Probleme recapitulative PS (pentru examenul scris) - soluții

Pentru determinarea cuantilelor din problemele de statistică se folosește programul Octave!

1. Se consideră vectorul aleator discret (U, V) cu distribuția dată sub formă tabelară:

•	U V U	-1	1	b
	1	0,25	0,05	a
	3	0,3	a	0,1

- a) Să se determine constantele reale a și b, știind că E(V) = 0.15.
- b) Sunt variabilele aleatoare U şi V independente?
- c) Să se calculeze valoarea medie a variabilei aleatoare $(U-3)^2$.

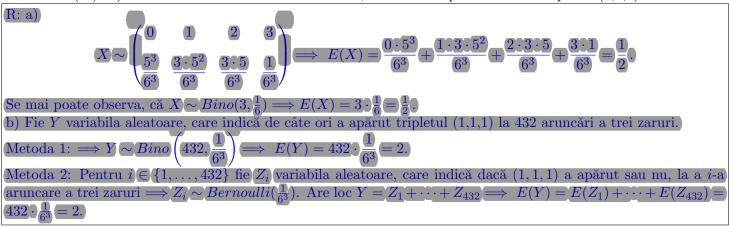
```
R: a) 2a + 0.4 + 0.3 = 1 \Rightarrow a = 0.15; P(V = -1) = 0.55, P(V = 1) = 0.2, P(V = b) = 0.25, E(V) = -0.55 + 0.2 + b \cdot 0.25 = 0.15 \Rightarrow b = 2.

b) P(U = 1) = 0.45, P(V = 1) = 0.2, P(U = 1, V = 1) = 0.05

\Rightarrow P(U = 1) \cdot P(V = 1) \neq P(U = 1, V = 1) pentru că 0.09 \neq 0.05

c) P(U = 1) = 0.45; P(U = 3) = 0.55 \Rightarrow E((U - 3)^2) = 0.45 \cdot 4 = 1.8.
```

2. a) Fie X variabila aleatoare care indică de câte ori a apărut numărul 1 la 3 aruncări ale unui zar. Să se calculeze E(X). b) Dacă se aruncă de 432 ori trei zaruri, de câte ori apare \hat{in} medie tripletul (1,1,1)?



3. Timpul de reacție (în secunde) al unui proces chimic este o variabilă aleatoare, notată cu X, care are funcția de densitate $f: \mathbb{R} \to \mathbb{R}$,

$$f(x) = \begin{cases} \frac{3a^3}{x^4}, & x > a \\ 0, & x \le a \end{cases} \quad \text{unde } a > 0 \text{ este parametru necunoscut.}$$

Să se calculeze $P(X \leq 2a)$ și E(X) în funcție de a.

R:
$$P(X \le 2a) = \int_a^{2a} \frac{3a^3}{x^4} dx = -\frac{a^3}{x^3} \Big|_a^{2a} = \frac{7}{8}$$
; $E(X) = \frac{3a}{2}$.

4. Fie variabilele aleatoare X, Y, care iau valori binare $(0 \le 1)$, astfel încât P(X = 1) = 0.4, iar valoarea variabilele aleatoare Y depinde doar de valoarea variabilele aleatoare X, astfel: P(Y = 1|X = 1) = 0.2; P(Y = 0|X = 0) = 0.3. Determinați: a) P(X = Y); b) E(Y); c) $E(X \cdot Y)$; d) valoarea medie a numărului (aleator) binar $YX_{(2)}$ trecut în baza 10.

```
R: a) P(X = 1) = 0.4 \Longrightarrow P(X = 0) = 0.6; P(X = Y = 1) + P(X = Y = 0) = P(X = 1)P(Y = 1|X = 1) + P(X = 0)P(Y = 0|X = 0) = 0.4 \cdot 0.2 + 0.6 \cdot 0.3 = 0.26. b) E(Y) = 1 \cdot P(Y = 1) + 0 \cdot P(Y = 0) = P(Y = 1) = P(Y = 1|X = 1) \cdot P(X = 1) + P(Y = 1|X = 0) \cdot P(X = 0) = P(Y = 1|X = 1) \cdot P(X = 1) + (1 - P(Y = 0|X = 0)) \cdot P(X = 0) = 0.2 \cdot 0.4 + 0.7 \cdot 0.6 = 0.5. c) Scriem: X \cdot Y = 1 \Longleftrightarrow X = Y = 1; E(X \cdot Y) = 1 \cdot P(X \cdot Y = 1) + 0 \cdot P(X \cdot Y = 0) = P(X = Y = 1) = 0.4 \cdot 0.2 = 0.08. d) E(2^1Y + 2^0X) = 2E(Y) + E(X) = 2P(Y = 1) + P(X = 1) = 2 \cdot 0.5 + 0.4 = 1.4.
```

5. Fie X v.a. care indică timpul de restartare al unui anumit sistem și are funcția de densitate $f_X: \mathbb{R} \to \mathbb{R}$

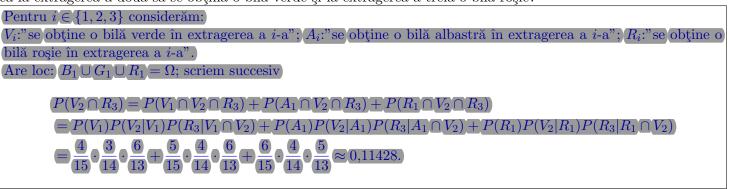
$$f_X(t) = \begin{cases} c (3-t)^2, & 0 < t < 3 \\ 0, & \text{altfel.} \end{cases}$$

- a) Determinați valoarea constantei c.
- b) Determinați funcția de repartiție F_X .
- c) Determinați P(1 < X < 2) și P(X < 2|X > 1).

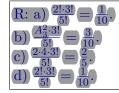
R: a)
$$\int_{-\infty}^{\infty} f_X(t)dt = 1 \Rightarrow c = \frac{1}{9}.$$
b)
$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \begin{cases} 0, & x < 0 \\ \frac{x^3}{27} - \frac{x^2}{3} + x, & 0 \le x < 3 \\ 1, & 3 \le x. \end{cases}$$
c)
$$P(1 \le X \le 2) = F_X(2) - F_X(1) = \frac{7}{27}$$

$$P(X < 2|X > 1) = \frac{P(1 \le X < 2)}{P(X > 1)} = \frac{F_X(2) - F_X(1)}{1 - F_X(1)} = \frac{7}{8}.$$

6. Într-o urnă sunt 4 bile verzi, 5 bile albastre, 6 bile roşii. Se extrag fără returnare 3 bile. Care este probabilitatea ca la extragerea a doua să se obțină o bilă verde și la extragerea a treia o bilă roșie?



- 7. 5 bile numerotate consecutiv de la 1 la 5 sunt așezate orizontal în mod aleator. Determinați:
- a) probabilitatea ca prima și ultima bilă să aibe numere pare;
- b) probabilitatea ca primele două bile să aibe numere impare;
- c) probabilitatea ca bilele cu numere pare să fie alăturate;
- d) probabilitatea să nu fie două bile cu numere de aceași paritate alăturate.



- 8. Un cod de 5 cifre este generat aleator. Care este probabilitatea ca
- a) toate cifrele să fie distincte?
- b) să conțină doar cifre pare distincte?
- c) trei cifre să fie egale și celelalte două cifre să fie egale (de exemplu: 12121, 35355, 44477, etc.)?
- d) să conțină doar cele trei cifre 1, 2, 3 (de exemplu: 12131, 22113, 31312, etc.)?



- 9. Se dau două urne. Prima urnă conține 2 bile negre și 3 bile roșii. A doua urnă conține 3 bile negre și 2 bile roșii. Se aruncă un zar. Dacă se obține un număr par, atunci se extrag două bile din prima urnă, cu repunerea bilei extrase în urnă. Dacă se obține un număr impar, atunci se extrag două bile din a doua urnă, fără repunerea bilei extrase în urnă. Fie X variabila aleatoare care indică numărul de bile roșii extrase. Determinați:
- a) distributia lui X:
- **b)** valoarea medie lui X;
- c) funcția de repartiție a lui X;
- d) valoarea medie a numărului de repetiții independente ale experimentului descris mai sus până la prima repetiție a experimentului în urma căreia se obțin două bile roșii.

R: a) Folosind modelul urnei cu două culori și bilă returnată, respectiv nereturnată (distribuția binomială, respectiv hipergeometrică) și formula probabilității totale, se obține: $P(X=k) = \frac{3}{6} \cdot C_2^k \left(\frac{3}{5}\right)^k \cdot \left(\frac{2}{5}\right)^{2-k} + \frac{3}{6} \cdot \frac{C_2^k C_3^{2-k}}{C_2^2}, k = 0, 1, 2.$

Deci,
$$X \sim \begin{pmatrix} 0 & 1 & 2 \\ \frac{23}{100} & \frac{54}{100} & \frac{23}{100} \end{pmatrix}$$
.
b) $E(X) = 0 \cdot \frac{23}{100} + 1 \cdot \frac{54}{100} + 2 \cdot \frac{23}{100} = 1$.
c) $F_X(x) = P(X \le x) = \begin{cases} 0, & x < 0 \\ \frac{23}{100}, & x \in [0, 1) \\ \frac{23}{100} + \frac{54}{100}, & x \in [1, 2) \end{cases} = \begin{cases} 0, & x < 0 \\ \frac{23}{100}, & x \in [0, 1) \\ \frac{23}{100}, & x \in [1, 2) \end{cases}$
d) Fig. Y =numěrul do rapatitiji independente ale experimentului pôně la pri

- d) Fie Y=numărul de repetiții independente ale experimentului până la prima repetiție în urma căreia se obțin două bile roşii. Atunci $Y \sim Geo(p)$, unde $p = P(X = 2) = \frac{23}{100}$. Avem (rezultat problemă de la seminar): $E(Y) = \frac{1-p}{p} = \frac{77}{23} \approx 3,35.$
- 10. Icsulescu face naveta cu microbusul. El ajunge în fiecare zi în autogară la ora 16:30 cu o întârziere T (în minute) care are funcția de repartiție $F_T(t) = \begin{cases} 0, & t < 0 \\ \frac{t}{10}, & t \in [0, 10]. \text{ Microbusul pornește la ora 16:39. Determinați:} \\ 1, & t > 10 \end{cases}$
- a) probabilitatea ca Icsulescu să piardă microbusul;
- b) valoarea medie a întârzierii lui Icsulescu;
- c) probabilitatea ca în 5 zile (cu întârzieri independente) Icsulescu să piardă microbuzul în cel puțin două zile.
- d) Fie Y numărul de zile succesive (cu întârzieri independente) când Icsulescu prinde microbusul până la prima zi când pierde microbusul. Ce distribuție are Y?

R: a)
$$P(T \ge 9) = 1 - F(9) = 1 - \frac{9}{10} = \frac{1}{10}$$
. b) $f_T(t) = \begin{cases} \frac{1}{10}, & t \in [0, 10] \\ 0, & t \notin [0, 10] \end{cases}$. $E(T) = \int_0^{10} \frac{t}{10} dt = 5$. c) Z =numărul de zile din 5 când Icsulescu pierde microbusul $\Longrightarrow Z \sim Bino(5, p)$. $P(Z \ge 2) = 1 - P(Z = 0) - P(Z = 1) = 1 - C_5^0 p^0 (1 - p)^5 - C_5^1 p^1 (1 - p)^4 = 1 - \frac{9^5}{10^5} - 5 \cdot \frac{1}{10} \cdot \frac{9^4}{10^4} = 0.08146 \approx 8\%$. d) $Y \sim Geo(p), p = P(T \ge 9) = \frac{1}{10}$.

- **11.** Se alege uniform aleator un punct în dreptunghiul $[1,2] \times [2,4] \subset \mathbb{R}^2$. Să se calculeze: a) probabilitatea ca punctul să fie în dreptunghiul $\left[\frac{4}{3},\frac{5}{3}\right] \times \left[\frac{5}{2},\frac{7}{2}\right]$;
- b) valoarea medie a pătratului distanței de la punctul ales la origine.

Fie
$$(X,Y)$$
 coordonatele punctului ales. Atunci (X,Y) are distribuţia uniformă pe dreptunghiul $[1,2] \times [2,4]$, adică are funcţia de densitate: $f_{(X,Y)}: \mathbb{R}^2 \to \mathbb{R}, f_{(X,Y)}(x,y) = \begin{cases} \frac{1}{(2-1)(4-2)}, & (x,y) \in [1,2] \times [2,4] \\ 0, & (x,y) \notin [1,2] \times [2,4] \end{cases} = \begin{cases} \frac{1}{2}, & (x,y) \in [1,2] \times [2,4] \\ 0, & (x,y) \notin [1,2] \times [2,4] \end{cases}$

$$\mathbb{R}: X \text{ are funcţia de densitate } f_X(x) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,y) dy = \begin{cases} 1, & x \in [1,2] \\ 0, & x \notin [1,2] \end{cases}$$

$$\mathbb{R}: X \text{ are funcţia de densitate } f_X(x) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,y) dy = \begin{cases} 1, & x \in [1,2] \\ 0, & x \notin [1,2] \end{cases}$$

$$\mathbb{R}: X \text{ are funcţia de densitate } f_X(x) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,y) dy = \begin{cases} 1, & x \in [1,2] \\ 0, & x \notin [1,2] \end{cases}$$

 $\int_{-\infty}^{\infty} f_{(X,Y)}(x,y)dx = \begin{cases} \frac{1}{2}, & y \in [2,4] \\ 0, & y \notin [2,4] \end{cases}; \text{ adică } X \sim Unif[1,2] \text{ si } Y \sim Unif[2,4].$ $\text{a) } P\left((X,Y) \in \left[\frac{4}{3}, \frac{5}{3}\right] \times \left[\frac{5}{2}, \frac{7}{2}\right]\right) = \iint_{\left[\frac{4}{3}, \frac{5}{3}\right] \times \left[\frac{5}{2}, \frac{7}{2}\right]} f_{(X,Y)}(x,y) \, dxdy = \iint_{\left[\frac{4}{3}, \frac{5}{3}\right] \times \left[\frac{5}{2}, \frac{7}{2}\right]} \frac{1}{2} \, dxdy = \frac{1}{6}.$ $b) E(X^{2} + Y^{2}) = E(X^{2}) + E(Y^{2}) = \int_{1}^{13} x^{2} \frac{1}{2 - 1} dx + \int_{2}^{4} y^{2} \frac{1}{4 - 2} dy = \frac{x^{3}}{3} \Big|_{1}^{2} + \frac{y^{3}}{3 \cdot 2} \Big|_{2}^{4} = \frac{7}{3} + \frac{28}{3} = \frac{35}{3}$

12. Fie $N \sim Unid(5)$. Se generează apoi un cod binar cu N cifre, în care probabilitatea de apariție a lui 0,

respectiv 1, este egală cu 0,5 (de exemplu: N=2, X=01; N=3, X=010; N=1, X=1). Să se calculeze:

- a) P(X = 1011|N = 4);
- b) probabilitatea ca suma S, a cifrelor lui X, să fie egală cu 3.

R: a)
$$P(X = 1011|N = 4) = \frac{1}{2^4}$$
.
b) $P(S = 3) = \sum_{i=1}^{5} P(S = 3|N = i)P(N = i) = \frac{1}{5} \sum_{i=3}^{5} \frac{C_i^3}{2^i}$.

13. Departamentul de resurse umane al unei companii mari suspectează că angajații iau în medie pauza de prânz (care cf. contractului durează maxim 1 oră) mult mai lungă. Pe baza unui eșantion format din 36 de angajați s-a obținut o medie de 70 minute și o abatere standard de 15 minute pentru durata pauzei. Formulați ipoteza nulă și ipoteza alternativă și testați cu un nivelul de semnificație de 5% dacă în medie durata pauzei este respectată sau este depășită. Construiți un interval de încredere bilateral pentru abaterea standard a duratei pauzei de prânz. Se presupune că durata pauzei este o variabilă aleatoare normal distribuită.

R: Test pentru medie când varianța este necunoscută (testul Student); $m_0 = 60, n = 36; \bar{x}_n = 70, s_n = 15, \alpha = 0.05, t_{1-\alpha} = \texttt{tinv}(0.95, 35) = 1.69,$

$$0.05, t_{1-\alpha} = \text{tinv}(0.95, 35) = 1.69,$$

$$H_0: m \le 60, \ H_1: m > 60, \ t = \frac{\bar{x}_n - m_0}{\frac{s_n}{\sqrt{n}}} = 4 > t_{1-\alpha} = 1.69$$

 $\Longrightarrow H_0$ este respinsă \Longrightarrow în medie durata pauzei nu este respectată.

Intervalul de încredere bilateral pentru abaterea standard a duratei pauzei de prânz este

$$\left(\sqrt{\frac{n-1}{c_{1-\frac{\alpha}{2}}}} \cdot s_n, \sqrt{\frac{n-1}{c_{\frac{\alpha}{2}}}} \cdot s_n\right) = \left(\sqrt{\frac{35}{53.20}} \cdot 15, \sqrt{\frac{35}{20.57}} \cdot 15\right) \approx (12.17, 19.57).$$

14. Fie $x_1, \ldots, x_{10} \in (0,1)$ date statistice pentru caracteristica X, a cărei funcție de densitate este $f: \mathbb{R} \to \mathbb{R}$

$$f(x) = \begin{cases} 2\theta x^{2\theta - 1}, \text{dacă } 0 < x \le 1\\ 0, \text{dacă } x \notin (0, 1] \end{cases},$$

unde $\theta > 0$ este parametru necunoscut. Să se estimeze θ cu ajutorul a) metodei verosimilității maxime; b) metodei momentelor.

R: a) Scriem funcția de verosimilitate

$$L(x_1, ..., x_{10}; \theta) = f(x_1) \cdot ... \cdot f(x_{10}) = 2^{10} \theta^{10} (x_1 \cdot ... \cdot x_{10})^{2\theta - 1}$$

Proprietățile funcției ln: $\ln(a \cdot b) = \ln(a) \cdot \ln(b)$, $\ln(\frac{a}{b}) = \frac{\ln(a)}{\ln(b)}$, $\ln(a^b) = b \ln(a)$, a, b > 0, $\ln(e) = 1$, $\ln(1) = 0$.

$$\Rightarrow \ln L(x_1, ..., x_{10}; \theta) = 10 \ln(2) + 10 \ln(\theta) + (2\theta - 1) \ln(x_1 \cdot ... \cdot x_{10})$$

$$\Rightarrow \frac{\partial \ln L}{\partial \theta} = \frac{10}{\theta} + 2\ln(x_1 \cdot \dots \cdot x_{10}) \Rightarrow \frac{\partial^2 \ln L}{\partial \theta^2} = -\frac{10}{\theta^2} < 0$$

$$\Rightarrow \hat{\theta}(x_1, ..., x_{10}) = -\frac{10}{2\ln(x_1 \cdot ... \cdot x_{10})} = -\frac{5}{\ln(x_1 \cdot ... \cdot x_{10})}.$$

b) Calculăm $E(X) = \frac{2\theta}{2\theta+1}$. Scriem $\frac{2\theta}{2\theta+1} = \frac{1}{10}(x_1 + ... + x_{10})$, unde $x_1, ..., x_{10}$ sunt valorile de selecție pen-

tru caracteristica X. Obţinem estimatorul $\hat{\theta}(X_1,...,X_{10}) = \frac{\bar{X}_{10}}{2(1-\bar{X}_{10})}$ şi valoarea estimată pentru parametrul

necunoscut θ : $\hat{\theta}(x_1, ..., x_{10}) = \frac{\bar{x}_{10}}{2(1 - \bar{x}_{10})}$.

15. Fie $X_1,...,X_n,...$ variabile aleatoare independente, care au aceeași distribuție

$$P(X_i = -1) = P(X_i = 1) = 0.5$$
 pentru fiecare $i \in \mathbb{N}^*$.

Fie $Y_i = \max\{X_i, X_{i+1}\}$ pentru $1 \le i \le n - 1$.

- (a) Să se determine distribuția de probabilitate pentru Y_i , $1 \le i \le n-1$.
- (b) Să se calculeze valoarea medie și varianța pentru Y_i , $1 \le i \le n-1$.
- (c) Fie $Z_n = \frac{1}{n}(X_1^5 + X_2^5 + ... + X_n^5), n \in \mathbb{N}^*$. Spre ce valoare converge aproape sigur şirul $(Z_n)_n$?

R: (a) X_i și X_{i+1} sunt variabile aleatoare independente

$$\Rightarrow P(Y_i = -1) = P(\max\{X_i, X_{i+1}\} = -1) = P(X_i = -1, X_{i+1} = -1) = P(X_i = -1)P(X_{i+1} = -1) = \frac{1}{4}$$

$$\Rightarrow P(Y_i = 1) = \frac{3}{4}.$$

$$\Rightarrow Y_i \sim \begin{pmatrix} -1 & 1\\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}, \ \forall \ i = \overline{1, n-1}.$$

- (b) $E(Y_i) = \frac{1}{2}, V(Y_i) = \frac{3}{4}$.
- (c) Se folosețe legea numerelor mari pentru șirul $(X_n^5)_n$

$$\implies Z_n = \frac{1}{n}(X_1^5 + X_2^5 + \dots + X_n^5) \xrightarrow{a.s.} E(X_1^5) = 0.$$

- 16. Timpul de servire a unui client la un anumit ghișeu la o bancă durează în medie 10 minute și poate fi descrisă de o variabilă aleatoare exponențială.
 - (a) Cu ce probabilitate servirea unui client durează cel mult un sfert de oră?
 - (b) În medie cât ar trebui să dureze timpul de servire a unui client, dacă probabilitatea ca timpul de servire a unui client să fie mai mare decât un sfert de oră este egală cu 0.1?

Indicație: Funcția de densitate a distribuției exponențiale $X \sim Exp(\lambda)$ este

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & : x > 0\\ 0 & : x \le 0. \end{cases}$$

R: (a) Calculăm valoarea medie $E(X) = \frac{1}{\lambda}$ și din E(X) = 10 rezultă $\lambda = \frac{1}{10}$. Are loc $P(X \le 15) = \sqrt{\int_0^{15} \lambda e^{-\lambda t} dt} = \frac{1}{\lambda}$ $1 - e^{-\lambda \cdot 15} = 1 - e^{-1.5}$. (b) $E(X) = \frac{1}{\lambda} = ?$ astfel încât P(X > 15) = 0.1. Scriem $e^{-\lambda \cdot 15} = 0.1 \Longrightarrow E(X) = \frac{1}{\lambda} = \frac{15}{\ln(10)}$

17. Se studiază caracteristica: timpul de producție a unui anumit produs. Această caracteristică se presupune a fi normal distribuită și s-a realizat un eșantion cu 25 de valori (în minute). Pe baza acestora s-a obținut o valoare medie de 8.1 minute și o abatere standard de 1.6 minute. Cu un nivel de semnificație $\alpha = 0.05$, testați dacă se poate afirma: a) abaterea standard a timpului de prelucrare este 1.5 minute; b) media timpului de prelucrare este 7.8 minute.

R: a)
$$H_0$$
: $\sigma = 1.5$, H_1 : $\sigma \neq 1.5$
$$n = 25, \sigma_0 = 1.5, s_{25} = 1.6, \alpha = 0.05, q = \frac{n-1}{\sigma_0^2} \cdot s_n^2 = \frac{24}{1.5^2} \cdot 1.6^2 = 21.094, q_{0.975} = \text{chi2inv}(0.975, 24) = 39.3641, q_{0.025} = \text{chi2inv}(0.025, 24) = 12.4012 \Longrightarrow H_0$$
 se acceptă, adică abaterea standard a timpului de prelucrare este 1.5 minute.

b)
$$H_0$$
: $m = 7.8$, H_1 : $m \neq 7.8$
 $n = 25$, $\sigma_0 = 7.8$, $\bar{x}_{25} = 8.1$, $s_{25} = 1.6$, $\alpha = 0.05$, $t = \frac{\bar{x}_n - m_0}{\frac{s_n}{\sqrt{n}}} = 0.93750$, $t_{0.975} = \text{tinv}(0.975, 24) = 2.0639 \Longrightarrow H_0$

se acceptă, adică media timpului de prelucrare este 7.8 minute.