

Seminarul 7

1. i) Estimați parametrul necunoscut $p \in (0, 1)$ pentru distribuția binomială a unei caracteristici cercetate: $X \sim Bino(N, p)$, unde $N \in \mathbb{N}^*$ este cunoscut, cu metoda momentelor, respectiv metoda verosimilității maxime. Sunt estimatorii obținuți nedeplasați, respectiv consistenti?

ii) Într-o urnă sunt bile albe și negre. Proporția de bile albe $p \in (0, 1)$ este necunoscută. În urma a $n = 6$ serii a către $N = 5$ extrageri cu returnarea bilei extrase în urnă s-au obținut: 3, 4, 2, 0, 2, respectiv 1, bile albe. Estimați valoarea lui p cu metoda momentelor, respectiv metoda verosimilității maxime.

2. i) O caracteristică cercetată X are funcția de densitate

$$f_X(x) = \begin{cases} \lambda^2 x e^{-\lambda x}, & x > 0, \\ 0, & x \leq 0, \end{cases}$$

unde $\lambda > 0$ este fixat. Estimați parametrul necunoscut λ cu metoda momentelor, respectiv metoda verosimilității maxime. Sunt estimatorii obținuți consistenti?

ii) Durata culorii roșii (în minute) X a unui anumit semafor are funcția de densitate f_X dată mai sus, cu parametrul $\lambda > 0$ necunoscut. Un taximetrist (curios din fire) a observat următoarele dure (în minute) ale culorii roșii pentru acest semafor: $1, \frac{3}{2}, 3, 2, 3, \frac{5}{2}, 1, 2$. Aplicați metoda momentelor, respectiv metoda verosimilității maxime, pentru a estima valoarea lui λ , folosind datele furnizate de taximetrist.

3. a) Un specialist IT instalează un anumit soft nou pe 64 de calculatoare și s-a constatat că: timpul mediu de instalare este de 25 de minute, iar abaterea standard empirică este de 5 minute. Calculați un interval de încredere de 95% pentru media timpului de instalare.

b) Se știe că timpul de instalare (în minute) al softului pe un calculator urmează distribuția $N(20, 8)$. Dacă un specialist instalează softul (independent) pe două calculatoare, să se aproximeze probabilitatea ca timpul total de instalare să se încadreze în intervalul [36, 44].

```
[1]: from scipy.stats import norm, t
```

```
[2]: round(norm.ppf([0.025, 0.05, 0.95, 0.975]), decimals=3)
```

```
[2]: array([-1.96, -1.645, 1.645, 1.96])
```

```
[3]: round(t.ppf([0.025, 0.05, 0.95, 0.975], 63), decimals=3)
```

```
[3]: array([-1.998, -1.669, 1.669, 1.998])
```

```
[2]: round(norm.cdf([-3, -2, -1, 0, 1, 2, 3]), decimals=3)
```

```
[2]: array([0.001, 0.023, 0.159, 0.841, 0.977, 0.999])
```

4. Un statistician studiază caracteristica X a unei persoane dintr-o populație dată: dacă o persoană aleasă aleator preferă ciocolata neagră, atunci X ia valoarea 1; dacă o persoană aleasă aleator preferă ciocolata albă, atunci X ia valoarea 0. Fie \bar{X}_{100} media de selecție a lui X pentru 100 de variabile de selecție. Folosind TLC, estimați $P(0,49 < \bar{X}_{100} < 0,51)$, știind că proporția de persoane care preferă ciocolata neagră este 50% din populația dată.

```
[1]: from scipy.stats import norm
```

```
[2]: round(norm.cdf([-1/3, -1/5, 0, 1/3, 1/5]), decimals=3)
```

```
[2]: array([0.37, 0.42, 0.5, 0.58, 0.63])
```

Pentru rezolvarea următoarelor probleme de statistică, puteți folosiți următoarele valori numerice calculate în Python:

```
[1]: from scipy.stats import norm, t, chi2
```

```
[2]: round(norm.ppf([0.025, 0.05, 0.95, 0.975]), decimals=2)
```

```
[2]: array([-1.96, -1.64, 1.64, 1.96])
```

```
[3]: round(t.ppf([0.025, 0.05, 0.95, 0.975], 4), decimals=2),
round(t.ppf([0.025, 0.05, 0.95, 0.975], 99), decimals=2)
```

```
[3]: (array([-2.78, -2.13, 2.13, 2.78]),
array([-1.98, -1.66, 1.66, 1.98]))
```

```
[4]: round(chi2.ppf([0.025, 0.05, 0.95, 0.975], 2), decimals=2),
round(chi2.ppf([0.025, 0.05, 0.95, 0.975], 4), decimals=2)
```

```
[4]: (array([0.05, 0.1, 5.99, 7.38]),
array([0.48, 0.71, 9.49, 11.14]))
```

5. Considerăm următoarele date statistice pentru masa corporală a persoanelor dintr-o anumită populație: 71 kg; 68 kg; 77 kg; 69 kg; 65 kg. Presupunem că masa corporală este o caracteristică ce urmează distribuția normală.

- a) Știind că varianța/dispersia masei corporale este 20, determinați un interval de încredere bilateral cu nivelul de încredere 95% pentru media masei corporale.
- b) Știind că varianța/dispersia masei corporale este necunoscută, determinați un interval de încredere bilateral cu nivelul de încredere 95% pentru media masei corporale.
- c) Determinați un interval de încredere bilateral cu nivelul de încredere 95% pentru varianța masei corporale.

6. Un provider de internet își asigură clienții că viteza conexiunii la internet este în medie 250 Mbps între orele 20:00 și 22:00. Pe de altă parte, providerul susține că în acest interval orar conexiunea nu este stabilă, având o abatere standard de 40 Mbps. În urma unei selecții de 100 de clienți s-a constatat că valoarea medie de selecție este 242 Mbps pentru viteza conexiunii între orele specificate. Să se construiască un interval de încredere unilateral stâng cu nivelul de încredere 95% pentru media vitezei conexiunii.

7. O companie dorește înlocuirea unui sistem de frânare pentru un anumit tip de mașină cu unul nou, care să reducă semnificativ distanța de frânare. Media distanței de frânare pentru vechiul sistem este mai mare sau egală decât 50 m, pentru o viteză de 80 km/h pe ploaie. În urma testării a 100 de mașini cu noul sistem de frânare instalat, pentru o viteză de 80 km/h pe ploaie, s-a constatat că valoarea medie de selecție este 49 m și că valoarea abaterii standard de selecție este 1 m pentru distanța de frânare a acestui eșantion. Să se construiască un interval de încredere unilateral drept cu nivelul de încredere 95% pentru media distanței de frânare a noului sistem.