	0.901	**1.62**	0.947	**1.95**	0.974
0.31	0.621	**0.64**	0.738	**0.97**	0.833	**1.30**	0.903	**1.63**	0.948	**1.96**	0.975
0.32	0.625	**0.65**	0.742	**0.98**	0.836	**1.31**	0.904	**1.64**	0.949	**1.97**	0.975

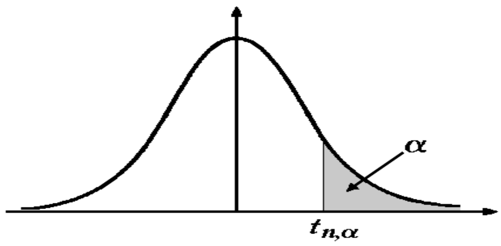


$$\Phi_0(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{t^2}{2}} dt$$

ფუნქციის ცხრილები										
	0.00	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981

2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

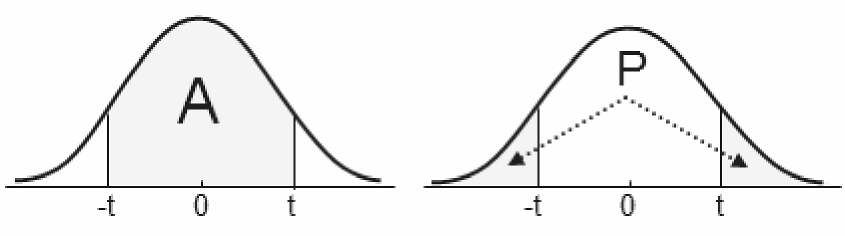
t (სტიუდენტის) განაწილების ზედა α -კრიტიკული წერტილები ($t_{n,\alpha}$)



	α						
n	**0.1**	**0.05**	**0.025**	**0.01**	**0.005**	**0.0025**	**0.001**
1	3.078	6.314	12.706	31.821	63.656	127.321	318.289
2	1.886	2.920	4.303	6.965	9.925	14.089	22.328
3	1.638	2.353	3.182	4.541	5.841	7.453	10.214
4	1.533	2.132	2.776	3.747	4.604	5.598	7.173
5	1.476	2.015	2.571	3.365	4.032	4.773	5.894
6	1.440	1.943	2.447	3.143	3.707	4.317	5.208
7	1.415	1.895	2.365	2.998	3.499	4.029	4.785
8	1.397	1.860	2.306	2.896	3.355	3.833	4.501
9	1.383	1.833	2.262	2.821	3.250	3.690	4.297
10	1.372	1.812	2.228	2.764	3.169	3.581	4.144
11	1.363	1.796	2.201	2.718	3.106	3.497	4.025
12	1.356	1.782	2.179	2.681	3.055	3.428	3.930
13	1.350	1.771	2.160	2.650	3.012	3.372	3.852
14	1.345	1.761	2.145	2.624	2.977	3.326	3.787
15	1.341	1.753	2.131	2.602	2.947	3.286	3.733
16	1.337	1.746	2.120	2.583	2.921	3.252	3.686
17	1.333	1.740	2.110	2.567	2.898	3.222	3.646
18	1.330	1.734	2.101	2.552	2.878	3.197	3.610

19	1.328	1.729	2.093	2.539	2.861	3.174	3.579
20	1.325	1.725	2.086	2.528	2.845	3.153	3.552
21	1.323	1.721	2.080	2.518	2.831	3.135	3.527
22	1.321	1.717	2.074	2.508	2.819	3.119	3.505
23	1.319	1.714	2.069	2.500	2.807	3.104	3.485
24	1.318	1.711	2.064	2.492	2.797	3.091	3.467
25	1.316	1.708	2.060	2.485	2.787	3.078	3.450
26	1.315	1.706	2.056	2.479	2.779	3.067	3.435
27	1.314	1.703	2.052	2.473	2.771	3.057	3.421
28	1.313	1.701	2.048	2.467	2.763	3.047	3.408
29	1.311	1.699	2.045	2.462	2.756	3.038	3.396
30	1.310	1.697	2.042	2.457	2.750	3.030	3.385

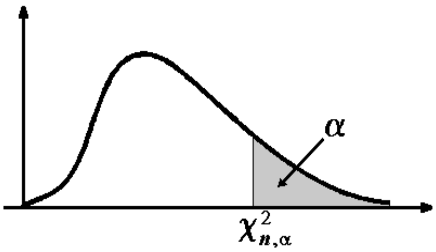
t განაწილების ზედა $\alpha/2$-კრიტიკული წერტილები $t_{n,\alpha/2}$ (ორკუდიანი)



n	**A**	**0.80**	**0.90**	**0.95**	**0.98**	**0.99**	**0.995**	**0.998**	**0.999**
	α	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**	**0.005**	**0.002**	**0.001**
1		3.078	6.314	12.706	31.820	63.657	127.321	318.309	636.619
2		1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.599
3		1.638	2.353	3.182	4.541	5.841	7.453	10.215	12.924

4		1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5		1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6		1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7		1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8		1.397	1.860	2.306	2.897	3.355	3.833	4.501	5.041
9		1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10		1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11		1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12		1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13		1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14		1.345	1.761	2.145	2.625	2.977	3.326	3.787	4.140
15		1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16		1.337	1.746	2.120	2.584	2.921	3.252	3.686	4.015
17		1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18		1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19		1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20		1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21		1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22		1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23		1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.768
24		1.318	1.711	2.064	2.492	2.797	3.090	3.467	3.745
25		1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26		1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27		1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28		1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29		1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30		1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
∞		1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

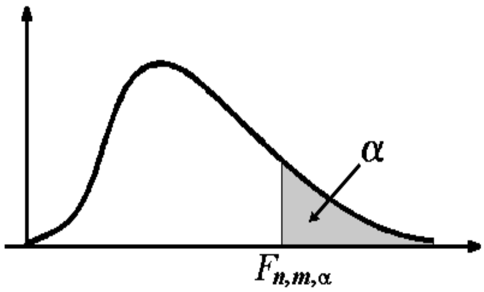
χ^2 (ხი კვადრატ) განაწილების ზედა α-კრიტიკული წერტილები ($\chi^2_{n,\alpha}$)



	α										
n	**0.995**	**0.975**	**0.20**	**0.10**	**0.05**	**0.025**	**0.02**	**0.01**	**0.005**	**0.002**	**0.001**
1	0.000039	0.00098	1.642	2.706	3.841	5.024	5.412	6.635	7.879	9.550	10.828
2	0.0100	0.0506	3.219	4.605	5.991	7.378	7.824	9.210	10.597	12.429	13.816
3	0.0717	0.216	4.642	6.251	7.815	9.348	9.837	11.345	12.838	14.796	16.266
4	0.207	0.484	5.989	7.779	9.488	11.143	11.668	13.277	14.860	16.924	18.467
5	0.412	0.831	7.289	9.236	11.070	12.833	13.388	15.086	16.750	18.907	20.515
6	0.676	1.237	8.558	10.645	12.592	14.449	15.033	16.812	18.548	20.791	22.458
7	0.989	1.690	9.803	12.017	14.067	16.013	16.622	18.475	20.278	22.601	24.322
8	1.344	2.180	11.030	13.362	15.507	17.535	18.168	20.090	21.955	24.352	26.124
9	1.735	2.700	12.242	14.684	16.919	19.023	19.679	21.666	23.589	26.056	27.877
10	2.156	3.247	13.442	15.987	18.307	20.483	21.161	23.209	25.188	27.722	29.588
11	2.603	3.816	14.631	17.275	19.675	21.920	22.618	24.725	26.757	29.354	31.264
12	3.074	4.404	15.812	18.549	21.026	23.337	24.054	26.217	28.300	30.957	32.909
13	3.565	5.009	16.985	19.812	22.362	24.736	25.472	27.688	29.819	32.535	34.528
14	4.075	5.629	18.151	21.064	23.685	26.119	26.873	29.141	31.319	34.091	36.123
15	4.601	6.262	19.311	22.307	24.996	27.488	28.259	30.578	32.801	35.628	37.697
16	5.142	6.908	20.465	23.542	26.296	28.845	29.633	32.000	34.267	37.146	39.252
17	5.697	7.564	21.615	24.769	27.587	30.191	30.995	33.409	35.718	38.648	40.790
18	6.265	8.231	22.760	25.989	28.869	31.526	32.346	34.805	37.156	40.136	42.312
19	6.844	8.907	23.900	27.204	30.144	32.852	33.687	36.191	38.582	41.610	43.820
20	7.434	9.591	25.038	28.412	31.410	34.170	35.020	37.566	39.997	43.072	45.315
21	8.034	10.283	26.171	29.615	32.671	35.479	36.343	38.932	41.401	44.522	46.797
22	8.643	10.982	27.301	30.813	33.924	36.781	37.659	40.289	42.796	45.962	48.268
23	9.260	11.689	28.429	32.007	35.172	38.076	38.968	41.638	44.181	47.391	49.728
24	9.886	12.401	29.553	33.196	36.415	39.364	40.270	42.980	45.559	48.812	51.179
25	10.520	13.120	30.675	34.382	37.652	40.646	41.566	44.314	46.928	50.223	52.620
26	11.160	13.844	31.795	35.563	38.885	41.923	42.856	45.642	48.290	51.627	54.052
27	11.808	14.573	32.912	36.741	40.113	43.195	44.140	46.963	49.645	53.023	55.476
28	12.461	15.308	34.027	37.916	41.337	44.461	45.419	48.278	50.993	54.411	56.892
29	13.121	16.047	35.139	39.087	42.557	45.722	46.693	49.588	52.336	55.792	58.301
30	13.787	16.791	36.250	40.256	43.773	46.979	47.962	50.892	53.672	57.167	59.703
31	14.458	17.539	37.359	41.422	44.985	48.232	49.226	52.191	55.003	58.536	61.098
32	15.134	18.291	38.466	42.585	46.194	49.480	50.487	53.486	56.328	59.899	62.487
33	15.815	19.047	39.572	43.745	47.400	50.725	51.743	54.776	57.648	61.256	63.870

34	16.501	19.806	40.676	44.903	48.602	51.966	52.995	56.061	58.964	62.608	65.247
35	17.192	20.569	41.778	46.059	49.802	53.203	54.244	57.342	60.275	63.955	66.619

$F(n,m)$ (ფიშერის) განაწილების ზედა α -კრიტიკული წერტილები ($F_{n,m,\alpha}$)



	$n \quad \alpha = 0.1$								
m	**1**	**2**	**3**	**4**	**5**	**7**	**10**	**15**	**20**
1	39.864	49.500	53.593	55.833	57.240	58.906	60.195	61.220	61.740
2	8.5264	8.9999	9.1618	9.2434	9.2926	9.3491	9.3915	9.4248	9.4413
3	5.5384	5.4624	5.3907	5.3426	5.3092	5.2661	5.2304	5.2003	5.1845
4	4.5448	4.3245	4.1909	4.1073	4.0505	3.9790	3.9198	3.8704	3.8443
5	4.0605	3.7798	3.6194	3.5202	3.4530	3.3679	3.2974	3.2379	3.2067
7	3.5895	3.2575	3.0740	2.9605	2.8833	2.7850	2.7025	2.6322	2.5947
10	3.2850	2.9244	2.7277	2.6054	2.5216	2.4139	2.3226	2.2434	2.2007
15	3.0731	2.6951	2.4898	2.3615	2.2729	2.1582	2.0593	1.9722	1.9243
20	2.9746	2.5893	2.3801	2.2490	2.1582	2.0397	1.9368	1.8450	1.7939
30	2.8808	2.4887	2.2761	2.1423	2.0493	1.9269	1.8195	1.7222	1.6674
60	2.7911	2.3932	2.1774	2.0409	1.9457	1.8194	1.7070	1.6034	1.5435

	$n \quad \alpha = 0.05$								
m	**1**	**2**	**3**	**4**	**5**	**7**	**10**	**15**	**20**
1	161.45	199.50	215.71	224.58	230.16	236.77	241.88	245.95	248.01
2	18.513	19.000	19.164	19.247	19.296	19.353	19.396	19.429	19.446

3	10.128	9.5522	9.2766	9.1172	9.0135	8.8867	8.7855	8.7028	8.6602
4	7.7086	6.9443	6.5915	6.3882	6.2560	6.0942	5.9644	5.8579	5.8026
5	6.6078	5.7862	5.4095	5.1922	5.0504	4.8759	4.7351	4.6187	4.5582
7	5.5914	4.7375	4.3469	4.1202	3.9715	3.7871	3.6366	3.5108	3.4445
10	4.9645	4.1028	3.7082	3.4780	3.3259	3.1354	2.9782	2.8450	2.7741
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7066	2.5437	2.4035	2.3275
20	4.3512	3.4928	3.0983	2.8660	2.7109	2.5140	2.3479	2.2032	2.1241
30	4.1709	3.3159	2.9223	2.6896	2.5336	2.3343	2.1646	2.0149	1.9317
60	4.0012	3.1505	2.7581	2.5252	2.3683	2.1666	1.9927	1.8365	1.7480

$F(n,m)$ (ფიშერის) განაწილების ზედა α -კრიტიკული წერტილები ($F_{n,m,\alpha}$)

	n $\alpha = 0.01$								
m	**1**	**2**	**3**	**4**	**5**	**7**	**10**	**15**	**20**
1	4052.2	4999.5	5403.4	5624.6	5763.6	5928.4	6055.8	6157.3	6208.7
2	98.503	99.000	99.166	99.249	99.299	99.356	99.399	99.433	99.449
3	34.116	30.817	29.457	28.710	28.237	27.672	27.229	26.872	26.690
4	21.198	18.000	16.694	15.977	15.522	14.976	14.546	14.198	14.020
5	16.258	13.274	12.060	11.392	10.967	10.455	10.051	9.7222	9.5526
7	12.246	9.5467	8.4513	7.8466	7.4605	6.9929	6.6201	6.3143	6.1554
10	10.044	7.5594	6.5523	5.9944	5.6363	5.2001	4.8492	4.5582	4.4055
15	8.6831	6.3588	5.4169	4.8932	4.5557	4.1416	3.8049	3.5223	3.3719
20	8.0960	5.8489	4.9382	4.4306	4.1027	3.6987	3.3682	3.0880	2.9377
30	7.5624	5.3903	4.5098	4.0179	3.6990	3.3046	2.9791	2.7002	2.5486
60	7.0771	4.9774	4.1259	3.6491	3.3388	2.9530	2.6318	2.3522	2.1978

	n $\alpha = 0.005$								
m	**1**	**2**	**3**	**4**	**5**	**7**	**10**	**15**	**20**
1	16211	19999	21615	22500	23056	23715	24224	24630	24836
2	198.50	199.00	199.17	199.25	199.30	199.36	199.40	199.43	199.45

3	55.552	49.799	47.467	46.195	45.392	44.434	43.686	43.085	42.777
4	31.333	26.284	24.259	23.155	22.456	21.622	20.967	20.438	20.167
5	22.785	18.314	16.530	15.556	14.940	14.200	13.618	13.146	12.903
7	16.235	12.404	10.882	10.050	9.5221	8.8853	8.3803	7.9677	7.7539
10	12.826	9.4270	8.0807	7.3428	6.8723	6.3025	5.8467	5.4706	5.2740
15	10.798	7.7007	6.4760	5.8029	5.3721	4.8473	4.4235	4.0697	3.8826
20	9.9439	6.9865	5.8176	5.1744	4.7616	4.2569	3.8470	3.5020	3.3178
30	9.1796	6.3547	5.2387	4.6233	4.2275	3.7416	3.3439	3.0058	2.8231
60	8.4946	5.7950	4.7290	4.1399	3.7599	3.2911	2.9042	2.5705	2.3872

	n $\alpha = 0.025$									
m	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	647.789	799.500	864.163	899.583	921.847	937.111	948.216	956.656	963.284	968.627
2	38.5063	39.0000	39.1655	39.2484	39.2982	39.3315	39.3552	39.3730	39.3869	39.3980
3	17.4434	16.0441	15.4392	15.1010	14.8848	14.7347	14.6244	14.5399	14.4731	14.4189
4	12.2179	10.6491	9.9792	9.6045	9.3645	9.1973	9.0741	8.9796	8.9047	8.8439
5	10.0070	8.4336	7.7636	7.3879	7.1464	6.9777	6.8531	6.7572	6.6811	6.6192
6	8.8131	7.2599	6.5988	6.2272	5.9876	5.8198	5.6955	5.5996	5.5234	5.4613
7	8.0727	6.5415	5.8898	5.5226	5.2852	5.1186	4.9949	4.8993	4.8232	4.7611
8	7.5709	6.0595	5.4160	5.0526	4.8173	4.6517	4.5286	4.4333	4.3572	4.2951
9	7.2093	5.7147	5.0781	4.7181	4.4844	4.3197	4.1970	4.1020	4.0260	3.9639
10	6.9367	5.4564	4.8256	4.4683	4.2361	4.0721	3.9498	3.8549	3.7790	3.7168
11	6.7241	5.2559	4.6300	4.2751	4.0440	3.8807	3.7586	3.6638	3.5879	3.5257
12	6.5538	5.0959	4.4742	4.1212	3.8911	3.7283	3.6065	3.5118	3.4358	3.3736
13	6.4143	4.9653	4.3472	3.9959	3.7667	3.6043	3.4827	3.3880	3.3120	3.2497
14	6.2979	4.8567	4.2417	3.8919	3.6634	3.5014	3.3799	3.2853	3.2093	3.1469
15	6.1995	4.7650	4.1528	3.8043	3.5764	3.4147	3.2934	3.1987	3.1227	3.0602

დანართი 3 (ამოცანების პასუხები)

სტატისტიკური დასკვნების თეორია

თავი I

1. 2.58; 2.33; 1.96; 1.65; 1.88. **2.** 32.03; 11.01; (28.1, 35.97); 2.01. **3.** 46.9709; 14.3582; (41.8329, 52.1089). **4.** (77, 87); (75, 89); II, ვინაიდან საიმედოობა უფრო დიდია. **5.** 196; 617.3; (52, 340). **6.** (11.9,13.3); ეს ძალიან მცირე ალბათობის მქონეა, რადგან 30 ბევრჯერ აღემატება 13.3-ს. **7.** (184, 188); (185, 187); II, ვინაიდან 100 > 40. **8.** (18.13, 18.87). **9.** (5.6, 6.2). **10.** 3222.4; 3480.1; (2341.5, 4130.3). **11.** (37, 39); (35, 41); ვინაიდან $8 > 4$. **12.** (59.5, 62.9). **13.** 106. **14.** 45. **15.** (57.4, 58.6). **16.** 25. **17.** 10. **18.** 5. **19.** 2.898; 2.624; 2.093; 2.074; 1.833. **20.** 563.2; 87.9; (500.4, 626); 62.8. **21.** 81.72; 11.58; (75.96, 87.48). **22.** (15, 17). **23.** (17, 19). **24.** 33.4; 28.7; (21.2, 45.6); მონაცემი 132 არაჩვეულებრივად დიდია („ამოვარდნილი“ – „outlier“ მონაცემია) **25.** (8, 9.2). **26.** (11990, 12410). **27.** 58.9; 5.1; (55.5, 62.3). **28.** (8.7, 9.9). **29.** (13, 23). **30.** (17.29, 19.77). **31.** (123, 129). **32.** (109, 121). **33.** (38.4, 44.8). **34.** (7.9, 165.9). **35.** (94, 98). **36.** (58197, 58241).

თავი II

1. (0.365, 0.415) ანუ (36.5%,41.5%). **2.** (0.197, 0.343). **3.** (0.557, 0.743). **4.** (0.109, 0.151). **5.** (0.797, 0.883). **6.** (0.086, 0.214). **7.** (0.153, 0.307); ეს შედეგი ეთანხმება ჯანდაცვის სამინისტროს მონაცემებს. **8.** (0.467, 0.683). **9.** (0.337, 0.543). **10.** (0.413, 0.487). **11.** (0.149, 0.351). **12.** 577. **13.** 3097; 4129. **14.** 225; 273. **15.** 99; 273. **16.** 1448; 1509. **17.** 95%. **18.** 96%.

თავი III

1. 6.262, 27.488; 0.711, 9.488; 8.643, 42.796; 15.308, 44.461; 5.892, 22.362. **2.** (1.7, 6.2); (1.3, 2.5). **3.** (30.9, 78.2); (5.6, 8.8). **4.** (1.4, 11.7); (1.2, 3.4). **5.** (0.4, 2.25); (0.63, 1.5). **6.** (3.5, 9.3); (1.9, 3). **7.** (1.2, 7.7); (1.1, 2.8). **8.** (259.343, 772.724); (16.104, 27.798). **9.** (4.1, 7.1).

10. (0.001, 0.008); (0.032, 0.089). **11.** (16.2, 19.8). **12.** (42.31, 45.09); 54 სთ გაცილებით დიდია, ვიდრე მოსალოდნელი ინტერვალი. **13.** (2.5, 2.7). **14.** (2494, 2760). **15.** (7.46, 7.54). **16.** (16.99, 19.47). **17.** (25, 31). **18.** 106. **19.** 28. **20.** (0.395, 0.445) ანუ (39.5%, 44.5%). **21.** (0.343, 0.457). **22.** (0.791, 0.909). **23.** 1418. **24.** 8487. **25.** (0.25, 0.51). **26.** (1.7, 4.7). **27.** (5.1, 18.3). **28.** (1.447, 2.491). **29.** (22.79, 24.11). **30.** (43.15, 46.45). **31.** (3954, 4346). **32.** (45.7, 51.5). **33.** (418, 458). **34.** (26, 36). **35.** 180. **36.** 25. **37.** (0.604, 0.810). **38.** (0.295, 0.425). **39.** (0.342, 0.547). **40.** 545. **41.** (7, 13). **42.** (30.9, 78.2); (5.6, 8.8). **43.** (1.8, 3.2).

თავი IV

1. ± 2.58; 1.65; -1.65; -1.28; ± 1.96; 1.75; -2.33; ± 1.65; 2.05; ± 2.33. **2.** $H_0: E\xi = 39$, $H_1: E\xi \neq 39$; $H_0: E\xi = 25000$, $H_1: E\xi \neq 25000$; $H_0: E\xi \leq 25$, $H_1: E\xi > 25$; $H_0: E\xi \geq 85$, $H_1: E\xi < 85$; $H_0: E\xi \geq 56$, $H_1: E\xi < 56$; $H_0: E\xi \leq 250$, $H_1: E\xi > 250$; $H_0: E\xi \leq 75$, $H_1: E\xi > 75$; $H_0: E\xi \geq 950$, $H_1: E\xi < 950$. **3.** $H_0: E\xi = 25000$, $H_1: E\xi \neq 25000$; $C.V. = \pm 1.96$; $T.V. \equiv z = -1.59$; H_0. **4.** $H_0: E\xi = 69.21$, $H_1: E\xi \neq 69.21$; $C.V. = \pm 1.96$; $T.V. \equiv z = -1.15$; H_0. **5.** $H_0: E\xi \geq 9.78$, $H_1: E\xi < 9.78$; $C.V. = -1.28$; $T.V. \equiv z = -0.54$; H_0. **6.** $H_0: E\xi \leq 24$, $H_1: E\xi > 24$; $C.V. = 1.65$; $T.V. \equiv z = 1.85$; H_1. **7.** $H_0: E\xi = 800$, $H_1: E\xi \neq 800$; $C.V. = \pm 2.33$; $T.V. \equiv z = 10.12$; H_1. **8.** $H_0: E\xi \geq 14$, $H_1: E\xi < 14$; $C.V. = -2.33$; $T.V. \equiv z = -4.89$; H_1. **9.** $H_0: E\xi = 24$, $H_1: E\xi \neq 24$; $C.V. = \pm 1.96$; $T.V. \equiv z = -9.73$; H_1. **10.** $H_0: E\xi = 70$, $H_1: E\xi \neq 70$; $C.V. = \pm 1.9$; $T.V. \equiv z = -2.59$; H_1. **11.** $H_0: E\xi = 1660$, $H_1: E\xi \neq 1660$; $C.V. = \pm 1.96$; $T.V. \equiv z = -25.06$; H_1. **12.** $H_0: E\xi = 36$, $H_1: E\xi \neq 36$; $C.V. = \pm 2.58$; $T.V. \equiv z = -3.54$; H_1. **13.** $H_0: E\xi = 60000$, $H_1: E\xi \neq 60000$; $C.V. = \pm 1.96$; $T.V. \equiv z = 1.78$; H_0. **14.** $H_0: E\xi \geq 240$, $H_1: E\xi < 240$; $C.V. = -2.33$; $T.V. \equiv z = -3.87$; H_1. **15.** არა; არა; არა; კი; კი. **16.** $H_0: E\xi \leq 27.5$, $H_1: E\xi > 27.5$; $T.V. \equiv z = 2.55$; P-მნიშვნელობა = 0.0054; კი.

17. $H_0: E\xi \ge 264$, $H_1: E\xi < 264$; $T.V. \equiv z = -2.53$; P -მნიშვნელობა = =0.0057; კი. 18. $H_0: E\xi \ge 40$, $H_1: E\xi < 40$; $T.V. \equiv z = -2.45$; P -მნიშვნელობა = 0.0071; კი. 19. $H_0: E\xi \le 84$, $H_1: E\xi > 84$; $T.V. \equiv z = 1.1$; P -მნიშვნელობა=0.1357; არა. 20. $H_0: E\xi = 800$, $H_1: E\xi \ne 800$; $T.V. \equiv z = -2.61$; P -მნიშვნელობა = 0.009; კი. 21. $H_0: E\xi = 6.32$, $H_1: E\xi \ne 6.32$; $T.V. \equiv z = 2.49$; P -მნიშვნელობა = 0.0128; H_1. 22. $H_0: E\xi = 30000$, $H_1: E\xi \ne 30000$; $T.V. \equiv z = 1.71$; P -მნიშვნელობა= = 0.0872; H_1; კი. 23. $H_0: E\xi = 60$, $H_1: E\xi \ne 60$; $T.V. \equiv z = -0.03$; P -მნიშვნელობა = 0.976; არა.

თავი V

1. 1.833; ± 1.740; -3.365; 2.306; ± 2.145; -2.819; ± 2.771; ± 2.583. 2. (0.01,0.025); (0.05, 0.1); (0.1, 0.25); (0.1, 0.2); P -მნიშვნელობა < <0.005); (0.1, 0.25); P -მნიშვნელობა = 0.05; P -მნიშვნელობა > 0.25. 3. $H_0: E\xi \ge 11.52$, $H_1: E\xi < 11.52$; $C.V. = -1.83$; $T.V. \equiv t = -9.97$; კი. 4. $H_0: E\xi \ge 2000$, $H_1: E\xi < 2000$; $C.V. = -3.747$; $T.V. \equiv t = -0.104$; არა. 5. $H_0: E\xi = 800$, $H_1: E\xi \ne 800$; $C.V. = \pm 2.262$; $T.V. \equiv t = 9.96$; კი. 6. $H_0: E\xi \ge 48$, $H_1: E\xi < 48$; $C.V. = -2.567$; $T.V. \equiv t = -5.94$; კი. 7. $H_0: E\xi \ge 700$, $H_1: E\xi < 700$; $C.V. = -2.262$; $T.V. \equiv t = -2.71$; კი. 8. $H_0: E\xi = 17.63$, $H_1: E\xi \ne 17.63$; $C.V. = \pm 2.145$; $T.V. \equiv t = 1.16$; H_0. 9. $H_0: E\xi = 750$, $H_1: E\xi \ne 750$; $C.V. = \pm 3.106$; $T.V. \equiv t = -3.67$; კი. 10. $H_0: E\xi \ge 6.62$, $H_1: E\xi < 6.62$; $n = 10$; $C.V. = -1.383$; $T.V. \equiv t = -2.62$; კი. 11. $H_0: E\xi \le 350$, $H_1: E\xi > 350$; $C.V. = 1.796$; $T.V. \equiv t = 1.732$; H_0. 12. $H_0: E\xi = 3$, $H_1: E\xi \ne 3$; $C.V. = \pm 2.776$; $T.V. \equiv t = 5.22$; H_1. 13. $H_0: E\xi = 37$, $H_1: E\xi \ne 37$; $C.V. = \pm 2.048$; $T.V. \equiv t = -1.88$; 0.05 < P -მნიშვნელობა; რადგანაც $P > \alpha$, ამიტომ H_0 არ უნდა უკუვაგდოთ. 14. $H_0: E\xi \le 30$, $H_1: E\xi > 30$; $T.V. \equiv t = 0.85$; 0.2< P -მნიშვნელობა < 0.5; რადგანაც $P > \alpha$, ამიტომ H_0 არ უნდა უკუ-

ვაგდოთ. **15.** $H_0: E\xi = 75$, $H_1: E\xi \neq 75$; $T.V. \equiv t = -2.83$; 0.01< P -მნიშვნელობა<0.02; რადგანაც $P > \alpha$, ამიტომ H_0 არ უნდა უკუვაგდოთ.

თავი VI

1. $H_0: D\xi \leq 225$, $H_1: D\xi > 225$, $C.V. = 27.587$; $H_0: D\xi \geq 225$, $H_1: D\xi < 225$, $C.V. = 14.042$; $H_0: D\xi = 225$, $H_1: D\xi \neq 225$, $C.V. = 5.629$, 26.119; $H_0: D\xi = 225$, $H_1: D\xi \neq 225$, $C.V. = 2.167$, 14.067; $H_0: D\xi \leq 225$, $H_1: D\xi > 225$, $C.V. = 32$; $H_0: D\xi \geq 225$, $H_1: D\xi < 225$, $C.V. = 8.907$; $H_0: D\xi = 225$, $H_1: D\xi \neq 225$, $C.V. = 3.074$, 28.299; $H_0: D\xi \geq 225$, $H_1: D\xi < 225$, $C.V. = 15.308$. **2.** (0.01, 0.025); (0.005, 0.01); (0.01, 0.025): P -მნიშვნელობა<0.005; (0.025, 0.05); (0.1, 0.2); (0.05, 0.1); P -მნიშვნელობა<0.01. **3.** $H_0: \sigma = 60$, $H_1: \sigma \neq 60$; $C.V. = 8.672$, 27.587; $T.V. \equiv \chi^2 = 19.707$; H_0. **4.** $H_0: D\xi \leq 6.2$, $H_1: D\xi > 6.2$; $C.V. = 33.409$; $T.V. \equiv \chi^2 = 17.823$; H_0. **5.** $H_0: D\xi \leq 25$, $H_1: D\xi > 25$; $C.V. = 27.204$; $T.V. \equiv \chi^2 = 27.36$; კი. **6.** $H_0: D\xi \leq 1.6$, $H_1: D\xi > 1.6$; $C.V. = 55.758$; $T.V. \equiv \chi^2 = 70.438$; კი. **7.** $H_0: \sigma \leq 1.2$, $H_1: \sigma > 1.2$; $T.V. \equiv \chi^2 = 31.5$; P -მნიშვნელობა < 0.005 < α; H_1. **8.** $H_0: \sigma \leq 0.03$, $H_1: \sigma > 0.03$; $T.V. \equiv \chi^2 = 14.381$; 0.025 < P -მნიშვნელობა < 0.05 = α; არა. **9.** $H_0: \sigma \leq 2$, $H_1: \sigma > 2$; $C.V. = 24.725$; $T.V. \equiv \chi^2 = 22.02$; არა.

თავი VII

1. $H_0: p = 0.4$, $H_1: p \neq 0.4$; $C.V. = \pm 2.58$; $T.V. \equiv z = -0.612$; არა.
2. $H_0: p \geq 0.23$, $H_1: p < 0.23$; $C.V. = -1.65$; $T.V. \equiv z = -0.83$; არა.
3. $H_0: p = 0.37$, $H_1: p \neq 0.37$; $C.V. = \pm 2.58$; $T.V. \equiv z = -2.9$; კი.
4. $H_0: p \geq 0.4$, $H_1: p < 0.4$; $C.V. = -1.28$; $T.V. \equiv z = -0.462$; არა.
5. $H_0: p \leq 0.32$, $H_1: p > 0.32$; $C.V. = +1.65$; $T.V. \equiv z = +1.29$; არა.

6. $H_0: p=0.63$, $H_1: p\neq 0.63$; $C.V.=\pm 1.96$; $T.V.\equiv z=-0.88$; H_0. 7. $H_0: p=0.17$, $H_1: p\neq 0.17$; $C.V.=\pm 1.96$; $T.V.\equiv z=1.76$; H_0. 8. $H_0: p\geq 0.15$, $H_1: p<0.15$; $C.V.=-1.65$; $T.V.\equiv z=-0.94$; H_0. 9. $H_0: p\leq 0.25$, $H_1: p>0.25$; $T.V.\equiv z=1.15$; P -მნიშვნელობა = 0.1251 > α; H_0. 10. $H_0: p\leq 0.3$, $H_1: p>0.3$; $T.V.\equiv z=1.85$; P -მნიშვნელობა=0.0322< α; H_1. 11. $H_0: p=0.1$, $H_1: p\neq 0.1$; $T.V.\equiv z=1.34$; P -მნიშვნელობა=0.1802> α; H_0. 12. არა. 13. არა.

თავი VIII

1. $H_0: E\xi=1800$, $H_1: E\xi\neq 1800$; $C.V.=\pm 1.96$; $T.V.\equiv z=0.47$; H_0; $1706.04<E\xi< \ <1953.96$; $1800\in(1706.04,1953.96)$. 2. $H_0: E\xi=42$, $H_1: E\xi\neq 42$; $C.V.=\pm 1.65$; $T.V.\equiv z=0.47$; H_1; $43.83<E\xi<52.17$; $42\notin(43.83,52.17)$. 3. $H_0: E\xi=86$, $H_1: E\xi\neq 86$; $C.V.=\pm 2.58$; $T.V.\equiv z=-1.29$; H_0; $80<E\xi<88$; $86\in(80,88)$. 4. $H_0: E\xi=47$, $H_1: E\xi\neq 47$; $C.V.=\pm 1.65$; $T.V.\equiv z=-2.26$; H_1; $38.35<E\xi<45.65$; $47\notin(38.35,45.65)$. 5. $H_0: E\xi=22$, $H_1: E\xi\neq 22$; $C.V.=\pm 2.58$; $T.V.\equiv z=-2.32$; H_0; $19.47<E\xi<22.13$; $22\in(19.47,22.13)$. 6. $H_0: E\xi=98$, $H_1: E\xi\neq 98$; $C.V.=\pm 1.96$; $T.V.\equiv z=-2.02$; H_1. 7. $H_0: E\xi=14$, $H_1: E\xi\neq 14$; $C.V.=\pm 1.96$; $T.V.\equiv z=4.22$; H_1. 8. $H_0: E\xi=16.3$, $H_1: E\xi\neq 16.3$; $T.V.\equiv z=11.3$; P -მნიშვნელობა<0.01; კი. 9. $H_0: E\xi=61.2$, $H_1: E\xi\neq 61.2$; $C.V.=\pm 2.831$; $T.V.\equiv t=-4.378$; კი. 10. $H_0: E\xi\leq 67$, $H_1: E\xi>67$; $C.V.=1.383$; $T.V.\equiv t=7.47$; არა. 11. $H_0: E\xi\leq 23.2$, $H_1: E\xi>23.2$; $T.V.\equiv t=-1.27$; 0.75< P -მნიშვნე-ლობა<0.9; არა. 12. $H_0: E\xi=6$, $H_1: E\xi\neq 6$; $C.V.=\pm 2.821$; $T.V.\equiv t=1.835$; არა. 13. $H_0: p\geq 0.3$, $H_1: p<0.3$; $C.V.=-1.65$; $T.V.\equiv z=-0.85$; H_0. 14. $H_0: p\geq 0.6$, $H_1: p<0.6$; $C.V.=-1.28$; $T.V.\equiv z=-1.22$; კი. 15. $H_0: p=0.8$, $H_1: p\neq 0.8$; $C.V.=\pm 2.33$; $T.V.\equiv z=-1.92$; H_0. 16. $H_0: p=0.65$, $H_1: p\neq 0.65$; $T.V.\equiv z=1.17$; P -მნიშვნელობა =

=0.242 > α; H_0. 17. $H_0: E\xi = 225$, $H_1: E\xi \neq 225$; $T.V. \equiv z = 2.36$; P-მნიშვნელობა = 0.0182; H_0. 18. $H_0: \sigma = 3.4$, $H_1: \sigma \neq 3.4$; $C.V. = 11.689$ და 38.076; $T.V. \equiv \chi^2 = 35.1$; არა. 19. $H_0: \sigma \geq 4.3$, $H_1: \sigma < 4.3$; $T.V. \equiv \chi^2 = 6.95$; 0.005< P-მნიშვნელობა<0.01< α; კი. 20. $H_0: \sigma = 95$, $H_1: \sigma \neq 95$; $C.V. = 6.408$ და 33.409; $T.V. \equiv \chi^2 = 15.0212$; არა.
21. $H_0: \sigma = 18$, $H_1: \sigma \neq 18$; $C.V. = 11.143$ და 0.484; $T.V. \equiv \chi^2 = 5.44$; H_0.
22. $H_0: E\xi = 28.6$, $H_1: E\xi \neq 28.6$; $C.V. = \pm 1.96$; $T.V. \equiv z = 2.14$; H_1.
23. $H_0: E\xi = 6500$, $H_1: E\xi \neq 6500$; $C.V. = \pm 1.96$; $T.V. \equiv z = 5.27$; H_1.
24. $H_0: E\xi = 21$, $H_1: E\xi \neq 21$; $C.V. = \pm 2.921$; $T.V. \equiv t = -2.06$; H_0.
25. $H_0: E\xi \geq 12.4$, $H_1: E\xi < 12.4$; $C.V. = -1.345$; $T.V. \equiv t = -2.324$; კი.

თავი IX

1. $H_0: a_1 = a_2$, $H_1: a_1 \neq a_2$; $C.V. = \pm 2.58$; $T.V. \equiv z = -0.856$; კი.
2. $H_0: a_1 - a_2 = 0$, $H_1: a_1 - a_2 \neq 0$; $C.V. = \pm 1.65$; $T.V. \equiv z = 2.27$; კი.
3. $H_0: a_1 \leq a_2$, $H_1: a_1 > a_2$; $C.V. = +1.65$; $T.V. \equiv z = 2.56$; კი.
4. $H_0: a_1 \leq a_2$, $H_1: a_1 > a_2$; $C.V. = +2.33$; $T.V. \equiv z = 1$; არა.
5. $H_0: a_1 \leq a_2$, $H_1: a_1 > a_2$; $C.V. = +2.05$; $T.V. \equiv z = 1.12$; H_0.
6. $H_0: a_1 = a_2$, $H_1: a_1 \neq a_2$; $C.V. = \pm 2.58$; $T.V. \equiv z = 3.44$; კი.
7. $H_0: a_1 \leq a_2$, $H_1: a_1 > a_2$; $C.V. = +1.65$; $T.V. \equiv z = -2.01$; არა.
8. $H_0: a_1 \leq a_2$, $H_1: a_1 > a_2$; $C.V. = +1.65$; $T.V. \equiv z = 3.3$; არა.
9. $H_0: a_1 \leq a_2$, $H_1: a_1 > a_2$; $C.V. = +2.33$; $T.V. \equiv z = 1.09$; არა.
10. $H_0: a_1 = a_2$, $H_1: a_1 \neq a_2$; $T.V. \equiv z = 10.36$; P-მნიშვნელობა < <0.002< α; კი. 11. $H_0: a_1 = a_2$, $H_1: a_1 \neq a_2$; $T.V. \equiv z = 1.01$; P-მნიშვნელობა = 0.3124; არა. 12. $H_0: a_1 = a_2$, $H_1: a_1 \neq a_2$; $T.V. \equiv z = -0.76$; P-მნიშვნელობა=0.4472; H_0. 13. $2.8 < a_1 - a_2 < 6$. 14. $-7.3 < a_1 - a_2 < -1.3$.
15. $0.3 < a_1 - a_2 < 0.5$.

თავი X

1. I) $H_0:\sigma_1^2=\sigma_2^2$, $H_1:\sigma_1^2\neq\sigma_2^2$; $C.V.=4.03$; $\overline{f}=1.93$; H_0; II) $H_0:a_1=a_2$, $H_1:a_1\neq a_2$; $C.V.=t_{18,0.025}=\pm2.101$; $t=-4.2$; კი – H_1; III) $-7762<a_1-a_2<-2434$. **2.** I) $H_0:\sigma_1^2=\sigma_2^2$, $H_1:\sigma_1^2\neq\sigma_2^2$; $C.V.=2.62$; $\overline{f}=1.44$; H_0; II) $H_0:a_1\leq a_2$, $H_1:a_1>a_2$; $C.V.=t_{34,0.05}=1.65$; $t=0.55$; არა – H_0. **3.** I) $H_0:\sigma_1^2=\sigma_2^2$, $H_1:\sigma_1^2\neq\sigma_2^2$; $C.V.=2.72$; $\overline{f}=5.06$; H_1; II) $H_0:a_1\geq a_2$, $H_1:a_1<a_2$; $C.V.=-t_{25,0.1}=-1.316$; $t=-1.24$; არა – H_0. **4.** I) $H_0:\sigma_1^2=\sigma_2^2$, $H_1:\sigma_1^2\neq\sigma_2^2$; $C.V.=3.88$; $\overline{f}=1.78$; H_0;
II) $H_0:a_1\leq a_2$, $H_1:a_1>a_2$; $C.V.=t_{36,0.01}=2.33$; $t=1.26$; არა – H_0.
5. I) $H_0:\sigma_1^2=\sigma_2^2$, $H_1:\sigma_1^2\neq\sigma_2^2$; $C.V.=14.94$; $\overline{f}=1.41$; H_0; II) $H_0:a_1\leq a_2$, $H_1:a_1>a_2$; $C.V.=t_{10,0.01}=2.764$; $t=1.45$; არა – H_0.
6. I) $H_0:\sigma_1^2=\sigma_2^2$, $H_1:\sigma_1^2\neq\sigma_2^2$; $C.V.=6.39$; $\overline{f}=1.139$; H_0; II) $H_0:a_1\leq a_2$, $H_1:a_1>a_2$; $C.V.=t_{8,0.1}=1.397$; $t=16.252$; კი – H_1.
7. I) $H_0:\sigma_1^2=\sigma_2^2$, $H_1:\sigma_1^2\neq\sigma_2^2$; $C.V.=4.19$; $\overline{f}=1.7$; H_0; II) $H_0:a_1=a_2$, $H_1:a_1\neq a_2$; $C.V.=t_{22,0.01}=\pm2.508$; $t=-2.97$; კი – H_1; III) $-11.1<a_1-a_2<-0.93$. **8.** I) $H_0:\sigma_1^2=\sigma_2^2$, $H_1:\sigma_1^2\neq\sigma_2^2$; $C.V.=4.65$; $\overline{f}=4.46$; H_0; II) $H_0:a_1\geq a_2$, $H_1:a_1<a_2$; $C.V.=t_{14,0.05}=-1.761$; $t=-4.2$; კი – H_1. **9.** I) $H_0:\sigma_1^2=\sigma_2^2$, $H_1:\sigma_1^2\neq\sigma_2^2$; $\overline{f}=3.67$; 0.02 < P-მნიშვნელობა < 0.05; H_1; II) $H_0:a_1\geq a_2$, $H_1:a_1<a_2$; $d.f.=11$; $t=-4.98$; P-მნიშვნელობა < 0.005 < α; კი – H_1. **10.** I) $H_0:\sigma_1^2=\sigma_2^2$, $H_1:\sigma_1^2\neq\sigma_2^2$; $\overline{f}=1$; P-მნიშვნელობა > 0.2 > $\alpha=0.0$; H_0; II) $H_0:a_1=a_2$, $H_1:a_1\neq a_2$; $d.f.=30$; $t=1.89$; 0.05 < P-მნიშვნელობა < 0.1; H_0; III) $-0.15<a_1-a_2<0.95$. **11.** I) $H_0:\sigma_1^2=\sigma_2^2$, $H_1:\sigma_1^2\neq\sigma_2^2$; $C.V.=7.15$; $\overline{f}=1.23$; H_0; II) $H_0:a_1=a_2$, $H_1:a_1\neq a_2$; $C.V.=t_{10,0.025}=\pm2.228$; $t=0.119$; არა – H_0; III) $-5.9<a_1-a_2<6.5$.

თავი XI

1. $C.V. = F_{15,22,0.005} = 3.36$; $C.V. = F_{24,13,0.01} = 3.59$; $C.V. = F_{45,29,0.025} = 2.03$; $C.V. = F_{20,16,0.05} = 2.28$; $C.V. = F_{10,10,0.05} = 2.98$. 2. 0.025 < P -მნიშვნელობა< <0.05; 0.05 < P -მნიშვნელობა < 0.1; P -მნიშვნელობა = 0.05; 0.005< < P -მნიშვნელობა < 0.01; P -მნიშვნელობა = 0.05; P -მნიშვნელობა>0.1; 0.05< P -მნიშვნელობა<0.1; 0.01< P -მნიშვნელობა<0.02. 3. $H_0 : \sigma_1^2 \le \sigma_2^2$, $H_1 : \sigma_1^2 > \sigma_2^2$; $C.V. = 2.23$; $\overline{f} = 1.41$; არა. 4. $H_0 : \sigma_1^2 = \sigma_2^2$, $H_1 : \sigma_1^2 \ne \sigma_2^2$; $C.V. = 3.18$; $\overline{f} = 1.67$; კი. 5. $H_0 : \sigma_1^2 = \sigma_2^2$, $H_1 : \sigma_1^2 \ne \sigma_2^2$; $C.V. = 2.86$; $\overline{f} = 7.85$; კი. 6. $H_0 : \sigma_1 = \sigma_2$, $H_1 : \sigma_1 \ne \sigma_2$; $C.V. = 2.51$; $\overline{f} = 3.346$; კი. 7. $H_0 : \sigma_1^2 = \sigma_2^2$, $H_1 : \sigma_1^2 \ne \sigma_2^2$; $C.V. = 2.53$; $\overline{f} = 2.09$; არა. 8. $H_0 : \sigma_1^2 = \sigma_2^2$, $H_1 : \sigma_1^2 \ne \sigma_2^2$; $C.V. = 3.53$; $\overline{f} = 1.88$; არა. 9. $H_0 : \sigma_1^2 \le \sigma_2^2$, $H_1 : \sigma_1^2 > \sigma_2^2$; $C.V. = 2.66$; $\overline{f} = 5.27$; H_1. 10. $H_0 : \sigma_1^2 \le \sigma_2^2$, $H_1 : \sigma_1^2 > \sigma_2^2$; $C.V. = 3.05$; $\overline{f} = 2.17$; არა. 11. $H_0 : \sigma_1^2 \le \sigma_2^2$, $H_1 : \sigma_1^2 > \sigma_2^2$; $C.V. = 2.9$; $\overline{f} = 1.72$; არა. 12. $H_0 : \sigma_1^2 \ge \sigma_2^2$, $H_1 : \sigma_1^2 < \sigma_2^2$; $C.V. = 3.15$; $\overline{f} = 1.45$; არა. 13. $H_0 : \sigma_1^2 = \sigma_2^2$, $H_1 : \sigma_1^2 \ne \sigma_2^2$; $C.V. = 3.87$; $\overline{f} = 3.18$; H_0. 14. $H_0 : \sigma_1^2 = \sigma_2^2$, $H_1 : \sigma_1^2 \ne \sigma_2^2$; $\overline{f} = 5.32$; P -მნიშვნელობა<0.01< $\alpha = 0.05$; H_1. 15. $H_0 : \sigma_1^2 = \sigma_2^2$, $H_1 : \sigma_1^2 \ne \sigma_2^2$; $C.V. = 3.5$; $\overline{f} = 3.773$; H_1.

თავი XII

1. $H_0 : a_D \le 0$, $H_1 : a_D > 0$; $C.V. = t_{9,0.05} = 1.833$; $t = 1.56$; არა – H_0. 2. $H_0 : a_D \ge 0$, $H_1 : a_D < 0$; $C.V. = t_{8,0.1} = -1.397$; $t = -2.8$; კი – H_1. 3. $H_0 : a_D = 0$, $H_1 : a_D \ne 0$; $C.V. = \pm t_{9,0.025} = \pm 2.262$; $t = 0.176$; არა – H_0; $(-1.18, 1.38)$. 4. $H_0 : a_D \le 0$, $H_1 : a_D > 0$; $C.V. = t_{9,0.05} = 1.833$; $t = 5.435$; არა – H_1. 5. $H_0 : a_D \ge 0$, $H_1 : a_D < 0$; $C.V. = t_{9,0.1} = -1.383$; $t = -3.4$; კი – H_1. 6. $H_0 : a_D \le 0$, $H_1 : a_D > 0$; $C.V. = t_{5,0.025} = 2.571$; $t = 2.24$; კი – H_0. 7. $H_0 : a_D \ge 0$, $H_1 : a_D < 0$; $C.V. = t_{7,0.025} = -2.365$; $t = 0.765$; არა – H_0. 8. $H_0 : a_D = 0$, $H_1 : a_D \ne 0$; $d.f. = 7$; $t = 0.978$; α < 0.2 < P -მნიშ-

ვნელობა < 0.5; H_0. **9.** $H_0: a_D = 0$, $H_1: a_D \neq 0$; $d.f. = 15$; $t = -0.5$; P -მნიშვნელობა > 0.5 > α; არა – H_0.

თავი XIII

1. $H_0: p_1 = p_2$, $H_1: p_1 \neq p_2$; $C.V. = \pm z_{0.025} = \pm 1.96$; $z = 3.64$; კი – H_1. **2.** $H_0: p_1 \leq p_2$, $H_1: p_1 > p_2$; $C.V. = z_{0.05} = 1.65$; $z = 1.36$; არა – H_0. **3.** $H_0: p_1 = p_2$, $H_1: p_1 \neq p_2$; $C.V. = \pm z_{0.025} = \pm 1.96$; $z = -2.12$; კი – H_1. **4.** $H_0: p_1 = p_2$, $H_1: p_1 \neq p_2$; $C.V. = \pm z_{0.025} = \pm 1.96$; $z = -0.99$; არა – H_0; (-0.181, 0.055). **5.** $H_0: p_1 = p_2$, $H_1: p_1 \neq p_2$; $C.V. = \pm z_{0.025} = \pm 1.96$; $z = 1.39$; H_0; (-0.032, 0.192). **6.** $H_0: p_1 = p_2$, $H_1: p_1 \neq p_2$; $C.V. = \pm z_{0.05} = \pm 1.65$; $z = -1.56$; არა – H_0; (-0.123, 0.003). **7.** $H_0: p_1 = p_2$, $H_1: p_1 \neq p_2$; $C.V. = \pm z_{0.005} = \pm 2.58$; $z = 1.302$; არა – H_0; (-0.097, 0.297). **8.** $H_0: p_1 \leq p_2$, $H_1: p_1 > p_2$; $z = 9.9$; P -მნიშვნელობა<0.001< $\alpha = 0.01$; H_1. **9.** $H_0: p_1 = p_2$, $H_1: p_1 \neq p_2$; $C.V. = \pm z_{0.025} = \pm 1.96$; $z = 0.521$; არა – H_0; (-0.103, 0.178). **10.** $H_0: p_1 = p_2$, $H_1: p_1 \neq p_2$; $z = 0.96$; P -მნიშვნელობა 0.337> $\alpha = 0.02$; არა – H_0. **11.** $H_0: p_1 = p_2$, $H_1: p_1 \neq p_2$; $C.V. = \pm z_{0.005} = \pm 2.58$; $z = -1.45$; არა – H_0; (-0.165, 0.045). **12.** $H_0: p_1 \geq p_2$, $H_1: p_1 < p_2$; $z = -1.41$; P -მნიშვნელობა = 0.0793 > α; არა – H_0. **13.** $H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$; $C.V. = \pm z_{0.05} = \pm 1.65$; $z = -3.76$; H_1; (-4328.91,-4291.1). **14.** $H_0: \mu_1 \leq \mu_2$, $H_1: \mu_1 > \mu_2$; $C.V. = z_{0.01} = 2.33$; $z = 0.59$; H_0. **15.** $H_0: \sigma_1^2 = \sigma_2^2$, $H_1: \sigma_1^2 \neq \sigma_2^2$; $C.V. = F_{17,15,0.025} = 2.86$; $T.V. = \overline{f} = 2.24$; არა – H_0. **16.** $H_0: \sigma_1^2 = \sigma_2^2$, $H_1: \sigma_1^2 \neq \sigma_2^2$; $C.V. = F_{24,24,0.05} = 1.98$; $T.V. = \overline{f} = 2.16$; კი – H_1. **17.** $H_0: \sigma_1 = \sigma_2$, $H_1: \sigma_1 \neq \sigma_2$; $C.V. = F_{23,10,0.05} = 2.77$; $T.V. = \overline{f} = 9.88$; კი – H_1. **18.** $H_0: \sigma_1 = \sigma_2$, $H_1: \sigma_1 \neq \sigma_2$; $C.V. = F_{29,29,0.005} = 2.76$; $T.V. = \overline{f} = 3.84$; კი – H_1. **19.** $H_0: \sigma_1^2 \leq \sigma_2^2$, $H_1: \sigma_1^2 > \sigma_2^2$; $C.V. = F_{64,41,0.1} = 1.47$; $T.V. = \overline{f} = 2.32$; H_1. **20.** I) $H_0: \sigma_1^2 = \sigma_2^2$, $H_1: \sigma_1^2 \neq \sigma_2^2$; $C.V. = F_{14,23,0.005} = 3.47$; $T.V. = \overline{f} = 2.78$; H_0; II) $H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$; $C.V. = \pm t_{37,0.005} = \pm 2.74$; $t = -14.09$; H_1.**21.** I) $H_0: \sigma_1^2 = \sigma_2^2$,

$H_1:\sigma_1^2 \neq \sigma_2^2$; $C.V.=1.98$; $T.V.=\overline{f}=1.11$; H_0; II) $H_0:\mu_1 \leq \mu_2$, $H_1:\mu_1 > \mu_2$; $C.V.=1.28$; $t=1.31$; H_1. **22.** I) $H_0:\sigma_1^2=\sigma_2^2$, $H_1:\sigma_1^2 \neq \sigma_2^2$; $C.V.=2.87$; $T.V.=\overline{f}=5.17$; H_1; II) $H_0:\mu_1=\mu_2$, $H_1:\mu_1 \neq \mu_2$; $C.V.=\pm 2.624$; $d.f.=14$; $t=6.54$; H_1; III) (3447.8, 8068.2). **23.** I) $H_0:\sigma_1^2=\sigma_2^2$, $H_1:\sigma_1^2 \neq \sigma_2^2$; $T.V.=\overline{f}=39.7$; P-მნიშვნელობა < 0.05; H_1; II) $H_0:\mu_1 \leq \mu_2$, $H_1:\mu_1 > \mu_2$; $d.f.=11$; $t=3.74$; P-მნიშვნელობა < 0.005; H_1. **24.** $H_0:a_D \geq 0$, $H_1:a_D < 0$; $C.V.=t_{9,0.01}=-2.821$; $T.V.\equiv t=-4.17$; კი – H_1. **25.** $H_0:a_D \geq 0$, $H_1:a_D < 0$; $C.V.=t_{7,0.05}=-1.895$; $T.V.\equiv t=-2.73$; კი – H_1. **26.** $H_0:p_1=p_2$, $H_1:p_1 \neq p_2$; $C.V.=\pm z_{0.01}=\pm 2.33$; $z=3.03$; კი – H_1; (0.027, 0.213). **27.** $H_0:p_1=p_2$, $H_1:p_1 \neq p_2$; $C.V.=\pm z_{0.025}=\pm 1.96$; $z=2.51$; არა – H_1; (0.058, 0.422).

თავი XIV

1. H_0: თანაბარია, H_1: არათანაბარია; $C.V.=\chi^2_{6,0.05}=12.59$; $T.V.=\chi^2=28.887$; H_1. **2.** H_0 : არ ენიჭება, H_1 : ენიჭება; $C.V.=\chi^2_{3,0.1}=6.251$; $T.V.=\chi^2=3.28$; H_0. **3.** H_0 : არ ენიჭება (თანაბარია), H_1 : ენიჭება (თანაბარია); $C.V.=\chi^2_{3,0.01}=11.345$; $T.V.=\chi^2=8.625$; კი – H_0. **4.** H_0 : არ ანიჭებენ, H_1 : ანიჭებენ; $C.V.=\chi^2_{5,0.05}=11.071$; $T.V.=\chi^2=12.067$; კი – H_1. **5.** H_0 : იგივეა, H_1 : განსხვავებულია; $C.V.=\chi^2_{3,0.1}=6.251$; $T.V.=\chi^2=100.275$; H_1. **6.** H_0: იგივეა, H_1 : განსხვავებულია; $C.V.=\chi^2_{4,0.05}=9.488$; $T.V.=\chi^2=7.872$; არა -- H_0. **7.** H_0 : ემთხვევა, H_1 : არ ემთხვევა; $C.V.=\chi^2_{3,0.01}=11.345$; $T.V.=\chi^2=36.8897$; H_1. **8.** H_0 : არა აქვს უპირატესობა, H_1 : აქვს უპირატესობა; $C.V.=\chi^2_{3,0.05}=7.815$; $T.V.=\chi^2=4.67$; H_0. **9.** H_0 : სწორია, H_1 : არასწორია; $T.V.=\chi^2=0.6$;

P-მნიშვნელობა > 0.1; კი -- H_0. **10.** H_0 : წესიერია, H_1 : არა-წესიერია; $C.V.=\chi^2_{3,0.05}=7.815$; $T.V.=\chi^2=139.4$; H_1.

თავი XV

1. H_0: დამოუკიდ., H_1: დამოკიდ.; $C.V.=\chi^2_{2,0.05}=5.991$; $T.V.\equiv\chi^2=6.789$; H_1. **2.** H_0: დამოუკიდ., H_1: დამოკიდ.; $C.V.=\chi^2_{6,0.01}=16.812$; $T.V.\equiv\chi^2=15.824$; არა – H_0. **3.** H_0: დამოუკიდ., H_1: დამოკიდ.; $C.V.=\chi^2_{6,0.05}=12.592$; $T.V.\equiv\chi^2=24.004$; არა – H_1. **4.** H_0: დამო-უკიდ., H_1: დამოკიდ.; $C.V.=\chi^2_{2,0.05}=5.991$; $T.V.\equiv\chi^2=21.347$; H_1. **5.** H_0: დამოუკიდ., H_1: დამოკიდ.; $C.V.=\chi^2_{4,0.1}=7.779$; $T.V.\equiv\chi^2=19.507$; კი – H_1. **6.** H_0: დამოუკიდ., H_1: დამოკიდ.; $C.V.=\chi^2_{3,0.05}=7.815$; $T.V.\equiv\chi^2=2.218$; კი – H_1. **7.** H_0: დამოუკიდ., H_1: დამოკიდ.; $C.V.=\chi^2_{8,0.1}=13.362$; $T.V.\equiv\chi^2=46.733$; არა – H_0. **8.** H_0: დამოუკიდ., H_1: დამოკიდ.; $C.V.=\chi^2_{1,0.05}=3.841$; $T.V.\equiv\chi^2=11.441$; H_1. **9.** H_0: დამოუკიდ., H_1: დამოკიდ.; $C.V.=\chi^2_{2,0.1}=4.605$; $T.V.\equiv\chi^2=6.342$; არა – H_1. **10.** H_0: არაე-ფექტ., H_1: ეფექტ.; $T.V.\equiv\chi^2=10.643$; P-მნიშვნელობა < 0.005; H_1. **11.** H_0: დამოუკიდ., H_1: დამოკიდ.; $T.V.\equiv\chi^2=19.43$; P-მნიშვნელობა < 0.005; H_1.

თავი XVI

1. $H_0: p_1=p_2=p_3=p_4$, H_1: ერთი მაინც განსხ.; $C.V.=\chi^2_{3,0.05}=7.815$; $T.V.\equiv\chi^2=5.317$; H_0. **2.** $H_0: p_1=p_2=p_3$, H_1: ერთი მაინც განსხ.; $C.V.=\chi^2_{2,0.01}=9.210$; $T.V.\equiv\chi^2=8.046$; H_0. **3.** $H_0: p_1=p_2=p_3=p_4$, H_1: ერთი მაინც განსხ.; $C.V.=\chi^2_{3,0.05}=7.815$; $T.V.\equiv\chi^2=3.4$; H_0.

4. $H_0: p_1 = p_2 = p_3$, H_1: ერთი მაინც განსხ.; $C.V. = \chi^2_{2,0.1} = 4.605$; $T.V. \equiv \chi^2 = 18.06$; H_1. **5.** $H_0: p_1 = p_2 = p_3 = p_4$, H_1: ერთი მაინც განსხ.; $C.V. = \chi^2_{3,0.1} = 6.251$; $T.V. \equiv \chi^2 = 12.755$; H_1. **6.** $H_0: p_1 = = p_2 = p_3 = p_4$, H_1: ერთი მაინც განსხ.; $C.V. = \chi^2_{3,0.05} = 7.815$; $T.V. \equiv \chi^2 = 5$; არა – H_0. **7.** $H_0: p_1 = p_2 = p_3 = p_4$, H_1: ერთი მაინც განსხ.; $T.V. \equiv \chi^2 = 1.734$; P-მნიშვნელობა > 0.1; H_0. **8.** $H_0: p_1 = p_2 = p_3$, H_1: ერთი მაინც განსხ.; $C.V. = \chi^2_{2,0.1} = 4.605$; $T.V. \equiv \chi^2 = 2.401$; H_0.

თავი XVII

1. $r = -0.98$, ძლიერი უარყოფითი კორელაცია. **2.** $r = 0.98$, ძლიერი დადებითი კორელაცია. **3.** $r = 0.62$. **4.** $H_0: \rho = 0$, $H_1: \rho \neq 0$; $r = 0.988$; $C.V. = \pm t_{n-2,\alpha/2} = \pm t_{4,0.025} = \quad = \pm 2.776$; H_1. **5.** $H_0: \rho = 0$, $H_1: \rho \neq 0$; $r = -0.832$; H_1. **6.** $H_0: \rho = 0$, $H_1: \rho \neq 0$; $r = 0.952$; H_1. **7.** $H_0: \rho = 0$, $H_1: \rho \neq 0$; $r = -0.979$; H_1. **8.** $H_0: \rho = 0$, $H_1: \rho \neq 0$; $r = 0.716$; H_1. **9.** $H_0: \rho = 0$, $H_1: \rho \neq 0$; $r = 0.814$; H_0. **10.** $H_0: \rho = 0$, $H_1: \rho \neq 0$; $r = 0.956$; H_1; $y = 1.969x - 10.944$; $y(30) = 48$. **11.** $H_0: \rho = 0$, $H_1: \rho \neq 0$; $r = -0.984$; H_1; $y = -0.14x + 10.199$. **12.** $H_0: \rho = 0$, $H_1: \rho \neq 0$; $r = 0.963$; H_1; $y = 0.86x + 10.251$. **13.** $H_0: \rho = 0$, $H_1: \rho \neq 0$; $r = -0.981$; H_1; $y = -2.668x + 96.784$. **14.** $H_0: \rho = 0$, $H_1: \rho \neq 0$; $r = -0.306$; 0.2 < P-მნიშვნელობა < 0.5; H_0; r არამნიშვნელოვანია. **15.** $H_0: \rho = 0$, $H_1: \rho \neq 0$; $r = 0.491$; 0.2< P-მნიშვნელობა<0.5; H_0; r არამნიშვნელოვანია. **16.** $H_0: \rho = 0$, $H_1: \rho \neq 0$; $r = 0.857$; H_1; $y = 0.19x - 2.818$; $y(35) = 3.84 \approx 4$. **17.** $H_0: \rho = 0$, $H_1: \rho \neq 0$; $r = -0.078$; H_0; კორელაცია არამნიშვნე-ლოვანია. **18.** $H_0: \rho = 0$, $H_1: \rho \neq 0$; $r = 0.842$; H_1; $y = 0.551x - 1.918$; $y(11) = 4.14 \approx 4$. **19.** $H_0: \rho = 0$, $H_1: \rho \neq 0$; $r = 0.602$; H_0; კორელა-ცია არამნიშვნელოვანია. **20.** 1.129. **21.** 29.5. **22.** (0, 5). **23.** 217.5 – დაკვირვებულ მნიშვნელობათა საშუალო.

თავი XVIII

1. $\rho = 0.714$. 2. 40. 3. 0.75.

თავი XIX

1. H_0: $a_1 = a_2 = a_3$, H_1: ერთი მაინც განსხ.; $C.V. = F_{2,16,0.05} = 3.63$; $T.V. \equiv f = 5.96$; კი – H_1. 2. H_0: $a_1 = a_2 = a_3$, H_1: ერთი მაინც განსხ.; $C.V. = F_{2,27,0.05} = 3.35$; $T.V. \equiv f = 7.456$; კი -- H_1. 3. H_0: $a_1 = a_2 = a_3$, H_1: ერთი მაინც განსხ.; $C.V. = F_{2,12,0.1} = 2.81$; $T.V. \equiv f = 8.448$; კი – H_1. 4. H_0: $a_1 = a_2 = a_3$, H_1: ერთი მაინც განსხ.; $C.V. = F_{2,18,0.05} = 3.55$; $T.V. \equiv f = 19.05$; კი – H_1. 5. H_0: $a_1 = a_2 = a_3$, H_1: ერთი მაინც განსხ.; $C.V. = F_{2,11,0.05} = 3.98$; $T.V. \equiv f = 7.75$; კი – H_1. 6. H_0: $a_1 = a_2 = a_3$, H_1: ერთი მაინც განსხ.; $C.V. = F_{2,17,0.1} = 2.64$; $T.V. \equiv f = 1.28$; არა – H_0. 7. H_0: $a_1 = a_2 = a_3 = a_4$, H_1: ერთი მაინც განსხ.; $C.V. = F_{3,24,0.05} = 3.01$; $T.V. \equiv f = 6.974$; არა – H_1. 8. H_0: $a_1 = a_2 = a_3$, H_1: ერთი მაინც განსხ.; $C.V. = F_{2,12,0.1} = 2.81$; $T.V. \equiv f = 9.16$; კი – H_1. 9. H_0: $a_1 = a_2 = a_3 = a_4$, H_1: ერთი მაინც განსხ.; $C.V. = F_{3,31,0.1} = 2.28$; $T.V. \equiv f = 234.5$; კი – H_1. 10. H_0: $a_1 = a_2 = a_3$, H_1: ერთი მაინც განსხ.; $C.V. = F_{2,12,0.05} = 3.89$; $T.V. \equiv f = 5.877$; კი – H_1. 11. H_0: ერთიდაიგივეა, H_1: განსხვავებულია; $C.V. = \chi^2_{4,0.05} = 9.488$; $T.V. \equiv \chi^2 = 87.14$; არა – H_1. 12. H_0: ერთიდაიგივეა, H_1: განსხვავებულია; $C.V. = \chi^2_{3,0.01} = 11.345$; $T.V. \equiv \chi^2 = 13.38$; არა – H_1. 13. H_0: დამოუკიდ., H_1: დამოკიდ.; $C.V. = \chi^2_{2,0.1} = 4.605$; $T.V. \equiv \chi^2 = 6.163$; კი – H_1. 14. H_0: დამოუკიდ., H_1: დამოკიდ.; $C.V. = \chi^2_{2,0.1} = 4.605$; $T.V. \equiv \chi^2 = 3.05$; არა – H_0. 15. H_0: დამოუკიდ., H_1: დამოკიდ.; $C.V. = \chi^2_{4,0.05} = 9.488$; $T.V. \equiv \chi^2 = 28$; კი – H_1. 16. H_0: დამოუკიდ., H_1: დამოკიდ.; $C.V. = \chi^2_{2,0.1} = 4.605$; $T.V. \equiv \chi^2 = 7.674$; H_1. 17. H_0: $p_1 = p_2 = p_3$, H_1: ერთი მაინც განსხ.; $C.V. = \chi^2_{2,0.01} = 9.21$; $T.V. \equiv \chi^2 = 23.89$;

არა – H_1. **18.** H_0: $a_1 = a_2 = a_3$, H_1: ერთი მაინც განსხ.; $C.V. = F_{2,18,0.01} = 6.01$; $T.V. \equiv f = 27.02$; H_1. **19.** H_0: $a_1 = a_2 = a_3$, H_1: ერთი მაინც განსხ.; $C.V. = F_{2,13,0.05} = 3.81$; $T.V. \equiv f = 0.533$; არა – H_0. **20.** H_0: $a_1 = a_2 = a_3$, H_1: ერთი მაინც განსხ.; $C.V. = F_{2,12,0.05} = 3.89$; $T.V. \equiv f = 6.141$; კი – H_1. **21.** H_0: $a_1 = a_2 = a_3$, H_1: ერთი მაინც განსხ.; $C.V. = F_{2,12,0.05} = 3.89$; $T.V. \equiv f = 65.263$; კი – H_1. **22.** H_0: $a_1 = a_2 = a_3$, H_1: ერთი მაინც განსხ.; $C.V. = F_{2,12,0.05} = 3.89$; $T.V. \equiv f = 3.673$; არა – H_0.

ლიტერატურა

1. გ. მანია. ალბათობის თეორია და მათემატიკური სტატისტიკა. სახელმძღვანელო ეკონომიკური ფაკულტეტის სტუდენტებისათვის. თსუ გამომცემლობა, თბილისი, 1976
2. ნ. ანთელავა, ა. ედიბერიძე, გ. მანია. ალბათობის თეორიისა და მათემატიკური სტატისტიკის ამოცანათა კრებული. თსუ გამომცემლობა, თბილისი, 1980.
3. გ. მარი, ა. მოსიძე, ზ. ციგროშვილი. სტატისტიკა. დამხმარე სახელმძღვანელო ESM-თბილისის სტუდენტებისათვის, ESM-თბილისი, 1996.
4. ნ. ლაზრიევა, მ. მანია, გ. მარი, ა. მოსიძე, ა. ტორონჯაძე, თ. ტორონჯაძე, თ. შერვაშიძე. ალბათობის თეორია და მათემატიკური სტატისტიკა ეკონომისტებისათვის. ფონდი «ევრაზია», თბილისი, 2000.
5. ე. ნადარაია, რ. აბსავა, მ. ფაცაცია. ალბათობის თეორია. თბილისი, 2005.
6. გ. მარი, ა. მოსიძე. ალბათობის თეორია და მათემატიკური სტატისტიკა. დამხმარე სახელმძღვანელო ESM-თბილისის სტუდენტებისათვის, ESM-თბილისი, 2007.
7. ო. ფურთუხია. ალბათობის თეორია და მათემატიკური სტატისტიკა. თბილისი, 2007.
8. ო.ფურთუხია. აღწერითი სტატისტიკა, ალბათობა, სტატისტიკური დასკვნების თეორია. თბილისი, 2008.
9. ე. ნადარაია, ბ. დოჭვირი, პ. ბაბილუა, მ. ბერაძე, გ. ლომინაშვილი, მ. მნაცაკანიანი. ალბათობის თეორის ამოცანათა კრებული. ქუთაისი, 2008.
10. ე. ნადარაია, ბ. დოჭვირი, თ. ბოკელავაძე, გ. ლომინაშვილი, მ. მნაცაკანიანი, მ. ფაცაცია. ალბათობის თეორია და მათემატიკური სტატისტიკა (ამოცანათა კრებული). ქუთაისი, 2008.

11. ო. ფურთუხია. ალბათობა და სტატისტიკა მაგალითებსა და ამოცანებში. თბილისი, 2009.
12. ო. ფურთუხია, ზ. ციგროშვილი, ქ. მანჯგალაძე. უმაღლესი მათემატიკა, ნაწილი III, ალბათობა და მათემატიკური სტატისტიკა. თსუ გამომცემლობა, თბილისი, 2009.
13. ლ. ალექსიძე, ზ. ზერაკიძე, ო. ფურთუხია, ზ. ხეჩინაშვილი. ალბათობის თეორია და მათემატიკური სტატისტიკა. თბ., 2009.
14. Sheldon Ross. A first Course in Probability. PRENTICE HALL, 1997.
15. Allan G. Bluman. Ementary Statistics: a brief version, second edition. Published by McGraw-Hill, New York, 2003.
16. Peter Olofsson. Probability, Statistics, and Stoichastic Processes. WILEY-INTERSCIENCE, 2005.
17. P. Newbold, W. L. Carlson, B. M. Thorne. Statistics for Business and Economics, sixth edition. Prentice Hall, Upper Saddle River, New Jersey, 2007.
18. R.A. Barnett, M.R.Ziegler, K.E.Byleen. Finite Mathematics for Business, Economics, Life Sciences, and Social Sciences. Twelfth edition, Pearson, 2009.
19. В. Феллер. ВВедение в теорию вероятностей и ее приложения. Москва, 1967.
20. УА. Г. Дьячков. Теория вероятностей. Москва, 1980.
21. А. Н. Колмогоров, И. Г. Журбенко, А. В. Прохоров. ВВедение в теорию вероятностей. Москва, 1982.
22. В. К. Захаров, Б. А. Севастьянов, В. П. Чистяков. Теория вероятностей. Москва, 1988.
23. В. Е. Гмурман. Руководство к решению задач по теории вероятностей и математической статистике. Москва, 2003.