Vjerojatnost

Bilješke s predavanja

Hrvoje Planinić

Prirodoslovno matematički fakultet – Matematički odsjek
Sveučilište u Zagrebu

E-mail adresa: planinic@math.hr

Sadržaj

Predgovor

Ove bilješke su napisane kao nadopuna (a ne zamjena) predavanjima – tako da neke stvari mogu ostaviti studentima da sami pogledaju te da studenti mogu pogledati nešto što na predavanjima nisu stigli napisati (ili nisu uspjeli pročitati). Stvari koje se nalaze u uglatim zagradama su nešto što tipično samo ispričam. *Uglavnom* koristim **podebljani** font za stvari koje definiramo, a *kurziv* za stvari koje želim samo naglasiti.

Posljednja izmjena napravljena je 30.1.2024.

POGLAVLJE 1

Vjerojatnost

1.1. Uvod

Primjeri nekih ”vjerojatnosnih” tvrdnji:

- vjerojatnost da će simetričan novčić pokazati pismo je $\frac{1}{2}$
  $\rightsquigarrow$ **klasična vjerojatnost** (simetrija)
- vjerojatnost da će nesimetričan novčić pokazati pismo je 0.61
  $\rightsquigarrow$ **asimptotski pristup** (zakon velikih brojeva)
- vjerojatnost da će sutra u Zagrebu biti orkanska oluja je 1%
  $\rightsquigarrow$ **subjekivna vjerojatnost** (statistika)

Cilj ovog kolegija je dati *preciznu* matematičku teoriju vjerojatnosti [koja je

- u skladu s intuicijom u jednostavnim primjerima (npr. u slučaju simetrije),
- daje nam odgovore i kada nemamo dobru intuiciju, te
- je važan element u drugim disciplinama kao npr. u statistici ili optimizaciji, te dokazivanju teorema u analizi, teoriji grafova ...]

Neki primjeri:

- **Lažno pozitivni testovi**
  [Laboratorijski krvni test je 95% učinkovit u otkrivanju jedne rijetke bolesti kada je ta bolest prisutna. Test također ukazuje na bolest i u 0.05% slučajeva kada je osoba zdrava (*false positive*). Poznato je da samo 0.001% populacije ima tu bolest. Test je pokazao da imate bolest, kolika je vjerojatnost da je zaista imate?
  Odgovor je **0.2%**!? Tu se radi o pojmu uvjetne vjerojatnosti i Bayesove formule.]
- **Problem rođendana**
  [U grupi od 23 slučajno odabranih ljudi, kolika je vjerojatnost da su barem dvoje rođeni na isti dan u godini? $\rightsquigarrow$ **50.7%**. U grupi od 70 ljudi? $\rightsquigarrow$ **99.9%**!]
- **Zakon arkus sinusa i duga vodstva**
  [Dva igrača bacaju novčić za koji ne znamo je li simetričan ili ne. Kada padne pismo, prvi igrač daje drugom jednu kunu, a kada padne glava, obrnuto. Nakon 20 bacanja, prvi igrač je ukupno 16 trenutaka bio u vodstvu. Je li to dovoljno dokaza da bi zaključili da je novčić nesimetričan?
  Ne nužno! Naime, ako je novčić simetričan, vjerojatnost da će jedan od igrača voditi barem 16 trenutaka je otprilike **68.5%**!]

1.2. Vjerojatnosni prostor $(\Omega, \mathcal{F}, \mathbb{P})$

Neformalno, *(slučajni) pokus* je svaka procedura ("mehanizam") čiji ishod nije jednoznačno određen. Na primjer, (i) bacanje novčića, (ii) cijene dionica na burzi sutra ...

Vjerojatnosni prostor predstavlja matematički model za dani pokus, a sastoji se od tri komponente:

(i) skupa svih mogućih ishoda pokusa (oznaka obično Ω)
$\rightsquigarrow$ elemente od Ω (tj. ishode pokusa) obično označavamo s ω te ih još zovemo **elementarni događaji**, a Ω **prostor elementarnih događaja**.

(ii) familije podskupova od Ω čije elemente zovemo **događaji** (oznaka obično $\mathcal{F}$, pri čemu je dakle $\mathcal{F} \subseteq \mathcal{P}(\Omega)$[1])

(iii) **vjerojatnosti** koju pridružujemo svakom događaju (za $A \in \mathcal{F}$, oznaka obično $\mathbb{P}(A)$).

PRIMJER 1.1. Bacanje kocke.

(i) za Ω je prirodno uzeti

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

(ii) primjeri događaja su[2]

- $A = \{\text{pao je paran broj}\} = \{2, 4, 6\} \subseteq \Omega$
- $B = \{\text{pao je broj } 6\} = \{6\} \subseteq \Omega$

(iii) vjerojatnosti su

- $\mathbb{P}(A) = [\text{simetrija}] = \frac{3}{6} = \frac{1}{2}$
- $\mathbb{P}(B) = [\text{simetrija}] = \frac{1}{6}$

1.2.1. Događaji. Kažemo da se događaj $A \in \mathcal{F}$ **dogodio** ako se dogodio ishod $\omega \in \Omega$ takav da je $\boldsymbol{\omega \in A}$. Također, za $A, B \in \mathcal{F}$:

- $A^c = \Omega \setminus A$ predstavlja događaj "A se nije dogodio";
- $A \cup B$ predstavlja događaj "dogodio se A ili B";
- $A \cap B$ predstavlja događaj "dogodili su se i A i B";
- $A \setminus B$ predstavlja događaj "dogodio se A, ali ne i B";
- $A \subseteq B$ (tj. $\omega \in A$ povlači $\omega \in B$) interpretiramo kao "ako se dogodio A, dogodio se i B";
- $A \cap B = \emptyset$ [tj. A i B su disjunktni] interpetiramo kao "ne mogu se istovremeno dogoditi i A i B";
- Ω nazivamo "siguran događaj", a $\emptyset$ "nemoguć događaj".

[Bitno je naviknuti se na ovu vezu između događaj i skupova/skupovnih operacija.]

NAPOMENA. Iz *tehničkih* razloga, neće uvijek moći vrijediti $\mathcal{F} = \mathcal{P}(\Omega)$ [tj. da su *svi* $A \subseteq \Omega$ događaji]. Ipak, uvijek ćemo tražiti da $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ zadovoljava svojstva tzv. **σ-algebre**. $\square$

[1] $\mathcal{P}(\Omega)$ je oznaka za partitivni skup od Ω, tj. skup svih podskupova skupa Ω.
[2] Događaje često opisujemo riječima.

DEFINICIJA 1.2. Neka je Ω neprazan skup. Familija podskupova $\mathcal{F}$ od Ω zove se **σ-algebra** (događaja), ako vrijede sljedeća tri svojstva:

(i) $\Omega \in \mathcal{F}$;

(ii) Ako je $A \in \mathcal{F}$, onda je i $A^c \in \mathcal{F}$ [zatvorenost na uzimanje komplementa];

(iii) Ako su $A_j \in \mathcal{F}$, $j \in \mathbb{N}$, onda je i $\bigcup_{j=1}^{\infty} A_j \in \mathcal{F}$ [zatvorenost na *prebrojive* unije].

Uređen par $(\Omega, \mathcal{F})$ zove se **izmjeriv prostor**.

NAPOMENA 1.3. Ako je $\mathcal{F}$ σ-algebra na Ω, vrijedi i

(a) $\emptyset \in \mathcal{F}$;

(b) Za svaki $n \in \mathbb{N}$, ako su $A_1, A_2, \ldots, A_n \in \mathcal{F}$, onda je i $\bigcup_{j=1}^{n} A_j \in \mathcal{F}$ [zatvorenost na *konačne* unije].

(c) Ako su $A_j \in \mathcal{F}$, $j \in \mathbb{N}$, onda je i $\bigcap_{j=1}^{\infty} A_j \in \mathcal{F}$ [zatvorenost na prebrojive *presjeke*].

DOKAZ. (a) $\emptyset = \Omega^c$ pa tvrdnja slijedi iz (i) i (ii);

(b) ako stavimo $A_j := \emptyset$ za sve $j \geqslant n+1$, imamo $\bigcup_{j=1}^{n} A_j = \bigcup_{j=1}^{\infty} A_j$ pa tvrdnja slijedi iz (iii) i (a);

(c) $\bigcap_{j=1}^{\infty} A_j = \left(\bigcup_{j=1}^{\infty} A_j^c\right)^c$ pa tvrdnja slijedi iz (ii) i (iii).

$\square$

ZADATAK. (a) Ako je $A, B \in \mathcal{F}$, pokažite da je i $A \setminus B, A \triangle B \in \mathcal{F}$. (b) Za svaki $n \in \mathbb{N}$, ako su $A_1, A_2, \ldots, A_n \in \mathcal{F}$, onda je i $\bigcap_{j=1}^{n} A_j \in \mathcal{F}$.

Poanta: $\mathcal{F}$ je zatvorena na (gotovo) sve konačne ili prebrojive operacije nad događajima, tj. to će opet biti događaji!

PRIMJER 1.4. Ako je $\Omega \neq \emptyset$,

(a) $\mathcal{P}(\Omega)$ je najveća, a $\{\emptyset, \Omega\}$ najmanja σ-algebra na Ω.

(b) Za $A \subseteq \Omega$, $\{\emptyset, \Omega, A\}$ *nije*, dok $\{\emptyset, \Omega, A, A^c\}$ *jest* σ-algebra na Ω [i to najmanja σ-algebra koja sadrži A].

Ako je Ω konačan ili prebrojiv skup (tzv. *diskretan* slučaj) najčešće uzimamo $\mathcal{F} = \mathcal{P}(\Omega)$.

[NAPOMENA. σ-algebre zapravo igraju važnu ulogu u općenitoj teoriji vjerojatnosti. Za dani Ω, razne σ-algebre na Ω predstavljaju *dostupne informacije*. Na primjer, kod bacanja kocke imamo $\Omega = \{1, 2, \ldots, 6\}$ te

(i) $\mathcal{F} = \mathcal{P}(\Omega)$ predstavlja slučaj kada imamo sve informacije, dok

(ii) $\mathcal{F}' = \{\emptyset, \Omega, \{1, 3, 5\}, \{2, 4, 6\}\}$ predstavlja slučaj kada samo znamo je li pao paran ili neparan broj.

Ipak, u ovom kolegiju σ-algebre *neće* igrati važnu ulogu.]

1.2.2. Definicija vjerojatnosti i osnovna svojstva.

Definicija 1.5 (Kolmogorov, 1933.). Neka je $(\Omega, \mathcal{F})$ izmjeriv prostor. **Vjerojatnost** na $(\Omega, \mathcal{F})$ je svaka funkcija $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ koja zadovoljava sljedeća tri aksioma:

(A1) (*nenegativnost)* Za sve $A \in \mathcal{F}$, $\mathbb{P}(A) \geqslant 0$;

(A2) (*normiranost*) $\mathbb{P}(\Omega) = 1$;

(A3) (σ-*aditivnost*) ako su $A_1, A_2, \dots \in \mathcal{F}$[3] u parovima disjunktni (tj. $A_i \cap A_j = \emptyset$ za $i \neq j$), vrijedi

$$\mathbb{P}\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mathbb{P}(A_j)\,.$$

Uređena trojka $(\Omega, \mathcal{F}, \mathbb{P})$ zove se **vjerojatnosni prostor**.

[Napomena. "Aksiomatski pristup": Izgradit ćemo cijelu teoriju koju onda možemo primijeniti kadgod imamo model $(\Omega, \mathcal{F}, \mathbb{P})$ gdje je $\mathcal{F}$ σ-algebra, a $\mathbb{P}$ zadovoljava samo aksiome (A1)-(A3). Primjerice, pokazat ćemo da uvijek vrijedi tzv. *zakon velikih brojeva* te *centralni granični teorem.*]

Napomena 1.6. Vidjet ćemo da iz (A1)-(A3) slijedi i

(i) $\mathbb{P}(\emptyset) = 0$ te $\mathbb{P}(A) \in [0, 1]$ za sve $A \in \mathcal{F}$ [što bismo intuitivno i htjeli];

(ii) $\mathbb{P}$ je **konačno aditivna**: za svaki $n \in \mathbb{N}$ i svaki konačan niz $A_1, \dots, A_n$ po parovima disjunktnih događaja iz $\mathcal{F}$ vrijedi

$$\mathbb{P}\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} \mathbb{P}(A_j)\,. \tag{1.1}$$

Zadatak. Pokažite da za proizvoljnu funkciju $\mathbb{P} : \mathcal{F} \to \mathbb{R}$,

(a) vrijedi (1.1) akko za sve disjunktne $A, B \in \mathcal{F}$ vrijedi

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) \tag{1.2}$$

(b) ako je $\mathbb{P}(\emptyset) = 0$ i $|\mathcal{F}| < \infty$,[4] (A3) vrijedi čim pokažemo da vrijedi (1.2).

Primjer 1.7 (**Laplaceov model**). Ako je $|\Omega| < \infty$ te su zbog simetrije svi $\omega \in \Omega$ "jednako vjerojatni", prirodno je uzeti

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|}\,,\ A \in \mathcal{F}\,, \tag{1.3}$$

uz $\mathcal{F} = \mathcal{P}(\Omega)$. Lako se vidi da $\mathbb{P}$ zaista zadovoljava (A1)-(A3), detalje ostavljamo za DZ.[5]

Primjer 1.8 (**Geometrijska vjerojatnost**). Na "slučajan" način biramo točku iz kvadrata $[-1, 1]^2$. Dakle, $\Omega = [-1, 1]^2$ pri čemu $(x, y) \in \Omega$ predstavlja ishod kada smo izabrali točku

[3] Ima ih *prebrojivo* mnogo.
[4] $|\mathcal{F}|$ označava broj elemenata skupa $\mathcal{F}$.
[5] Uočite da (1.2) slijedi jer za disjunktne $A, B \in \mathcal{F}$ vrijedi $|A \cup B| = |A| + |B|$.

(x, y), a za $\mathbb{P}$ je prirodno uzeti[6]

$$\mathbb{P}(A) = \frac{\lambda(A)}{\lambda(\Omega)}\,,\ A \in \mathcal{F}\,. \tag{1.4}$$

Na primjer,

- za $A = \{$točka je upala u krug upisan $[-1,1]^2\} = \{(x,y) \in \Omega : \sqrt{x^2+y^2} \leqslant 1\}$ imamo
$$\mathbb{P}(A) = \frac{1^2\pi}{2^2} = \frac{\pi}{4}\,.$$
- za $B = \{$izabrali smo točku $(0,0)\} = \{(0,0)\}$ imamo[7]
$$\mathbb{P}(B) = \frac{\lambda(\{(0,0)\})}{2^2} = \frac{0}{4} = 0\,.$$

Što je $\mathcal{F}$?* Može se pokazati da *ne postoji* funkcija $\mathbb{P} : \mathcal{P}(\Omega) \to \mathbb{R}$ koja zadovoljava (A1)-(A3), a da zadovoljava (1.4) za sve pravokutnike $A = [a,b] \times [c,d] \subseteq \Omega$. Ipak, to se može postići na najmanjoj σ-algebri koja sadrži sve takve pravokutnike, a nazivamo je **Borelova σ-algebra** na $[-1,1]^2$.[8]

PROPOZICIJA 1.9 (Osnovna svojstva vjerojatnosti). *Iz (A1)-(A3) slijedi i*

(a) $\mathbb{P}(\emptyset) = 0$*;*

(b) $\mathbb{P}$ *je konačno aditivna;*

(c) $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$*;*

(d) $A \subseteq B$ *povlači da je* $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$*, te da je* $\mathbb{P}(A) \leqslant \mathbb{P}(B)$ *(monotonost od* $\mathbb{P}$*)*[9]*;*

(e) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$*;*

*(f) Za **sve*** $A, B \in \mathcal{F}$ *vrijedi*

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)\,.\ \textit{(formula uključivanja-isključivanja)}$$

DOKAZ. (a) Stavimo $p := \mathbb{P}(\emptyset) \geqslant 0$ (zbog (A1)) te $A_j := \emptyset$, $j \in \mathbb{N}$. Tada je $(A_j)_{j\in\mathbb{N}}$ niz po parovima disjunktnih događaja pa iz (A3) slijedi
$$p = \mathbb{P}(\emptyset) = \mathbb{P}\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mathbb{P}(A_j) = \sum_{j=1}^{\infty} \mathbb{P}(\emptyset) = \sum_{j=1}^{\infty} p\,.$$
Odavde očito slijedi da je $p = 0$.

(b) DZ.

(c) Intuitivno, tvrdnju lako provjerimo crtajući **Vennov dijagram** i razmišljajući kao da je $\mathbb{P}(A) = \frac{\lambda(A)}{\lambda(\Omega)}$. Formalno, tvrdnja slijedi jer iz $B = (B \setminus A) \cup (A \cap B)$, disjunktnosti zadnja dva skupa, te konačne aditivnosti slijedi $\mathbb{P}(B) = \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B)$.

(d) Slijedi iz (c) jer je $A \cap B = A$ te je $\mathbb{P}(B \setminus A) \geqslant 0$ (zbog (A1)).

---

[6] $\lambda(A)$ je oznaka za "površinu" skupa A.

[7] Zapravo je $\mathbb{P}(\{(x,y)\}) = 0$ za sve $(x,y) \in \Omega$!

[8] Napomenimo da postoje podskupovi od $[-1,1]^2$ koji se ne nalaze u Borelovoj σ-algebri, ali da ona zapravo sadrži sve skupove koje nas tipično zanimaju, te da za $\mathbb{P}$ u (1.4) (A3) u principu slijedi jer za disjunktne pravokutnike $A, B \subseteq \Omega$ vrijedi $\lambda(A \cup B) = \lambda(A) + \lambda(B)$.

[9] Specijalno, za sve $A \in \mathcal{F}$ vrijedi $\mathbb{P}(A) \leqslant \mathbb{P}(\Omega) = 1$, tj. $\mathbb{P}(A) \in [0,1]$.

(e) i (f) DZ.

$\square$

1.2.3. Diskretan vjerojatnosni prostor. Neka je Ω konačan ili prebrojiv, $\mathcal{F} = \mathcal{P}(\Omega)$ te $\mathbb{P}$ vjerojatnost na $(\Omega, \mathcal{F})$. Tada za svaki $A \in \mathcal{F}$ imamo[10]

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{\omega \in A} \{\omega\}\right) = [\text{konačna}/\sigma\text{-aditivnost}] = \sum_{\omega \in A} \mathbb{P}(\{\omega\}) . \tag{1.5}$$

Dakle, svaka vjerojatnost $\mathbb{P}$ na $(\Omega, \mathcal{F})$ je u potpunosti određena s vrijednostima $\mathbb{P}(\omega), \omega \in \Omega$.[11]

PRIMJER 1.10. Ako je $|\Omega| < \infty$, tj. $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ za $n \in \mathbb{N}$, te vrijedi $\mathbb{P}(\omega_i) = \mathbb{P}(\omega_j)$ za sve i, j, imamo

$$1 = \mathbb{P}(\Omega) = [(1.5)] = \sum_{i=1}^{n} \mathbb{P}(\omega_i) = n\mathbb{P}(\omega_1) .$$

Dakle, vrijedi $\mathbb{P}(\omega_1) = \frac{1}{n} = \mathbb{P}(\omega_i)$, za sve $i = 1, \dots, n$, te

$$\mathbb{P}(A) = [(1.5)] = \frac{|A|}{|\Omega|} , \ A \in \mathcal{F} ,$$

tj. imamo Laplaceov model. $\square$

PRIMJER 1.11. Ako je $\Omega = \{\omega_1, \omega_2, \dots\}$ (dakle, Ω prebrojivo beskonačan), iz $\sum_{i=1}^{\infty} \mathbb{P}(\omega_i) = 1$ i nenegativnosti vjerojatnosti, slijedi da *ne postoji* vjerojatnost $\mathbb{P}$ na $(\Omega, \mathcal{F})$ takva da vrijedi $\mathbb{P}(\omega_i) = \mathbb{P}(\omega_j)$ za sve i, j.

Zbog toga nije odmah jasno što znače tvrdnje poput

$$\mathbb{P}(\{\text{slučajno odabrani broj u } \mathbb{N} \text{ je paran}\}) = \frac{1}{2} .$$

Formalno opravdanje može biti sljedeće. Za $n \in \mathbb{N}$ fiksan, slučajno odaberemo broj iz $\Omega_n := \{1, \dots, n\}$. Ovdje je dobro definirana vjerojatnost

$$\mathbb{P}_n(A) = \frac{|A|}{|\Omega_n|} , \ A \subseteq \Omega_n ,$$

te vrijedi

$$p_n := \mathbb{P}_n(\{\text{slučajno odabrani broj u } \Omega_n \text{ je paran}\}) = \begin{cases} \frac{1}{2} , & \text{ako je } n \text{ paran,} \\ \frac{(n-1)/2}{n} , & \text{ako je } n \text{ neparan.} \end{cases}$$

Sada možemo staviti

$$\mathbb{P}(\{\text{slučajno odabrani broj u } \mathbb{N} \text{ je paran}\}) := \lim_{n \to \infty} p_n = \frac{1}{2} .$$

$\square$

[Sljedeći rezultat je zapravo obrat tvrdnje proizašle iz (1.5).]

[10] U nastavku ćemo koristiti (nepreciznu) oznaku $\mathbb{P}(\omega) := \mathbb{P}(\{\omega\})$.
[11] Usporedite s $\mathbb{P}$ iz Primjera 1.8.

PROPOZICIJA 1.12. *Neka je* $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ *za* $n \in \mathbb{N}$ *te* $\mathcal{F} = \mathcal{P}(\Omega)$. *Za svaki vektor* $(p_1, \dots, p_n) \in \mathbb{R}^n$ *koji zadovoljava*[12]

$$p_i \geqslant 0\,, \forall i = 1, \dots, n\,, \quad te \quad \sum_{i=1}^{n} p_i = 1\,, \tag{1.6}$$

postoji vjerojatnost $\mathbb{P}$ *na* $(\Omega, \mathcal{F})$ *takva da vrijedi* $\mathbb{P}(\omega_i) = p_i$, *za sve* $i = 1, \dots, n$.

DOKAZ. Iz (1.5) slijedi da je *jedini* kandidat za $\mathbb{P}$ funkcija $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ definirana sa

$$\mathbb{P}(A) := \sum_{l=1}^{k} p_{i_l}\,, \text{ za } A = \{w_{i_1}, \dots, w_{i_k}\} \in \mathcal{F}\,, \tag{1.7}$$

uz $\mathbb{P}(\emptyset) := 0$. Sada uvjeti (1.6) osiguravaju da $\mathbb{P}$ iz (1.7) zadovoljava uvjete (A1)-(A3). Detalje ostavljamo za vježbu. □

NAPOMENA 1.13. Tvrdnja prethodne propozicije vrijedi i kada je $\Omega = \{\omega_1, \omega_2, \dots\}$ prebrojivo beskonačan [pri čemu sada imamo niz $(p_1, p_2, \dots)$ nenegativnih brojeva takvih da je $\sum_{i=1}^{\infty} p_i = 1$)]. Ipak, nećemo ulaziti u detalje. □

1.3. Laplaceov model – primjeri

Ako je Ω konačan te su svi elementarni ishodi zbog simetrije jednako vjerojatni, imamo dakle

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|}\,,$$

pa pri računanju gornjih vjerojatnosti često koristimo tehnike **prebrojavanja**.

PRIMJER 1.14. Kutija sadrži n bijelih i n crnih kuglica. Na slučajan način izvučemo iz kutije dvije kuglice, jednu za drugom (bez vraćanja).

(a) Odredite prostor elementarnih događaja.
(b) Odredite vjerojatnost da su izvučene kuglice različite boje.
(c) Odredite vjerojatnost da je 1. kuglica bijela.
(d) Odredite vjerojatnost da je 2. kuglica bijela.

RJEŠENJE. (a) Možemo kuglice u kutiji "numerirati" brojevima od 1 do $2n$ (pri čemu su npr. crne kuglice označene brojevima $1, 2, \dots, n$). Izbor dvije kuglice iz kutije tada odgovara izboru uređenog para različitih brojeva, što vodi na sljedeći prostor elementarnih događaja:

$$\Omega = \{(i,j) : 1 \leqslant i, j \leqslant 2n, i \neq j\}\,.$$

Ovdje elementarni događaj (i,j) interpretiramo kao: prva izvučan kuglica ima broj i, a druga broj j. Iz simetrije slijedi da su svi elementarni događaji jednako vjerojatni, odnosno nalazimo se u Laplaceovom modelu, pri čemu je

$$|\Omega| = 2n(2n-1)\,.$$

[12] Svaki takav vektor $(p_1, \dots, p_n)$ naziva se distribucija ili razdioba na Ω.

(b) Zanima nas vjerojatnost događaja $A = \{\text{izvučene kuglice su različite boje}\}$. Stavimo

$$A_1 = \{\text{1. izvučena kuglica je crna, 2. izvučena kuglica je bijela}\}$$
$$A_2 = \{\text{1. izvučena kuglica je bijela, 2. izvučena kuglica je crna}\}\,.$$

Budući da su A_1 i A_2 disjunktni te $A = A_1 \cup A_2$, vrijedi $\mathbb{P}(A) = \mathbb{P}(A_1) + \mathbb{P}(A_2)$. Očito vrijedi da je $|A_1| = n \cdot n = n^2$ (broj uređenih parova (i, j) kod kojih je na prvoj koordinati crna, a na drugoj bijela kuglica je n^2). Slično, $|A_2| = n^2$. Slijedi da je[13]

$$\mathbb{P}(A) = \mathbb{P}(A_1) + \mathbb{P}(A_2) = \frac{n^2}{2n(2n-1)} + \frac{n^2}{2n(2n-1)} = \frac{n}{2n-1}\,.$$

Napomena.

(i) Za prostor elementarnih događaja mogli smo uzeti $\Omega = \{BB, BC, CB, CC\}$ [gdje, npr. BC označava da je prva izvučena kuglica bijela, a druga crna], ali tu imamo

$$\mathbb{P}(BC) = [(b)] = \frac{n^2}{2n(2n-1)} > \frac{n(n-1)}{2n(2n-1)} = [\text{DZ}] = \mathbb{P}(BB)\,,$$

to jest, ne nalazimo se u Laplaceovom modelu.

(ii) S druge strane, mogli smo pretpostaviti da su obje kuglice izvučene *istovremeno* budući da se tražena vjerojatnost očito *neće promijeniti*. Umjesto uređenog para, elementarni događaj je dvočlani *podskup* skupa $\{1, 2, \ldots, 2n\}$, tj. imamo

$$\Omega = \{\{i, j\} : 1 \leqslant i, j \leqslant 2n, i \neq j\}\,.$$

Opet su svi $\omega \in \Omega$ jednako vjerojatni, te sada imamo

$$|\Omega| = \binom{2n}{2} = n(2n-1)\,.$$

Sada je $A = \{\{i, j\} : i \in \{1, \ldots, n\}, j \in \{n+1, \ldots, 2n\}\}$. Dakle, vrijedi $|A| = \binom{n}{1}\binom{n}{1} = n^2$ te je

$$\mathbb{P}(A) = \frac{n^2}{n(2n-1)} = \frac{n}{2n-1}\,,$$

kao i prije.

Poanta je da izbor prostora elementarnih događaja *nije jedinstven*! U pravilu biramo onaj Ω uz koji lakše možemo riješiti dani problem. □

(c) Uočimo da događaj $B_1 := \{\text{1. izvučena kuglica je bijela}\}$ ovisi *samo* o prvom izvlačenju. Zato je dovoljno gledati

$$\Omega_1 := \{\text{svi ishodi 1. izvlačenja}\} = \{1, 2, \ldots, 2n\}\,.$$

Budući da smo očito u Laplaceovom modelu, imamo $\mathbb{P}(B_1) = \frac{n}{2n} = \frac{1}{2}$.

[13] Uočimo da je $\frac{n}{2n-1} > \frac{1}{2}$ za sve $n \in \mathbb{N}$, ali da $\frac{n}{2n-1} \to \frac{1}{2}$ kada $n \to \infty$.

(d) [Intuitivno?] Neka je $B_2 := \{2.\text{ izvučena kuglica je bijela}\}$. Ako gledamo samo ishod drugog izvlačenja, očito svaka od $2n$ kuglica ima jednaku vjerojatnost da bude izvučena. Dakle, ako gledamo

$$\Omega_2 := \{\text{svi ishodi 2. izvlačenja}\} = \{1, 2, \ldots, 2n\},$$

opet smo u Laplacevom modelu, te je opet $\mathbb{P}(B_2) = \frac{n}{2n} = \frac{1}{2}$.
Alternativno, ako Ω uzmemo kao u (a) dijelu, B_2 možemo rastaviti ovisno o tome je li prva izvučena kuglica bila bijela ili crna te zaključiti da je

$$\mathbb{P}(B_2) = \mathbb{P}(BB) + \mathbb{P}(BC) = \frac{n(n-1)}{2n(2n-1)} + \frac{n^2}{2n(2n-1)} = \frac{1}{2}.$$

□

Primjer 1.15. Imamo špil od 53 karte (4 boje po 13 jačina).

(a) Ako promiješamo špil, kolika je vjerojatnost da je prva karta jačine as?
(b) Ako podijelimo slučajno 13 karata svakom od četiri igrača, kolika je vjerojatnost da je svaki igrač dobio asa?

Rješenje. (a) Označimo karte brojevima $1, \ldots, 52$ te uzmimo

$$\Omega = \{\text{svi rasporedi špila}\} = \{\text{sve permutacije skupa } \{1, 2, \ldots, 52\}\}$$

Očito su svi elementarni ishodi jednako vjerojatni te je $|\Omega| = 52!$. Traženi događaj je

$$A = \{\text{sve permutacije t.d. je prvi element jedan od 4 asa}\},$$

pa je $|A| = 51! \cdot 4$ [51! je broj permutacija u kojima je na prvom mjestu as *fiksne* boje], te

$$\mathbb{P}(A) = \frac{51! \cdot 4}{52!} = \frac{1}{13}.$$

Alternativno (i bolje), uzmimo

$$\Omega = \{\text{sve moguće jačine prve karte}\} = \{\text{as, dvojka,} \ldots, \text{ kralj}\}.$$

Budući da svaka od 13 jačina ima *točno* 4 boje, ovo je opet Laplaceov model. Sada je $A = \{\text{as}\}$ te $\mathbb{P}(A) = \frac{1}{13}$.

(b) [Što bi ovdje mogli uzeti za Ω?] Uočimo da se tražena vjerojatnost ne mijenja ako prvih 13 karata (promiješanog) špila ide prvom igraču, drugih 13 drugom igraču, itd. Dakle, opet možemo uzeti

$$\Omega = \{\text{svi rasporedi špila}\} = \{\text{sve permutacije skupa } \{1, 2, \ldots, 52\}\}$$

pri čemu je traženi događaj

$$B = \{\text{permutacije t.d. je u svakom od 4 bloka po jedan as}\}.$$

Imamo

$$|B| = 48! \cdot 13^4 \cdot 4!\,, \text{ te } \mathbb{P}(B) = \frac{48! \cdot 13^4 \cdot 4!}{52!}\,.$$

[48! je broj permutacija kada *fiksiramo* 4 asa na po jednu *fiksnu* poziciju unutar svakog bloka, 13^4 je broj različitih kombinacija pozicija za aseve unutar svakog bloka *nakon* što smo fiksirali koji as ide u koji blok, 4! je broj rasporeda 4 asa na 4 bloka.]
Alternativno, uočimo da događaj B ovisi samo o pozicijama aseva, pri čemu nije bitno koji se točno as nalazi na kojoj poziciji. Dakle, možemo uzeti

$$\begin{aligned}\Omega &= \{\text{sve kombinacije pozicija za 4 asa}\}\\ &= \{\text{svi 4-člani podksupovi od } \{1{,}2{,}\dots,52\}\}\,.\end{aligned}$$

Zbog simetrije opet smo u Laplacevom modelu te vrijedi $|\Omega| = \binom{52}{4}$. Sada je $B = \{\text{po jedna pozicija za as u svakom od 4 bloka}\}$ pa imamo

$$|B| = \binom{13}{1}^4, \text{ te } \mathbb{P}(B) = \frac{13^4}{\binom{52}{4}}\,,$$

što je naravno isto kao i u prvom rješenju.

□

[**Poanta** je da vjerojatnost nije isto što i prebrojavanje jer nas zanima *omjer* $|A|/|\Omega|$, a njega često možemo jednostavnije odrediti ako iskoristimo dodatnu simetriju problema.]

1.3.1. Formula uljučivanja-isključivanja.

Propozicija 1.16. *Neka je $n \in \mathbb{N}$ te $A_1, A_2, \dots, A_n$ događaji u $\mathcal{F}$. Tada vrijedi*[14]

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mathbb{P}(A_i) - \sum_{1\leqslant i_1<i_2\leqslant n} \mathbb{P}(A_{i_1}\cap A_{i_2}) + \sum_{1\leqslant i_1<i_2<i_3\leqslant n} \mathbb{P}(A_{i_1}\cap A_{i_2}\cap A_{i_3}) - \cdots + (-1)^{n+1}\mathbb{P}(A_1\cap A_2\cap\cdots\cap A_n)\,. \quad \text{(FUI)}$$

Dokaz. Indukcijom po n. Za $n = 1$ tvrdnja očito vrijedi, a za $n = 2$ to je točno Propozicija 1.9(f). Pretpostavimo da tvrdnja vrijedi za n. Za $n + 1$ imamo

$$\begin{aligned}\mathbb{P}\left(\bigcup_{i=1}^{n+1} A_i\right) &= \mathbb{P}\left(\left(\bigcup_{i=1}^{n} A_i\right)\bigcup A_{n+1}\right)\\ &\overset{\text{Prop.1.9}(f)}{=} \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) + \mathbb{P}(A_{n+1}) - \mathbb{P}\left(\left(\bigcup_{i=1}^{n} A_i\right)\bigcap A_{n+1}\right)\\ &= \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) + \mathbb{P}(A_{n+1}) - \mathbb{P}\left(\bigcup_{i=1}^{n}\left(A_i\bigcap A_{n+1}\right)\right)\,.\end{aligned}$$

Tvrdnja sada slijedi kada iskoristimo pretpostavku indukcije za prvi i treći član. Detalje ostavljamo za vježbu. □

[14] Na primjer, $\mathbb{P}(A\cup B\cup C) = \mathbb{P}(A)+\mathbb{P}(B)+\mathbb{P}(C)-\mathbb{P}(A\cap B)-\mathbb{P}(A\cap C)-\mathbb{P}(B\cap C)+\mathbb{P}(A\cap B\cap C)$.

PRIMJER 1.17 (Deranžmani). Kažemo da permutacija π skupa $\{1, 2, \ldots, n\}$[15] ima *fiksnu točku* u i ako je $\pi(i) = i$ [tj. ako je i ostao na istom mjestu]. Odredite vjerojatnost p_n da *slučajno* odabrana permutacija skupa $\{1, \ldots, n\}$ nema fiksnu točku te nađite $\lim_{n\to\infty} p_n$.

[Ekvivalentno pitanje je na primjer: Promiješamo dva ista špila od po n različtih karata. Nakon toga otvaramo jednu po jednu kartu iz svakog špila. Kolika je vjerojatnost da niti jednom nećemo otvoriti dvije iste karte?]

RJEŠENJE. Neka je n fiksan te $\Omega = \{\pi : \pi \text{ je permutacija skupa } \{1, \ldots, n\}\}$; dakle $|\Omega| = n!$. Tada je traženi događaj $A = \{\pi \in \Omega : \pi \text{ nema fiksnu točku}\}$. Ako za $i = 1, \ldots, n$ stavimo $A_i = \{\pi : \ \pi(i) = i\}$, imamo

$$\bigcup_{i=1}^{n} A_i = \{\pi : \pi \text{ ima barem jednu fiksnu točku}\} = A^c\,.$$

Događaji $A_1, \ldots, A_n$ nisu u parovima disjunktni pa moramo koristiti FUI. Za to nam trebaju sljedeće vjerojatnosti[16]

$$\begin{aligned}
\mathbb{P}(A_i) &= \frac{|A_i|}{|\Omega|} = \frac{(n-1)!}{n!} = \frac{1}{n}\,,\ \forall i\\
\mathbb{P}(A_{i_1} \cap A_{i_2}) &= \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}\,,\ \forall i_1 < i_2\\
\mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k}) &= \frac{(n-k)!}{n!} = \frac{1}{n(n-1)\cdots(n-k+1)}\,,\ \forall i_1 < i_2 < \cdots < i_k\\
\mathbb{P}(A_1 \cap \cdots \cap A_n) &= \frac{1}{n!}\,.
\end{aligned}$$

Također, uočimo da je k-članih podksupova $\{i_1, i_2, \ldots, i_k\}$ skupa $\{1, 2 \ldots, n\}$ ima $\binom{n}{k}$. Dakle, iz (FUI) slijedi da je

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) &= n\mathbb{P}(A_1) - \binom{n}{2}\mathbb{P}(A_1 \cap A_2) + \cdots + (-1)^{n+1}\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n)\\
&= \sum_{k=1}^{n} \binom{n}{k}\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_k) \cdot (-1)^{k+1}\\
&= \sum_{k=1}^{n} \frac{n!}{k!(n-k)!}\frac{(n-k)!}{n!} \cdot (-1)^{k+1}\\
&= \sum_{k=1}^{n} (-1)^{k+1}\frac{1}{k!}\,.
\end{aligned}$$

Sada imamo

$$p_n = \mathbb{P}(A^c) = 1 - \mathbb{P}(A) = \sum_{k=0}^{n} (-1)^k \frac{1}{k!}\,,\ \forall n \in \mathbb{N}\,.$$

[15] Dakle, π je bijekcija s $\{1, 2, \ldots, n\}$ u $\{1, 2, \ldots, n\}$.

[16] Zašto je $\mathbb{P}(A_i) = \frac{1}{n}$? Uočimo da A_i ovisi samo o $\pi(i)$, tj. što se nalazi na i-tom mjestu permutacije. Budući da svaki od n brojeva ima jednaku vjerojatnost doći na to mjesto, odmah slijedi da je $\mathbb{P}(A_i) = \frac{1}{n}$. Analogno možemo odrediti i ostale vjerojatnosti.

Konačno,

$$\lim_{n\to\infty} p_n = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} = e^{-1} \approx 0.3679 \,.$$

[Vjerojatnosno, posljednji rezultat vezan je uz tzv. *zakon rijetkih događaja* i *Poissonovu aproksimaciju.*]

□

1.3.2. Primjer kada je Ω neprebrojiv.

PRIMJER 1.18. Igrači A,B,C bacaju kocku tim redoslijedom. Svaki igrač ispada iz igre čim prvi put dobije šesticu. Odredite vjerojatnost događaja

$$A_i = \{\text{A je ispao } i\text{-ti po redu}\}\,,\; i = 1, 2, 3\,.$$

RJEŠENJE. Pretpostavit ćemo da svaki igrač nastavlja bacati kocku i *nakon* što prvi put dobije šesticu, tj. uzeti

$$\Omega = \{(x_1, x_2, x_3, \dots) : x_i \in \{1, \dots, 6\}\,,\, \forall i \in \mathbb{N}\} = \{1, \dots, 6\}^{\mathbb{N}}\,,$$

pri čemu su npr. $x_1, x_4, x_7, \dots$ bacanja prvog igrača.

Napomenimo da je Ω neprebrojiv skup te da je formalna konstrukcija vjerojatnosti na $(\Omega, \mathcal{F})$ netrivijalan problem. [Ipak, u nastavku dajemu osnovnu ideju.]

Odredimo prvo $\mathbb{P}(A_1)$. Ako stavimo $D_k := \{\text{A je ispao nakon točno } k \text{ svojih bacanja}\}$, $k \in \mathbb{N}$, imamo da je

$$D := \bigcup_{k=1}^{\infty} D_k = \{\text{A je u nekom trenutku ispao}\}.$$

Nadalje,

$$\begin{aligned}\mathbb{P}(A_1) &= \mathbb{P}(A_1 \cap D) + \overbrace{\mathbb{P}(A_1 \cap D^c)}^{=0} = \mathbb{P}(A_1 \cap (\cup_{k=1}^{\infty} D_k)) = \mathbb{P}(\cup_{k=1}^{\infty}(A_1 \cap D_k)) \\ &= [A_1 \cap D_1, A_1 \cap D_2, \dots \text{ u parovima disjunktni}] = \sum_{k=1}^{\infty} \mathbb{P}(A_1 \cap D_k)\,.\end{aligned}$$

[U praksi zapravo odmah zaključujemo da je $\mathbb{P}(A_1) = \sum_{k=1}^{\infty} \mathbb{P}(A \cap D_k)$, a ovo gore je formalan dokaz.]

Uočimo da za $k \in \mathbb{N}$, D_k (pa onda i $A_1 \cap D_k$) ovisi samo o prvih $3(k-1)+1 = 3k-2$ bacanja. Zbog toga $A_1 \cap D_k$ možemo promatrati kao događaj u

$$\Omega_{3k-2} = \{\text{svi ishodi prvih } 3k-2 \text{ bacanja}\} = \{1, 2, \dots, 6\}^{3k-2}\,,$$

a budući da su svi $\omega \in \Omega_{3k-2}$ jednako vjerojatni, imamo

$$\mathbb{P}(A_1 \cap D_k) = \frac{|A_1 \cap D_k|}{|\Omega_{3k-2}|}\,.$$

U gornjoj formuli $A_1 \cap D_k$ dakle predstavlja podskup od Ω_{3k-2}. [Ovo je zapravo glavna ideja kako formalno konstruirati $\mathbb{P}$ na $(\Omega, \mathcal{F})$.]

Imamo

$$|A_1 \cap D_k| = |\{(x_1, \ldots, x_{3k-2}) \in \Omega_{3k-2} : x_{3k-2} = 6, x_j \neq 6,\ j = 1, \ldots, 3k-3\}| = 1 \cdot 5^{3k-3}$$

pa je

$$\mathbb{P}(A_1 \cap D_k) = \frac{5^{3k-3}}{6^{3k-2}} = \left(\frac{5}{6}\right)^{3k-3} \frac{1}{6}\,,\ k \in \mathbb{N}\,.$$

[Kasnije ćemo ovo odmah zaključiti iz "nezavisnosti" bacanja.]

Dakle,

$$\begin{aligned}\mathbb{P}(A_1) &= \sum_{k=1}^{\infty} \mathbb{P}(A_1 \cap D_k) = \sum_{k=1}^{\infty} \left(\frac{5}{6}\right)^{3k-3} \frac{1}{6} = \frac{1}{6} \sum_{m=0}^{\infty} \left[\left(\frac{5}{6}\right)^3\right]^m \\ &= \left[\sum_{m=0}^{\infty} x^m = \frac{1}{1-x}\,,\ |x| < 1\right] = \frac{1}{6} \cdot \frac{1}{1 - \left(\frac{5}{6}\right)^3} = \frac{36}{91}\,.\end{aligned}$$

Za vježbu ostavljamo pokazati da je $\mathbb{P}(A_2) = \frac{300}{1001}$ i $\mathbb{P}(A_3) = \frac{305}{1001}$. Dakle, imamo

$$\mathbb{P}(A_1) = \frac{396}{1001} > \mathbb{P}(A_3) > \mathbb{P}(A_2)\,.$$

Zašto je $\mathbb{P}(A_3) > \mathbb{P}(A_2)$?! [Ako znamo da A nije prvi ispao, sigurno nije ispao u svom prvom bacanju pa zapravo postaje "treći" igrač. Kasnije ćemo ovo formalizirati koristeći pojam uvjetne vjerojatnosti.]

[Rješenje za $\mathbb{P}(A_2)$: Neka je $B = \{\text{B ispao prvi}\}$ i $C = \{\text{C ispao prvi}\}$. Budući da je $A_2 \subseteq B \cup C$, te su B i C disjunktni, imamo

$$\mathbb{P}(A_2 \cap D_k) = \mathbb{P}(A_2 \cap D_k \cap B) + \mathbb{P}(A_2 \cap D_k \cap C)\,,\ k \in \mathbb{N}\,.$$

Nadalje, $A_2 \cap D_k \cap B$ ponovno ovisi samo prvih $3k-2$ bacanja te se kao podskup od Ω_{3k-2} sastoji od svih $(x_1, \ldots, x_{3k-2}) \in \Omega_{3k-2}$ takvih da je $x_{3k-2} = 6$, postoji barem jedna 6ica u prvih $k-1$ bacanja igrača B, te niti jedna 6ica u prvih $k-1$ bacanja igrača A i C. Dakle,

$$|A_2 \cap D_k \cap B| = 1 \cdot (6^{k-1} - 5^{k-1}) \cdot 5^{k-1} \cdot 5^{k-1} = (6^{k-1} - 5^{k-1}) \cdot 5^{2k-2}\,.$$

Analogno, $|A_2 \cap D_k \cap C| = (6^{k-1} - 5^{k-1}) \cdot 5^{2k-2}$. Dakle,

$$\mathbb{P}(A_2 \cap D_k) = 2 \cdot \frac{(6^{k-1} - 5^{k-1}) \cdot 5^{2k-2}}{6^{3k-2}} = 2 \cdot \left\{\frac{1}{6}\left(\frac{5}{6}\right)^{2(k-1)} - \frac{1}{6}\left(\frac{5}{6}\right)^{3(k-1)}\right\}\,.$$

Sada lako slijedi da je $\mathbb{P}(A_2) = \sum_{k=1}^{\infty} \mathbb{P}(A_1 \cap D_k) = \frac{300}{1001}$.] $\square$

1.4. Nizovi događaja

1.4.1. Subaditivnost. [Booleova nejednakost]

PROPOZICIJA 1.19. *Za svaki niz događaja $A_1, A_2, \dots \in \mathcal{F}$, vrijedi*

$$\mathbb{P}\left(\bigcup_{j=1}^{\infty} A_j\right) \leqslant \sum_{j=1}^{\infty} \mathbb{P}(A_j)\,. \tag{σ-subaditivnost}$$

Specijalno, za sve $n \in \mathbb{N}$,

$$\mathbb{P}\left(\bigcup_{j=1}^{n} A_j\right) \leqslant \sum_{j=1}^{n} \mathbb{P}(A_j)\,. \tag{konačna subaditivnost}$$

DOKAZ. Definiramo novi niz $(B_j)_{j\in\mathbb{N}}$ po parovima disjunktnih događaja na sljedeći način. Stavimo $B_1 := A_1$ te za $j \geqslant 2$, $B_j := A_j \setminus (A_1 \cup \cdots \cup A_{j-1})$ [nacrtati sliku]. Vrijedi $B_j \in \mathcal{F}$ te (a) $\bigcup_{j=1}^{n} B_j = \bigcup_{j=1}^{n} A_j$ za sve $n \in \mathbb{N}$ i $\bigcup_{j=1}^{\infty} B_j = \bigcup_{j=1}^{\infty} A_j$, (b) $B_j \subseteq A_j$ za sve j, te (c) $(B_j)_{j\in\mathbb{N}}$ su po parovima disjunktni. Iz σ-aditivnosti i monotonosti vjerojatnosti slijedi da je

$$\mathbb{P}\left(\bigcup_{j=1}^{\infty} A_j\right) \overset{(a)}{=} \mathbb{P}\left(\bigcup_{j=1}^{\infty} B_j\right) \overset{(c)}{=} \sum_{j=1}^{\infty} \mathbb{P}(B_j) \overset{(b)}{\leqslant} \sum_{j=1}^{\infty} \mathbb{P}(A_j)\,.$$

Konačnu subaditivnost ostavljamo kao jednostavnu vježbu. □

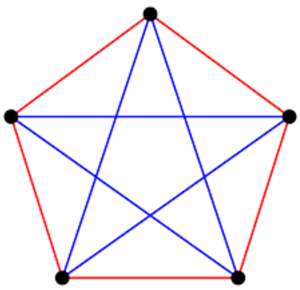
PRIMJER 1.20. (Erdős – "vjerojatnosna metoda" u kombinatorici)

Gledamo *potpun* graf s n vrhova.[17] Ako $k \leqslant n$ zadovoljava

$$\binom{n}{k} 2^{1-\binom{k}{2}} < 1\,, \tag{1.8}$$

tada je moguće obojati bridove u crveno (C) i plavo (P) tako da **ne** postoji k vrhova čiji su svi međusobni bridovi iste boje (tzv. *monokromatski* podgraf veličine k).

Na primjer, parovi $n = 5, k = 4$ (vidi Sliku 1) te $n = 20, k = 12$ zadovoljavaju (1.8).



SLIKA 1. Primjer bojanja peterokuta za koji ne postoji monokromatski podgraf veličine 4 (zapravo niti veličine 3, tj. trokut s bridovima iste boje).

[17] Dakle, *svaka* dva vrha su spojena bridom.

DOKAZ. Obojimo bridove *slučajno* tako da svaki s jednakom vjerojatnosti bude plavi ili crveni. Dakle, imamo

$$\Omega_n = \{\text{sva moguća bojanja potpunog grafa s } n \text{ vrhova}\},$$

a budući da je broj bridova jednak $\binom{n}{2}$, imamo $|\Omega_n| = 2^{\binom{n}{2}}$.

Imamo $\binom{n}{k}$ različitih podgrafova veličine k, te za *svaki* $i = 1, \dots, \binom{n}{k}$, vrijedi

$$\mathbb{P}(A_i) := \mathbb{P}(\{i\text{-ti podgraf je monokromatski}\}) = 2 \cdot \frac{1}{2^{\binom{k}{2}}}.$$

[2 dolazi od crvene ili plave boje, a $\binom{k}{2} = |\Omega_k|$]

Dakle, za $A = \{\text{postoji monokromatski podgraf}\}$, iz konačne subaditivnosti dobivamo

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{i=1}^{\binom{n}{k}} A_i\right) \leqslant \sum_{i=1}^{\binom{n}{k}} \mathbb{P}(A_i) = \binom{n}{k} 2^{1-\binom{k}{2}} \overset{(1.8)}{<} 1.$$

Dakle $|A| < |\Omega_n|$, odnosno $\{\text{postoji monokromatski podgraf}\} = A^c \neq \emptyset$. □

NAPOMENA 1.21. Za $k \in \mathbb{N}$, $R(k,k)$ definira se kao najmanji $n \geqslant k$ takav da za *svako* bojanje potpunog grafa s n vrhova postoji monokromatski podgraf s k vrhova, te ga zovemo *Ramseyev broj*. U prethodnom primjeru smo pokazali da ako par (n,k) zadovoljava (1.8), vrijedi $R(k,k) \geqslant n+1$. Na primjer, može se pokazati da je $R(4,4) = 18$, a $R(12,12) \geqslant 1640$. □

ZADATAK. Neka su $A_1, A_2, \dots \in \mathcal{F}$. Pokažite

(i) Ako je $\mathbb{P}(A_j) = 0$, za sve $j \in \mathbb{N}$, onda je i $\mathbb{P}\left(\bigcup_{j=1}^{\infty} A_j\right) = 0$.

(ii) Ako je $\mathbb{P}(A_j) = 1$, za sve $j \in \mathbb{N}$, onda je i $\mathbb{P}\left(\bigcap_{j=1}^{\infty} A_j\right) = 1$.

[Ovdje je bitno da je događaja (najviše) prebrojivo mnogo!]

NAPOMENA. Ako za događaj $A \in \mathcal{F}$ vrijedi $\mathbb{P}(A) = 0$ (odnosno, $\mathbb{P}(A) = 1$), kažemo da se A *gotovo sigurno* (*g.s.*) neće (odnosno, hoće) dogoditi. Bitno je uočiti da tada nije nužno $A = \emptyset$ (odnosno, $A = \Omega$) – npr. ako slučajno biramo broj iz $[-1,1]^2$, za događaj $B = \{\text{odabrali točku } (0,0)\} \neq \emptyset$ imamo $\mathbb{P}(B) = 0$. □

1.4.2. Neprekidnost vjerojatnosti. Za niz događaja $A_1, A_2, \dots \in \mathcal{F}$ kažemo da je *neopadajući* (odnosno, *nerastući*) ako je $A_n \subseteq A_{n+1}$ (odnosno, $A_{n+1} \subseteq A_n$) za sve $n \in \mathbb{N}$.

PROPOZICIJA 1.22. *(Neprekidnost vjerojatnosti s obzirom na neopadajući/nerastući niz događaja)*

(a) Ako je niz $A_1, A_2, \dots \in \mathcal{F}$ neopadajući, vrijedi

$$\mathbb{P}\left(\bigcup_{j=1}^{\infty} A_j\right) = \lim_{n\to\infty} \mathbb{P}(A_n). \tag{1.9}$$

(b) *Ako je niz $A_1, A_2, \dots \in \mathcal{F}$ nerestući, vrijedi*

$$\mathbb{P}\left(\bigcap_{j=1}^{\infty} A_j\right) = \lim_{n\to\infty} \mathbb{P}(A_n)\,. \tag{1.10}$$

DOKAZ. (a) Konstruiramo niz $B_1, B_2, \dots \in \mathcal{F}$ kao u Propoziciji 1.19. Budući da je $A_j \subseteq A_{j+1}$, za sve $n \in \mathbb{N}$ sada imamo

$$\bigcup_{j=1}^{n} B_j = \bigcup_{j=1}^{n} A_j = A_n\,.$$

Koristeći da su $B_1, B_2, \dots$ u parovima diskjunktni dobivamo

$$\begin{aligned}\mathbb{P}\left(\cup_{j=1}^{\infty} A_j\right) = \mathbb{P}\left(\cup_{j=1}^{\infty} B_j\right) &= \sum_{j=1}^{\infty} \mathbb{P}(B_j) = \lim_{n\to\infty} \sum_{j=1}^{n} \mathbb{P}(B_j) \\ &= \lim_{n\to\infty} \mathbb{P}(\cup_{j=1}^{n} B_j) = \lim_{n\to\infty} \mathbb{P}(A_n)\,.\end{aligned}$$

(b) DZ [Budući da su $A_1^c, A_2^c, \dots$ neopadajući, iz (a) dijela dobivamo

$$\begin{aligned}\mathbb{P}\left(\cap_{j=1}^{\infty} A_j\right) &= 1 - \mathbb{P}\left(\cup_{j=1}^{\infty} A_j^c\right) = 1 - \lim_{n\to\infty} \mathbb{P}(A_n^c) \\ &= 1 - \lim_{n\to\infty} \left(1 - \mathbb{P}(A_n)\right) = \lim_{n\to\infty} \mathbb{P}(A_n)\,.]\end{aligned}$$

□

PRIMJER 1.23. Bacamo simetričan novčić (beskonačno mnogo puta).

(a) Odredite vjerojatnost događaja $A = \{$P je palo barem jedan put$\}$.

(b) Ako je s bilo koji niz slova P,G duljine $r \in \mathbb{N}$ (npr. za $r = 3$, $s =$ PPG), odredite vjerojatnost događaja $B = \{$niz s se pojavio barem jedan put$\}$.

RJEŠENJE. Imamo $\Omega = \{\text{P,G}\}^{\mathbb{N}}$.

(a) Niz događaja $A_n := \{$P je palo barem jedan put u prvih n bacanja$\}$, $n \in \mathbb{N}$, je *neopadajući* te vrijedi $A = \cup_{n=1}^{\infty} A_n$. Neprekidnost vjerojatnosti sad povlači da je

$$\begin{aligned}\mathbb{P}(A) &= \lim_{n\to\infty} \mathbb{P}(A_n) = \lim_{n\to\infty} 1 - \mathbb{P}(A_n^c) \\ &= \lim_{n\to\infty} 1 - \mathbb{P}\{\text{samo G u prvih } n \text{ bacanja}\}) = \lim_{n\to\infty} \left(1 - \tfrac{1}{2^n}\right) = 1\,.\end{aligned}$$

[Dakle, vjerojatnost da pismo neće nikada pasti je 0.]

(b) Ponovno, neka je

$$B_n := \{\text{niz } s \text{ se pojavio barem jedan put unutar prvih } n \text{ bacanja}\}\,,\ n \in \mathbb{N}\,,$$

[Problem je sada što se $\mathbb{P}(B_n^c)$ ne može tako jednostavno odrediti budući da se se različiti blokovi duljine $r \geqslant 2$ mogu sadržavati ista bacanja.]

Podijelimo bacanja u *disjunktne* blokove od po r bacanja. Dakle, unutar n bacanja imamo $k_n = \lfloor n/r_n \rfloor$ cijelih blokova. Sada je

$$B_n \subseteq \{s \text{ se nije pojavio niti u jednom od prvih } k_n \text{ blokova}\} =: C_n\,,\ n \in \mathbb{N}\,,$$

a budući da C_n ovisi samo o prvih $k_n r$ bacanja, imamo

$$\mathbb{P}(C_n) = \frac{(2^r - 1)^{k_n}}{2^{k_n r}} = \left(1 - \frac{1}{2^r}\right)^{k_n}$$

[Zapravo, ovdje koristimo "nezavisnost" blokova.]
Sada, budući da $k_n \to \infty$ i $|1 - \frac{1}{2^r}| < 1$, imamo

$$0 \leqslant \mathbb{P}(B_n^c) \leqslant \mathbb{P}(C_n) \to 0\,,$$

pa specijalno i $\lim_{n\to\infty} \mathbb{P}(B_n^c) = 0$, tj.

$$\mathbb{P}(B) = \lim_{n\to\infty} \mathbb{P}(B_n) = 1\,.$$

□

Zadatak. Pokažite da neprekidnost vjerojatnosti povlači da **za svaki** niz $A_1, A_2, \dots \in \mathcal{F}$ vrijedi

$$\mathbb{P}\left(\bigcup_{j=1}^{\infty} A_j\right) = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{j=1}^{n} A_n\right)\,, \tag{1.11}$$

te

$$\mathbb{P}\left(\bigcap_{j=1}^{\infty} A_j\right) = \lim_{n\to\infty} \mathbb{P}\left(\bigcap_{j=1}^{n} A_n\right)\,. \tag{1.12}$$

[Sljedeći rezultat pokazuje da konačna aditivnost i neprekidnost povlače σ-aditivnost. To je nekad korisno jer je prva dva svojstva često lakše provjeriti.]

Propozicija 1.24. *Neka je $\mathbb{P} : \mathcal{F} \to [0,1]$ funkcija koja zadovoljava $\mathbb{P}(\Omega) = 1$ te je konačno aditivna. Ako vrijedi (a) (1.9), ili (b) (1.10) u slučaju $\cap_{j=1}^{\infty} A_j = \emptyset$, tada je $\mathbb{P}$ i σ-aditivna (dakle, $\mathbb{P}$ je vjerojatnost).*

Dokaz. Pogledati sami.

[(a) Neka je $(A_n)_{n\in\mathbb{N}}$ niz po parovima disjunktnih događaja. Iz (1.9) lako slijedi da je $\mathbb{P}\left(\cup_{j=1}^{\infty} A_j\right) = \lim_{n\to\infty} \mathbb{P}\left(\cup_{j=1}^{n} A_j\right)$ (vidi (1.11)), a onda iz konačne aditivnosti imamo

$$\mathbb{P}\left(\cup_{j=1}^{\infty} A_j\right) = \lim_{n\to\infty} \sum_{j=1}^{n} \mathbb{P}\left(A_j\right) = \sum_{j=1}^{\infty} \mathbb{P}\left(A_j\right)\,.$$

(b) Pokažimo da vrijedi (1.9) za proizvoljan neopadajući niz događaja $(A_n)_{n\in\mathbb{N}}$, pa tvrdnja slijedi iz (a). Neka je $A = \cup_{n=1}^{\infty} A_n$ te $B_n := A \setminus A_n = A_n^c \cap A \in \mathcal{F}$, $n \in \mathbb{N}$ [nacrtati sliku]. Tada je $B_1 \supseteq B_2 \supseteq \dots$ te

$$\bigcap_{n=1}^{\infty} B_n = \bigcap_{n=1}^{\infty} (A_n^c \cap A) = \left(\bigcap_{n=1}^{\infty} A_n^c\right) \cap A = \left(\bigcup_{n=1}^{\infty} A_n\right)^c \cap A = \emptyset\,.$$

Iz pretpostavke (b) sada slijedi

$$0 = \mathbb{P}(\emptyset) = \lim_{n\to\infty} \mathbb{P}(B_n) = \lim_{n\to\infty} \mathbb{P}(A \setminus A_n) = \lim_{n\to\infty} \left(\mathbb{P}(A) - \mathbb{P}(A_n)\right)\,,$$

otkud dobivamo da je $\mathbb{P}(A) = \lim_{n\to\infty} \mathbb{P}(A_n)$. Uočite da treća jednakost slijedi iz $\mathbb{P}(C \setminus D) = \mathbb{P}(C) - \mathbb{P}(D)$ za $C \supseteq D$, što je posljedica konačne aditivnosti.]

$\square$

POGLAVLJE 2

Uvjetna vjerojatnost i nezavisnost

2.1. Uvjetna vjerojatnost

PRIMJER 2.1. [Motivacija za definiciju uvjetne vjerojatnosti] Bacamo dvije kocke, dakle

$$\Omega = \{(i,j) : i,j \in \{1,2,\ldots,6\}\}.$$

- Ako je $A = \{\text{zbroj je } 4\} = \{(1,3),(2,2),(3,1)\}$, imamo

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{3}{6^2} = \frac{1}{12}.$$

- Ako *znamo* da se dogodio

$$B = \{\text{na prvoj kocki pala dvojka}\} = \{(2,1),\ldots,(2,6)\},$$

kolika je sada vjerojatnost događaja A?
$\rightsquigarrow$ sada B postaje "novi" Ω te su opet svi događaji jednako vjerojatni, pa budući da je od svih $\omega \in B$, jedino ishod $(2,2)$ povoljan za A, imamo

$$\begin{aligned}\mathbb{P}(A \mid B) &:= \textbf{"uvjetna vjerojatnost od } A \textbf{ uz dano } B\textbf{"} \\ &= \frac{|\{(2,2)\}|}{|B|} = \frac{1}{6} \left(> \frac{1}{12} = \mathbb{P}(A)\right).\end{aligned}$$

Uočimo, vrijedi

$$\mathbb{P}(A \mid B) = \frac{|A\cap B|}{|B|} = \frac{|A\cap B|/|\Omega|}{|B|/|\Omega|} = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}.$$

DEFINICIJA 2.2. Neka je $(\Omega,\mathcal{F},\mathbb{P})$ proizvoljan vjerojatnosni prostor. Ako $B \in \mathcal{F}$ zadovoljava $\mathbb{P}(B) > 0$, za sve $A \in \mathcal{F}$ definiramo

$$\mathbb{P}(A \mid B) := \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}. \tag{2.1}$$

Vrijedi

$$\mathbb{P}(A\cap B) = \mathbb{P}(A \mid B)\mathbb{P}(B). \tag{2.2}$$

NAPOMENA 2.3. Ako je $\mathbb{P}(B) = 0$, $\mathbb{P}(A \mid B)$ nije definirana, ali svejedno koristimo (2.2), tj. u tom slučaju definiramo

$$\mathbb{P}(A \mid B)\mathbb{P}(B) := 0 \ \ (= \mathbb{P}(A\cap B)).$$

□

2.1.1. Formula potpune vjerojatnosti (FPV). Iz (2.2) za sve $A, B \in \mathcal{F}$ slijedi

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) = \mathbb{P}(A \mid B)\mathbb{P}(B) + \mathbb{P}(A \mid B^c)\mathbb{P}(B^c)\,. \tag{2.3}$$

Općenito, za konačnu ili prebrojivu familiju događaja $(H_i)_{i\in I}$ (dakle, $I = \{1, 2, \ldots, n\}$ ili $I = \mathbb{N}$), kažemo da je **potpun sistem događaja** (PSD) ako vrijedi (a) $H_i \cap H_j$, za sve $i \neq j$, te (b) $\cup_{i\in I} H_i = \Omega$.[1]

Propozicija 2.4. (**FPV**) *Za sve $A \in \mathcal{F}$ i bilo koji PSD $(H_i)_{i\in I}$ vrijedi*

$$\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}(A \mid H_i)\mathbb{P}(H_i)\,. \tag{2.4}$$

Dokaz. Budući da je $A \subseteq \Omega$, imamo

$$\begin{aligned}\mathbb{P}(A) = \mathbb{P}(A \cap \Omega) &= \mathbb{P}\left(A \cap \Big(\bigcup_{i\in I} H_i\Big)\right) = \mathbb{P}\left(\bigcup_{i\in I} (A \cap H_i)\right) \\ &= \sum_{i\in I} \mathbb{P}(A \cap H_i) = \sum_{i\in I} \mathbb{P}(A \mid H_i)\mathbb{P}(H_i)\,,\end{aligned}$$

pri čemu četvrta jednakost slijedi iz konačne/σ-aditivnosti [tu koristimo da je I konačan ili prebrojiv, te da su H_i u parovima disjunktni], a zadnja jednakost iz (2.2). $\square$

Napomena 2.5. Za fiksan $A \in \mathcal{F}$, za (2.4) je dovoljno da vrijedi $H_i \cap H_j$, za sve $i \neq j$, te $A \subseteq \cup_{i\in I} H_i$ (umjesto $\cup_{i\in I} H_i = \Omega$). $\square$

Primjer 2.6. (Primjer 1.18) Igrači A,B,C naizmjence bacaju kocku sve dok ne dobiju šesticu. Ako je

$$\begin{aligned} A_1 &= \{\text{A je prvi dobio 6}\}\,,\ a := \mathbb{P}(A_1)\,, \\ B_1 &= \{\text{B je prvi dobio 6}\}\,,\ b := \mathbb{P}(B_1)\,, \\ C_1 &= \{\text{C je prvi dobio 6}\}\,,\ c := \mathbb{P}(C_1)\,, \end{aligned}$$

odredite a, b i c.

Rješenje. Koristeći PSD $H_1 := \{\text{A je dobio 6 u prvom bacanju}\}$, $H_2 := H_1^c$, iz FPV dobivamo

$$\begin{aligned} a = \mathbb{P}(A_1) &= \mathbb{P}(A_1 \mid H_1)\mathbb{P}(H_1) + \mathbb{P}(A \mid H_1^c)\mathbb{P}(H_1^c) \\ &= 1 \cdot \frac{1}{6} + \mathbb{P}(C_1) \cdot \frac{5}{6} = \frac{1}{6} + \frac{5}{6}c\,. \end{aligned}$$

Jednakost $\mathbb{P}(A \mid H_1^c) = \mathbb{P}(C_1)$ slijedi jer ako se dogodio H_1^c, A "postaje" treći igrač po redu!

Nadalje, kako bi odredili c, primijetimo da vrijedi

$$\begin{aligned} C_1 \subseteq &\{\text{u prva dva bacanja nema 6ice, u trećem bacanju pala 6ica}\} \\ &\cup \{\text{u prva dva bacanja nema 6ice, u trećem bacanju nije pala 6ica}\} \\ &=: H_1' \cup H_2'\,. \end{aligned}$$

[1] Drugim riječima $(H_i)_{i\in I}$ je *particija* skupa Ω.

Budući da su H_1' i H_2' još i disjunktni, iz FPV (vidi Napomenu 2.5) dobivamo

$$\begin{aligned} c &= \mathbb{P}(C_1 \mid H_1')\mathbb{P}(H_1') + \mathbb{P}(C_1 \mid H_2')\mathbb{P}(H_2') \\ &= 1 \cdot \frac{5^2 \cdot 1}{6^3} + \mathbb{P}(C_1) \cdot \frac{5^3}{6^3} = \frac{25}{216} + \frac{125}{216} c \,. \end{aligned}$$

Iz gornjeg slijedi $c = \frac{25}{91}$, te $a = \frac{1}{6} + \frac{5}{6}c = \frac{36}{91}$ [Usporedite s rješenjem iz Primjera 1.18?!]

Za vježbu pokažite da vrijedi $b = \frac{5}{6^2} + \frac{5^2}{6^2}c$, te specijalno, $b = \frac{30}{91}$.[2] □

[Napomena. Nekad nas baš zanima $\mathbb{P}(A \mid B)$ te ju računamo po definiciji (2.1), ali često "znamo" $\mathbb{P}(A \mid B)$ i to koristimo kako bi izračunali npr. $\mathbb{P}(A)$ ili $\mathbb{P}(A \cap B)$, kao u prethodnom primjeru.]

2.1.2. Bayesov teorem.

Teorem 2.7. (**Bayes**) *Ako je $A \in \mathcal{F}$ takav da je $\mathbb{P}(A) > 0$ te $(H_i)_{i\in I}$ PSD, za sve $j \in I$ vrijedi*

$$\mathbb{P}(H_j \mid A) = \frac{\mathbb{P}(A \mid H_j)\mathbb{P}(H_j)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A \mid H_j)\mathbb{P}(H_j)}{\sum_{i\in I} \mathbb{P}(A \mid H_i)\mathbb{P}(H_i)} \,. \tag{2.5}$$

Dokaz. Za vježbu. □

Primjer 2.8. Laboratorijski krvni test je 95% učinkovit u otkrivanju jedne rijetke bolesti kada je ta bolest prisutna. Test također ukazuje na bolest i u 0.05% slučajeva kada je osoba zdrava (*false positive*). Poznato je da samo 0.001% populacije ima tu bolest. Test je pokazao da imate bolest, kolika je vjerojatnost da je zaista imate?

Rješenje. Gledamo PSD $H_1 := \{\text{imate bolest}\}$, $H_2 := H_1^c$, te vrijedi

$$\mathbb{P}(H_1) = \frac{1}{100000} \,,\ \mathbb{P}(H_2) = \frac{99999}{100000} \,.$$

Neka je $T = \{\text{test je pozitivan}\}$; dakle imamo

$$\mathbb{P}(T \mid H_1) = 0.95 \,,\ \mathbb{P}(T \mid H_2) = 0.005 \,,$$

te nas zanima $\mathbb{P}(H_1 \mid T)$. Koristeći Bayesovu formulu dobivamo

$$\begin{aligned} \mathbb{P}(H_1 \mid T) &= \frac{\mathbb{P}(T \mid H_1)\mathbb{P}(H_1)}{\mathbb{P}(T \mid H_1)\mathbb{P}(H_1) + \mathbb{P}(T \mid H_2)\mathbb{P}(H_2)} = \frac{0.95 \cdot \frac{1}{100000}}{0.95 \cdot \frac{1}{100000} + 0.005 \cdot \frac{99999}{100000}} \\ &= \frac{0.95}{0.95 + 499.95} \approx 0.002 = 0.2\% \,, \end{aligned}$$

iako je $\mathbb{P}(T \mid H_1) = 0.95$!?

[Uočite da je gornja vjerojatnost iznimno mala zbog člana 499.95, koji je velik jer je $\mathbb{P}(H_2)$ jako velika!]

[2] Ako iskoristimo činjenicu da će 6ica pasti gotovo sigurno, onda je naravno najednostavnije iskoristiti da je $b = 1 - a - c$.

[Ipak, ovaj test nije nužno beskoristan jer se praksi testiraju *samo* osobe kod kojih, na temelju nekih drugih indikatora, već sumnja da imaju tu bolest. Za njih je $\mathbb{P}(H_2)$ *puno manja* nego u općenitoj populaciji.]

□

2.1.3. Ostala svojstva uvjetne vjerojatnosti. Za proizvoljne $A_1, A_2, \dots, A_n \in \mathcal{F}$, vrijedi

$$\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2 \mid A_1)\mathbb{P}(A_3 \mid A_1 \cap A_2) \cdots \mathbb{P}(A_n \mid A_1 \cap \cdots \cap A_{n-1}), \quad (2.6)$$

pri čemu desnu stranu shvaćamo kao 0 ukoliko je $\mathbb{P}(A_1 \cap \cdots \cap A_k) = 0$ za neki $k = 1, \dots, n-1$; dokaz ostavljamo za vježbu.

[Raspišemo desnu stranu po definiciji i dobijemo

$$\mathbb{P}(A_1)\frac{\mathbb{P}(A_2 \cap A_1)}{P(A_1)}\frac{\mathbb{P}(A_3 \cap A_1 \cap A_2)}{\mathbb{P}(A_1 \cap A_2)} \cdots \frac{\mathbb{P}(A_n \cap A_1 \cap \cdots \cap A_{n-1})}{\mathbb{P}(A_1 \cap \cdots \cap A_{n-1})}.$$

Svi članovi se pokrate osim brojnika u zadnjem faktoru. Ukoliko je $\mathbb{P}(A_1 \cap \cdots \cap A_k) = 0$ za neki $k = 1, \dots, n-1$, lijeva strana je nužno jednaka 0.]

Primjer 2.9. U kutiji je n bijelih i n crnih kuglica. SLučajno izvučemo $k \leqslant n$ kuglica, jednu za drugom bez vraćanja. Kolika je vjerojatnost da su izvučene samo bijele kuglice?

Rješenje. Neka je B_k traženi događaj, te

$$A_i := \{\text{izvukli smo bijelu kuglicu u } i\text{-tom izvlačenju}\}, \ i = 1, \dots, k.$$

Tada je

$$\mathbb{P}(B_k) = \mathbb{P}(A_1 \cap A_k) \overset{(2.6)}{=} \mathbb{P}(A_1)\mathbb{P}(A_2 \mid A_1)\mathbb{P}(A_3 \mid A_1 \cap A_2) \cdots \mathbb{P}(A_k \mid A_1 \cap \cdots \cap A_{k-1}).$$

Računamo

- $\mathbb{P}(A_1) = \frac{n}{2n} = \frac{1}{2}$ [$= \mathbb{P}(A_i)$, za sve i],
- $\mathbb{P}(A_2 \mid A_1) =$ [u kutiji ostalo $n-1$ B i n C kuglica] $= \frac{n-1}{2n-1}$, to jest općenito
- $\mathbb{P}(A_l \mid A_{l-1} \cap \cdots \cap A_1) = \frac{n-(l-1)}{2n-(l-1)} = \frac{n-l+1}{2n-l+1}$, $l = 1, \dots, k$.

Dakle,

$$\mathbb{P}(B_k) = \frac{n}{2n} \cdot \frac{n-1}{2n-1} \cdots \frac{n-k+1}{2n-k+1}.$$

[Ovo je dakle formalan dokaz tvrdnje koja je intuitivno jasna!] □

Za proizvoljan $B \in \mathcal{F}$ takav da je $\mathbb{P}(B) > 0$, definiramo funkciju $\mathbb{P}_B : \mathcal{F} \to \mathbb{R}$ sa

$$\mathbb{P}_B(A) := \mathbb{P}(A \mid B), \ A \in \mathcal{F}. \quad (2.7)$$

Propozicija 2.10. *$\mathbb{P}_B$ je **vjerojatnost** na $(\Omega, \mathcal{F})$.*

Dokaz prethodne propozicije ostavljamo za vježbu.

[DOKAZ.

(A1) $\mathbb{P}_B(A) = \mathbb{P}(A \mid B) = \mathbb{P}(A \cap B)/\mathbb{P}(B) \geqslant 0$, za sve $A \in \mathcal{F}$.

(A2) $\mathbb{P}_B(\Omega) = \mathbb{P}(\Omega \cap B)/\mathbb{P}(B) = \mathbb{P}(B)/\mathbb{P}(B) = 1$.

(A3) Ako su $A_1, A_2, \ldots \in \mathcal{F}$ u parovima disjunktni, imamo i da su $A_1 \cap B, A_2 \cap B, \ldots \in \mathcal{F}$ u parovima disjunktni, pa vrijedi

$$\mathbb{P}_B\left(\bigcup_{j=1}^{\infty} A_j\right) = \frac{\mathbb{P}_B\left((\bigcup_{j=1}^{\infty} A_j) \cap B\right)}{\mathbb{P}(B)} = \frac{\mathbb{P}_B\left(\bigcup_{j=1}^{\infty}(A_j \cap B)\right)}{\mathbb{P}(B)} =$$

$$= \frac{\sum_{j=1}^{\infty} \mathbb{P}_B\left(A_j \cap B\right)}{\mathbb{P}(B)} = \sum_{j=1}^{\infty} \mathbb{P}(A_j \mid B) = \sum_{j=1}^{\infty} \mathbb{P}_B(A_j)\,,$$

pri čemu smo u trećoj jednakosti iskoristili σ-aditivnost od $\mathbb{P}$.]

Prethodna propozicija kaže da $\mathbb{P}_B(\,\cdot\,)$ ima ista svojstva kao i "obična" vjerojatnost, npr.

- $\mathbb{P}(A^c \mid B) = 1 - \mathbb{P}(A \mid B)$;
- Ako su A i C disjunktni, $\mathbb{P}(A \cup C \mid B) = \mathbb{P}(A \mid B) + \mathbb{P}(A \mid C)$.

Nadalje, lako se pokaže da za sve $A, C \in \mathcal{F}$ takve da je $\mathbb{P}(A \mid B) > 0$, vrijedi

$$\mathbb{P}_B(C \mid A) = \mathbb{P}(C \mid B \cap A)\,, \tag{2.8}$$

iz čega dalje slijedi

$$\mathbb{P}(A \cap C \mid B) = [(2.2) \text{ za } \mathbb{P}_B] = \mathbb{P}_B(A)\mathbb{P}_B(C \mid A) \overset{(2.8)}{=} \mathbb{P}(A \mid B)\mathbb{P}(C \mid A \cap B)\,,$$

te

$$\mathbb{P}(A \mid B) = [\text{FPV za } \mathbb{P}_B] = \sum_{i \in I} \mathbb{P}_B(H_i)\mathbb{P}_B(A \mid H_i) \overset{(2.8)}{=} \sum_{i \in I} \mathbb{P}(H_i \mid B)\mathbb{P}(A \mid H_i \mid B)\,,$$

za svaki PSD $(H_i)_{i \in I}$.

[Glavna poruka je da gornja svojstva (osim (2.8)) ne treba posebno "pamtiti", već jednostavno zapamtiti da slijede iz činjenice da je $\mathbb{P}_B$ "obična" vjerojatnost.]

2.2. Nezavisnost

DEFINICIJA 2.11. (i) Događaji $A, B \in \mathcal{F}$ su **nezavisni** ako vrijedi

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)\,. \tag{2.9}$$

(ii) Događaji $A, B \in \mathcal{F}$ su **uvjetno nezavisni uz dani** $C \in \mathcal{F}$, $\mathbb{P}(C) > 0$, ako vrijedi

$$\mathbb{P}(A \cap B \mid C) = \mathbb{P}(A \mid C)\mathbb{P}(B \mid C)\,, \tag{2.10}$$

to jest ako su **nezavisni** s obzirom na $\mathbb{P}_C$!

NAPOMENA 2.12. Provjeru sljedećih tvrdnji ostavljamo za vježbu.

(i) Ako je $\mathbb{P}(B) > 0$ i $A \in \mathcal{F}$ proizvoljan, (2.9) je ekvivalentno s

$$\mathbb{P}(A \mid B) = \mathbb{P}(A)\,, \tag{2.11}$$

što daje intuitivniju definiciju nezavisnosti.

(ii) Ako je $\mathbb{P}(B) = 0$ ili 1, A i B su nezavisni **za svaki** $A \in \mathcal{F}$.

(iii) Ako je $\mathbb{P}(B \cap C) > 0$ (tj. $\mathbb{P}_C(B) > 0$) i $A \in \mathcal{F}$ proizvoljan, (2.10) je ekvivalentno s

$$\mathbb{P}(A \mid B \cap C) = \mathbb{P}(A \mid C)\,. \tag{2.12}$$

(iv) Ako je $\mathbb{P}(B \mid C) = 0$ ili 1, (2.10) vrijedi **za svaki** $A \in \mathcal{F}$.

[(iii) i (iv) slijede direktno iz (i) i (ii) kada se primjene na $\mathbb{P}_C$.] □

Primjer 2.13. Bacamo dvije kocke te neka je

$$\begin{aligned} A &= \{\text{na prvoj kocki } 3\}\,, \\ B &= \{\text{na drugoj kocki } 4\}\,, \\ C &= \{\text{zbroj je } 7\}\,. \end{aligned}$$

Tada vrijedi

(a) A i B su nezavisni [inače očito imamo problem u definiciji];

(b) A i C su nezavisni, B i C su nezavisni [intuicija?];

(c) uz dano C, A i B **nisu** nezavisni (tj. kažemo da su **zavisni**).

Dokaz. (a) Imamo $\Omega = \{1, 2, \ldots, 6\}^2$ te

$$\mathbb{P}(A \cap B) = \frac{|A \cap B|}{|\Omega|} = \frac{1}{6^2} = \frac{1}{6} \cdot \frac{1}{6} = \mathbb{P}(A)\mathbb{P}(B)\,.$$

[U nastavku, u ovom i sličnim slučajevima mi ćemo koristiti "činjenicu" da su A i B nezavisni, tj. računat ćemo $\mathbb{P}(A \cap B) = [\text{nezavisnost}] = \mathbb{P}(A)\mathbb{P}(B) = \frac{1}{6} \cdot \frac{1}{6}$.]

(b) Budući da smo u Laplaceovom modelu,

$$\mathbb{P}(A \mid C) = \frac{|A \cap C|}{|C|} = \frac{|\{(3,4)\}|}{|\{(1,6),(2,5),\ldots,(6,1)\}|} = \frac{1}{6} = \mathbb{P}(A)\,,$$

te analogno $\mathbb{P}(B \mid C) = \mathbb{P}(B)$. Dakle, pojam nezavisnosti je različit od pojma disjunktnosti. Za vježbu pokažite da su disjunktni događaji $A, B \in \mathcal{F}$ nezavisni akko je $\mathbb{P}(A) = 0$ ili $\mathbb{P}(B) = 0$.[3]

(c) Intuitivno, imamo zavisnost jer ako *znamo* da je zbroj 7, to da je na drugoj kocki pao broj 4 nam odmah govori da je na prvoj kocki morao pasti broj 3. Formalno, imamo dakle $\mathbb{P}(A \mid B \cap C) = 1$, dok je

$$\mathbb{P}(A \mid C) \overset{(b)}{=} \frac{1}{6} \neq \mathbb{P}(A \mid B \cap C)\,.$$

□

[3] Na primjer, ako je $\mathbb{P}(A) \in (0,1)$, A^c i A su nužno zavisni jer je $\mathbb{P}(A^c \mid A) = 0 \neq \mathbb{P}(A)$.

PRIMJER 2.14. U prvoj kutiji nalazi se 9 bijelih i 1 crna, a u drugoj 3 bijele i 7 crnih kuglica. Slučajno izaberemo jednu kutiju te iz nje izvučemo 2 kuglice s vraćanjem. Neka su

$$B_i = \{i\text{-ta izvučena kuglica je bijela}\}\,,\ i = 1, 2\,,$$
$$A = \{\text{u početku smo izabrali prvu kutiju}\}\,.$$

Tada vrijedi

(a) B_1 i B_2 su uvjetno nezavisni uz dano A.

(b) B_1 i B_2 su zavisni.

DOKAZ. (a) Vrijedi

$$\mathbb{P}(B_1 \cap B_2 \mid A) = [\text{izvlačimo 2 kuglice iz 1. kutije}] = \frac{9 \cdot 9}{10 \cdot 10} = \frac{9}{10} \cdot \frac{9}{10} = \mathbb{P}(B_1 \mid A)\mathbb{P}(B_2 \mid A)\,.$$

[Ovdje je uvjetna nezavisnost zapravo pretpostavka modela.]

(b) Intuicija? Budući da je u 1. kutiji veći omjer bijelih kuglica, očekujemo da je $\mathbb{P}(A \mid B_1) > \mathbb{P}(A)$ pa onda i $\mathbb{P}(B_2 \mid B_1) > \mathbb{P}(B_2)$. Zaista,

$$\mathbb{P}(B_2 \mid B_1) = [\text{FPV za } \mathbb{P}_{B_1}] = \mathbb{P}(B_2 \mid B_1 \cap A)\mathbb{P}(A \mid B_1) + \mathbb{P}(B_2 \mid B_1 \cap A^c)\mathbb{P}(A^c \mid B_1)\,,$$

pri čemu je

- $\mathbb{P}(B_2 \mid B_1 \cap A) \overset{(a)}{=} \mathbb{P}(B_2 \mid A) = \frac{9}{10}$;
- analogno, $\mathbb{P}(B_2 \mid B_1 \cap A^c) = \mathbb{P}(B_2 \mid A^c) = \frac{3}{10}$;
- koristeći Bayesovu formulu,

$$\mathbb{P}(A \mid B_1) = \frac{\mathbb{P}(B_1 \mid A)\mathbb{P}(A)}{\mathbb{P}(B_1 \mid A)\mathbb{P}(A) + \mathbb{P}(B_1 \mid A^c)\mathbb{P}(A^c)} = \frac{\frac{9}{10} \cdot \frac{1}{2}}{\frac{9}{10} \cdot \frac{1}{2} + \frac{3}{10} \cdot \frac{1}{2}} = \frac{3}{4}\,,$$

što je zaista veće nego $\frac{1}{2} = \mathbb{P}(A)$;
- $\mathbb{P}(A^c \mid B_1) = 1 - \mathbb{P}(A \mid B_1) = \frac{1}{4}$.

Dakle,

$$\mathbb{P}(B_2 \mid B_1) = \frac{9}{10} \cdot \frac{3}{4} + \frac{3}{10} \cdot \frac{1}{4} = \frac{3}{4}\,.$$

S druge strane,

$$\begin{aligned}\mathbb{P}(B_2) &= \mathbb{P}(B_2 \mid A)\mathbb{P}(A) + \mathbb{P}(B_2 \mid A^c)\mathbb{P}(A^c)\\ &= \frac{9}{10} \cdot \frac{1}{2} + \frac{3}{10} \cdot \frac{1}{2} = \frac{3}{5} < \mathbb{P}(B_2 \mid B_1)\,.\end{aligned}$$

□

Poruka prethodna dva primjera je da, općenito, nezavisnost ne povlači uvjetnu nezavisnost, i obratno.

LEMA 2.15. *Ako $A, B \in \mathcal{F}$ nezavisni, onda su nezavisni i parovi događaja:*

$$A^c \ i\ B\,,\ A\ i\ B^c\,,\ A^c\ i\ B^c\,.$$

Dokaz. Imamo

$$\begin{aligned}\mathbb{P}(A^c \cap B) &= \mathbb{P}(B) - \mathbb{P}(A \cap B) = [\text{nezavisnost}] = \mathbb{P}(B) - \mathbb{P}(A)\mathbb{P}(B) \\ &= \mathbb{P}(B)(1 - \mathbb{P}(A)) = \mathbb{P}(B)\mathbb{P}(A^c)\,,\end{aligned}$$

ili

$$\mathbb{P}(A^c \mid B) = 1 - \mathbb{P}(A \mid B) = [\text{nezavisnost}] = 1 - \mathbb{P}(A) = \mathbb{P}(A^c)\,,$$

ako je $\mathbb{P}(B) > 0$ (inače su trivijalno nezavisni). Ostale tvrdnje se dokazuju analogno. □

2.2.1. Nezavisnost familije događaja.

Definicija 2.16. Familija događaja $(A_i)_{i\in I}$ pri čemu I može biti proizvoljan skup, je

(i) **nezavisna** ako vrijedi

$$\mathbb{P}\left(\bigcap_{i\in F} A_i\right) = \prod_{i\in F} \mathbb{P}(A_i)\,,$$

za svaki **konačan** i neprazan podskup $F \subseteq I$.

(ii) **uvjetno nezavisna** uz dano $C \in \mathcal{F}$ ($\mathbb{P}(C) > 0$), ako je nezavisna [u smislu definicije (i)] s obzirom na $\mathbb{P}_C$.

(iii) u **parovima nezavisna** ako su A_i i A_j nezavisni za sve $i, j \in I$, $i \neq j$.

Na primjer, $A, B, C \in \mathcal{F}$ su nezavisni ako vrijedi

- $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$, $\mathbb{P}(A \cap C) = \mathbb{P}(A)\mathbb{P}(C)$, $\mathbb{P}(B \cap C) = \mathbb{P}(B)\mathbb{P}(C)$, te još i
- $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$.

[Zašto ne tražiti samo $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$? Uzmite na primjer $C = \emptyset$, onda bi za sve A, B imali da su A, B, C nezavisni što očito nema smisla. Nadalje, imamo iduću napomenu koja pokazuje da naša definicija i intuitivno ima smisla.]

Napomena 2.17. Ako su $(A_i)_{i\in I}$ nezavisni, vrijedi

$$\mathbb{P}\left(\bigcap_{i\in F_2} A_i \,\middle|\, \bigcap_{i\in F_1} A_i\right) = \mathbb{P}\left(\bigcap_{i\in F_2} A_i\right) = \prod_{i\in F_2} \mathbb{P}(A_i)\,, \tag{2.13}$$

za sve konačne, neprazne i **disjunktne** podskupove $F_1, F_2 \subseteq I$ (pri čemu pretpostavljamo da je $\mathbb{P}\left(\bigcap_{i\in F_1} A_i\right) > 0$). Provjeru ove tvrdnje ostavljamo za vježbu. □

Primjer 2.18. Neka su događaji A, B, C kao u Primjeru 2.13. Pokazali smo da su A, B, C u parovima nezavisni, ali tvrdimo da *nisu nezavisni*. To slijedi iz

$$\mathbb{P}(A \mid B \cap C) = 1 \neq \frac{1}{6} = \mathbb{P}(A)\,,$$

što je kontradikcija s (2.13), ili iz

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\{(3,4)\}) = \frac{1}{36} \neq \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{6}{36} = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)\,.$$

[Dakle, općenito u parovima nezavisnost **ne povlači** nezavisnost.] □

Napomena 2.19. Ako je familija događaja $(A_i)_{i\in I}$ nezavisna, tada je nezavisna i svaka familija događaja $(B_i)_{i\in I}$ pri čemu je $B_i = A_i$ ili A_i^c, za sve $i \in I$ (bez dokaza). $\square$

Primjer 2.20. Igrači A i B igraju niz *nezavisnih* igara, pri čemu svaka igra završava

- pobjedom igrača A s vjerojatnošću p,
- pobjedom igrača B s vjerojatnošću q,
- neriješeno s vjerojatnošću r,

za neke $0 < p, q, r < 1$ takve da vrijedi $p + q + r = 1$. Odredite vjerojatnost događaja $A := \{\text{A je prvi došao do pobjede}\}$.

Rješenje. Za svaki $n \in \mathbb{N}$, definiramo događaje

$$H_n := \{\text{prva pobjeda bilo kojeg igrača se dogodila u } n\text{-toj igri}\},$$
$$A_n := A \cap H_n\,.$$

Tada vrijedi $A = \bigcup_{n\in\mathbb{N}} A_n$, te su događaji $A_1, A_2, \ldots$ u parovima disjunktni.

Nadalje, pretpostavljamo da se igre "nastavljaju" igrati i nakon što je jedan od igrača prvi došao do pobjede, te za svaki $n \in \mathbb{N}$ definiramo događaje

$$D_n := \{\text{neriješeno u } n\text{-toj igri}\},$$
$$P_n := \{\text{A pobijedio u } n\text{-toj igri}\}.$$

1. rješenje: Imamo

$$\begin{aligned}
\mathbb{P}(A) &= \sum_{n=1}^{\infty} \mathbb{P}(A_n) = \sum_{n=1}^{\infty} \mathbb{P}(D_1 \cap \cdots \cap D_{n-1} \cap P_n) \\
&= [\text{nezavisnost igara}] = \sum_{n=1}^{\infty} \mathbb{P}(D_1)\ldots\mathbb{P}(D_{n-1})\mathbb{P}(P_n) \\
&= \sum_{n=1}^{\infty} r^{n-1}p = [|r| < 1] = p \cdot \frac{1}{1-r} = \frac{p}{p+q}\,.
\end{aligned}$$

2. rješenje: Uvjetovanjem na rezultat 1. igre dobivamo (detalji za vježbu)

$$\mathbb{P}(A) = p \cdot 1 + r \cdot \mathbb{P}(A) + q \cdot 0 = p + r\mathbb{P}(A)\,,$$

iz čega slijedi da je $\mathbb{P}(A) = \frac{1}{1-r}$.

3. rješenje: Zašto je rješenje baš $\frac{p}{p+q}$? Ključno je sljedeće: vrijedi

$$\begin{aligned}
\mathbb{P}(A \mid H_n) &= \mathbb{P}(P_n \mid H_n) = \mathbb{P}(P_n \mid D_1 \cap \cdots \cap D_{n-1} \cap D_n^c) = \mathbb{P}(P_n \mid D_n^c) \\
&= \frac{\mathbb{P}(P_n)}{\mathbb{P}(D_n^c)} = \frac{p}{1-r} = \frac{p}{p+q}\,,
\end{aligned}$$

neovisno o $n \in \mathbb{N}$; formalno opravdanje prve, treće i četvrte jednakosti ostavljamo za vježbu [iskoristite redom $A \cap H_n = P_n \cap H_n$, nezavisnost igara, te $P_n \subseteq D_n^c$]. Sada iz

FPV dobivamo

$$\mathbb{P}(A) = \sum_{n=1}^{\infty} \mathbb{P}(A \mid H_n)\mathbb{P}(H_n) = \frac{p}{p+q} \sum_{n=1}^{\infty} \mathbb{P}(H_n)$$

pri čemu se lako provjeri da iz

$$\mathbb{P}(\{\text{netko je pobijedio u } n\text{-toj igri}\}) = \mathbb{P}(D_n^c) = p + q = 1 - r > 0\,,\ n \in \mathbb{N}$$

te nezavisnosti bacanja slijedi

$$\sum_{n=1}^{\infty} \mathbb{P}(H_n) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} H_n\right) = \mathbb{P}(\{\text{netko je pobijedio}\}) = 1\,.$$

[ili iskoristi 2. BC lemu dolje niže.]

□

2.2.2. Borel-Cantellijeve leme.

Lema 2.21 (Borel-Cantelli). *Ako za niz događaja $A_1, A_2, \dots \in \mathcal{F}$ vrijedi*

$$\sum_{i=1}^{\infty} \mathbb{P}(A_i) < \infty\,,$$

onda je

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right) = 0\,.$$

Interpretacija: za ishod $\omega \in \Omega$, imamo $\omega \in \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$ akko za svaki $n \in \mathbb{N}$ postoji $k = k(\omega) \geqslant n$ takav da je $\omega \in A_k$, tj.

$$\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \{\text{dogodilo se beskonačno mnogo } A_n\text{-ova}\} =: \{A_n \text{ b.m.p.}\}\,.$$

Zadatak. Koja je interpretacija događaja $\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$?

Dokaz. Neka je $B_n := \bigcup_{k=n}^{\infty} A_k, n \in \mathbb{N}$. Budući da su $B_1, B_2, \dots$ nerastući, iz neprekidnosti vjerojatnosti imamo

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} \mathbb{P}(B_n) = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k=n}^{\infty} A_k\right)$$
$$\leqslant [\sigma\text{-subaditivnost}] = \lim_{n\to\infty} \sum_{k=n}^{\infty} \mathbb{P}(A_k) = 0\,,$$

pri čemu zadnja jednakost slijedi zbog $\sum_{i=1}^{\infty} \mathbb{P}(A_i) < \infty$ [ostatak konvergentnog reda]. □

Primjer 2.22. Bacamo niz nesimetričnih novčića pri čemu vrijedi

$$p_n := \mathbb{P}(\{\text{palo pismo na } n\text{-tom novčiću}\}) = \frac{1}{n^2}\,.$$

Budući da je $\sum_{n=1}^{\infty} p_n < \infty$, BC lema povlači da je

$$\mathbb{P}(\{\text{pismo palo b.m.p.}\}) = 0\,,$$

to jest

$$\mathbb{P}(\{\text{nakon nekog trenutka su padale samo glave}\}) = 1\,.$$

Napomenimo da s druge strane, u slučaju $p_n = \frac{1}{n}$, $n \in \mathbb{N}$, zbog $\sum_{n=1}^{\infty} p_n = \infty$ i "nezavisnosti" bacanja vrijedi $\mathbb{P}(\{\text{pismo palo b.m.p.}\}) = 1$, iako $p_n \to 0$! [to je tzv. druga BC lema] $\square$

Lema 2.23. (2. BC lema) *Ako je $A_1, A_2, \dots \in \mathcal{F}$ niz* ***nezavisnih*** *događaja takvih da vrijedi*

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty\,,$$

onda je

$$\mathbb{P}(\{A_n \text{ b.m.p.}\}) = 1\,.$$

Napomena. Nezavisnost je **bitna**. Na primjer, ako je $B \in \mathcal{F}$ takav da vrijedi $0 < \mathbb{P}(B) < 1$, za niz $A_n := B$, $n \in \mathbb{N}$, imamo $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$, ali

$$\mathbb{P}(\{A_n \text{ b.m.p.}\}) = \mathbb{P}(B) < 1\,.$$

$\square$

Dokaz. Neka je ponovno

$$A := \{A_n \text{ b.m.p.}\} = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k\,,$$

te $B_n := \cup_{k=n}^{\infty} A_k$. Tada je $(B_n)_{n\in\mathbb{N}}$ nerastući niz događaja i vrijedi $A = \cap_{n=1}^{\infty} B_n$, pa neprekidnost vjerojatnosti povlači da je $\mathbb{P}(A) =\searrow \lim_{n\to\infty} \mathbb{P}(B_n)$. Dakle, imamo $\mathbb{P}(A) = 1$ akko $\mathbb{P}(B_n) = 1$ za svaki $n \in \mathbb{N}$ akko $\mathbb{P}(B_n^c) = 0$ za svaki $n \in \mathbb{N}$ [idemo na komplement da iskoristimo nezavisnost].

Za proizvoljan $n \in \mathbb{N}$ imamo

$$\begin{aligned}\mathbb{P}(B_n^c) &= \mathbb{P}\left(\bigcap_{k=n}^{\infty} A_k^c\right) = [\text{neprekidnost, vidi (1.12)}] = \lim_{m\to\infty} \mathbb{P}\left(\bigcap_{k=n}^{m} A_k^c\right)\\ &= [\text{nezavisnost}] = \lim_{m\to\infty} \prod_{k=n}^{m} \mathbb{P}(A_k^c) = \lim_{m\to\infty} \prod_{k=n}^{m} (1 - \mathbb{P}(A_k))\,.\end{aligned}$$

Iz nejednakosti $1 - x \leqslant e^{-x}$ za sve $x \geqslant 0$, slijedi

$$\prod_{k=n}^{m} (1 - \mathbb{P}(A_k)) \leqslant \prod_{k=n}^{m} e^{-\mathbb{P}(A_k)} = e^{-\sum_{k=n}^{m} \mathbb{P}(A_k)}\,,\ \forall m > n\,.$$

Iz pretpostavke $\sum_{k=1}^{\infty} \mathbb{P}(A_k) = \infty$ slijedi da je $\lim_{m\to\infty} \sum_{k=n}^{m} \mathbb{P}(A_k) = \sum_{k=n}^{\infty} \mathbb{P}(A_k) = \infty$, pa imamo

$$0 \leqslant \mathbb{P}(B_n^c) \leqslant \lim_{m\to\infty} e^{-\sum_{k=n}^{m} \mathbb{P}(A_k)} = 0$$

to jest $\mathbb{P}(B_n^c) = 0$. $\square$

Primjer 2.24. Pretpostavimo da su događaji $A_1, A_2, \dots \in \mathcal{F}$ **nezavisni** te takvi da je

$$\mathbb{P}(A_n) = p > 0\,,\ \forall n \in \mathbb{N}\,.$$

[npr. bacanje novčića, nezavisne igre ...]

Tada iz $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$ i 2. BC leme odmah slijedi $\mathbb{P}(\{A_n \text{ b.m.p.}\}) = 1$. Specijalno,

$$\mathbb{P}(\{\text{dogodit će se barem jedan } A_n\}) = 1\,,$$

i sve to **bez obzira** koliko je p blizu 0; vidi "Infinite monkey theorem". □

2.3. Paradoksi

PRIMJER 2.25. (**Bertrand**) Imamo tri kutije: u prvoj su 2 crne, u drugoj 2 bijele, a u trećoj po jedna bijela i crna kuglica. Slučajno izaberemo kutiju te zatim jednu kuglicu iz te kutije. Ako je izvučena crvena kuglica, kolika je vjerojatnost da je i druga kuglica iz te kutije također crna?

RJEŠENJE. Neka je

$$C_i = \{i\text{-ta izvučena kuglica je crna}\}\,,\ i = 1, 2$$

$$H_i = \{\text{u početku izabrana } i\text{-ta kutija}\}\,,\ i = 1, 2, 3\,.$$

Očito je $\mathbb{P}(H_i) = \frac{1}{3}$ za sve i, a zanima nas

$$\mathbb{P}(C_2 \mid C_1) = \mathbb{P}(C_2 \cap C_1 \mid C_1) = \mathbb{P}(H_1 \mid C_1)\,.$$

Prvo rješenje. Koristeći Bayesa dobivamo

$$\mathbb{P}(H_1 \mid C_1) = \frac{\mathbb{P}(C_1 \mid H_1)\mathbb{P}(H_1)}{\sum_{i=1}^{3} \mathbb{P}(C_1 \mid H_i)\mathbb{P}(H_i)} = \frac{\mathbb{P}(C_1 \mid H_1)}{\sum_{i=1}^{3} \mathbb{P}(C_1 \mid H_i)} = \frac{1}{1 + 0 + \frac{1}{2}} = \frac{2}{3}\,,$$

te $\mathbb{P}(H_2 \mid C_1) = 0, \mathbb{P}(H_3 \mid C_1) = \frac{1}{3}$. [Dakle, $\mathbb{P}(H_1 \mid C_1) \neq \frac{1}{2}$!]

Drugo rješenje. Neka je[4]

$$\Omega = \{\text{koja od 6 kuglica je izvučena}\} = \{C_1^{(1)}, C_2^{(1)}, C_1^{(3)}, B_1^{(2)}, B_2^{(2)}, B_1^{(3)}\}\,.$$

Budući da je *za sve* $\omega \in \Omega$, $\mathbb{P}(\omega) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$, nalazimo se u Laplacevom modelu. Dakle, imamo

$$\mathbb{P}(H_1 \mid C_1) = \frac{|H_1 \cap C_1|}{|C_1|} = \frac{|\{C_1^{(1)}, C_2^{(1)}\}|}{|\{C_1^{(1)}, C_2^{(1)}, C_1^{(3)}\}|} = \frac{2}{3}\,,$$

te analogno $\mathbb{P}(H_3 \mid C_1) = \frac{1}{3}$. Ovaj fenomen naziva se "pristranost po veličini" (engl. "size-biasing"). [Jednostavno je: prva kutija ima dva puta veću vjerojatnost da smo je izabrali jer ima *dva puta više crnih kuglica* od treće kutije. Ipak, ovaj fenomen se javlja u raznim kontekstima te nije ga uvijek trivijalno prepoznati ("paradoks vremena čekanja").] □

PRIMJER 2.26. (**Monty Hall**) Imate izbor od troja vrata – iza jednih se nalazi auto, a iza druga dvoja vrate su koze. Izabrali ste prva, a nakon toga je voditelj otvorio treća vrata te je iza njih bila koza. Ako možete, trebate li promijeniti vaš izbor vrata?

[4] Na primjer, $C_2^{(1)}$ je druga crna kuglica iz prve kutije.

Rješenje. ["Paradoks" je u tome što problem nije jednoznačno zadan, tj. postoji više rješenja.]

Neka je

$$
\begin{aligned}
H_i &= \{\text{izabrali ste } i\text{-ta vrata}\}\,,\ i = 1,2,3\\
A_i &= \{\text{auto se nalazi iza } i\text{-tih vrata}\}\,,\ i = 1,2,3\,,\\
V_i &= \{\text{voditelj otvara } i\text{-ta vrata}\}\,,\ i = 1,2,3\,,\\
K &= V_3 \cap A_3^c\,.
\end{aligned}
$$

Ako definiramo

$$p_i := \mathbb{P}(A_i \mid H_1 \cap K) = \mathbb{P}_{H_1}(A_i \mid K) =: \mathbb{P}_1(A_i \mid K)\,,\ i = 1,2,3\,,$$

zanima nas je li $p_2 > p_1$ (očito je $p_3 = 0$). Koristeći Bayesa dobivamo

$$p_2 = \frac{\mathbb{P}_1(K \mid A_2)\mathbb{P}_1(A_2)}{\sum_{j=1}^{3} \mathbb{P}_1(K \mid A_j)\mathbb{P}_1(A_j)} = [\mathbb{P}_1(A_j) = \tfrac{1}{3}\,, \forall j\,, \mathbb{P}(K \mid A_3) = 0] = \frac{\mathbb{P}_1(K \mid A_2)}{\mathbb{P}_1(K \mid A_2) + \mathbb{P}_2(K \mid A_2)}\,.$$

Budući da se uvjetno na A_j, $j = 1,2$, sigurno dogodio A_3^c, vrijedi $\mathbb{P}_1(K \mid A_j) = \mathbb{P}_1(V_3 \mid A_j)$, $j = 1,2$, tj.

$$p_2 = \frac{\mathbb{P}_1(V_3 \mid A_2)}{\mathbb{P}_1(V_3 \mid A_2) + \mathbb{P}_2(V_3 \mid A_2)}\,.$$

Dakle, zanimaju nas vjerojatnosti $\mathbb{P}_1(V_3 \mid A_j)$, $j = 1,2$. Ipak, one ovise o "protokolu" na temelju kojeg voditelj odlučuje koja će vrata otvoriti! [Drugim riječima, mi ne znamo sve postavke ovog slučajno pokusa, već sam ovidimo jedan njegov ishod.]

Prvi protokol: Voditelj otvara ona vrata iza kojih je koza (bira slučajno ako su iza oboja vrata koze). Tada je

$$\mathbb{P}_1(V_3 \mid A_1) = \frac{1}{2}\,,\ \mathbb{P}_1(V_3 \mid A_2) = 1\,,$$

te

$$p_2 = \frac{1}{\frac{1}{2} + 1} = \frac{2}{3} > \frac{1}{3} = p_1\,.$$

Drugi protokol: Voditelj slučajno otvara jedna od preostala dvoja vrata. Tada je

$$\mathbb{P}_1(V_3 \mid A_1) = \frac{1}{2}\,,\ \mathbb{P}_1(V_3 \mid A_2) = \frac{1}{2}\,,$$

te $p_2 = p_1 = \frac{1}{2}$, itd. [Probajte smisliti protokol u kojem je $p_2 < p_1$.] □

Primjer 2.27. (**Simpsonov paradoks**) Promatramo podatke o uspješnosti dvije vrste operacija bubrežnih kamenaca:

	Broj	Uspješnost
Tretman A	350	0.78
Tretman B	350	0.83

Čini se da je tretman B bolji od tretmana A. Ipak, ako gledamo uspješnost s obzirom na veličinu kamenca:

	kamenac $>$ 2cm		kamenac $>$ 2cm	
	Broj	Uspješnost	Broj	Uspješnost
Tretman A	87	0.93	263	0.73
Tretman B	270	0.87	80	0.69

Dakle, A se čini bolji u oba slučaja! Problem je u startu bio što nismo uzeli u obzir *veličinu kamenca* (**"confounding"** varijabla) – veće kamence je teže tretirati, a tretman A se češće dodjeljivao baš tim težim slučajevima.

Vjerojatnosno objašnjenje: Moguće je da vrijedi

$$\mathbb{P}(\text{uspjeh} \mid A \cap \text{"} < 2\text{"}) > \mathbb{P}(\text{uspjeh} \mid B \cap \text{"} < 2\text{"}) \,,$$
$$\mathbb{P}(\text{uspjeh} \mid A \cap \text{"} > 2\text{"}) > \mathbb{P}(\text{uspjeh} \mid B \cap \text{"} > 2\text{"}) \,,$$

ali $\mathbb{P}(\text{uspjeh} \mid A) < \mathbb{P}(\text{uspjeh} \mid B)$. [Drugi poznati primjer je pitanje spolne diskriminacije pri upisu na fakultet.] □

2.4. Konstrukcija vjerojatnosnog prostora za nezavisne pokuse

Pretpostavimo da imamo dva "nezavisna" pokusa $(\Omega_i, \mathcal{F}_i, \mathbb{P}_i)$, $i = 1, 2$, pri čemu je $|\Omega_i| < \infty$, $\mathcal{F}_i = \mathcal{P}(\Omega_i)$, $i = 1, 2$ [npr. jedan pokus je bacanje 3 simetrične kocke, a drugi je bacanje 2 simetrična novčića]. Cilj je formalno definirati vjerojatnosni prostor koji sadrži oba pokusa pri čemu bi događaji vezani uz ta dva pokusa zaista bili nezavisni.

Na izmjerivom prostoru

$$(\Omega, \mathcal{F}) := (\Omega_1 \times \Omega_2, \mathcal{P}(\Omega_1 \times \Omega_2)) \,,$$

definiramo potencijalnu vjerojatnost s

$$\mathbb{P}\left((\omega_1, \omega_2)\right) := \mathbb{P}(\{(\omega_1, \omega_2)\}) := \mathbb{P}_1(\omega_1)\mathbb{P}_2(\omega_2) \,, \ \forall (\omega_1, \omega_2) \in \Omega \,. \tag{2.14}$$

Budući da je $|\Omega| < \infty$, a lako se provjeri da je $\mathbb{P}(\omega) \geqslant 0$, $\forall \omega \in \Omega$ i $\sum_{\omega \in \Omega} \mathbb{P}(\omega) = 1$, Propozicija 1.12 povlači da se $\mathbb{P}$ može jedinstveno proširiti do vjerojatnosti na $(\Omega, \mathcal{F})$. Nadalje, iz (1.5) i (2.14) slijedi da za sve $A \subseteq \Omega$ nužno vrijedi

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega) = \sum_{(\omega_1, \omega_2) \in A} \mathbb{P}_1(\omega_1)\mathbb{P}_2(\omega_2) \,,$$

pa za $A_1 \subseteq \Omega_1, A_2 \subseteq \Omega_2$ specijalno vrijedi

$$\mathbb{P}(A_1 \times A_2) = \mathbb{P}_1(A_1)\mathbb{P}_2(A_2) \,. \tag{2.15}$$

Sada za $A_1 \subseteq \Omega_1$, ali kao događaj u Ω, imamo

$$\{\text{dogodio se } A_1\} = A_1 \times \Omega_2 =: \widetilde{A}_1 \subseteq \Omega \,.$$

Analogno, za $A_2 \subseteq \Omega_2$ definiramo $\widetilde{A}_2 := \Omega_1 \times A_2$.

Za sve $A_1 \subseteq \Omega_1, A_2 \subseteq \Omega_2$ sada vrijedi

$$\mathbb{P}(\widetilde{A}_1 \cap \widetilde{A}_2) = \mathbb{P}(A_1 \times A_2) \overset{(2.15)}{=} \mathbb{P}_1(A_1)\mathbb{P}_2(A_2)\,. \tag{2.16}$$

Ako u gornju jednakost uvrstimo $A_2 := \Omega_2$, dobivamo $\mathbb{P}(\widetilde{A}_1) = \mathbb{P}_1(A_1)$, te analogno da je $\mathbb{P}(\widetilde{A}_2) = \mathbb{P}_2(A_2)$. Uvrštavajući u (2.16) slijedi da je

$$\mathbb{P}(\widetilde{A}_1 \cap \widetilde{A}_2) = \mathbb{P}(\widetilde{A}_1)\mathbb{P}(\widetilde{A}_2)\,,$$

to jest, $\widetilde{A}_1$ i $\widetilde{A}_2$ su zaista nezavisni.

Napomena. (i) Analogno se konstruira vjerojatnost na $(\Omega_1 \times \cdots \times \Omega_n, \mathcal{P}(\Omega_1 \times \cdots \times \Omega_n))$, te (ii) se slično može napraviti i na $(\Omega_1 \times \Omega_2 \times \ldots, \mathcal{F})$ pri čemu sada $\mathcal{F} \neq \mathcal{P}(\Omega_1 \times \Omega_2 \times \ldots)$ (bez detalja). $\square$

POGLAVLJE 3

Diskretne slučajne varijable

3.1. Definicija i osnovna svojstva

U nastavku $(\Omega, \mathcal{F}, \mathbb{P})$ predstavlja proizvoljan vjerojatnosni prostor.

DEFINICIJA 3.1. **Slučajna varijabla** je svaka funkcija $X : \Omega \to \mathbb{R}$ takva da $\forall x \in \mathbb{R}$ vrijedi

$$\{X \leqslant x\} := \{\omega \in \Omega : X(\omega) \leqslant x\} = X^{-1}(\langle -\infty, x]) \in \mathcal{F}. \tag{3.1}$$

NAPOMENA 3.2. Za slučajnu varijablu X zanimat će nas vjerojatnosti

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

za razne skupove $B \subseteq \mathbb{R}$ [npr. $B = \langle -\infty, x], [a, b], \{x\}$]. Može se pokazati da (3.1) povlači puno više, tj. da je u tom slučaju $\{X \in B\} \in \mathcal{F}$ za "većinu" $B \subseteq \mathbb{R}$, pa je vjerojatnost $\mathbb{P}(X \in B)$ **dobro definirana**. U nastavku uvjet (3.1) nećemo provjeravati [tj. pretpostavit ćemo da je zadovoljen]. Uočimo da je taj uvjet trivijalno zadovoljen ako je $\mathcal{F} = \mathcal{P}(\Omega)$. □

Za slučajnu varijablu $X : \Omega \to \mathbb{R}$, neka je

$$D_X := \mathrm{Im}(X) = \{X(\omega) : \omega \in \Omega\},$$

skup svih vrijednosti koje X može poprimiti.

DEFINICIJA 3.3. Slučajnu varijabla $X : \Omega \to \mathbb{R}$ je **diskretna** ako je D_X konačan ili prebrojiv.

Ako je $D_X \subseteq B$ za neki $B \subseteq \mathbb{R}$, često to označavamo s "$X \in B$" (npr. $X \geqslant 0$ znači da je $D_X \subseteq [0, \infty\rangle$).

PRIMJER 3.4. Bacamo kocku te neka je

$$X := \text{broj koji je pao},$$

tj. $\Omega = \{1, 2, \dots, 6\}$, $X(\omega) := \omega$, $\forall \omega \in \Omega$. Budući da je $X : \Omega \to D_X := \{1, 2, \dots, 6\} \subseteq \mathbb{R}$, X je diskretna slučajna varijabla. □

PRIMJER 3.5. Slučajno biramo broj iz $\langle 0, 10\rangle$, te stavimo

$$X := \text{taj broj}\,, Y := \text{najveće cijelo tog broja}\,,$$

tj. $\Omega = \langle 0, 10\rangle$, te $X(\omega) := \omega$, $Y(\omega) := \lfloor \omega \rfloor$, $\forall \omega$. Tada su i X i Y slučajne varijable, li samo je Y diskretna ($D_X = \langle 0, 10\rangle$, $D_Y = \{0, 1, \dots, 9\}$).

Notacija:

(i) Slučajne varijable tipično označavamo velikim slovima: $X, Y, U, \ldots$, a njihove vrijednosti ("realizacije") s malim slovima: $x, y, u, \ldots$ (npr. $x = X(\omega)$ za neki ω).

(ii) Ako je X *diskretna* slučajna varijabla, često ćemo elemente od D_X označiti s

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

gdje je $I = \{1, 2, \ldots, n\}$ za neki $n \in \mathbb{N}$, ili $I = \mathbb{N}$ [I dakle također ovisi o X, ali radi jednostavnosti to nećemo pisati.] □

3.1.1. Distribucija diskretne slučajne varijable. Za diskretnu slučajnu varijablu X, skup $D_X = \{a_i : i \in I\}$ zajedno s vjerojatnostima

$$p_i := \mathbb{P}(X = a_i)\,,\ i \in I\,, \tag{3.2}$$

zovemo **distribucijom** (ili **razdiobom**) slučajne varijable X. Distribuciju često zapisujemo u obliku **tablice**:

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

[Gornju relaciju čitamo kao "X ima distribuciju ..."]

Primjer 3.6. Ako je X broj koji je pao na simetričnoj kocki,

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

što predstavlja tzv. diskretnu **uniformnu** razdiobu na $\{1, 2, \ldots, 6\}$. □

Iz (3.2) imamo da nužno vrijedi (DZ)

$$p_i \geqslant 0\,, \forall i \in I\,, \quad \text{te} \quad \sum_{i \in I} p_i = 1\,, \tag{3.3}$$

tj. $(p_i : i \in I)$ je *distribucija* na D_X (vidi Prop. 1.12).

[Zaista, prva tvrdnja je očita, a druga slijedi jer

$$\begin{aligned} \sum_{i \in I} p_i &= \sum_{i \in I} \mathbb{P}(X = a_i) = [\text{konačna ili } \sigma\text{-aditivnost}] \\ &= \mathbb{P}\left(\bigcup_{i \in I} \{X = a_i\}\right) = \mathbb{P}(X \in D_X) = 1\,.] \end{aligned}$$

[U nastavku pokazujemo da za proizvoljnu distribuciju na nekom konačnom ili prebrojivom skupu možemo konstruirati vjerojatnosti prostor i slučajnu varijablu na tom prostoru s upravo tom distribucijom.]

Propozicija 3.7. *Neka je $(p_i : i \in I)$ proizvoljna distribucija na konačnom ili prebrojivom skupu $D = \{a_i : i \in I\} \subseteq \mathbb{R}$. Tada postoji vjerojatnosni prostor $(\Omega, \mathcal{F}, \mathbb{P})$ i diskretna slučajna varijabla $X : \Omega \to D$ t.d. $\mathbb{P}(X = a_i) = p_i$, $\forall i \in I$.*

DOKAZ. [Dokaz ovakvih tvrdnji uvijek ima istu ideju.]

Neka je

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

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

  Egzistencija slijedi zbog uvjeta (3.3) te Prop. 1.12 i napomene nakon nje.
- $X(a) := a$, $\forall a \in \Omega = D$.

Sada očito vrijedi $D_X = D$ te

$$\mathbb{P}(X = a_i) = \mathbb{P}(\{a \in \Omega : X(a) = a_i\}) = \mathbb{P}(\{a_i\}) = p_i \,, \, \forall i \in I \,.$$

□

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

$$\mathbb{P}(X \in B) = \mathbb{P}\left(\bigcup_{i \in I : a_i \in B} \{X = a_i\}\right) = \sum_{a_i \in B} \mathbb{P}(X = a_i) = \sum_{a_i \in B} p_i \,. \tag{3.4}$$

pri čemu je "$a_i \in B$" (ovdje, ali i u nastavku) skraćena oznaka za "$i \in I : a_i \in B$".

[Iz (3.4) slijedi da ako slučajna varijabla X definirana na $(\Omega, \mathcal{F}, \mathbb{P})$ i slučajna varijabla Y definirana na $(\Omega', \mathcal{F}', \mathbb{P}')$ imaju *istu distribuciju*, vrijedi

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

Najčešće nas zanimaju stvari koje ovise *samo* o distribuciji slučajne varijable, tj. najčešće nam sam vjerojatnosti prostor te konkretna konstrukcija slučajne varijable *nisu bitni*. Bitno je samo da *postoji* slučajna varijabla s danom distribucijom, a to slijedi iz Prop. 3.7!]

Za diskretnu slučajnu varijablu X definiramo njenu **vjerojatnosnu funkciju mase**[1] ili **diskretnu (vjerojatnosnu) funkciju gustoće** $f_X := \mathbb{R} \to [0,1]$ sa

$$f_X(x) := \mathbb{P}(X = x) \,, \; x \in \mathbb{R} \,. \tag{3.5}$$

Očito vrijedi

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

[1] Engleski izraz je **probability mass function** ili skraćeno **p.m.f.**. Često se koristi i intuitivnija oznaka p_X umjesto f_X.

akko

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

Drugim riječima, f_X jedinstveno određuje distribuciju od X. [U ovim bilješkama nećemo koristiti f_X, ali to je samo stvar notacije.]

3.1.2. Neke poznate distribucije. U nastavku neka je $p \in [0,1]$, $q := 1-p$, te pretpostavimo da imamo niz *nezavisnih* pokusa pri čemu je

$$\mathbb{P}(A_i) := \mathbb{P}(\{\text{uspjeh u } i\text{-tom pokusu}\}) = p\,,\ \forall i \in \mathbb{N}\,.$$

Npr. bacamo niz (potencijalno nesimetričnih) novčića gdje je vjerojatnost za pismo ("uspjeh") upravo p.

Primjer 3.8. Ako je

$$X_i := \begin{cases} 1, & \text{uspjeh u } i\text{-tom pokusu}\,, \\ 0, & \text{inače} \end{cases}$$

imamo

$$X_i \sim \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix}\,,\ \forall i \in \mathbb{N}\,.$$

Ova razdioba naziva se **Bernoullijeva razdioba** s parametrom (uspjeha) p (oznaka je "$X_i \sim \mathrm{B}(p)$"). [Ovakva slučajna varijabla često se naziva jednostavno "coin toss".]

Primjer 3.9. Općenito, za događaj $A \in \mathcal{F}$ definiramo funkciju $\mathbb{1}_A : \Omega \to \{0,1\}$ sa

$$\mathbb{1}_A(\omega) := \begin{cases} 1, & \omega \in A\,, \\ 0, & \omega \notin A\,. \end{cases}$$

Ovu funkciju (slučajnu varijablu) nazivamo **indikator** događaja A.

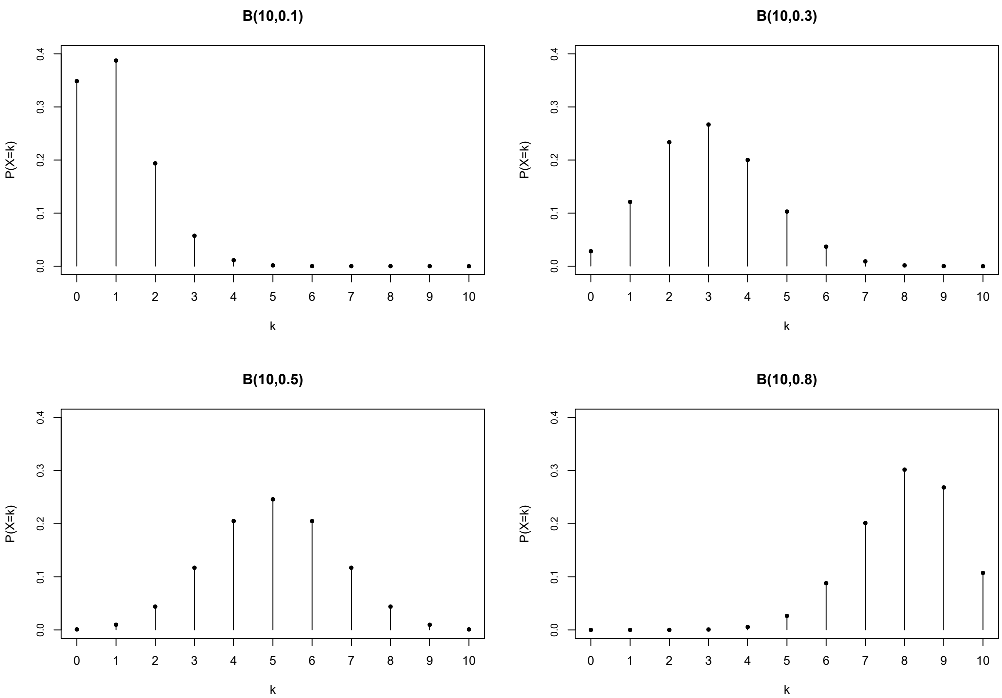
- uvijek vrijedi $\mathbb{1}_A \sim \mathrm{B}(p)$ uz
$$p = \mathbb{P}(\mathbb{1}_A = 1) = \mathbb{P}(\{\omega : \omega \in A\}) = \mathbb{P}(A)\,.$$
- u prethonom primjeru vrijedi $X_i = \mathbb{1}_{A_i}$, $\forall i$.

Primjer 3.10. Ako je $n \in \mathbb{N}$ fiksan te

$$X := \sum_{i=1}^{n} \mathbb{1}_{A_i} = \text{ukupan broj uspjeha u } n \text{ pokusa}\,.$$

imamo $X \in \{0,1,\ldots,n\}$ te

$$\mathbb{P}(X = k) = \binom{n}{k} p^k q^{n-k}\,,\ k \in \{0,1,\ldots,n\}\,.$$



SLIKA 2. Grafički prikaz binomne razdiobe za $n = 10$ i razne parametre p.

Ova razdioba naziva se **binomna** razdioba s parametrima n i p (oznaka "$X \sim \mathrm{B}(n, p)$"); vidi Sliku 2 za grafički prikaz ove razdiobe. Specijalno, vrijedi $\mathrm{B}(1, p) = \mathrm{B}(p)$.

PRIMJER 3.11. Ako je $p > 0$ te

$$T := \text{ukupan broj pokusa do pojave prvog uspjeha}\,,$$

imamo $T \in \{1, 2, \dots\} = \mathbb{N}$ te

$$\mathbb{P}(T = k) = q^{k-1}p\,,\ k \in \mathbb{N}\,.$$

Ova razdioba naziva se **geometrijska** razdioba na $\mathbb{N}$ s parametrom p (oznaka "$T \sim \mathrm{G}(p)$"); vidi Sliku 3. Slično, ako je

$$\widetilde{T} := \text{ukupan broj } \textit{neuspjeha} \text{ do pojave prvog uspjeha}\,,$$

vrijedi $\widetilde{T} = T - 1$, $\widetilde{T} \in \mathbb{N}_0$, te

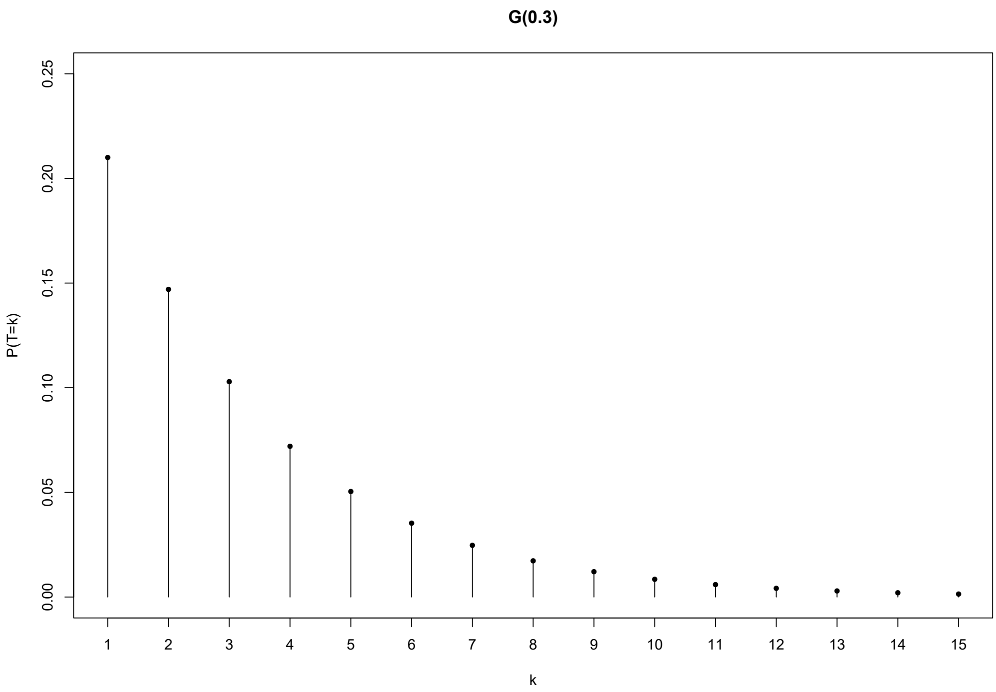
$$\mathbb{P}(T = k) = q^k p\,,\ k \in \mathbb{N}_0\,.$$

Ovo je tzv. geometrijska razdioba na $\mathbb{N}_0$ (oznaka "$\widetilde{T} \sim \mathrm{G}(p)$").

PRIMJER 3.12. Ako je $r \in \mathbb{N}$, te

(i) $Y :=$ broj pokusa do ukupno r uspjeha,

$$\mathbb{P}(Y = k) \overset{\text{DZ}}{=} \binom{k-1}{r-1} q^{k-r} p^r\,,\ k = r, r+1, \dots\,. \tag{3.6}$$



Slika 3. Grafički geometrijske razdiobe.

(ii) $Z := Y - r =$ broj neuspjeha do ukupno r uspjeha, iz (3.6) uz $k = m + r$ slijedi

$$\mathbb{P}(Z = m) = \binom{m+r-1}{r-1} q^m p^r \,,\ m = 0, 1, \ldots \,. \tag{3.7}$$

Ova razdioba naziva se **negativna binomna** razdioba s parametrima r i p (oznaka "NB(r, p)").

[Za $r = 1$, $Y \sim$ G(p), $Z \sim$ G$_0(p)$. NB razdioba se često koristi u modeliranju kada modeliramo veličinu koja poprima vrijednosti u $\mathbb{N}_0$.]

3.1.3. Funkcija slučajne varijable. Ako je X diskretna slučajna varijabla i $g : D_X \to \mathbb{R}$ prizvoljna funkcija, tada je funkcija $g(X) : \Omega \to \mathbb{R}$ definirana sa

$$g(X)(\omega) := g(X(\omega)) \,,\ \omega \in \Omega \,,$$

ponovno diskretna slučajna varijabla [jer je $D_{g(X)} = g(D_X)$ ponovno najviše prebrojiv skup].

Primjer 3.13. Ako je

$$X \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \end{pmatrix}$$

te $g(x) = x^+ := \max\{0, x\}$, imamo da je $X^+ = g(X) \in \{0, 1\}$ te

$$\mathbb{P}(X^+ = 0) = \mathbb{P}(X \leqslant 0) = \mathbb{P}(X = -1) + \mathbb{P}(X = 0) = \frac{1}{2} + \frac{1}{3} = \frac{5}{6} \,,$$

$$\mathbb{P}(X^+ = 1) = \mathbb{P}(X = 1) = \frac{1}{6} \,.$$

Dakle,

$$X^+ \sim \begin{pmatrix} 0 & 1 \\ \frac{5}{6} & \frac{1}{6} \end{pmatrix}$$

[Općenito, imamo...]

PROPOZICIJA 3.14. *Ako je* $D_X = \{a_i : i \in I\}$ *te* $D_{g(X)} = \{g(a_i) : i \in I\} =: \{b_j : j \in J\}$, *distribucija od* $g(X)$ *dana je s*

$$\mathbb{P}(g(X) = b_j) = \textstyle\sum_{i \in I : g(a_i) = b_j} \mathbb{P}(X = a_i)\,,\ \forall j \in J\,.$$

DOKAZ. Pogledati sami.

[Budući da je

$$\{\omega : g(X(\omega)) = b_j\} = \{\omega : X(\omega) \in g^{-1}(\{b_j\})\}$$

imamo

$$\mathbb{P}(g(X) = b_j) = \mathbb{P}(X \in g^{-1}(\{b_j\})) \overset{(3.4)}{=} \textstyle\sum_{a_i \in g^{-1}(\{b_j\})} \mathbb{P}(X = a_i) = \sum_{g(a_i) = b_j} \mathbb{P}(X = a_i)\,.]$$

□

3.2. Očekivanje

DEFINICIJA 3.15. Neka je X diskretna slučajna varijabla s distribucijom $X \sim \left(\begin{smallmatrix} a_1 & a_2 & \cdots\cdots \\ p_1 & p_2 & \cdots\cdots \end{smallmatrix}\right)$. Ako vrijedi

$$\sum_{i \in I} |a_i| p_i < \infty\,, \tag{3.8}$$

kažemo da X ima **(matematičko) očekivanje** koje onda definiramo kao

$$\mathbb{E}[X] := \sum_{i \in I} a_i p_i\,. \tag{3.9}$$

[Oznaka $\mathbb{E}$ dolazi od "expectation", a očekivanje često nazivamo i "prosječna vrijedost" što je vezano uz tzv. zakon velikih brojeva.]

NAPOMENA 3.16. (i) Ako je $|I| < \infty$, (3.8) uvijek vrijedi [pa dakle očekivanje uvijek postoji].

(ii) Ako je X nenegativna slučajna varijabla (tj. $a_i \geqslant 0$, $\forall i$), *uvijek* definiramo $\mathbb{E}[X]$ kao u (3.9), pri čemu dopuštamo da je $\mathbb{E}[X] = +\infty$; uz (3.8) je nužno $\mathbb{E}[X] \in \mathbb{R}$.

(iii) **Zašto tražimo (3.8)?** Ako je $(b_n)_{n \in \mathbb{N}}$ niz realnih brojeva te (a) $\sum_{n \in \mathbb{N}} |b_n| < \infty$, ili (b) $b_n \geqslant 0$, $\forall n$, tada $\sum_{n \in \mathbb{N}} b_n$ *ne ovisi* o poretku sumacije, tj. za svaku bijekciju $\pi : \mathbb{N} \to \mathbb{N}$ vrijedi

$$\sum_{n \in \mathbb{N}} b_n := \lim_{k \to \infty} \sum_{n=1}^{k} b_n = \lim_{k \to \infty} \sum_{n=1}^{k} b_{\pi(n)} = \sum_{n \in \mathbb{N}} b_{\pi(n)}\,.$$

Primjer 3.17. Ako je $X \sim \left(\begin{smallmatrix} 0 & 1 \\ q & p \end{smallmatrix}\right)$, $p \in [0,1]$, $q := 1-p$, očekivanje očito postoji te je

$$\mathbb{E}[X] = 0 \cdot q + 1 \cdot p = p\,.$$

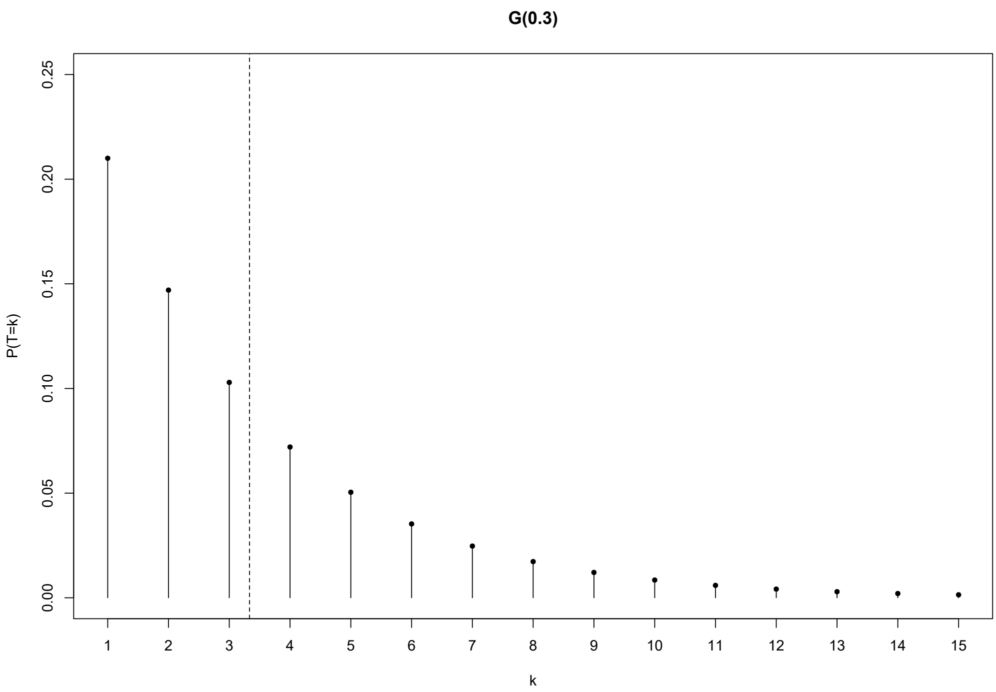
Uočimo da općenito (tj. kada je $p \neq 0, 1$), $\mathbb{E}[X] \notin \{0,1\}$. Specijalno, $\forall A \in \mathcal{F}$ vrijedi

$$\mathbb{E}\left[\mathbb{1}_A\right] = \mathbb{P}(A)\,.$$

Primjer 3.18. Neka je $T \sim \mathrm{G}(p)$ za $p \in (0,1]$ (tj. $\mathbb{P}(T=k) = q^{k-1}p$, $k \in \mathbb{N}$). Očekivanje postoji jer je $T \geqslant 0$, te iznosi[2]

$$\mathbb{E}[T] = \sum_{k=1}^{\infty} k\mathbb{P}(T=k) = p\sum_{k=1}^{\infty} kq^{k-1} = [q = 1-p \in [0,1)] = p\frac{1}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}\,.$$

Vidi Sliku 4. □



Slika 4. Geometrijska razdioba s parametrom $1/p = 0.3$. Isprekidana linija označava njeno očekivanje $p = 1/0.3 \approx 0.33$.

Primjer 3.19. Ako je X slučajna varijabla koja poprima vrijednosti u $\mathbb{N}$ (dakle $X \geqslant 0$) te[3]

$$\mathbb{P}(X = n) = \frac{1}{n(n+1)}\,,\ \forall n \in \mathbb{N}\,,$$

imamo

$$\mathbb{E}[X] = \sum_{n=1}^{\infty} n\frac{1}{n(n+1)} = \sum_{n=1}^{\infty} \frac{1}{n+1} = \infty\,.$$

[U ovom slučaju X s relativno velikim vjerojatnostima poprima velike vrijednosti.]

[2] Derivacijom jednakosti $\sum_{m=0}^{\infty} x^m = (1-x)^{-1}$ dobivamo da je $\sum_{m=1}^{\infty} mx^{m-1} = (1-x)^{-2}$ kadgod je $|x| < 1$.

[3] Provjerite da je ovo zaista distribucija na $\mathbb{N}$, tj. da je $\sum_{n\in\mathbb{N}} \frac{1}{n(n+1)} = 1$.

Zadatak. Nađite primjer diskretne slučajne varijable $X \in \mathbb{Z}$ za koju $\mathbb{E}[X]$ *ne postoji* (tj. ne vrijedi (3.8).

[Na primjer, uzmite $\mathbb{P}(X = n) = \mathbb{P}(X = -n) = \frac{1}{2}\frac{1}{n(n+1)}$, $\forall n \in N$. Tada je

$$\sum_{i \in I} |a_i| p_i = \frac{1}{2} \sum_{n \leqslant -1} |-n| \frac{1}{n(n+1)} + \frac{1}{2} \sum_{n \geqslant 1} |n| \frac{1}{n(n+1)} = \sum_{n=1}^{\infty} n \frac{1}{n(n+1)} = \infty \,. \text{]}$$

3.2.1. Osnovna svojstva.

Napomena 3.20. U nastavku ćemo koristiti tzv. **Fubinijev teorem**: Ako je $(a_{i,j} : i, j \in \mathbb{N})$ dvostruko indeksiran niz realnih brojeva, vrijedi

$$\sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} a_{i,j} = \sum_{j \in \mathbb{N}} \sum_{i \in \mathbb{N}} a_{i,j} \,,$$

tj. možemo zamijeniti poredak sumacije, ako je (a) $a_{i,j} \geqslant 0$, $\forall i, j$, ili (b) $\sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} |a_{i,j}| < \infty$. U slučaju (a) obje strane mogu biti $+\infty$.

Teorem 3.21 (!). *Neka je X diskretna slučajna varijabla s distribucijom $X \sim \left(\begin{smallmatrix} a_1 & a_2 & \cdots\cdots \\ p_1 & p_2 & \cdots\cdots \end{smallmatrix}\right)$, i $g : D_X \to \mathbb{R}$ funkcija. Ako vrijedi (a) $g(X) \geqslant 0$, ili (b) $\sum_{i \in I} |g(a_i)| p_i < \infty$, imamo da je*

$$\mathbb{E}[g(X)] = \sum_{i \in I} g(a_i) p_i \,. \tag{3.10}$$

[U slučaju (b) tvrdimo dakle i da postoji $\mathbb{E}[g(X)]$.]

Prethodni teorem je koristan jer daje način za određivanje $\mathbb{E}[g(X)]$ bez da računamo distribuciju sl. varijable $g(X)$.

Dokaz. Označimo $g(D_X) =: \{b_j : j \in J\}$ te $I_j := \{i \in I : g(a_i) = b_j\}$, $j \in J$. Iz Prop. 3.14 imamo

$$\mathbb{P}(g(X) = b_j) = \sum_{i \in I_j} p_i \,, \ \forall j \in J \,.$$

- Ako je $g(X) \geqslant 0$, tj. $b_j \geqslant 0$, $\forall j \in J$, očekivanje postoji uvijek te iznosi

$$\begin{aligned} \mathbb{E}[g(X)] &= \sum_{j \in J} b_j \mathbb{P}(g(X) = b_j) = \sum_{j \in J} b_j \sum_{i \in I_j} p_i = \sum_{j \in J} \sum_{i \in I_j} p_i b_j \\ &= \sum_{j \in J} \sum_{i \in I_j} p_i g(a_i) = \sum_{i \in I} p_i g(a_i) \in [0, \infty] \,. \end{aligned} \tag{3.11}$$

  pri čemu smo u zadnjem koraku iskoristili Fubinijev teorem uz

$$a_{i,j} := \begin{cases} p_i g(a_i), & \text{ako } i \in I_j \\ 0, & \text{ako } i \notin I_j. \end{cases}$$

  i činjenicu da skupovi I_j, $j \in J$ čine particiju skupa I [pa dakle $\sum_{j \in J} a_{i,j} = p_i g(a_i)$.]
- Ako je $\sum_{i \in I} p_i |g(a_i)| < \infty$, (3.11) povlači da je

$$\sum_{j \in J} |b_j| \mathbb{P}(g(X) = b_j) < \infty \,,$$

tj. da postoji $\mathbb{E}[g(X)]$, te onda analogno kao u (3.11) [uz primjenu Fubinijevog teorema] dobijemo (3.10) [pri čemu je dakle nužno $\mathbb{E}[g(X)] \in \mathbb{R}$].

$\square$

Napomena 3.22. Za $X \sim \left(\begin{smallmatrix} a_1 & a_2 & \cdots\cdots \\ p_1 & p_2 & \cdots\cdots \end{smallmatrix}\right)$ prethodni teorem uz $g(x)) = |x|$ daje

$$\mathbb{E}\Big[|X|\Big] = \sum_{i \in I} |a_i| p_i \,.$$

Specijalno, vrijedi (3.8) (tj. postoji $\mathbb{E}[X]$) akko $\mathbb{E}\Big[|X|\Big] < \infty$.

Propozicija 3.23. *Ako je* $\mathbb{E}\Big[|X|\Big] < \infty$, *za sve* $a, b \in \mathbb{R}$ *vrijedi*

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b \,. \tag{3.12}$$

[Tvrdnja vrijedi i ako je samo $X \geqslant 0$ *i* $a \geqslant 0$*.]*

Dokaz. Ako je $X \sim \left(\begin{smallmatrix} a_1 & a_2 & \cdots\cdots \\ p_1 & p_2 & \cdots\cdots \end{smallmatrix}\right)$, (3.10) nam daje

$$\sum_{i \in I} \underbrace{|a \cdot a_i + b|}_{\leqslant |a| \cdot |a_i| + |b|} p_i \leqslant |a| \sum_{i \in I} |a_i| p_i + b = |a| \mathbb{E}\Big[|X|\Big] + |b| < \infty \,.$$

Dakle, postoji $\mathbb{E}[X]$ te se analogno dokaže $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$. $\square$

Sljedeći teorem dokazujemo u poglavlju o slučajnim vektorima.

Propozicija 3.24. *Neka su* $X_1, \ldots, X_n$ *diskretne slučajne varijable definirane na istom* $(\Omega, \mathcal{F}, \mathbb{P})$ *[pa je dakle* $X_1 + \cdots + X_n$ *ponovno (diskretna) slučajna varijabla]. Ako vrijedi (a)* $X_i \geqslant 0$, $\forall i$, *ili (b)* $\mathbb{E}\Big[|X_i|\Big] < \infty$, $\forall i = 1, \ldots, n$, *tada je*

$$\mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i] \,. \tag{3.13}$$

Svojstva (3.12) i (3.13) zajedno nazivamo **linearnost očekivanja**.

Primjer 3.25. Neka je $X \sim \mathrm{B}(n, p)$, a zanima nas $\mathbb{E}[X]$.

1. Po definiciji imamo

$$\begin{aligned}
\mathbb{E}[X] &:= \sum_{k=0}^{n} k \mathbb{P}(X = k) = \sum_{k=1}^{n} k \binom{n}{k} p^k q^{n-k} \\
&= \sum_{k=1}^{n} k \cdot \frac{n(n-1)\cdots(n-k+1)}{k!} p^k q^{n-k} \\
&= np \sum_{k=1}^{n} \frac{(n-1)\cdots((n-1)-(k-1)+1)}{(k-1)!} p^{k-1} q^{(n-1)-(k-1)} \\
&= [l := k-1] = np \sum_{l=0}^{n-1} \binom{n-1}{l} p^l q^{n-1-l} = np(p+q)^{n-1} = np \,.
\end{aligned}$$

2. Neka su $A_1, \ldots, A_n \in \mathcal{F}$ nezavisni događaji t.d. $\mathbb{P}(A_i) = p$, $\forall i = 1, \ldots, n$. Ako stavimo $X := \sum_{i=1}^{n} \mathbb{1}_{A_i}$, vrijedi $X \sim \mathrm{B}(n, p)$ [Ovo je u redu jer očekivanje ovisi samo o distibuciji

slučajne varijable]. Dakle,

$$\mathbb{E}[X] := \mathbb{E}\left[\sum_{i=1}^n \mathbb{1}_{A_i}\right] = [\text{linearnost oč.}] = \sum_{i=1}^n \mathbb{E}\left[\mathbb{1}_{A_i}\right] = \sum_{i=1}^n \mathbb{P}(A_i) = np\,.$$

NAPOMENA 3.26. Ako vrijedi $\mathbb{P}(|X| \leqslant c) = 1$ za neki $c \geqslant 0$ [kažemo da je X ograničena g.s.], vrijedi

$$\mathbb{E}\Big[|X|\Big] \leqslant c \ \ (< \infty)\,.$$

Zaista, ako je $X \sim \left(\begin{smallmatrix} a_1 & a_2 & \cdots\cdots \\ p_1 & p_2 & \cdots\cdots \end{smallmatrix}\right)$, imamo $|a_i| > c \Rightarrow p_i = 0$, pa je

$$\begin{aligned} \mathbb{E}\Big[|X|\Big] &= \sum_{i\in I} |a_i| p_i = \sum_{|a_i|\leqslant c} |a_i| p_i \\ &= c \sum_{|a_i|\leqslant c} p_i = c \sum_{i\in I} p_i = c\,. \qquad \square \end{aligned}$$

Dokaz iduće propozicije je sličan pa ga ostavljamo za vježbu.

PROPOZICIJA 3.27. *Neka je X diskretna slučajna varijabla. Tada vrijedi:*

(a) $\mathbb{P}(X \geqslant 0) = 1$ *povlači* $\mathbb{E}[X] \geqslant 0$*;*

(b) $\mathbb{P}(X \geqslant 0) = 1$ *i* $\mathbb{E}[X] = 0$ *povlače* $\mathbb{P}(X = 0) = 1$*;*

(c) $\mathbb{P}(X = c) = 1$ *za neki* $c \in \mathbb{R}$, *povlači* $\mathbb{E}[X] = c$*;*

(d) $\mathbb{P}(a \leqslant X \leqslant b) = 1$ *za* $a \leqslant b$, *povlači* $a \leqslant \mathbb{E}[X] \leqslant b$.

TEOREM 3.28. *Neka je X slučajna varijabla koja poprima vrijednosti u $\mathbb{N}_0$. Tada vrijedi*

$$\mathbb{E}[X] = \sum_{n=1}^{\infty} \mathbb{P}(X \geqslant n) = \sum_{n=0}^{\infty} \mathbb{P}(X > n) \ \Big(\in [0, \infty]\Big)\,. \tag{3.14}$$

DOKAZ. Vrijedi

$$\begin{aligned} \sum_{n=1}^{\infty} \mathbb{P}(X \geqslant n) &= \sum_{n=1}^{\infty}\sum_{k=n}^{\infty} \mathbb{P}(X = k) = [\text{Fubini, nacrtati sliku}] \\ &= \sum_{k=1}^{\infty}\sum_{n=1}^{k} \mathbb{P}(X = k) = \sum_{k=1}^{\infty} k\mathbb{P}(X = k) = \mathbb{E}[X]\,, \end{aligned}$$

pri čemu smo Fubinijev teorem (sve je nenegativno) iskoristili uz

$$a_{n,k} := \begin{cases} \mathbb{P}(X = k), & k \geqslant n\,, \\ 0, & k < n\,. \end{cases}.$$

$\square$

PRIMJER 3.29. Za $T \sim \mathrm{G}(p)$ uz $p > 0$ imamo

$$\mathbb{P}(T > n) = \mathbb{P}(\{\text{neuspjeh u prvih } n \text{ pokusa}\}) = q^n\,,\ \forall n \in \mathbb{N}_0\,.$$

Iz (3.14) slijedi

$$\mathbb{E}[T] = \sum_{n=0}^{\infty} \mathbb{P}(T > n) = \sum_{n=0}^{\infty} q^n$$
$$= [q < 1] = \frac{1}{1-q} = \frac{1}{p}\,.$$

Primjer 3.30. Pretpostavimo da imamo niz nezavisnih događaja $A_1, A_2, \dots \in \mathcal{F}$ te da vrijedi $p_j := \mathbb{P}(A_j) \in (0,1), \forall j \in \mathbb{N}$. Definiramo

$$T := \inf\{n \geqslant 1 : \mathbb{1}_{A_n} = 1\} = \text{prvi trenutak kada se dogodio neki od } A_n\text{-ova}\,,$$

uz konvenciju $\inf \emptyset := \infty$. Uočimo, slučajna varijabla T moguće poprima vrijednost $+\infty$, te ju zovemo **proširena** (diskretna) slučajna varijabla. U ovom slučaju, ako je $\mathbb{P}(T = \infty) > 0$, stavljamo $\mathbb{E}[T] := +\infty$, a u suprotnom $\mathbb{E}[T] := \sum_{k=1}^{\infty} k\mathbb{P}(T = k)$. Kraće to možemo zapisati kao

$$\mathbb{E}[T] := \sum_{k=1}^{\infty} k\mathbb{P}(T = k) + \infty \cdot \mathbb{P}(T = \infty)\,.$$

Zanimaju nas kako vrijednosti $\mathbb{P}(T = \infty)$ i $\mathbb{E}[T]$ ovise o nizu $p_1, p_2, \dots$.

- Uočimo, za sve $n \geqslant 0$,

$$\mathbb{P}(T > n) = \mathbb{P}\left(\mathbb{1}_{A_1} = 0, \dots, \mathbb{1}_{A_n} = 0\right) = \mathbb{P}(A_1^c \cap \dots \cap A_n^c) =$$
$$= [\text{nezavisnost}] = \prod_{j=1}^{n} \mathbb{P}(A_j^c) = \prod_{j=1}^{n} (1 - p_j)\,.$$

  Budući da je $\{T = \infty\} = \cap_{n=0}^{\infty} \{T > n\}$ te su događaji $\{T > n\}, n \in \mathbb{N}_0$, nerastući, neprekidnost vjerojatnosti daje

$$\mathbb{P}(T = \infty) = \lim_{n\to\infty} \mathbb{P}(T > n) = \lim_{n\to\infty} \prod_{j=1}^{n} (1 - p_j) = \prod_{j=1}^{\infty} (1 - p_j)\,.$$

- Nije teško za pokazati da vrijedi sljedeća (općenita) tvrdnja:

$$\prod_{j=1}^{\infty} (1 - p_j) > 0 \iff \sum_{j=1}^{\infty} p_j < \infty\,.$$

  [Zaista,

$$\lim_{n\to\infty} \prod_{j=1}^{n} (1 - p_j) > 0 \iff \lim_{n\to\infty} \sum_{j=1}^{n} \log(1 - p_j) > -\infty \iff \sum_{j=1}^{\infty} b_j < \infty\,,$$

  uz $b_j := -\log(1 - p_j) \in (0, \infty)$. U svim gornji slučajevima nužno imamo $\lim_{j\to\infty} p_j = 0$, pa iz $\lim_{x\to 0} \frac{\log(1+x)}{x} = 1$ slijedi $\lim_{j\to\infty} \frac{b_j}{p_j} = 1$, tj.

$$\sum_{j=1}^{\infty} b_j < \infty \iff \sum_{j=1}^{\infty} p_j < \infty \quad]$$

  Dakle, imamo

$$\mathbb{P}(T = \infty) > 0 \iff \sum_{j=1}^{\infty} p_j < \infty\,.$$

- Nadalje, uočimo da formula (3.14) vrijedi i u slučaju kada je $\mathbb{P}(T=\infty)>0$ (pa dakle i $\mathbb{E}[X]=\infty$) jer je u tom slučaju

$$\mathbb{P}(T>n)\geqslant \mathbb{P}(T=\infty)>0\,,\;\forall n\geqslant 0\,.$$

Dakle,

$$\mathbb{E}[T]=\sum_{n=0}^{\infty}\mathbb{P}(T>n)=\sum_{n=0}^{\infty}\prod_{j=1}^{n}(1-p_j)\,.$$

- Na primjer,
  - $p_j=\frac{1}{(j+1)^2},\ \forall j\Rightarrow \mathbb{P}(T=\infty)>0$ te $\mathbb{E}[T]=\infty$;
  - $p_j=\frac{1}{j+1},\ \forall j\Rightarrow \mathbb{P}(T=\infty)=0$, ali $\mathbb{E}[T]=\infty$;
  - $p_j=p>0,\ \forall j\Rightarrow \mathbb{P}(T=\infty)=0$ te $\mathbb{E}[T]=\frac{1}{p}$ (jer $T\sim \mathrm{G}(p)$). $\square$

3.3. Varijanca

Definicija 3.31. Neka je X diskretna slučajna varijabla t.d. $\mathbb{E}\Big[|X|\Big]<\infty$. **Varijanca** od X definira se kao

$$\mathrm{Var}(X):=\mathbb{E}[(X-\mathbb{E}[X])^2]\in[0,\infty]\,. \tag{3.15}$$

Riječima, $\mathrm{Var}(X)$ je prosječno *kvadratno* odstupanje X-a od $\mathbb{E}[X]$. Broj $\sigma(X):=\sqrt{\mathrm{Var}(X)}$ naziva se **standardna devijacija** od X. [Tipično nas u statistici zanima $\sigma(X)$ jer je u istim mjernim jedinicama kao i X.]

Primjer 3.32. Ako je

$$X\sim\begin{pmatrix}-1 & 1\\ \frac12 & \frac12\end{pmatrix}\,,\quad Y\sim\begin{pmatrix}-100 & 100\\ \frac12 & \frac12\end{pmatrix}$$

Imamo $\mathbb{E}[X]=\mathbb{E}[Y]=0$, ali $\mathrm{Var}(X)=1,\mathrm{Var}(Y)=100^2$, tj. $\sigma(Y)=100\gg 1=\sigma(X)$.

Propozicija 3.33 (**Svojstva varijance**). *Neka je X slučajna varijabla t.d. $\mathbb{E}\Big[|X|\Big]<\infty$. Vrijedi*

(i) $\mathrm{Var}(X)=0$ *povlači* $\mathbb{P}(X=\mathbb{E}[X])=1$*;*
(ii) $\forall a,b\in\mathbb{R}$, $\mathrm{Var}(aX+b)=a^2\mathrm{Var}(X)$*;*
(iii) $\mathrm{Var}(X)=\mathbb{E}[X^2]-(\mathbb{E}[X])^2$ *[ovako najčešće računamo* $\mathrm{Var}(X)$*].*

Dokaz. (i) Slijedi iz Prop. 3.27(b).

(ii) Vrijedi

$$\begin{aligned}\mathrm{Var}(aX+b)&=\mathbb{E}[(aX+b-\mathbb{E}[aX+b])^2]=\mathbb{E}[(aX+b-a\mathbb{E}[X]-b)^2]\\&=\mathbb{E}[a^2(X-\mathbb{E}[X])^2]=a^2\mathrm{Var}(X)\,.\end{aligned}$$

(iii) Vrijedi

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}[(X-\mathbb{E}[X])^2] = \mathbb{E}[X^2 - 2X\mathbb{E}[X] + (\mathbb{E}[X])^2] \\ &= \mathbb{E}[X^2] - 2\mathbb{E}[X]\mathbb{E}[X] + (\mathbb{E}[X])^2 = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \,,\end{aligned}$$

pri čemu smo u trećoj jednakosti koristili linearnost očekivanja i činjenicu da je $\mathbb{E}[X]$ konstanta.

$\square$

Primjer 3.34. (i) Ako je $T \sim \mathrm{G}(p)$ za $p > 0$, imamo $\mathbb{E}[T] = \frac{1}{p}$ te uz $q := 1 - p$,

$$\begin{aligned}\mathbb{E}[T(T-1)] &= \sum_{k=1}^{\infty} k(k-1)\mathbb{P}(T=k) = \sum_{k=2}^{\infty} k(k-1)q^{k-1}p \\ &= pq\sum_{k=2}^{\infty} k(k-1)q^{k-2} = [\text{2. derivacija sume geometrijskog reda}] \\ &= pq \cdot \frac{2}{(1-q)^3} = \frac{2q}{p^2}\,.\end{aligned}$$

Sada slijedi

$$\mathbb{E}[T^2] = \mathbb{E}[T(T-1)] + \mathbb{E}[T] = \frac{2q}{p^2} + \frac{1}{p} = \frac{2}{p^2} - \frac{1}{p}\,,$$

te

$$\mathrm{Var}(T) = \mathbb{E}[T^2] - (\mathbb{E}[T])^2 = \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2} = \frac{q}{p^2}\,.$$

[Uočimo da $\mathrm{Var}(T) \to 0$ kada $p \to 1$, te $\mathrm{Var}(T) \to \infty$ kada $p \to 0$.]
Napomenimo da za $\widetilde{T} := T - 1 \sim \mathrm{G}_0(p)$ vrijedi

$$\begin{aligned}\mathbb{E}[\widetilde{T}] &= \mathbb{E}[T] - 1 = \frac{1}{p} - 1 = \frac{q}{p}\,, \\ \mathrm{Var}(\widetilde{T}) &= \mathrm{Var}(T-1) = \mathrm{Var}(T) = \frac{q}{p^2}\,.\end{aligned}$$

(ii) Ako je $X \sim \mathrm{B}(p)$, budući da je $X^2 = X$, vrijedi

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \mathbb{E}[X] - (\mathbb{E}[X])^2 = p - p^2 = pq\,.$$

[Maksimalna varijanca postiže se za $p = \frac{1}{2}$.]

(iii) [Pogledati sami.] Ako je $X \sim \mathrm{B}(n,p)$, vrijedi $\mathbb{E}[X] = np$ te

$$\begin{aligned}\mathbb{E}[X(X-1)] &= \sum_{k=1}^{n} k(k-1)\binom{n}{k}p^k q^{n-k} = n(n-1)p^2 \sum_{k=2}^{n}\binom{n-2}{k-2}p^{k-2}q^{(n-2)-(k-2)} \\ &= n(n-1)p^2(p+q)^{n-2} = n(n-1)p^2\,,\end{aligned}$$

pa je dakle

$$\mathrm{Var}(X) = \mathbb{E}[X(X-1)] + \mathbb{E}[X] - (\mathbb{E}[X])^2 = \cdots = npq\,.$$

Uočimo da je, slično kao kod očekivanja, $\mathrm{Var}(X) = n\mathrm{Var}(Y)$ za $Y \sim \mathrm{B}(p)$ [kasnije ćemo vidjeti zašto].

TEOREM 3.35. *Ako su slučajne varijable X i Y definirane na istom $(\Omega, \mathcal{F}, \mathbb{P})$, te vrijedi $0 \leqslant X \leqslant Y$, nužno je*

$$(0 \leqslant)\ \mathbb{E}[X] \leqslant \mathbb{E}[Y]\,. \qquad \text{(monotonost)}$$

DOKAZ.
- ako je $\mathbb{E}[X], \mathbb{E}[Y] < \infty$, iz $Y - X \geqslant 0$ i linearnosti očekivanja slijedi[4]

$$0 \leqslant \mathbb{E}[Y - X] = \mathbb{E}[Y] - \mathbb{E}[X]$$

- dokaz u općenitom slučaju (kada je moguće $\mathbb{E}[X] = +\infty$) ćemo dati u poglavlju o slučajnim vektorima.

□

PROPOZICIJA 3.36. *Ako je X slučajna varijabla t.d. vrijedi $\mathbb{E}\Big[|X|^m\Big] < \infty$ za neki $m \in \mathbb{N}$, tada je i*

$$\mathbb{E}\Big[|X|^n\Big] < \infty\,, \quad \forall n \leqslant m\,.$$

*Broj $\mathbb{E}\Big[|X|^m\Big] < \infty$ zovemo n**-ti moment** slučajne varijable X.*

DOKAZ. Za proizvoljan realan broj $z \in \mathbb{R}$, vrijedi

$$|z|^n \leqslant \begin{cases} |z|^m\,, & \text{ako } |z| \geqslant 1\,, \\ 1\,, & \text{ako } |z| < 1, \end{cases} \leqslant 1 + |z|^m\,.$$

Sada koristeći linearnost i monotonost očekivanja odmah dobivamo

$$\mathbb{E}\Big[|X|^n\Big] \leqslant \mathbb{E}\Big[1 + |X|^m\Big] \leqslant 1 + \mathbb{E}\Big[|X|^m\Big] < \infty\,.$$

□

NAPOMENA 3.37. Iz identiteta $\mathbb{E}[X^2] = \mathrm{Var}(X) + (\mathbb{E}[X])^2$ i prethodne propozicije, slijedi

$$\exists \mathrm{Var}(X) \text{ (tj. } \mathbb{E}\Big[|X|\Big] < \infty) \text{ i } \mathrm{Var}(X) < \infty \iff \mathbb{E}[X^2] < \infty\,.$$

3.4. Poissonova razdioba

PRIMJER 3.38. Slučajna varijabla X ima **Poissonovu** razdiobu s parametrom $\lambda > 0$ (oznaka je "$X \sim \mathrm{P}(\lambda)$") ako je $D_X = \mathbb{N}_0$ te

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}\,, \quad \forall k \in \mathbb{N}_0\,.$$

[Provjerite da zaista vrijedi $\sum_{k \in \mathbb{N}_0} \mathbb{P}(X = k) = 1$.]

[4] Ovdje imamo sumu nenegativne slučajne varijable Y i nepozitivne slučajne varijable $-X$, pa za korištenje linearnosti trebamo da vrijedi $\mathbb{E}\big[|-X|\big], \mathbb{E}\big[|Y|\big] < \infty$.

TEOREM 3.39 (**Zakon rijetkih događaja**). *Neka je za sve $n \in \mathbb{N}$, $X_n \sim \mathrm{B}(n, p_n)$, pri čemu postoji*

$$\lambda := \lim_{n\to\infty} np_n = \lim_{n\to\infty} \mathbb{E}[X_n]\,, \tag{3.16}$$

te vrijedi $\lambda \in (0, \infty)$. Tada za sve $k \in \mathbb{N}_0$ vrijedi

$$\lim_{n\to\infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k)\,,$$

za $X \sim \mathrm{P}(\lambda)$.

NAPOMENA. Teorem zovemo zakon *rijetkih događaja* jer uz uvjet (3.16) nužno vrijedi $p_n \to 0$, kada $n \to \infty$; dakle, za veliki n, X_n je ukupan broj uspjeha u n nezavisnih pokusa s *jako malom* vjerojatnosti uspjeha u svakom pokusu.

DOKAZ. Budući da vrijedi $np_n \to \lambda$ i $p_n \to 0$ kada $n \to \infty$, za svaki $k \in \mathbb{N}_0$ imamo

$$\begin{aligned}\mathbb{P}(X_n = k) &= \binom{n}{k} p_n^k q_n^{n-k} = \frac{n(n-1)\cdots(n-k+1)}{k!} p_n^k q_n^{n-k} \cdot \frac{n^k}{n^k} \\ &= \frac{n(n-1)\cdots(n-k+1)}{n^k} \cdot \frac{(np_n)^k}{k!} \cdot \left(1 - \frac{np_n}{n}\right)^n \cdot (1-p_n)^{-k} \\ &\xrightarrow{n\to\infty} 1^k \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot 1^{-k} = \frac{\lambda^k}{k!} e^{-\lambda}\,.\end{aligned}$$

□

Prethodni teorem koristimo tako da za $X \sim \mathrm{B}(n, p)$ uz "veliki" n i "mali" p, aproksimiramo

$$\mathbb{P}(X = k) \approx \frac{(np)^k}{k!} e^{-np}\,, \quad k = 0, \ldots, n\,. \tag{3.17}$$

NAPOMENA 3.40. Zapravo vrijedi i puno više: Za sve n, p, za $X \sim \mathrm{B}(n, p)$ i $Y \sim \mathrm{P}(np)$ vrijedi

$$\sup_{B \in \mathbb{N}_0} |\mathbb{P}(X \in B) - \mathbb{P}(Y \in B)| \leqslant \min\{p, np^2\} \leqslant p\,. \tag{3.18}$$

Dakle, (3.17) i općenito (3.17) su dobre aproksimacije čim je p malen! [Ovi rezultati su razlog zašto je Poissonova razdioba često korišten model u statistici; npr. pri modeliranju dnevnog broja smrtnih slučajeva u nekom gradu.]

PRIMJER 3.41. Vlasnik web-stranice analizira broj posjetitelja. Svaki dan $n = 10^6$ ljudi nezavisno i slučajno posjećuje tu stranicu s vjerojatnošću $p = 2 \cdot 10^{-6}$. Ako je X ukupan broj posjeta stranici u jednom danu, vlasnika zanima $\mathbb{P}(X \geqslant 3)$.

RJEŠENJE.
- Očito je $X \sim \mathrm{B}(n, p)$, pa je

$$\mathbb{P}(X \geqslant 3) = 1 - \mathbb{P}(X \leqslant 2) = 1 - \sum_{k=0}^{2} \binom{n}{k} p^k q^{n-k}\,.$$

  Budući da je n jako velik, a p jako blizu 0, moguće je da ćemo naići na numeričke poteškoće pri računanju gornje i sličnih vjerojatnosti.

- Ipak, budući da je p jako malen, možemo koristiti Poissonovu aproksimaciju: $\mathbb{E}[X] = np = 2$ pa uz $Y \sim \mathrm{P}(2)$ aproksimiramo

$$\begin{aligned}\mathbb{P}(X \geqslant 3) &\overset{(3.18)}{\approx} \mathbb{P}(Y \geqslant 3) = 1 - \mathbb{P}(Y \leqslant 2) \\ &= 1 - e^{-2} - \frac{2^1}{1!}e^{-2} - \frac{2^2}{2!}e^{-2} = 1 - 5e^{-2} \approx 0.3233\,.\end{aligned}$$

[Dakle, greška je manja od $p = 2 \cdot 10^{-6}$.]

$\square$

Primjer 3.42. Za $X \sim \mathrm{P}(\lambda)$, $\lambda > 0$, vrijedi

$$\mathbb{E}[X] = \mathrm{Var}(X) = \lambda\,. \tag{3.19}$$

Zaista,

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \lambda e^{-\lambda} \underbrace{\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}}_{=e^{\lambda}} = \lambda\,.$$

Izvod varijance pogledati sami.

[Slično kao gore,

$$\mathbb{E}[X(X-1)] = \sum_{k=2}^{\infty} k(k-1)\frac{\lambda^k}{k!}e^{-\lambda} = \lambda^2 e^{-\lambda} \underbrace{\sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!}}_{=e^{\lambda}} = \lambda^2\,,$$

pa je $\mathrm{Var}(X) = \mathbb{E}[X(X-1)] + \mathbb{E}[X] - (\mathbb{E}[X])^2 = \cdots = \lambda$.]

Uočimo da sm ove izraze mogli naslutiti iz zakona rijetkih događaja: za $X_n \sim \mathrm{B}(n,p)$ i $\lambda := \lim_{n\to\infty} np_n$, imamo

$$\mathbb{E}[X_n] = np_n \longrightarrow \lambda\,, \quad \mathrm{Var}(X_n) = np_n \cdot (1-p_n) \longrightarrow \lambda \cdot 1 = \lambda\,. \qquad \square$$

Primjer 3.43 (Poissonova aproksimacija može vrijediti i u slučajevima kada gledamo ukupan broj uspjeha u n *zavisnih* pokusa.). Za $n \in \mathbb{N}$, neka je $I_n = \{1, \dots, n\}$ te

$$A_i^{(n)} := \{i \text{ je fiksna točka slučajno odabrane permutacije skupa } I_n\}\,,\ i = 1, \dots, n\,.$$

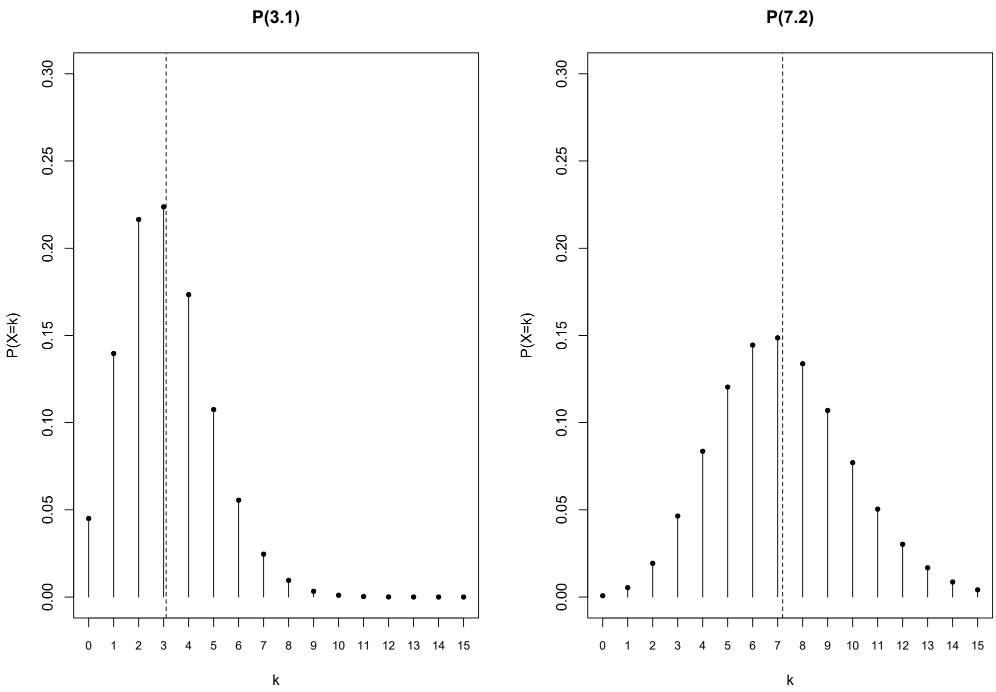
Ako je

$$X_n := \sum_{i=1}^{n} \mathbb{1}_{A_i^{(n)}} = \text{ukupan broj fiksnih točaka sluč. perm. skupa } I_n\,,$$

u Primjeru 1.17 smo pokazali da vrijedi

$$\mathbb{P}(X_n = 0) = \sum_{k=0}^{n} (-1)^k \frac{1}{k!} \longrightarrow e^{-1}\,, \quad n \to \infty\,.$$

Uočimo da je za $X \sim \mathrm{P}(1)$, $\mathbb{P}(X = 0) = e^{-1}$ te da je $\mathbb{E}[X_n] = n\mathbb{P}(A_1^{(n)}) = n \cdot \frac{1}{n} = 1$ za sve $n \in \mathbb{N}$ (linearnost očekivanja), pa specijalno i $\mathbb{E}[X_n] \to 1$ kada $n \to \infty$. Pokažimo da vrijedi i



Slika 5. Poissonova distribucija s parametrima 3.1 i 7.2. Isprekidana linija označava očekivanje, vidimo da se $\mathbb{P}(X=k)$ maksimizira u $k=3$, odnosno $k=7$ – općenito, za $X \sim \mathrm{P}(\lambda)$ maksimum se postiže za $k=\lfloor \lambda \rfloor$.

$\mathbb{P}(X_n = r) \to \mathbb{P}(X=k)$ za sve $k \in \mathbb{N}$: čim je $k \leqslant n$, zbog simetrije vrijedi

$$\begin{aligned}
\mathbb{P}(X_n = r) &= \binom{n}{r}\mathbb{P}(\overbrace{A_1^{(n)} \cap \dots A_r^{(n)}}^{=:A} \cap \overbrace{A_{r+1}^{(n)} \cap \dots A_n^{(n)}}^{=:B}) = \binom{n}{r}\mathbb{P}(A)\mathbb{P}(B \mid A) \\
&= \binom{n}{r}\frac{1^r \cdot (n-r)!}{n!} \cdot \mathbb{P}(\{\text{sluč. odabrana perm. skupa } \{r+1, \dots, n\} \text{ nema fiksnu točku}\}) \\
&= \frac{1}{r!}\mathbb{P}(X_{n-r} = 0) \xrightarrow{n\to\infty} \frac{1}{r!}e^{-1} = \mathbb{P}(X=r)\,. \qquad (3.20)
\end{aligned}$$

Ipak, (3.20) ne slijedi iz Teorema 3.39 jer X_n *nema* $\mathrm{B}(n, \frac{1}{n})$ razdiobu budući da događaji $A_1^{(n)}, \dots, A_n^{(n)}$ *nisu* nezavisni.[5]

3.5. Indikatori

Prisjetimo se, za događaj $A \in \mathcal{F}$ definiramo funkciju $\mathbb{1}_A : \Omega \to \{0,1\}$ sa

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

te je $\mathbb{E}[\mathbb{1}_A] = \mathbb{P}(A)$.

Propozicija 3.44. *Za sve $A, B \in \mathcal{F}$ vrijedi*

(a) $\mathbb{1}_{A^c} = 1 - \mathbb{1}_A$;

[5] Za više o Poissonovoj aproksimaciji u zavisnom slučaju vidi tzv. Chen-Steinovu metodu.

(b) $\mathbb{1}_{A\cap B} = \mathbb{1}_A \cdot \mathbb{1}_B$;

(c) $\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_{A\cap B}$.

Dokaz. Na primjer, (b) slijedi jer

$$\mathbb{1}_{A\cap B}(\omega) = \begin{cases} 1, & \omega \in A \text{ i } \omega \in B\,, \\ 0, & \text{inače}\,. \end{cases} = \mathbb{1}_A(\omega) \cdot \mathbb{1}_B(\omega)\,,\ \forall \omega \in \Omega\,.$$

□

Primjer 3.45. Pretpostavimo da $n \geqslant 2$ bračnih parova na slučajan način sjednu oko okruglog stola pri čemu muškarci i žene alterniraju. Ako je N ukupan broj muževa koji sjede do svojih žena, odredite $\mathbb{E}[N]$ i $\mathrm{Var}(N)$.[6]

Rješenje. Ako definiramo događaje

$$A_i := \{i\text{-ti par sjedi zajedno}\}\,,\ i = 1, \ldots, n\,,$$

vrijedi $N = \sum_{i=1}^n \mathbb{1}_{A_i}$. Sada koristeći linearnost očekivanja dobivamo

$$\mathbb{E}[N] = \sum_{i=1}^n \mathbb{P}(A_i) = \sum_{i=1}^n (\tfrac{1}{n} + \tfrac{1}{n}) = n \cdot \frac{2}{n} = 2\,.$$

Nadalje,

$$\begin{aligned} \mathbb{E}[N^2] &= \mathbb{E}\left[(\textstyle\sum_{i=1}^n \mathbb{1}_{A_i})^2\right] = \mathbb{E}\left[\textstyle\sum_{i=1}^n \mathbb{1}_{A_i}^2 + 2\sum_{i<j} \mathbb{1}_{A_i}\mathbb{1}_{A_j}\right] \\ &= \mathbb{E}\left[\textstyle\sum_{i=1}^n \mathbb{1}_{A_i} + 2\sum_{i<j} \mathbb{1}_{A_i\cap A_j}\right] = n\mathbb{P}(A_1) + 2\binom{n}{2}\mathbb{P}(A_1 \cap A_2)\,. \end{aligned}$$

Preostaje izračunati

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2 \mid A_1) = \frac{2}{n} \cdot \left(\frac{1}{n-1} \cdot \frac{1}{n-1} + \frac{(n-1)-1}{n-1} \cdot \frac{2}{n-1}\right) = \frac{4n-6}{n(n-1)^2}\,.$$

Uvjetnu vjerojatnost $\mathbb{P}(A_2 \mid A_1)$ smo izračunali tako da smo dodatno uvjetovali na to sjedi li muškarac iz 2. para odmah do 1. para ili ne. Dakle,

$$\mathbb{E}[N^2] = n \cdot \frac{2}{n} + 2\frac{n(n-1)}{2}\frac{4n-6}{n(n-1)^2} = 2 + \frac{4n-6}{n-1}\,.$$

te konačno

$$\mathrm{Var}(N) = \mathbb{E}[N^2] - (\mathbb{E}[N])^2 = 2 + \frac{4n-6}{n-1} - 2^2 = \frac{2(n-2)}{n-1}\,.$$

[Uočimo, $\lim_{n\to\infty} \mathrm{Var}(N) = 2$, a može se i pokazati da N konvergira prema P(2) distribuciji kada $n \to \infty$.] □

Primjer 3.46 (**Hipergeometrijska razdioba**). U kutiji je m_1 bijelih i m_2 crnih kuglica. Izaberemo slučajno $n \leqslant m_1 + m_2$ kuglica. Ako je X ukupan broj izvučenih bijelih kuglica, očito

[6] Ovo je zapravo varijacija problema broja fiksnih točaka slučajne permutacije.

je $X \in \{0, \ldots, n\}$, te preciznije

$$0 \leqslant X \leqslant \min\{n, m_1\}\,, \quad 0 \leqslant n - X \leqslant \min\{n, m_2\}\,.$$

Nadalje, distribucija je dana s

$$\mathbb{P}(X = k) = \frac{\binom{m_1}{k}\binom{m_2}{n-k}}{\binom{m_1+m_2}{n}}\,, \quad \max\{0, n - m_2\} \leqslant k \leqslant \min\{n, m_1\}\,.$$

Ovu razdiobu nazivamo **hipergeometrijska** razdioba (oznaka "HG(m_1, m_2, n)"). Zanima nas $\mathbb{E}[X]$. Ako definiramo događaje

$$A_i := \{\text{izvučena je } i\text{-ta bijela kuglica}\}\,, \ i = 1, \ldots, m_1\,,$$

imamo[7]

$$\begin{aligned}\mathbb{E}[X] &= \mathbb{E}\left[\textstyle\sum_{i=1}^{m_1} \mathbb{1}_{A_i}\right] = \textstyle\sum_{i=1}^{m_1} \mathbb{P}(A_i) = m_1\mathbb{P}(A_1)\\ &= m_1\frac{1 \cdot \binom{m_1+m_2-1}{n-1}}{\binom{m_1+m_2}{n}} = m_1 \cdot \frac{n}{m_1 + m_2}\,.\end{aligned}$$

Alternativno, istu razdiobu dobijemo ako n kuglica izvlačimo jednu za drugom, ali **bez vraćanja**. Ako definiramo

$$B_i := \{\text{u } i\text{-tom izvlačenju izvukli bijelu kuglicu}\}\,, \ i = 1, \ldots, n\,,$$

vrijedi $X = \sum_{i=1}^{n} \mathbb{1}_{B_i}$, pa iz $\mathbb{P}(B_i) = \frac{m_1}{m_1+m_2}$, $\forall i$, slijedi

$$\mathbb{E}[X] = n \cdot \frac{m_1}{m_1 + m_2}\,.$$

Ako je $p := \frac{m_1}{m_1+m_2}$, zašto X nema $\mathrm{B}(n, p)$ razdiobu? Vidi Sliku 6 za usporedbu ove dvije razdiobe za jedan konkretan izbor parametara – čini se da hipergeometrijska ima manju varijancu, je li to intuitivno jasno?

[Hipergeometrijska razdioba se javlja npr. u tzv. *Fisherovom egzaktnom testu* u statistici.]

Primjer 3.47. Za proizvoljne $A_1, \ldots, A_n \in \mathcal{F}$ vrijedi [Zašto?]

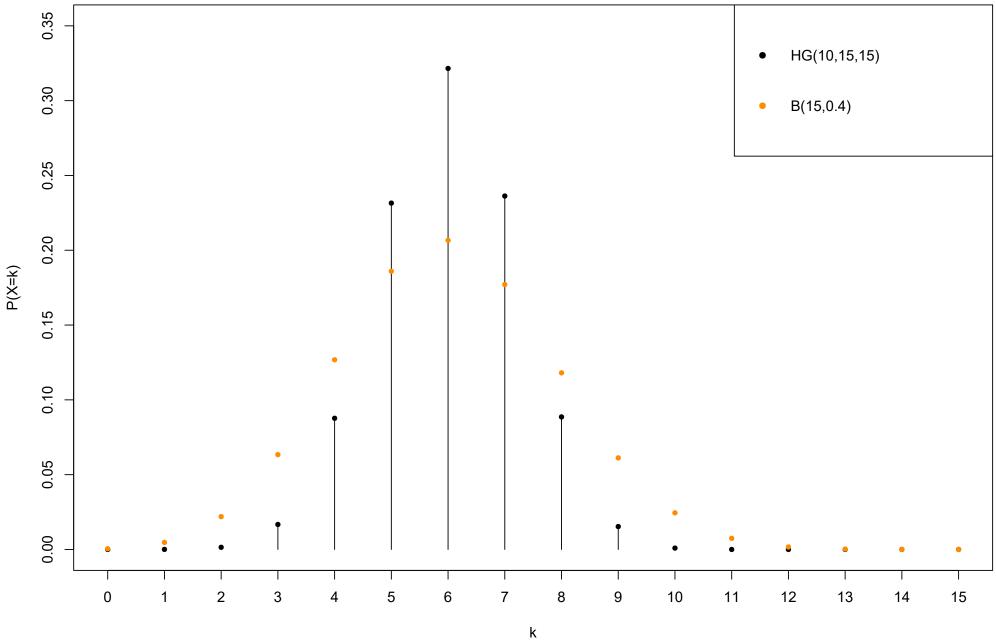
$$\mathbb{1}_{A_1 \cup \cdots \cup A_n} \leqslant \mathbb{1}_{A_1} + \cdots + \mathbb{1}_{A_n}\,.$$

Sada iz linearnosti i monotonosti očekivanja odmah slijedi subaditivnost

$$\mathbb{P}(A_1 \cup \cdots \cup A_n) \leqslant \sum_{i=1}^{n} \mathbb{P}(A_i)\,.$$

[Slično se alternativno može dokazati FUI.]

[7] Jesmo li korištenjem dodatnih indikatora mogli doći do $\mathbb{P}(A_1) = \frac{n}{m_1+m_2}$?



SLIKA 6. Usporedba hipergeometrijske i binomne razdiobe.

3.6. Uvjetne razdiobe

Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ vjerojatnosni prostor i X diskretna slučajna varijbla t.d. $D_X = \{a_i : i \in I\}$. Distribucija od X (s obzirom na $\mathbb{P}$) je $\left(\begin{smallmatrix} a_1 & a_2 & \cdots\cdots \\ p_1 & p_2 & \cdots\cdots \end{smallmatrix}\right)$ za $p_i := \mathbb{P}(X = a_i)$, $i \in I$.

Za događaj $B \in \mathcal{F}$ takav da vrijedi $\mathbb{P}(B) > 0$, **uvjetna distribucija** od X uz dano B je $\begin{pmatrix} a_1 & a_2 & \cdots\cdots \\ p_1^B & p_2^B & \cdots\cdots \end{pmatrix}$ za

$$p_i^B := \mathbb{P}(X = a_i \mid B) = \mathbb{P}_B(X = a_i)\,, \quad i \in I\,.$$

Drugim riječima, to je distribucija od X, ali s obzirom na vjerojatnost $\mathbb{P}_B$! **Uvjetno** očekivanje i varijancu od X uz dano B (oznake $\mathbb{E}[X \mid B]$ i $\mathrm{Var}(X \mid B)$) definiramo jednostavno kao očekivanje i varijancu od X s obzirom na vjerojatnost $\mathbb{P}_B$; na primjer, ako je $\mathbb{P}(X \geqslant 0 \mid B) = 1$,

$$\mathbb{E}[X \mid B] := \sum_{i \in I} a_i \mathbb{P}(X = a_1 \mid B) \in [0, \infty]\,.$$

TEOREM 3.48 (**Formula potpunog očekivanja**). *Neka je $(H_j)_{j \in J}$ PSD (t.d. vrijedi $\mathbb{P}(H_j) > 0$, $\forall j \in J$) te X diskretna slučajna varijabla takva da vrijedi (a) $X \geqslant 0$, ili (b) $\mathbb{E}\big[|X|\big] < \infty$. Tada vrijedi*

$$\mathbb{E}[X] = \sum_{j \in J} \mathbb{P}(H_j) \mathbb{E}[X \mid H_j]\,. \tag{3.21}$$

DOKAZ. Neka je $D_X = \{a_i : i \in I\}$. (a) Ako je $X \geqslant 0$, nužno je $\mathbb{P}(X \geqslant 0 \mid H_j) = 1$, $\forall j \in J$, te je

$$
\begin{aligned}
\mathbb{E}[X] &= \sum_{i\in I} a_i \mathbb{P}(X = a_i) \overset{\text{FPV}}{=} \sum_{i\in I} a_i \sum_{j\in J} \mathbb{P}(H_j)\mathbb{P}(X = a_i \mid H_j) \\
&= [\text{Fubini (sve je nenegativno)}] = \sum_{j\in J} \mathbb{P}(H_j) \sum_{i\in I} a_i \mathbb{P}(X = a_i \mid H_j) \\
&= \sum_{j\in J} \mathbb{P}(H_j)\mathbb{E}[X \mid H_j] \in [0, \infty] \,. \qquad (3.22)
\end{aligned}
$$

(b) Budući da je $|X| \geqslant 0$, iz (3.22) dobivamo

$$+\infty > \mathbb{E}\Big[|X|\Big] = \sum_{j\in J} \mathbb{P}(H_j)\mathbb{E}\Big[|X| \mid H_j\Big] \,.$$

Specijalno, nužno je $\mathbb{E}\Big[|X| \mid H_j\Big] < \infty$, $\forall j$, tj. postoji $\mathbb{E}[X \mid H_j]$, te (3.21) slijedi analogno kao u (3.22) [pri čemu Fubinijev teorem možemo koristiti jer vrijedi

$$\sum_{j\in J} \mathbb{P}(H_j) \sum_{i\in I} |a_i| \mathbb{P}(X = a_i \mid H_j) = \sum_{j\in J} \mathbb{P}(H_j)\mathbb{E}\Big[|X| \mid H_j\Big] < \infty \,. \;]$$

$\square$

PRIMJER 3.49. Bacamo novčić na kojem je vjeorjatnost za glavu jednaka $p \in (0, 1)$. Ako je R duljina prvog niza (to može biti niz glava ili niz pisama), odredite $\mathbb{E}[R]$.

RJEŠENJE. Neka je $q := 1 - p$, $H_G = \{\text{u prvom bacanju pala glava}\}$, $H_P := H_G^c$. Budući da je $R \geqslant 1 \geqslant 0$, (3.21) nam daje

$$\mathbb{E}[R] = \mathbb{P}(H_G)\mathbb{E}[R \mid H_G] + \mathbb{P}(H_P)\mathbb{E}[X \mid H_P] = p\mathbb{E}[R \mid H_G] + q\mathbb{E}[X \mid H_P] \,.$$

Uvjetno na H_G, prvi niz je dakle niz glava, te R s obzirom na $\mathbb{P}_{H_1}$ ima istu razdiobu kao sl. varijabla $1 + X$ gdje je $X \sim \mathrm{G}_0(q)$. Dakle,

$$\mathbb{E}[R \mid H_G] = \mathbb{E}[1 + X] = 1 + \frac{p}{q} = \frac{1}{q} \,.$$

Analogno, $\mathbb{E}[R \mid H_P] = 1 + \frac{q}{p} = \frac{1}{p}$. Dakle,

$$\mathbb{E}[R] = p \cdot \frac{1}{q} + q \cdot \frac{1}{p} = \frac{p}{q} + \frac{q}{p} \,.$$

$\square$

ZADATAK. Ako je R_n duljina n-tog niza (dakle, $R_1 = R$), pokažite da vrijedi

(i) $\mathbb{E}[R_{2k-1}] = \mathbb{E}[R_1]$, $\mathbb{E}[R_{2k}] = 2$, za sve $k \geqslant 1$, te

(ii) vrijedi $\mathbb{E}[R_1] \geqslant 2$ uz jednakost akko $p = q = \frac{1}{2}$.

3.7. Očekivanje na diskretnom vjerojatnosnom prostoru

[Ključna svojstva očekivanja su linearnost i monotonost. Cilj ovog poglavlja je dati intuiciju zašto očekivanje ima ta svojstva.]

Ako je $(\Omega, \mathcal{F}, \mathbb{P})$ diskretan vjerojatnosni prostor (tj. Ω najviše prebrojiv i $\mathcal{F} = \mathcal{P}(\Omega)$), svaka slučajna varijabla $X : \Omega \to \mathbb{R}$ je nužno diskretna [jer je D_X najviše prebrojiv skup].

PROPOZICIJA 3.50. *Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ diskretan vjerojatnosni prostor te $X : \Omega \to \mathbb{R}$ slučajna varijabla na njemu. Ako je (a) $X \geqslant 0$, ili (b) $\mathbb{E}\Big[|X|\Big] < \infty$, vrijedi*

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\omega) . \tag{3.23}$$

DOKAZ. Dokaz je jako sličan dokazu Teorema 3.21 – tvrdnja slijedi jer je $\sum_{i \in I} a_i \mathbb{P}(X = a_i)$ samo alternativan način sumiranja članova $X(\omega)\mathbb{P}(\omega), \omega \in \Omega$. Pogledati sami.

[Neka je $D_X = \{a_i : i \in I\}$ te $I_i := \{\omega \in \Omega : X(\omega) = a_i\} = \{X = a_i\}$, $i \in I$.

(a) Ako je $X \geqslant 0$, budući da je svaki I_i najviše prebrojiv skup, vrijedi

$$\begin{aligned}\mathbb{E}[X] &= \sum_{i \in I} a_i \mathbb{P}(X = a_i) = \sum_{i \in I} a_i \sum_{\omega \in I_i} \mathbb{P}(\omega) = \sum_{i \in I} \sum_{\omega \in I_i} X(\omega)\mathbb{P}(\omega) \\ &= [\text{Fubini} + (I_i)_{i \in I} \text{ je particija skupa } \Omega] = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\omega) \in [0, \infty] .\end{aligned}$$

(b) Ako je $\mathbb{E}\Big[|X|\Big] < \infty$, iz (a) dijela slijedi da je i $\sum_{\omega \in \Omega} |X(\omega)|\mathbb{P}(\omega) < \infty$, pa analogno kao u (a) dijelu uz upotrebnu Fubinijevog teorema zaključujemo da vrijedi (3.23).]

□

[Sada dajemo potpuni dokaz linearnosti i monotonosti očekivanja na diskretnom vjerojatnosnom prostoru.]

KOROLAR 3.51. *Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ diskretan vjerojatnosni prostor te $X, Y : \Omega \to \mathbb{R}$ slučajne varijable na njemu. Tada vrijedi*

(i) ako je $0 \leqslant X \leqslant Y$, nužno je $\mathbb{E}[X] \leqslant \mathbb{E}[Y]$ [ovdje nema nikakvih restrikcija na konačnost očekivanja].

(ii) ako je (a) $X, Y \geqslant 0$, ili (b) $\mathbb{E}\Big[|X|\Big], \mathbb{E}\Big[|Y|\Big] < \infty$, vrijedi

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y] .$$

DOKAZ. (i) Vrijedi

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} \overbrace{X(\omega)}^{\leqslant Y(\omega), \forall \omega} \mathbb{P}(\omega) \leqslant \sum_{\omega \in \Omega} Y(\omega)\mathbb{P}(\omega) = \mathbb{E}[Y] .$$

(ii) (a) Ako su $X, Y \geqslant 0$, vrijedi

$$\mathbb{E}[X + Y] = \sum_{\omega \in \Omega} \overbrace{(X + Y)(\omega)}^{= X(\omega) + Y(\omega), \forall \omega} \mathbb{P}(\omega) = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\omega) + \sum_{\omega \in \Omega} Y(\omega)\mathbb{P}(\omega) = \mathbb{E}[X] + \mathbb{E}[Y] .$$

(b) Ako je $\mathbb{E}\Big[|X|\Big], \mathbb{E}\Big[|Y|\Big] < \infty$, monotonost i linearnost iz (b) dijela povlače

$$\mathbb{E}\Big[|X + Y|\Big] \leqslant \mathbb{E}\Big[|X| + |Y|\Big] = \mathbb{E}\Big[|X|\Big] + \mathbb{E}\Big[|Y|\Big] < \infty .$$

Dakle, postoji $\mathbb{E}[X+Y]$ te se $\mathbb{E}[X+Y]=\mathbb{E}[X]+\mathbb{E}[Y]$ pokaže analogno kao u (a) dijelu. $\square$

NAPOMENA. Na općenitom vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ očekivanje se definira kao $\mathbb{E}[X] = \int_{\mathbb{R}} X(\omega)\mathbb{P}(\mathrm{d}\omega)$ gdje je ovo tzv. *Lebesgueov integral*, a onda linearnost i monotonost očekivanja slijede direktno iz svojstava tog intervala. U diskretnom slučaju je upravo $\int_{\mathbb{R}} X(\omega)\mathbb{P}(\mathrm{d}\omega) = \sum_{\omega\in\Omega} X(\omega)\mathbb{P}(\omega)$.

POGLAVLJE 4

Diskretni slučajni vektori – (ne)zavisnost

[Za primjene npr. u statistici i/ili strojnom učenju ključno je promatrati više slučajnih varijabli koje su ishodi **istog** vjerojatnosnog modela, tj. koje su definirane na **istom** vjerojatnosnom prostoru. Na primjer, ako su

- Y indikator događaja da osoba koju promatramo ima ili nema određenu bolest ($Y \in \{0, 1\}$), te
- za $p \in \mathbb{N}$, $X_1, \ldots, X_p$ razna mjerenja (tlak, puls i slično) za tu istu osobu,

može nas zanimati kako određene kombinacije mjerenja $X_1, \ldots, X_p$ utječu na vjerojatnost da je $Y = 1$. Uočite, iako u samom problemu nema nikakve slučajnosti, u statistici pretpostavljamo da su $X_1, \ldots, X_p$ i Y slučajne varijable, te gornji problem postaje problem određivanja **zajedničke razdiobe** slučajnih varijabli $X_1, \ldots, X_p, Y$, tj. **slučajnog vektora** $(X_1, \ldots, X_p, Y)$.]

4.1. Zajednička distribucija

Primjer 4.1. Bacamo 2 simetrična novčića te neka je

$$A_i := \{\text{na } i\text{-tom novčiću palo pismo}\}\,, \quad i = 1, 2\,.$$

Tada su $X := \mathbb{1}_{A_1}$ i $Y := \mathbb{1}_{A_2}$ slučajne varijable definirane na **istom** vjer. prostoru $(\Omega, \mathcal{F}, \mathbb{P})$, te obje imaju B(1/2) razdiobu. Ako za sve $\omega \in \Omega$, definiramo

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

funkcija $(X, Y) : \Omega \to \mathbb{R}^2$ je primjer **slučajnog vektora** čija je distribucija:

- $(X, Y)(\omega) \in \{(0, 0), (0, 1), (1, 0), (1, 1)\}$, $\forall \omega \in \Omega$, te
- $\mathbb{P}((X, Y) = (a, b)) = \mathbb{P}(X = a, Y = b) = \frac{1}{4}$, za sve $a, b \in \{0, 1\}$,

pri čemu događaji $\{(X, Y) = (a, b)\}$ i $\{X = a, Y = b\}$ predstavljaju događaj

$$\{\omega \in \Omega : (X, Y)(\omega) = (a, b)\} = \{\omega \in \Omega : X(\omega) = a\,,\ Y(\omega) = b\}\,.$$

Gornju distribuciju možemo zapisati tablično:

$X\backslash Y$	0	1
0	$\frac{1}{4}$	$\frac{1}{4}$
1	$\frac{1}{4}$	$\frac{1}{4}$

S druge strane, ako definiramo

$$Z := 1 - X = \mathbb{1}_{\{\text{na prvom novčiću pala } \textbf{glava}\}}\,,$$

opet vrijedi $Z \sim \mathrm{B}(1/2)$, ali razdioba vektora (X, Z) je drugačija:

$X \backslash Z$	0	1
0	0	$\frac{1}{2}$
1	$\frac{1}{2}$	0

[Dakle, distribucije koordinata **ne** određuju zajedničku distribuciju!]

Definicija 4.2. Ako je $d \in \mathbb{N}$ te $X_1, \dots, X_d$ slučajne varijable na vj. prostoru $(\Omega, \mathcal{F}, \mathbb{P})$, funkciju $(X_1, \dots, X_d) : \Omega \to \mathbb{R}^d$ zovemo **(d-dimenzionalan) slučajni vektor**. 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 $D_i := D_{X_i} = \mathrm{Im}(X_i)$, on poprima vrijednosti u skupu $D_1 \times \cdots \times D_d =: D$, koji je ponovno najviše prebrojiv. Skup D zajedno s vjerojatnostima

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

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 $f_{X_1,\dots,X_d} = f : \mathbb{R}^d \to [0, 1]$, 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 $(X_1, \dots, X_d)$.

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

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

[Sljedeća svojstva dokazuju se slično kao u slučaju $d = 1$, te se lako proširuju na slučaj $d \geqslant 3$ – oboje ostavljamo za vježbu.]

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

$$\mathbb{P}((X_1, X_2) \in C) = \textstyle\sum_{i,j:(a_i,b_j)\in C} \mathbb{P}(X_1 = a_i, X_2 = b_j) = \sum_{(a_i,b_j)\in C} p_{ij}\,. \tag{4.1}$$

  [Specijalno,

$$\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} = \sum_{(a_i,b_j)\in D} p_{ij} = \mathbb{P}((X_1, X_2) \in D) = 1\,.$$

  Nadalje, potpuno isto kao za slučajne varijable (tj. slučaj $d = 1$), pokaže se iduća tvrdnja: Za proizvoljan prebrojiv skup $D \subseteq \mathbb{R}^2$ i distribuciju $(p_{ij} : (i, j) \in D)$ na D[1], postoji vjerojatnosni prostor $(\Omega, \mathcal{F}, \mathbb{P})$ i slučajni vektor $(X_1, X_2) : \Omega \to D$ takav da vrijedi $\mathbb{P}(X_1 = i, X_2 = j) = p_{ij}$ za sve $(i, j) \in D$.]

[1] Dakle, (i) $p_{ij} \geqslant 0$, za sve i, j, te (ii) $\sum_{(i,j)\in D} p_{ij} = 1$.

- Iz zajedničke razdiobe vektora (X_1, X_2) lako dobijemo tzv. **marginalne** razdiobe:

$$\mathbb{P}(X_1 = a_i) = \mathbb{P}((X_1, X_2) \in \{a_i\} \times D_2) \stackrel{(4.1)}{=} \sum_{b_j \in D_2} p_{ij} =: p_i^{(1)}\,,\ i \in I$$

$$\mathbb{P}(X_2 = b_j) = \cdots = \sum_{a_i \in D_1} p_{ij} =: p_j^{(2)}\,,\ j \in J\,.$$

- Ako je $g : D \to \mathbb{R}^k$ proizvoljna funkcija, tada je $g(X_1, X_2) : \Omega \to \mathbb{R}^k$ slučajni vektor koja poprima vrijednosti u skupu $g(D)$ te vrijedi

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

Ako je dodatno $D_1 = \{a_1, \dots, a_n\}$, $D_2 = \{b_1, \dots, b_m\}$, distribuciju od (X_1, X_2) možemo zapisati tablično:

$X_1 \backslash X_2$	b_1	b_2	$\cdots$	b_j	$\cdots$	b_m	$\mathbb{P}(X_1 = \cdot)$
a_1				$\vdots$			$\vdots$
a_2				$\vdots$			$\vdots$
$\vdots$				$\vdots$			$\vdots$
a_i	$\cdots$	$\cdots$	$\cdots$	p_{ij}	$\cdots$	$\cdots$	$p_i^{(1)}$
$\vdots$				$\vdots$			$\vdots$
a_n				$\vdots$			$\vdots$
$\mathbb{P}(X_2 = \cdot)$	$\cdots$	$\cdots$	$\cdots$	$p_j^{(2)}$	$\cdots$	$\cdots$	1

Primjer 4.3. U kutiji su 2 crne, 5 bijelih i 3 zelene kuglice. Slučajno izvučemo 3 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) $\mathbb{P}(X \geqslant Y)$;

(d) razdiobu slučajne varijable $X + Y$ [=ukupan broj crnih i bijelih kuglica].

Rješenje. (a) Sl. varijabla X (ond. Y) poprima vrijednosti u skupu $\{0, 1, 2\}$ (odn. $\{0, 1, 2, 3\}$). Vrijedi

$$\mathbb{P}(X = 0, Y = 0) = \mathbb{P}(\{\text{izvukli samo zelene kuglice}\}) = \frac{\binom{3}{3}}{\binom{10}{3}} = \frac{1}{120}\,,$$

$$\mathbb{P}(X = 0, Y = 1) = \mathbb{P}(\{\text{1 bijela i 2 zelene}\}) = \frac{\binom{5}{1} \cdot \binom{3}{2}}{\binom{10}{3}} = \frac{5}{120}\,,$$

i tako dalje. Sličnim računom dobijemo da je distribucija

$X\backslash Y$	0	1	2	3	$\mathbb{P}(X=\cdot)$
0	$\frac{1}{120}$	$\frac{15}{120}$	$\frac{30}{120}$	$\frac{10}{120}$	$\frac{56}{120}$
1	$\frac{6}{120}$	$\frac{30}{120}$	$\frac{20}{120}$	0	$\frac{56}{120}$
2	$\frac{3}{120}$	$\frac{5}{120}$	0	0	$\frac{8}{120}$
$\mathbb{P}(Y=\cdot)$	$\frac{10}{120}$	$\frac{50}{120}$	$\frac{50}{120}$	$\frac{10}{120}$	1

(b) Iz tablice u (a) dijelu odmah slijedi

$$X \sim \begin{pmatrix} 0 & 1 & 2 \\ \frac{56}{120} & \frac{56}{120} & \frac{8}{120} \end{pmatrix}, \quad Y \sim \begin{pmatrix} 0 & 1 & 2 & 3 \\ \frac{10}{120} & \frac{50}{120} & \frac{50}{120} & \frac{10}{120} \end{pmatrix}$$

(c) Ako primijenimo (4.1) uz $C = \{(a,b) : a \geqslant b\}$, dobivamo

$$\mathbb{P}(X \geqslant Y) = \sum_{a=0}^{2}\sum_{b=0}^{a} \mathbb{P}(X=a, Y=b) = \frac{1}{120}\cdot(1+6+30+3+5+0) = \frac{3}{8}\,.$$

(d) Budući da je $X+Y = g(X,Y)$, lako se vidi da je $\mathbb{P}(X+Y \in \{0,1,2,3\}) = 1$ [nacrtati tablicu], a iz (4.2) slijedi

$$\mathbb{P}(X+Y=0) = \mathbb{P}(X=0, Y=0) = \frac{1}{120}\,,$$
$$\mathbb{P}(X+Y=1) = \mathbb{P}(X=1, Y=0) + \mathbb{P}(X=0, Y=1) = \frac{6}{120} + \frac{15}{120} = \frac{21}{120}\,,$$

itd. Analognim računom dobijemo da je distribucija

$$X+Y \sim \begin{pmatrix} 0 & 1 & 2 & 3 \\ \frac{1}{120} & \frac{21}{120} & \frac{63}{120} & \frac{35}{120} \end{pmatrix}$$

□

4.2. Nezavisnost

Definicija 4.4. Neka su $X_1, \ldots, X_d$ diskretne slučajne varijable definirane na $(\Omega, \mathcal{F}, \mathbb{P})$ te neka je $D_i := D_{X_i}$, $i = 1, \ldots, d$. Kažemo da su $X_1, \ldots, X_d$ **nezavisne** ako vrijedi

$$\mathbb{P}(X_1 = x_1, \ldots, X_d = x_d) = \mathbb{P}(X_1 = x_1)\cdots\mathbb{P}(X_d = x_d)\,, \tag{4.3}$$

za sve $x_1, \ldots, x_d$ takve da je $x_i \in D_i$, $i = 1, \ldots, d$.

[Napomena.

- Uočite razliku u odnosu na definiciju nezavisnosti događaja – ovdje ne gledamo sve moguće podskupove od $\{1, \ldots, d\}$, ali gledamo sve moguće $x_1, \ldots, x_d$.
- Iz (4.3) odmah slijedi da kada su varijable $X_1, \ldots, X_d$ nezavisne, njihove marginalne razdiobe jedinstveno određuju distribuciju vektora $(X_1, \ldots, X_d)$.
- Uvjetna nezavisnost (s obzirom na događaj A t.d. $\mathbb{P}(A) > 0$) definira se kao nezavisnost s obzirom na vjerojatnost $\mathbb{P}_A$.]

PRIMJER 4.5. Varijable X i Y iz Primjera 4.3 *nisu* nezavisne [tj. zavisne su] jer npr. vrijedi

$$\mathbb{P}(X = 2, Y = 2) = 0 \neq \mathbb{P}(X = 2)\mathbb{P}(Y = 2) > 0\,.$$

[Je li to intuitivno jasno?]

PRIMJER 4.6. Bacamo simetričnu kocku n puta ne neka je X_i rezultat na i-toj kocki. Tada za sve $x_1, \ldots, x_n \in \{1, \ldots, 6\}$ vrijedi

$$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) = \frac{1}{6^n} = \left(\frac{1}{6}\right)^n = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i)\,,$$

tj. $X_1, \ldots, X_n$ su nezavisne [kao što bismo i htjeli da budu].

[U nastavku u ovom ili sličnim slučajevima, nezavisnost ne provjeravamo, tj. pretpostavljamo da "očito" vrijedi.]

NAPOMENA 4.7. Dokaze sljedećih tvrdnji ostavljamo za vježbu.

(a) Slučajne varijable $X_1, \ldots, X_d$ su nezavisne akko za njihovu diskretnu funkciju gustoće vrijedi

$$f_{X_1,\ldots,X_d}(x_1, \ldots, x_d) = f_{X_1}(x_1) \cdots f_{X_d}(x_d)\,,\ \forall x_1, \ldots, x_d \in \mathbb{R}\,.$$

(b) Događaji $A_1, \ldots, A_d$ su nezavisni akko su $\mathbb{1}_{A_1}, \ldots, \mathbb{1}_{A_d}$ nezavisne slučajne varijable.

PRIMJER 4.8 (**Metoda maksimalne vjerodostojnosti**[2]). [Pretpostavimo da želimo procijeniti postotak populacije koji podržava stranku ABC; označimo taj (nepoznat) postotak s $p_0 \in [0, 1]$. U tu svrhu uzmemo *slučajan uzorak* od n ljudi te stavimo

$$X_i := \mathbb{1}_{\{i\text{-ta osoba podržava stranku ABC}\}}\,.\,]$$

- Radi jednostavnosti, pretpostavimo da su $X_1, \ldots, X_n$ nezavisne te da sve imaju $\mathrm{B}(p_0)$ distribuciju; kraće kažemo da su $X_1, \ldots, X_n$ **nezavisne i jednako distribuirane** ("njd") sa zajedničkom distribucijom $\mathrm{B}(p_0)$.
- Ako smo dobili rezultate $X_i(\omega) = x_i$ za konkretne $x_i \in \{0, 1\}$, $i = 1, \ldots, n$, **procjenitelj maksimalne vjerodostojnosti** za p_0 je

  $$\widehat{p}_n := \arg\max_{p \in [0,1]} L(p)\,,$$

  gdje je

  $$L(p) := f_p(x_1, \ldots, x_n) := f_{X_1,\ldots,X_n}(x_1, \ldots, x_n; p)\,,$$

  tzv. **vjerodostojnost** podataka $x_1, \ldots, x_n$ za parametar p; f_p je dakle oznaka za gustoću vektora $(X_1, \ldots, X_n)$ ukoliko su oni njd s distribucijom $\mathrm{B}(p)$.[3]

[2] engl. *maximum likelihood*
[3] Intuitivno, očekujemo da će vjerodostojnost biti velika za p-ove koji su blizu p_0.

- Budući da je

$$f_{X_i}(x_i;p) = \begin{cases} p\,, & \text{ako je } x_i = 1, \\ 1-p\,, & \text{ako je } x_i = 0 \end{cases} = p^{x_i}(1-p)^{1-x_i}\,,$$

zbog nezavisnosti vrijedi

$$L(p) = \prod_{i=1}^{n} f_{X_i}(x_i;p) = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i} =: p^s(1-p)^{n-s}\,.$$

Budući da nas zanima samo koji p maksimizira $L(p)$, dovoljno je naći maksimizator tzv. **log-vjerodostojnosti**

$$l(p) := \log L(p) = s\log(p) + (n-s)\log(1-p)\,,$$

te se lako pokaže (DZ) da vrijedi

$$\widehat{p}_n := \arg\max_{p\in[0,1]} l(p) = \frac{s}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i\,.$$

[Što je u ovom slučaju i prirodan procjenitelj. Ipak, metoda maksimalne vjerodostojnosti je vrlo općenita te iza nje postoji značajna teorija.]

□

[Za općenite slučajne varijable $X_1,\ldots,X_d$, nezavisnost se definira kao u (4.4) dolje, a sljedeći rezultat govori da se u slučaju diskretnih slučajnih varaijabli to svodi na (jednostavniji) uvjet (4.3).]

PROPOZICIJA 4.9. *Diskretne slučajne varijable $X_1,\ldots,X_d$ su nezavisne akko vrijedi*

$$\mathbb{P}(X_1\in B_1,\ldots,X_d\in B_d) = \mathbb{P}(X_1\in B_1)\cdots\mathbb{P}(X_d\in B_d)\,, \tag{4.4}$$

za sve $B_i\subseteq\mathbb{R}$, $i=1,\ldots,d$.

DOKAZ. "$\Leftarrow$" Slijedi odmah ako uzmemo $B_i := \{x_i\}$, $i=1,\ldots,d$.

"$\Rightarrow$" Vrijedi

$$\begin{aligned}\mathbb{P}(X_1\in B_1,\ldots,X_d\in B_d) &= \mathbb{P}((X_1,\ldots,X_d)\in B_1\times\cdots\times B_d)\\ &\overset{(4.1)}{=} \textstyle\sum_{x_i\in B_i\cap D_i,\forall i}\mathbb{P}(X_1=x_1,\ldots,X_d=x_d)\\ &= [\text{nezavisnost, tj. (4.3)}] = \textstyle\sum_{x_i\in B_i\cap D_i,\forall i}\mathbb{P}(X_1=x_1)\cdots\mathbb{P}(X_d=x_d)\\ &= \left(\textstyle\sum_{x_1\in B_1\cap D_1}\mathbb{P}(X_1=x_1)\right)\cdots\left(\textstyle\sum_{x_d\in B_d\cap D_d}\mathbb{P}(X_d=x_d)\right)\\ &\overset{(4.1)}{=} \mathbb{P}(X_1\in B_1)\cdots\mathbb{P}(X_d\in B_d)\,.\end{aligned}$$

□

NAPOMENA 4.10. Neka su $X_1,\ldots,X_d$ nezavisne. Dokaze sljedećih tvrdnji ostavljamo za vježbu.

(a) Za svaki podskup $F \subseteq \{1, \dots, d\}$, $|F| \geqslant 2$, slučajne varijable $(X_i : i \in F)$ su također nezavisne; u (4.4) stavi $B_i := \mathbb{R}$ za $i \neq F$). [Na primjer, za svaki $i \neq j$, X_i i X_j su nezavisne slučajne varijable.]

(b) Za sve $m \in \{1, \dots, d\}$ i sve $A \subseteq \mathbb{R}^m, B \subseteq \mathbb{R}^{d-m}$ vrijedi

$$\begin{aligned}&\mathbb{P}((X_1, \dots, X_m) \in A, (X_{m+1}, \dots, X_d) \in B) \\ &\qquad = \mathbb{P}((X_1, \dots, X_m) \in A)\mathbb{P}((X_{m+1}, \dots, X_d) \in B)\,, \end{aligned} \tag{4.5}$$

te kažemo da su vektori $(X_1, \dots, X_m)$ i $(X_{m+1}, \dots, X_d)$ nezavisni. Specijalno, za sve funkcije $g : \mathbb{R}^m \to \mathbb{R}, h : \mathbb{R}^{d-m} \to \mathbb{R}$, slučajne varijable $g(X_1, \dots, X_m)$ i $h(X_{m+1}, \dots, X_d)$ su nezavisne.

Primjer 4.11. Neka su $p_1, p_2 \in (0, 1]$, te $X \sim \mathrm{G}(p_1)$ i $Y \sim \mathrm{G}(p_2)$ nezavisne. Odredite distribuciju slučajne varijable $Z := \min\{X, Y\}$.

Rješenje. Budući da su $X, Y \in \mathbb{N}$, očito isto vrijedi i za Z. Uočimo, za $n \in \mathbb{N}_0$,

$$\begin{aligned}\mathbb{P}(Z > n) &= \mathbb{P}(X > n, Y > n) = [\mathbb{P}(X \in \{n, n+1, \dots\}, Y \in \{n, n+1, \dots\})] \\ &= [\text{nezavisnost, tj. } (4.4)] = \mathbb{P}(X > n)\mathbb{P}(Y > n) \\ &= [X \sim \mathrm{G}(p_1), Y \sim \mathrm{G}(p_2)] = q_1^n q_2^n =: q^n\,,\end{aligned}$$

gdje je $q_i := 1 - p_i$ i $q := q_1 q_2$. Sada odmah slijedi da je $Z \sim \mathrm{G}(1 - q)$ (DZ).

[Zaista, za $n \in \mathbb{N}$,

$$\mathbb{P}(Z = n) = \mathbb{P}(Z > n - 1) - \mathbb{P}(Z > n) = q^{n-1} - q^n = q^{n-1}(1 - q)\,.\,]$$

Alternativno (i bolje), pretpostavimo da imamo dva nezavisna niza nezavisnih pokusa t.d. je vjerojatnost uspjeha u svakom pokusu niza i jednaka baš p_i. Ako stavimo

$$X_n := \mathbb{1}_{\{\text{uspjeh u } n\text{-tom pokusu 1. niza}\}}\,, \quad Y_n := \mathbb{1}_{\{\text{uspjeh u } n\text{-tom pokusu 2. niza}\}}\,,$$

te definiramo

$$X := \min\{n \geqslant 1 : X_n = 1\}\,,\ Y := \min\{n \geqslant 1 : Y_n = 1\}$$

vrijedi $X \sim \mathrm{G}(p_1)$ i $Y \sim \mathrm{G}(p_2)$, te su one nezavisne [nacrtati sliku]. Nadalje, budući da je

$$Z = \min\{X, Y\} = \min\{n \geqslant 1 : X_n = 1 \text{ ili } Y_n = 1\}\,,$$

slijedi da je i $Z \sim \mathrm{G}(p)$ uz parametar

$$\begin{aligned}p &= \mathbb{P}(X_n = 1 \text{ ili } Y_n = 1) = 1 - \mathbb{P}(X_n = 0, Y_n = 0) = [\text{nezavisnost}] \\ &= 1 - \mathbb{P}(X_n = 0)\mathbb{P}(Y_n = 0) = 1 - q_1 q_2 = 1 - q\,.\end{aligned}$$

□

Napomena 4.12. Formalno, za niz slučajnih varijabli $X_1, X_2, \dots,$ kažemo

(i) da su nezavisne, ako su $X_1,\dots,X_n$ nezavisne za svaki $n\in\mathbb{N}$;

(ii) da su nezavisne od drugog niza $Y_1,Y_2,\dots,$ ako su vektori $(X_1,\dots,X_n)$ i $(Y_1,\dots,Y_m)$ nezavisni za sve $n,m\in\mathbb{N}$ (u smislu (4.5)).

4.3. Linearnost i monotonost očekivanja

[Sljedeći teorem je generalizacija Teorema 3.21 (slučaj $d=1$), te se dokazuje potpuno analogno. Koristeći donji rezultat napokon možemo formalno pokazati da je očekivanje linearno i monotono.]

Teorem 4.13. *Neka su $X_1,\dots,X_d$ diskretne slučajne variable na $(\Omega,\mathcal{F},\mathbb{P})$, te $D_i := D_{X_i}$, $i=1,\dots,d$, a $g:\mathbb{R}^d\to\mathbb{R}$ funkcija. Tada vrijedi*

$$\mathbb{E}[g(X_1,\dots,X_d)] = \textstyle\sum_{x_i\in D_i,\forall i} g(x_1,\dots,x_d)\mathbb{P}(X_1=x_1,\dots,X_d=x_d)\,, \tag{4.6}$$

ukoliko (a) $g(X_1,\dots,X_d)\geqslant 0$, ili (b) red na desnoj strani od (4.6) apsolutno konvergentan.

Teorem 4.14. *Ako su $X_1,\dots,X_n$ diskretne slučajne variable na $(\Omega,\mathcal{F},\mathbb{P})$ t.d. vrijedi (a) $X_i\geqslant 0$, $\forall i$, ili (b) $\mathbb{E}\big[|X_i|\big]<\infty$, $\forall i$, tada je*

$$\mathbb{E}\left[\sum_{i=1}^n \alpha_i X_i\right] = \sum_{i=1}^n \alpha_i \mathbb{E}[X_i]\,, \tag{4.7}$$

za sve $\alpha_1,\dots,\alpha_n\in\mathbb{R}$, tj. $\alpha_1,\dots,\alpha_n\geqslant 0$ u slučaju (a).

Dokaz. Neka je $n=2$, te $D_1=\{a_i : i\in I\}$, $D_2=\{b_j : j\in J\}$; slučaj $n\in\mathbb{N}$ dokazuje se lagano matematičkom indukcijom.

(a) Ako je $X_1,X_2\geqslant 0$ i $\alpha_1,\alpha_2\geqslant 0$, budući da je $\alpha_1X_1+\alpha_2X_2=g(X_1,X_2)\geqslant 0$, (4.6) nam daje

$$\begin{aligned}\mathbb{E}[\alpha_1X_1+\alpha_2X_2] &= \sum_{i\in I}\sum_{j\in J}(\alpha_1a_i+\alpha_2b_j)\mathbb{P}(X_1=a_i,X_2=b_j)\\ &= [\text{Fubini}] = \alpha_1\sum_{i\in I}a_i\underbrace{\sum_{j\in J}\mathbb{P}(X_1=a_i,X_2=b_j)}_{=\mathbb{P}(X_1=a_i)}+\alpha_2\sum_{j\in J}b_j\underbrace{\sum_{i\in I}\mathbb{P}(X_1=a_i,X_2=b_j)}_{=\mathbb{P}(X_2=b_j)}\\ &= \alpha_1\sum_{i\in I}a_i\mathbb{P}(X_1=a_i)+\alpha_2\sum_{j\in J}b_j\mathbb{P}(X_2=b_j) = \alpha_1\mathbb{E}[X_1]+\alpha_2\mathbb{E}[X_2]\in[0,\infty]\,.\end{aligned}$$

(b) Ako su $X_1,X_2\in\mathbb{R}$ te $\alpha_1,\alpha_2\in\mathbb{R}$, monotonost očekivanja (vidi Teorem 4.17 dolje) povlači

$$\begin{aligned}\mathbb{E}\Big[|\alpha_1X_1+\alpha_2X_2|\Big] &\leqslant \mathbb{E}\Big[|\alpha_1||X_1|+|\alpha_2||X_2|\Big]\\ &= [\text{(a) dio}] = |\alpha_1|\cdot\mathbb{E}\Big[|X_1|\Big]+|\alpha_2|\cdot\Big[|X_2|\Big]<\infty\,.\end{aligned}$$

Dakle, postoji $\mathbb{E}[\alpha_1X_1+\alpha_2X_2]$ te se (4.7) dokazuje analogno kao u (a) dijelu. □

Napomena 4.15. Jednakost (4.7) vrijedi **bez obzira** na zavisnost među X_i-evima.

PRIMJER 4.16 (**"Coupon collector's problem"**). Postoji n različitih kupona. Izvlačimo kupone na slučajan način *s vraćanjem* sve dok ne izvučemo svaki od kupona barem jednom. Ako je T ukupan broj izvlačenja, odredite $\mathbb{E}[T]$.

RJEŠENJE. [Prva ideja bila bi tražiti distribuciju slučajne varijable T, ali to nije najsretniji pristup.]

Neka je $T_1 := 1$ broj izvlačenja potreban do prvog kupona kojeg nismo vidjeli, T_2 *dodatni* broj izvlačenja do drugog kupona kojeg nismo vidjeli itd. [nacrtati primjer]. Tada očito vrijedi $T = T_1 + \cdots + T_n$, pri čemu [zbog *nezavisnosti izvlačenja*] vrijedi $T_i \sim \mathrm{G}(p_i)$ uz

$$p_i = \frac{n-(i-1)}{n} = \frac{n-i+1}{n}\,,\; i = 1, \ldots, n\,,$$

jer nakon što smo izvukli $i-1$ različitih kupona, preostaje $n-(i-1)$ kupona koje još nismo izvukli. [Dodatno su $T_1, \ldots, T_n$ i nezavisne, ali to nam za očekivanje nije bitno.] Linearnost očekivanja nam sada odmah daje

$$\mathbb{E}[T] = \mathbb{E}\left[\sum_{i=1}^{n} T_i\right] = \sum_{i=1}^{n} \mathbb{E}[T_i] = \sum_{i=1}^{n} \frac{1}{p_i} = \sum_{i=1}^{n} \frac{n}{n-i+1} = n \sum_{k=1}^{n} \frac{1}{k}\,.$$

Npr. za $n = 50$, $\mathbb{E}[T] = 225$, a može se pokazati da za velike n vrijedi

$$\mathbb{E}[T] \approx n \log(n) + 0.577n + 0.5\,.$$

□

TEOREM 4.17. *Ako su X i Y diskretne slučajne varijable takve da je $\mathbb{P}(0 \leqslant X \leqslant Y) = 1$, vrijedi $\mathbb{E}[X] \leqslant \mathbb{E}[Y]$, pri čemu obje strane mogu biti $+\infty$.*

DOKAZ. Pogledati sami.

[Neka je $D_X = \{a_i : i \in I\}, D_Y = \{b_j : j \in J\}$, te $p_{ij} = \mathbb{P}(X = a_i, Y = b_j)$. Pretpostavka zadatka povlači da za sve i, j,

$$p_{ij} > 0 \implies a_i \leqslant b_j\,,$$

jer bi inače imali $\mathbb{P}(X > Y) \geqslant p_{ij} > 0$.

Sada slijedi

$$\begin{aligned}\mathbb{E}[X] &= \sum_{i\in I} a_i \underbrace{\mathbb{P}(X = a_i)}_{=\sum_{j\in J} p_{ij}} = \sum_{i\in I}\sum_{j\in J} a_i p_{ij} \leqslant \sum_{i\in I}\sum_{j\in J} b_j p_{ij} \\ &= [\text{Fubini}] = \sum_{j\in J} b_j \sum_{i\in I} p_{ij} = \sum_{j\in J} b_j \mathbb{P}(Y = b_j) = \mathbb{E}[Y]\,.\end{aligned}$$

Gornju implikaciju smo iskoristili da bi dobili nejednakost u trećem koraku budući da zapravo sumiramo samo po parovima i, j za koje je $p_{ij} > 0$.] □

4.4. Kovarijanca i korelacija

[Pitanje je kako mjeriti zavisnost između dvije slučajne varijable.]

DEFINICIJA 4.18. Neka su X i Y slučajne varijable definirane na $(\Omega, \mathcal{F}, \mathbb{P})$ takve da vrijedi $\mathbb{E}[X^2], \mathbb{E}[Y^2] < \infty$. **Kovarijanca** između X i Y definira se kao[4]

$$\mathrm{Cov}(X, Y) := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]\,. \tag{4.8}$$

NAPOMENA. Kovarijanca $\mathrm{Cov}(X, Y)$ mjeri **"linearnu zavisnost"** između X i Y. [Intuitivno, ako u slučajevima kada je vrijednost od X veća (manja) od $\mathbb{E}[X]$, tipično vrijedi i da je vrijednost od Y veća (manja) od $\mathbb{E}[Y]$, kovarijanca će biti pozitivna].

Dokaz idućih osnovnih svojstava kovarijance slijedi direktno iz definicije, tako da ga ostavljamo za vježbu.

PROPOZICIJA 4.19 (**Svojstva kovarijance**). *Vrijedi*

(a) $\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$ *[ovako tipično računamo* $\mathrm{Cov}(X, Y)$*];*

(b) $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$*;*

(c) $\mathrm{Cov}(aX + b, cY + d) = ac \cdot \mathrm{Cov}(X, Y)$*,* $\forall a, b, c, d \in \mathbb{R}$*.*

PRIMJER 4.20. Bacamo simetričan novčić 2 puta te s X označimo broj pisama, a s Y broj glava koje su pale. Tada X i Y imaju $\mathrm{B}(2, \frac{1}{2})$ razdiobu, a zajednička razdioba im je

$X\backslash Y$	0	1	2
0	0	0	$\frac{1}{4}$
1	0	$\frac{1}{2}$	0
2	$\frac{1}{4}$	0	0

Specijalno, imamo $\mathbb{E}[X] = \mathbb{E}[Y] = 2 \cdot \frac{1}{2} = 1$, a budući da je $X \cdot Y = g(X, Y)$ iz (4.2) povlači

$$\mathbb{E}[XY] = \frac{1}{2} \cdot 1 \cdot 1 + 0 = \frac{1}{2}\,,$$

pa je $\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = -\frac{1}{2}$. U ovom slučaju je $\mathrm{Cov}(X, Y) < 0$ jer je zapravo $Y = 2 - X$.

TEOREM 4.21. *Ako su X i Y nezavisne te je (a) $X, Y \geqslant 0$, ili (b) $\mathbb{E}\big[|X|\big], \mathbb{E}\big[|Y|\big] < \infty$, vrijedi*

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]\,. \tag{4.9}$$

DOKAZ. Budući da je $XY = g(X, Y)$, dokaz je jednostavna primjena tvrdnje (4.2), pa ga ostavljamo studentima da pogledaju sami.

[Neka je $D_X = \{a_i : i \in I\}$, $D_Y = \{b_j : j \in J\}$.

[4] Kasnije ćemo pokazati da uvjeti $\mathbb{E}[X^2], \mathbb{E}[Y^2] < \infty$ osiguravaju da je $\mathrm{Cov}(X, Y)$ dobro definirana.

(a) Budući da je $XY \geqslant 0$, vrijedi

$$\begin{aligned}\mathbb{E}[XY] &= \sum_{i\in I}\sum_{j\in J} a_i b_j \mathbb{P}(X = a_i, Y = b_j) = [\text{nez.}] = \sum_{i\in I}\sum_{j\in J} a_i b_j \mathbb{P}(X = a_i)\mathbb{P}(Y = b_j)\\ &= \left(\sum_{i\in I} a_i \mathbb{P}(X = a_i)\right)\cdot\left(\sum_{j\in J} b_j \mathbb{P}(Y = b_j)\right) = \mathbb{E}[X]\cdot\mathbb{E}[Y] \in [0,\infty]\,.\end{aligned}$$

(b) Iz (a) dijela slijedi da je nužno

$$\mathbb{E}\Big[|XY|\Big] = \mathbb{E}\Big[|X|\cdot|Y|\Big] = \mathbb{E}\Big[|X|\Big]\cdot\mathbb{E}\Big[|Y|\Big] < \infty\,.$$

Dakle, postoji $\mathbb{E}[XY]$ te se (4.9) pokaže analogno kao u (a) dijelu.]

□

Uočimo, ako su X i Y nezavisne (te vrijedi $\mathbb{E}[X^2], \mathbb{E}[Y^2] < \infty$), nužno je

$$\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \stackrel{(4.9)}{=} 0\,.$$

Definicija 4.22. Ako vrijedi $\mathrm{Cov}(X,Y) = 0$, kažemo da su X i Y **nekorelirane**.

[Dakle, pokazali smo da nezavisnost povlači nekoreliranost.]

Primjer 4.23 (**Nekoreliranost ne povlači nezavisnost**). Neka je $X \sim \begin{pmatrix} -1 & 0 & 1\\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}\end{pmatrix}$ te $Y := X^2$. Vrijedi $\mathbb{E}[X] = 0$, a budući da je $XY = X^3 = X$, vrijedi i $\mathbb{E}[XY] = \mathbb{E}[X] = 0$, iz čega odmah slijedi da je $\mathrm{Cov}(X,Y) = 0$. Dakle, X i Y su nekorelirane iako su (potpuno) zavisne! [Problem je u tome što je zavisnost u ovom slučaju nelinearna.] □

Propozicija 4.24 (**Varijanca sume zavisnih varijabli**). *Vrijedi*

(a) $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y)$*;*

(b) $\mathrm{Var}(X_1 + \cdots + X_n) = \sum_{i=1}^n \mathrm{Var}(X_i) + 2\sum_{1\leqslant i<j\leqslant n}\mathrm{Cov}(X_i,X_j)$.

[Ovdje pretpostavljamo da svaka slučajna varijabla ima konačan drugi moment tako da su sve kovarijance dobro definirane.]

Dokaz. Dokaz ostavljamo za vježbu – (a) slijedi direktno iz definicija varijance i kovarijance, a (b) dio slijedi iz (a) koristeći indukciju te jednakost $\mathrm{Cov}(X+Y,Z) = \mathrm{Cov}(X,Z) + \mathrm{Cov}(Y,Z)$.

[(a) Vrijedi

$$\begin{aligned}\mathrm{Var}(X+Y) &= \mathbb{E}[(X+Y-\mathbb{E}[X+Y])^2] = [\text{lin.}] = \mathbb{E}[((X-\mathbb{E}[X]) + (Y-\mathbb{E}[Y]))^2]\\ &= [\text{lin.}] = \mathbb{E}[(X-\mathbb{E}[X])^2] + 2\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] + \mathbb{E}[(Y-\mathbb{E}[Y])^2]\\ &= \mathrm{Var}(X) + 2\mathrm{Cov}(X,Y) + \mathrm{Var}(Y)\,.\end{aligned}$$

(b) Vrijedi

$$\begin{aligned}
\mathrm{Var}(X_1+\cdots+X_n) &= \mathrm{Var}((X_1+\cdots+X_{n-1})+X_n)\\
&\overset{(a)}{=} \mathrm{Var}(X_1+\cdots+X_{n-1})+\mathrm{Var}(X_n)+2\mathrm{Cov}(X_1+\cdots+X_{n-1},X_n)\\
&\overset{\mathrm{PI}}{=} \sum_{i=1}^{n-1}\mathrm{Var}(X_i)+2\sum_{1\leqslant i<j\leqslant n-1}\mathrm{Cov}(X_i,X_j)+\mathrm{Var}(X_n)+2\sum_{i=1}^{n-1}\mathrm{Cov}(X_i,X_n)\\
&= \sum_{i=1}^{n}\mathrm{Var}(X_i)+2\sum_{1\leqslant i<j\leqslant n}\mathrm{Cov}(X_i,X_j)\,.\,]
\end{aligned}$$

□

Specijalno, ako su $X_1,\dots,X_n$ *u parovima nezavisne* (ili samo vrijedi $\mathrm{Cov}(X_i,X_j)=0$ za sve $i\neq j$) [i t.d. $\mathbb{E}[X_i^2]<\infty$, $\forall i$], vrijedi

$$\mathrm{Var}\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}\mathrm{Var}(X_i)\,. \tag{4.10}$$

Primjer 4.25. Bacamo novčić na kojem je vjerojatnost za pismo jednaka $p\in(0,1]$. Neke je $n\in\mathbb{N}$ i T broj bacanja potreban da dobijemo ukupno n pisama. Odredite $\mathbb{E}[T]$ i $\mathrm{Var}(T)$.

Rješenje. [Ovdje je ideja slična kao kod problema sakupljača kupona.]

Neka je T_1 broj bacanja potreban da se dobije prvo pismo, T_2 *dodatni* broj bacanja potreban da se dobije drugo pismo itd. Tada očito vrijedi $T=T_1+\cdots+T_n$, a iz nezavisnosti bacanja slijedi:[5]

- $T_1,\dots,T_n$ su nezavisne;
- $T_i\sim \mathrm{G}(p)$, za sve $i=1,\dots,n$.

Sada odmah imamo da je

$$\mathbb{E}[T]=\sum_{i=1}^{n}\mathbb{E}[T_i]=\sum_{i=1}^{n}\frac{1}{p}=\frac{n}{p}\,,$$

a zbog *nezavisnosti* imamo i

$$\mathrm{Var}(T)=\sum_{i=1}^{n}\mathrm{Var}(T_i)=\sum_{i=1}^{n}\frac{1-p}{p^2}=n\cdot\frac{1-p}{p^2}\,.$$

□

Primjer 4.26 (**Varijanca** $\mathrm{B}(n,p)$ **razdiobe**). Neka su $X_1,\dots,X_n$ njd sa zajedničkom distribucijom $\mathrm{B}(p)$ za $p\in[0,1]$. Tada za $X:=\sum_{i=1}^n X_i$ vrijedi $X\sim\mathrm{B}(n,p)$, te je zbog nezavisnosti X_i-eva

$$\mathrm{Var}(X)=\sum_{i=1}^{n}\mathrm{Var}(X_i)=n\mathrm{Var}(X_1)=np(1-p)\,.$$

Primjer 4.27. Neka je $n\in\mathbb{N}$,

$$A_i:=\{i \text{ je fiksna točka slučajne permutacije skupa } \{1,\dots,n\}\}\,,\quad i=1,\dots,n\,.$$

[5] Probajte ovo formalno pokazati.

te $X := \sum_{i=1}^n \mathbb{1}_{A_i}$ ukupan broj fiksnih točaka. Odredite $\mathrm{Var}(X)$.

Rješenje. Ranije smo pokazali da je $\mathbb{P}(A_i) = \frac{1}{n}$, $\mathbb{P}(A_i \cap A_j) = \frac{1}{n(n-1)}$, za sve $i \neq j$. [Specijalno, indikatori $\mathbb{1}_{A_1}, \ldots, \mathbb{1}_{A_n}$ su zavisni, tako da za $\mathrm{Var}(X)$ ne možemo koristiti (4.10).] Imamo

$$\mathrm{Var}(X) = \sum_{i=1}^n \mathrm{Var}(\mathbb{1}_{A_i}) + 2 \sum_{1 \leqslant i < j \leqslant n} \mathrm{Cov}(\mathbb{1}_{A_i}, \mathbb{1}_{A_j}).$$

Računamo

- $\mathrm{Var}(\mathbb{1}_{A_i}) = \mathbb{P}(A_i)(1 - \mathbb{P}(A_i)) = \frac{n-1}{n^2}$, $\forall i$;
- Za sve $i \neq j$,

$$\begin{aligned}\mathrm{Cov}(\mathbb{1}_{A_i}, \mathbb{1}_{A_j}) &= \mathbb{E}[\underbrace{\mathbb{1}_{A_i} \cdot \mathbb{1}_{A_j}}_{=\mathbb{1}_{A_i \cap A_j}}] - \mathbb{E}[\mathbb{1}_{A_i}]\mathbb{E}[\mathbb{1}_{A_j}] = \mathbb{P}(A_i \cap A_j) - \mathbb{P}(A_i)\mathbb{P}(A_j) \\ &= \frac{1}{n(n-1)} - \frac{1}{n^2} = \frac{1}{n^2(n-1)}\end{aligned}$$

Dakle,

$$\mathrm{Var}(X) = n \cdot \frac{n-1}{n^2} - 2 \cdot \binom{n}{2} \frac{1}{n^2(n-1)} = \cdots = 1, \ \forall n \in \mathbb{N}.$$

Uočimo da je za fiksne $i \neq j$, $\mathrm{Cov}(\mathbb{1}_{A_i}, \mathbb{1}_{A_j}) > 0$ za sve $n \in \mathbb{N}$, ali da $\lim_{n\to\infty} \mathrm{Cov}(\mathbb{1}_{A_i}, \mathbb{1}_{A_j}) = 0$. [Je li to intuitivno jasno?] □

4.4.1. Korelacija. Kovarijanca $\mathrm{Cov}(X, Y)$ zapravo nije dobra mjera zavisnosti jer npr. za svaki $a > 0$, $\mathrm{Cov}(aX, Y) = a \cdot \mathrm{Cov}(X, Y)$. [Dakle, nije svejedno mjerimo li visinu X u metrima ili centimetrima.] Zbog toga uvodimo pojam **korelacije**.

Napomena 4.28. Ako je $\mathbb{E}\big[|X|\big] < \infty$, vrijedi

$$\Big|\mathbb{E}[X]\Big| \leqslant \mathbb{E}\big[|X|\big]. \tag{4.11}$$

Zaista, to slijedi iz monotonosti očekivanja te činjenice da su $|X| - X$ te $|X| + X$ nenegativne slučajne varijable.

Teorem 4.29. *Neka su X i Y diskretne slučajne varijable takve da je $\mathbb{E}[X^2], \mathbb{E}[Y^2] < \infty$. Tada vrijedi*

(a) $\mathbb{E}\big[|XY|\big] < \infty$;

*(b) (**Cauchy-Schwartzova nejednakost**)* $|\mathbb{E}[XY]| \leqslant \sqrt{\mathbb{E}[X^2]} \cdot \sqrt{\mathbb{E}[Y^2]}$;

(c) Ako je $\mathbb{E}[X^2], \mathbb{E}[Y^2] > 0$, U CS nejednakosti vrijedi jednakost akko postoji $\lambda \in \mathbb{R}$, $\lambda \neq 0$, takav da je $\mathbb{P}(Y = \lambda X) = 1$.

Dokaz. Neka je $D_X = \{a_i : i \in I\}, D_Y = \{b_j : j \in J\}$, te $p_{ij} = \mathbb{P}(X = a_i, Y = b_j)$.

(a) Pogledati sami – primjena CS nejednakosti za konačne nizove.

[Vrijedi

$$\mathbb{E}\Big[|XY|\Big] = \sum_{i\in I, j\in J} |a_i \cdot b_j| p_{ij} = \sum_{i,j} \underbrace{|a_i| p_{ij}^{\frac{1}{2}}}_{=:A_{i,j}} \underbrace{|b_j| p_{ij}^{\frac{1}{2}}}_{=:B_{i,j}} .$$

Budući da je $\{(A_{i,j}, B_{i,j} : i \in I\,,\ j \in J)\}$ najviše prebrojiv skup, njegove elemente možemo poredati u konačan ili beskonačan niz $(A_1, B_1), (A_2, B_2), \dots$, pa upotrebom CS nejednakosti na vektore $(A_1, \dots, A_n)$ i $(B_1, \dots, B_n)$ dobivamo

$$\begin{aligned}
\mathbb{E}\Big[|XY|\Big] &= \lim_{n\to\infty} \sum_{k=1}^{n} A_k \cdot B_k \leqslant \lim_{n\to\infty} \left(\sum_{k=1}^{n} A_k^2\right)^{\frac{1}{2}} \cdot \left(\sum_{k=1}^{n} B_k^2\right)^{\frac{1}{2}} \\
&= \left(\sum_{k=1}^{\infty} A_k^2\right)^{\frac{1}{2}} \cdot \left(\sum_{k=1}^{\infty} B_k^2\right)^{\frac{1}{2}} = \left(\sum_{i\in I, j\in J}^{\infty} a_i^2 p_{ij}\right)^{\frac{1}{2}} \cdot \left(\sum_{i\in I, j\in J}^{\infty} b_j^2 p_{ij}\right)^{\frac{1}{2}} \\
&= \left(\sum_{i\in I}^{\infty} a_i^2 \mathbb{P}(X = a_i)\right)^{\frac{1}{2}} \cdot \left(\sum_{j\in J}^{\infty} b_j^2 \mathbb{P}(Y = b_j)\right)^{\frac{1}{2}} = \mathbb{E}[X^2]^{\frac{1}{2}} \cdot \mathbb{E}[Y^2]^{\frac{1}{2}} .
\end{aligned}$$

(b) [Ovo slijedi (4.11) i dokaza (a) dijela. Ipak, dajemo alternativan dokaz koji će nam odmah dati i (c) dio.]

Za sve $\lambda \in \mathbb{R}$ definiramo $W_\lambda := X + \lambda Y$. Za sve $\lambda \in \mathbb{R}$ vrijedi

$$\begin{aligned}
0 &\leqslant \mathbb{E}[W_\lambda^2] = \mathbb{E}[X^2 + 2\lambda XY + \lambda^2 Y^2] \\
&= \mathbb{E}[X^2] + 2\lambda \mathbb{E}[XY] + \lambda^2 \mathbb{E}[Y^2] =: c + \lambda b + \lambda^2 a =: g(\lambda) .
\end{aligned}$$

U drugoj jednakosti smo mogli koristiti linearnost jer po pretpostavci i (a) dijelu sva očekivanja postoje.

U slučaju da je $a = \mathbb{E}[Y^2] = 0$, vrijedi $\mathbb{P}(Y = 0) = 1$, pa je i $\mathbb{P}(XY = 0) = 1$ te CS nejednakost trivijalno vrijedi. Pretpostavimo da je $a = \mathbb{E}[Y^2] > 0$. Tada nenegativnost kvadratnog polinoma g povlači da nema dvostruku realnu nulotočku, tj. da je nužno

$$0 \geqslant D = b^2 - 4ac = (2\mathbb{E}[XY])^2 - 4\mathbb{E}[X^2]\mathbb{E}[Y^2] ,$$

što je upravo CS nejednakost.

(c) Iz (b) dijela slijedi da u CS nejednakosti imamo jednakost akko je $D = 0$, tj. akko postoji $\lambda \in \mathbb{R}$ takav da je $\mathbb{E}[W_\lambda^2] = g(\lambda) = 0$. U tom slučaju vrijedi

$$1 = \mathbb{P}((X + \lambda Y)^2 = 0) = \mathbb{P}(X + \lambda Y = 0) = \mathbb{P}(x = -\lambda Y) .$$

Uočimo samo da je nužno $\lambda \neq 0$ jer bi inače bilo $\mathbb{P}(X = 0) = 1$, ali to ne može vrijediti jer smo pretpostavili da je $\mathbb{E}[X^2] > 0$.] □

Napomena 4.30. Ako je $\mathbb{E}[X^2], \mathbb{E}[Y^2] < \infty$, znamo da je i $\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2], \mathrm{Var}(Y) < \infty$.

- Primjenom Teorema 4.29(a) na varijable $X-\mathbb{E}[X]$ i $Y-\mathbb{E}[Y]$ slijedi da je i
$$\mathbb{E}\Big[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])\Big]<\infty\,, \tag{4.12}$$
tj. $\mathrm{Cov}(X,Y)=\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])]$ je dobro definirana.
- Slično, iz CS nejednakosti dobivamo da je
$$|\mathrm{Cov}(X,Y)|\leqslant\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}\,. \tag{4.13}$$

Definicija 4.31. Za slučajne varijable X,Y takve da je $0<\mathrm{Var}(X),\mathrm{Var}(Y)<\infty$, **koeficijent korelacije** između X i Y definira se kao
$$\rho(X,Y):=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}\,.$$

Propozicija 4.32 (**Svojstva korelacije**). *(a) $\rho(aX+b,cY+d)=\rho(X,Y)$, za sve $a,b,c,d\in\mathbb{R}$ $(a,b\neq 0)$;*

(b) $\rho(X,Y)\in[-1,1]$ [ovo je CS nejednakost (4.13)];

(c) $|\rho(X,Y)|=1$ akko postoje $a,b\in\mathbb{R}$, $a\neq 0$, takvi da je $\mathbb{P}(Y=aX+b)=1$. U tom slučaju,
$$\rho(X,Y)=\begin{cases}1\,, & a>0\\ -1\,, & a<0\end{cases}$$

(d) Ako su X i Y nezavisne, nužno je $\rho(X,Y)=0$.

Dokaz. Dokazujemo samo (c) dio. Iz Teorema 4.29 i (4.13) slijedi da je $|\rho(X,Y)|=1$ akko postoji $\lambda\neq 0$ takav da vrijedi
$$1=\mathbb{P}(Y-\mathbb{E}[Y]=\lambda\cdot(X-\mathbb{E}[X]))=\mathbb{P}(Y=\lambda X-\lambda\mathbb{E}[X]+\mathbb{E}[Y])=:\mathbb{P}(Y=aX+b)\,.$$
Nadalje,
$$\mathrm{Cov}(X,Y)=\mathrm{Cov}(X,aX+b)=a\mathrm{Cov}(X,X)=a\mathrm{Var}(X)\,,$$
pa budući da je $\mathrm{Var}(X)>0$, slijedi da $\rho(X,Y)\in\{-1,1\}$ ima isti predznak kao i a. □

Primjer 4.33. Bacamo dva simetrična novčića te označimo s X ukupan broj pisama, a s Y ukupan broj glava koje su pale. Već smo pokazali da je $\mathrm{Cov}(X,Y)=-\frac{1}{2}$, pa budući da je
$$\mathrm{Var}(X)=[X\sim\mathrm{B}(2,\tfrac{1}{2})]=2\cdot\tfrac{1}{2}\cdot(1-\tfrac{1}{2})=\tfrac{1}{2}=\mathrm{Var}(Y)\,,$$
imamo
$$\rho(X,Y)=\frac{-\frac{1}{2}}{\sqrt{\frac{1}{2}}\cdot\sqrt{\frac{1}{2}}}=-1\,.$$
To smo naravno mogli odmah zaključiti iz prethodne propozicije, budući da je $Y=2-X$.

Primjer 4.34. (a) **Monte Carlo metoda.** Pretpostavimo da je parametar $\theta\in\mathbb{R}$ nepoznat, ali da ga možemo zapisati kao $\theta=\mathbb{E}[X]$ pri čemu je X slučajna varijabla koju znamo simulirati, te da vrijedi $\sigma^2 := \mathrm{Var}(X) < \infty$.[6] Ako su $X_1, \ldots, X_n$ njd slučajne varijable s istom distribucijom kao i X, procjenitelj

$$\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i \,, \tag{4.14}$$

za θ ima sljedeća svojstva:

- $\mathbb{E}[\overline{X}_n] = \mathbb{E}[X_1] = \theta$ [dakle, u prosjeku će $\overline{X}_n$ biti približno jednak θ – kažemo da je *nepristran*];
- $\mathrm{Var}(\overline{X}_n) = \frac{\mathrm{Var}(X_1)}{n} = \frac{\sigma^2}{n}$ [dakle, odstupanje od θ je manje što je varijanca σ^2 manja, te što je n veći].

(b) **Kontrolne varijate*.** Pretpostavimo da postoji slučajna varijabla Z takva da je $\mathrm{Cov}(X, Z) < 0$, te da **znamo** njeno očekivanje – BSOMP da je $\mathbb{E}[Z] = 0$.[7] Možemo li tu informaciju iskoristiti kako bi dobili bolji procjenitelj za θ?

Neka je $\sigma_Z := \mathrm{Var}(Z) < \infty$, te $(X_1, Z_1), \ldots, (X_n, Z_n)$ njd slučajni vektori s istom distribucijom kao i (X, Z). Za proizvoljan $\lambda \in \mathbb{R}$, procjenitelj

$$\widehat{\theta}_{n,\lambda} := \frac{1}{n}\sum_{i=1}^{n}(X_i + \lambda Z_i) = \overline{X}_n + \lambda \cdot \overline{Z}_n \,,$$

ima sljedeća svojstva:

- $\mathbb{E}[\widehat{\theta}_{n,\lambda}] = \theta$ jer je $\mathbb{E}[\overline{Z}_n] = \mathbb{E}[Z_1] = 0$;
- $\mathrm{Var}(\widehat{\theta}_{n,\lambda}) = \frac{1}{n} \cdot \mathrm{Var}(X_1 + \lambda Z_1) = \frac{1}{n} \cdot (\sigma^2 + \lambda^2 \cdot \sigma_Z^2 + \lambda \cdot 2\mathrm{Cov}(X, Z))$.

Lako se pokaže da se gornja varijanca minimizira za

$$\lambda^* = -\frac{\mathrm{Cov}(X, Z)}{\sigma_Z^2} \quad (> 0) \,,$$

te da je u tom slučaju

$$\mathrm{Var}(\widehat{\theta}_{n,\lambda^*}) = \frac{\sigma^2}{n} \cdot (1 - \rho(X, Z)^2) \,,$$

što je strogo manje od $\mathrm{Var}(\overline{X}_n) = \frac{\sigma^2}{n}$ budući da je $\rho(X, Z)^2 > 0$, te je smanjenje varijance veće što je korelacija bliže 1 po apsolutnoj vrijednosti.

Uočimo, ako je $\overline{Z}_n > 0$, zbog $\lambda^* > 0$ imamo $\widehat{\theta}_{n,\lambda^*} > \overline{X}_n$. Je li to intuitivno jasno?

[Zamislimo da je X maksimalna temperatura, a Z količina kiše, u istom danu, pri čemu smo Z centrirali tako da vrijednost $Z = 0$ odgovara slučaju kada je količina kiše bila jednaka prosječnoj vrijednosti. Očekujemo da je $\mathrm{Cov}(X, Z) < 0$, a slučaj $\overline{Z}_n > 0$ znači da smo u promatranom periodu imali n iznadprosječno kišnih dana, što zbog negativne korelacije povlači

[6] Slučajna varijabla X može biti na primjer funkcija velikog broja međusobno zavisnih slučajnih varijabli tako da računanje njenog očekivanja po definiciji nije jednostavan problem.
[7] Inače samo uzmemo slučajnu varijablu $Z - \mathbb{E}[Z]$.

da smo vjerojatno u istom periodu imali ispodprosječno tople dane, tj. $\overline{X}_n < \theta$. Dakle, kako bismo dobili bolju procjenu trebamo korigirati procjenu $\overline{X}_n$ prema gore.][8]

4.5. Zbroj nezavisnih slučajnih varijabli

PROPOZICIJA 4.35. *Ako su $X \sim \mathrm{B}(n,p)$ i $Y \sim \mathrm{B}(m,p)$ nezavisne, vrijedi $X+Y \sim \mathrm{B}(n+m,p)$.*

DOKAZ. Neka su $X_1, \ldots, X_{n+m}$ njd slučajne varijable sa zajedničkom razdiobom $\mathrm{B}(p)$. Ako stavimo

$$X := \sum_{i=1}^{n} X_i\,, \quad Y := \sum_{i=n+1}^{n+m} X_i\,,$$

vrijedi

- $X \sim \mathrm{B}(n,p)$, $Y \sim \mathrm{B}(m,p)$;
- X i Y su nezavisne jer $X = g(X_1, \ldots, X_n)$, a $Y = g(X_{n+1}, \ldots, X_{n+m})$ [tj. X i Y ovise o različitim skupovima X_i-eva].

Nadalje,

$$X + Y = \sum_{i=1}^{n+m} X_i \sim rB(n+m,p)\,.$$

□

PROPOZICIJA 4.36. *Ako su $X \sim \mathrm{P}(\lambda)$ i $Y \sim \mathrm{P}(\mu)$ nezavisne, vrijedi $X+Y \sim \mathrm{P}(\lambda+\mu)$.*

DOKAZ. [Intuitivno, rezultat slijedi iz zakona rijetkih događaja. Naime, ako su $X_n \sim \mathrm{B}(\lfloor n\lambda \rfloor, \frac{1}{n})$ i $Y_n \sim \mathrm{B}(\lfloor n\mu \rfloor, \frac{1}{n})$ nezavisne, X_n konvergira po distribuciji prema Poissonovoj slučajnoj varijabli s parametrom $\lim_{n\to\infty} \lfloor n\lambda \rfloor \cdot \frac{1}{n} = \lambda$, Y_n prema Poissonovoj s parametrom μ. S druge strane, po prethodnom primjeru, $X_n + Y_n \sim \mathrm{B}(\lfloor n\lambda \rfloor + \lfloor n\mu \rfloor, \frac{1}{n})$, pa $X_n + Y_n$ konvergira po distribuciji prema Poissonovoj razdiobi s parametrom $\lambda + \mu$. Uz malo dodatnih rezultata o konvergenciji po distribuciji, ovaj argument je zapravo i formalan dokaz.]

Formalno, $X+Y$ očito poprima vrijednosti u $\mathbb{N}_0$, te za sve $k \in \mathbb{N}_0$, vrijedi

$$\begin{aligned}
\mathbb{P}(X+Y=k) &= \sum_{i=0}^{\infty} \mathbb{P}(X=i, X+Y=k) = \sum_{i=0}^{\infty} \mathbb{P}(X=i, Y=k-i) \\
&= \sum_{i=0}^{k} \mathbb{P}(X=i, Y=k-i) = [\text{nezavisnost}] = \sum_{i=0}^{k} \mathbb{P}(X=i)\mathbb{P}(Y=k-i) \\
&= \sum_{i=0}^{k} \frac{\lambda^k}{k!} e^{-\lambda} \cdot \frac{\mu^{k-i}}{(k-i)!} e^{\mu} \cdot \frac{k!}{k!} = e^{-(\lambda+\mu)} \cdot \frac{1}{k!} \underbrace{\sum_{i=0}^{k} \binom{k}{i} \lambda^i \mu^{k-i}}_{=(\lambda+\mu)^k} \\
&= e^{-(\lambda+\mu)} \cdot \frac{(\lambda+\mu)^k}{k!}\,.
\end{aligned}$$

□

[8] Slučajnu varijablu Z nazivamo *kontrolna varijata* (engl. *control variate*) jer nam omogućava da kontroliramo našu procjenu parametra θ. Metoda kontrolnih varijata je jedna od osnovnih metoda smanjenja varijance u Monte Carlo metodi.

4.6. Uvjetne razdiobe

Neka je (X,Y) slučajni vektor na $(\Omega,\mathcal{F},\mathbb{P})$. Za svaki $y\in\mathbb{R}$ t.d. je $\mathbb{P}(Y=y)>0$ (BSMOP da je $y\in D_Y$), **uvjetna razdioba** od X uz dano $Y=y$ je uvjetna razdioba od X uz dano $A=\{Y=y\}\in\mathcal{F}$, te $\mathbb{E}[X\mid Y=y]:=\mathbb{E}[X\mid A]$ (vidi Poglavlje 3.6).

PROPOZICIJA 4.37. *Ako su $X\sim\mathrm{B}(n,p)$ i $Y\sim\mathrm{B}(m,p)$ nezavisne, tada uvjetno na $X+Y=k$ (za $k\leqslant n+m$), X ima $\mathrm{HG}(n,m,k)$ razdiobu (vidi Primjer 3.46). [Uočimo da ova uvjetna razdioba ne ovisi p, što ćemo iskoristiti u idućem primjeru.]*

DOKAZ. [Prvo dajemo intuitivan dokaz. Pretpostavimo da imamo n bijelih i m crnih kuglica, te da svaku od $n+m$ kuglica, nezavisno od drugih kuglica, s vjerojatnošću p označimo nekim slovom. Tada očito $X:=$ broj označenih bijelih i $Y:=$ broj označenih crnih kuglica, očito imaju tražene razdiobe te su nezavisne. Ako dodatno znamo da je

$$\text{ukupan broj označenih kuglica}=X+Y=k\,,$$

taj skup od k označenih kuglica po konstrukciji s jednakom vjerojatnosti može biti bilo koji k-člani podskup od $n+m$ kuglica – specijalno, slučajna varijabla X je tada ukupan broj bijelih kuglica u tom odabranom skupu od k označenih kuglica, pa ima točno $\mathrm{HG}(n,m,k)$ razdiobu.]

Formalno,

$$\begin{aligned}\mathbb{P}(X=i\mid X+Y=k)&=\frac{\mathbb{P}(X=i,Y=k-i)}{\mathbb{P}(X+Y=k)}\overset{\text{nez.}}{=}\frac{\mathbb{P}(X=i)\mathbb{P}(Y=k-i)}{\mathbb{P}(X+Y=k)}\\&=[X+Y\sim\mathrm{B}(n+m,p)]=\frac{\binom{n}{i}p^i(1-p)^{n-i}\cdot\binom{m}{k-i}p^{k-i}(1-p)^{m-(k-i)}}{\binom{n+m}{k}p^k(1-p)^{n+m-k}}=\frac{\binom{n}{i}\cdot\binom{m}{k-i}}{\binom{n+m}{k}}\,,\end{aligned}$$

pri čemu to vrijedi za sve $i\in\{0,1,\dots,k\}$ takve da je $i\leqslant n$ i $k-i\leqslant m$.

□

PRIMJER 4.38 (**Fisherov egzaktni test**). [Želimo provjeriti imaju li žene veći rizik obolijevanja od određene bolesti od muškaraca. U tu svrhu uzmemo slučajan uzorak od 30 žena i 25 muškaraca, te označimo s X broj oboljelih žena, a s Y broj oboljelih muškaraca u uzorcima.]

Pretpostavimo da su $X\sim\mathrm{B}(30,p_1)$ i $Y\sim\mathrm{B}(25,p_2)$ nezavisne [pri čemu su postoci p_1 i p_2 nepoznati]. Ako je $X+Y=10$, a $X=7$, imamo li dovoljno dokaza u koristi $p_1>p_2$?

Uočimo, kada bi vrijedilo $p_1=p_2=p$ (tzv. *nulta hipoteza*),

$$\mathbb{P}(X\geqslant 7\mid X+Y=10)=[\text{Prop. }4.37]=\mathbb{P}(\mathrm{HG}(30,25,10)\geqslant 7)\approx 0.233\,.$$

Riječima, kada bi rizik od obolijevanja bio isti, uvjetno da je ukupan broj oboljelih 10, vjerojatnost da ćemo među tih 10 oboljelih imati 7 ili više žena je ≈ 0.233 (tzv. *p-vrijednost testa*). Zaključak?

ZADATAK. Ako su $X, Y \sim \mathrm{G}_0(p)$ nezavisne, pokažite da za sve $m \in \mathbb{N}_0$, uvjetno na $X+Y=m$, X ima uniformnu razdiobu na skupu $\{0,1,\ldots,m\}$. [Intuicija? Što predstavlja $X+Y$?]

4.6.1. Uvjetno očekivanje. Za vektor (X, Y) definiramo funkciju $g : D_Y \to \mathbb{R}$ sa

$$g(y) := \mathbb{E}[X \mid Y = y]\,, \quad y \in D_Y\,. \tag{4.15}$$

Slučajnu varijablu (!) $\mathbb{E}[X \mid Y]$ definiranu s

$$\mathbb{E}[X \mid Y] := g(Y)\,,$$

zovemo **uvjetno očekivanje** od X uz dano Y.

Sljedeće svojstvo uvjetnog očekivanja $\mathbb{E}[X \mid Y]$ je samo reformulacija formule potpunog očekivanja iz Teorema 3.48.

PROPOZICIJA 4.39. *Ako oba očekivanja postoje, vrijedi*

$$\mathbb{E}[\mathbb{E}[X \mid Y]] = \mathbb{E}[X]$$

DOKAZ. Budući da je $\mathbb{E}[X \mid Y] = g(Y)$ za g definiranu u (4.15), vrijedi

$$\mathbb{E}[\mathbb{E}[X \mid Y]] = \sum_{y \in D_Y} g(y)\mathbb{P}(Y = y) = \sum_{y \in D_Y} \mathbb{E}[X \mid Y = y]\mathbb{P}(Y = y) = \mathbb{E}[X]\,,$$

gdje smo u zadnjoj jednakosti iskoristili formulu potpunog očekivanja za potpuni sistem događaja $\{\{Y = y\},\, y \in D_Y\}$. □

PRIMJER 4.40 (**Slučajne sume**). Neka je $X_1, X_2, \ldots$ niz njd slučajnih varijabli za koje postoji zajedničko očekivanje $\mu := \mathbb{E}[X_1] \in \mathbb{R}$, te neka je $N \in \mathbb{N}_0$ slučajna varijabla nezavisna od niza X_i-eva. Ako je $X := \sum_{i=1}^{N} X_i$ (uz $\sum_{i=1}^{0} X_i := 0$), odredite $\mathbb{E}[X]$.

RJEŠENJE. Za sve $n \in \mathbb{N}_0$, vrijedi

$$\begin{aligned} g(n) := \mathbb{E}[X \mid N = n] = \mathbb{E}\left[\sum_{i=1}^{N} X_i \mid N = n\right] &= \mathbb{E}\left[\sum_{i=1}^{n} X_i \mid N = n\right] = \sum_{i=1}^{n} \mathbb{E}[X_i \mid N = n] \\ &= [X_i \text{ i } N \text{ nez.}] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n \cdot \mu\,, \end{aligned}$$

pri čemu formula zaista vrijedi i za $n = 0$.

Slijedi da je

$$\mathbb{E}[X \mid N] = g(N) = N \cdot \mu\,,$$

pa je

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid N]] = \mathbb{E}[N \cdot \mu] = \mathbb{E}[N] \cdot \mu = \mathbb{E}[N] \cdot \mathbb{E}[X_1]\,, \tag{4.16}$$

što intuitivno ima smisla, te se naziva **Waldova jednakost**. □

POGLAVLJE 5

Neprekidne slučajne varijable

5.1. Funkcija gustoće

DEFINICIJA 5.1. Za slučajnu varijablu $X : \Omega \to \mathbb{R}$ kažemo da je **(apsolutno) neprekidna** ako postoji funkcija $f : \mathbb{R} \to [0, \infty)$ takva da za sve $x \in \mathbb{R}$ vrijedi[1]

$$\mathbb{P}(X \leqslant x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t\,. \tag{5.1}$$

U tom slučaju f zovemo **funkcija gustoće** (neprekidne) slučajne varijable X te često pišemo $f = f_X$. □

Uvjet (5.1) povlači da za sve $a < b$, vrijedi

$$\mathbb{P}(a < X \leqslant b) = \mathbb{P}(\{X \leqslant b\} \setminus \{X \leqslant a\}) = \mathbb{P}(X \leqslant b) - \mathbb{P}(X \leqslant a) = \int_{a}^{b} f(t)\,\mathrm{d}t\,. \tag{5.2}$$

[nacrtati sliku]

Također, **za sve** $x \in \mathbb{R}$ vrijedi

$$\mathbb{P}(X = x) = \lim_{n\to\infty} \mathbb{P}(x - \tfrac{1}{n} < X \leqslant x + \tfrac{1}{n}) = \lim_{n\to\infty} \int_{x-\frac{1}{n}}^{x+\frac{1}{n}} f(t)\,\mathrm{d}t = 0\,. \tag{5.3}$$

Zadnju jednakost gore navodimo bez dokaza, ali intuitivno slijedi jer je gornji limes jednak "$\int_x^x f(t)\,\mathrm{d}t$". [Uočimo razliku u odnosu na diskretne slučajne varijable!] Specijalno, budući da je $\mathbb{P}(X = a) = \mathbb{P}(X = b) = 0$, npr. vrijedi

$$\mathbb{P}(a \leqslant X < b) = \mathbb{P}(a < X \leqslant b) = \int_{a}^{b} f(t)\,\mathrm{d}t\,,$$

to jest, kod neprekidnih slučajnih varijabli *ne moramo* paziti jesu li rubovi uključeni ili ne.

[1] Integral u (5.1) je zapravo tzv. *Lebesgueov integral* koji ćete raditi na kolegiju *Mjera i integral.* On je općenitiji od klasičnog Riemannovog integrala, te ima neka poželjnija teorijska svojstva. Ipak, u ovom kolegiju integrirat ćemo samo nenegativne funkcije f koje su neprekidne osim u najviše konačno mnogo točaka. Tada, za proizvoljne $-\infty \leqslant a < b \leqslant \infty$, ako su $t_1 < t_2 < \cdots < t_{n-1}$ točke prekida funkcije f koje su unutar segmenta $[a, b]$ (te označimo $t_0 := a, t_n := b$), svojstva Lebesguevog integrala povlače da vrijedi

$$\int_{a}^{b} f(t)\,\mathrm{d}t = \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} f(t)\,\mathrm{d}t\,,$$

a integrali $\int_{t_i}^{t_{i+1}} f(t)\,\mathrm{d}t$, $i = 1, \ldots, n-1$, se podudaraju s (moguće nepravim) Riemannovim integralom funkcije f na intervalu (t_i, t_{i+1}); vidi Primjer 5.17 kasnije.

Općenito, za "skoro sve"[2] $B \subseteq \mathbb{R}$ vrijedi

$$\mathbb{P}(X \in B) = \int_B f(t)\,\mathrm{d}t =: \int_{-\infty}^{\infty} f(t) \cdot \mathbb{1}_{t \in B}\,\mathrm{d}t\,.$$

Napomena 5.2 (**Intuicija o funkciji gustoće**). Ako je $f = f_X$, za sve $x \in \mathbb{R}$ dakle **ne vrijedi** $f(x) = \mathbb{P}(X = x)$, ali za *mali* $\Delta x > 0$ vrijedi [nacrtati sliku]

$$\mathbb{P}(X \in [x - \Delta x, x + \Delta x]) = \int_{x-\Delta x}^{x+\Delta x} f(t)\,\mathrm{d}t \approx 2\Delta x \cdot f(x)\,.$$

Primjer 5.3. Ako je X slučajno odabran broj iz $[0, 1]$, vrijedi[3]

$$\mathbb{P}(X \leqslant x) = \begin{cases} 0\,, & x < 0 \\ x\,, & x \in [0,1] \\ 1\,, & x > 1\,. \end{cases}$$

Uočimo, ako definiramo funkciju [nacrtati sliku]

$$f(t) := \begin{cases} 1\,, & t \in [0,1] \\ 0\,, & \text{inače}\,, \end{cases}$$

za sve $x \in \mathbb{R}$ vrijedi $\mathbb{P}(X \leqslant x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$. Dakle, "slučajno odabrani broj" X je zapravo neprekidna slučajna varijabla s gustoćom f, te kažemo da X ima **uniformu razdiobu** na $[0, 1]$ (oznaka $X \sim \mathrm{U}(0, 1)$). Uočimo da je zapravo $\mathbb{P}(X \in (0, 1)) = 1$. □

Ako je $f : \mathbb{R} \to [0, \infty)$ funkcija gustoće neke neprekidne slučajne varijable X, koristeći neprekidnost vjerojatnosti nužno vrijedi

$$\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = \lim_{n\to\infty} \int_{-\infty}^{n} f(t)\,\mathrm{d}t = \lim_{n\to\infty} \mathbb{P}(X \leqslant n) = \mathbb{P}(X \in \mathbb{R}) = 1\,.$$

[Napomena. (bez dokaza)

- Vrijedi i obrat: svaka funkcija $f : \mathbb{R} \to [0, \infty)$ takva da je $\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1$ je funkcija gustoće neke neprekidne slučajne varijable, tj. postoji vjerojatnosni prostor i neprekidna slučajna varijabla X na njemu takva da je $f = f_X$.
- Funkcija gustoće jedinstveno određuje "distribuciju" neprekidne slučajne varijable X, tj. jedinstveno određuje familiju vjerojatnosti $\mathbb{P}(X \in B)$, za skoro sve $B \subseteq \mathbb{R}$.
- Ipak, funkcija gustoće **nije jedinstvena**: npr. ako gustoći f_X promijnimo vrijednost u konačno mnogo točaka, to će ponovno biti funkcija gustoće od X jer opet vrijedi (5.1).]

[2] To su tzv. *Borelovi* podskupovi od $\mathbb{R}$.
[3] Za $x \in [0, 1]$, $\mathbb{P}(X \leqslant x) = \mathbb{P}(X \in [0, x]) = \frac{x}{1} = x$.

5.2. Funkcija distribucije

DEFINICIJA 5.4. Ako je X slučajna varijabla (bilo kakva!), njena **funkcija distribucije**[4] je funkcija $F : \mathbb{R} \to [0,1]$ definirana s

$$F(x) := \mathbb{P}(X \leqslant x)\,, \quad x \in \mathbb{R}\,. \tag{5.4}$$

Kao i kod funkcije gustoće, često pišemo $F = F_X$. □

Uočimo, ako je $F = F_X$, za sve $a < b$ uvijek vrijedi

$$\mathbb{P}(a < X \leqslant b) = F(b) - F(a)\,. \tag{5.5}$$

Gore je bitno koji rub je uključen, a koji ne (osim ako se ne radi o neprekidnoj slučajnoj varijabli)!

PRIMJER 5.5. Ako je $X \sim \mathrm{B}(\frac{3}{4})$, lako slijedi da je njena funkcija distribucije

$$F_X(x) = \begin{cases} 0\,, & x < 0 \\ \frac{1}{4}\,, & x \in [0,1) \\ 1\,, & x \geqslant 0 \end{cases}$$

Dakle, to je po dijelovima konstantna neopadajuća funkcija koja ima dva skoka – jedan u 0 koji iznosi $\mathbb{P}(X = 0) = \frac{1}{4}$, te drugi u 1 koji iznosi $\mathbb{P}(X = 1) = \frac{3}{4}$ [nacrtati sliku].

PRIMJER 5.6. Ako je $X \sim \mathrm{Unif}(0,1)$, iz Primjera 5.3 slijedi da je njena funkcija distribucije

$$F_X(x) = \begin{cases} 0\,, & x < 0 \\ x\,, & x \in [0,1] \\ 1\,, & x > 1\,. \end{cases}$$

Uočimo, za razliku od prethodnog primjera ova funkcija distribucije je neprekidna [nacrtati sliku].

[Iako vidimo da se funkcija distribucije izgleda dosta drugačije kod neprekidnih u odnosu na diskretne slučajne varijable, u idućem rezultatu navodimo neka zajednička svojstva.]

TEOREM 5.7 (**Svojstva funkcije distribucije**). *Neka je X slučajna varijabla (bilo kakva) i F njena funkcija distribucije. Tada vrijedi:*

(a) F je neopadajuća funkcija, tj. ako je $x \leqslant y$, vrijedi $F(x) \leqslant F(y)$;

(b) $F(-\infty) := \lim_{x\to-\infty} F(x) = 0$, $F(\infty) := \lim_{x\to\infty} F(x) = 1$;

(c) Za sve $x \in \mathbb{R}$, F je ***neprekidna zdesna*** *u x, tj. $F(x^+) := \lim_{y\to x^+} F(y) = F(x)$, te vrijedi*

$$F(x^-) := \lim_{y\to x^-} F(y) = \mathbb{P}(X < x)\,. \tag{5.6}$$

[4] U literaturi se uglavnom koristi izraz **cumulative distribution function** (CDF).

DOKAZ. Dijelove (a) i (b) pogledati sami.

(a) Za sve $x \leqslant y$, vrijedi $\{X \leqslant x\} \subseteq \{X \leqslant y\}$ pa monotonost vjerojatnosti povlači

$$F(x) = \mathbb{P}(X \leqslant x) \leqslant \mathbb{P}(X \leqslant y) = F(y)\,.$$

(b) Budući da je F neopadajuća, limesi $F(-\infty)$ i $F(+\infty)$ postoje, te zbog neprekidnosti vjerojatnosti vrijedi

$$F(+\infty) = \lim_{n\to\infty} F(n) = \mathbb{P}\left(\cup_{n=1}^{\infty}\{X \leqslant n\}\right) = \mathbb{P}(X \in \mathbb{R}) = 1\,.$$

Slično dobivamo da je i $F(-\infty) = \mathbb{P}(\emptyset) = 0$.

(c) Opet, budući da je F neopadajuća, limesi $F(x^-)$ i $F(x^+)$ postoje, te zbog neprekidnosti vjerojatnosti vrijedi

$$\begin{aligned} F(x^+) &= \lim_{n\to\infty} F(x+\tfrac{1}{n}) = \lim_{n\to\infty} \mathbb{P}(X \leqslant x+\tfrac{1}{n}) \\ &= \mathbb{P}\left(\bigcap_{n=1}^{\infty}\{X \leqslant x+\tfrac{1}{n}\}\right) = \mathbb{P}(X \leqslant x) = F(x)\,. \end{aligned}$$

Tvrdnja (5.6) slijedi slično jer je

$$\bigcup_{n=1}^{\infty}\{X \leqslant x-\tfrac{1}{n}\} = \{X < x\}\,.$$

□

Uočimo, ako je X slučajna varijabla [bilo kakva] i F njena funkcija distribucije, (5.6) povlači da za sve $x \in \mathbb{R}$ vrijedi

$$\mathbb{P}(X = x) = \mathbb{P}(X \leqslant x) - \mathbb{P}(X < x) = F(x) - F(x^-)\,, \tag{5.7}$$

tj. $\mathbb{P}(X = x)$ je točno veličina skoka funkcije F u x. Specijalno, ako je X neprekidna slučajna varijabla, zbog $\mathbb{P}(X = x) = 0$, $\forall x \in \mathbb{R}$, slijedi da je F_X uvijek *neprekidna* na $\mathbb{R}$ [odatle i ime "neprekidna slučajna varijabla"].

Tvrdnje iz sljedeće napomene ne dokazujemo.

NAPOMENA 5.8. (a) Za svaku funkciju $F : \mathbb{R} \to [0,1]$ koja zadovoljava (a)-(c) iz Teorema 5.7, nužno postoji vjerojatnosni prostor te slučajna varijabla X na to prostoru čija je funkcija distribucije upravo F.

(b) Ako slučajne varijable X i Y imaju iste funkcije distribucije, tada one imaju i istu distribuciju, tj. vrijedi $\mathbb{P}(X \in B) = \mathbb{P}(Y \in B)$ za "skoro sve" $B \subseteq \mathbb{R}$. [Drugim riječima, funkcija distribucije jedinstveno određuje distribuciju slučajne varijable.]

(c) (**Kako odrediti gustoću iz funkcije distribucije?**) Ako je X neprekidna slučajna varijabla te je $F = F_X$, $f = f_X$, za *većinu* točaka $t \in \mathbb{R}$ vrijedi

$$f(t) = F'(t)\,.$$

Vidi Napomenu 5.18(b) dolje za više detalja. Na primjer, ako je $X \sim \mathrm{Unif}(0,1)$ te gustoća f zadana kao u Primjeru 5.3), $F'(t)$ postoji za sve $t \in \mathbb{R} \setminus \{0,1\}$ te vrijedi $f(t) = F'(t)$ za sve takve t.

(d) Neka je X slučajna varijabla s funkcijom distribucije [nacrtati sliku]

$$F_X(x) = \begin{cases} 0\,, & x < 0 \\ x\,, & x \in [0, \frac{1}{2}] \\ \frac{1}{2}\,, & x \in (\frac{1}{2}, 1) \\ 1\,, & x \geqslant 1\,. \end{cases}$$

Uočimo, vrijedi $\mathbb{P}(X = 1) = F_X(1) - F_X(1^-) = \frac{1}{2}$ te $\mathbb{P}(X = x) = 0$ za sve $x \neq 1$. Specijalno, X nije niti diskretna niti neprekidna slučajna varijabla! [**Kako možemo generirati X?**] Poanta je da funkcija distribucije **uvijek** postoji, tj. ona je općenitiji objekt i od vjerojatnosne funkcije mase (koja određuje razdiobu diskretne slučajne varijable) i od funkcije gustoće (neprekidne slučajne varijable). □

Primjer 5.9 (**Uniformna razdioba**). Za slučajnu varijablu X kažemo da ima **uniformnu** razdiobu na $[a, b]$ (oznaka $X \sim \mathrm{Unif}(a, b)$) za $a < b$, ako je

(a) neprekidna s gustoćom

$$f(t) = \begin{cases} \frac{1}{b-a}\,, & t \in [a, b] \\ 0\,, & \text{inače}, \end{cases} \tag{5.8}$$

[dakle, $f(t) = f(s)$, za sve $s, t \in [a, b]$] ili ekvivalentno,

(b) ako joj je funkcija distribucije dana s

$$F(x) = \begin{cases} 0\,, & x < a \\ \frac{x-a}{b-a}\,, & x \in [a, b] \\ 1\,, & x > b\,. \end{cases} \tag{5.9}$$

Ekvivalencija između (a) i (b) slijedi jer za funkcije f i F definirane u (5.8) i (5.9) vrijedi $F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t$ za sve $x \in \mathbb{R}$. Uočimo, ako je $X \sim \mathrm{Unif}(a, b)$, za sve $[c, d] \subseteq [a, b]$ vrijedi

$$\begin{aligned} \mathbb{P}(X \in [c, d]) &= [\mathbb{P}(X = c) = 0] = \mathbb{P}(X \in (c, d]) = F(d) - F(c) \\ &= \frac{d - c}{b - a} = \frac{\mathrm{duljina}([c, d])}{\mathrm{duljina}([a, b])}\,. \end{aligned}$$

Primjer 5.10 (**Eksponencijalna razdioba**). Kažemo da X ima **eksponencijalnu razdiobu** s parametrom $\lambda > 0$ (oznaka $X \sim \mathrm{Exp}(\lambda)$) ako je [nacrtati slike]

(a) neprekidna s gustoćom

$$f(t) = \begin{cases} \lambda e^{-\lambda t}\,, & t \geqslant 0 \\ 0\,, & \text{inače}, \end{cases} \tag{5.10}$$

ili ekvivalentno,

(b) ako joj je funkcija distribucije dana s

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geqslant 0 \\ 0, & x < 0. \end{cases} \tag{5.11}$$

Možda je najlakše pamtiti ovu razdiobu po obliku njene **repne funkcije distribucije**: $X \sim \text{Exp}(\lambda)$ akko vrijedi

$$\mathbb{P}(X > x) = 1 - F_X(x) = e^{-\lambda x}, \ \ \forall x \geqslant 0.$$

Uočimo, ako je $X \sim \text{Exp}(\lambda)$, očito vrijedi $\mathbb{P}(X \geqslant 0) = 1$, te nadalje X ima tzv. *svojstvo zaboravljivosti*: za sve $z, x \geqslant 0$,

$$\mathbb{P}(X - z \geqslant x \mid X \geqslant z) = \frac{\mathbb{P}(X - z \geqslant x, X \geqslant z)}{\mathbb{P}(X \geqslant z)} = \frac{\mathbb{P}(X \geqslant z + x)}{\mathbb{P}(X \geqslant z)} = \frac{e^{-\lambda(z+x)}}{e^{-\lambda z}} = e^{-\lambda x},$$

to jest, uvjetno na $X \geqslant z$, $X - z$ ponovno ima $\text{Exp}(\lambda)$ razdiobu. □

[Exp(λ) je neprekidna verzija geometrijske razdiobe – čekamo prvi uspjeh ("događaj") u neprekidnom vremenu, pri čemu je prosječan broj događaja po jedinici vremena jednak λ. Malo preciznije...]

Napomena 5.11 (*). **Poissonov proces sa intezitetom** $\lambda > 0$ je niz slučajnih varijabli $0 < T_1 < T_2 < \dots$ [T_i predstavlja vrijeme i-tog događaja], takvih da za

$$N(\langle a, b]) := \sum_{i=1}^{\infty} \mathbb{1}_{\{T_i \in \langle a,b]\}} = \text{broj događaja u intervalu } \langle a, b], \ \ a < b,$$

vrijedi

(a) $N(\langle a, b]) \sim \text{P}(\lambda(b - a))$ za sve $a < b$, te
(b) $N(I_1), \dots, N(I_n)$ su nezavisne za sve n i sve u parovima disjunktne intervale $I_1, \dots, I_n \subseteq (0, \infty)$.

[Poissonov proces koristi se za modeliranje događaja koji se događaju vremenu – npr. dolazaka klijenata u banku ili pojave potresa na nekom području. Ipak, ovaj model je često prejednostavan, ali se koristi kao baza za fleksibilnije modele, vidi npr. *Hawkesove procese.*]

Uočimo, iz (a) slijedi

$$\frac{\mathbb{E}[N(\langle a, b])]}{b - a} = \frac{\lambda(b - a)}{b - a} = \lambda,$$

to jest, intezitet λ predstavlja očekivani broj događaja po jedinici vremena. Nadalje, vrijeme čekanja do prvog događaja T_1 zadovoljava

$$\mathbb{P}(T_1 > t) = \mathbb{P}(N(\langle 0, t]) = 0) \overset{(a)}{=} e^{-\lambda(t-0)} = e^{-\lambda t}, \ \ \forall t > 0,$$

to jest T_1 ima točno $\text{Exp}(\lambda)$ razdiobu [tako možemo razmišljati o $\text{Exp}(\lambda)$ razdiobi i parametru λ]! Zapravo, može se pokazati da za vremena čekanja $E_i := T_i - T_{i-1}$, $i \geqslant 1$, vrijedi da su nezavisna jednako distribuirana sa zajedničkom razdiobom $\mathrm{Exp}(\lambda)$ [što daje jednostavnu metodu za simuliranje Poissonovog procesa]. $\square$

NAPOMENA 5.12 (**Zatvorenost eksponencijalne razdiobe na skaliranje**). Za $X \sim \mathrm{Exp}(1)$[5] i sve $\lambda > 0$ vrijedi $\frac{X}{\lambda} \sim \mathrm{Exp}(\lambda)$. Zaista,

$$\mathbb{P}(X/\lambda \leqslant t) = \mathbb{P}(X \leqslant \lambda t) = [X \sim \mathrm{Exp}(1)] = 1 - e^{1\cdot\lambda t} = 1 - e^{\lambda t}\,,\ t \geqslant 0\,.$$

5.3. Funkcije neprekidne slučajne varijable

Ako je X neprekidna slučajna varijabla, te $g : \mathbb{R} \to \mathbb{R}$ funkcija, slučajna varijabla $g(X)$ *ne mora* nužno biti neprekidna.

PRIMJER 5.13. Neka je $U \sim \mathrm{Unif}(0,1)$, $p \in [0,1]$, te $X := \mathbb{1}_{\{U \geqslant 1-p\}}$. Tada Y očito poprima vrijednosti u $\{0,1\}$ te je

$$\mathbb{P}(X = 1) = \mathbb{P}(U \geqslant 1-p) = \mathbb{P}(U \in [1-p,1]) = \frac{p}{1} = p\,.$$

Dakle, $X \sim \mathrm{B}(p)$.

[Zapravo, konstrukcijom sličnom kao u prethodnom primjeru možemo koristeći jednu $\mathrm{Unif}(0,1)$ slučajnu varijablu generirati proizvoljnu diskretnu slučajnu varijablu.]

PRIMJER 5.14. Neka je $U \sim \mathrm{Unif}(0,1)$, $\lambda > 0$, te

$$Y := -\frac{\log(U)}{\lambda}\,.$$

Odredimo F_Y. Budući da je $\mathbb{P}(U \in (0,1)) = 1$, Y je dobro definirana te je $\mathbb{P}(Y > 0) = 1$.

- za $y < 0$ je $F_Y(y) = \mathbb{P}(Y \leqslant y) = 0$.
- za $y \geqslant 0$,

$$F_Y(y) = \mathbb{P}\left(-\frac{\log(U)}{\lambda} \leqslant y\right) = \mathbb{P}(U \geqslant e^{-\lambda y}) = \mathbb{P}(U \in [e^{-\lambda y},1]) = 1 - e^{-\lambda y}\,.$$

Dakle, $Y \sim \mathrm{Exp}(\lambda)$. $\square$

[Sada ćemo poopćiti prethodni primjer.]

PROPOZICIJA 5.15 (**Simuliranje slučajnih varijabli**). *Neka je F neprekidna funkcija distribucije takva da za neke $a, b \in [-\infty, +\infty]$, $a < b$, vrijedi*

(i) $F(a) = 0$, $F(b) = 1$,[6] *te*

(ii) F je strogo rastuća na (a,b).

Tada postoji inverz $F^{-1} : (0,1) \to (a,b)$ te za $U \sim \mathrm{Unif}(0,1)$, slučajna varijabla $X := F^{-1}(U)$ ima funkciju distribucije točno F.

[5] Ovu razdiobu često nazivamo **standardna** eksponencijalna razdioba.
[6] Ovo zapravo znači da ako je $F = F_X$, $\mathbb{P}(X \in (a,b)) = 1$.

[Prethodni rezultat koristan je za simuliranje slučajnih varijabli čije funkcije distribucije zadovoljavaju gornje pretpostavke – to npr. isključuje sve diskretne razdiobe, ali čak i dosta neprekidnih razdioba. Ipak, gornji rezultat se poopćiti na proizvoljnu funkciju distribucije F tako da se inverz F^{-1} zamijeni s tzv. *generaliziranim inverzom*

$$F^{\leftarrow}(u) := \inf\{x \in \mathbb{R} : F(x) \geqslant u\}\,, \quad u \in (0,1)\,.\,]$$

DOKAZ. Uvjeti (i) i (ii) osiguravaju da je F bijekcija sa (a,b) u $(0,1)$, pa dakle postoji i F^{-1} na $(0,1)$. Pokažimo da je $F_X = F$.

- za $x \notin (a,b)$, budući da je $F^{-1}(u) \in (a,b)$ za sve $u \in (0,1)$,

$$F_X(x) = \mathbb{P}(F^{-1}(U) \leqslant x) = \begin{cases} 0\,, & x \leqslant a \\ 1\,, & x \geqslant b \end{cases} = F(x)\,.$$

- za $x \in (a,b)$,

$$\begin{aligned} F_X(x) &= \mathbb{P}(F^{-1}(U) \leqslant x) = [F \text{ strogo rastuća}] = \mathbb{P}(F(F^{-1}(U)) \leqslant F(x)) \\ &= \mathbb{P}(U \leqslant F(x)) = [F(x) \in (0,1),\, U \sim \mathrm{Unif}(0,1)] = F(x)\,. \end{aligned}$$

□

ZADATAK. Neka F zadovoljava uvjete prethodne propozicije. Ako je X slučajna varijabla s funkcijom distribucije F, pokažite da slučajna varijabla $Y := F(X)$ ima $\mathrm{Unif}(0,1)$ razdiobu.

PRIMJER 5.16. Neka je $F = F_Z$ za $Z \sim \mathrm{Exp}(\lambda)$ – dakle $F(x) = 1 - e^{-\lambda x}$ za $x \geqslant 0$ i $F(x) = 0$ za $x < 0$. Tada F zadovoljava uvjete Prop. 5.15 uz $a = 0, b = +\infty$, te je

$$F^{-1}(u) = -\frac{\log(1-u)}{\lambda}\,, \quad u \in (0,1).$$

Dakle, ako je $U \sim \mathrm{Unif}(0,1)$, slučajna varijabla $X := -\frac{\log(1-U)}{\lambda}$ ima $\mathrm{Exp}(\lambda)$ razdiobu. Uočimo, budući da $1-U$ također ima $\mathrm{Unif}(0,1)$ razdiobu (DZ), slijedi i

$$-\frac{\log(U)}{\lambda} \sim \mathrm{Exp}(\lambda)\,,$$

što smo pokazali u Primjeru 5.14.

ZADATAK.(**Zatvorenost uniformne razdiobe na linearne transformacije**) Ako je $U \sim \mathrm{Unif}(0,1)$, pokažite da za sve $a < b$, slučajna varijabla $X := (b-a)U + a$ ima $\mathrm{Unif}(a,b)$ razdiobu.

[Ipak, kao što smo već vidjeli, nelinearne transformacije uniformne slučajne varijable neće biti uniformne!]

PRIMJER 5.17. Neka je $X \sim \mathrm{Unif}(-1,1)$ te $Y := X^2$. Tada Y poprima vrijednosti u $(0,1)$ – odredimo joj razdiobu.

- očito je $F_Y(y) = 0$ za $y \leqslant 0$, $F_Y(y) = 1$ za $y \geqslant 1$.

- za $y \in (0,1)$,

$$F_Y(y) = \mathbb{P}(X^2 \leqslant y) = \mathbb{P}(X \in [-\sqrt{y}, \sqrt{y}]) = [X \sim \mathrm{Unif}(-1,1)]$$
$$= \frac{2\sqrt{y}}{2} = \sqrt{y}\,.$$

[nacrtati F_Y]

Je li Y neprekidna slučajna varijabla? DA – nije teško provjeriti da za funkciju

$$f(t) := \begin{cases} F_Y'(t)\,, & t \in \mathbb{R} \setminus \{0,1\} \\ 0\,, & t \in \{0,1\} \end{cases} = \begin{cases} \frac{1}{2\sqrt{t}}\,, & t \in (0,1) \\ 0\,, & t \notin (0,1) \end{cases}$$

vrijedi $F_Y(y) = \int_{-\infty}^y f(t)\,\mathrm{d}t$,[7] za sve $y \in \mathbb{R}$. Dakle, Y je neprekidna s gustoćom f – specijalno, Y nema $\mathrm{Unif}(0,1)$ razdiobu.

Napomena 5.18 (**Kako prepoznati funkciju distribucije neprekidne razdiobe?**).

(a) Ako je $F = F_X$ diferencijabilna osim u konačno mnogo točaka, X ne mora nužno biti neprekidna slučajna varijabla.
Na primjer, ako je $X \sim \mathrm{B}(1/2)$, $F_X'(t)$ postoji za sve $t \neq 0, 1$ (te je $F_X'(t) = 0$).

(b) Ipak, ako je (i) $F = F_X$ *neprekidna* na $\mathbb{R}$, (ii) diferencijabilna osim u najviše konačno mnogo točaka $S \subseteq \mathbb{R}$, te (iii) F' neprekidna na $\mathbb{R} \setminus S$ [kao u prethodnom primjeru][8], tada je slučajna varijabla X neprekidna, a za njenu gustoću možemo uzeti

$$f(t) := \begin{cases} F'(t)\,, & t \notin S \\ 0\,, & t \in S\,. \end{cases}$$

[Dokaz.* Neka su $a < b$ proizvoljni, te neka je $a = t_0 < t_1 < \cdots < t_n = b$ subdivizija segmenta $[a,b]$, ali takva je F diferencijabilna za sve $t \in (a,b)$, $t \notin \{t_1, \ldots, t_{n-1}\}$. Po Newton-Leibnizovom teoremu, za sve $i = 1, \ldots, n$, vrijedi

$$\int_{t_{i-1}}^{t_i} f(t)\,\mathrm{d}t = \lim_{\epsilon \to 0} \int_{t_{i-1}+\epsilon}^{t_i-\epsilon} f(t)\,\mathrm{d}t = \lim_{\epsilon \to 0} F(t_i - \epsilon) - F(t_{i-1} + \epsilon) = F(t_i) - F(t_{i-1})\,,$$

gdje smo u zadnjem koraku iskoristili neprekidnost funkcije F. Specijalno, vrijedi

$$\mathbb{P}(X \in (a,b]) = F(b) - F(a) = \sum_{i=1}^n F(t_i) - F(t_{i-1}) = \sum_{i=1}^n \int_{t_{i-1}}^{t_i} f(t)\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t\,.$$

Sada dobivamo da je za proizvoljan $b \in \mathbb{R}$,

$$F(b) = \mathbb{P}(X \leqslant b) = \lim_{a \to -\infty} \mathbb{P}(X \in (a,b]) = \int_{-\infty}^b f(t)\,\mathrm{d}t\,,$$

to jest, X je neprekidna s gustoćom f.] □

---

[7] Budući da f ima vertikalnu asimptotu u 0, ovdje za $y \in (0,1)$, integral $\int_{-\infty}^y f(t)\,\mathrm{d}t$ formalno računamo kao

$$\int_{-\infty}^y f(t)\,\mathrm{d}t = \int_{-\infty}^0 f(t)\,\mathrm{d}t + \int_0^y f(t)\,\mathrm{d}t = \int_0^y f(t)\,\mathrm{d}t = \lim_{a \to 0^-} \int_a^y f(t)\,\mathrm{d}t$$
$$= \lim_{a \to 0^+} F_Y(y) - F_Y(a) = F_Y(y) - F_Y(0) = F_Y(y)\,.$$

[8] Napomenimo ipak da rezultat vrijedi i bez pretpostavke (iii).

Propozicija 5.19 (**Zamjena varijabli**). *Pretpostavimo da je*

- *X neprekidna slučajna varijabla;*
- *I otvoren interval t.d. $\mathbb{P}(X \in I) = 1$, te*
- *$g : I \to (m, M)$ strogo rastuća i neprekidna bijekcija takva da postoji $(g^{-1})'(y)$, za sve $y \in (m, M)$.*

Tada je $Y := g(X)$ neprekidna slučajna varijabla s gustoćom

$$f_Y(y) = f_X(g^{-1}(y))(g^{-1})'(y)\,,\ y \in (m, M) \tag{5.12}$$

te $f_Y(y) = 0$ za $y \notin (m, M)$.

Dokaz prethodne propozicije slijedi deriviranjem kompozicije funkcije budući da je

$$F_Y(y) = \mathbb{P}(g(X) \leqslant y) = [\text{i } g^{-1} \text{ strogo rastuća}] = \mathbb{P}(X \leqslant g^{-1}(y)) = F_X(g^{-1}(y))\,,\ \ y \in (m, M)\,.$$

Detalje ostavljamo za vježbu (vidi dolje) budući da je u praksi lakše svaki put ponovno provesti dokaz nego pamtiti iskaz, pogotovo jer često uvjeti propozicije nisu zadovoljeni (kao u Primjeru 5.17). Ipak, **poanta** je sljedeća – kada je X diskretna i g bijekcija, tada za sve y, vjerojatnosne funkcije mase zadovoljavaju

$$f_{g(X)}(y) = \mathbb{P}(g(X) = y) = \mathbb{P}(X = g^{-1}(y)) = f_X(g^{-1}(y))\,.$$

S druge strane, kada je X neprekidna, općenito za gustoće imamo

$$f_{g(X)}(y) \neq f_X(g^{-1}(y))$$

osim ako je $(g^{-1})'(y) = 1$. Dakle, vrijedit će $f_{g(X)}(y) = f_X(g^{-1}(y))$ za sve $y \in (m, M)$ akko je $g^{-1}(y) = y + c$, tj. akko je $g(x) = x - c$, za neki $c \in \mathbb{R}$ (tj. ako je g **translacija**).

[Dokaz Propozicije 5.19. Očito je $F_Y(y) = 0$ za $y \leqslant m$ i $F_Y(y) = 1$ za $y \geqslant M$. Za $y \in (m, M)$ je dakle $F_Y(y) = F_X(g^{-1}(y))$. Budući da je F_X neprekidna na $\mathbb{R}$ te $g^{-1} : (m, M) \to I$ neprekidna i strogo rastuća, lako se provjeri da je F_Y također neprekidna na $\mathbb{R}$ te diferencijabilna osim eventualno u $y \in \{m, M\}$. Tvrdnja sada slijedi iz Napomene 5.18(b). □]

Napomena 5.20. Ukoliko je u Prop. 5.19 funkcija g strogo **padajuća** umjesto rastuća, $Y = g(X)$ je ponovno neprekidna, ali s gustoćom

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \left|(g^{-1})'(y)\right|,\ y \in (m, M)\,. \tag{5.13}$$

[Dokaz. Za $y \in (m, M)$ vrijedi

$$\begin{aligned} F_Y(y) &= \mathbb{P}(g(X) \leqslant y) = \mathbb{P}(X \geqslant g^{-1}(y)) = [X \text{ neprekidna}] \\ &= \mathbb{P}(X > g^{-1}(y)) = 1 - F_X(g^{-1}(y))\,, \end{aligned}$$

pa tvrdnja slijedi slično kao u Prop. 5.19 budući da je $(g^{-1})'(y) < 0$, pa je $-(g^{-1})'(y) = |(g^{-1})'(y)|$.]

Zadatak. (a) Riješite Primjer 5.14 pomoću (5.13) (ovdje je $I = (0, 1)$, $(m, M) = (0, \infty)$).

(b) Ako je X neprekidna slučajna varijabla i $Y := aX + b$ za $a \neq 0, b \in \mathbb{R}$, pokažite da je Y neprekidna s gustoćom koja zadovoljava

$$f_Y(y) = f_X((y-b)/a) \cdot \frac{1}{|a|}\,,$$

za sve $y \in a \cdot D_X + b$. $\square$

5.4. Očekivanje i varijanca

Definicija 5.21. Neka je X neprekidna slučajna varijabla i $f = f_X$. Ako je (a) $X \geqslant 0$ (tj. $f(t) = 0$ za $t < 0$), ili (b) $\int_{-\infty}^{+\infty} |t| f(t)\,\mathrm{d}t < \infty$, kažemo da postoji očekivanje od X koje definiramo kao[9]

$$\mathbb{E}[X] := \int_{-\infty}^{+\infty} t f(t)\,\mathrm{d}t\,. \tag{5.14}$$

Primjer 5.22. Ako je $U \sim \mathrm{Unif}(0,1)$,

$$\mathbb{E}[U] = \int_{-\infty}^{+\infty} t f_U(t)\,\mathrm{d}t = \int_0^1 t \cdot 1\,\mathrm{d}t = \frac{t^2}{2}\Big|_0^1 = \frac{1}{2}\,,$$

što smo i očekivali.

[Sljedeće dokazujemo neprekidni analogon Teorema 3.21, koji je kao i u diskretnom slučaju dosta važan pri računanju očekivanja.]

Teorem 5.23. *Neka je X neprekidna slučajna varijabla, $f = f_X$ te $g : \mathbb{R} \to \mathbb{R}$ proizvoljna funkcija. Ako je (a) $g(X) \geqslant 0$, ili (b) $\int_{-\infty}^{+\infty} |g(t)| f(t)\,\mathrm{d}t < \infty$, tada vrijedi*

$$\mathbb{E}[g(X)] = \int_{-\infty}^{+\infty} g(t) f(t)\,\mathrm{d}t \tag{5.15}$$

Važno je napomenuti da (5.15) vrijedi *bez obzira* na to kakva je $g(X)$ slučajna varijabla (neprekidna, diskretna, ili nešto treće).

Dokaz. Tvrdnju dokazujemo samo u specijalnim slučajevima.

- Pretpostavimo da g zadovoljava uvjete Prop. 5.19 – dakle $g : I \to (m, M)$. Tada je dakle $Y := g(X)$ neprekidna s gustoćom danom u (5.12). Ako je $Y \geqslant 0$ (tj. $m \geqslant 0$), tada koristeći zamjenu varijabli dobivamo

$$\begin{aligned}
\mathbb{E}[Y] &= \int_{-\infty}^{+\infty} y f_Y(y)\,\mathrm{d}y = \int_m^M y f_X(g^{-1}(y))(g^{-1})'(y)\,\mathrm{d}y \\
&= [x = g^{-1}(y),\ \mathrm{d}x = (g^{-1})'(y)\,\mathrm{d}y,\ g^{-1}(\langle m, M\rangle) = I] \\
&= \int_I g(x) f_X(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} g(x) f_X(x)\,\mathrm{d}x \ \in [0,\infty]\,.
\end{aligned}$$

[9] Slično kao u diskretnom slučaju, uvjeti (a) i (b) povlače

$$\lim_{M\to\infty}\lim_{m\to\infty}\int_{-m}^{M} t f(t)\,\mathrm{d}t = \lim_{m\to\infty}\lim_{M\to\infty}\int_{-m}^{M} t f(t)\,\mathrm{d}t =: \int_{-\infty}^{+\infty} t f(t)\,\mathrm{d}t\,.$$

Uz uvjet (b), gornji integral je nužno realan broj, dok u slučaju (a) je iz $[0,\infty]$.

U općenitom slučaju, tj. ako vrijedi (b), gornji račun pokazuje da je $\int_{-\infty}^{+\infty} |y| f_Y(y)\,\mathrm{d}y < \infty$, pa postoji $\mathbb{E}[Y]$, te se (5.15) pokazuje analogno.

- (Pogledati sami*) Pretpostavimo sada da je $g(\mathbb{R}) = \{a_1, a_2, \dots\} = \{a_i : i \in I\} \subseteq [0, \infty)$. Sada je dakle $Y := g(X)$ diskretna i nenegativna slučajna varijabla, pa je

$$\mathbb{E}[Y] = \sum_{i \in I} a_i \mathbb{P}(Y = a_i)\,.$$

Sada uočimo da je

$$\mathbb{P}(Y = a_i) = \mathbb{P}(X \in g^{-1}(\{a_i\})) = \int_{g^{-1}(\{a_i\})} f_X(t)\,\mathrm{d}t\,,$$

pa budući da familija skupova $\{g^{-1}(\{a_i\}),\ i \in I\}$ čini particiju skupa $\mathbb{R}$, imamo

$$\mathbb{E}[Y] = \sum_{i \in I} \int_{g^{-1}(\{a_i\})} \underbrace{a_i}_{=g(t)} f_X(t)\,\mathrm{d}t = \int_{-\infty}^{\infty} g(t) f_X(t)\,\mathrm{d}t\,.$$

□

Iz (5.15) slijedi da i za X neprekidnu vrijedi da postoji $\mathbb{E}[X]$ te je $\mathbb{E}[X] \in \mathbb{R}$ akko vrijedi

$$\mathbb{E}\Big[|X|\Big] = \int_{-\infty}^{+\infty} |t| f_X(t)\,\mathrm{d}t < \infty\,.$$

Dokaz iduće propozicije ostavljamo za vježbu.

Propozicija 5.24. *Neka je X neprekidna slučajna varijabla, te $a, b \in \mathbb{R}$ konstante.*

(a) Ako je $X \geqslant 0$ (i $a \geqslant 0$) ili $\mathbb{E}\Big[|X|\Big] < \infty$, vrijedi

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b\,. \tag{5.16}$$

(b) Ako je $X \in [a, b]$, tj. $f_X(t) = 0$ za $a \notin [a, b]$, vrijedi $\mathbb{E}[X] \in [a, b]$.

[Dokaz. (a) Ako je $X \geqslant 0$ i $a \geqslant 0$,

$$\begin{aligned}\mathbb{E}[aX + b] &\overset{(5.15)}{=} \int_{-\infty}^{\infty} (at + b) \cdot f_X(t)\,\mathrm{d}t = a \int_{-\infty}^{\infty} t \cdot f_X(t)\,\mathrm{d}t + b \int_{-\infty}^{\infty} f_X(t)\,\mathrm{d}t \\ &= a\mathbb{E}[X] + b \cdot 1\,.\end{aligned}$$

Općeniti slučaj (postoji $\mathbb{E}[aX + b]$ i vrijedi (5.16)) dokazuje se analogno.

(b) Budući da je $f_X(t) = 0$ za $a \notin [a, b]$,

$$\begin{aligned}\mathbb{E}[X] &= \int_{-\infty}^{+\infty} t f_X(t)\,\mathrm{d}t = \int_a^b t f_X(t)\,\mathrm{d}t \\ &\leqslant b \cdot \int_a^b f_X(t)\,\mathrm{d}t = b \cdot \int_{-\infty}^{\infty} f_X(t)\,\mathrm{d}t = b \cdot 1 = b\,,\end{aligned}$$

te se analogno dobije $\mathbb{E}[X] \geqslant a$.]

Primjer 5.25 (**Očekivanje uniformne razdiobe**). Ako je $U \sim \mathrm{Unif}(0,1)$, znamo da je $X := (b-a)U + a \sim \mathrm{Unif}(a,b)$, pa budući da je $\mathbb{E}[U] = \frac{1}{2}$,

$$\mathbb{E}[X] \overset{(5.16)}{=} (b-a)\mathbb{E}[U] + a = \frac{a+b}{2}\,.$$

[Zapravo...]

Napomena 5.26. (!) Svojstva **linearnosti** i **monotonosti** očekivanja (Teoremi 4.14 i 4.17) vrijede i za neprekidne slučajne varijable (bez dokaza).

Propozicija 5.27. *Ako je X neprekidna **nenegativna** slučajna varijabla, vrijedi*

$$\mathbb{E}[X] = \int_0^\infty \mathbb{P}(X > t)\,\mathrm{d}t \;\in [0,\infty]\,. \tag{5.17}$$

Dokaz. Vrijedi

$$\int_0^\infty \mathbb{P}(X>t)\,\mathrm{d}t = \int_0^\infty \left(\int_t^\infty f_X(z)\,\mathrm{d}z\right)\,\mathrm{d}t\,.$$

U gornjem integralu integriramo po području

$$\{(t,z) : t \geqslant 0, z \geqslant t\} = \{(t,z) : z \geqslant 0, t \in [0,z]\}\,,$$

pa koristeći neprekidnu verziju Fubinijevog teorema možemo zamijeniti poredak integracije

$$\int_0^\infty \mathbb{P}(X>t)\,\mathrm{d}t = \int_0^\infty \left(\int_0^z f_X(z)\,\mathrm{d}t\right)\,\mathrm{d}z = \int_0^\infty f_X(z) z\,\mathrm{d}z = \mathbb{E}[X]$$

gdje zadnja jednakost slijedi jer je $X \geqslant 0$, tj. $f_X(t) = 0$ za $t < 0$. □

Primjer 5.28. Ako je $X \sim \mathrm{Exp}(\lambda)$, tj. $\mathbb{P}(X > t) = e^{-\lambda t}$ za $t \geqslant 0$, vrijedi

$$\mathbb{E}[X] = \int_0^\infty \mathbb{P}(X>t)\,\mathrm{d}t = \int_0^\infty e^{-\lambda t}\,\mathrm{d}t = -\frac{e^{-\lambda t}}{\lambda}\Big|_0^\infty = (-0 + \tfrac{1}{\lambda}) = \frac{1}{\lambda}\,.$$

Primjer 5.29 (**Paretova razdioba**). Kažemo da slučajna varijabla Y ima **Paretovu** razdiobu s parametrom $\alpha > 0$ (oznaka $Y \sim \mathrm{Pareto}(\alpha)$) ako vrijedi

$$F_Y(y) = 1 - y^{-\alpha}\,,\; y \geqslant 1$$

te $F_Y(y) = 0$ za $y < 1$. Dakle, vrijedi $Y \geqslant 1$ te repna funkcija distribucije opada kao potencija: za $y \geqslant 1$, $\mathbb{P}(Y > y) = y^{-\alpha}$. Nadalje, to je neprekidna slučajna varijabla s gustoćom

$$f_Y(y) = F_Y'(y) = \alpha y^{-\alpha-1}\,,\; y > 1\,.$$

Koristeći (5.17) dobivamo da je

$$\mathbb{E}[Y] = \int_0^\infty \mathbb{P}(Y>y)\,\mathrm{d}y = \int_0^1 1\,\mathrm{d}y + \int_1^\infty y^{-\alpha}\,\mathrm{d}y = 1 + \int_1^\infty y^{-\alpha}\,\mathrm{d}y\,.$$

Sada lako slijedi da je

$$\mathbb{E}[Y] = \begin{cases} +\infty\,, & \text{za } \alpha \leqslant 1 \\ \frac{\alpha}{\alpha-1}\,, & \text{za } \alpha > 1 \end{cases}$$

[Paretova razdioba primjer je razdiobe **"teškog repa"**: vjerojatnost $\mathbb{P}(Y > y)$ puno *sporije* opada prema 0 kada $y \to \infty$ nego što je to slučaj npr. kod eksponencijalne razdiobe. Rep je "teži" što je parametar α manji, te za $\alpha \leqslant 1$ imamo čak i da je $\mathbb{E}[Y] = \infty$. Razdiobe teškog repa često se koriste pri procjeni rizika npr. u aktuarstvu i financijskoj industriji, te pri modeliranju raznih klimatoloških podataka.]

[NAPOMENA.* Formula (5.17) zapravo vrijedi za proizvoljnu nenegativnu slučajnu varijablu (dakle, i za diskretne, ali i sve druge nenegativne slučajne varijable). Na primjer, kada je $X \in \mathbb{N}_0$, imamo

$$\int_0^\infty \mathbb{P}(X > t)\,\mathrm{d}t = \sum_{n=0}^\infty \int_n^{n+1} \underbrace{\mathbb{P}(X > t)}_{=\mathbb{P}(X>n)}\,\mathrm{d}t = \sum_{n=0}^\infty \mathbb{P}(X > n)\,,$$

što je po Teoremu 3.28 upravo $\mathbb{E}[X]$. $\square$]

Ako je X neprekidna slučajna varijabla takva da je $\mathbb{E}\big[|X|\big] < \infty$, na isti način definiramo **varijancu** $\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$, te ponovno vrijede ista osnovna svojstva (Prop. 3.33)(ii) i (iii)) jer se u dokazu koristi isključivo linearnost očekivanja.

PRIMJER 5.30 (**Varijanca uniformne razdiobe**). Ako je $U \sim \mathrm{Unif}(0,1)$, imamo

$$\mathbb{E}[U^2] = \int_{-\infty}^\infty t^2 f_U(t)\,\mathrm{d}t = \int_0^1 t^2\,\mathrm{d}t = \frac{1}{3}\,,$$

pa je

$$\mathrm{Var}(U) = \mathbb{E}[U^2] - \mathbb{E}[U]^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}\,.$$

Nadalje, za $X := (b-a)U + a \sim \mathrm{Unif}(a,b)$, vrijedi

$$\mathrm{Var}(X) = (b-a)^2\mathrm{Var}(U) = \frac{(b-a)^2}{12}\,.$$

NAPOMENA 5.31 (**Nezavisnost**). Za neprekidne slučajne varijable $X_1, \dots, X_n$ kažemo da su **nezavisne** ako za sve $t_1, \dots, t_n \in \mathbb{R}$ vrijedi

$$\mathbb{P}(X_1 \leqslant t_1, \dots, X_n \leqslant t_n) = \prod_{i=1}^n \mathbb{P}(X_i \leqslant t_i)\,.$$

Može se pokazati da je gornje ekvivalentno s

$$\mathbb{P}(X_1 \in B_1, \dots, X_n \in B_n) = \prod_{i=1}^n \mathbb{P}(X_i \in B_i)\,,$$

za "skoro sve" $B_1, \dots, B_n \subseteq \mathbb{R}$, te da u tom slučaju opet vrijedi [uz uvjet da je sve dobro definirano]

- $\mathrm{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathrm{Var}(X_i)$;
- $\mathbb{E}[\prod_{i=1}^n X_i] = \prod_{i=1}^n \mathbb{E}[X_i]$.

PRIMJER 5.32 (**Minimum nezavisnih** Exp **varijabli**). Ako su $X \sim \mathrm{Exp}(\lambda)$ i $Y \sim \mathrm{Exp}(\mu)$ nezavisne, vrijedi $Z := \min\{X, Y\} \sim \mathrm{Exp}(\lambda + \mu)$.

DOKAZ. [Inutitivno, ovo slijedi iz priče o Poissonovom procesu jer je "zbroj" dva nezavisna Poissonova procesa s intezitetima (prosječan broj događaja po jedinici vremena) λ i μ ponovno Poissonov proces, ali s intezitetom $\lambda + \mu$.]

Formalno, za sve $t \geqslant 0$,

$$\mathbb{P}(Z > t) = \mathbb{P}(X > t, Y > t) = [\text{nez.}] = \mathbb{P}(X > t)\mathbb{P}(Y > t) = e^{-\lambda t}e^{-\mu t} = e^{-(\lambda+\mu)t} . \qquad \square$$

5.5. Normalna razdioba

[Normalna ili Gaussova razdiova je vjerojatno najpoznatija i najvažnija distribucija u vjerojatnosti. Razlog tomu je tzv. **centralni granični teorem** koji otprilike kaže da je zbroj velikog broja nezavisnih i jednako distribuiranih slučajnih varijabli (koje imaju konačnu varijancu) približno normalno distribuiran, *bez obzira* na inicijalnu razdiobu tih slučajnih varijabli.]

Kažemo da slučajna varijabla X ima **normalnu** (ili **Gaussovu**) razdiobu s parametrima μ i σ^2 (oznaka $X \sim \mathrm{N}(\mu, \sigma^2)$) za $\mu \in \mathbb{R}$, $\sigma > 0$, ako je neprekidna s gustoćom

$$f(t) = f_{\mathrm{N}(\mu,\sigma^2)}(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} , \quad t \in \mathbb{R} . \tag{5.18}$$

Razdiobu $\mathrm{N}(0, 1)$ nazivamo **standardna** normalna razdioba, te njenu gustoću/funkciju distribucije tipično označavamo s

$$\varphi(t) := f_{\mathrm{N}(0,1)}(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} , \quad t \in \mathbb{R}$$
$$\Phi(t) := F_{\mathrm{N}(0,1)}(x) = \int_{-\infty}^{x} \varphi(t)\, \mathrm{d}t \quad x \in \mathbb{R} .$$

[Nažalost, ne postoji jednostavni zatvoreni oblik za Φ, tj. nije ju moguće zapisati kao konačnu sumu nekih poznatih funkcija.]

Uočimo neka specijalna svojstva simetrije standardne normalne razdiobe:

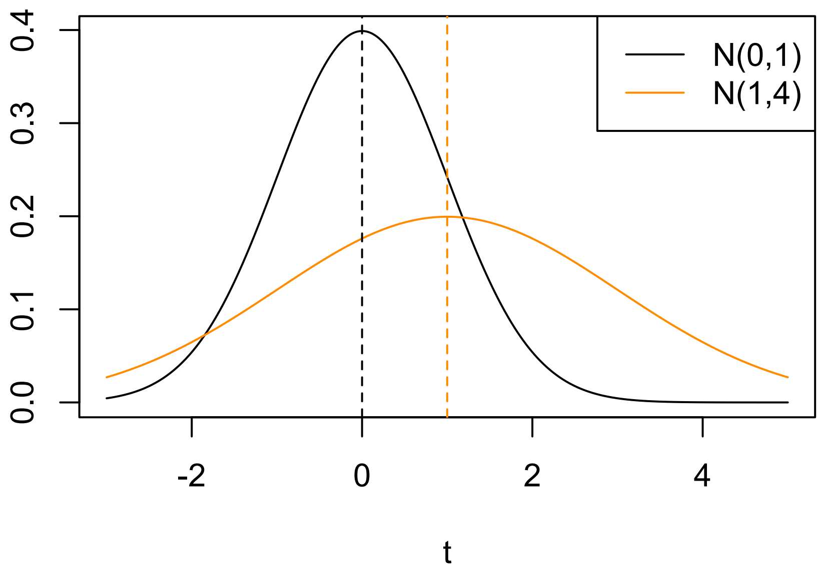
- $\varphi(-t) = \varphi(t)$, za sve $t \in \mathbb{R}$ [tj. φ je simetrična oko 0];
- ako je $Z \sim \mathrm{N}(0, 1)$, za sve $t_0 > 0$ vrijedi [skiciraj na grafu funkcije φ]

$$\Phi(-t_0) = \mathbb{P}(Z \leqslant -t_0) = \mathbb{P}(Z \geqslant t_0) = \mathbb{P}(Z > t_0) = 1 - \Phi(t_0) .$$

PROPOZICIJA 5.33. *Zaista vrijedi*

$$\int_{-\infty}^{\infty} f_{\mathrm{N}(\mu,\sigma^2)}(t)\, \mathrm{d}t = 1 , \quad \forall \mu \in \mathbb{R}, \sigma > 0 .$$

Dokaz koristi teorem o zamjeni varijabli u dvodimenzionalnom integralu pa se ne ispituje, te ga ostavljamo studentima da pogledaju sami.



SLIKA 7. Gustoće N$(0,1)$ i N$(1,4)$ razdiobe. One su simetrične oko parametra μ, a parametar σ kontrolira brzinu opadanja prema 0 kako t ide u $\pm\infty$.

DOKAZ.(*) Označimo gornji integral s I_{μ,σ^2}. Uz zamjenu varijabli $z = (t-\mu)/\sigma$ ($\mathrm{d}z = \frac{\mathrm{d}t}{\sigma}$) dobivamo $I_{\mu,\sigma^2} = I_{0,1}$ pa je dovoljno pokazati da je $I := I_{0,1} = 1$. Trik je gledati I^2:

$$\begin{aligned} I^2 &= \frac{1}{2\pi}\left(\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}\,\mathrm{d}z\right)\cdot\left(\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\,\mathrm{d}y\right) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{z^2+y^2}{2}}\,\mathrm{d}z\,\mathrm{d}y \\ &= [\text{zamjena u tzv. polarne koordinate} - r = \sqrt{z^2+y^2},\ \theta = \text{kut između } (z,y) \text{ i } x\text{-osi}] \\ &= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} e^{-\frac{r^2}{2}}\cdot r\,\mathrm{d}r\,\mathrm{d}\theta = \frac{1}{2\pi}\cdot\int_0^{2\pi}\mathrm{d}\theta\cdot\int_0^{\infty} e^{-\frac{r^2}{2}}\cdot r\,\mathrm{d}r = \frac{1}{2\pi}\cdot 2\pi\cdot 1 = 1\,. \qquad \square \end{aligned}$$

Odredimo sada očekivanje i varijancu standardne normalne razdiobe. Najprije, uočimo da je za $Z \sim \mathrm{N}(0,1)$,

$$\mathbb{E}\big[|Z|\big] = \int_{-\infty}^{\infty} |t|\frac{1}{2\pi}e^{-\frac{t^2}{2}}\,\mathrm{d}t = \frac{2}{2\pi}\int_0^{\infty} te^{-\frac{t^2}{2}}\,\mathrm{d}t = \frac{2}{2\pi}\cdot 1 < \infty$$

pa dakle postoji $\mathbb{E}[Z]$, te očito iznosi

$$\mathbb{E}[Z] = \int_{-\infty}^{\infty} t\cdot\varphi(t)\,\mathrm{d}t = 0$$

jer je $t \mapsto t\cdot\varphi(t)$ neparna funkcija. Nadalje,

$$\mathrm{Var}(Z) = \mathbb{E}[Z^2] = \int_{-\infty}^{\infty} t^2\frac{1}{2\pi}e^{-\frac{t^2}{2}}\,\mathrm{d}t = \frac{2}{\sqrt{2\pi}}\int_0^{\infty} t^2\frac{1}{2\pi}e^{-\frac{t^2}{2}}\,\mathrm{d}t$$

pa korištenjem parcijalne integracije uz $u = t$ i $\mathrm{d}v = te^{-\frac{t^2}{2}}$ ($\mathrm{d}u = 1$, $v = -e^{-\frac{t^2}{2}}$) dobivamo

$$\mathrm{Var}(Z) = \frac{2}{\sqrt{2\pi}} \cdot \left(-te^{-\frac{t^2}{2}} \Big|_0^\infty + \int_0^\infty e^{-\frac{t^2}{2}} \,\mathrm{d}t \right) = \frac{2}{\sqrt{2\pi}} \cdot \left(0 + \sqrt{2\pi} \underbrace{\int_0^\infty \varphi(t)\,\mathrm{d}t}_{=\frac{1}{2}} \right) = 1\,.$$

Propozicija 5.34 (**Zatvorenost normalne razdiobe na lin. transformacije**). *Za $Z \sim \mathrm{N}(0,1)$ i sve $\mu \in \mathbb{R}, \sigma > 0$, slučajna varijabla $X := \mu + \sigma \cdot Z$ ima $\mathrm{N}(\mu, \sigma^2)$ razdiobu.*

Dokaz. Slučajna varijabla X je neprekidna [npr. jer $g(z) = \mu + \sigma z$ zadovoljava uvjete Prop. 5.19]. Njena funkcija distribucije je

$$F_X(x) = \mathbb{P}(\mu + \sigma \cdot Z \leqslant x) = \mathbb{P}\left(Z \leqslant \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right), \quad x \in \mathbb{R}\,,$$

pa je gustoća

$$f_X(x) = F_X'(x) = \Phi'\left(\frac{x-\mu}{\sigma}\right) \cdot \left(\frac{x-\mu}{\sigma}\right)' = \varphi\left(\frac{x-\mu}{\sigma}\right) \cdot \frac{1}{\sigma}\,,$$

što je po (5.18) upravo $f_{\mathrm{N}(\mu,\sigma^2)}(x)$, za sve $x \in \mathbb{R}$. □

Iz prethodnog odmah slijedi da za $X \sim \mathrm{N}(\mu, \sigma^2)$ vrijedi

$$\begin{aligned} \mathbb{E}[X] &= \mu + \sigma\mathbb{E}[Z] = \mu\,, \\ \mathrm{Var}(X) &= \sigma^2 \mathrm{Var}(Z) = \sigma^2\,, \end{aligned}$$

tj. parametri μ i σ^2 su upravo očekivanje i varijanca. Nadalje, za sve $a \neq 0$, $b \in \mathbb{R}$, vrijedi

$$aX + b \sim \mathrm{N}(a\mu + b, a^2\sigma^2)\,,$$

te specijalno

$$\frac{X - \mu}{\sigma} \sim \mathrm{N}(0,1)\,.$$

Zadnju slučajnu varijablu često zovemo **standardizirana** verzija od X.

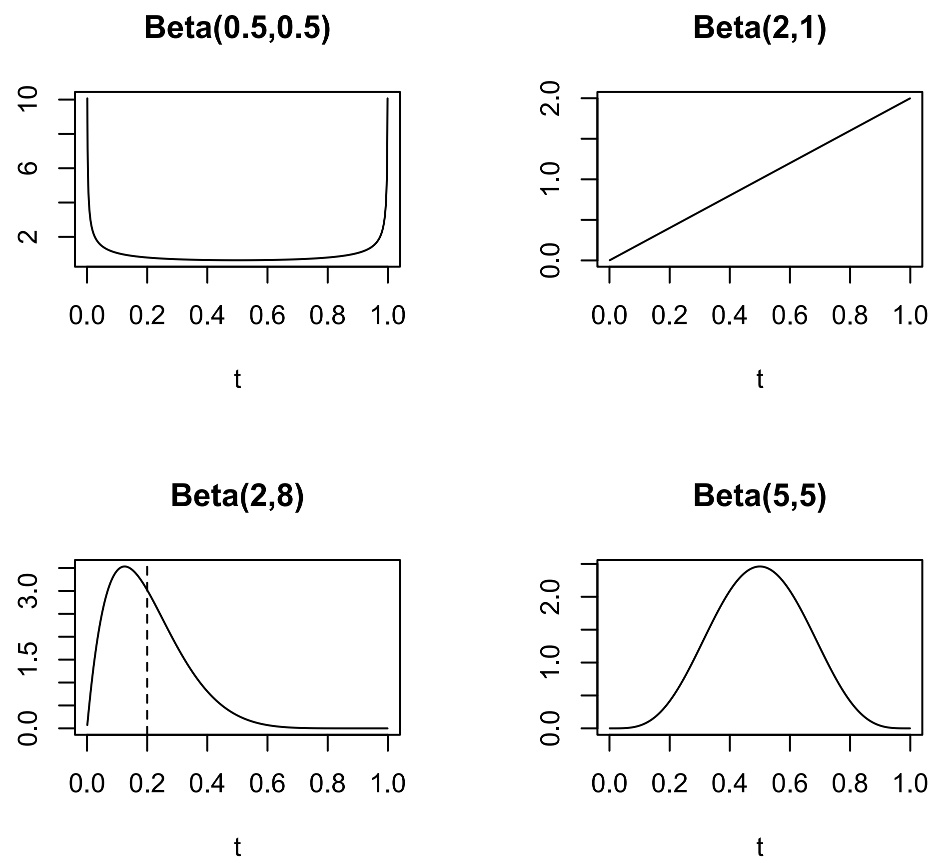
5.6. Beta razdioba

Kažemo da slučajna varijabla X ima **beta** razdiobu s parametrima $a, b > 0$ (oznaka $X \sim \mathrm{Beta}(a,b)$) ako je neprekidna s gustoćom

$$f(t) = \frac{1}{\beta(a,b)} t^{a-1}(1-t)^{b-1}\,, \quad t \in (0,1)\,, \tag{5.19}$$

uz $f(t) = 0$ za $t \notin (0,1)$, gdje je $\beta(a,b) := \int_0^1 t^{a-1}(1-t)^{b-1}\,\mathrm{d}t$ normalizirajuća konstanta [β je tzv. beta funkcija]. Ovo je dakle slučajna varijabla koja poprima vrijednosti u $(0,1)$, te je očito $\mathrm{Beta}(1,1) = \mathrm{Unif}(0,1)$. [Variranjem parametara možemo dobiti vrlo različite gustoće, vidi Sliku 8.] Može se pokazati da za $X \sim \mathrm{Beta}(a,b)$ vrijedi

$$\mathbb{E}[X] = \frac{a}{a+b}.$$



Slika 8. Gustoća $\mathrm{Beta}(a, b)$ za nekoliko kombinacija parametara $a, b > 0$. U slučaju $a = 2, b = 8$, vertikalnom linijom je naznačeno očekivanje $a/(a+b) = 0.2$.

Primjer 5.35 (**Bayesovska statstika***). [Bacamo novčić n puta te želimo procijeniti vjerojatnost $p \in [0, 1]$ za pojavu pisma, na temelju ukupnog broja pisama koji su pali N_n.]

Klasični ("frekvencionistički") pristup: Parametar p je fiksan i nepoznat, $N_n \sim \mathrm{B}(n, p)$, te npr. metodom maksimalne vjerodostojnosti dođemo do procjenitelja $\widehat{p}_n := \frac{N_n}{n}$; vidi Primjer 4.8.

Bayesovski pristup: Parametar p je *slučajan*, npr. pretpostavljamo da je

- $P \sim \mathrm{Beta}(a, b)$ za neke $a, b > 0$ (tzv. *apriorna* distribucija)[10], te
- uvjetno na $P = p$, N_n ima $\mathrm{B}(n, p)$ razdiobu.[11]

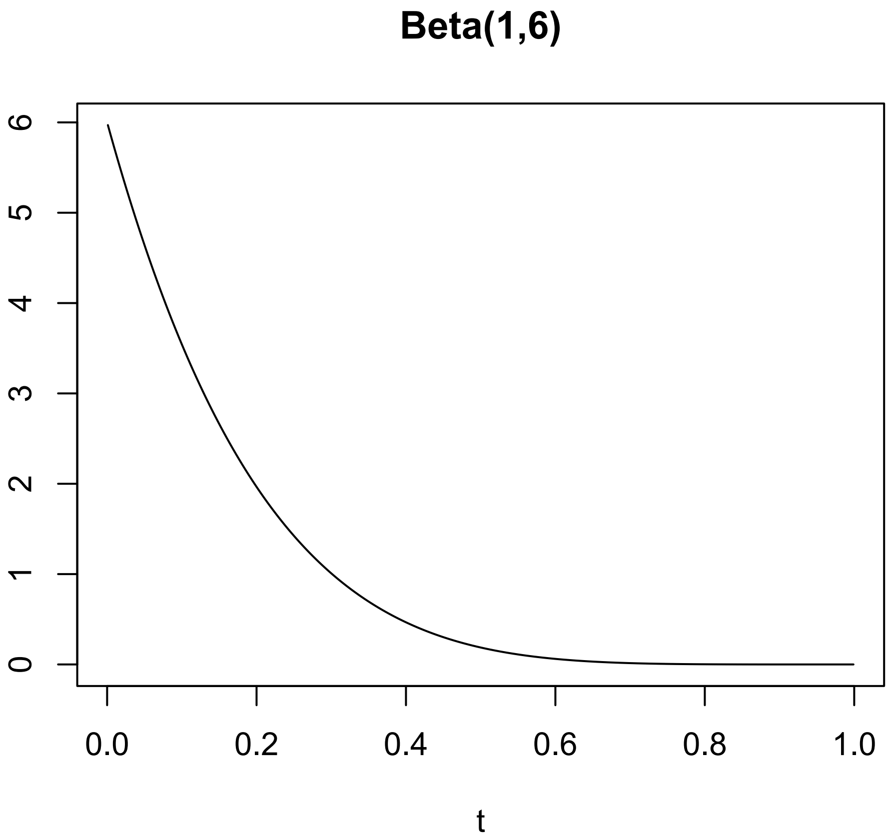
Ako je $N_n = k$, koristeći prikladnu verziju Bayesovog teorema dobivamo da je tzv. *aposteriorna* gustoća za P jednaka

$$\begin{aligned} f_P(p \mid N_n = k) &= \frac{\mathbb{P}(N_n = k \mid P = p) f_P(p)}{\mathbb{P}(N_n = k)} = \frac{\binom{n}{k} p^k (1-p)^{n-k} \cdot \frac{1}{\beta(a,b)} p^{a-1}(1-p)^{b-1}}{\mathbb{P}(N_n = k)} \\ &= c \cdot p^{a+k-1}(1-p)^{b+n-k-1}, \quad p \in (0, 1), \end{aligned}$$

[10] Apriorna razdioba predstavlja našu početnu informaciju o problemu – npr. ako ništa ne znamo, možemo staviti da je to $\mathrm{Beta}(1, 1) = \mathrm{Unif}(0, 1)$ razdioba.

[11] Intutivno je jasno što ovo znači, ali formalno nismo definirali ovakvo uvjetovanje jer se radi o uvjetovanju na događaj vjerojatnosti 0!

pri čemu je c konstanta koja ne ovisi o p. Dakle, dobili smo da uvjetno na $N_n = k$, P ima Beta$(a+k, b+n-k)$ razdiobu![12] Na primjer, ako je $P \sim \text{Beta}(1,1) = \text{Unif}(0,1)$ te je palo $N_n = 0$ pisama, aposteriorna razdioba je Beta$(1, n+1)$ – ako želimo procjenitelj za p, možemo uzeti očekivanje ove razdiobe, što dakle iznosi $\frac{1}{n+2}$. Uočimo da je klasični procjenitelj u ovom slučaju jednak $\frac{N_n}{n} = 0$.



Slika 9. Aposteriorna razdioba za P u Primjeru 5.35, i to u slučaju uniformne apriorne razdiobe, $n = 5$ i $N_n = 0$. Dakle, u Bayesovskom pristupu dobivamo cijelu razdiobu za parametra koji procjenjujemo!

[12] To da aposteriorna razdioba ostaje u istoj familiji razdioba kao i apriorna značajno olakšava računanje, te kažemo da je beta razdioba *konjugirana* apriorna razdioba za binomnu distribuciju. U mnogim drugim situacijama, određivanje aposteriorne razdiobe je netrivijalan problem te se za to koriste tzv. **Markov Chain Monte Carlo** (MCMC) metode.

POGLAVLJE 6

Funkcije izvodnice

[Funkcije izvodnice (FI) su analitički alat za baratanje sa slučajnim varijablama – prije svega, korisne su analizi zbroja *nezavisnih* slučajnih varijabli. FI se javljaju u različitim oblicima

- ako je $X \in \mathbb{N}_0$ – možemo koristiti FI *vjerojatnosti*;
- ako je $X \geqslant 0$ – možemo koristiti *Laplaceov funkcional*;
- ako je $X \in \mathbb{R}$ te vrijede neki dodatni uvjeti na momente od X – možemo koristiti FI *momenata*;
- za proizvoljnu $X \in \mathbb{R}$ – možemo koristiti *karakteristične funkcije.*

Ipak, ideja je uvijek vrlo slična.]

6.1. Funkcije izvodnice vjerojatnosti

U ovom poglavlju promatramo samo slučajne varijable koje poprimaju vrijednosti u $\mathbb{N}_0 = \{0, 1, \dots\}$.

DEFINICIJA 6.1. Neka je $X \in \mathbb{N}_0$ slučajna varijabla, te $p_n := \mathbb{P}(X = n)$, $n \in \mathbb{N}_0$ [njena distribucija].[1] **Funkcija izvodnica (vjerojatnosti)** od X je funkcija

$$G(s) := \mathbb{E}[s^X] = \sum_{n=0}^{\infty} p_n s^n \,, \tag{6.1}$$

definirana za sve

$$s \in S := \{t \in \mathbb{R} : \sum_{n=0}^{\infty} p_n |t|^n < \infty\} \,. \tag{6.2}$$

U nastavku ćemo pisati $G_X := G$ i $S_X := S$.

NAPOMENA 6.2. (a) Ako je $|s| \leqslant 1$,

$$|G(s)| \leqslant \sum_{n=0}^{\infty} p_n |s|^n \leqslant \sum_{n=0}^{\infty} p_n = 1 \,. \tag{6.3}$$

Dakle, uvijek je $[-1, 1] \subseteq S$, te vrijedi $G(1) = 1$ i $G(0) = p_0$.

(b) (!) Funkcija G je zapravo *red potencija* (oko 0) s radijusom konvergencije $s_0 := \sup S$ za koji znamo da vrijedi $s_0 \geqslant 1$. Specijalno, znamo da je $G \in \mathrm{C}^\infty(\langle -s_0, s_0 \rangle)$ te da derivacije od G možemo dobiti derivirajući "član po član", tj. za sve $k \geqslant 0$ vrijedi

$$G^{(k)}(s) = \sum_{n=k}^{\infty} p_n n(n-1) \cdots (n-k+1) s^{n-k} \,, \quad \forall |s| < s_0 \,. \tag{6.4}$$

[1] Ovo je notacija koju ćemo koristiti i u nastavku (ako ne kažemo drugačije).

Specijalno, $\forall k \in \mathbb{N}_0$, $G^{(k)}(0) = p_k k!$, tj.

$$\mathbb{P}(X = k) = p_k = \frac{G^{(k)}(0)}{k!}\,. \tag{6.5}$$

Teorem 6.3 (**FI jedinstveno određuje distribuciju**). *Ako su X i Y slučajne varijable u $\mathbb{N}_0$ t.d. vrijedi $G_X = G_Y$ [dovoljno na nekoj okolini oko 0], tada je $\mathbb{P}(X = k) = \mathbb{P}(Y = k)$, $\forall k \in \mathbb{N}_0$, tj. X i Y imaju istu distribuciju.*

Dokaz. Slijedi iz (6.5) jer $G_X^{(k)}(0) = G_Y^{(k)}(0)$ za sve $k \in \mathbb{N}_0$. □

Primjer 6.4 (**FI vjerojatnosti za poznate razdiobe**). (a) Ako je $X \sim \mathrm{B}(p)$,

$$G_X(s) = [p_0 = 1 - p =: q, p_1 = p, p_n = 0 \text{ za } n \geqslant 2] = qs^0 + ps^1 = q + ps\,, \quad \forall s \in \mathbb{R}\,.$$

(b) Ako je $X \sim \mathrm{B}(m, p)$,

$$G_x(s) = \sum_{n=0}^{m} \binom{m}{n} p^n q^{n-m} \cdot s^n = \sum_{n=0}^{m} \binom{m}{n} (ps)^n q^{n-m} = (q + ps)^m\,, \quad \forall s \in \mathbb{R}\,.$$

(c) Ako je $X \sim \mathrm{G}_0(p)$ za $p > 0$,

$$G_x(s) = \sum_{n=0}^{\infty} q^n p \cdot s^n = p \sum_{n=0}^{\infty} (qs)^n = p \cdot \frac{1}{1 - qs} = \frac{p}{1 - qs}\,,$$

za sve s t.d. $|qs| < 1$, tj. $|s| < \frac{1}{q}$.

[(d) i (e) za vježbu]

(d) Ako je $X \sim \mathrm{G}(p)$ za $p > 0$, $X - 1 \sim \mathrm{G}_0(p)$ pa je

$$G_x(s) = \mathbb{E}[s^{X-1+1}] = s\mathbb{E}[s^{X-1}] = \frac{ps}{1 - qs}\,, \quad \forall |s| \leqslant \frac{1}{q}\,.$$

(e) Ako je $X \sim \mathrm{P}(\lambda)$,

$$G_x(s) = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda} \cdot s^n = e^{-\lambda} \sum_{n=0}^{\infty} \frac{(\lambda s)^n}{n!} = e^{-\lambda} \cdot e^{-\lambda s} = e^{\lambda(s-1)}\,, \quad \forall s \in \mathbb{R}\,.$$

6.1.1. Momenti.

Teorem 6.5. *Ako je X slučajna varijabla u $\mathbb{N}_0$, za sve $k \in \mathbb{N}$ vrijedi*

$$G_X^{(k)}(1) = \mathbb{E}[X(X-1)\cdots(X-k+1)]\,, \tag{6.6}$$

pri čemu u slučaju $s_0 = 1$ definiramo[2]

$$G_X^{(k)}(1) := \lim_{s \to 1^-} G_X^{(k)}(s) \quad \in [0, \infty]\,. \tag{6.7}$$

Dokaz. Uvijek vrijedi

$$\begin{aligned}\mathbb{E}[X(X-1)\cdots(X-k+1)] &= \sum_{n=0}^{\infty} p_n n(n-1)\cdots(n-k+1) \\ &= \sum_{n=k}^{\infty} p_n n(n-1)\cdots(n-k+1) \quad \in [0, \infty]\,.\end{aligned}$$

[2] Zapravo, jednakost u (6.7) vrijedi i ako je $s_0 > 1$ jer je u tom slučaju G_X neprekidna na $\langle -s_0, s_0 \rangle$.

pa ako je $s_0 > 1$, (6.6) slijedi iz (6.4) jer je $1 < s_0$ (te je gornje očekivanje nužno konačno). Ako je $s_0 = 1$, rezultat slijedi iz Abelovog teorema[3] koji kaže da je

$$\lim_{s\to 1^-} \sum_{n=k}^{\infty} p_n n(n-1)\cdots(n-k+1)\cdot s^n = \sum_{n=k}^{\infty} p_n n(n-1)\cdots(n-k+1) \quad \in [0,\infty]\,.$$

□

Primjer 6.6. Za $X \sim \mathrm{P}(\lambda)$ znamo da je $G_x(s) = e^{\lambda(s-1)}$, $s \in \mathbb{R}$, pa iz

$$G_X'(s) = \lambda e^{\lambda(s-1)}\,, \quad G_X''(s) = \lambda^2 e^{\lambda(s-1)}\,, \quad s \in \mathbb{R}\,,$$

slijedi

- $\mathbb{E}[X] = G'(1) = \lambda$,
- $\mathbb{E}[X^2] - \mathbb{E}[X] = \mathbb{E}[X(X-1)] = G''(1) = \lambda^2$, te
- $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = (\lambda^2+\lambda) - \lambda^2 = \lambda$.

6.1.2. Zbroj nezavisnih slučajnih varijabli.

Teorem 6.7 (**FI zbroja nezavisnih varijabli**). *Ako su $X_1, \dots, X_n$ nezavisne slučajne varijable u $\mathbb{N}_0$, za sve $s \in \cap_{i=1}^n S_{X_i} \supseteq [-1,1]$ vrijedi*

$$G_{X_1+\cdots+X_n}(s) = G_{X_1}(s)\cdots G_{X_n}(s)\,. \tag{6.8}$$

Dokaz. Za sve $s \in \cap_{i=1}^n S_{X_i} \supseteq [-1,1]$ imamo

$$\begin{aligned} G_{X_1+\cdots+X_n}(s) &= \mathbb{E}[s^{X_1+\cdots+X_n}] = \mathbb{E}[s^{X_1}\cdots s^{X_n}] = [\text{nezavisnost}] \\ &= \mathbb{E}[s^{X_1}]\cdots\mathbb{E}[s^{X_n}] = G_{X_1}(s)\cdots G_{X_n}(s)\,. \end{aligned}$$

□

Primjer 6.8. Ako su $X \sim \mathrm{P}(\lambda)$ i $Y \sim \mathrm{P}(\mu)$ nezavisne, vrijedi

$$G_{X+Y}(s) = G_X(s)\cdot G_Y(s) = e^{\lambda(s-1)}e^{\mu(s-1)} = e^{(\lambda+\mu)\cdot(s-1)} = G_{\mathrm{P}(\lambda+\mu)}(s)\,,\ s \in \mathbb{R}\,.$$

Teorem 6.3 povlači da je nužno $X + Y \sim \mathrm{P}(\lambda+\mu)$ [usporedi s dokazom Prop. 4.36].

Primjer 6.9. Ako su $X_1, \dots, X_n$ njd t.d. $X_i \sim \mathrm{B}(p)$, budući da je $X := X_1 + \cdots + X_n \sim \mathrm{B}(n,p)$ odmah dobivamo da je FI $\mathrm{B}(n,p)$ razdiobe dana s

$$G_X(s) = [X_1,\dots,X_n \text{ su njd}] = (G_{X_1}(s))^n = (q+ps)^n\,, \quad s \in \mathbb{R}\,.$$

Teorem 6.10 (**FI slučajne sume**). *Neka je*

- *$X_1, X_2, \dots$ niz njd slučajnih varijabli u $\mathbb{N}_0$ te $S_n := \sum_{i=1}^n X_i$, $n \geqslant 1$, uz $S_0 := 0$, te*
- *N slučajna varijabla u $\mathbb{N}_0$ nezavisna od niza $X_1, X_2, \dots$.*

[3] **Abelov teorem**: Ako je $u_0, u_1, \dots$ nenegativan niz t.d. red $\sum_{n\geqslant 0} u_n s^n$ apsolutno konvergira za sve $|s| < 1$, tada vrijedi

$$\lim_{s\to 1^-} \sum_{n\geqslant 0} u_n s^n = \sum_{n\geqslant 0} u_n \ \in [0,\infty]\,.$$

Tada za FI od S_N (defin. s $S_N(\omega) := S_{N(\omega)}$, $\omega \in \Omega$) vrijedi

$$G_{S_N}(s) = G_N(G_{X_1}(s)), \quad \forall |s| \leqslant 1. \tag{6.9}$$

Specijalno, vrijedi

$$\mathbb{E}[S_N] = \mathbb{E}[N]\mathbb{E}[X_1] \in [0, \infty]. \tag{6.10}$$

Dokaz. Za sve $|s| \leqslant 1$ [barem] imamo

$$\begin{aligned} G_{S_N}(s) = \mathbb{E}[s^{S_N}] &= \sum_{n=0}^{\infty} \overbrace{\mathbb{E}[s^{S_N} \mid N = n]}^{=\mathbb{E}[s^{S_n}|N=n]} \mathbb{P}(N = n) = [N \text{ i } S_n \text{ nezavisne!}] \\ &= \sum_{n=0}^{\infty} \mathbb{E}[s^{S_n}]\mathbb{P}(N = n) = [X_1, \ldots, X_n \text{ su njd}] \\ &= \sum_{n=0}^{\infty} (G_{X_1}(s))^n \mathbb{P}(N = n) = G_N(G_{X_1}(s)). \end{aligned}$$

Specijalno,

$$\begin{aligned} \mathbb{E}[S_N] &= G'_{S_N}(1) = (G_N(G_{X_1}(s)))'\Big|_{s=1} = G'_N(G_{X_1}(s)) \cdot G'_{X_1}(s)\Big|_{s=1} \\ &= G'_N(1) \cdot G'_{X_1}(1) = \mathbb{E}[N] \cdot \mathbb{E}[X_1]. \end{aligned}$$

[Gornji argument zapravo pretpostavlja da su radijusi konvergencije od G_N i G_{X_1} strogo veći od 1. Općenito,

$$\begin{aligned} \mathbb{E}[S_N] &= \lim_{s \to 1^-} G'_{S_N}(s) = \lim_{s \to 1^-} G'_N(G_{X_1}(s)) \cdot G'_{X_1}(s) \\ &= [\lim_{s \to 1^-} G_{X_1}(s) = G_{X_1}(1) = 1] = \mathbb{E}[N] \cdot \mathbb{E}[X_1]. \end{aligned}$$

pri čemu bilo koje od očekivanja u zadnjem izrazu može biti $+\infty$.] □

Zadatak. Ako je $\mathbb{E}[S_N] = \mathbb{E}[N]\mathbb{E}[X_1] < \infty$, pokažite da vrijedi

$$\mathrm{Var}(S_N) = \mathbb{E}[N]\mathrm{Var}(X_1) + \mathbb{E}[X_1]^2 \mathrm{Var}(N). \tag{6.11}$$

[Intuicija?] Pogledajmo na šte svodi desna strana u dva specijalna slučaja:

- ako je $N \equiv k \in \mathbb{N}_0$, desna strana je $k\mathrm{Var}(X_1) + 0 = \mathrm{Var}(S_k) = \mathrm{Var}(S_N)$;
- ako je $X_i \equiv c \in \mathbb{N}_0$, desna strana je $0 + c^2\mathrm{Var}(N) = \mathrm{Var}(cN) = \mathrm{Var}(S_N)$.

Primjer 6.11 (**Stanjivanje Poissonove slučajne varijable**). Pretpostavimo da imamo $N \sim \mathrm{P}(\lambda)$ kuglica, te da svaku kuglicu, nezavisno od drugih, obojamo u plavo s vjerojatnošću p ili u crveno s vjerojatnošću $q := 1 - p$ ($p \in (0, 1)$). Ako je N_P ukupan broj plavih, a N_C ukupan broj crvenih kuglica, odredite razdiobe od N_P i N_C.

Rješenje. Neka je

$$X_i := \begin{cases} 1\,, & \text{ako je } i\text{-ta kuglica obojana u plavo,} \\ 0\,, & \text{inače} \end{cases}, \quad i = 1, 2, \ldots$$

[Treba nam beskonačan niz X_i-eva jer ne znamo unaprijed koliko je N.]

Po konstrukciji su $X_1, X_2, \ldots$ njd slučajne varijable t.d. $X_i \sim \mathrm{B}(p)$, te vrijedi

$$N_P = \sum_{i=1}^{N} X_i\,.$$

Iz (6.9) slijedi

$$\begin{aligned} G_{N_P}(s) &= G_N(G_{X_1})(s) = G_N(q+ps) = [G_N(t) = e^{\lambda(t-1)}] \\ &= e^{\lambda(q+ps-1)} = [q = 1-p] = e^{\lambda p(s-1)}\,, \ |s| \leqslant 1\,. \end{aligned}$$

Dakle, N_P ima $\mathrm{P}(\lambda p)$ razdiobu, pa je specijalno

$$\mathbb{E}[N_P] = \lambda p \ (= \mathbb{E}[N]\mathbb{E}[X_1])\,.$$

Analogno se pokaže $N_C \sim \mathrm{P}(\lambda q)$. To da nakon što za svaku od $N \sim \mathrm{P}(\lambda)$ kuglica bacimo novčić te s vjerojatnošću p obojimo u plavo, te nakon toga ukupan broj plavih kuglica i dalje ima Poissonovu razdiobu, zove se svojstvo "stanjivanja" (engl. *thinning*) Poissonove razdiobe. □

Ono što može pokazati (DZ) jest da su prethodnom primjeru N_P i N_C zapravo *nezavisne*. To je iznenađujuće budući da je $N_P + N_C = N$ – ovdje je ključno da je N slučajan te da ima baš Poissonovu razdiobu.

6.2. Funkcija izvodnica momenata

Definicija 6.12. Neka je X (bilo kakva) slučajna varijabla. Ako postoji $t_0 > 0$ za koji vrijedi

$$\mathbb{E}[e^{tX}] < \infty\,, \quad \forall |t| \leqslant t_0\,, \tag{6.12}$$

funkcija $M : [-t_0, t_0] \to [0, \infty)$ definirana sa

$$M(t) := \mathbb{E}[e^{tX}]\,, \tag{6.13}$$

zove se **funkcija izvodnica momenata** slučajne varijable X.

U nastavku ćemo često pisati $M_X := M$ kako bi naglasili da se radi o FI momenata za slučajnu varijablu X.

Napomena 6.13. Uvjet (6.12) ekvivalentan je uvjetu

$$\mathbb{E}[e^{t_0|X|}] < \infty\,. \tag{6.14}$$

To slijedi jer za sve $|t| \leqslant t_0$ vrijedi

$$e^{tX} \leqslant e^{t_0|X|} \leqslant \max\{e^{t_0X}, e^{-t_0X}\} \leqslant e^{t_0X} + e^{t_0X}\,,$$

pa specijalno i

$$\mathbb{E}[e^{tX}] \leqslant \mathbb{E}[e^{t_0|X|}] \leqslant \mathbb{E}[e^{t_0X}] + \mathbb{E}[e^{t_0X}]\,.$$

Teorem 6.14. *Pretpostavimo da postoji $t_0 > 0$ takav da vrijedi (6.12).*

*(a) Za sve $k \in \mathbb{N}_0$ je $\mathbb{E}\Big[|X|^k\Big] < \infty$, pa specijalno postoji **k-ti moment od X** definiran s $\mu_k := \mathbb{E}[X^k]$.*

(b) Za sve $|t| \leqslant t_0$ vrijedi[4]

$$M_X(t) = \sum_{k=0}^{\infty} \frac{\mu_k}{k!} t^k\,.$$

(c) Za sve $k \in \mathbb{N}_0$ vrijedi $M^{(k)}(0) = \mu_k$.

Dokaz. [Dokaz je baziran na jednakosti $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$, $x \in \mathbb{R}$.]

(a) Imamo

$$e^{t_0|X|} = \sum_{n=0}^{\infty} \frac{t_0^n|X|^n}{n!} \geqslant \frac{t_0^k}{k!}|X|^k\,, \quad \forall k \in \mathbb{N}_0\,. \tag{6.15}$$

Monotonost vjerojatnosti povlači da je

$$\frac{t_0^k}{k!}\mathbb{E}\Big[|X|^k\Big] \leqslant \mathbb{E}[e^{t_0|X|}] \overset{(6.14)}{<} \infty\,, \quad \forall k \in \mathbb{N}_0\,.$$

Tvrdnja sada slijedi budući da je $t_0 > 0$.

(b) Slično, za sve $|t| \leqslant t_0$ vrijedi

$$M_X(t) = \mathbb{E}[e^{tX}] = \mathbb{E}\left[\sum_{n=0}^{\infty} \frac{t^n X^n}{n!}\right].$$

Napomena*. Ono što bismo htjeli ovdje je zamijeniti očekivanje i sumu. To nam omogućava prikladna verzija Fubinijevog teorema[5] budući da je

$$\mathbb{E}\left[\sum_{n=0}^{\infty} \frac{|t|^n|X|^n}{n!}\right] = \mathbb{E}[e^{|t|\cdot|X|}] \leqslant \mathbb{E}[e^{t_0\cdot|X|}] \overset{(6.14)}{<} \infty\,. \qquad \square$$

[4] Zato se M_X zove funkcija izvodnica **momenata**.

[5] **Fubinijev teorem:** Neka je $X_1, X_2, \ldots$ niz slučajnih varijabli definiran na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$. Ako je $X_i \geqslant 0$, za sve $i \geqslant 1$, vrijedi

$$\mathbb{E}\left[\sum_{n=1}^{\infty} X_n\right] = \sum_{n=1}^{\infty} \mathbb{E}\left[X_n\right]\,,$$

pri čemu obje strane mogu biti $+\infty$. Općenito, gornja jednakost vrijedi ako imamo

$$\mathbb{E}\left[\sum_{n=1}^{\infty} |X_n|\right] = \sum_{n=1}^{\infty} \mathbb{E}[|X_n|] < \infty\,.$$

Dakle,

$$M_X(t) = \sum_{n=0}^{\infty} \mathbb{E}\left[\frac{t^n X^n}{n!}\right] = \sum_{n=0}^{\infty} \frac{\mu_n}{n!} t^n, \quad \forall |t| \leqslant t_0 . \tag{6.16}$$

(c) Ovo slijedi iz (6.16) – detalje ostavljamo za vježbu.

[Budući da je M red potencija (oko 0), slijedi da je M klase C^∞ na $(-t_0, t_0)$, te da je za sve $k \geqslant 0$,

$$M_X^{(k)}(t) = \sum_{n=k}^{\infty} \frac{\mu_n}{(n-k)!} t^{n-k}, \quad \forall |t| < t_0 . \tag{6.17}$$

Specijalno,

$$M_X^{(k)}(0) = \mu_k \cdot 1 + 0 = \mu_k .]$$

$\square$

Primjer 6.15 (**Eksponencijalna razdioba**). Za $X \sim \mathrm{Exp}(\lambda)$ vrijedi

$$\begin{aligned} M_X(t) = \mathbb{E}[e^{tX}] &= \int_{-\infty}^{\infty} e^{tx} f_X(x) \mathrm{d}x = \int_0^{\infty} e^{tx} \lambda \mathrm{d}x = \lambda \int_0^{\infty} e^{-(\lambda - t)x} \mathrm{d}x \\ &= [\text{ako je } \lambda > t] = \frac{\lambda}{\lambda - t} \int_0^{\infty} \underbrace{(\lambda - t) e^{-(\lambda - t)x}}_{= f_{\mathrm{Exp}(\lambda - t)}(x)} \mathrm{d}x = \frac{\lambda}{\lambda - t}, \quad t < \lambda := t_0 . \end{aligned}$$

[Formalno bi trebali uzeti $t_0 := \lambda - \epsilon > 0$ neki mali ϵ tako da vrijedi (6.12).] Specijalno, iz $M_X'(t) = \frac{\lambda}{(\lambda - t)^2}$ slijedi da je

$$\mathbb{E}[X] = M_X'(0) = \frac{1}{\lambda} .$$

Daljnjim deriviranjem možemo dobiti i $\mathbb{E}[X^k]$ za $k = 2, 3, \dots$, ali pogledajmo puno elegantniji pristup.

Ako je $E \sim \mathrm{Exp}(1)$, za sve $|t| < 1$ (specijalno je dakle $t < 1$) imamo

$$M_E(t) = \frac{1}{1 - t} = \sum_{n=0}^{\infty} t^n = \sum_{n=0}^{\infty} \frac{t^n}{n!} \cdot n! .$$

Teorem 6.14 povlači da je $\mathbb{E}[E^n] = n!$, za sve $n \geqslant 0$. Općenito, budući da $X \sim \mathrm{Exp}(\lambda)$ ima istu razdiobu kao i $\frac{E}{\lambda}$, odmah dobivamo da je

$$\mathbb{E}[X^n] = \frac{n!}{\lambda^n}, \quad \forall n \in \mathbb{N}_0 .$$

Primjer 6.16 (**Normalna razdioba**). Za $Z \sim \mathrm{N}(0, 1)$ imamo

$$\begin{aligned} M_Z(t) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \cdot e^{tz} \mathrm{d}z = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z^2 - 2tz + t^2) + \frac{t^2}{2}} \cdot e^{tz} \mathrm{d}z \\ &= e^{\frac{t^2}{2}} \int_{-\infty}^{\infty} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{(z-t)^2}{2}}}_{= f_{\mathrm{N}(t,1)}(z)} \mathrm{d}z = e^{\frac{t^2}{2}}, \quad \forall t \in \mathbb{R} . \end{aligned} \tag{6.18}$$

Za $X \sim \mathrm{N}(\mu, \sigma^2)$, budući da vrijedi $X \sim \sigma Z + \mu$, imamo

$$M_X(t) = \mathbb{E}[e^{tX}] = \mathbb{E}[e^{t\sigma Z + t\mu}] = e^{t\mu}\mathbb{E}[e^{t\sigma Z}] = e^{t\mu}M_Z(t\sigma) \overset{(6.18)}{=} e^{\mu t + \frac{\sigma^2}{2}t^2}, \quad \forall t \in \mathbb{R}. \tag{6.19}$$

Primjer 6.17 (**FI momenata ne postoji uvijek**). Ako je $X \sim \mathrm{Pareto}(\alpha)$ za $\alpha > 0$ proizvoljan (vidi Primjer 5.29), lako se pokaže da vrijedi $\mathbb{E}[X^\alpha] = \infty$. Specijalno, k-ti moment $\mathbb{E}[X^k]$ je $+\infty$ za sve $k \in \mathbb{N}, k \geqslant \alpha$, pa Teorem 6.14(a) povlači da ne postoji $t_0 > 0$ t.d. vrijedi (6.12), tj. ne postoji M_X.

[Napomena*. Ako za X ne postoji M_X, možemo koristiti tzv. *karakterističnu funkciju* $t \mapsto \mathbb{E}[e^{itX}] \in \mathbb{C}$ koja *uvijek* postoji. Karakteristična funkcija ima slična svojstva kao i FI momenata – npr. jedinstveno određuje distribuciju slučajne varijable te se lijepo ponaša na zbroj nezavisnih slučajnih varijabli. Ipak, kao što smo već vidjeli, FI momenata korisne su za određivanje momenata, ali dodatno, ako postoji M_X možemo dobiti eksponencijalne ocjene na vjerojatnosti oblika $\mathbb{P}(X > a)$; vidi vježbe (Chernoffova ograda), te tzv. *teoriju velikih devijacija* (engl. *large deviations*).]

Teorem 6.18. *Ako su slučajne varijable X i Y nezavisne te postoji $t_0 > 0$ tako da su M_X i M_Y dobro definirane na $[-t_0, t_0]$, M_{X+Y} je također dobro definirana na $[-t_0, t_0]$ te vrijedi*

$$M_{X+Y}(t) = M_X(t)M_Y(t), \quad |t| \leqslant t_0. \tag{6.20}$$

Dokaz je analogan dokazu u slučaju FI vjerojatnosti – detalje ostavljamo za vježbu.

[Dokaz. Za sve $|t| \leqslant t_0$ zbog nezavisnosti imamo

$$M_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX}e^{tY}] = [\text{nez.}] = \mathbb{E}[e^{tX}]\mathbb{E}[e^{tY}] = M_X(t)M_Y(t)\ (< \infty).\ ]$$

Napomena 6.19 (**FI momenata određuje distribuciju**). Može se pokazati da ako vrijedi $M_X = M_Y$ na nekom otvorenom intervalu oko 0, tada slučajne varijable X i Y imaju istu distribuciju (bez dokaza).

Primjer 6.20 (**Zbroj nezavisnih normalnih varijabli**). Koristeći (6.20) i prethodnu napomenu lako se pokaže da za $X \sim \mathrm{N}(\mu_1, \sigma_1^2)$ i $Y \sim \mathrm{N}(\mu_2, \sigma_2^2)$ nezavisne vrijedi $X + Y \sim \mathrm{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

[To da je $\mathbb{E}[X + Y] = \mu_1 + \mu_2$ i $\mathrm{Var}(X + Y) = \sigma_1^2 + \sigma_2^2$ slijedi iz nezavisnosti bez obzira na distribuciju – poanta ovdje je da distribucija zbroja ostaje normalna. Što se tiče dokaza, za sve $t \in \mathbb{R}$ imamo

$$M_{X+Y}(t) = [\text{nez.}] = M_X(t)M_Y(t) \overset{(6.19)}{=} e^{(\mu_1+\mu_2)t + \frac{1}{2}t^2(\sigma_1^2+\sigma_2^2)} \overset{(6.19)}{=} M_{\mathrm{N}(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)}(t).\ ]$$

Primjer 6.21 (**Crowdsourcing**). Pretpostavimo da su $X_i \sim \mathrm{N}(\mu, \sigma_i^2)$, $i = 1, \ldots, n$ nezavisne, za neko zajedničko očekivanje $\mu \in \mathbb{R}$, te (moguće različite) varijance $\sigma_1, \ldots, \sigma_n > 0$. Želimo procijeniti μ.

[Motivacija za ovaj problem može biti sljedeća. Imamo proizvod čiju (nepoznatu) kvalitetu želimo procijeniti. Parametar μ predstavlja kvalitetu proizvoda, a X_i predstavlja ocjenu koji

i-ti korisnik daje tom proizvodu – veći σ_i odgovara korisniku s manjom ekspertizom; vidi tzv. *crowdsourcing*.]

Osnovni procjenitelj koji možemo koristiti je $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$, a njegovu "grešku" možemo definirati kao $\mathbb{E}\Big[|\overline{X}_n - \mu|\Big]$. Po prethodnom primjeru znamo da $\sum_{i=1}^n X_i$ ima normalnu razdiobu, pa nadalje slijedi da i $\overline{X}_n - \mu$ ima normalnu razdiobu s parametrima

$$\mathbb{E}[\overline{X}_n - \mu] = 0\,, \quad \mathrm{Var}(\overline{X}_n - \mu) = \mathrm{Var}(\overline{X}_n) = \frac{\sum_{i=1}^n \sigma_i^2}{n^2}\,.$$

Specijalno, ako je $Z \sim \mathrm{N}(0,1)$, vrijedi

$$\mathbb{E}\Big[|\overline{X}_n - \mu|\Big] = \frac{\sqrt{\sum_{i=1}^n \sigma_i^2}}{n}\mathbb{E}\Big[|Z|\Big] = \sqrt{\frac{2}{\pi}}\cdot\frac{\sqrt{\sum_{i=1}^n \sigma_i^2}}{n}.$$

Na primjer, ako dopustimo da varijance σ_i ovise o n, tj. $\sigma_i = \sigma_i^{(n)}$, te ako za neke $c, \epsilon > 0$ vrijedi

$$\max_{i=1,\dots,n} \sigma_i^{(n)} = cn^{1+\epsilon}\,, \quad \forall n \geqslant 1\,,$$

imamo

$$\frac{\sqrt{\sum_{i=1}^n \sigma_i^2}}{n} \geqslant c^2\frac{n^{1+\epsilon}}{n} = c^2 n^\epsilon \longrightarrow +\infty\,, \quad n \to \infty\,.$$

Riječima, ukoliko je barem jedna varijanca dovoljno velika [tj. imamo barem jednog korisnika koji ima jako slabu ekspertizu], $\overline{X}_n$ neće biti dobar procjenitelj za μ. Pitanje je možemo li bolje [vidi vježbe]? Što ako su dodatno varijance σ_i^2 nepoznate [što je zapravo realna pretpostavka]? [Zadnje pitanje i dalje nije u potpunosti riješeno te je predmet aktivnog istraživanja.]

Primjer 6.22 (**Gama razdioba**). Kažemo da slučajna varijabla X ima **gama** razdiobu s parametrima $\alpha, \lambda > 0$ (oznaka $X \sim \mathrm{Gama}(\alpha, \lambda)$) ako je neprekidna s gustoćom

$$f(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}\,, \quad t > 0\,,$$

uz $f(t) = 0$ za $t \leqslant 0$, gdje je $\Gamma(\alpha) := \int_0^\infty y^{\alpha-1}e^{-y}\mathrm{d}y$ tzv. gama funkcija (vrijedi $\Gamma(n) = (n-1)!$ za sve $n \in \mathbb{N}$).

Nije teško pokazati da vrijedi

- $\mathrm{Gama}(1, \lambda)$ je upravo $\mathrm{Exp}(\lambda)$ razdioba;
- $X \sim \mathrm{Gama}(\alpha, 1)$ povlači da je $\frac{X}{\lambda} \sim \mathrm{Gama}(\alpha, \lambda)$, za sve $\lambda > 0$;
- $X \sim \mathrm{Gama}(\alpha, \lambda)$ povlači $M_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^\alpha$, za $t < \lambda$.

Propozicija 6.23 (**Zbroj nezavisnih eksponencijalnih varijabli**). *Ako su $X_1, \dots, X_n$ njd takve da je $X_i \sim \mathrm{Exp}(\lambda)$, vrijedi*

$$T_n := X_1 + \cdots + X_n \sim \mathrm{Gama}(n, \lambda)\,.$$

[U Poissonovom procesu s intenzitetom $\lambda > 0$, gornji T_n je upravo vrijeme kada se dogodio n-ti događaj.]

Dokaz. Iz Primjera 6.15 dobivamo

$$M_{T_n}(t) = [X_1, \ldots, X_n \text{ njd}] = (M_{X_1}(t))^n = \left(\frac{\lambda}{\lambda - t}\right)^n = M_{\text{Gama}(n,\lambda)}(t), \quad t < \lambda\,.$$

□

[Gama razdioba se često koristi u Bayesovskoj statistici kao apriorna distribucija za neke parametre. Na primjer, to je konjugirana apriorna razdioba za parametar λ u $\mathrm{P}(\lambda)$ razdiobi.]

POGLAVLJE 7

Nejednakosti i granični teoremi

7.1. Markovljeva i Čebiševljeva nejednakost

[U teorijskoj i primijenjenoj vjerojatnosti, vrlo često je slučaj da vjerojatnosti ne možemo odrediti egzaktno. Ipak, koristeći razne *nejednakosti* često možemo dobiti korisne gornje i/ili donje ograde. Mi ćemo spomenuti dvije osnovne nejednakosti, ali napominjemo da je područje tzv. *koncentracijskih nejednakosti* jedno od najvažnijih u modernoj primjeni teorije vjerojatnosti.]

Teorem 7.1 (**Markov**). *Za svaku slučajnu varijablu X, za sve $a > 0$ vrijedi*

$$\mathbb{P}\left(|X| \geqslant a\right) \leqslant \frac{\mathbb{E}\Big[|X|\Big]}{a}\,.$$

Dokaz. Za sve $a > 0$,

$$|X| \geqslant |X| \mathbb{1}_{\{|X| \geqslant a\}} \geqslant a \mathbb{1}_{\{|X| \geqslant a\}}\,.$$

[Gornje nejednakosti vrijede za svaki $\omega \in \Omega$.] Zaista, izraz u sredini jednak je $|X|$ ako je $|X| \geqslant a$, a 0 inače. Prva nejednakost slijedi jer je $|X| \geqslant 0$, a druga je očita.

Monotonost očekivanja sada povlači

$$\mathbb{E}\Big[|X|\Big] \geqslant \mathbb{E}\Big[a \mathbb{1}_{\{|X| \geqslant a\}}\Big] = a \mathbb{P}\left(|X| \geqslant a\right)\,.$$

□

Teorem 7.2 (**Čebišev**). *Ako je X slučajna varijabla t.d. je $\mathbb{E}\Big[|X|\Big] < \infty$, za sve $a > 0$ vrijedi*

$$\mathbb{P}\left(|X - \mathbb{E}[X]| \geqslant a\right) \leqslant \frac{\mathrm{Var}(X)}{a^2}\,.$$

Dokaz. Imamo

$$\begin{aligned}\mathbb{P}\left(|X - \mathbb{E}[X]| \geqslant a\right) &= \mathbb{P}\Big((X - \mathbb{E}[X])^2 \geqslant a^2\Big) = [\text{Markovljeva nej. za } (X - \mathbb{E}[X])^2] \\ &\leqslant \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{a^2} = \frac{\mathrm{Var}(X)}{a^2}\,.\end{aligned}$$

□

Primjer 7.3. Ako je X slučajna varijabla t.d. je $\sigma^2 := \mathrm{Var}(X) < \infty$, Čebiševljeva nejednakost daje da za sve $c > 0$ vrijedi

$$\mathbb{P}\left(|X - \mathbb{E}[X]| \geqslant c\sigma\right) \leqslant \frac{\mathrm{Var}(X)}{c^2\sigma^2} = \frac{1}{c^2}\,,$$

to jest

$$\mathbb{P}\left(|X - \mathbb{E}[X]| < c\sigma\right) \geqslant 1 - \frac{1}{c^2}\,,$$

pri čemu dakle gornje ograde *ne ovise* o X! Na primjer, za $c = 3$ vrijedi $1 - \frac{1}{c^2} = \frac{8}{9} \approx 0.89$ – dakle, *svaka* slučajna varijabla će biti najviše 3 standardne devijacije udaljena od svog očekivanja s vjerojatnošću od barem 0.89. [Za konkretne razdiobe ta vjerojatnost je često veća – vjerojatno najvažniji primjer je normalna razdioba.]

Primjer 7.4 (**"68–95–99.7% rule"**). Za $X \sim \mathrm{N}(\mu, \sigma^2)$ vrijedi

$$\begin{aligned}
\mathbb{P}\left(|X - \mathbb{E}[X]| < c\sigma\right) &= \left[Z := \frac{X-\mu}{\sigma}\right] = \mathbb{P}\left(|Z| < c\right) = [Z \sim \mathrm{N}(0,1)] \\
&= \Phi(c) - \Phi(-c) = 2\Phi(c) - 1 \approx \begin{cases} 0.6827\,, & c = 1 \\ 0.9545\,, & c = 2 \\ \mathbf{0.9973}\,, & c = 3 \end{cases}
\end{aligned}$$

7.2. Zakoni velikih brojeva

[Alternativa nejednakostima je korištenje aproksimacija iz općenitih graničnih teorema. U teoriji vjeorojatnosti dva najvažnija tipa graničnih teorema bez sumnje su **zakoni velikih brojeva** (ZVB) i **centralni granični teorem** (CGT).]

Definicija 7.5. Za niz slučajnih varijabli $X_1, X_2, \ldots$, kažemo da **konvergira po vjerojatnosti** prema slučajnoj varijabli X (oznaka $X_n \xrightarrow[n\to\infty]{\mathbb{P}} X$)[1] ako

$$\lim_{n\to\infty} \mathbb{P}\left(|X_n - X| > \epsilon\right) = 0\,, \quad \forall \epsilon > 0\,. \tag{7.1}$$

Teorem 7.6 (**Slabi ZVB**). *Neka je $X_1, X_2, \ldots$ niz nezavisnih slučajnih varijabli t.d. za neke $\mu \in \mathbb{R}$ i $\sigma^2 < \infty$, vrijedi $\mathbb{E}[X_i] = \mu$, $\mathrm{Var}(X_i) = \sigma^2$, $\forall i \geqslant 1$ [dakle, $X_1, X_2, \ldots$ nisu nužno jednako distribuirane]. Ako za sve $n \geqslant 1$ definiramo*

$$S_n := \sum_{i=1}^{n} X_i\,, \quad \overline{X}_n := \frac{S_n}{n}\,, \tag{7.2}$$

vrijedi

$$\overline{X}_n \xrightarrow[n\to\infty]{\mathbb{P}} \mu\,. \tag{7.3}$$

Dokaz. Koristimo Čebiševljevu nejednakost. Uočimo, vrijedi

$$\mathbb{E}[\overline{X}_n] = \frac{1}{n}\sum_{i=1}^{n} \overbrace{\mathbb{E}[X_i]}^{=\mu,\forall i} = \mu\,,$$

[1] Često ćemo pisati samo $X_n \xrightarrow{\mathbb{P}} X$.

te

$$\mathrm{Var}(\overline{X}_n) = \frac{1}{n^2}\mathrm{Var}(S_n) = [\text{nezavisnost}] = \frac{1}{n^2}\sum_{i=1}^{n}\overbrace{\mathrm{Var}(X_i)}^{=\sigma^2,\forall i} = \frac{\sigma^2}{n}\,.$$

Čebiševljeva nejednakost sada povlači

$$\mathbb{P}\Big(|\overline{X}_n - \mu| > \epsilon\Big) \leqslant \frac{\sigma^2/n}{\epsilon^2} = \frac{\sigma^2}{\epsilon^2}\cdot\frac{1}{n} \xrightarrow{n\to\infty} 0\,,$$

pri čemu smo na kraju iskoristili da je $\sigma^2 < \infty$ [u suprotnom je gornja ograda jednaka $+\infty$ za sve n]. □

Napomena 7.7. (a) Ako su $X_1, X_2, \ldots$ njd, slabi ZVB (tj. (7.3)) vrijedi i ako je samo $\mathbb{E}\big[|X_1|\big] < \infty$ (tzv. *Hinčinov* ZVB) – bez dokaza.[2]

(b) Nezavisnost je bitna![3] Na primjer, neka je $X \sim \mathrm{B}(1/2)$ te $X_i := X$, $\forall i \in \mathbb{N}$. Tada je

$$\overline{X}_n = X\,,\quad \forall n\,,$$

ali budući da je $X \in \{0,1\}$, uvijek je $|\overline{X}_n - \mathbb{E}[X_1]| = |X - \frac{1}{2}| = \frac{1}{2}$, pa za $\epsilon \in (0,1)$ vrijedi

$$\mathbb{P}\Big(|\overline{X}_n - \mathbb{E}[X_1]| > \epsilon\Big) = 1\,,\quad \forall n\,.$$

Specijalno, $\overline{X}_n$ ne konvergira po vjerojatnosti prema X. [Naravno, u ovom slučaju vrijedi $\overline{X}_n \xrightarrow{\mathbb{P}} X$.]

Primjer 7.8. Bacamo novčić na kojemu je vjerojatnost da će pasti pismo jednaka $p \in [0,1]$. Ako definiramo

$$X_i := \mathbb{1}_{\{\text{u } i\text{-tom bacanju palo pismo}\}}\,,\quad i \in \mathbb{N}\,,$$

zapravo je

$$\overline{X}_n = \text{postotak pisama koji su pali u prvih } n \text{ bacanja}\,,$$

pa SZVB ($X_1, X_2, \ldots$ su njd uz $\mu := \mathbb{E}[X_1] = p$) povlači da $\overline{X}_n \xrightarrow{\mathbb{P}} p$ [što odgovara našem intuitivnom shvaćanju "vjerojatnosti"].

7.2.1. Jaki ZVB.

Definicija 7.9. Za niz slučajnih varijabli $X_1, X_2, \ldots$ kažemo da **konvergira gotovo sigurno** ("g.s.") prema slučajnoj varijabli X (oznaka $X_n \xrightarrow[n\to\infty]{\text{g.s.}} X$) ako vrijedi

$$\mathbb{P}\Big(\lim_{n\to\infty} X_n = X\Big) = \mathbb{P}\Big(\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}\Big) = 1\,. \tag{7.4}$$

Napomena. Kasnije ćemo pokazati da $X_n \xrightarrow{\text{g.s.}} X$ povlači $X_n \xrightarrow{\mathbb{P}} X$.

[2] *Općenito, može se pokazati sljedeći rezultat – ako su $X_1, X_2, \ldots$ njd, nužan i dovoljan uvjet za postojanje niza konstanti $\mu_1, \mu_2, \ldots$ takvih da vrijedi $\overline{X}_n - \mu_n \xrightarrow{\mathbb{P}} 0$ je da $n\mathbb{P}(|X| > n) \xrightarrow{n\to\infty} 0$. U tom slučaju za μ_n se može uzeti $\mu_n = \mathbb{E}[X_1 \mathbb{1}_{|X_1|\leqslant n}]$ – ako je $\mathbb{E}\big[|X_1|\big] < \infty$, može se pokazati da vrijedi $n\mathbb{P}(|X_1| > n) \xrightarrow{n\to\infty} 0$ te da $\mu_n = \mathbb{E}[X_1\mathbb{1}_{\{|X_1|\leqslant n\}}] \xrightarrow{n\to\infty} \mathbb{E}[X_1] = \mu$.

[3] Ipak, nezavisnost nije nužna – u dokazu je jedino bilo bitno da je $\mathbb{E}[\overline{X}_n] = \mu$ te da $\mathrm{Var}(\overline{X}_n) \xrightarrow{n\to\infty} 0$.

TEOREM 7.10 (**Jaki ZVB**). *Neka je $X_1, X_2, \ldots$ niz njd slučajnih varijabli t.d. vrijedi $\mathbb{E}\Big[|X_1|\Big] < \infty$, te neka je $\mu := \mathbb{E}[X_1]$. Tada vrijedi*

$$\overline{X}_n \xrightarrow[n\to\infty]{g.s.} \mu\,.$$

DOKAZ. Dokaz ćemo provesti uz jaču pretpostavku – pretp. ćemo da je četvrti moment konačan, tj. $\mathbb{E}[X_1^4] =: K < \infty$; dokaz u općenitom slučaju izlazi van okvira ovog kolegija. Pretpostavimo najprije da je $\mu = 0$.

1. korak Pokazujemo da za sve $n \geqslant 1$ vrijedi $\mathbb{E}[S_n^4] \leqslant 4Kn^2$.

Za sve $n \geqslant 1$, $\mathbb{E}[S_n^4] = \mathbb{E}[(X_1 + \cdots + X_n)^4]$ je suma koja se sastoji od članova oblika

- $\mathbb{E}[X_i^4] = [X_1, X_2, \ldots$ su njd$] = \mathbb{E}[X_1^4] = K$, $\forall i$;
- $\mathbb{E}[X_i^2 X_j^2] = [$nezavisnost$] = \mathbb{E}[X_i^2]\mathbb{E}[X_j^2] = [X_1, X_2, \ldots$ su njd$] = (\mathbb{E}[X_1^2])^2$, $\forall i \neq j$;
- $\mathbb{E}[X_i^3 X_j] = \mathbb{E}[X_i^3]\mathbb{E}[X_j] = \mathbb{E}[X_i^3] \cdot \mu = 0$, $\forall i \neq j$;
- $\mathbb{E}[X_i^2 X_j X_k] = 0$, $\forall i, j, k$ različite;
- $\mathbb{E}[X_i X_j X_k X_l] = 0$, $\forall i, j, k, l$ različite.

Dakle,

$$\begin{aligned}\mathbb{E}[S_n^4] &= n\mathbb{E}[X_1^4] + \binom{n}{2}\binom{4}{2}\mathbb{E}[X_1^2 X_2^2] + 0 = nK + 3n(n-1)(\mathbb{E}[X_1^2])^2 \\ &\leqslant nK + 3n(n-1)K \leqslant 4Kn^2\,, \quad \forall n \geqslant 1\,,\end{aligned}$$

pri čemu smo u trećem koraku iskoristili nejednakost $\mathbb{E}[X_1^2]^2 \leqslant \mathbb{E}[X_1^4] = K$ koja npr. slijedi jer $0 \leqslant \mathrm{Var}(X_1^2) = \mathbb{E}[X_1^4] - \mathbb{E}[X_1^2]^2$.

2. korak Korištenjem Fubinijevog teorema (sve je nenegativno) imamo

$$\mathbb{E}\left[\sum_{n=1}^{\infty} \frac{S_n^4}{n^4}\right] = \sum_{n=1}^{\infty} \frac{\mathbb{E}[S_n^4]}{n^4} \leqslant \sum_{n=1}^{\infty} \frac{4K}{n^2} < \infty\,.$$

Specijalno, imamo da je $\mathbb{P}(\sum_{n=1}^{\infty} \frac{S_n^4}{n^4} < \infty) = 1$ jer bi u suprotnom bilo $\mathbb{E}[\sum_{n=1}^{\infty} \frac{S_n^4}{n^4}] = +\infty$. Nadalje, iz nužnog uvjeta za konvergenciju reda, dobivamo da je[4]

$$1 = \mathbb{P}\left(\lim_{n\to\infty} \frac{S_n^4}{n^4} = 0\right) = \mathbb{P}\left(\lim_{n\to\infty} \frac{S_n}{n} = 0\right) = \mathbb{P}\left(\lim_{n\to\infty} \overline{X}_n = \mu\right)\,.$$

[4] Alternativno, korištenjem Markovljeve nejednakosti imamo

$$\mathbb{P}(|\overline{X}_n| \geqslant \epsilon) = \mathbb{P}(S_n^4 \geqslant \epsilon^4 n^4) \leqslant \frac{\mathbb{E}[S_n^4]}{\epsilon^4 n^4} \leqslant \frac{4K}{\epsilon^4} \cdot \frac{1}{n^2}\,, \quad \forall \epsilon > 0\,.$$

Budući da je $\sum_{n=1}^{\infty} \mathbb{P}(|\overline{X}_n| > \epsilon) < \infty$, BC lema povlači da je

$$0 = \mathbb{P}(|\overline{X}_n| \geqslant \epsilon \text{ b.m.p.}) = \mathbb{P}(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} \{|\overline{X}_n| \geqslant \epsilon\}) =: \mathbb{P}(A_\epsilon^c)\,, \quad \forall \epsilon > 0\,.$$

Dakle, $1 = \mathbb{P}(A_\epsilon) = \mathbb{P}(\cup_{n=1}^{\infty} \cap_{k=n}^{\infty} \{|\overline{X}_n| < \epsilon\})$, $\forall \epsilon > 0$. Tvrdnja sada slijedi iz neprekidnosti vjerojatnosti budući da je

$$\left\{\lim_{n\to\infty} \overline{X}_n = 0\right\} = \bigcap_{k=1}^{\infty} A_{1/k}\,.$$

Na kraju, ako je $\mu \in \mathbb{R}$, za $Y_i := X_i - \mu$, $i \in \mathbb{N}$, imamo da je $Y_1, Y_2, \ldots$ njd niz uz $\mathbb{E}[Y_1] = 0$, pa prethodni dokaz povlači

$$1 = \mathbb{P}\left(\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} Y_i = 0\right) = \mathbb{P}\left(\lim_{n\to\infty} \frac{S_n - n\mu}{n} = 0\right) = \mathbb{P}\left(\lim_{n\to\infty} \overline{X}_n = \mu\right) .$$

□

[Napomena*. U prethodnom teoremu pretpostavka o postojanju očekivanja je bitna! Naime, može se pokazati da u slučaju $\mathbb{E}\Big[|X_1|\Big] = \infty$ vrijedi

$$\mathbb{P}(\text{postoji } \lim_{n\to\infty} \overline{X}_n \text{ i poprima vrijednost u } (-\infty, \infty)) = 0 .$$

Na primjer, ako je $X_1, X_2, \ldots$ njd niz takav da je $X_1 \sim -X_1$ (tj. X_1 simetrična) te vrijedi $\mathbb{E}\Big[|X_1|\Big] = \infty$, ali $n\mathbb{P}(|X_1| > n) \xrightarrow{n\to\infty} 0$, imamo $\overline{X}_n \xrightarrow{\mathbb{P}} 0$ (jer je $\mu_n = \mathbb{E}[X_1 \mathbb{1}_{\{|X_1| \leqslant n\}}] = 0$), ali ne i $\overline{X}_n \xrightarrow{\text{g.s.}} 0$. □]

Primjer 7.11 (**Neparametarska statistika**). Pretpostavimo da je F proizvoljna (nepoznata) funkcija distribucije i $X_1, \ldots, X_n$ njd niz takav da je $F_{X_i} = F$. Pitanje je možemo li iz jedne realizacije uzorka $X_1, \ldots, X_n$ procijeniti F.

Empirijsku funkciju distribucije uzorka $X_1, \ldots, X_n$ definiramo sa

$$\widehat{F}_n(t) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{X_i \leqslant t\}} , \quad t \in \mathbb{R} .$$

Uočimo, $\widehat{F}_n(t)$ je slučajna varijabla za sve $t \in \mathbb{R}$ [dakle, na $\widehat{F}_n$ možemo gledati kao na slučajnu funkciju (distribucije)]; vidi Sliku 10. Zapravo, ako imamo realizaciju $X_i(\omega) = x_i$, $i = 1, \ldots, n$, te je $x_i \neq x_j$, $\widehat{F}_n(\omega)$ je zapravo funkcija distribucije uniformne razdiobe na skupu $\{x_1, \ldots, x_n\}$.

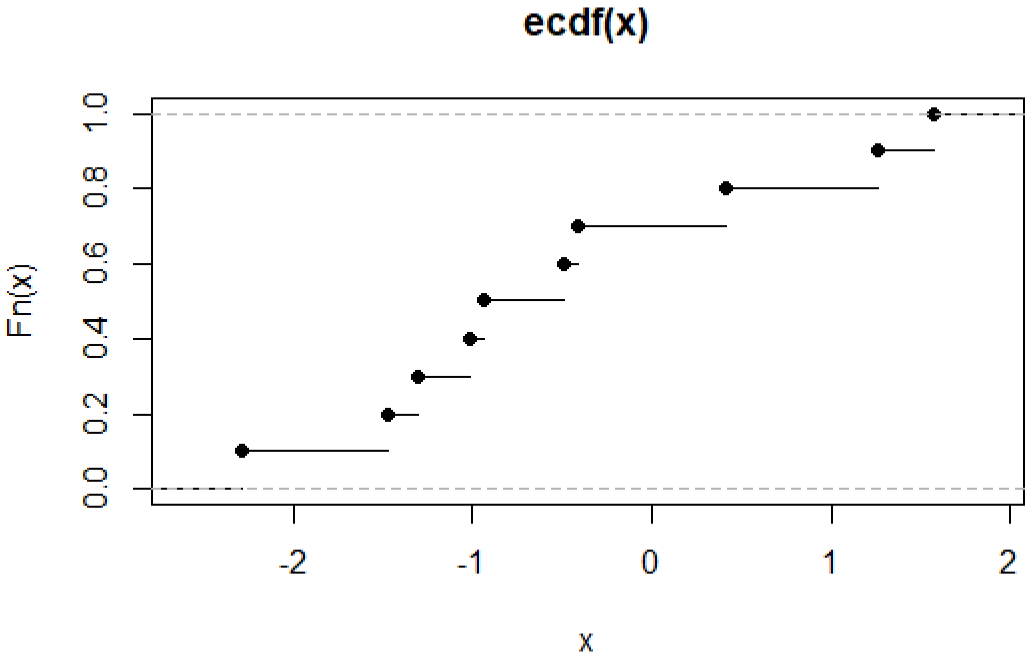
Iz JZVB-a za njd niz $Y_i := \mathbb{1}_{\{X_i \leqslant t\}}$, $i \in \mathbb{N}$, slijedi da za sve $t \in \mathbb{R}$ vrijedi

$$\widehat{F}_n(t) \xrightarrow[n\to\infty]{\text{g.s.}} \mathbb{E}[\mathbb{1}_{\{X_1 \leqslant t\}}] = \mathbb{P}(X_1 \leqslant t) = [F_{X_1} = F] = F(t) .$$

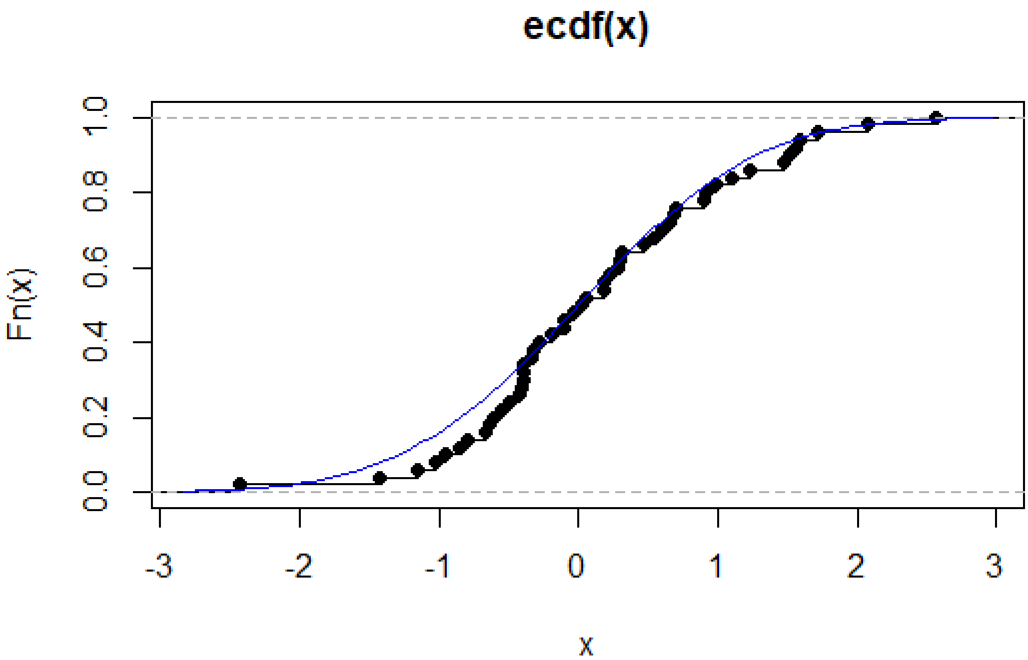
Zapravo, vrijedi i više: tzv. *Glivenko-Cantellijev* teorem daje

$$\sup_{t\in\mathbb{R}} |\widehat{F}_n(t) - F(t)| \xrightarrow[n\to\infty]{\text{g.s.}} 0 ,$$

vidi Sliku 11 za ilustraciju. Prethodni rezultat mam omogućava da (nepoznatu) razdiobu F (ili neki njezin funkcional, primjerice očekivanje) procjenjujemo s $\widehat{F}_n$, te je baza tzv. *neparametarske statistike* (jer u ovom slučaju ne pretpostavljamo da F dolazi iz neke unaprijed određene (parametarske) familije funkcija distribucije).



SLIKA 10. Empirijska funkcija distribucije $\widehat{F}_n$ kada je $X_1 \sim \mathrm{N}(0,1)$ i $n = 10$ – uočimo da je $\widehat{F}_n$ po dijelovima konstantna funkcija sa skokovima veličine $\frac{1}{n} = \frac{1}{10}$ (ako su svi X_i-evi različiti).



SLIKA 11. Empirijska funkcija distribucije $\widehat{F}_n$ kada je $X_1 \sim \mathrm{N}(0,1)$ i $n = 50$. Plavom linijom je označena funkcija distribucije standardne normalne razdiobe.

7.3. Centralni granični teorem

Ako su $X_1, X_2, \ldots$ njd te $\mu := \mathbb{E}[X_1]$, JZVB povlači da je gotovo sigurno $\frac{S_n}{n} \approx \mu$, tj. $S_n \approx n\mu$, za velike n. Intuitivno, ukoliko je $\sigma^2 := \mathrm{Var}(X_1) < \infty$, centralni granični teorem (CGT) povlači da S_n približno ima $\mathrm{N}(n\mu, n\sigma^2)$ razdiobu, te specijalno da je odstupanje $|S_n - n\mu|$ reda veličine $\sqrt{n}$.

[Preciznije...]

Uočimo, ako je $\sigma^2 := \mathrm{Var}(X_1) \in (0, \infty)$, za standardizirane varijable

$$Z_n := \frac{S_n - n\mu}{\sigma\sqrt{n}}\,, \quad n \in \mathbb{N}\,, \tag{7.5}$$

vrijedi $\mathbb{E}[Z_n] = 0$, $\mathrm{Var}(Z_n) = 1$, $\forall n \in \mathbb{N}$.

Teorem 7.12 (**CGT**). *Neka su $X_1, X_2, \ldots$ njd slučajne varijable t.d. je $\sigma^2 := \mathrm{Var}(X_1) \in (0, \infty)$, te neka je $\mu := \mathbb{E}[X_1]$ i Z_n definiran u (7.5). Tada vrijedi*

$$\mathbb{P}(Z_n \leqslant x) \xrightarrow{n\to\infty} \mathbb{P}(Z \leqslant x)\,, \quad \forall x \in \mathbb{R}\,, \tag{7.6}$$

pri čemu je $\underline{Z \sim \mathrm{N}(0,1)}$.

Dakle, za dovoljno veliki n, i sve $x \in \mathbb{R}$, vrijedi

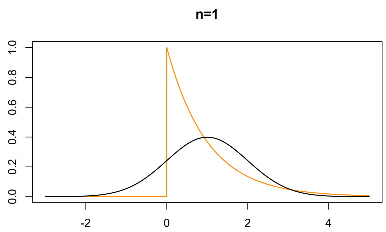
$$\mathbb{P}(S_n \leqslant x) = \mathbb{P}\left(Z_n \leqslant \frac{x - n\mu}{\sigma\sqrt{n}}\right) \overset{(7.6)}{\approx} \mathbb{P}\left(Z \leqslant \frac{x - n\mu}{\sigma\sqrt{n}}\right) = \mathbb{P}(\sigma\sqrt{n}Z + n\mu \leqslant x)\,.$$

Budući da je $\sigma\sqrt{n}Z + n\mu \sim \mathrm{N}(n\mu, n\sigma^2)$, CGT možemo neformalno iskazati kao

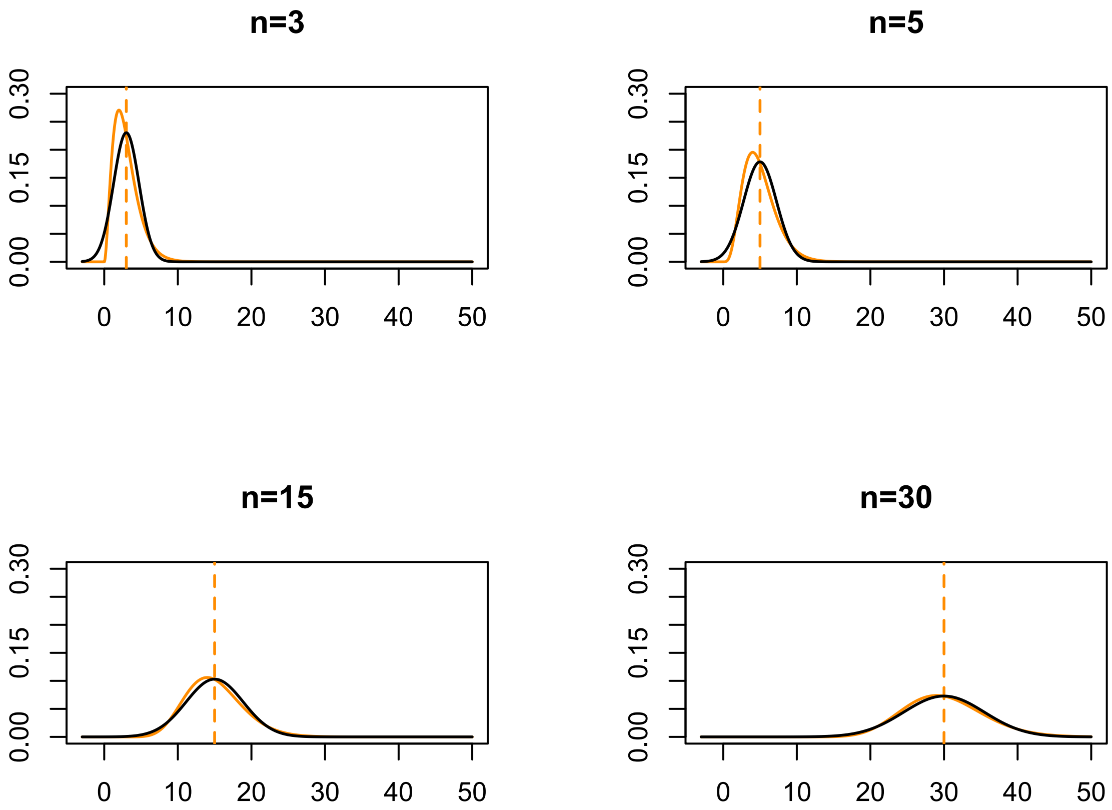
$$S_n \overset{d}{\approx} \mathrm{N}(n\mu, n\sigma^2)\,, \quad \text{za velike } n, \tag{7.7}$$

pri čemu gornja "normalna" aproksimacija (uočimo da je $n\mu = \mathbb{E}[S_n], n\sigma^2 = \mathrm{Var}(S_n)$) vrijedi *bez obzira* na distribuciju od X_1 (ali uz pretpostavku $\mathrm{Var}(X_1) \in (0, \infty)$)!

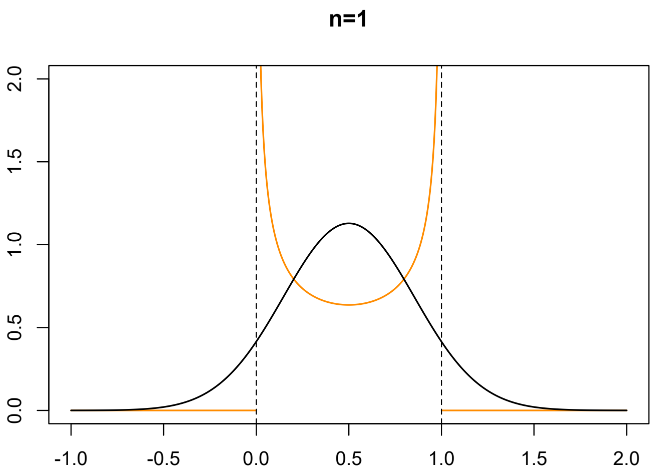
Za ilustraciju CGT-a, tj. aproksimacije (7.7), za razne distribucije od X_i, vidi Slike 12-16. U nekim slučajevima koristimo tzv. *histogram* dobiven na temelju velikog broja simulacije slučajne varijable S_n – o tome ćete više učiti na kolegiju Statistika, ali za naše potrebe dovoljno je znati da histogram predstavlja procjenu gustoće neprekidne razdiobe koja najbolje odgovara razdiobi od S_n.



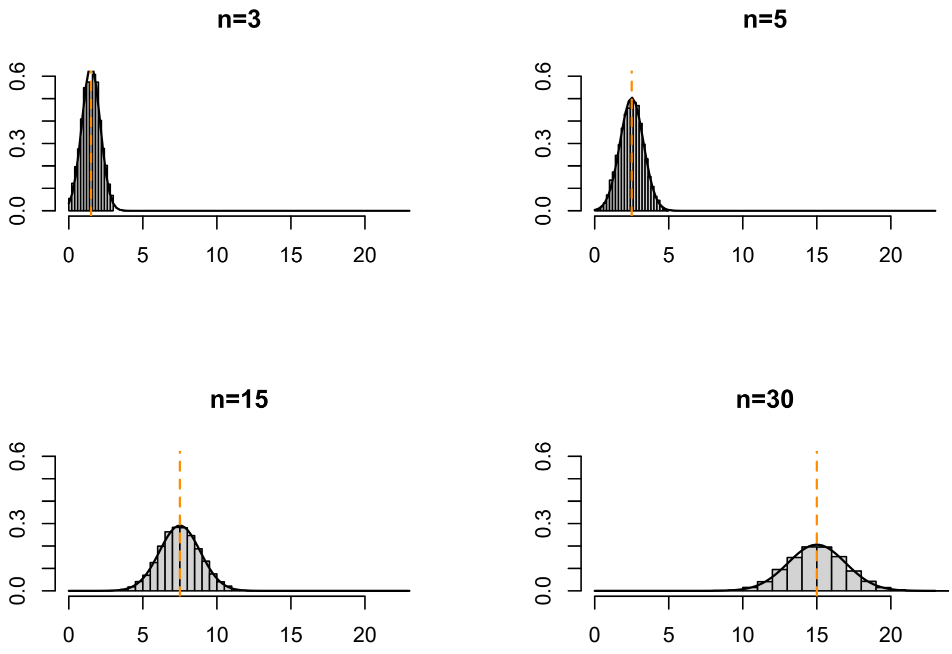
Slika 12. Gustoća Exp(1) razdiobe (narančasta linija), zajedno s gustoćom normalne razdiobe s istom očekivanjem i varijancom.



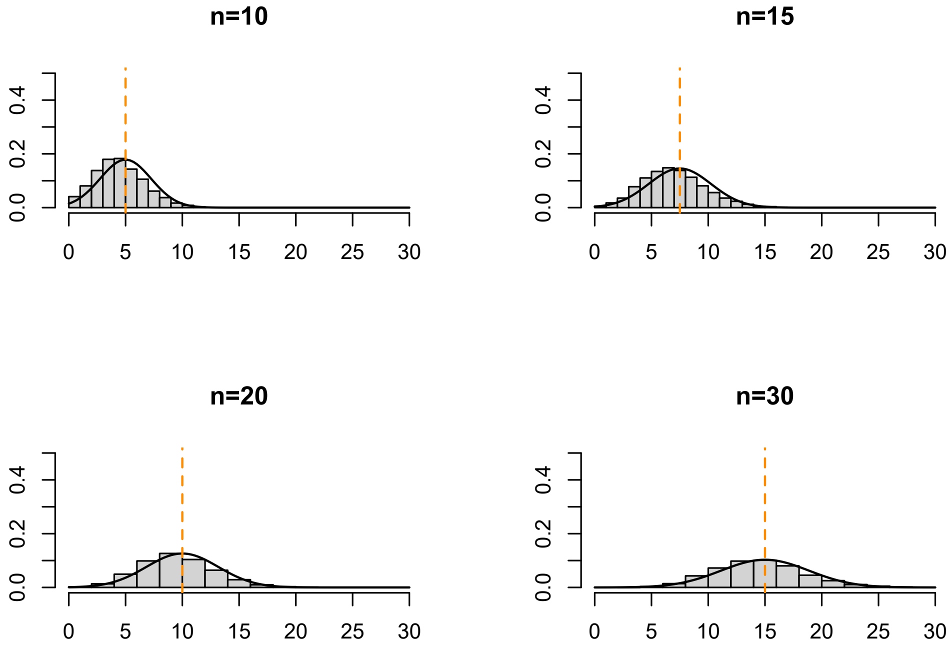
SLIKA 13. Gustoća od $S_n = \sum_{i=1}^{n} X_i$ (narančasta linija) u slučaju kada su $X_1, X_2, \ldots$ njd iz Exp(1) razdiobe ($S_n \sim \text{Gama}(n,1)$), zajedno s gustoćom normalne razdiobe s istom očekivanjem (vertikalna linija) i varijancom, za razne n-ove.



SLIKA 14. Gustoće Beta(1/2, 1/2) razdiobe, zajedno s gustoćom normalne razdiobe s istom očekivanjem i varijancom.



Slika 15. Histogram od $S_n = \sum_{i=1}^{n} X_i$ u slučaju kada su $X_1, X_2, \ldots$ njd iz Beta$(1/2, 1/2)$ razdiobe, zajedno s gustoćom normalne razdiobe s istom očekivanjem i varijancom, za razne n-ove. Iako je gustoća Beta$(1/2, 1/2)$ razdiobe dosta drugačija od bilo koje normalne razdiobe, aproksimacija (7.7) se čini dosta dobra i za male n-ove.



Slika 16. Histogram od $S_n = \sum_{i=1}^{n} X_i$ u slučaju kada su $X_1, X_2, \ldots$ njd iz Poissonove razdiobe s parametrom $\lambda = 1/2$, zajedno s gustoćom normalne razdiobe s istom očekivanjem i varijancom, za razne n-ove. Dakle, aproksimacija s normalnom (dakle, neprekidnom) razdiobom vrijedi i za diskretne varijable.

[Za dokaz CGT-a trebamo dva pomoćna rezultata. Dokaz prvog ostavljamo kao jednostavnu vježbu iz analize, a drugi navodimo bez dokaza.]

LEMA 7.13. *Ako je $b_1, b_2, \ldots$ niz realnih brojeva t.d. je $\lim_{n\to\infty} b_n = b \in \mathbb{R}$, vrijedi*

$$\left(1 + \frac{b_n}{n}\right)^n \xrightarrow{n\to\infty} e^b .$$

[DOKAZ. Uzimanjem logaritma slijedi da je dovoljno dokazati da vrijedi

$$n \log\left(1 + \frac{b_n}{n}\right) \xrightarrow{n\to\infty} b .$$

Ključna je aproksimacija $\lim_{x\to 0} \frac{\log(1+x)}{x} = 1$ budući da vrijedi $b_n/n \xrightarrow{n\to\infty} 0$. Pretpostavimo da je $b \neq 0$ – budući da je $b_n/n \neq 0$ za dovoljno veliki n, imamo

$$n \log\left(1 + \frac{b_n}{n}\right) = b_n \frac{\log\left(1 + \frac{b_n}{n}\right)}{\frac{b_n}{n}} \xrightarrow{n\to\infty} b \cdot 1 = b .$$

U slučaju $b = 0$, samo treba komentirati da za sve n takve da je $b_n = 0$, vrijedi $n \log\left(1 + \frac{b_n}{n}\right) = 0 = b$.]

TEOREM 7.14 (**Teorem neprekidnosti**). *Neka su $Y, Y_1, Y_2, \ldots$ slučajne varijable te pretpostavimo da postoji $t_0 > 0$ t.d. su FI momenata $M_Y, M_{Y_1}, M_{Y_2}, \ldots$ dobro definirane na $(-t_0, t_0)$ [dakle, $M_Y(t), M_{Y_1}, \cdots < \infty$ za sve $|t| < t_0$]. Tada, ako FI momenata konvergiraju, tj.*

$$M_{Y_n}(t) \xrightarrow{n\to\infty} M_Y(t) , \quad \forall |t| < t_0 ,$$

vrijedi i

$$\mathbb{P}(Y_n \leqslant x) = F_{Y_n}(x) \xrightarrow{n\to\infty} F_Y(x) = \mathbb{P}(Y \leqslant x) ,$$

za sve $x \in \mathbb{R}$ u kojima je F_Y neprekidna [kasnije ćemo ovo definirati kao konvergenciju po distribuciji].

DOKAZ TEOREMA 7.12. Teorem ćemo dokazati uz jaču pretpostavku: pretpostavit ćemo da X_1 ima FI momenata, tj. da postoji $t_0 > 0$ t.d. je $M_{X_1} < \infty$ za sve $|t| \leqslant t_0$ [vidi napomenu nakon dokaza za dokaz u općenitom slučaju].

Ako stavimo $Y_i := X_i - \mu$, $i \in \mathbb{N}$ [tzv. centriranje], vrijedi

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{n} Y_i =: \frac{1}{\sigma\sqrt{n}} S_n^Y , \tag{7.8}$$

te

$$M(t) := M_{Y_1}(t) = e^{-t\mu} M_{X_1}(t) < \infty , \quad t \in [-t_0, t_0] .$$

Budući da su $Y_1, Y_2, \ldots$ njd, iz (7.8) slijedi

$$M_{Z_n}(t) = \mathbb{E}\left[e^{\frac{t}{\sigma\sqrt{n}} S_n^Y}\right] = M_{S_n^Y}\left(\frac{t}{\sigma\sqrt{n}}\right) = \left(M\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n ,$$

za sve $\frac{|t|}{\sigma\sqrt{n}} \leqslant t_0$, tj. za sve $t \in \mathbb{R}$ takve da vrijedi $|t| \leqslant t_0\sigma\sqrt{n}$ – zadnje vrijedi za svaki fiksan $t_0 \in \mathbb{R}$ čim je n dovoljno velik.

Taylorov razvoj funkcije M (oko 0) [M beskonačno puta diferencijabilna na $(-t_0, t_0)$] povlači da je

$$M(x) = M(0) + M'(0)x + \frac{M''(x)}{2}x^2 + R_2(x)\,, \quad \forall |x| < t_0\,,$$

pri čemu je ostatak oblika

$$R_2(x) = \frac{M^{(3)}(c_x)}{3!}x^3 =: x^2 h(x)\,, \quad \text{za neki } |c_x| \in (0, x)\,,$$

te je dakle $\lim_{x\to 0} h(x) = 0$. Budući da je $M = M_{Y_1}$ te je po konstrukciji $\mathbb{E}[Y_1] = 0$, $\mathrm{Var}(Y_1) = \mathrm{Var}(X_1) = \sigma^2$, imamo

$$M(x) = 1 + \mathbb{E}[Y_1] + \frac{\mathbb{E}[Y_1^2]}{2}x^2 + x^2h(x) = 1 + \frac{\sigma^2}{2}x^2 + x^2h(x)\,, \quad \forall |x| < t_0\,. \tag{7.9}$$

Sada, za svaki fiksni $t \in \mathbb{R}$ te n dovoljno velik (tako da je $|t| < t_0\sigma\sqrt{n}$) imamo

$$\begin{aligned} M_{Z_n}(t) &= \left(M\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n \overset{(7.9)}{=} \left(1 + \frac{t^2}{2n} + \frac{t^2}{\sigma^2 n}h\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n \\ &= \left(1 + \frac{1}{n}\left[\frac{t^2}{2} + \frac{t^2}{\sigma^2}h\left(\frac{t}{\sigma\sqrt{n}}\right)\right]\right)^n =: \left(1 + \frac{b_n(t)}{n}\right)^n. \end{aligned}$$

Budući da vrijedi $b_n(t) \xrightarrow{n\to\infty} \frac{t^2}{2}$, iz Leme 7.13 dobivamo

$$M_{Z_n}(t) \xrightarrow{n\to\infty} e^{\frac{t^2}{2}} \overset{(6.18)}{=} M_Z(t)\,, \quad \forall t \in \mathbb{R}\,.$$

Budući da je $F_Z = \Phi$ neprekidna na $\mathbb{R}$, (7.6) slijedi iz Teorema 7.14.

□

Napomena 7.15. U slučaju da ne postoji FI momenata M_{X_1} (vidi Primjer 6.17)), dokaz CGT-a se na vrlo sličan način može provesti koristeći tzv. **karakterističnu funkciju**

$$t \mapsto \mathbb{E}[e^{itX_1}] = \mathbb{E}[\cos(tX)] + i\mathbb{E}[\sin(tX)] \in \mathbb{C}\,, \quad t \in \mathbb{R}\,,$$

koja je uvijek dobro definirana.

Napomena 7.16. Može se pokazati da konvergencija u (7.6) povlači i više: tj. za sve $-\infty \leqslant a < b \leqslant +\infty$ i proizvoljan interval $I \in \{(a,b), [a,b), (a,b], [a,b]\}$, vrijedi

$$\mathbb{P}(Z_n \in I) \xrightarrow{n\to\infty} \mathbb{P}(Z \in I) = [Z \text{ neprekidna}] = \Phi(b) - \Phi(a)\,. \tag{7.10}$$

Primjer 7.17 (**Normalna aproksimacija za binomnu**). Neka je $Y_n \sim \mathrm{B}(n,p)$, $n \in \mathbb{N}$, za $p \in (0,1)$ fiksan. Ako u Teoremu 7.12 uzmemo $X_i \sim \mathrm{B}(p)$, imamo $\mu = p$, $\sigma^2 = p(1-p) =: pq$ te $S_n \sim Y_n$, pa CGT povlači tzv. **de Moivre-Laplaceov teorem**:

$$\mathbb{P}\left(\frac{Y_n - np}{\sqrt{npq}} \leqslant x\right) \xrightarrow{n\to\infty} \Phi(x)\,, \quad \forall x \in \mathbb{R}\,, \tag{7.11}$$

to jest

$$Y_n \overset{d}{\approx} \mathrm{N}(np, npq)\,, \quad \text{za velike } n.$$

Napomena 7.18 (**Alternativna formulacija CGT-a**). Uočimo, budući da je

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} \cdot \frac{1/n}{1/n} = \frac{\sqrt{n}}{\sigma} \cdot (\overline{X}_n - \mu)\,, \tag{7.12}$$

CGT povlači

$$\overline{X}_n \overset{d}{\approx} \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)\,, \quad \text{za velike } n,$$

zbog čega se normalna razdioba često javlja u statistici.

Primjer 7.19. Nepoznati broj glasača, $p \in (0,1)$, glasat će za jednu stranku. Slučajno je ispitano n ljudi te neka je S_n broj ispitanika koji će glasati za tu stranku. Koliki treba biti n takav da s vjerojatnošću od barem 0.95, $\overline{X}_n = \frac{S_n}{n}$ bude najviše 3% udaljeno od p?

Rješenje. Pretpostavimo da je $S_n \sim \mathrm{B}(n,p)$. CGT povlači da je

$$\overline{X}_n \sim \mathrm{N}\left(p, \frac{\sigma_p^2}{n}\right)$$

gdje je $\sigma_p^2 := p(1-p)$. Budući da za $Y \sim \mathrm{N}(\mu, \sigma^2)$, za μ, σ proizvoljne, vrijedi

$$\mathbb{P}\Big(|Y - \mu| \leqslant 1.96\sigma\Big) = 2\Phi(1.96) - 1 \approx 0.95\,,$$

ako je $0.03 \geqslant 1.96 \cdot \frac{\sigma_p^2}{n}$, tj. $n \geqslant \left(\frac{1.96}{0.03}\right)^2 \cdot \sigma_p^2$, vrijedi

$$\mathbb{P}\Big(|\overline{X}_n - p| \leqslant 0.03\Big) = \mathbb{P}\Big(|\overline{X}_n - p| \leqslant 1.96\tfrac{\sigma_p}{\sqrt{n}}\Big) \overset{\text{CGT}}{\approx} 2\Phi(1.96) - 1 \approx 0.95\,.$$

Jedini problem se čini u tome što ne znamo p, pa time ni $\sigma_p^2 = p(1-p)$. Ipak, bez obzira na vrijednost od $p \in (0,1)$, vrijedi

$$\sigma_p^2 \leqslant \frac{1}{2}\left(1 - \frac{1}{2}\right) = \frac{1}{4}\,,$$

pa slijedi da je dovoljno uzeti

$$n \geqslant \left(\frac{1.96}{0.03}\right)^2 \cdot \frac{1}{4} = 1067.11\,,$$

što je tipična veličina uzorka u ovakvim ispitivanjima. □

7.4. Različiti oblici konvergencija slučajnih varijabli

U nastavku, ako ne kažemo drugačije, $X, X_1, X_2, \dots$ predstavlja niz slučajnih varijabli definiranih na istom vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$.

Teorem 7.20. *Ako vrijedi* $X_n \xrightarrow{g.s.} X$, *nužno vrijedi i* $X_n \xrightarrow{\mathbb{P}} X$.

Dokaz. Neka je $\epsilon > 0$ proizvoljan. Želimo pokazati da vrijedi

$$\mathbb{P}\Big(|X_n - X| > \epsilon\Big) =: \mathbb{P}(A_n(\epsilon)) \xrightarrow{n\to\infty} 0\,.$$

Imamo da je

$$1 = \mathbb{P}(\lim_{n\to\infty} X_n = X) \leqslant \mathbb{P}\Big(\bigcup_{n=1}^{\infty}\bigcap_{k\geqslant n} A_k(\epsilon)^c\Big)\,,$$

pa slijedi i da je vjerojatnost na desnoj strani jednaka 1. Specijalno, uzimanjem komplementa i korištenjem neprekidnosti vjerojatnosti dobivamo

$$0 = \mathbb{P}\Big(\bigcap_{n=1}^{\infty}\bigcup_{k\geqslant n} A_k(\epsilon)\Big) = \lim_{n\to\infty}\mathbb{P}\Big(\bigcup_{k\geqslant n} A_k(\epsilon)\Big) \geqslant \lim_{n\to\infty}\mathbb{P}\Big(A_n(\epsilon)\Big)\,,$$

pa iz nenegativnostni vjerojatnosti slijedi $\mathbb{P}(A_n(\epsilon)) \xrightarrow{n\to\infty} 0$. □

Primjer 7.21. Obrat u prethodnom teoremu općenito ne vrijedi![5] Na primjer, neka su $X_1, X_2, \ldots$ nezavisne slučajne varijable takve da je $X_n \sim \mathrm{B}(1/n)$ [nacrtati skicu distribucije], $n \in \mathbb{N}$. Ako postoji limes niza $X_1, X_2, \ldots$, očekujemo da bi to trebala biti konstanta 0. Pogledajmo prvo konvergenciju po vjerojatnosti.

Ako je $\epsilon \in (0, 1)$, budući da je $X_n \in \{0, 1\}$,

$$\mathbb{P}\Big(|X_n - 0| > \epsilon\Big) = \mathbb{P}\Big(X_n > \epsilon\Big) = \mathbb{P}(X_n = 1) = \frac{1}{n} \xrightarrow{n\to\infty} 0\,.$$

Za svaki $\epsilon \geqslant 1$ očito imamo

$$\mathbb{P}\Big(|X_n| > \epsilon\Big) \leqslant \mathbb{P}\Big(|X_n| > \tfrac{1}{2}\Big) \xrightarrow{n\to\infty} 0\,,$$

pa smo dakle pokazali da vrijedi $X_n \xrightarrow{\mathbb{P}} 0$.

Pogledajmo sada konvergenciju gotovo sigurno. Prvo, ako $X_n \xrightarrow{\text{g.s.}} X$ za neku slučajnu varijablu X, iz Teorema 7.20 i $X_n \xrightarrow{\mathbb{P}} 0$ slijedi da je nužno $X \equiv 0$. Ipak, budući da je

$$\sum_{n=1}^{\infty}\mathbb{P}(X_n = 1) = \sum_{n=1}^{\infty}\frac{1}{n} = \infty\,,$$

te su događaji $\{X_n = 1\}$, $n \in \mathbb{N}$, nezavisni, druga BC lema povlači da je

$$1 = \mathbb{P}(X_n = 1 \text{ b.m.p.}) \leqslant \mathbb{P}(\lim_{n\to\infty} X_n \neq 0)\,,$$

pri čemu nejednakost slijedi jer na događaju $\{X_n = 1 \text{ b.m.p.}\}$, niz (X_n) ima gomilište u 1 pa $\lim_{n\to\infty} X_n$ (čak i ako postoji) ne može biti 0. Specijalno vjerojatnost na desnoj strani je također jednaka 1. Dakle, niz (X_n) ne konvergira gotovo sigurno prema 0 (niti prema bilo kojoj drugoj slučajnoj varijabli).

Definicija 7.22. Za niz slučajnih varijabli $X_1, X_2, \ldots$, kažemo da **konvergira po distribuciji** prema slučajnoj varijabli X (oznaka $X_n \xrightarrow{d} X$) ako vrijedi

$$F_{X_n}(t) \xrightarrow{n\to\infty} F_X(t)\,, \quad \forall t \in C_{F_X}\,, \tag{7.13}$$

[5] Ipak, može se pokazati da uvijek postoji podniz $(n_k)_{k\in\mathbb{N}}$ takav da $X_{n_k} \xrightarrow{\text{g.s.}} X$ kada $k \to \infty$.

gdje je $C_{F_X} \subseteq \mathbb{R}$ skup svih točaka neprekidnosti funkcije (distribucije) F_X.[6]

NAPOMENA 7.23. (a) **Zašto ne tražimo da (7.13) vrijedi za sve $t \in \mathbb{R}$?** Na primjer, ako je $X_n := \frac{1}{n}$ [dakle, konstanta], $n \in \mathbb{N}$, očekivali bismo da niz (X_n) konvergira po distribuciji prema $X := 0$. To zaista i je slučaj [nacrtati pripadne funkcije distribucije] jer za sve $t \neq 0$ (tj. $t \in C_{F_X}$) vrijedi

- ako je $t < 0$, imamo $F_{X_n}(t) = 0 = F_X(t)$, za sve n;
- ako je $t > 0$, čim je $t \geqslant \frac{1}{n}$, tj. $n \geqslant \frac{1}{t}$, imamo $F_{X_n}(t) = 1 = F_X(t)$.

Ipak, u ovom slučaju $F_{X_n}(0) = 0 \nrightarrow 1 = F_X(0)$ kada $n \to \infty$.

(b)* Može se pokazati da je (7.13) dovoljno provjeriti tako da zapravo vrijedi i puno više: u tom slučaju imamo

$$\mathbb{P}(X_n \in B) \xrightarrow{n\to\infty} \mathbb{P}(X \in B)\,,$$

za sve $B \subseteq \mathbb{R}$ za koje vrijedi $\mathbb{P}(X \in \partial B) = 0$, gdje ∂B označava tzv. *rub* skupa B; na primjer, $\partial(-\infty, x] = \{x\}, \partial[a, b) = \{a, b\}$.

(c) **(!)** CGT (tj. Teorem 7.12) možemo ekvivalentno iskazati kao

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} Z\,,$$

za $Z \sim \mathrm{N}(0, 1)$, jer je F_Z je neprekidna na $\mathbb{R}$ (tj. $C_{F_Z} = \mathbb{R}$).

ZADATAK. Ako su $X, X_1, X_2, \ldots$ diskretne slučajne varijable koje poprimaju vrijednosti u $\mathbb{N}_0$, onda

$$\mathbb{P}(X_n = k) \xrightarrow{n\to\infty} \mathbb{P}(X = k)\,, \quad \forall k \in \mathbb{N}_0\,,$$

povlači $F_{X_n}(t) \xrightarrow{n\to\infty} F_X(t)$, $\forall t \in \mathbb{R}$, pa specijalno vrijedi i $X_n \xrightarrow{d} X$. Primjerice, u zakonu rijetkih događaja (Teorem 3.39) dakle imamo da $\mathrm{B}(n, p_n) \sim X_n \xrightarrow{d} X \sim \mathrm{P}(\lambda)$ ako $np_n \xrightarrow{n\to\infty} \lambda \in (0, \infty)$.

TEOREM 7.24. *Ako vrijedi $X_n \xrightarrow{\mathbb{P}} X$, nužno vrijedi i $X_n \xrightarrow{d} X$.*

DOKAZ. Neka je $t \in C_{F_X}$ proizvoljan. Za sve $\epsilon > 0$, i sve $n \in \mathbb{N}$ imamo

$$\begin{aligned} F_{X_n}(t) &= \mathbb{P}(X_n \leqslant t, |X_n - X| \leqslant \epsilon) + \mathbb{P}(X_n \leqslant t, |X_n - X| > \epsilon) \\ &\leqslant \mathbb{P}(X \leqslant t + \epsilon) + \mathbb{P}\Big(|X_n - X| > \epsilon\Big)\,. \end{aligned}$$

Sada budući da $X_n \xrightarrow{\mathbb{P}} X$, imamo

$$\limsup_{n\to\infty} F_{X_n} \leqslant F_X(t + \epsilon) + 0 = F_X(t + \epsilon)\,,$$

[6] Uočimo, $X_n \xrightarrow{d} X$ ovisi samo o distribucijama slučajnih varijabli $X, X_1, X_2, \ldots$. Specijalno, one uopće ne moraju biti definirane na istom vjerojatnosnom prostoru, kao što je to slučaj s konvergencijama po vjerojatnosti i gotovo sigurno.

pa budući da je $\epsilon > 0$ bio proizvoljan, puštanjem $\epsilon \to 0$, dobivamo da je

$$\limsup_{n\to\infty} F_{X_n} \leqslant \lim_{\epsilon\to 0^+} F_X(t+\epsilon) = F_X(t)\,,$$

jer je svaka funkcija distribucije neprekidna zdesna na $\mathbb{R}$.

Slično se pokaže da za sve $\epsilon > 0$, $n \in \mathbb{N}$, vrijedi

$$F_X(t-\epsilon) \leqslant F_{X_n}(t) + \mathbb{P}\Big(|X_n - X| > \epsilon\Big)$$

pa budući da $X_n \xrightarrow{\mathbb{P}} X$ dobivamo da je

$$\liminf_{n\to\infty} F_{X_n}(t) \geqslant \lim_{\epsilon\to 0^+} F_{X_n}(t-\epsilon) = F_{X_n}(t)\,,$$

pri čemu zadnja jednakost vrijedi jer je $t \in C_{F_X}$. Dakle, pokazali samo da vrijedi $F_{X_n}(t) \xrightarrow{n\to\infty} F_X(t)$, $\forall t \in C_{F_X}$. $\square$

PRIMJER 7.25. Obrat u prethodnom teoremu općenito ne vrijedi! Na primjer, ako je $Z \sim \mathrm{B}(1/2)$, $X_n := Z$ za $n \in \mathbb{N}$, te $X := 1 - Z$. Ovdje je ključno da i $1 - Z$ ima $\mathrm{B}(1/2)$ razdiobu, pa budući da je $F_{X_n} = F_X$, za sve $n \in \mathbb{N}$, trivijalno vrijedi da $X_n \xrightarrow{d} X$. Ipak, po konstrukciji je $|X_n - X| = |2Z - 1| = 1$, pa za bilo koji $\epsilon \in (0,1)$ imamo

$$\mathbb{P}\Big(|X_n - X| > \epsilon\Big) = \mathbb{P}(1 > \epsilon) = 1\,,\quad \forall n \in \mathbb{N}\,.$$

Dakle, niz $(X_n)_n$ ne konvergira po vjerojatnosti prema X.

[Ipak...]

TEOREM 7.26. *Ako za neku konstantu $c \in \mathbb{R}$ vrijedi $X_n \xrightarrow{d} c$, onda vrijedi i $X_n \xrightarrow{\mathbb{P}} c$.*

Dokaz ostavljamo studentima da pogledaju (probaju) sami.

DOKAZ. Za sve $\epsilon > 0, n \in \mathbb{N}$, vrijedi

$$\begin{aligned}\mathbb{P}\Big(|X_n - c| > \epsilon\Big) &= \mathbb{P}(X_n > c+\epsilon) + \mathbb{P}(X_n < c-\epsilon)\\ &\leqslant \mathbb{P}(X_n > c+\epsilon) + \mathbb{P}(X_n \leqslant c-\epsilon) = 1 - F_{X_n}(c+\epsilon) + F_{X_n}(c-\epsilon)\,.\end{aligned}$$

Budući da $X_n \xrightarrow{d} X$ te su $c+\epsilon, c-\epsilon \in C_{F_c}$, imamo

$$0 \leqslant \mathbb{P}\Big(|X_n - c| > \epsilon\Big) \xrightarrow{n\to\infty} 1 - F_c(c+\epsilon) + F_c(c-\epsilon) = 1 - 1 + 0 = 0\,.$$

$\square$

Sve skupa, pokazali smo da

$$X_n \xrightarrow{\text{g.s.}} X \;\;\Rightarrow\;\; X_n \xrightarrow{\mathbb{P}} X \;\;\Rightarrow\;\; X_n \xrightarrow{d} X\,,$$

pri čemu između $\xrightarrow{\mathbb{P}}$ i $\xrightarrow{d}$ imamo ekvivalenciju ako je X konstanta.