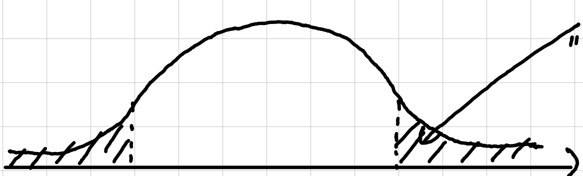
Cursul # 11

4) Legea numerelor mari și teorema limită centrală.

• Inegalitatea lui Markov

Fie $X \sim p\,dx$ v.a. cont. cu $\mathbb{E}[|X|] < +\infty$ și $a > 0$.

$$\mathbb{P}(|X| \geq a) \leq \frac{\mathbb{E}[|X|]}{a}$$



"Coada unei v. a. cu medie finită descrește cel puțin liniar"

Demonstrație:

$$\mathbb{E}[|X|] = \int_{-\infty}^{+\infty} |x|\, p(x)\, dx =$$

$$\int_{-\infty}^{-a} \underbrace{|x|}_{\geq a} p(x)\, dx + \underbrace{\int_{-a}^{a} |x|\, p(x)\, dx}_{\geq 0} + \int_{a}^{+\infty} \underbrace{|x|}_{\geq a} p(x)\, dx$$

$$\geq a \left[\int_{-\infty}^{-a} p(x)\, dx + \int_{a}^{+\infty} p(x)\, dx \right] = a\, \mathbb{P}(|X| \geq a) \quad \square$$

- Inegalitatea lui Cebâșev

Fie X v.a. cont. cu $\mathbb{E}[X^2] < +\infty$ și $a>0$

$$\mathbb{P}(|X - E[X]| \geq a) \leq \frac{Var[X]}{a^2}$$

„Abaterea de la medie descrește cel puțin pătratic"

Demonstrație:

Aplicăm inegalitatea lui Markov pentru $(X - \mathbb{E}[X])^2$ și a^2:

$$\mathbb{P}(|X - E[X]|^2 \geq a^2) \leq \frac{\mathbb{E}[(X - E[X])^2]}{a^2}$$

$$\mathbb{P}(|X - E[X]| \geq a) \qquad\qquad \frac{Var[X]}{a^2} \qquad \square$$

- Inegalitatea lui Chernoff (Opțională)

Fie X v.a. cont. cu $\mathbb{E}[e^{\alpha X}] < +\infty \; \forall \alpha > 0$ și $a > 0$.

$$P(X \geq a) \leq \inf_{\alpha > 0} \mathbb{E}[e^{\alpha X}] e^{-\alpha a}$$

„Descreștere exponențială a cozii"

• Exemplu

Numărul de studenți care termină facultatea de matematică este o v.a. de medie 60.

a) Ce putem spune de probabilitatea de a avea cel puțin 100 de absolvenți într-un an?

b) Ce putem spune de probabilitatea de a avea între 35 și 85 de absolvenți într-un an, știind că varianța este de 100?

Rezolvare:

a) $\mathbb{P}(X > 100) \leq \frac{60}{100} = \frac{3}{5}$ (60%)

b) $\mathbb{P}(35 < X < 85) = \mathbb{P}(|X - 60| < 25)$

$\geq 1 - \frac{10^2}{25^2} = \frac{25^2 - 10^2}{25^2} = \frac{35 \cdot 15}{25^2} = \frac{21}{25}$ (84%)

• Definiție

Fie $X_1, \dots X_n$ v.a. i.i.d (independente și identic distribuite) cu $\mathbb{E}[X_1^2] < +\infty$

$$\Updownarrow$$

$$\left[\mathbb{P}(X_1 \in A) = \mathbb{P}(X_i \in A) \quad \forall i = \overline{1,n}\right]$$

$$\overline{S}_n := \frac{X_1 + \dots + X_n}{n}$$ s.n. media empirică a v.a. $X_1 \dots X_n$

• Proprietăți

i) $\mathbb{E}[\overline{S}_n] = \dfrac{n \cdot \mathbb{E}[X_1]}{n} = \mathbb{E}[X_1] := \mu$ (media teoretică)

ii) $\mathbb{V}ar[\overline{S}_n] = \dfrac{n \cdot \mathbb{V}ar[X_1]}{n^2} = \dfrac{\mathbb{V}ar[X_1]}{n} := \dfrac{\sigma^2}{n}$

• Interpretare:

X_1 aproximează μ cu „risc" σ^2

$\dfrac{X_1 + X_2}{2}$ aproximează μ cu „risc" $\dfrac{\sigma^2}{2}$

. . .

$\overline{S}_n$ aproximează μ cu „risc" $\dfrac{\sigma^2}{n} \xrightarrow[n \to +\infty]{} 0$

• Legea numerelor mari: Ex. de „jucărie"

Primim o monedă măsluită

$X : \Omega \to \{0, 1\}$ „Cap" $\equiv 1$, „Pajură" $\equiv 0$

$X \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$, $p \in (0, 1)$ necunoscut

Q_1: Cum determinăm p?

Aruncăm moneda de n ori: $X_i \overset{i.i.d.}{\sim} \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$

$$\overline{S}_n = \frac{X_1 + \dots + X_n}{n} = \frac{\text{\# Monede care pică cap}}{\text{\# Total de aruncări}}$$

$\overline{S}_n$ aproximează $\mathbb{E}[X_1] = p$ cu $\dfrac{Var[X_1]}{n} = \dfrac{p(1-p)}{n}$

Q_2: Putem cuantifica cât de bine aproxi-
mează $\overline{S}_n$ pe p?

$$\mathbb{P}(\underbrace{|\overline{S}_n - p| \geq \varepsilon}) \leq \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}$$

Eroarea de aproximare

e mai mare ca $\varepsilon > 0$

$$\mathbb{P}(|\overline{S}_n - p| \leq \varepsilon) \geq \boxed{1 - \frac{1}{4n\varepsilon^2}}$$

Limită inferioară pentru probabilitatea de a nu greși cu mai mult de ε aproximarea (Nivelul de încredere)

Q_3: De câte ori trebuie să aruncăm moneda pentru a estima p cu o eroare de maxim $\varepsilon = 0.1$ și un nivel de încredere de 95%?

$$\mathbb{P}(|\overline{S}_n - p| \geq 0.1) \leq \frac{1}{4 \cdot n \cdot 0.01}$$

$$\mathbb{P}(|\overline{S}_n - p| < 0.1) \geq 1 - \frac{1}{4 \cdot n \cdot 0.01} \geq 0.95$$

$$\Leftrightarrow 1 - \frac{25}{n} \geq 0.95 \Leftrightarrow \frac{25}{n} \leq 0.05 = \frac{1}{20}$$

$\Leftrightarrow n \geq 500$, i.e. trebuie să aruncăm de cel puțin 500 de ori moneda.

• Teoremă (Legea numerelor mari)

Fie $X_1 \dots X_n$ v.a. i.i.d cu $\mathbb{E}[X_i] = \mu$ şi $Var[X_i] = \sigma^2 < +\infty$ $\forall i = \overline{1,n}$. Atunci

$$\overline{S_n} = \frac{\sum_{i=1}^{n} X_i}{n} \longrightarrow \mu$$ în probabilitate, i.e.

$$\mathbb{P}(|\overline{S_n} - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2} \xrightarrow[n \to +\infty]{} 0, \forall \varepsilon > 0$$