

Seminar 6

Variabile aleatoare discrete. Vectori aleatori discreți

În acest seminar ne interesează să aplicăm următoarele noțiuni legate de variabile aleatoare discrete:

- Determinarea distribuției de probabilitate. Calculul probabilității unor evenimente;
- Determinarea funcției de repartiție a unei variabile aleatoare discrete;
- Funcții de o variabilă aleatoare;
- Medie. Dispersie.

Pentru vectori aleatori discreți ne interesează:

- Distribuția comună de probabilitate a unui vector (X, Y) ;
- Distribuții marginale;
- Variabile aleatoare condiționate;
- Independența variabilelor aleatoare;
- Funcții de două variabile aleatoare.

6.1 Probleme rezolvate

1. Se dă variabila aleatoare ce descrie numărul de puncte acordate într-un joc de noroc:
$$X = \begin{pmatrix} -3 & -1 & 0 & 1 & 3 & 5 \\ \frac{1}{12} & \frac{1}{6} & \frac{1}{12} & \frac{1}{3} & \frac{1}{12} & p \end{pmatrix}.$$
 Să se determine:

- a) probabilitatea de a primi 5 puncte;
- b) funcția de repartiție a variabilei X ;
- c) probabilitatea de a primi mai mult de 1 punct;
- d) distribuția variabilei X^2 .

2SEMINAR 6. VARIABILE ALEATOARE DISCRETE. VECTORI ALEATORI DISCREȚI

e) numărul mediu de puncte care se acordă și dispersia variabilei X ;

Rezolvare:

a) Se știe că $\sum_{i=0}^6 p_i = 1$, deci $p = P(X = 5) = \frac{1}{4}$.

b) $F_X(x) := P(X \leq x)$. Pentru o variabilă aleatoare discretă avem $F_X(x) = \sum_{i: x_i \leq x} p_i$.

$$\text{Deci, } F_X(x) = \begin{cases} 0 & \text{pentru } x < -3 \\ 1/12 & \text{pentru } -3 \leq x < -1 \\ 3/12 & \text{pentru } -1 \leq x < 0 \\ 4/12 & \text{pentru } 0 \leq x < 1 \\ 8/12 & \text{pentru } 1 \leq x < 3 \\ 9/12 & 3 \leq x \leq 5 \\ 1 & x > 5 \end{cases}$$

c) $P(X > 1) = P(X = 3) + P(X = 5) = \frac{4}{12}$ sau
 $P(X > 1) = 1 - P(X \leq 1) = 1 - F(1) = 1 - \frac{8}{12} = \frac{4}{12}$.

d) **Observație:** : In general, dacă X o variabilă aleatoare cu distribuția de probabilitate:

$$X = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix},$$

iar $\varphi : E \supset D_X \rightarrow \mathbb{R}$ este o funcție reală continuă sau cu discontinuități de prima speță, atunci dacă $Y = \varphi \circ X$, rezultă că Y este de asemenea o variabilă aleatoare discretă, având valorile $y_i = \varphi(x_i)$, $i \in I$. Pentru ca să determinăm distribuția de probabilitate a variabilei aleatoare $Y = \varphi(X)$, avem:

$$P(Y = y_i) = P(\varphi(X) = y_i) = P(X \in \varphi^{-1}(\{y_i\})), \quad (6.1)$$

unde $\varphi^{-1}(\{y_i\}) = \{x \in D_X \mid \varphi(x) = y_i\}$ este preimaginea mulțimii $\{y_i\}$, adică mulțimea elementelor din D_X ce sunt aplicate în y_i .

Considerând funcția $\varphi : \mathbb{R} \rightarrow \mathbb{R}$, $\varphi(x) = x^2$, variabila aleatoare $Y = X^2 = \varphi(X)$ ia valorile $D_Y = \{(-3)^2, (-1)^2, 0, 1, 9, 25\} = \{0, 1, 9, 25\}$. Fie $y \in \{0, 1, 9, 25\}$.

$$P(Y = y) = P(X^2 = y) = P(X \in \varphi^{-1}(y)) = P((X = -\sqrt{y}) \cup (X = \sqrt{y})) = P(X = -\sqrt{y}) + P(X = \sqrt{y}).$$

Astfel avem:

$$P(Y = 0) = P(X^2 = 0) = P(X = 0) = 1/12$$

$$P(Y = 1) = P(X^2 = 1) = P(X = -1) + P(X = 1) = 1/6 + 1/3 = 1/2$$

$$P(Y = 9) = P(X^2 = 9) = P(X = -3) + P(X = 3) = 1/12 + 1/12 = 1/6.$$

$$P(Y = 25) = P(X^2 = 25) = P(X = -5) + P(X = 5) = 0 + 1/4 = 1/4.$$

Deci, distribuția de probabilitate a variabilei $Y = X^2$ este:

$$Y = X^2 = \begin{pmatrix} 0 & 1 & 9 & 25 \\ \frac{1}{12} & \frac{1}{2} & \frac{1}{6} & \frac{1}{4} \end{pmatrix}.$$

$$\text{e) } M(X) = \sum_i x_i p_i = \frac{17}{12}.$$

Dispersia se determină cu formula $\sigma^2(X) = M(X^2) - [M(X)]^2$, unde $M(X^2) = 0 \cdot \frac{1}{12} +$

$1 \cdot \frac{1}{2} + 9 \cdot \frac{1}{6} + 25 \cdot \frac{1}{4} = \frac{99}{12}$ Avem, $\sigma^2(X) = 6.25$.

Observație: (Legea LOTUS pentru varb. discrete="law of the unconscious statistician"). Dacă $Y = \varphi(X)$, atunci

$$M(Y) = M(\varphi(X)) = \sum_{x_i \in D_X} \varphi(x_i) p_i$$

Cu LOTUS, avem că $M(X^2) = \sum_{x \in D_X} x_i^2 p_i = \frac{99}{12}$

2. Fie o variabila aleatoare X ce înregistrează numărul de apeluri la o centrală telefonică pe oră. Se știe că această centrală deservește până la 20 de apeluri/h. Dacă X este o variabilă aleatoare discretă și are funcția de repartiție: $F_X : \mathbb{R} \rightarrow [0, 1]$,

$$F_X(x) = \begin{cases} 0 & \text{pentru } x < 1 \\ 1/10 & \text{pentru } 1 \leq x < 3 \\ 3/10 & \text{pentru } 3 \leq x < 10 \\ 5/10 & \text{pentru } 10 \leq x < 15 \\ 9/10 & \text{pentru } 15 \leq x < 20 \\ 1 & x > 20 \end{cases}, \text{ să se determine:}$$

a) distribuția de probabilitate a variabilei X ;

b) $P(2 \leq X < 8)$.

Rezolvare:

a) Din definiția funcției de repartiție avem că $D_X = \{1, 3, 10, 15, 20\}$.

În plus, avem că $P(X = x_0) = F_X(x_0) - \lim_{x \rightarrow x_0, x < x_0} F_X(x)$.

Deci, $P(X = 1) = F(1) - \lim_{x \rightarrow 1, x < 1} F_X(x) = \frac{1}{10} - 0 = \frac{1}{10}$.

Obținem $X = \begin{pmatrix} 1 & 3 & 10 & 15 & 20 \\ \frac{1}{10} & \frac{2}{10} & \frac{2}{10} & \frac{4}{10} & \frac{1}{10} \end{pmatrix}$.

b) $P(2 \leq X < 8) = P(X = 3) = \frac{2}{10}$.

3. Un blogger realizează 0, 1 sau 2 postări, sâmbăta și 0 sau 1 duminică. Notăm cu S , respectiv D variabilele aleatoare ce dau numărul de postări făcute sâmbăta, respectiv, duminică. Vectorul aleator (S, D) are distribuția de probabilitate:

		D	
		0	1
	0	0.1	0.1
S	1	0.3	0.2
	2	0.1	0.2

a) Să se determine distribuția numărului de postări pentru sâmbătă, respectiv duminică.

b) Este numărul de postări afișate sâmbăta independent de numărul celor afișate duminică?

c) Să se calculeze probabilitatea ca în cele două zile să afișeze două postări.

4SEMINAR 6. VARIABLE ALEATOARE DISCRETE. VECTORI ALEATORI DISCREȚI

d) Să se determine distribuția v.a. ce dă numărul maxim de postări din cele două zile $Z = \max(S, D)$.

Rezolvare: a) Vom calcula probabilitățile marginale $P(S = i)$, $i = 0, 1, 2$, $P(D = j)$, $j = 0, 1$ și acest lucru se obține însumând pe linie, respectiv pe coloană probabilitățile din tabel. Avem:

		D		
		0	1	$p_{i.}$
	0	0.1	0.1	0.2
S	1	0.3	0.2	0.5
	2	0.1	0.2	0.3
	$p_{.j}$	0.5	0.5	

Obținem: $S = \begin{pmatrix} 0 & 1 & 2 \\ 0.2 & 0.5 & 0.3 \end{pmatrix}$, $D = \begin{pmatrix} 0 & 1 \\ 0.5 & 0.5 \end{pmatrix}$.

b) Să verificăm dacă $P(S = i, D = j) = P(S = i)P(D = j)$, $\forall i = 0, 1, 2, \forall j = 0, 1$. Se observă că pentru $i = 0, j = 0$, avem $P(S = 0, D = 0) = 0.1 = P(S = 0)P(D = 0) = 0.2 \cdot 0.5$, la fel se verifică și pentru $i = 0, j = 1$, avem $0.1 = 0.2 \cdot 0.5$.

Dar, pentru $i = 1, j = 0$, $P(S = 1, D = 0) = 0.3 \neq P(S = 1)P(D = 0) = 0.5 \cdot 0.5$.

În concluzie, variabilele S și D nu sunt independente.

c) Trebuie să calculăm $P(S + D = 2)$. Evenimentul $(S + D = 2) = (S = 1, D = 1) \cup (S = 2, D = 0)$. Deci $P(S + D = 2) = P(S = 1, D = 1) + P(S = 2, D = 0) = 0.2 + 0.1 = 0.3$.

d) Variabila $Z = \max(S, D)$ are valorile $D_Z = \{0, 1, 2\}$.

Avem: $P(Z = 0) = P(\max(S, D) = 0) = P(S = 0, D = 0) = 0.1$.

În plus, $P(Z = 1) = P(\max(S, D) = 1) = P((S = 0, D = 1) \cup (S = 1, D = 0) \cup (S = 1, D = 1)) = 0.1 + 0.3 + 0.2 = 0.6$, iar $P(Z = 2) = P(\max(S, D) = 2) = P((S = 2, D = 0) \cup (S = 2, D = 1)) = 0.1 + 0.2 = 0.3$ Avem

$$Z = \begin{pmatrix} 0 & 1 & 2 \\ 0.1 & 0.6 & 0.3 \end{pmatrix}.$$

4. Firma ta acordă asistență telefonică clienților. Evenimentele pe care le raportează clienții sunt:

A: "PC-ul este virusat" și B: "un fișier sistem este corupt".

Probabilitățile producerii **simultane** a combinațiilor dintre aceste două evenimente și opusele lor sunt date în tabloul:

	A	\bar{A}
B	0.01	0.19
\bar{B}	0.02	0.78

De exemplu, numărul 0.19 din tabelul de mai sus reprezintă probabilitatea evenimentului $\bar{A} \cap B$, adică $P(\bar{A} \cap B) = 0.19$.

Fie (X, Y) vectorul aleator ale cărui coordonate sunt variabile aleatoare Bernoulli. X ia valoarea 1 dacă se produce evenimentul A și valoarea 0 dacă se produce evenimentul \bar{A} , iar Y ia valoare 1 dacă se produce B și 0 în caz contrar.

- a) Să se deducă distribuțiile marginale ale lui X și Y
 b) Să se calculeze $P(X = 1|Y = 0)$

Rezolvare: a) Din datele problemei rezultă că vectorul aleator (X, Y) are distribuția din tabelul:

		X	
		1	0
Y	1	0.01	0.19
	0	0.02	0.78

De exemplu,

$$P(X = 1, Y = 0) = P(A \cap \bar{B}) = 0.02.$$

Variabila aleatoare X are distribuția

$$X = \begin{pmatrix} 1 & 0 \\ 0.01 + 0.02 & 0.19 + 0.78 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0.03 & 0.97 \end{pmatrix},$$

iar distribuția variabilei Y este

$$Y = \begin{pmatrix} 1 & 0 \\ 0.01 + 0.19 & 0.02 + 0.78 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0.2 & 0.8 \end{pmatrix}.$$

Vom calcula desfășurat $P(X = 1)$. Avem

$$\begin{aligned} P(X = 1) &= P((X = 1, Y = 0) \cup (X = 1, Y = 1)) \\ &= P(X = 1, Y = 0) + P(X = 1, Y = 1) \\ &= 0.01 + 0.02 = 0.03. \end{aligned}$$

b) Avem:

$$P(X = 1|Y = 0) = \frac{P(X = 1, Y = 0)}{P(Y = 0)} = \frac{0.02}{0.8} = \frac{1}{40}.$$

5. Fie X, Y două variabile aleatoare discrete ce pot lua valorile $\{1, 2, 3, 4\}$. Distribuția de probabilitate comună a celor două variabile este dată în tabloul:

			Y		
		1	2	3	4
	1	0.03	0.05	0.1	0.12
X	2	0.05	0.06	0.08	0.07
	3	0.07	0.06	0.06	0.02
	4	0.07	0.09	0.05	0.02

6SEMINAR 6. VARIABILE ALEATOARE DISCRETE. VECTORI ALEATORI DISCREȚI

- a) Să se determine distribuțiile marginale;
 b) Să se determine distribuția de probabilitate a variabilei condiționate ($X|Y = 2$).

Rezolvare: a)

			Y			
		1	2	3	4	p_X
	1	0.03	0.05	0.1	0.12	0.3
X	2	0.05	0.06	0.08	0.07	0.26
	3	0.07	0.06	0.06	0.02	0.21
	4	0.07	0.09	0.05	0.02	0.23
	p_Y	0.22	0.26	0.29	0.23	

b) Variabila ($X|Y = 2$) are valorile $\{1, 2, 3, 4\}$ și trebuie să calculăm probabilitatea cu care sunt luate aceste valori. Avem că: $P(X = 1|Y = 2) = \frac{P(X=1,Y=2)}{P(Y=2)} = \frac{0.05}{0.26}$. Distribuția de probabilitate a variabilei condiționate ($X|Y = 2$) este:

$$\begin{aligned}
 (X|Y = 2) &= \left(\frac{1}{\frac{p_{12}}{p_Y(2)}} \quad \frac{2}{\frac{p_{22}}{p_Y(2)}} \quad \frac{3}{\frac{p_{32}}{p_Y(2)}} \quad \frac{4}{\frac{p_{42}}{p_Y(2)}} \right) = \\
 &= \left(\frac{1}{\frac{0.05}{0.26}} \quad \frac{2}{\frac{0.06}{0.26}} \quad \frac{3}{\frac{0.06}{0.26}} \quad \frac{4}{\frac{0.09}{0.26}} \right)
 \end{aligned}$$

6.2 Probleme propuse

6. Fie X o v.a. discretă a cărei distribuție de probabilitate p_X , definită prin

$$p_X(k) = P(X = k),$$

cu valorile: $p_X(-4) = 1/4$, $p_X(-\pi) = 1/4$, $p_X(0) = 1/4$, $p_X(4) = 1/8$, $p_X(\pi) = 1/8$. Funcția de repartiție a v.a. X este F .

- a) Să se calculeze $F(-5)$, $F(-3)$, $F(\pi)$ și $F(2\pi)$;
 b) Să se calculeze $P(X \geq 0)$;
 c) Să se calculeze media v. a. X .

7. O v.a. discretă X are funcția de repartiție:

$$F(x) = \begin{cases} 0 & \text{dacă } x < 0 \\ \frac{1}{3} & \text{dacă } 0 \leq x < 1 \\ \frac{2}{3} & \text{dacă } 1 \leq x < 2 \\ 1 & \text{dacă } x \geq 2 \end{cases}$$

Să se determine distribuția de probabilitate a v.a. X , media $M(X)$ și dispersia $\sigma^2(X)$.

8. Studenților din anul I li se oferă 3 șanse de a promova un examen. Fie X v.a. ce asociază fiecărui student promovat numărul 1, 2 și 3, al prezentării în care acesta a promovat examenul. Distribuția de probabilitate a variabilei X , dedusă din anii precedenți este

$$p_k = P(X = k) = \frac{0.4^{k-1}0.6}{0.936}, k = 1, 2, 3$$

Să se determine procentul de studenți care au promovat acest examen în cel puțin două încercări?

9. Distribuția de probabilitate a vectorului aleator (X, Y) este:

			Y		
		1	2	3	5
	-1	0.06	0.02	0.06	0.12
X	0	0.08	0.12	0.20	0.10
	1	0.16	0.02	0.02	0.04

a) Să se afle distribuția marginală alui X și Y și să se calculeze probabilitățile $P(Y \geq 2)$, $P(X \leq 0)$.

b) Sunt variabilele X și Y independente?

c) Să se determine distribuția v.a. XY .

10. Vectorul aleator (X, Y) are distribuția de probabilitate:

			Y		
		1	2	3	4
	1	0.03	0.05	0.1	0.12
X	2	0.05	0.06	0.08	0.07
	3	0.07	0.06	0.06	0.02
	4	0.07	0.09	0.05	0.02

a) Să se calculeze $P(X \leq 2, Y \leq 2)$;

b) Să se determine distribuțiile v.a. condiționate $(X|Y = 3)$ și $(Y|X = 4)$.

c) Să se determine distribuția v.a. $X + Y, \min(X, Y)$.