Notițe curs 2023-2024

Introducerea în probabilități

- experiență aleatoare = rezultatul nu poate fi anticipat cu certitudine
  ex: aruncarea unui zar
- probă/eveniment = rezultatul unei experiențe aleatoare
- eveniment aleator = orice rezultat potențial al unei experiențe aleatoare, a cărui realizare poate fi confirmată sau infirmată de o probă.
- eveniment sigur = toate cazurile posibile ale experienței sunt cazuri favorabile
- eveniment imposibil = nu se poate realiza ($\emptyset$)
- evenimente → elementare $\Rightarrow nf = 1$
  → compuse $\Rightarrow nf > 1$
- eveniment incompatibil = nu se pot realiza împreună în nicio experiență
  $A \cap B = \emptyset$
- evenimente compatibile = au cel puțin un caz favorabil comun
- $\Omega \rightarrow$ mulțimea tuturor rezultatelor posibile
- $P(\Omega) \rightarrow$ familia părților (toate submulț. $\Omega \rightarrow 2^n$ elem)

spațiu de selecție → mulțimea tuturor rezultatelor posibile

σ-algebră → o submulțime a părților → $\begin{cases} A \in K \Rightarrow \bar{A} \in K \\ \{A_n\}_{n \geq 1} \in K, \bigcup_{n \geq 1} A_n \in K \end{cases}$

||

(Ω, K, P) câmp de probabilitate

Proprietăți → $P(A) \in [0,1]$
→ $P(\Omega) = 1$
→ $P(\emptyset) = 0$

$A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

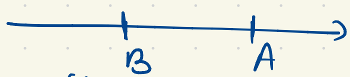
$A \cap B \neq \emptyset \Rightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$$

evenimente independente → nu depind unul de altul

$$P(A \cap B) = P(A) \cdot P(B)$$

evenimente condiționate



$$P(A|B) = P_B(A) = \frac{P(A \cap B)}{P(B)}$$

$\Rightarrow$ intersecția ev. cond.

$$P(A \cap B) = P(B) \cdot P(A|B)$$

A, B evenimente → incompatibile $A \cap B = \emptyset$, $P(A \cap B) = 0$

→ compatibile $A \cap B \neq \emptyset$, $P(A \cap B) \neq 0$

→ independente $P(A \cap B) = P(A) \cdot P(B)$

→ condiționate $P(A \cap B) = P(B) \cdot P(A|B)$

Probabilități totale

$$P(B) = \sum_{i=1}^{m} P(A_i) \cdot P(B/A_i)$$

$P(B) = P(A_1) \cdot P(B|A_1) + \dots$

n urne cu bile

B: „extrag o bilă albă"

$P(B) =$ (prob să extrag din prima urnă ∩ prob să fie albă) ∪ ———

Formula lui Bayes (teorema ipotezelor)

ex: Știm că am extras o anumită bilă, care este prob să o fi extras dintr-o anumită urnă.

! am proba/rezultatul, de unde a fost extrasă

! utilizare: la controlul de calitate ⇒ formula își zice locul

$$P(A_i / B) = \frac{P(B/A_i) \cdot P(A_i)}{P(B)}$$

Distribuții

Schema lui Bernoulli (binomială) → prob de a avea succes

- avem o sg urnă, n extrageri, cu returnare
- prob de a extrage o bilă albă este p și $q = 1 - p$

↳ adică succes

$X \sim Bernoulli(p)$

$$X \sim \begin{pmatrix} 0 & 1 \\ \underbrace{1 - P(A)}_{\bar{A} \text{ insucces}} & P(A) \end{pmatrix} \xrightarrow{n \text{ extrageri}} X \sim Bino(p)$$

$$X \sim \begin{pmatrix} k \\ C_n^k p^k (1-p)^{n-k} \end{pmatrix}$$

$$(a+b)^n = \sum_{k=0}^{n} C_n^k a^k b^{n-k}$$

k → nr. succese

Distribuția hipergeometrică

- pt n extrageri fără returnare
- $P = \frac{C_a^k \cdot C_b^{n-k}}{C_{a+b}^n}$, k-bile albe din urnă (cu bile albe și negre)

Variabile aleatoare

- var aleatoare = funcție matem. (leagă $\underbrace{\text{spațiul de selecție}}_{\text{evenimente}}$ și mult. $\mathbb{R}$)

o mulțime

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

(Ω, K, P) spațiu de selecție

funcț prob.

submult. Ω

$X(\Omega) \subseteq \mathbb{R}$

→ nu e tot codomeniul acoperit

- clasificare
  - → v.a. discrete
    - → $X(\Omega)$ finită
    - → $X(\Omega)$ infinită, dar numărabilă → să pot spune asta e prima, a doua val.
  - → v.a. continue → $X(\Omega)$ infinită

ex: $\Omega = \{1, 2, 3, 4, 5, 6\} \Rightarrow X(\Omega)$ – v.a. discretă

Funcția de repartiție

→ stabilesc o corespondență între Ω și $\mathbb{R}$

(Ω, K, P) spațiu probabilitate

$X : \Omega \to \mathbb{R}$ var. aleatoare

$F_X : \mathbb{R} \to [0, 1]$

$F_X(x) = P(X < x)$

> În practică, e dificil să găsim valorile acestor corespondențe, dar putem determina cât de des apar valorile (cu ce prob.)

→ tablou de repartiție

$$X = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix} \qquad p_i \geq 0 \qquad \sum_{i=1}^{n} p_i = 1.$$

$p_i = P(A_i) = P\{X = x_i\}$

ex: $\begin{pmatrix} 1 & 2 & 3 & \dots & 6 \\ 1/6 & 1/6 & 1/6 & \dots & 1/6 \end{pmatrix}$

Funcția de probabilitate → prob cu care fiecare val. este luată

$$f(x_i) = p_i = P(X = x_i), \quad i = \overline{1,m}$$

Funcția de repartiție

→ prob ca var. mea să ia val mai mici ca x

se plimbă printre val. lui X

$$F_X(x) = \begin{cases} 0, & x < x_1 \\ p_1, & x_1 \le x < x_2 \\ p_1 + p_2, & x_2 \le x < x_3 \\ \dots \\ \sum_{j=1}^{i-1} p_j, & x_{i-1} \le x < x_i \\ \dots \\ 1, & x \ge x_n \end{cases}$$

Variabile aleatoare discrete independente

$$X = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix} \qquad Y = \begin{pmatrix} y_1 & y_2 & \dots & y_m \\ q_1 & q_2 & \dots & q_m \end{pmatrix} \qquad \sum p_i = \sum q_j = 1$$

Def: $\left.\begin{matrix}(X = x_i) \\ (Y = y_j)\end{matrix}\right\}$ independente

$$P[(X = x_i), (Y = y_j)] = P[(X = x_i) \cap (Y = y_j)] = p_i \cdot q_j$$

Operații - var. aleatoare discrete

$$a + X = \begin{pmatrix} a + x_i \\ p_i \end{pmatrix}_{i=\overline{1,m}} \qquad a \cdot X = \begin{pmatrix} a \cdot x_i \\ p_i \end{pmatrix} \qquad a \in \mathbb{R} \to ct.$$

$$X + Y = \begin{pmatrix} x_i + y_j \\ p_{ij} \end{pmatrix} \qquad X \cdot Y = \begin{pmatrix} x_i \cdot y_j \\ p_{ij} \end{pmatrix} \qquad X^k = \begin{pmatrix} x_i^k \\ p_i \end{pmatrix}$$

p_{ij} ↳ prob să se întâmple simultan (dacă e indep. ⇒ produs)

Variabila aleatoare continue

densitate de repartiție → fcț pt a descrie repartizarea valorilor.

(Ω, K, P) câmp de probabilitate

$X \to$ clasă de var aleatoare $\quad f: \mathbb{R} \to [0, +\infty)$

$$P(X < x) = F(x) = \int_{-\infty}^{x} f(t)\,dt \qquad F'(x) = f(x)$$

$F(x)$ ↓ funcția de repartiție a var. X.

$$X \sim \begin{pmatrix} x \\ f(x) \end{pmatrix}$$

Proprietăți
- $f(x) \geq 0$ pt $\forall x \in \mathbb{R}$
- $\int_{-\infty}^{\infty} f(x)\,dx = 1$
- $P(a \leq X < b) = \int_a^b f(x)\,dx = F(b) - F(a)$

Obs $\quad F_X(x) = P(X < x) \xRightarrow[\forall a, b \in \mathbb{R}]{}$
- $P(X = a) = 0$
- $P(a \leq X < b) = P(a < X < b) = P(a < X \leq b) = P(a \leq X \leq b) = \int_a^b f(x)\,dx = F(b) - F(a)$

Caracteristici numerice ale variabilelor aleatoare

Media → o măsură a tendinței centrale a valorilor sale
- ↳ v.a. discretă $\quad M[X] = \sum_{i=1}^{n} x_i \cdot p_i$
- ↳ v.a. continuă $\quad M[X] = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$
- ↳ v.a. discretă cu inf. de val $\quad M[X] = \sum_{n \geq 0} x_n \cdot p_n$
- ↳ proprietăți
  - $X : \begin{pmatrix} c \\ 1 \end{pmatrix}$, $M[X] = c$
  - $M[a + X] = a + M[X]$
  - $M[a \cdot X] = a \cdot M[X]$
  - $\min x_i < M[X] < \max x_i$
  - $M[X + Y + ..] = M[X] + M[Y] + ..$ analog produs

Obs $M[g(X)] = \sum_{i=1}^{m} g(x_i)\cdot p_i$ $\qquad M(g(X)) = \int_{-\infty}^{\infty} g(x)\cdot f(x)\,dx$

Dispersia → caract. cum sunt repartizate val unei var. aleatoare

$$D^2[X] = M[(X-m)^2] \qquad m = M[X]$$

$$D^2[X] = \sum_{i=1}^{m} (x_i - m)^2 \cdot p_i \qquad D^2[X] = \int_{-\infty}^{+\infty} (x-m)^2 \cdot f(x)\,dx$$

→ proprietăți
- $D^2[C] = 0$, c-ct.
- $D^2[a+X] = D^2[X]$
- $D^2[aX] = a^2 D^2[X]$
- $D^2[X+Y+\dots] = D^2[X] + D^2[Y] + \dots$

Abaterea (Deviația standard) $\sigma = \sqrt{D^2[X]}$

Covarianța

$$\mathrm{Cov}[X,Y] = M[(X-M[X])(Y-M[Y])] \text{ sau } M[XY] - M[X]M[Y]$$

↓ cât de împrăștiate sunt împreună

→ proprietăți
- $\mathrm{Cov}[X,Y] = \mathrm{Cov}[Y,X]$
- $\mathrm{Cov}[X,X] = D^2[X]$
- $\mathrm{Cov}[a+X, b+Y] = \mathrm{Cov}[X,Y]$
- $D^2[X \pm Y] = D^2[X] + D^2[Y] \pm 2\,\mathrm{Cov}[X,Y]$
- X, Y - var indepen. $\mathrm{Cov}[X,Y] = 0$. ($\nexists$ reciproca)

Coeficient de corelație

$$\rho(X,Y) = \frac{\mathrm{Cov}[X,Y]}{\sigma_X \cdot \sigma_Y}$$

Obs: X, Y - indepen.
→ $\rho(X,Y) = 0$

prop
- $-1 \le \rho(X,Y) \le 1$
- $\rho[X,Y] = \begin{cases} 1, & a>0 \\ -1, & a<0 \end{cases} \qquad Y = aX + b$

Distribuțiile de probabilitate continue

Distribuția uniformă pe intervalul $[a,b]$

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b] \\ 0, & x \in \mathbb{R} \setminus [a,b] \end{cases}$$

Distribuția normală (Gauss)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Distribuția exponențială

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

Clasificarea naivă a lui Bayes

(Ω, K, P) spațiu de probabilitate

$A, B \in K$ — evenimente condițional indepen.

$$P(A \cap B \mid C) = P(A \mid C) \cdot P(B \mid C)$$

Repartiții clasice

↳ Repartiții v.a. discrete

1. Bernoulli (numai 2 val, modelarea unui fenomen)

- experiența se efectuează o sg. dată

$$X \sim \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix} \qquad p+q=1 \qquad q = 1-p \text{ (insucces)}$$

- media $E[X] = p$
- Dispersia (varianța) $D^2[X] = p(1-p)$
- fct. generatoare de momente $g_X(t) = (1-p) + p \cdot e^t$
- fct. caract. $\varphi(t) = (1-p) + p \cdot e^{it}$

2. Repartiția Binomială – Bino(p)

– generalizarea lui Bernoulli (pt N ori de repetare)

$$X \sim \begin{pmatrix} 0 & 1 & \dots & k & \dots & n \\ C_n^0 p^0 q^n & C_n^1 p^1 q^{n-1} & \dots & C_n^k p^k q^{n-k} & \dots & C_n^n p^n q^0 \end{pmatrix}$$

$$= \begin{pmatrix} k \\ C_n^k p^k q^{n-k} \end{pmatrix}$$

$p+q=1$

↳ $1-p$

- $E[X] = n \cdot p$
- $D^2[X] = n \cdot p \cdot (1-p)$
- $g_X(t) = (p \cdot e^t + 1-p)^n$
- $\varphi_X(t) = (p \cdot e^{it} + 1-p)^n$

3. Repartiția Poisson (pt prob. ff mică)

$$X \sim \begin{pmatrix} k \\ e^{-\lambda} \frac{\lambda^k}{k!} \end{pmatrix}, \quad \lambda > 0$$

$n \to \infty$

$p \to 0$

$n \cdot p = \lambda$

a → nr apariții în unitatea de timp

$\lambda = a \cdot t$ media aparițiilor

↳ interval timp

$$P(X=k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

- $E(X) = \lambda$
- $D^2[X] = \lambda$
- $g_X(t) = e^{\lambda(e^t - 1)}$
- $\varphi_X(t) = e^{\lambda(e^{it} - 1)}$

4. Repartiția uniformă discretă (nr. finit de experiențe)

- $p = \frac{1}{n}$ (ex. zar)

$$X \sim \begin{pmatrix} 1 & 2 & \dots & k & \dots & n \\ 1/n & 1/n & \dots & 1/n & \dots & 1/n \end{pmatrix}$$

- $E[X] = \frac{n+1}{2}$
- $D^2(X) = \frac{n^2-1}{12}$

↳ Repartiții v.a. continue

1. Repartiția uniformă continuă   Unif([a,b])

$n \to \infty$

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in (a,b) \\ 0, & x \notin (a,b) \end{cases} \qquad b > a.$$

↓
densitate de repartiție

- $E(X) = \frac{a+b}{2}$
- $D^2(X) = \frac{(b-a)^2}{12}$

2. Repartiția exponențială → fenomene cu timpi de așteptare
↳ param. continuu

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

- $E(X) = \frac{1}{\lambda}$
- $D^2(X) = 1/\lambda^2$

λ → constanta de proporționalitate
↓
prob ca X să ia val într-un interval de timp direct proporțională cu lungimea acelui interval

crește timpul ⇒ crește probabilitatea

3. Repartiția normală (Gauss)   Norm(m, σ)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-m)^2}{2\sigma^2}}$$

- suma a 2 v.a. normale $X_1 \pm X_2 \in Norm(m_1 \pm m_2, \sqrt{\sigma_1^2 + \sigma_2^2})$
- v.a. indepen. și identic repartizate

$$Y = \frac{X_1 + X_2 + \dots + X_n}{n}, \quad Y \in Norm\left(m, \frac{\sigma}{\sqrt{n}}\right)$$

- $E(X) = m$
- $D^2(X) = \sigma^2$

4. Repartiția normală standard $(E(X) = 0,\ D^2(X) = 1)$

$m = 0 \quad \sigma = 1$

$$f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$$

Funcț. de repartiție

$$F_X(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt$$

Repartiții clasice

pt. v. a. discrete

Bernoulli

$$X \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$$

$p+q=1$

$E(X)=p$, $D^2(X)=p(1-p)$

Binomială $Bino(p)$

$$X \sim \begin{pmatrix} k \\ C_n^k \, p^k (1-p)^{n-k} \end{pmatrix}$$

$p+q=1$

$E(X)=n \cdot p$, $D^2(X)=n \cdot p \cdot (1-p)$

Uniformă $(p = 1/n)$

$$X \sim \begin{pmatrix} 1 & 2 & \dots & n \\ 1/n & 1/n & \dots & 1/n \end{pmatrix}$$

$E(X) = \frac{n+1}{2}$, $D^2(X) = \frac{n^2-1}{12}$

pt. v. a. continue

Uniformă $Unif([a,b])$

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in (a,b) \\ 0, & x \notin (a,b) \end{cases}$$

densitate de repartiție

$E(X) = \frac{a+b}{2}$, $D^2(X) = \frac{(b-a)^2}{12}$

Exponențială $Exp(\lambda)$

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

$E(X) = 1/\lambda$, $D^2(X) = 1/\lambda^2$

Normală (Gauss) $N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$E(X) = \mu$, $D^2(x) = \sigma^2$

$N(0,1) \to$ **standard normală**

$$f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$$

Alte formule

$E(X)$:
- v.a. discretă $\to \sum_{i=1}^{n} x_i \cdot p_i$
- v.a. discretă cu inf. val. $\to \sum_{n \ge 0} x_n \cdot p_n$
- v.a. continuă $\to \int_{-\infty}^{\infty} x \cdot f(x)\,dx$

$\sigma = \sqrt{D^2(X)}$

↳ Abaterea/Dev. standard

$D^2(x)$:
- v.a. discretă $\to \sum_{i=1}^{n} (x_i - E(X))^2 \cdot P(X=x_i)$
- v.a. continuă $\to E(x^2) - E(X)^2$

$$Cov(X,Y) = E\left[(X - E(X))(Y - Y(X))\right]$$

Vectori aleatori

(Ω, K, P) spațiu de probabilitate

↳ o funcție reală

$U : \Omega \to \mathbb{R}^2$, $U = (X, Y) \to$ vector aleator

$F_U(x, y) = P(X < x, Y < y)$

↓ funcție de repartiție / distribuție

1. Vectori aleatori discreți

$X \backslash Y$	y_1	y_2	$- - -$	y_n	
x_1	p_{11}	p_{12}	$- - -$		$\sum_{j=1}^{n} p_{1j}$
x_2	p_{21}	p_{22}	$---$		
$\vdots$		$\vdots$			
x_m	p_{m1}	p_{m2}	$- - -$		
	$\sum_{i=1}^{m} p_{i1}$		$- - -$		1

$$p_{ij} = P((X = x_i) \cap (Y = y_j)) = P(X = x_i, Y = y_j)$$

$X, Y \to$ indepen. $p_{ij} = p_i \cdot q_j$

↳ fcț. de repartiție și probabilitate

$$f(x_i, y_j) = p_{ij} \qquad F_U(x, y) = P(X < x, Y < y) = \sum_{l=1}^{j-1} \sum_{k=1}^{i-1} p_{kl}$$

2. Vectori aleatori continui

$$F_U = (x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v)\, du\, dv$$

↳ fct. repartiție ↳ densitate

Legea numerelor mari (LTNM)

↳ studiază leg dintre frecvența de apariție a unor evenimente și prob. (experimentul se repetă de n mare de ori)

$$\frac{1}{n} \sum_{k=1}^{n} (X_k - E(X_k)) \xrightarrow{a.s.} 0$$

Elemente de statistică

Introducere

- populația statistică = indivizi/unități statistice
- caracteristica = trăsătura comună
- medie de selecție $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$
- media empirică $\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$

Metoda momentelor

↳ estimatori absoluți corecți pt momentele teoretice