Curs 1 - Statistică

Notare → Proiect (30%)
(potențial)

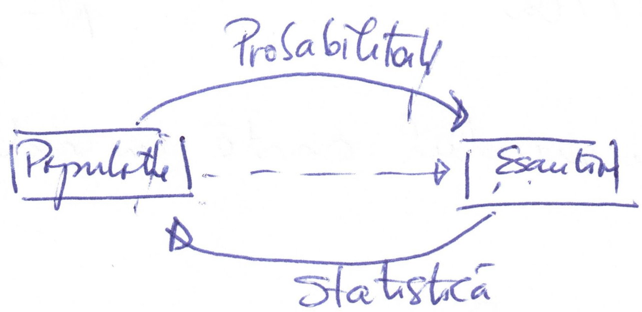
+1p Bonus

→ Activitatea de la seminar + teme (20%)

→ Examen (scris) - 50% $\boxed{40\%}$

1p Oficiu

Structură
→ 1) Simulare
→ 2) Estimare punctuală
→ 3) Estimare prin interval
→ 4) Testare de ipoteze



Exp: O urnă cu bile roșii și verzi într-o proporție necunoscută $\theta \in [0,1]$. Efectuăm n extrageri și notăm culorile extrase.

a) Pb prob: știm că 17% (θ) dintre bilele din urnă sunt de culoare roșie, extragem 20 (n) de bile (cu revenire) și ne întrebăm care este probabilitatea să obținem 5 bile verzi?

6) Pb. statistică: am extras din urnă 20 de bile cu întoarcere și am observat că 5 sunt verzi. În ce măsură această informație ne ajută să aproximăm (estimăm) θ?

Def.: (Model statistic)

Numim model statistic un spațiu măsurabil $(\Omega, \mathcal{F})$ și o familie de măsuri de probabilitate pe $(\Omega, \mathcal{F})$, $\mathbb{P}_\theta$, $\theta \in \Theta$
Θ s.n. spațiul parametrilor.

$(\Omega, \mathcal{F})$ — Ω: sp. stărilor (mulț. ev. elementare); $\mathcal{F}$: $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ mulț. ev. posibile asociate exp. aleatoare

θ - proporția necunoscută a bilelor roșii

$$\Theta = \begin{cases} \left\{ \frac{k}{N} \mid 0 \le k \le N \right\} & , N - \text{bile în urnă (nr. total de bile din urnă)} \\ [0,1] \cap \mathbb{Q} & , N - \text{nu este cunoscut} \end{cases}$$

Cine este Ω? Experimentul constă în extragerea a n bile cu întoarcere:

$$\Omega = \Omega_n = \{ (u_1, u_2, \dots u_n) \mid u_i \in \{r, v\} \}$$
$$= \{r, v\}^n$$

$$\mathcal{F} = \mathcal{F}_n = \mathcal{P}(\Omega_n)$$

Pentru fiecare $\theta \in \Theta$ alegem $\boxed{\mathbb{P}_\theta}$ pe $(\Omega, \mathcal{F})$ iar extragerile să fie independente și pentru fiecare i, $\overline{1,n}$ prob. ca la extragerea i să avem bile roșii (r) să fie θ

$$\mathbb{P}_\theta(\underbrace{(u_1, u_2, \dots u_n)}_{\omega}) = \theta^j (1-\theta)^{n-j}, \quad j = \text{card}\{i \le n \text{ și } u_i = r\}$$

Definim șirul $(X_i)_{i=\overline{1,n}}$

$$X_i(\omega) = X_i((u_1, \dots, u_n)) = \begin{cases} 1, & u_i = r \\ 0, & \text{altfel} \end{cases}$$

X_i sunt indep., $X_i \sim \mathcal{B}(\theta)$

Recap: X, Y indep $\quad X \perp\!\!\!\perp Y$

$$X, Y : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$$

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B), \quad \forall A, B \in \mathcal{B}_{\mathbb{R}}$$

$$X \perp\!\!\!\perp Y \iff \mathbb{P}(X \le x, Y \le y) = \mathbb{P}(X \le x)\mathbb{P}(Y \le y), \quad \forall x, y \in \mathbb{R}$$

$$X_i \sim \mathcal{B}(\theta) \Rightarrow \mathbb{P}_\theta(X_i = 1) = \theta \text{ și } \mathbb{P}_\theta(X_i = 0) = 1-\theta$$

$$S_n = X_1 + X_2 + \dots + X_n \sim \mathcal{B}(n, \theta)$$

$$\mathbb{P}_\theta(S_n = k) = C_n^k \theta^k (1-\theta)^{n-k}$$

$M_n = \frac{S_n}{n} \rightarrow$ proporția de bile roșii în cele n bile extrase

$$\mathbb{P}_\theta\left(M_n = \frac{k}{n}\right) = C_n^k \theta^k (1-\theta)^{n-k}$$

$$\mathbb{E}_\theta[M_n] = \frac{1}{n}\mathbb{E}_\theta[S_n] = \frac{1}{n} \cdot (n\theta) = \theta$$

Pentru n suficient de mare, rep lui M_n este concentrată în jurul lui θ.

Elemente de recapitulare

X v.a. / $(\Omega, \mathcal{F}, \mathbb{P})$

$$X : \Omega \longrightarrow \mathbb{R}$$

$$X^{-1}(A) \in \mathcal{F}, \quad \forall A \in \mathcal{B}_{\mathbb{R}}$$

X v.a. $\Leftrightarrow$ $X^{-1}((-\infty, x]) \in \mathcal{F}, \ \forall x \in \mathbb{R}$

$$X^{-1}(A) = \{w \in \Omega \mid X(w) \in A\}$$

$$\{X \in A\}$$

$\boxed{\text{repartiția} \equiv \text{distribuția} \equiv \text{legea}}$

Repartiția unei v.a. $X : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ este

$$Q = \mathbb{P} \circ X^{-1} : \mathcal{B}_{\mathbb{R}} \longrightarrow [0, 1]$$

$$Q(A) = (\mathbb{P} \circ X^{-1})(A)$$

$$= \mathbb{P}(X^{-1}(A)) = \mathbb{P}(X \in A)$$

$\cancel{E_Q(X^2)}$