Facultatea de Matematică
Teoria Probabilităților, Semestrul IV
Conf. dr. Lucian MATICIUC

Cursurile 4 – 6

Capitolul II. Variabile aleatoare discrete

II.1 Variabile aleatoare. Generalități

O mărime legată de o experiență aleatoare $\mathscr{E}$ și care ia valori la întâmplare, în funcție de rezultatele experienței, se numește variabilă aleatoare.

Pentru o definiție mai riguroasă, să considerăm un spațiu de probabilitate $(\Omega, \mathscr{F}, \mathbb{P})$.

Definiţia II.1.1 *Funcția $X : \Omega \to \mathbb{R}$ se numește* **variabilă aleatoare** *(pe scurt* **v.a.***) dacă*

$$\{X \le x\} \in \mathscr{F}, \quad \text{pentru orice } x \in \mathbb{R}, \tag{II.1.1}$$

unde, pentru simplitate, s-a notat $\{X \le x\} \overset{\text{def}}{=} \{\omega : X(\omega) \le x\}$.

Definiţia II.1.2 *De fapt, spunem că $X : (\Omega, \mathscr{F}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ este o v.a. dacă este o* **funcție măsurabilă** *între cele două spații măsurabile, i.e.*

$$X^{-1}(B) \in \mathscr{F}, \quad \text{pentru orice } B \in \mathscr{B}(\mathbb{R}).$$

Remarca II.1.3 Într-adevăr, are loc scrierea

$$\{X \le x\} = \{\omega : X(\omega) \le x\} = \{\omega : X(\omega) \in (-\infty, x]\} = X^{-1}((-\infty, x]),$$

unde notația $X^{-1}(A)$ desemnează **preimaginea** (sau **imaginea inversă** a) mulțimii Borel $A \in \mathscr{B}(\mathbb{R})$ prin X.

Prin urmare, Definiția II.1.1 devine: funcția $X : \Omega \to \mathbb{R}$ este o v.a. dacă și numai dacă

$$\text{pentru orice } x \in \mathbb{R}, \quad X^{-1}((-\infty, x]) \in \mathscr{F}.$$

Aceasta înseamnă, echivalent, că $X^{-1}(B) \in \mathscr{F}$, pentru orice $B \in \mathscr{B}(\mathbb{R})$, adică X este o funcție măsurabilă, adică Definiția II.1.1 este echivalentă cu Definiția II.1.2. $\diamond$

Mulțimea $X(\Omega) \subseteq \mathbb{R}$ este mulțimea valorilor v.a. Vom clasifica o v.a. X astfel:

- X este o **v.a. de tip discret** dacă mulțimea $X(\Omega)$ este finită, sau infinită dar numărabilă[1].
- X este o **v.a. de tip continuu** dacă mulțimea $X(\Omega)$ este un interval (posibil nemărginit) sau o reuniune de intervale din $\mathbb{R}$ (un interval poate fi nemărginit).

Exemple de v.a. discrete: v.a. care are drept valori numărul de puncte apărute pe o față a zarului, la aruncarea unui zar; v.a. care are drept valori numărul de apeluri zilnice primite de un Serviciu Clienți al unei firme; v.a. care are drept valori numărul de cutremure apărute într-o anumită zonă seismică.

Exemple de v.a. de tip continuu: v.a. ale cărei valori reprezintă timpul de funcționare a unui aparat, până la prima defectare; v.a. ale cărei valori reprezintă durata de viață a unui organism; v.a. ale cărei valori reprezintă mărimile erorilor comise la efectuarea măsurătorilor.

Definiţia II.1.4 *Fie $X : \Omega \to \mathbb{R}$ o v.a. Funcția*

$$\mathbb{P}_X : \mathscr{B}(\mathbb{R}) \to [0,1],$$

definită prin

$$\mathbb{P}_X(B) \stackrel{\text{def}}{=} \mathbb{P}(X \in B) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in B\}), \quad B \in \mathscr{B}(\mathbb{R}), \tag{II.1.2}$$

se numește **legea** *sau* **repartiția** *sau* **distribuția** *v.a. X.*

Remarca II.1.5 Putem scrie și sub forma

$$\mathbb{P}_X(B) = \mathbb{P}\left(X^{-1}(B)\right) = \left(\mathbb{P} \circ X^{-1}\right)(B), \quad B \in \mathscr{B}(\mathbb{R}).$$

$\diamond$

Remarca II.1.6 În *Teoria probabilităților* suntem interesați nu de valorile pe care le poate lua o v.a., ci de probabilitățile ca X să ia valori într-o anumită mulțime $B \in \mathscr{B}(\mathbb{R})$, adică $\mathbb{P}(X \in B)$, pentru $B \in \mathscr{B}(\mathbb{R})$, deci de legea v.a. X. $\diamond$

Remarca II.1.7 Se poate demonstra că dacă $\mathbb{P}_X$ este legea unei v.a. X, atunci $\mathbb{P}_X : (\mathbb{R}, \mathscr{B}(\mathbb{R})) \to \mathbb{R}$ este o probabilitate, deci tripletul $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \mathbb{P}_X)$ devine un spațiu de probabilitate. $\diamond$

Având în vedere că $\mathscr{B}(\mathbb{R})$ poate fi generată de intervale de tipul $(-\infty, x]$, adică

$$\mathscr{B}(\mathbb{R}) = \sigma(\{(-\infty, x] : x \in \mathbb{R}\}),$$

a cunoaște legea unei v.a. este echivalent cu a cunoaște $\mathbb{P}_X\big((-\infty, x]\big)$. Astfel obținem următoarea noțiune.

[1] Spunem că o mulțime este numărabilă dacă are cardinalul mai mic sau egal cu $\aleph_0$. Prin urmare, o v.a. X spunem că este discretă dacă are $\operatorname{card}(X(\Omega))$ finit sau dacă are $\operatorname{card}(X(\Omega)) = \aleph_0$.

Se poate arăta: $\operatorname{card}(\mathbb{N}) = \operatorname{card}(\mathbb{N}^*) = \operatorname{card}(\mathbb{Z}) = \operatorname{card}(\mathbb{Q}) = \aleph_0$.

Definiţia II.1.8 *Fie $X:\Omega\to\mathbb{R}$ o v.a. Funcția $F_X:\mathbb{R}\to[0,1]$, definită prin*

$$\boxed{F_X(x)\stackrel{\text{def}}{=}\mathbb{P}(X\le x),\quad x\in\mathbb{R},}\tag{II.1.3}$$

se numește **funcția de repartiție** *(sau* **funcția de distribuție** *sau* **funcția de distribuție cumulativă***) asociată v.a. X.*

Remarca II.1.9 Putem scrie și sub forma

$$F_X(x)=\mathbb{P}\big(X^{-1}((-\infty,x])\big)=\mathbb{P}_X\big((-\infty,x]\big).$$

$\diamond$

Remarca II.1.10 Prin urmare, determinarea funcției de repartiție F_X, asociată v.a. X, calculată în punctul $x\in\mathbb{R}$, înseamnă determinarea probabilității evenimentului ca v.a. X să ia valori mai mici decât x. $\diamond$

Remarca II.1.11 *Teoria probabilităților* se ocupă cu studiul legilor v.a. sau, echivalent, al funcțiilor de repartiție, și al proprietăților acestora. Tocmai această concentrare asupra studiului repartițiilor este ceea ce distinge *Teoria probabilităților* de *Teoria măsurii.* $\diamond$

Propoziţia II.1.12 *Au loc următoarele* **proprietăți caracteristice ale funcției de repartiție***:*

(i) *Funcția F_X este monoton nedescrescătoare (i.e. crescătoare dar nu neapărat strict crescătoare).*

(ii) *Funcția F_X este continuă la dreapta în orice punct $x\in\mathbb{R}$, i.e.*

$$F_X(x)=F_X(x+0),\quad \textit{pentru orice } x\in\mathbb{R},$$

unde $F_X(x+0)\stackrel{\text{def}}{=}\lim_{\substack{y\to x\\ y>x}}F_X(y)$.

(iii) *Au loc și*

$$\boxed{F_X(-\infty)=0\quad \textit{și}\quad F_X(+\infty)=1,}\tag{II.1.4}$$

unde $F_X(-\infty)\stackrel{\text{def}}{=}\lim_{x\to-\infty}F_X(x)$ *și* $F_X(+\infty)\stackrel{\text{def}}{=}\lim_{x\to+\infty}F_X(x)$.

Remarca II.1.13 Se poate demonstra că dacă $F:\mathbb{R}\to[0,1]$ este o funcție reală care satisface cele trei proprietăți de mai sus, atunci există un spațiu de probabilitate $(\Omega,\mathscr{F},\mathbb{P})$ și o v.a. X definită pe acesta astfel încât F este funcția de repartiție asociată lui X, i.e. $F=F_X$. $\diamond$

Demonstraţie. (i) Într-adevăr[2], dacă $x<y$, atunci $\{X\le x\}\subseteq\{X\le y\}$, deci $F_X(x)=\mathbb{P}(X\le x)\le\mathbb{P}(X\le y)=F_X(y)$.

[2] Aceste demonstrații **nu sunt pentru Examen !**

(ii) Fie șirul $(x_n)_{n\in\mathbb{N}^*}$ astfel încât $x_n \searrow x$, pentru $n \to \infty$. Atunci, pentru orice n, avem $\{X \le x_n\} \supseteq \{X \le x_{n+1}\}$ iar

$$\bigcap_{n\in\mathbb{N}^*} \{X \le x_n\} = \{X \le x\},$$

adică

$$\{X \le x_n\} \searrow \{X \le x\}.$$

Folosind continuitatea probabilității $\mathbb{P}$ în raport cu șiruri monotone de evenimente obținem $F_X(x_n) = \mathbb{P}(X \le x_n) \searrow \mathbb{P}(X \le x) = F_X(x)$.

(iii) Fie șirul $(x_n)_{n\in\mathbb{N}^*}$ astfel încât $x_n \to -\infty$, pentru $n \to \infty$. Atunci șirul de evenimente $(\{X \le x_n\})_{n\in\mathbb{N}^*}$ este necrescător iar $\bigcap_{n\in\mathbb{N}^*} \{X \le x_n\} = \varnothing$, adică

$$\{X \le x_n\} \searrow \varnothing.$$

Astfel obținem $F_X(x_n) = \mathbb{P}(X \le x_n) \searrow \mathbb{P}(\varnothing) = 0$.

Fie șirul $(y_n)_{n\in\mathbb{N}^*}$ astfel încât $y_n \to +\infty$, pentru $n \to \infty$. Atunci șirul de evenimente $(\{X \le y_n\})_{n\in\mathbb{N}^*}$ este nedescrescător iar $\bigcup_{n\in\mathbb{N}^*} \{X \le y_n\} = \Omega$, adică

$$\{X \le y_n\} \nearrow \Omega.$$

Astfel obținem $F_X(y_n) = \mathbb{P}(X \le y_n) \nearrow \mathbb{P}(\Omega) = 1$. ■

Remarca II.1.14 Egalitatea

$$\mathbf{P}\big((-\infty, x]\big) = F(x), \quad \text{pentru orice } x \in \mathbb{R},$$

stabilește o corespondență biunivocă între măsurile de probabilitate $\mathbf{P}$ pe $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ și funcțiile de repartiție F (definite de cele trei proprietăți din Propoziția II.1.12). ◇

Propoziţia II.1.15 *Au loc următoarele* **proprietăți suplimentare ale funcției de repartiție***:*

(iv) *Funcția F_X admite o mulțime cel mult numărabilă de puncte de discontinuitate (care sunt doar de specia întâi).*

(v) *Funcția F_X admite limită la stânga în orice punct $x \in \mathbb{R}$, i.e.*

$$\text{există } F_X(x-0) \stackrel{\text{def}}{=\!=} \lim_{\substack{y\to x\\ y<x}} F_X(y).$$

(vi) *Au loc următoarele formule, pentru orice $a \in \mathbb{R}$:*

$$\boxed{\begin{aligned} &\mathbb{P}(X < a) = F_X(a-0), \\ &\mathbb{P}(X > a) = 1 - F_X(a), \\ &\mathbb{P}(X \ge a) = 1 - F_X(a-0). \end{aligned}}$$

(vii) *Pentru orice* $a \in \mathbb{R}$:

$$\mathbb{P}(X = a) = F_X(a) - F_X(a - 0). \tag{II.1.5}$$

Prin urmare,

$$\mathbb{P}(X = a) = 0 \quad \Leftrightarrow \quad F_X \text{ este continuă în } a \tag{II.1.6}$$

și, deoarece F *admite un număr cel mult numărabil de salturi, deducem că mulțimea punctelor* a *pentru care* $\mathbb{P}(X = a) \neq 0$ *este și ea cel mult numărabilă:*

$$\{a \in \mathbb{R} : \mathbb{P}(X = a) > 0\} \quad \text{este cel mult numărabilă.}$$

$(viii)$ *Au loc următoarele formule, pentru orice* $-\infty \leq a < b \leq +\infty$:

$$\begin{aligned} \mathbb{P}(a < X \leq b) &= F_X(b) - F_X(a), \\ \mathbb{P}(a < X < b) &= F_X(b - 0) - F_X(a), \\ \mathbb{P}(a \leq X < b) &= F_X(b - 0) - F_X(a - 0), \\ \mathbb{P}(a \leq X \leq b) &= F_X(b) - F_X(a - 0). \end{aligned} \tag{II.1.7}$$

Evident, formulele precedente se pot scrie utilizând legea v.a. $\mathbb{P}_X$ *a unei v.a.* X*; astfel obținem legătura dintre legea* $\mathbb{P}_X$ *și funcția de repartiție, valabilă pentru orice* $-\infty \leq a < b \leq +\infty$:

$$\mathbb{P}_X\big((a, b]\big) = F_X(b) - F_X(a). \tag{II.1.8}$$

(ix) *Prin urmare, dacă* F *este continuă, atunci, pentru orice* $-\infty \leq a < b \leq +\infty$,

$$\begin{aligned} F_X(b) - F_X(a) &= \mathbb{P}(a < X \leq b) \\ &= \mathbb{P}(a < X < b) \\ &= \mathbb{P}(a \leq X < b) \\ &= \mathbb{P}(a \leq X \leq b). \end{aligned} \tag{II.1.9}$$

Demonstraţie. (iv) Având în vedere că funcția F_X este nedescrescătoare, obținem că mulțimea punctelor de discontinuitate este cel mult numărabilă.

(v) Fie șirul $(x_n)_{n\in\mathbb{N}^*}$ astfel încât $x_n \nearrow x$, pentru $n \to \infty$. Atunci, pentru orice n, avem $\{X \leq x_n\} \subseteq \{X \leq x_{n+1}\}$ iar

$$\bigcup\nolimits_{n\in\mathbb{N}^*} \{X \leq x_n\} = \{X < x\},$$

adică

$$\{X \leq x_n\} \nearrow \{X < x\}.$$

Folosind continuitatea probabilității $\mathbb{P}$ în raport cu șiruri monotone de evenimente obținem $F_X(x_n) = \mathbb{P}(X \leq x_n) \nearrow \mathbb{P}(X < x)$, deci există $F_X(x-0)$ iar $F_X(x-0) = \mathbb{P}(X < x)$.

(vi) Pentru demonstrație sunt utile egalitățile:

$$\{X \leq a\} \cup \{X > a\} = \Omega, \quad \{X < a\} \cup \{X \geq a\} = \Omega.$$

(vii) Pentru demonstrație este utilă egalitatea:

$$\{X < a\} \cup \{X = a\} = \{X \leq a\}.$$

$(viii)$ Pentru demonstrație sunt utile egalitățile:

$$\{X \leq a\} \cup \{a < X \leq b\} = \{X \leq b\}, \quad \{X \leq a\} \cup \{a < X < b\} = \{X < b\},$$

$$\{X < a\} \cup \{a \leq X < b\} = \{X < b\}, \quad \{X < a\} \cup \{a \leq X \leq b\} = \{X \leq b\}.$$

■

Definiţia II.1.16 *Spunem că v.a.* $X, Y : \Omega \to \mathbb{R}$ *sunt* **identic distribuite** *sau* **identic repartizate** *sau* **egale în lege**, *notat*

$$X \overset{\mathrm{d}}{=} Y \quad \textit{sau} \quad X \overset{\mathrm{lege}}{=\!=} Y,$$

dacă au aceași lege, i.e.

$$\mathbb{P}_X = \mathbb{P}_Y, \quad \textit{pe } \mathscr{B}(\mathbb{R}),$$

sau, echivalent, dacă au aceeași funcția de repartiție, i.e.

$$F_X = F_Y, \quad \textit{pe } \mathbb{R},$$

sau, echivalent, dacă urmează aceeași distribuție.

Remarca II.1.17 Evident, în cazul a două v.a. discrete, cele două v.a. sunt identic distribuite dacă au același tablou de repartiție. ◇

Definiţia II.1.18 *Spunem că v.a.* X, Y *sunt* **egale aproape sigur în raport cu măsura de probabilitate** $\mathbb{P}$ *și scriem*

$$X \overset{\mathrm{a.s.}}{=} Y \qquad \textit{sau} \qquad X = Y, \quad \mathbb{P}-\textit{a.s.},$$

dacă[3] *există evenimentul* $N \in \mathscr{F}$ *cu* $\mathbb{P}(N) = 0$ *astfel încât* $X(\omega) = Y(\omega)$, *pentru orice* $\omega \in \Omega \setminus N$.

[3] Fie spațiul de probabilitate $(\Omega, \mathscr{F}, \mathbb{P})$. Spunem că o proprietate $\mathscr{A}$ are loc $\mathbb{P}-$**aproape sigur** (notat $\mathbb{P}-$**a.s.**) dacă există evenimentul $N \in \mathscr{F}$ cu $\mathbb{P}(N) = 0$ astfel încât proprietatea $\mathscr{A}$ are loc pentru orice $\omega \in \Omega \setminus N$.

Să observăm că mulțimea $M \subseteq N$ pentru care proprietatea $\mathscr{A}$ nu are loc nu este neapărat neglijabilă deoarece M nu trebuie, în mod necesar, să fie eveniment, i.e. $M \in \mathscr{F}$. Prin urmare, știm că există M, N astfel încât $M \subseteq N$, $N \in \mathscr{F}$, cu $\mathbb{P}(N) = 0$, iar proprietatea $\mathscr{A}$ are loc pentru orice $\omega \in (\Omega \setminus N) \cup (N \setminus M)$ și nu are loc pentru orice $\omega \in M$.

Astfel, $X = Y, \quad \mathbb{P}-$*a.s., înseamnă*

$$\mathbb{P}(X = Y) = \mathbb{P}(\{\omega : X(\omega) = Y(\omega)\}) = 1 \quad \text{sau, echivalent,}$$

$$\mathbb{P}(X \neq Y) = \mathbb{P}(\{\omega : X(\omega) \neq Y(\omega)\}) = 0.$$

Lema II.1.19 *Fie două v.a.* $X, Y : \Omega \to \mathbb{R}$. *Dacă* X, Y *sunt egale* $\mathbb{P}-$*a.s., atunci cele două v.a. sunt identic distribuite:*

$$\boxed{X \overset{\text{a.s.}}{=\!=} Y \quad \Rightarrow \quad X \overset{\text{d}}{=} Y}$$

sau

$$X = Y, \quad \mathbb{P} - a.s. \quad \Rightarrow \quad F_X = F_Y\,, \quad \text{pe } \mathbb{R}.$$

Evident, reciproca nu este adevărată: putem să avem două v.a. diferite $\mathbb{P}-$*a.s. (pot fi diferite chiar în orice* $\omega \in \Omega$*) dar care să aibă aceeași lege.*

Definiţia II.1.20 *Dacă* $X : \Omega \to \mathbb{R}$ *este o v.a., atunci definim* $\sigma-$*algebra generată de v.a.* X *prin*

$$\sigma(X) \overset{\text{def}}{=\!=} \left\{X^{-1}(B) : B \in \mathscr{B}(\mathbb{R})\right\}.$$

Aceasta este cea mai mică $\sigma-$*algebră în raport cu care* X *este funcție măsurabilă.*

Remarca II.1.21 Se poate arăta că $\sigma-$algebra generată de v.a. X este dată de

$$\sigma(X) = \sigma\left(\left\{X^{-1}((-\infty, x]) : x \in \mathbb{R}\right\}\right).$$

$\diamond$

Definiţia II.1.22 *Spunem că v.a.* $X_i : \Omega \to \mathbb{R}$, $i = \overline{1,n}$, *sunt independente în ansamblu dacă* $\sigma-$*algebrele* $\sigma(X_i)\,, i = \overline{1,n}$, *generate de v.a. date, sunt* $\sigma-$*algebre independente în ansamblu, i.e.*

$$\mathbb{P}\left(\bigcap\nolimits_{i=1}^{n} C_i\right) = \prod\nolimits_{i=1}^{n} \mathbb{P}(C_i)\,,$$

pentru orice evenimente $C_i \in \sigma(X_i)\,, i = \overline{1,n}$.

Având în vedere că $\sigma(X_i) = \{X_i^{-1}(B) : B \in \mathscr{B}(\mathbb{R})\} = \{X_i \in B : B \in \mathscr{B}(\mathbb{R})\}$, deducem că $C_i \in \sigma(X_i)$ dacă există $B_i \in \mathscr{B}(\mathbb{R})$ astfel încât $C_i = \{X_i \in B_i\}$, prin urmare putem defini și în modul următor.

Definiţia II.1.23 *Spunem că v.a.* $X_i : \Omega \to \mathbb{R}$, $i = \overline{1,n}$, *sunt* **independente în ansamblu** *dacă*

$$\mathbb{P}(X_1 \in B_1, \ldots, X_n \in B_n) = \mathbb{P}(X_1 \in B_1) \ldots \mathbb{P}(X_n \in B_n)\,,$$

pentru orice mulțime Borel $B_i \in \mathscr{B}(\mathbb{R})\,, i = \overline{1,n}$.

Definiţia II.1.24 *Spunem că v.a.* $X_i : \Omega \to \mathbb{R}$, $i = \overline{1,n}$, *sunt* **independente două câte două** *dacă*

$$\mathbb{P}(X_i \in B_1, X_j \in B_2) = \mathbb{P}(X_i \in B_1)\ \mathbb{P}(X_j \in B_2),$$

pentru orice $i, j \in \{1, 2, \ldots, n\}$ *și pentru orice mulțimi Borel* $B_1, B_2 \in \mathscr{B}(\mathbb{R})$.

Remarca II.1.25 Evident, dacă v.a. X_i , $i = \overline{1,n}$, sunt independente în sensul Definiției II.1.23, atunci ele sunt independente două câte două. Reciproca acestei afirmații nu este adevărată. Există exemple în care v.a. sunt independente două câte două dar nu sunt independente în ansamblu. $\diamond$

Remarca II.1.26 Alegând $B_i = (-\infty, a_i]$ obținem caracterizarea: v.a. $X_i : \Omega \to \mathbb{R}$, $i = \overline{1,n}$, sunt independente în ansamblu dacă și numai dacă

$$\mathbb{P}(X_1 \le a_1, \ldots, X_n \le a_n) = \prod\nolimits_{i=1}^{n} \mathbb{P}(X_i \le a_i),$$

pentru orice $a_1, a_2, \ldots, a_n \in \mathbb{R}$. $\diamond$

II.2 Variabile aleatoare discrete cu un număr finit de valori

Cunoașterea legii $\mathbb{P}_X : \mathscr{B}(\mathbb{R}) \to [0,1]$, dată de $\mathbb{P}_X(B) = \mathbb{P}(X \in B)$, unde $B \in \mathscr{B}(\mathbb{R})$, asociată unei v.a. discrete X, cu un număr finit sau infinit de valori, este echivalentă cu cunoașterea valorile pe care le poate lua v.a. (adică a mulțimii $X(\Omega)$) și a probabilităților cu care este luată fiecare valoare din $X(\Omega)$.

Putem reprezenta aceste informații sub forma unui **tabloul de repartiție**:

$$X : \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ p_1 & p_2 & \ldots & p_n \end{pmatrix} \quad \text{sau} \quad X : \begin{pmatrix} x_i \\ p_i \end{pmatrix}_{i=\overline{1,n}}, \tag{II.2.1}$$

unde $X(\Omega) = \{x_1, x_2, \ldots, x_n\}$ iar $p_i = \mathbb{P}(X = x_i)$, cu $i = \overline{1,n}$, deci

$$p_i \ge 0, \quad \text{pentru } i = \overline{1,n} \quad \text{și} \quad \sum\nolimits_{i=1}^{n} p_i = 1$$

(prin convenție, am scris valorile $x_1, x_2, \ldots, x_n$ ale v.a. în ordine strict crescătoare).

Pe scurt, putem scrie că o v.a. discretă X, cu un numar finit de valori, este descrisă complet prin cunoașterea familiei finite de valori

$$\boxed{(p_i)_{i=\overline{1,n}} \overset{\text{not}}{=\!=} (\mathbb{P}(X = x_i))_{i=\overline{1,n}}\, .}$$

Remarca II.2.1 Să menționăm că o funcție discretă $X : \Omega \to \mathbb{R}$, definită pe un spațiu finit de probabilitate $(\Omega, \mathscr{F}, \mathbb{P})$, este v.a. (vezi Definiția II.1.1 sau Definiția II.1.2) dacă și numai dacă $\{X = x\} \in \mathscr{F}$, pentru orice $x \in X(\Omega)$.

Prin urmare, în cazul unei funcții discrete $X : \Omega \to \mathbb{R}$ definită pe un spațiu finit de probabilitate $(\Omega, \mathscr{P}(\Omega), \mathbb{P})$, condiția din Definiția II.1.1 are loc întotdeauna.

Prin urmare, dacă avem $\mathrm{card}(\Omega)$ finit, atunci

$$\text{orice funcție } X : (\Omega, \mathscr{P}(\Omega)) \to (\mathbb{R}, \mathscr{B}(\mathbb{R})) \text{ este v.a.,}$$

adică măsurabilă. $\diamond$

Exemplul II.2.2 Dacă luăm $\mathscr{F} \subsetneq \mathscr{P}(\Omega)$, afirmația precedentă nu mai este adevărată.

Într-adevăr, fie spațiul finit de probabilitate $(\Omega, \mathscr{F}, \mathbb{P})$ cu $\Omega = \{1, 2, 3\}$ și $\mathscr{F} = \{\varnothing, \Omega, \{1, 2\}, \{3\}\}$. Atunci funcția identitate $X : \Omega \to \mathbb{R}$, dată de $X(1) = 1$, $X(2) = 2$, $X(3) = 3$, nu este măsurabilă, deoarece, de exemplu, $\{1\}$ nu este eveniment, mai precis $X^{-1}(\{1\}) = \{1\} \notin \mathscr{F}$. $\diamond$

Exemplul II.2.3 În cazul unei v.a. constante funcția de repartiție se calculează foarte ușor. Mai precis, dacă $X = c$, $\mathbb{P}-$a.s., atunci se obține:

$$F_X(x) = \mathbb{P}(X \leq x) = \begin{cases} 0, & x < c, \\ 1, & x \geq c. \end{cases}$$

$\diamond$

De fapt, funcția de repartiție, dată de Definiția II.1.8, asociată oricărei v.a. discrete, cu un număr finit de valori, este o funcția în scară (adică e constantă pe porțiuni și nedescrescătoare). Deci

$$F_X(x) = \mathbb{P}(X \leq x) = \sum\nolimits_{\substack{x_i \in X(\Omega) \\ x_i \leq x}} \mathbb{P}(X = x_i), \quad \text{pentru orice } x \in \mathbb{R}.$$

De exemplu, dacă $x < x_1$ avem $F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(\varnothing) = 0$.

Apoi, pentru $x_1 \leq x < x_2$, avem

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X \leq x_1) = \mathbb{P}(X = x_1) = p_1.$$

Pentru $x_2 \leq x < x_3$ avem

$$\begin{aligned} F_X(x) &= \mathbb{P}(X \leq x) = \mathbb{P}(X \leq x_2) = \mathbb{P}(X \in \{x_1, x_2\}) \\ &= \mathbb{P}(X = x_1) + \mathbb{P}(X = x_2) = p_1 + p_2. \end{aligned}$$

Astfel, se obține

$$F_X(x) = \begin{cases} 0, & x < x_1\,, \\ p_1\,, & x_1 \leq x < x_2\,, \\ p_1 + p_2\,, & x_2 \leq x < x_3, \\ \vdots & \\ p_1 + \ldots + p_{n-1}\,, & x_{n-1} \leq x < x_n\,, \\ 1 = p_1 + \ldots + p_n\,, & x_n \leq x. \end{cases} \tag{II.2.2}$$

Observăm că, în acest caz, funcția de repartiție F_X este o funcție continuă în orice punct cu excepția $\{x_k\}_{k=\overline{1,n}}$, mai precis, singurele puncte de discontinuitate sunt punctele de salt $\{x_k\}_{k=\overline{1,n}}$.

Pe de altă parte, dacă se dă o funcție de tipul (II.2.2), atunci se poate arăta că ea este o funcție de repartiție și, în plus, putem determina v.a. X căreia funcția îi este asociată. Astfel, probabilitățile $\mathbb{P}(X = x)$ sunt date de (II.1.5):

$$\mathbb{P}(X = x) = F_X(x) - F_X(x - 0)\,, \quad \text{pentru orice } x \in X(\Omega)\,.$$

Prin urmare, în punctul x în care F_X este continuă $\mathbb{P}(X = x) = 0$ iar în punctul x în care F_X are un salt $\mathbb{P}(X = x) \neq 0$, adică X ia doar valorile x în care F_X este discontinuă.

Remarca II.2.4 Deci X este v.a. discretă dacă și numai dacă funcția de repartiție F_X asociată ei verifică proprietățile date de Propoziția II.1.12 și este constantă pe porțiuni.

Exemplul II.2.5 **Funcția indicatoare $\mathbb{1}_A$ a evenimentului** $A \in \mathscr{F}$, dată de

$$\mathbb{1}_A(\omega) \overset{\text{def}}{=\!=} \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A, \end{cases}$$

este o v.a. discretă.

Mai mult, este evident că dacă $A \in \mathscr{P}(\Omega)$ este o mulțime, atunci funcția $\mathbb{1}_A$ este v.a. (adică funcție măsurabilă) dacă și numai dacă A este eveniment, adică $A \in \mathscr{F}$. ◇

Exemplul II.2.6 Să considerăm experiența care constă în aruncarea unui zar. Fie X v.a. ale cărei valori reprezintă numărul de puncte apărute la aruncarea unui zar. Să se scrie tabloul de repartiție și funcția de repartiție.

Mulțimea evenimentelor este $\{1, 2, 3, 4, 5, 6\}$ și ea coincide cu Ω, adică

$$\Omega = \{\omega = i : i \in \{1, 2, \ldots, 6\}\}\,.$$

Atunci $X : \Omega \to \mathbb{R}$ este dată de $X : \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}$.

Funcția de repartiție este $F(x) = \begin{cases} 0, & x < 1, \\ 1/6\,, & 1 \leq x < 2, \\ 2/6\,, & 2 \leq x < 3, \\ \vdots & \\ 5/6\,, & 5 \leq x < 6, \\ 1\,, & 6 \leq x. \end{cases}$

$\diamond$

II.2.1 Operații cu variabile aleatoare[4]

Având în vedere că v.a. sunt funcții definite pe mulțimea evenimentelor elementare și cu valori reale, putem defini suma și produsul a două v.a (ca și la funcții). Astfel, dacă ω este un eveniment elementar și X, Y au același domeniu de definiție, atunci definim (suma, produsul și produsul cu un scalar α)

$$\begin{aligned} &(X+Y)(\omega) = X(\omega) + Y(\omega)\,, \\ &(X \cdot Y)(\omega) = X(\omega)\,Y(\omega)\,, \\ &(\alpha X)(\omega) = \alpha\,X(\omega)\,. \end{aligned}$$

Fie v.a. discrete X, Y cu tablourile

$$X : \begin{pmatrix} x_1 & x_2 & \dots & x_m \\ p_1 & p_2 & \dots & p_m \end{pmatrix}, \quad Y : \begin{pmatrix} y_1 & y_2 & \dots & y_n \\ q_1 & q_2 & \dots & q_n \end{pmatrix},$$

unde $p_i, q_j \in [0,1]$, cu $i = \overline{1,m}\,,\ j = \overline{1,n}\,,$ și $\sum_{i=1}^{m} p_i = 1 = \sum_{j=1}^{n} q_j\,.$

Pentru $i = \overline{1,m}\,,\ j = \overline{1,n}$ să notăm cu

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

Atunci tabloul de repartiție al v.a. $X + Y$ este dat de

$$X + Y : \begin{pmatrix} x_1 + y_1 & x_1 + y_2 & \dots & x_1 + y_n & \dots & x_m + y_n \\ p_{11} & p_{12} & \dots & p_{1n} & \dots & p_{mn} \end{pmatrix},$$

iar pentru produsul $X \cdot Y$ avem

$$X \cdot Y : \begin{pmatrix} x_1 y_1 & x_1 y_2 & \dots & x_1 y_n & \dots & x_m y_n \\ p_{11} & p_{12} & \dots & p_{1n} & \dots & p_{mn} \end{pmatrix}.$$

[4] Operaţiile cu variabile aleatoare **nu sunt prezentate în cadrul Cursului ci doar în cadrul Seminarului (prin exemple concrete)**.

În particular, dacă $Y = a$ este o v.a. constantă, atunci

$$a + X : \begin{pmatrix} a + x_1 & a + x_2 & \ldots & a + x_m \\ p_1 & p_2 & \ldots & p_m \end{pmatrix},$$

$$aX : \begin{pmatrix} ax_1 & ax_2 & \ldots & ax_m \\ p_1 & p_2 & \ldots & p_m \end{pmatrix}.$$

În particular, adunând v.a. aX și b, se obține și

$$aX + b : \begin{pmatrix} ax_1 + b & ax_2 + b & \ldots & ax_m + b \\ p_1 & p_2 & \ldots & p_m \end{pmatrix}.$$

De asemenea, dacă luăm $X = Y$ atunci obținem pătratul unei v.a.

$$X^2 \overset{\text{def}}{=\!=} X \cdot X : \begin{pmatrix} x_1^2 & x_2^2 & \ldots & x_m^2 \\ p_1 & p_2 & \ldots & p_m \end{pmatrix},$$

deoarece în acest caz $X = Y$ și, pentru $i \neq j$,

$$p_{ij} = \mathbb{P}\left(X = x_i, X = x_j\right) = \mathbb{P}\left(\{X = x_i\} \cap \{X = x_j\}\right) = 0,$$

iar pentru $i = j$,

$$p_{ii} = \mathbb{P}\left(X^2 = x_i^2\right) = \mathbb{P}\left(\{X = x_i\} \cap \{X = x_i\}\right) = \mathbb{P}\left(X = x_i\right) = p_i\,.$$

II.2.2 Caracteristici numerice

Fie v.a. X cu tabloul de repartiție

$$X : \begin{pmatrix} x_1 & x_2 & \ldots & x_m \\ p_1 & p_2 & \ldots & p_m \end{pmatrix},$$

unde $p_i \in [0,1]$, cu $i = \overline{1,m}$ și $\sum_{i=1}^m p_i = 1$.

Definiţia II.2.7 *Vom numi* **media** *v.a. discrete X, numărul*

$$\mathbb{E}\left(X\right) \overset{\text{def}}{=\!=} x_1p_1 + x_2p_2 + \ldots + x_mp_m = \sum\nolimits_{i=1}^m x_ip_i\,. \tag{II.2.3}$$

Remarca II.2.8 Definiția precedentă o putem scrie și sub forma

$$\mathbb{E}\left(X\right) = \sum\nolimits_{x \in X(\Omega)} x \cdot \mathbb{P}\left(X = x\right).$$

◇

Remarca II.2.9 Media unei v.a. X se numește și: **valoarea medie** sau **valoarea așteptată** sau **speranța**. ◇

Denumirea de medie este îndreptățită dacă ținem seama de sensul ei practic. Presupunem că am repetat de N ori o experiență care ne-a condus la v.a. X. Dacă fiecare valoare x_i este luată de n_i ori astfel încât $\sum_{i=1}^{m} n_i = N$. Atunci suma valorilor luate de v.a. X este $n_1 x_1 + \ldots + n_m x_m$ iar media aritmetică a valorilor luate de X va fi deci

$$\frac{n_1 x_1 + \ldots + n_m x_m}{N} = \frac{n_1}{N} x_1 + \frac{n_2}{N} x_2 + \frac{n_3}{N} x_3 + \ldots + \frac{n_m}{N} x_m .$$

Dar $\frac{n_i}{N}$ reprezintă raportul dintre numărul cazurilor în care s-a luat valoarea x_i și numărul total de experimentări, adică $p_i \overset{\text{def}}{=\!=} \frac{n_i}{N}$ reprezintă probabilitatea de apariție a valorii x_i .

Media aritmetică a valorilor luate de X este, prin urmare $p_1 x_1 + \ldots + p_m x_m$, adică media lui X.

Media lui X ne arată la ce valoare putem să ne așteptăm pentru media aritmetică a unui mare număr de valori ale lui X, obținute în urma repetării experienței date.

Remarca II.2.10 În cazul în care nu este specificat explicit, următoarele definiții și proprietăți și formule de calcul sunt **valabile în cazul unor v.a. X oarecare (nu neapărat cu un număr numărabil de valori).** $\diamond$

Propoziția II.2.11 (Proprietăți ale mediei)

(i) *Valoarea medie a unei constante*[5] *este egală cu constanta. Mai precis*[6], *pentru* $c \in \mathbb{R}$,

$$X = c, \quad \mathbb{P} - a.s. \quad \Rightarrow \quad \mathbb{E}(X) = c.$$

(ii) *Pentru orice v.a. X, Y care admit medie*[7] *și pentru orice $a, b \in \mathbb{R}$,*

$$\mathbb{E}(aX + bY) = a\,\mathbb{E}(X) + b\,\mathbb{E}(Y) .$$

(iii) *V.a. dată de funcția indicatoare $\mathbb{1}_A$ a evenimentului $A \in \mathscr{F}$ are media*

$$\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A) .$$

(iv) *Pentru orice v.a. X care admite medie, are loc*

$$\big|\mathbb{E}(X)\big| \leq \mathbb{E}\big(|X|\big). \tag{II.2.4}$$

(vi) *Dacă X, Y sunt două v.a. astfel încât X admite medie, atunci*

$$X = Y, \quad \mathbb{P} - a.s. \quad \Rightarrow \quad \textit{există și } \mathbb{E}(Y) \textit{ iar } \quad \mathbb{E}(X) = \mathbb{E}(Y) .$$

[5] De fapt, putem vedea c ca pe o v.a. $X : \begin{pmatrix} c \\ 1 \end{pmatrix}$; prin urmare $\mathbb{E}(X) = c \cdot 1 = c$.

[6] Pentru semnificația lui $X = c$, $\mathbb{P}-$a.s., vezi Nota 3.

[7] În cazul a două v.a. oarecare (nu neapărat cu un număr finit de valori), trebuie impusă condiția ca cele două v.a. să admită medie finită.

(vii) *Dacă v.a. X admite medie, atunci*

$$X \geq 0, \quad \mathbb{P}-a.s. \quad \Rightarrow \quad \mathbb{E}\left(X\right) \geq 0$$

și, în plus, dacă X are valori nenegative $\mathbb{P}$–a.s., atunci

$$\mathbb{E}\left(X\right) = 0 \quad \Leftrightarrow \quad X = 0, \quad \mathbb{P}-a.s.. \tag{II.2.5}$$

Prin urmare, dacă X^2 admite medie, atunci

$$\mathbb{E}\left(X^2\right) = 0 \quad \Leftrightarrow \quad X = 0, \quad \mathbb{P}-a.s.$$

(vezi și punctul (ii) din Propoziția II.2.31).

$(viii)$ *Evident, conform (vii), dacă X, Y sunt două v.a. care admit medie, atunci*

$$X \leq Y, \quad \mathbb{P}-a.s. \quad \Rightarrow \quad \mathbb{E}\left(X\right) \leq \mathbb{E}\left(Y\right).$$

(xii) *Dacă două v.a. sunt independente care admit medie, atunci media produsului este produsul mediilor, i.e.*

$$X, Y \text{ independente} \quad \Rightarrow \quad \mathbb{E}\left(XY\right) = \mathbb{E}\left(X\right)\mathbb{E}\left(Y\right). \tag{II.2.6}$$

Remarca II.2.12 Relația (II.2.6) oferă o condiție necesară pentru independența a două v.a. dar ea nu este și suficientă. În acest sens există exemple (vezi și Remarca II.2.28). $\diamond$

Următorul rezultat arată că independența a două v.a. implică independența oricăror funcții aplicate acestor v.a..

Propoziția II.2.13 *Dacă $X, Y : \Omega \to \mathbb{R}$ sunt două v.a., atunci*

$$X, Y \text{ independente} \quad \Leftrightarrow \quad g\left(X\right), h\left(Y\right) \text{ independente,} \\ \text{pentru orice } g, h : \mathbb{R} \to \mathbb{R}.$$

Definiţia II.2.14 *Se numește* **moment de ordin** $r > 0$ *al v.a. X, media v.a. X^r (dacă este bine definită). Vom nota*

$$\mu_r \stackrel{\text{not}}{=\!=} \mathbb{E}(X^r).$$

Deci, în cazul unei v.a. discrete X, cu un număr finit de valori,

$$\mu_r = \sum\nolimits_{i=1}^{m} x_i^r\, p_i = \sum\nolimits_{x \in X(\Omega)} x^r \cdot \mathbb{P}\left(X = x\right).$$

Evident, momentul de ordin 1 este exact media v.a. X, i.e. $\mu_1 = \mathbb{E}\left(X\right).$

Definiţia II.2.15 *Se numește* **moment central de ordin** $r > 0$ *al v.a.* X*, media v.a.* $(X - \mu_1)^r$ *(dacă este bine definită). Vom nota*

$$\nu_r \overset{\text{not}}{=} \mathbb{E}\big[\,(X - \mu_1)^r\,\big] = \mathbb{E}\big[\,(X - \mathbb{E}\,(X))^r\,\big].$$

Deci, în cazul unei v.a. discrete X*, cu un număr finit de valori,*

$$\nu_r = \sum\nolimits_{i=1}^{m} (x_i - \mu_1)^r\, p_i = \sum\nolimits_{x \in X(\Omega)} (x - \mathbb{E}(X))^r \cdot \mathbb{P}\,(X = x)\ .$$

Definiţia II.2.16 *V.a.* $X - \mathbb{E}\,(X)$ *se numește v.a. centrată. Evident, media v.a. centrate* $X - \mathbb{E}\,(X)$ *este nulă, i.e.* $\mathbb{E}\,(X - \mathbb{E}\,(X)) = \mathbb{E}\,(X) - \mathbb{E}\,(X) = 0.$

Definiţia II.2.17 *Se numește* **moment absolut de ordin** $r > 0$ *al v.a.* X*, media v.a.* $|X|^r$*. Deci, în cazul unei v.a. discrete* X*, cu un număr finit de valori,*

$$\mathbb{E}(|X|^r) = \sum\nolimits_{i=1}^{m} |x_i|^r p_i = \sum\nolimits_{x \in X(\Omega)} |x|^r \cdot \mathbb{P}\,(X = x)\ .$$

Media unei funcții aplicată unei v.a. poate fi calculată folosind următorul rezultat.

Teorema II.2.18 (Formula de transfer) *Fie* X *o .v.a. discretă cu tabloul de repartiție*

$$X : \begin{pmatrix} x_1 & x_2 & \ldots & x_m \\ p_1 & p_2 & \ldots & p_m \end{pmatrix}.$$

Fie $h : \mathbb{R} \to \mathbb{R}$ *o funcție măsurabilă. Atunci* $h\,(X) : \Omega \to \mathbb{R}$ *este și ea o v.a. și are media dată de:*

$$\mathbb{E}\,(h\,(X)) = \sum\nolimits_{i=1}^{m} h\,(x_i)\ p_i\,. \tag{II.2.7}$$

Remarca II.2.19 Formula precedentă o putem scrie și sub forma

$$\mathbb{E}\,(h(X)) = \sum\nolimits_{x \in X(\Omega)} h(x) \cdot \mathbb{P}\,(X = x)\,.$$

$\diamond$

De multe ori la o v.a. ne interesează cât de mult se abat valorile variabilei de la valoarea medie. Trebuie să stabilim un indicator al împrăștierii valorilor v.a. în jurul valorii medii. Valoarea medie a abaterii este zero deci nu poate caracteriza această împrăștiere. Foarte utilă va fi folosirea cantității $\mathbb{E}\big(\,(X - \mu_1)^2\,\big)$.

Definiţia II.2.20 *Se numește* **dispersia** *(sau* **varianța***) v.a.* X*, notată cu* $\mathrm{D}^2\,(X)$ *(sau cu* $\mathrm{Var}\,(X)$ *), momentul central de ordin* 2 *al v.a.* X*, i.e.*[8]

$$\mathrm{D}^2\,(X) = \mathrm{Var}(X) \overset{\text{def}}{=} \nu_2 = \mathbb{E}\,(X - \mu_1)^2\,, \quad \text{unde } \mu_1 = \mathbb{E}\,(X)\,.$$

[8] În cazul unei v.a. oarecare (nu neapărat cu un număr finit de valori), trebuie impusă condiția ca v.a. X^2 să admită medie, i.e. $\mathbb{E}\,(X^2) < \infty$.

Deci, în cazul unei v.a. discrete X,

$$\mathrm{D}^2(X) = \sum\nolimits_{i=1}^{m} (x_i - \mu_1)^2 p_i = \sum\nolimits_{x \in X(\Omega)} (x - \mathbb{E}(X))^2 \cdot \mathbb{P}(X = x) \ .$$

Menționăm că dispersia unei v.a. X este, într-un anume sens, cea mai bună valoare care caracterizează împrăștierea valorilor $x_1, \ldots, x_m$ față de medie.

Remarca II.2.21 Deoarece v.a. $X - \mathbb{E}(X)$ are media zero, obținem și proprietatea

$$\mathrm{D}^2(X - \mathbb{E}(X)) = \mathbb{E}\left[(X - \mathbb{E}(X)) - \mathbb{E}(X - \mathbb{E}(X))\right]^2 = \mathrm{D}^2(X)$$

(dispersia v.a. centrate coincide cu dispersia v.a.). $\diamond$

Remarca II.2.22 Având în vedere formula de transfer (II.2.7) obținem

$$\begin{aligned}\mathrm{D}^2(X) &= \mathbb{E}(X - \mu_1)^2 = \sum\nolimits_{i=1}^{m} p_i (x_i - \mu_1)^2 = \sum\nolimits_{i=1}^{m} p_i x_i^2 - 2\sum\nolimits_{i=1}^{m} p_i x_i \mu_1 + \sum\nolimits_{i=1}^{m} p_i \mu_1^2 \\ &= \sum\nolimits_{i=1}^{m} p_i x_i^2 - 2\mu_1 \sum\nolimits_{i=1}^{m} p_i x_i + \mu_1^2 \sum\nolimits_{i=1}^{m} p_i = \sum\nolimits_{i=1}^{m} p_i x_i^2 - 2\mu_1\mu_1 + \mu_1^2 \cdot 1 = \sum\nolimits_{i=1}^{m} p_i x_i^2 - \mu_1^2 ,\end{aligned}$$

deci are loc **formula de calcul a dispersiei**:

$$\mathrm{D}^2(X) = \mathbb{E}\left(X^2\right) - (\mathbb{E}(X))^2 .$$

$\diamond$

Remarca II.2.23 Formula precedentă este adevărată și în cazul unei v.a. X oarecare (nu neapărat discretă). Într-adevăr,

$$\begin{aligned}\mathrm{D}^2(X) &= \mathbb{E}(X - \mathbb{E}(X))^2 = \mathbb{E}\left[X^2 - 2\mathbb{E}(X) \cdot X + (\mathbb{E}(X))^2\right] \\ &= \mathbb{E}\left(X^2\right) - 2\mathbb{E}\left[\mathbb{E}(X) \cdot X\right] + \mathbb{E}\left[(\mathbb{E}(X))^2\right] \\ &= \mathbb{E}\left(X^2\right) - 2\mathbb{E}(X) \cdot \mathbb{E}(X) + (\mathbb{E}(X))^2 = \mathbb{E}\left(X^2\right) - (\mathbb{E}(X))^2 .\end{aligned}$$

$\diamond$

Definiţia II.2.24 *Se numește* **covarianța** *v.a. X și Y, notată* $\mathrm{Cov}(X, Y)$, *media*

$$\mathrm{Cov}(X, Y) \stackrel{\text{def}}{=\!=} \mathbb{E}\left[(X - \mathbb{E}(X)) \cdot (Y - \mathbb{E}(Y))\right] . \tag{II.2.8}$$

Remarca II.2.25 Dacă facem calculele, obținem **formula de calcul a covarianței**:

$$\mathrm{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) . \tag{II.2.9}$$

$\diamond$

Remarca II.2.26 Evident, covarianța dintre v.a. X cu ea însăși este chiar dispersia, deci putem lua drept definiție:

$$\mathrm{D}^2(X) \stackrel{\text{def}}{=} \mathrm{Cov}(X, X).$$

$\diamond$

Definiţia II.2.27 *Dacă* $\mathrm{D}^2(X) > 0$ *și* $\mathrm{D}^2(Y) > 0$, *atunci*

$$\mathrm{Cor}(X, Y) = \rho(X, Y) \stackrel{\text{def}}{=} \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{D}^2(X) \cdot \mathrm{D}^2(Y)}}$$

se numește **corelația** *(sau* **coeficientul de corelație** *al) v.a.* X *și* Y*.*

Remarca II.2.28 Dacă două v.a. X, Y sunt independente, atunci, folosind (II.2.6), obținem covarianța (sau, echivalent, corelația) nulă, i.e. $\mathrm{Cov}(X, Y) = 0 = \rho(X, Y)$. În acest caz spunem că cele două v.a. sunt **necorelate**.

Deci

$$\text{dacă două v.a. sunt independente, atunci sunt necorelate.} \tag{II.2.10}$$

Reciproca nu este adevărată; în acest sens vezi Exemplul II.2.29. $\diamond$

Exemplul II.2.29 Se poate arăta că v.a.

$$X : \begin{pmatrix} -1 & 0 & 1 \\ 1/3 & 1/3 & 1/3 \end{pmatrix} \quad \text{și} \quad Y = \begin{cases} 1, & \text{dacă } X = 0, \\ 0, & \text{în rest} \end{cases}$$

sunt două v.a. necorelate, dar ele sunt, evident, dependente.

Într-adevăr, obținem

$$Y : \begin{pmatrix} 0 & 1 \\ 2/3 & 1/3 \end{pmatrix}, \quad \text{deci} \quad X \cdot Y : \begin{pmatrix} -1 & 0 & 1 \\ 1/9 & 7/9 & 1/9 \end{pmatrix}$$

și $\mathbb{E}(X) = 0$, $\mathbb{E}(Y) = 1/3$ iar $\mathbb{E}(X \cdot Y) = 0$. $\diamond$

Remarca II.2.30 Se poate arăta că

$$|\rho(X, Y)| \leq 1$$

și că

$$|\rho(X, Y)| = 1 \quad \text{dacă și numai dacă}$$

$$\text{exista } a, b, c \in \mathbb{R}, \quad \text{cu } a, b \neq 0, \quad \text{astfel încât } aX + bY = c, \quad \mathbb{P} - \text{a.s.},$$

adică $\mathbb{P}(aX + bY = c) = 1$.

Să remarcăm faptul că dependența liniară obținută nu are loc pentru toți ω ci are loc aproape sigur în raport cu măsura de probabilitate $\mathbb{P}$, adică există evenimentul neglijabil N, i.e. $N \in \mathscr{F}$ cu $\mathbb{P}(N) = 0$, astfel încât $aX(\omega) + bY(\omega) = c$, pentru orice $\omega \in \Omega \setminus N$.

Astfel putem spune că ρ este o **măsură a gradului de dependență liniară dintre cele două v.a.**. Dacă $|\rho| = 1$, atunci dependența este liniară. Dacă $\rho \in (-1, 1)$, atunci dependența nu este complet liniară, dar poate fi o dependență (de tip neliniar).

Știm că dacă cele două v.a. sunt independente atunci ele sunt necorelate (i.e. $\rho = 0$), dar dacă $\rho = 0$, atunci aceasta nu înseamnă că cele două v.a. sunt independente, ci doar indică o absență totală a unei dependențe liniare. Prin urmare, două variabile pot fi necorelate, dar pot fi totuși dependente într-un mod neliniar. $\diamond$

Propoziţia II.2.31 (Proprietăți ale dispersiei)

(i) *Dispersia unei constante este nulă. Mai precis*[9]*, pentru* $c \in \mathbb{R}$,

$$X = c, \quad \mathbb{P} - a.s. \quad \Rightarrow \quad \mathrm{D}^2(X) = 0.$$

(ii) *Dispersia unei v.a. este nulă dacă și numai dacă v.a. este constantă* $\mathbb{P}$*–a.s., i.e.*

$$\mathrm{D}^2(X) = 0 \quad \Leftrightarrow \quad X = c, \quad \mathbb{P} - a.s. \quad (\text{unde } c = \mathbb{E}(X)).$$

(iii)

$$\mathrm{D}^2(aX) = a^2 \mathrm{D}^2(X), \quad a \in \mathbb{R}.$$

(iv)

$$\mathrm{D}^2(a + bX) = b^2 \mathrm{D}^2(X), \quad a, b \in \mathbb{R}.$$

(v) *Dacă două v.a. sunt necorelate, atunci dispersia sumei este suma dispersiilor*[10]*, i.e.*

$$X, Y \text{ necorelate}^{11} \quad \Rightarrow \quad \mathrm{D}^2(X + Y) = \mathrm{D}^2(X) + \mathrm{D}^2(Y).$$

Reciproca nu este adevărată.

În cazul a n *v.a.* $(X_i)_{i=\overline{1,n}}$ *necorelate două câte două dispersia sumei este suma dispersiilor, i.e.*

$$(X_i)_{i=\overline{1,n}} \text{ necorelate} \quad \Rightarrow \quad \mathrm{D}^2\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{D}^2(X_i).$$

[9] Pentru semnificația lui $X = c$, $\mathbb{P}$–a.s., vezi Nota 3.

[10] În cazul a două v.a. oarecare (nu neapărat cu un număr finit de valori), trebuie impusă condiția ca cele două v.a. să aibă dispersie finită.

[11] Evident, vezi (II.2.10), dacă v.a. sunt independente, atunci obținem aceeași concluzie.

(vi) *Are loc*[12] *următoarea formulă de calcul a dispersiei sumei a două v.a. în cazul general (în care v.a. nu sunt neapărat necorelate):*

$$\boxed{\mathrm{D}^2(X+Y) = \mathrm{D}^2(X) + \mathrm{D}^2(Y) + 2\,\mathrm{Cov}(X,Y)\,.} \tag{II.2.11}$$

Demonstraţie. (i) Avem $\mathrm{D}^2(c) = \mathbb{E}(c^2) - (\mathbb{E}(c))^2 = c^2 - c^2 = 0$.

(ii) În cazul unei v.a. X discrete să arătăm mai întâi că $\mathbb{E}(X^2) = 0$ implică $X = 0$, $\mathbb{P}-$a.s..

Într-adevăr, deoarece avem că o sumă de numere nenegative este nulă, i.e.

$$0 = \mathbb{E}(X^2) = \sum\nolimits_{i=1} x_i^2\,\mathbb{P}(X = x_i)\,,$$

deducem că pentru orice $x_i \neq 0$, $\mathbb{P}(X = x_i) = 0$.

Prin urmare, singura probabilitate diferită de zero este $\mathbb{P}(X = 0) = 1$.

Apoi, dacă $\mathrm{D}^2(X) = \mathbb{E}(X - \mu_1)^2 = 0$, deducem, folosind pasul precedent, că $X = \mu_1$, $\mathbb{P}-$a.s..

(iii) Avem

$$\mathrm{D}^2(aX) = \sum\nolimits_{i=1}^{m} p_i\,(ax_i - a\mu_1)^2 = a^2 \sum\nolimits_{i=1}^{m} p_i\,(x_i - \mu_1)^2 = a^2\mathrm{D}^2(X)\,.$$

Evident, formula este adevărată și în cazul unei v.a. X oarecare (nu neapărat discretă):

$$\mathrm{D}^2(aX) = \mathbb{E}(aX)^2 - (\mathbb{E}(aX))^2 = a^2\big(\mathbb{E}(X)^2 - (\mathbb{E}X)^2\big) = a^2\mathrm{D}^2(X)\,.$$

(iv) Avem

$$\begin{aligned}\mathrm{D}^2(a+bX) &= \mathbb{E}\big((a+bX)^2\big) - (\mathbb{E}(a+bX))^2\\ &= a^2 + 2ab\mathbb{E}(X) + b^2\mathbb{E}(X^2) - a^2 - 2ab\mathbb{E}(X) - b^2(\mathbb{E}(X))^2 = b^2\mathrm{D}^2(X)\,.\end{aligned}$$

(v) Avem

$$\begin{aligned}\mathrm{D}^2(X+Y) &= \mathbb{E}\big((X+Y)^2\big) - (\mathbb{E}(X+Y))^2\\ &= \mathbb{E}(X^2) + 2\mathbb{E}(XY) + \mathbb{E}(Y^2) - (\mathbb{E}X)^2 - 2\mathbb{E}(X)\,\mathbb{E}(Y) - (\mathbb{E}Y)^2\\ &= \mathbb{E}(X^2) + \mathbb{E}(Y^2) - (\mathbb{E}X)^2 - (\mathbb{E}Y)^2 = \mathrm{D}^2(X) + \mathrm{D}^2(Y)\,.\end{aligned}$$

(vi) Avem

$$\begin{aligned}\mathrm{D}^2(X+Y) &= \mathbb{E}\big((X+Y)^2\big) - (\mathbb{E}(X+Y))^2\\ &= \mathbb{E}(X^2) + 2\mathbb{E}(XY) + \mathbb{E}(Y^2) - (\mathbb{E}X)^2 - 2\mathbb{E}(X)\,\mathbb{E}(Y) - (\mathbb{E}Y)^2\\ &= \mathrm{D}^2(X) + \mathrm{D}^2(Y) + 2\,\mathrm{Cov}(X,Y)\,.\end{aligned}$$

■

[12] În cazul a două v.a. oarecare (nu neapărat cu un număr finit de valori), trebuie impusă condiția ca cele două v.a. să aibă dispersie finită.

Remarca II.2.32 Folosind inegalitatea lui Liapunov și definiția dispersiei se poate arăta că momentul $\mathbb{E}\left(X^2\right)$ de ordin 2 al unei v.a. X există și este finit dacă și numai dacă media $\mathbb{E}(X)$ și dispersia $\mathrm{D}^2(X)$ există și sunt finite:

$$\text{există } \mathbb{E}\left(X^2\right) < \infty \quad \Leftrightarrow \quad \text{există } \mathbb{E}(X) < \infty \quad \text{și} \quad \mathrm{D}^2(X) < \infty.$$

$\diamond$

Definiţia II.2.33 *De obicei, gradul de împrăștiere a valorilor unei v.a. X se exprimă nu prin dispersie ci prin* **deviația standard** *(sau* **abaterea standard***) notată $\mathrm{D}(X)$ și definită de*

$$\mathrm{D}(X) \stackrel{\text{def}}{=\!=} \sqrt{\mathrm{D}^2(X)}.$$

Această noțiune are avantajul că se exprimă prin aceleași unități de măsură ca și valorile v.a. X.

Remarca II.2.34 Momentul central de ordin 3 este[13] de asemenea important. El este folosit pentru a măsura lipsa de simetrie a distribuției. Pe de altă parte $\nu_3 = \mathbb{E}\left[(X-\mu)^3\right]$ depinde de unitatea de măsură utilizată pentru valorile lui X. Pentru a deveni independentă în raport cu unitatea de măsură vom considera momentul de ordin 3 al v.a. standardizate asociată v.a. X. Astfel, dacă $\mu = \mathbb{E}(X)$ iar $\sigma = \sqrt{\mathrm{D}^2(X)}$, definim **coeficientul de asimetrie** (sau ***skewness***)

$$S_3 = \mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right].$$

Distribuția normală standard și orice altă distribuție simetrică (cu momentul de ordin 3 finit) are coeficientul $S_3 = 0$. Astfel, distribuția simetrică, respectiv $S_3 = 0$, devine un element de comparație pentru alte distribuții.

Momentul central de ordin 4 este utilizat pentru a descrie gradul de apropiere față de densitatea distribuției normale standard. Dacă $\mu = \mathbb{E}(X)$ iar $\sigma = \sqrt{\mathrm{D}^2(X)}$, definim **coeficientul de exces** (sau ***kurtosis***)

$$E_4 = \mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^4\right].$$

Distribuția normală standard are coeficientul $E_4 = 3$. Astfel, distribuția normală standard, respectiv $E_4 = 3$, devine un element de comparație pentru alte distribuții. $\diamond$

Prezentăm în continuare o serie de inegalități importante care folosesc media.

Teorema II.2.35 (Inegalitatea lui Markov) *Fie $g : \mathbb{R} \to [0, \infty)$ o funcție crescătoare și măsurabilă și X o v.a..*

Atunci[14] *are loc, pentru orice $\epsilon > 0$,*

$$\mathbb{P}(X \geq \epsilon) \leq \frac{\mathbb{E}(g(X))}{g(\epsilon)}. \tag{II.2.12}$$

[13] Aceste definiții **nu sunt pentru Examen !**

[14] Inegalitatea lui Markov **nu este pentru Examen !**

Demonstraţie. Având în vedere că
$$\mathbb{E}\left(\mathbb{1}_A\right) = \mathbb{P}\left(A\right),$$
obținem $\mathbb{E}\left(\mathbb{1}_{\{X \geq \epsilon\}}\right) = \mathbb{P}\left(X \geq \epsilon\right)$, deci
$$\mathbb{E}\left(g\left(X\right)\right) \geq \mathbb{E}\left(g\left(X\right)\, \mathbb{1}_{[\epsilon,\infty)}\left(X\right)\right) \geq g\left(\epsilon\right)\, \mathbb{E}\left(\mathbb{1}_{[\epsilon,\infty)}\left(X\right)\right) = g\left(\epsilon\right)\, \mathbb{P}\left(X \geq \epsilon\right),$$
pentru orice $\epsilon > 0$. ■

Corolarul II.2.36

(d) *Dacă X este v.a. și $g\left(x\right) = x^2\, \mathbb{1}_{[0,\infty)}\left(x\right)$, obținem* **inegalitatea lui Cebâșev**
$$\mathbb{P}\left(|X - \mathbb{E}\left(X\right)| \geq \epsilon\right) \leq \frac{\mathrm{D}^2\left(X\right)}{\epsilon^2}, \quad \text{pentru orice } \epsilon > 0. \tag{II.2.13}$$
sau, echivalent,
$$\mathbb{P}\left(|X - \mathbb{E}\left(X\right)| < \epsilon\right) \geq 1 - \frac{\mathrm{D}^2\left(X\right)}{\epsilon^2}, \quad \text{pentru orice } \epsilon > 0. \tag{II.2.14}$$

(e) *În cazul particular $\mathbb{E}\left(X\right) = m$, $\mathrm{D}^2\left(X\right) = \sigma^2 < \infty$ și $\epsilon = k\sigma$, inegalitatea (II.2.13) a lui Cebâșev devine*
$$\mathbb{P}\left(|X - m| \geq k\sigma\right) \leq \frac{1}{k^2}. \tag{II.2.15}$$

Remarca II.2.37 Inegalitățile lui Markov și Cebâșev (precum și celelalte cazuri particulare) au loc pentru orice tip de v.a. ◇

Exemplul II.2.38 Fie v.a.
$$X : \begin{pmatrix} 0.2 & 0.3 & 0.4 & 0.5 \\ 0.1 & 0.2 & 0.3 & 0.4 \end{pmatrix}.$$
Să se obțină o estimare a probabilității $\mathbb{P}\left(|X - \mu_1| < 0.2\right)$.

Media și dispersia sunt date de
$$\mathbb{E}\left(X\right) = 0.4, \quad \mathbb{E}\left(X^2\right) = 0.17, \quad \mathrm{D}^2\left(X\right) = 0.17 - 0.16 = 0.01.$$
Avem deci, aplicând inegalitatea (II.2.14) a lui Cebâșev,
$$\mathbb{P}\left(|X - 0.4| < 0.2\right) \geq 0.75.$$
◇

II.2.3 Exemple de v.a. discrete

II.2.3.1 Distribuția Bernoulli

Să considerăm distribuția Poisson în cazul particular $U_i = U$ și deci $p_i = p$, $q_i = q$ cu $i = 1$.

Mai precis, spunem că v.a. X spunem că urmează o **distribuție de tip Bernoulli**, și scriem $X \sim \mathrm{Bernoulli}(p)$, dacă tabloul de distribuție este dat de

$$X : \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}, \quad \text{unde } q = 1 - p.$$

Remarca II.2.39 Valoarea p se numește parametrul distribuției. $\diamond$

Remarca II.2.40 Semnificația distribuției este următoarea: fie o experiență și A un eveniment legat de ea care se realizează cu probabilitatea $p = \mathbb{P}(A)$. Atunci X **reprezintă v.a. care ia drept valori numărul de realizări ale evenimentului A dacă se efectuează o singură dată experiența (adică numărul de apariții ale *Succesului* (apariția evenimentului A) la o singură efectuare a experienței).** $\diamond$

Remarca II.2.41 Să observăm că v.a. indicatoare $\mathbb{1}_A$ a evenimentului $A \in \mathscr{F}$ este chiar v.a. de tip $\mathrm{Bernoulli}(p)$, adică

$$\mathbb{1}_A : \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}, \quad \text{unde } p \stackrel{\text{def}}{=} \mathbb{P}(A).$$

Într-adevăr, dacă v.a. este $\mathbb{1}_A(\omega) \stackrel{\text{def}}{=} \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A \end{cases}$, atunci

$$p = \mathbb{P}(\mathbb{1}_A = 1) = \mathbb{P}(\{\omega : \mathbb{1}_A(\omega) = 1\}) = \mathbb{P}(A).$$

Se obține imediat că

$$\boxed{\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)} \tag{II.2.16}$$

și apoi $\mathbb{E}(\mathbb{1}_A^2) = \mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$, deci

$$\mathrm{D}^2(\mathbb{1}_A) = \mathbb{P}(A)(1 - \mathbb{P}(A)),$$

adică: $\diamond$

Propoziţia II.2.42 *Dacă $X \sim \mathrm{Bernoulli}(p)$, atunci media și dispersia sunt date de:*

$$\boxed{\mathbb{E}(X) = p \qquad \textit{și} \qquad \mathrm{D}^2(X) = pq.} \tag{II.2.17}$$

II.2.3.2 Distribuția binomială

Să considerăm distribuția Poisson în cazul particular $U_i = U$ și deci $p_i = p$, $q_i = q$.

Deci v.a. X spunem că urmează o distribuție binomială, și scriem $X \sim \mathscr{B}(n, p)$, dacă tabloul de distribuție este dat de

$$X : \begin{pmatrix} 0 & 1 & 2 & \dots & k & \dots & n \\ C_n^0 p^0 q^n & C_n^1 p^1 q^{n-1} & C_n^2 p^2 q^{n-2} & \dots & C_n^k p^k q^{n-k} & \dots & C_n^n p^n q^0 \end{pmatrix}$$

sau, scris pe scurt,

$$X : \begin{pmatrix} k \\ C_n^k p^k q^{n-k} \end{pmatrix}_{k=\overline{0,n}}, \quad \text{unde } q \stackrel{\text{def}}{=} 1 - p.$$

Evident, avem $\mathbb{P}(X = k) = C_n^k p^k q^{n-k} \geq 0$.

Pentru a arăta că suma probabilităților este 1 folosim binomul lui Newton[15]:

$$\sum_{k=0}^{n} p_k = \sum_{k=0}^{n} \mathbb{P}(X = k) = \sum_{k=0}^{n} C_n^k p^k q^{n-k} = (p + q)^n = 1$$

Remarca II.2.43 Valorile n și p se numesc parametrii distribuției. $\diamond$

Remarca II.2.44 Semnificația distribuției este următoarea: fie o experiență și A un eveniment legat de ea care se poate realiza cu probabilitatea $p = \mathbb{P}(A)$. Se repetă experiența de n ori în aceleași condiții. Atunci X **reprezintă v.a. care ia drept valori numărul de realizări ale evenimentului A la n efectuări ale experienței, adică numărul de apariții ale *Succesului* (apariția evenimentului A) la n efectuări ale experienței.** $\diamond$

Remarca II.2.45 În cazul particular $n = 1$ obținem distribuția Bernoulli, adică

$$\mathscr{B}(1, p) = \text{Bernoulli}(p),$$

deci $X \sim \mathscr{B}(1, p)$ înseamnă

$$X : \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}.$$

$\diamond$

Propoziția II.2.46 *Dacă $X \sim \mathscr{B}(n, p)$, atunci media și dispersia sunt date de:*

$$\mathbb{E}(X) = n \cdot p \qquad \text{și} \qquad \mathrm{D}^2(X) = n \cdot pq. \tag{II.2.18}$$

[15] **Binomul lui Newton**:

$$(a + b)^r = \sum_{i=0}^{r} C_r^i a^i b^{r-i} = b^r + C_r^1 a b^{r-1} + C_r^2 a^2 b^{r-2} + \dots + C_r^{r-1} a^{r-1} b + a^r,$$

pentru orice $a, b \in \mathbb{R}$ și orice $r \in \mathbb{N}^*$.

Demonstraţie. Într-adevăr, mai întâi avem că

$$\begin{aligned}\mathbb{E}(X) &= \sum_{k=0}^{n} k \cdot C_n^k p^k q^{n-k} = \sum_{k=1}^{n} \frac{n!}{(k-1)!\,(n-k)!}\, p^k q^{n-k} \\ &= np \sum_{k'=0}^{n-1} \frac{(n-1)!}{(k')!\,((n-1)-k')!}\, p^{k'} q^{(n-1)-k'} = np\,(p+q)^{n-1} = np\end{aligned}$$

și

$$\begin{aligned}\mathbb{E}\left(X^2\right) &= \sum_{k=0}^{n} k^2 \cdot C_n^k p^k q^{n-k} = \sum_{k=1}^{n} k\, \frac{n!}{(k-1)!\,(n-k)!}\, p^k q^{n-k} \\ &= \sum_{k=1}^{n} (k-1)\, \frac{n!}{(k-1)!\,(n-k)!}\, p^k q^{n-k} + \sum_{k=1}^{n} \frac{n!}{(k-1)!\,(n-k)!}\, p^k q^{n-k} \\ &= n\,(n-1)\,p^2 \sum_{k=2}^{n} \frac{(n-2)!}{(k-2)!\,(n-k)!}\, p^{k-2} q^{n-k} + np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!\,(n-k)!}\, p^{k-1} q^{n-k} \\ &= n\,(n-1)\,p^2 \sum_{k'=0}^{n-2} \frac{(n-2)!}{(k')!\,((n-2)-k')!}\, p^{k'} q^{(n-2)-k'} \\ &\quad + np \sum_{k'=0}^{n-1} \frac{(n-1)!}{(k')!\,((n-1)-k')!}\, p^{k'} q^{(n-1)-k'} \\ &= n\,(n-1)\,p^2\,(p+q)^{n-2} + np\,(p+q)^{n-1} = n\,(n-1)\,p^2 + np.\end{aligned}$$

Deci

$$\mathrm{D}^2(X) = \mathbb{E}\left(X^2\right) - (\mathbb{E}(X))^2 = n^2p^2 - np^2 + np - n^2p^2 = np - np^2 = np\,(1-p)\,.$$

■

Propoziţia II.2.47 *Suma a două v.a. independente care urmează o distribuție de tip binomial este o v.a care urmează o distribuție tot de tip binomial. Mai precis*[16],

$$X_i \sim \mathscr{B}(n_i, p)\,,\ i = \overline{1,2}\,, \quad \textit{independente} \quad \Rightarrow \quad X_1 + X_2 \sim \mathscr{B}(n_1+n_2, p)\,.$$

Demonstraţie. Valorile pe care le poate lua $X_1 + X_2$ sunt $k \in \{0, 1, \ldots, n_1 + n_2\}$. Atunci

$$\mathbb{P}(X_1 + X_2 = k) = \sum_{i\in\mathbb{N}} \mathbb{P}(X_1 = i) \cdot \mathbb{P}(X_2 = k - i)\,,$$

unde indicele $i \in \mathbb{N}$ satisface resticțiile: $0 \le i \le n_1$ și $0 \le k - i \le n_2$ sau, echivalent, $k - n_2 \le i \le k$.

Avem, de exemplu, în cazul $k \le n_1 \wedge n_2$,

$$\begin{aligned}\mathbb{P}(X_1 + X_2 = k) &= \sum_{i=0}^{k} \mathbb{P}(X_1 = i) \cdot \mathbb{P}(X_2 = k - i) \\ &= \sum_{i=0}^{k} C_{n_1}^i p^i q^{n_1-i} \cdot C_{n_2}^{k-i} p^{k-i} q^{n_2-k+i} \\ &= p^k q^{n_1+n_2-k} \cdot \sum_{i=0}^{k} C_{n_1}^i \cdot C_{n_2}^{k-i} = p^k q^{n_1+n_2-k} C_{n_1+n_2}^k\,.\end{aligned}$$

Celelalte cazuri se studiază similar.

Deci v.a. $X_1 + X_2 \sim \mathscr{B}(n_1 + n_2, p)$. ■

[16] Acest rezultat se poate demonstra și cu ajutorul funcției caracteristice.

Remarca II.2.48 Ca o consecință a rezultatului anterior se obține următoarea proprietate.

Dacă avem n v.a. independente, distribuite Bernoulli de același parametru p, adică n v.a. independente $X_i \sim \text{Bernoulli}(p) = \mathscr{B}(1,p)$, atunci suma lor este distribuită binomial, mai precis,

$$\boxed{X = X_1 + X_2 + \ldots + X_n = \sum_{i=1}^{n} X_i \sim \mathscr{B}(n,p)\,.}$$

Aceasta este evident, având în vedere că dacă X_i reprezintă numărul de succese obținute la încercarea $i = \overline{1,n}$ (adică X_i ia doar valorile 0 sau 1), atunci numărul natural aleator $X = X_1 + X_2 + \ldots + X_n$ reprezintă numărul total de succese obținute în n încercări, adică $X \sim \mathscr{B}(n,p)\,.$

Deci

orice v.a. $X \sim \mathscr{B}(n,p)$ poate fi văzută ca o sumă de n v.a. **independente**, de tip $\text{Bernoulli}(p)$.

În plus, știm că $\mathbb{E}(X_i) = p$ iar $\mathrm{D}^2(X_i) = pq$, deci **este foarte ușor să obținem și să reținem media și dispersia unei v.a. distribuite binomial**, adică formulele (II.2.18):

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{E}(X_i) = \sum_{i=1}^{n} p = n \cdot p$$

și

$$\mathrm{D}^2(X) = \mathrm{D}^2\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{D}^2(X_i) = \sum_{i=1}^{n} pq = n \cdot pq.$$

◇

II.2.3.3 Distribuția hipergeometrică

V.a. X spunem că urmează o distribuție de tip hipergeometric dacă tabloul ei de distribuție este dat de

$$X : \begin{pmatrix} k \\ p_k \end{pmatrix}_{k=\overline{0,n}}, \tag{II.2.19}$$

unde

$$\mathbb{P}(X = k) = p_k = \frac{C_a^k C_b^{n-k}}{C_{a+b}^n}\,, \quad \text{pentru } k = \overline{0,n}\,.$$

Evident, avem $\mathbb{P}(X = k) = \frac{C_a^k C_b^{n-k}}{C_{a+b}^n} \geq 0.$

Pentru a arăta că suma probabilităților este 1 folosim identitatea demonstrată în cadrul Seminarului și obținem că $\sum_{k=0}^{n} p_k = 1.$

Remarca II.2.49 Semnificația distribuției este următoarea: fie o urnă cu a bile albe și b bile negre din care **se extrag pe rând, dar fără a pune bila la loc** (sau **se extrag simultan**), n bile, cu $n \leq a + b$. Atunci X **reprezintă v.a. care ia drept valori numărul de apariții ale *Succesului* (apariția unei bile albe) în cele** n **extrageri.** ◇

Remarca II.2.50 La fel ca și în cazul unei v.a. distribuite binomial (vezi Remarca II.2.48), să definim v.a. X_i ca fiind numărul de bile albe obținute la extragerea $i = \overline{1,n}$ (adică X_i ia doar valorile 0 sau 1, deci X_i sunt v.a. care urmează o lege de tip Bernoulli) în cadrul de lucru dat de Remarca II.2.49: din urna cu a bile albe și b bile negre se extrag pe rând $n \leq a+b$ bile, dar fără a pune bila la loc (sau se extrag simultan n bile). Prin urmare, X_i sunt v.a. care nu sunt independente între ele.

Pe de altă parte, numărul natural aleator $X_1 + X_2 + \ldots + X_n$ reprezintă numărul total de bile albe obținute după cele n extrageri, prin urmare

$$X = X_1 + X_2 + \ldots + X_n$$

reprezintă o v.a. distribuită hipergeometric.

Deci

orice v.a. X distribuită hipergeometric poate fi văzută ca

o sumă de n v.a. **dependente**, de tip Bernoulli.

Pe de altă parte, evident, $X_1 : \begin{pmatrix} 0 & 1 \\ \frac{b}{a+b} & \frac{a}{a+b} \end{pmatrix}$, deci $\mathbb{E}(X_1) = \dfrac{a}{a+b}$.

Apoi obținem și tabloul v.a. $X_2 : \begin{pmatrix} 0 & 1 \\ \frac{b}{a+b} & \frac{a}{a+b} \end{pmatrix}$, deoarece avem o structură de tip arbore și, folosind formula probabilității totale, putem calcula:

$$\begin{aligned}
\mathbb{P}(X_2 = 0) &= \mathbb{P}(X_2 = 0 | X_1 = 1) \cdot \mathbb{P}(X_1 = 1) + \mathbb{P}(X_2 = 0 | X_1 = 0) \cdot \mathbb{P}(X_1 = 0) \\
&= \frac{b}{a+b-1} \cdot \frac{a}{a+b} + \frac{b-1}{a+b-1} \cdot \frac{b}{a+b} = \frac{b}{a+b}, \\
\mathbb{P}(X_2 = 1) &= \mathbb{P}(X_2 = 1 | X_1 = 1) \cdot \mathbb{P}(X_1 = 1) + \mathbb{P}(X_2 = 1 | X_1 = 0) \cdot \mathbb{P}(X_1 = 0) \\
&= \frac{a-1}{a+b-1} \cdot \frac{a}{a+b} + \frac{a}{a+b-1} \cdot \frac{b}{a+b} = \frac{a}{a+b}.
\end{aligned}$$

Tabloul v.a. $X_3 : \begin{pmatrix} 0 & 1 \\ \frac{b}{a+b} & \frac{a}{a+b} \end{pmatrix}$, deoarece avem o structură de tip arbore și, folosind formula probabilității totale, putem calcula:

$$\begin{aligned}
\mathbb{P}(X_3=0) &= \mathbb{P}(X_3=0|X_2=1,X_1=1)\cdot\mathbb{P}(X_2=1|X_1=1)\cdot\mathbb{P}(X_1=1)\\
&\quad+\mathbb{P}(X_3=0|X_2=0,X_1=1)\cdot\mathbb{P}(X_2=0|X_1=1)\cdot\mathbb{P}(X_1=1)\\
&\quad+\mathbb{P}(X_3=0|X_2=1,X_1=0)\cdot\mathbb{P}(X_2=1|X_1=0)\cdot\mathbb{P}(X_1=0)\\
&\quad+\mathbb{P}(X_3=0|X_2=0,X_1=0)\cdot\mathbb{P}(X_2=0|X_1=0)\cdot\mathbb{P}(X_1=0)\\
&=\frac{b}{a+b-2}\cdot\frac{a-1}{a+b-1}\cdot\frac{a}{a+b}+\frac{b-1}{a+b-2}\cdot\frac{b}{a+b-1}\cdot\frac{a}{a+b}\\
&\quad+\frac{b-1}{a+b-2}\cdot\frac{a}{a+b-1}\cdot\frac{b}{a+b}+\frac{b-2}{a+b-2}\cdot\frac{b-1}{a+b-1}\cdot\frac{b}{a+b}\\
&=\frac{ab(a+b-2)}{(a+b)(a+b-1)(a+b-2)}+\frac{b(b-1)(a+b-2)}{(a+b)(a+b-1)(a+b-2)}\\
&=\frac{b(a+b-1)}{(a+b)(a+b-1)}=\frac{b}{a+b}\,,
\end{aligned}$$

prin urmare $\mathbb{P}(X_3=1)=\dfrac{a}{a+b}$.

Similar,

$$X_i:\begin{pmatrix}0 & 1\\ \frac{b}{a+b} & \frac{a}{a+b}\end{pmatrix},\quad \text{deci } \mathbb{E}(X_i)=\frac{a}{a+b}\,,\quad \text{pentru } i=\overline{1,n}\,.$$

Deci **este foarte ușor să obținem și să reținem media unei v.a. distribuite hipergeometric**:

$$\mathbb{E}(X)=\mathbb{E}\left(\sum\nolimits_{i=1}^{n}X_i\right)=\sum\nolimits_{i=1}^{n}\mathbb{E}(X_i)=\sum\nolimits_{i=1}^{n}\frac{a}{a+b}=n\cdot\frac{a}{a+b}\,.$$

În ceea ce privește dispersia, $(X_i)_{i=\overline{0,n}}$ nefiind independente, este mai greu de calculat $\mathrm{D}^2(X)$ folosind doar $\mathrm{D}^2(X_i)$. ◇

Remarca II.2.51 Dintr-o populație statistică de mărime N s-a făcut o selecție de n indivizi, nerepetată (individul extras nu se întoarce în populație) în vederea cercetării unei caracteristici A (sau se poate considera că la n indivizi se face simultan cercetarea caracteristicii A). Dacă presupunem că o parte dintre indivizii populației posedă caracteristica A iar restul nu o posedă, atunci numărul de indivizi cercetați posedând proprietatea A urmează schema hipergeometrică (sau schema bilei nerevenite). ◇

II.3 Variabile aleatoare discrete cu un număr infinit de valori

Ca și în cazul v.a. cu un număr finit de valori (vezi pagina 8), cunoașterea legii dată de $\mathbb{P}_X(B)=\mathbb{P}(X\in B)$, unde $B\in\mathscr{B}(\mathbb{R})$, asociată unei v.a. discrete X, cu un număr infinit de valori, este echivalentă cu cunoașterea valorile pe care le poate lua v.a. (adică a mulțimii $X(\Omega)$) și a probabilităților cu care este luată fiecare valoare din $X(\Omega)$.

Putem reprezenta aceste informații sub forma unui **tabloul de repartiție**:

$$X:\begin{pmatrix} x_1 & x_2 & \dots & x_i & \dots \\ p_1 & p_2 & \dots & p_i & \dots \end{pmatrix} \quad \text{sau} \quad X:\begin{pmatrix} x_i \\ p_i \end{pmatrix}_{i\in I}, \tag{II.3.1}$$

unde I este o mulțime cel mult numărabilă (prin convenție, am scris valorile $x_1, x_2, \dots, x_n$ ale v.a. în ordine strict crescătoare).

Deci $X(\Omega) = \{x_1, x_2, \dots, x_n, \dots\}$ iar $p_i = \mathbb{P}(X = x_i)$, cu $i \in I$, astfel încât

$$p_i \geq 0, \quad \text{pentru } i \in I, \quad \text{și} \quad \sum\nolimits_{i\in I} p_i = 1$$

(seria $\sum_{i\in I} p_i$ este convergentă și are suma 1).

Pe scurt, putem scrie că o v.a. discretă X este descrisă complet prin cunoașterea familiei cel mult numărabile de valori

$$(p_i)_{i\in I} \overset{\text{not}}{=\!=} (\mathbb{P}(X = x_i))_{i\in I}.$$

Are loc, și în acest caz, scrierea unei v.a. discrete dată de Remarca **??**:

$$X = \sum\nolimits_{i\in I} x_i \, \mathbb{1}_{A_i},$$

unde $A_i = X^{-1}(\{x_i\}) = \{X = x_i\}$, pentru $i \in I$.

Remarca II.3.1 Să menționăm că o funcție discretă $X : \Omega \to \mathbb{R}$, definită pe un spațiu de probabilitate $(\Omega, \mathscr{F}, \mathbb{P})$ cu $X(\Omega)$ cel mult numarabila, este v.a. (vezi Definiția II.1.1 sau Definiția II.1.2) dacă și numai dacă $\{X = x\} \in \mathscr{F}$, pentru orice $x \in X(\Omega)$. $\diamond$

Funcția de repartiție este dată de Definiția II.1.8 și este o funcția în scară:

$$F(x) = \mathbb{P}(X \leq x) = \sum\nolimits_{\substack{x_i \in X(\Omega) \\ x_i \leq x}} \mathbb{P}(X = x_i), \quad \text{pentru orice } x \in \mathbb{R}.$$

Operațiile cu variabile cu un număr infinit de valori se definesc ca și în cazul v.a. discrete cu un număr finit de valori (vezi Secțiunea II.2.1).

Definiția II.2.7 se poate da și în cazul unei v.a. discrete cu un număr infinit de valori. Pentru simplitatea scrierii vom lua $I = \mathbb{N}^*$ (vezi Nota 1).

Definiţia II.3.2 *Fie v.a. discretă* $X:\begin{pmatrix} x_i \\ p_i \end{pmatrix}_{i\in\mathbb{N}^*}$. *Spunem că v.a.* X **admite medie** *(sau că* **este integrabilă***), dacă următoarea serie este convergentă:*

$$\sum\nolimits_{i\in\mathbb{N}^*} |x_i| \, p_i < \infty.$$

În caz că v.a. X admite medie, atunci suma seriei $\sum_{i\in\mathbb{N}^*} x_i\, p_i$ se va numi **media** *v.a. discrete X, i.e.*

$$\mathbb{E}\left(X\right) \stackrel{\text{def}}{=} x_1p_1 + x_2p_2 + \ldots + x_ip_i + \ldots = \sum\nolimits_{i\in\mathbb{N}^*} x_i\, p_i\,. \tag{II.3.2}$$

Remarca II.3.3 Dacă v.a. X admite medie, atunci putem scrie definiția precedentă și sub forma

$$\mathbb{E}\left(X\right) = \sum\nolimits_{x\in X(\Omega)} x\cdot\mathbb{P}\left(X=x\right).$$

◇

Remarca II.3.4 Media unei v.a. discrete cu un număr finit de valori există întotdeauna, fiind o sumă finită, pe când media unei v.a. discrete cu un număr infinit de valori nu; astfel, seria modulelor care apare în definiția mediei este o serie cu termeni nenegativi și poate să fie convergentă (are suma finită) sau nu (are suma infinită).

Să observăm că dacă se impune ca seria care definește media să fie absolut convergentă, atunci seria fără module $\sum_{i\in\mathbb{N}^*} x_i\, p_i$ este convergentă, prin urmare media este bine definită și este suma acelei serii. ◇

Remarca II.3.5 Se impune ca seria $\sum_{i\in\mathbb{N}^*} x_i\, p_i$ să fie absolut convergentă deoarece se dorește ca **prin orice permutare a termenilor seriei să se obțină o serie tot convergentă și cu aceeași sumă**[17].

Astfel, se impune absoluta convergentă deoarece, atunci când se calculează media v.a., să nu conteze ordinea de scriere a valorilor v.a. și respectiv a termenilor sumei din definiția mediei. ◇

Media unei funcții aplicată unei v.a. poate fi calculată folosind următorul rezultat.

Teorema II.3.6 (Formula de transfer) *Fie X o .v.a. discretă cu tabloul de repartiție*

$$X : \begin{pmatrix} x_i \\ p_i \end{pmatrix}_{i\in\mathbb{N}^*}.$$

Fie $h:\mathbb{R}\to\mathbb{R}$ o funcție măsurabilă. Atunci $h\left(X\right):\Omega\to\mathbb{R}$ este și ea o v.a. și această v.a. admite medie dacă și numai dacă

$$\sum\nolimits_{i\in\mathbb{N}^*} \left|h\left(x_i\right)\right|\, p_i < \infty.$$

În cazul în care $h\left(X\right)$ admite medie, aceasta este dată de:

$$\mathbb{E}\left(h\left(X\right)\right) = \sum\nolimits_{i\in\mathbb{N}^*} h\left(x_i\right)\; p_i\,.$$

[17] Dacă seria $\sum_{i\in\mathbb{N}^*} a_i$ este absolut convergentă, atunci orice altă serie obținută printr-o permutare a termenilor este de asemenea convergentă și are aceeași sumă cu seria inițială.

Dacă seria $\sum_{i\in\mathbb{N}^*} a_i$ nu ar fi absolut convergentă (ci doar convergentă), atunci, conform Teormei lui Riemann, pentru orice $A\in\bar{\mathbb{R}}$, există o permutare a termenilor astfel încât noua serie să aibă suma A.

Remarca II.3.7 Dacă v.a. $h(X)$ admite medie, atunci formula precedentă o putem scrie și sub forma

$$\mathbb{E}\left(h(X)\right) = \sum_{x\in X(\Omega)} h(x)\cdot \mathbb{P}\left(X=x\right).$$

$\diamond$

Remarca II.3.8 Dispersia, covarianța, corelația, precum și momentele inițiale, centrale sau absolute de ordin $r>0$ se definesc cu ajutorul seriilor numerice și similar cazului v.a. cu un număr finit de valori. $\diamond$

Remarca II.3.9 Au loc și în cazul v.a. discrete cu un număr infinit de valori rezultatele stabilite de Propozițiile II.2.11 și II.2.31 și inegalitățile lui Markov și Cebâșev date de (II.2.12) și respectiv (II.2.13). $\diamond$

II.3.1 Exemple de v.a. discrete

II.3.1.1 Distribuția Poisson (cu un număr infinit de valori)

Spunem că X este o v.a. este de tip Poisson de parametru $\lambda\in(0,\infty)$, și scriem $X\sim\mathscr{P}(\lambda)$, dacă are tabloul

$$X:\begin{pmatrix} 0 & 1 & 2 & \ldots & k & \ldots \\ e^{-\lambda} & \dfrac{\lambda}{1!}\,e^{-\lambda} & \dfrac{\lambda^2}{2!}\,e^{-\lambda} & \ldots & \dfrac{\lambda^k}{k!}\,e^{-\lambda} & \ldots \end{pmatrix}.$$

Evident, avem $\mathbb{P}\left(X=k\right)=\frac{\lambda^k}{k!}\,e^{-\lambda}>0$, unde $k\in\mathbb{N}$.

Pentru a arăta că suma probabilităților este 1 folosim dezvoltarea în serie de puteri a exponențialei[18]:

$$\sum_{k=0}^{\infty} p_k = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}\,e^{-\lambda} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda}=1.$$

Propoziţia II.3.10 *Media și dispersia v.a. distribuite Poisson X sunt date de*

$$\mathbb{E}\left(X\right)=\lambda \qquad \text{și} \qquad \mathrm{D}^2\left(X\right)=\lambda. \tag{II.3.3}$$

Remarca II.3.11 Deci, în cazul unei v.a. de tip $\mathscr{P}(\lambda)$, parametrul ei este chiar media (precum și dispersia) v.a.. $\diamond$

Demonstraţie. Într-adevăr,

$$\mathbb{E}\left(X\right)=\sum_{k=0}^{\infty} kp_k = \sum_{k=0}^{\infty} k\,\frac{\lambda^k}{k!}\,e^{-\lambda} = \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}=\lambda$$

[18] Dezvoltarea în serie de puteri a exponențialei, i.e. dezvoltarea Taylor în jurul lui $a=0$, adică seria Maclaurin, asociată funcției e^x :

$$e^x=\sum_{k=0}^{\infty}\frac{x^k}{k!}=1+\frac{x}{1!}+\frac{x^2}{2!}+\ldots+\frac{x^n}{n!}+\ldots, \quad \text{pentru orice } x\in\mathbb{R}.$$

și

$$\mathbb{E}\left(X^2\right) = \sum\nolimits_{k=0}^{\infty} k^2 p_k = \sum\nolimits_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum\nolimits_{k=1}^{\infty} k \frac{\lambda^k}{(k-1)!} = e^{-\lambda} \sum\nolimits_{k=1}^{\infty} (k-1+1) \frac{\lambda^k}{(k-1)!}$$
$$= \lambda^2 e^{-\lambda} \sum\nolimits_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda e^{-\lambda} \sum\nolimits_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda^2 e^{-\lambda} e^{\lambda} + \lambda e^{-\lambda} e^{\lambda} = \lambda^2 + \lambda,$$

deci

$$\mathrm{D}^2(X) = \mathbb{E}\left(X^2\right) - \left(\mathbb{E}(X)\right)^2 = \lambda.$$

■

Remarca II.3.12 Distribuția Poisson mai poartă numele de **legea evenimentelor rare**.

Distribuția Poisson este utilizată când ne interesează numărul de evenimente care apar într-un interval de timp precizat, dacă se observă că aceste evenimente apar cu o anumită medie (rată de apariție sau intensitate) și faptul că ele apar independent unele față de altele.

De exemplu, v.a. distribuite Poisson pot fi cele care reprezintă numărul de accidente ce apar într-un an, într-o anumită secțiune a unei străzi; sau numărul persoanelor care mor într-un an din cauza bolii Lyme; sau numărul de cutremure ce apar într-un an, într-o anumită zonă geografică; sau numărul de apeluri zilnice primite de un Serviciu Clienți al unei firme.

Deoarece $\mathbb{E}(X) = \lambda$, parametrul $\lambda > 0$ poate fi văzut ca **rata** sau **intensitatea medie de apariție a evenimentului** pe unitatea de timp. ◇

Propoziţia II.3.13 *Suma a două v.a. independente care urmează o distribuiție de tip Poisson este o v.a care urmează o distribuiție tot de tip Poisson. Mai precis*[19],

$$X_i \sim \mathscr{P}(\lambda_i),\ i = \overline{1,2}, \quad \textit{independente} \quad \Rightarrow \quad X_1 + X_2 \sim \mathscr{P}(\lambda_1 + \lambda_2).$$

Demonstraţie. Valorile pe care le poate lua $X_1 + X_2$ sunt $k \in \mathbb{N}$. Atunci

$$\mathbb{P}(X_1 + X_2 = k) = \sum\nolimits_{i=0}^{k} \mathbb{P}(X_1 = i) \cdot \mathbb{P}(X_2 = k - i) = \sum\nolimits_{i=0}^{k} \frac{\lambda_1^i}{i!} e^{-\lambda_1} \cdot \frac{\lambda_2^{k-i}}{(k-i)!} e^{-\lambda_2}$$
$$= e^{-(\lambda_1+\lambda_2)} \sum\nolimits_{i=0}^{k} \frac{\lambda_1^i}{i!} \cdot \frac{\lambda_2^{k-i}}{(k-i)!} = \frac{1}{k!} e^{-(\lambda_1+\lambda_2)} \sum\nolimits_{i=0}^{k} C_k^i \lambda_2^i \lambda_2^{k-i}$$
$$= \frac{1}{k!} e^{-(\lambda_1+\lambda_2)} (\lambda_1 + \lambda_2)^k.$$

Deci v.a. $X_1 + X_2 \sim \mathscr{P}(\lambda_1 + \lambda_2)$. ■

Folosind Propoziția II.3.13 obținem:

Propoziţia II.3.14 *Dacă $X_i \sim \mathscr{P}(\lambda)$, pentru $i = \overline{1,n}$, sunt v.a. independente, atunci*

$$S_n \overset{\text{not}}{=\!=} X_1 + X_2 + \ldots + X_n \sim \mathscr{P}(n\lambda).$$

[19] Acest rezultat se poate demonstra și cu ajutorul funcției caracteristice.

II.3.1.2 Distribuția geometrică

V.a. X spunem că urmează o distribuție de tip geometric de parametru $p \in (0,1)$, și scriem $X \sim \mathscr{G}(p)$, dacă tabloul ei de distribuție este dat de

$$X : \begin{pmatrix} k \\ pq^{k-1} \end{pmatrix}_{k \in \mathbb{N}^*} .$$

Evident, avem $\mathbb{P}(X = k) = pq^{k-1} \geq 0$.

Pentru a arăta că suma probabilităților este 1 folosim seria geometrică[20] (ceea ce justifică și numele distribuției):

$$\sum_{k \in \mathbb{N}^*} p_k = \sum_{k=1}^{\infty} pq^{k-1} = p \sum_{k=1}^{\infty} q^{k-1} = p \sum_{k=0}^{\infty} q^k = p \, \frac{1}{1-q} = 1.$$

Remarca II.3.15 Semnificația distribuției este următoarea: fie o experiență și A un eveniment legat de ea care se poate realiza cu probabilitatea $p = \mathbb{P}(A)$. Se repetă experiența de mai multe ori și în aceleași condiții. Atunci X reprezintă v.a. care ia drept valori **numărul de încercări până când are loc prima realizare a evenimentului** A, **i.e. primul *Succes* (apariția evenimentului** A **).**

Deci, evenimentul $\{X = k\}$ este evenimentul ca în primele $(k-1)$ efectuări ale experienței, evenimentul A nu se produce deloc (spunem că s-au produs $(k-1)$ eșecuri) și în efectuarea k a experienței evenimentul A se produce (spunem că se produce primul succes).

Obținem $\mathbb{P}(X = k) = \underbrace{q \cdot q \cdot \ldots \cdot q}_{\text{de } (k-1)\text{-ori}} \cdot p = q^{k-1}p.$ ◇

Remarca II.3.16 Putem defini distribuția geometrică și în modul următor: fie o experiență și A un eveniment legat de ea care se realizează cu probabilitatea $p = \mathbb{P}(A)$. Se repetă experiența de mai multe ori și în aceleași condiții. Să definim X ca fiind v.a. care ia drept valori **numărul de eșecuri până când are loc prima realizare a evenimentului** A**, i.e. primul *Succes* (apariția evenimentului** A **)**.

Atunci evenimentul $\{X = k\}$ este evenimentul ca în primele k efectuări ale experienței, evenimentul A nu se produce deloc și în efectuarea $(k+1)$ a experienței evenimentul A se produce.

Deci $\mathbb{P}(X = k) = \underbrace{q \cdot q \cdot \ldots \cdot q}_{\text{de } k\text{-ori}} \cdot p = q^k p$ iar tabloul este dat de

$$X : \begin{pmatrix} k \\ pq^k \end{pmatrix}_{k \in \mathbb{N}} .$$

[20] **Seria geometrică** este suma termenilor unei progresii geometrice și este dată de următoarea dezvoltare în serie de puteri, i.e. dezvoltarea Taylor în jurul lui $a = 0$, adică seria Maclaurin, asociată funcției $\frac{1}{1-x}$:

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + \ldots + x^k + \ldots , \quad \text{pentru orice } |x| < 1.$$

$\diamond$

Propoziţia II.3.17 *Media și dispersia unei v.a.* $X \sim \mathscr{G}(p)$ *sunt date de*

$$\boxed{\mathbb{E}(X) = \frac{1}{p} \quad \textit{și} \quad \mathrm{D}^2(X) = \frac{q}{p^2}\,.} \tag{II.3.4}$$

Demonstraţie. Folosim dezvoltări cunoscute, plecând[21] de la dezvoltarea funcției $\dfrac{1}{1-x}$ în serii de puteri (i.e. de la seria geometrică). Obținem

$$\mathbb{E}(X) = \sum\nolimits_{k=1}^{\infty} k\,p_k = p\sum\nolimits_{k=1}^{\infty} k\,q^{k-1} = p\,\frac{1}{(1-q)^2} = \frac{1}{p}$$

și

$$\mathbb{E}\left(X^2\right) = \sum\nolimits_{k=1}^{\infty} k^2\,p_k = p\sum\nolimits_{k=1}^{\infty} k^2\,q^{k-1} = p\,\frac{1+q}{(1-q)^3} = \frac{2-p}{p^2}\,,$$

deci

$$\mathrm{D}^2(X) = \mathbb{E}\left(X^2\right) - (\mathbb{E}(X))^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}\,.$$

$\blacksquare$

Exerciţiul II.3.18 *Să se calculeze media și dispersia în cazul în care* X *numără eșecurile până la apariția primului* Succes.

Propoziţia II.3.19 *Dacă* $X \sim \mathscr{G}(p)$*, atunci, pentru orice* $n \in \mathbb{N}^*$*, funcția de repartiție este dată de*

$$\boxed{F(n) = 1 - q^n \qquad \textit{sau, echivalent,} \qquad \mathbb{P}(X > n) = q^n,} \tag{II.3.5}$$

pentru orice $n \in \mathbb{N}^*$.

Demonstraţie. Într-adevăr,

$$F(n) = \mathbb{P}(X \le n) = \sum\nolimits_{k=1}^{n} pq^{k-1} = p\sum\nolimits_{k=0}^{n-1} q^k = p\cdot\frac{1-q^n}{1-q} = 1 - q^n\,.$$

$\blacksquare$

---

[21] Derivând o dată dezvoltarea funcției $\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$ obținem dezvoltarea

$$\frac{1}{(1-x)^2} = \sum\nolimits_{k=1}^{\infty} kx^{k-1}, \quad \text{pentru orice } |x| < 1.$$

Se obține imediat și

$$\frac{x}{(1-x)^2} = \sum\nolimits_{k=1}^{\infty} kx^{k}, \quad \text{pentru orice } |x| < 1.$$

Derivând încă o dată obținem dezvoltarea

$$\frac{2}{(1-x)^3} = \sum\nolimits_{k=2}^{\infty} k(k-1)\,x^{k-2}, \quad \text{pentru orice } |x| < 1.$$

precum și

$$\sum\nolimits_{k=1}^{\infty} k^2x^{k-1} = \left(\frac{x}{(1-x)^2}\right)' = \frac{1+x}{(1-x)^3}\,, \quad \text{pentru orice } |x| < 1.$$

Integrând dezvoltarea funcției $\frac{1}{1-x}$ obținem dezvoltarea

$$\ln\left(\frac{1}{1-x}\right) = \sum\nolimits_{k=0}^{\infty} \frac{1}{k+1}\,x^{k+1}, \quad \text{pentru orice } |x| < 1.$$

Exercițiul II.3.20 (a) *Să se determine probabilitatea de a obține o dublă atunci când se aruncă, o singură dată, două zaruri.*

(b) *Se aruncă două zaruri în mod repetat, până când se obține prima dublă. Fie X v.a. care reprezintă numărul de încercări. Să se determine legea v.a. X.*

(c) *Să se determine probabilitatea de a nu avea nici o dublă în primele 5 aruncări (i.e. $\mathbb{P}(X > 5)$).*

II.3.1.3 Distribuția binomială cu exponent negativ

V.a. X spunem că urmează o distribuție de tip binomial cu exponent negativ de parametri $r \in \mathbb{N}^*$, $p \in (0,1)$, și scriem $X \sim \mathscr{NegB}(r,p)$, dacă tabloul ei de distribuție este dat de

$$X : \begin{pmatrix} k \\ C_{k-1}^{r-1}\, p^r\, q^{k-r} \end{pmatrix}_{k\in\mathbb{N},\, k\geq r}.$$

Remarca II.3.21 Semnificația distribuției este următoarea: fie o experiență și A un eveniment legat de ea care se poate realiza cu probabilitatea $p = \mathbb{P}(A)$. Se repetă experiența de mai multe ori și în aceleași condiții. Atunci X reprezintă v.a. care ia drept valori **numărul de încercări până când evenimentul A se realizează de r ori, i.e. se produce *Succesul* (apariția evenimentului A) de r ori.**

Într-adevăr, evenimentul $\{X = k\}$, unde $k \geq r$, este intersecția a două evenimente independente: în primele $(k-1)$ efectuări ale experienței, evenimentul A se produce de $r-1$ ori și la efectuarea k a experienței evenimentul A se produce.

Probabilitatea primului eveniment este, conform schemei binomiale, dat de $C_{k-1}^{r-1}\, p^{r-1}\, q^{k-r}$ iar probabilitatea celui de-al doilea eveniment este p. Deci probabilitatea intersecției este

$$\mathbb{P}(X = k) = C_{k-1}^{r-1}\, p^{r-1} q^{k-r} \cdot p = C_{k-1}^{r-1}\, p^r q^{k-r}.$$

$\diamond$

Remarca II.3.22 Evident, în cazul particular $r = 1$ distribuția binomială cu exponent negativ devine distribuția geometrică, mai precis,

$$\mathscr{NegB}(1,p) = \mathscr{G}(p).$$

$\diamond$

Remarca II.3.23 Dacă vedem $X \sim \mathscr{NegB}(r,p)$ ca fiind **numărul de încercări până la apariția succesului de ordin r (sau a primelor r succese)**, atunci X este o sumă de r v.a. X_i independente date de: X_1 reprezintă numărul de încercări necesare până la apariția primului succes, X_2 reprezintă numărul de încercări necesare de după apariția primului succes, până la apariția celui de-al doilea succes, X_3 reprezintă numărul de încercări necesare de după apariția celui de-al doilea succes, până la apariția celui de-al treilea succes ș.a.m.d.

În definițiile de mai sus X_i reprezintă r v.a. independente distribuite geometric de același parametru p, i.e. $X_i \sim \mathscr{G}(p)$, deci

orice v.a. X distribuită negativ binomial poate fi văzută ca o sumă de r v.a. independente, de tip geometric,

i.e.

$$X = X_1 + X_2 + \ldots + X_r = \sum\nolimits_{i=0}^{r} X_i \sim \mathscr{NegB}(r,p) \ .$$

În plus, știm că $\mathbb{E}(X_i) = \frac{1}{p}$ iar $\mathrm{D}^2(X_i) = \frac{q}{p^2}$, deci **este foarte ușor să obținem și să reținem media și dispersia unei v.a. distribuite binomial cu exponent negativ** $\mathscr{NegB}(r,p)$:

$$\mathbb{E}(X) = \mathbb{E}\left(\sum\nolimits_{i=1}^{r} X_i\right) = \sum\nolimits_{i=1}^{r} \mathbb{E}(X_i) = \sum\nolimits_{i=1}^{r} \frac{1}{p} = r \cdot \frac{1}{p} \tag{II.3.6}$$

și

$$\mathrm{D}^2(X) = \mathrm{D}^2\left(\sum\nolimits_{i=1}^{r} X_i\right) = \sum\nolimits_{i=1}^{r} \mathrm{D}^2(X_i) = \sum\nolimits_{i=1}^{r} \frac{q}{p^2} = r \cdot \frac{q}{p^2} \ . \tag{II.3.7}$$

$\diamond$

Legătura precedentă dintre o v.a. distribuită binomial cu exponent negativ și o sumă de v.a. distribuite geometric poate fi demonstrată și formal.

Propoziţia II.3.24 *Suma a două v.a. independente care urmează o distribuție de tip geometric este o v.a care urmează o distribuție de tip negativ binomial. Mai precis,*

$$X_1, X_2 \sim \mathscr{G}(p) \quad \textit{independente} \quad \Rightarrow \quad X_1 + X_2 \sim \mathscr{NegB}(2,p) \, .$$

Demonstraţie. Valorile pe care le poate lua suma X_1+X_2 sunt $k \in \{2,3,\ldots\}$, deoarece $X_1(\Omega) = X_2(\Omega) = \mathbb{N}^*$. Atunci

$$\begin{aligned} \mathbb{P}(X_1 + X_2 = k) &= \sum\nolimits_{i=1}^{k-1} \mathbb{P}(X_1 = i) \cdot \mathbb{P}(X_2 = k-i) \\ &= \sum\nolimits_{i=1}^{k-1} pq^{i-1} \cdot pq^{k-i-1} = \sum\nolimits_{i=1}^{k-1} p^2 q^{k-2} = (k-1)\, p^2 q^{k-2} \\ &= C_{k-1}^{2-1} p^2 q^{k-2} , \quad \text{pentru orice } k \geq 2, \end{aligned}$$

adică $X_1 + X_2 \sim \mathscr{NegB}(2,p)$. ■

Exerciţiul II.3.25 *Un arcaș trage asupra unei ținte. Probabilitatea de a atinge ținta este de 2/3. Să se determine tabloul v.a. X care are drept valori numărul de încercări înregistrate până când ținta este atinsă de trei ori.*

Să se determine și tabloul v.a. Y care are drept valori numărul de eșecuri până la primele trei nimeriri ale țintei (vezi și Remarca II.3.16).

Pentru rezolvare se poate folosi semnificația distribuției și tipul de calcul dat de Remarca II.3.21 sau tipul de calcul dat de Propoziția II.3.24.