Curs 6

Variabile aleatoare continue

6.1 Funcţia de repartiţie

În cele ce urmează $\{\mathbf{F}, \Omega, P\}$ este un câmp de probabilitate.

Definiţia 6.1.1 *Fie X o v. a. Se numeşte* ***funcţie de repartiţie*** *a lui X (sau* ***funcţie de distribuţie cumulativă****) funcţia reală*

$F_X : \mathbb{R} \to [0,1]\,,\; F_X(x) = P(\{X < x\})$.

Rezultă că oricărui număr real x i se asociază prin funcţia de repartiţie F probabilitatea ca valorile lui X să fie mai mici decât x.

Pentru o variabilă aleatoare discretătă cu matricea de repartiţie dată prin:

$X : \begin{pmatrix} x_1 & x_2 & ... & x_n \\ p_1 & p_2 & ... & p_n \end{pmatrix}$, unde am presupus, pentru simplitate, $x_1 < x_2 < ... < x_n$,

se poate introduce funcţia de repartiţie care este o funcţie scară dată de relaţa:

$$F_X(x) = \begin{cases} 0, & \text{dacă} \quad x \le x_1, \\ p_1, & \text{dacă} \quad x_1 < x \le x_2, \\ p_1 + p_2, & \text{dacă} \quad x_2 < x \le x_3, \\ \dots & \\ p_1 + \dots + p_{n-1}, & \text{dacă} \quad x_{n-1} < x \le x_n, \\ 1, & \text{dacă} \quad x_n < x. \end{cases}$$

Graficul acestei funcţii este o funcţie scară. În acest caz funcţia de repartiţie mai poate fi scrisă $F_X(x) = \sum\limits_{i=1}^{n} p_i \theta(x - x_i)$, unde θ este funcţia treaptă unitate a lui Heaviside,

$\theta(x) = \begin{cases} 1, & x \ge 0 \\ 0, & x < 0. \end{cases}$

Teorema 6.1.2 *Următoarele afirmaţii sunt adevărate:*

1. Dacă $x_1 \le x_2$, atunci $F_X(x_1) \le F_X(x_2)$, $\;\forall x_1, x_2 \in \mathbb{R}$;

2. $F_X(x - 0) = F_X(x)$, $\forall x \in \mathbb{R}$ (F_X este continuă la stânga);

3. $\lim\limits_{x \to -\infty} F_X(x) = 0$;

4. $\lim\limits_{x \to +\infty} F_X(x) = 1$.

Demonstraţie. 1.Observăm că dacă $x_1 < x_2$ atunci $\{X < x_1\} \subset \{X < x_2\}$ şi aplicând monotonia funcţiei de probabilitate rezultă

$$F_X(x_1) = P(\{X < x_1\}) \leq P(\{X < x_2\}) = F_X(x_2).$$

2. Pentru orice $x \in \mathbb{R}$ avem $F_X(x-0) = \lim\limits_{n\to\infty} F_X(x - \frac{1}{n}) = \lim\limits_{n\to\infty} P(\left\{X < x - \frac{1}{n}\right\}) = \lim\limits_{n\to\infty} P\left(\bigcap\limits_{n\geq 1} \left\{X < x - \frac{1}{n}\right\}\right) = P(\{X < x\}) = F_X(x)$. (Am folosit şi propoziţia **??** din paragraful Completări şi exerciţii).

3. Fie $B_n = \{X < -n\}$. Atunci $\lim\limits_{x\to-\infty} F_X(x) = \lim\limits_{n\to\infty} F_X(-n) = \lim\limits_{n\to\infty} P(\{X < -n\}) = \lim\limits_{n\to\infty} P(B_n)$. Dar (B_n) este un şir descrescător de evenimente şi $\bigcap\limits_{n\geq 1} B_n = \emptyset$. Rezultă, aplicând propoziţia **??**, $\lim\limits_{x\to-\infty} F_X(x) = \lim\limits_{n\to\infty} P(B_n) = \lim\limits_{n\to\infty} P(\bigcap\limits_{n\geq 1} B_n) = P(\emptyset) = 0$.

4. Fie $A_n = \{X < n\}$. Observăm că (A_n) este un şir crescător de evenimente, $\bigcup\limits_{n\geq 1} A_n = \Omega$ (epuizează toate valorile v. a. X) şi datorită continuităţii lui P (propoziţia **??**) obţinem $\lim\limits_{x\to+\infty} F_X(x) = \lim\limits_{n\to+\infty} F_X(n) = \lim\limits_{n\to\infty} P(A_n) = \lim\limits_{n\to\infty} P(\bigcup\limits_{n\geq 1} A_n) = P(\Omega) = 1$. □

Observaţia 6.1.3 *Variabila aleatore nu este definită explicit, nu se vede; cunoaşterea funcţiei reale F_X permite obţinerea de informaţii de tip statistic relativ la comportarea lui X. De reţinut că v. a. distincte pot avea aceeaşi funcţie de repartiţie.*

Teorema 6.1.4 *Fie $F_X : \mathbb{R} \to [0,1]$ funcţia de repartiţie a unei v. a. X. Atunci pentru orice numere reale $a < b$ avem:*

a) $P(\{X \geq a\}) = 1 - F_X(a)$,
$P(\{X \leq a\}) = F_X(a+0)$,
$P(\{X > a\}) = 1 - F_X(a+0)$.
b) $P(\{a \leq X < b\}) = F_X(b) - F_X(a)$,
$P(\{a < X < b\}) = F_X(b) - F_X(a+0)$,
$P(\{a < X \leq b\}) = F_X(b+0) - F_X(a+0)$,
$P(\{a \leq X \leq b\}) = F_X(b+0) - F_X(a)$.

Demonstraţie. a)

$$P(\{X \geq a\}) = P(\overline{\{X < a\}}) = 1 - P(\{X < a\}) = 1 - F_X(a),$$
$$P(\{X \leq a\}) = P(\bigcap_{i\geq 1}\left\{X < a + \frac{1}{n}\right\}) = F_X(a+0),$$
$$P(\{X > a\}) = P(\overline{\{X \leq a\}}) = 1 - P(\{X \leq a\}) = 1 - F_X(a+0).$$

b)

$$\{a \leq X < b\} = \{X < b\} \setminus \{X < a\}, \{X < a\} \subset \{X < b\}$$
$$\Rightarrow P\left(\{a \leq X < b\}\right) = F_X(b) - F_X(a).$$

Analog rezultă şi celelalte relaţii. □

Exerciţiul 1 Să presupunem că pentru o v. a. Y funcţia de repartiţie

$$F_Y(x) = \begin{cases} 0 & \text{pentru } x \leq 0 \\ \dfrac{x^2}{2} & \text{pentru } 0 < x < 1 \\ 1 & \text{pentru } x \geq 1. \end{cases}.$$

Să se calculeze $P(\left\{Y < \frac{1}{2}\right\})$ şi $P(\left\{Y^2 \geq \frac{1}{2}\right\})$.

Soluţie.

$$P(\left\{Y < \frac{1}{2}\right\}) = F_Y(\frac{1}{2}) = \frac{1}{8},$$
$$P(\left\{Y^2 \geq \frac{1}{2}\right\}) = 1 - P(\left\{Y^2 < \frac{1}{2}\right\}) = 1 - P(\left\{-\frac{1}{\sqrt{2}} < Y < \frac{1}{\sqrt{2}}\right\})$$
$$= 1 - F_Y(\frac{1}{\sqrt{2}}) + F_Y(-\frac{1}{\sqrt{2}} + 0) = 1 - \frac{1}{4} + 0 = \frac{3}{4}.$$

6.2 Densitatea de probabilitate

Definiţia 6.2.1 *Spunem că **v. a.** X este de **tip continuu** cu **densitatea de probabilitate** $f_X : \mathbb{R} \to \mathbb{R}$ dacă funcţia de repartiţie $F_X : \mathbb{R} \to \mathbb{R}$ este continuă şi derivabilă pe porţiuni şi $f_X(x) = F'_X(x)$ în orice punct x în care F_X este dervabilă.*

Proprietăţile densităţii de probabilitate

Teorema 6.2.2 *Fie X o v. a. continuă cu densitatea de probabilitate f_X. Atunci:*

$$a)\ f_X(x) \geq 0, \forall x \in \mathbb{R},$$

$$b)\ F_X(x) = \int_{-\infty}^{x} f_X(t)dt, \tag{6.1}$$

$$c)\ \int_{-\infty}^{\infty} f_X(x)dx = 1, \tag{6.2}$$

d) pentru orice numere reale $a < b$ avem:

$$P(\{a \leq X < b\}) = P(\{a < X < b\}) = P(\{a < X \leq b\}) = P(\{a \leq X \leq b\}) = \int_{a}^{b} f_X(x)dx. \tag{6.3}$$

Demonstraţie. a) Deoarece F_X este monoton crescătoare, conform teoremei 6.1.2, deci derivata ei este pozitivă adică $f_X(x) \geq 0, \forall x \in \mathbb{R}$.

b) $f_X(x) = F'_X(x) \Rightarrow \int_{-\infty}^{x} F'_X(x)dt = \int_{-\infty}^{x} f_X(t)dt \Rightarrow F_X(x) - \lim_{x\to-\infty} F_X(x) = \int_{-\infty}^{x} f_X(t)dt \Rightarrow$ $F_X(x) = \int_{-\infty}^{x} f_X(t)dt$.

c) $F_X(x) = \int_{-\infty}^{x} f_X(t)dt \Rightarrow \lim_{x\to+\infty} F_X(x) = \lim_{x\to+\infty} \int_{-\infty}^{x} f_X(t)dt \Rightarrow 1 = \int_{-\infty}^{\infty} f_X(x)dx.$

d) Demonstrăm $P(\{a \leq X < b\}) = \int_{a}^{b} f_X(x)dx.$

Într-avevăr, $P(\{a \leq X < b\}) = F_X(b) - F_X(a) = \int\limits_{-\infty}^{b} f_X(t)dt - \int\limits_{-\infty}^{a} f_X(t)dt = \int\limits_{a}^{b} f_X(t)dt.$

Celelalte egalităţi rezultă folosind teorema 6.1.4, b) şi ţinând seama de continuitatea funcţiei F_X. □

Observaţia 6.2.3 *Densitatea de probabilitate este folosită pentru a calcula aria care reprezintă probabilitatea ca v. a. X să ia valori în intervalul $[a, b]$. De exemplu, în cazul măsurării intensităţii curentului, probabilitatea ca v. a. X (ale căerei valori reprezintă intensitatea curentului) să aibă ca rezultat o valoare care să aparţină intervalului $[14\ mA,\ 15\ mA]$ este integrala densităţii de probabilitate a lui X calculată pe acest interval.*

Exerciţiul 2 Se spune că v. a. X este **repartizată uniform pe intervalul** $[a, b]$, $a < b$, notat $X \sim Uniform\,[a, b]$, dacă ea are densitatea de probabilitate

$$f_X : \mathbb{R} \to \mathbb{R},\ f_X(x) = \begin{cases} \dfrac{1}{b-a}, & \text{dacă } x \in [a, b] \\ 0, & \text{în caz contrar.} \end{cases}.$$

Să se verifice că f_X are proprietăţile densităţii de probabilitate şi să se determine funcţia de repartiţie.

Soluţie. Verificăm proprietăţile a) şi b) ale teoremei 6.2.2. $f_X(x) \geq 0, \forall x \in \mathbb{R}$, este evidentă şi $\int\limits_{-\infty}^{\infty} f_X(x)dx = \int\limits_{a}^{b} \frac{1}{b-a}dx = 1.$

Calculăm funcţia de repartiţie. Pentru $x < a \Rightarrow F_X(x) = \int\limits_{-\infty}^{x} f_X(t)dt = 0,$

Pentu $x \in [a, b] \Rightarrow F_X(x) = \int\limits_{-\infty}^{x} f_X(t)dt = \int\limits_{a}^{x} \frac{1}{b-a}dt = \frac{x-a}{b-a},$

Pentru $x > b \Rightarrow F_X(x) = \int\limits_{-\infty}^{x} f_X(t)dt = \int\limits_{a}^{b} \frac{1}{b-a}dt = 1$, deci

$$F_X(x) = \begin{cases} 0, & \text{dacă } x < a \\ \dfrac{x-a}{b-a} & \text{dacă } x \in [a, b] \\ 1 & \text{dacă } x > b \end{cases}. \blacktriangledown$$

Exerciţiul 3 Se spune că v. a. X este **repartizată exponenţial,** cu notaţia $X \sim Exponential\,[\lambda]$, dacă ea are densitatea de probabilitate

$$f_X : \mathbb{R} \to \mathbb{R},\ f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{dacă } x \geq 0 \\ 0, & \text{dacă } x < 0. \end{cases}.$$

Să se verifice că f_X are proprietăţile densităţii de probabilitate şi să se determine funcţia de repartiţie.

6.3 Caracteristici numerice ale v. a. continue

Definiţia 6.3.1 *Fie X o v. a. având densitatea de probabilitate f. Definim:*

Media $M[X] = \int\limits_{-\infty}^{\infty} xf(x)dx,$

Dispersia $D[X] = \int\limits_{-\infty}^{\infty} (x - M[X])^2 f(x)dx,$

în ipoteza că integralele respective sunt convergente. La fel se definesc, în ipoteza că integralele respective sunt convergente:

Abaterea medie pătratică $\sigma = \sqrt{D[X]},$

Momentul de ordin $k, k \geq 1, M\left[X^k\right] = \int\limits_{-\infty}^{\infty} x^k f(x)dx,$

Momentul centrat de ordin $k, k \geq 1, D\left[X^k\right] = \int\limits_{-\infty}^{\infty} (x - M[X])^k f(x)dx.$

Observaţia 6.3.2 *Dacă* X *este o v. a. continuă şi* h *este o funcţie continuă, ştim că* $Y = h(X)$ *este o v. a. şi* $M[h(X)] = \int\limits_{-\infty}^{\infty} h(x)f(x)dx.$

Observaţia 6.3.3 *Momentul centrat de ordinul trei,* $D[X^3]$*, caracterizează* ***asimetria*** *graficului densităţii de probabilitate a v. a.* X *în raport cu dreapta de ecuaţie* $x = M[X]$*. Momentul centrat de ordinul patru,* $D[X^4]$*, îndică modul în care este aplatizat graficul densităţii de repartiţie a lui* X*, adică dacă acest grafic este mai ascuţit sau mai plat.*

Definiţia 6.3.4 *Se numeşte* ***mod*** *al v. a.* X *acea valoare a v. a. pentru care densitatea de probabilitate are valoare maximă.*

Observaţia 6.3.5 *Dacă* X *este o v. a. de tip continuu cu densitatea de probabilitate* f*, atunci* $f'(x_m) = 0, f''(x_m) < 0.$

Definiţia 6.3.6 *Numim* ***mediana*** *v. a.* X *acea valoare* $x_{0.5}$ *a lui* X *pentru care*
$P(\{\omega \in \Omega \mid X(\omega) \leq x_{0.5}\}) = 0.5.$

Această relaţie se poate scrie, dacă cunoştem densitatea de probabilitate, de forma $F(x_{0.5}) = \int\limits_{-\infty}^{x_{0.5}} f(x)dx = 0.5$, adică dreapta $x = x_{0.5}$ împarte aria cuprinsă între graficul lui f şi axa Ox în două părţi egale.

Propoziţia 6.3.7 (Inegalitatea lui Cebâşev) *Dacă variabila aleatoare* X *este de tip continuu cu media* m *şi dispersia* σ^2*, atunci are loc*

$$P(\{|X - m| < \varepsilon\}) \geq 1 - \frac{\sigma^2}{\varepsilon^2}, \quad \forall \varepsilon > 0, \tag{6.4}$$

sau echivalent

$$P(\{|X - m| \geq \varepsilon\}) < \frac{\sigma^2}{\varepsilon^2} \qquad \forall \varepsilon > 0.$$

Demonstraţie. Dispersia este dată de

$$\sigma^2 = \int_{-\infty}^{\infty} (x - m)^2 f(x)dx = \int_{-\infty}^{m-\varepsilon} (x - m)^2 f(x)dx + \int_{m-\varepsilon}^{m+\varepsilon} (x - m)^2 f(x)dx +$$

$$+ \int_{m+\varepsilon}^{+\infty} (x - m)^2 f(x)dx.$$

Din faptul că $x \notin (m-\varepsilon, m+\varepsilon)$, rezultă $(x-m)^2 \geq \varepsilon^2$ şi dispersia poate fi minorată prin

$$\sigma^2 \geq \varepsilon^2 \left(\int_{-\infty}^{m-\varepsilon} f(x)dx + \int_{m+\varepsilon}^{\infty} f(x)dx \right) =$$

$$= \varepsilon^2 \left(1 - \int_{m-\varepsilon}^{m+\epsilon} f(x)dx \right),$$

deoarece $\int_{-\infty}^{\infty} f(x)dx = 1$. Ultima paranteză se poate scrie

$$\sigma^2 \geq \varepsilon^2(1 - P(\{|X-m| < \varepsilon\})),$$

care este echivalentă cu relaţia (6.4). □

Exerciţiul 4 Timpul mediu de răspuns al unui calculator este de 15 secunde pentru o anumită operaţie, cu abaterea medie pătratică 3 secunde. Estimaţi probabilitatea ca timpul de răspuns să se abată cu mai puţin de 5 secunde faţă de medie. Vom folosi inegalitatea lui Cebâşev. Avem $P(\{|X-15| < 5\}) \geq 1 - \frac{9}{25} = 0.64$.

Definiţia 6.3.8 *Şirul* $\{X_n\}_{n\in\mathbb{N}}$ ***converge în probabilitate*** *către* X *(vom folosi notaţia* $\{X_n\} \xrightarrow{prob} X$*) dacă*

$$\forall \varepsilon > 0 : \lim_{n\to\infty} P(\{\omega \in \Omega \mid |X_n(\omega) - X(\omega)| > \varepsilon\}) = 0 \tag{6.5}$$

sau, echivalent,

$$\forall \varepsilon > 0 : \lim_{n\to\infty} P(\{\omega \in \Omega | |X_n(\omega) - X(\omega)| \leq \varepsilon\}) = 1. \tag{6.6}$$

Corolarul 6.3.9 *Fie* X_n *un şir de v. a. având medie şi dispersie astfel încât* $\lim\limits_{n\to\infty} D[X_n] = 0$. *Atunci* $\{X_n - M[X_n]\} \xrightarrow{prob} 0$ *(convergenţă în probabilitate).*

Demonstraţie. Se spune că şirul de v. a. (X_n) conveerge în probabilitate către X dacă $\forall\varepsilon > 0$, $\lim\limits_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$.

În cazul acestui corolar avem aplicând inegalitatea lui Cebâşev,

$\varepsilon > 0$, $\lim\limits_{n\to\infty} P(|X_n - M[X_n]| \geq \varepsilon) \leq \frac{D[X_n]}{\varepsilon^2} \to 0$. Rezultatul se obţine aplicând lema cleştelui. □

Teorema 6.3.10 (Legea numerelor mari a lui Bernoulli) *Fie* X_n *un şir de v. a. având aceeaşi medie şi dispersiile egal majorate (adică există* $C > 0$ *astfel încât* $\forall n \geq 1, D[X_n] \leq C$*). Atunci* $\left\{ \frac{1}{n}(X_1 + ... + X_n) \right\} \xrightarrow{prob} m$ *(în probabilitate).*

Demonstraţie. Ştim că $M[X_i] = m, D[X_i] \leq C$. Notăm $Y_n = \dfrac{1}{n}(X_1 + ... + X_n)$ şi deoarece v. a. sunt independente avem $M[Y_n] = m, D[Y_n] = \dfrac{1}{n^2} D[X_1 + ... + X_n] \leq \dfrac{C}{n} \to 0$. Aplicând corolarul 6.3.9 rezultă că $Y_n - M[Y_n] = \dfrac{1}{n}(X_1 + ... + X_n) - m \to 0$ (convergenţă în probabilitate). □

Exemplul 6.3.11 Presupunem că se efectuează serii de experienţe Bernoulli cu $P(A) = p$. Pentru orice n notăm

$$X_n = \begin{cases} 1 & \text{dacă la experimentul } n \text{ se produce } A \\ 0 & \text{dacă la experimentul } n \text{ nu se produce } A \end{cases}.$$

Matricea de repartiţie a lui X_n este
$X_n : \begin{pmatrix} 1 & 0 \\ p & 1-p \end{pmatrix}$, deci $M\left[X_n\right] = p, D\left[X_n\right] = p(1-p)$.

Conform legii numerelor mari a lui Bernoulli rezultă că $\frac{1}{n}\left(X_1 + ... + X_n\right) \rightarrow p$. Dar $X_1 + ... + X_n$ reprezintă numărul de apartţii ale lui A după primele n experimente, adică frecvenţa relativă lui A, notată în statistică cu $f_n(A)$, şi deci am arătat că $f_n(A) \xrightarrow{prob} p$, deci un rezultat de stabilitate statistică.