ფინალური გამოცდის ბილეთის ნიმუშის პასუხი ალბათობის თეორიასა და მათემატიკურ სტატისტიკაში

თეორიული საკითხები

1. ხდომილებათა დამოუკიდებლობის ორი განმარტება, წყვილ-წყვილად და ერთობლივად დამოუკიდებლობა.

A და B ხდომილებას ეწოდება დამოუკიდებელი თუ:

I) $P(A|B)=P(A)$; **II)** $P(A\cap B)=P(A)P(B)$.

ხდომილებათა ერთობლიობას $A_1, A_2, \dots, A_n$ ეწოდება **წყვილ-წყვილად დამოუკიდებელი** თუ: $P(A_i \cap A_j)=P(A_i)\cdot P(A_j), \quad \forall i \neq j$.

ხდომილებათა ერთობლიობას $A_1, A_2, \dots, A_n$ ეწოდება **ერთობლივად დამოუკიდებელი** თუ $\forall \ 2 \le k \le n$, $i_1 \ne i_2 \ne \dots \ne i_k$: $P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k})=P(A_{i_1})\cdot P(A_{i_2}) \cdots P(A_{i_k})$.

2. პუასონის განაწილების კანონი (განმარტება) და რიცხვითი მახასიათებლები (ლოდინი, დისპერსია, მოდა).

პუასონის განაწილება პარამეტრით λ აღინიშნება სიმბოლოთი $Po(\lambda)$ და განიმარტება:

$$P\{Po(\lambda)=k\}=\frac{\lambda^k}{k!}e^{-\lambda}, \ k=0,1,2,\dots \ ; \quad EPo(\lambda)=\lambda; \quad DPo(\lambda)=\lambda; \quad MoPo(\lambda)=[\lambda]-1.$$

3. დაწერეთ ნდობის ინტერვალი ნორმალური პოპულაციის საშუალოსათვის უცნობი დის-პერსიის შემთხვევაში (თითოეული სიდიდის მითითებით).

$1-\alpha$ საიმედოობის (α მნიშვნელოვნების დონის მქონე) ნდობის ინტერვალს ნორმალური პოპულაციის საშუალოსათვის უცნობი დისპერსიის შემთხვევაში აქვს სახე:

$$(\bar{x}-t_{n-1,\alpha/2}\cdot\frac{s'}{\sqrt{n}}, \bar{x}+t_{n-1,\alpha/2}\cdot\frac{s'}{\sqrt{n}}), \qquad (1)$$

სადაც n არის შერჩევის მოცულობა, s' არის შესწორებული შერჩევითი სტანდარტული გადახრა – $s'=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}$ (რომელშიც $\bar{x}$ შერჩევითი საშუალოა – $\bar{x}=\frac{x_1+\dots+x_n}{n}$), ხო-ლო $t_{n-1,\alpha/2}$ არის თავისუფლების $n-1$ ხარისხის მქონე სტიუდენტის განაწილების ზედა $\alpha/2$ -კრიტიკული წერტილი.

4. ჰიპოთეზის შემოწმება ნორმალური პოპულაციის საშუალოსათვის ცნობილი დისპერსიის შემთვევაში (კრიტერიუმი მარჯვენა ცალმხრივია).

პოპულაცია ნორმალურია $\xi \cong N(a,\sigma^2)$, $D\xi=\sigma^2$ ცნობილია, $E\xi$ უცნობია. ძირითადი ჰიპოთეზა: $H_0: E\xi=a_0$, ალტერნატიული ჰიპოთეზა: $H_1: E\xi=a_1>a_0$. კრიტერიუმის სტატისტიკა და მისი განაწილების კანონი: $Z=\frac{\overline{X}-a_0}{\sigma/\sqrt{n}}\cong N(0,1)$. მნიშვნელოვნების დონე: α. კრიტიკული წერტილი: $C.V.=z_\alpha$. აქ z_α – სტანდარტული ნორმალური განაწილების ზედა α-კრიტიკული წერტილია. კრიტიკული არე: C.R. = $[z_\alpha,+\infty)$.

გადაწყვეტილების მიღების წესი: თუ კრიტერიუმის დაკვირვებული მნიშვნელობა **T.V.**$\in$ **C.R.**, მაშინ H_0 ჰიპოთეზას უკუვაგდებთ, წინააღმდეგ შემთხვევაში ამის საფუძველი არა გვაქვს.

პრაქტიკული ამოცანები

1. ჩანთაში დევს 6 წითელი და 4 მწვანე კალკულატორი. იღებენ ორ კალკულატორს დაბრუ-ნების გარეშე. იპოვეთ ალბათობა იმისა, რომ: ა) ორივე კალკულატორი წითელია; ბ) ორივე მწვანეა; გ) ზუსტად ერთი კალკულატორი წითელია; დ) ერთი კალკულატორი მაინც წითელია; ე) მეორე კალკულატორი წითელია.

ამოხსნა. შემოვიღოთ ხდომილებები: A – პირველი კალკულატორი წითელია, B – მეორე კალკულატორი წითელია. ამ შემთხვევაში $|\Omega| = A_{10}^2 = 90$.

ა) $P(AB) = A_6^2 / A_{10}^2 = 1/3$;

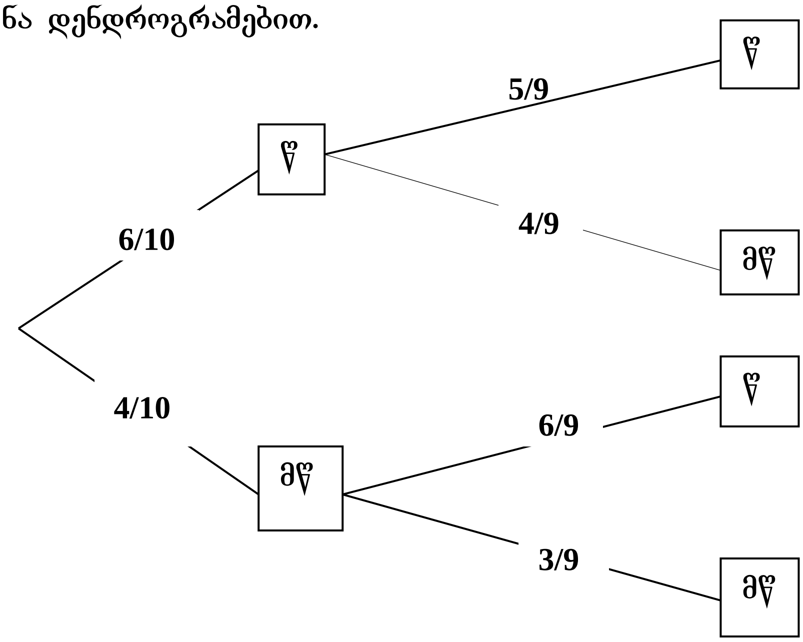
ბ) $P(\overline{A}\,\overline{B}) = A_4^2 / A_{10}^2 = 2/15$;

გ) $P(A\overline{B} \cup \overline{A}B) = P(A\overline{B}) + P(\overline{A}B) = (6\cdot 4 + 4\cdot 6)/A_{10}^2 = 8/15$;

დ) $P[\overline{(\overline{A}\,\overline{B})}] = 1 - P(\overline{A}\,\overline{B}) = 1 - A_4^2 / A_{10}^2 = 13/15$;

ე) $P(B) = P(A)P(B\,|\,A) + P(\overline{A})P(B\,|\,\overline{A}) = \frac{6}{10}\cdot\frac{5}{9} + \frac{4}{10}\cdot\frac{6}{9} = 3/5$.

ამოხსნა დენდროგრამებით.



ა) $P(\text{I - წ;II - წ}) = \frac{6}{10}\cdot\frac{5}{9} = \frac{1}{3}$;

ბ) $P(\text{I - მწ;II - მწ}) = \frac{4}{10}\cdot\frac{3}{9} = \frac{2}{15}$;

გ) $P(\text{I - წ;II - მწ ან I - მწ;II - წ}) = \frac{6}{10}\cdot\frac{4}{9} + \frac{4}{10}\cdot\frac{6}{9} = \frac{8}{15}$;

დ) $P(\text{ერთი მაინც წითელია}) = 1 - P(\text{ორივე მწვანეა}) = 1 - P(\text{I - მწ;II - მწ}) = 1 - \frac{4}{10}\cdot\frac{3}{9} = \frac{13}{15}$;

ე) $P(\text{მეორე წითელია}) = P(\text{I - მწ;II - წ ან I - წ;II - წ}) = \frac{4}{10}\cdot\frac{6}{9} + \frac{6}{10}\cdot\frac{5}{9} = \frac{3}{5}$.

1*. აგდებენ ორ სათამაშო კამათელს. თუ ცნობილია, რომ ერთ კამათელზე აღმოჩნდა 4 ქულა, მაშინ რას უდრის ალბათობა იმისა, რომ: ა) მეორე კამათელზე აღმოჩნდება 5 ქულა? ბ) მეორე კამათელზე მოსული ქულა ნაკლები იქნება 4-ზე? გ) ორივე კამათელზე მოსული ჯამური ქულა მეტი იქნება 7-ზე?

ამოხსნა. ამ შემთხვევაში $|\Omega| = 6\cdot 6 = 36$. შემოვიღოთ ხდომილებები: A – ერთ კამათელზე აღმოჩნდა 4 ქულა, B – მეორე კამათელზე აღმოჩნდება 5 ქულა, C – მეორე კამათელზე მოსული ქულა ნაკლებია 4-ზე, D – ორივე კამათელზე მოსული ჯამური ქულა მეტია 7-ზე. ცხადია, რომ:

$A = \{(4,1);(4,2);\ldots;(4,6);(1,4);(2,4);(3,4);(5,4);(6,4)\}$, $|A| = 11$;

$AB = \{(5,4);(4,5)\}$, $|AB| = 2$;

$AC = \{(4,1);(4,2);(4,3);(1,4);(2,4);(3,4)\}$, $|AC| = 6$;

$AD = \{(4,4);(4,5);(4,6);(5,4);(6,4)\}$, $|AD| = 5$.

ა) $P(B\,|\,A) = P(AB)/P(A) = (2/36)/(11/36) = 2/11$;

ბ) $P(C \mid A) = P(AC)/P(A) = (6/36)/(11/36) = 6/11$;

გ) $P(D \mid A) = P(AD)/P(A) = (5/36)/(11/36) = 5/11$.

2. მოყვარული სინოპტიკოსის თეორიის თანახმად თუ ერთ წელს იყო წყალდიდობა, მაშინ ალბათობა იმისა, რომ იგი განმეორდება მომდევნო წელს არის 0.7, ხოლო თუ ერთ წელს არ იყო წყალდიდობა, მაშინ ალბათობა იმისა, რომ იგი არ იქნება მომდევნო წელს არის 0.6. გასულ წელს წყალდიდობა არ ყოფილა. იპოვეთ ალბათობა იმისა, რომ წყალდიდობა იქნება: ა) მომდევნო სამ წელიწადს ზედიზედ; ბ) ზუსტად ერთჯერ მომდევნო სამი წლის განმავლობაში.

ამოხსნა. შემოვიღოთ ხდომილებები: F_i – წყალდიდობა იყო i-ურ წელიწადს. ცნობილია, რომ $P(F_{i+1} \mid F_i) = 0.7$ და $P(\overline{F}_{i+1} \mid \overline{F_i}) = 0.6$. შესაბამისად, საწინააღმდეგო ხდომილების პირობითი ალბათობები იქნება: $P(\overline{F}_{i+1} \mid F_i) = 0.3$ და $P(F_{i+1} \mid \overline{F_i}) = 0.4$. საძიებელია: ა) $P(F_{i+1}F_{i+2}F_{i+3} \mid \overline{F_i}) = ?$ და ბ) $P[(F_{i+1}\overline{F}_{i+2}\overline{F}_{i+3}) \cup (\overline{F}_{i+1}F_{i+2}\overline{F}_{i+3}) \cup (\overline{F}_{i+1}\overline{F}_{i+2}F_{i+3}) \mid \overline{F_i}] = ?$. ნამრავლის ალბათობის ფორმულის თანახმად გვაქვს:

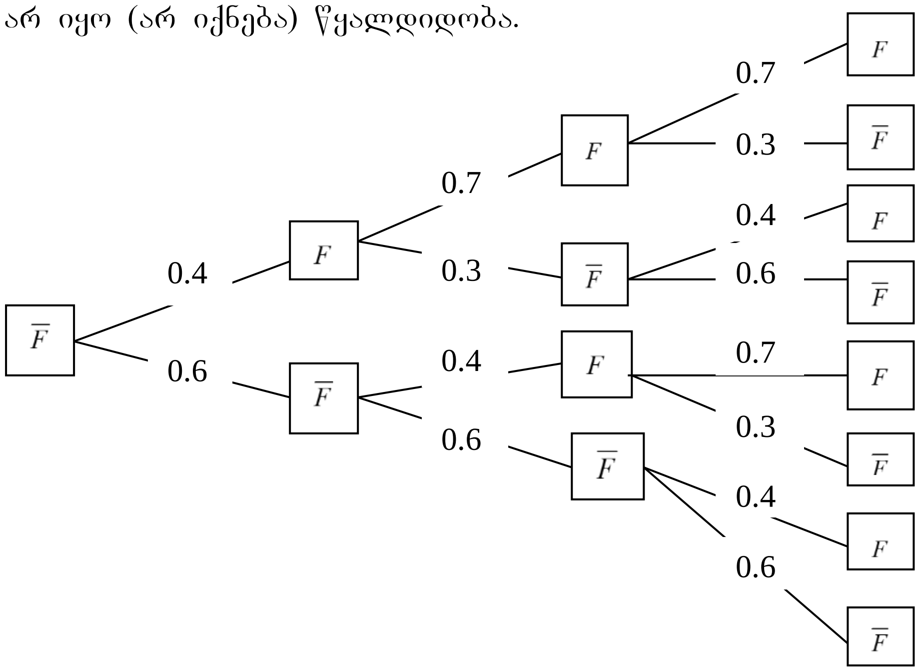
ა) $P(F_{i+1}F_{i+2}F_{i+3} \mid \overline{F_i}) = P(F_{i+1} \mid \overline{F_i})P(F_{i+2} \mid \overline{F_i}F_{i+1})P(F_{i+3} \mid \overline{F_i}F_{i+1}F_{i+2}) = 0.4 \cdot 0.7 \cdot 0.7 = 0.196$;

ანალოგიურად,

ბ) $P[(F_{i+1}\overline{F}_{i+2}\overline{F}_{i+3}) \cup (\overline{F}_{i+1}F_{i+2}\overline{F}_{i+3}) \cup (\overline{F}_{i+1}\overline{F}_{i+2}F_{i+3}) \mid \overline{F_i}] = P[(F_{i+1}\overline{F}_{i+2}\overline{F}_{i+3}) \mid \overline{F_i}] +$
$+ P[(\overline{F}_{i+1}F_{i+2}\overline{F}_{i+3}) \mid \overline{F_i}] + P[(\overline{F}_{i+1}\overline{F}_{i+2}F_{i+3}) \mid \overline{F_i}] = 0.4 \cdot 0.3 \cdot 0.6 + 0.6 \cdot 0.4 \cdot 0.3 + 0.6 \cdot 0.6 \cdot 0.4 = 0.288$.

შენიშვნა. შეგიძლიათ არ დაწეროთ არცერთი ფორმულა და ამოხსნათ დენდროგრამების გამოყენებით.

ამოხსნა დენდროგრამებით. შემოვიღოთ ხდომილებები: F – იყო (იქნება) წყალდიდობა, $\overline{F}$ – არ იყო (არ იქნება) წყალდიდობა.



ა) $P = 0.4 \cdot 0.7 \cdot 0.7 = 0.196$;

ბ) $P = 0.4 \cdot 0.3 \cdot 0.6 + 0.6 \cdot 0.4 \cdot 0.3 + 0.6 \cdot 0.6 \cdot 0.4 = 0.288$.

2*. თესლის მოცემული პარტიიდან დათესეს 10000 თესლი და ნახეს, რომ აღმოცენდა 8498 თესლი. ამის შედეგად გაკეთდა დასკვნა, რომ თესლის მოცემული პარტიის აღმოცენებადობა შეადგენს 85%-ს ($p = 0.85$). ვიპოვოთ ალბათობა იმისა, რომ მოცემული პარტიიდან დათესილი 100 თესლიდან აღმოცენდება: ა) ზუსტად 85 თესლი? ბ) 75-დან 90 თესლამდე? გ) არანაკლებ 80 თესლი? დ) არაუმეტეს 92 თესლი?

ამოხსნა. ცხადია, რომ საქმე გვაქვს დამოუკიდებელ ცდათა სქემასთან (ბერნულის სქემასთან) და ვინაიდან $np = 100 \cdot 0.85 = 85 > 15$, ამიტომ შეგვიძლია ვისარგებლოთ მუავრ-ლაპლასის ლოკალური და ინტეგრალური ფორმულებით. 100 თესლიდან აღმოცენებული თესლის რაოდენობა აღვნიშნოთ S_{100}-ით.

ა) $P\{S_{100}=85\}=\frac{1}{\sqrt{100\cdot 0.85\cdot 0.15}}\varphi(\frac{85-100\cdot 0.85}{\sqrt{100\cdot 0.85\cdot 0.15}})=\frac{1}{\sqrt{12.75}}\varphi(\frac{0}{\sqrt{12.75}})=$

$$=\frac{1}{3.57}\varphi(0)=0.28\cdot 0.399=0.11;$$

ბ) $P\{75\le S_{100}\le 90\}=\Phi(\frac{90-100\cdot 0.85}{\sqrt{100\cdot 0.85\cdot 0.15}})-\Phi(\frac{75-100\cdot 0.85}{\sqrt{100\cdot 0.85\cdot 0.15}})=\Phi(\frac{5}{3.57})-\Phi(\frac{-10}{3.57})=$

$=\Phi(1.4)-\Phi(-2.8)=0.5+\Phi_0(1.4)-[0.5-\Phi_0(2.8)]=\Phi_0(1.4)+\Phi_0(2.8)=0.4192+0.4974=0.9166;$

გ) $P\{S_{100}\ge 80\}=1-\Phi(\frac{80-100\cdot 0.85}{\sqrt{100\cdot 0.85\cdot 0.15}})=1-\Phi(\frac{-5}{3.57})=1-\Phi(-1.4)=1-[0.5-\Phi_0(1.4)]=$

$$=0.5+0.4192=0.9192;$$

დ) $P\{S_{100}\le 92\}=\Phi(\frac{92-100\cdot 0.85}{\sqrt{100\cdot 0.85\cdot 0.15}})=\Phi(\frac{7}{3.57})=\Phi(1.96)=0.5+\Phi_0(1.96)]=0.5+0.475=0.975.$

2.** სუპერმარკეტში დღის განმავლობაში გაყიდული ციტრუსების რაოდენობა მოდელირდება ნორმალური განაწილებით. აღმოჩნდა, რომ ხანგრძლივი პერიოდის მანძილზე დღეში საშუალოდ იყიდებოდა 35 კგ. ციტრუსი, ხოლო 15 კგ-ზე ნაკლები გაყიდულ იქნა საშუალოდ ყოველი 20 დღიდან ერთ დღეში. ა) გამოთვალეთ გაყიდვების სტანდარტული გადახრა σ; ბ) ცნობილია, რომ კონკრეტულ დღეს გაიყიდა 53 კგ-ზე მეტი ციტრუსი. იპოვეთ ალბათობა იმისა, რომ ამ დღეს გაიყიდა 56 კგ-ზე მეტი ციტრუსი.

ამოხსნა. დღეში გაყიდული ციტრუსების რაოდენობა აღვნიშნოთ ξ-თი. მაშინ: $\xi\cong N(35,\sigma^2)$ და $P\{\xi<15\}=\frac{1}{20}=0.05$.

ა) $P\{\xi<15\}=0.05\Rightarrow\Phi(\frac{15-35}{\sigma})=0.05\Rightarrow\Phi_0(-\frac{15-35}{\sigma})=0.5-0.05=0.45\Rightarrow$

$\Rightarrow-\frac{15-35}{\sigma}=1.65\Rightarrow\sigma=12.12$;

ბ) $P\{\xi>56\,|\,\xi>53\}=\frac{P\{(\xi>56)\bigcap(\xi>53)\}}{P\{\xi>53\}}=\frac{P\{\xi>56\}}{P\{\xi>53\}}=$

$$=\frac{1-\Phi(\frac{56-35}{12.12})}{1-\Phi(\frac{53-35}{12.12})}=\frac{1-\Phi(1.73)}{1-\Phi(1.49)}=\frac{1-[0.5+\Phi_0(1.73)]}{1-[0.5+\Phi_0(1.49)]}=\frac{0.5-0.4582}{0.5-0.4319}=\frac{0.0418}{0.0681}=0.614.$$

3. შემთხვევითი სიდიდის განაწილების კანონია $P\{\xi=k\}=C_{20}^k 0.3^k 0.7^{20-k}$, $k=0,1,...,20$. რას უდრის: ა) ლოდინი; ბ) დისპერსია; გ) მედიანა; დ) მოდა; ე) ამ შემთხვევით სიდიდეზე 300 დაკვირვებისას 0.4 კვანტილზე მეტი მნიშვნელობის მოსვლის უალბათესი რიცხვი.

ამოხსნა. ცხადია, რომ ეს შემთხვევითი სიდიდე განაწილებულია ბინომიალური კანონით პარამეტრებით $n=20$ და $p=0.3$, ანუ $\xi\cong Bi(20,0.3)$. ამიტომ, როგორც ცნობილია:

ა) $E\xi=20\cdot 0.3$;

ბ) $D\xi=20\cdot 0.3\cdot 0.7$;

გ) $Me=[20\cdot 0.3]=6$;

დ) $Mo=[21\cdot 0.3]=6$;

ე) $n=300$, $p=P(\xi>x_{0.4})=1-P(\xi\le x_{0.4})=0.6$ და შესაბამისად, უალბათესი რიცხვი წარმოადგენს შემდეგი უტოლობის მთელ ამონახსნს:

$$300\cdot 0.6-0.4\le k_0\le 300\cdot 0.6+0.6,$$

ანუ $179.6\le k_0\le 180.6$, ე. ი. $k_0=180$.

3*. შემთხვევითი სიდიდის განაწილების სიმკვრივეა $f_\xi(x)=\begin{cases}0, & x<0;\\ 0.4e^{-0.4x}, & x\ge 0.\end{cases}$ რას უდრის: ა) ლოდინი; ბ) დისპერსია; გ) მედიანა; დ) პირველი კვარტილი; ე) ამ შემთხვევით სიდიდეზე 400

დაკვირვებისას პირველ კვარტილზე ნაკლები მნიშვნელობის მოსვლის მოსალოდნელი რიცხვი.

ამოხსნა. ცხადია, რომ ეს შემთხვევითი სიდიდე განაწილებულია ექსპონენციალური (ანუ მაჩვენებლიანი) კანონით პარამეტრით $\lambda = 0.4$, ანუ $\xi \cong Exp(0.4)$. ამიტომ, როგორც ცნობილია:

ა) $E\xi = 1/0.4$;

ბ) $D\xi = 1/0.4^2$;

გ) $Me = \ln 2/0.4$;

დ) $Q_1 = x_{0.25} = -\ln(1-0.25)/0.4$;

ე) $n = 400$, $p = P(A) = P(\xi < x_{0.25}) = 0.25$ და ამიტომ ალბათობის სტატისტიკური განმარტების თანახმად:

$$0.25 \approx \frac{n_A}{400}, \text{ საიდანაც } n_A \approx 400 \cdot 0.25 = 100.$$

3.** მოცემულია: $\xi \cong \chi^2(16)$. იპოვეთ:**ა)** იპოვეთ ისეთი მარცხენა ცალმხრივი ინტერვალი, რომელშიც ξ-ს მოხვედრის ალბათობაა 0.975; **ბ)** ამ შემთხვევით სიდიდეზე 300 დაკვირვებისას 0.8-კვანტილზე ნაკლები მნიშვნელობის მოსვლის მოსალოდნელი რიცხვი; **გ)** ამ შემთხვევით სიდიდეზე 100 დაკვირვებიდან მედიანაზე მეტი მნიშვნელობის მოსვლის უალბათესი რიცხვი; **დ)** ალბათობა იმისა, რომ ამ შემთხვევით სიდიდეზე 10000 დაკვირვებისას 0.8 -კვანტილზე მეტი მნიშვნელობა მოვა არანაკლებ 1950 და არაუმეტეს 2050; **ე)** დაწერეთ ნორმალური განაწილების სიმკვრივე, რომლის რიცხვითი მახასიათებლები ემთხვევა $U([-2,3])$-ის შესაბამის რიცხვით მახასიათებლებს.

ამოხსნა. ა) $0.975 = P\{\xi \in (0,x)\} = 1 - P\{\xi \ge x\}$, $P\{\xi \ge x\} = 0.025$, $x = \chi^2_{16,0.025} = 28.845$;

ბ) $n = 300$, $p = P(A) = P(\xi < x_{0.8}) = 0.8$ და ამიტომ ალბათობის სტატისტიკური განმარტების თანახმად:

$$0.8 \approx \frac{n_A}{300}, \text{ საიდანაც } n_A \approx 300 \cdot 0.8 = 240.$$

გ) $n = 100$, $p = P(\xi > Me) = 0.5$ და შესაბამისად, უალბათესი რიცხვი წარმოადგენს შემდეგი უტოლობის მთელ ამონახსნს:

$$100 \cdot 0.5 - 0.5 \le k_0 \le 100 \cdot 0.5 + 0.5,$$

$$\text{ანუ } 49.5 \le k_0 \le 50.5, \text{ ე. ი. } k_0 = 50.$$

დ) $n = 10000$, $p = P(\xi > x_{0.8}) = 1 - P(\xi \le x_{0.8}) = 0.2$, $a = 1950$, $b = 2050$, $np > 15$ და ამიტომ მუავრ-ლაპლასის ინტეგრალური ფორმულის თანახმად:

$$P(1950 \le S_{10000} \le 2050) \approx \Phi\left(\frac{2050 - 10000 \cdot 0.2}{\sqrt{10000 \cdot 0.2 \cdot 0.8}}\right) - \Phi\left(\frac{1950 - 10000 \cdot 0.2}{\sqrt{10000 \cdot 0.2 \cdot 0.8}}\right) =$$

$$= \Phi(1.25) - \Phi(-1.25) = 2\Phi(1.25) - 1 = 2 \cdot [0.5 + \Phi_0(1.25)] - 1 = 2 \cdot 0.8944 - 1 = 0.7888.$$

ე) ვინაიდან $EU([-2,3]) = (-2+3)/2 = 0.5$ და $DU([-2,3]) = (-2-3)^2/12 \approx 2.08$, ამიტომ გვაქვს:

$$f(x) = \frac{1}{\sqrt{2\pi \cdot 2.08}} e^{-\frac{(x-0.5)^2}{2 \cdot 2.08}}.$$

4. 250 გამოკითხული მოსწავლიდან 60% ყოველ საღამოს 1 საათს მაინც მეცადინეობს. **ა)** ააგეთ 95%-იანი ნდობის ინტერვალი იმ მოსწავლეთა რეალური პროპორციისათვის, რომლებიც ყოველ საღამოს 1 საათს მაინც მეცადინეობენ. მასწავლებელს სურს 98%-იანი საიმედოობითა და 0.03-ის სიზუსტით შეაფასოს მაღალი კლასის იმ მოსწავლეების რეალური პროპორცია, რომლებიც ყოველ საღამოს 1 საათს მაინც მეცადინეობენ. **ბ)** რა მოცულობის უნდა იყოს საჭირო შერჩევა; **გ)** რა მოცულობის შერჩევა უნდა ავიღოთ იმ შემთხვევაში, როცა არ იქნება ხელმისაწვდომი შერჩევითი პროპორცია?

ამოხსნა. ა) $n = 250$, $S_n = 250 \cdot 60/100 = 150$, $w_n = 150/250 = 0.6$, $\alpha = 1 - 0.95 = 0.05$, $z_{\alpha/2} = z_{0.0.25} = 1.96$. ამიტომ საძიებელი ნდობის ინტერვალი იქნება:

$$(0.6-1.96\sqrt{\frac{0.6\cdot(1-0.6)}{250}},\, 0.6+1.96\sqrt{\frac{0.6\cdot(1-0.6)}{250}}),$$

ბ) $\alpha=1-0.98=0.02$, $z_{\alpha/2}=z_{0.0.1}=2.33$, $l=0.03$, $w_n=0.6$, $1-w_n=0.4$. ამიტომ შერჩევის მინიმალური მოცულობა იქნება: $n^*=[0.6\cdot 0.4\cdot(\frac{2.33}{0.03})^2]+1=[1447.83]+1=1448$.

გ) როცა არ იქნება ხელმისაწვდომი შერჩევითი პროპორცია, უნდა ავიღოთ $w_n=0.5$ და გვექნება: $n^*=[0.5\cdot 0.5\cdot(\frac{2.33}{0.03})^2]+1=[1508.16]+1=1509$.

4*. წინა კვლევების თანახმად მკვლევარი თვლის, რომ პირველკურსელ სტუდენტთა წლოვანებების დისპერსია შეადგენს 1.6-ს. შემთხვევით შერჩეული 50 პირველკურსელის წლოვანებების შერჩევითი დისპერსია აღმოჩნდა 2.5. **ა)** $\alpha=0.05$ მნიშვნელოვნების დონით არის თუ არა პირველკურსელ სტუდენტთა წლოვანებების დისპერსია 1.6-ზე მეტი? **ბ)** ააგეთ 99%-იანი ნდობის ინტერვალი პოპულაციის სტანდარტული გადახრისათვის.

ამოხსნა. **ა)** $H_0: D\xi\le 1.6$, $H_1: D\xi>1.6$, $n=50$, $s^2=2.5$, $\alpha=0.05$, $T.V.\equiv\chi^2=\frac{50\cdot 2.5}{1.6}=78.13$, $C.V.=\chi^2_{49,0.05}=55.758$, $C.R.=[\chi^2_{49,0.05},+\infty)=[55.758,+\infty)$. ვინაიდან $T.V.\in C.R.$, ამიტომ პირველკურსელ სტუდენტთა წლოვანებების დისპერსია მეტია 1.6-ზე.

ბ) $\alpha=1-0.99=0.01$, $\alpha/2=0.005$, $1-\alpha/2=0.995$, $s^2=2.5$, $\chi^2_{49,0.005}=78.23$, $\chi^2_{49,0.995}=27.25$. საძიებელი ნდობის ინტერვალი იქნება:

$$(\sqrt{\frac{50\cdot 2.5}{78.23}},\sqrt{\frac{50\cdot 2.5}{27.25}}).$$

5. არის თუ არა შემთხვევით შერჩეული კვირის განმავლობაში პირველი კომპანიის ტაქსების მიერ გავლილი მანძილების საშუალო მეტი ვიდრე მეორე კომპანიის ტაქსების მიერ გავლილი მანძილების საშუალო, თუ შესაბამისი შერჩევების მახასიათებლებია: $\overline{x}_{10}=2837$, $s_1'=30$; $\overline{y}_{11}=2753$, $s_2'=66$. ჩათვალეთ, რომ პოპულაციები ნორმალურია; მნიშვნელოვნების დონედ აიღეთ $\alpha=0.05$. ააგეთ 95%-იანი ნდობის ინტერვალები საშუალოთა სხვაობისათვის და დისპერსიათა შეფარდებისათვის.

მითითება: წინასწარ შეამოწმეთ ჰიპოთეზა დისპერსიების ტოლობის შესახებ.

ამოხსნა. I) ეტაპი: შევამოწმოთ ჰიპოთეზა დისპერსიების შესახებ.

$H_0:\sigma_1^2=\sigma_2^2$, $H_1:\sigma_1^2\ne\sigma_2^2$, $\alpha=0.05$, $\alpha/2=0.025$, $1-\alpha/2=0.975$, $C.V._2=F_{9,10,0.025}=3.779$, $C.V._1=F_{9,10,0.975}=1/F_{10,9,0.025}=1/3.9639=0.252$ (**<u>მიაქციეთ ყურადღება $F_{n,m,1-\alpha}$-ს გამოთვლის წესს:</u>** $F_{n,m,1-\alpha}=1/F_{m,n,\alpha}$), $C.R.=(0,\,0.252]\cup[3.779,+\infty)$, $T.V.\equiv F(9,10)=\frac{30^2}{66^2}=0.21$. ვინაიდან $T.V.\in C.R.$, ამიტომ სამართლიანია ალტერნატიული ჰიპოთეზა, ანუ პოპულაციების უცნობი დისპერსიები განსხვავებულია. შესაბამისად, მეორე ეტაპზე ვისარგებლებთ სატერტვაიტის მეთოდით (**<u>წინააღმდეგ შემთხვევაში გამოცდაზე ისარგებლეთ დისპერსიების გაერთიანებული შეფასების მეთოდით!!!</u>**)

II) ეტაპი: დავუბრუნდეთ ამოცანის ძირითად კითხვას. გვაქვს:

$$H_0: a_1\le a_2,\ H_1: a_1>a_2,\ \alpha=0.05,\ C.V.=t_{[c],0.05},\ \text{სადაც}$$

$$c=\frac{(30^2/10+66^2/11)^2}{(30^2/10)^2/9+(66^2/11)^2/10}=\frac{(90+396)^2}{8100/9+156816/10}=\frac{236196}{16581.6}=14.24.$$

შესაბამისად, გვაქვს: $C.V.=t_{[c],0.05}=t_{14,0.05}=1.761$ და ამიტომ $C.R.=(1.761,+\infty)$.

$$T.V.=\frac{2837-2753}{\sqrt{30^2/10+66^2/11}}=\frac{84}{\sqrt{90+396}}=\frac{84}{20.05}=4.19.$$

ვინაიდან $T.V. \in C.R.$, ამიტომ სამართლიანია ალტერნატიული ჰიპოთეზა, ანუ პირველი პოპულაციის საშუალო მეტია მეორე პოპულაციის საშუალოზე.

ნდობის ინტერვალი საშუალოთა სხვაობისათვის: აქ $t_{[c],\alpha/2} = t_{14,0.025} = 2.145$ და შესაბამისად, გვაქვს:

$$(2837 - 2753 - 2.145 \cdot \sqrt{30^2/10 + 66^2/11},\ 2837 - 2753 + 2.145 \cdot \sqrt{30^2/10 + 66^2/11})$$

$$(84 - 2.145 \cdot 20.05,\ 84 + 2.145 \cdot 20.05)\ ,\quad (84 - 43,\ 84 + 43)\ ,\quad (41, 127)\,.$$

ნდობის ინტერვალი დისპერსიათა შეფარდებისათვის: აქ $F_{m-1,n-1,\alpha/2} = F_{10,9,0.025} = 3.9639$; $F_{m-1,n-1,1-\alpha/2} = F_{10,9,0.975} = 1/F_{9,10,0.025} = 1/3.779 = 0.2646$ და შესაბამისად, გვაქვს:

$$(0.2646 \cdot 30^2/66^2,\ 3.9639 \cdot 30^2/66^2)\,;$$

$$(0.055,\ 0.819)\,.$$

6. სამი ქარხანა უშვებს ერთიდაიგივე დეტალს. ამ ქარხნებიდან აღებულია შესაბამისად 250, 200 და 150 დეტალი, რომელთაგან არასტანდარტული აღმოჩნდა შესაბამისად 10, 9 და 11 დეტალი. შევამოწმოთ ჰიპოთეზა აღნიშნულ შერჩევათა ერთგვაროვნების შესახებ $\alpha = 0.1$ მნიშვნელოვნების დონით.

ამოხსნა. H_0 :პოპულაციები ერთგვაროვანია,

H_1 :ორი პოპულაცია მაინც განსხვავებულია,

$$\alpha = 0.1\,,\ C.V. = \chi^2_{(3-1)\cdot(2-1),0.1} = \chi^2_{2,0.1} = 4.605\,,\ C.R. = (4.605, +\infty)\,.$$

გამოვთვალოთ მოსალოდნელი სიხშირეები:

$$e_{1,1} = 250 \cdot 570/600 = 237.5\,,\ e_{1,2} = 250 \cdot 30/600 = 12.5\,,\ e_{2,1} = 200 \cdot 570/600 = 190\,,$$

$$e_{2,2} = 200 \cdot 30/600 = 10\,,\ e_{3,1} = 150 \cdot 570/600 = 142.5\,,\ e_{3,2} = 150 \cdot 30/600 = 7.5\,,$$

$$T.V. \equiv \chi^2 = \frac{(240 - 237.5)^2}{237.5} + \frac{(10 - 12.5)^2}{12.5} + \frac{(191 - 190)^2}{190} +$$

$$+ \frac{(9 - 10)^2}{10} + \frac{(139 - 142.5)^2}{142.5} + \frac{(11 - 7.5)^2}{7.5} = 2.36\,.$$

რადგანაც $2.36 < 4.605$, ამიტომ ნულოვანი ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს, ანუ პოპულაციები ერთგვაროვანია.

6*. 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. H_0 :ალკოჰოლის რაოდენობა დამოუკიდებელია სქესისაგან,

H_1 : ალკოჰოლის რაოდენობა დამოკიდებელია სქესზე,

$$\alpha = 0.1\,,\ C.V. = \chi^2_{(2-1)\cdot(3-1),0.1} = \chi^2_{2,0.1} = 4.605\,, C.R. = (4.605, +\infty)$$

გამოვთვალოთ მოსალოდნელი სიხშირეები:

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

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

$$T.V. \equiv \chi^2 = \frac{(10 - 9.13)^2}{9.13} + \frac{(9 - 9.93)^2}{9.93} + \frac{(8 - 7.94)^2}{7.94} +$$

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

რადგანაც $0.283 < 4.605$, ამიტომ ნულოვანი ჰიპოთეზის უარყოფის საფუძველი არა გვაქვს, ანუ ალკოჰოლის მოხმარების რაოდენობა არაა დამოკიდებელი სქესზე.

ბონუს ქულები

1*). შეამოწმეთ ხდომილებათა დამოუკიდებლობის ორი განმარტების ეკვივალენტურობა. **2.5 ქულა**

ა) ვთქვათ, $P(A|B)=P(A)$. მაშინ პირობითი ალბათობის განმარტების თანახმად გვაქვს $\frac{P(A\cap B)}{P(B)}=P(A)$, საიდანაც გვაქვს: $P(A\cap B)=P(A)P(B)$.

ბ) ვთქვათ, $P(A\cap B)=P(A)P(B)$. მაშინ პირობითი ალბათობის განმარტება გვაძლევს: $P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{P(A)\cdot P(B)}{P(B)}=P(A)$.

2*). გამოთვალეთ პუასონის განაწილების ლოდინი. **2.5 ქულა**

თუ ვისარგებლებთ წარმოდგენით $e^x=\sum_{m=0}^{\infty}\frac{x^m}{m!}$, მაშინ გვექნება:

$$EPo(\lambda)=\sum_{k=0}^{\infty}k\frac{\lambda^k}{k!}e^{-\lambda}=e^{-\lambda}\sum_{k=1}^{\infty}k\frac{\lambda^k}{k!}=\lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}=\lambda e^{-\lambda}e^{\lambda}=\lambda.$$

3*). ახსენით რატომაა 3-ში ნდობის ინტერვალი ასეთი სახის. **2.5 ქულა**

როგორც ცნობილია, ნორმალური $N(a,\sigma^2)$ პოპულაციისათვის, უცნობი დისპერსიის შემთხვევაში $\frac{\overline{X}-a}{S'/\sqrt{n}}$-ს გააჩნია სტიუდენტის განაწილება თავისუფლების ხარისხით $n-1$ ($\frac{\overline{X}-a}{S'/\sqrt{n}}\cong t(n-1)$). სტიუდენტის განაწილების ზედა $\alpha/2$-კრიტიკული წერტილის განმარტების თანახმად $P\{-t_{n-1,\alpha/2}<\frac{\overline{X}-a}{S'/\sqrt{n}}<t_{n-1,\alpha/2}\}=1-\alpha$. ეს უკანასკნელი ტოლფასია თანაფარდობის $P\{\overline{X}-t_{n-1,\alpha/2}\cdot\frac{S'}{\sqrt{n}}<a<\overline{X}+t_{n-1,\alpha/2}\cdot\frac{S'}{\sqrt{n}}\}=1-\alpha$, რაც იმას ნიშნავს, რომ (1) საძიებელი ნდობის ინტერვალია.

4*). ახსენით რატომაა 4-შიკრიტიკული არე ასეთი სახის. **2.5 ქულა**

H_1 ჰიპოთეზის სამართლიანობის შემთხვევაში $\overline{X}$-ის მნიშვნელობა ახლოს იქნება a_1-თთან, რომელიც მეტია a_0-ზე, ამიტომ Z-ის დაკვირვებული მნიშვნელობა შეიძლება გადაიხაროს ნულიდან მარჯვნივ და ვინაიდან კრიტიკული არე წარმოადგენს ალტერნატიული ჰიპოთეზის მიღების არეს, გასაგებია, რომ ის უნდა იყოს მარჯვენა ცალმხრივი ინტერვალი C.R. = $[x_{კრ},+\infty)$. I გვარის შეცდომის განმარტების თანახმად

$$\alpha=P(H_1|H_0)=P(Z\ge x_{კრ}|H_0)=1-\Phi(x_{კრ}),$$

ანუ $\Phi(x_{კრ})=1-\alpha$, რაც ზედა კრიტიკული წერტილის განმარტების თანახმად ნიშნავს, რომ $x_{კრ}=z_\alpha$, ანუ C.R. = $[z_\alpha,+\infty)$.

შენიშვნა! ა) ჰიპოთეზების შემოწმების პრაქტიკულ ამოცანაში: ბოლომდე უნდა გამოთვალოთ კრიტერიუმის მნიშვნელობა და იპოვოთ კრიტიკული არე და აქედან გამომდინარე გამოიტანოთ დასკვნა.

ბ) შეგიძლიათ ისარგებლოთ ქვემოთმოყვანილი გამოსახულებებით:

$$T_{n,m} = [(\overline{X}_n - \overline{Y}_m) - (a_1 - a_2)]/S'_{n,m}\sqrt{1/n + 1/m} \cong t(n+m-2). \qquad k = \min\{n-1, m-1\}.$$

$$\overline{x}_n - \overline{y}_m \mp z_{\alpha/2}\sqrt{\sigma_1^2/n + \sigma_2^2/m}. \qquad Z = [(\overline{X}_n - \overline{Y}_m) - (a_1 - a_2)]/\sqrt{\sigma_1^2/n + \sigma_2^2/m} \cong N(0,1).$$

$$\overline{x}_n - \overline{y}_m \mp t_{n+m-2,\alpha/2}s'_{n,m}\sqrt{1/n + 1/m}. \qquad n_i = \sum_{j=1}^{r} n_{ij}. \qquad n_{i*} = \sum_{j=1}^{r} n_{ij}. \qquad n_{*j} = \sum_{i=1}^{k} n_{ij}.$$

$$\overline{x}_n - \overline{y}_m \mp t_{k,\alpha/2}\sqrt{s_{1n}'^2/n + s_{2m}'^2/m}. \quad Z = [(\overline{X}_n - \overline{Y}_m) - (a_1 - a_2)]/\sqrt{S_{1n}'^2/n + S_{2m}'^2/m} \overset{as}{\cong} N(0,1).$$

$$T = [(\overline{X}_n - \overline{Y}_m) - (a_1 - a_2)]/\sqrt{S_{1n}'^2/n + S_{2m}'^2/m} \cong t(k). \qquad n\overline{p}_1, n\overline{q}_1, m\overline{p}_2, m\overline{q}_2 \geq 5.$$

$$\overline{x}_n - \overline{y}_m \mp z_{\alpha/2}\sqrt{s_{1n}'^2/n + s_{2m}'^2}. \quad T_{n,m} \equiv (\overline{X}_n - \overline{Y}_m)/\sqrt{S_{1n}'^2/n + S_{2m}'^2/m} \cong t([c]). \qquad \overline{Q} = 1 - \overline{P}.$$

$$(n+m-2)S_{n,m}'^2 = (n-1)S_{1n}'^2 + (m-1)S_{2m}'^2. \qquad \overline{p}_1 - \overline{p}_2 \mp z_{\alpha/2}\sqrt{\overline{p}_1\overline{q}_1/n + \overline{p}_2\overline{q}_2/m}.$$

$$[(s_{1n}'^2/n + s_{2m}'^2/m)^2]/[(s_{1n}'^2/n)^2/(n-1) + (s_{2m}'^2/m)^2/(m-1)]. \qquad \overline{P}_1 = S_1/n, \quad \overline{P}_2 = S_2/m.$$

$$\overline{P} = \frac{n}{n+m}\overline{P}_1 + \frac{m}{n+m}\overline{P}_2. \quad n_{ij}^* = \frac{n_{i*} \cdot n_{*j}}{n}. \quad \chi^2 = \sum_{i=1}^{s}\frac{(n_i - n_i')^2}{n_i'}. \quad F = \frac{S_{1n}'^2/\sigma_1^2}{S_{2m}'^2/\sigma_2^2} \cong F(n-1, m-1).$$

$$\overline{f} = \frac{\max\{s_{1n}'^2, s_{2m}'^2\}}{\min\{s_{1n}'^2, s_{2m}'^2\}}. \quad \hat{\chi}^2 = \sum_{i=1}^{k}\sum_{j=1}^{r}\frac{(n_{ij} - n_{ij}^*)^2}{n_{ij}^*}. \quad \overline{x}_n - \overline{y}_m \mp t_{[c],\alpha/2}\sqrt{s_{1n}'^2/n + s_{2m}'^2/m}.$$

$$Z = [(\overline{P}_1 - \overline{P}_2) - (p_1 - p_2)]/\sqrt{\overline{P}\overline{Q}(1/n + 1/m)} \overset{as}{\cong} N(0,1).$$

$$(F_{m-1,n-1,1-\alpha/2} \cdot s_1'^2/s_2'^2, F_{m-1,n-1,\alpha/2} \cdot s_1'^2/s_2'^2).$$

P.S. ბილეთს დართული ექნება ყველა აუცილებელი ცხრილი.