Soluţii la examenul de Statistică - semianul B, iunie 2017

1. (3 p) Care dintre următoarele statistici este estimator nedeplasat şi pentru ce parametru al populaţiei?

☒ media de selecţie	☐ prima cvartilă	☐ deviaţia standard a eşantionului
☐ deviaţia standard a populaţiei	☒ media populaţiei	☐ dispersia populaţiei

2. (6 p) Zece studenţi obţin la *Probabilităţi şi statistică* următoarele credite: $4, 3, 4, 2, 2, 4, 4, 0, 3, 3$.
(a) (3p) Calculaţi media şi modul. (b) (3p) Există valori aberante (folosind regula 1.5 IQR)?

(a) Media $(4+3+4+2+2+4+4+0+3+3)/10 = 2.9$, modul $= 4$.
(b) Valorile ordonate sunt: $0, 2, 2, 3, 3, 3, 4, 4, 4, 4$. $Q_1 = 2, Q_3 = 4, IQR = 4 - 2 = 2$. Valorile care nu sunt aberante se găsesc în intervalul: $[-1, 7]$. Astfel, nu există valori aberante.

3. (6 p) Se consideră o variabilă aleatoare continuă cu funcţia de densitate de probabilitate

$$f(x) = \begin{cases} \alpha x(2-x), & x \in [0,2] \\ 0, & \text{altfel} \end{cases}$$

. Determinaţi $\alpha \in \mathbb{R}$, funcţia de repartiţie a lui X şi $P(1 \leqslant X \leqslant 2)$.

$1 = \int_{-\infty}^{+\infty} f(x)\,dx = \int_0^2 f(x)\,dx = \alpha\left(x^2 - \frac{x^3}{3}\right)\Big|_0^2 = \frac{4\alpha}{3}, \alpha = 3/4$. $F(a) = \begin{cases} 0, & a < 0 \\ \frac{3a^2 - a^3}{4}, & a \in [0,2] \\ 1, & a > 2 \end{cases}$, $P(1 \leqslant X \leqslant 2) = \int_1^2 f(x)\,dx = F(2) - F(1) = 1 - \frac{1}{2} = \frac{1}{2}$.

4. (5 p) Se consideră un algoritm aleator Monte Carlo, A, pentru rezolvarea unei probleme de decizie P (pentru o intrare x, algoritmul returnează $A(x)$, în timp ce răspunsul corect este $P(x)$). Se ştie că probabilitatea ca algoritmul să greşească este cel mult $1/2$. De câte ori trebuie rulat algoritmul, în mod independent, astfel ca probabilitatea de a greşi să fie cel mult $1/10$?

$P(A_k(x) \neq P(x)) = [P(A(x) =\neq P(x))]^k \leqslant \frac{1}{2^k} \leqslant \frac{1}{10} \Leftrightarrow k \geqslant [\log_2 10] \Leftrightarrow k \geqslant 4$.

5. (12 p) *Handedness Research Institute* susţine că 4 din 5 americani adulţi folosesc mâna dreaptă. Pentru a verifica această afirmaţie se alege un eşantion (aleator simplu) reprezentativ la nivel naţional format din 900 de americani adulţi. Dintre aceştia 154 sunt stângaci.
(a) (8p) În ce mod ar putea fi atacată afirmaţia institutului? Formulaţi ipotezele statistice şi întreprindeţi un test corespunzător de semnificaţie ($\alpha = 1\%$). $-\text{qnorm}(0.005) = 2.576$, $\text{qt}(0.01, 899) = -2.330$, $-\text{qnorm}(0.01) = -2.326$

Se aplică testul Z al proporţiilor $H_0 : p = 0.2 \quad H_a : p < 0.2$, scorul este $z = \frac{154/900 - 0.2}{\sqrt{0.2 \cdot 0.8/900}} = -\frac{26}{12} = -2.166$, valoarea critică este $z^* = \text{qnorm}(0.01) = -2.326$. Deoarece $z > z^*$ ipoteza nulă nu se poate respinge.

((b) (4p) Probabilitatea de a nu respinge afirmaţia institutului atunci când aceasta este de fapt falsă este

☐ puterea testului ☒ probabilitatea erorii de tip II ☐ probabilitatea erorii de tip I

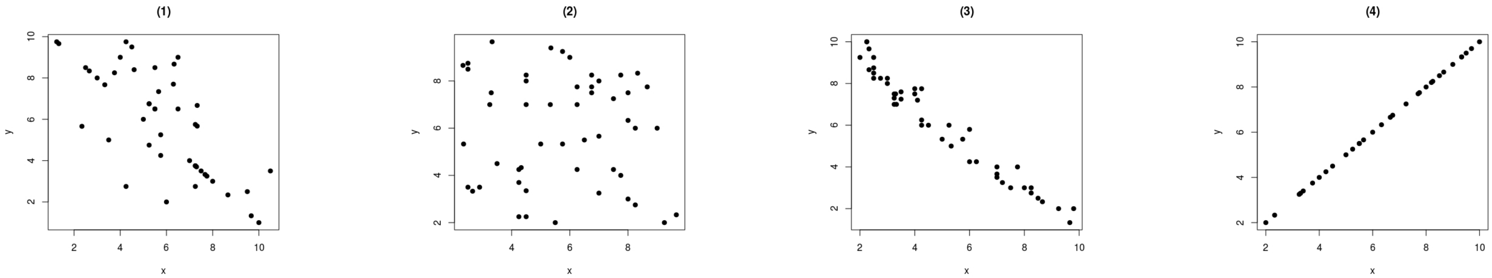
6. (5 p) Pentru o populaţie distribuită normal cu deviaţia standard $\sigma = 3$ se doreşte găsirea unui interval de încredere de 95% pentru media necunoscută a populaţiei. Cât de mare trebuie să fie dimensiunea eşantionului pentru ca intervalul rezultat să aibă o lungime de cel mult 0.98? ($-\text{qnorm}(0.05) = 1.645$, $-\text{qnorm}(0.005) = 2.576$, $\text{qnorm}(0.025) = -1.960$)

Pentru o lungime a intervalului de $w = 0.98$ avem $n = \left[(2z^*\sigma)^2/w^2\right] = \left[(2 \cdot 1.960 \cdot 3)^2/(0.98)^2\right] = 144$.

7. (10 p) O companie de asigurări auto susţine că viaţa medie a unui autoturism este $\mu = 6$ ani cu deviaţia standard $\sigma = 1.1$ ani. Pentru un eşantion aleator simplu de 25 autoturisme se găseşte o medie de $\overline{x}_{25} = 6.5$ ani. Întreprindeţi un test de semnificaţie corespunzător acestor date (1%). (Se ştie că viaţa unui autoturism, în ani, urmează o lege normală.)
$\text{qt}(0.01, 15) = -2.602$, $-\text{qt}(0.05, 15) = 1.753$, $\text{qnorm}(0.01) = -2.326$, $-\text{qt}(0.005, 15) = 2.947$

Se aplică testul Z pentru media unei populaţii cu dispersia cunoscută: $H_0 : \mu = 6 \quad H_a : \mu > 6$, scorul este $z = \frac{6.5 - 6}{1.1/\sqrt{25}} = 2.272$, iar valoarea critică este $z^* = \text{qnorm}(0.99) = -\text{qnorm}(0.01) = 2.326$. Deoarece $z < z^*$ ipoteza nulă nu se poate respinge.

8. (8 p) Pentru diagramele de mai jos coeficienţii de corelaţie sunt $1, -0.975, -0.722$ şi -0.084
(a) (4p) Găsiţi coeficientul corespunzător fiecărei diagrame în parte: (1): -0.722 (2): -0.084 (3): -0.975 (4): 1
(b) (4p) Care dintre diagrame corespunde asocierii pozitive perfecte şi care este dependenţa liniară a acestei asocieri?
Diagrama (4) corespunde unei asocieri liniare perfecte şi anume $X = Y$.


(1)
(2)
(3)
(4)
x
y