

4. Diskretni slučajni vektori – (ne)zavisnost

Motivacija

- Za primjene ključno je promatrati više slučajnih varijabli koje su ishodi **istog** vjerojatnosnog modela, tj. koje su definirane na **istom** $(\Omega, \mathcal{F}, \mathbb{P})$.
- Na primjer, ako su
 - (a) Y indikator događaja da osoba koju promatramo ima ili nema određenu bolest ($Y \in \{0, 1\}$), te
 - (b) za $p \in \mathbb{N}$, $X_1, \dots, X_p$ razna mjerenja (tlak, puls i slično) za tu **istu osobu**,zanima nas kako mjerenja $X_1, \dots, X_p$ utječu na vjerojatnost da je $Y = 1$.
- Zapravo, zanima nas **zajednička razdioba** slučajnih varijabli $X_1, \dots, X_p, Y$, tj. razdioba **slučajnog vektora** $(X_1, \dots, X_p, Y)$.

4.1. Zajednička razdioba

Definicija 4.1. Ako je $d \in \mathbb{N}$ te $X_1, \dots, X_d$ slučajne varijable na istom vj. prostoru $(\Omega, \mathcal{F}, \mathbb{P})$, funkciju $(X_1, \dots, X_d) : \Omega \rightarrow \mathbb{R}^d$ zovemo **(d -dimenzionalan) slučajni vektor**.¹ Kažemo da je **diskretan** ako su $X_1, \dots, X_d$ diskretne slučajne varijable. □

Ako je $(X_1, \dots, X_d)$ diskretan slučajan vektor te $D_i := D_{X_i} = \text{Im}(X_i)$, on poprima vrijednosti u skupu $D_1 \times \dots \times D_d =: D$, koji je ponovno najviše prebrojiv.

$\rightsquigarrow (X_1, \dots, X_d)$ je **diskretan** slučajni element u $\mathbb{R}^d$

$\rightsquigarrow$ **Distribuciju** diskretnog slučajnog vektora definiramo analogno kao u jednodimenzionalnom (diskretnom) slučaju, te vrijede ista osnovna svojstva.

¹Za $\omega \in \Omega$, definiramo $(X_1, \dots, X_d)(\omega) := (X_1(\omega), \dots, X_d(\omega))$. 

Distribucija slučajnog vektora

Skup D zajedno s vjerojatnostima²

$$\mathbb{P}(X_1 = x_1, \dots, X_d = x_d), (x_1, \dots, x_d) \in D,$$

naziva se **zajednička distribucija** slučajnih varijabli $X_1, \dots, X_d$ ili **distribucija** slučajnog vektora $(X_1, \dots, X_d)$.

Diskretna funkcija gustoće [ili vjerojatnosna funkcija mase] vektora $(X_1, \dots, X_d)$ je funkcija $f_{X_1, \dots, X_d} = f : \mathbb{R}^d \rightarrow [0, 1]$, definirana s

$$f(x_1, \dots, x_d) := \mathbb{P}(X_1 = x_1, X_2 = x_2, \dots, X_d = x_d), x_1, \dots, x_d \in \mathbb{R},$$

te ona jedinstveno određuje distribuciju vektora $(X_1, \dots, X_d)$.

²Događaj $\{X_1 = x_1, \dots, X_d = x_d\}$ je oznaka za događaj

$$\{\omega \in \Omega : X_1(\omega) = x_1, \dots, X_d(\omega) = x_d\},$$

a nekad ćemo koristiti i oznaku $\{(X_1, \dots, X_d) = (x_1, \dots, x_d)\}$.

4.1.1. Osnovna svojstva

Pretpostavimo da je $d = 2$ te $D_1 = \{a_i : i \in I\}$, $D_2 = \{b_j : j \in J\}$, te neka je

$$p_{ij} := \mathbb{P}(X_1 = a_i, X_2 = b_j), \text{ za sve } i, j.$$

- Za sve $C \subseteq \mathbb{R}^2$ vrijedi

$$\mathbb{P}((X_1, X_2) \in C) = \sum_{(a_i, b_j) \in C} \mathbb{P}(X_1 = a_i, X_2 = b_j) = \sum_{(a_i, b_j) \in C} p_{ij}. \quad (4.1)$$

- tzv. **marginalne** razdiobe od X_1 i X_2 su:

$$\mathbb{P}(X_1 = a_i) = \mathbb{P}((X_1, X_2) \in \{a_i\} \times D_2) \stackrel{(4.1)}{=} \sum_{b_j \in D_2} p_{ij}, \quad i \in I$$

$$\mathbb{P}(X_2 = b_j) = \sum_{a_i \in D_1} p_{ij}, \quad j \in J.$$

[Marginalne razdiobe općenito **ne određuju** zajedničku distribuciju!]

- Za svaki $b \in \mathbb{R}$ t.d. je $\mathbb{P}(X_2 = b) > 0$, **uvjetna razdioba** od X_1 uz dano $X_2 = b$ je uvjetna razdioba od X_1 uz dano $A = \{X_2 = b\} \in \mathcal{F}$, tj.

$$\mathbb{P}(X_1 = a_i \mid X_2 = b) = \frac{\mathbb{P}(X_1 = a_i, X_2 = b)}{\mathbb{P}(X_2 = b)}, \quad \forall i \in I.$$

- Ako je $g : D \rightarrow \mathbb{R}^k$ proizvoljna funkcija, tada je $g(X_1, X_2) : \Omega \rightarrow \mathbb{R}^k$ diskretan slučajni vektor s distribucijom

$$\mathbb{P}(g(X_1, X_2) = c) = \sum_{g(a_i, b_j) = c} p_{ij}, \quad \forall c \in g(D). \quad (4.2)$$

Primjer 4.2

U kutiji su 1 crna, 2 bijele i 3 zelene kuglice. Slučajno izvučemo 2 kuglice te označimo s X ukupan broj crnih, a s Y ukupan broj bijelih kuglica koje smo izvukli. Odredite

- (a) razdiobu vektora (X, Y) ;
- (b) marginalne razdiobe;
- (c) razdiobu od Y uvjetno na $X = 1$;
- (d) $\mathbb{P}(X \geq Y)$;
- (e) razdiobu slučajne varijable $X + Y$, tj. ukupnog broja crnih i bijelih kuglica.

Rješenje (a) razdioba vektora (X, Y)

Imamo $X \in \{0, 1\}$, $Y \in \{0, 1, 2\}$ te

$$\mathbb{P}(X = 0, Y = 0) = \mathbb{P}(\{\text{izvukli samo zelene kuglice}\}) = \frac{\binom{3}{2}}{\binom{6}{2}} = \frac{3}{15},$$

$$\mathbb{P}(X = 0, Y = 1) = \mathbb{P}(\{1 \text{ bijela i } 1 \text{ zelena}\}) = \frac{\binom{2}{1} \cdot \binom{3}{1}}{\binom{6}{2}} = \frac{6}{15}, \text{ itd.}$$

U **dvodimenzionalnom** slučaju, distribuciju možemo zapisati i **tablično**

$X \backslash Y$	0	1	2
0	$\frac{3}{15}$	$\frac{6}{15}$	$\frac{1}{15}$
1	$\frac{3}{15}$	$\frac{2}{15}$	0

Rješenje (b) marginalne razdiobe od X i Y

$X \backslash Y$	0	1	2
0	$\frac{3}{15}$	$\frac{6}{15}$	$\frac{1}{15}$
1	$\frac{3}{15}$	$\frac{2}{15}$	0

Ako označimo $p_{ij} := \mathbb{P}(X = i, Y = j)$, imamo npr.

$$\mathbb{P}(X = 0) = p_{00} + p_{01} + p_{02} = \frac{3}{15} + \frac{6}{15} + \frac{1}{15} = \frac{10}{15} = \frac{2}{3}.$$

Iz tablice

$X \backslash Y$	0	1	2	$\mathbb{P}(X = \cdot)$
0	$\frac{3}{15}$	$\frac{6}{15}$	$\frac{1}{15}$	$\frac{10}{15}$
1	$\frac{3}{15}$	$\frac{2}{15}$	0	$\frac{5}{15}$
$\mathbb{P}(Y = \cdot)$	$\frac{6}{15}$	$\frac{8}{15}$	$\frac{1}{15}$	1

dakle slijedi

$$X \sim \begin{pmatrix} 0 & 1 \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}, \quad Y \sim \begin{pmatrix} 0 & 1 & 2 \\ \frac{6}{15} & \frac{8}{15} & \frac{1}{15} \end{pmatrix}$$

Rješenje (c) $\mathbb{P}(Y = \cdot \mid X = 1)$

$X \backslash Y$	0	1	2	$\mathbb{P}(X = \cdot)$
0	$\frac{3}{15}$	$\frac{6}{15}$	$\frac{1}{15}$	$\frac{10}{15}$
1	$\frac{3}{15}$	$\frac{2}{15}$	0	$\frac{5}{15}$
$\mathbb{P}(Y = \cdot)$	$\frac{6}{15}$	$\frac{8}{15}$	$\frac{1}{15}$	1

Imamo

$$\mathbb{P}(Y = j \mid X = 1) = \frac{p_{1j}}{\mathbb{P}(X = 1)} \propto 3, 2, 0, \quad j = 0, 1, 2.$$

Dakle,

$$\mathbb{P}(Y \in \cdot \mid X = 1) = \begin{pmatrix} 0 & 1 & 2 \\ \frac{3}{5} & \frac{2}{5} & 0. \end{pmatrix}$$

$[\mathbb{P}(Y \in \cdot \mid X = 1) \neq \mathbb{P}(Y \in \cdot)] \rightsquigarrow X$ i Y su **zavisne!**

Rješenje (d) $\mathbb{P}(Y = \cdot \mid X = 1)$

$X \backslash Y$	0	1	2
0	$\frac{3}{15}$	$\frac{6}{15}$	$\frac{1}{15}$
1	$\frac{3}{15}$	$\frac{2}{15}$	0

Uz $C = \{(a, b) : a \geq b\}$, dobivamo

$$\begin{aligned}\mathbb{P}(X \geq Y) &= [\mathbb{P}((X, Y) \in C)] = p_{00} + p_{10} + p_{11} \\ &= \frac{1}{15} \cdot (3 + 3 + 2) = \frac{8}{15}.\end{aligned}$$

Rješenje (e) razdioba od $X + Y$

Imamo $X + Y = g(X, Y)$ pa iz

$X \backslash Y$	0	1	2
0	$\frac{3}{15}$	$\frac{6}{15}$	$\frac{1}{15}$
1	$\frac{3}{15}$	$\frac{2}{15}$	0

$X \backslash Y$	0	1	2
0	0	1	2
1	1	2	3

sljedeći $X + Y \in \{0, 1, 2\}$ te da je distribucija od $X + Y$ dana s

$$\mathbb{P}(X + Y = 0) = p_{00} = \frac{3}{15},$$

$$\mathbb{P}(X + Y = 1) = p_{10} + p_{01} = \frac{3}{15} + \frac{6}{15} = \frac{9}{15},$$

$$\mathbb{P}(X + Y = 2) = 1 - \frac{3}{15} - \frac{9}{15} = \frac{3}{15}.$$