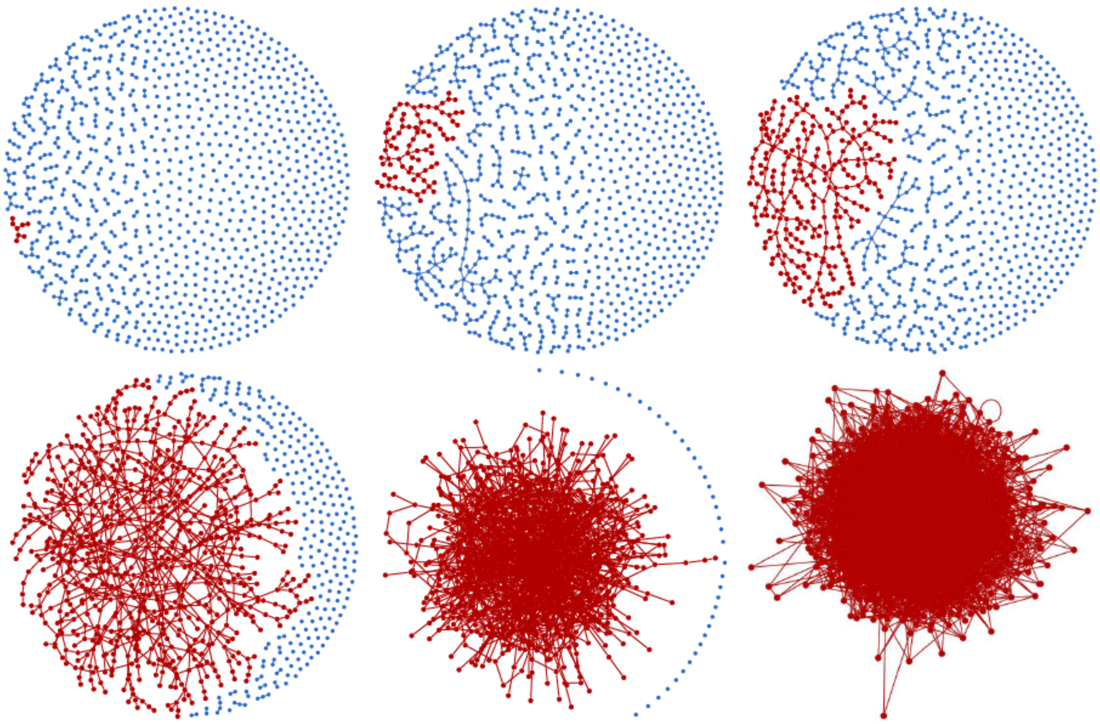
Vjerojatnost

Bilješke s predavanja

Hrvoje Planinić

Prirodoslovno matematički fakultet – Matematički odsjek
Sveučilište u Zagrebu

E-mail adresa: planinic@math.hr



Realizacije tzv. *Erdős–Rényi* slučajnog grafa – svaki par od ukupno n vrhova je spojen bridom s vjerojatnosti p. Gore je $n = 1000$, a p je oblika $p = \frac{c}{n}$ – $c > 0$ (približno) predstavlja prosječan broj susjeda svakog vrha. Vrijednosti za c su, s lijeva na desno, 0.1, 0.5, 1, 1.1, $\frac{\log(n)}{2}$, $2\log(n)$. Vrhovi označeni crvenom bojom predstavljaju najveću povezanu komponentu grafa. Poznati rezultati iz teorije slučajnih grafova kažu da čim je $c > 1$ s velikom vjerojatnosti graf sadrži jednu divovsku komponentu reda veličine $\rho_c n$, za $\rho_c \in (0,1)$, a čim je $c > \log(n)$ da je graf s velikom vjerojatnosti povezan, što je u skladu s gornjim simulacijama. Područje slučajnih grafova je trenutno jedno od najvažnijih i najzanimljivijih područja teorije vjerojatnosti. Slika preuzeta iz [**Cur24**].

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* ili podcrtam stvari koje želim samo naglasiti.

Postoji veliki broj materijala (knjiga i *lecture notesa*) koji se mogu koristiti kao uvod u vjerojatnost – primjerice, klasične knjige su [**GW14**] ili [**Sti03**]. Ipak, daleko najbolje što sam (do sada) vidio je knjiga [**BH19**] – što se tiče teorije, tu nema neke razlike u odnosu na ostale knjige, ali količina motivacije, intuicije, te (modernih i zanimljivih) primjera i zadataka koji se mogu naći u [**BH19**] je nešto što (po mom mišljenju) izdvaja ovu knjigu od ostalih. Još bi ovdje spomenuo jako dobre materijale sa sličnog kolegija na Cambridgeu [**Web**].

Sve uočene greške slobodno javite na `planinic@math.hr`.

Posljednja izmjena napravljena je 22.9.2025.

Izmjene u odnosu na 24./25.

- dodan je motivacijski primjer o Erdős–Rényi slučajnom grafu (vidi par stranica gore);
- Primjeri 1.18 i 1.25(b) su prebačeni na vježbe (u Poglavlje 2), a ideja o kontrukciji vjerojatnosnog prostora za beskonačan niz bacanja novčića dodana je kroz Potpoglavlje 1.4.3.
- termini ”neopadajući”/”nerastući” su (nadamo se) svugdje zamijenjeni s ”rastući”/”padajući”. Ukoliko budemo htjeli naglasiti strogi rast ili pad, koristit ćemo termine ”strogo rastući”/”strogo padajući”.
- reorganizacija Poglavlja 4 – dodano zasebno poglavlje o varijanci sume nezavisnih varijabli, primjer s kontrolnim varijatama prebačen na vježbe, a kovarijanca/korelacija prebačena pred kraj poglavlja.
- Primjer 3.30 je izmijenjen i pojednostavljen. Dio je prebačen na vježbe.
- Zadaci iz Pogl. 5.3 prebačeni na vježbe te Propozicija 5.18 odmah iskazana općenitije tako da je izbačena napomena nakon nje.
- Poglvalja 7 i 8 su zamijenjena. Stari Primjer 8.23 se nije ispitivao – sada se ispituje te je pomaknut ranije kao Primjer 7.10. Dodan je Primjer 8.13.

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 ...[1]]

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 svake minute 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 minuta bio u vodstvu. Je li to dovoljno dokaza da bi zaključili da je novčić nesimetričan?

[1] U zadnje vrijeme najprestižnije nagrade za doprinost matematici vrlo često dobivaju matematičari čiji je rad direktno ili indirektno vezan uz teoriju vjerojatnosti. Primjerice, Abelovu nagradu za 2024. godinu dobio je Michel Talagrand, između ostalog za svoj rad na tzv. *koncentracijskim nejednakostima*, a 2022. Fieldsovu medalju dobio je Hugo Duminil-Copin za svoj rad u *teoriji perkolacija*.

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 $\mathcal{F} \subseteq \mathcal{P}(\Omega)$[2])

(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[3]

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

(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 $\omega \in A$. 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".

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

[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**. □

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 $A_1, A_2, \dots, A_n \in \mathcal{F}$, 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).

□

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 $A_1, A_2, \dots, A_n \in \mathcal{F}$, 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 $\Omega = \{1, 2, \dots, 6\}$ 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 $A_1, A_2, \dots \in \mathcal{F}$[4] 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 $|\mathcal{F}| < \infty$,[5] (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.[6]

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

[4] Ima ih *prebrojivo* mnogo.
[5] $|\mathcal{F}|$ označava broj elemenata skupa $\mathcal{F}$.
[6] 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[7]

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

Na primjer,

- za $A = \{\text{točka je upala u krug upisan } [-1,1]^2\} = \{(x,y) \in \Omega : \sqrt{x^2+y^2} \leqslant 1\}$ imamo
$$\mathbb{P}(A) = \frac{1^2\pi}{2^2} = \frac{\pi}{4}.$$
- za $B = \{\text{izabrali smo točku } (0,0)\} = \{(0,0)\}$ imamo[8]
$$\mathbb{P}(B) = \frac{\lambda(\{(0,0)\})}{2^2} = \frac{0}{4} = 0.$$

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 $[-1,1]^2$.[9]

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* $\mathbb{P}$*)*[10]*;*

(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)).

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

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

[9] 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 $\lambda(A \cup B) = \lambda(A) + \lambda(B)$.

[10] 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[11]

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

Dakle, svaka vjerojatnost $\mathbb{P}$ na $(\Omega, \mathcal{F})$ je u potpunosti određena s vrijednostima $\mathbb{P}(\omega), \omega \in \Omega$.[12]

Primjer 1.10. Ako je $|\Omega| < \infty$, tj. $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ 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 $i = 1, \ldots, n$, te

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

tj. imamo Laplaceov model. $\square$

Primjer 1.11. Ako je $\Omega = \{\omega_1, \omega_2, \ldots\}$ (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 $\Omega_n := \{1, \ldots, n\}$. 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

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

[Ipak, treba uvijek biti svjestan da za ovaj model zapravo ne postoji dobro definiran vjerojatnosni prostor, vidi Primjer 1.21.] $\square$

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

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

Propozicija 1.12. *Neka je* $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ *za* $n \in \mathbb{N}$ *te* $\mathcal{F} = \mathcal{P}(\Omega)$. *Za svaki vektor* $(p_1, \ldots, p_n) \in \mathbb{R}^n$ *koji zadovoljava*[13]

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

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

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

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

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

[Napomena. Diskretan model koji smo promatrali u ovom potpoglavlju je generalizacija Laplaceovog modela. S druge strane, (vrlo važnu) generalizaciju "slučajnog odabira točke" iz neprebrojivog skupa $\Omega \subseteq \mathbb{R}^2$ iz Primjera 1.8 dobijemo tako da stavimo

$$\mathbb{P}(A) = \int_A f(x)\,\mathrm{d}x\,, \quad A \subseteq \Omega\,.$$

za nenegativnu funkciju $f : \Omega \to \mathbb{R}$. Drugim riječima, $\mathbb{P}(A)$ je površina ispod grafa funkcije f – model iz Primjera 1.8 dobijemo tako da za f uzmemo konstantnu funkciju na (omeđenom) skupu Ω. Takvim modelima ćemo se baviti u poglavljima o neprekidnim slučajnim varijablama i vektorima. Opet ponavljamo glavnu poruku – apstraktan aksiomatski pristup vjerojatnosti koji smo opisali omogućuje nam da ove dvije fundamentalno različite klase primjera (kao i mnoge druge) tretiramo kao specijalne slučajeve jedne te iste teorije.]

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.

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

(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

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

(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[14]

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

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

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

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 $\{1, 2, \dots, 2n\}$, 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

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

[14] 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 $n \to \infty$.

Sada je $A = \{\{i,j\} : i \in \{1,\dots,n\}, j \in \{n+1,\dots,2n\}\}$. 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. $\square$

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

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

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

(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

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

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

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

$\square$

[Idući primjer je sličan te ostavljamo da ga studenti sami pogledaju.]

PRIMJER 1.15. Imamo špil od 52 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 $1,\dots,52$ te uzmimo

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

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

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

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

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

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.

$\square$

[**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. Također, bitno je uvijek provjeriti nalazimo li se zaista u Laplaceovom modelu, tj. jesu li zaista svi elementarni događaji jednako vjerojatni.]

1.3.1. Formula uljučivanja-isključivanja.

PROPOZICIJA 1.16. *Neka je $n \in \mathbb{N}$ te $A_1, A_2, \ldots, A_n$ događaji u $\mathcal{F}$. Tada vrijedi*[15]

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

DOKAZ. Indukcijom po n. Za $n = 1$ tvrdnja očito vrijedi, a za $n = 2$ to je točno Propozicija 1.9(f). Ostatak za DZ.[16] [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.] □

PRIMJER 1.17 (**Deranžmani**). Kažemo da permutacija π skupa $\{1, 2, \ldots, n\}$[17] 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 $\{1, \ldots, n\}$[18] 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

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

[15] Na primjer, $\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)$.
[16] Kasnije ćete na vježbama isti rezultat dokazati puno elegantnije koristeći tzv. indikatorske slučajne varijable.
[17] Dakle, π je bijekcija s $\{1, 2, \ldots, n\}$ u $\{1, 2, \ldots, n\}$.
[18] Kako biste u praksi generirali ovu slučajnu permutaciju?

Događaji $A_1, \dots, A_n$ nisu u parovima disjunktni pa moramo koristiti FUI. Za to nam trebaju sljedeće vjerojatnosti[19]

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

Također, uočimo da je k-članih podksupova $\{i_1, i_2, \dots, i_k\}$ skupa $\{1, 2\dots, n\}$ 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

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

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.*]

□

1.4. Nizovi događaja

1.4.1. Subaditivnost. [Booleova nejednakost]

PROPOZICIJA 1.18. *Za svaki niz događaja $A_1, A_2, \dots \in \mathcal{F}$, vrijedi*

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

[19] 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.

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

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

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. □

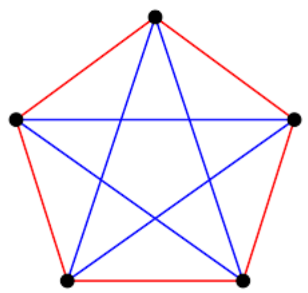
Primjer 1.19. (**Erdős – ”vjerojatnosna metoda” u kombinatorici**)

Gledamo *potpun* graf s n vrhova.[20] 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).

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* $i = 1, \ldots, \binom{n}{k}$, vrijedi

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

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

[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 1.20. 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 $R(k,k) \geqslant n+1$. Ipak ta ograda je dosta gruba – 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 $\mathbb{P}(B) = 0$. □

Primjer 1.21. Slučajno biramo bilo koji prirodan broj. Ako je $A_k := \{\text{izabrali smo broj } k\}$, kao ranije mogli bismo definirati

$$\mathbb{P}(A_k) := \lim_{n\to\infty} \mathbb{P}(\{\text{izabrali } k \text{ iz skupa } \{1,\dots,n\}\}) = \lim_{n\to\infty} \frac{1}{n} = 0\,, \quad \forall k \in \mathbb{N}\,.$$

Ipak,

$$\mathbb{P}(\cup_{k=1}^{\infty} A_k) = \mathbb{P}(\{\text{izabrali smo broj iz } \mathbb{N}\}) = 1\,.$$

Je li ovo kontradikcija s prethodnim zadatkom?[21]

1.4.2. Neprekidnost vjerojatnosti. Za niz događaja $A_1, A_2, \dots \in \mathcal{F}$ kažemo da je *rastući* (odnosno, *padajući*) 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 rastući/padajući niz događaja)*

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

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

[21] Ovo pitanje postavio je student Matej Vojvodić.

(b) Ako je niz $A_1, A_2, \dots \in \mathcal{F}$ padajući, 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 1.18. 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 $A_1^c, A_2^c, \dots$ rastući, iz (a) dijela dobivamo

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

□

[Prethodni rezultat tipično koristimo na $(\Omega, \mathcal{F}, \mathbb{P})$ gdje je $|\Omega| = \infty$ (i tipično neprebrojiv).]

1.4.3. $(\Omega, \mathcal{F}, \mathbb{P})$ za beskonačan niz bacanja novčića. [Želimo konstruirati vjerojatnosni model za beskonačan niz bacanja simetričnog novčića.]

Dakle

$$\Omega := \{(a_1, a_2, \dots) : a_i \in \{P, G\}, \forall i \in \mathbb{N}\} = \{P, G\}^{\mathbb{N}} .$$

Zanimati nas mogu događaji poput

- $B_1 := \{\text{padala su samo pisma}\} = \{PPP\cdots\}$;
- $B_2 = \{\text{palo je beskonačno mnogo glava}\}$;
- $B_3 = \{(a_1, a_2, \dots) \in \Omega : \lim_{n\to\infty} \frac{|\{i \leqslant n : a_i = P\}|}{n} = \frac{1}{2}\}$ [tj. asimptotski je broj P i G jednak].

Uočimo da je Ω neprebrojiv skup – dakle, nije dovoljno zadati $\mathbb{P}(\omega)$ za svaki niz ω [zapravo, imat ćemo $\mathbb{P}(\omega) = 0$ za sve $\omega \in \Omega$!].

Za sve $n \in \mathbb{N}$ neka $\Omega_n := \{P, G\}^n$ predstavlja sve moguće ishode u prvih n bacanja. Kažemo da $A \subseteq \Omega$ ovisi o prvih n bacanja (oznaka $A \in \mathcal{F}_n$) ako postoji $A_n \subseteq \Omega_n$ t.d. vrijedi

$$A = \{(a_1, a_2, \dots) \in \Omega : (a_1, \dots, a_n) \in A_n\} .$$

[Dakle, ishod $(a_1, \dots, a_n)$ je u A_n, a $a_{n+1}, a_{n+2}, \dots$ mogu biti bilo što.]

Na primjer,

- ako je $A \subseteq \Omega$ skup svih nizova u kojima smo imali točno jedno pismo u prva 3 bacanja, očito je $A \in \mathcal{F}_3$ pri čemu je $A_3 = \{PGG, GPG, GGP\}$, ali i $A \in \mathcal{F}_n$ za sve $n \geqslant 4$;
- očito je $A := \Omega \in \mathcal{F}_1$ pri čemu je $A_1 = \{P, G\}$.

Ako je $A \in \mathcal{F}_n$ za neki $n \geqslant 1$, budući da na Ω_n imamo Laplaceov model, prirodno je definirati

$$\mathbb{P}(A) := \frac{|A_n|}{|\Omega_n|} = \frac{|A_n|}{2^n}.$$

[Budući da ja $A \in \mathcal{F}_k$ za sve $k \geqslant n$, nije apriori jasno da je gornja definicija konzistenta, ali to se lako provjeri (DZ)]. Time smo pokrili sve skupove iz familije $\mathcal{F}_\infty := \cup_{n=1}^\infty \mathcal{F}_n$. Ipak, $\mathcal{F}_\infty$ nije σ-algebra – na primjer, za sve n, $B^{(n)} := \{\text{u prvih } n \text{ bacanja pala samo pisma}\}$ je u $\mathcal{F}_\infty$, ali $\cap_{n=1}^\infty B^{(n)} = \{PPPPP\cdots\} = B_1$ nije u $\mathcal{F}_n$ niti za jedan n.

Ipak, netrivijalan rezultat iz mjere i integrala kaže da se $\mathbb{P}$ može jedinstveno proširiti do vjerojatnosti na $(\Omega, \mathcal{F})$ pri čemu je $\mathcal{F}$ <u>najmanja</u> σ-algebra koja sadrži $\mathcal{F}_\infty$. Na primjer,

$$B_1 = \{PPPPP\cdots\} = \cap_{n=1}^\infty B^{(n)} \in \mathcal{F},$$

jer je za svaki n, $A^{(n)} \in \mathcal{F}_n \subseteq \mathcal{F}$ i $\mathcal{F}$ je *sigma*-algebra – dakle, $\mathbb{P}(B_1)$ je dobro definirana!

Vjerojatnosti $\mathbb{P}(A)$ za $A \in \mathcal{F} \setminus \mathcal{F}_\infty$ tipično računamo koristeći svojstva od $\mathbb{P}$ [neprekidnost σ-aditivnost itd.].

PRIMJER 1.23. Odredimo $\mathbb{P}(B_1)$.

RJEŠENJE. *1.* Budući da je $B^{(n)}$, $n \geqslant 1$, *padajući* niz događaja, neprekidnost od $\mathbb{P}$ povlači

$$\mathbb{P}(B_1) = \mathbb{P}(\cap_{n=1}^\infty A^{(n)}) = \lim_{n\to\infty} \mathbb{P}(A^{(n)}) = \lim_{n\to\infty} \frac{1}{2^n} = 0.$$

2. Imamo

$$B_1^c = \{\text{pala barem jedna G}\} = \cup_{n=1}^\infty \underbrace{\{\text{G prvi put pala u } n\text{-tom bacanju}\}}_{=:G_n}$$

pa po σ-aditivnosti (G_n-ovi su u parovima disjunktni) imamo

$$\mathbb{P}(B_1^c) = \sum_{n=1}^\infty \mathbb{P}(G_n) = \sum_{n=1}^\infty \frac{1}{2^n} = \frac{1}{2} \cdot \frac{1}{1 - \frac{1}{2}} = 1.$$

Dakle, $\mathbb{P}(B_1) = 1 - \mathbb{P}(B_1^c) = 0$.

[Uočite da mala modifikacija gornjih dokaza pokazuje da je $\mathbb{P}(\{\omega\}) = 0$ za svaki fiksni niz $\omega \in \Omega$.] $\square$

[Prethodni rezultat u skladu da našom intuicijom i čini se potpuno očit. Ipak, poanta ovdje je da smo to dokazali u sklopu naše aksiomatske teorije gdje nije apriori jasno da to mora vrijediti. Ključna stvar ovdje je da je zaista moguće konstruirati vjerojatnosni model beskonačnog niza bacanja novčića.]

ZADATAK. Pokažite da su i B_2 i B_3 (definirani gore) također elementi of $\mathcal{F}$. [Napomenimo da vrijedi $\mathbb{P}(B_2) = \mathbb{P}(B_3) = 1$ – za B_2 to će slijediti iz tzv. Borel-Cantellijeve leme, a za B_3 to je tzv. jaki zakon velikih brojeva.]

POGLAVLJE 2

Uvjetna vjerojatnost i nezavisnost

2.1. Uvjetna vjerojatnost

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

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

- 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

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

  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, $I = \{1, 2, \dots, n\}$ ili $I = \mathbb{N}$), kažemo da je **potpun sistem događaja** (PSD) ako vrijedi (a) $H_i \cap H_j = \emptyset$, 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 $H_i \cap H_j = \emptyset$, 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. 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 $a = \frac{1}{6} + \frac{5}{6}c = \frac{36}{91}$

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 0.5% 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.]

$\square$

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

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

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] $= \frac{n-1}{2n-1}$ ($< \frac{1}{2} = \mathbb{P}(A_2)$), 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!] $\square$

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

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

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

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

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). □

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 $A_1, A_2, \dots$ 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

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

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 $\mathbb{P}(A) = \frac{p}{1-r}$.

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.

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

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

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

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 koju sada dokazujemo.]

□

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

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

□

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 $(B_n)_{n\in\mathbb{N}}$ padajući 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

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

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 $\mathbb{P}(B_n^c) = 0$. □

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"[4]. □

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 crna 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[5]

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

Budući da je *za sve* $\omega \in \Omega$, $\mathbb{P}(\omega) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$, nalazimo se u Laplaceovom 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").] □

[4] https://en.wikipedia.org/wiki/Infinite_monkey_theorem

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

[Sljedeći klasični primjer ostavljamo studentima da sami pogledaju. "Paradoks" nastaje zbog toga što vjerojatnosni prostor nije jednoznačno zadan.]

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?

RJEŠENJE. 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.] $\square$

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.

(a) Potpuno analogno se konstruira vjerojatnost $\mathbb{P}_n$ na $(\Omega_1 \times \cdots \times \Omega_n, \mathcal{P}(\Omega_1 \times \cdots \times \Omega_n))$.

(b) *Konstrukcija vjerojatnosti na $\Omega := \Omega_1 \times \Omega_2 \times \ldots$, tj. za <u>beskonačan</u> niz nezavisnih pokusa može se napraviti koristeći $\mathbb{P}_n$-ove iz (a) dijela, vrlo slično kao u Potpoglavlju 1.4.3.

[Za $A \subseteq \Omega$ koji je za neki n i $A_n \subseteq \Omega_1 \times \cdots \times \Omega_n$ oblika

$$A = \{(\omega_1, \omega_2, \ldots) \in \Omega : (\omega_1, \ldots, \omega_n) \in A_n\}$$

stavljamo $\mathbb{P}(A) := \mathbb{P}_n(A_n)$, te se onda $\mathbb{P}$ može proširiti do funkcije na $\mathcal{F}$ koje je najmanja σ-algebra na Ω koje sadrži sve takve A-ove. Napomenimo da je ovdje $\mathcal{F} \subsetneq \mathcal{P}(\Omega_1 \times \Omega_2 \times \ldots)$.]

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 $\mathcal{F} = \mathcal{P}(\Omega)$. $\square$

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. (a) Bacamo dvije kocke te neka je

$$X := \text{zbroj brojeva na kockama}\,,$$

tj. $\Omega = \{1, 2, \ldots, 6\}^2$, $X(\omega) := \omega_1 + \omega_2$, $\forall \omega = (\omega_1, \omega_2) \in \Omega$. Budući da je $X : \Omega \to D_X := \{2, \ldots, 12\} \subseteq \mathbb{R}$, X je diskretna slučajna varijabla [kao i svaka druga slučajna varijabla na ovom vjerojatnosnom prostoru].

(b) 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, ali samo je Y diskretna ($D_X = \langle 0, 10\rangle$, $D_Y = \{0, 1, \ldots, 9\}$).

□

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.] □

NAPOMENA 3.5 (**Slučajni element**). Općenito, ako je D proizvoljan skup, funkciju $X : \Omega \to D$ nazivamo **slučajni element** u D. Na primjer,

- $D = \mathbb{R} \rightsquigarrow X$ je slučajna varijabla;
- $D = \mathbb{R}^k \rightsquigarrow X$ je **slučajni vektor** [iduće poglavlje];
- $D =$ skup svih grafova s n vrhova $\rightsquigarrow X$ je **slučajan graf** [npr. Erdős-Rényijev graf.[1] Kako bi tu konstruirali $(\Omega, \mathcal{F}, \mathbb{P})$ i $X : \Omega \to D$?]
- $D = \mathrm{C}[0, 1] \rightsquigarrow X$ je **slučajna funkcija** [npr. Brownovo gibanje[2]];

[Slučajni element X je diskretan ako je slika $D_X \subseteq D$ najviše prebrojiv skup; to mogu biti i relativno komplicirani slučajni elementi kao što je npr. slučajni graf. Većina definicija i rezultata koje ćemo navesti u ovom poglavlju vrijede i za *diskretne* slučajne elemente uz potpuno iste dokaze – jednostavno se elementi od D_X poredaju u niz $a_1, a_2, \ldots$; to ćemo direktno koristiti u poglavlju o diskretnim slučajnim vektorima. Iznimka su rezultati gdje se zaista koristi struktura prostora $\mathbb{R}$ – npr. kad govorimo o *očekivanju* slučajne varijable.]

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**:

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

[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 $\{1, 2, \ldots, 6\}$. □

[1] https://en.wikipedia.org/wiki/Erdős-Rényi_model
[2] https://en.wikipedia.org/wiki/Wiener_process

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

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

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**[3] 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

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

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 previše koristiti f_X, ali to je samo stvar notacije. Ona je korisna npr. u statistici jer možemo na unificiran način tretirati diskretne i neprekidne slučajne varijable, vidi Primjere 4.8 i 8.7]

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

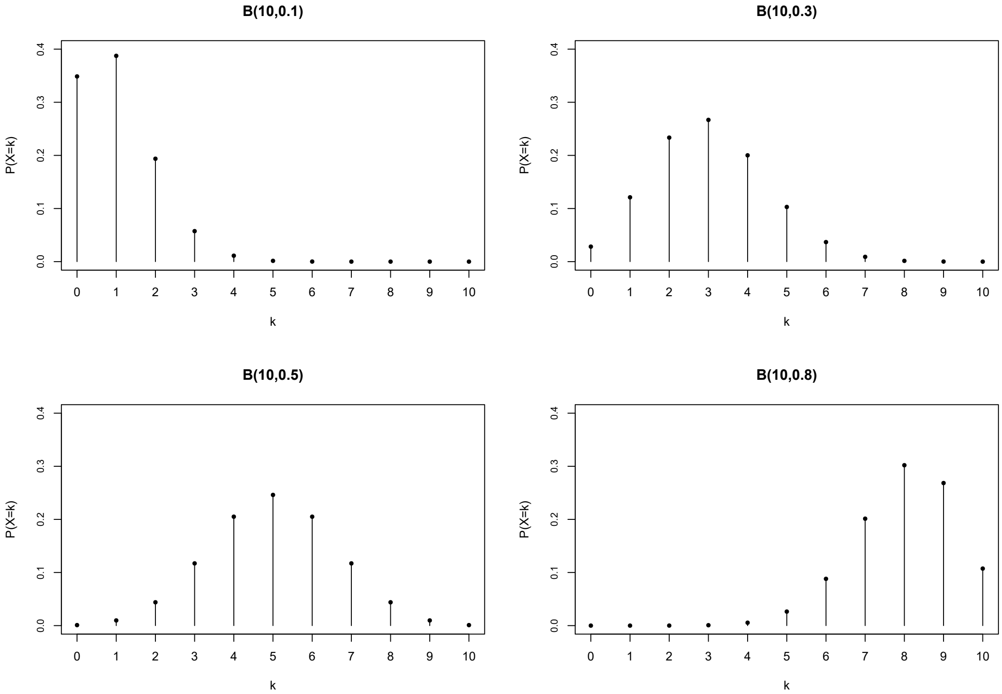
$$X_i := \begin{cases} 1, & \text{uspjeh u } i\text{-tom pokusu}\,, \\ 0, & \text{inače} \end{cases}$$

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".]

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



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

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 prethodnom primjeru vrijedi $X_i = \mathbb{1}_{A_i}$, $\forall i$.

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

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

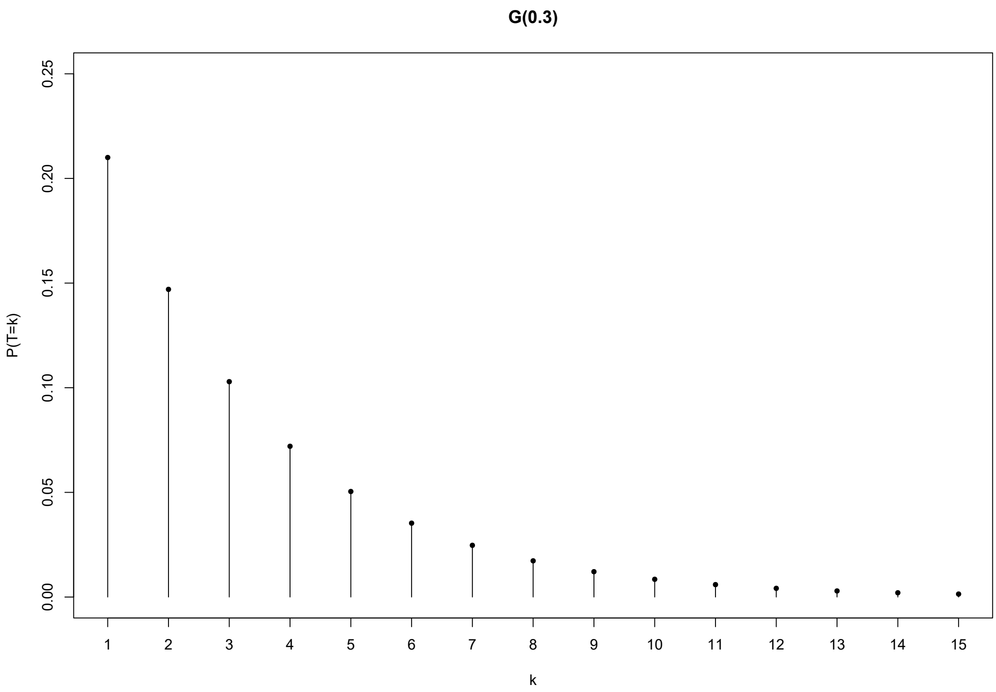
imamo $X \in \{0, 1, \ldots, n\}$ te

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

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

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



Slika 3. Grafički prikaz geometrijske razdiobe.

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,

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

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

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

Ova razdioba naziva se **negativna binomna** razdioba s parametrima r i p (oznaka "$\mathrm{NB}(r, p)$").

[Za $r = 1$, $Y \sim \mathrm{G}(p)$, $Z \sim \mathrm{G}_0(p)$. 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 (a) $X \geqslant 0$ (tj. $a_i \geqslant 0$, $\forall i \in I$), ili

$$\text{(b)}\quad \sum_{i \in I} |a_i| p_i < \infty\,, \tag{3.8}$$

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 $X \geqslant 0$ moguće imamo $\mathbb{E}[X] = +\infty$, a ako vrijedi (3.8) nužno vrijedi $\mathbb{E}[X] \in \mathbb{R}$.

(iii) **Zašto tražimo nenegativnost ili (3.8)?** Ako je $(b_n)_{n\in\mathbb{N}}$ niz realnih brojeva te vrijedi (a) $b_n \geqslant 0$, $\forall n$, ili (b) $\sum_{n\in\mathbb{N}} |b_n| < \infty$, 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[4]

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

Vidi Sliku 4. □

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

$$\mathbb{P}(X = n) = \frac{c}{n^2} \,, \ \forall n \in \mathbb{N} \,,$$

gdje je $c > 0$ normalizirajuća konstanta, imamo

$$\mathbb{E}[X] = \sum_{n=1}^{\infty} n\frac{c}{n^2} = c\sum_{n=1}^{\infty} \frac{1}{n} = +\infty \,.$$

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

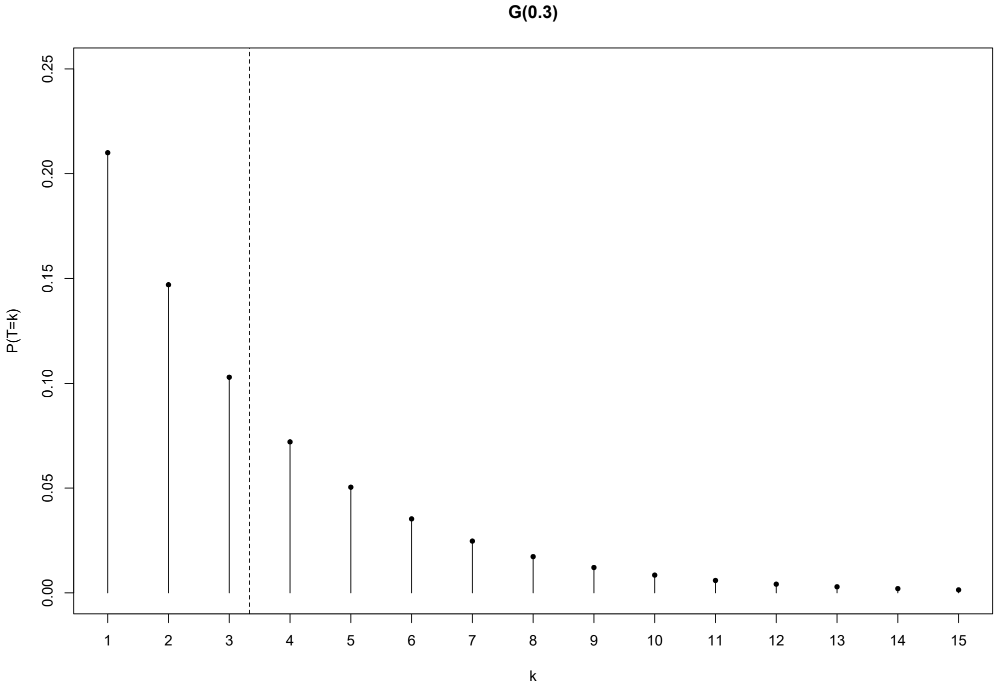
PRIMJER 3.20. Neka je $X \in \mathbb{Z}$ s distribucijom

$$\mathbb{P}(X = n) = \mathbb{P}(X = -n) = \frac{c}{n^2} \,, \forall n \in \mathbb{N} \,,$$

uz $\mathbb{P}(X = 0) = 0$, pri čemu je $c > 0$ normalizirajuća konstanta. Tada je

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

[4] 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 $|x| < 1$.



SLIKA 4. Geometrijska razdioba s parametrom $p = 0.3$. Isprekidana linija označava njeno očekivanje $\mathbb{E}[X] = 1/0.3 \approx 3.33$.

Dakle, $\mathbb{E}[X]$ *ne postoji* jer ne vrijedi (3.8) [te X nije nenegativna].

[Zbog simetrije se ovdje čini prirodno staviti $\mathbb{E}[X] := 0$. Ipak, pojam očekivanja je usko vezan uz (jaki) zakon velikih brojeva (JZVB) koji kaže da će prosječna vrijednost velikog broja (nezavisnih) varijabli s istom distribucijom kao X, biti približno jednaka $\mathbb{E}[X]$. Nužan i dovoljan uvjet za JZVB (ako je $X \in \mathbb{R}$) je to da vrijedi (3.8), te se može pokazati za slučajnu varijablu iz prethodnog primjera prosječna vrijednost neće konvergirati niti jednoj konstanti, pa tako ni konstanti 0. Vidi napomenu iza Teorema 7.11.]

3.2.1. Osnovna svojstva.

NAPOMENA 3.21. 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 3.22 (!). *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

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

  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 $\mathbb{E}[g(X)] \in \mathbb{R}$].

□

Napomena 3.23. Za $X \sim \left(\begin{smallmatrix} a_1 & a_2 & \cdots \\ p_1 & p_2 & \cdots \end{smallmatrix}\right)$ 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 $\mathbb{E}[X]$ i $\mathbb{E}[X] \in \mathbb{R}$) akko $\mathbb{E}\Big[|X|\Big] < \infty$.

Propozicija 3.24. *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 $X \sim \left(\begin{smallmatrix} a_1 & a_2 & \cdots \\ p_1 & p_2 & \cdots \end{smallmatrix}\right)$, (3.10) nam daje

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

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

Sljedeći teorem dokazujemo u poglavlju o slučajnim vektorima.

PROPOZICIJA 3.25. *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**. [Ovo je jedno od najvažnijih svojstava očekivanja te ono zapravo vrijedi za proizvoljne (dakle, ne nužno diskretne) slučajne varijable.]

PRIMJER 3.26 (**Očekivanje** $\mathrm{B}(n,p)$ **razdiobe**). Neka je $X \sim \mathrm{B}(n,p)$.

1. [Pogledati sami] 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, tj. nije bitno kako je ona konstruirana]. Dakle,

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

[Navodimo još neka svojstva očekivanja koja slijede direktno iz definicije, pri čemu u iskazima tražimo samo da određeno svojstvo vrijedi na događaju vjerojatnosti 1.]

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* $a \leqslant \mathbb{E}[X] \leqslant b$.

DOKAZ. (a) Ako je $X \sim \left(\begin{smallmatrix} a_1 & a_2 & \cdots\cdots \\ p_1 & p_2 & \cdots\cdots \end{smallmatrix}\right)$, imamo $a_i < 0 \Rightarrow p_i = 0$, pa je

$$\mathbb{E}[X] = \sum_{i \in I} a_i p_i = \sum_{a_i \geqslant 0} a_i p_i \geqslant 0\,.$$

Ostatak za DZ. □

3.2.2. Očekivanje za varijable u $\mathbb{N}_0$.

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

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

□

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

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

[U sljedećem primjeru promatramo generalizaciju geometrijske distribucije – čekamo prvi uspjeh u nizu nezavisnih pokusa, ali pri čemu se vjerojatnost uspjeha može mijenjati.]

PRIMJER 3.30 (**Proširena slučajna varijabla**). Pretpostavimo da imamo niz nezavisnih novčića te neka je

$$A_n := \{\text{palo pismo na } n\text{-tom novčiću}\} , \quad n \geqslant 1 ,$$

uz pretpostavku $p_n := \mathbb{P}(A_n) < 1, \forall n \in \mathbb{N}$. Definiramo

$$T := \inf\{n \geqslant 1 : \mathbb{1}_{A_n} = 1\} = \text{trenutak kada je palo prvo pismo} ,$$

uz konvenciju $\inf \emptyset := \infty$. Dakle, slučajna varijabla T moguće poprima vrijednost $+\infty$ [tj. moguće da nikada neće pasti pismo], te ju zovemo **proširena** (diskretna) slučajna varijabla.

Ako je $\sum_{n=1}^{\infty} p_n = \infty$ (primjerice $p_n \sim \frac{1}{n}$ kada $n \to \infty$), 2. BC lema povlači da je $\mathbb{P}(T = \infty) = 0$ jer g.s. imamo beskonačno mnogo pisama. Što ako je $\sum_{n=1}^{\infty} p_n < \infty$ [dakle, ako p_n dovoljno brzo teži u 0]?

Za sve $n \geqslant 0$,

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

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

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

Iz $\lim_{x\to 0} \frac{\log(1+x)}{x} = 1$ lako slijedi da

$$\sum_{k=1}^{\infty} p_k < \infty \;\Rightarrow\; \prod_{k=1}^{\infty} (1 - p_k) > 0\,.$$

[Zapravo vrijedi ekvivalencija ako je $p_k > 0$ za svaki $k \geqslant 1$. Zaista,

$$\lim_{n\to\infty} \prod_{k=1}^{n} (1 - p_k) > 0 \iff \lim_{n\to\infty} \sum_{k=1}^{n} \log(1 - p_k) > -\infty \iff \sum_{k=1}^{\infty} b_k < \infty\,,$$

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

$$\sum_{j=1}^{\infty} b_k < \infty \iff \sum_{j=1}^{\infty} p_k < \infty \;\;]$$

Sve skupa, pokazali smo da

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

Na primjer, ako je $p_n \sim \frac{1}{n^2}$ i $p_n < 1$ za sve $n \geqslant 1$, vrijedi $\mathbb{P}(T = \infty) > 0$.

Napomena. Ako je $\mathbb{P}(T = \infty) > 0$, stavljamo $\mathbb{E}[T] := +\infty$. Nadalje, uočimo da formula (3.14) vrijedi i u dalje jer je

$$\sum_{n\geqslant 0} \mathbb{P}(T > n) \geqslant \sum_{n\geqslant 0} \mathbb{P}(T = \infty) = +\infty = \mathbb{E}[T]\,.$$

Oprez: S druge strane, moguće je da u tom slučaju vrijedi

$$+\infty > \sum_{n=0}^{\infty} n\mathbb{P}(T = n) \neq \mathbb{E}[T]\,.$$

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

$$X \sim \begin{pmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad Y \sim \begin{pmatrix} -100 & 100 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

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[5] i činjenicu da je $\mathbb{E}[X]$ konstanta.

□

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

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

[5] Napomenimo ovdje da formalno ne možemo primijeniti (3.13) ukoliko je $\mathbb{E}[X^2] = +\infty$. Ipak, može se pokazati da tvrdnja vrijedi i u tom slučaju jer smo pretpostavili da je $\mathbb{E}\big[|X|\big] < \infty$.

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

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

(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].

[Uz linearnost, drugo fundamentalno svojstvo očekivanja je monotonost. Kao i linearnost, ovo svojstvo vrijedi za proizvoljne (dakle, ne nužno samo diskretne) slučajne varijable.]

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*

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

Dokaz.

- ako je $\mathbb{E}[X], \mathbb{E}[Y] < \infty$, iz $Y - X \geqslant 0$ i linearnosti očekivanja slijedi[6]

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

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

□

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

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 $\mathbb{E}\Big[|X|^n\Big]$ 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

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

$\square$

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 $\sum_{k \in \mathbb{N}_0} \mathbb{P}(X = k) = 1$.]

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

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

pri čemu je $\underline{X \sim \mathrm{P}(\lambda)}$.

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

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

pri čemu smo za konvergenciju trećeg člana koristili Lemu 7.16 (koju dakle navodimo tek kasnije). □

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

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

Napomena 3.40 (**Le Camov teorem**). 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 (3.18) 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 $p = 2 \cdot 10^{-6}$.]

□

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.

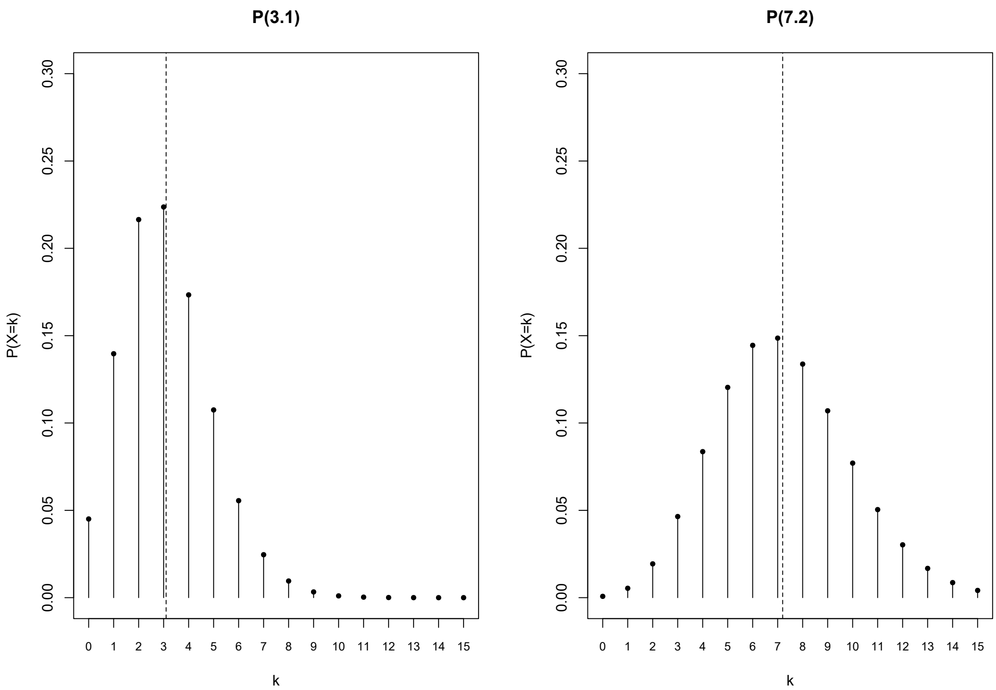
[Slično kao gore,

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

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

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



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 $k = \lfloor \lambda \rfloor$.

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 više, tj. da vrijedi

$$\mathbb{P}(X_n = r) \to \mathbb{P}(X = r)\,, \quad \forall r \in \mathbb{N}_0\,. \tag{3.20}$$

Zaista, za fiksan $r \geqslant 1$, čim je $n \geqslant r$, zbog simetrije vrijedi

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

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.[7]

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

(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

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

□

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

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 $\mathrm{Var}(N)$.[8]

Rješenje. Ako definiramo događaje

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

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 $\mathrm{P}(2)$ distribuciji kada $n \to \infty$. Napomenimo da ćemo kasnije uvesti pojam kovarijance koji će olakšati račun varijance u gornjem primjeru.] □

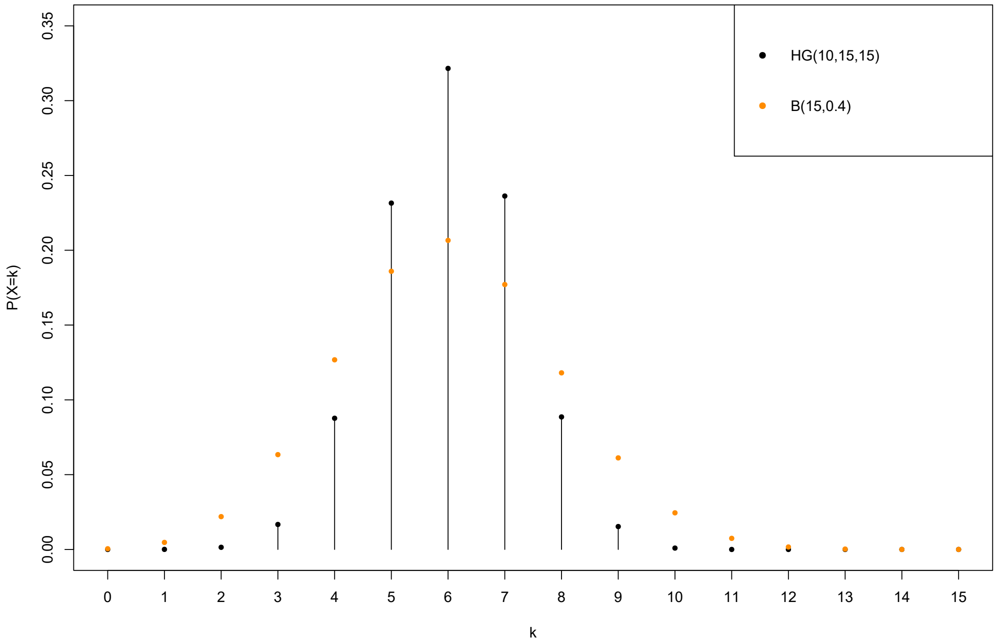
Primjer 3.46 (**Hipergeometrijska razdioba**). U kutiji je m_1 bijelih i m_2 crnih kuglica. Izaberemo slučajno $n \leqslant m_1 + m_2$ kuglica. Ako je X ukupan broj izvučenih bijelih kuglica, očito je $X \in \{0, \dots, n\}$, te preciznije

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

Nadalje, distribucija je dana s

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

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



Slika 6. Usporedba hipergeometrijske i binomne razdiobe.

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

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

imamo[9]

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

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

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

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.]

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

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

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

Sada iz linearnosti i monotonosti očekivanja odmah slijedi subaditivnost

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

[Slično se alternativno može dokazati FUI.]

3.6. Uvjetne razdiobe

Neka je $(\Omega, \mathcal{F}, \mathbb{P})$ vjerojatnosni prostor i $X : \Omega \to \mathbb{R}$ 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 $\left(\begin{smallmatrix} a_1 & a_2 & \cdots\cdots \\ p_1^B & p_2^B & \cdots\cdots \end{smallmatrix}\right)$ 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

$$\begin{aligned} \mathbb{E}[X] &= \sum_{i \in I} a_i \mathbb{P}(X = a_i) \overset{\text{FPV}}{=} \sum_{i \in I} a_i \sum_{j \in J} \mathbb{P}(H_j) \mathbb{P}(X = a_i \mid H_j) \\ &= [\text{Fubini (sve je nenegativno)}] = \sum_{j \in J} \mathbb{P}(H_j) \sum_{i \in I} a_i \mathbb{P}(X = a_i \mid H_j) \\ &= \sum_{j \in J} \mathbb{P}(H_j) \mathbb{E}[X \mid H_j] \in [0, \infty] . \end{aligned} \tag{3.22}$$

(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

$$\sum_{j\in J}\mathbb{P}(H_j)\sum_{i\in I}|a_i|\mathbb{P}(X=a_i \mid H_j) = \sum_{j\in J}\mathbb{P}(H_j)\mathbb{E}\Big[|X| \mid H_j\Big] < \infty\,.\ ]$$

□

Primjer 3.49. Bacamo novčić na kojem je vjerojatnost 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,

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

□

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. Linearnost i monotonost očekivanja 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 3.22 – 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).]

$\square$

[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

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

Ako je $(X_1, \ldots, X_d)$ 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 – dakle, $(X_1, ..., X_d)$ je diskretan slučajni element u $\mathbb{R}^d$. [Sada distribuciju diskretnog slučajnog vektora definiramo analogno kao u jednodimenzionalnom (diskretnom) slučaju, te vrijede ista osnovna svojstva, koja navodimo niže dolje.]

Skup D zajedno s vjerojatnostima[2]

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

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

[1] Za $\omega \in \Omega$, definiramo $(X_1, \ldots, X_d)(\omega) := (X_1(\omega), \ldots, X_d(\omega))$.
[2] Događaj $\{X_1 = x_1, \ldots, X_d = x_d\}$ je oznaka za događaj

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

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

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

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

[Donja svojstva se lako proširuju na slučaj $d \geqslant 3$.]

- 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,

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} p_{ij} = \sum_{(a_i,b_j)\in D} p_{ij} = \mathbb{P}((X_1, X_2) \in D) = 1\,.$$

  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[3], 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 $(i,j) \in D$.]
- Iz zajedničke razdiobe vektora (X_1, X_2) lako dobijemo tzv. **marginalne** razdiobe od X_1 i X_2:

$$\begin{aligned}\mathbb{P}(X_1 = a_i) &= \mathbb{P}((X_1, X_2) \in \{a_i\} \times D_2) \overset{(4.1)}{=} \sum_{b_j \in D_2} p_{ij}\,,\ i \in I\\ \mathbb{P}(X_2 = b_j) &= \sum_{a_i \in D_1} p_{ij}\,,\ j \in J\,.\end{aligned}$$

  [Važno je napomenuti da marginalne razdiobe ne određuju zajedničku distribuciju.]
- Za svaki $b \in \mathbb{R}$ t.d. je $\mathbb{P}(X_2 = b) > 0$, **uvjetna razdioba** od X_1 uz dano $X_2 = b$ je uvjetna razdioba od X_1 uz dano $A = \{X_2 = b\} \in \mathcal{F}$ (vidi Poglavlje 3.6), tj.

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

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

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

Primjer 4.2. U kutiji su 1 crna, 2 bijele i 3 zelene kuglice. Slučajno izvučemo 2 kuglice te označimo s X ukupan broj crnih, a s Y ukupan broj bijelih kuglica koje smo izvukli. Odredite

[3] Dakle, (i) $p_{ij} \geqslant 0$, za sve i, j, te (ii) $\sum_{(i,j)\in D} p_{ij} = 1$.

(a) razdiobu vektora (X, Y);

(b) marginalne razdiobe;

(c) razdiobu od Y uvjetno na $X = 1$;

(d) $\mathbb{P}(X \geqslant Y)$;

(e) razdiobu slučajne varijable $X + Y$ [=ukupan broj crnih i bijelih kuglica].

Rješenje. (a) Sl. varijabla X (ond. Y) poprima vrijednosti u skupu $\{0, 1\}$ (odn. $\{0, 1, 2\}$). Vrijedi

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

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

i tako dalje. U dvodimenzionalnom slučaju, distribuciju možemo zapisati tablično

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

(b) Iz tablice u (a) dijelu odmah slijedi

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

(c) Normalizacijom drugog retka tablice dobivamo da je uvjetna razdioba

$$Y \mid X = 1 \sim \begin{pmatrix} 0 & 1 & 2 \\ \frac{3}{5} & \frac{2}{5} & 0 \end{pmatrix}$$

[To da se uvjetno na ishod varijable X, razdioba od Y mijenja, povlači da su X i Y *zavisne*, vidi iduće potpoglavlje.]

(d) Ako primijenimo (4.1) uz $C = \{(a, b) : a \geqslant b\}$, dobivamo

$$\mathbb{P}(X \geqslant Y) = \mathbb{P}(X = 0, Y = 0) + \mathbb{P}(X = 1, Y = 0) + \mathbb{P}(X = 1, Y = 1) = \frac{1}{15} \cdot (3 + 3 + 2) = \frac{8}{15}.$$

(e) Budući da je $X + Y = g(X, Y)$, lako se vidi da je $\mathbb{P}(X + Y \in \{0, 1, 2\}) = 1$ [nacrtati tablicu], a iz (4.2) slijedi

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

$$\mathbb{P}(X + Y = 1) = \mathbb{P}(X = 1, Y = 0) + \mathbb{P}(X = 0, Y = 1) = \frac{3}{15} + \frac{6}{15} = \frac{9}{15},$$

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

□

4.2. Nezavisnost

DEFINICIJA 4.3. Neka su $X_1, \dots, X_d$ diskretne slučajne varijable definirane na $(\Omega, \mathcal{F}, \mathbb{P})$ te neka je $D_i := D_{X_i}$, $i = 1, \dots, d$. Kažemo da su $X_1, \dots, X_d$ **nezavisne** ako vrijedi

$$\mathbb{P}(X_1 = x_1, \dots, X_d = x_d) = \mathbb{P}(X_1 = x_1) \cdots \mathbb{P}(X_d = x_d)\,, \tag{4.3}$$

za sve $x_1, \dots, x_d$ takve da je $x_i \in D_i$, $i = 1, \dots, d$.

[NAPOMENA.

- Uočite razliku u odnosu na definiciju nezavisnosti događaja – ovdje ne gledamo sve moguće podskupove od $\{1, \dots, d\}$, ali gledamo sve moguće $x_1, \dots, x_d$.
- Iz (4.3) odmah slijedi da kada su varijable $X_1, \dots, X_d$ nezavisne, njihove marginalne razdiobe jedinstveno određuju distribuciju vektora $(X_1, \dots, X_d)$.
- 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 4.4. Varijable X i Y iz Primjera 4.2 *nisu* nezavisne [tj. zavisne su] jer npr. vrijedi

$$\mathbb{P}(X = 1, Y = 2) = 0 \neq \mathbb{P}(X = 1)\mathbb{P}(Y = 2) > 0\,.$$

Alternativno, to slijedi iz

$$0 < \mathbb{P}(Y = 2) \neq \mathbb{P}(Y = 2 \mid X = 1) = 0\,.$$

PRIMJER 4.5. Bacamo simetričnu kocku n puta te neka je X_i rezultat na i-toj kocki. Tada za sve $x_1, \dots, x_n \in \{1, \dots, 6\}$ vrijedi

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \frac{1}{6^n} = \left(\frac{1}{6}\right)^n = \prod_{i=1}^{n} \mathbb{P}(X_i = x_i)\,,$$

tj. $X_1, \dots, X_n$ su nezavisne [kao što bismo i htjeli da budu].

[U nastavku u ovom ili sličnim slučajevima, nezavisnost ne provjeravamo, tj. pretpostavljamo da "očito" vrijedi.]

PRIMJER 4.6 (**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 $q := 1-p$ ($p \in (0,1)$). Neka je N_P ukupan broj plavih, a N_C ukupan broj crvenih kuglica. Odredite zajedničku i marginalne razdiobe od N_P i N_C. Jesu li one nezavisne?

[Intuicija o nezavisnosti?]

Rješenje. Za svaki $k, l \in \mathbb{N}_0$ imamo

$$\begin{aligned}\mathbb{P}(N_P = k, N_C = l) &= [N_P + N_C = N] = \mathbb{P}(N_P = k, N = k + l)\\ &= \mathbb{P}(N_P = k \mid N = k + l)\mathbb{P}(N = k + l)\\ &= [(N_P \mid N = k + l) \sim \mathrm{B}(k + l, p)]\\ &= \binom{k+l}{k} p^k q^l \cdot \frac{\lambda^{k+l}}{(k+l)!} e^{-\lambda} = \frac{(\lambda p)^k}{k!} e^{-\lambda p} \cdot \frac{(\lambda q)^l}{l!} e^{-\lambda q}\\ &= \mathbb{P}(\mathrm{P}(\lambda p) = k)\mathbb{P}(\mathrm{P}(\lambda q) = l)\,.\end{aligned}$$

Sada odmah slijedi

(i) $N_P \sim \mathrm{P}(\lambda p)$, $N_C \sim \mathrm{P}(\lambda q)$, te

(ii) N_P i N_C su **nezavisne**.

[Zaista, za sve $k \in \mathbb{N}_0$,

$$\mathbb{P}(N_P = k) = \sum_{l=0}^{\infty} \mathbb{P}(N_P = k, N_C = l) = \mathbb{P}(\mathrm{P}(\lambda p) = k) \underbrace{\sum_{l=0}^{\infty} \mathbb{P}(\mathrm{P}(\lambda q) = l)}_{=1} = \mathbb{P}(\mathrm{P}(\lambda p) = k)\,,$$

pa uvrštavajući natrag dobivamo

$$\mathbb{P}(N_P = k, N_C = l) = \mathbb{P}(N_P = k)\mathbb{P}(N_C = l)\,, \quad \forall k, l \in \mathbb{N}_0\,. \ ]$$

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.

Nezavisnost od N_P i N_C je pomalo iznenađujuća budući da je $N_P + N_C = N$ – ovdje je ključno da je N slučajan te da ima baš Poissonovu razdiobu. □

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**[4]). [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 $\mathrm{B}(p_0)$.

[4] engl. *maximum likelihood*

- 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** parametra p ukoliko su podaci $x_1, \ldots, x_n$; f_p je dakle oznaka za gustoću vektora $(X_1, \ldots, X_n)$ ukoliko su oni njd s distribucijom $\mathrm{B}(p)$.[5]
- 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.]

$\square$

[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 varijabli 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 $B_i := \{x_i\}$, $i = 1, \ldots, d$.

[5] Intuitivno, očekujemo da će vjerodostojnost biti velika za p-ove koji su blizu p_0.

"$\Rightarrow$" Vrijedi

$$\begin{aligned}
\mathbb{P}(X_1 \in B_1, \dots, X_d \in B_d) &= \mathbb{P}((X_1, \dots, X_d) \in B_1 \times \cdots \times B_d) \\
&\overset{(4.1)}{=} \sum_{x_1 \in B_1 \cap D_1} \cdots \sum_{x_d \in B_d \cap D_d} \mathbb{P}(X_1 = x_1, \dots, X_d = x_d) \\
&= [\text{nezavisnost, tj. } (4.3)] = \sum_{x_1 \in B_1 \cap D_1} \cdots \sum_{x_d \in B_d \cap D_d} \mathbb{P}(X_1 = x_1) \cdots \mathbb{P}(X_d = x_d) \\
&= \left(\textstyle\sum_{x_1 \in B_1 \cap D_1} \mathbb{P}(X_1 = x_1)\right) \cdots \left(\textstyle\sum_{x_d \in B_d \cap D_d} \mathbb{P}(X_d = x_d)\right) \\
&\overset{(4.1)}{=} \mathbb{P}(X_1 \in B_1) \cdots \mathbb{P}(X_d \in B_d) .
\end{aligned}$$

$\square$

Napomena 4.10. Neka su $X_1, \dots, X_d$ 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 (**Minimum nezavisnih geometrijskih varijabli**). 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 $X_1, X_2, \ldots$, kažemo

(i) da su nezavisne, ako su $X_1, \ldots, X_n$ nezavisne za svaki $n \in \mathbb{N}$;

(ii) da su nezavisne od drugog niza $Y_1, Y_2, \ldots$, ako su vektori $(X_1, \ldots, X_n)$ i $(Y_1, \ldots, Y_m)$ nezavisni za sve $n, m \in \mathbb{N}$ (u smislu (4.5)).

4.2.1. Zbroj nezavisnih binomnih i Poissonovih varijabli.

Propozicija 4.13. *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,

$$X + Y = \sum_{i=1}^{n+m} X_i \sim \mathrm{B}(n + m, p)\,.$$

□

Propozicija 4.14. *Ako su* $X \sim \mathrm{P}(\lambda)$ *i* $Y \sim \mathrm{P}(\mu)$ *nezavisne, vrijedi* $X + Y \sim \mathrm{P}(\lambda + \mu)$.

Dokaz. *1. rješenje*

Varijabla $X+Y$ očito poprima vrijednosti u $\mathbb{N}_0$, te za sve $k \in \mathbb{N}_0$, vrijedi

$$\begin{aligned}
\mathbb{P}(X+Y=k) &= \sum_{i=0}^{\infty} \mathbb{P}(X=i, X+Y=k) = \sum_{i=0}^{\infty} \mathbb{P}(X=i, Y=k-i) \\
&= \sum_{i=0}^{k} \mathbb{P}(X=i, Y=k-i) = [\text{nezavisnost}] = \sum_{i=0}^{k} \mathbb{P}(X=i)\mathbb{P}(Y=k-i) \\
&= \sum_{i=0}^{k} \frac{\lambda^k}{k!} e^{-\lambda} \cdot \frac{\mu^{k-i}}{(k-i)!} e^{\mu} \cdot \frac{k!}{k!} = e^{-(\lambda+\mu)} \cdot \frac{1}{k!} \underbrace{\sum_{i=0}^{k} \binom{k}{i} \lambda^i \mu^{k-i}}_{=(\lambda+\mu)^k} \\
&= e^{-(\lambda+\mu)} \cdot \frac{(\lambda+\mu)^k}{k!} .
\end{aligned}$$

2. rješenje (zakon rijetkih događaja)

Pogledati sami [samo nacrtati skicu].

[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, budući da je $\lim_{n\to\infty} \lfloor n\lambda \rfloor \cdot \frac{1}{n} = \lambda$, imamo

$$\mathbb{P}(X_n = k) \longrightarrow \mathbb{P}(\mathrm{P}(\lambda) = k) = \mathbb{P}(X=k) , \quad \forall k \in \mathbb{N}_0 ,$$

te analogno $\mathbb{P}(Y_n = k) \longrightarrow \mathbb{P}(Y=k)$. Nadalje, po prethodnom primjeru,

$$X_n + Y_n \sim \mathrm{B}(\lfloor n\lambda \rfloor + \lfloor n\mu \rfloor, \frac{1}{n}) ,$$

pa $\lim_{n\to\infty} \mathbb{P}(X_n + Y_n = k) = \mathbb{P}(\mathrm{P}(\lambda+\mu) = k)$, a s druge strane (čim je $n \geqslant k$)

$$\begin{aligned}
\mathbb{P}(X_n + Y_n = k) &= \sum_{l=0}^{k} \mathbb{P}(X_n = l, Y_n = k-l) = [\text{nezavisnost}] = \sum_{l=0}^{k} \mathbb{P}(X_n = l)\mathbb{P}(Y_n = k-l) \\
&\longrightarrow \sum_{l=0}^{k} \mathbb{P}(X=l)\mathbb{P}(Y=k-l) = [\text{nezavisnost}] = \mathbb{P}(X+Y=k) . \ ]
\end{aligned}$$

3. rješenje

Tvrdnja zapravo slijedi iz Primjera 4.6 o stanjivanju Poissonove slučajne varijable. Zaista, ako uzmemo $N \sim \mathrm{P}(\lambda+\mu)$ te $p := \frac{\lambda}{\lambda+\mu}$, $q = \frac{\mu}{\lambda+\mu}$, tada je $X := N_P \sim \mathrm{P}(\lambda)$ i $Y := N_C \sim \mathrm{P}(\mu)$ te su one nezavisne. Dakle, konstruirali smo X i Y kao u pretpostavci propozicije, a za njih vrijedi

$$X + Y = N \sim \mathrm{P}(\lambda+\mu) .$$

□

4.2.2. Veza između binomne i hipergeometrijske razdiobe.

PROPOZICIJA 4.15. *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. 1. rješenje. Po definiciji uvjetne vjerojatnosti

$$\mathbb{P}(X=i \mid X+Y=k) = \frac{\mathbb{P}(X=i, Y=k-i)}{\mathbb{P}(X+Y=k)} \overset{\text{nez.}}{=} \frac{\mathbb{P}(X=i)\mathbb{P}(Y=k-i)}{\mathbb{P}(X+Y=k)}$$
$$= [X+Y \sim \mathrm{B}(n+m,p)] = \frac{\binom{n}{i}p^i(1-p)^{n-i} \cdot \binom{m}{k-i}p^{k-i}(1-p)^{m-(k-i)}}{\binom{n+m}{k}p^k(1-p)^{n+m-k}} = \frac{\binom{n}{i} \cdot \binom{m}{k-i}}{\binom{n+m}{k}},$$

pri čemu to vrijedi za sve $i \in \{0,1,\dots,k\}$ takve da je $i \leqslant n$ i $k-i \leqslant m$.

2. rješenje. [Zašto se javlja HG razdioba?] Pretpostavimo da u kutiji imamo n bijelih i m crnih kuglica, te da svaku od $n+m$ kuglica, nezavisno od drugih kuglica, s vjerojatnošću p izvučemo iz kutije, tj. s vjerojatnošću $1-p$ ostavimo u kutiji. Tada $X :=$ broj izvučenih bijelih i $Y :=$ broj izvučenih crnih kuglica očito imaju tražene razdiobe te su nezavisne. Ako uvjetujemo na

$$\text{ukupan broj izvučenih kuglica} = X+Y = k\,,$$

taj skup od k izvučenih kuglica je s **jednakom vjerojatnosti** bilo koji k-člani podskup od $n+m$ kuglica. Specijalno, slučajna varijabla X je tada ukupan broj bijelih kuglica u tom skupu od k izvučenih kuglica, pa ima točno $\mathrm{HG}(n,m,k)$ razdiobu.

[Tvrdnja da su uvjetno na $X+Y=k$, svi k-člani podskupovi od $n+m$ kuglica jednako vjerojatni je intuitivno jasna. Formalni dokaz je sljedeći: neka su $X_1,\dots,X_{n+m}$ njd $\mathrm{B}(p)$ slučajne varijable, te $X := \sum_{i=1}^n X_i$, $Y := \sum_{i=n+1}^{n+m} X_i$ – X_i označava je li i-ta kuglica izvučena, pri čemu su prvih n kuglica bijele. Za fiksni $k \in \{0,\dots,n+m\}$, neka je $S_k := \{(a_1,\dots,a_{n+m}) \in \{0,1\}^{n+m} : \sum_{i=1}^{n+m} a_i = k\}$, tj. S_k sadrži sve moguće ishode s točno k izvučenih kuglica. Zbog njd pretpostavke za sve $a \in S_k$ vrijedi

$$\mathbb{P}((X_1,\dots,X_{n+m}) = a) = p^k(1-p)^{n+m-k}\,,$$

pri čemu dakle gornja vjerojatnost ne ovisi o konkretnoj permutaciji 0-la i 1-ca! Sada odmah slijedi da je za sve $a, a' \in S_k$,

$$\mathbb{P}((X_1,\dots,X_{n+m}) = a \mid X+Y=k) = \mathbb{P}((X_1,\dots,X_{n+m}) = a' \mid X+Y=k) = \frac{1}{|S_k|} = \frac{1}{\binom{n+m}{k}}\,.]$$

□

PRIMJER 4.16 (**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 korist $p_1 > p_2$?

Uočimo, kada bi vrijedilo $p_1 = p_2 = p$ (tzv. *nulta hipoteza*),

$$\mathbb{P}(X \geqslant 7 \mid X+Y=10) = [\text{Prop. } 4.15] = \mathbb{P}(\mathrm{HG}(30,25,10) \geqslant 7) \approx 0.233\,.$$

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?

4.3. Linearnost i monotonost očekivanja za diskretne varijable

[Sljedeći teorem je generalizacija Teorema 3.22 (slučaj $d = 1$), te se dokazuje potpuno analogno. Koristeći donji rezultat napokon možemo formalno pokazati da je za diskretne slučajne varijable na općenitom vjerojatnosnom prostoru, očekivanje linearno i monotono.]

TEOREM 4.17. *Neka su $X_1, \ldots, X_d$ diskretne slučajne variable na $(\Omega, \mathcal{F}, \mathbb{P})$, te $D_i := D_{X_i}$, $i = 1, \ldots, d$, a $g : \mathbb{R}^d \to \mathbb{R}$ funkcija. Tada vrijedi*

$$\mathbb{E}[g(X_1, \ldots, X_d)] = \textstyle\sum_{x_i \in D_i, \forall i} g(x_1, \ldots, x_d)\mathbb{P}(X_1 = x_1, \ldots, X_d = x_d)\,, \tag{4.6}$$

ukoliko (a) $g(X_1, \ldots, X_d) \geqslant 0$, ili (b) red na desnoj strani od (4.6) apsolutno konvergentan.

TEOREM 4.18. *Ako su $X_1, \ldots, X_n$ 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 $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$, tj. $\alpha_1, \ldots, \alpha_n \geqslant 0$ 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 $\alpha_1 X_1 + \alpha_2 X_2 = g(X_1, X_2) \geqslant 0$, (4.6) nam daje

$$\begin{aligned}
\mathbb{E}[\alpha_1 X_1 + \alpha_2 X_2] &= \sum_{i \in I} \sum_{j \in J} (\alpha_1 a_i + \alpha_2 b_j)\mathbb{P}(X_1 = a_i, X_2 = b_j) \\
&= [\text{Fubini}] = \alpha_1 \sum_{i \in I} a_i \underbrace{\sum_{j \in J} \mathbb{P}(X_1 = a_i, X_2 = b_j)}_{=\mathbb{P}(X_1 = a_i)} + \alpha_2 \sum_{j \in J} b_j \underbrace{\sum_{i \in I} \mathbb{P}(X_1 = a_i, X_2 = b_j)}_{=\mathbb{P}(X_2 = b_j)} \\
&= \alpha_1 \sum_{i \in I} a_i \mathbb{P}(X_1 = a_i) + \alpha_2 \sum_{j \in J} b_j \mathbb{P}(X_2 = b_j) = \alpha_1 \mathbb{E}[X_1] + \alpha_2 \mathbb{E}[X_2] \in [0, \infty]\,.
\end{aligned}$$

(b) Ako su $X_1, X_2 \in \mathbb{R}$ te $\alpha_1, \alpha_2 \in \mathbb{R}$, monotonost očekivanja (vidi Teorem 4.21 dolje) povlači

$$\begin{aligned}
\mathbb{E}\Big[|\alpha_1 X_1 + \alpha_2 X_2|\Big] &\leqslant \mathbb{E}\Big[|\alpha_1||X_1| + |\alpha_2||X_2|\Big] \\
&= [\text{(a) dio}] = |\alpha_1| \cdot \mathbb{E}\Big[|X_1|\Big] + |\alpha_2| \cdot \Big[|X_2|\Big] < \infty\,.
\end{aligned}$$

Dakle, postoji $\mathbb{E}[\alpha_1 X_1 + \alpha_2 X_2]$ te se (4.7) dokazuje analogno kao u (a) dijelu. $\square$

NAPOMENA 4.19. Jednakost (4.7) vrijedi **bez obzira** na zavisnost među X_i-evima.

PRIMJER 4.20 (**"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

$$\mathbb{E}[T] \approx n\log(n) + 0.577n + 0.5\,.$$

$\square$

TEOREM 4.21. *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 $p_{ij} > 0$.] $\square$

4.4. Varijanca sume nezavisnih varijabli

TEOREM 4.22. *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*

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]\,. \tag{4.8}$$

DOKAZ. Budući da je $XY = g(X, Y)$, dokaz je jednostavna primjena tvrdnje (4.2), pa ga ostavljamo studentima da pogledaju sami.

[Neka je $D_X = \{a_i : i \in I\}$, $D_Y = \{b_j : j \in J\}$.

(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 (4.8) pokaže analogno kao u (a) dijelu.]

□

PROPOZICIJA 4.23. *Ako su $X_1, \dots, X_n$ u parovima nezavisne slučajne varijable [i t.d. je* $\mathrm{Var}(X_i) < \infty$ *za sve $i = 1, \dots, n$], vrijedi*

$$\mathrm{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathrm{Var}(X_i)\,. \tag{4.9}$$

DOKAZ. Imamo

$$\begin{aligned}\mathrm{Var}\left(\textstyle\sum_{i=1}^n X_i\right) &= \mathbb{E}\left[(\textstyle\sum_{i=1}^n X_i - \mathbb{E}[\sum_{i=1}^n X_i])^2\right] = [\text{lin. oček.}] = \mathbb{E}\left[(\textstyle\sum_{i=1}^n (X_i - \mathbb{E}[X_i]))^2\right] \\ &= [\text{lin. oček.}] = \sum_{i=1}^n \mathbb{E}[(X_i - \mathbb{E}[X_i])^2] + 2\sum_{1\leqslant i<j\leqslant n} \underbrace{\mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])]}_{=[(4.8)]=\mathbb{E}[(X_i-\mathbb{E}[X_i])]\mathbb{E}[(X_j-\mathbb{E}[X_j])]=0} \\ &= \sum_{i=1}^n \mathrm{Var}(X_i)\end{aligned}$$

□

PRIMJER 4.24. 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:[6]

- $T_1, \ldots, T_n$ su nezavisne;
- $T_i \sim \mathrm{G}(p)$, za sve $i = 1, \ldots, n$.

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 4.25 (**Varijanca** $\mathrm{B}(n,p)$ **razdiobe**). Neka su $X_1, \ldots, X_n$ 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 4.26 (**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 $\sigma^2 := \mathrm{Var}(X) < \infty$.[7] Ako su $X_1, \ldots, X_n$ njd slučajne varijable s istom distribucijom kao i X, klasični *procjenitelj*

$$\overline{X}_n := \frac{1}{n} \sum_{i=1}^{n} X_i \,,$$

za θ ima sljedeća svojstva:

- $\mathbb{E}[\overline{X}_n] = [\text{lin.}] = \frac{1}{n} n\mathbb{E}[X_1] = \mathbb{E}[X_1] = \theta$ [dakle, u prosjeku će $\overline{X}_n$ biti približno jednak θ – kažemo da je *nepristran*];
- $\mathrm{Var}(\overline{X}_n) = [\text{nez.}] = \frac{1}{n^2} n\mathrm{Var}(X_1) = \frac{\sigma^2}{n}$ [dakle, odstupanje od θ je manje što je varijanca σ^2 manja, te što je n veći].

[Vrlo važno je napomenuti da gornji postupak možemo provesti za *bilo koju* slučajnu varijablu X za koju vrijedi $\mathbb{E}[X] = \theta$ (te je znamo simulirati). Ipak, cilj je naći varijablu X sa što manjom varijancom $\mathrm{Var}(X) = \sigma^2$. Jednu takvu metodu *smanjenja varijance* ćete vidjeti na vježbama.]

4.5. Slučajne sume

Primjer 4.27 (**Očekivanje i varijanca sume njd varijabli**). Neka su $X_1, X_2, \ldots$ njd slučajne varijable s očekivanjem $\mu := \mathbb{E}[X_1]$ i $\sigma^2 := \mathrm{Var}(X_1) < \infty$, te neka je $S_n := X_1 + \cdots + X_n$,

[6] Probajte ovo formalno pokazati.

[7] Slučajna varijabla X može biti na primjer funkcija velikog broja međusobno zavisnih slučajnih varijabli tako da računanje njenog očekivanja po definiciji nije jednostavan problem.

$n \in \mathbb{N}$, uz $S_0 := 0$.[8] Tada imamo

$$\mathbb{E}[S_n] = [\text{linearnost oč.}] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n\mu\,,$$
$$\mathrm{Var}(S_n) = [\text{nezavisnost}] = \sum_{i=1}^{n} \mathrm{Var}[X_i] = n\sigma^2\,.$$

[Tvrdnje očito vrijede i kada je $n = 0$, a uočimo da prva tvrdnja vrijedi i kada su X_i-evi samo nenegativni, tj. ako je u tom slučaju $\mu = \infty$ vrijedi i $\mathbb{E}[S_n] = \infty = n\mu$.]

PRIMJER 4.28 (**Slučajne sume**). Neka je S_n kao u prethodnom primjeru, te neka je $N \in \mathbb{N}_0$ slučajna varijabla nezavisna od niza X_i-eva. Odredite $\mathbb{E}[S_N]$ i $\mathrm{Var}(S_N)$. [Intuicija?]

RJEŠENJE. Pretpostavimo prvo da je X_i nenegativna slučajna varijabla, pa dakle $\mathbb{E}[S_N]$ uvijek postoji. Za sve $n \in \mathbb{N}_0$, vrijedi

$$\mathbb{E}[S_N \mid N = n] = \mathbb{E}[S_n \mid N = n] = [S_n \text{ i } N \text{ nez.}] = \mathbb{E}[S_n] = n\mu\,,$$

pa je

$$\mathbb{E}[S_N] = \sum_{n=0}^{\infty} \mathbb{E}[S_N \mid N = n]\mathbb{P}(N = n) = \mu \sum_{n=0}^{\infty} n\mathbb{P}(N = n) = \mathbb{E}[N]\mu\,, \tag{4.10}$$

što intuitivno ima smisla te se naziva **Waldova jednakost** [jednakost vrijedi i ako je $\mu = \infty$ i/ili $\mathbb{E}[N] = \infty$].

Ako je $X_i \in \mathbb{R}$ te $\mathbb{E}\Big[|X_i|\Big] < \infty$, iz prethodnog imamo

$$\mathbb{E}\Big[|S_N|\Big] \leqslant \mathbb{E}\Big[\textstyle\sum_{i=1}^{N} |X_i|\Big] = \mathbb{E}[N]\mathbb{E}\Big[|X_1|\Big]\,,$$

pa ako je $\mathbb{E}[N] < \infty$, postoji $\mathbb{E}[S_N]$ te se kao gore pokaže da vrijedi (4.10). U tom slučaju postoji i $\mathrm{Var}(S_N)$, te za vježbu pokažite da vrijedi

$$\mathrm{Var}(S_N) = \mathbb{E}[N]\sigma^2 + \mu^2\mathrm{Var}(N)\,.$$

[Intuicija? Zašto nemamo samo prvi član?] Pogledajmo na što se 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 $X_i \equiv c \in \mathbb{R}$, desna strana je $0 + c^2\mathrm{Var}(N) = \mathrm{Var}(cN) = \mathrm{Var}(S_N)$.

[Dakle, u obzir moramo uzeti i varijabilnost koja dolazi od toga što je broj elemenata N slučajan.] □

PRIMJER 4.29 (**Stanjivanje Poissonove slučajne varijable, Primjer 4.6 nastavak**). [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 $q := 1 - p$ ($p \in (0, 1)$). Neka je N_P ukupan broj plavih kuglica.] Odredite $\mathbb{E}[N_P]$ i $\mathrm{Var}(N_P)$.

[8] Niz slučajnih varijabli $S_0, S_1, \dots$ zovemo *slučajna šetnja* i osnovni je primjer tzv. *slučajnog procesa*.

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

$$N_P = \sum_{i=1}^{N} X_i = S_N\,.$$

Iz prethodnog primjera slijedi

$$\mathbb{E}[N_P] = \mathbb{E}[N]\mathbb{E}[X_1] = \lambda p\,,$$
$$\mathrm{Var}(N_P) = \lambda p(1-p) + p^2\lambda = \lambda p\,.$$

[U Primjeru 4.6 smo pokazali i puno više, tj. da N_P ima točno $\mathrm{P}(\lambda p)$ razdiobu, ali uočimo da Primjer 4.29 možemo jednako riješiti i u slučaju kada N nema nužno Poissonovu razdiobu.] □

4.6. Kovarijanca i korelacija

[Pitanje je kako mjeriti zavisnost između dvije slučajne varijable.]

Definicija 4.30. 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[9]

$$\mathrm{Cov}(X, Y) := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]\,. \tag{4.11}$$

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 i linearnosti očekivanja, tako da ga ostavljamo za vježbu.

Propozicija 4.31 (**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 4.32. 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

[9] Kasnije ćemo pokazati da uvjeti $\mathbb{E}[X^2], \mathbb{E}[Y^2] < \infty$ osiguravaju da je $\mathrm{Cov}(X, Y)$ dobro definirana.

$X\backslash Y$	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 $Y = 2 - X$.

Uočimo, ako su X i Y nezavisne (te vrijedi $\mathbb{E}[X^2], \mathbb{E}[Y^2] < \infty$), nužno je

$$\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \overset{(4.8)}{=} 0 .$$

DEFINICIJA 4.33. Ako vrijedi $\mathrm{Cov}(X, Y) = 0$, kažemo da su X i Y **nekorelirane**.

[Dakle, pokazali smo da nezavisnost povlači nekoreliranost.]

PRIMJER 4.34 (**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 4.35 (**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 slijedi odmah iz dokaza Propozicije 4.23 i definicije kovarijance.

□

[Uočimo da je za $\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$ dovoljno samo da X_i i X_j nekorelirane za sve $i \neq j$.]

PRIMJER 4.36. Neka je $n \in \mathbb{N}$,

$$A_i := \{i \text{ je fiksna točka slučajne permutacije skupa } \{1, \ldots, n\}\}, \quad i = 1, \ldots, n .$$

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 $\mathbb{1}_{A_1}, \dots, \mathbb{1}_{A_n}$ su zavisni, tako da za $\mathrm{Var}(X)$ ne možemo koristiti (4.9).] 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?] □

4.6.1. 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 4.37. Ako je $\mathbb{E}\big[|X|\big] < \infty$, vrijedi

$$\big|\mathbb{E}[X]\big| \leqslant \mathbb{E}\big[|X|\big]\,. \tag{4.12}$$

Zaista, to slijedi iz monotonosti očekivanja te činjenice da su $|X| - X$ te $|X| + X$ nenegativne slučajne varijable.

Teorem 4.38. *Neka su X i Y 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. (a) Za sve $a, b \in \mathbb{R}$, vrijedi $|ab| \leqslant a^2 + b^2$, pa monotonost i linearnost očekivanja povlače

$$\mathbb{E}\big[|XY|\big] \leqslant \mathbb{E}[X^2] + \mathbb{E}[Y^2] < \infty\,.$$

(b) 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 ne postoje dvije različite realne nultočke, 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 u slučaju $\mathbb{E}[Y^2] > 0$ imamo jednakost akko je $D = 0$, tj. akko postoji $\lambda \in \mathbb{R}$ takav da je $\mathbb{E}[W_\lambda^2] = g(\lambda) = 0$. Budući da je W_λ^2 nenegativna, zadnje je ekvivalentno postojanjem $\lambda \in \mathbb{R}$ takvim da je

$$1 = \mathbb{P}((X + \lambda Y)^2 = 0) = \mathbb{P}(X + \lambda Y = 0) = \mathbb{P}(X = -\lambda Y)\,.$$

Uočimo da je u zadnjem slučaju nužno $\lambda \neq 0$ jer bi inače bilo $\mathbb{P}(X = 0) = 1$, ali to ne može vrijediti jer smo pretpostavili da je $\mathbb{E}[X^2] > 0$. $\square$

Napomena 4.39. 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 4.38(a) na varijable
$$\overline{X} := X - \mathbb{E}[X]\,,\ \ \overline{Y} := Y - \mathbb{E}[Y]\,,$$
slijedi da je i $\mathbb{E}\Big[|\overline{X} \cdot \overline{Y}|\Big] < \infty$, tj. $\mathrm{Cov}(X, Y) = \mathbb{E}[\overline{X} \cdot \overline{Y}]$ je dobro definirana.
- Slično, iz CS nejednakosti za $\overline{X}$ i $\overline{Y}$ dobivamo da je
$$|\mathrm{Cov}(X, Y)| \leqslant \sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}\,. \tag{4.13}$$

Definicija 4.40. 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 4.41 (**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 4.38 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 4.42. 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 $Y = 2 - X$.

[Napomena. Za proizvoljne (dakle, ne nužno diskretne) slučajne varijable, kovarijanca i korelacija se definiraju na potpuno isti način (nakon što se definira pojam očekivanja), te svi rezultati iz ovog poglavlja vrijede i u tom općenitom slučaju – većina dokaza je zapravo i ista jer se koristi samo linearnost očekivanja.]

4.7. Multinomna razdioba

[Multinomna razdioba je jedna od najvažnih diskretnih višedimenzionalnih razdioba. Predstavlja generalizaciju binomne razdiobe – umjesto da brojimo uspjehe u nizu od n pokusa u kojem postoje samo dva ishoda, multinomnu razdiobu dobijemo kada broj ishoda može biti 3 ili više. Multinomna razdioba prirodno se javlja u statistici/strojnom učenju u tzv. *problemu klasifikacije.*]

Imamo pokus koji ima ukupno k mogućih različitih ishoda $1, \ldots, k$, pri čemu je vjerojatnost svakog ishoda $p_i \in [0,1]$, $i = 1, ..., k$ ($\sum_{i=1}^{n} p_i = 1$). Ako s X_i označimo broj pojavljivanja ishoda i u n nezavisnih ponavljanja pokusa, slučajni vektor $(X_1, ..., X_k)$ ima tzv. **multinomnu distribuciju** s parametrima $(n, p_1, ..., p_k)$, oznaka $\mathrm{Mult}(n, p_1, ..., p_k)$.

PROPOZICIJA 4.43. *Neka je* $(X_1, ..., X_k) \sim \mathrm{Mult}(n, p_1, ..., p_k)$.

(a) (**diskretna funkcija gustoće**) *Vrijedi*

$$\mathbb{P}(X_1 = n_1, \ldots, X_k = n_k) = \frac{n!}{n_1! n_2! \cdots n_k!} \cdot p_1^{n_1} \cdots p_k^{n_k}$$

za sve $n_1, \ldots, n_k \geqslant 0$ *takve da je* $n_1 + \cdots + n_k = n$.

(b) (**marginalne razdiobe**) $X_i \sim \mathrm{B}(n, p_i)$, $\forall i$.

(c) (**spajanje ishoda**) $(X_1 + X_2, X_3, \ldots, X_k) \sim \mathrm{Mult}(n, p_1 + p_2, p_3, \ldots, p_k)$.

(d) (**uvjetne razdiobe**) *Vrijedi*

$$(X_2, \ldots, X_k) \mid X_1 = n_1 \sim \mathrm{Mult}(n - n_1, p_2', p_3', \ldots, p_k')$$

gdje je $p_i' = \frac{p_i}{p_2 + \cdots + p_k}$, $i = 2, \ldots, k$.

(e) (**kovarijanca**) *Za sve* $\imath \neq j$ *vrijedi*

$$\mathrm{Cov}(X_i, X_j) = -np_i p_j \,, \quad \rho(X_i, X_j) = -\sqrt{\frac{p_i p_j}{(1 - p_i)(1 - p_j)}} \,.$$

Specijalno, $X_1, \ldots, X_k$ *su zavisne.*[10] *[Je li negativna korelacija intuitivno jasna? Kada će korelacija biti blizu* -1*?]*

DOKAZ. (a)-(c) sami.

[(a) Ako fiksiramo n_1 mjesta na kojima smo dobili ishod 1, n_2 mjesta za ishod 2 itd., svaki takav raspored ima vjerojatnost $p_1^{n_1} \cdots p_k^{n_k}$. Tvrdnja sada slijedi budući da takvih rasporeda ima točno $\frac{n!}{n_1! n_2! \cdots n_k!}$.

(b) i (c) Slijedi direktno iz konstrukcije multinomnog vektora, ili ako netko baš želi, može se dokazati sumirajući vjerojatnosti iz (a).]

(d) Ako znamo da je točno n_1 ishoda bilo 1, na preostalih $n - n_1$ pokusa imamo samo ishode $2, \ldots, k$, ti su pokusi nezavisni te je vjerojatnost da u jednom pokusu dobijemo ishod i jednaka

$$\mathbb{P}(\text{ishod je } i \mid \text{ishod nije } 1) = \frac{\mathbb{P}(\text{ishod je } i)}{\mathbb{P}(\text{ishod nije } 1)} = \frac{p_i}{1 - p_1} = \frac{p_i}{p_2 + \cdots + p_k} \,, \quad i = 2, \ldots, k \,.$$

[Uz malo truda, gornji argument se može u potpunosti formalizirati.]

Alternativno, ako je $n_2 + \cdots + n_k = n - n_1$, imamo

$$\begin{aligned} \mathbb{P}(X_2 = n_2, \ldots, X_k = n_k \mid X_1 = n_1) &= \frac{\mathbb{P}(X_1 = n_1, \ldots, X_k = n_k)}{\mathbb{P}(X_1 = n_1)} = \frac{\frac{n!}{n_1! n_2! \cdots n_k!} \cdot p_1^{n_1} \cdots p_k^{n_k}}{\binom{n}{n_1} p_1^{n_1} (1 - p_1)^{n - n_1}} \\ &= \frac{(n - n_1)!}{n_2! \cdots n_k!} \cdot \left(\frac{p_2}{1 - p_1}\right)^{n_2} \cdots \left(\frac{p_k}{1 - p_1}\right)^{n_k} . \end{aligned}$$

[10] Uočimo da je zapravo $X_1 + X_2 + \cdots + X_k = n$.

(e) Neka je $A_{m,i} = \{\text{u } m\text{-tom pokusu dogodio se ishod } i\}$, $i = 1, \ldots, k$, $m = 1, \ldots, n$. Tada imamo

$$X_i = \sum_{m=1}^{n} 1_{A_{m,i}}\,,\ \forall i = 1, \ldots, k$$

pa je za $i \neq j$,

$$\begin{aligned}\mathrm{Cov}(X_i, X_j) &= \mathrm{Cov}\left(\sum_{m=1}^{n} 1_{A_{m,i}}, \sum_{l=1}^{n} 1_{A_{l,j}}\right) = \sum_{m=1}^{n}\sum_{l=1}^{n} \mathrm{Cov}(1_{A_{m,i}}, 1_{A_{l,j}}) \\ &= [\text{nezavisnost pokusa}] = \sum_{m=1}^{n} \mathrm{Cov}(1_{A_{m,i}}, 1_{A_{m,j}}) = n\mathrm{Cov}(1_{A_{1,i}}, 1_{A_{1,j}}) \\ &= n(\mathbb{P}(A_{1,i} \cap A_{1,j}) - \mathbb{P}(A_{1,i})\mathbb{P}(A_{1,j})) = [i \neq j] = n(0 - p_i p_j) = -np_i p_j\,.\end{aligned}$$

Izraz za $\rho(X_i, X_j)$ slijedi uvrštavanjem u definiciju uz $\mathrm{Var}(X_i) = np_i(1 - p_i)$ i $\mathrm{Var}(X_j) = np_j(1 - p_j)$. □

Primjer 4.44. Bacamo simetričnu kocku 100 puta, te s X označimo broj jedinica, a s Y broj šestica koje su pale. Odredite $\rho(X, Y)$.

Rješenje. Ako stavimo $Z := 100 - X - Y$ [tj. ukupan broj bacanja gdje smo dobili $2, \ldots, 5$], imamo

$$(X, Y, Z) \sim \mathrm{Mult}(100, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{4}{6})\,,$$

pa iz prethodne propozicije odmah slijedi

$$\rho(X, Y) = -\sqrt{\frac{\frac{1}{6} \cdot \frac{1}{6}}{\frac{5}{6} \cdot \frac{5}{6}}} = -\frac{1}{5}\,.$$

□

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 5.14 kasnije.

Općenito, za "skoro sve"[2] $B \subseteq \mathbb{R}$ vrijedi

$$\mathbb{P}(X \in B) = \int_{-\infty}^{\infty} f(t) \cdot \mathbb{1}_{\{t \in B\}} \, \mathrm{d}t =: \int_B f(t) \, \mathrm{d}t \, .$$

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 rastuća 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 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 rastuća 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 rastuća, 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 rastuća, 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

$$\bigcup_{n=1}^{\infty}\{X \leqslant x - \tfrac{1}{n}\} = \{X < x\}\,.$$

$\square$

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 $\mathbb{R}$.

[Ipak, postoje slučajne varijable čije su funkcije distribucije neprekidne, ali nisu oblika $F(x) = \int_{-\infty}^{x} f(t)\,dt$, niti za jednu funkciju f – takve slučajne varijable *nisu* (apsolutno) neprekidne u smislu Definicije 5.1. Funkcije koje su tog oblika zovu se upravo *apsolutno neprekidne* funkcije i odatle dolazi ime "**apsolutno** 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 5.15(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**: $X \sim \mathrm{Exp}(\lambda)$ akko vrijedi

$$\mathbb{P}(X > x) = 1 - F_X(x) = e^{-\lambda x}, \quad \forall x \geqslant 0.$$

Uočimo, ako je $X \sim \mathrm{Exp}(\lambda)$, 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 $\mathrm{Exp}(\lambda)$ razdiobu. □

[$\mathrm{Exp}(\lambda)$ 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 intenzitetom** $\lambda > 0$ je niz slučajnih varijabli $0 < T_1 < T_2 < \ldots$ [T_i predstavlja vrijeme i-tog događaja], takvih da za

$$N(\langle a, b]) := \sum_{i=1}^{\infty} \mathbb{1}_{\{T_i \in \langle a, b]\}} = \text{broj događaja u intervalu } \langle a, b], \quad a < b,$$

vrijedi

(a) $N(\langle a, b]) \sim \mathrm{P}(\lambda(b - a))$ za sve $a < b$, te

(b) $N(I_1), \ldots, N(I_n)$ su nezavisne za sve n i sve u parovima disjunktne intervale $I_1, \ldots, I_n \subseteq (0, \infty)$.

[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, intenzitet λ predstavlja očekivani broj događaja po jedinici vremena. Nadalje, vrijeme čekanja do prvog događaja T_1 zadovoljava

$$\mathbb{P}(T_1 > t) = \mathbb{P}(N(\langle 0,t]) = 0) \overset{(a)}{=} e^{-\lambda(t-0)} = e^{-\lambda t}\,, \quad \forall t > 0\,,$$

to jest T_1 ima točno Exp(λ) razdiobu [tako možemo razmišljati o Exp(λ) 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 Exp(λ) [što daje jednostavnu metodu za simuliranje Poissonovog procesa]. □

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

Poruka ovog potpoglavlja: ako je X neprekidna slučajna varijabla, te $g : \mathbb{R} \to \mathbb{R}$ funkcija,

- $g(X)$ ne mora nužno biti neprekidna;
- čak i kada je $g(X)$ neprekidna, i g bijekcija, tipično ne vrijedi

$$f_{g(X)}(g(x)) = f_X(x)\,.$$

[kao što je to bilo u diskretnom slučaju]

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 Unif$(0,1)$ slučajnu varijablu generirati proizvoljnu diskretnu slučajnu varijablu.]

Primjer 5.14. Neka je $X \sim \mathrm{Unif}(-1,1)$ te $Y := X^2$. Tada Y poprima vrijednosti u $(0,1)$ – odredimo F_Y [tj. njenu razdiobu].

- očito je $F_Y(y) = 0$ za $y \leqslant 0$, $F_Y(y) = 1$ za $y \geqslant 1$.
- za $y \in (0,1)$,

$$\begin{aligned} F_Y(y) &= \mathbb{P}(X^2 \leqslant y) = \mathbb{P}(X \in [-\sqrt{y},\sqrt{y}]) = [X \sim \mathrm{Unif}(-1,1)] \\ &= \frac{2\sqrt{y}}{2} = \sqrt{y}\,. \end{aligned}$$

[nacrtati F_Y]

[5] Ovu razdiobu često nazivamo **standardna** eksponencijalna razdioba.

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 $F_Y(y) = \int_{-\infty}^{y} f(t)\,\mathrm{d}t$,[6] za sve $y \in \mathbb{R}$. Dakle, Y je neprekidna s gustoćom f – specijalno, Y nema $\mathrm{Unif}(0,1)$ razdiobu.

NAPOMENA 5.15 (**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][7], 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.] □

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!]

[6] 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

$$\begin{aligned} \int_{-\infty}^{y} f(t)\,\mathrm{d}t &= \int_{-\infty}^{0} f(t)\,\mathrm{d}t + \int_0^y f(t)\,\mathrm{d}t = \int_0^y f(t)\,\mathrm{d}t = \lim_{a \to 0^-} \int_a^y f(t)\,\mathrm{d}t \\ &= \lim_{a \to 0^+} F_Y(y) - F_Y(a) = F_Y(y) - F_Y(0) = F_Y(y)\,. \end{aligned}$$

[7] Napomenimo ipak da rezultat vrijedi i bez pretpostavke (iii).

PROPOZICIJA 5.16 (**Simuliranje slučajnih varijabli**). *Neka je F funkcija distribucije takva da za neke $a, b \in [-\infty, +\infty]$, $a < b$, vrijedi da je F neprekidna i strogo rastuća funkcija na (a, b) t.d. je $F(a) = 0, F(b) = 1$ – specijalno, F je bijekcija sa (a, b) u $(0, 1)$ [nacrtati primjer].*[8] *Tada za $U \sim \mathrm{Unif}(0, 1)$, slučajna varijabla $X := F^{-1}(U) \in (0, 1)$ ima funkciju distribucije točno F.*

[Ako znamo simulirati $\mathrm{Unif}(0, 1)$ slučajnu varijablu, prethodni rezultat kaže da možemo pomoću nje simulirati bilo koju varijablu čija funkcija distribucije zadovoljavaja gornje pretpostavke – to npr. isključuje sve diskretne razdiobe, ali čak i dosta neprekidnih razdioba. Ipak, gornji rezultat se lako može 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. 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 $x \in (a, b)$,

$$\begin{aligned} F_X(x) &= \mathbb{P}(F^{-1}(U) \leqslant x) = [F \text{ strogo rastuća}] = \mathbb{P}(F(F^{-1}(U)) \leqslant F(x)) \\ &= \mathbb{P}(U \leqslant F(x)) = [F(x) \in (0, 1),\, U \sim \mathrm{Unif}(0, 1)] = F(x)\,. \end{aligned}$$

□

[**Napomena.** Obratno, ako F zadovoljava uvjete prethodne propozicije i X je t.d. je $F_X = F$, slučajna varijabla $Y := F(X)$ ima $\mathrm{Unif}(0, 1)$ razdiobu (vježbe). Ovaj rezultat vrijedi i općenitije – dovoljno je da je F samo neprekidna (bez dokaza), ali bez tog uvjeta rezultat ne mora vrijedi.]

PRIMJER 5.17. 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. 5.16 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

$$-\frac{\log(U)}{\lambda} \sim \mathrm{Exp}(\lambda)\,.$$

PROPOZICIJA 5.18 (**Zamjena varijabli**). *Pretpostavimo da je*

- *X neprekidna slučajna varijabla;*
- *I otvoren interval t.d. $\mathbb{P}(X \in I) = 1$, te*

[8] Uvjet $F(a) = 0, F(b) = 1$ povlači da je $\mathbb{P}(X \in (a, b)) = 1$.

- *$g : I \to (m, M)$ strogo monotona 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, npr. ako je g strogo rastuća,

$$F_Y(y) = \mathbb{P}(g(X) \leqslant y) = [\text{i } g^{-1} \text{ strogo rastuća}] = \mathbb{P}(X \leqslant g^{-1}(y)) = F_X(g^{-1}(y))\,,\quad y \in (m, M)\,.$$

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 5.14). 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 5.18. Neka je g strogo rastuća. 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 5.15(b).

Ako je g strogo padajuća, 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 analogno budući da je $(g^{-1})'(y) < 0$, pa je $-(g^{-1})'(y) = |(g^{-1})'(y)|$. □]

5.4. Očekivanje i varijanca

Definicija 5.19. 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]

$$\mathbb{E}[X] := \int_{-\infty}^{+\infty} t f(t)\,\mathrm{d}t\,. \tag{5.13}$$

Primjer 5.20. 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 3.22, koji je kao i u diskretnom slučaju dosta važan pri računanju očekivanja.]

Teorem 5.21. *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*

$$\mathbb{E}[g(X)] = \int_{-\infty}^{+\infty} g(t) f(t)\,\mathrm{d}t \tag{5.14}$$

Važno je napomenuti da (5.14) 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. 5.18 – 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
$$\begin{aligned}\mathbb{E}[Y] &= \int_{-\infty}^{+\infty} y f_Y(y)\,\mathrm{d}y = \int_m^M y f_X(g^{-1}(y))(g^{-1})'(y)\,\mathrm{d}y \\ &= [x = g^{-1}(y),\ \mathrm{d}x = (g^{-1})'(y)\,\mathrm{d}y,\ g^{-1}(\langle m, M\rangle) = I] \\ &= \int_I g(x) f_X(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} g(x) f_X(x)\,\mathrm{d}x \ \in [0,\infty]\,.\end{aligned}$$
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 (5.14) 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
$$\mathbb{P}(Y = a_i) = \mathbb{P}(X \in g^{-1}(\{a_i\})) = \int_{g^{-1}(\{a_i\})} f_X(t)\,\mathrm{d}t\,,$$

[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 $[0,\infty]$.

pa budući da familija skupova $\{g^{-1}(\{a_i\}),\ i \in I\}$ čini particiju skupa $\mathbb{R}$, imamo

$$\mathbb{E}[Y] = \sum_{i \in I} \int_{g^{-1}(\{a_i\})} \underbrace{a_i}_{=g(t)} f_X(t)\,\mathrm{d}t = \int_{-\infty}^{\infty} g(t) f_X(t)\,\mathrm{d}t\,.$$

$\square$

Iz (5.14) 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 5.22. *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

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b\,. \tag{5.15}$$

(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 $a \geqslant 0$,

$$\begin{aligned}\mathbb{E}[aX + b] &\overset{(5.14)}{=} \int_{-\infty}^{\infty} (at + b) \cdot f_X(t)\,\mathrm{d}t = a \int_{-\infty}^{\infty} t \cdot f_X(t)\,\mathrm{d}t + b \int_{-\infty}^{\infty} f_X(t)\,\mathrm{d}t \\ &= a\mathbb{E}[X] + b \cdot 1\,.\end{aligned}$$

Općeniti slučaj (postoji $\mathbb{E}[aX + b]$ i vrijedi (5.15)) 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 5.23 (**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 $\mathbb{E}[U] = \frac{1}{2}$,

$$\mathbb{E}[X] \overset{(5.15)}{=} (b - a)\mathbb{E}[U] + a = \frac{a + b}{2}\,.$$

[Zapravo...]

Napomena 5.24. (!) Svojstva **linearnosti** i **monotonosti** očekivanja (Teoremi 4.18 i 4.21) vrijede i za neprekidne slučajne varijable [bez dokaza].

Propozicija 5.25. *Ako je X neprekidna* ***nenegativna*** *slučajna varijabla, vrijedi*

$$\mathbb{E}[X] = \int_0^{\infty} \mathbb{P}(X > t)\,\mathrm{d}t \ \in [0, \infty]\,. \tag{5.16}$$

Dokaz. Vrijedi

$$\int_0^\infty \mathbb{P}(X > t)\,\mathrm{d}t = \int_0^\infty \Big(\int_t^\infty f_X(z)\,\mathrm{d}z\Big)\,\mathrm{d}t\,.$$

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

$$\int_0^\infty \mathbb{P}(X > t)\,\mathrm{d}t = \int_0^\infty \Big(\int_0^z f_X(z)\,\mathrm{d}t\Big)\,\mathrm{d}z = \int_0^\infty f_X(z) z\,\mathrm{d}z = \mathbb{E}[X]$$

gdje zadnja jednakost slijedi jer je $X \geqslant 0$, tj. $f_X(t) = 0$ za $t < 0$. □

Primjer 5.26. 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 5.27 (**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 (5.16) 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 (5.16) 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 $\mathbb{E}[X]$. □]

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 5.28 (**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 5.29 (**Nezavisnost**). Za neprekidne slučajne varijable $X_1, \ldots, X_n$ kažemo da su **nezavisne** ako za sve $t_1, \ldots, t_n \in \mathbb{R}$ vrijedi

$$\mathbb{P}(X_1 \leqslant t_1, \ldots, X_n \leqslant t_n) = \prod_{i=1}^n \mathbb{P}(X_i \leqslant t_i)\,.$$

Može se pokazati da je gornje ekvivalentno s

$$\mathbb{P}(X_1 \in B_1, \ldots, X_n \in B_n) = \prod_{i=1}^n \mathbb{P}(X_i \in B_i)\,,$$

za "skoro sve" $B_1, \ldots, B_n \subseteq \mathbb{R}$, te da u tom slučaju opet vrijedi [uz uvjet da je sve dobro definirano]

- $\mathbb{E}[\prod_{i=1}^n X_i] = \prod_{i=1}^n \mathbb{E}[X_i]$ [dokaz u poglavlju o neprekidnim slučajnim vektorima];
- $\mathrm{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathrm{Var}(X_i)$ [slijedi iz prethodnog svojstva i linearnosti očekivanja].

Primjer 5.30 (**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. [Intuitivno, ovo slijedi iz priče o Poissonovom procesu jer je "zbroj" dva nezavisna Poissonova procesa s intenzitetima (prosječan broj događaja po jedinici vremena) λ i μ ponovno Poissonov proces, ali s intenzitetom $\lambda + \mu$.]

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.]

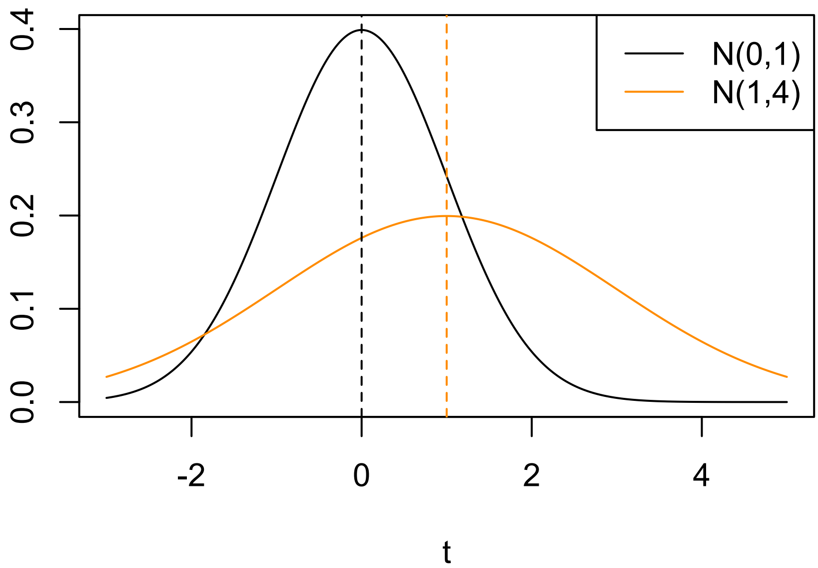
Kažemo da slučajna varijabla X ima **normalnu** (ili **Gaussovu**) razdiobu s parametrima μ i σ^2 (oznaka $X \sim \mathrm{N}(\mu, \sigma^2)$) za $\mu \in \mathbb{R}$, $\sigma > 0$, ako je neprekidna s gustoćom

$$f(t) = f_{\mathrm{N}(\mu,\sigma^2)}(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}\,, \quad t \in \mathbb{R}\,. \tag{5.17}$$

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.]



SLIKA 7. Gustoće $\mathrm{N}(0,1)$ i $\mathrm{N}(1,4)$ razdiobe. One su simetrične oko parametra μ, a parametar σ kontrolira brzinu opadanja prema 0 kako t ide u $\pm\infty$.

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 5.31. *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.

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 $Z \sim \mathrm{N}(0,1)$,

$$\mathbb{E}\big[|Z|\big] = \int_{-\infty}^{\infty} |t|\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\,\mathrm{d}t = \frac{2}{\sqrt{2\pi}}\int_0^{\infty} te^{-\frac{t^2}{2}}\,\mathrm{d}t = \sqrt{\frac{2}{\pi}}\cdot 1 < \infty$$

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 5.32 (**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. 5.18]. Njena funkcija distribucije je

$$F_X(x) = \mathbb{P}(\mu + \sigma\cdot Z \leqslant x) = \mathbb{P}\left(Z \leqslant \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)\,,\ x \in \mathbb{R}\,,$$

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 (5.17) 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

$$\mathbb{E}[X] = \mu + \sigma\mathbb{E}[Z] = \mu\,,$$
$$\mathrm{Var}(X) = \sigma^2\mathrm{Var}(Z) = \sigma^2\,,$$

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.

POGLAVLJE 6

Funkcije izvodnice

[Funkcije izvodnice (FI) su analitički alat za baratanje sa slučajnim varijablama – prije svega, korisne su u 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 $|s| \leqslant 1$,

$$|G(s)| \leqslant \sum_{n=0}^{\infty} p_n |s|^n \leqslant \sum_{n=0}^{\infty} p_n = 1 \,.$$

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

$$G^{(k)}(s) = \sum_{n=k}^{\infty} p_n n(n-1) \cdots (n-k+1) s^{n-k} \,, \quad \forall |s| < s_0 \,. \tag{6.3}$$

[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.

$$\mathbb{P}(X = k) = p_k = \frac{G^{(k)}(0)}{k!}\,. \tag{6.4}$$

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 (6.4) 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*

$$G_X^{(k)}(1) = \mathbb{E}[X(X-1)\cdots(X-k+1)]\,, \tag{6.5}$$

pri čemu u slučaju $s_0 = 1$ definiramo[2]

$$G_X^{(k)}(1) := \lim_{s \to 1^-} G_X^{(k)}(s) \quad \in [0, \infty]\,. \tag{6.6}$$

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 (6.6) 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 $s_0 > 1$, (6.5) slijedi iz (6.3) 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 $X_1, \ldots, X_n$ nezavisne slučajne varijable u $\mathbb{N}_0$, za sve $s \in \cap_{i=1}^n S_{X_i} \supseteq [-1,1]$ vrijedi*

$$G_{X_1+\cdots+X_n}(s) = G_{X_1}(s)\cdots G_{X_n}(s)\,. \tag{6.7}$$

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. 4.14].

PRIMJER 6.9. Ako su $X_1, \ldots, X_n$ 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

$$G_X(s) = [X_1, \ldots, X_n \text{ su njd}] = (G_{X_1}(s))^n = (q+ps)^n\,, \quad s \in \mathbb{R}\,.$$

TEOREM 6.10 (**FI slučajne sume**). *Neka je*

- *$X_1, X_2, \ldots$ 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 $X_1, X_2, \ldots$.*

[3] **Abelov teorem**: Ako je $u_0, u_1, \ldots$ 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 vrijedi

$$G_{S_N}(s) = G_N(G_{X_1}(s))\,, \quad \forall |s| \leqslant 1\,. \tag{6.8}$$

Dokaz. Za sve $|s| \leqslant 1$ [barem] imamo

$$\begin{aligned} G_{S_N}(s) = \mathbb{E}[s^{S_N}] &= \sum_{n=0}^{\infty} \overbrace{\mathbb{E}[s^{S_N} \mid N = n]}^{=\mathbb{E}[s^{S_n} | N=n]} \mathbb{P}(N = n) = [N \text{ i } S_n \text{ nezavisne!}] \\ &= \sum_{n=0}^{\infty} \mathbb{E}[s^{S_n}]\mathbb{P}(N = n) = [X_1, \dots, X_n \text{ su njd}] \\ &= \sum_{n=0}^{\infty} (G_{X_1}(s))^n \mathbb{P}(N = n) = G_N(G_{X_1}(s))\,. \end{aligned}$$

□

Primjer 6.11 (**Stanjivanje Poissonove slučajne varijable, Primjer 4.29** nastavak). [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 $q := 1 - p$ $(p \in (0, 1))$. Neka je N_P ukupan broj plavih kuglica.]. Koristeći (6.8) pokažite da N_P ima $\mathrm{P}(\lambda p)$ razdiobu.

Rješenje. Vrijedi $N_P = S_N$ gdje su $X_1, X_2, \dots$ njd $\mathrm{B}(p)$ slučajne varijable [i nezavisne od N]. Iz (6.8) 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. □

[Napomenimo za kraj da je klasična primjena gornjeg rezultata o FI vjerojatnosti slučajne sume je na tzv. *procese grananja* – to ćete raditi na kolegiju Markovljevi lanci.]

6.2. Funkcija izvodnica momenata

Definicija 6.12. Neka je X (bilo kakva) slučajna varijabla. Ako postoji $t_0 > 0$ za koji vrijedi

$$\mathbb{E}[e^{tX}] < \infty\,, \quad \forall |t| \leqslant t_0\,, \tag{6.9}$$

funkcija $M : [-t_0, t_0] \to [0, \infty)$ definirana sa

$$M(t) := \mathbb{E}[e^{tX}]\,, \tag{6.10}$$

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 (6.9) ekvivalentan je uvjetu

$$\mathbb{E}[e^{t_0|X|}] < \infty\,. \tag{6.11}$$

To slijedi jer za sve $|t| \leqslant t_0$ vrijedi

$$e^{tX} \leqslant e^{|tX|} \leqslant e^{t_0|X|} \leqslant \max\{e^{t_0X}, e^{-t_0X}\} \leqslant e^{t_0X} + e^{-t_0X}\,,$$

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 (6.9).*

*(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

$$e^{t_0|X|} = \sum_{n=0}^{\infty} \frac{t_0^n |X|^n}{n!} \geqslant \frac{t_0^k}{k!}|X|^k\,, \quad \forall k \in \mathbb{N}_0\,. \tag{6.12}$$

Monotonost očekivanja povlači da je

$$\frac{t_0^k}{k!}\mathbb{E}\big[|X|^k\big] \leqslant \mathbb{E}[e^{t_0|X|}] \overset{(6.11)}{<} \infty\,, \quad \forall k \in \mathbb{N}_0\,.$$

Tvrdnja sada slijedi budući da je $t_0 > 0$.

(b) Slično, za sve $|t| \leqslant t_0$ vrijedi

$$M_X(t) = \mathbb{E}[e^{tX}] = \mathbb{E}\left[\sum_{n=0}^{\infty} \frac{t^n X^n}{n!}\right]\,.$$

[4] Zato se M_X zove funkcija izvodnica **momenata**.

NAPOMENA*. Ono što bismo htjeli ovdje je zamijeniti očekivanje i sumu. To nam omogućava prikladna verzija Fubinijevog teorema[5] budući da je

$$\mathbb{E}\left[\sum_{n=0}^{\infty}\frac{|t|^n|X|^n}{n!}\right] = \mathbb{E}[e^{|t|\cdot|X|}] \leqslant \mathbb{E}[e^{t_0\cdot|X|}] \overset{(6.11)}{<} \infty\,. \qquad \square$$

Dakle,

$$M_X(t) = \sum_{n=0}^{\infty}\mathbb{E}\left[\frac{t^nX^n}{n!}\right] = \sum_{n=0}^{\infty}\frac{\mu_n}{n!}t^n\,,\quad \forall|t|\leqslant t_0\,. \tag{6.13}$$

(c) Ovo slijedi iz (6.13) – 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 $k \geqslant 0$,

$$M_X^{(k)}(t) = \sum_{n=k}^{\infty}\frac{\mu_n}{(n-k)!}t^{n-k}\,,\quad \forall|t| < t_0\,. \tag{6.14}$$

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

$$\begin{aligned} M_X(t) = \mathbb{E}[e^{tX}] &= \int_{-\infty}^{\infty} e^{tx}f_X(x)\mathrm{d}x = \int_0^{\infty} e^{tx}\lambda e^{-\lambda x}\mathrm{d}x = \lambda\int_0^{\infty}e^{-(\lambda-t)x}\mathrm{d}x \\ &= [\text{ako je } \lambda > t] = \frac{\lambda}{\lambda-t}\int_0^{\infty}\underbrace{(\lambda-t)e^{-(\lambda-t)x}}_{=f_{\mathrm{Exp}(\lambda-t)}(x)}\mathrm{d}x = \frac{\lambda}{\lambda-t}\,,\quad t<\lambda := t_0\,. \end{aligned}$$

[Formalno bi trebali uzeti $t_0 := \lambda - \epsilon > 0$ neki mali ϵ tako da vrijedi (6.9).] 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 $k = 2, 3, \ldots$, ali pogledajmo puno elegantniji pristup.

Ako je $E \sim \mathrm{Exp}(1)$, za sve $|t| < 1$ (specijalno je dakle $t < 1$) imamo

$$M_E(t) = \frac{1}{1-t} = \sum_{n=0}^{\infty} t^n = \sum_{n=0}^{\infty}\frac{t^n}{n!}\cdot n!\,.$$

[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

$$\mathbb{E}\left[\sum_{n=1}^{\infty}X_n\right] = \sum_{n=1}^{\infty}\mathbb{E}[X_n]\,,$$

pri čemu obje strane mogu biti $+\infty$. Općenito, gornja jednakost vrijedi ako imamo

$$\mathbb{E}\left[\sum_{n=1}^{\infty}|X_n|\right] = \sum_{n=1}^{\infty}\mathbb{E}[|X_n|] < \infty\,.$$

Teorem 6.14 povlači da je $\mathbb{E}[E^n] = n!$, za sve $n \geqslant 0$. Općenito, budući da $X \sim \text{Exp}(\lambda)$ 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

$$\begin{aligned} M_Z(t) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \cdot e^{tz} \mathrm{d}z = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(z^2 - 2tz + t^2\right) + \frac{t^2}{2}} \mathrm{d}z \\ &= e^{\frac{t^2}{2}} \int_{-\infty}^{\infty} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{(z-t)^2}{2}}}_{= f_{\mathrm{N}(t,1)}(z)} \mathrm{d}z = e^{\frac{t^2}{2}}, \quad \forall t \in \mathbb{R} . \end{aligned} \tag{6.15}$$

Za $X \sim \mathrm{N}(\mu, \sigma^2)$, budući da vrijedi $X \sim \sigma Z + \mu$, imamo

$$M_X(t) = \mathbb{E}[e^{tX}] = \mathbb{E}[e^{t\sigma Z + t\mu}] = e^{t\mu}\mathbb{E}[e^{t\sigma Z}] = e^{t\mu} M_Z(t\sigma) \overset{(6.15)}{=} e^{\mu t + \frac{\sigma^2}{2}t^2}, \quad \forall t \in \mathbb{R} . \tag{6.16}$$

Primjer 6.17 (**FI momenata ne postoji uvijek**). Ako je $X \sim \text{Pareto}(\alpha)$ za $\alpha > 0$ proizvoljan (vidi Primjer 5.27), 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 (6.9), 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*

$$M_{X+Y}(t) = M_X(t) M_Y(t), \quad |t| \leqslant t_0 . \tag{6.17}$$

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 (6.17) 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

$$M_{X+Y}(t) = [\text{nez.}] = M_X(t)M_Y(t) \overset{(6.16)}{=} e^{(\mu_1+\mu_2)t + \frac{1}{2}t^2(\sigma_1^2+\sigma_2^2)} \overset{(6.16)}{=} M_{\mathrm{N}(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)}(t)\,.\ ]$$

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

$$\max_{i=1,\ldots,n} \sigma_i^{(n)} = cn^{1+\epsilon}\,, \quad \forall n \geqslant 1\,,$$

imamo

$$\frac{\sqrt{\sum_{i=1}^n \sigma_i^2}}{n} \geqslant c^2\frac{n^{1+\epsilon}}{n} = cn^{\epsilon} \longrightarrow +\infty\,, \quad n \to \infty\,.$$

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 Primjer 8.7 i 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* $X_1, \ldots, X_n$ *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

$$M_{T_n}(t) = [X_1, \ldots, X_n \text{ njd}] = (M_{X_1}(t))^n = \left(\frac{\lambda}{\lambda - t}\right)^n = M_{\mathrm{Gama}(n,\lambda)}(t), \quad t < \lambda\,.$$

□

[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 modernim primjenama 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

$$\mathbb{P}\left(|\overline{X}_n - \mu| > \epsilon\right) \leqslant \frac{\sigma^2/n}{\epsilon^2} = \frac{\sigma^2}{\epsilon^2}\cdot\frac{1}{n} \xrightarrow{n\to\infty} 0\,,$$

pri čemu smo na kraju iskoristili da je $\sigma^2 < \infty$ [u suprotnom je gornja ograda jednaka $+\infty$ za sve n]. $\square$

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

$$\mathbb{P}\left(|\overline{X}_n - \mathbb{E}[X_1]| > \epsilon\right) = 1\,, \quad \forall n\,.$$

Specijalno, $\overline{X}_n$ ne konvergira po vjerojatnosti prema $\mathbb{E}[X_1]$. [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"].

[Ovdje nije loše malo zastati i razmisliti što smo pokazali. To da će postotak pisama koji su pali u velikom broju bacanja s velikom vjerojatnosti biti blizu p odgovara našem intuitivnom shvaćanju "vjerojatnosti" i onome što vidimo u stvarnome svijetu. Ipak, ovaj rezultat je čisto teorijski rezultat baziran na našoj aksiomatskoj definiciji vjerojatnosti, definiciji slučajne varijable, definiciji nezavisnosti itd. Ključna je ovdje činjenica da je zaista moguće formalno konstruirati vjerojatnosni prostor koji bi predstavljao model za beskonačan niz bacanja nezavisnih novčića.]

[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 $\mathrm{Var}(\overline{X}_n) \xrightarrow{n\to\infty} 0$.

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. Može se pokazati da $X_n \xrightarrow{\text{g.s.}} X$ povlači $X_n \xrightarrow{\mathbb{P}} X$. □

[Prije jakog ZVB-a, na jednostavnom primjeru ćemo ilustrirati konvergencije $\xrightarrow{\text{g.s.}}$ i $\xrightarrow{\mathbb{P}}$, te usput pokazati da $\xrightarrow{\mathbb{P}}$ nužno ne povlači $\xrightarrow{\text{g.s.}}$.]

Primjer 7.10. Na primjer, neka su $X_1, X_2, \ldots$ nezavisne slučajne varijable takve da je za $n \in \mathbb{N}$, $X_n \sim \mathrm{B}(1/n)$ [nacrtati skicu distribucije]. 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 $X_n \in \{0,1\}$,

$$\mathbb{P}\Big(|X_n - 0| > \epsilon\Big) = \mathbb{P}\Big(X_n > \epsilon\Big) = \mathbb{P}(X_n = n) = \frac{1}{n} \xrightarrow{n\to\infty} 0\,.$$

Za svaki $\epsilon \geqslant 1$ očito imamo

$$\mathbb{P}\Big(|X_n| > \epsilon\Big) = 0 \xrightarrow{n\to\infty} 0\,,$$

pa smo dakle pokazali da vrijedi $X_n \xrightarrow{\mathbb{P}} 0$. [Zapravo je uvijek dovoljno gledati ϵ-e na proizvoljno malom intervalu oko 0.]

Pogledajmo sada konvergenciju gotovo sigurno. 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

$$\mathbb{P}(\underbrace{X_n = 1 \text{ b.m.p.}}_{=:A_1}) = 1$$

te na analogan način zaključujemo da je i

$$\mathbb{P}(\underbrace{X_n = 0 \text{ b.m.p.}}_{=:A_0}) = 1\,.$$

Specijalno, $\mathbb{P}(A_0 \cap A_1) = 1$ pri čemu za sve $\omega \in A_1 \cap A_2$, $X_1(\omega), X_2(\omega), \ldots$ je (deterministički) niz u $\{0,1\}$ koji ima beskonačno 0-a i beskonačno mnogo 1-a (tj. ima dva gomilišta), pa taj niz nije konvergentan, tj. ne postoji $X(\omega) \in \mathbb{R}$ t.d. $X_n(\omega) \to X(\omega)$. Sada odmah slijedi

$$\mathbb{P}(\{\exists \lim_{n\to\infty} X_n\}) \leqslant \mathbb{P}((A_0 \cap A_1)^c) = 0\,,$$

pa specijalno ne postoji X t.d. $X_n \xrightarrow{\text{g.s.}} X$. □

TEOREM 7.11 (**Jaki ZVB**). *Neka je $X_1, X_2, \dots$ 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

- $\mathbb{E}[X_i^4] = [X_1, X_2, \dots$ su njd$] = \mathbb{E}[X_1^4] = K$, $\forall i$;
- $\mathbb{E}[X_i^2 X_j^2] =$ [nezavisnost] $= \mathbb{E}[X_i^2]\mathbb{E}[X_j^2] = [X_1, X_2, \dots$ 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 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

$$\mathbb{P}\left(\lim_{n\to\infty} \overline{X}_n = 0\right) = \mathbb{P}\left(\bigcap_{k=1}^{\infty} A_{1/k}\right) = \lim_{k\to\infty} \mathbb{P}(A_{1/k}) = 1\,.$$

[Uočimo da je ključno bilo da je ocjena na $\mathbb{P}(|\overline{X}_n| \geqslant \epsilon)$ bila reda veličine $O(n^{-2})$, što smo mogli postići upravo zbog pretpostavke o konačnom četvrtom momentu. Na primjer, kada bi imali samo konačan drugi moment, koristeći Čebiševljevu nejednakost dobili bi gornju ogradu reda veličine $O(n^{-1})$, što ne bi bilo dovoljno za korištenje prve BC leme.]

Na kraju, ako je $\mu \in \mathbb{R}$, za $Y_i := X_i - \mu$, $i \in \mathbb{N}$, imamo da je $Y_1, Y_2, \dots$ njd niz uz $\mathbb{E}[Y_1] = 0$, pa prethodni dokaz povlači

$$1 = \mathbb{P}\left(\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} Y_i = 0\right) = \mathbb{P}\left(\lim_{n\to\infty} \frac{S_n - n\mu}{n} = 0\right) = \mathbb{P}\left(\lim_{n\to\infty} \overline{X}_n = \mu\right).$$

$\square$

[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 $X_1, X_2, \dots$ 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 $\overline{X}_n \xrightarrow{\text{g.s.}} 0$. $\square$]

Primjer 7.12 (**Neparametarska statistika**). Pretpostavimo da je F proizvoljna (nepoznata) funkcija distribucije i $X_1, \dots, X_n$ njd niz takav da je $F_{X_i} = F$. Pitanje je možemo li iz jedne realizacije uzorka $X_1, \dots, X_n$ procijeniti F.

Empirijsku funkciju distribucije uzorka $X_1, \dots, X_n$ 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 8. Zapravo, ako imamo realizaciju $X_i(\omega) = x_i$, $i = 1, \dots, n$, te je $x_i \neq x_j$, $\widehat{F}_n(\omega)$ je zapravo funkcija distribucije uniformne razdiobe na skupu $\{x_1, \dots, x_n\}$.

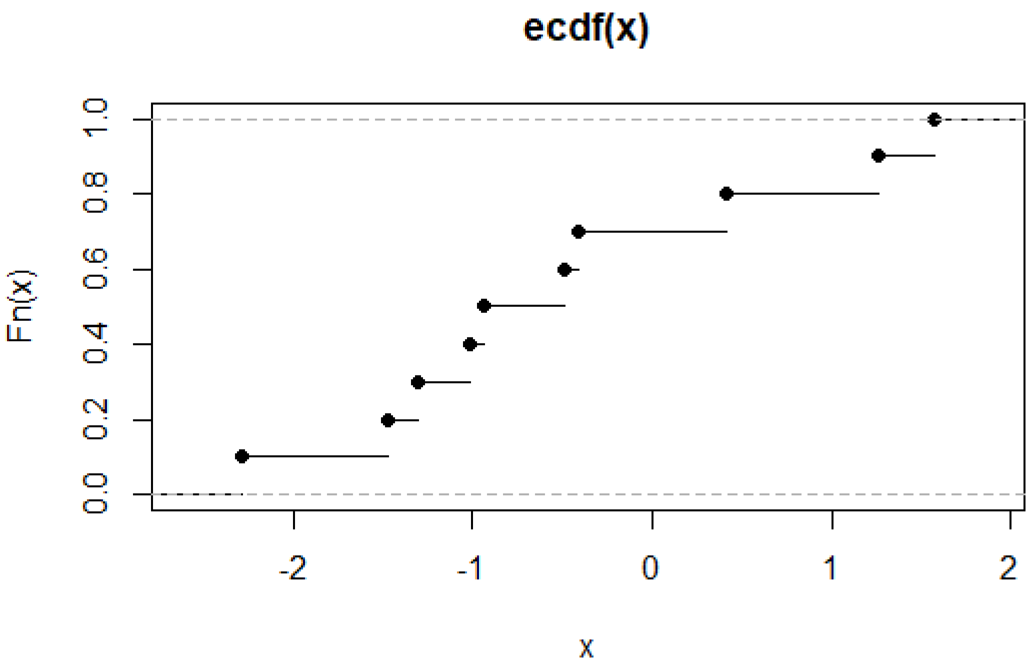
Iz JZVB-a za njd niz $Y_i := \mathbb{1}_{\{X_i \leqslant t\}}$, $i \in \mathbb{N}$, slijedi da za sve $t \in \mathbb{R}$ vrijedi

$$\widehat{F}_n(t) \xrightarrow[n\to\infty]{\text{g.s.}} \mathbb{E}[\mathbb{1}_{\{X_1 \leqslant t\}}] = \mathbb{P}(X_1 \leqslant t) = [F_{X_1} = F] = F(t)\,.$$

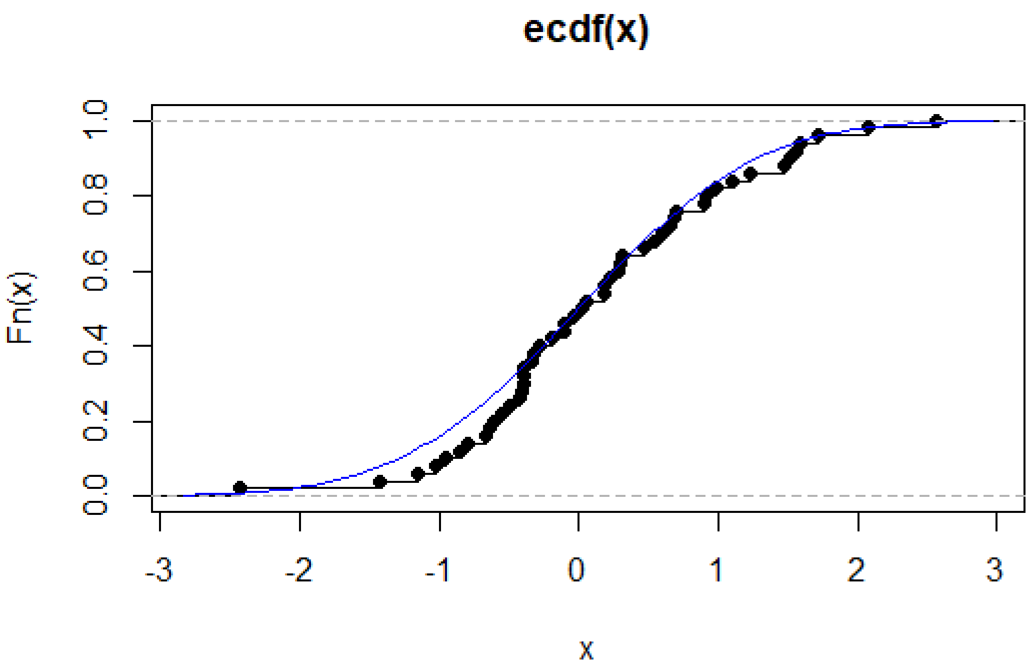
Zapravo, vrijedi i više: tzv. *Glivenko-Cantellijev* teorem daje

$$\sup_{t\in\mathbb{R}} |\widehat{F}_n(t) - F(t)| \xrightarrow[n\to\infty]{\text{g.s.}} 0\,,$$

vidi Sliku 9 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 8. 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 9. 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

[Kako bi iskazali centralni granični teorem, potreban nam je treći oblik konvergencije slučajnih varijabli.]

DEFINICIJA 7.13. Za niz slučajnih varijabli $Y_1, Y_2, \ldots$, kažemo da **konvergira po distribuciji** prema slučajnoj varijabli Y (oznaka $Y_n \xrightarrow{d} Y$) ako vrijedi

$$F_{Y_n}(t) \xrightarrow{n\to\infty} F_Y(t)\,, \quad \forall t \in C_{F_Y}\,, \tag{7.5}$$

gdje je $C_{F_Y} \subseteq \mathbb{R}$ skup svih točaka neprekidnosti funkcije (distribucije) F_Y.[4]

NAPOMENA 7.14. (a) **Zašto ne tražimo da (7.5) vrijedi za sve $t \in \mathbb{R}$?** Na primjer, ako je $Y_n := \frac{1}{n}$ [dakle, konstanta], $n \in \mathbb{N}$, očekivali bismo da niz (Y_n) konvergira po distribuciji prema $Y := 0$. To zaista i je slučaj [nacrtati pripadne funkcije distribucije] jer za sve $t \neq 0$ (tj. $t \in C_{F_Y}$) vrijedi

- ako je $t < 0$, imamo $F_{Y_n}(t) = 0 = F_Y(t)$, za sve n;
- ako je $t > 0$, čim je $t \geqslant \frac{1}{n}$, tj. $n \geqslant \frac{1}{t}$, imamo $F_{Y_n}(t) = 1 = F_Y(t)$.

Ipak, u ovom slučaju $F_{Y_n}(0) = 0 \not\to 1 = F_Y(0)$ kada $n \to \infty$.

(b) (!) Može se pokazati da je (7.5) povlači puno više: u tom slučaju imamo

$$\mathbb{P}(Y_n \in B) \xrightarrow{n\to\infty} \mathbb{P}(Y \in B)\,,$$

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, $\partial(-\infty, x] = \{x\}, \partial[a,b) = \partial(a,b) = \{a,b\}$. Zbog ovoga (7.5) neformalno pišemo kao $Y_n \overset{d}{\approx} Y$ za velike n.

□

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

$$Z_n := \frac{S_n - n\mu}{\sigma\sqrt{n}}\,, \quad n \in \mathbb{N}\,, \tag{7.6}$$

vrijedi $\mathbb{E}[Z_n] = 0$, $\mathrm{Var}(Z_n) = 1$, $\forall n \in \mathbb{N}$.

---

[4] Uočimo, $Y_n \xrightarrow{d} Y$ 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.

Teorem 7.15 (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 (7.6). Tada vrijedi*

$$Z_n \xrightarrow{d} Z \sim \mathrm{N}(0,1)\,. \tag{7.7}$$

Dakle, za dovoljno veliki n, i sve $x \in \mathbb{R}$, vrijedi

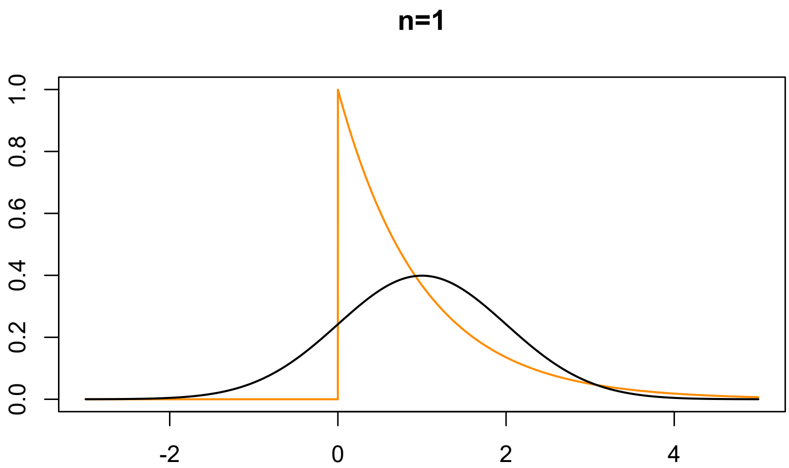
$$\mathbb{P}(S_n \leqslant x) = \mathbb{P}\left(Z_n \leqslant \frac{x - n\mu}{\sigma\sqrt{n}}\right) \overset{(7.7)}{\approx} \mathbb{P}\left(Z \leqslant \frac{x - n\mu}{\sigma\sqrt{n}}\right) = \mathbb{P}(\sigma\sqrt{n}Z + n\mu \leqslant x)\,.$$

Budući da je $\sigma\sqrt{n}Z + n\mu \sim \mathrm{N}(n\mu, n\sigma^2)$, CGT možemo neformalno iskazati kao

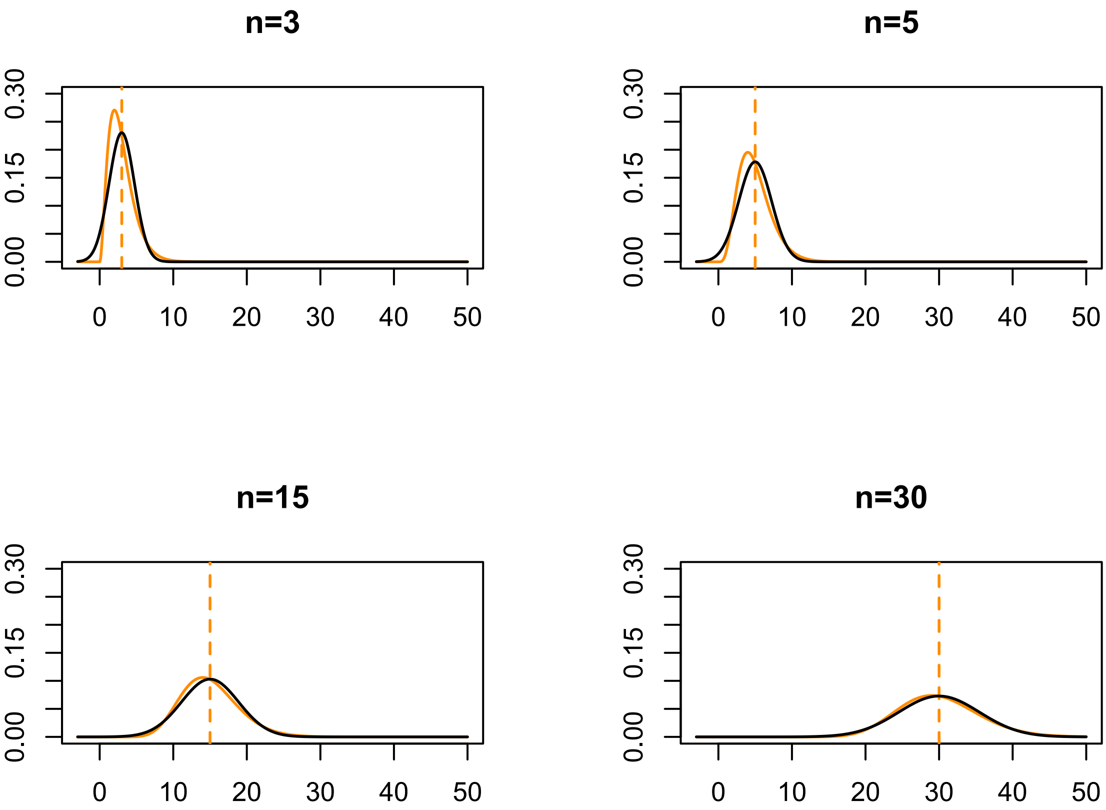
$$S_n \overset{d}{\approx} \mathrm{N}(n\mu, n\sigma^2)\,, \quad \text{za velike } n, \tag{7.8}$$

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 $\mathrm{Var}(X_1) \in (0, \infty)$)!

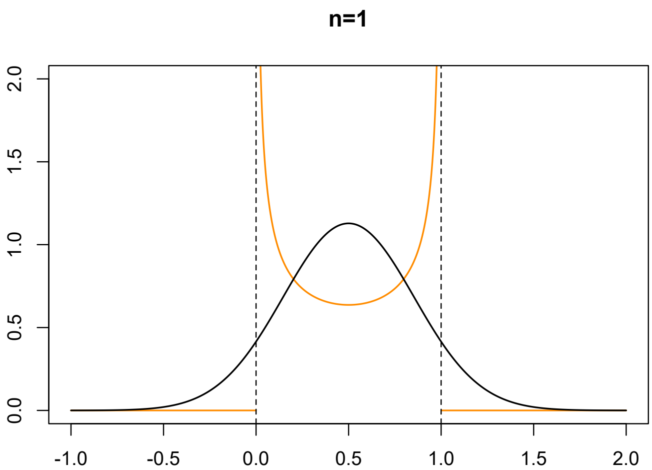
Za ilustraciju CGT-a, tj. aproksimacije (7.8), za razne distribucije od X_i, vidi Slike 10-14. U nekim slučajevima koristimo tzv. *histogram* dobiven na temelju velikog broja simulacije slučajne varijable S_n – o tome ćete više učiti na kolegiju Statistika, ali za naše potrebe dovoljno je znati da histogram predstavlja procjenu gustoće neprekidne razdiobe koja najbolje odgovara razdiobi od S_n.



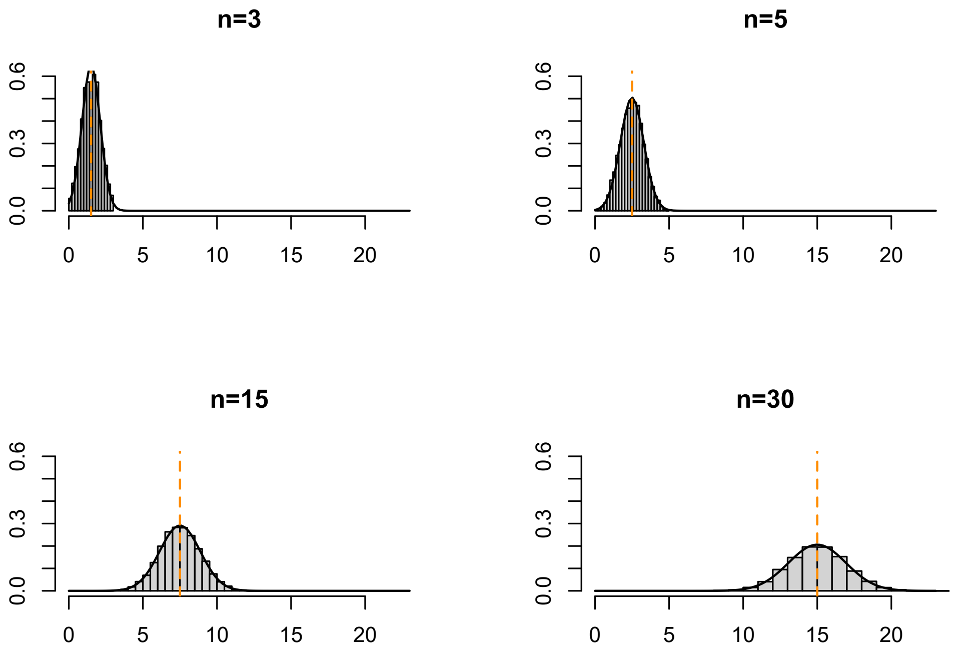
Slika 10. Gustoća Exp(1) razdiobe (narančasta linija), zajedno s gustoćom normalne razdiobe s istom očekivanjem i varijancom.



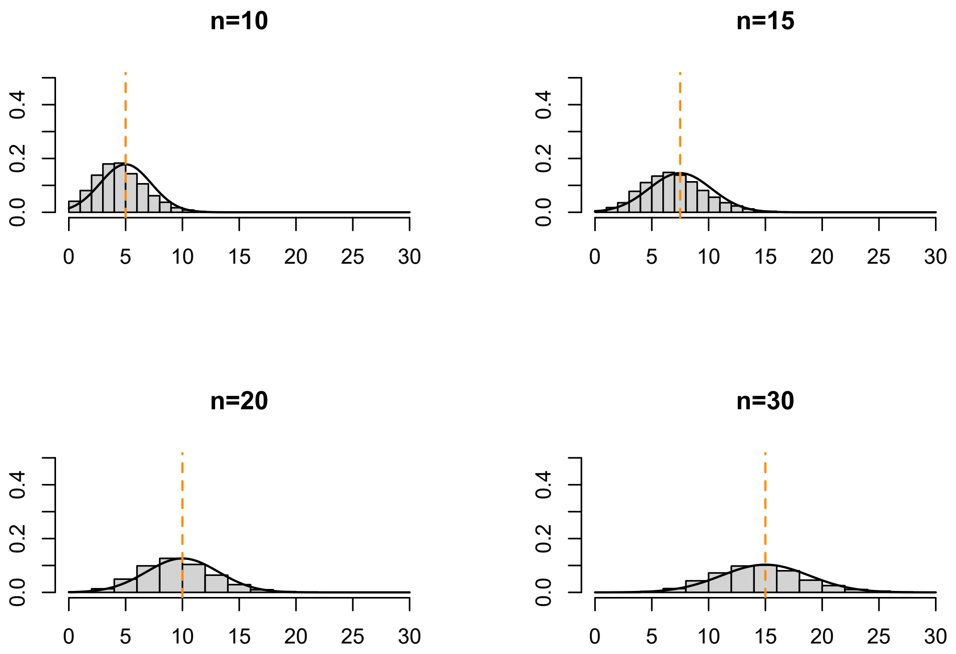
Slika 11. Gustoća od $S_n = \sum_{i=1}^{n} X_i$ (narančasta linija) u slučaju kada su $X_1, X_2, \ldots$ njd iz Exp(1) razdiobe ($S_n \sim \text{Gama}(n, 1)$), zajedno s gustoćom normalne razdiobe s istom očekivanjem (vertikalna linija) i varijancom, za razne n-ove.



Slika 12. Gustoće Beta(1/2, 1/2) razdiobe (vidi iduće poglavlje za preciznu definiciju), zajedno s gustoćom normalne razdiobe s istom očekivanjem i varijancom.



Slika 13. Histogram od $S_n = \sum_{i=1}^{n} X_i$ u slučaju kada su $X_1, X_2, \dots$ 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 (7.8) se čini dosta dobra i za male n-ove.



Slika 14. Histogram od $S_n = \sum_{i=1}^{n} X_i$ u slučaju kada su $X_1, X_2, \dots$ 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 7.16. *Ako je $b_1, b_2, \dots$ 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 7.17 (**Teorem neprekidnosti**). *Neka su $Y, Y_1, Y_2, \dots$ slučajne varijable te pretpostavimo da postoji $t_0 > 0$ t.d. su FI momenata $M_Y, M_{Y_1}, M_{Y_2}, \dots$ dobro definirane na $(-t_0, t_0)$ [dakle, $M_Y(t), M_{Y_1}, \dots < \infty$ 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 $Y_n \xrightarrow{d} Y$.

DOKAZ TEOREMA 7.15. 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

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^n Y_i =: \frac{1}{\sigma\sqrt{n}} S_n^Y , \tag{7.9}$$

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 $Y_1, Y_2, \dots$ njd, iz (7.9) 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

$$M(x) = M(0) + M'(0)x + \frac{M''(0)}{2}x^2 + R_2(x)\,, \quad \forall |x| < t_0\,,$$

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

$$M(x) = 1 + \mathbb{E}[Y_1] + \frac{\mathbb{E}[Y_1^2]}{2}x^2 + x^2 h(x) = 1 + \frac{\sigma^2}{2}x^2 + x^2 h(x)\,, \quad \forall |x| < t_0\,. \tag{7.10}$$

Sada, za svaki fiksni $t \in \mathbb{R}$ te n dovoljno velik (tako da je $|t| < t_0 \sigma \sqrt{n}$) imamo

$$\begin{aligned} M_{Z_n}(t) &= \left(M\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n \overset{(7.10)}{=} \left(1 + \frac{t^2}{2n} + \frac{t^2}{\sigma^2 n} h\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n \\ &= \left(1 + \frac{1}{n}\left[\frac{t^2}{2} + \frac{t^2}{\sigma^2} h\left(\frac{t}{\sigma\sqrt{n}}\right)\right]\right)^n =: \left(1 + \frac{b_n(t)}{n}\right)^n . \end{aligned}$$

Budući da vrijedi $b_n(t) \xrightarrow{n\to\infty} \frac{t^2}{2}$, iz Leme 7.16 dobivamo

$$M_{Z_n}(t) \xrightarrow{n\to\infty} e^{\frac{t^2}{2}} \overset{(6.15)}{=} M_Z(t)\,, \quad \forall t \in \mathbb{R}\,,$$

pa (7.7) slijedi iz Teorema 7.17.

$\square$

Napomena 7.18. 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 7.19. Konvergencija u (7.7) povlači da za sve $-\infty \leqslant a < b \leqslant +\infty$ i proizvoljan interval $I \in \{(a,b), [a,b), (a,b], [a,b]\}$, vrijedi

$$\mathbb{P}(Z_n \in I) \xrightarrow{n\to\infty} \mathbb{P}(Z \in I) = \Phi(b) - \Phi(a)\,. \tag{7.11}$$

jer je $\partial I \subseteq \{a, b\}$, a $\mathbb{P}(Z \in \{a,b\}) = [Z \text{ neprekidna}] = 0$; vidi Napomenu 7.14(b).

Primjer 7.20 (**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 7.15 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**:

$$\frac{Y_n - np}{\sqrt{npq}} \xrightarrow{d} Z \sim \mathrm{N}(0,1)\,, \tag{7.12}$$

to jest

$$Y_n \overset{d}{\approx} \mathrm{N}(np, npq)\,, \quad \text{za velike } n.$$

Napomena 7.21 (**Alternativna formulacija CGT-a**). Uočimo, budući da je

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} \cdot \frac{1/n}{1/n} = \frac{\sqrt{n}}{\sigma} \cdot (\overline{X}_n - \mu)\,, \tag{7.13}$$

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 7.22. 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

$$\overline{X}_n \overset{d}{\approx} \mathrm{N}\left(p, \frac{\sigma_p^2}{n}\right)$$

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 $0.03 \geqslant 1.96 \cdot \frac{\sigma_p}{\sqrt{n}}$, tj. $n \geqslant \left(\frac{1.96}{0.03}\right)^2 \cdot \sigma_p^2$, vrijedi

$$\mathbb{P}\Big(|\overline{X}_n - p| \leqslant 0.03\Big) = \mathbb{P}\Big(|\overline{X}_n - p| \leqslant 1.96\tfrac{\sigma_p}{\sqrt{n}}\Big) \overset{\mathrm{CGT}}{\approx} 2\Phi(1.96) - 1 \approx 0.95\,.$$

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*

Napomena. Cilj ovog poglavlja je dokazati odnose između svih oblika konvergencija slučajnih varijabli koje smo do sada spomenuli. Dokazi samih rezultata se *neće* ispitivati, ali potrebno je znati da vrijedi sljedeći odnos:

$$X_n \xrightarrow{\text{g.s.}} X \;\Rightarrow\; X_n \xrightarrow{\mathbb{P}} X \;\Rightarrow\; X_n \xrightarrow{d} X\,,$$

pri čemu između $\xrightarrow{\mathbb{P}}$ i $\xrightarrow{d}$ imamo ekvivalenciju ako je limes X konstanta, tj. $\mathbb{P}(X = c) = 1$ za neki $c \in \mathbb{R}$.

U nastavku, ako ne kažemo drugačije, $X, X_1, X_2, \ldots$ predstavlja niz slučajnih varijabli definiranih na istom vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$.

Teorem 7.23. *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 $\mathbb{P}(A_n(\epsilon)) \xrightarrow{n\to\infty} 0$. □

U Primjeru 7.10 smo pokazali da u prethdnom teoremu obrat općenito ne vrijedi!.[5]

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

$$\limsup_{n\to\infty} F_{X_n} \leqslant F_X(t+\epsilon) + 0 = F_X(t+\epsilon)\,,$$

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

$$F_X(t-\epsilon) \leqslant F_{X_n}(t) + \mathbb{P}\Big(|X_n - X| > \epsilon\Big)$$

[5] Ipak, može se pokazati da $X_n \xrightarrow{\mathbb{P}}$ povlači da postoji podniz $(n_k)_{k\in\mathbb{N}}$ takav da $X_{n_k} \xrightarrow{\text{g.s.}} X$ kada $k \to \infty$.

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 $F_{X_n}(t) \xrightarrow{n\to\infty} F_X(t)$, $\forall t \in C_{F_X}$. □

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

$$0 \leqslant \mathbb{P}\Big(|X_n - c| > \epsilon\Big) \xrightarrow{n\to\infty} 1 - F_c(c + \epsilon) + F_c(c - \epsilon) = 1 - 1 + 0 = 0\,.$$

□

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.

Zadatak. Ako su $X, X_1, X_2, \ldots$ diskretne slučajne varijable koje poprimaju vrijednosti u $\mathbb{N}_0$, pokažite da 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 $np_n \xrightarrow{n\to\infty} \lambda \in (0, \infty)$.

POGLAVLJE 8

Neprekidni slučajni vektori

[U poglavlju o diskretnim slučajnim vektorima smo promatrali dvije ili više zavisnih diskretnih slučajnih varijabli. Za primjene je ipak daleko zanimljiviji neprekidan slučaj, te ćemo se njime baviti u nastavku. U odnosu na diskretan slučaj, neprekidan je tehnički nešto zahtjevniji – primjerice, (i) trebat će nam pojam višedimenzionalnog integrala (vidi Dodatak A) koji ćete raditi tek na kasnijim kolegijima, te (ii) uvjetovanje na ishod neprekidne slučajne varijable je jako bitno, ali to je uvjetovanje na događaj vjerojatnosti 0 te zahtijeva poseban oprez.]

8.1. Zajednička funkcija gustoće

Za [proizvoljne] slučajne varijable X i Y [na istom vjerojatnosnom prostoru] definiramo njihovu **zajedničku funkciju distribucije** $F_{X,Y} = F : \mathbb{R}^2 \to [0,1]$ sa

$$F(a,b) := \mathbb{P}(X \leqslant a, Y \leqslant b)\,, \quad a,b \in \mathbb{R}\,. \tag{8.1}$$

Za slučajni vektor (X,Y) kažemo da je **neprekidan** ako postoji funkcija $f_{X,Y} = f : \mathbb{R}^2 \to [0,\infty)$ takva da za sve $a,b \in \mathbb{R}$ vrijedi

$$F(a,b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f(x,y)\,\mathrm{d}y\mathrm{d}x\,. \tag{8.2}$$

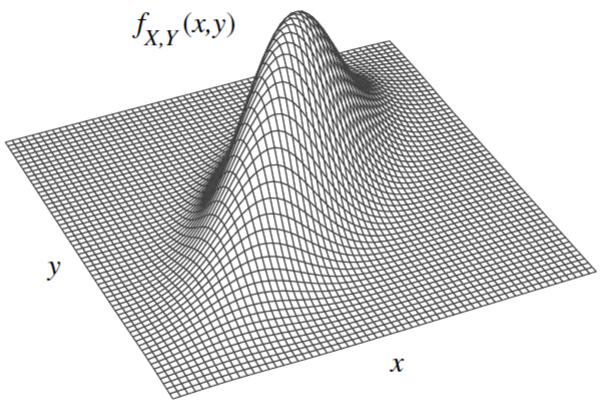
U tom slučaju funkciju f nazivamo funkcijom gustoće vektora (X,Y) ili **zajedničkom funkcijom gustoće** varijabli X i Y. Nadalje, za sve skupove $A \subseteq \mathbb{R}^2$ vrijedi

$$\mathbb{P}((X,Y) \in A) = \int_A f(x,y)\,\mathrm{d}x\mathrm{d}y\,. \tag{8.3}$$

Na primjer,

$$\mathbb{P}(X \geqslant 4, 2 < Y \leqslant 3) = \int_2^3 \int_4^\infty f(x,y)\,\mathrm{d}x\mathrm{d}y\,.$$

Napomena 8.1. Gustoća f nije jedinstvena – ako funkciji f promijenimo vrijednost na skupu površine 0 (npr. konačna ili prebrojiva unija točaka i/ili pravaca), to će i dalje biti gustoća vektora (X,Y). Ipak, može se pokazati da ako su f i g gustoće vektora (X,Y) tada postoji skup $A \subseteq \mathbb{R}^2$ takav da je A^c površine 0 te vrijedi $f(x,y) = g(x,y)$ za sve $(x,y) \in A$ – kažemo da su f i g jednake *gotovo svuda*. U nastavku ćemo radi jednostavnosti ovu činjenicu implicitno podrazumijevati. □



Slika 15. Primjer zajedničke gustoće – za svaki $A \subseteq \mathbb{R}^2$, $\mathbb{P}((X,Y) \in A)$ je volumen skupa $\{(x,y,z) \in \mathbb{R}^3 : (x,y) \in A, 0 \leqslant z \leqslant f(x,y)\}$. Preuzeto iz [BH19].

Osnovna svojstva. Neka je (X,Y) neprekidan slučajan vektor s gustoćom f.

1. Varijable X i Y su nužno neprekidne – gustoća od X dana je s
$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y\,, \quad x \in \mathbb{R}\,, \tag{8.4}$$
te ju zovemo **marginalna gustoća** od X. Zaista,
$$\mathbb{P}(X \leqslant a) = \mathbb{P}(X \leqslant a, Y \in \mathbb{R}) = \int_{-\infty}^{a} \left(\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y \right) \mathrm{d}x\,.$$
Vidi Sliku 17 za ilustraciju.
2. Za svaki skup $A \subseteq \mathbb{R}^2$ površine 0 (npr. točka ili pravac) vrijedi
$$\mathbb{P}((X,Y) \in A) = 0\,.$$
Specijalno, budući da je $A = \{(x,y) : x = y\}$ pravac, uvijek vrijedi
$$\mathbb{P}(X = Y) = 0\,.$$
3. **Interpretacija gustoće**: ako je $B_\epsilon(x,y) = (x,y) + [-\epsilon,\epsilon]^2$ kvadrat površine $(2\epsilon)^2$ centriran oko (x,y), za male $\epsilon > 0$ vrijedi
$$\mathbb{P}((X,Y) \in B_\epsilon(x,y)) = \int_{x-\epsilon}^{x+\epsilon} \int_{y-\epsilon}^{y+\epsilon} f(s,t)\,\mathrm{d}t\mathrm{d}s \approx f(x,y) \cdot (2\epsilon)^2\,.$$
4. Uvijek vrijedi
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x\mathrm{d}y = \mathbb{P}(X \in \mathbb{R}, Y \in \mathbb{R}) = 1\,.$$
5. Za gotovo sve $(x,y) \in \mathbb{R}^2$ vrijedi
$$f(x,y) = \frac{\partial^2}{\partial x \partial y} F(x,y)\,.$$

Nezavisnost. Za proizvoljne slučajne varijable X i Y [definirane na istom vjerojatnosnom prostoru] kažemo da su nezavisne ako vrijedi

$$F_{X,Y}(a,b) = F_X(a)F_Y(b)\,, \quad \forall a,b \in \mathbb{R}\,.$$

Može se pokazati da je to ekvivalentno s jačim uvjetom

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

Propozicija 8.2. *Ako je (X,Y) neprekidan slučajan vektor, X i Y su nezavisne akko je*

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)\,, \quad \forall x,y \in \mathbb{R}\,. \tag{8.5}$$

Dokaz. Ako su X i Y nezavisne, budući da su neprekidne s gustoćama f_X i f_Y, imamo

$$F_{X,Y}(a,b) = F_X(a)F_Y(b) = \left(\int_{-\infty}^{a} f_X(x)\,\mathrm{d}x\right)\cdot\left(\int_{-\infty}^{b} f_Y(y)\,\mathrm{d}y\right) = \int_{-\infty}^{a}\int_{-\infty}^{b} f_X(x)f_Y(y)\mathrm{d}y\mathrm{d}x\,.$$

Dakle, $f(x,y) := f_X(x)f_Y(y)$, $x,y \in \mathbb{R}$, je gustoća od (X,Y).

Obratno, za proizvoljne $a,b \in \mathbb{R}$ iz (8.5) slijedi

$$\begin{aligned} F_{X,Y}(a,b) &= \int_{-\infty}^{a}\int_{-\infty}^{b} f_{X,Y}(x,y)\mathrm{d}y\mathrm{d}x = \int_{-\infty}^{a}\int_{-\infty}^{b} f_X(x)f_Y(y)\mathrm{d}y\mathrm{d}x \\ &= \left(\int_{-\infty}^{a} f_X(x)\,\mathrm{d}x\right)\cdot\left(\int_{-\infty}^{b} f_Y(y)\,\mathrm{d}y\right) = F_X(a)F_Y(b)\,. \end{aligned}$$

□

[Napomena. U vidu Napomene 8.1, jednakost u (8.5) shvaćamo kao: postoji gustoća $f_{X,Y}$ od (X,Y) te gustoće f_X i f_Y of X i Y tako da vrijedi (8.5).]

Primjer 8.3 (**Dvodimenzionalna uniformna razdioba**). Neka je $A \subseteq \mathbb{R}^2$ proizvoljan skup pozitivne i konačne površine [dakle, A npr. ne može biti skup koji se sastoji samo od konačno mnogo točaka]. Kažemo da vektor (X,Y) ima uniformnu razdiobu na A (oznaka $(X,Y) \sim \mathrm{Unif}(A)$) ako je neprekidan s gustoćom

$$f(x,y) = \begin{cases} 1/\lambda(A)\,, & (x,y) \in A\,, \\ 0\,, & (x,y) \notin A\,. \end{cases}$$

[Kada kažemo da je (X,Y) "slučajno odabrana točka iz A", tipično pretpostavljamo da je $(X,Y) \sim \mathrm{Unif}(A)$.]

(a) Neka je $A = [-1,1]^2$; dakle, $\lambda(A) = 4$. Intuitivno je jasno da bi marginalne razdiobe od X i Y trebale biti uniformne na $[-1,1]$, te da su koordinate X i Y nezavisne. Pokažimo to i formalno. Za $x \in [-1,1]$,

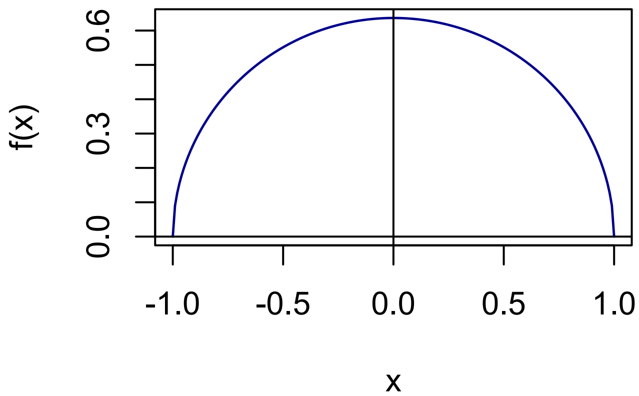
$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y = \int_{-1}^{1} 1/4\,\mathrm{d}y = 1/2\,,$$

te $f_X(x) = 0$ za $x \notin [-1,1]$. Dakle, $X \sim \mathrm{Unif}(-1,1)$, a zbog simetrije isto vrijedi i za Y. Nezavisnost slijedi odmah jer imamo $f(x,y) = f_X(x)f_Y(y)$ za sve $x,y \in \mathbb{R}$.

(b) Neka je sada A krug radijusa 1 oko ishodišta; dakle, $\lambda(A) = \pi$. [Intuitivno, vrijedi li sada $X \sim \mathrm{Unif}(-1,1)$, te jesu li X i Y nezavisne? NE!]
Prvo, budući da za $x \in [-1,1]$ imamo $(x,y) \in A$ akko $x^2 + y^2 \leqslant 1$, tj. ako $y \in [-\sqrt{1-x^2}, \sqrt{1-x^2}]$ (nacrtati sliku), vrijedi

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}y = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 1/\pi\,\mathrm{d}y = \frac{2}{\pi}\sqrt{1-x^2}\,.$$

Za $x \notin [-1,1]$ je $f_X(x) = 0$. Zbog simetrije vrijedi $f_Y(y) = f_X(y)$ za sve $y \in \mathbb{R}$. Uočimo da X očito nema $\mathrm{Unif}(-1,1)$ razdiobu, te da s većom vjerojatnosti poprima vrijednosti blizu 0 nego blizu ± 1, vidi Sliku 16.



Slika 16. Gustoća $f(x) = \frac{2}{\pi}\sqrt{1-x^2}$ za $x \in [-1,1]$.

Nadalje, X i Y nisu nezavisne jer, intuitivno, ako znamo da je $X = x$, to nužno povlači da je $Y \in [-\sqrt{1-x^2}, \sqrt{1-x^2}]$. Formalno, ako je $B = (a,b) \times (c,d) \subseteq [-1,1]^2 \setminus A$ proizvoljan i takav da je $\lambda(B) > 0$, imamo

$$\mathbb{P}(X \in (a,b), Y \in (c,d)) = \mathbb{P}((X,Y) \in B) = 0\,,$$

dok s druge strane vrijedi

$$\mathbb{P}(X \in (a,b)) \cdot \mathbb{P}(Y \in (c,d)) > 0\,.$$

Alternativno, vidimo da (8.5) sigurno ne vrijedi za sve $(x,y) \in [-1,1]^2 \setminus A$, pa slijedi da X i Y nisu nezavisne; ovdje je bitno da skup na kojem (8.5) ne vrijedi ima *pozitivnu površinu* jer bi inače mogli modificirati $f_{X,Y}$ tako da na tom skupu vrijedi (8.5), a i dalje imati gustoću od (X,Y).

□

Propozicija 8.4 (!). *Ako su X i Y nezavisne neprekidne slučajne varijable, tada je slučajni vektor (X,Y) nužno neprekidan s gustoćom $f(x,y) := f_X(x) f_Y(y)$, $x, y \in \mathbb{R}$.*

Dokaz. Slijedi iz dokaza Propozicije 8.2.

□

Napomena 8.5. U prethodnoj propoziciji nezavisnost je bitna [tj. ako su X i Y neprekidne i zavisne, (X,Y) ne mora nužno biti neprekidan slučajan vektor; uočite razliku u odnosu na diskretan slučaj]. Na primjer, ako je X neprekidna slučajna varijabla i $Y := X$, tada (X,Y) *nije* neprekidan slučajan vektor jer je očito $\mathbb{P}(X = Y) = 1 \neq 0$.

Primjer 8.6 (**Nezavisne eksponencijalne varijable s različitim parametrima**). Neka su $X \sim \text{Exp}(\lambda_1)$ i $Y \sim \text{Exp}(\lambda_2)$ nezavisne. Pokažimo da vrijedi $\mathbb{P}(X < Y) = \frac{\lambda_1}{\lambda_1+\lambda_2}$. [1]

Rješenje. Budući da je (X,Y) neprekidan slučajni vektor te vrijedi

$$A = \{(x,y) : x, y \geqslant 0, x < y\} = \{(x,y) : x \geqslant 0,\ y > x\}\,,$$

imamo

$$\begin{aligned}
\mathbb{P}(X < Y) = \mathbb{P}((X,Y) \in A) &= \int_0^\infty \int_x^\infty f_X(x) f_Y(y)\,\mathrm{d}y\mathrm{d}x \\
&= \int_0^\infty \lambda_1 e^{-\lambda_1 x} \left(\int_x^\infty \lambda_2 e^{-\lambda_2 y}\,\mathrm{d}y \right) \mathrm{d}x = \int_0^\infty \lambda_1 e^{-\lambda_1 x} e^{-\lambda_2 x} \mathrm{d}x \\
&= \frac{\lambda_1}{\lambda_1+\lambda_2} \int_0^\infty (\lambda_1+\lambda_2) e^{-(\lambda_1+\lambda_2)x} \mathrm{d}x = \frac{\lambda_1}{\lambda_1+\lambda_2}\,.
\end{aligned}$$

□

Sve gornje definicije i rezultati za dvodimenzionalne neprekidne slučajne vektore se prirodno poopćavaju na n-dimenzonalan slučaj za $n \geqslant 3$. Na primjer, ako je (X,Y,Z,W) neprekidan slučajan vektor u $\mathbb{R}^4$ s gustoćom $f : \mathbb{R}^4 \to [0,\infty)$, slučajni vektor (X,Z) je također neprekidan s gustoćom

$$f_{X,Z}(x,z) = \int_{-\infty}^\infty \int_{-\infty}^\infty f(x,y,z,w)\,\mathrm{d}y\,\mathrm{d}w\,,\ x,z \in \mathbb{R}\,.$$

Primjer 8.7 (**Maksimalna vjerodostojnost u neprekidnom slučaju**). Pretpostavimo da su $X_1,\ldots,X_n$ nezavisne slučajne varijable takve da je $X_i \sim \mathrm{N}(\mu^*, \sigma_i^2)$, $i = 1,\ldots,n$, pri čemu znamo varijance $\sigma_1^2,\ldots,\sigma_n^2 > 0$, ali ne znamo zajedničko očekivanje $\mu^* \in \mathbb{R}$ (ovime smo se već bavili u Primjeru 6.21).

Pretpostavimo da smo dobili realizacije $X_1 = x_1,\ldots,X_n = x_n$. Za svaki $\mu \in \mathbb{R}$ definiramo njegovu **log-vjerodostojnost** sa

$$l(\mu) := \log(f(x_1,\ldots,x_n;\mu))\,,$$

pri čemu je $f(\cdot\,;\mu)$ gustoća slučajnog vektora $(X_1^\mu,\ldots,X_n^\mu)$ gdje su $X_i^\mu \sim \mathrm{N}(\mu,\sigma_i^2)$, $i = 1,\ldots,n$ nezavisne. Procjenitelj maksimalne vjerodostojnosti je parametar $\widehat{\mu} = \widehat{\mu}(x_1,\ldots,x_n) \in \mathbb{R}$ za

[1] Znamo da vrijedi $m := \min\{X,Y\} \sim \text{Exp}(\lambda_1+\lambda_2)$, a ovdje računamo $\mathbb{P}(X < Y) = \mathbb{P}(m = X)$.

koji vrijedi

$$l(\widehat{\mu}) = \max_{\mu \in \mathbb{R}} l(\mu)\,.$$

Zbog nezavisnosti je

$$f(x_1, \ldots, x_n; \mu) = \prod_{i=1}^{n} f_{\mathrm{N}(\mu, \sigma_i^2)}(x_i)\,,$$

pa je

$$l(\mu) = \sum_{i=1}^{n} \log(f_{\mathrm{N}(\mu,\sigma_i^2)}(x_i)) = \sum_{i=1}^{n} \log\left(\frac{1}{\sqrt{2\pi}\sigma_i}\right) - \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma_i^2}\,.$$

Sada vidimo da je

$$\widehat{\mu} = \arg\max_{\mu \in \mathbb{R}} l(\mu) = \arg\min_{\mu \in \mathbb{R}} \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma_i^2} \;=\; \mathrm{DZ} \;= \sum_{i=1}^{n} w_i x_i\,,$$

pri čemu su težine dane s

$$w_i = \frac{1/\sigma_i^2}{\sum_{j=1}^{n} 1/\sigma_j^2}\,, \quad i = 1, \ldots, n\,.$$

Uočimo da ovaj procjenitelj intuitivno ima smisla – veću težinu dajemo onim x_i-evima s manjom varijancom σ_i^2 jer će oni (u prosjeku) manje biti udaljeni od očekivanja μ^*. Ukoliko je $\sigma_i = \sigma$, za sve $i = 1, \ldots, n$, $\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$ – uočite da u ovom slučaju ne trebamo znati vrijednost varijance σ^2. □

Bez dokaza navodimo generalizaciju Teorema 5.21.

Propozicija 8.8. *Ako je (X, Y) neprekidan slučajni vektor te $g : \mathbb{R}^2 \to \mathbb{R}$ funkcija, vrijedi*

$$\mathbb{E}[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y)\, \mathrm{d}x\mathrm{d}y\,, \tag{8.6}$$

pod uvjetom da je g nenegativna ili integral na desnoj strani apsolutno konvergira.

Primjer 8.9. Neka su X i Y dvije nezavisne Unif$(0, 1)$ slučajne varijable. Odredimo očekivanu udaljenost između njih, tj. $\mathbb{E}\big[|X - Y|\big]$.

Rješenje. Budući da je (X, Y) neprekidan s gustoćom $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ koja je jednaka 1 za $(x, y) \in [0, 1]^2$, a 0 inače, imamo

$$\begin{aligned} \mathbb{E}\big[|X - Y|\big] &= \int_0^1 \int_0^1 |x - y|\, \mathrm{d}x\mathrm{d}y = \int_0^1 \int_y^1 (x - y)\, \mathrm{d}x\mathrm{d}y + \int_0^1 \int_0^y (y - x)\, \mathrm{d}x\mathrm{d}y \\ &= 2\int_0^1 \int_y^1 (x - y)\, \mathrm{d}x\mathrm{d}y = 2\int_0^1 (1/2 - y) - (y^2/2 - y^2)\mathrm{d}y = 2 \cdot \frac{1}{6} = \frac{1}{3}\,. \end{aligned}$$

□

[Sada smo u stanju pokazati da i za (zajednički) neprekidne slučajne varijable vrijede osnovna svojstva matematičkog očekivanja.]

Zadatak. Ako je (X, Y) neprekidan slučajan vektor, te $X, Y \geqslant 0$, koristeći (8.6) pokažite da vrijedi

(a) $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$;
(b) $\mathbb{E}[X] \leqslant \mathbb{E}[Y]$ ako je $X \leqslant Y$.
(c) $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$ ako su X i Y nezavisne.

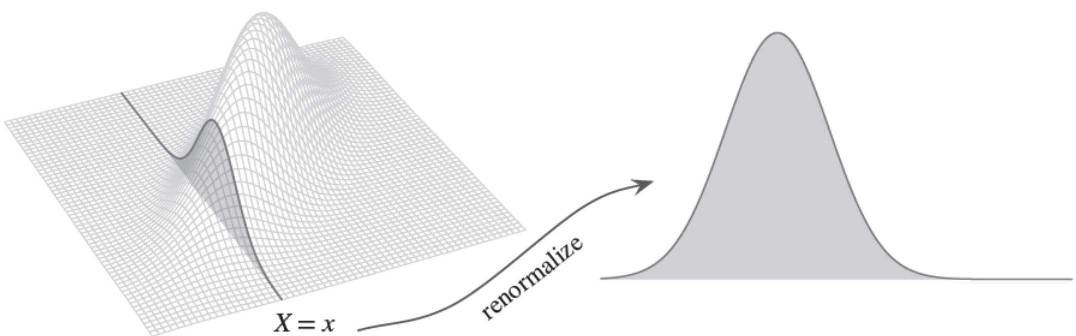
8.2. Uvjetovanje u neprekidnom slučaju

[Kod neprekidnog slučajnog vektora (X, Y), uvjetovanje na događaj $\{X = x\}$ je problematično jer imamo $\mathbb{P}(X = x) = 0$ za sve x. Taj problem probat ćemo riješiti tako što ćemo uvjetovanje raditi direktno s funkcijama gustoće.]

Definicija 8.10. Ako je (X, Y) neprekidan slučajan vektor te $x \in \mathbb{R}$ takav da vrijedi $f_X(x) > 0$, definiramo **uvjetnu funkciju gustoće** od Y uz dano $X = x$ sa

$$f_{Y|X=x}(y) := \frac{f_{X,Y}(x, y)}{f_X(x)}, \quad y \in \mathbb{R}. \tag{8.7}$$

[Lako se vidi da je $\int_{-\infty}^{\infty} f_{Y|X=x}(y)\,\mathrm{d}y = 1$ (vidi (8.4)), tako da je s (8.7) dobro definirana gustoća neke neprekidne slučajne varijable.]



Slika 17. Integral osjenčanog dijela (po y) na lijevoj slici je upravo vrijednost marginalne gustoće $f_X(x)$ iz (8.4). S druge strane, gustoću od Y uz dano $X = x$ dobijemo tako da funkciju $y \mapsto f_{X,Y}(x, y)$ normaliziramo tako da postane dobro definirana gustoća. Preuzeto iz [**BH19**].

Veza s klasičnim uvjetovanjem.* Budući da je $\mathbb{P}(X = x) = 0$, nismo mogli direktno uvjetovati na $X = x$. Ipak, gornja definicija motivirana je sljedećom idejom: za sve $A \subseteq \mathbb{R}^2$ definiramo

$$\mathbb{P}(Y \in A \mid X = x) := \lim_{\epsilon \to 0} \mathbb{P}(Y \in A \mid X \in (x - \epsilon, x + \epsilon)),$$

uz pretpostavku da taj limes postoji. Tada je za $b \in \mathbb{R}$,

$$\begin{aligned}
\mathbb{P}(Y \leqslant b \mid X = x) &= \lim_{\epsilon \to 0} \frac{\mathbb{P}((X,Y) \in (x-\epsilon, x+\epsilon) \times (-\infty, b]))}{\mathbb{P}(X \in (x-\epsilon, x+\epsilon))} \\
&= \lim_{\epsilon \to 0} \frac{\int_{x-\epsilon}^{x+\epsilon} \int_{-\infty}^{b} f_{X,Y}(s,y)\,\mathrm{d}y\mathrm{d}s}{\int_{x-\epsilon}^{x+\epsilon} f_X(s)\,\mathrm{d}s} \cdot \frac{1/2\epsilon}{1/2\epsilon} \\
&= \frac{\int_{-\infty}^{b} f_{X,Y}(x,y)\,\mathrm{d}y}{f_X(x)} = \int_{-\infty}^{b} f_{Y|X=x}(y)\,\mathrm{d}y\,.
\end{aligned}$$

Dakle, "uvjetno na $X = x$", Y je neprekidna s gustoćom $f_{Y|X=x}$. $\square$

Tehnička napomena. Ako je (X,Y) neprekidan slučajan vektor s gustoćom f, tada je nužno i

$$\widetilde{f}(x,y) := \begin{cases} f(x,y)\,, & f_X(x) > 0 \text{ i } f_Y(y) > 0\,, \\ 0\,, & f_X(x) = 0 \text{ ili } f_Y(y) = 0 \end{cases}$$

gustoća od (X,Y). [To slijedi budući da je $\mathbb{P}(f_X(X) > 0) = \mathbb{P}(f_Y(Y) > 0) = 1$, pa je za sve $a, b \in \mathbb{R}$,

$$\begin{aligned}
\mathbb{P}(X \leqslant a, Y \leqslant b) &= \mathbb{P}(X \leqslant a, Y \leqslant b, f_X(X) > 0, f_Y(Y) > 0) \\
&= \mathbb{P}((X,Y) \in \{(x,y) : x \leqslant a, Y \leqslant b, f_X(x) > 0, f_Y(y) > 0\}) \\
&= \int_{-\infty}^{a} \int_{-\infty}^{b} \widetilde{f}(x,y)\,\mathrm{d}y\mathrm{d}x\,. \ ]
\end{aligned}$$

Zbog ovoga možemo (i hoćemo) u nastavku pretpostaviti da za gustoću $f = f_{X,Y}$ vrijedi

$$(f_X(x) = 0 \text{ ili } f_Y(y) = 0) \ \Rightarrow \ f(x,y) = 0\,. \tag{8.8}$$

$\square$

Iz definicije (i prethodne napomene) odmah slijedi

(a) Varijable X i Y su nezavisne akko vrijedi

$$f_{Y|X=x}(y) = f_Y(y)\,,\ y \in \mathbb{R}\,,$$

za sve x t.d. je $f_X(x) > 0$.

(b) Zajedničku gustoću uvijek možemo dobiti iz formule

$$f_{X,Y}(x,y) = f_X(x) f_{Y|X=x}(y)\,,$$

za sve (x,y) t.d. je $f_X(x) > 0$ [u ostalim slučajevima imamo $f_{X,Y}(x,y) = 0$].

Primjer 8.11. Ako je (X,Y) uniforman slučajni vektor na $A = \{(x,y) : x^2 + y^2 \leqslant 1\}$, te $x \in (-1,1)$ fiksan. Iz Primjera 8.3 slijedi da za sve $|y| \leqslant \sqrt{1-x^2}$ imamo

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{\frac{1}{\pi}}{\frac{2}{\pi}\sqrt{1-x^2}} = \frac{1}{2\sqrt{1-x^2}}\,.$$

Dakle, uvjetno na $X = x$, Y ima uniformnu razdiobu na $[-\sqrt{1-x^2}, \sqrt{1-x^2}]$. Budući da $f_{Y|X=x}(y)$ ovisi o x, ovo daje alternativan dokaz zavisnosti između X i Y. Također, ovo daje metodu za simuliranje slučajnog vektora (X, Y):

1. Simuliraj $X = x$ iz gustoće f_X;
2. Simuliraj $Y = y$ iz uniformne razdiobe na $[-\sqrt{1-x^2}, \sqrt{1-x^2}]$;
3. Vrati (x, y).

□

Uvjetno očekivanje. Budući da je $f_{Y|X=x}$ (za $f_X(x) > 0$) dobro definirana (jednodimenzionalna) gustoća, za proizvoljnu funkciju $g : \mathbb{R}^2 \to \mathbb{R}$ prirodno je definirati

$$\mathbb{E}[g(X,Y) \mid X = x] := \int_{-\infty}^{\infty} g(x,y) f_{Y|X=x}(y)\, \mathrm{d}y\,,$$

ukoliko je g nenegativna ili gornji integral apsolutno konvergira; uz konvenciju $\mathbb{E}[g(X,Y) \mid X = x] := 0$ ukoliko je $f_X(x) = 0$. Uočimo, ako je za sve x takve da je $f_X(x) > 0$, Y_x neprekidna slučajna varijabla s gustoćom $f_{Y|X=x}$, zapravo je

$$\mathbb{E}[g(X,Y) \mid X = x] = \mathbb{E}[g(x, Y_x)]\,.$$

Iz (8.6) (i (8.8)) dobivamo **formulu potpunog očekivanja**

$$\mathbb{E}[g(X,Y)] = \int_{-\infty}^{\infty} \mathbb{E}[g(X,Y) \mid X = x] \cdot f_X(x)\, \mathrm{d}x\,. \tag{8.9}$$

[U praksi, distribuciju vektora (X, Y) često zadajemo tako da zadamo marginalnu gustoću od X te gustoću od Y uvjetno na $X = x$.]

Primjer 8.12. Štap duljine 1 prelomimo na uniformno odabranom mjestu X. Uvjetno na $X = x \in (0,1)$, dio štapa $[0, x]$ ponovno prelomimo na uniformno odabranom mjestu Y. Odredite $\mathbb{E}[Y]$ i $f_{X,Y}$. Ima li (X, Y) uniformnu razdiobu na $B := \{(x,y) \in [0,1]^2 : y \leqslant x\}$?

Rješenje. Po pretpostavci je $X \sim \mathrm{Unif}(0,1)$, a uvjetno na $X = x \in (0,1)$, Y ima $\mathrm{Unif}(0,x)$ razdiobu. Dakle,

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} \mathbb{E}[Y \mid X = x] \cdot f_X(x)\, \mathrm{d}x = \int_0^1 \mathbb{E}[Y \mid X = x] \cdot 1\, \mathrm{d}x = \int_0^1 \frac{x}{2}\, \mathrm{d}x = \mathbb{E}[X/2] = \frac{1}{4}\,.$$

Nadalje,

$$f_{X,Y}(x,y) = f_{Y|X=x}(y) f_X(x) = \begin{cases} \frac{1}{x} \cdot 1\,, & x \in [0,1],\ y \in [0,x]\,, \\ 0\,, & \text{inače.} \end{cases}$$

Specijalno, $f_{X,Y} \neq f_{\mathrm{Unif}(B)}$, pa (X, Y) nije uniforman na B. □

Primjer 8.13. Neka je $x \in [0,1]$ te $N_x := \min\{n \geqslant 1 : U_1 + \cdots + U_n > x\}$ pri čemu je $U_1, U_2, \ldots$ niz njd $\mathrm{Unif}(0,1)$ slučajnih varijabli. Odredite $m(x) := \mathbb{E}[N_x]$ [možete pretpostaviti da je m neprekidna na $[0,1]$].

Rješenje. Uvjetujemo na ishod $U_1 = u$ te razlikujemo dva slučaja. Ako je $u > x$ imamo $N_x = 1$, a ako je $u \leqslant x$, zbog pretpostavke da su U_i-evi njd "krećemo ispočetka" i gledamo koliko nam uniformnih treba da "preskočimo" $x - u$. Formalno, za sve $x \in [0,1]$ imamo

$$\begin{aligned} m(x) &= \int_0^1 \mathbb{E}[N_x \mid U_1 = u] f_{U_1}(u) \mathrm{d}u = [U_1, U_2, \dots \text{ su njd}] \\ &= \int_0^x (1 + m(x-u)) \mathrm{d}u + \int_x^1 1 \mathrm{d}u = 1 + \int_0^x m(x-u) \mathrm{d}u \\ &= [t = x - u] = 1 + \int_0^x m(t) \mathrm{d}t \,. \end{aligned}$$

Sada slijedi da je $m'(x) = m(x)$, $x \in (0,1)$ te je $m(0) = 1$, iz čega slijedi da je $m(x) = e^x$. □

8.2.1. Neprekidni miješani model. Pretpostavimo da je Y neprekidna slučajna varijabla s gustoćom f_Y na $(\Omega, \mathcal{F})$, te da za svaki $y \in \mathbb{R}$ t.d. $f_Y(y) > 0$ imamo zadanu vjerojatnost $\mathbb{P}_y$ (i pripadno očekivanje $\mathbb{E}_y$) na $(\Omega, \mathcal{F})$ koju interpetiramo i pišemo kao

$$\mathbb{P}_y(A) = \mathbb{P}(A \mid Y = y) \,.$$

Može se pokazati da je u tom slučaju funkcija $\mathbb{P} : \mathcal{F} \to [0,1]$ definirana s

$$\mathbb{P}(A) := \int_{-\infty}^{\infty} \mathbb{P}(A \mid Y = y) f_Y(y) \, \mathrm{d}y \,, \quad A \in \mathcal{F} \,, \tag{8.10}$$

dobro definirana vjerojatnost, te da za svaku slučajnu varijablu X [ne nužno samo neprekidnu] na $(\Omega, \mathcal{F})$ vrijedi [ako postoji $\mathbb{E}[X]$] formula potpunog očekivanja

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} \mathbb{E}[X \mid Y = y] \cdot f_Y(y) \, \mathrm{d}y \,, \tag{8.11}$$

pri čemu je $\mathbb{E}[X \mid Y = y] := \mathbb{E}_y[X]$. Kažemo da je $\mathbb{P}$ **neprekidna mješavina (engl. *mixture*) razdioba** $\mathbb{P}_y, y \in \mathbb{R}$, s težinama $f_Y(y), y \in \mathbb{R}$.

Primjer 8.14 (**Poissonova sa slučajnim parametrom**). Neka je Λ nenegativna neprekidna slučajna varijabla s gustoćom f, a X Poissonova slučajna varijabla s parametrom Λ, tj. za svaki $\lambda > 0$, uvjetno na $\Lambda = \lambda$, X ima $\mathrm{P}(\lambda)$ distribuciju. U ovom slučaju kažemo da X ima **miješanu razdiobu** (engl. *mixture distribution*). Odredimo $\mathbb{E}[X]$ i $\mathrm{Var}(X)$.

Rješenje. Prisjetimo se, za $Z \sim \mathrm{P}(\lambda)$, vrijedi $\mathbb{E}[Z] = \mathrm{Var}(Z) = \lambda$, te $\mathbb{E}[Z^2] = \mathrm{Var}(Z) + (\mathbb{E}[Z])^2 = \lambda + \lambda^2$. Sada imamo

$$\mathbb{E}[X] \overset{(8.11)}{=} \int_0^{\infty} \mathbb{E}[X \mid \Lambda = \lambda] f(\lambda) \, \mathrm{d}\lambda = \int_0^{\infty} \lambda f(\lambda) \, \mathrm{d}\lambda = \mathbb{E}[\Lambda] \,,$$

te

$$\begin{aligned} \mathbb{E}[X^2] &= [(8.11) \text{ za } X^2] = \int_0^{\infty} \mathbb{E}[X^2 \mid \Lambda = \lambda] f(\lambda) \, \mathrm{d}\lambda \\ &= \int_0^{\infty} (\lambda + \lambda^2) f(\lambda) \, \mathrm{d}\lambda = \mathbb{E}[\Lambda] + \mathbb{E}[\Lambda^2] \,. \end{aligned}$$

Dakle,[2]

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \mathbb{E}[\Lambda] + \mathbb{E}[\Lambda^2] - (\mathbb{E}[\Lambda])^2 = \mathbb{E}[\Lambda] + \mathrm{Var}(\Lambda)\,,$$

pa je specijalno $\mathrm{Var}(X) > \mathbb{E}[X]$.

[Dakle, uvođenje slučajnog parametra povećava varijabilnost. U praksi se Poissonova razdioba sa slučajnim parametrom koristi u statističkom modeliranju u slučajevima kada standardna Poissonova razdioba nije adekvatan model za dane podatke, tj. preciznije, kada je varijanca podataka veća od varijance Poissonovog modela – taj slučaj nazivamo *prekomjerna disperzija* (engl. *overdispersion*).] □

[Napomena.* Kada je Y neprekidna slučajna varijabla na proizvoljnom vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$, nije apriori jasno kako bi definirali $\mathbb{P}(A \mid Y = y)$ za proizvoljan događaj $A \in \mathcal{F}$, te općenitije $\mathbb{E}[X \mid Y = y]$ za proizvoljnu slučajnu varijablu X na $(\Omega, \mathcal{F}, \mathbb{P})$. Zapravo, ovo se ispostavlja kao netrivijalan teorijski problem, vidi npr. *Borel-Kolmogorovljev paradoks.* Poanta je da za fiksni $y \in \mathbb{R}$, $\mathbb{E}[X \mid Y = y]$ ne možemo definirati jednoznačno, ali da, uz neke dodatne pretpostavke na vjerojatnosni prostor, postoji familija vjerojatnosnih funkcija $\{\mathbb{P}_y, y \in \mathbb{R}\}$ na $(\Omega, \mathcal{F})$ takva da vrijedi formula potpunog očekivanja (8.11); ovo se u teoriji naziva *teorem disintegracije.* Slično kao kod definicije funkcije gustoće, izbor ove familije nije jedinstven, ali za bilo koje dvije takve familije $\{\mathbb{P}_y : y \in \mathbb{R}\}$ i $\{\mathbb{P}'_y : y \in \mathbb{R}\}$, za skup $N = \{y \in \mathbb{R} : \mathbb{P}_y \neq \mathbb{P}'_y\}$ vrijedi $\mathbb{P}(Y \in N) = 0$.]

8.2.2. Bayesova formula.

Propozicija 8.15. *Za slučajni vektor (X, Y) imamo sljedeće oblike Bayesove formule.*

(a) ako su X i Y diskretne,

$$\mathbb{P}(Y = y \mid X = x) = \frac{\mathbb{P}(X = x \mid Y = y)\mathbb{P}(Y = y)}{\mathbb{P}(X = x)}\,,$$

u slučaju kada je $\mathbb{P}(X = x), \mathbb{P}(Y = y) > 0$.

(b) ako je (X, Y) neprekidan,

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{f_{X|Y=y}(x) f_Y(y)}{f_X(x)}\,, \quad f_X(x), f_Y(y) > 0\,.$$

(c) Ako je X diskretan, a Y neprekidna,

$$f_Y(y \mid X = x) = \frac{\mathbb{P}(X = x \mid Y = y) f_Y(y)}{\mathbb{P}(X = x)}\,, \tag{8.12}$$

te

$$\mathbb{P}(X = x \mid Y = y) = \frac{f_Y(y \mid X = x)\mathbb{P}(X = x)}{f_Y(y)}\,, \tag{8.13}$$

ukoliko je $f_Y(y) > 0$ i $\mathbb{P}(X = x) > 0$.

[2] Jesmo li mogli jednostavno računati $\mathrm{Var}(X) = \int_0^\infty \mathrm{Var}(X \mid \Lambda = \lambda) f(\lambda)\,\mathrm{d}\lambda$? Zašto?

Dokaz. Jedino treba pokazati dio (c), a to ćemo učiniti uz pretpostavku da je vjerojatnost $\mathbb{P}$ konstruirana kao u miješanom modelu (8.10). [Formula vrijedi i općenito, dakle kada su diskretna varijabla X i neprekidna varijabla Y definirane na proizvoljnom vjerojatnosnom prostoru.]

Za sve $z \in \mathbb{R}$ imamo

$$\mathbb{P}(Y \leqslant z \mid X = x) = \frac{\mathbb{P}(Y \leqslant z, X = x)}{\mathbb{P}(X = x)}\,,$$

te nadalje vrijedi

$$\mathbb{P}(Y \leqslant z, X = x) = \int_{-\infty}^{\infty} \mathbb{P}(Y \leqslant z, X = x \mid Y = y) f_Y(y)\,\mathrm{d}y = \int_{-\infty}^{z} \mathbb{P}(X = x \mid Y = y) f_Y(y)\,\mathrm{d}y\,,$$

pri čemu zadnja jednakost slijedi iz

$$\mathbb{P}(Y \leqslant z, X = x \mid Y = y) = \begin{cases} \mathbb{P}(X = x \mid Y = y)\,, & \text{ako je } y \leqslant z \\ 0\,, & \text{ako je } y > z. \end{cases}$$

Sada imamo

$$\mathbb{P}(Y \leqslant z \mid X = x) = \int_{-\infty}^{z} \frac{\mathbb{P}(X = x \mid Y = y) f_Y(y)}{\mathbb{P}(X = x)}\,\mathrm{d}y\,,\ z \in \mathbb{R}\,.$$

Dakle, po definiciji imamo da je, uvjetno na $X = x$, Y neprekidna s gustoćom danom u (8.12). Formula (8.13) slijedi direktno iz (8.12). □

Primjer 8.16 (**Diskretna miješana razdioba**). Na tržištu postoje dva tipa žarulja – trajanje svake žarulje je slučajno te ima $\mathrm{Exp}(\lambda_i)$ razdiobu ako je žarulja tipa i, $i = 1, 2$. Udio žarulja tipa 1 na tržištu je $p \in (0, 1)$. Neka je T trajanje slučajno odabrane žarulje sa tržišta, a $I \in \{1, 2\}$ indikator kojeg je tipa odabrana žarulja.

(a) Odredite funkciju distribucije i funkciju gustoće od T.

(b) Odredite razdiobu od I uvjetno na $T = t$ za $t > 0$. Što se događa kada $t \to \infty$?

Rješenje. Ključno je primijetiti da je

$$(T \mid I = i) \sim \mathrm{Exp}(\lambda_i)\,,\quad i = 1, 2\,.$$

Kažemo da je T (*diskretna*) mješavina dvije eksponencijalne razdiobe.

(a) Za svaki $t > 0$, koristeći FPV dobivamo

$$\begin{aligned} F_T(t) &= \mathbb{P}(T \leqslant t) = \mathbb{P}(T \leqslant t \mid I = 1)\mathbb{P}(I = 1) + \mathbb{P}(T \leqslant t \mid I = 2)\mathbb{P}(I = 2) \\ &= (1 - e^{-\lambda_1 t}) \cdot p + (1 - e^{-\lambda_2 t}) \cdot q = 1 - pe^{-\lambda_1 t} - qe^{-\lambda_2 t}\,. \end{aligned}$$

[Za $t \leqslant 0$, očito je $F_T(t) = 0$. Uočimo da je F_T neprekidna na $\mathbb{R}$, te (neprekidno) diferencijabilna svugdje osim u 0 pa je dakle T neprekidna slučajna varijabla.] Gustoću dobivamo deriviranjem,

$$f_T(t) = F_T'(t) = p\lambda_1 e^{-\lambda_1 t} + q\lambda_2 e^{-\lambda_2 t}\ [= p f_{\mathrm{Exp}(\lambda_1)}(t) + q f_{\mathrm{Exp}(\lambda_2)}(t)]\,,\quad t > 0\,.$$

[Uz naravno $f_T(t) = 0$ za $t \leqslant 0$.]

(b) Koristeći Bayesovu formulu (8.13) dobivamo

$$\begin{aligned}\mathbb{P}(I=1 \mid T=t) &= \frac{f_T(t \mid I=1)\mathbb{P}(I=1)}{f_T(t)} = \frac{\lambda_1 e^{-\lambda_1 t} p}{p\lambda_1 e^{-\lambda_1 t} + q\lambda_2 e^{-\lambda_2 t}} \\ &= \frac{\lambda_1 p}{p\lambda_1 + q\lambda_2 e^{-(\lambda_2-\lambda_1)t}}\,, \quad t > 0\,.\end{aligned}$$

Dakle, ako je $\lambda_1 = \lambda_2$ imamo $\mathbb{P}(I=1 \mid T=t) = p$, a u suprotnom

$$\lim_{t\to\infty} \mathbb{P}(I=1 \mid T=t) = \begin{cases} 1\,, & \text{ako je } \lambda_1 < \lambda_2 \\ 0\,, & \text{ako je } \lambda_1 > \lambda_2. \end{cases}$$

[Je li rezultat intuitivno jasan? Da – slučaju $\lambda_1 < \lambda_2$, za velike t vrijedi

$$\mathbb{P}(T \geqslant t \mid I=1) = e^{-\lambda_1 t} \gg e^{-\lambda_2 t} = \mathbb{P}(T \geqslant t \mid I=2)\,.$$

Kažemo da $\mathrm{Exp}(\lambda_1)$ radioba ima *teži rep* od $\mathrm{Exp}(\lambda_2)$ razdiobe.]

□

8.3. 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

$$f(t) = \frac{1}{\beta(a,b)} t^{a-1}(1-t)^{b-1}\,, \quad t \in (0,1)\,, \tag{8.14}$$

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 18.] Može se pokazati da za $X \sim \mathrm{Beta}(a,b)$ vrijedi

$$\mathbb{E}[X] = \frac{a}{a+b}.$$

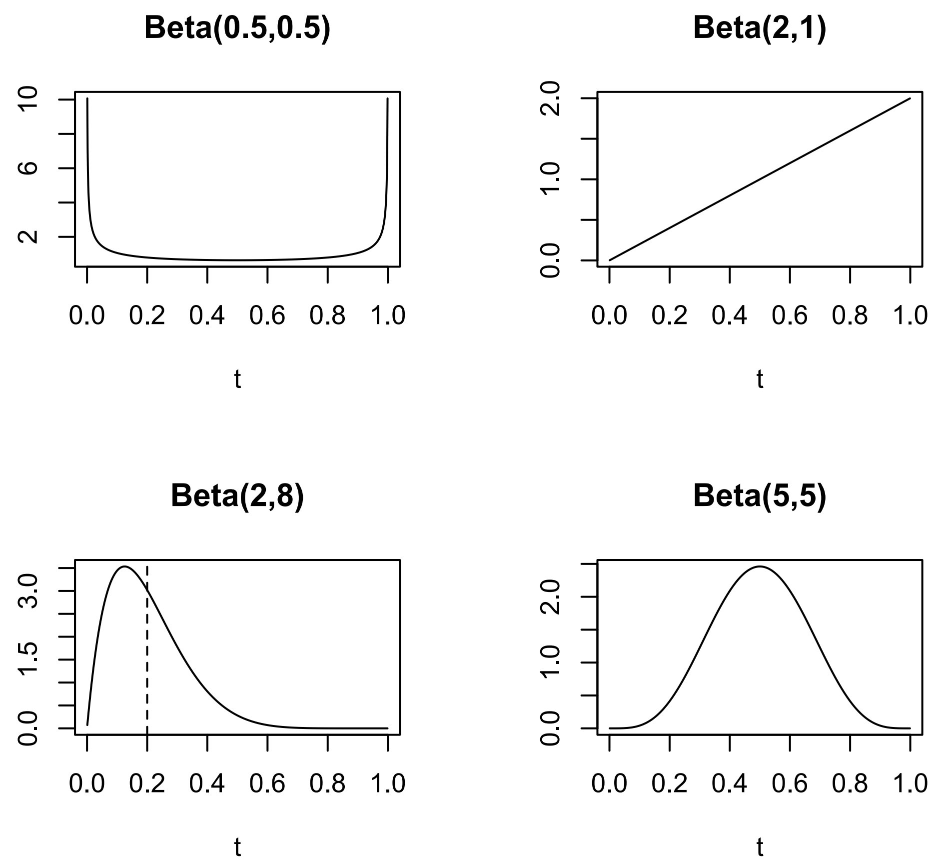
Primjer 8.17 (**Bayesovska statistika i Beta-binomna razdioba**). [Bacamo novčić n puta te želimo procijeniti vjerojatnost $p \in [0,1]$ za pojavu pisma, na temelju ukupnog broja pisama koji su pali N_n.]

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)[3], te
- uvjetno na $P = p$, N_n ima $\mathrm{B}(n,p)$ razdiobu.[4]

[3] Apriorna razdioba predstavlja našu početnu informaciju o problemu – npr. ako ništa ne znamo, možemo staviti da je to $\mathrm{Beta}(1,1) = \mathrm{Unif}(0,1)$ razdoiba.
[4] Dakle, N_n je neprekidna mješavina binomnih razdioba – kažemo da ima beta-binomnu razdiobu.



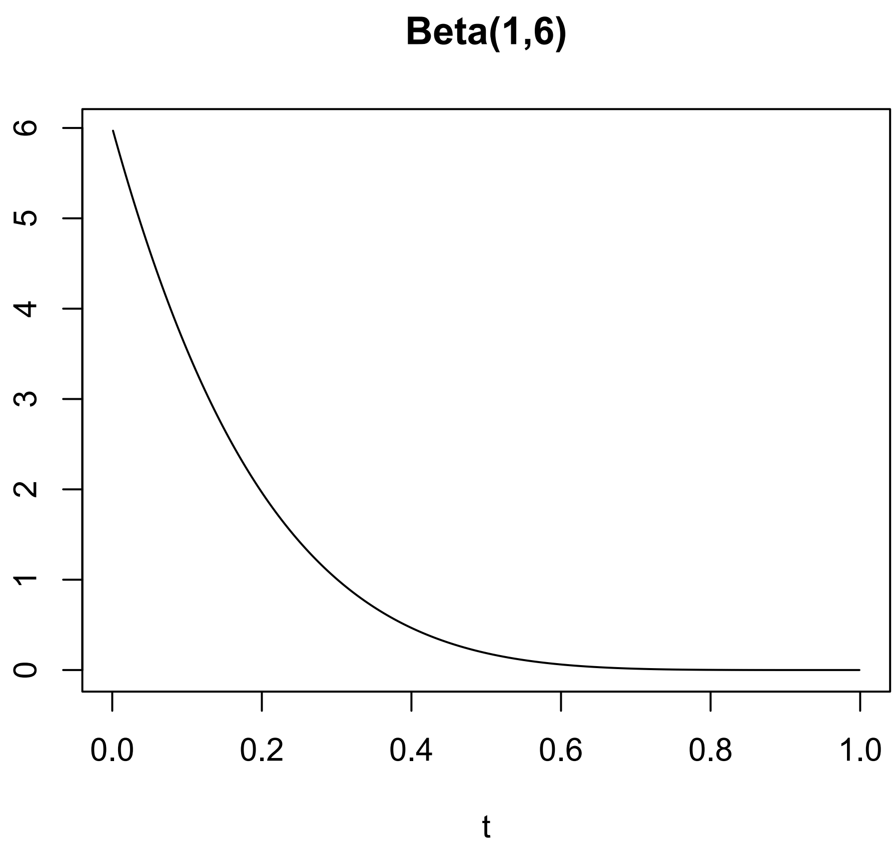
SLIKA 18. Gustoća Beta(a, b) za nekoliko kombinacija parametara $a, b > 0$. U slučaju $a = 2, b = 8$, vertikalnom linijom je naznačeno očekivanje $a/(a+b) = 0.2$.

Ako je $N_n = k$, koristeći Bayesovu formulu dobivamo da je tzv. *aposteriorna* gustoća za P jednaka

$$f_P(p \mid N_n = k) = \frac{\mathbb{P}(N_n = k \mid P = p) f_P(p)}{\mathbb{P}(N_n = k)} = \frac{\binom{n}{k} p^k (1-p)^{n-k} \cdot \frac{1}{\beta(a,b)} p^{a-1} (1-p)^{b-1}}{\mathbb{P}(N_n = k)}$$
$$= c \cdot p^{a+k-1} (1-p)^{b+n-k-1}, \quad p \in (0,1),$$

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![5] Na primjer, ako je $P \sim$ Beta$(1, 1) =$ 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 $\frac{N_n}{n} = 0$.

[5] 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.



SLIKA 19. Aposteriorna razdioba za P u Primjeru 8.17, i to u slučaju uniformne apriorne razdiobe, $n = 5$ i $N_n = 0$. Dakle, u Bayesovskom pristupu dobivamo cijelu razdiobu za parametra koji procjenjujemo!

8.4. Kovarijacijska matrica

[Kako bi mogli definirati višedimenzionalnu normalnu razdiobu, prvo se bavimo pojmom kovarijacijske matrice općenitog slučajnog vektora.]

Ako je $X = (X_1, \ldots, X_p)$ p-dimenzionalan slučajni vektor [ne nužno samo diskretan ili neprekidan], definiramo

(i) njegovo očekivanje s $\mathbb{E}[X] := (\mathbb{E}[X_1], \ldots, \mathbb{E}[X_p]) \in \mathbb{R}^p$;

(ii) njegovu **kovarijacijsku matricu** $\mathrm{Cov}(X) = \Sigma_X = (\sigma_{i,j} : i, j = 1, \ldots, p)$ kao $p \times p$ matricu s elementima

$$\sigma_{i,j} := \mathrm{Cov}(X_i, X_j)\,.$$

Specijalno, dijagonalni elementi su

$$\sigma_{i,i} = \mathrm{Var}(X_i)\,.$$

[Formalno, očekivanje i kovarijacijsku matricu definiramo jedino u slučaju kada su svi njihovi elementi dobro definirani.]

PRIMJER 8.18 (**njd slučaj**). Ako su $X_1, \dots, X_p$ njd slučajne varijable sa zajedničkom varijancom $\sigma^2 < \infty$, tada je

$$\mathrm{Cov}(X_1, \dots, X_p) = \sigma^2 I_p \,,$$

gdje je I_p jedinična $p \times p$ matrica.

Konvencija. U nastavku p-dimenzionalne vektore (slučajne ili determinističke) shvaćamo kao vektor-stupce, tj. matrice dimenzija $p \times 1$. Na primjer, ako je $X = (X_1, \dots, X_p)$ slučajan vektor te $A = (a_{i,j})_{i,j} \in \mathbb{R}^{k \times p}$ (deterministička) matrica, AX je k-dimenzionalan slučajan vektor čiji je i-ti element jednak

$$(AX)_i = \sum_{j=1}^{p} a_{i,j} X_j \,, \quad i = 1, \dots, k \,.$$

PROPOZICIJA 8.19. *Neka su X i Y p-dimenzionalni slučajni vektori, te $A = (a_{i,j})_{i,j} \in \mathbb{R}^{k \times p}$ matrica. Tada vrijedi*

(a) $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$;

(b) $\mathbb{E}[AX] = A\mathbb{E}[X]$;

(c) $\Sigma_X = \mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^\tau]$;

(d) $\mathrm{Cov}(AX) = A\Sigma_X A^\tau$.

Specijalno, ako je $a = (a_1, \dots, a_p)$ vektor, imamo

$$\mathbb{E}[a^\tau X] = a^\tau \mathbb{E}[X] \,, \quad \mathrm{Var}(a^\tau X) = a^\tau \Sigma_X a = \sum_{i,j=1}^{p} a_i a_j \mathrm{Cov}(X_i, X_j) \,. \tag{8.15}$$

PRIMJER 8.20. Kao specijalan slučaj uz $a := (1, \dots, 1)$, (8.15) postaje

$$\mathrm{Var}(X_1 + \cdots + X_p) = \sum_{i=1}^{p} \mathrm{Var}(X_i) + 2 \sum_{i<j} \mathrm{Cov}(X_i, X_j) \,.$$

DOKAZ. (a) Slijedi direktno iz definicije te linearnosti očekivanja.

(b) Za sve $i = 1, \dots, k$, iz linearnosti očekivanja slijedi

$$\mathbb{E}[(AX)_i] = \mathbb{E}[\textstyle\sum_{j=1}^{p} a_{i,j} X_j] = \sum_{j=1}^{p} a_{i,j} \mathbb{E}[X_j] = (A\mathbb{E}[X])_i \,.$$

(c) Na mjestu (i, j) matrice $\mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^\tau]$ nalazi se

$$\mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])] = \mathrm{Cov}(X_i, X_j) = \sigma_{i,j} \,.$$

(d) Za sve $i, j = 1, \dots, p$ koristeći kolinearnost kovarijance dobivamo

$$\begin{aligned} \mathrm{Cov}((AX)_i, (AX)_j) &= \mathrm{Cov}(\textstyle\sum_{k=1}^{p} a_{i,k} X_k, \sum_{k'=1}^{p} a_{i,k'} X_{k'}) \\ &= \sum_{k=1}^{p} a_{i,k} \mathrm{Cov}(X_k, \textstyle\sum_{k'=1}^{p} a_{i,k'} X_{k'}) \\ &= \sum_{k,k'=1}^{p} a_{i,k} a_{i,k'} \sigma_{k,k'} \overset{DZ}{=} (A\Sigma_X A)_{i,j} \,. \end{aligned}$$

□

PROPOZICIJA 8.21. *Svaka kovarijacijska matrica* $\Sigma_X = (\sigma_{i,j} : i, j = 1, \ldots, p)$ *je simetrična i pozitivno semidefinitna.*

DOKAZ. Simetričnost slijedi iz simetričnosti kovarijance, a iz (8.15) za $a \in \mathbb{R}^{p\times 1}$ proizvoljan slijedi

$$a^\tau \Sigma_X a = \mathrm{Var}(a^\tau X) \geqslant 0\,.$$

□

[Vrijedi i obrat: za proizvoljnu simetričnu i pozitivno semidefinitnu matricu Σ postoji slučajan vektor X takav da je $\mathrm{Cov}(X) = \Sigma$.]

8.5. Višedimenzionalna normalna razdioba

[Višedimenzionalna normalna razdioba generalizacije je jednodimenzionalne normalne razdiobe, te je najčešće korišten model u višedimenzionalnoj statistici, tj. kada modeliramo podatke $x_1, \ldots, x_n$ gdje je svaki x_i p-dimenzionalan vektor, npr. x_i može sadržavati razne karakteristike i-te osobe. Definicija je motivirana generalizacijom CGT-a koji kaže da se ta razdioba jako dobro aproksimira razdiobu sume velikog broja njd slučajnih vektora za koje je dobro definirana kovarijacijska matrica. Cilj ovog poglavlja je samo definirati ovu razdiobu te dati neka njezina osnovna svojstva.]

Neka su $Z_1, \ldots, Z_p$ njd $\mathrm{N}(0,1)$ slučajne varijable. Tada je vektor $Z = (Z_1, \ldots, Z_p)$ neprekidan s gustoćom

$$\begin{aligned} g(z_1, \ldots, z_n) &= \prod_{j=1}^{p} \varphi(z_j) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}\sum_{j=1}^{p} z_j^2} \\ &= \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2} z^\tau z} = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}\|z\|^2}\,, \quad z = (z_1, \ldots, z_p) \in \mathbb{R}^p\,. \end{aligned} \tag{8.16}$$

gdje je $\|z\|$ (Euklidska) norma vektora z. Kažemo da Z ima **standardnu višedimenzionalnu normalnu razdiobu**.

[Kao i u jednodimenzionalnom slučaju, općenitu normalnu razdiobu možemo dobiti linearnim transformacijama standardnog normalnog slučajnog vektora.]

Ako definiramo slučajan vektor

$$X = (X_1, \ldots, X_p) := \mu + AZ\,, \tag{8.17}$$

pri čemu je $\mu \in \mathbb{R}$, a A invertibilna $p \times p$ matrica, Propozicija 8.19 povlači da je $\mathbb{E}[X] = \mu + A\mathbb{E}[Z] = \mu$, te je kovarijacijska matrica jednaka

$$\mathrm{Cov}(X) = A\Sigma_Z A^\tau = [\Sigma_Z = I_p] = AA^\tau =: \Sigma\,.$$

Budući da je A punog ranga, Σ je pozitivno definitna.[6] Nadalje, budući da je $Z = A^{-1}(X-\mu)$, koristeći višedimenzionalni teorem o zamjeni varijabli može se pokazati da je X neprekidan s gustoćom

$$\begin{aligned} f(x_1,\dots,x_n) &= \frac{1}{(2\pi)^{n/2}} \cdot \det(A^{-1}) e^{-\frac{1}{2}(A^{-1}(x-\mu))^\tau (A^{-1}(x-\mu))} \\ &= \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} e^{-\frac{1}{2}(x-\mu)^\tau \Sigma^{-1}(x-\mu)}, \; x=(x_1,\dots,x_n)\in\mathbb{R}^p. \end{aligned} \tag{8.18}$$

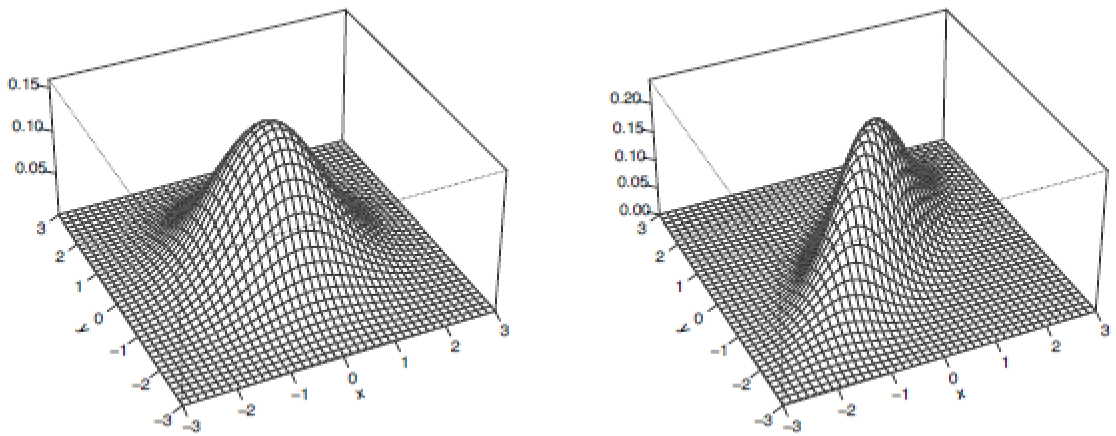
Definicija 8.22. Neka je $\mu \in \mathbb{R}^p$ te Σ pozitivno definitna $p\times p$ matrica. Kažemo da p-dimenzionalan slučajni vektor $X=(X_1,\dots,X_p)$ ima **višedimenzionalnu normalnu razdiobu** s očekivanjem μ i kovarijacijskom matricom Σ (oznaka $X \sim \mathrm{MVN}(\mu,\Sigma)$)[7] ako je neprekidan s gustoćom (8.18).

Uočimo, gustoća vektora $X=(X_1,\dots,X_p)\sim \mathrm{MVN}(0,I_p)$ svodi se na gustoću (8.16), tj. X je standarni normalni vektor [dakle, $X_1,\dots,X_p$ su njd standardne normalne varijable].

Primjer 8.23 (**Dvodimenzionalna normalna**). U slučaju $p=2$, za $(X_1,X_2)\sim \mathrm{MVN}(\mu,\Sigma)$ imamo $\mu_j = \mathbb{E}[X_j]$, a kovarijacijska matrica je oblika

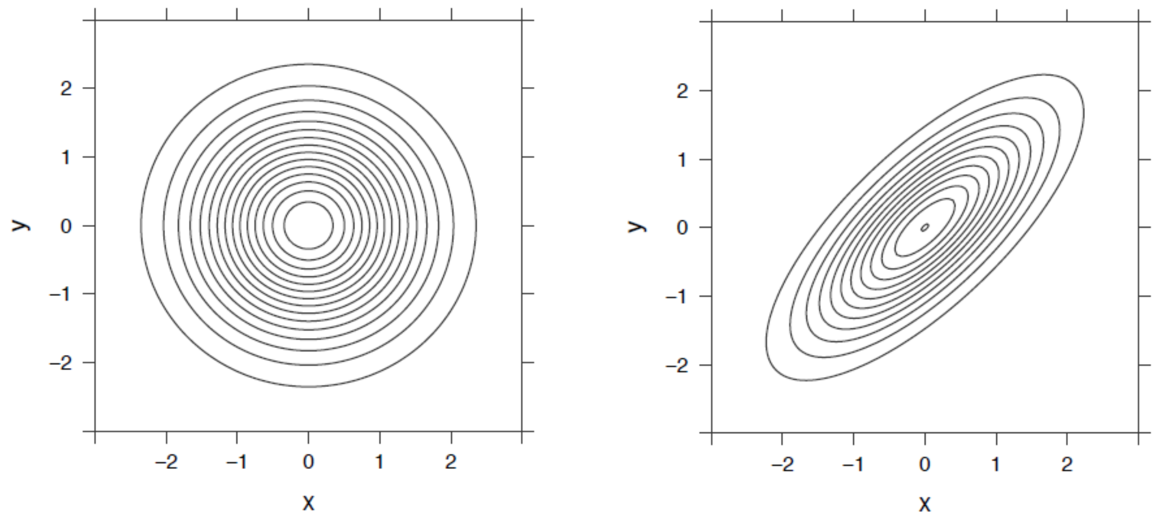
$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

gdje je $\sigma_j^2 = \mathrm{Var}(X_j) > 0$, te $\rho = \rho(X_1,X_2)\in(-1,1)$[8] korelacija između X_1 i X_2; vidi Slike 20 i 21.



Slika 20. Grafovi gustoće dvodimenzionalne normalne razdiobe pri čemu je $\mu_1=\mu_2=0$ te $\sigma_1=\sigma_2=1$, a korelacije $\rho=0$ (lijevo) i $\rho=0.75$ (desno). Na lijevoj slici dakle imamo gustoću standardne normalne razdiobe iz (8.16). Preuzeto iz [**BH19**].

[6] To slijedi jer za svaki $x\neq 0\in\mathbb{R}^p$, imamo $Ax\neq 0$, pa je i $x^\tau\Sigma x = (Ax)^\tau Ax = \|Ax\|^2 > 0$.
[7] MVN dolazi od *MultiVariate Normal*.
[8] Isključujemo slučaj $|\rho|=1$ jer tada imamo singularnu (dakle, samo pozitivno semidefinitnu) matricu Σ (stupci su linearno zavisni); vidi Napomenu 8.28 kasnije.



Slika 21. Alternativan prikaz gustoća sa Slike 20 koristeći *konture* konstantne gustoće. Ako je $f : \mathbb{R}^2 \to [0, \infty)$ gustoća, gornje linije (*konture*) su skupovi $\{(x, y) : f(x, y) = c\}$ za nekoliko različitih izbora konstante $c > 0$. Drugim riječima, konture dobijemo tako da graf funkcije f presječemo s ravninom $z = c$. Konture gustoće standardne dvodimenzionalne normalne razdiobe (lijevo) su kružnice oko ishodišta jer $f(x, y)$ ovisi samo o normi $\|(x, y)\|$ (tj. udaljenosti od ishodišta). Općenito, konture konstantne gustoće su rotirane elipse (vidi desno). Preuzeto iz [**BH19**].

[**Napomena.** Ako je $X \sim \mathrm{MVN}(\mu, \Sigma)$ uvijek se može naći invertibilna $p \times p$ matrica A takva da je $\Sigma = AA^\tau$ (npr. koristeći faktorizaciju Choleskog), s tim da izbor matrice A nije uvijek jedinstven. Nadalje, slično kao gore pokaže da se tada $Z = (Z_1, \dots, Z_p) := A^{-1}(X - \mu)$ sastoji od njd $\mathrm{N}(0, 1)$ slučajnih varijabli, te po konstrukciji vrijedi $X = \mu + AZ$. Dakle, uvijek možemo pretpostaviti da je X konstruiran kao u (8.17).]

Primjer 8.24 (**Simulacija dvodimenzionalne normalne razdiobe**). Pretpostavimo da imamo dvije nezavisne standardne normalne slučajne varijable V i W. Kako iz njih možemo dobiti MVN vektor (X, Y) tako da su X i Y standarne normalne, a $\rho(X, Y) = \rho \in (-1, 1)$?

Rješenje. Dovoljno je naći matricu

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

takvu da je

$$AA^\tau = A = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

te staviti

$$\begin{bmatrix} X \\ Y \end{bmatrix} := A \begin{bmatrix} V \\ W \end{bmatrix}.$$

Drugim riječima,

$$X := aV + bW$$
$$Y := cV + dW$$

pri čemu koeficijenti moraju zadovoljavati

$$\mathrm{Var}(X) = a^2 + b^2 = 1$$
$$\mathrm{Var}(Y) = c^2 + b^2 = 1$$
$$\rho(X,Y) = \mathrm{Cov}(X,Y) = ac + bd = \rho\,.$$

Imamo tri jednadžbe s 4 nepoznanice, ali nama je dovoljno samo jedno rješenje, pa možemo staviti $b = 0$. Rješenje je tada $a = 1, c = \rho, d = \sqrt{1-\rho^2}$. Dakle,

$$X := V$$
$$Y := \rho V + \sqrt{1-\rho^2}W$$

daje vektor s traženom distribucijom. □

Propozicija 8.25 (**Linearne kombinacije su normalne**). *Ako je* $X \sim \mathrm{MVN}(\mu, \Sigma)$*, za svaki* $b \in \mathbb{R}^p$ $(b \neq 0)$*, vrijedi*

$$b^\tau X = b_1 X_1 + \cdots + b_p X_p \sim \mathrm{N}\left(b^\tau \mu,\, b^\tau \Sigma b\right)\,.$$

Dokaz. Neka je $X = \mu + AZ$ kao u (8.17). Tada je

$$b^\tau X = b^\tau \mu + (b^\tau A) Z\,.$$

Budući da je $b^\tau A$ $1 \times p$ matrica, $(b^\tau A)Z$ je linearna kombinacija *nezavisnih* normalnih varijabli. Slijedi da je $(b^\tau A)Z$ (pa i onda i $b^\tau X$) također ima normalnu distribuciju. Izrazi za $\mathbb{E}[b^\tau X]$ i $\mathrm{Cov}(b^\tau X)$ slijede direktno iz Propozicije 8.19.

□

Korolar 8.26 (**Marginalne razdiobe su normalne**). *Ako je* $X = (X_1, \ldots, X_p) \sim \mathrm{MVN}(\mu, \Sigma)$*, vrijedi*

$$X_j \sim \mathrm{N}(\mu_j, \Sigma_{j,j})\,, \quad \forall j = 1, \ldots, p\,.$$

Dokaz. U prethodnoj propoziciji uzmi $b = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^p$ gdje je jedinica na j-tom mjestu. □

Bez dokaza navodimo generalizaciju Propozicije 8.25.

Propozicija 8.27 (**Afine transformacije su višedimenzionalne normalne**). *Neka je* $X \sim \mathrm{MVN}(\mu, \Sigma)$*, te* $B \in \mathbb{R}^{k \times p}$ *matrica ranga* k *uz* $k \leqslant p$*. Tada za* k*-dimenzionalan vektor* BX *vrijedi*

$$BX \sim \mathrm{MVN}(B\mu, B\Sigma B^\tau)\,.$$

Napomena 8.28 (**Singularna kovarijacijska matrica**). U prethodnom rezultatu bilo je bitno da je B ranga k. Na primjer, ako je $(X_1, X_2) \sim \mathrm{MVN}(0, I_2)$ i $B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ (dakle, $k = p = 2$ te je B ranga 1), tada je $BX = (X_1, X_1)^\tau$, što ne može biti neprekidan slučajni vektor.

[U ovom slučaju je kovarijacijska matrica

$$BI_2B^\tau = BB^\tau = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

singularna. Općenito se višedimenzionalna normalna razdioba može definirati i u tom slučaju koristeći istu konstrukciju kao u (8.17) uz ne nužno invertibilne matrice A (kao što je to slučaj u ovom primjeru). U tom slučaju nemamo neprekidan slučajan vektor već on poprima vrijednosti na nekom afinom potprostoru od $\mathbb{R}^p$ (u ovom primjeru to je pravac u $\mathbb{R}^2$). Ipak sva svojstva normalne razdiobe iz ovog poglavlja zapravo vrijede i u tom općenitom slučaju, uz konvenciju da konstantu $\mu \in \mathbb{R}$ poistovjećujemo s $\mathrm{N}(\mu, 0)$ razdiobom.]

[Općenito smo vidjeli da nekoreliranost ne povlači nezavisnost. Ipak, kada znamo da varijable zajednički imaju višedimenzionalnu normalnu razdiobu, za nezavisnost je dovoljno provjeriti samo da su varijable u parovima nekorelirane. Drugim riječima, kod normalnih vektora, zavisnost je u potpunosti određena samo s korelacijama.]

Propozicija 8.29 (**Nekoreliranost povlači nezavisnost**). *Ako je* $X = (X_1, \dots, X_p) \sim \mathrm{MVN}(\mu, \Sigma)$ *te je* $\mathrm{Cov}(X_i, X_j) = 0$*, za sve* $i \neq j$*, tj.*

$$\Sigma = \mathrm{diag}(\sigma_1^2, \dots, \sigma_p^2)\,,$$

varijable $X_1, \dots, X_p$ *su nužno nezavisne.*

Dokaz. Imamo da je $\Sigma^{-1} = \mathrm{diag}(\sigma_1^{-2}, \dots, \sigma_p^{-2})$[9] pa je gustoća od X oblika

$$f(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2} \cdot (\sigma_1^2 \cdots \sigma_p^2)^{1/2}} e^{-\frac{1}{2}\sum_{j=1}^p x_j^2/\sigma_j^2} = \prod_{j=1}^p \frac{1}{(2\pi)^{1/2}\sigma_j} e^{-\frac{x_j^2}{2\sigma_j^2}} = \prod_{j=1}^p f_{X_j}(x_j)\,,$$

za sve $(x_1, \dots, x_p) \in \mathbb{R}^p$. □

Primjer 8.30. Neka su X, Y nezavisne standardne normalne varijable. Pokažimo da su $X+Y$ i $X-Y$ nezavisne varijable.

Rješenje. Imamo da je $(X, Y) \sim \mathrm{MVN}(0, I_2)$, a

$$\begin{bmatrix} X+Y \\ X-Y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \end{bmatrix}.$$

[9] Pozitivna definitnost od Σ povlači da je $\sigma_j > 0$ za sve j.

Budući da je gornja transformacija regularna [stupci su očito linearno nezavisni], vektor $(X + Y, X - Y)$ ima dvodimenzionalnu normalnu razdiobu, te je

$$\begin{aligned}\mathrm{Cov}(X+Y, X-Y) &= \mathrm{Cov}(X, X-Y) + \mathrm{Cov}(Y, X-Y) \\ &= \mathrm{Var}(X) - \mathrm{Cov}(X,Y) + \mathrm{Cov}(Y,X) - \mathrm{Var}(Y) = 0\end{aligned}$$

jer $\mathrm{Var}(X) = 1 = \mathrm{Var}(Y)$ te $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$. Tvrdnja sada slijedi iz prethodne propozicije. □

Primjer 8.31 (**Analiza glavnih komponenti***). [U praksi često želimo analizirati višedimenzionalne podatke, npr. za svaku osobu imamo skup od $p \in \mathbb{N}$ različitih karakteristika kao što su visina, težina, dob, itd. U modernim primjenama dimenzija p može biti jako velika što značajno otežava analizu – npr. kako uopće vizualizirati podatke? Iz tog razloga htjeli bismo naše originalne podatke zamijeniti s podacima koji su puno manje dimenzije. Ovo nazivamo *redukcijom dimenzije* problema, a jedna od najpopularnijih metoda za tu svrhu je tzv. **analiza glavnih komponenti** (engl. **principal component analysis**, tj. **PCA**).]

Neka je $X = (X_1, \dots, X_p) \in \mathbb{R}^p$ slučajni vektor s kovarijacijskom matricom Σ [X predstavlja model za naše podatke]. Za proizvoljan jedinični vektor $v \in \mathbb{R}^p$, varijabla $v^\tau X \in \mathbb{R}$ je koeficijent projekcije vektora X na vektor v. Za $i = 1, 2, \dots, p$, definiramo i-tu **glavnu komponentu** (engl. *principal component*) od X kao projekciju $Y_i = v_i^\tau X$ takvu da smjer v_i zadovoljava

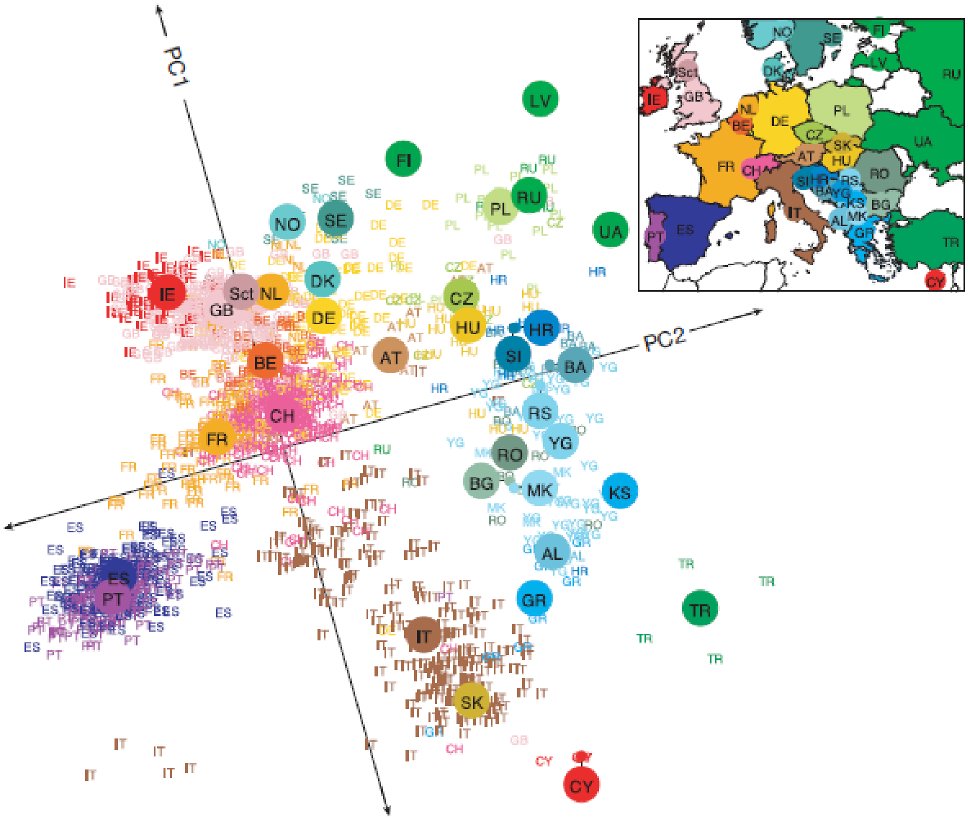
$$v_i = \arg \max_{v \in \mathbb{R}^p, \|v\|=1} \mathrm{Var}(v^\tau X)\ , \quad \text{uz uvjet } v_k^\tau v = 0 \text{ za } k = 1, \dots, i-1\,.$$

[Drugim riječima, v_1 je smjer u kojem projekcija $v^\tau X$ ima najveću varijancuu, v_2 je smjer okomit na v_1 u kojem $v_2^\tau X$ ima najveću varijancu, itd.]

Budući da je $\mathrm{Var}(v^\tau X) = v^\tau \Sigma v$, standarni rezultat linearne algebre povlači da kao $v_1, \dots, v_p$ možemo uzeti (normalizirane) *svojstvene vektore* kovarijacijske matrice Σ pri čemu su oni poredani silazno s obzirom na pripadne svojstvene vrijednosti, tj. $\lambda_1 \geqslant \lambda_2 \geqslant \dots \geqslant \lambda_p > 0$.

Primjerice, kod dvodimenzionalne normalne razdiobe konture konstantne gustoće su rotirane elipse (vidi Sliku 21), a nije teško za pokazati da smjerovi glavnih komponenti leže upravo na osima tih elipsi [nacrtati sliku].

U praksi imamo skup podataka $x^{(i)} \in \mathbb{R}^p$, $i = 1, \dots, n$, koje modeliramo kao njd uzorak za neki konkretan slučajni vektor $X \in \mathbb{R}^p$ s kovarijacijskom matricom Σ. **Uzoračke glavne komponente** dobijemo tako da računamo svojstvene vektore $v_1, \dots, v_p$ nekog procjenitelja $\widehat{\Sigma}$ od Σ. Redukciju dimenzije problema tada dobivamo tako da originalne podatke zamijenimo s projekcijama na samo prvih nekoliko glavnih komponenti – najčešće samo prve dvije $(v_1^\tau x^{(i)}, v_2^\tau x^{(i)}) \in \mathbb{R}^2$ radi lakše vizualizacije. Vidi Sliku 22 za lijepu primjenu PCA metode.



SLIKA 22. U studiji objavljenoj u [ea08], za $n \approx 1400$ Europljana prikupljeni su podaci o njihovim genomima – za svaku osobu to je skup od čak $p \approx 200000$ podataka! Gore je prikaz podataka u prve dvije glavne komponente (uz dodatnu rotaciju). Svaka točka odgovara jednoj osobi, a boja/naziv odgovara geografskom podrijetlu te osobe. Zanimljivo je da ovaj prikaz više-manje odgovara stvarnoj karti Europe, iako u podacima nemamo podataka o geografskom podrijetlu osoba. To upućuje da koristeći samo DNA pojedine osobe, možemo s velikom preciznosti odrediti njezino geografsko podrijetlo. Za više detalja, vidjeti [ea08].

DODATAK A

Višedimenzionalni integral

Ako je $f : \mathbb{R}^2 \to \mathbb{R}$ funkcija i $A \subseteq \mathbb{R}^2$, (Riemannov) integral funkcije f po A moguće je definirati slično kao u jednodimenzionalnom slučaju te se označava s $\int_A f(x,y)\,\mathrm{d}x\mathrm{d}y$ ili $\int\int_{(x,y)\in A} f(x,y)\,\mathrm{d}x\mathrm{d}y$. Ukoliko je f nenegativna funkcija $\int_A f(x,y)\,\mathrm{d}x\mathrm{d}y$ predstavlja *volumen* ispod grafa funkcije f te iznad područja A – na primjer, ako je $f \equiv 1$, taj integral je upravo *površina* skupa A.

U praksi, višedimenzionalni integral računamo koristeći tzv. **Fubinijev teorem**: ako je $A = \{(x,y) : x \in [a,b], y \in [\varphi(x), \psi(x)]\}$, vrijedi

$$\int_A f(x,y)\,\mathrm{d}x\mathrm{d}y = \int_a^b \left(\int_{\varphi(x)}^{\psi(x)} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x\,.$$

Na primjer, ako je $A = [a,b] \times [c,d]$ pravokutnik,

$$\int_A f(x,y)\,\mathrm{d}x\mathrm{d}y = \int_a^b \left(\int_c^d f(x,y)\,\mathrm{d}y \right) \mathrm{d}x\,,$$

pri čemu možemo zamijeniti poredak varijabli (u Fubinijevom teoremu samo zamijenimo ulogu varijabli x i y), tj. vrijedi

$$\int_A f(x,y)\,\mathrm{d}x\mathrm{d}y = \int_c^d \left(\int_a^b f(x,y)\,\mathrm{d}x \right) \mathrm{d}y\,.$$

U praksi tipično izostavljamo pisanje zagrada u gornjim integralima.

Primjer A.1. Ako je $A = \{(x,y) : 0 \leqslant x \leqslant 1\,,\ x \leqslant y \leqslant 1\}$, imamo

$$\begin{aligned}\int_A (x-y)^2\,\mathrm{d}x\mathrm{d}y &= \int_0^1 \left(\int_x^1 (x-y)^2\,\mathrm{d}y \right) \mathrm{d}x = \int_0^1 \left(-\frac{(x-y)^3}{3}\Big|_{y=x}^{1} \right) \mathrm{d}x \\ &= \int_0^1 \left(-\frac{(x-1)^3}{3} + 0 \right) \mathrm{d}x = -\frac{(x-1)^4}{3\cdot 4}\Big|_{x=0}^{1} = 0 + \frac{1}{3\cdot 4} = \frac{1}{12}\,.\end{aligned}$$

Alternativno, budući da je $A = \{(x,y) : 0 \leqslant y \leqslant 1\,,\ 0 \leqslant x \leqslant y\}$, imamo

$$\begin{aligned}\int_A (x-y)^2\,\mathrm{d}x\mathrm{d}y &= \int_0^1 \left(\int_0^y (x-y)^2\,\mathrm{d}x \right) \mathrm{d}y = \int_0^1 \left(\frac{(x-y)^3}{3}\Big|_{x=0}^{y} \right) \mathrm{d}x \\ &= \int_0^1 \left(0 + \frac{y^3}{3} \right) \mathrm{d}y = \frac{y^4}{3\cdot 4}\Big|_{y=0}^{1} = \frac{1}{3\cdot 4} - 0 = \frac{1}{12}\,.\end{aligned}$$

$\square$

Integral funkcije $f : \mathbb{R}^d \to \mathbb{R}$ za $d \geqslant 3$ se također može definirati kao u slučaju $d = 2$, te vrijedi analogon Fubinijevog teorema. Na kraju napominjemo da smo ovdje očito preskočili razne tehničke detalje. Rigoroznu teoriju višedimenzionalnog Riemannovog integrala raditi ćete na kolegiju INTRAF, a teoriju općenitijeg (jednodimenzionalnog i višedimenzionalnog) Lebesgueovog integrala (koji se koristi u teoriji vjerojatnosti) na kolegiju Mjera i integral.

Bibliografija

[BH19] Joseph K Blitzstein i Jessica Hwang: *Introduction to probability*. Chapman and Hall/CRC, 2019.

[Cur24] Nicolas Curien: *A random walk among random graphs*. arXiv preprint arXiv:2412.19752, 2024.

[ea08] Novembre et al.: *Genes mirror geography within Europe*. Nature, 456(7218):98–101, 2008.

[GW14] Geoffrey Grimmett i Dominic JA Welsh: *Probability: an introduction*. Oxford University Press, 2014.

[Sti03] D Stirzaker: *Elementary Probability*. Cambridge University Press, 2003.

[Web] Richard Weber: *Probability (lecture notes, Cambridge)*. `http://www.statslab.cam.ac.uk/~rrw1/prob/prob-weber.pdf`.