4) Diskretne slučajne varijable

Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ vjerojatnosni prostor.

Def. 4.1 Slučajna varijabla je svaka f-ja $\boxed{X : \Omega \to \mathbb{R}}$ t.d.

$$\{X \le x\} := \{\omega \in \Omega : X(\omega) \le x\}$$

("izmjerivost")

$$= X^{-1}(\langle -\infty, x]) \in \mathcal{F}, \qquad (4.1)$$

za sve $x \in \mathbb{R}$ ∎

Nap. 4.2 Za sluč. varijablu X tipično nas zanimaju vjerojatnosti

$$\mathbb{P}(X \in B) := \mathbb{P}(\{\omega \in \Omega : X(\omega) \in B\}),$$

za razne $B \subseteq \mathbb{R}$; npr.

- $B = \langle -\infty, x] \Rightarrow \mathbb{P}(X \le x)$
- $B = \langle x, +\infty \rangle \Rightarrow \mathbb{P}(X > x)$
- $B = \{x\} \Rightarrow \mathbb{P}(X = x)$.

npr. to vrijedi $\forall B \subseteq \mathbb{R}$ ako je $\mathcal{F} = \mathcal{P}(\Omega)$

Uvjet (4.1) osigurava da je $\boxed{\{X \in B\} \in \mathcal{F}}$ za "većinu" $B \subseteq \mathbb{R}$ (uključujući gore navedene) pa je za te B vjerojatnost $\mathbb{P}(X \in B)$ dobro definirana.

U nastavku uvjet (4.1) nećemo provjeravati, tj. pretpostavit ćemo da vrijedi. ∎

Def. 4.3 Slučajna varijabla $X : \Omega \to \mathbb{R}$ je diskretna ako postoji konačan ili prebrojiv skup $D \subseteq \mathbb{R}$ t.d.

$$\boxed{X(\omega) \in D, \ \forall \omega \in \Omega} \qquad ("X \in D")$$ ∎

Primjer 4.4 Bacamo kocku te neka je X broj koji je pao,

tj., $\Omega = \{1,2,3,4,5,6\}$, $X(\omega) := \omega$, $\forall \omega \in \Omega$.

Uočimo, $X : \Omega \to \mathbb{R}$ je f-ja koja poprima vrijednosti u skupu

$D = \{1,2,\dots,6\} \subseteq \mathbb{R}$ $\underset{|D|<\infty}{\Longrightarrow}$ X je diskretna sluč. var. ■

Primjer 4.5 Slučajno biramo broj iz $\langle 0,10\rangle$ te neka je

X = taj broj,

Y = najveće cijelo tog broja,

tj. $\Omega = \langle 0,10\rangle$, i za sve $\omega \in \Omega$,

$$X(\omega) := \omega,$$

$$Y(\omega) := \lfloor \omega \rfloor.$$

$\Rightarrow$ X i Y su slučajne varijable, ali samo Y je diskretna sl. var ($X \in [0,10]$, $Y \in \{0,1,\dots,9\}$) ■

NOTACIJA:

(i) Slučajne varijable obično označavamo velikim slovima, npr. $\boxed{X, Y, U}$, a konkretne vrijednosti koje sluč. varijable poprima s malim slovima, npr. $\boxed{x, y, u}$.

(ii) Ako je X diskretna sluč. varijabla i $D \subseteq \mathbb{R}$ konačan ili prebrojiv t.d. $X \in D$, često ćemo pisati

$$D = \{a_i : i \in I\},$$

gdje je $I = \{1,2,\dots,n\}$ za neki $n \in \mathbb{N}$, ili $I = \mathbb{N}$. ■

Distribucija diskretne slučajne varijable

Neka je X diskretna sluč. varijabla te $D = \{a_i : i \in I\}$ t.d. $X \in D$. Skup D zajedno s vjerojatnostima

$$p_i := \mathbb{P}(X = a_i), \quad i \in I, \qquad (4.2)$$

zovemo **distribucijom** (ili **razdiobom**) sluč. varijable X.

Distribuciju najčešće zapisujemo u obliku **tablice**:

$$X \sim \begin{pmatrix} a_1 & a_2 & \cdots \\ p_1 & p_2 & \end{pmatrix}$$

["X ima distribuciju ..."]

Primjer 4.6 Ako je X rezultat bacanja simetrične kocke,

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix} \quad \boxtimes$$

Iz (4.2) slijedi da je

(i) $p_i \geq 0, \ \forall i \in I$; $\qquad (4.3)$

(ii) $\sum_{i \in I} p_i = 1 \qquad (4.4)$

$\left(\sum_{i \in I} p_i = \sum_{i \in I} \mathbb{P}(X = a_i) = [\{X = a_i\}, i \in I, \text{ međ. disjunktni}] = \mathbb{P}(\bigcup_{i \in I} \{X = a_i\}) = \mathbb{P}(\underbrace{X \in D}_{= \Omega}) = 1 \right)$

Svaka familija brojeva $(p_i : i \in I)$ koja zadovoljava (4.3) i (4.4) naziva se **distribucija** (ili **razdioba**) na

$D = \{a_i : i \in I\}$.

<u>Prop. 4.7</u> Neka je $(p_i : i \in I)$ <u>proizvoljna</u> distribucija na $D = \{a_i : i \in I\}$. Tada postoji vjerojatnosni prostor $(\Omega, \mathcal{F}, \mathbb{P})$ i diskretna slučajna varijabla X na $(\Omega, \mathcal{F}, \mathbb{P})$ t.d.

$$X \sim \begin{pmatrix} a_1 & a_2 & \cdots \\ p_1 & p_2 & \cdots \end{pmatrix}.$$

<u>Dokaz</u> Neka je

tu koristimo (4.3) i (4.4)

- $\Omega := D$
- $\mathcal{F} := \mathcal{P}(D)$
- $\mathbb{P} : \mathcal{F} \to [0,1]$ (jedinstvena) vjerojatnost t.d.

$$\boxed{\mathbb{P}(\{a_i\}) = p_i, \ \forall i \in I}$$

(vidi Prop. 1.8 i napomenu nakon nje)

- $\boxed{X(\omega) := \omega, \ \forall \omega \in D.}$

Očito je $X \in D$, te

$$\mathbb{P}(X = a_i) = \mathbb{P}(\{\omega \in \Omega : X(\omega) = a_i\})$$

$$= \mathbb{P}(\{a_i\}) = p_i, \quad \forall i \in I \quad \blacksquare$$

Ako vrijedi $X \sim \begin{pmatrix} a_1 & a_2 & \cdots \\ p_1 & p_2 & \cdots \end{pmatrix}$, za sve $B \in \mathbb{R}$ imamo

$$\underline{\mathbb{P}(X \in B)} = \mathbb{P}(X \in B \cap D) = \mathbb{P}\Big(\bigcup_{\substack{i \in I: \\ a_i \in B}} \{X = a_i\}\Big)$$

$$= \sum_{\substack{i \in I: \\ a_i \in B}} \mathbb{P}(X = a_i) = \sum_{\substack{i \in I: \\ a_i \in B}} p_i =: \underline{\sum_{a_i \in B} p_i} \qquad (4.5)$$

Uočimo, ako je Y diskretna slučajna varijabla (moguće na nekom drugom vjerojatnosnom prostoru) koja ima <u>istu distribuciju</u> kao i X (oznaka "$X \sim Y$"), vrijedi

$$\mathbb{P}(X \in B) = \mathbb{P}(Y \in B), \quad \forall B \subseteq \mathbb{R}.$$

(Najčešće nas zanima samo distribucija slučajne varijable, tj. nije nam bitan vjerojatnosni prostor i način na koji je ona definirana (bitno je samo da postoji takva sluč. varijabla, što slijedi iz Prop. 4.7!).

<u>Neke poznate distribucije</u>

[često se pojavljuju u različitim kontekstima pa zaslužuju posebno ime!]

<u>Primjer 4.8</u> | Ako je

često $q := 1-p = \mathbb{P}(X=0)$

$$X \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$$

za neki $p \in [0,1]$, X nazivamo <u>Bernoullijeva</u> slučajna varijabla s parametrom p ($p = \mathbb{P}(X=1)$). $\rightarrow$ oznaka $\boxed{X \sim B(p)}$.

Npr., bacamo novčić te stavimo

$$X := \begin{cases} 1, & \text{ako je palo } P \\ 0, & \text{— " — } G \end{cases}$$

$$\Rightarrow X \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \text{ tj. } X \sim B\left(\tfrac{1}{2}\right). \blacksquare$$

Nap. 4.9 Ako je $A \in \mathcal{F}$ proizvoljan događaj te $\mathbb{1}_A : \Omega \to \{0,1\}$ t.d.

$$\mathbb{1}_A(\omega) := \begin{cases} 1, & \text{ako je } \omega \in A \\ 0, & \text{--- '' ---}\ \omega \notin A \end{cases},$$

"indikator" događaja A

u Pr. 4.8, $X = \mathbb{1}_{\{\text{palo je } P\}}$

tada je $\mathbb{1}_A$ Bernoullijeva sluč. var. s parametrom

$$\boxed{p = \mathbb{P}(\mathbb{1}_A = 1) = \mathbb{P}(\{\omega \in \Omega : \omega \in A\}) = \mathbb{P}(A)}$$ ■

Primjer 4.10 Neka su $n \in \mathbb{N}$ i $p \in [0,1]$ proizvoljni. Pretpost. da bacamo novčić na kojem je vjerojatnost pojave P jednaka p (u svakom bacanju), te su bacanja nezavisna. Neka je

$X :=$ ukupan broj P koja su pala u n bacanja.

Očito je $X(\omega) \in \{0,1,2,\dots,n\}$, $\forall \omega \in \Omega$, te

$$\boxed{\mathbb{P}(X = k) = \binom{n}{k} \underbrace{p^k q^{n-k}}_{\text{vjerojatnost pojave } k \text{ P i } n-k \text{ G}}, \quad k = 0,1,2,\dots,n} \quad (4.6)$$

mjesta za P

ako tiksiramo k mjesta za P, (a samim time i za $n-k$ G)

Slučajna varijabla X s distribucijom (4.7) naziva se **binomna** slučajna varijabla s parametrima n i p (oznaka $\boxed{X \sim B(n,p)}$).

Uočimo, ako je $A_i := \{\text{u } i\text{-tom bacanju palo } P\}$, $i = 1, \dots, n$,

tada je

$$\mathbb{1}_{A_i} \sim \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}, \quad \forall i = 1, \dots, n.$$

te

$$\boxed{X = \sum_{i=1}^{n} \mathbb{1}_{A_i} \sim B(n,p)}.$$

Specijalno, $B(1,p)$ je upravo Bernoullijeva razdioba s parametrom p.

<u>Primjer 4.11</u> Sve kao u Pr. 4.10, s tim da je $p \in \langle 0, 1]$.

Neka je $[T = +\infty ?]$

$$T := \text{broj bacanja do pojave prvog } P.$$

Očito je $T(\omega) \in \{1, 2, \dots\} = \mathbb{N}$, $\forall \omega \in \Omega$, te

$$\boxed{\mathbb{P}(T = k) = q^{k-1} p, \quad k \in \mathbb{N}} \qquad (4.7)$$

Sluč. varijabla T s distribucijom (4.7) naziva se <u>geometrijska</u> slučajna varijabla <u>na $\mathbb{N}$</u> s parametrom p (oznaka $\boxed{T \sim G(p)}$).

Ako je

$$\tilde{T} := \text{broj } G \text{ prije prve pojave } P,$$

imamo $\tilde{T} = T - 1$, pa je $\tilde{T} \in \{0,1,2,\dots\} = \mathbb{N}_0$ te

$$\boxed{\mathbb{P}(\tilde{T} = k) = q^{k} \cdot p \,, \; k \in \mathbb{N}_0 .} \qquad (4.8)$$

→ geometrijska slučajna varijabla na $\mathbb{N}_0$,

oznaka $\tilde{T} \sim G_0(p)$

Nap. 4.12 U Pr. 4.10 i Pr. 4.11 nije bilo bitno što bacamo novčić. Bitno je da imamo niz nezavisnih "pokusa" pri čemu je vjerojatnost "uspjeha" u svakom pokusu ista i jednaka $p \in [0,1]$. Tada vrijedi

$X :=$ broj uspjeha u n pokusa $\sim B(n,p)$

$T :=$ broj pokusa do prvog uspjeha $\sim G(p)$ $[p>0]$

$\tilde{T} :=$ broj neuspjeha do prvog uspjeha $\sim G_0(p)$ $[p>0]$

Nap. 4.13 Za diskretnu slučajnu varijablu X s distr.

$$X \sim \begin{pmatrix} a_1 & a_2 & \dots \\ p_1 & p_2 & \dots \end{pmatrix}$$

često se koristi njena diskretna (vjerojatnosna) f-ja gustoće (ili f-ja mase, engl. probability mass function). To je f-ja $f : \mathbb{R} \to [0,1]$ def. s

$$f(x) := \mathbb{P}(X = x) \,, \; x \in \mathbb{R} . \qquad (4.9)$$

Uočimo,

$$f(x) = \begin{cases} p_i, & \text{ako } x = a_i \text{ za neki } i \in I \\ 0, & \text{inače} \end{cases},$$

te f jedinstveno određuje distribuciju sluč. varijable X.

Mi nećemo koristiti f da izbjegnemo izraze poput

$$\sum_{x \in \mathbb{R}} f(x)$$ ← [koristi se u Sarapa, Vondraček]

Funkcija slučajne varijable

Neka je X diskretna slučajna var. t.d. $X \in D = \{a_i : i \in I\}$,

te $\boxed{g : D \to \mathbb{R}}$ neka f-ja.

Definiramo $\boxed{Y : \Omega \to \mathbb{R}}$ t.s. $\boxed{Y := g(X)}$, tj.

$$Y(\omega) := g(X(\omega)), \ \omega \in \Omega. \quad (Y = g \circ X)$$

$\Rightarrow$ $Y \in g(D) = \{g(a_i) : i \in I\} =: D'$, te je

D' najviše prebrojiv.

$\Rightarrow$ Y je diskretna sluč. varijabla.

$J = \{1, 2, \ldots, n\}$ ili $J = \mathbb{N}$

Prop. 4.14 Ako je $D' = \{b_j : j \in J\}$, distribucija sluč. var. Y dana je s

$$\mathbb{P}(Y = b_j) = \sum_{\substack{i \in I: \\ \boxed{g(a_i) = b_j}}} \mathbb{P}(X = a_i), \ j \in J. \quad (4.10)$$

[Izgleda kompliciranije nego što jest!]

Dokaz: $\forall j \in J$,

$$\mathbb{P}(Y = b_j) = \mathbb{P}(g(X) = b_j) = \mathbb{P}(\{\omega \in \Omega : g(X(\omega)) = b_j\})$$

$$= \mathbb{P}(\{\omega\in\Omega : X(\omega) \in \underbrace{g^{-1}(\{b_j\})}_{\in \mathcal{D}}\}) = \mathbb{P}(X \in g^{-1}(\{b_j\}))$$

$$\overset{(4.5)}{=} \sum_{\substack{i\in I:\\ a_i \in g^{-1}(\{b_j\})}} \mathbb{P}(X=a_i) = \sum_{\substack{i\in I:\\ \boxed{g(a_i)=b_j}}} \mathbb{P}(X=a_i) \qquad \blacksquare$$

$\boxed{\text{Pr. } 4.15}$ Neka je

$$X \sim \begin{pmatrix} -1 & 0 & 1 & 2 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & \frac{1}{8} \end{pmatrix}$$

te $g(x) = x^2$. Tada $Y := g(X)$ poprima vrijednosti u

$$g(\{-1,0,1,2\}) = \{0,1,4\}$$

te je

ili odmah (4.10)

$$\mathbb{P}(Y=0) = \mathbb{P}(X^2=0) = \mathbb{P}(X=0) = \frac{1}{2}$$

$$\mathbb{P}(Y=1) = \mathbb{P}(X^2=1) = \mathbb{P}(X\in\{-1,1\})$$

$$\overset{(4.10)}{=} \mathbb{P}(X=-1) + \mathbb{P}(X=1) = \frac{1}{4}+\frac{1}{8} = \frac{3}{8}$$

$$\mathbb{P}(Y=4) = \mathbb{P}(X^2=4) = \mathbb{P}(X\in\{-2,2\}) \underset{X\neq -2}{=} \mathbb{P}(X=2) = \frac{1}{8}.$$

(4.10)

tj.

$$Y \sim \begin{pmatrix} 0 & 1 & 4 \\ \frac{1}{2} & \underbrace{\frac{1}{4}+\frac{1}{8}}_{=\frac{3}{8}} & \frac{1}{8} \end{pmatrix} \qquad \blacksquare$$