Matematici speciale

(curs 3 - S3)

07.03.2024

Recapitulare

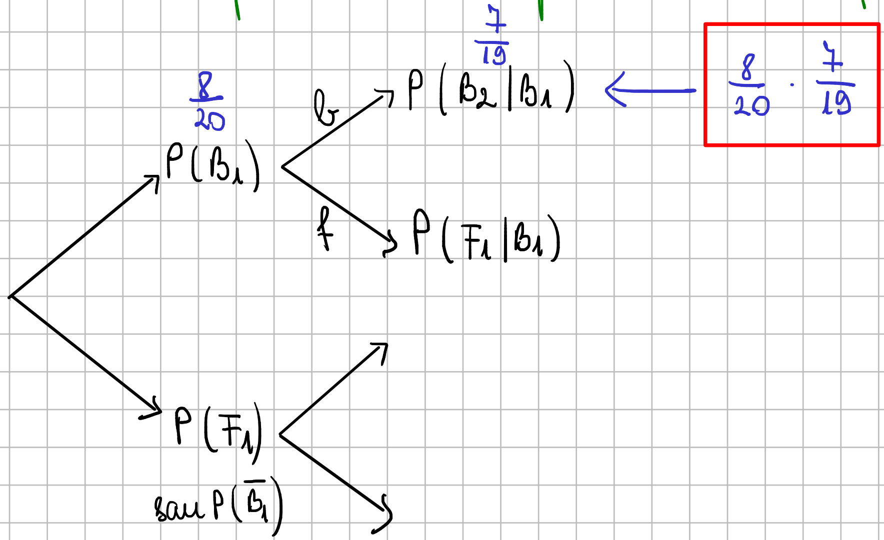
Eveniment	Mulțime	Formulă
A	A	$P(A) \in [0,1]$
ev $(A \cup B)$	Mulțime $(A \cup B)$	$P(A \cup B) = P(A) + P(B) \setminus P(A \cap B)$
ev. imposibil	$\emptyset$	$P(\emptyset) = 0$
ev. sigur	Ω	$P(\Omega) = 1$
ev. $C_A \equiv \bar{A}$	CA / $\bar{A}$	$P(CA) = 1 - P(A)$
ev. $A \cap B$	$A \cap B$	$P(A \cap B) = P(A) \cdot P(B \mid A)$

$$+ \quad P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2 \mid A_1) \cdot P(A_3 \mid A_1 \cap A_2)$$

ex: O echipă are 20 de persoane → 8 băieți, → 12 fete

Se aleg 2 persoane la întâmplare

Care este probabilitatea ca persoanele alese să fie băieți?



Seminar S3 → temă

8. Intr-un birou se cumpără un nou computer. Firma producătoare menţionează în certificatul de garanţie că există probabilitatea de 5% ca acest calculator să se defecteze în primul an. Dacă nu s-a defectat în primul an, atunci cu probabilitate de 10% se poate defecta în al doilea an. Dacă nu s-a defectat în primii doi ani de funcţionare, atunci cu probabilitate de 30% s-ar putea defecta în al treilea an. Să se calculeze:
a) probabilitatea să nu se defecteze în primii doi ani;
b) probabilitatea să nu se defecteze în primii trei ani.

$P(\overline{A_1}) = \frac{5}{100}$ $\quad P(\overline{A_2}) = \frac{10}{100}$ $\quad P(\overline{A_3}) = \frac{30}{100}$

a) $P(A_1 \cap A_2) = P(A_1) \cdot P(A_2 | A_1)$

$$\boxed{P(A_1) = 1 - P(\overline{A_1})} = \frac{95}{100}$$

$$\boxed{P(A_2|A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)}} = \frac{P(A_1) \cdot P(A_2)}{P(A_1)} = P(A_2) = 1 - P(\overline{A_2}) = \frac{90}{100}$$

$$P(A_1 \cap A_2) = 0{,}95 \cdot 0{,}90 = 0{,}855$$

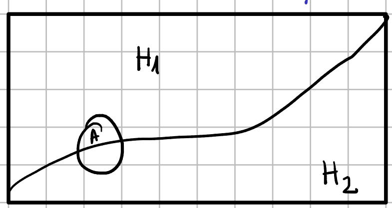
b) $P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2 | A_1) \cdot P(A_3 | A_1 \cap A_2)$

$$P(A_3 | A_1 \cap A_2) = \frac{P(A_1) \cdot P(A_2) \cdot P(A_3)}{P(A_1) \cdot P(A_2)} = P(A_3) = 1 - P(\overline{A_3}) = \frac{70}{100}$$

$$P(A_1 \cap A_2 \cap A_3) = 0{,}95 \cdot 0{,}9 \cdot 0{,}7 = 0{,}59$$

$P(CA | B) = 1 - P(A|B)$, B - fixat

Formula probabilităţii totale. Formula lui Bayes



Ω = "sample space"

H_1, H_2 - evenimente de tip ipoteză

A - ev. de tip informaţie

$$\begin{cases} H_1 \cup H_2 = \Omega \\ H_1 \cap H_2 = \emptyset \text{ (mutual exclusive)} \end{cases}$$

Formula probabilităţii totale: $P(A) = \sum_{i=1}^{2} P(H_i) \cdot P(A|H_i) = P(H_1) \cdot P(A|H_1) + P(H_2) \cdot P(A|H_2)$

Formula lui Bayes: $P(H_i | A) = \frac{P(A|H_i) \cdot P(H_i)}{P(A)}$

ex: 30% din email-uri → spam

5% din email-uri spam conţin cuv. "loterie"

0,5% din email-uri non-spam conţin cuv. "loterie"

Care e probabilitatea ca un email care conţine cuv "loterie" să fie în spam?

$P(H_1)=\frac{30}{100} \Rightarrow P(H_2)=\frac{70}{100}$ (cele care nu sunt în spam)

$P(A|H_1)=\frac{5}{100}$

$P(A|H_2)=\frac{0,5}{100}$

$P(H_1|A)=?$

$$P(A)=P(H_1)\cdot P(A|H_1)+P(H_2)\cdot P(A|H_2)=0,3\cdot 0,05+0,7\cdot 0,005=0,019$$

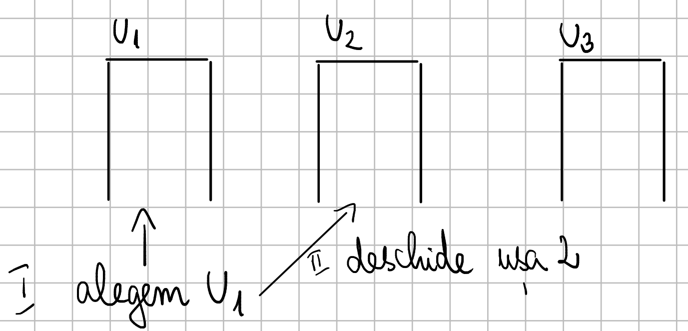
$$P(H_i|A)=\frac{P(A|H_i)\cdot P(H_i)}{P(A)}$$

$$P(A_1|A)=\frac{\frac{5}{100}\cdot\frac{30}{100}}{0,019}=0,789$$

ex: Problema Monty-Hall

8. Monty Hall problem

Prezentăm acum una din cele mai cunoscute probleme din teorie probabilităţii, fiind legată de un concurs de televiziune din anii '70, Monty Hall: Într-un concurs televizat ţi se oferă posibilitatea alegerii dintre trei uşi, in spatele unei aflându-se un automobil, iar în spatele celorlalte, capre. Tu alegi o uşă, să zicem nr.1, iar gazda emisiunii, care ştie ce se află în spatele uşilor, deschide uşa numărul 2, în spatele căreia se află o capră. Apoi, te întreabă: "vrei să alegi uşa cu numărul 3?". Este în avantajul tău să schimbi alegerea iniţială? *Pentru mai multe informaţii va invităm sa citiţi pagina Wikipedia dedicată acestui subiect.*



H_i – ev. maşina se află în spatele uşii i

U_i – ev. se deschide uşa i după alegere

$$P(H_1|U_2) \longrightarrow \text{păstrează alegerea} = \frac{\overset{\frac{1}{3}}{P(H_1)}\cdot \overset{\frac{1}{2}}{P(U_2|H_1)}}{P(U_2)} = \frac{1}{3}$$

$$P(H_3|U_2) \longrightarrow \text{schimbă alegerea} = \frac{\overset{\frac{1}{3}}{P(H_3)}\cdot \overset{1}{P(U_2|H_3)}}{P(U_2)} = \frac{2}{3}$$

$$P(U_2)=P(H_1)\cdot P(U_2|H_1)+P(H_2)\cdot P(U_2|H_2)+P(H_3)\cdot P(U_2|H_3)=$$
$$=\frac{1}{2}$$

⇒ mai bine schimbă alegerea

Variabile aleatoare

O variabilă aleatoare (v.a.) este o funcţie ce atribuie un număr real fiecărui rezultat din spaţiul tuturor realizărilor unui experiment aleator.

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

Variabilă aleatoare discretă: are valori într-o mulţime finită sau infinit numărabilă

- număr de biţi transmişi cu eroare într-un canal de comunicaţie
- proporţia de componente defecte din cele 1000 testate;

Variabilă aleatoare continuă: poate lua orice valoare dintr-un interval din $\mathbb{R}$ (mărginit sau nu)

- temperatura, greutatea, presiunea

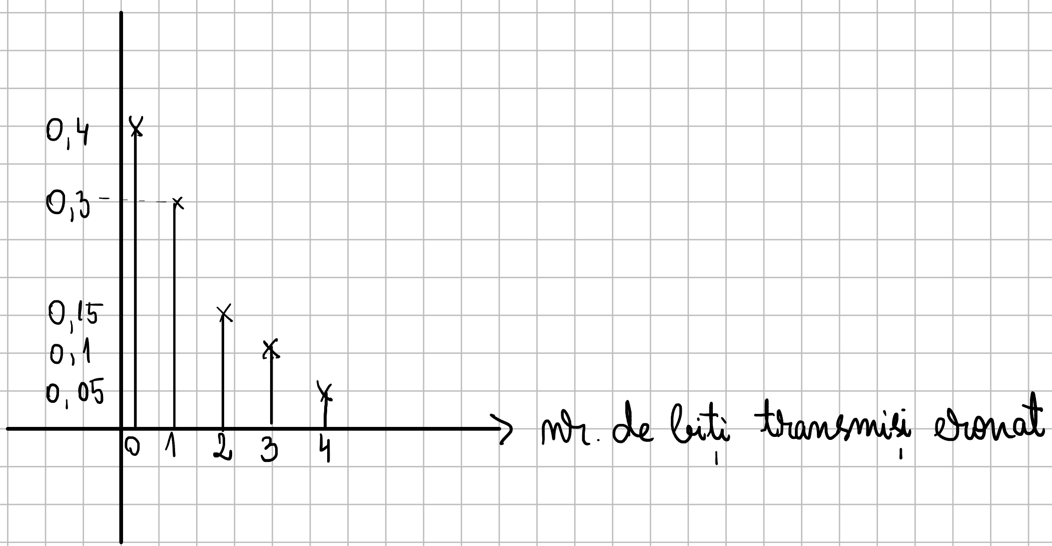
Intr-un canal de transmitere digitală biţii pot fi transmişi eronat cu o anumită probabilitate.

- Se transmit 4 biţi şi notăm cu X numărul de biţi transmişi eronat.
- Valorile posibile ale lui X sunt $\{0, 1, 2, 3, 4\}$.
- Presupunem că probabilităţile de a transmite eronat aceşti biţi sunt $P(X = 0) = 0.4$, $P(X = 1) = 0.3$, $P(X = 2) = 0.15$, $P(X = 3) = 0.1$, $P(X = 4) = 0.05$

$$X = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 0,4 & 0,3 & 0,15 & 0,1 & 0,05 \end{pmatrix}$$

$D_X = \{0,1,2,3,4\}$

$\boxed{p_k = P(X = k)}$ $\sum p_k = 1 \Rightarrow$ variabilă aleatoare discretă $X = \begin{pmatrix} x_i \\ p_i \end{pmatrix}$



Reprezentare alternativă

$P_X : B_X \to [0,1]$ $\quad P_X(x_k) = p_k$

$$P_X(x) = \begin{cases} 0{,}4, & x=0 \\ 0{,}3, & x=1 \\ 0{,}15, & x=2 \\ 0{,}1, & x=3 \\ 0{,}05, & x=4 \end{cases}$$

Obs: Cu variabile aleatoare putem calcula probabilitatea unor evenimente

$$P(X \le 3) = P(X=0 \cup X=1 \cup X=2 \cup X=3) =$$
$$= 0{,}4 + 0{,}3 + 0{,}15 + 0{,}1 =$$
$$= 0{,}95$$

$$P(X\le 3) = 1 - P(C(X\le 3)) = 1 - P(X=4) = 1 - 0{,}05 = 0{,}95$$

• Funcția de repartiție $F_X : \mathbb{R} \to [0,1]$, $F_X(x) : P(X \le x)$

Avem:

$$F_X(x) = \sum_{i : x_i \le x} p_i$$

$$F_X(x) = \begin{cases} 0 & \text{pentru } x < x_1 \\ p_1 & \text{pentru } x_1 \le x < x_2 \\ p_1 + p_2 & \text{pentru } x_2 \le x < x_3 \\ \vdots & \\ p_1 + p_2 + \cdots + p_k & \text{pentru } x_k \le x < x_{k+1} \\ \vdots & \\ 1 & x \ge x_n \end{cases}$$

pt $x < 0$ $\quad F_X(x) = P(X \le x) = 0$

pt. $0 \le x < 1 \Rightarrow F_X(x) = P(X \le x) = P(x = 0) = 0{,}4$

pt. $1 \le x < 2 \Rightarrow F_X(x) = P(X < 2) = P(X=0 \cup X=1) = P(X=0) + P(X=1) =$

$$= 0{,}4 + 0{,}3 = 0{,}7$$

:

Dacă $x \geq 4$, atunci $F(x) = P(X \leq x) = P(X = 0) + P(X = 1) + P(X = 2) + P(X + 3) + P(X = 4) = 1$.

In final avem : $F_X(x) = \begin{cases} 0 & \text{pentru } x < 0 \\ 0.4 & \text{pentru } 0 \leq x < 1 \\ 0.7 & \text{pentru } 1 \leq x < 2 \\ 0.85 & \text{pentru } 2 \leq x < 3 \\ 0.95 & \text{pentru } 3 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$

$P(X \leq 3) = F(3) = 0.95$(din definiția funcției de repartiție)

Proprietăți

- crescătoare: $a < b \Rightarrow F(a) \leq F(b)$
- $\lim_{x \to -\infty} F_X(x) = 0$ și $\lim_{x \to \infty} F_X(x) = 1$
- orice punct de discontinuitate este de spețа întâi
- este continuă la dreapta în orice punct $x_0 \in \mathbb{R} \to P(X = x_0) = F_X(x_0) - \lim_{x \to x_0, x < x_0} F_X(x)$

Media unei variabile aleatoare

Fie X o variabilă aleatoare discretă de valori x_i, și $p_i = P(X = x_i)$, $i \in 1, n$, probabilitățile cu care X ia aceste valori.

$M(X) = E(X) = \sum_{i \geq 1} x_i p_i$

O variabilă aleatoare discretă X ce ia o singură valoare, $a \in \mathbb{R}$ și cu distribuția de probabilitate:

$$X = \begin{pmatrix} a \\ 1 \end{pmatrix}$$

are valoarea medie $M(X) = a$.

Media este liniară.

Exemplu: Variabilă aleatoare discretă cu medie infinită
Fie X o variabilă aleatoare

$$X = \begin{pmatrix} 2^i \\ 1/2^i \end{pmatrix}, i = 1, 2, \ldots$$

Avem : $M(X) = \sum_{i \geq 1} 2^i \frac{1}{2^i} = \sum_{i \geq 1} 1 = \infty$

Teoremă

Fie $X_1, X_2, \ldots X_2$ o mulțime de variabile aleatoare discrete (cu medie finită). Atunci,

$$M(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} M(X_i)$$

Dispersia unei variabile aleatoare ne dă informații despre cât de împrăștiate valorile lui X față de medie.

$$\sigma^2(X) = \sum_{i=1}^{n} (x_i - m)^2 \cdot p_i$$

Abatere standard : $\sigma(X) = \sqrt{\sigma^2(X)}$

$X = a \Rightarrow \sigma^2(X) = 0 \quad (X = ct.)$

Rezultate:

- $M(aX + b) = aM(X) + b, \forall a, b \in \mathbb{R}$
- $D^2(aX + b) = a^2 D^2(X), \forall a, b \in \mathbb{R}$
- $\sigma^2(X) = M(X^2) - (M(X))^2$
- $M(X^2) \geq (M(X))^2$
- Inegalitatea Jensen: dacă f este o funcție convexă, atunci $M(f(X)) \geq f(M(X))$

Pentru ultimele relații trebuie sa introducem: "Funcții de o variabilă aleatoare"[1]

Exemplu Fie X variabila aleatoare ce înregistrează rezultatul aruncării unui zar si Notăm cu $Y = X^2$. Să se determine distribuția lui Y și $M(Y)$.
Rezolvare:

X	1	2	3	4	5	6
Y=X^2	1	4	9	16	25	36
prob	1/6	1/6	1/6	1/6	1/6	1/6

Atunci, $M(X) = \sum_{i=1}^{6} x_i p_i = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \ldots + 6 \cdot \frac{1}{6} = 3.5$

$M(Y) = \sum_{i=1}^{6} x_i^2 p_i = 1^2 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + \ldots + 36 \cdot \frac{1}{6} = 15.16$

Observație: In general, $M(f(X)) = \sum_i f(x_i) p_i$.

Atenție: $M(f(X)) \neq f(M(X))$.

- cea mai simplă distribuție cu un număr finit de valori, egal probabile
- fie $x_1, \ldots, x_n$ valorile, iar $p_i = \frac{1}{n}$ frecvența de apariție
- variabila X se spune că are o **distribuție discretă uniformă** dacă

$$X = \begin{pmatrix} x_i \\ 1/n \end{pmatrix}, i = 1, 2, \ldots, n \Leftrightarrow p_i = p(x_i) = P(X = x_i) = \frac{1}{n}, \forall i$$

- Dacă X are valorile consecutive întregi $a, a+1, a+2, \ldots, b$, atunci

$$M(X) = \frac{b+a}{2}, \sigma^2(X) = \frac{(b-a+1)^2 - 1}{12}$$