Sadržaj

1 Vjerojatnost

U vjerojatnosti nas zanimaju pitanja oblika:

(a) Koja je vjerojatnost da, ako bacimo igraću kocku, padne paran broj?

(b) Kolika je vjerojatnost da, ako bacamo dvije igraće kocke, zbroj brojeva koji padne na kockama bude manji ili jednak 7?

(c) Koji je očekivani broj bacanja simetričnog novčića prije nego što dva puta zaredom padne pismo?

(d) Koja je vjerojatnost da dva nasumično odabrana studenta iz predavaonice imaju rođendan na isti datum?

Prije nego što pokažemo kako izračunati odgovore na ova pitanja, definirajmo objekte koje ćemo koristiti.

Skup svih ishoda nekog slučajnog pokusa označavamo s Ω i zovemo ga **prostor elementarnih događaja**.

Primjer 1.1. Primjeri prostora elementarnih događaja Ω za pokuse:

(a) bacanje simetrične kocke: $\Omega = \{1, 2, 3, 4, 5, 6\}$,

(b) bacanje simetričnog novčića: $\Omega = \{P, G\}$,

(c) bacamo dva simetrična novčića: $\Omega = \{PP, PG, GP, GG\}$.

Prostor elementarnih događaja može sadržavati i beskonačno mnogo elemenata:

(d) brojimo koliko ljudi uđe na PMF između 9 i 10 sati: $\Omega = \{0, 1, 2, 3, ...\} = \mathbb{N}_0$,

(e) mjerimo vrijeme čekanja do dolaska tramvaja: $\Omega = [0, +\infty)$,

(f) bacamo simetričan novčić sve dok prvi put ne padne pismo:

$$\Omega = \{P, GP, GGP, GGGP, \dots\} \cup \{GGGG \cdots\}.$$

Napomenimo da je ishod $GGGG\cdots$ teoretski moguć, ali ima "vjerojatnost 0" pa se zapravo može zanemariti.

Podskup $A \subseteq \Omega$ zovemo **događaj**. Na primjer, kod bacanja simetrične kocke podskup

$$A = \{\text{pao je paran broj}\} = \{2, 4, 6\}$$

predstavlja jedan događaj. Ako je ishod slučajnog pokusa $\omega \in \Omega$ i vrijedi $\omega \in A$, kažemo da se događaj A *dogodio*. Na primjer, ako je ishod bacanja kocke da je pao broj 2, tada je $\omega = 2$, a s obzirom da je 2 paran, $2 \in A = \{2, 4, 6\}$ i kažemo da se događaj A dogodio.

Zadatak 1.2. Slučajni pokus sastoji se od bacanja dvije igraće kocke. Odredite prostor elementarnih događaja Ω te prikažite pomoću elementarnih događaja sljedeće događaje:

$$A = \{\text{barem jedna kocka je pala na } 1\},$$

$$B = \{\text{zbroj brojeva na kockama je } 5\},$$

$$C = \{\text{zbroj brojeva na kockama je } 5, \text{ ali nijedna kocka nije pala na } 1\}.$$

Rješenje. Prostor elementarnih događaja je

$$\Omega = \{(i,j) : \ i,j \in \{1,2,3,4,5,6\}\} = \{1,2,3,4,5,6\}^2,$$

pri čemu npr. $(5,3) \in \Omega$ predstavlja ishod kada je na prvoj kocki pao broj 5, a na drugoj broj 3. Tada je

$$\begin{aligned} A &= \{(1,1),(1,2),(2,1),(1,3),(3,1),(1,4),(4,1),(1,5),(5,1),(1,6),(6,1)\} \\ &= \{(1,j) : j \in \{1,\ldots,6\}\} \cup \{(i,1) : i \in \{1,\ldots,6\}\}, \\ B &= \{(1,4),(4,1),(2,3),(3,2)\}, \\ C &= B \setminus A = \{(2,3),(3,2)\}. \end{aligned}$$

□

S obzirom na to da ćemo raditi sa skupovima, korisno je prisjetiti se osnovnih operacija sa skupovima i pripadnih svojstava.

Zadatak 1.3. Ako su A, B i C događaji, prikažite pomoću A, B i C sljedeće događaje:

(a) dogodio se barem jedan gornji događaj,

(b) dogodio se točno jedan gornji događaj,

(c) dogodila su se točno dva gornja događaja,

(d) nisu se dogodila sva tri gornja događaja.

Rješenje. Traženi događaji su:

(a) $A \cup B \cup C$,

(b) $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C)$,

(c) $(A \cap B \cap C^c) \cup (A \cap B^c \cap C) \cup (A^c \cap B \cap C)$,

(d) $(A \cap B \cap C)^c = A^c \cup B^c \cup C^c$.

□

Općenito, nećemo sve podskupove od Ω zvati događajima, već samo one koji pripadaju familiji $\mathcal{F}$ podskupova od Ω koja zadovoljava sljedeća svojstva.

Definicija 1.4. Familija $\mathcal{F}$ podskupova od Ω je **σ-algebra** na Ω ako je:

(F1) $\emptyset \in \mathcal{F}$,

(F2) $A \in \mathcal{F} \Rightarrow A^c = \Omega \setminus A \in \mathcal{F}$,

(F3) $A_n \in \mathcal{F}, n \in \mathbb{N} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Generalno, σ-algebra treba sadržavati sve događaje za koje želimo znati odrediti vjerojatnost.

Napomena 1.5. (a) Uvijek je $\Omega = \emptyset^c \in \mathcal{F}$.

(b) Najmanja σ-algebra na Ω je $\mathcal{F} = \{\emptyset, \Omega\}$, a najveća $\mathcal{F} = \mathcal{P}(\Omega)$, gdje je $\mathcal{P}(\Omega)$ partitivni skup od Ω, tj. familija svih podskupova od Ω.

(c) Najmanja σ-algebra na Ω koja sadrži $A \subseteq \Omega$ je $\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$.

(d) Iz (F3) specijalno slijedi: $A, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}$. Posljedično, ista tvrdnja vrijedi i za sve konačne unije.

Zadatak 1.6. Neka je $\mathcal{F}$ σ-algebra na Ω i $A, B \in \mathcal{F}$. Dokažite da je tada

$$A \cap B,\ A \setminus B,\ B \setminus A,\ A \triangle B \in \mathcal{F}.$$

Rješenje.

- Iz svojstava (F2) i (F3) slijedi da je $A \cap B = (A^c \cup B^c)^c \in \mathcal{F}$.
- Koristeći prethodno slijedi da je $A \setminus B = A \cap B^c \in \mathcal{F}$. Analogno je i $B \setminus A \in F$.
- Po prethodnom, $A \triangle B = (A \setminus B) \cup (B \setminus A) \in \mathcal{F}$.

$\square$

Zadatak 1.7. Neka je $\Omega = \{1, 2, 3, 4\}$. Koje od sljedećih familija podskupova od Ω su σ-algebre na Ω:

(a) $\mathcal{F}_1 = \{\emptyset, \{1, 2\}, \{3, 4\}\}$,

(b) $\mathcal{F}_2 = \{\emptyset, \Omega, \{1\}, \{2, 3, 4\}, \{1, 2\}, \{3, 4\}\}$,

(c) $\mathcal{F}_3 = \{\emptyset, \Omega, \{1\}, \{2\}, \{1, 2\}, \{3, 4\}, \{2, 3, 4\}, \{1, 3, 4\}\}$?

Rješenje.

(a) $\mathcal{F}_1$ nije σ-algebra jer $\Omega \notin \mathcal{F}_1$.

(b) $\mathcal{F}_2$ nije σ-algebra jer $\{1\} \cup \{3, 4\} = \{1, 3, 4\} \notin \mathcal{F}_2$.

(c) $\mathcal{F}_3$ je σ-algebra jer zadovoljava uvjete iz Definicije 1.4. Uočite da su skupovi $\{1\}, \{2\}, \{3,4\} \in \mathcal{F}_3$ međusobno disjunktni i da se svaki skup iz $\mathcal{F}_3$ može prikazati kao neka unija ova tri skupa (kažemo da su ti skupovi *atomi* σ-algebre $\mathcal{F}_3$).

$\square$

Zadatak 1.8. Neka je $\Omega = \{1,2,3,4,5\}$. Nađite najmanju σ-algebru na Ω koja sadrži skupove

$$\{1\},\ \{1,2\},\ \{1,2,3\},\ \{1,2,3,4\}.$$

Rješenje. Neka je $\mathcal{F}$ tražena σ-algebra. Tada $\mathcal{F}$ mora sadržavati i sljedeće skupove:

$$\{1\},\ \{2\} = \{1,2\} \setminus \{1\},\ \{3\} = \{1,2,3\} \setminus \{1,2\},\ \{4\} = \{1,2,3,4\} \setminus \{1,2,3\} \text{ i } \{5\} = \{1,2,3,4\}^c.$$

Budući da $\mathcal{F}$ onda mora sadržavati i sve moguće unije gornjih skupova slijedi da je svaki podskup od Ω sadržan u $\mathcal{F}$, tj. $\mathcal{P}(\Omega) \subseteq \mathcal{F}$, pa slijedi $\mathcal{F} = \mathcal{P}(\Omega)$.

$\square$

Napomena 1.9. Ako je Ω konačan ili najviše prebrojivo beskonačan, najčešće uzimamo da je $\mathcal{F} = \mathcal{P}(\Omega)$.

Primjer 1.10. Bacamo simetričnu kocku i gledamo je li pao paran ili neparan broj. Kako izgleda skup elementarnih događaja Ω i pripadna σ-algebra $\mathcal{F}$? Pokušajte odrediti vjerojatnosti traženih ishoda.

Rješenje. Skup elementarnih događaja su svi ishodi bacanja kocke, dakle svi brojevi od 1 do 6

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

$\mathcal{F}$ kao σ-algebra mora sadržavati $\emptyset$ i Ω. Osim toga, znamo da σ-algebra mora sadržavati događaje kojima želimo odrediti vjerojatnost pa će $\mathcal{F}$ sadržavati skupove $\{\text{pao je paran broj}\} = \{2,4,6\}$ i $\{\text{pao je neparan broj}\} = \{1,3,5\}$. Ako definiramo

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

$\mathcal{F}$ je zaista σ-algebra (sami provjerite da zadovoljava uvjete iz definicije). S obzirom da je kocka simetrična, svaki broj je jednako vjerojatan pa je, koristeći Laplaceovu vjerojatnost,

$$\mathbb{P}\left(\{\text{pao je paran broj}\}\right) = 3/6 = 1/2$$

$$\mathbb{P}\left(\{\text{pao je neparan broj}\}\right) = 3/6 = 1/2$$

Dodatno, možemo odrediti i vjerojatnosti preostalih elemenata iz $\mathcal{F}$, $\mathbb{P}(\emptyset) = 0$ (vjerojatnost da ništa nije palo nakon bacanja je 0), i $\mathbb{P}(\Omega) = 1$ (vjerojatnost da je pao broj između 1 i 6 je 1).

$\square$

Definicija 1.11. **Vjerojatnost** je funkcija $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ sa sljedećim svojstvima:

(P1) $0 \leq \mathbb{P}(A) \leq 1$, za svaki $A \in \mathcal{F}$,

(P2) $\mathbb{P}(\Omega) = 1$,

(P3) Ako su $A_n \in \mathcal{F}, n \in \mathbb{N}$ međusobno disjunktni, tada je

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

Uređenu trojku $(\Omega, \mathcal{F}, \mathbb{P})$ zovemo **vjerojatnosni prostor**.

Kada je Ω konačan skup i svi elementarni događaji su jednako vjerojatni, koristimo **Laplaceov model**, tj.

$$\mathbb{P}(A) = \frac{k(A)}{k(\Omega)} = \frac{\text{broj elementarnih događaja iz } A}{\text{broj svih elementarnih događaja}}.$$

Nije teško za provjeriti da ovako definirana funkcija $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ zaista vjerojatnost.

Zadatak 1.12. Bacamo dvije simetrične kocke. Izračunajte vjerojatnost sljedećih događaja:

(a) $A = \{\text{suma brojeva na kockama je 7 ili 11}\}$,

(b) $B = \{\text{brojevi na kockama su relativno prosti}\}$,

(c) $C = \{\text{produkt brojeva na kockama je neparan}\}$,

(d) $D = \{\text{jedan broj na kocki dijeli drugi}\}$.

Rješenje. Prostor elementarnih događaja je

$$\Omega = \{(i, j) : \ i, j \in \{1, 2, 3, 4, 5, 6\}\} = \{1, 2, 3, 4, 5, 6\}^2,$$

pa je

$$k(\Omega) = 6 \cdot 6 = 36.$$

(a) Budući da je

$$A = \{(1, 6), (6, 1), (2, 5), (5, 2), (3, 4), (4, 3), (5, 6), (6, 5)\},$$

slijedi da je $k(A) = 8$ i $\mathbb{P}(A) = \frac{k(A)}{k(\Omega)} = \frac{8}{36} = \frac{2}{9}$.

(b) Uočimo da je umjesto $k(B)$ lakše odrediti $k(B^c)$ jer B^c ima manje elemenata. Budući da je

$$B^c = \{(2, 2), (2, 4), (4, 2), (2, 6), (6, 2), (3, 3), (3, 6), (6, 3), (4, 4), (4, 6), (6, 4), (5, 5), (6, 6)\},$$

slijedi da je $k(B^c) = 13$ i $\mathbb{P}(B) = \frac{k(B)}{k(\Omega)} = \frac{k(\Omega)-k(B^c)}{k(\Omega)} = \frac{36-13}{36} = \frac{23}{36}$.

(c) Uočimo da je zapravo $C = \{\text{oba broja su neparna}\}$ pa je

$$C = \{(1,1),(1,3),(3,1),(1,5),(5,1),(3,3),(3,5),(5,3),(5,5)\}.$$

Dakle, $k(C) = 9$ i $\mathbb{P}(C) = \frac{k(C)}{k(\Omega)} = \frac{9}{36} = \frac{1}{4}$. Ovaj zadatak smo lakše mogli riješiti ako primijetimo da je za određivanje $\mathbb{P}(C)$ dovoljno gledati samo koje su parnosti brojevi na kockama. Budući da je tada ukupno 4 (jednako vjerojatna) ishoda, a samo je jedan povoljan za C, slijedi $\mathbb{P}(C) = \frac{1}{4}$.

(d) Uočimo ponovno da je lakše odrediti $k(D^c)$ nego $k(D^c)$. Budući da je

$$D^c = \{(2,3),(3,2),(2,5),(5,2),(3,4),(4,3),(3,5),(5,3),(4,5),(5,4),(4,6),(6,4),(5,6),(6,5)\},$$

slijedi da je $k(D^c) = 14$ i $\mathbb{P}(D) = \frac{k(D)}{k(\Omega)} = \frac{k(\Omega)-k(D^c)}{k(\Omega)} = \frac{22}{36} = \frac{11}{18}$.

$\square$

Zadatak 1.13. U nekoj školi učenici mogu učiti 3 strana jezika: engleski, njemački i francuski. Od 100 učenika, 28 uči engleski, 16 francuski, 26 njemački, 4 engleski i francuski, 12 engleski i njemački, 6 njemački i francuski, a 2 uče sva tri jezika. Izračunajte vjerojatnost da slučajno odabrani učenik

(a) ne uči niti jedan strani jezik,

(b) uči samo engleski ili samo francuski jezik,

(c) uči engleski,

(d) uči samo engleski,

(e) uči točno 2 jezika,

(f) ne uči francuski.

Rješenje. Prostor elementarnih događaja je

$$\Omega = \{\text{svi učenici}\}$$

pa je $k(\Omega) = 100$. Definiramo događaje

$$E = \{\text{odabrani učenik uči engleski jezik}\},$$

$$F = \{\text{odabrani učenik uči francuski jezik}\},$$

$$G = \{\text{odabrani učenik uči njemački jezik}\}.$$

(a) Tražimo vjerojatnost

$$\mathbb{P}(E^c \cap F^c \cap G^c) = \frac{k(E^c \cap F^c \cap G^c)}{k(\Omega)}.$$

Na temelju danih informacija, koristeći Vennov diagram, nije teško za odrediti da je $k(E \cup F \cup G) = 50$ pa je

$$k(E^c \cap F^c \cap G^c) = k((E \cup F \cup G)^c) = k(\Omega) - k(E \cup F \cup G) = 100 - 50 = 50$$

te je tražena vjerojatnost $\frac{50}{100} = \frac{1}{2}$.

(b) Slično kao u (a), imamo da je

$$k\left((E \cap F^c \cap G^c) \cup (E^c \cap F \cap G^c)\right) = k\left(E \cap F^c \cap G^c\right) + k\left(E^c \cap F \cap G^c\right) = 14 + 8 = 22$$

pa je

$$\mathbb{P}\left((E \cap F^c \cap G^c) \cup (E^c \cap F \cap G^c)\right) = \frac{22}{100} = \frac{11}{50}\,.$$

(c) $\frac{7}{25}$

(d) $\frac{7}{50}$

(e) $\frac{4}{25}$

(f) $\frac{21}{25}$

$\square$

Napomena 1.14 (Svojstva vjerojatnosti). Neka su $A, B, C \in \mathcal{F}$, tada vrijedi:

(a) $A \subseteq B \Rightarrow \mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$. Specijalno, $A \subseteq B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$.

(b) $\mathbb{P}(A^c) = \mathbb{P}(\Omega \setminus A) = \mathbb{P}(\Omega) - \mathbb{P}(A) = 1 - \mathbb{P}(A)$.

(c) $\mathbb{P}(\emptyset) = 1 - \mathbb{P}(\Omega) = 0$.

(d) (**Formula uključivanja-isključivanja**)
$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ i
$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)$.

Zadatak 1.15. Set za čaj se sastoji od tri šalice i tri tanjurića u tri boje. Šalice su slučajno raspoređene na tanjuriće. Kolika je vjerojatnost da ni jedna šalica nije na tanjuriću iste boje?

Rješenje. Neka su boje C=crvena, P=plava i Z=zelena. Nekako fiksiramo položaj tanjurića, npr. crveni, plavi pa zeleni. Prostor elementarnih događaja je

$$\Omega = \{\text{sve moguće permutacije šalica na tanjuriće}\} = \{PCZ, PZC, CPZ, CZP, ZCP, ZPC\}$$

pa je $k(\Omega) = 3! = 6$.

Lako vidimo da je traženi događaj

$$A = \{PZC, ZCP\}$$

pa je

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

Ipak, ako je broj šalica i tanjurića veći, rješenje će se lakše naći koristeći (poopćenu) formulu uključivanja-isključivanja. Ako definiramo događaje

$$A_x = \{\text{šalica boje } x \text{ je na tanjuriću boje } x\},$$

za $x \in \{P, C, Z\}$, tražimo vjerojatnost

$$\begin{aligned}\mathbb{P}(A_P^c \cap A_C^c \cap A_Z^c) &= 1 - \mathbb{P}(A_P \cup A_C \cup A_Z) \\ &= 1 - \mathbb{P}(A_P) - \mathbb{P}(A_C) - \mathbb{P}(A_Z) + \mathbb{P}(A_P \cap A_C) + \mathbb{P}(A_P \cap A_Z) \\ &+ \mathbb{P}(A_C \cap A_Z) - \mathbb{P}(A_P \cap A_C \cap A_Z) \\ &= 1 - 3 \cdot \frac{2}{6} + 3 \cdot \frac{1}{6} - \frac{1}{6} = \frac{1}{3}.\end{aligned}$$

□

Zadatak 1.16. Bacamo 5 simetričnih kocki. Izračunajte vjerojatnost da produkt dobivenih brojeva

(a) bude djeljiv s 5,

(b) ima zadnju znamenku 0,

(c) ima zadnju znamenku 5.

Rješenje. Prostor elementarnih događaja je

$$\Omega = \{(i_1, \ldots, i_5) : i_k \in \{1, 2, \ldots, 6\}, k = 1, 2, \ldots, 5\} = \{1, 2, \ldots, 6\}^5$$

pa je $k(\Omega) = 6^5$.

(a) Uočite da zapravo tražimo vjerojatnost događaja

$$A = \{\text{pala je barem jedna petica}\}.$$

Umjesto da ovaj događaj rastavljamo ovisno o točnom broju petica, primijetimo da je

$$A^c = \{\text{nije pala niti jedna petica}\} = \{(i_1, \ldots, i_5) : i_k \neq 5, k = 1, \ldots, 5\} = \{1, 2, 3, 4, 6\}^5$$

pa je $k(A^c) = 5^5$ i $\mathbb{P}(A^c) = \frac{k(A^c)}{k(\Omega)} = \frac{5^5}{6^5}$. Dakle, $\mathbb{P}(A) = 1 - \mathbb{P}(A^c) = 1 - \frac{5^5}{6^5}$.

(b) Traženi događaj je

$$B = \{\text{produkt djeljiv i s dva i s pet}\}.$$

U tu svrhu definiramo događaje

$$B_1 = \{\text{pao je barem jedan paran broj}\}$$

i

$$B_2 = \{\text{pala je barem jedna petica}\} = A.$$

Tada je $B = B_1 \cap B_2$ te kao i u (a) dijelu računamo

$$\begin{aligned}
\mathbb{P}(B) &= \mathbb{P}(B_1 \cap B_2) = 1 - \mathbb{P}((B_1 \cap B_2)^c) = 1 - \mathbb{P}(B_1^c \cup B_2^c) \\
&= [\text{FUI}] = 1 - \mathbb{P}(B_1^c) - \mathbb{P}(B_2^c) + \mathbb{P}(B_1^c \cap B_2^c) \\
&= 1 - \frac{k(B_1^c)}{k(\Omega)} - \frac{k(B_2^c)}{k(\Omega)} + \frac{k(B_1^c \cap B_2^c)}{k(\Omega)} = 1 - \frac{3^5}{6^5} - \frac{5^5}{6^5} + \frac{2^5}{6^5} \\
&= 1 - \frac{1}{2^5} - \frac{5^5}{6^5} + \frac{1}{3^5}.
\end{aligned}$$

(c) Traženi događaj je zapravo

$$C = \{\text{produkt djeljiv s pet, ali ne i s dva}\} = B_2 \setminus B_1.$$

Slijedi da je

$$\begin{aligned}
\mathbb{P}(C) &= \mathbb{P}(B_2 \setminus B_1) = \mathbb{P}(B_2 \setminus (B_1 \cap B_2)) \\
&= [B_1 \cap B_2 \subseteq B_2] = \mathbb{P}(B_2) - \mathbb{P}(B_1 \cap B_2) \\
&= \mathbb{P}(A) - \mathbb{P}(B) \\
&= \left(1 - \frac{5^5}{6^5}\right) - \left(1 - \frac{1}{2^5} - \frac{5^5}{6^5} + \frac{1}{3^5}\right) \\
&= \frac{1}{2^5} - \frac{1}{3^5}.
\end{aligned}$$

□

Zadatak 1.17. Neka su $A, B, C \in \mathcal{F}$. Dokažite:

(a) $\mathbb{P}(A \cup B \cup C) \leq \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C)$,

(b) $\mathbb{P}(A \cap B \cap C) \geq \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - 2$,

(c) $\mathbb{P}(A \cup B)\mathbb{P}(A \cap B) \leq \mathbb{P}(A)\mathbb{P}(B)$.

Rješenje.

(a) Definirajmo događaje

$$D_1 = A,\ D_2 = B \setminus A,\ D_3 = C \setminus (A \cup B)\,.$$

Očito je $A \cup B \cup C = D_1 \cup D_2 \cup D_3$, ali skupovi D_1, D_2 i D_3 su međusobno disjunktni pa slijedi

$$\begin{aligned}
\mathbb{P}(A \cup B \cup C) &= \mathbb{P}(D_1 \cup D_2 \cup D_3) = \mathbb{P}(D_1) + \mathbb{P}(D_2) + \mathbb{P}(D_3) \\
&\leq [D_1 \subseteq A,\ D_2 \subseteq B,\ D_3 \subseteq C] \leq \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C).
\end{aligned}$$

(b) Koristeći prethodni dio imamo

$$\begin{aligned}\mathbb{P}(A \cap B \cap C) &= 1 - \mathbb{P}(A^c \cup B^c \cup C^c) \geq 1 - \mathbb{P}(A^c) - \mathbb{P}(B^c) - \mathbb{P}(C^c) \\ &= 1 - (1 - \mathbb{P}(A)) - (1 - \mathbb{P}(B)) - (1 - \mathbb{P}(C)) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - 2.\end{aligned}$$

(c) Budući da je $A \cup B$ disjunktna unija skupova A i $B \setminus A$ slijedi da je

$$\begin{aligned}\mathbb{P}(A \cup B)\mathbb{P}(A \cap B) &= (\mathbb{P}(A) + \mathbb{P}(B \setminus A))\,\mathbb{P}(A \cap B) \\ &= \mathbb{P}(A)\mathbb{P}(A \cap B) + \mathbb{P}(B \setminus A)\underbrace{\mathbb{P}(A \cap B)}_{\leq \mathbb{P}(A)} \\ &\leq \mathbb{P}(A)\underbrace{(\mathbb{P}(A \cap B) + \mathbb{P}(B \setminus A))}_{A \cap B \text{ i } B \setminus A \text{ disjunktni}} \\ &= \mathbb{P}(A)\mathbb{P}((A \cap B) \cup (B \setminus A)) = \mathbb{P}(A)\mathbb{P}(B).\end{aligned}$$

$\square$

Napomena 1.18. Iz prethodnog zadatka slijedi:

- $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = 0 \Rightarrow \mathbb{P}(A \cup B \cup C) = 0$. Zaista,

$$0 \leq \mathbb{P}(A \cup B \cup C) \leq \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) = 0.$$

- $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = 1 \Rightarrow \mathbb{P}(A \cap B \cap C) = 1$. Zaista,

$$1 \geq \mathbb{P}(A \cap B \cap C) \geq \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - 2 = 3 - 2 = 1.$$

Zadaci za vježbu

Zadatak 1.19. Nadopunite familije podskupova iz zadatka 1.7 tako da čine σ-algebre.

Zadatak 1.20. Neka su $\mathcal{F}, \mathcal{G}$ σ-algebre na Ω:

(a) Je li $\mathcal{F} \cap \mathcal{G}$ σ-algebra?

(b) Je li $\mathcal{F} \cup \mathcal{G}$ σ-algebra?

Ako je vaš odgovor potvrdan, dokažite to, inače nađite protuprimjer.

Zadatak 1.21. Neka je $\mathcal{F}$ σ-algebra na Ω i $B \in \mathcal{F}$. Dokažite da je tada i

$$\mathcal{G} = \{C \subseteq B : \text{ postoji } A \in \mathcal{F} \text{ takav da je } C = A \cap B\}$$

σ-algebra na B.

Zadatak 1.22. Pokažite da vrijedi formula uključivanja-isključivanja iz napomene 1.14.
Uputa: zapišite $A \cup B$ kao uniju disjunktnih skupova $A \backslash (A \cap B)$, $B \backslash (A \cap B)$ i $A \cap B$.

Zadatak 1.23. Neka su $A, B \in \mathcal{F}$.

(a) Ako je $\mathbb{P}(A \cup B) = 0.8, \mathbb{P}(A \cap B) = 0.2$ i $\mathbb{P}(B^c) = 0.6$, koliko je $\mathbb{P}(A), \mathbb{P}(B)$ i $\mathbb{P}(B \setminus A)$?

(b) Ako je $\mathbb{P}(A \cup B^c) = 0$, koliko je $\mathbb{P}(B)$ i $\mathbb{P}(A \cap B)$?

Zadatak 1.24. Bacamo dvije simetrične kocke i dobivene ishode označimo s A i B. Odredite prostor elementarnih događaja i izračunajte vjerojatnost da jednadžba $x^2 + Ax + B = 0$ ima realna rješenja?

2 Prebrojavanje

Zadatak 2.1. U vrećici imamo 550 jabuka, od čega je 2% trulih. Kolika je vjerojatnost da u slučajnom uzorku od 25 jabuka budu točno 2 trule?

Rješenje. Prostor elementarnih događaja je

$$\Omega = \{\text{svi 25-člani podskupovi od 550 jabuka}\}$$

pa je $k(\Omega) = \binom{550}{25}$. Traženi događaj je

$$A = \{\text{točno su 2 jabuke u uzorku trule}\}.$$

Budući da je u vreći $0.02 \cdot 550 = 11$ trulih jabuka slijedi da je

$$k(A) = \binom{11}{2} \cdot \binom{550-11}{25-2} = \binom{11}{2} \cdot \binom{539}{23}.$$

Dakle,

$$\mathbb{P}(A) = \frac{k(A)}{k(\Omega)} = \frac{\binom{11}{2} \cdot \binom{539}{23}}{\binom{550}{25}} \approx 0.074.$$

□

Zadatak 2.2. Magnus na šahovsku ploču postavlja tri topa, pri čemu ne može staviti dva topa na isto polje. Kolika je vjerojatnost da se nikoja dva topa ne napadaju?

Rješenje. Šahovska ploča ima $8 \cdot 8 = 64$ polja, označimo ih brojevima $1, 2, \dots, 64$. Za prostor elementarnih događaja možemo uzeti

$$\Omega = \{(i, j, k) : i, j, k \in \{1, \dots, 64\}, i \neq j \neq k\}$$

pri čemu prva, druga odnosno treća koordinata predstavljaju pozicije prvog, drugog odnosno trećeg topa. Dakle, $k(\Omega) = 64 \cdot 63 \cdot 62$. Kada stavimo prvog topa na bilo koje od 64 polja, ostaje $(8-1) \cdot (8-1) = 7^2$ mogućih polja koja prvi top ne napada. Isto tako, ako stavimo drugog topa na neko od tih polja, ostaje 6^2 mogućih polja na koje možemo staviti trećeg topa tako da ne napada prva dva. Dakle, ako je A traženi događaj, imamo da je

$$\mathbb{P}(A) = \frac{k(A)}{k(\Omega)} = \frac{8^2 \cdot 7^2 \cdot 6^2}{64 \cdot 63 \cdot 62} = \frac{14}{31} \approx 0.4516.$$

Uočite da nismo trebali razlikovati topove, tj. za prostor elementarnih događaja mogli smo uzeti

$$\Omega' = \{\text{svi 3-člani podskupovi skupa od 64 polja}\}.$$

U tom slučaju je $k(\Omega') = \binom{64}{3} = \frac{64 \cdot 63 \cdot 62}{3!}$, ali i $k(A) = \frac{8^2 \cdot 7^2 \cdot 6^2}{3!}$ pa je dakle vjerojatnost traženog događaja ista kao i u prvom slučaju. Vidi Napomenu 2.4 dolje. □

Zadatak 2.3. Špil se sastoji od 52 karte od kojih svaka ima neku od 13 jačina (2,..,10,J,Q,K,A) i neku od 4 boje.

(a) Ako slučajno izaberete dvije karte, kolika je vjerojatnost da ste dobili

  (a1) dvije karte iste boje,

  (a2) dvije karte različitih jačina,

  (a3) dvije karte različitih jačina, a istih boja,

  (a4) dvije karte istih boja koja su "susjednih" jačina (npr. 10 i J).

(b) U igri *poker* svaki od igrača na sreću dobije 5 karata iz špila. Kolika je vjerojatnost da igrač dobije

  (b1) jedan par, tj. dvije karte iste jačine i tri karte koje nisu te jačine niti istih jačina (npr. Q Q J 3 4),

  (b2) tri karte iste jačine i jedan par (npr. 10 10 10 K K)?

Rješenje.

(a) Prostor elementarnih događaja je $\Omega = \{\text{svi 2-člani podskupovi od 52 karte}\}$ pa je $k(\Omega) = \binom{52}{2}$.

  (a1) Neka je A_1 traženi događaj. Budući da boju možemo izabrati na 4 načina, a dvije karte u toj boji na $\binom{13}{2}$ načina slijedi da je

$$\mathbb{P}(A_1) = \frac{k(A_1)}{k(\Omega)} = \frac{4 \cdot \binom{13}{2}}{\binom{52}{2}} \approx 0.23529.$$

  (a2) Neka je A_2 traženi događaj. Budući da dvije jačine možemo izabrati na $\binom{13}{2}$ načina, a po jednu karte svake jačine na 4 načina slijedi da je

$$\mathbb{P}(A_2) = \frac{k(A_2)}{k(\Omega)} = \frac{\binom{13}{2} \cdot 4 \cdot 4}{\binom{52}{2}} \approx 0.94118.$$

  (a3) Ako su istih boja uvijek će biti različitih jačina, tako da je vjerojatnost ista kao vjerojatnost da su karte iste boje (a1).

  (a4) $\frac{4 \cdot 12}{\binom{52}{2}}$

(b) Prostor elementarnih događaja je $\Omega = \{\text{svi 5-člani podskupovi od 52 karte}\}$ pa je $k(\Omega) = \binom{52}{5}$.

  (b1) Neka je B_1 traženi događaj. Dvije karte iste jačine možemo izabrati na $\binom{13}{1} \cdot \binom{4}{2}$ načina. Budući da preostale tri karte moraju biti različite jačine, njih možemo izabrati na $\binom{12}{3} \cdot 4 \cdot 4 \cdot 4$ načina. Dakle,

$$\mathbb{P}(B_1) = \frac{k(B_1)}{k(\Omega)} = \frac{13 \cdot \binom{4}{2} \cdot \binom{12}{3} \cdot 64}{\binom{52}{5}} \approx 0.42257.$$

(b2) Neka je B_2 traženi događaj. Budući da 3 karte iste jačine možemo izabrati na $\binom{13}{1}\cdot\binom{4}{3}$ načina, a nakon toga par karata iste jačine možemo odabrati na $\binom{12}{1}\cdot\binom{4}{2}$ načina, slijedi da je

$$\mathbb{P}(B_2) = \frac{k(B_2)}{k(\Omega)} = \frac{13\cdot\binom{4}{3}\cdot 12\cdot\binom{4}{2}}{\binom{52}{5}} \approx 0.01441.$$

□

Napomena 2.4. Pri računanju vjerojatnosti u prethodnom zadatku nije bitno uzimamo li u obzir poredak izvučenih karata ili ne, ali onda treba paziti da se broj kombinacija računa na isti način. Na primjer, u (a) dijelu smo za prostor elementarnih događaja Ω mogli uzeti sva moguća izvlačenja dvije karte iz špila, ali pazeći na to koju smo kartu izvukli prvu, a koju drugu. Tada je $k(\Omega) = 52\cdot 51$ i

$$k(A_1) = 52\cdot(13-1), \quad \text{i} \quad k(A_2) = 52\cdot 48$$

što daje iste vjerojatnosti za ova dva događaja.

Nije uvijek svejedno uzimamo li u obzir poredak ili ne. Na primjer, bacamo 2 kocke i označimo sa B događaj da su pale jedinica i petica. Ako razlikujemo kocke (tj. poredak kojim su pali brojevi), $\Omega_1 = \{(i,j) : i,j = 1,2,...,6\}$, $k(\Omega_1) = 36$, $k(B) = 2$ i $\mathbb{P}(B) = 1/18$. Ako bismo to htjeli izračunati bez da razlikujemo kocke $\Omega_2 = \{\{i,j\}, i,j = 1,2,...,6\} \cup \{\{i\}, i = 1,2,...,6\}$, $k(\Omega_2) = \binom{6}{2} + \binom{6}{1} = 21$, $k(B) = 1$ i $\mathbb{P}(B) = 1/21$. Gdje smo pogriješili? U drugom slučaju je jednako vjerojatno da smo dobili bilo koji rezultat bacanja. Intuitivno, jasno je da bi vjerojatnost jedinice i petice od npr. dvije petice trebala biti veća. Gledajući ishode bacanja kao skupove smo "pobrisali" informaciju da se jedinica i petica mogu dobiti na duplo više načina nego dvije petice.

Općenito, prostor elementarnih događaja bira se tako da se najlakše riješi dani problem.

Zadatak 2.5. (Problem rođendana) Izračunajte vjerojatnost da je u grupi od n ljudi barem dvoje rođeno istog dana.

Rješenje. Prostor elementarnih događaja je

$$\Omega = \{(x_1,\ldots,x_n) : x_k \in \{1,\ldots,365\}, k = 1,\ldots,n\} = \{1,\ldots,365\}^n$$

pri čemu x_k predstavlja dan u godini kada je rođena k-ta osoba. Dakle, $k(\Omega) = 365^n$. Uočimo da ako je A traženi događaj, tada je

$$A^c = \{\text{sve osobe rođene na različite dane}\}$$

pa je

$$k(A^c) = 365\cdot(365-1)\cdot\cdots\cdot(365-n+1).$$

Dakle,

$$\mathbb{P}(A) = 1 - \mathbb{P}(A^c) = 1 - \frac{365\cdot 364\cdot\cdots\cdot(365-n+1)}{365^n}.$$

U tablici ispod su vrijednosti gornje vjerojatnosti za neke n-ove:

n	10	20	22	**23**	40	**70**
$\mathbb{P}(A)$	0.117	0.411	0.476	**0.507**	0.891	**0.999**

$\square$

Zadatak 2.6. Za okruglim stolom sjedi n osoba na n stolica. Ako je raspored sjedenja slučajan, kolika je vjerojatnost da Ante ne sjedi ni do Marka ni do Luke?

Rješenje. Prostor elementarnih događaja je

$$\Omega = \{\text{svi rasporedi sjedenja } n \text{ ljudi za okruglim stolom}\}.$$

Želimo odrediti $k(\Omega)$. Uočimo da je ovaj problem različit od problema: "Na koliko načina možemo n osoba posložiti u red?" ($= n!$) jer kada "spojimo" krajeve reda, više različitih rasporeda (permutacija) može dati isti raspored sjedenja za okruglim stolom. Budući da svakom rasporedu sjedenja oko okruglog stola odgovara točno n različitih permutacija ljudi u red, slijedi da je

$$k(\Omega) = \frac{n!}{n} = (n-1)!.$$

Alternativno, mogli smo izabrati jednu osobu i nju smatrati početkom, pa je jasno da preostale osobe možemo posložiti na $(n-1)!$ načina. Definirajmo događaje

$$M = \{\text{Ante sjedi do Marka}\},$$

$$L = \{\text{Ante sjedi do Luke}\}.$$

Tada je tražena vjerojatnost

$$\mathbb{P}(M^c \cap L^c) = 1 - \mathbb{P}(M \cup L) = 1 - \mathbb{P}(M) - \mathbb{P}(L) + \mathbb{P}(M \cap L).$$

Ako pretpostavimo da Ante sjedi do Marka, možemo ih smatrati jednim blokom te ujedno početkom. Unutar bloka ih rasporediti na $2!$ načina, a ostale ljude rasporediti na $(n-2)!$ načina, pa je

$$\mathbb{P}(M) = \frac{k(M)}{k(\Omega)} = \frac{2 \cdot (n-2)!}{(n-1)!} = \frac{2}{n-1} = \mathbb{P}(L).$$

Isto tako je

$$\mathbb{P}(M \cap L) = \frac{2! \cdot 1 \cdot (n-3)!}{(n-1)!} = \frac{2}{(n-1)(n-2)}$$

pa je

$$\mathbb{P}(M^c \cap L^c) = 1 - \frac{4}{n-1} + \frac{2}{(n-1)(n-2)} = \frac{n^2 - 7n + 12}{(n-1)(n-2)} = \frac{(n-3)(n-4)}{(n-1)(n-2)}.$$

Alternativno (i bolje), uočimo da traženi događaj ovisi jedino o tome koje su dvije osobe sjele pokraj Ante. Ako za prostor elementarnih događaja Ω uzmemo sve moguće parove susjeda (pazeći tko je zdesna a tko slijeva Anti), imamo da je $k(\Omega) = (n-1)(n-2)$, a budući da su svi ishodi jednako vjerojatni, imamo da je

$$\mathbb{P}(M^c \cap L^c) = \frac{k(M^c \cap L^c)}{k(\Omega)} = \frac{(n-3)(n-4)}{(n-1)(n-2)}\ .$$

$\square$

Zadatak 2.7. Luster ima 4 grla za žarulje, od kojih su 2 ispravna, a od 7 žarulja koje imamo, 4 su ispravne. Ako odaberemo na sreću 4 žarulje i stavimo ih u grla, kolika je vjerojatnost da ćemo uključivanjem lustera u struju dobiti svjetlo?

Rješenje. Prvo primijetimo da tražena vjerojatnost ne ovisi o tome kojim redoslijedom stavljamo žarulje na grla tako da možemo pretpostaviti da su prva dva grla ispravna. Prostor elementarnih događaja je

$$\Omega = \{\text{svi rasporedi sedam žarulja na četiri grla}\}$$

pa je $k(\Omega) = 7 \cdot 6 \cdot 5 \cdot 4$, te definiramo

$$A = \{\text{dobili smo svjetlo}\}.$$

Budući da je

$$A^c = \{\text{na prva dva grla stavljene neispravne žarulje}\}$$

slijedi da je

$$\mathbb{P}(A) = 1 - \mathbb{P}(A^c) = 1 - \frac{3 \cdot 2 \cdot 5 \cdot 4}{7 \cdot 6 \cdot 5 \cdot 4} = \frac{6}{7}.$$

Uočite da je dovoljno bilo gledati samo što se događa na dva ispravna grla.

Alternativno, ako definiramo događaje

$$A_1 = \{\text{dobili smo svjetlo iz prvog grla}\},$$

$$A_2 = \{\text{dobili smo svjetlo iz drugog grla}\},$$

tada je

$$\mathbb{P}(A) = \mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2).$$

Imamo da je $\mathbb{P}(A_1) = \frac{4}{7}$, a budući da na svako grlo s jednakom vjerojatnošću može doći bilo koja žarulja, mora biti i $\mathbb{P}(A_2) = \frac{4}{7}$[1]. Nadalje,

$$\mathbb{P}(A_1 \cap A_2) = \frac{4 \cdot 3}{7 \cdot 6} = \frac{2}{7}.$$

Dakle,

$$\mathbb{P}(A) = 2 \cdot \frac{4}{7} - \frac{2}{7} = \frac{6}{7}.$$

□

Zadatak 2.8. U kutiji se nalazi 5 crnih, 6 bijelih i 7 zelenih kuglica. Na slučajan način izvučemo 4 kuglice. Izračunajte vjerojatnost da među izvučenim kuglicama nisu zastupljene sve tri boje.

[1] Formalno, ako rastavimo A_2 na slučajeve ovisno o tome je li na prvo grlo došla ispravna ili neispravna žarulja, zaista dobijemo $\mathbb{P}(A_2) = \frac{4 \cdot 3 + 3 \cdot 4}{7 \cdot 6} = \frac{4}{7}$

Rješenje. Prostor elementarnih događaja je

$$\Omega = \{\text{svi 4-člani podskupovi od } 5+6+7=18 \text{ kuglica}\}$$

pa je $k(\Omega) = \binom{18}{4}$. Ako definiramo događaje

$$C = \{\text{nije zastupljena crna boja}\},$$

$$B = \{\text{nije zastupljena bijela boja}\},$$

$$Z = \{\text{nije zastupljena zelena boja}\},$$

tražena vjerojatnost je

$$\begin{aligned}\mathbb{P}(C \cup B \cup Z) &= \mathbb{P}(C) + \mathbb{P}(B) + \mathbb{P}(Z) - \mathbb{P}(C \cap B) - \mathbb{P}(C \cap Z) \\ &\quad - \mathbb{P}(B \cap Z) + \mathbb{P}(C \cap B \cap Z).\end{aligned}$$

Računamo

$$\mathbb{P}(C) = \frac{\binom{6+7}{4}}{\binom{18}{4}} = \frac{\binom{13}{4}}{\binom{18}{4}},\ \mathbb{P}(B) = \frac{\binom{5+7}{4}}{\binom{18}{4}} = \frac{\binom{12}{4}}{\binom{18}{4}},\ \mathbb{P}(Z) = \frac{\binom{5+6}{4}}{\binom{18}{4}} = \frac{\binom{11}{4}}{\binom{18}{4}},$$

$$\mathbb{P}(C \cap B) = \frac{\binom{7}{4}}{\binom{18}{4}},\ \mathbb{P}(C \cap Z) = \frac{\binom{6}{4}}{\binom{18}{4}},\ \mathbb{P}(B \cap Z) = \frac{\binom{5}{4}}{\binom{18}{4}},$$

$$\mathbb{P}(C \cap B \cap Z) = 0.$$

Uvrštavanjem dobivamo da je

$$\mathbb{P}(C \cup B \cup Z) = \frac{1485}{3060} = \frac{33}{68} = \approx 0.4853.$$

Alternativno, primijetimo da je

$$\begin{aligned}(C \cup B \cup Z)^c &= \{\text{zastupljene sve tri boje}\} \\ &= \{\text{2C,1B,1Z}\} \cup \{\text{1C,2B,1Z}\} \cup \{\text{1C,1B,2Z}\}.\end{aligned}$$

Budući da su ti događaji disjunktni,

$$\mathbb{P}(C \cup B \cup Z) = 1 - \frac{1}{\binom{18}{4}} \left[\binom{5}{2}\binom{6}{1}\binom{7}{1} + \binom{5}{1}\binom{6}{2}\binom{7}{1} + \binom{5}{1}\binom{6}{1}\binom{7}{2} \right] = \frac{33}{68}.$$

□

Napomena 2.9. Za proizvoljan $n \in \mathbb{N}$ i događaje $A_1, \ldots, A_n \in \mathcal{F}$ vrijedi sljedeće poopćenje FUI (Silvesterova formula)

$$\begin{aligned}\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) &= \sum_{i=1}^{n} \mathbb{P}(A_i) - \sum_{1 \le i < j \le n} \mathbb{P}(A_i \cap A_j) + \sum_{1 \le i < j < k \le n} \mathbb{P}(A_i \cap A_j \cap A_k) \\ &\quad - \cdots + (-1)^{n+1} \mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right).\end{aligned}$$

Zadatak 2.10. Jednog radnog tjedna u nekom gradu 10 ljudi je pozvalo električara u svoje kuće radi popravka nekih električnih uređaja. Svaki čovjek je slučajno odabrao neki radni dan u tom tjednu i pozvao električara. Izračunajte vjerojatnost da električar nema posla bar jedan radni dan.

Rješenje. Prostor elementarnih događaja je

$$\Omega = \{(x_1, \dots, x_{10}) : x_k \in \{1,2,3,4,5\}\} = \{1,2,3,4,5\}^{10}$$

pri čemu x_k predstavlja dan u tjednu u koji je k-ta osoba pozvala električara. Dakle, $k(\Omega) = 5^{10}$. Neka je A traženi događaj te

$$A_i = \{\text{ i-ti radni dan nitko nije zvao električara}\},\ i = 1, \dots 5\,.$$

Zbog simetrije za sve različite i, j, k, l vrijedi

$$\mathbb{P}(A_i) = \frac{4^{10}}{5^{10}},\ \mathbb{P}(A_i \cap A_j) = \frac{3^{10}}{5^{10}},\ \mathbb{P}(A_i \cap A_j \cap A_k) = \frac{2^{10}}{5^{10}},$$

$$\mathbb{P}(A_i \cap A_j \cap A_k \cap A_l) = \frac{1^{10}}{5^{10}},\ \mathbb{P}(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5) = 0.$$

Koristeći poopćenje FUI za uniju 5 događaja slijedi da je

$$\begin{aligned}
\mathbb{P}(A) &= \mathbb{P}(A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5) \\
&= \binom{5}{1} \cdot \mathbb{P}(A_1) - \binom{5}{2} \cdot \mathbb{P}(A_1 \cap A_2) + \binom{5}{3} \cdot \mathbb{P}(A_1 \cap A_2 \cap A_3) \\
&- \binom{5}{4} \cdot \mathbb{P}(A_1 \cap A_2 \cap A_3 \cap A_4) + \mathbb{P}(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5) \\
&= \frac{5 \cdot 4^{10} - 10 \cdot 3^{10} + 10 \cdot 2^{10} - 5}{5^{10}} \approx 0.4775.
\end{aligned}$$

□

2.1 Beskonačni vjerojatnosni prostor

Zadatak 2.11. Da bi počeli igrati s čovječuljkom u igri "Čovječe ne ljuti se" morate prvo dobiti šesticu na kocki. Izračunajte vjerojatnost

(a) da u trećem pokušaju prvi puta dobijete šesticu,

(b) da vam treba više od tri pokušaja da prvi puta dobijete šesticu.

Rješenje. Prostor elementarnih događaja je

$$\begin{aligned}
\Omega = &\{(x_1, \dots, x_{n-1}, 6) : x_1, \dots, x_{n-1} \in \{1, 2, \dots, 5\}, n \in \mathbb{N}\} \\
&\cup \{(x_1, x_2, \dots) : x_i \in \{1, 2, \dots, 5\} \text{ za sve } i \in \mathbb{N}\}.
\end{aligned}$$

Uočimo, $k(\Omega) = \infty$.

(a) Traži se vjerojatnost događaja

$$A = \{(x_1, x_2, 6) : x_1, x_2 \in \{1, 2, \dots, 5\}\}.$$

Budući da događaj A ovisi samo prva tri bacanja, dovoljno je promatrati samo njih. Svi mogući ishodi u prva tri bacanja su

$$\tilde{\Omega} = \{(x_1, x_2, x_3) : x_1, x_2, x_3 \in \{1, 2, \dots, 6\}\} = \{1, 2, \dots, 6\}^3$$

pa je

$$\mathbb{P}(A) = \frac{k(A)}{k(\tilde{\Omega})} = \frac{5 \cdot 5 \cdot 1}{6 \cdot 6 \cdot 6} = \frac{25}{216} \approx 0.1157.$$

(b) Ako za $n \in \mathbb{N}$ definiramo događaje

$$\begin{aligned} B_n &= \{\text{potrebno točno } n \text{ pokušaja da se dobije šestica}\} \\ &= \{(x_1, \dots, x_{n-1}, 6) : x_k \in \{1, 2, \dots, 5\}, k = 1, \dots, n-1\}, \end{aligned}$$

tada se traži vjerojatnost događaja

$$B = \bigcup_{n=4}^{\infty} B_n.$$

Analogno kao u (a) dijelu zadatka, za $n \in \mathbb{N}$ je

$$\mathbb{P}(B_n) = \frac{5^{n-1} \cdot 1}{6^n}$$

pa budući da su događaji B_n međusobno disjunktni slijedi da je

$$\begin{aligned} \mathbb{P}(B) &= \mathbb{P}\Big(\bigcup_{n=4}^{\infty} B_n\Big) = \sum_{n=4}^{\infty} \mathbb{P}(B_n) = \sum_{n=4}^{\infty} \frac{5^{n-1}}{6^n} = \frac{1}{6} \cdot \left(\frac{5}{6}\right)^3 \cdot \sum_{k=0}^{\infty} \left(\frac{5}{6}\right)^k \\ &= \frac{1}{6} \cdot \left(\frac{5}{6}\right)^3 \frac{1}{1 - \frac{5}{6}} = \frac{125}{216} \approx 0.5787. \end{aligned}$$

□

Zadatak 2.12. U kutiji se nalazi 10 crvenih, 8 bijelih i 5 plavih kuglica. Na slučajan način izvučemo jednu kuglicu, pogledamo joj boju i potom je vratimo u kutiju. Postupak nastavljamo sve dok ne izvučemo crvenu ili bijelu kuglicu.

(a) Kolika je vjerojatnost da ćemo izvlačenje završiti s crvenom kuglicom?

(b) Kolika je vjerojatnost da igra nikad neće završiti?

Rješenje.

(a) Prostor elementarnih događaja je

$$\Omega = \{B, C, PB, PC, PPB, PPC, \dots\} \cup \{PPPPP\cdots\}.$$

Tražimo

$$\mathbb{P}\left(\{\text{crvena izvučena prije bijele}\}\right) = \mathbb{P}\left(\{C, PC, PPC, \dots\}\right) = \sum_{n=0}^{\infty} \mathbb{P}\left(\{\underbrace{PP\cdots P}_{n \text{ puta}}C\}\right),$$

pri čemu zadnja jednakost slijedi jer su imamo prebrojivo mnogo međusobno disjunktnih događaja. Računamo,

$$\mathbb{P}\left(\{C\}\right) = \frac{10}{23}, \; \mathbb{P}\left(\{PC\}\right) = \frac{5 \cdot 10}{23 \cdot 23}, \; \mathbb{P}\left(\{PPC\}\right) = \frac{5 \cdot 5 \cdot 10}{23 \cdot 23 \cdot 23} = \left(\frac{5}{23}\right)^2 \frac{10}{23}$$

i općenito

$$\mathbb{P}\left(\{\underbrace{PP\cdots P}_{n \text{ puta}}C\}\right) = \left(\frac{5}{23}\right)^n \frac{10}{23}, \quad n \in \mathbb{N}_0.$$

Dakle,

$$\mathbb{P}\left(\{\text{crvena izvučena prije bijele}\}\right) = \sum_{n=0}^{\infty} \left(\frac{5}{23}\right)^n \frac{10}{23} = \frac{10}{23} \cdot \frac{1}{1 - \frac{5}{23}} = \frac{5}{9}.$$

Ako malo razmislimo, uz pretpostavku da znamo da ćemo izvlačenje sigurno završiti u konačnom broju koraka, rezultat iz (a) dijela je intuitivno jasan jer "ako znamo" da smo u nekom trenutku izvukli crvenu ili bijelu kuglicu, vjerojatnost da smo baš izvukli crvenu je uvijek $\frac{10}{10+8} = \frac{5}{9}$.

(b) Označimo sa $A_1 = \{\text{crvena izvučena prije bijele}\}$, $A_2 = \{\text{bijela izvučena prije crvene}\}$. Analogno se dobije $\mathbb{P}\left(\{A_2\}\right) = \frac{4}{9}$. Nas zanima vjerojatnost događaja $\{PPPP\cdots\} = (A_1 \cup A_2)^C$ pa je

$$\mathbb{P}\left(\{PPPP\cdots\}\right) = \mathbb{P}\left((A_1 \cup A_2)^C\right) = 1 - \mathbb{P}\left(A_1 \cup A_2\right) = 1 - \mathbb{P}\left(A_1\right) - \mathbb{P}\left(A_2\right) = 0$$

□

Zadatak 2.13. Dva igrača naizmjence bacaju novčić, a pobjeđuje onaj kod kojeg se prvog pojavi grb. Modelirajte odgovarajući vjerojatnosni prostor i izračunajte

(a) vjerojatnost pobjede za svakog igrača,

(b) vjerojatnost da igra nikada ne stane.

Rješenje. Prostor elementarnih događaja je

$$\Omega = \{G, PG, PPG, \dots\} \cup \{PPPPP\cdots\}.$$

Uočimo,

$$\mathbb{P}(\{G\}) = \frac{1}{2},\ \mathbb{P}(\{PG\}) = \frac{1\cdot 1}{2\cdot 2},\ \mathbb{P}(\{PPG\}) = \frac{1\cdot 1\cdot 1}{2\cdot 2\cdot 2} = \frac{1}{2^3}$$

i općenito

$$\mathbb{P}(\{\underbrace{PP\cdots P}_{n \text{ puta}}G\}) = \frac{1}{2^{n+1}}\,, \quad n \in \mathbb{N}_0$$

(a) Neka je $A_i = \{\text{pobijedio je i-ti igrač}\}$, $i = 1, 2$. Tada je

$$\begin{aligned}\mathbb{P}(A_1) &= \mathbb{P}(\{G, PPG, PPPPG, \dots\}) = \sum_{n=0}^{\infty} \mathbb{P}(\{\underbrace{PP\cdots P}_{2n \text{ puta}}G\}) \\ &= \sum_{n=0}^{\infty} \frac{1}{2^{2n+1}} = \frac{1}{2}\sum_{n=0}^{\infty}\frac{1}{4^n} = \frac{1}{2}\cdot\frac{1}{1-\frac{1}{4}} = \frac{2}{3}\,.\end{aligned}$$

Na sličan način dobijemo da je $\mathbb{P}(A_2) = \frac{1}{3}$.

(b) Koristeći dio (a) dobivamo

$$\begin{aligned}\mathbb{P}(\{\text{igra nikada ne stane}\}) &= \mathbb{P}((A_1 \cup A_2)^c) = 1 - \mathbb{P}(A_1 \cup A_2) \\ &= [A_1 \text{ i } A_2 \text{ disjunktni}] = 1 - \mathbb{P}(A_1) - \mathbb{P}(A_2) = 0\,.\end{aligned}$$

□

Napomena 2.14. Ako za neki događaj $A \in \mathcal{F}$ vrijedi $\mathbb{P}(A) = 1$ (odnosno $\mathbb{P}(A) = 0$), kažemo da će se događaj A *gotovo sigurno* (g.s.) dogoditi (odnosno neće dogoditi).

Zadaci za vježbu

Zadatak 2.15. Leon bira šifru na mobitelu. Ako se šifra sastoji od 4 broja koja je vjerojatnost:

a) da su sve znamenke neparne?

b) da su sve znamenke različite?

c) da su sve znamenke neparne ili sve znamenke različite?

d) da su sve znamenke neparne i sve znamenke različite?

Zadatak 2.16. Iz snopa od 52 karte biramo na sreću 3 karte. Odredite vjerojatnost da dobijemo:

(a) točno jednog kralja,

(b) barem jednog asa,

(c) sve karte različite boje,

(d) dvije karte iste boje i dvije karte iste jačine.

Zadatak 2.17. Na peronu je vlak koji se sastoji od 15 vagona. Ako 7 putnika nasumce bira vagon, izračunajte vjerojatnost

(a) da je u svakom vagonu najviše jedan putnik,

(b) da je u zadnjem vagonu točno jedan putnik.

Zadatak 2.18. U nekom se kraljevstvu organizira viteški turnir. Dan prije na turnir je došlo 5 vitezova. Netko je preko noći na slučajan način vitezovima izmiješao koplja. Kolika je vjerojatnost da je barem jedan vitez na turniru nastupio sa svojim kopljem?

Zadatak 2.19. U liftu zgrade s 5 katova nalazi se 7 osoba. Izračunajte vjerojatnost da

(a) na prvom katu izađu točno 3 osobe,

(b) na svakom katu izađe barem jedna osoba.

Zadatak 2.20. Simetričnu kocku bacamo dok se prvi put ne pojavi 1 ili 6. Odredite odgovarajući vjerojatnosni prostor i izračunajte vjerojatnost da će pokus završiti u parnom broju koraka.

Zadatak 2.21. Oskar je za rođendan dobio sva tri nastavka "Gospodara prstenova". Kako je bio u žurbi, samo ih je nasumce stavio na svoju praznu policu zajedno s četiri stare knjige. Kolika je vjerojatnost da tri nastavka "Gospodara prstenova" čine jedan blok na polici (tj. da između njih nema knjiga)?

3 Geometrijska vjerojatnost

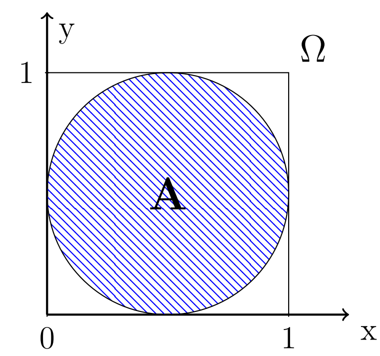
Želimo modelirati vjerojatnosni prostor motiviran Laplaceovim modelom koji odgovara biranju točke na slučajan način iz nekog ograničenog skupa $\Omega \subseteq \mathbb{R}^n$ (pretpostavimo $n = 1$, 2 ili 3). U tu svrhu, za ograničen skup $A \subseteq \mathbb{R}^n$ označimo s $\lambda(A)$ njegovu duljinu, površinu, tj. volumen, ovisno o kojem je $n \in \{1, 2, 3\}$ riječ. Imamo za

$$\begin{array}{lrl} n = 1 & \text{duljinu} & \lambda([0,1]) = 1 - 0 = 1 \\ n = 2 & \text{površinu} & \lambda([0,1] \times [2,3]) = (1-0)(3-2) = 1 \\ n = 3 & \text{volumen} & \lambda([0,1]^3) = (1-0)^3 = 1. \end{array}$$

Modeliramo vjerojatnosni prostor $(\Omega, \mathcal{F}, \mathbb{P})$ koji odgovara slučajnom odabiru točke iz $\Omega \subseteq \mathbb{R}^n$ (umjesto ograničenosti od Ω, zapravo je potrebno zahtijevati da je $\lambda(\Omega) < \infty$). Dakle, Ω je prostor elementarnih događaja, $\mathcal{F}$ će biti tzv. *Borelova* σ-algebra na Ω, tj. σ-algebra koja sadrži sve podskupove od Ω kojima "možemo izmjeriti" površinu (jer ne možemo svakom), a $\mathbb{P}$ je vjerojatnost definirana tako da za svaki $A \in \mathcal{F}$ vrijedi

$$\mathbb{P}(A) = \frac{\lambda(A)}{\lambda(\Omega)}.$$

Primjer 3.1. Ako slučajno odaberemo točku iz kvadrata stranice duljine 1, kolika je vjerojatnost da ta točka upadne u krug upisan tom kvadratu?



Slika 1: Slika za Primjer 3.1

U ovom slučaju možemo uzeti

$$\Omega = [0,1]^2$$

te staviti $A = \{\text{krug upisan u } [0,1]^2\} \subseteq \Omega$, pa je tražena vjerojatnost

$$\mathbb{P}(A) = \frac{\lambda(A)}{\lambda(\Omega)} = \frac{\left(\frac{1}{2}\right)^2 \pi}{1^2} = \frac{\pi}{4}.$$

Napomena 3.2. Više o funkciji λ koja računa površinu i σ-algebri $\mathcal{F}$ koja sadrži skupove koji imaju površinu možete naučiti na kolegiju *Mjera i integral.* Istaknimo samo da je u ovom slučaju $\mathcal{F} \subsetneq \mathcal{P}(\Omega)$.

Za razliku od od Laplacevog modela, u ovom slučaju postoje neprazni događaji $B \in \mathcal{F}$ takvi da je $\mathbb{P}(B) = 0$.

Zadatak 3.3. Iz segmenta $[0, 1]$ slučajno se i nezavisno biraju dva broja x i y. Ako su $a, b, c, d \in [0, 1]$ fiksni, izračunajte vjerojatnost događaja:

(a) $A = \{$izabrani su brojevi $x = a$ i $y = b\}$,

(b) $B = \{$izabrani su brojevi takvi da je $x = a$ ili $x = c$ i $y = b$ ili $y = d\}$

(c) $C = \{$izabran je broj $x = a\}$,

(d) $D = \{$izabrani su isti brojevi$\}$,

(e) $E = \{$prvi broj je manji ili jednak od drugoga$\}$.

Rješenje. Prostor elementarnih događaja (tj. skup svih ishoda ovog pokusa) je $\Omega = \{(x, y) \,|\, x, y \in [0, 1]\} = [0, 1]^2$. Budući da točke $x, y \in [0, 1]$ biramo "nezavisno" jednu od druge to zapravo odgovara slučajnom odabiru točke iz kvadrata $\Omega = [0, 1]^2$. Dakle, nalazimo se u situaciji opisanoj u uvodu ovog poglavlja. Uočimo da je $\lambda(\Omega) = \lambda([0, 1]^2) = 1$.

(a) Imamo $A = \{(a, b)\}$ pa budući da svaka točka u ravnini ima površinu 0, vrijedi

$$\mathbb{P}(A) = \frac{\lambda(A)}{\lambda(\Omega)} = \frac{\lambda(\{(a, b)\})}{\lambda(\Omega)} = \frac{0}{1} = 0\,.$$

Dakle, vjerojatnost da ćemo izabrati unaprijed fiksiranu točku (a, b) je jednaka 0.

(b) Imamo $B = \{(a, b), (a, d), (c, b), (c, d)\}$. Budući da svaka točka u ravnini ima površinu 0 i četiri točke će imati površinu 0 pa je $\mathbb{P}(B) = 0$

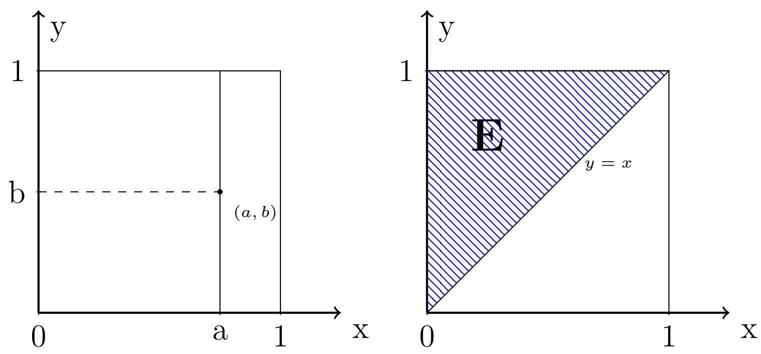
(c) i (d) Imamo $C = \{(x, y) \in \Omega \,|\, x = a\}$ i $D = \{(x, y) \in \Omega \,|\, x = y\}$ pa budući da i svaka dužina u ravnini također ima površinu 0 ponovno je

$$\mathbb{P}(C) = \frac{\lambda(C)}{\lambda(\Omega)} = \frac{0}{1} = 0 = \mathbb{P}(D)\ .$$

(e) Događaj $E = \{(x, y) \in \Omega \,|\, x \leq y\}$ nacrtan je desno na Slici 2 i očito je

$$\mathbb{P}(E) = \frac{\lambda(E)}{\lambda(\Omega)} = \frac{1/2}{1} = \frac{1}{2}\,,$$

što je intuitivno odmah bilo jasno zbog simetrije.



Slika 2: Slika Zadatka 3.3, prva slika za a), b) i c) dio, druga za d) i e) dio.

$\square$

Zadatak 3.4. Izračunajte vjerojatnost da je slučajno odabran broj iz segmenta $[0, 1]$

(a) racionalan,

(b) iracionalan.

Rješenje.

(a) Imamo da je $\Omega = [0, 1]$, a traženi događaj $A = \mathbb{Q} \cap [0, 1]$. Zbog prebrojivosti skupa $\mathbb{Q}$, postoji niz $(q_n)_n$ takav da je $A = \{q_1, q_2, q_3, \dots\}$ pa je

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{n=1}^{\infty}\{q_n\}\right) \overset{disjun.}{=} \sum_{n=1}^{\infty}\mathbb{P}(\{q_n\}) = \sum_{n=1}^{\infty} 0 = 0.$$

Primijetimo da je ključna stvar u računu bila da je $\mathbb{Q}$ prebrojiv.

(b) 1

$\square$

Zadatak 3.5. Dana je jednadžba $x^2 + 2bx + c = 0$ pri čemu su b i c slučajno odabrani brojevi takvi da je $|b| \leq 3$, a $|c| \leq 12$. Kolika je vjerojatnost da su rješenja gornje jednadžbe realna?

Rješenje. Prostor elementarnih događaja je

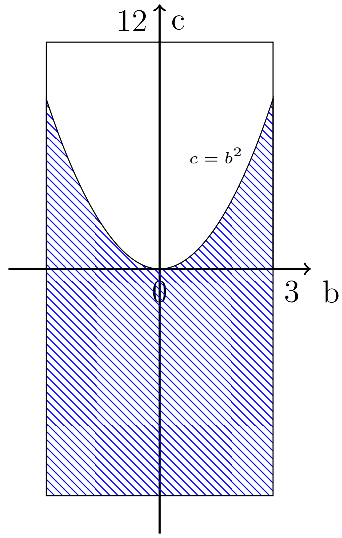
$$\Omega = \{(b, c) \,|\, -3 \leq b \leq 3, -12 \leq c \leq 12\} = [-3, 3] \times [-12, 12]\,,$$

pa je $\lambda(\Omega) = 6 \cdot 24 = 144$. Tražimo vjerojatnost događaja

$$\begin{aligned} A &= \{\text{rješenja jednadžbe su realna}\} \\ &= \{(b, c) \in \Omega \,|\, 4b^2 - 4c \geq 0\} \\ &= \{(b, c) \in \Omega \,|\, b^2 \geq c\}\,, \end{aligned}$$

jer kvadratna jednadžba ima realna rješenja ako i samo ako joj je diskriminanta nenegativna. Sada lako računamo da je

$$\mathbb{P}(A) = \frac{\lambda(A)}{\lambda(\Omega)} = \frac{6 \cdot 12 + \int_{-3}^{3} b^2 \, \mathrm{d}b}{144} = \frac{5}{8}.$$



Slika 3: Slika za Zadatak 3.5

□

Zadatak 3.6. Profesor i student dolaze u učionicu s jednakom vjerojatnosti bilo kada između 8 i 8:30. Student će pričekati profesora najviše 15 minuta prije nego napusti učionicu, a profesor će zatvoriti vrata učionice 2 minute nakon svog dolaska u učionicu i više se neće moći ući i slušati to iznimno bitno profesorovo predavanje. Koja je vjerojatnost da student nazoči profesorovom predavanju?

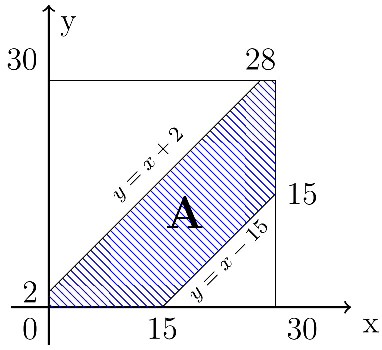
Rješenje. Za prostor elementarnih događaja možemo uzeti

$$\Omega = \{(x, y) \,|\, x, y \in [0, 30]\} = [0, 30]^2 \,,$$

gdje x smatramo vremenom dolaska profesora, a y vremenom dolaska studenta (oboje u minutama nakon 8). Očito je $\lambda(\Omega) = 30^2$.

Neka je $A = \{\text{student nazoči predavanju}\}$ traženi događaj. Student će biti na predavanju ako i samo ako je došao najkasnije 2 minute nakon profesora, a najranije 15 minuta prije njega, tj. $(x, y) \in A$ ako samo ako je $x - 15 \leq y \leq x + 2$. Dakle, $A = \{(x, y) \in \Omega \,|\, x - 15 \leq y \leq x + 2\}$ te

$$\begin{aligned} \mathbb{P}\,(A) &= \frac{\lambda(A)}{\lambda(\Omega)} = \frac{\lambda(\Omega) - \lambda(A^c)}{\lambda(\Omega)} \\ &= \frac{\lambda(\Omega) - \frac{15 \cdot 15}{2} - \frac{28 \cdot 28}{2}}{\lambda(\Omega)} \approx 0.439. \end{aligned}$$



Slika 4: Slika za Zadatak 3.6

□

Zadatak 3.7. Potrošnja Markovog auta je slučajno odabrana vrijednost između 5 i 10 litara benzina po 100 km, a trenutno mu je ostalo samo 2.5 litre u rezervoaru. Ako je udaljenost do najbliže benzinske postaje slučajna veličina između 20 km i 50 km, kolika je vjerojatnost da će s trenutnom količinom benzina uspjeti stići do benzinske postaje?

Rješenje. Neka je x predstavlja potrošnju auta po 100 km (dakle $x \in [5, 10]$), a y udaljenost do najbliže benzinske postaje (dakle $y \in [20, 50]$). Prostor elementarnih događaja je

$$\Omega = \{(x, y) \,|\, x \in [5, 10], y \in [20, 50]\} = [5, 10] \times [20, 50]$$

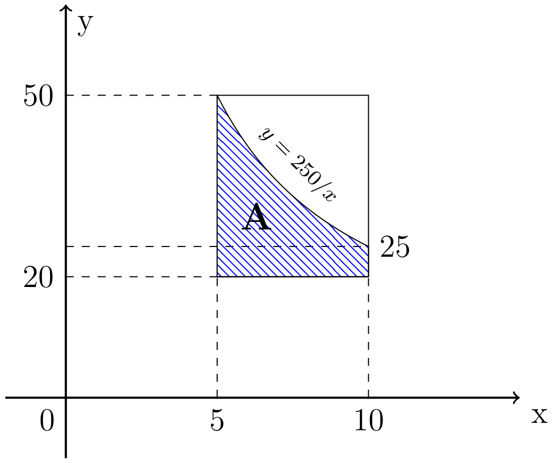
i očito je $\lambda(\Omega) = 5 \cdot 30 = 150$.

Primijetimo da s $\frac{x}{100}$ litara benzina auto može prijeći 1 km pa je potrebno $\frac{x}{100} \cdot y$ litara do benzinske. Budući da Marko ima 2.5 litre u rezervoaru, imamo

$$\begin{aligned} A &= \{\text{Marko je došao do benzinske}\} \\ &= \{(x, y) \in \Omega \,|\, \tfrac{x}{100} \cdot y \leq 2.5\} \\ &= \{(x, y) \in \Omega \,|\, y \leq \tfrac{250}{x}\}. \end{aligned}$$

Dakle,

$$\begin{aligned} \mathbb{P}(A) &= \frac{\lambda(A)}{\lambda(\Omega)} = \frac{\int_5^{10} \frac{250}{x} \, dx - 5 \cdot 20}{150} \\ &= \frac{250 \cdot \ln(x)|_5^{10} - 100}{150} = \frac{250 \cdot \ln 2 - 100}{150} \approx 0.48858. \end{aligned}$$



□

Napomena 3.8 (*)**.** Primijetimo da u gornjem zadatku imamo slučajnu vrijednost $z = \frac{1}{x}$ koja poprima vrijednosti u intervalu $[1/10, 1/5]$. Netko bi možda pokušao riješiti zadatak tako da za prostor elementarnih događaja uzme $\Omega' = \{(z, y) \,|z \in [1/10, 1/5], y \in [20, 50]\}$ jer je u tom slučaju površinu skupa (događaja) $A = \{(z, y) \in \Omega' \,|\, y \leq 250z\}$ jednostavno odrediti. Ipak z "ne poprima sa **istom vjerojatnosti** svaku vrijednost" iz $[1/10, 1/5]$ (kasnije u kolegiju ćemo preciznije reći da z nema uniformnu razdiobu na $[1/10, 1/5]$) pa u ovom slučaju ne možemo koristiti vjerojatnosni model opisan u uvodu ovog poglavlja. Zaista, kada bi z imala uniformu razdiobu, vjerojatnost da z upadne u $[1/10, 3/20]$, tj. prvu polovicu segmenta $[1/10, 1/5]$, bila bi $1/2$. Ipak, budući da je $z = 1/x$ ta vjerojatnost jednaka je vjerojatnosti da x upadne u segment $[20/3, 10]$ pa zapravo iznosi $2/3$.

Zadatak 3.9. Unutar intervala $[0, 1]$ biramo na sreću dva broja: x i y. Odredite vjerojatnost da je broj $\lfloor 5x + y \rfloor$ djeljiv s 3 gdje je $\lfloor x \rfloor$ najveći cijeli broj koji je manji ili jednak od x.

Rješenje. Imamo $\Omega = \{(x, y) \,|\, x, y \in [0, 1]\} = [0, 1]^2$ te uočimo da vrijedi

$$(x, y) \in \Omega \Rightarrow 5x + y \in [0, 6] \Rightarrow \lfloor 5x + y \rfloor \in \{0, 1, \ldots, 6\}\,.$$

Iz ovoga slijedi da je traženi događaj

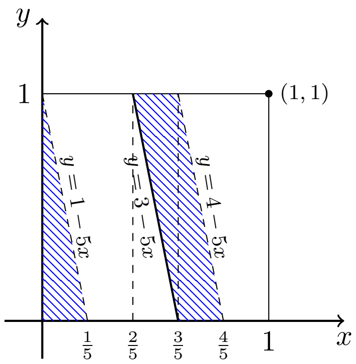
$$A = \{(x, y) \in \Omega \,|\, \lfloor 5x + y \rfloor \text{ djeljiv s } 3\} = \{(x, y) \in \Omega \,|\, \lfloor 5x + y \rfloor \in \{0, 3, 6\}\}$$

pa je

$$\begin{aligned} A &= \{(x, y) \in \Omega \,|\, 0 \leq 5x + y < 1 \text{ ili } 3 \leq 5x + y < 4 \text{ ili } 5x + y = 6\} \\ &= \{(x, y) \in \Omega \,|\, 0 \leq y < 1 - 5x \text{ ili } 3 - 5x \leq y < 4 - 5x \text{ ili } x = y = 1\}\,, \end{aligned}$$

gdje smo u prvoj nejednakosti u zadnjem retku izbacili $-5x$ jer je za sve $(x, y) \in \Omega$ uvijek $-5x \leq 0 \leq y$. Sada slijedi da je

$$\mathbb{P}\,(A) = \frac{\lambda(A)}{\lambda(\Omega)} = \frac{\frac{1 \cdot \frac{1}{5}}{2} + \frac{1}{5} \cdot 1 + 0}{1} = \frac{3}{10}$$



Uočite da nije bitno koji su rubovi uključeni u područje, a koji nisu jer svaki rub ima površinu 0. □

Zadatak 3.10. Štap duljine L razlomljen je na tri dijela tako da smo slučajno i nezavisno izabrali dvije točke prijeloma. Izračunajte vjerojatnost da je duljina najduljeg dijela veća od $\frac{2}{5}L$.

Rješenje. Biranje dvije točke prijeloma $x, y \in [0, L]$ odgovara slučajnom izboru uređenog para $(x, y) \in [0, L]^2$. Kako bi riješili zadatak trebamo rastaviti na slučajeve ovisno o tome koja je od točaka prvi prijelom štapa, tj. trebali bi rastaviti na slučajeve $x \leq y$ i $y < x$. Ipak, zbog simetrije možemo pretpostaviti npr. da je $x \leq y$, tj. dovoljno je izračunati vjerojatnost traženog događaja ako je prostor elementarnih događaja

$$\Omega = \{(x, y) \,|\, 0 \leq x \leq y \leq L\} .$$

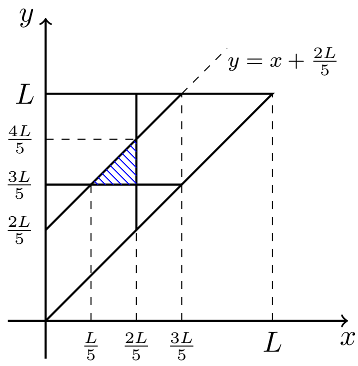
Očito je $\lambda(\Omega) = \frac{L^2}{2}$. Nadalje, zbog $x \leq y$ imamo da je prvi odlomljeni dio štapa duljine x, drugi duljine $y - x$, a treći $L - y$. Ako je $A = \{\text{najdulji odlomljeni dio dulji je od } \frac{2}{5}L\}$ traženi događaj, primijetimo da je

$$\begin{aligned} A^c &= \{\text{najdulji odlomljeni dio kraći je od } \tfrac{2}{5}L\} = \{\text{svi dijelovi kraći su od } \tfrac{2}{5}L\} \\ &= \{(x, y) \in \Omega \,|\, x \leq \tfrac{2}{5}L, y - x \leq \tfrac{2}{5}L, L - y \leq \tfrac{2}{5}L\} . \end{aligned}$$

Slijedi

$$\mathbb{P}(A^c) = \frac{\lambda(A^c)}{\lambda(\Omega)} = \frac{(L/5)^2/2}{L^2/2} = \frac{1}{25} .$$

Dakle, $\mathbb{P}(A) = 1 - \mathbb{P}(A^c) = \frac{24}{25}$.



Uvjerimo se na kraju da je ovo zaista točno rješenje. Kada bi za prostor elementarnih događaja uzeli $\Omega' = [0, L]^2$ te ako je $A' = \{\text{najdulji odlomljeni dio dulji je od } \frac{2}{5}L\} \subseteq \Omega'$, tada je očito $\lambda(\Omega') = 2\lambda(\Omega)$, a zbog simetrije $\lambda(A') = 2\lambda(A)$ (provjerite!). Dakle $\mathbb{P}'(A') = \frac{\lambda(A')}{\lambda(\Omega')} = \frac{2\lambda(A)}{2\lambda(\Omega)} = \frac{\lambda(A)}{\lambda(\Omega)} = \mathbb{P}(A)$. $\square$

Zadatak 3.11. Na kružnici polumjera r slučajno i nezavisno odabrane su tri točke: A, B i C. Izračunajte vjerojatnost da je trokut određen tim točkama: (a) šiljastokutan, (b) pravokutan, (c) tupi.

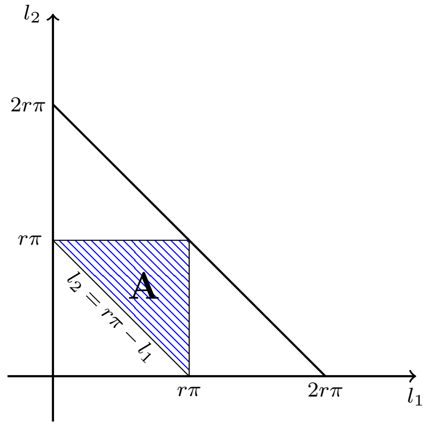
Rješenje. Ako fiksiramo neku (bilo koju) točku kružnice kao početnu, slučajan izbor točaka A, B i C možemo dobiti kao slučajan izbor uređene trojke iz kocke $[0, 2r\pi]^3$ pri čemu svaka koordinata predstavlja duljinu luka (npr. obrnuto od smjera kazaljke na satu) od početne točke kružnice do točaka A, B i C. Ipak, zbog simetrije, možemo fiksirati jednu od točaka za početnu, npr. A, i gledati samo koje su duljine lukova do B i C. Neka je $l_1 \in [0, 2r\pi]$ duljina luka do B obrnuto od smjera kazaljke na satu, a $l_2 \in [0, 2r\pi]$ duljina luka do C u smjeru kazaljke na satu. Slično kao i u Zadatku 3.10, zbog simetrije dovoljno je gledati samo slučaj $l_1 + l_2 \leq 2r\pi$, tj. kada, gledajući suprotno od smjera kazaljke na satu, nakon točke A prvo dolazi točka B (nacrtajte!). Dakle, za prostor elementarnih događaja možemo uzeti

$$\Omega = \{(l_1, l_2) \in [0, 2r\pi]^2 \mid l_1 + l_2 \leq 2\pi r\}.$$

Uočite da je duljina luka od B do C jednaka $2r\pi - l_1 - l_2$.

(a) Definiramo događaj $A = \{\text{trokut } \triangle ABC \text{ je šiljastokutan}\}$. Znamo da je trokut šiljastokutan ako i samo ako je svaki luk na njemu opisanoj kružnici manji od poluopsega pa je $A = \{(l_1, l_2) \in \Omega \mid l_1 < r\pi, l_2 < r\pi, 2r\pi - l_1 - l_2 < r\pi\}$ te

$$\mathbb{P}(A) = \frac{\lambda(A)}{\lambda(\Omega)} = \frac{\frac{(r\pi)^2}{2}}{\frac{(2r\pi)^2}{2}} = \frac{1}{4}.$$



(b) Traženi događaj je

$$\begin{aligned} B &= \{\text{trokut } \triangle ABC \text{ je pravokutan}\} \\ &= \{(l_1, l_2) \in \Omega \,|\, l_1 = r\pi \text{ ili } l_2 = r\pi \text{ ili } 2r\pi - l_1 - l_2 = r\pi\}, \end{aligned}$$

tj. $B = \partial A$, pa je $\lambda(B) = 0 = \mathbb{P}(B)$.

(c) Koristeći prethodna dva dijela dobivamo da je $\mathbb{P}(\{\text{trokut } \triangle ABC \text{ je tupi}\}) = 1 - \mathbb{P}(A) - \mathbb{P}(B) = \frac{3}{4}$.

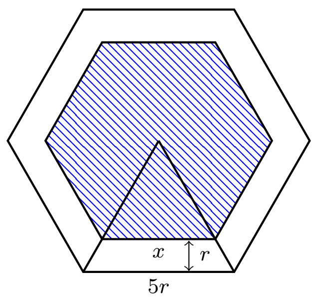
□

Zadatak 3.12. Novčić polumjera r bačen je na slučajan način na parkiralište popločano pravilnim šesterokutima čije su stranice duge $5r$. Izračunajte vjerojatnost da novčić ne siječe rub niti jednog šesterokuta.

Rješenje. Uočimo, gdje god da je novčić pao, ishod je da je njegovo središte u nekom od šesterokuta stranice $5r$. Zbog simetrije dovoljno je promatrati samo jedan šesterokut i reći da je slučajno odabrana točka unutar tog šesterokuta središte slučajno bačenog novčića. Dakle, za prostor elementarnih događaja možemo uzeti $\Omega = \{\text{šesterokut stranice } 5r\}$. Tražimo vjerojatnost događaja $A = \{\text{novčić ne siječe rub šesterokuta}\}$.

Uočimo, novčić ne siječe rub šesterokuta ako i samo ako je središte novčića udaljeno za najmanje r od ruba šesterokuta. Dakle, područje A je novi manji šesterokut u kojem je udaljenost središta do stranice za r manja nego u velikom šesterokutu. Ako sa x označimo duljinu stranice manjeg šesterokuta, nakon kraćeg računa dobijemo da je tražena vjerojatnost

$$\begin{aligned} \mathbb{P}(A) &= \frac{\lambda(A)}{\lambda(\Omega)} = \frac{6 \cdot x^2\sqrt{3}/4}{6 \cdot (5r)^2\sqrt{3}/4} = \left(\frac{x}{5r}\right)^2 \\ &= \left(\frac{5r\sqrt{3}/2 - r}{5r\sqrt{3}/2}\right)^2 = \left(1 - \frac{2\sqrt{3}}{15}\right)^2 \approx 0.59145. \end{aligned}$$



$\square$

Sljedeći problem je jedan od najstarijih problema iz područja geometrijske vjerojatnosti

Zadatak 3.13 (**Buffonov problem***)**.** Igla duljine l bačena je na stol koji je podijeljen paralelnim linijama međusobno udaljenim za d, pri čemu je $l \leq d$. Izračunajte vjerojatnost da igla siječe neku od paralelnih linija.

Rješenje. Za svaki ishod bacanja igle, označimo sa $\theta \in [0, \frac{\pi}{2}]$ šiljasti kut koji igla zatvara s paralelnim linijama, a sa $y \in [0, \frac{d}{2}]$ udaljenost centra igle do najbliže linije. Uočimo, igla siječe neku od linija ako i samo ako je

$$y \leq \frac{l}{2} \sin\theta \,.$$

Dakle, ako stavimo $\Omega = \{(\theta, y) \,|\, \theta \in [0, \frac{\pi}{2}],\ y \in [0, \frac{d}{2}]\}$, tada je

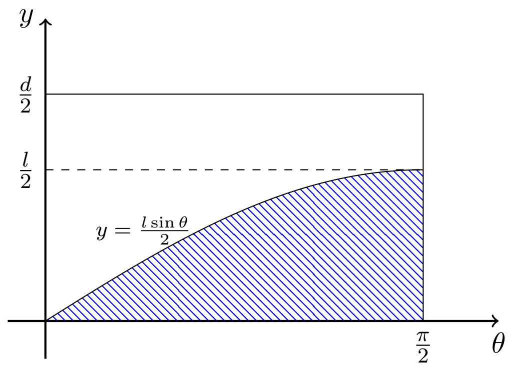
$$A = \{\text{igla siječe neku od linija}\} = \{(\theta, y) \in \Omega \,|\, y \leq \frac{l\sin\theta}{2}\}$$

pa je

$$\lambda(A) = \int_0^{\pi/2} \frac{l\sin\theta}{2} d\theta = \frac{l}{2} \,.$$

Napomenimo na kraju da, budući da je igla bačena slučajno, kut θ i udaljenost y su izabrani nezavisno (tj. kut igle nam ništa ne govori o udaljenosti do linija, i obratno) pa možemo koristiti model opisan u uvodu ovog poglavlja. Dakle,

$$\mathbb{P}(A) = \frac{\lambda(A)}{\lambda(\Omega)} = \frac{2}{\pi} \cdot \frac{l}{d} \,.$$



$\square$

Zadaci za vježbu

Zadatak 3.14. Dva broda moraju stići u isto pristanište. Vremena dolaska brodova su nezavisna i jednako vjerojatna u toku dana. Vrijeme zadržavanja prvog broda u pristaništu je 1 sat, drugog 2 sata.

(a) Odredite vjerojatnost da će jedan od brodova morati čekati na oslobađanje pristaništa.

(b) Odredite vjerojatnost da će brodovi doći u isto vrijeme.

Zadatak 3.15. Na žicu duljine 20 m izmedu dva telefonska stupa su slučajno i nezavisno sletjela 2 vrapca. Izračunajte vjerojatnost da je udaljenost vrabaca od stupova, kao i njihova međusobna udaljenost, barem 2 m.

Zadatak 3.16. Unutar intervala $[-2, 2]$ biramo na sreću dva broja. Odredite vjerojatnost da su apsolutna vrijednost njihove razlike i suma kvadrata veći od 1.

Rješenje. $\Omega = \{(x, y) \,|\, x, y \in [-2, 2]\} = [-2, 2]^2$.
$A = \{(x, y) \in \Omega \,|\, |x - y| > 1, x^2 + y^2 > 1\}$.
$\mathbb{P}(A) = \frac{\lambda(A)}{\lambda(\Omega)} = \frac{2\cdot(2^2 - \frac{1^2\cdot\pi}{4}) + 4\cdot\frac{1\cdot 1}{2}}{4^2} = \frac{5}{8} - \frac{\pi}{32} \approx 0.5268.$ □

Zadatak 3.17. Iz segmenta $[0, 1]$ slučajno su odabrani brojevi x i y. Izračunajte vjerojatnost da je $x + y \leq 1$ i $x \cdot y \leq \frac{2}{9}$.

Rješenje. $\Omega = [0, 1]^2$ pa je $\lambda(\Omega) = 1$, i $A = \{(x, y) \in \Omega \,|\, x + y \leq 1, x \cdot y \leq \frac{2}{9}\} = \{(x, y) \in \Omega \,|\, y \leq 1 - x, y \leq \frac{2}{9x}\}$. Sada tražimo točke A i B, tj. rješavamo jednadžbu $1 - x = \frac{2}{9x} \iff 9x^2 + 9x + 2 = 0 \iff (3x - 1)(3x - 2) = 0 \iff x = \frac{1}{3}$ ili $x = \frac{2}{3}$. Dakle,

$$\mathbb{P}(A) = \frac{\lambda(A)}{\lambda(\Omega)} = \int_0^{1/3} (1 - x)\,\mathrm{d}x + \int_{1/3}^{2/3} \frac{2}{9x}\,\mathrm{d}x + \int_{2/3}^{1} (1 - x)\,\mathrm{d}x = \frac{1}{3} + \frac{2}{9}\ln 2 \approx 0.4874.$$

□

Zadatak 3.18. Na kvadratično ispletenu mrežicu pada okomito s visine metalna kuglica. Ako je stranica kvadrata mrežice duga 10 mm, a promjer kuglice 5 mm, kolika je vjerojatnost da će kuglica proći kroz mrežicu, a da ne dotakne njezine niti?

Zadatak 3.19. Dva vlaka duljine 200 m kreću se brzinom od 1200 metara po minuti po prugama koje se međusobno sijeku. Vrijeme ulaska svakog od vlakova u raskrižje je slučajno, između 20h i 20 : 30h. Izračunajte vjerojatnost da se vlakovi sudare.

Zadatak 3.20 **(*).** Štap duljine 1 slučajno je razlomljen na 3 dijela. Kolika je vjerojatnost da je duljina najduljeg dijela veća od $\frac{2}{5}$ ako smo štap najprije prelomili na jednom mjestu pa zatim

(a) ostatak štapa (desno od prvog prijeloma) prelomili na jednom mjestu;

(b) uzeli dulji od dva dobivena dijela te njega prelomili na jednom mjestu?

(**Hint:** Za drugu varijabli uzmite postotak duljine drugog dijela kojeg smo prelomili. Usporedite s Zadatkom 3.10.)

Rješenje.

(a) Označimo s $x \in [0,1]$ poziciju gdje je prva točka prelomila štap, a s $y \in [0,1]$ poziciju gdje je druga točka prelomila štap. Dakle, prostor elementarnih događaja je $\Omega = \{(x,y)|0 \leq x \leq y \leq 1\}$. Nažalost, odabir točaka x i y na ovaj način ne odgovara slučajnom izboru uređenog para iz Ω. Zaista, budući da je x slučajan broj iz $[0,1]$, vjerojatnost da upadne u segment $[0,1/2]$ je očito $1/2$, a kada bi (x,y) bio slučajan par iz Ω imali bi da je vjerojatnost da x upadne u $[0,1/2]$ jednaka

$$\mathbb{P}\left(\{(x,y) \in \Omega | x \leq 1/2\}\right) = \frac{\lambda(\{(x,y) \in \Omega | x \leq 1/2\})}{\lambda(\Omega)} = \frac{3/8}{1/2} = \frac{3}{4}.$$

Dakle, ne možemo koristiti model opisan u uvodu ovog poglavlja.

Ipak, ako sa $p \in [0,1]$ označimo postotak duljine koji smo odlomili dijela štapa desno od prvog prijeloma, tada je izbor točaka x i p nezavisan. U tom slučaju su duljine dijelova štapa $x, p(1-x)$ te $(1-p)(1-x)$.

Ako je A traženi događaj, na isti se način, ali s malo težim računom, kao u Zadatku 3.10 izračuna da je $\lambda(A^c) = \frac{4}{5}\ln(\frac{4}{3}) - \frac{1}{5} = \mathbb{P}(A^c)$ iz čega slijedi da je $\mathbb{P}(A) = 1 - (\frac{4}{5}\ln(\frac{4}{3}) - \frac{1}{5}) \approx 0.9698$. Dakle, vjerojatnost traženog događaja je (malo) veća nego u Zadatku 3.10.

(b) U ovom slučaju zbog simetrije možemo pretpostaviti da je $x \leq 1/2$ te za prostor elementarnih događaja možemo uzeti $\Omega = \{(x,p)|x \in [0,1/2], p \in [0,1]\} = [0,1/2] \times [0,1]$. Ako je A traženi događaj, iz prethodnog dijela slijedi da je $\lambda(A^c) = \frac{4}{5}\ln(\frac{4}{3}) - \frac{1}{5}$ iz čega slijedi da je $\mathbb{P}(A) = 1 - 2 \cdot (\frac{4}{5}\ln(\frac{4}{3}) - \frac{1}{5}) \approx 0.9397$. Dakle, vjerojatnost traženog događaja je manja nego u Zadatku 3.10 što je intuitivno bilo i za očekivati.

$\square$

4 Uvjetna vjerojatnost i nezavisnost

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

Definicija 4.1. Neka su $A, B \in \mathcal{F}$ takvi da je $\mathbb{P}(B) > 0$. **Uvjetna vjerojatnost** od A uz dano B je

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

Primjer 4.2. Iz $\Omega = \{1, 2, \ldots, 20\}$ slučajno se izabire jedan broj. Ako znamo da je broj djeljiv s 3, kolika je vjerojatnost da je broj paran?

Intuitivno, ako znamo da smo izabrali neki od brojeva $3, 6, 9, 12, 15, 18$, izabrani će broj biti paran ako smo izabrali neki od brojeva $6, 12, 18$. Budući da su u početku svi izbori bili jednako vjerojatni tražena vjerojatnost trebala bi biti $\frac{3}{6} = \frac{1}{2}$. Pokažimo da je naša definicija u skladu s intuicijom.

Ako definiramo događaje

$$A = \{\text{izabran broj je paran}\} \quad \text{i} \quad B = \{\text{izabran broj djeljiv je s 3}\},$$

tražimo $\mathbb{P}(A \mid B)$. Budući da se nalazimo u Laplaceovom modelu, imamo

$$\begin{aligned}\mathbb{P}(A \mid B) &= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \\ &= \frac{k(A \cap B)/k(\Omega)}{k(B)/k(\Omega)} = \frac{k(A \cap B)}{k(B)}\,.\end{aligned}$$

Budući da je $k(B) = k(\{3, 6, 9, 12, 15, 18\}) = 6$, a $k(A \cap B) = k(\{6, 12, 18\}) = 3$, slijedi da je

$$\mathbb{P}(A \mid B) = \frac{3}{6} = \frac{1}{2}\,.$$

Uočimo da za svaka dva događaja $A, B \in \mathcal{F}$ vrijedi

$$\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B)\,\mathbb{P}(B) = \mathbb{P}(B \mid A)\,\mathbb{P}(A)\,. \tag{1}$$

Dakle, ako na primjer znamo $\mathbb{P}(B)$ i $\mathbb{P}(A \mid B)$ lako možemo odrediti $\mathbb{P}(A \cap B)$, tj. vjerojatnost da će se dogoditi i A i B.

Zadatak 4.3. U kutiji se nalazi 8 bijelih i 10 crnih kuglica. Dva puta uzastopno izvlačimo po jednu kuglicu bez vraćanja. Izračunajte vjerojatnost da su obje izvučene kuglice bijele boje.

Rješenje. Ako definiramo događaje

$$A = \{\text{prva izvučena kuglica je bijela}\}\,, \; B = \{\text{druga izvučena kuglica je bijela}\}\,,$$

tražimo $\mathbb{P}(A \cap B)$. Očito je $\mathbb{P}(A) = 8/18$, a budući da drugo izvlačenje ovisi o rezultatu prvog izvlačenja, $\mathbb{P}(B \mid A) = 7/17$. Zaista, ako prvo izvučemo bijelu kuglicu, u kutiji ostaje 7 bijelih i 10 crnih kuglica. Dakle,

$$\mathbb{P}(A \cap B) = \mathbb{P}(B \mid A)\,\mathbb{P}(A) = \frac{8}{18} \cdot \frac{7}{17}\,.$$

Alternativno, rezultat $\mathbb{P}(A \cap B) = \frac{8 \cdot 7}{18 \cdot 17}$ mogli smo dobiti i direktnim prebrojavanjem. $\square$

Ako je $0 < \mathbb{P}(B) < 1$, onda vrijedi

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}(A \cap \Omega) = \mathbb{P}(A \cap (B \cup B^c)) = \mathbb{P}((A \cap B)) \cup (A \cap B^c) = \\ &\overset{\text{disjun.}}{=} \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) = \mathbb{P}(A \mid B)\,\mathbb{P}(B) + \mathbb{P}(A \mid B^c)\,\mathbb{P}(B^c), \end{aligned}$$

pri čemu smo u 4. jednakosti iskoristili činjenicu da su $A \cap B$ i $A \cap B^c$ disjunktni. Prethodni račun se lako poopćuje.

Definicija 4.4. Niz događaja $H_1, H_2, \dots, H_n$ zovemo **potpun sistem događaja** ako je $H_1 \cup H_2 \cup \dots \cup H_n = \Omega$, $H_i \cap H_j = \emptyset$ za svaki $i \neq j$ (tj. ovi događaji su u parovima disjunktni), i $\mathbb{P}(H_i) > 0$ za svaki i.

Na primjer, ako je $0 < \mathbb{P}(B) < 1$, $H_1 = B, H_2 = B^c$ je jedan potpun sistem događaja. Nadalje, za potpun sistem događaja $(H_i)_{i=1}^n$ i za svaki $A \in \mathcal{F}$ vrijedi

$$\begin{aligned} \mathbb{P}(A) = \mathbb{P}(A \cap \Omega) = \mathbb{P}\left(A \cap \left(\bigcup_{i=1}^n H_i\right)\right) &= \mathbb{P}\left(\bigcup_{i=1}^n (A \cap H_i)\right) \\ = \sum_{i=1}^n \mathbb{P}(A \cap H_i) &= \sum_{i=1}^n \mathbb{P}(A \mid H_i)\,\mathbb{P}(H_i). \end{aligned}$$

Formulu

$$\mathbb{P}(A) = \sum_{i=1}^n \mathbb{P}(A \mid H_i)\,\mathbb{P}(H_i),$$

nazivamo **formulom potpune vjerojatnosti**. Napomenimo da ona vrijedi i u slučaju prebrojivog potpunog sistema događaja $(H_i)_{i \in \mathbb{N}}$.

Definicija 4.5. Kažemo da su $A, B \in \mathcal{F}$ su **nezavisni** ako je $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$.

Uočimo da ako su A i B nezavisni (te $\mathbb{P}(B) > 0$) vrijedi

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)\,\mathbb{P}(B)}{\mathbb{P}(B)} = \mathbb{P}(A),$$

tj. znanje o tome da se dogodio B ne govori nam ništa dodatno o vjerojatnosti događaja A. Obratno, ako vrijedi $\mathbb{P}(A \mid B) = \mathbb{P}(A)$ lako slijedi da su A i B nezavisni pa nam ova formula zapravo daje (intuitivniju) karakterizaciju nezavisnosti.

Napomena 4.6. Ako su A i B nezavisni, onda su nezavisni i A i B^c, A^c i B, i A^c i B^c.

Dokaz. Npr. za prvi slučaj $\mathbb{P}(A \cap B^c) = \mathbb{P}(A \backslash (A \cap B)) = \mathbb{P}(A) - \mathbb{P}(A \cap B) \overset{\text{nez.}}{=} \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A)(1 - \mathbb{P}(B)) = \mathbb{P}(A)\mathbb{P}(B^c)$. Ostali slučajevi se dokazuju slično. $\square$

Zadatak 4.7. Vjerojatnost da strijelac pogodi metu ako puše vjetar iznosi 0.4, a ako vjetar ne puše 0.7. Gađanja mete su nezavisna i u bilo kojem gađanju vjerojatnost pojave vjetra je 0.3. Nađite vjerojatnost

(a) da se u promatranom gađanju javi vjetar i strijelac pogodi metu,

(b) da u promatranom gađanju pogodi metu,

(c) da je puhao vjetar ako je u promatranome gađanju strijelac pogodio metu

(d) pogodi metu točno jednom u dva gađanja.

Rješenje.

(a) i (b) Neka je $A = \{\text{u promatranom gađanju pogođena meta}\}$. Vjerojatnost pogotka ovisi o vjetru pa stavimo

$$H_1 = \{\text{vjetar puše prilikom gađanja}\}\,,\ H_2 = \{\text{vjetar ne puše prilikom gađanja}\}\,.$$

Uočimo da je H_1, H_2 očito potpun sistem događaja ($H_2 = H_1^c$).

Iz zadatka znamo $\mathbb{P}(H_1) = 0.3$, $\mathbb{P}(H_2) = 0.7$, $\mathbb{P}(A \mid H_1) = 0.4$ i $\mathbb{P}(A \mid H_2) = 0.7$ pa je tražena vjerojatnost u (a) dijelu

$$\mathbb{P}(A \cap H_1) = \mathbb{P}(A \mid H_1)\mathbb{P}(H_1) = 0.4 \cdot 0.3 = 0.12\,.$$

Koristeći formulu potpune vjerojatnosti tražena vjerojatnost u (b) dijelu je

$$\mathbb{P}(A) = \mathbb{P}(A \mid H_1)\mathbb{P}(H_1) + \mathbb{P}(A \mid H_2)\mathbb{P}(H_2) = 0.4 \cdot 0.3 + 0.7 \cdot 0.7 = 0.61\,.$$

(c) Zanima nas $\mathbb{P}(H_1|A) = \frac{\mathbb{P}(A\cap H_1)}{\mathbb{P}(A)}$. U prvom dijelu smo izračunali oboje pa je $\mathbb{P}(H_1|A) = \frac{0.12}{0.61}$.

(d) Ako stavimo $A_j = \{\text{u j-tom gađanju meta je pogođena}\}$, za $j = 1, 2$, tada su A_1 i A_2 nezavisni i $\mathbb{P}(A_j) = \mathbb{P}(A) = 0.61$ i $\mathbb{P}(A_j^c) = 1 - \mathbb{P}(A) = 0.39$, $j = 1, 2$. Tražimo

$$\begin{aligned}\mathbb{P}((A_1 \cap A_2^c) \cup (A_1^c \cap A_2)) &\overset{\text{disjun.}}{=} \mathbb{P}(A_1 \cap A_2^c) + \mathbb{P}(A_1^c \cap A_2) \overset{\text{nez.}}{=} \mathbb{P}(A_1)\mathbb{P}(A_2^c) + \mathbb{P}(A_1^c)\mathbb{P}(A_2)\\ &= 0.61 \cdot 0.39 + 0.39 \cdot 0.61 = 0.4758.\end{aligned}$$

□

Za potpun sistem događaja $H_1, \ldots, H_n$ vrijedi **Bayesova formula**:

$$\mathbb{P}(H_j|A) = \frac{\mathbb{P}(A \mid H_j)\mathbb{P}(H_j)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A \mid H_j)\,\mathbb{P}(H_j)}{\sum_{i=1}^{n} \mathbb{P}(A \mid H_i)\,\mathbb{P}(H_i)}.$$

Napomena 4.8. Za fiksni $B \in \mathcal{F}$, funkcija $A \mapsto \mathbb{P}(A \mid B)$ je ponovno vjerojatnost. Na primjer, vrijedi $\mathbb{P}(A^c \mid B) = 1 - \mathbb{P}(A \mid B)$.

Zadatak 4.9. U nekom gradu na aerodromu vozi 30% taksija plave boje, 20% taksija zelene boje i 50% taksija žute boje. Oni putnike dovedu prekasno na let s vjerojatnosti 0.1, 0.2 i 0.3 (redom). Jednog dana žureći na aerodrom neki je putnik zaustavio taksi na ulici i rekao vozaču da vozi na aerodrom. Na kraju taj putnik nije zakasnio na let. Koja je vjerojatnost da se vozio u žutom taksiju?

Rješenje. Budući da dolazak na let na vrijeme ovisi o boji taksija, definiramo potpun sistem događaja

$$\begin{aligned} H_1 &= \{\text{zaustavljen plavi taksi}\}\,, \\ H_2 &= \{\text{zaustavljen zeleni taksi}\}\,, \\ H_3 &= \{\text{zaustavljen žuti taksi}\}\,. \end{aligned}$$

Iz zadatka znamo da je $\mathbb{P}(H_1) = 0.3$, $\mathbb{P}(H_2) = 0.2$ i $\mathbb{P}(H_3) = 0.5$.

Ako definiramo događaj $A = \{\text{putnik nije zakasnio na let}\}$, tražimo $\mathbb{P}(H_3 \mid A)$, a zadano nam je

$$\mathbb{P}(A \mid H_1) = 1 - \mathbb{P}(A^c \mid H_1) = 1 - 0.1 = 0.9\,,$$

te slično $\mathbb{P}(A \mid H_2) = 0.8$ i $\mathbb{P}(A \mid H_3) = 0.7$. Dakle, možemo iskoristiti Bayesovu formulu. Imamo

$$\mathbb{P}(A) = \sum_{i=1}^{3} \mathbb{P}(A \mid H_i)\mathbb{P}(H_i) = 0.9 \cdot 0.3 + 0.8 \cdot 0.2 + 0.7 \cdot 0.5 = 0.78$$

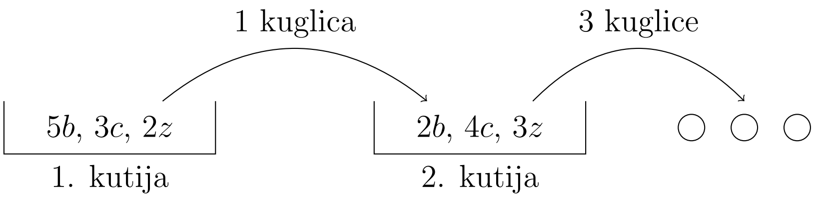
pa je

$$\mathbb{P}(H_3 \mid A) = \frac{\mathbb{P}(A \mid H_3)\mathbb{P}(H_3)}{\mathbb{P}(A)} = \frac{35}{78}\,.$$

Uočimo da je $\mathbb{P}(H_3 \mid A) = \frac{35}{78} < 0.5 = \mathbb{P}(H_3)$, tj. ako znamo da je putnik stigao na let vjerojatnost da ga je dovezao žuti taksi se smanjuje, što je intuitivno jasno jer vožnja žutim taksijem daje najmanju vjerojatnost da putnik stigne na vrijeme. Formalno, to možemo vidjeti i iz Bayesove formule i činjenice da je $\mathbb{P}(A \mid H_3) < \mathbb{P}(A)$. Suprotan zaključak vrijedi za događaje H_1 i H_2 , tj. plavi i zeleni taksi. $\square$

Zadatak 4.10. U prvoj kutiji je 5 bijelih, 3 crvene i 2 zelene kuglice, a u drugoj kutiji su 2 bijele, 4 crvene i 3 zelene kuglice. Iz prve kutije izvučemo jednu kuglcu i prebacimo ju u drugu kutiju, a zatim iz druge kutije izvučemo 3 kuglice. Ako su sve izvučene kuglice bile različitih boja, kolika je vjerojatnost da smo iz prve kutije u drugu prebacili crvenu kuglicu?

Rješenje.



Neka su

$$\begin{aligned} H_1 &= \{\text{prebacili smo bijelu kuglicu iz prve kutije u drugu}\} \quad \Rightarrow \quad \mathbb{P}(H_1) = \frac{5}{5+3+2} = \frac{1}{2}, \\ H_2 &= \{\text{prebacili smo crvenu kuglicu iz prve kutije u drugu}\} \quad \Rightarrow \quad \mathbb{P}(H_2) = \frac{3}{10}, \\ H_3 &= \{\text{prebacili smo zelenu kuglicu iz prve kutije u drugu}\} \quad \Rightarrow \quad \mathbb{P}(H_3) = \frac{1}{5}. \end{aligned}$$

Ako je $A = \{$iz druge kutije izvukli smo tri kuglice različite boje$\}$, slijedi da je

$$\mathbb{P}(A \mid H_1) = [\text{u 2. kutiji je } 3b,\ 4c,\ 3z] = \frac{\binom{3}{1}\binom{4}{1}\binom{3}{1}}{\binom{10}{3}},$$
$$\mathbb{P}(A \mid H_2) = [\text{u 2. kutiji je } 2b,\ 5c,\ 3z] = \frac{\binom{2}{1}\binom{5}{1}\binom{3}{1}}{\binom{10}{3}},$$
$$\mathbb{P}(A \mid H_3) = [\text{u 2. kutiji je } 2b,\ 4c,\ 4z] = \frac{\binom{2}{1}\binom{4}{1}\binom{4}{1}}{\binom{10}{3}}.$$

Tražimo $\mathbb{P}(H_2 \mid A)$ pa koristeći Bayesovu formulu imamo

$$\mathbb{P}(H_2 \mid A) = \frac{\mathbb{P}(A \mid H_2)\mathbb{P}(H_2)}{\sum_{i=1}^{3} \mathbb{P}(A \mid H_i)\mathbb{P}(H_i)} = \cdots = \frac{45}{167} \approx 0.2695.$$

□

Zadatak 4.11. Bacamo 5 simetričnih novčića. Nakon prvog bacanja ponovo bacamo novčiće na kojima je u prvom bacanju pala glava. Kolika je vjerojatnost da će nakon drugog bacanja ukupno pasti barem 3 pisma (dakle, brojeći i pisma koja su pala u prvom bacanju)?

Rješenje. Očito, broj pisama dobivenih u drugom bacanju ovisi o broju pisama u prvom bacanju. Neka su

$$H_i = \{\text{u 1. bacanju palo je } i \text{ pisama}\},\ i = 0, 1, \ldots, 5\,.$$

Tada je $(H_i)_i$ potpun sistem događaja i vrijedi

$$\mathbb{P}(H_i) = \frac{\binom{5}{i}}{2^5},\ i = 0, 1, \ldots, 5\,.$$

Zaista, ukupno ima 2^5 ishoda, a na $\binom{5}{i}$ načina biramo i od 5 novčića na kojima je palo pismo pa je to upravo broj ishoda s točno i pisama (i $5 - i$ glava).

Definirajmo

$$A = \{\text{palo je ukupno barem 3 pisma}\} = \{\text{palo je 3, 4 ili 5 pisama}\}$$

Mi ćemo prvo odrediti vjerojatnost događaja $A^c = \{$palo je 0,1 ili 2 pisma$\}$.

Računamo

$$\begin{aligned}
\mathbb{P}(A^c \mid H_0) &= [\text{palo je ukupno 0,1 ili 2 pisma ako je isprva palo 0}] \\
&= [\text{na pet novčića bacanih u 2. bacanju palo 0,1 ili 2 pisma}] \\
&= \mathbb{P}(\{\text{na 5 novčića palo je 0,1 ili 2 pisma}\}) \\
&= \frac{\binom{5}{0}}{2^5} + \frac{\binom{5}{1}}{2^5} + \frac{\binom{5}{2}}{2^5} = \frac{1}{2}.
\end{aligned}$$

Nadalje,

$$\begin{aligned}\mathbb{P}\left(A^c \mid H_1\right) &= [\text{palo je 0,1 ili 2 pisma ako je palo jedno pismo u prvom bacanju}] \\ &= [\text{na četiri novčića bacanih u 2. bacanju palo 0 ili 1 pismo}] \\ &= \mathbb{P}\left(\{\text{palo je 0 ili 1 pismo na 4 novčića}\}\right) \\ &= \frac{\binom{4}{0}}{2^4} + \frac{\binom{4}{1}}{2^4} = \frac{5}{16}.\end{aligned}$$

Slično, $\mathbb{P}\left(A^c \mid H_2\right) = \mathbb{P}\left(\{\text{palo je 0 pisma na 3 novčića}\}\right) = 1/8$. Nadalje, očito je $0 = \mathbb{P}\left(A^c \mid H_3\right) = \mathbb{P}\left(A^c \mid H_4\right) = \mathbb{P}\left(A^c \mid H_5\right)$ jer ne možemo ukupno imati manje od 3 pisma ako smo već u prvom bacanju dobili najmanje 3 pisma. Dakle,

$$\mathbb{P}\left(A^c\right) = \sum_{i=0}^{5} \mathbb{P}\left(A^c \mid H_i\right)\mathbb{P}\left(H_i\right) = \frac{1}{2} \cdot \frac{\binom{5}{0}}{2^5} + \frac{5}{16} \cdot \frac{\binom{5}{1}}{2^5} + \frac{1}{8} \cdot \frac{\binom{5}{2}}{2^5} = \frac{53}{512},$$

pa je

$$\mathbb{P}\left(A\right) = 1 - \mathbb{P}\left(A^c\right) = 1 - \frac{53}{512} = \frac{459}{512} \approx 0.896\,.$$

□

Zadatak 4.12. U kutiji je 5 kuglica od kojih svaka može biti bijela ili crna s jednakom vjerojatnosti. Ako smo izvukli bijelu kuglicu, koji je najvjerojatniji broj crnih kuglica u kutiji na početku pokusa.

Rješenje. Neka su

$$H_i = \{\text{u kutiji se nalazi } i \text{ crnih kuglica}\}\,,\ i = 0, 1, \ldots, 5\,,$$

te kao i u prethodnom zadatku (umjesto pismo-glava imamo bijela-crna boja) je $\mathbb{P}\left(H_i\right) = \frac{\binom{5}{i}}{2^5}$ za sve $i = 0, 1, \ldots, 5$. Ako stavimo $A = \{\text{izvukli smo bijelu kuglicu}\}$ tražimo i za koji je $\mathbb{P}\left(H_i \mid A\right)$ najveća. Uočimo da je

$$\mathbb{P}\left(A \mid H_i\right) = \frac{5-i}{5}\,,\ i = 0, 1, \ldots, 5\,,$$

jer ako vrijedi H_i u kutiji ima $5-i$ bijelih od ukupno 5 kuglica pa koristeći Bayesovu formulu imamo

$$\mathbb{P}\left(H_i \mid A\right) = \frac{\mathbb{P}\left(A \mid H_i\right)\mathbb{P}\left(H_i\right)}{\mathbb{P}\left(A\right)} = \frac{1}{\mathbb{P}\left(A\right)} \cdot \frac{5-i}{5} \cdot \frac{\binom{5}{i}}{2^5} = \underbrace{\frac{1}{\mathbb{P}\left(A\right) \cdot 5 \cdot 2^5}}_{\text{ne ovisi o } i} \cdot (5-i) \cdot \binom{5}{i}.$$

Dakle, tražimo za koji i se postiže $\max\limits_{i=0,\ldots,5} (5-i) \cdot \binom{5}{i}$.

i	0	1	2	3	4	5
$(5-i) \cdot \binom{5}{i}$	5	20	30	20	5	0

Dakle, maksimum od $\mathbb{P}(H_i \,|\, A)$ se postiže za $i = 2$, tj. najvjerojatnije je da smo imali 2 crne kuglice na početku.

Uočimo da je $\mathbb{P}(A \,|\, H_0) = 1$, tj. ako su u kutiji sve bijele kuglice na početku, sigurno ćemo izvući bijelu kuglicu. Ipak, to nije najvjerojatniji raspored kuglica na početku ako znamo da smo izvukli bijelu kuglicu jer sama početna vjerojatnost $\mathbb{P}(H_0)$ nije velika u odnosu na ostale. $\square$

Napomena 4.13. Za $A_1, \ldots, A_n \in \mathcal{F}$ vrijedi poopćenje formule (1):

$$\mathbb{P}(\cap_{i=1}^n A_i) = \mathbb{P}(A_1)\mathbb{P}(A_2 \,|\, A_1)\mathbb{P}(A_3 \,|\, A_2 \cap A_1) \cdots \mathbb{P}(A_n \,|\, A_{n-1} \cap \cdots \cap A_1).$$

Zadatak 4.14 (*)**.** Na predavanju sa 100 ljudi profesor je zamolio studente da jedan po jedan kažu na glas u koji su dan u godini rođeni, a prvi student čiji se rođendan bude poklapao s nekim od prethodno izgovorenih (ako takav uopće postoji), dobiva simboličnu nagradu. Ako pretpostavimo da svaki student s jednakom vjerojatnosti i nezavisno od ostalih može imati rođendan na bilo koji od 365 dana, pokažite da najveću šansu da dobije nagradu ima dvadeseta osoba po redu.

Rješenje. Označimo s p_n vjerojatnost da je n-ta osoba osvojila nagradu za $n = 1, \ldots, 100$. Ako definiramo događaje

$$A_n = \{\text{neka od prvih } (n-1) \text{ osoba ima isti rođendan kao i } n\text{–ta osoba}\}, \ n = 1, \ldots, 100,$$

tada je

$$\begin{aligned}
p_n &= \mathbb{P}(A_1^c \cap A_2^c \cap \cdots \cap A_{n-1}^c \cap A_n) \\
&= \mathbb{P}(A_1^c)\mathbb{P}(A_2^c \,|\, A_1^c)\mathbb{P}(A_3^c \,|\, A_2^c \cap A_1^c) \cdots \mathbb{P}(A_{n-1}^c \,|\, A_{n-2}^c \cap \cdots \cap A_1^c)\mathbb{P}(A_n \,|\, A_{n-1}^c \cap \cdots \cap A_1^c) \\
&= 1 \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{365-(n-2)}{365} \cdot \frac{n-1}{365} \\
&= \frac{365 \cdot 364 \cdots (365-n+2)}{365^n}(n-1).
\end{aligned}$$

Zaista, ako znamo da se na primjer dogodio $A_2^c \cap A_1^c$ to znači da prve dvije osobe imaju rođendan na dva različita dana pa trećoj "preostaje" $365 - 2 = 363$ različitih dana na koje može imati rođendan, a da nije jednak nekom od prethodna dva. Alternativno, gornju vjerojatnost smo mogli izračunati i direktnim prebrojavanje kao u jednom od zadataka iz prethodnog poglavlja ("problem rođendana").

Pokažimo da funkcija $n \mapsto p_n$ postiže maksimum za $n = 20$. Uočimo,

$$\frac{p_{n+1}}{p_n} = \frac{365-n+1}{365} \cdot \frac{n}{n-1},$$

pa imamo da je $\frac{p_{n+1}}{p_n} \gtreqless 1$ ako i samo ako je

$$n^2 - n - 365 \lesseqgtr 0.$$

Budući da polinom $x^2 - x - 365$ ima nultočke $\frac{1 \pm \sqrt{1461}}{2} \approx -18.61, 19.61$, imamo da je $\frac{p_{n+1}}{p_n} > 1$ za $n \leq 19$, a $\frac{p_{n+1}}{p_n} < 1$ za $n \geq 20$. Dakle, vjerojatnost p_n je najveća za $n = 20$. $\square$

Zadaci za vježbu

Zadatak 4.15. Tri prijateljice, Dunja, Lidija i Tina, prijavile su se kao ekipa na kviz znanja. One se za odgovor na pitanje prijavljuju s vjerojatnostima $\frac{1}{2}$, $\frac{1}{3}$ i $\frac{1}{6}$. Na pitanje odgovaraju točno s vjerojatnostima $\frac{4}{5}$, $\frac{3}{5}$ i $\frac{3}{5}$.

(a) Kolika je vjerojatnost da na pitanje neće točno odgovoriti?

(b) Ako pitanje nije točno odgovoreno, nađite vjerojatnost da je na pitanje odgovarala Tina.

Zadatak 4.16. 3 nesimetrična novčića C_1, C_2, C_3 leže na stolu. Vjerojatnosti da na njima padnu glave redom su $\frac{1}{3}$, $\frac{2}{3}$ i 1. Na slučajan način uzmemo jedan novčić, bacimo ga i uočimo da je pala glava. Izračunajte vjerojatnost da je to novčić C_i, za $i = 1, 2, 3$.

Rješenje. $A = \{\text{pala je glava na novčić}\}$. Neka su $H_i = \{\text{izabrali smo novčić } C_i\}$, za $i = 1, 2, 3$. Tada iz teksta zadatka znamo da

$$\begin{aligned} \mathbb{P}(A \mid H_1) &= \frac{1}{3}, \\ \mathbb{P}(A \mid H_2) &= \frac{2}{3}, \\ \mathbb{P}(A \mid H_3) &= 1. \end{aligned}$$

U zadatku se traže vjerojatnosti $\mathbb{P}(H_i \mid A)$, za $i = 1, 2, 3$. Koristimo Bayesovu formulu, tj.

$$\mathbb{P}(H_i \mid A) = \frac{\mathbb{P}(A \mid H_i)\mathbb{P}(H_i)}{\mathbb{P}(A)}. \tag{$\star$}$$

Za dovršetak zadatka treba još izračunati

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}(A \mid H_1)\mathbb{P}(H_1) + \mathbb{P}(A \mid H_2)\mathbb{P}(H_2) + \mathbb{P}(A \mid H_3)\mathbb{P}(H_3) \\ &= \frac{1}{3} \cdot \frac{1}{3} + \frac{2}{3} \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = \frac{2}{3}. \end{aligned}$$

Dakle, po ($\star$) $\mathbb{P}(H_1 \mid A) = \frac{1}{6}$, $\mathbb{P}(H_2 \mid A) = \frac{1}{3}$ i $\mathbb{P}(H_3 \mid A) = \frac{1}{2}$. □

Zadatak 4.17. U ponoć su na parkiralištu bila 2 siva i 1 crni Ford, 3 siva i 4 crna BMW-a i 3 sive i 1 crna Toyota. Te noći je kradljivac automobila nasumce odabrao automobil i ukrao ga. Ako je ukraden automobil sive boje, koja je vjerojatnost da je to bio BMW?

Zadatak 4.18. Iz kutije u kojoj je 10 bijelih i 8 crnih kuglica je izgubljena jedna kuglica nepoznate boje. Ako su iz kutije izvučene dvije bijele kuglice, izračunajte vjerojatnost da je izgubljena kuglica bila bijela.

Zadatak 4.19. U restoranu rade Ivica, Marica i zla vještica. Pri ulasku bacate dvije kocke. Ako je zbroj na kockama neparan, poslužit će Vas Ivica i pritom će Vas s vjerojatnosti 0.2 zabunom otrovati. Ako je zbroj paran, ali različit od 12, onda će Vas posluživati Marica i pritom Vas otrovati s vjerojatnosti 0.1. Ako ste dobili 12, poslužit će Vas zla vještica i pritom će Vas sigurno otrovati. Koja je vjerojatnost da ćete preživjeti? Ako ste preživjeli, koja je vjerojatnost da Vas je poslužila Marica?

Rješenje. Neka je $A = \{\text{preživjeli ste}\}$ i neka su

$$
\begin{aligned}
H_1 &= \{\text{poslužio Vas je Ivica}\} && \Rightarrow \quad \mathbb{P}(H_1) = 1/2, \\
H_2 &= \{\text{poslužila Vas je Marica}\} && \Rightarrow \quad \mathbb{P}(H_2) = 17/36, \\
H_3 &= \{\text{poslužila Vas je zla vještica}\} && \Rightarrow \quad \mathbb{P}(H_3) = 1/36.
\end{aligned}
$$

Iz uvjeta zadatka imamo

$$
\begin{aligned}
\mathbb{P}(A \mid H_1) &= 1 - \mathbb{P}(A^c \mid H_1) = 1 - 0.2 = 0.8, \\
\mathbb{P}(A \mid H_2) &= 1 - \mathbb{P}(A^c \mid H_2) = 1 - 0.1 = 0.9, \\
\mathbb{P}(A \mid H_3) &= 1 - \mathbb{P}(A^c \mid H_3) = 1 - 1 = 0.
\end{aligned}
$$

$\mathbb{P}(A) = \sum_{i=1}^{3} \mathbb{P}(A \mid H_i)\mathbb{P}(H_i) \approx 0.825.$
$\mathbb{P}(H_2 \mid A) = \frac{\mathbb{P}(A \mid H_2)\mathbb{P}(H_2)}{\mathbb{P}(A)} \approx 0.515.$ □

Zadatak 4.20. U prvoj su kutiji 2 bijele i 2 crne kuglice, a u drugoj kutiji je 5 bijelih i 7 crnih kuglica. Iz prve kutije izvučemo dvije kuglice i prebacimo ih u drugu kutiju, a zatim iz druge kutije izvučemo dvije kuglice i prebacimo ih u prvu kutiju. Ako su u prvoj kutiji sve kuglice iste boje, kolika je vjerojatnost da su one crne?

Zadatak 4.21 (Pólya's urn)**.** U kutiji imamo b bijelih i c crvenih kuglica. Na slučajan način izvlačimo kuglicu iz kutije, zabilježimo njenu boju i potom je vratimo nazad u kutiju s još $d \geq 1$ kuglica iste boje. Tako ponavljamo postupak. Koja je vjerojatnost

(a) da je druga izvučena kuglica crvena,

(b) da je prva izvučena kuglica crvena ako je u druga izvučena kuglica isto crvena,

(c)* da je n-ta izvučena kuglica crvena? (*Uputa:* Ako ste izračunali vjerojatnost za $n = 1, 2$, probajte pogoditi rješenje za općeniti n i dokazati ga pomoću indukcije, uvjetujući na to je li u prvom koraku izvučena crvena ili bijela kuglica.)

Rješenje. Neka je $C_n = \{n\text{-ta izvučena kuglica je crvena}\}$ za $n = 1, 2, \dots$.

(a) $H_1 = C_1$, $H_2 = C_1^c$. Tražimo pomoću formule potpune vjerojatnosti $\mathbb{P}(C_2) = \mathbb{P}(C_2 \mid C_1)\,\mathbb{P}(C_1) + \mathbb{P}(C_2 \mid C_1^c)\,\mathbb{P}(C_1^c) = \frac{c}{b+c}$.

(b) Pomoću Bayesove formule $\mathbb{P}(C_1 \mid C_2) = \frac{\mathbb{P}(C_2 \mid C_1)\mathbb{P}(C_1)}{\mathbb{P}(C_2)} = \mathbb{P}(C_2 \mid C_1) = \frac{c+d}{c+b+d}$.

(c) Tvrdimo da je $\mathbb{P}(C_n) = \frac{c}{b+c}$ za sve $n \geq 1$ i proizvoljne početne brojeve kuglica b i c. Za $n = 1$ je to jasno, a za $n = 2$ smo to pokazali u (a) dijelu. Pretpostavimo da je tvrdnja istinita u slučaju $(n-1)$-vog izvlačenja za proizvoljne b i c. Imamo

$$\mathbb{P}(C_n) = \mathbb{P}(C_n \mid C_1)\,\mathbb{P}(C_1) + \mathbb{P}(C_n \mid C_1^c)\mathbb{P}(C_1^c)\,.$$

Uočimo da je vjerojatnost da ćemo izvući crvenu kuglicu u n-tom izvlačenju ako smo u prvom izvukli crvenu kuglicu ista kao i vjerojatnost da ćemo u $(n-1)$-vom izvlačenju izvući crvenu kuglicu ako smo krenuli s $c+d$ crvenih i b bijelih kuglica na početku, i slično u slučaju da smo u prvom izvlačenju izvukli bijelu kuglicu. Dakle, koristeći pretpostavku indukcije dobivamo

$$\mathbb{P}\left(C_n\right) = \frac{c+d}{b+c+d} \cdot \frac{c}{b+c} + \frac{c}{b+c+d} \cdot \frac{b}{b+c} = \frac{c}{b+c}\,.$$

$\square$

5 Slučajne varijable

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

Definicija 5.1. Funkciju $X : \Omega \to \mathbb{R}$ koja poprima najviše prebrojivo mnogo različitih vrijednosti zovemo **diskretna slučajna varijabla**[2].

Za $a \in \mathbb{R}$ zanimat će nas vjerojatnost da slučajna varijabla X poprimi vrijednost a, tj. vjerojatnost događaja $\{\omega \in \Omega : X(\omega) = a\}$ kojeg ćemo jednostavnije označavati s $\{X = a\}$. Malo općenitije, zanimat će nas vjerojatnost događaja $\{X \in B\} := \{\omega \in \Omega : X(\omega) \in B\}$ za $B \subseteq \mathbb{R}$.

Primjer 5.2. Bacamo dva simetrična novčića te neka je X broj palih pisama. Ovaj broj je slučajan jer ovisi o ishodu bacanja novčića i primjer je jedne diskretne slučajne varijable. Preciznije, možemo staviti $\Omega = \{GG, GP, PG, PP\}$ pri čemu je $\mathbb{P}(\{\omega\}) = \frac{1}{4}$ za sve $\omega \in \Omega$ i definirati $X : \Omega \to \mathbb{R}$ na sljedeći način:

$$X(GG) = 0,\ X(GP) = X(PG) = 1,\ X(PP) = 2\,.$$

Uočimo, X može poprimiti vrijednosti $0, 1$ ili 2 i to s vjerojatnostima

$$\mathbb{P}(X = 0) = \mathbb{P}(\{GG\}) = \frac{1}{4},\ \mathbb{P}(X = 1) = \mathbb{P}(\{GP, PG\}) = \frac{1}{2} \text{ i } \mathbb{P}(X = 2) = \mathbb{P}(\{PP\}) = \frac{1}{4}\,.$$

To kraće zapisujemo u **tablicu distribucije**

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

Općenito, ako diskretna slučajna varijabla X može poprimiti vrijednosti $a_1, a_2, \ldots$ i to s vjerojatnostima $p_i = \mathbb{P}(X = a_i)$, $i = 1, 2, \ldots$, brojeve $a_1, a_2, \ldots$ i pripadne vjerojatnosti $p_1, p_2, \ldots$ zovemo **distribucijom** (ili **razdiobom**) slučajne varijable X i pišemo

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

Uočimo, nužno je $p_i \geq 0$ za sve $i = 1, 2, \ldots$ i $\sum_{i=1}^{\infty} p_i = 1$.

Obratno, ako su zadani brojevi a_i i brojevi $p_i \geq 0$ takvi da je $\sum_{i=1}^{\infty} p_i = 1$, može se pokazati da uvijek postoji vjerojatnosni prostor i slučajna varijabla X čija je to distribucija. Često nam sam vjerojatnosni prostor neće biti važan, već samo distribucija slučajne varijable X.

Zadatak 5.3. Slučajna varijabla X ima sljedeću distribuciju

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ c & 2c & 2c & 3c & c^2 & 2c^2 & 7c^2 + c \end{pmatrix}$$

za neki $c \in \mathbb{R}$.

[2]Formalno, da bi funkciju X zvali slučajnom varijablom, potrebno je da bude "izmjeriva", tj. da skupovi $\{X = a\}$ uvedeni ispod definicije budu pravi događaji, tj. elementi od $\mathcal{F}$. U nastavku ćemo ovo zanemariti.

(a) Odredite konstantu c.

(b) Izračunajte vjerojatnost da X poprimi vrijednost između 2 i 5 (uključivo).

(c) Odredite najmanji $k \in \mathbb{N}$ takav da vrijedi $\mathbb{P}(X \leq k) \geq \frac{2}{5}$.

Rješenje.

(a) Mora vrijediti $c + 2c + 2c + 3c + c^2 + 2c^2 + 7c^2 + c = 1$, tj. $10c^2 + 9c - 1 = 0$. Budući da je očito $c > 0$ lako izračunamo da je $c = \frac{1}{10}$.

(b) Tražimo $\mathbb{P}(2 \leq X \leq 5)$, a budući da X poprima vrijednosti u skupu $\{1, 2, \ldots, 7\}$ imamo

$$\begin{aligned}\mathbb{P}(2 \leq X \leq 5) &= \mathbb{P}(\overbrace{X \in \{2,3,4,5\}}^{=\bigcup_{i=2}^{5}\{X=i\}}) = [\text{disjunktnost}] \\ &= \mathbb{P}(X=2) + \mathbb{P}(X=3) + \mathbb{P}(X=4) + \mathbb{P}(X=5) \\ &= \frac{2}{10} + \frac{2}{10} + \frac{3}{10} + \frac{1}{100} = \frac{71}{100},\end{aligned}$$

pri čemu gornje vjerojatnosti iščitavamo iz tablice distribucije od X uz $c = \frac{1}{10}$.

(c) Računamo

$$\begin{aligned}\mathbb{P}(X \leq 1) &= \mathbb{P}(X = 1) = \frac{1}{10} < \frac{2}{5}, \\ \mathbb{P}(X \leq 2) &= \mathbb{P}(X = 1) + \mathbb{P}(X = 2) = \frac{3}{10} < \frac{2}{5}, \\ \mathbb{P}(X \leq 3) &= \mathbb{P}(X \leq 2) + \mathbb{P}(X = 3) = \frac{1}{2} \geq \frac{2}{5},\end{aligned}$$

pri čemu smo u zadnjem retku iskoristili činjenicu da su događaji $\{X \leq 2\}$ i $\{X = 3\}$ disjunktni. Dakle, najmanji k je $k = 3$.

□

Napomena 5.4. Ako je X diskretna slučajna varijabla i $g : \mathbb{R} \to \mathbb{R}$ funkcija, tada s $g(X)$ označavamo funkciju $g(X) : \Omega \to \mathbb{R}$ definiranu s $g(X)(\omega) := g(X(\omega))$ za sve $\omega \in \Omega$, tj. $g(X) := g \circ X$. Budući da $g(X)$ očito može poprimiti najviše prebrojivo mnogo vrijednosti, ona je također diskretna slučajna varijabla.

Zadatak 5.5. Bacamo simetričnu kocku i neka X označava broj koji je pao na kocki. Odredite distribucije slučajnih varijabli X, X^2 i $|X - 3|$.

Rješenje. Očito,

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

Nadalje, imamo sljedeću tablicu s vrijednostima X-a i odgovarajućim vrijednostima od X^2 i $|X - 3|$

X	X^2	$\|X-3\|$
1	1	2
2	4	1
3	9	0
4	16	1
5	25	2
6	36	3

Malo preciznije, za sve $\omega \in \Omega$ za koje je $X(\omega) = 1$ imamo $X^2(\omega) = X(\omega)^2 = 1$ i $|X-3|(\omega) = |X(\omega) - 3| = 2$ itd. Sada lako slijedi

$$X^2 \sim \begin{pmatrix} 1 & 4 & 9 & 16 & 25 & 36 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}.$$

Na primjer,

$$\begin{aligned} \mathbb{P}\left(X^2 = 1\right) &= \left[\mathbb{P}\left(\{\omega \in \Omega : X(\omega)^2 = 1\}\right) = \mathbb{P}\left(\{\omega \in \Omega : X(\omega) \in \{-1, 1\}\}\right)\right] \\ &= \mathbb{P}\left(X \in \{-1, 1\}\right) = \mathbb{P}\left(X = 1\right) = \frac{1}{6}, \end{aligned}$$

pri čemu smo u predzadnjoj jednakosti iskoristili da X ne može poprimiti vrijednost -1. Slično,

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

Na primjer,

$$\mathbb{P}\left(|X-3| = 1\right) = \mathbb{P}\left(X = 2 \text{ ili } 4\right) = \mathbb{P}\left(X = 2\right) + \mathbb{P}\left(X = 4\right) = \frac{2}{6} = \frac{1}{3}$$

□

Općenito, ako je $X \sim \begin{pmatrix} a_1 & a_2 & \cdots\cdots \\ p_1 & p_2 & \cdots\cdots \end{pmatrix}$, $g : \mathbb{R} \to \mathbb{R}$ i $Y = g(X)$, tada Y poprima vrijednosti u skupu $\{b_1, b_2, \ldots\} = g(\{a_1, a_2, \ldots\})$ s vjerojatnostima

$$\mathbb{P}\left(Y = b_j\right) = \sum_{i : g(a_i) = b_j} p_i, \; j = 1, 2, \ldots$$

Zadatak 5.6. Ako slučajna varijabla X poprima vrijednosti u skupu prirodnih brojeva i vrijedi $\mathbb{P}\left(X = n\right) = \frac{1}{2^n}$, $n \in \mathbb{N}$, odredite distribuciju slučajne varijable $Y = \sin\left(\frac{\pi X}{2}\right)$.

Rješenje. Uočimo da je

$$\sin\left(\frac{n\pi}{2}\right) = \begin{cases} 0 & \text{za } n \equiv 0 \pmod 2 \\ 1 & \text{za } n \equiv 1 \pmod 4 \\ -1 & \text{za } n \equiv 3 \pmod 4 \end{cases}, \; n \in \mathbb{N}.$$

Dakle, budući da X poprima vrijednosti u $\mathbb{N}$, Y može poprimiti vrijednosti $-1, 0$ i 1 i to s vjerojatnostima

$$\mathbb{P}(Y=0) = \sum_{k=1}^{\infty} \mathbb{P}(X=2k) = \sum_{k=1}^{\infty} \frac{1}{2^{2k}} = \frac{1}{4}\sum_{k=0}^{\infty} \frac{1}{4^k} = \frac{1}{4}\cdot\frac{1}{1-\frac{1}{4}} = \frac{1}{3},$$
$$\mathbb{P}(Y=1) = \sum_{k=0}^{\infty} \mathbb{P}(X=4k+1) = \sum_{k=0}^{\infty} \frac{1}{2^{4k+1}} = \frac{1}{2}\sum_{k=0}^{\infty} \frac{1}{16^k} = \frac{1}{2}\cdot\frac{1}{1-\frac{1}{16}} = \frac{8}{15},$$

a preostalu vjerojatnost najlakše možemo dobiti kao $\mathbb{P}(Y=-1) = 1 - \mathbb{P}(Y=0) - \mathbb{P}(Y=1) = \frac{2}{15}$. Dakle,

$$Y \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{2}{15} & \frac{1}{3} & \frac{8}{15} \end{pmatrix}.$$

$\square$

Primjer 5.7. Slučajnu varijablu X s distribucijom

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

za $p \in [0,1]$ zovemo Bernoullijeva slučajna varijabla s parametrom p. Slučajna varijabla X predstavlja rezultat slučajnog pokusa s dva ishoda pri čemu je vjerojatnost uspjeha $p = \mathbb{P}(X = 1)$.

Primjer 5.8. Ako je X broj uspjeha u nizu n nezavisnih pokusa pri čemu je vjerojatnost uspjeha u svakom pokusu jednaka $p \in [0,1]$, tada je, uz $q = 1-p$,

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

Slučajnu varijablu X s ovakvom distribucijom zovemo binomna slučajna varijabla s parametrima n i p te označavamo $X \sim B(n,p)$.

Zadatak 5.9. Marko svaki dan (nezavisno od ostalih dana) kasni na nastavu s vjerojatnošću 0.2.

(a) Kolika je vjerojatnost da je u jednom tjednu (tj. u 5 radnih dana) zakasnio najviše jednom?

(b) Za koji najveći broj dana u tjednu možemo biti barem 90% sigurni da je barem toliko puta došao na vrijeme?

Rješenje. Neka je X broj dana u kojima je Marko zakasnio na nastavu. Budući da imamo 5 dana (5 "pokusa") te u svakom danu vjerojatnost kašnjenja (vjerojatnost "uspjeha") 0.2 i to nezavisno od ostalih dana, očito je $X \sim B(5, 0.2)$.

(a) Tražimo

$$\begin{aligned}\mathbb{P}(X \le 1) = \mathbb{P}(X=0) + \mathbb{P}(X=1) &= \binom{5}{0} 0.2^0 0.8^5 + \binom{5}{1} 0.2^1 0.8^4 \\ &= 0.32768 + 0.4096 = 0.73728.\end{aligned}$$

(b) Broj dana u koje je Marko došao na vrijeme je naravno $5 - X$, tako da tražimo najveći k takav da je $\mathbb{P}(5 - X \geq k) = \mathbb{P}(X \leq 5 - k) \geq 0.9$.

Već smo prethodno vidjeli da za $k = 4$ imamo $\mathbb{P}(X \leq 1) < 0.9$. S druge strane, za $k = 3$ imamo

$$\begin{aligned}\mathbb{P}(X \leq 2) = \mathbb{P}(X \leq 1) + \mathbb{P}(X = 2) &= 0.73728 + \binom{5}{2} 0.2^2 0.8^3 \\ &= 0.73728 + 0.2048 = 0.94208 > 0.9.\end{aligned}$$

Dakle, traženi broj dana je 3. Uočimo da traženu nejednakost za $k = 5$ nismo ni provjeravali jer je očito $\mathbb{P}(X \leq 0) \leq \mathbb{P}(X \leq 1) < 0.9$.

□

Primjer 5.10. Neka je T broj pokušaja do pojave prvog uspjeha u nizu nezavisnih pokusa pri čemu je vjerojatnost uspjeha u svakom pokusu jednaka $p \in (0, 1]$. Tada je

$$\mathbb{P}(T = k) = q^{k-1} p\,,\ k = 1, 2, \dots$$

Slučajnu varijablu T s ovom distribucijom zovemo geometrijska slučajna varijabla na $\mathbb{N}$ s parametrom (uspjeha) p te označavamo $T \sim G(p)$. Uočimo, zaista je $\sum_{k=1}^{\infty} \mathbb{P}(T = k) = \sum_{k=1}^{\infty} q^{k-1} p = 1$. Specijalno, vjerojatnost da se nikada neće pojaviti uspjeh za bilo koji $p \in (0, 1]$ je $\mathbb{P}(T = \infty) = 1 - \sum_{k=1}^{\infty} \mathbb{P}(T = k) = 0$.

Zadatak 5.11. U 18. stoljeću lutrija se igrala tako da se izvlačila jedna od 32 različite kuglice, a dobitnici su dobivali isplatu 28 puta veću od uloga. Ljudi su zaključili da ovakva oklada pogoduje organizatorima[3] pa su oni, želeći dokazati poštenje, tvrdili da će proizvoljno odabrana kuglica biti izvučena barem jednom u 22 pokušaja i za to su nudili okladu s omjerom 1:1, tj. ukoliko se to ne dogodi dobijete isplatu dvostruko veću od uloga. Pokažite da takva oklada i dalje pogoduje organizatorima.

Rješenje. Fiksirajmo proizvoljnu kuglicu te neka je T broj igara potreban dok se ona prvi put ne izvuče. Želimo pokazati da je

$$\mathbb{P}(T \leq 22) > 0.5\,.$$

U svakoj igri vjerojatnost da izvučemo fiksiranu kuglicu jednaka je $\frac{1}{32}$ pa budući da su izvlačenja nezavisna slijedi da je

$$T \sim G(\tfrac{1}{32})\,.$$

Odredimo sada $\mathbb{P}(T > 22)$ (pa time i $\mathbb{P}(T \leq 22)$). Uočimo da je za sve $n \geq 0$

$$\begin{aligned}\mathbb{P}(T > n) = \sum_{k=n+1}^{\infty} \mathbb{P}(T = k) &= \sum_{k=n+1}^{\infty} \left(\frac{31}{32}\right)^{k-1} \frac{1}{32} \\ &= \frac{1}{32} \left(\frac{31}{32}\right)^{n} \sum_{k=0}^{\infty} \left(\frac{31}{32}\right)^{k} \\ &= \frac{1}{32} \left(\frac{31}{32}\right)^{n} \frac{1}{1 - \frac{31}{32}} = \left(\frac{31}{32}\right)^{n}.\end{aligned}$$

[3] To je zaista istina jer ćete "u prosjeku" dobiti jedanput u 32 igre (vidi pojam očekivanja za geometrijsku slučajnu varijablu u sljedećem potpoglavlju.) pa bi fer isplata trebala biti 32 puta veća od uloga.

Alternativno, i puno bolje,

$$\mathbb{P}\left(T > n\right) = \mathbb{P}\left(\text{u prvih } n \text{ pokušaja nije izvučena odabrana kuglica}\right) = \left(1 - \frac{1}{32}\right)^n = \left(\frac{31}{32}\right)^n$$

za svaki $n \geq 0$. Specijalno,

$$\mathbb{P}\left(T > 22\right) = \left(\frac{31}{32}\right)^{22} \approx 0.4973 < 0.5\,,$$

pa je vjerojatnost dobitka organizatora $\mathbb{P}\left(T \leq 22\right) > 0.5$. □

Napomena 5.12. Nekada se geometrijskom slučajnom varijablom naziva slučajna varijabla Y s distribucijom

$$\mathbb{P}\left(Y = k\right) = q^k p\,,\ k = 0, 1, 2, \ldots$$

Za razliku od prethodno definirane geometrijske slučajne varijable, slučajna varijabla Y predstavlja broj neuspjeha prije pojave prvog uspjeha te vrijedi $Y + 1 \sim G(p)$. Distribuciju slučajne varijable Y zvat ćemo geometrijska distribucija s parametrom p na $\mathbb{N}_0$ i označavati $Y \sim G_0(p)$.

Primjer 5.13. Neka je X slučajna varijabla s distribucijom

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

pri čemu je $\lambda > 0$. Lako se vidi da je X dobro definirana, a nazivamo je Poissonovom slučajna varijabla s parametrom λ i označavamo je s $X \sim P(\lambda)$. Ova distribucija nema direktnu interpretaciju kao na primjer binomna ili geometrijska, ali zbog tzv. *zakona rijetkih događaja* jedna je od najvažnijih distribucija u vjerojatnosti.

5.1 Matematičko očekivanje i varijanca

Motivacija: Neka je

$$X \sim \begin{pmatrix} 1 & 2 & 3 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

Koja je "očekivana" (ili "prosječna") vrijednost koju će poprimiti X? Logično bi bilo da je to 2. Ako je pak

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

onda za "očekivanu" vrijednost ima smisla uzeti $\frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{2}{3} \cdot 3 = \frac{5}{2} > 2$.

Definicija 5.14. Neka je X slučajna varijabla s distribucijom $X \sim \begin{pmatrix} a_1 & a_2 & \ldots \\ p_1 & p_2 & \ldots \end{pmatrix}$. **Matematičko očekivanje** od X je broj

$$\mathbb{E}\left[X\right] := \sum_{n=1}^{\infty} a_n p_n$$

ako taj red apsolutno konvergira (tj. ako $\sum_{n=1}^{\infty} |a_n| p_n < \infty$). U suprotnom kažemo da X nema očekivanje.

Uočite da svaka slučajna varijabla koja poprima konačno mnogo različitih vrijednosti nužno ima matematičko očekivanje.

Zadatak 5.15. Slučajna varijabla X ima razdiobu

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

Odredite a tako da X ima očekivanje $\mathbb{E}\,X = \frac{1}{3}$.

Rješenje. Prema definiciji očekivanja znamo da je

$$\mathbb{E}\,X = -1 \cdot \tfrac{1}{2} + 0 \cdot \tfrac{1}{3} + a \cdot \tfrac{1}{6} = -\tfrac{1}{2} + \tfrac{a}{6},$$

pa dobivamo da je tražena vrijednost $a = 5$. ☐

Napomena 5.16 (Svojstva očekivanja)**.** (a) Ako je $X = c$, tj. $X \sim \binom{c}{1}$, imamo $\mathbb{E}\,[X] = c$.

(b) Ako je $X \geq 0$, tj. $a_n \geq 0$ za sve n, onda je $\mathbb{E}\,[X] \geq 0$.

(c) Za $\alpha, \beta \in \mathbb{R}$ i slučajne varijable X i Y je $\mathbb{E}\,[\alpha X + \beta Y] = \alpha\,\mathbb{E}\,[X] + \beta\,\mathbb{E}\,[Y]$. Ovo svojstvo nekada nazivamo linearnost očekivanja.

(d) Ako je $X \leq Y$, onda je i $\mathbb{E}\,[X] \leq \mathbb{E}\,[Y]$.

(e) Ako je $X \sim \begin{pmatrix} a_1 & a_2 & \cdots \\ p_1 & p_2 & \cdots \end{pmatrix}$ i $g : \mathbb{R} \to \mathbb{R}$ neka funkcija, onda je

$$\mathbb{E}\,[g(X)] = \sum_{n=1}^{\infty} g(a_n) p_n = \sum_{n=1}^{\infty} g(a_n) \mathbb{P}\,(X = a_n)$$

pod uvjetom da je $\sum\limits_{n=1}^{\infty} |g(a_n)| p_n < \infty$. Ova formula je dosta korisna jer ne moramo računati distribuciju slučajne varijable $g(X)$.

Zadatak 5.17. Bacamo simetričnu kocku. Ako slučajna varijabla X predstavlja broj koji je pao na kocki, nađite matematičko očekivanje slučajne varijable $Y = \frac{1}{X+1}$.

Rješenje. Očito je

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

Neka je $g \colon \mathbb{R} \to \mathbb{R}$ s pravilom pridruživanja $g(x) = \frac{1}{x+1}$. Tada je $Y = g(X)$, pa po prošloj napomeni imamo da je

$$\mathbb{E}\,[Y] = \sum_{i=1}^{6} g(i) \mathbb{P}\,(X = i) = \frac{1}{2} \cdot \frac{1}{6} + \frac{1}{3} \cdot \frac{1}{6} + \frac{1}{4} \cdot \frac{1}{6} + \frac{1}{5} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{7} \cdot \frac{1}{6} = \frac{223}{840}.$$

☐

Zadatak 5.18. Marko je na piknik ponio 5 konzervi: 2 graha, 2 paprike i 1 tunu. Nakon što ga je ulovila kiša i oprala naljepnice s konzervi, Marko je odlučio otvarati konzerve sve dok ne dobije sva tri jela. Odredite očekivani broj otvaranja konzervi do sva tri jela.

Rješenje. Neka je X broj otvaranja do sva tri jela. Kako bismo izračunali $\mathbb{E}[X]$, prvo moramo pronaći distribuciju od X. Za početak, X jedino može poprimiti vrijednosti $3, 4$ ili 5.

- Odredimo $\mathbb{P}(X = 3)$. Uočimo da je $X = 3$ akko smo u prva tri otvaranja izvukli tri različite konzerve. Na primjer, vjerojatnost da izvučemo grah, papriku pa tunu je

$$\mathbb{P}(\{GPT\}) = \frac{2 \cdot 2 \cdot 1}{5 \cdot 4 \cdot 3} = \frac{1}{15}\,.$$

  Nadalje, vjerojatnost da izvučemo na primjer tunu, grah pa papriku je

$$\mathbb{P}(\{TGP\}) = \frac{1 \cdot 2 \cdot 2}{5 \cdot 4 \cdot 3} = \mathbb{P}(\{GPT\})\,.$$

  Dakle, u brojniku se samo zamijenio poredak kojim množimo brojeve 2,2 i 1 pa je dakle vjerojatnost ostala ista. Budući da ovo vrijedi za svih $3! = 6$ permutacija graha, paprike i tune, imamo da je

$$\mathbb{P}(X = 3) = 6 \cdot \frac{1}{15} = \frac{2}{5}\,.$$

- Odredimo $\mathbb{P}(X = 5)$ jer je lakše nego $\mathbb{P}(X = 4)$. Uočimo da je $X = 5$ akko je tuna izvučena zadnja. Ukupan broj poredaka u kojima se to dogodi je $\binom{4}{2} = 6$ (biramo 2 mjesta od prva 4 na kojima smo izvukli npr. grah), a svaki takav poredak ima vjerojatnost

$$\frac{2 \cdot 1 \cdot 2 \cdot 1 \cdot 1}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = \frac{1}{30}\,.$$

  Dakle,

$$\mathbb{P}(X = 5) = 6 \cdot \frac{1}{30} = \frac{1}{5}\,.$$

- Preostaje $\mathbb{P}(X = 4) = 1 - \mathbb{P}(X = 3) - \mathbb{P}(X = 5) = \frac{2}{5}$.

Dakle,

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

pa je

$$\mathbb{E}[X] = 3 \cdot \frac{2}{5} + 4 \cdot \frac{2}{5} + 5 \cdot \frac{1}{5} = 3.8\,.$$

Kao dobru vježbu iz prebrojavanja probajmo direktno izračunati $\mathbb{P}(X = 4)$. Imamo da je $X = 4$ akko smo u prva četiri izvlačenja na prva tri mjesta izvukli dvije iste konzerve (dakle grah ili paprika), a na neko od preostala dva mjesta tunu. Ako je fiksiramo koju smo konzervu izvukli dva puta (npr. G), svaki takav poredak (npr. $GGPT$) ima vjerojatnost

$$\mathbb{P}(\{GGPT\}) = \frac{2 \cdot 1 \cdot 2 \cdot 1}{5 \cdot 4 \cdot 3 \cdot 2} = \frac{1}{30}\,.$$

Takvih poredaka ima $\binom{3}{2} \cdot 2! = 6$ (biramo mjesta za grah i permutiramo papriku i tunu na ostala dva mjesta), a isto vrijedi ako smo fiksirali papriku pa zaključujemo da je

$$\mathbb{P}(X = 4) = 2 \cdot \binom{3}{2} \cdot 2! \cdot \frac{1}{30} = \frac{2}{5}.$$

□

Definicija 5.19. Ako slučajna varijabla X ima očekivanje, **varijanca** od X je

$$\mathrm{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right].$$

Varijanca nam daje dodatnu informaciju od slučajnoj varijabli. Intuitivno, varijanca mjeri koliko u prosjeku slučajna varijabla odstupa od svoje očekivane vrijednosti.

Primjer 5.20. Neka su

$$X \sim \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \; Y \sim \begin{pmatrix} -1 & 1 \\ 1/2 & 1/2 \end{pmatrix}, \; Z \sim \begin{pmatrix} -100 & 100 \\ 1/2 & 1/2 \end{pmatrix}.$$

Tada je

$$\mathbb{E}[X] = \mathbb{E}[Y] = \mathbb{E}[Z] = 0,$$

ali

$$\begin{aligned} \mathrm{Var}(X) &= \mathbb{E}\left[(X-0)^2\right] = 0^2 \cdot 1 = 0, \\ \mathrm{Var}(Y) &= \mathbb{E}\left[Y^2\right] = (-1)^2 \cdot 1/2 + 1^2 \cdot 1/2 = 1, \\ \mathrm{Var}(Z) &= \mathbb{E}\left[Z^2\right] = 10000. \end{aligned}$$

Dakle, sve slučajne varijable imaju isto očekivanje, ali Z ima najveće odstupanje.

Napomena 5.21. Varijancu najčešće računamo pomoću formule

$$\mathrm{Var}(X) = \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2.$$

Također, za $a, b \in \mathbb{R}$ vrijedi

$$\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X).$$

Zadatak 5.22. Bacamo dvije simetrične kocke te neka je X apsolutna vrijednost razlike brojeva na kockama. Pronađite distribuciju, očekivanje i varijancu od X.

Rješenje. Sljedeća tablica prikazuje vrijednosti slučajne varijable X za sve moguće ishode bacanja dvije kocke.

1.kocka \ 2.kocka	1	2	3	4	5	6
1	0	1	2	3	4	5
2	1	0	1	2	3	4
3	2	1	0	1	2	3
4	3	2	1	0	1	2
5	4	3	2	1	0	1
6	5	4	3	2	1	0

Budući da svaki ishod ima vjerojatnost $\frac{1}{36}$, prebrojavanjem dobivamo da je

$$X \sim \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ \frac{6}{36} & \frac{10}{36} & \frac{8}{36} & \frac{6}{36} & \frac{4}{36} & \frac{2}{36} \end{pmatrix}.$$

Sada računamo

$$\mathbb{E}[X] = 0 \cdot \frac{6}{36} + 1 \cdot \frac{10}{36} + 2 \cdot \frac{8}{36} + 3 \cdot \frac{6}{36} + 4 \cdot \frac{4}{36} + 5 \cdot \frac{2}{36} = \frac{35}{18} \approx 1.94$$

i

$$\begin{aligned} \operatorname{Var}(X) &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\ &= 0^2 \cdot \frac{6}{36} + 1^2 \cdot \frac{10}{36} + 2^2 \cdot \frac{8}{36} + 3^2 \cdot \frac{6}{36} + 4^2 \cdot \frac{4}{36} + 5^2 \cdot \frac{2}{36} - \left(\frac{35}{18}\right)^2 \\ &= \frac{105}{18} - \left(\frac{35}{18}\right)^2 \approx 2.052\,. \end{aligned}$$

□

Napomena 5.23. Budući da se često javljaju u primjenama, korisno je znati očekivanje i varijancu sljedećih distribucija (za dokaze vidi predavanja):

(a) $X \sim B(1,p) \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix} \Rightarrow \mathbb{E}[X] = p$, $\operatorname{Var}(X) = pq$, pri čemu je ovdje i u sljedećim primjerima $q = 1-p$.

(b) $X \sim B(n,p) \Rightarrow \mathbb{E}[X] = np$ i $\operatorname{Var}(X) = npq$.

(c) $X \sim G(p) \Rightarrow \mathbb{E}[X] = \frac{1}{p}$ i $\operatorname{Var}(X) = \frac{q}{p^2}$.

(d) $X \sim P(\lambda) \Rightarrow \mathbb{E}[X] = \operatorname{Var}(X) = \lambda$.

Napomena 5.24. Ukoliko je $X \sim G_0(p)$, tada je $X+1 \sim G(p)$ pa slijedi da je

$$\mathbb{E}[X] = \mathbb{E}[X+1] - 1 = \frac{1}{p} - 1 = \frac{q}{p}\,,$$

i

$$\operatorname{Var}(X) = \operatorname{Var}(X+1) = \frac{q}{p^2}\,.$$

Zadatak 5.25. U jednakokračnom trokutu osnovice duljine 1 cm i krakova duljine 2 cm slučajno i nezavisno biramo 1000 točaka. Izračunajte očekivanje i varijancu broja točaka koje se nalaze unutar kruga upisanog tom trokutu.

Rješenje. Neka je X broj točaka koje su upale u krug upisan tom trokutu i $r > 0$ radijus tog kruga. Tada je očito $X \sim B(1000, p)$ pri čemu je vjerojatnost "uspjeha" u svakom od pokusa (izbor točke iz trokuta)

$$p = \mathbb{P}(\{\text{točka upala u krug}\}) = \frac{\lambda(\text{krug})}{\lambda(\text{trokut})} = \frac{r^2\pi}{\frac{1}{2} \cdot 1 \cdot \sqrt{2^2 - (\frac{1}{2})^2}} = \frac{4r^2\pi}{\sqrt{15}}\,.$$

Budući da je $\frac{\sqrt{15}}{4} = \lambda(\text{trokut}) = \frac{1\cdot r}{2} + 2 \cdot \frac{2\cdot r}{2}$, slijedi da je $r = \frac{\sqrt{15}}{10}$ te

$$p = \tfrac{\sqrt{15}\pi}{25} \approx 0.4867\,.$$

Dakle,

$$\begin{aligned}\mathbb{E}\,[X] &= 1000 \cdot p \approx 486.7\\ \mathrm{Var}\,(X) &= 1000 \cdot pq \approx 249.8\,.\end{aligned}$$

□

Napomena 5.26 (*)**.** Prethodni zadatak daje metodu kako simulacijom procijeniti površinu npr. jediničnog kruga (znamo da je njegova površina $1^2\pi = \pi$). Naime, ako gledamo jedinični krug oko ishodišta, on je upisan kvadratu $\Omega = [-1, 1]^2$. Sada, ako znamo simulirati slučajan izbor velikog broja točaka iz kvadrata Ω (što se može), po tzv. zakonu velikih brojeva, postotak točaka koje će upasti u krug približno će biti jednak

$$p = \mathbb{P}\left(\{\text{jedna točka upadne u krug}\}\right) = \frac{\pi}{4}\,.$$

Dakle, ako taj postotak pomnožimo s 4 dobit ćemo procjenu za površinu kruga. Ovaj jednostavan primjer ilustrira glavnu ideju tzv. *Monte-Carlo metoda* koje se danas naširoko koriste u primjenama.

5.2 Nezavisnost

Definicija 5.27. Neka su X i Y diskretne slučajne varijable (definirane na istom vjerojatnosnom prostoru) koje poprimaju vrijednosti $(a_i)_{i\in\mathbb{N}}$, odnosno $(b_j)_{j\in\mathbb{N}}$. Kažemo da su X i Y nezavisne ako vrijedi

$$\mathbb{P}\left(X = a_i, Y = b_j\right) = \mathbb{P}\left(X = a_i\right)\mathbb{P}\left(Y = b_j\right)\,,\ \text{za sve } i, j \in \mathbb{N}\,, \tag{2}$$

pri čemu je $\{X = a_i, Y = b_j\}$ oznaka za događaj $\{X = a_i\} \cap \{Y = b_j\}$.

Analogno se definira nezavisnost n diskretnih slučajnih varijabli $X_1, \ldots, X_n$: treba vrijediti

$$\mathbb{P}\left(X_1 = x_1, \ldots, X_n = x_n\right) = \mathbb{P}\left(X_1 = x_1\right) \cdots \mathbb{P}\left(X_n = x_n\right)$$

za sve $x_1, \ldots, x_n \in \mathbb{R}$ takve da X_i može poprimiti vrijednost x_i za sve $i = 1, \ldots, n$.

Napomena 5.28. Ako su X i Y nezavisne, tj. vrijedi (2), tada i za sve $A, B \subseteq \mathbb{R}$ vrijedi

$$\mathbb{P}\left(X \in A, Y \in B\right) = \mathbb{P}\left(X \in A\right)\mathbb{P}\left(Y \in B\right)\,.$$

Napomena 5.29. (a) Ako su X i Y **nezavisne** slučajne varijable takve da očekivanja od X i Y postoje, tada je

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

Obrat općenito ne vrijedi!

(b) Ako su X i Y **nezavisne** slučajne varijable takve da $\mathrm{Var}(X), \mathrm{Var}(Y) < \infty$, onda je

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) .$$

Gornje tvrdnje se lako poopćuju na slučaj n nezavisnih slučajnih varijabli $X_1, \ldots, X_n$. Na primjer, u tom slučaju

$$\mathrm{Var}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \mathrm{Var}(X_k) .$$

Zadatak 5.30. Neka su X i Y nezavisne Poissonove slučajne varijable s parametrom $\lambda = 3$. Izračunajte $\mathbb{P}(X+Y=1)$ i $\mathbb{E}[(X-Y)^2]$.

Rješenje. Budući da X i Y poprimaju vrijednosti u $\mathbb{N}_0$, $\{X+Y=1\} = \{X=0, Y=1\} \cup \{X=1, Y=0\}$ i pri tome su ova dva događaja očito disjunktna (ne može istovremeno biti $X=0$ i $X=1$). Dakle[4], koristeći nezavisnost u drugom retku imamo

$$\begin{aligned}
\mathbb{P}(X+Y=1) &= \mathbb{P}(X=0, Y=1) + \mathbb{P}(X=1, Y=0) \\
&\overset{\text{nez.}}{=} \mathbb{P}(X=0)\mathbb{P}(Y=1) + \mathbb{P}(X=1)\mathbb{P}(Y=0) \\
&= e^{-3}\frac{3^0}{0!}e^{-3}\frac{3^1}{1!} + e^{-3}\frac{3^1}{1!}e^{-3}\frac{3^0}{0!} = 6e^{-6}.
\end{aligned}$$

Nadalje,

$$\begin{aligned}
\mathbb{E}[(X-Y)^2] &= \mathbb{E}\left[X^2 - 2XY + Y^2\right] = \mathbb{E}[X^2] - 2\mathbb{E}[XY] + \mathbb{E}[Y^2] \\
&\overset{\text{nez.}}{=} \mathbb{E}[X^2] - 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[Y^2],
\end{aligned}$$

pri čemu smo u drugoj jednakosti iskoristili linearnost očekivanja. Budući da je $\mathbb{E}[X] = \mathbb{E}[Y] = \lambda = 3$ i $\mathrm{Var}(X) = \mathrm{Var}(Y) = \lambda = 3$, iz formule $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$ slijedi da je

$$\mathbb{E}[X^2] = \mathrm{Var}(X) + (\mathbb{E}[X])^2 = 3 + 3^2 = 12 = \mathbb{E}[Y^2]$$

pa zaključujemo

$$\mathbb{E}[(X-Y)^2] = 2 \cdot 12 - 2 \cdot 3 \cdot 3 = 6 .$$

Alternativno, koristeći nezavisnost i svojstva varijance

$$\begin{aligned}
\mathbb{E}[(X-Y)^2] &= \mathrm{Var}(X-Y) + \mathbb{E}[X-Y]^2 = \left[\mathbb{E}[X-Y] = \mathbb{E}[X] - \mathbb{E}[Y] = 0\right] \\
&= \mathrm{Var}(X-Y) \overset{\text{nez.}}{=} \mathrm{Var}(X) + \mathrm{Var}(-Y) = \mathrm{Var}(X) + (-1)^2\mathrm{Var}(Y) \\
&= 2 \cdot 3 = 6 .
\end{aligned}$$

□

[4] Može se pokazati da za dvije nezavisne Poissonove slučajne varijable X i Y s parametrima λ i μ, vrijedi da je $X+Y$ ponovno Poissonova slučajna varijabla s parametrom $\lambda+\mu$ (DZ*). U našem slučaju je dakle $X+Y \sim P(6)$ što objašnjava dobiveni rezultat.

Napomena 5.31. Ako su $X_1, X_2, \dots, X_n$ nezavisne Bernoullijeve slučajne varijable s parametrom p, tj. $X_i \sim B(1,p)$ za svaki $i = 1, \dots, n$, intuitivno je jasno, a nije teško ni dokazati, da je

$$X := X_1 + X_2 + \cdots + X_n \sim B(n,p)\,.$$

Koristeći svojstva očekivanja i varijance te prethodnu tvrdnju, možemo na jednostavan način izračunati (ili pamtiti) očekivanje i varijancu binomne slučajne varijable. Naime, budući da je $\mathbb{E}[X_i] = p$ i $\mathrm{Var}(X_i) = pq$ za sve $i = 1, \dots, n$, slijedi da za $X \sim B(n,p)$ imamo $\mathbb{E}[X] = \mathbb{E}[X_1 + X_2 + \cdots + X_n] = \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n] = np$ (zbog linearnosti očekivanja) te $\mathrm{Var}(X) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) = npq$ (zbog nezavisnosti).

Napomena 5.32. Ako slučajna varijabla X poprima vrijednosti u $\mathbb{N}_0$, tada vrijedi

$$\mathbb{E}[X] = \sum_{n=0}^{\infty} \mathbb{P}(X > n) = \sum_{n=1}^{\infty} \mathbb{P}(X \geq n)\,. \tag{3}$$

Na primjer, za $X \sim G(p)$, $p > 0$, smo već vidjeli da je $\mathbb{P}(X > n) = (1-p)^n = q^n$ za sve $n \in \mathbb{N}_0$ pa je zbog $q < 1$

$$\mathbb{E}[X] = \sum_{n=0}^{\infty} q^n = \frac{1}{1-q} = \frac{1}{p}\,.$$

Zadatak 5.33. Neka su X i Y nezavisne slučajne varijable takve da je $X \sim G(p_1)$, $Y \sim G(p_2)$, $p_1, p_2 \in (0,1)$. Izračunajte $\mathbb{E}[\min\{X,Y\}]$. (*Uputa:* Iskoristite (3).)

Rješenje. Uočimo da slučajna varijabla $\min\{X,Y\}$ poprima vrijednosti u $\mathbb{N}$ te da je

$$\mathbb{P}(\min\{X,Y\} > n) = \mathbb{P}(X > n, Y > n) \overset{nez.}{=} \mathbb{P}(X > n)\,\mathbb{P}(Y > n) = (1-p_1)^n(1-p_2)^n$$

za sve $n \geq 0$. Po prethodnoj napomeni sada slijedi da je

$$\begin{aligned}
\mathbb{E}[\min\{X,Y\}] &= \sum_{n=0}^{\infty} \mathbb{P}(\min\{X,Y\} > n) = \sum_{n=0}^{\infty}((1-p_1)(1-p_2))^n \\
&= \frac{1}{1-(1-p_1)(1-p_2)} = \frac{1}{p_1 + p_2 - p_1 p_2}\,.
\end{aligned}$$

□

Napomena 5.34 (*)**.** Uočimo, iz prethodnog zadatka slijedi da je za sve $n \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{P}(\min\{X,Y\} = n) &= \mathbb{P}(\min\{X,Y\} > n-1) - \mathbb{P}(\min\{X,Y\} > n) \\
&= ((1-p_1)(1-p_2))^{n-1}(\underbrace{1-(1-p_1)(1-p_2)}_{=:p'}) = (1-p')^{n-1}p'\,.
\end{aligned}$$

Dakle, $\min\{X,Y\}$ je ponovno geometrijska slučajna varijabla i to s parametrom uspjeha $p' = 1 - (1-p_1)(1-p_2)$. Intuitivno objašnjenje za ovo je sljedeće. Zamislimo da nezavisno provodimo dva niza pokusa, pri čemu su vjerojatnosti uspjeha p_1 i p_2, te neka su X i Y vremena prvog uspjeha u svakom od pokusa. Tada su X i Y nezavisne geometrijske slučajne varijable s parametrima

p_1, odnosno p_2, a $\min\{X, Y\}$ je prvo vrijeme kada će se u barem jednom od dva pokusa pojaviti uspjeh. Budući da je vjerojatnost barem jednog uspjeha u svakom pokuse jednaka

$$p' := 1 - \mathbb{P}\left(\{\text{oba neuspjeha}\}\right) \overset{nez.}{=} 1 - (1 - p_1)(1 - p_2)\,,$$

slijedi da je $\min\{X, Y\}$ geometrijska s parametrom p'.

Zadatak 5.35. Bacamo simetričnu kocku dok se šestica ne pojavi po drugi put. Ako X označava potreban broj bacanja, odredite (a) $\mathbb{P}(X \leq 4)$, (b) $\mathbb{E}[X]$ i (c)* $\mathrm{Var}(X)$

Rješenje.

(a) Za svaki $n \geq 2$ uočimo da je $X = n$ ako i samo ako smo šesticu dobili u n-tom bacanju i točno jednom u prvih $n - 1$ bacanja. Zbog nezavisnosti bacanja, svaki ishod n bacanja kocke u kojem imamo ukupno dvije šestice, a na ostalim mjestima bilo što osim šestice, ima vjerojatnost

$$\left(\frac{5}{6}\right)^{n-2}\left(\frac{1}{6}\right)^{2}.$$

Budući da takvih ishoda sa šesticom na kraju ima točno $\binom{n-1}{1} = n - 1$ (biramo mjesto za drugu šesticu), slijedi da je

$$\mathbb{P}(X = n) = (n-1)\left(\frac{5}{6}\right)^{n-2}\left(\frac{1}{6}\right)^{2}, \ n \geq 2\,.$$

Dakle,

$$\mathbb{P}(X \leq 4) = \mathbb{P}(X = 2) + \mathbb{P}(X = 3) + \mathbb{P}(X = 4) = \left(\frac{1}{6}\right)^{2} + 2\cdot\frac{5}{6}\left(\frac{1}{6}\right)^{2} + 3\cdot\left(\frac{5}{6}\right)^{2}\left(\frac{1}{6}\right)^{2} \approx 0.104\,.$$

(b) Računamo

$$\mathbb{E}[X] = \sum_{n=2}^{\infty} n\mathbb{P}(X = n) = \left(\frac{1}{6}\right)^{2}\sum_{n=2}^{\infty} n(n-1)\left(\frac{5}{6}\right)^{n-2}.$$

Koristeći derivaciju geometrijskog reda[5] dobivamo da je

$$\mathbb{E}[X] = \left(\frac{1}{6}\right)^{2}\frac{2}{(1-\frac{5}{6})^3} = 12\,.$$

(c) Kako bi odredili varijancu, ispada da je lakše najprije odrediti $\mathbb{E}[X(X+1)]$ koristeći treću derivaciju geometrijskog reda:

$$\begin{aligned}\mathbb{E}[X(X+1)] &= \sum_{n=2}^{\infty} n(n+1)\mathbb{P}(X = n) = \left(\frac{1}{6}\right)^{2}\sum_{n=2}^{\infty}(n-1)n(n+1)\left(\frac{5}{6}\right)^{(n+1)-3} \\ &= \left(\frac{1}{6}\right)^{2}\cdot\frac{6}{(1-\frac{5}{6})^4} = 6^3\,,\end{aligned}$$

[5]Deriviranjem identiteta $\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n$, dobivamo da je $\frac{1}{(1-x)^2} = \sum_{n=1}^{\infty} nx^{n-1}$ i $\frac{2}{(1-x)^3} = \sum_{n=2}^{\infty} n(n-1)x^{n-2}$.

iz čega onda slijedi

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}[X(X+1)] - \mathbb{E}[X] - (\mathbb{E}[X])^2 \\ &= 6^3 - 12 - 12^2 = 60\,.\end{aligned}$$

Ipak, varijancu možemo izračunati puno elegantnije. Naime, ako s T_1 označimo broj bacanja do pojave prve šestice, a T_2 broj bacanja nakon toga dok ne padne druga šestica, onda je očito $X = T_1 + T_2$. Nadalje, T_1 i T_2 imaju geometrijsku distribuciju s parametrom $\frac{1}{6}$ i intutitvno je jasno da su T_1 i T_2 nezavisne. Sada slijedi da je

$$\mathrm{Var}(X) = \mathrm{Var}(T_1 + T_2) \overset{nez.}{=} \mathrm{Var}(T_1) + \mathrm{Var}(T_2) = 2 \cdot \frac{\frac{5}{6}}{(\frac{1}{6})^2} = 60\,.$$

Na isti način smo mogli izračunati $\mathbb{E}[X]$, s tim da tu nezavisnost nije bitna. Uočite, ovaj argument direktno možemo poopćiti na slučaj kada čekamo dok šestica padne po n-ti put za proizvoljan $n \in \mathbb{N}$.

$\square$

Napomena 5.36. Ako slučajna varijabla $X \sim \begin{pmatrix} a_1 & a_2 & \cdots \\ p_1 & p_2 & \cdots \end{pmatrix}$ poprima samo nenegativne vrijednosti, tj. $a_i \geq 0$ za sve $i \in \mathbb{N}$, te ako je $\sum_{i=1}^{\infty} a_i p_i = +\infty$, stavljamo $\mathbb{E}[X] := +\infty$.

Zadaci za vježbu

Zadatak 5.37. Odredite očekivanje i varijancu slučajnih varijabli X, X^2 i $|X-3|$ iz zadatka 5.5.

Zadatak 5.38. Svaki dan vozite se bez karte busom kojeg za vrijeme vaših vožnji posjećuje kontrola s vjerojatnosti $\frac{1}{5}$. Neka je X redni broj vožnje kada vas prvi put ulovi kontrola. Odredite vjerojatnost da vas kontrola ne uhvati barem prva dva puta.

Zadatak 5.39. Slučajna varijabla X ima razdiobu

$$X \sim \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} & \cdots \\ \frac{c}{4} & \frac{c}{4^2} & \frac{c}{4^3} & \cdots & \frac{c}{4^n} & \cdots \end{pmatrix}.$$

(a) Odredite c.

(b) Odredite razdiobu slučajne varijable $Y = \sqrt{1 + \cos \frac{\pi}{X}}$.

(c) Izračunajte $\mathbb{E}[\frac{1}{X}]$.

Zadatak 5.40. Strastveni igrač Lota 6/45 odlučio je igrati kombinaciju od 6 brojeva sve dok ona ne bude izvučena. Označimo sa X redni broj izvlačenja u kojem je ta kombinacija izvučena.

(a) Odredite distribuciju slučajne varijable X.

(b) Koliko izvlačenja očekujemo da će proći, a da strastveni igrač pogodi izvučenu kombinaciju?

Zadatak 5.41. Konobar počinje smjenu s 0 kn. Od svakog gosta dobije napojnicu i to od 10 kn ili 5 kn, pri čemu je manja napojnica dva puta vjerojatnija. To jutro je konobar poslužio 4 gosta.

(a) Izračunajte vjerojatnost da je konobar dobio manje od 30 kn od napojnica.

(b) Koliki je očekivani iznos koji je konobar dobio od napojnica?

Zadatak 5.42. Na raspolaganju nam je 6 žarulja, od kojih su 2 ispravne i 4 neispravne. Žarulje isprobavamo na lampi jednu za drugom, te s X označimo broj pokušaja do pojave svjetlosti. Odredite $\mathbb{E}[X]$ i $\operatorname{Var}[X]$.

Zadatak 5.43. (a) Bacaju se dvije simetrične kocke i rezultati se zbroje. Izračunajte razdiobu, očekivanje i varijancu dobivenog zbroja.

(b) Baca se deset simetričnih kocki i rezultati se zbroje. Izračunajte očekivanje i varijancu dobivenog zbroja.

Zadatak 5.44 (*). Kockar se kladi na igru u kojoj je vjerojatnost dobitka $p \leq 1/2$. Ako dobije u igri, dobitak je jednak iznosu uloga (ako izgubi, gubi ulog). Kockar prestaje igrati čim dobije igru, a ako izgubi, u sljedećoj se kladi sa dvostruko većim ulogom. Prva oklada je 1 kuna, sljedeća (ako izgubi prvu) je 2 kune, itd. – u n-toj igri ulog je 2^{n-1} kuna.

(a) Pokažite da je kockarov ukupni dobitak 1 kuna s vjerojatnošću 1. (Zvuči super, zar ne? Problem je što u praksi nitko nema neograničeno bogatstvo. Vidi (b) dio.)

(b) Koliko će očekivano kockar uložiti novaca prije nego dobije igru?

Rješenje.

(a) Neka je T broj oklada koje će kockar odigrati do prvog dobitka. Tada je očito $\mathbb{P}(T = n) = (1-p)^{n-1}p$ za svaki $n \in \mathbb{N}$, odnosno $T \sim G(p)$. S slučaju da je $T = n$, tada je u n-toj igri pobijedio i dobio 2^{n-1} kuna, ali je do tada već izgubio $1 + 2 + \cdots + 2^{n-2} = 2^{n-1} - 1$ kuna, što znači da mu je ukupni dobitak točno 1 kuna. Dakle $\mathbb{P}(X = 1 | T = n) = 1$ za sve $n \in \mathbb{N}$. Budući da znamo da je $\mathbb{P}(T < \infty) = \mathbb{P}(\cup_{n=1}^{\infty}\{T = n\}) = 1$, tj. kockar će sigurno u nekom trenutku pobijediti, po formuli potpune vjerojatnosti sada dobivamo da je

$$\mathbb{P}(X = 1) = \sum_{n=1}^{\infty} \mathbb{P}(X = 1 | T = n)\, \mathbb{P}(T = n) = \sum_{n=1}^{\infty} \mathbb{P}(T = n) = 1\,.$$

(b) Budući da će kockar dobiti T-toj igri, do tada će (uključujući i zadnju igru) uložiti $1 + 2 + \cdots + 2^{T-1} = 2^T - 1$ kuna. Dakle, očekivani ulog je

$$\mathbb{E}\left[2^T - 1\right] = \mathbb{E}\left[2^T\right] - 1 = \sum_{n=1}^{\infty} 2^n (1-p)^{n-1} p - 1 = 2p \sum_{n=1}^{\infty} (2(1-p))^{n-1} - 1 = +\infty$$

pri čemu posljednja jednakost slijedi jer je, zbog pretpostavke $p \leq 1/2$, $2(1-p) \geq 1$.

□

6 Slučajni vektori - zavisnost

Primjer 6.1. Neka je $\Omega = \{\omega_1, \omega_2, \omega_3\}$, $\mathbb{P}(\{\omega_i\}) = 1/3$. Definirajmo slučajne varijable: $X(\omega_1) = 1$, $X(\omega_2) = 2$, $X(\omega_3) = 3$, $Y(\omega_1) = 2$, $Y(\omega_2) = 3$, $Y(\omega_3) = 1$. Odmah slijedi da je

$$X, Y \sim \begin{pmatrix} 1 & 2 & 3 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

Dakle, X i Y imaju jednaku distribuciju. Pogledajmo sada distribuciju od $X + Y$. Na primjer, imamo da je $(X+Y)(\omega_1) = X(\omega_1) + Y(\omega_1) = 1 + 2 = 3$. Slični račun za elementarne događaje ω_2 i ω_3 nam onda daje

$$X + Y \sim \begin{pmatrix} 3 & 4 & 5 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}.$$

S druge strane,

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

dakle, $X+Y$ i $X+X$ nemaju istu distribuciju iako $X \sim Y$. Općenito, za distribuciju od $g(X,Y)$ potrebno je znati zajedničku distribuciju od X i Y.

Definicija 6.2. Za diskretne slučajne varijable X i Y definirane na $(\Omega, \mathcal{F}, \mathbb{P})$, funkciju $(X,Y) : \Omega \to \mathbb{R}^2$ zovemo diskretan slučajni vektor.

Ako X može poprimiti vrijednosti a_i, $i \in \mathbb{N}$, a Y vrijednosti b_j, $j \in \mathbb{N}$, parove (a_i, b_j), $i, j \in \mathbb{N}$ i vjerojatnosti

$$p_{i,j} := \mathbb{P}(X = a_i, Y = b_j)\ ,\ i, j \in \mathbb{N}$$

zovemo distribucijom slučajnog vektora (X,Y) i zapisujemo je preko tablice distribucije

$X \backslash Y$	b_1	b_2	$\cdots$	$\cdots$	$\sum$
a_1	$p_{1,1}$	$p_{1,2}$	$\cdots$	$\cdots$	q_1
a_2	$p_{2,1}$	$p_{2,2}$	$\cdots$	$\cdots$	q_2
$\vdots$	$\vdots$	$\vdots$	$\ddots$		$\vdots$
$\vdots$	$\vdots$	$\vdots$		$\ddots$	$\vdots$
$\sum$	r_1	r_2	$\cdots$	$\cdots$	1

pri čemu je npr.

$$q_1 := \sum_{j=1}^{\infty} p_{1,j} = \sum_{j=1}^{\infty} \mathbb{P}(X = a_1, Y = b_j) = \mathbb{P}\left(\{X = a_1\} \cap \underbrace{\left(\cup_{j=1}^{\infty} \{Y = b_j\}\right)}_{=\Omega} \right) = \mathbb{P}(X = a_1)\ .$$

Iz tablice lako iščitamo da vrijedi

$$X \sim \begin{pmatrix} a_1 & a_2 & \cdots\cdots \\ q_1 & q_2 & \cdots\cdots \end{pmatrix} \text{ i } Y \sim \begin{pmatrix} b_1 & b_2 & \cdots\cdots \\ r_1 & r_2 & \cdots\cdots \end{pmatrix}.$$

Napomena 6.3. (i) Uočimo, X i Y su nezavisne ako i samo ako je

$$p_{i,j} = \mathbb{P}(X = a_i, Y = b_j) = \mathbb{P}(X = a_i)\,\mathbb{P}(Y = b_j) = q_i r_j$$

za sve $i, j \in \mathbb{N}$.

(ii) Ako je $g : \mathbb{R}^2 \to \mathbb{R}$ funkcija, tada je za svaki $c \in \mathbb{R}$

$$\mathbb{P}(g(X,Y) = c) = \sum_{\substack{i,j=1 \\ g(a_i,b_j)=c}}^{\infty} \underbrace{\mathbb{P}(X = a_i, Y = b_j)}_{=p_{i,j}}$$

te vrijedi

$$\mathbb{E}[g(X,Y)] = \sum_{i,j=1}^{\infty} g(a_i, b_j) p_{i,j}\,. \tag{4}$$

Definicija 6.4. Neka su X i Y slučajne varijable takve da je $\mathrm{Var}(X), \mathrm{Var}(Y) < \infty$. **Kovarijanca** od X i Y je

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

Ukoliko je $\mathrm{Cov}(X,Y) = 0$, kažemo da su slučajne varijable X i Y nekorelirane, a u suprotnom kažemo da su korelirane. Prisjetimo se, za nezavisne slučajne varijable smo rekli da vrijedi $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$, tj. nezavisne slučajne varijable su uvijek nekorelirane. Obrat ne vrijedi, dat ćemo primjer na kraju poglavlja.

Nadalje, **koeficijent korelacije** slučajnih varijabli X i Y je

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}} \in [-1,1]\,.$$

Zadatak 6.5. Bacamo dvije simetrične kocke. Označimo s X broj šestica, a s Y broj jedinica koje su pale.

(a) Odredite distribuciju slučajnog vektora (X,Y) te distribucije slučajnih varijabli X i Y. Jesu li X i Y nezavisne?

(b) Izračunajte koeficijent korelacije slučajnih varijabli X i Y.

(c) Odredite distribuciju slučajne varijable $X + 2Y$.

Rješenje.

(a) Napravimo tablicu sa svim mogućim ishodima bacanja dvije kocke i odgovarajućim vrijednostima vektora (X,Y). Budući da svaki ishod ima vjerojatnost $\frac{1}{36}$, jednostavnim prebrojavanjem iz te tablice dobivamo da je distribucija slučajnog vektora (X,Y) dana s

$X\backslash Y$	0	1	2	$\sum$
0	$\frac{4}{9}$	$\frac{2}{9}$	$\frac{1}{36}$	$\frac{25}{36}$
1	$\frac{2}{9}$	$\frac{1}{18}$	0	$\frac{5}{18}$
2	$\frac{1}{36}$	0	0	$\frac{1}{36}$
$\sum$	$\frac{25}{36}$	$\frac{5}{18}$	$\frac{1}{36}$	1

,

Na primjer, $\mathbb{P}(X=0,Y=0)=4\cdot 4\cdot \frac{1}{36}=\frac{4}{9}$, $\mathbb{P}(X=0,Y=1)=2\cdot 4\cdot \frac{1}{36}=\frac{2}{9}$ i tako dalje.
Alternativno, distribuciju od (X,Y) dobivamo prebrojavanjem. Odredimo, u svrhu ilustracije, $\mathbb{P}(X=1,Y=1)$. Promatrani događaj se dogodio jedino ako je na prvoj kocki pala šestica, a na drugoj jedinica, ili obratno. Zato je

$$\mathbb{P}(X=1,Y=1)=\frac{1}{6}\cdot\frac{1}{6}+\frac{1}{6}\cdot\frac{1}{6}=\frac{1}{18}.$$

Nadalje, iz tablice distribucije odmah iščitavamo da vrijedi

$$X,Y\sim\begin{pmatrix} 0 & 1 & 2\\ \frac{25}{36} & \frac{5}{18} & \frac{1}{36}\end{pmatrix}.$$

Slučajne varijable X i Y nisu nezavisne jer je npr.

$$\mathbb{P}(X=1)\,\mathbb{P}(Y=2)=\frac{5}{18}\cdot\frac{1}{36}\neq 0=\mathbb{P}(X=1,Y=2).$$

(b) Očito je $\mathbb{E}[X]=\mathbb{E}[Y]=\frac{5}{18}+2\cdot\frac{1}{36}=\frac{1}{3}$, a koristeći (4) slijedi da je

$$\begin{aligned}\mathbb{E}[XY]&=0\cdot 0\cdot\frac{4}{9}+0\cdot 1\cdot\frac{2}{9}+0\cdot 2\cdot\frac{1}{36}\\&+1\cdot 0\cdot\frac{2}{9}+1\cdot 1\cdot\frac{1}{18}+1\cdot 2\cdot 0\\&+2\cdot 0\cdot\frac{1}{36}+2\cdot 1\cdot 0+2\cdot 2\cdot 0=\frac{1}{18}.\end{aligned}$$

Gornji račun smo proveli radi ilustracije primjene formule (4), ali u praksi bi odmah zaključili da je $\mathbb{E}[XY]=1\cdot 1\cdot\frac{1}{18}=\frac{1}{18}$ jer je to jedini sumand koji nije 0.
Sada zaključujemo da je

$$\operatorname{Cov}(X,Y)=\mathbb{E}[XY]-\mathbb{E}[X]\,\mathbb{E}[Y]=-\frac{1}{18}.$$

S druge strane, $\mathbb{E}[X^2]=\mathbb{E}[Y^2]=\frac{5}{18}+4\cdot\frac{1}{36}=\frac{7}{18}$, pa je $\operatorname{Var}(X)=\operatorname{Var}(Y)=\mathbb{E}[Y^2]-(\mathbb{E}[Y])^2=\frac{5}{18}$. Tako dobivamo da je traženi koeficijent korelacije jednak

$$\rho(X,Y)=\frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)}\sqrt{\operatorname{Var}(Y)}}=-\frac{1}{5}.$$

(c) Imamo sljedeću tablicu sa svim ishodima vektora (X,Y) i odgovarajućim vrijednostima od $X+2Y$.

$X\backslash Y$	0	1	2
0	0	2	4
1	1	3	5
2	2	4	6

Vidimo da slučajna varijabla $X+2Y$ može poprimiti vrijednosti $0,1,2,3,4,5,6$. Ipak, odmah uočimo da je

$$\mathbb{P}(X+2Y=5)=\mathbb{P}(X=1,Y=2)=0$$

te također

$$\mathbb{P}(X+2Y=6)=\mathbb{P}(X=2,Y=2)=0\,.$$

Dakle, $X+2Y$ može poprimiti samo vrijednosti $0,1,2,3,4$. Sada lako slijedi da je tražena distribucija

$$X+2Y\sim\begin{pmatrix}0 & 1 & 2 & 3 & 4\\ \frac{4}{9} & \frac{2}{9} & \frac{1}{4} & \frac{1}{18} & \frac{1}{36}\end{pmatrix}.$$

Na primjer, iz prethodne tablice slijedi

$$\mathbb{P}(X+2Y=0)=\mathbb{P}(X=0,Y=0)=\tfrac{4}{9}$$

ili

$$\mathbb{P}(X+2Y=2)=\mathbb{P}(X=0,Y=1)+\mathbb{P}(X=2,Y=0)=\tfrac{2}{9}+\tfrac{1}{36}=\tfrac{1}{4}\,.$$

□

Napomena 6.6. Koeficijent korelacije $\rho(X,Y)$ je jedan pokazatelj jačine zavisnosti varijabli X i Y. Intuitivno, vrijednost od $\rho(X,Y)$ blizu -1 i 1 indicira jaku zavisnost između X i Y. Uočimo, ako su X i Y nezavisne, onda je $\rho(X,Y)=0$, tj. one su nekorelirane. Drugim riječima, ako je $\rho(X,Y)\neq 0$ (kao u prethodnom zadatku), odmah slijedi da X i Y nisu nezavisne.

Nadalje, ukoliko je $\rho(X,Y)\neq 0$, predznak od $\rho(X,Y)$ nam govori nešto o načinu na koji su X i Y zavisne. Na primjer, u prethodnom zadatku imamo $\rho(X,Y)<0$ jer, intuitivno, veći broj šestica (dakle, veći X) povlači da je broj jedinica manji (dakle, da je Y manji).

Primjer 6.7. Neka je $X\sim\begin{pmatrix}-1 & 0 & 1\\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}\end{pmatrix}$ i $Y=X^2$. U ovom slučaju X i Y su potpuno zavisne, ali

$$\mathrm{Cov}(X,Y)=\mathbb{E}[XY]-\mathbb{E}[X]\,\mathbb{E}[Y]=\underbrace{\mathbb{E}\left[X^3\right]}_{=0}-0\cdot\mathbb{E}[Y]=0\,,$$

tj. X i Y su nekorelirane. Dakle, $\mathrm{Cov}(X,Y)=0$ općenito ne implicira da su X i Y nezavisne. To je zato što kovarijanca mjeri samo jedan oblik zavisnosti, tzv. "linearnu" zavisnost.

Zadaci za vježbu

Zadatak 6.8. Iz skupa $\{1,2,3\}$ dva puta nezavisno izaberemo jedan broj (moguće isti). Ako je X veći, a Y manji od ta dva broja, izračunajte koeficijent korelacije od X i Y.

Zadatak 6.9. Slučajni vektor (X,Y) ima distribuciju

$X\setminus Y$	$-a$	0	a
$-a$	0	1/4	0
0	1/4	0	1/4
a	0	1/4	0

za neki $a>0$.

(a) Jesu li slučajne varijable X i Y nezavisne?

(b) Jesu li slučajne varijable $X + Y$ i $X - Y$ nezavisne?

Zadatak 6.10. U vreći se nalaze 4 novčanice od 100 kuna i 6 novčanica od 10 kuna. Marko slučajno izvlači prvu novčanicu iz vreće i stavlja je u svoj novčanik. Označimo vrijednost prve novčanice sa X_1. Nakon toga Marko izvlači neku od preostalih novčanica iz vreće i stavlja je u svoj novčanik. Označimo vrijednost druge novčanice sa X_2.

(a) Odredite distribuciju slučajnog vektora (X_1, X_2).

(b) Odredite koeficijent korelacije slučajnih varijabli X_1 i X_2.

(c) Odredite očekivani iznos novca u Markovom novčaniku.

7 Neprekidne slučajne varijable

Definicija 7.1. Slučajna varijabla X na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ je **neprekidna slučajna varijabla** ako postoji funkcija $f : \mathbb{R} \to [0, \infty\rangle$ takva da je

$$\mathbb{P}(X \leq a) = \int_{-\infty}^{a} f(t)dt, \quad \forall a \in \mathbb{R}.$$

Funkciju f zovemo **funkcija gustoće** slučajne varijable X.

Napomena 7.2. Neka je X neprekidna slučajna varijabla s gustoćom f.

(a) Za sve $a, b \in \mathbb{R}$, $a \leq b$ vrijedi

$$\begin{aligned}\mathbb{P}(a < X \leq b) &= \mathbb{P}(\{X \leq b\} \setminus \{X \leq a\}) = \mathbb{P}(X \leq b) - \mathbb{P}(X \leq a) \\ &= \int_a^b f(t)dt\,.\end{aligned}$$

(b) Vrijedi

$$\mathbb{P}(X > a) = 1 - \mathbb{P}(X \leq a) = \int_a^{\infty} f(t)dt$$

te

$$\int_{-\infty}^{\infty} f(t)dt = \mathbb{P}(X \in \mathbb{R}) = 1\,.$$

(c) Razlika u odnosu na diskretne slučajne varijable je da za sve $a \in \mathbb{R}$ vrijedi

$$\mathbb{P}(X = a) = \int_a^a f(t)dt = 0\,.$$

Iz ovoga slijedi da je

$$\mathbb{P}(X \leq a) = \mathbb{P}(X < a) + \mathbb{P}(X = a) = \mathbb{P}(X < a)\,.$$

Na isti način zaključujemo da vrijedi

$$\mathbb{P}(a < X \leq b) = \mathbb{P}(a < X < b) = \mathbb{P}(a \leq X < b) = \mathbb{P}(a \leq X \leq b)\,,$$

tj. kod neprekidnih slučajnih varijabli nije bitno je li rub uključen ili ne.

(d) Ako je f jednaka nuli osim eventualno na intervalu $\langle a, b\rangle$ tada je

$$\mathbb{P}(a < X < b) = \int_a^b f(t)dt = \int_{-\infty}^{\infty} f(t)dt = 1\,.$$

Dakle, u tom slučaju X poprima vrijednosti samo iz intervala $\langle a, b\rangle$.

Napomena 7.3. Očekivanje neprekidne slučajne varijable X definira se kao

$$\mathbb{E}[X] := \int_{-\infty}^{\infty} tf(t)dt.$$

Općenito, ako je $g : \mathbb{R} \to \mathbb{R}$ funkcija, definiramo[6]

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(t)f(t)dt.$$

Posebno, varijancu definiramo kao i u diskretnom slučaju,

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

te lako slijedi da ponovno vrijedi $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$.

Zadatak 7.4. Neka je X neprekidna slučajna varijabla s gustoćom

$$f(t) = \begin{cases} c \cdot t^4, & 0 < t < 1, \\ 0, & \text{inače.} \end{cases}$$

Odredite vrijednost konstante c i izračunajte $\mathbb{P}(X > \frac{1}{2})$, $\mathbb{E}[X]$ i $\mathrm{Var}(X)$.

Rješenje. Znamo da je

$$1 = \int_{-\infty}^{+\infty} f(t)\,dt = \int_0^1 c \cdot t^4\,dt = c\left(\frac{t^5}{5}\right)\Big|_0^1 = c \cdot \frac{1}{5},$$

pa je $c = 5$. Zato je

$$\mathbb{P}\left(X > \frac{1}{2}\right) = \int_{\frac{1}{2}}^{+\infty} f(t)\,dt = \int_{\frac{1}{2}}^{1} 5t^4\,dt = \left(t^5\right)\Big|_{\frac{1}{2}}^{1} = 1 - \frac{1}{2^5} \approx 0.969\,,$$

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} tf(t)\,dt = \int_0^1 5t^5\,dt = \frac{5}{6}\,,$$

te

$$\begin{aligned}\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 &= \int_{-\infty}^{+\infty} t^2 f(t)\,dt - \left(\frac{5}{6}\right)^2 \\ &= \int_0^1 5t^6\,dt - \frac{25}{36} = \frac{5}{7} - \frac{25}{36} = \frac{2}{252} \approx 0.0198\,.\end{aligned}$$

□

[6]Zapravo, ako je g "izmjeriva" funkcija, tada je $g(X)$ slučajna varijabla (ne mora biti ni diskretna ni neprekidna) i vrijedi ova formula za računanje očekivanja od $g(X)$.

Definicija 7.5. Funkcija distribucije slučajne varijable X (diskretne ili neprekidne) je funkcija F definirana s

$$F(x) = \mathbb{P}(X \leq x), \quad \forall x \in \mathbb{R}.$$

Posebno, ako je X neprekidna slučajna varijabla s gustoćom f, onda joj je funkcija distribucije

$$F(x) = \int\limits_{-\infty}^{x} f(t)dt, \quad \forall x \in \mathbb{R}.$$

Napomena 7.6. Ako je F funkcija distribucije slučajne varijable X vrijedi

$$\mathbb{P}(a < X \leq b) = \mathbb{P}(X \leq b) - \mathbb{P}(X \leq a) = F(b) - F(a)$$

i

$$\mathbb{P}(X > x) = 1 - \mathbb{P}(X \leq x) = 1 - F(x).$$

Dakle, ako znamo F možemo računati i vjerojatnosti vezane uz X.

Zadatak 7.7. X ima neprekidnu **uniformnu** distribuciju na intervalu $\langle a, b\rangle$ ako joj je funkcija gustoće dana s

$$f(t) = \begin{cases} \frac{1}{b-a}, & a < t < b, \\ 0, & \text{inače}. \end{cases}$$

Pišemo $X \sim U(a, b)$. Odredite joj funkciju distribucije i izračunajte $\mathbb{E}\, X$.

Rješenje. Budući da je f nula svugdje osim na intervalu $\langle a, b\rangle$ imamo da je

$$F(x) = \int\limits_{-\infty}^{x} f(t)\,dt = \begin{cases} 0, & x \leq a, \\ \int_a^x \frac{1}{b-a}, & a < x < b, \\ \int_a^b \frac{1}{b-a}, & x \geq b. \end{cases}$$
$$= \begin{cases} 0, & x \leq a, \\ \frac{x-a}{b-a}, & a < x < b, \\ 1, & x \geq b. \end{cases}$$

Računamo i očekivanje

$$\mathbb{E}[X] = \int\limits_{-\infty}^{\infty} t\, f(t)\,dt = \int\limits_{a}^{b} \frac{t}{b-a}dt = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}.$$

□

Napomena 7.8. Ako je $X \sim U(a, b)$, njena gustoća f je konstantna funkcija na $\langle a, b\rangle$. Intuitivno, to znači da će X poprimiti bilo koju vrijednost iz $\langle a, b\rangle$ s jednakom vjerojatnosti. Zapravo, budući da za svaki $A := \langle c, d\rangle \subseteq \langle a, b\rangle := \Omega$ imamo

$$\mathbb{P}(X \in A) = \mathbb{P}(c < X < d) = F(d) - F(c)$$
$$= \frac{d-c}{b-a} = \frac{\lambda(A)}{\lambda(\Omega)},$$

izbor točke X točno odgovara "slučajnom izboru točke iz $\langle a, b\rangle$" iz poglavlja Geometrijska vjerojatnost.

U nastavku ćemo se baviti jednom od najvažnijih distribucija u vjerojatnosti – normalnom (ili Gaussovom) distribucijom. Njena važnost posljedica je tzv. centralnog graničnog teorema kojim ćemo se baviti u sljedećem potpoglavlju.

Definicija 7.9. Slučajna varijabla X ima **normalnu distribuciju** s parametrima $\mu \in \mathbb{R}$ i $\sigma^2 > 0$ ako joj je funkcija gustoće dana s

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

U tom slučaju pišemo $X \sim N(\mu, \sigma^2)$.

Napomena 7.10. (a) Ako je $X \sim N(\mu, \sigma^2)$, njena gustoća zadovoljava $f(t) > 0$ za sve $t \in \mathbb{R}$, tj. X može poprimiti bilo koju vrijednost iz $\mathbb{R}$.

(b) Ako je $X \sim N(\mu, \sigma^2)$, onda je $\mathbb{E}[X] = \mu$ i $\mathrm{Var}(X) = \sigma^2$. Broj $\sigma = \sqrt{\mathrm{Var}(X)}$ nazivamo standardna devijacija.

(c) Ako je $X \sim N(\mu, \sigma^2)$ tada je $Z := \frac{X-\mu}{\sigma} \sim N(0,1)$.

(d) Slučajnu varijablu $Z \sim N(0,1)$ nazivamo standardna normalna slučajna varijabla, a njenu funkciju distribucije označavamo s Φ, tj.

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt\,,\ x \in \mathbb{R}\,.$$

Budući da se Φ ne može eksplicitno izračunati, njene vrijednosti računaju se numerički, a mi ih čitamo iz tablice koju možete naći na webu kolegija. Uočimo, u ovom slučaju vrijedi $f(-t) = f(t)$ za sve $t \geq 0$ pa je

$$\begin{aligned}\mathbb{P}(Z \leq 0) = \Phi(0) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-\frac{t^2}{2}} dt = [\text{zamjena varijabli } t \to -t] \\ &= \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-\frac{t^2}{2}} dt = 1 - \Phi(0) = \mathbb{P}(Z > 0)\end{aligned}$$

pa iz $\mathbb{P}(Z \leq 0) + \mathbb{P}(Z > 0) = 1$ slijedi da je

$$\mathbb{P}(Z \leq 0) = \Phi(0) = \frac{1}{2} = \mathbb{P}(Z > 0)\ .$$

Općenitije, za sve $x \leq 0$ vrijedi

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} \int_{-x}^{\infty} e^{-\frac{t^2}{2}} dt = 1 - \Phi(-x)\,, \tag{5}$$

dakle dovoljno je znati vrijednosti $\Phi(x)$ za $x \geq 0$.

Primjer 7.11. Neka je $X \sim N(1, 4)$. Pitamo se koliko je $\mathbb{P}(1 < X < 3/2)$. Imamo dakle $\mu = 1$ i $\sigma^2 = 4$ pa je po prethodnoj napomeni $Z = \frac{X-\mu}{\sigma} = \frac{X-1}{2} \sim N(0,1)$. Sada računamo

$$\begin{aligned}\mathbb{P}(1 < X < \frac{3}{2}) &= \mathbb{P}\left(\frac{1-1}{2} < \frac{X-1}{2} < \frac{3/2-1}{2}\right) \\ &= \mathbb{P}(0 < Z < 1/4) \\ &= \Phi(1/4) - \Phi(0) = 0.5987 - 0.5 = 0.0987\,,\end{aligned}$$

pri čemu smo vrijednosti $\Phi(1/4)$ i $\Phi(0)$ iščitali iz tablice.

Zadatak 7.12. Vijek trajanja neke automobilske gume je normalno distribuiran s očekivanjem 34000 km i standardnom devijacijom od 4000km.

(a) Izračunajte vjerojatnost da guma traje više od 40000 km.

(b) Ako je guma prešla 30000 km, izračunajte vjerojatnost da će trajati još 10000 km.

Rješenje. Neka je X vijek trajanja automobilske gume koji po zadatku slijedi distribuciju $N(34000, 4000^2)$. Po prethodnoj napomeni znamo da je $Z := \frac{X-\mu}{\sigma} = \frac{X-34000}{4000} \sim N(0,1)$.

(a) Imamo

$$\mathbb{P}(X \geq 40000) = \mathbb{P}\left(Z \geq \frac{3}{2}\right) = 1 - \Phi(1.5) \approx 0.0668\,.$$

(b) Slično,

$$\begin{aligned}\mathbb{P}(X \geq 40000 | X \geq 30000) &= \frac{\mathbb{P}(X \geq 40000, X \geq 30000)}{\mathbb{P}(X \geq 30000)} = \frac{\mathbb{P}(X \geq 40000)}{\mathbb{P}(X \geq 30000)} \\ &= \frac{\mathbb{P}\left(Z \geq \frac{3}{2}\right)}{\mathbb{P}(Z \geq -1)} = \frac{1-\Phi(1.5)}{1-\Phi(-1)} = \frac{1-\Phi(1.5)}{\Phi(1)} \approx 0.0794\,,\end{aligned}$$

pri čemu smo u predzadnjoj jednakosti iskoristili da je $\Phi(-1) = 1 - \Phi(1)$ (vidi (5)).

$\square$

Zadatak 7.13. Pretpostavimo da je vrijeme putovanja nekog studenta od kuće do fakulteta približno normalno distribuirano s očekivanjem 40 minuta i standardnom devijacijom od 7 minuta. Student želi stići na predavanje koje počinje u 12:15 sati.

(a) Ako je student krenuo od kuće u 11:40, izračunajte vjerojatnost da stići na vrijeme na predavanje.

(b) Kada bi najkasnije student trebao krenuti od kuće da s vjerojatnošću od barem 0.95 bude siguran da će stići na vrijeme na predavanje?

Rješenje. Neka je X vrijeme putovanja u minutama za koje po zadatku ima distribuciju $N(40, 49)$.

(a) Tražena vjerojatnost jednaka je

$$\mathbb{P}(X \leq 35) = \mathbb{P}\left(\frac{X-40}{7} \leq -\frac{5}{7}\right) = \Phi(-\tfrac{5}{7}) = 1 - \Phi(\tfrac{5}{7}) \approx 0.2389,$$

pri čemu smo zaokružili $\Phi(\frac{5}{7}) \approx \Phi(0.71) = 0.7611$.

(b) Neka je a vrijeme (minutama) od polaska do početka predavanja. Dakle, student će stići na vrijeme akko $X \leq a$. Tražimo najmanji a takav da je

$$0.95 \leq \mathbb{P}(X \leq a) = \mathbb{P}\left(\frac{X-40}{7} \leq \frac{a-40}{7}\right) = \Phi\left(\frac{a-40}{7}\right).$$

Budući da je Φ rastuća funkcija, a $\Phi(1.65) = 0.9505 > 0.95$ i $\Phi(1.64) < 0.95$, zaključujemo da je a najmanji broj takav da je $\frac{a-40}{7} \geq 1.65$, odnosno $a \geq 51.55$. Dakle $a = 52$, tj. student bi od kuće trebao krenuti najkasnije u 11:23.

$\square$

Zadaci za vježbu

Zadatak 7.14. Neka je $X \sim N(70, 4)$. Izračunajte $\mathbb{P}((X-68)^2 \geq 9)$. (Uputa: $\{(X-68)^2 \geq 9\} = \{|X-68| \geq 3\} = \{|X-68| < 3\}^c$.)

Zadatak 7.15. Na nekom ispitu bodovi slučajno odabranog učenika su normalno distribuirani s očekivanjem 76 i standardnom devijacijom 15. Odredite:

(a) Najmanji broj bodova za ocjenu 5 takav da je vjerojatnost da slučajno odabrani učenik dobije peticu najviše 0.15.

(b) Najveći broj bodova za ocjenu 2 takav da je vjerojatnost da slučajno odabrani učenik ne položi ispit najviše 0.1.

Zadatak 7.16. Neka je $X \sim N(\mu, \sigma^2)$. Izračunajte

(a) $\mathbb{P}(\mu - 3\sigma \leq X \leq \mu + 3\sigma)$,

(b) $\mathbb{P}(X^2 - 2\mu X \geq \sigma^2 - \mu^2)$.

Zadatak 7.17. Pretpostavimo da je duljina telefonskog poziva X (u minutama) neprekidna slučajna varijabla s funkcijom gustoće

$$f(x) = \begin{cases} ce^{-\frac{x}{10}}, & x \geq 0 \\ 0, & x < 0 \end{cases},$$

gdje je c neka konstanta.

(a) Odredite c.

(b) Odredite funkciju distribucije od X.

(c) Ako je netko taman prije vas došao u telefonsku govornicu i započeo razgovor, izračunajte vjerojatnost da ćete čekati između 10 i 20 minuta da dođete na red.

(d) Odredite $\mathbb{E}[X]$.

(Napomena: Slučajna varijabla s ovom funkcijom distribucije se zove **eksponencijalna slučajna varijabla** s parametrom 1/10.)

7.1 Centralni granični teorem

Teorem 7.18 (de Moivre-Laplace CGT)**.** Za svaki $n \geq 1$, neka je $X_n \sim B(n,p)$ pri čemu je $p \in (0,1)$. Ako je $Z \sim N(0,1)$, tada za svaki $x \in \mathbb{R}$ vrijedi

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{X_n - np}{\sqrt{npq}} \leq x\right) = \mathbb{P}\left(Z \leq x\right) = \Phi(x)\,.$$

Uočite da je $np = \mathbb{E}\left[X_n\right]$, a $\sqrt{npq} = \sqrt{\mathrm{Var}\left(X_n\right)}$.

Napomena 7.19. Prethodni teorem najčešće koristimo u sljedećem obliku. Za $b \in \mathbb{R}$ i veliki n

$$\mathbb{P}\left(X_n \leq b\right) = \mathbb{P}\left(\frac{X_n - np}{\sqrt{npq}} \leq \frac{b - np}{\sqrt{npq}}\right) \approx \Phi\left(\frac{b-np}{\sqrt{npq}}\right)\,.$$

Zadatak 7.20. Ako je za svako novorođenče jednako vjerojatno da bude dječak ili djevojčica, približno izračunajte vjerojatnost da među 1000 novorođenčadi bude najviše 490 djevojčica.

Rješenje. Neka X označava broj rođenih djevojčica. Tada je očito $X \sim B(1000, \frac{1}{2})$, pa primjenom CGT-a slijedi da je

$$\begin{aligned}\mathbb{P}\left(X \leq 490\right) &= \mathbb{P}\left(\tfrac{X - 1000\cdot\frac{1}{2}}{\sqrt{1000\cdot\frac{1}{2}\cdot\frac{1}{2}}} \leq \tfrac{490 - 1000\cdot\frac{1}{2}}{\sqrt{1000\cdot\frac{1}{2}\cdot\frac{1}{2}}}\right)\\ &\approx \Phi(-0.63) = 1 - \Phi(0.63) = 1 - 0.7357 = 0.2643\,.\end{aligned}$$

Napomenimo da je točna vrijednost

$$\mathbb{P}\left(X \leq 490\right) = 0.2739864\,,$$

dakle pomoću CGT-a smo na jednostavan način dobili dosta dobru aproksimaciju. U suprotnom bi trebali računati

$$\mathbb{P}\left(X \leq 490\right) = \sum_{k=0}^{490} \mathbb{P}\left(X = k\right)\,.$$

□

Napomena 7.21. Za $X \sim B(n,p)$ i veliki n i $a \leq b$ također vrijedi

$$\mathbb{P}\left(X \geq a\right) \approx \mathbb{P}\left(Z \geq \tfrac{a-np}{\sqrt{npq}}\right) = 1 - \Phi(\tfrac{a-np}{\sqrt{npq}})$$

i

$$\mathbb{P}\left(a \leq X \leq b\right) \approx \mathbb{P}\left(\tfrac{a-np}{\sqrt{npq}} \leq Z \leq \tfrac{b-np}{\sqrt{npq}}\right) = \Phi(\tfrac{b-np}{\sqrt{npq}}) - \Phi(\tfrac{a-np}{\sqrt{npq}})\,.$$

Zadatak 7.22. Na ispitu je 40 zadataka i za svaki su ponuđena četiri odgovora od kojih je samo jedan točan. Za točno zaokružen odgovor dobije se 15 bodova, a za pogrešno zaokružen gubi se 5 bodova. Približno izračunajte vjerojatnost da student koji slučajnim odabirom bira odgovore ostvari barem 120 bodova dovoljnih za prolaznu ocjenu?

Rješenje. Neka je X broj točno riješenih zadataka. Tada je $X \sim B(40, \frac{1}{4})$, a broj ostvarenih bodova je

$$Y = 15X - 5(40 - X) = 20X - 200\,.$$

Primjenom CGT-a dobivamo da je tražena vjerojatnost

$$\begin{aligned}\mathbb{P}\left(Y \geq 120\right) &= \mathbb{P}\left(20X - 200 \geq 120\right)\\ &= \mathbb{P}\left(X \geq 16\right) \approx 1 - \Phi\left(\frac{16 - 40\cdot\frac{1}{4}}{\sqrt{40\cdot\frac{1}{4}\cdot\frac{3}{4}}}\right)\\ &= 1 - \Phi(2.19) = 1 - 0.9857 = 0.0143\,.\end{aligned}$$

□

Zadatak 7.23. Prosjak dobije novčić od prolaznika s vjerojatnošću 0.05. Koliko najmanje prolaznika treba proći ulicom da bi prosjak skupio barem 150 novčića s vjerojatnošću od barem 0.95?

Rješenje. Ako je n broj prolaznika, a X broj novčića koje prosjak dobije, tada je $X \sim B(n, 0.05)$. Tražimo najmanji n takav da vrijedi

$$\mathbb{P}\left(X \geq 150\right) \geq 0.95\,.$$

Koristeći CGT imamo

$$\mathbb{P}\left(X \geq 150\right) \approx 1 - \Phi\left(\tfrac{150 - n\cdot 0.05}{\sqrt{n\cdot 0.05\cdot 0.95}}\right),$$

pa dakle mora vrijediti (uz množenje brojnika i nazivnika s 20)

$$\Phi\left(\tfrac{3000 - n}{\sqrt{19n}}\right) \leq 1 - 0.95 = 0.05\,.$$

Budući da je $\Phi(-1.65) \approx 0.05$, mora dakle vrijediti

$$\tfrac{3000 - n}{\sqrt{19n}} \leq -1.65\,.$$

Rješavanjem ove kvadratne nejednadžbe (u terminima $\sqrt{n}$), dobivamo da je $\sqrt{n} \geq 58.5$, odnosno $n \geq 3420.7$. Zaključujemo ulicom mora proći barem 3421 prolaznik. □

Zadatak 7.24. Kazalište koje prima 529 gledatelja ima 2 ulaza i pored svakog se nalazi garderoba sa k vješalica. Ako svaki posjetitelj s jednakom vjerojatnosti bira bilo koji od ulaza te zatim ostavlja kaput u garderobi, koliki je najmanji k takav da s vjerojatnošću od barem 0.95 bude mjesta za kapute svih 529 posjetitelja?

Rješenje. Neka je X broj posjetitelja koji su ušli na prvi ulaz, dakle $X \sim B(529, \frac{1}{2})$. Budući da je onda $529 - X$ broj posjetitelja koji su ušli na drugi ulaz, mjesta će biti za sve kapute ako i samo ako je $X \leq k$ i $529 - X \leq k$, odnosno $529 - k \leq X \leq k$.

Dakle, tražimo najmanji k takav da je

$$\mathbb{P}\left(529 - k \leq X \leq k\right) \geq 0.95\,.$$

Koristeći CGT dobivamo

$$\begin{aligned}\mathbb{P}\left(529 - k \leq X \leq k\right) &\approx \Phi\left(\frac{k - \frac{529}{2}}{\sqrt{\frac{529}{4}}}\right) - \Phi\left(\frac{529 - k - \frac{529}{2}}{\sqrt{\frac{529}{4}}}\right) \\ &= \Phi\left(\frac{k - \frac{529}{2}}{\sqrt{\frac{529}{4}}}\right) - \Phi\left(\frac{\frac{529}{2} - k}{\sqrt{\frac{529}{4}}}\right)\,.\end{aligned}$$

Uočimo, budući da je ukupan broj vješalica $2k$, sigurno k mora biti barem $529/2$, tj. $592/2 - k \leq 0$ pa, koristeći $\Phi(x) = 1 - \Phi(-x)$ za $x \leq 0$,

$$\mathbb{P}\left(529 - k \leq X \leq k\right) \approx 2\Phi\left(\frac{k - \frac{529}{2}}{\sqrt{\frac{529}{4}}}\right) - 1\,.$$

Dakle, mora vrijediti

$$\Phi\left(\frac{k - \frac{529}{2}}{\sqrt{\frac{529}{4}}}\right) \geq 0.975\,,$$

pa budući da je $\Phi(1.96) = 0.975$, mora vrijediti

$$\frac{k - \frac{529}{2}}{\sqrt{\frac{529}{4}}} \geq 1.96$$

odnosno $k \geq 287.04$. Zaključujemo da svaka garderoba mora imati najmanje 288 vješalica. $\square$

Zadaci za vježbu

Zadatak 7.25. Rulet ima 18 crvenih polja, 18 crnih polja i jednu nulu. Pretpostavite da u svakoj igri igrate na crveno. Ako kuglica padne na crveno, onda dobijete 1 kn, a inače gubite 1 kn. Približno odredite vjerojatnost da ste nakon 1000 igara na dobitku.

Zadatak 7.26. 1920. godine su u Chicagu bila dva vlaka koja su se borila za 1000 putnika koji su u isti sat kretali iz Chicaga u Los Angeles. Pretpostavimo da putnici s jednakom vjerojatnošću odabiru svaki od vlakova. Koliko minimalno sjedala mora imati svaki vlak da bi s vjerojatnošću 0.99 bili sigurni da će svaki putnik imati svoje sjedalo?

8 Statistika i statistički testovi

Do sada smo pretpostavljali da poznajemo distribuciju slučajne varijable i ovisno o tome donosili zaključke o vjerojatnostima pojedinih ishoda. U praksi, često nećemo unaprijed znati distribuciju iz koje naši podaci dolaze, već ćemo iz podataka pokušati rekonstruirati distribucije iz kojih dolaze, njihova svojstva i donositi neke zaključke.

Definicija 8.1. *Slučajan uzorak* duljine $n \in \mathbb{N}$ za slučajnu varijablu X je niz nezavisnih i jednako distribuiranih slučajnih varijabli $X_1, X_2, \ldots, X_n$ koji imaju istu distribuciju kao i X.

Uzorak duljine $n \in \mathbb{N}$ za slučajnu varijablu X je jedna realizacija slučajnog uzorka i označavat ćemo ga s $x_1, x_2, ..., x_n$.

Primjetite da je slučajan uzorak niz slučajnih varijabli, a uzorak niz brojeva. Najčešće ćemo imati uzorak i htjeti nešto zaključiti o distribuciji iz koje dolazi (odnosno o distribuciji od X).

Pretpostavimo da je prosječna visina svih stanovnika Hrvatske 173cm. Ovaj podatak daje nam srednju vrijednost visine Hrvata. Ponekad je srednja vrijednost dovoljna informacija da zaključimo ono što nas zanima o našem setu podataka. Postoji više načina kako srednju vrijednost možemo definirati, a najčešće se koristi aritmetička sredina.[7]

Definicija 8.2. Za uzorak $x_1, x_2, \ldots, x_n$ aritmetičku sredinu označavamo s $\overline{x}_n$ i računamo po formuli $\overline{x}_n = \frac{x_1+\cdots+x_n}{n}$. Za slučajni uzorak $X_1, X_2, \ldots, X_n$ aritmetičku sredinu označavamo s $\overline{X}_n$ i jednaka je $\overline{X}_n = \frac{X_1+\cdots+X_n}{n}$.

Osim sredine, bitan podatak nam je i raspršenost uzorka koju mjerimo uzoračkom varijancom.

Definicija 8.3. Za uzorak $x_1, x_2, ..., x_n$ uzoračku varijancu definiramo kao

$$s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x}_n)^2,$$

Uzoračku varijancu slučajnog uzorka definiramo analogno, zamjenom uzoračkih vrijednosti x_i sa slučajnim varijablama X_i. Oznaka je S_n^2

Napomena 8.4. Za uzoračku varijancu vrijedi

$$s_n^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\overline{x}_n^2\right).$$

Primjer 8.5. Odredite aritmetičku sredinu i uzoračku varijancu za uzorke:

(a) $1, 1, 2, 2, 2, 3, 3, 3, 3, 4, 4$

(b) $-100, 1, 2, 2, 2, 2, 3, 3, 3, 10, 100$

Rješenje.

[7] druge česte mjere srednje vrijednosti su **medijan** i **mod**

(a) $\overline{x}_n = 2.5, s_n^2 = 1.166667$

(b) $\overline{x}_n = 2.6, s_n^2 = 2230.267$

$\square$

Ispostavlja se da aritmetičkom sredinom $\overline{X}_n$ i uzoračkom varijancom S_n^2 možemo dobro procijeniti očekivanje i varijancu slučajne varijable X iz čije distribucije uzorak dolazi. Što znači dobro procijeniti? Ako želimo procijeniti neki parametar distribucije od X-a koji ćemo označiti sa θ, kažemo da je funkcija slučajnog uzorka $X_1, ..., X_n$ od X, $f(X_1, ..., X_n)$ **nepristrani procjenitelj** za θ ako vrijedi $\mathbb{E}\left[f(X_1, ..., X_n)\right] = \theta$. Biti nepristrani procjenitelj je dobro svojstvo ako želimo procijeniti parametar θ jer je očekivana (prosječna) vrijednost te funkcije slučajnog uzorka baš θ.

Zadatak 8.6. Neka je $X_1, X_2, \ldots, X_n$ slučajni uzorak iz populacije s konačnim očekivanjem μ i varijancom σ^2. Pokažite da je

(a) aritmetička sredina $\overline{X}_n = \frac{X_1 + \cdots + X_n}{n}$ nepristrani procjenitelj za μ,

(b) uzoračka varijanca $S_n^2 = \frac{1}{n-1} \sum\limits_{i=1}^{n} \left(X_i - \overline{X}_n\right)^2$ nepristrani procjenitelj za σ^2.

Rješenje. Ponovimo još jednom pojmove iz teksta zadatka. Slučajni uzorak su slučajne varijable $X_1, X_2, \ldots, X_n$, što znači da su to nezavisne slučajne varijable s istom distribucijom. Napomenimo samo da (bez obzira na oznake) ta distribucija ne mora biti normalna distribucija, već je samo zadano da ima konačno očekivanje i varijancu. Trebamo odrediti očekivanje statistika $\overline{X}_n$ i S_n^2.

(a) Zbog linearnosti očekivanja i jednake distribuiranosti slučajnog uzorka imamo

$$\mathbb{E}\left[\overline{X}_n\right] = \frac{\mathbb{E}\left[X_1\right] + \cdots + \mathbb{E}\left[X_n\right]}{n} = \frac{n\mu}{n} = \mu$$

iz čega zaključujemo da je $\overline{X}_n$ nepristrani procjenitelj za μ.

(b) Slično, iz

$$
\begin{aligned}
\mathbb{E}\,[S_n^2] &= \frac{1}{n-1}\mathbb{E}\,[\sum_{i=1}^{n}\left(X_i-\overline{X}_n\right)^2] = \frac{1}{n-1}\mathbb{E}\,[\sum_{i=1}^{n}X_i^2 - 2\overline{X}_n\sum_{i=1}^{n}X_i + n\overline{X}_n^2] = \\
&= \frac{1}{n-1}\mathbb{E}\,[\sum_{i=1}^{n}X_i^2 - 2n\overline{X}_n^2 + n\overline{X}_n^2] = \frac{1}{n-1}\mathbb{E}\,[\sum_{i=1}^{n}X_i^2 - n\overline{X}_n^2] = \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\mathbb{E}\,[X_i^2] - \frac{1}{n(n-1)}\mathbb{E}\,[\left(\sum_{i=1}^{n}X_i\right)^2] = \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\left(\sigma^2+\mu^2\right) - \frac{1}{n(n-1)}\mathrm{Var}\left(\sum_{i=1}^{n}X_i\right) - \frac{1}{n(n-1)}\mathbb{E}\,[\sum_{i=1}^{n}X_i]^2 = \\
&= \frac{n}{n-1}\left(\sigma^2+\mu^2\right) - \frac{1}{n(n-1)}\sum_{i=1}^{n}\sigma^2 - \frac{1}{n(n-1)}\left(\sum_{i=1}^{n}\mu\right)^2 = \\
&= \left(\frac{n}{n-1} - \frac{n}{n(n-1)}\right)\sigma^2 + \left(\frac{n}{n-1} - \frac{n^2}{n(n-1)}\right)\mu^2 = \sigma^2
\end{aligned}
$$

zaključujemo da je S_n^2 nepristrani procjenitelj za σ^2.

$\square$

Pretpostavimo kako želimo ispitati je li prosječan student PMF-a viši od prosječne ljudske visine u Hrvatskoj. Neka X predstavlja visinu slučajno odabranog studenta na PMF-u i pretpostavimo da je X približno normalno distribuirana s nepoznatim očekivanjem μ i standardnom devijacijom $\sigma = 10$ cm (tj. varijanca je $\sigma^2 = 100$). Iz nekog razloga sumnjamo da je prosječna visina PMF-ovaca veća od prosjeka Hrvatske te bismo htjeli "testirati" tu *hipotezu*. Budući da zasad nemamo nikakvih realnih/opipljivih/stvarnih **dokaza** da je visina PMF-ovaca veća od prosjeka (imamo samo slutnju), ono što pretpostavljamo je da vrijedi tzv. **nulta hipoteza**

$$H_0 : \mu = 173 \quad \text{(prosječna visina PMF-ovca je 173cm).}$$

Hipotezu koju želimo dobiti, tj. ono što naslućujemo, jest **alternativna hipoteza**

$$H_1 : \mu > 173 \quad \text{(prosječna visina PMF-ovca veća je od 173cm).}$$

Budući da ne možemo izmjeriti visinu svih studenata (npr. zahtijeva previše vremena/novaca ili je jednostavno nemoguće), izmjerimo visine npr. 16 slučajno odabranih studenata. Ako te visine označimo s $X_1, X_2, \ldots, X_{16}$, to je točno slučajan uzorak duljine 16 za slučajnu varijablu X. Pretpostavimo da smo izmjerili prosječnu visinu od 175 cm, tj. $\overline{x}_n = 175$ je **ishod** (ili *realizacija*) slučajne varijable

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

Budući da je to više od 173 cm možda bi mogli doći u napast odmah zaključiti da je naša sumnja bila točna. Ipak, mi smo mjerili visine samo 16 slučajno odabranih studenata pa se zapravo moramo pitati jesmo li jednostavno kao plod slučajnosti izabrali 16 studenata malo viših od prosjeka populacije, ili je realizacija zaista plod veće visine studenata PMF-a? Malo preciznije – želimo znati koja je vjerojatnost da smo na 16 slučajno odabranih studenta dobili prosjek od barem 175 cm iako cijela populacija ima prosjek 173 cm? U tome će nam pomoći sljedeća napomena.

Napomena 8.7. Neka je $X_1, \dots, X_n$ slučajan uzorak za $X \sim N(\mu, \sigma^2)$ pri čemu su parametri $\mu \in \mathbb{R}, \sigma > 0$ proizvoljni te neka je

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

Tada vrijedi

$$Z := \frac{\overline{X}_n - \mu}{\sigma} \sqrt{n} \sim N(0, 1).$$

Vratimo se na PMF-ovce. Ako nije došlo do povećanja visine, tj. vrijedi nulta hipoteza H_0, tada je $X \sim N(173, 10^2)$ pa je po prethodnoj napomeni

$$Z := \frac{\overline{X}_n - 173}{10} \sqrt{16} \sim N(0, 1)$$

te vrijedi

$$\begin{aligned} \mathbb{P}\,(\overline{X}_n \geq 175) &= \mathbb{P}\,(Z \geq \overbrace{\tfrac{175-173}{10}\sqrt{16}}^{=0.8}) \\ &= 1 - \Phi(0.8) = 0.2119 \,. \end{aligned}$$

Dakle, i u slučaju da je prosječna visina populacije ostala 173, vjerojatnost da prosječna visina 16 slučajno odabranih studenata bude barem 175 nije baš mala pa nismo skloni odbaciti H_0. Da smo izmjerili prosječnu visinu $\overline{x} = 180$ cm, gornja vjerojatnost bila bi $\mathbb{P}\,(\overline{X} \geq 180) = 0.0228$ te bismo u tom slučaju bili skloni odbaciti hipotezu H_0 u korist H_1, tj. u tom slučaju bi nam prikupljeni uzorak bio "dovoljno dobar" dokaz da kažemo da H_0 ne vrijedi već da H_1 vrijedi.

Da bismo odredili što je "dovoljno" dobar dokaz, tj. kada trebamo donijeti odluku odbacujemo li nultu hipotezu ili ne, prvo se moramo odlučiti za **razinu značajnosti** $\alpha \in \langle 0, 1 \rangle$ uz koju donosimo odluku – tipično $\alpha = 0.05, 0.025, 0.01$. Što je α manji, prilikom odbacivanja nulte hipoteze bit ćemo sigurniji da nismo napravili pogrešku (tzv. **pogreška prve vrste** - odbacili smo H_0 mada je H_0 zaista točna). Kada smo odabrali razinu značajnosti α, odabiremo $z_\alpha \in \mathbb{R}$ takav da je $\Phi(z_\alpha) = 1 - \alpha$, tj. kada vrijedi H_0 imamo $Z \sim N(0, 1)$ te

$$\mathbb{P}\,(Z \geq z_\alpha) = \mathbb{P}\left(\frac{\overline{X}_n - 173}{10} \sqrt{25} \geq z_\alpha \right) = 1 - \Phi(z_\alpha) = \alpha$$

što nekad kraće označavamo s

$$\mathbb{P}\,(Z \geq z_\alpha \,|\, H_0) = \alpha \,.$$

Sada, ako je z realizacija tzv. **testne statistike** Z i vrijedi $z \geq z_\alpha$, kažemo da **na razini značajnosti** α **odbacujemo** H_0 **u korist** H_1, tj. imamo dovoljno dobar dokaz da odbacimo H_0, a u suprotnom kažemo da **na razini značajnosti** α **ne odbacujemo** H_0 **u korist** H_1, tj.

nemamo dovoljno dobar dokaz da odbacimo H_0. (N. B. u zadnjem slučaju nikada ne kažemo da prihvaćamo H_0!)

U našem primjeru je $z = 0.8$ i pretpostavimo da smo se odlučili za razinu značajnosti od 5%, tj. $\alpha = 0.05$ – iz tablice čitamo da je $z_\alpha = 1.65$. Budući da je $z = 0.8 < 1.65 = z_\alpha$, na razini značajnosti od 5% ne odbacujemo H_0 u korist H_1, tj. na razini značajnosti od 5% ne možemo tvrditi da PMF-ovci imaju veću prosječnu visinu od prosjeka Hrvatske.

Napomenimo da odabir razine značajnosti α ovisi o tipu problema, npr. ponekad u psihološkim istraživanjima razina značajnosti od 5% je i više nego dobra, dok fizičari u CERN-u ponekad zahtijevaju razinu značajnosti od 0.00001%.

Test koji smo opisali naziva se **Z-test** pa ponovimo kratko pretpostavke i korake ovog testa.

Primjer 8.8 (Z-test). Neka je $X \sim N(\mu, \sigma^2)$ gdje nam je μ nepoznati, a σ^2 poznati parametar, te neka je $\mu_0 \in \mathbb{R}$ fiksan. Nadalje, neka je $\alpha \in \langle 0, 1\rangle$ razina značajnosti.

(a) Pretpostavimo da testiramo sljedeće hipoteze:

$$\begin{aligned} H_0 &: \mu = \mu_0 \\ H_1 &: \mu > \mu_0 . \end{aligned}$$

H_0 nazivamo **nul-hipoteza**, a H_1 nazivamo **alternativa hipoteza**. Hipoteza H_1 dakle predstavlja tvrdnju koju želimo provjeriti i bitno ju je dobro postaviti!

Neka je $X_1, \ldots, X_n$ slučajan uzorak za X. Ako vrijedi H_0, onda po prethodnoj napomeni vrijedi

$$Z := \frac{\overline{X}_n - \mu_0}{\sigma}\sqrt{n} \sim N(0,1) .$$

Slučajnu varijablu Z nazivamo **testna statistika**. Definiramo **kritično područje**

$$C := [z_\alpha, \infty\rangle ,$$

gdje je z_α takav da je $\Phi(z_\alpha) = 1 - \alpha$, tj.

$$\mathbb{P}\left(Z \in C \,|\, H_0\right) = \mathbb{P}\left(Z \geq z_\alpha \,|\, H_0\right) = 1 - \Phi(z_\alpha) = \alpha .$$

Neka je z ishod slučajne varijable Z. Ako je $z \in C$, tj. $z \geq z_\alpha$, onda kažemo da **na razini značajnosti α odbacujemo H_0 u korist H_1**, a ako $z \notin C$, tj. $z < z_\alpha$, onda kažemo da **ne odbacujemo H_0 u korist H_1**.

(b) Pretpostavimo da testiramo sljedeće hipoteze:

$$\begin{aligned} H_0 &: \mu = \mu_0 \\ H_1 &: \mu < \mu_0 . \end{aligned}$$

Sve je isto kao i u (a) dijelu samo što je $C := \langle -\infty, -z_\alpha]$. Uočimo da je

$$\begin{aligned} \mathbb{P}\left(Z \in C \,|\, H_0\right) &= \mathbb{P}\left(Z \leq -z_\alpha \,|\, H_0\right) = \Phi(-z_\alpha) \\ &= 1 - \Phi(z_\alpha) = \alpha . \end{aligned}$$

Dakle, ako je $z \in C$, odbacujemo H_0 u korist H_1, a u suprotnom ne odbacujemo H_0 u korist H_1.

(c) Pretpostavimo da testiramo sljedeće hipoteze:

$$\begin{aligned} H_0 &: \mu = \mu_0 \\ H_1 &: \mu \neq \mu_0. \end{aligned}$$

Ovdje je $C := \langle -\infty, -z_{\frac{\alpha}{2}}] \cup [z_{\frac{\alpha}{2}}, \infty\rangle$. Vidimo

$$\begin{aligned} \mathbb{P}(Z \in C \,|\, H_0) &= \mathbb{P}(Z \leq -z_{\frac{\alpha}{2}} \,|\, H_0) + \mathbb{P}(Z \geq z_{\frac{\alpha}{2}} \,|\, H_0) \\ &= \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha\,. \end{aligned}$$

Napomena 8.9. Kritična područja C u prethodnom primjeru birali smo tako da vrijedi $\mathbb{P}(Z \in C \,|\, H_0) = \alpha$ te da, u slučaju da vrijedi H_1, s velikom vjerojatnosti Z "upada" u C, tj. da u tom slučaju s velikom vjerojatnosti test donese ispravnu odluku i odbaci nul-hipotezu.

Zadatak 8.10. Neki proizvođač proizvodi sajle čija je izdržljivost u prosjeku jednaka 1800 kg uz standardnu devijaciju od 100 kg. Nedavno je proizvođač uveo novu tehniku proizvodnje i tvrdi da se na taj način mogu dobiti sajle veće izdržljivosti. Odabran je slučajni uzorak od 50 sajli proizvedenih novom tehnikom i izračunata je prosječna izdržljivost od 1850 kg. Uz pretpostavku da je izdržljivost sajli normalno distribuirana sa standardnom devijacijom 100 kg, može li se na razini značajnosti od 1% zaključiti da se novom tehnikom mogu dobiti izdržljivije sajle?

Rješenje. Ako je X slučajna varijabla koja predstavlja izdržljivost novih sajli, znamo da je $X \sim N(\mu, 100^2)$ gdje nam je μ nepoznati parametar. Želimo testirati sljedeće hipoteze

$$\begin{aligned} H_0 &: \mu = 1800 \quad \text{(nema promjene u izdržljivosti)} \\ H_1 &: \mu > 1800 \quad \text{(izdržljivost se povećala)}, \end{aligned}$$

uz razinu značajnosti od $\alpha = 0.01$. Koristimo Z-test (uz $\mu_0 = 1800$).

Testna statistika je $Z := \frac{\overline{X}_n - \mu_0}{\sigma}\sqrt{n}$ uz $\sigma = 100$ i $n = 50$. Kritično područje je $C = [z_{0.01}, \infty\rangle$ pri čemu iz tablice za normalnu razdiobu odredimo $z_{0.01}$ takav da je $\Phi(z_{0.01}) = 1 - 0.01 = 0.99$ – dakle $z_{0.01} = 2.33$.

Ishod od $\overline{X}_n$ je $\overline{x}_n = 1850$ pa je ishod testne statistike Z jednak

$$z = \frac{\overline{x}_n - \mu_0}{\sigma}\sqrt{n} = \frac{1850 - 1800}{100}\sqrt{50} = 3.536 > 2.33\,.$$

Dakle, $z \geq z_\alpha$ pa odbacujemo H_0 u korist H_1 na razini značajnosti $\alpha = 0.01$, tj. na razini značajnosti od 1% podaci upućuju na to da su nove sajle izdržljivije. □

Zadatak 8.11. Mjerenjem mase 20 istovrsnih čokolada dobiveni su sljedeći rezultati u gramima:

97	99	98	96	98	101	98	95	97	99
98	96	97	98	98	100	99	97	101	98

Pretpostavimo da se mase čokolada podvrgavaju normalnoj razdiobi sa standardnom devijacijom 1.55. Ako na omotu čokolade piše da je njena masa 100 g, možemo li na razini značajnosti od 5% zaključiti da proizvođač vara?

Rješenje. Neka slučajna varijabla X predstavlja masu slučajno odabrane čokoladice, dakle $X \sim N(\mu, 1.55^2)$ gdje je μ nepoznati parametar. Želimo testirati sljedeće hipoteze

$$\begin{aligned} H_0 &: \mu = 100 \\ H_1 &: \mu < 100, \end{aligned}$$

uz $\alpha = 0.05$. Koristimo Z-test (uz $\mu_0 = 100$).

Testna statistika je $Z := \frac{\overline{X}_n - \mu_0}{\sigma}\sqrt{n}$ uz $n = 20$ i $\sigma = 1.55$, a kritično područje je $C = \langle -\infty, -z_{0.05}]$ uz $z_{0.05} = 1.65$.

Na temelju podataka računamo

$$\overline{x}_n = \frac{1}{20}(97 + 99 + 98 + \cdots + 101 + 98) = 98\,.$$

Dakle,

$$z = \frac{\overline{x}_n - \mu_0}{\sigma}\sqrt{n} = \frac{98 - 100}{1.55}\sqrt{20} = -5.7 < -1.65 = -z_{0.05}$$

pa odbacujemo H_0 u korist H_1 na razini značajnosti od 5%, tj. na razini značajnosti od 5% zaključujemo da proizvođač vara. $\square$

Primjer 8.12 (Asimptotski Z-test). Neka je

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

za $p \in \langle 0, 1\rangle$ i neka je $X_1, \ldots, X_n$ uzorak za X. Uočimo da je $p = \mathbb{E}[X]$.

Ako je

$$Z := \frac{\overline{X}_n - p}{\sqrt{p(1-p)}}\sqrt{n}$$

imamo

$$Z = \frac{\sum_{i=1}^n X_i - np}{\sqrt{npq}}\,,$$

pa budući da je $\sum_{i=1}^n X_i \sim B(n,p)$, iz centralnog graničnog teorema slijedi da je

$$\mathbb{P}(Z \geq x) \approx 1 - \Phi(x), \quad \text{za velike } n,$$

tj. kažemo da je $Z \sim AN(0,1)$ (asimptotski normalna). Uz pomoć ovog rezultata možemo testirati hipoteze

$$\begin{aligned} H_0 &: p = p_0 \\ H_1 &: p > p_0 \quad (\text{ili } p < p_0 \text{ ili } p \neq p_0). \end{aligned}$$

Za razinu značajnosti α i veliki n gledamo testnu statistiku

$$Z = \frac{\overline{X}_n - p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n}\,,$$

a za kritično područje uzimamo $C = [z_\alpha, \infty\rangle$ (ili $C = \langle -\infty, -z_\alpha]$ ili $C = \langle -\infty, -z_{\frac{\alpha}{2}}] \cup [z_{\frac{\alpha}{2}}, \infty\rangle$) , dakle isto kao u pravom Z-testu uz malo drugačiju testnu statistiku.

Zadatak 8.13. Proizvođač tvrdi da njegove pošiljke sadrže najviše 7% defektnih proizvoda. Uzet je slučajni uzorak od 200 proizvoda iz jedne velike pošiljke i ustanovljeno je da je u njemu 9% defektnih prozivoda. Možete li na razini značajnosti od $\alpha = 0.05$ zaključiti da nas proizvođač vara?

Rješenje. Promatramo slučajnu varijablu

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

koja označava je li slučajno odabrani proizvod defektan ($X = 1$) ili ne ($X = 0$) pri čemu je $p =$„vjerojatnost da je slučajno odabrani proizvod defektan". Testiramo hipoteze

$$\begin{aligned} H_0 &: p = 0.07 \\ H_1 &: p > 0.07. \end{aligned}$$

jer proizvođač tvrdi da njegove pošiljke sadrže najviše $p_0 = 0.07$ defektnih proizvoda.

Ovdje je $n = 200$, dakle uzorak je relativno velik pa testiramo asimptotskim Z-testom uz razinu značajnosti $\alpha = 0.05$. Testna statistika je $Z = \frac{\overline{X}_n - p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n}$, a kritično područje $C = [z_{0.05}, \infty\rangle$ uz $z_{0.05} = 1.65$.

Ishod od Z je

$$z = \frac{0.09 - 0.07}{\sqrt{0.07 \cdot 0.93}}\sqrt{200} = 1.108 < 1.65 = z_\alpha\,,$$

pa na razini značajnosti $\alpha = 0.05$ ne odbacujemo H_0 u korist H_1, tj. na razini značajnosti $\alpha = 0.05$ ne možemo tvrditi da nas proizvođač vara. □

Zadatak 8.14. Kocka je bacana 144 puta i palo je ukupno 32 šestice. Možemo li na razini značajnosti $\alpha = 0.05$ zaključiti

(a) da kocka nije simetrična, tj. da se vjerojatnost pada šestice na kocki razlikuje od $\frac{1}{6}$?

(b) da je vjerojatnost pada šestice veća od $\frac{1}{6}$?

Rješenje. Imamo slučajan uzorak duljine $n = 144$ za slučajnu varijablu $X \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$ gdje je $p =$„vjerojatnost da je pala 6-ica".

(a) Testiramo

$$\begin{aligned} H_0 &: p = 1/6 \\ H_1 &: p \neq 1/6. \end{aligned}$$

Budući da je $n = 144$, testiramo asimptotskim Z-testom. Testna statistika je

$$Z = \frac{\overline{X}_n - \frac{1}{6}}{\sqrt{\frac{1}{6} \cdot \frac{5}{6}}}\sqrt{144}\,,$$

a kritično područje $C = \langle -\infty, -z_{\alpha/2}] \cup [z_{\alpha/2}, \infty\rangle$ pri čemu je $z_{\alpha/2} = z_{0.025} = 1.96$.
Budući da je $\overline{x}_n = \frac{32}{144}$ imamo

$$z = \frac{\frac{32}{144} - \frac{1}{6}}{\sqrt{\frac{1}{6} \cdot \frac{5}{6}}} \sqrt{144} = 1.789 \notin C\,.$$

Dakle, ne odbacujemo H_0 u korist H_1 na razini značajnosti $\alpha = 0.05$, tj. iako je 32 više od "očekivanih" $144 \cdot \frac{1}{6} = 24$ šestice, to nije dovoljan dokaz na razini značajnosti 0.05 da zaključimo kako je kocka nesimetrična.

(b) U ovom slučaju testiramo hipoteze

$$H_0 : p = 1/6$$
$$H_1 : p > 1/6.$$

Koristimo istu testnu statistiku kao u prethodnom podzadatku, ali budući da promatramo drugačiju alternativnu hipotezu, kritično područje za istu razinu značajnosti $\alpha = 0.05$ je $C = [z_\alpha, \infty\rangle = [1.65, \infty\rangle$. Sada pak vrijedi $z = 1.789 \in C$ pa na razini značajnosti 0.05 možemo zaključiti da je vjerojatnost pada šestice na kocki veća od $\frac{1}{6}$.

□

Napomena 8.15 (Važno!). Kod testiranja hipoteza, na temelju onoga što želimo provjeriti najprije postavljamo hipoteze. Kasnije te hipoteze ne mijenjamo, tj. pogrešno je na temelju ishoda slučajnog uzorka mijenjati hipoteze. Na primjer, u prethodnom zadatku, nakon što smo vidjeli da je palo 32 šestice, što je više od očekivanih 24 u slučaju da je kocka simetrična, primamljivo je u (a) dijelu alternativnu hipotezu promijeniti iz $p \neq \frac{1}{6}$ u $p > \frac{1}{6}$, što bi dovelo do odbacivanja nulte hipoteze. Ipak, ovakav način je pogrešan i zapravo odgovara testiranju hipoteza $H_0 : p = \frac{1}{6}$, $H_1 : p \neq \frac{1}{6}$ uz dvostruko veću (dakle dvostruko slabiju) razinu značajnosti.

Zadaci za vježbu

Zadatak 8.16. Pokažite da vrijedi formula iz napomene 8.4

Zadatak 8.17. Ako $X \sim B(p)$, pokažite da je aritmetička sredina nepristrani procjenitelj za parametar p.

Zadatak 8.18. Neka kompanija koja proizvodi perece koristi stroj za pakiranje pereca u vrećice na kojima piše da je masa 454 g. Uzet je slučajni uzorak vrećica i dobiveni su rezultati

464 450 450 456 452 433

446 446 450 447 442 438

Pretpostavimo da je distribucija mase vrećice pereca normalno distribuirana s standardnom devijacijom $\sigma = 8.077$. Možemo li na razini značajnosti od $\alpha = 0.05$ zaključiti da stroj ne radi dobro? (Koristite dvostrani test.)

Zadatak 8.19. U grupi od 371 studenta, njih 45 je odabralo broj sedam kada su upitani da izaberu bilo koji broj od jedan do dvadeset. Možete li na temelju tih podataka, na razini značajnosti od $\alpha = 0.01$, zaključiti da je postotak studenata koji biraju broj sedam veći od $1/20 = 0.05$?

9 Rješenja zadataka za vježbu

Vjerojatnost

1.19 (a) dodati Ω (b) dodati $\{2\}$ i $\{1,3,4\}$ **1.20** (a) Da (b) Ne (nađite protuprimjer npr. na $\Omega = \{1,2,3\}$) **1.23** (a) $\mathbb{P}(A) = 0.6, \mathbb{P}(B) = 0.4$ i $\mathbb{P}(B \setminus A) = 0.2$. (b) $\mathbb{P}(B) = 1$ i $\mathbb{P}(A \cap B) = 0$. **1.24** $\Omega = \{(A,B) : A, B \in \{1,2,\ldots,6\}\}, 0.527778$

Prebrojavanje

2.15 (a) 0.0625 (b) 0.504 (c) 0.5545 (d) 0.012 **2.16** (a) 0.20416 (b) 0.21737 (c) 0.39765 (d) 0.0847 **2.17** (a) 0.1898 (b) 0.30848 **2.18** 0.63333 **2.19** (a) 0.1147 (b) 0.21504 **2.20** 0.4 **2.21** $\frac{5!\cdot 3!}{7!}$

Geometrijska vjerojatnost

3.14 (a) 0.12066 (b) 0 **3.15** 0.49 **3.16** 0.5268 **3.17** 0.4874 **3.18** 0.25 **3.19** 0.01108 **3.20** (a) 0.9698 (b) 0.9397

Uvjetna vjerojatnost i nezavisnost

4.15 (a) 0.3 (b) 0.22222 **4.16** $\frac{1}{6}, \frac{1}{3}, \frac{1}{2}$ **4.17** 0.375 **4.18** 0.5 **4.19** 0.825, 0.515 **4.20** 0.67742 **4.21** (a) $\frac{c}{b+c}$ (b) $\frac{c+d}{c+b+d}$ (c) $\frac{c}{b+c}$, za sve $n \geq 1$

Slučajne varijable

5.37 $\mathbb{E}[X] = 3.5, \mathrm{Var}(X) = 2.92, \mathbb{E}[X^2] = 15.17, \mathrm{Var}(X^2) = 149.14, \mathbb{E}[|X-3|] = 1.5, \mathrm{Var}(|X-3|) = 0.92$ **5.38** $\frac{16}{25}$ **5.39** (a) 3 (b) $\mathbb{P}(X=0) = 0.8, \mathbb{P}(X=\sqrt{2}) = 0.2$ (c) 1.3333 **5.40** (a) $X \sim G(1.23 \cdot 10^{-7})$ (b) 8145060 **5.41** (a) 0.5926 (b) 26.6666 **5.42** $\mathbb{E}[X] = 2.333$, $\mathrm{Var}[X] = 1.555$.
5.43 (a) $\begin{pmatrix} 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ \frac{1}{36} & \frac{2}{36} & \frac{3}{36} & \frac{4}{36} & \frac{5}{36} & \frac{6}{36} & \frac{5}{36} & \frac{4}{36} & \frac{3}{36} & \frac{2}{36} & \frac{1}{36} \end{pmatrix}$, 7, 5.833 (b) 35, 29.16666
5.44 (b) $+\infty$

Slučajni vektori

6.8 0.421 **6.9** (a) Ne (b) Da **6.10** (a)

$X_1 \setminus X_2$	100	10
100	2/15	4/15
10	4/15	1/3

(b) -0.1111 (c) 92

Neprekidne slučajne varijable

7.15 (a) 92 (b) 56 **7.14** 0.3147 **7.16** (a) 0.9973 (b)0.3174. **7.17** (a) 0.1 (b) $F(x) = \begin{cases} 1 - e^{-\frac{x}{10}}, & x \geq 0 \\ 0, & x < 0 \end{cases}$, (c) 0.233 (d) 10. **7.25** 0.1977 **7.26** 541

Statistički testovi

8.18 Možemo zaključiti da stroj ne radi dobro na danoj razini značajnosti jer je $z = -2.6591 < -1.96$ **8.19** Na danoj razini značajnosti možemo zaključiti da studenti 'malo češće' biraju broj sedam jer je $z = 6.19 > 2.33$