Curs 3 : Variabile aleatoare discrete: Definiție. Proprietăți

Conf.dr. Maria Jivulescu

Departamentul de Matematică
UPT

Cuprins

Variabile aleatoare: Definiție

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

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

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

- număr de biți transmiși cu eroare într-un canal de comunicație
- proporția 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

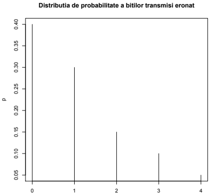
Variabile aleatoare discrete: exemplu

Intr-un canal de transmitere digitală biții pot fi transmiși eronat cu o anumită probabilitate.

- Se transmit 4 biți și notăm cu X numărul de biți transmiși eronat.
- Valorile posibile ale lui X sunt $\{0, 1, 2, 3, 4\}$.
- Presupunem că probabilitățile de a transmite eronat acești biți sunt $P(X = 0) = 0.4, P(X = 1) = 0.3, P(X = 2) = 0.15, P(X = 3) = 0.1, P(X = 4) = 0.05$

Putem descrie pe X cu un tabel $X = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 0.4 & 0.3 & 0.15 & 0.1 & 0.05 \end{pmatrix}$

Sau cu o funcție $p(k) = P(X = k), k \in \{0, 1, 2, 3, 4\}$ al cărei grafic este:



Distribuția de probabilitate

- Fie un experiment aleator cu n realizări, cuantificate numeric de valorile reale $x_1, x_2, \ldots, x_n$;
- Notăm cu X variabila ce ia aceste valori;
- Ev. elementare ale experimentului $E_i = (X = x_i), i \in 1, n$
- $(X = x_i)$ se citește "evenimentul ca X să ia valoarea x_i";
- Notăm cu $p_i = P(X = x_i)$ probabilitatea ev. $(X = x_i)$, $i \in 1, n$.
- Ev. sigur al experimentului este $\Omega = \cup_{i=1}^{n}(X = x_i)$
- Ev. $(X = x_1)$, $(X = x_2), \ldots, (X = x_n)$ sunt mutual exclusive, deci $P(\Omega) = P(X = x_1) + P(X = x_2) + \cdots P(X = x_n)$
- Avem $p_1 + p_2 + \cdots p_n = 1$

Distribuția de probabilitate

Tabloul (matricea cu două linii) ce înregistrează pe prima linie valorile variabilei X și pe a doua probabilitțile p_i asociate, tablou notat la fel, cu X:

$$X = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix}, \quad \sum_{i=1}^{n} p_i = 1$$

definește *distribuția de probabilitate* a variabilei aleatoare.

Alternativ

Distribuția de probabilitate printr-o funcție:

$$p_X : D \to [0,1], p_X(x_k) = P(X = x_k)$$

(probabilitatea ca variabila aleatoare X să ia valoarea k).

Această funcție are proprietățile: $p(x_k) \geq 0, \sum_{k=1}^{n} p(x_k) = 1.$

Ce putem face cu variabile aleatoare?

Fie X o variabilă aleatoare discretă X. Pornind de la evenimentele elementare $(X = x_i)$ ale experimentului aleator se pot construi evenimente de tipul:

- $(X < x)$,$(X \leq x)$
- $(X > x)$, $X \geq x)$
- $(a < X < b)$, $(a \leq X < b)$, $(a \leq X \leq b)$, $(a < X \leq b)$

Cum aceste evenimente se pot scrie ca reuniune de evenimentele mutual exclusive $(X = x_i)$, probabilitatea lor este suma probabilităților evenimentelor elementare componente.

Intr-adevăr, pentru $X = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 0.4 & 0.3 & 0.15 & 0.1 & 0.05 \end{pmatrix}$ avem:

$P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) =$
$0.4 + 0.3 + 0.15 + 0.1 + 0.05 = 0.95$
sau $P(X \leq 3) = 1 - P(X = 4) = 0.95$

Fie X o variabilă aleatoare discretă.
Funcția $F_X : \mathbb{R} \to [0,1]$, $F_X(x) = P(X \leq x)$ (probabilitatea ca variabila aleatoare X să ia valori mai mici sau egale cu x), se numește funcția de repartiție a variabilei X("cumulative distribution function"(c.d.f.))

Avem:

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

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

Funcția de repartiție

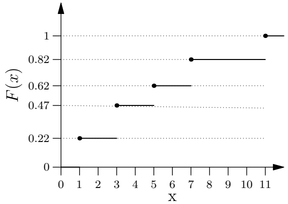


Figure: Graficul functiei de repartitie a variabilei aleatoare X

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 speța întâi
- este continuă la dreapta în orice punct
  $x_0 \in \mathbb{R} \to P(X = x_0) = F_X(x_0) - \lim_{x \to x_0, x < x_0} F_X(x)$

Funcția de repartiție

Exemplu: $X = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 0.4 & 0.3 & 0.15 & 0.1 & 0.05 \end{pmatrix}$

Avem: dacă $x < 0$, atunci $F(x) = P(X \leq x) = 0$.

Dacă $0 \leq x < 1$, atunci $F(x) = P(X \leq x) = P(X = 0) = 0.4$.

Dacă $1 \leq x < 2$, atunci

$F(x) = P(X \leq x) = P(X = 0) + P(X = 1) = 0.4 + 0.3 = 0.7$.

$\vdots$

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

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

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

Media și dispersia unei variabile aleatoare discrete

Fie X o variabilă aleatoare discretă de valori x_i, și $p_i = P(X = x_i)$, $i \in \overline{1, n}$, probabilitățile cu care X ia aceste valori.
Valoarea medie a variabilei X este numărul notat $M(X)$:

$$M(X) = \sum_{i=1}^{n} x_i p_i.$$

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 $M(X) = a$.

Liniaritatea mediei

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

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

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

Teoremă

Fie $X_1, X_2, \dots X_2$ 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 variabilei aleatoare

Fie X o variabilă aleatoare discretăX având valoarea medie m.
Dispersia variabilei aleatoare X, notată $D^2(X)$ sau $\sigma^2(X)$, este numărul real definit prin:

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

Abaterea standard a variabilei X este $\sqrt{\sigma^2(X)}$ și se notează prin $\sigma(X)$.
O variabilă aleatoare constantă, $X = a$, are dispersia $\sigma^2(X) = 0$.

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]

[1] Funcții de o variabilă aleatoare mai multe despre in Seminar 5

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: $M(f(X)) \neq f(M(X))$.

Concluzii: Distribuția discretă uniformă

- 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

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