Vjerojatnost i statistika

Diskretan slučajni vektor. Kovarijanca i koeficijent korelacije

5. prosinca 2023.

Diskretan slučajni vektor

Definicija 1. Neka je $(\Omega, \mathcal{F}, P)$ vjerojatnosni prostor pridružen slučajnom pokusu. Funkciju $(X_1, \dots, X_n)$ koja svakom ishodu slučajnog pokusa $\omega \in \Omega$ pridružuje uređenu n-torku realnih brojeva $(x_1, \dots, x_n)$ zovemo n-**dimenzionalni slučajni vektor** ako vrijedi

$$\{X_1 \leq x_1\} \cap \cdots \cap \{X_n \leq x_n\} \in \mathcal{F}$$

za sve $x_1, \dots, x_n \in \mathbb{R}$.

Definicija 2. Neka je $\mathbb{Z} = (X, Y)$ slučajni vektor. Distribucije njegovih komponenti dane su s

$$X \sim \begin{pmatrix} x_1 & x_2 & \dots \\ p_1 & p_2 & \dots \end{pmatrix}, \quad Y \sim \begin{pmatrix} y_1 & y_2 & \dots \\ q_1 & q_2 & \dots \end{pmatrix}.$$

Distribucija vektora $\mathbb{Z}$ zadana je s

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

pri čemu je $p_{ij} \geq 0$ i $\sum_{i,j} p_{ij} = 1$.

Uz pretpostavku da imamo konačno mnogo vrijednosti x_i i y_j, distribuciju vektora $\mathbb{Z}$ najčešće zapisujemo tablično

X/Y	y_1	y_2	y_3	$\dots$	y_m	
x_1	p_{11}	p_{12}	p_{13}	$\dots$	p_{1m}	p_1
x_2	p_{21}	p_{22}	p_{23}	$\dots$	p_{2m}	p_2
x_3	p_{31}	p_{32}	p_{33}	$\dots$	p_{3m}	p_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
x_n	p_{n1}	p_{n2}	p_{n3}	$\dots$	p_{nm}	p_n
	q_1	q_2	q_3	$\dots$	q_m	

Napomena 1. Distribucije slučajnih varijabli X i Y koje čine slučajni vektor $\mathbb{Z} = (X, Y)$ zovu se **marginalne** distribucije. One se lako iščitaju iz tablice distribucije od $\mathbb{Z}$

$$X \sim \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}, \quad Y \sim \begin{pmatrix} y_1 & y_2 & \dots & y_m \\ q_1 & q_2 & \dots & q_m \end{pmatrix}.$$

Definicija 3. Uvjetna distribucija od X uz uvjet da je $Y = y_j$ dana je tablicom

$$X|_{Y=y_j} \sim \begin{pmatrix} x_1 & x_2 & \dots \\ p_1|_{Y=y_j} & p_2|_{Y=y_j} & \dots \end{pmatrix},$$

pri čemu je

$$p_i|_{Y=y_j} = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{q_j}, \ i = 1, 2, \dots$$

Napomena 2. Neka su X i Y diskretne slučajne varijable na $(\Omega, \mathcal{F}, P)$ s distribucijama

$$X \sim \begin{pmatrix} x_1 & x_2 & \dots \\ p_1 & p_2 & \dots \end{pmatrix}, \quad Y \sim \begin{pmatrix} y_1 & y_2 & \dots \\ q_1 & q_2 & \dots \end{pmatrix}.$$

Slučajne varijable X i Y su **nezavisne** ako i samo ako

$$p_{ij} = P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j) = p_i q_j$$

za sve $i = 1, \dots, n$ te za sve $j = 1, \dots, m$.

Definicija 4. Neka je $\mathbb{X} = (X_1, \ldots, X_n)$ slučajni vektor na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, P)$. Ako je $E|X_i| < \infty$ za svaki i, kažemo da postoji očekivanje slučajnog vektora $\mathbb{X}$ koje je dano s

$$E\mathbb{X} = \begin{bmatrix} EX_1 \\ \vdots \\ EX_n \end{bmatrix}.$$

Teorem 1. *Ako su slučajne varijable X i Y nezavisne i ako postoje EX i EY, tada slučajna varijabla XY ima očekivanje i vrijedi*

$$E(XY) = EX \cdot EY.$$

Kovarijanca i koeficijent korelacije

Definicija 5. Neka su X i Y slučajne varijable takve da postoje EX, EY i $E(XY)$. Tada definiramo kovarijancu slučajnih varijabli X i Y s

$$\operatorname{Cov}(X, Y) = E\big((X - EX)(Y - EY)\big) = E(XY) - EX \cdot EY.$$

Teorem 2. *Neka je (X, Y) slučajni vektor za koji postoje EX i EY. Ako su slučajne varijable X i Y nezavisne, onda je* $\operatorname{Cov}(X, Y) = 0$.

Obrat ovog teorema općenito ne vrijedi. Ako je $\operatorname{Cov}(X, Y) = 0$, to ne znači nužno da su slučajne varijable X i Y nezavisne. Za takve slučajne varijable ćemo reći da su **nekorelirane**.

Definicija 6. Broj

$$\rho_{X,Y} = \frac{\operatorname{Cov}(X, Y)}{\sqrt{\operatorname{Var}(X) \operatorname{Var}(Y)}} \in [-1, 1]$$

zove se **koeficijent korelacije** slučajnih varijabli X i Y.

Zadatak 1. Slučajan pokus sastoji se od dva uzastopna bacanja simetričnog novčića. Neka je (X, Y) slučajni vektor pri čemu X označava broj glava, a Y broj pisama koja se realiziraju u ta dva bacanja. Odredite

a) distribuciju i marginalne distribucije slučajnog vektora (X, Y),

b) uvjetnu distribuciju $X|_{Y=1}$,

c) koeficijent korelacije $\rho_{X,Y}$.

Zadatak 2. Neka je (X, Y) slučajni vektor kod kojeg X predstavlja broj izlazaka trgovačkog putnika na teren, a Y broj prodanih knjiga. Distribucija slučajnog vektora zadana je tablicom

X/Y	0	1	2	3	4	5	6
1	1/32	1/16	1/32	0	0	0	0
2	0	1/8	1/16	1/32	1/32	0	0
3	0	1/16	1/8	1/8	1/8	1/16	0
4	0	1/128	1/64	1/32	1/32	1/32	1/128

Odredite

a) marginalne distribucije slučajnog vektora (X, Y),

b) vjerojatnost da je trgovački putnik prodao četiri knjige, ako je dva puta izašao na teren,

c) vjerojatnost da je trgovački putnik prodao barem četiri knjige, ako je četiri puta izašao na teren.

Zadatak 3. Promotrimo slučajan pokus koji se sastoji od bacanja igraće kockice dva puta zaredom. Neka je (X, Y) slučajni vektor u kojemu X predstavlja broj realiziranih trojki, a Y broj realiziranih dvojki. Odredite

a) distribuciju i marginalne distribucije slučajnog vektora (X, Y),

b) uvjetnu distribuciju $X|_{Y=1}$,

c) koeficijent korelacije $\rho_{X,Y}$.

Zadatak 4. Bacamo dvije igraće kockice. Neka je X manji, a Y veći od dva pojavljena broja. Odredite

a) distribuciju i marginalne distribucije slučajnog vektora (X, Y),

b) uvjetnu distribuciju $X|_{Y=4}$,

c) vjerojatnost događaja $A = (X \geq 2|Y = 4)$ i $B = (Y = 4|X \geq 2)$.

Zadatak 5. Promotrimo slučajan pokus koji se sastoji od bacanja dvaju novčića tri puta za redom. Novčić A je simetričan ($P_A(G) = P_A(P) = 0.5$), ali novčić B nije ($P_B(G) = 0.25$, $P_B(P) = 0.75$. Neka je (X, Y) slučajni vektor u kojem X predstavlja broj glava realiziranih bacanjem novčića A, a Y broj glava realiziranih bacanjem novčića B. Odredite

a) distribuciju i marginalne distribucije slučajnog vektora (X, Y),

b) uvjetnu distribuciju $X|_{Y=2}$,

c) koeficijent korelacije $\rho_{X,Y}$.

Zadatak 6. Distribucija slučajnog vektora (X, Y) zadana je s

$$P(X = x_i, Y = y_j) = \begin{cases} k(2x_i + y_j), & x_i \in \{1, 2\},\ y_j \in \{1, 2\}, \\ 0, & \text{inače.} \end{cases}$$

Odredite

a) konstantu k,

b) marginalne distribucije slučajnog vektora (X, Y),

c) $P(X = 2|Y = 2)$.

Rješenja

1.

a)

X/Y	0	1	2	
0	0	0	1/4	1/4
1	0	1/2	0	1/2
2	1/4	0	0	1/4
	1/4	1/2	1/4	

b) $X|_{Y=1} = 1$ g.s.

c) $\rho_{X,Y} = -1$

2.

a) $X \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1/8 & 1/4 & 1/2 & 1/8 \end{pmatrix}$,

$Y \sim \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 1/32 & 33/128 & 15/64 & 3/16 & 3/16 & 3/32 & 1/128 \end{pmatrix}$

b) $P(Y = 4|X = 2) = 1/8$

c) $P(Y \geq 4|X = 4) = 9/16$

3.

a)

X/Y	0	1	2	
0	4/9	2/9	1/36	25/36
1	2/9	1/18	0	5/18
2	1/36	0	0	1/36
	25/36	5/18	1/36	

b) $X|_{Y=1} \sim \begin{pmatrix} 0 & 1 \\ 4/5 & 1/5 \end{pmatrix}$

c) $\rho_{X,Y} = -1/5$

4.

a)

X/Y	1	2	3	4	5	6	
1	1/36	2/36	2/36	2/36	2/36	2/36	11/36
2	0	1/36	2/36	2/36	2/36	2/36	9/36
3	0	0	1/36	2/36	2/36	2/36	7/36
4	0	0	0	1/36	2/36	2/36	5/36
5	0	0	0	0	1/36	2/36	3/36
6	0	0	0	0	0	1/36	1/36
	1/36	3/36	5/36	7/36	9/36	11/36	

b) $X|_{Y=4} \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2/7 & 2/7 & 2/7 & 1/7 \end{pmatrix}$

c) $P(X \geq 2|Y = 4) = 5/7, \;\; P(Y = 4|X \geq 2) = 5/16$

5. $X \sim \mathcal{B}(3, 0.5)$, $Y \sim \mathcal{B}(3, 0.25)$

a) $X = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 1/8 & 3/8 & 3/8 & 1/8 \end{pmatrix}$, $\quad Y = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 27/64 & 27/64 & 9/64 & 1/64 \end{pmatrix}$,

X / Y	0	1	2	3	
0	27/512	27/512	9/512	1/512	64/512
1	81/512	81/512	27/512	3/512	192/512
2	81/512	81/512	27/512	3/512	192/512
3	27/512	27/512	9/512	1/512	64/512
	216/512	216/512	72/512	8/512	

b) $X|_{Y=2} \stackrel{d}{=} X$

c) $\rho_{X,Y} = 0$

6.

a)

X / Y	1	2	
1	3k	4k	7k
2	5k	6k	11k
	8k	10k	

, $\Rightarrow k = 1/18$

b) $X = \begin{pmatrix} 1 & 2 \\ 7/18 & 11/18 \end{pmatrix}$, $\quad Y = \begin{pmatrix} 1 & 2 \\ 4/9 & 5/9 \end{pmatrix}$

c) $P(X = 2|Y = 2) = 3/5$