

## Tema 3

### Exercițiul 1

- a) Nivelul de zgomot al unei mașini de spălat este o v.a. de medie 44 dB și de abatere standard 5 dB. Admițând aproximarea normală care este probabilitatea să găsim o medie a zgomotului superioară la 48 dB într-un eșantion de talie 10 mașini de spălat ?
- b) O telecabină are o capacitate de 100 de persoane. Știind că greutatea populației (țării) este o v.a. de medie 66.3 Kg și o abatere standard de 15.6 Kg și presupunând că persoanele care au urcat în telecabină au fost alese în mod aleator din populație, care este probabilitatea ca greutatea totală acestora să depășească 7000 Kg ?

### Exercițiul 2

Fie  $X_1, \dots, X_n$  un eșantion de talie  $n$  dintr-o populație de medie  $\mu$  și varianță  $\sigma^2$ . Arătați că varianța varianței eșantionului este:

$$\mathbb{V}(S^2) = \frac{1}{n} \left( \mu_4 - \frac{n-3}{n-1} \sigma^4 \right)$$

unde  $\mu_4 = \mathbb{E}[(X_i - \mu)^4]$  este momentul centrat de ordin 4. Ce revine această formulă în cazul Gaussian (normal) ?

### Exercițiul 3

Fie  $X_1, \dots, X_n$  un eșantion de talie  $n$  dintr-o populație de medie  $\mu$  și varianță  $\sigma^2$ . Arătați că

$$\text{Cov}(\bar{X}, S^2) = \frac{\mu_3}{n}$$

unde  $\mu_3 = \mathbb{E}[(X_i - \mu)^3]$  este momentul centrat de ordin 3. Acest rezultat ne arată că cele două statistici sunt asimptotic *necorelate*.

### Exercițiul 4

Fie  $X_1, X_2, \dots, X_n$  un eșantion de talie  $n$  dintr-o populație  $F$  cu  $\mathbb{E}[X_1^2] < \infty$ .

1. a) Arătați că  $\mathbb{E}[X_1] = \arg \min_{t \in \mathbb{R}} \mathbb{E}[(X_1 - t)^2]$

b) Determinați  $\arg \min_{t \in \mathbb{R}} \sum_{i=1}^n \frac{(X_i - t)^2}{n}$

2. Notăm cu  $x_{\frac{1}{2}} = F^{-1}(\frac{1}{2})$  mediana repartiției lui  $X_1$

- a) Arătați că dacă  $F$  este continuă pe  $\mathbb{R}$  și strict crescătoare pe o vecinătate a lui  $x_{\frac{1}{2}}$  atunci

$$x_{\frac{1}{2}} = \arg \min_{t \in \mathbb{R}} \mathbb{E}[|X_1 - t|]$$

- b) Determinați, în funcție de paritatea lui  $n$ ,  $\arg \min_{t \in \mathbb{R}} \sum_{i=1}^n \frac{|X_i - t|}{n}$ .

## Exercițiul 5

$X_1, \dots, X_n$  un eșantion de talie  $n$  cu funcția de repartiție  $F(x)$  și densitatea  $f(x)$  și  $(Y_1, \dots, Y_n)$  versiunea ordonată crescător a acestuia. Notăm cu  $H_k(x)$  și  $h_k(x)$  funcția de repartiție și densitatea v.a.  $Y_k$ . Fie  $Y_1 = \inf X_i$  și  $Y_n = \sup X_i$ .

- Care este funcția de repartiție și densitatea lui  $Y_1$  și  $Y_n$  ?
- Care este probabilitatea ca o observație dintr-o v.a. de lege  $\mathcal{N}(\mu, \sigma^2)$  să depășească  $\mu + 3\sigma$  ?
- Dar într-un eșantion de talie 100 cat este această probabilitate (i.e. probabilitatea ca o observație să depășească  $\mu + 3\sigma$ )?
- Dintr-un eșantion de talie 100 dintr-o populație repartizată  $\mathcal{N}(0, 1)$  ce valoare nu poate fi depășită cu o probabilitate de 99% ?
- O societate de analiză a calității apei și a mediului efectuează un sondaj în laboratoarele sale (50 la număr, repartizate pe tot teritoriul României) pentru a testa dacă efectuează măsurători corecte. Pentru aceasta, serviciul de calitate trimite la fiecare laborator un eșantion de apă care conține o anumită concentrație de crom și le cere să determine această concentrație de crom. Ținând cont de fluctuațiile care apar în prepararea soluției, precum și de imprecizia aparatelor de măsură, societatea presupune că repartiția concentrației de crom (mg/l) este  $\mathcal{N}(10, 1)$ .

Printre rezultatele obținute de la laboratoare, două dintre acestea au înregistrat măsurători mai diferite decât celelalte: laboratorul  $L_1$  a înregistrat o concentrație de 6 mg/l (cea mai mică valoare înregistrată) iar laboratorul  $L_2$  a măsurat o concentrație de 13 mg/l (cea mai mare dintre măsurători).

Puteți spune, cu o probabilitate de 99%, că aceste valori sunt coerente sau că valorile obținute sunt aberante (datorită erorilor de măsurare, de calibrare a aparatelor, etc.) ?

## Exercițiul 6

Fie  $X_1, X_2, \dots, X_n$  un eșantion de talie  $n$  dintr-o populație  $\mathcal{U}([0, \theta])$  cu  $\theta > 0$  necunoscut.

- Fie  $\hat{\theta}_n = \max \{X_1, \dots, X_n\}$ . Determinați funcția de repartiție a lui  $\hat{\theta}_n$ .
- Arătați că  $\hat{\theta}_n$  este un estimator consistent pentru  $\theta$ .
- Arătați că  $\hat{\theta}_n$  nu este un estimator nedepășat pentru  $\theta$  și construiți un asemenea estimator.

## Exercițiul 7

Fie  $X \sim B(10, \theta)$  cu  $\theta \in (0, 1)$  necunoscut. Fie  $\hat{\theta}_1 = \frac{X}{10}$  și  $\hat{\theta}_2 = \frac{X+1}{12}$  doi estimatori pentru  $\theta$ .

- Calculați  $\mathbb{E}_\theta[\hat{\theta}_1]$  și  $\mathbb{E}_\theta[\hat{\theta}_2]$ .
- Calculați erorile medii pătratice:  $MSE_\theta(\hat{\theta}_1)$  și  $MSE_\theta(\hat{\theta}_2)$ .
- Trasați pe același grafic erorile medii pătratice ale celor doi estimatori ca funcții de  $\theta$ . Pe care dintre cei doi estimatori îl preferați?