Sadržaj

1 Osnovni elementi kombinatorike

Često se u praksi susrećemo s problemima određivanja broja elemenata konačnog skupa S. Podsjetimo se da se broj elemenata od S označava sa $|S|$ ili $c(S)$. Kombinatorika je grana matematike koja se bavi prebrojavanjem konačnih skupova. Iako je prebrojavanje svakodnevan posao, prebrojavanje elemenata nekog skupa može biti iznimno težak posao.

Promotrimo neka pitanja na koja ćemo dati odgovore u ovom poglavlju. Koliko postoji različitih tekućih računa u nekoj banci, ako su brojevi deseteroznamenkasti, a prva znamenka nije jednaka nuli? Koliko ima šesteroznamenkastih brojeva kojima su sve znamenke različite? Na koliko načina može deset osoba sjesti na petnaest stolica? Na koliko se načina može izvući sedam brojeva u igri LOTO 7 od 39?

U prvoj točki ovog poglavlja naučit ćemo neka osnovna pravila prebrojavanja. To su "pravilo jednakosti", "pravilo zbroja" i "pravilo umnoška". Pojednostavljeno, prvo pravilo kaže da nije važno kako prebrojimo sve elemente, drugo kaže da je cjelina jednaka zbroju dijelova, a treće da je veličina umnoška jednaka umnošku faktora.

1.1 Osnovna pravila prebrojavanja

Broj elemenata nekog skupa dobivamo prebrojavajući njegove elemente. Međutim, u određenoj situaciji prebrojavanje možemo zamijeniti efikasnijim postupcima.

Primjer 1.1. *Grad A i grad B povezuju tri ceste, a grad B i grad C četiri ceste. Na koliko različitih načina možemo doći iz grada A u grad C, preko grada B?*

Rješenje: Neka su a, b i c ceste koje povezuju gradove A i B, te d, e, f, g ceste koje povezuju gradove B i C. Pogledajmo ponajprije koliko ima traženih putova koji sadrže cestu a. To su putovi ad, ae, af i ag. Dakle, iz grada A u grad C imamo četiri puta koji sadrže cestu a. Analogno, imamo po četiri puta koji sadrže ceste b i c. Dakle, ukupan broj traženih putova je $4 \cdot 3 = 12$. $\square$

Uočimo kako smo u prethodnom primjeru lagano mogli popisati sve putove od grada A do grada C. Naime, to su putovi ad, ae, af, ag, bd, be, bf, bg, cd, ce, cf, cg. Međutim, potpuno isto zaključivanje vrijedi i za skupove s velikim brojem elemenata, gdje bi takvo ispisivanje bilo nepraktično. Uočimo kako smo ovdje, ustvari, odredili Kartezijev umnožak skupova $\{a, b, c\}$ i $\{d, e, f, g\}$. Zato ćemo se prvo prisjetiti definicije Kartezijevog umnoška skupova.

Kartezijev umnožak skupova – pravilo umnoška

Neka su S_1 i S_2 dva neprazna skupa. Kartezijev umnožak skupova S_1 i S_2 je skup $S_1 \times S_2$ čiji su elementi uređeni parovi (a, b), pri čemu je $a \in S_1$, $b \in S_2$. Pišemo

$$S_1 \times S_2 = \{(a, b), a \in S_1, b \in S_2\}.$$

Dva uređena para (a, b) i (x, y) su jednaka ako i samo ako je $a = x$ i $b = y$.

Pogledajmo sada koliko ima elemenata u Kartezijevom umnošku dvaju skupova. Promotrimo skupove s malim brojem elemenata kako bismo mogli ispisati sve njegove članove. Neka je $S_1 = \{a_1, a_2, a_3\}$ i $S_2 = \{b_1, b_2\}$. Tada je

$$S_1 \times S_2 = \{(a_1, b_1), (a_1, b_2), (a_2, b_1), (a_2, b_2), (a_3, b_1), (a_3, b_2)\}.$$

Dakle, broj uređenih parova u danom Kartezijevom umnošku je $6 = 3 \cdot 2$, što prepoznajemo kao umnožak broja elemenata u skupu S_1 s brojem elemenata u skupu S_2. Slično vrijedi i za skupove s po volji mnogo elemenata.

Broj elemenata Kartezijevog umnoška
Ako skup S_1 ima s_1 elemenata, a skup S_2 s_2 elemenata, onda Kartezijev umnožak $S_1 \times S_2$ ima $s_1 s_2$ elemenata. Pišemo $|S_1 \times S_2| = |S_1| \cdot |S_2|$.

Poredak skupova u Kartezijevom umnošku je bitan, jer za različite skupove S_1 i S_2 vrijedi $S_1 \times S_2 \neq S_2 \times S_1$. Međutim, broj elemenata u oba ova skupa je jednak. Prema tome, ukoliko nas zanima samo broj elemenata u Kartezijevom umnošku, ne moramo paziti na poredak podskupova. Napomenimo još kako se prethodno pravilo u literaturi često naziva i pravilo umnoška ili pravilo produkta.

Na isti način možemo definirati i Kartezijev umnožak više skupova. Neka su S_1, S_2, ..., S_k zadani skupovi. Kartezijev umnožak tih skupova je skup uređenih k-torki:

$$S_1 \times S_2 \times \cdots \times S_k = \{(s_1, s_2, \ldots, s_k) : s_1 \in S_1, s_2 \in S_2, \ldots, s_k \in S_k\}.$$

Na potpuno analogan način, kao i u slučaju dvaju skupova, imamo pravilo za broj elemenata Kartezijevog umnoška.

Kartezijev umnožak više skupova
Broj elemenata u Kartezijevom umnošku k skupova je $|S_1 \times S_2 \times \cdots \times S_k| = |S_1| \cdot |S_2| \cdots |S_k|$.

Primjer 1.2. *Snježana u svom garderobnom ormaru ima* 127 *pari cipela,* 135 *haljina i* 75 *šešira. Na koliko se različitih načina Snježana može odjenuti?*

Rješenje: Povežimo zadani problem s Kartezijevim umnoškom odgovarajućih skupova. Ovdje se radi o umnošku triju skupova: skupa $C = \{c_1, c_2, \ldots, c_{127}\}$ iz kojeg Snježana bira cipele, skupa $H = \{h_1, h_2, \ldots, h_{135}\}$ iz kojeg Snježana bira haljinu te skupa $S = \{s_1, s_s, \ldots, s_{75}\}$ iz kojeg Snježana bira šešir. Svaka uređena trojka (c_i, h_j, s_k) određuje jednu odjevnu kombinaciju. Drugim riječima, trebamo naći broj elemenata Kartezijevog umnoška $C \times H \times S$.

Taj je broj jednak $127 \cdot 135 \cdot 75 = 1285875$, pa se Snježana može odjenuti na 1285875 načina. □

Uočimo kako smo u prethodnom primjeru iskoristili relaciju $|S_1 \times S_2 \times S_3| = |S_1| \cdot |S_2| \cdot |S_3|$. Slične ćemo zadatke ubuduće rješavati ne spominjući Kartezijeve umnoške, već samo način biranja pojedinih elemenata uređene $k-$torke.

Primjer 1.3. *Koliko postoji različitih tekućih računa u nekoj banci, ako su brojevi deseteroznamenkasti, a prva znamenka nije jednaka nuli?*

Rješenje: Broj tekućeg računa sastoji se od deset znamenki pri čemu prva znamenka ne smije biti jednaka nuli. Prema tome, prvu znamenku možemo odabrati iz skupa $\{1, 2, \ldots, 9\}$, odnosno na 9 načina, a svaku sljedeću iz skupa $\{0, 1, 2, \ldots, 9\}$, odnosno na 10 načina. Dakle, ukupan broj tekućih računa je

$$9 \cdot 10 \cdot 10 \cdot 10 \cdot 10 \cdot 10 \cdot 10 \cdot 10 \cdot 10 \cdot 10 = 9000000000.$$

□

Primjer 1.4. *Na nekoj visokoj školi studira* 80 *studenata prve godine,* 70 *studenata druge godine,* 60 *studenata treće godine i* 50 *studenata četvrte godine.*

(a) Na koliko se načina može odabrati delegacija od četiri predstavnika svake godine?

(b) Na koliko načina se nakon toga može odabrati druga delegacija, sastavljena od četiri predstavnika svake godine, u kojoj nije nitko iz prve delegacije?

(c) Na koliko se načina može odabrati delegacija od dva predstavnika s različitih godina?

Rješenje: (a) Prva delegacija se može odabrati na $80 \cdot 70 \cdot 60 \cdot 50 = 16800000$ načina.
(b) Nakon što je odabrana prva delegacija, drugu možemo odabrati na $79 \cdot 69 \cdot 59 \cdot 49 = 15758841$ načina.
(c) Ovaj je problem složeniji. Odabir studenata moramo provesti u dva koraka. Najprije se moramo odlučiti za godine studija, a zatim odabrati po jednog studenta sa svake godine. Imamo sljedećih šest mogućnosti:

I i II	$N_1 = 80 \cdot 70 = 5600$
I i III	$N_2 = 80 \cdot 60 = 4800$
I i IV	$N_3 = 80 \cdot 50 = 4000$
II i III	$N_4 = 70 \cdot 60 = 4200$
II i IV	$N_5 = 70 \cdot 50 = 3500$
III i IV	$N_6 = 60 \cdot 50 = 3000$

Ukupan broj izbora je $N = N_1 + N_2 + N_3 + N_4 + N_5 + N_6 = 25100$. □

U Kartezijevom umnošku neki od skupova mogu biti jednaki. Također, možemo promatrati Kartezijeve umnoške skupa sa samim sobom. To će nas dovesti do pojma varijacija s ponavljanjem.

Neka je $S = \{a_1, a_2, \ldots, a_n\}$ zadani skup. Koliko postoji različitih uređenih $k-$torki elemenata skupa S? Isto pitanje možemo postaviti i na ovaj način: Na koliko se različitih načina može izabrati k elemenata skupa S pazeći na njihov poredak, s tim da se elementi mogu ponavljati? Ovdje je očito riječ o broju elemenata u Kartezijevom umnošku k jednakih skupova $S \times S \times \cdots \times S$. Njihov ukupan broj je n^k.

Varijacije s ponavljanjem
Varijacija s ponavljanjem $k-$tog razreda u $n-$članom skupu S je svaka uređena $k-$torka Kartezijevog umnoška k skupova $S \times S \times \ldots \times S = S^k$. Broj varijacija s ponavljanjem označavamo $\overline{V}_n^k$, pa je

$$\overline{V}_n^k = |S \times S \times \ldots \times S| = |S|^k = n^k.$$

Na primjer, varijacije s ponavljanjima drugog razreda u skupu $S = \{a, b, c, d, e, f\}$ su

$$\begin{array}{cccccc} aa & ab & ac & ad & ae & af \\ ba & bb & bc & bd & be & bf \\ ca & cb & cc & cd & ce & cf \\ da & db & dc & dd & de & df \\ ea & eb & ec & ed & ee & ef \\ fa & fb & fc & fd & fe & ff \end{array}$$

Njihov je broj $\overline{V}_6^2 = 6^2 = 36$.

Uočimo kako smo ovdje, radi jednostavnosti, umjesto uređenog para (a, a) pisali aa. Slično vrijedi i za preostale parove. Često ćemo umjesto uređene $k-$torke govoriti o nizu elemenata.

Primjer 1.5. *U nekom gradu ima* 40 *semafora. Svaki semafor može svijetliti crveno, žuto ili zeleno. Na koliko načina mogu svi semafori svijetliti u određenom trenutku?*

Rješenje: Svaki semafor može svijetliti na tri različita načina. Ukupan broj različitih načina, na koje semafori mogu svijetliti, jednak je $3 \cdot 3 \cdots 3 = 3^{40}$. Dakle, riječ je o varijacijama s ponavljanjima razreda 40 u skupu od 3 elementa. □

Koliko podskupova ima konačan skup od n elemenata? Za odgovor na to pitanje pomoći će nam varijacije s ponavljanjem.

Primjer 1.6 (Broj podskupova zadanog skupa). *Koliki je broj podskupova skupa S koji ima n elemenata (uključujući prazan skup i cijeli skup)?*

Rješenje: Svakom elementu skupa S možemo pridružiti broj 0 ili 1, sa značenjem

0: taj se element ne uzima u podskup,

1: taj se element uzima u podskup.

Tako dobivamo niz duljine n koji se sastoji od nula i jedinica, a koji opisuje način izbora podskupa. Na primjer, ako je $S = \{a, b, c, d, e\}$, onda niz 10101 određuje podskup $\{a, c, e\}$, a niz 10000 određuje podskup $\{a\}$. Niz 00000 određuje prazan podskup, a niz 11111 cijeli skup.
Time smo pokazali da je broj podskupova jednak broju nizova duljine n koji se sastoje od nula i jedinica. Možemo reći kako smo svaki podskup od S kodirali nizom nula i jedinica. Prvu znamenku u tom nizu možemo odabrati na dva načina, drugu i sve ostale također na dva načina. Zato je ukupan broj različitih nizova, odnosno broj svih podskupova, jednak 2^n. □

Partitivni skup
Skup svih podskupova skupa S označavamo s $\mathcal{P}(S)$ i nazivamo partitivnim skupom skupa S. Ako je $|S| = n$, onda je $|\mathcal{P}(S)| = 2^n$.

U prethodnom primjeru smo koristili takozvano **pravilo bijekcije ili jednakosti**. Naime, sve elemente skupa $\mathcal{P}(S)$ smo bijektivno preslikali na binarne nizove duljine n, a taj skup smo vrlo jednostavno prebrojili. Pojednostavljeno govoreći, nije važno kako prebrajamo elemente nekog skupa.

Pravilo uzastopnog prebrojavanja

Prebrojavanje elemenata Kartezijevog umnoška možemo poopćiti i na slučaj kada promatramo broj elemenata u nekim njegovim podskupovima. Možemo odgovoriti na drugo pitanje u uvodu ovog poglavlja.

Primjer 1.7. *Koliko ima šesteroznamenkastih brojeva kojima su sve znamenke različite?*

Rješenje: Kako na prvom mjestu ne može biti nula, prvu znamenku možemo odabrati na 9 načina. Drugu znamenku također možemo odabrati na 9 načina zato jer ne možemo koristiti znamenku koja je odabrana na prvom mjestu, ali možemo koristiti nulu. Za svaku sljedeću znamenku imamo po jednu mogućnost manje, pa je broj šesteroznamenkastih brojeva s različitim znamenkama jednak

$$9 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 = 136080.$$

□

Primjer 1.8. *Koliko ima četveroznamenkastih brojeva koji su neparni i kojima su sve znamenke različite?*

Rješenje: Prvo trebamo odrediti mogućnosti za znamenke s dodatnim uvjetima, a to su posljednja i prva znamenka. Kako promatramo neparne brojeve, zadnju znamenku možemo odabrati na 5 načina jer je ona iz skupa $\{1, 3, 5, 7, 9\}$. Sada je jedna znamenka potrošena, pa prvu znamenku možemo odabrati na 8 načina (sve osim nule i zadnje znamenke). Preostale dvije znamenke su u ravnopravnom položaju, drugu biramo na 8 načina, treću na 7 načina, pa je ukupan broj traženih četveroznamenkastih brojeva jednak $8 \cdot 8 \cdot 7 \cdot 5 = 2240$. □

Ideje koje smo koristili u svim dosadašnjim primjerima možemo formulirati na sljedeći način:

Pravilo uzastopnog prebrojavanja
Ako element s_1 možemo izabrati iz skupa S_1 na n_1 različitih načina, nakon toga (bez obzira na to koji smo element već izabrali) element s_2 iz skupa S_2 na n_2 načina, nakon toga element s_3 iz skupa S_3 na n_3 načina itd., onda je ukupan broj načina izbora niza $s_1, s_2, \dots, s_k$ jednak

$$N = n_1 \cdot n_2 \cdots n_k.$$

Pravilo uzastopnog prebrojavanja obično iskazujemo u slobodnijoj formi ovim riječima: Ako se prvi dio posla može učiniti na n_1 način, drugi dio posla na n_2 načina, ..., posljednji na n_k načina, onda se cijeli posao može učiniti na $n_1 \cdot n_2 \cdots n_k$ načina.

Ukoliko poredak elemenata nije bitan, pri primjeni pravila o uzastopnom prebrojavanju moramo biti vrlo oprezni. Objasnit ćemo to kroz dva primjera.

Primjer 1.9. *Na nekom šahovskom turniru svaki je igrač odigrao sa svakim od preostalih po jednu partiju. Ukupno je odigrano* 190 *partija na turniru. Koliko je šahista sudjelovalo na turniru?*

Rješenje: Neka je n traženi broj šahista i označimo šahiste s $A_1, A_2, \dots, A_n$. Šahist A_1 odigrao je po jednu partiju s preostalim šahistima $A_2, A_3, \dots, A_n$, dakle, ukupno $n-1$ partiju. Isto vrijedi i za preostale šahiste, pa je ukupan broj (uređenih) parova partija jednak $n(n-1)$. Međutim, svaka partija je računata dva puta, na primjer između šahista A_1 i A_n kao uređeni par (A_1, A_n) i (A_n, A_1). Prema tome, ukupan broj odigranih partija je upola manji, odnosno jednak je $\frac{n(n-1)}{2}$.
Iz uvjeta zadatka vrijedi da je

$$\frac{n(n-1)}{2} = 190,$$

odakle sređivanjem dobivamo kvadratnu jednadžbu $n^2 - n - 380 = 0$ čija su rješenja $n_1 = -19$ i $n_2 = 20$. Dakako, negativno rješenje otpada, pa je na turniru sudjelovalo 20 šahista.

□

Primjer 1.10. *Na nekom međunarodnom kongresu sudjeluje* 20 *država i svaka država ima* 15 *predstavnika. Na koliko različitih načina možemo odabrati dva predstavnika iz iste države?*

Rješenje: Državu možemo odabrati na 20 načina, a prvog predstavnika te države na 15 načina. Nakon što smo odabrali prvog predstavnika, preostaje 14 mogućnosti za izbor drugog predstavnika. Ponovno je, kao i u prethodnom primjeru, svaki par brojen dva puta. Prema tome, ukupan broj izbora dvaju predstavnika iz jedne države je $20 \cdot 15 \cdot 14 \cdot \frac{1}{2} = 2100$. □

U trećoj točki ovog poglavlja naučit ćemo kako se prethodna dva primjera rješavaju na drugi način.

Kod varijacija s ponavljanjem promatrali smo Kartezijev umnožak skupa sa samim sobom. Pri tome su se elementi u uređenim $k-$torkama mogli ponavljati. Sada ćemo promatrati slučaj kada se ti elementi ne ponavljaju. Drugim riječima, na koliko načina možemo poredati k različitih elemenata iz skupa od n elemenata?

Prvi element možemo odabrati na n načina. Nakon toga, drugi element možemo odabrati na $n-1$ način, jer mora biti različit od prvog. Treći možemo odabrati na $n-2$ načina. Posljednji, $k-$ti na $n-(k-1) = n-k+1$ način. Zato je ukupan broj načina jednak $n(n-1)\cdots(n-k+1)$.

Varijacije bez ponavljanja
Uređena $k-$torka različitih elemenata istog skupa $S = \{a_1, a_2, \ldots, a_n\}$ naziva se varijacijom $k-$tog razreda u skupu od n elemenata. Pri tom mora biti $k \leq n$. Broj varijacija označavamo s V_n^k, pa je
$$V_n^k = n(n-1)\cdots(n-k+1).$$

Primjer 1.11. *U lift zgrade od* 15 *katova ušlo je* 5 *ljudi. Na koliko načina oni mogu izaći iz lifta tako da svaki čovjek izađe na različitom katu?*

Rješenje: Riječ je o varijacijama petog razreda u skupu od 15 elemenata. Zato je traženi broj jednak $V_{15}^5 = 15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 = 360360$.
Svakako je korisnije zapamtiti način na koji smo došli do ovog broja, nego samu formulu. Moramo biti sigurni da razumijemo pravilo uzastopnog prebrojavanja:

- prvi čovjek može izaći na 15 načina,
- drugi čovjek može izaći na 14 načina (među preostalih 14 katova),
- treći čovjek može izaći na 13 načina (među preostalih 13 katova),
- četvrti čovjek može izaći na 12 načina (među preostalih 12 katova),
- peti čovjek može izaći na 11 načina (među preostalih 11 katova).

Zato je broj različitih načina traženih izlazaka iz lifta jednak $15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 = 360360$. □

Primjer 1.12. *Na koliko načina možemo odabrati peteroznamenkasti broj sastavljen od znamenki* $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ *ako:*

(a) su mu sve znamenke različite,

(b) *se na parnim dekadskim mjestima nalaze parne znamenke, a na neparnim dekadskim mjestima neparne znamenke,*

(c) *isto kao i (b), ali su sve znamenke u broju različite.*

Rješenje: Redom dobivamo sljedeće rezultate.

(a) Na $9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 = 15120$ načina.

(b) Na $5 \cdot 4 \cdot 5 \cdot 4 \cdot 5 = 2000$ načina.

(c) Na $5 \cdot 4 \cdot 4 \cdot 3 \cdot 3 = 720$ načina.

$\square$

U nastavku ćemo promotriti još neka važna pravila koja koristimo kod prebrojavanja.

Pravilo zbroja

Krenimo od sljedećeg primjera:

Primjer 1.13. *U nekoj studentskoj grupi ima* 27 *djevojaka i* 23 *mladića. Na koliko načina možemo odabrati jednog predstavnika te grupe?*

Rješenje: Neka je A skup djevojaka, a B skup mladića. Očito su skupovi A i B disjunktni tj. $A \cap B = \emptyset$. Mi trebamo odabrati jedan element iz skupa svih studenata odnosno odrediti broj elemenata unije $A \cup B$. Kako su skupovi A i B disjunktni slijedi da je

$$|A \cup B| = |A| + |B| = 27 + 23 = 50.$$

$\square$

U prethodnom primjeru smo koristili jedno od najjednostavnijih pravila prebrojavanja, takozvano pravilo zbroja.

Pravilo zbroja
Ako su A i B konačni disjunktni skupovi, onda za broj elemenata unije tih dvaju skupova vrijedi
$$|A \cup B| = |A| + |B|.$$

Prethodno pravilo ne vrijedi za skupove koji nisu disjunktni. Primjerice, neka je $A = \{1, 3, 4, 5\}$ i $B = \{1, 4, 6, 7, 8\}$. Tada je $A \cup B = \{1, 3, 4, 5, 6, 7, 8\}$, pa je $|A \cup B| = 7$. Kako je $|A| + |B| = 4 + 5 = 9$, očito ne vrijedi prethodno pravilo zbroja. Razlog je u tome što smo zajedničke elemente skupova A i B brojali dva puta. Stoga, da bismo općenito odredili broj elemenata unije dvaju skupova moramo modificirati prethodnu formulu, tj. oduzeti broj zajedničkih elemenata tih dvaju skupova.

Broj elemenata unije dvaju skupova
Ako su A i B konačni skupovi, onda za broj elemenata unije tih dvaju skupova vrijedi
$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Primjer 1.14. *U razredu od* 30 *učenika svaki učenik uči barem jedan od dvaju stranih jezika: engleski ili njemački. Ukupno* 23 *učenika uče engleski jezik, a njih* 13 *njemački jezik. Koliko učenika uči oba strana jezika?*

Rješenje: Neka je A skup svih učenika koji uče engleski jezik, a B skup svih učenika koji uče njemački jezik. Prema uvjetu zadatka je $|A| = 23$, $|B| = 13$. Nadalje, u razredu je 30 učenika, pa je $|A \cup B| = 30$. Iz formule $|A \cup B| = |A| + |B| - |A \cap B|$ imamo da je $30 = 23 + 13 - |A \cap B|$, odakle je $|A \cap B| = 23 + 13 - 30 = 6$. Prema tome, šest učenika uči oba strana jezika. □

Pravilo komplementa

Neka je U neki skup kojeg promatramo te neka je A njegov podskup, tj. $A \subseteq U$. Komplement skupa A (u skupu U) je skup $U \backslash A$. Komplement skupa A označavamo s $\overline{A}$, odnosno vrijedi $\overline{A} = U \backslash A$. Skup U obično nazivamo univerzalnim skupom.

Očito, za skup i njegov komplement vrijedi $A \cup \overline{A} = U$ i $A \cap \overline{A} = \emptyset$, pa zbog pravila zbroja vrijedi $|U| = |A \cup \overline{A}| = |A| + |\overline{A}|$ iz čega dobivamo:

Pravilo komplementa
Neka je U univerzalni skup te neka je $A \subseteq U$. Tada je broj elemenata komplementa skupa A

$$|\overline{A}| = |U| - |A|.$$

Ponekad je lakše odrediti broj elemenata komplementa nekog skupa nego samog skupa. U to ćemo se uvjeriti na sljedećem primjeru.

Primjer 1.15. *Koliko ima četveroznamenkastih brojeva koji imaju barem jednu znamenku* 7*?*

Rješenje: Traženi brojevi mogu imati jednu, dvije, tri ili četiri znamenke 7. Dakako, vrlo je nepraktično pristupiti ovom problemu na takav način pa ćemo stoga iskoristiti pravilo komplementa.
Neka je U skup svih četveroznamenkastih brojeva, te neka je A skup svih četveroznamenkastih brojeva koji imaju barem jednu znamenku 7. Tada je $\overline{A}$ skup svih četveroznamenkastih brojeva koji nemaju znamenku 7. Lakše nam je odrediti broj elemenata skupa $\overline{A}$. Naime, $|\overline{A}| = 8 \cdot 9 \cdot 9 \cdot 9 = 5832$ jer za prvu znamenku imamo 8 mogućnosti (sve osim 0 i 7), a za preostale po devet mogućnosti (sve osim 7). Slično je $|U| = 9 \cdot 10 \cdot 10 \cdot 10 = 9000$, jer na prvom mjestu ne može biti znamenka 0. Konačno, zbog pravila komplementa je $|A| = |U| - |\overline{A}| = 9000 - 5832 = 3168$. □

1.2 Permutacije

Permutacija skupa $S = \{a_1, a_2, \ldots, a_n\}$ od n različitih elemenata uređena je $n-$torka svih njegovih elemenata.

Tako su na primjer, permutacije skupa $S_1 = \{a, b, c\}$: *abc*, *acb*, *bac*, *bca*, *cab*, *cba*, i to su sve moguće permutacije. Jedna permutacija skupa $S_2 = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ je 0246813579. Dakako, praktički je nemoguće popisati sve permutacije skupa S_2 zato jer je taj broj jako velik. Međutim, u ovoj točki ćemo odgovoriti na pitanje koliko permutacija ima taj skup.

Označimo s P_n broj različitih permutacija skupa od n elemenata. Općenito, broj permutacija skupa od n elemenata dobivamo ovako:

- prvi element možemo izabrati na n načina,
- drugi element možemo izabrati nakon toga na $n - 1$ načina,
- pretposljednji element možemo izabrati na dva načina (jer su samo dva elementa preostala),

- posljednji element biramo samo na jedan način, jer je jedini preostao.

Broj permutacija
Broj različitih permutacija skupa od n elemenata je $P_n = n \cdot (n-1) \cdots 2 \cdot 1 = n!$.

Prisjetimo se kako izraz $n!$ čitamo kao 'n faktorijel'. Nadalje, po definiciji je $0! = 1$.
Primijetimo da je permutacija zapravo varijacija $n-$tog razreda u skupu od n elemenata. Zato je $V_n^n = P_n$. Uočimo nadalje kako broj varijacija $k-$tog razreda u skupu od n elemenata možemo zapisati u obliku $V_n^k = \dfrac{n!}{(n-k)!}$. Zato za $n = k$ vrijedi $V_n^n = \dfrac{n!}{(n-n)!} = \dfrac{n!}{0!} = n! = P_n$.

Primjer 1.16. *Zadan je skup $S = \{1, 2, 3, 4, 5, 6\}$.*

(a) Koliko ima permutacija skupa S?

(b) Koliko permutacija skupa S počinje brojem 1?

(c) Koliko permutacija skupa S ne završava brojem 5?

(d) Koliko permutacija skupa S počinje sa 36?

Rješenje:

(a) Skup S ima šest elemenata pa je ukupan broj permutacija jednak $6!$.

(b) Broj 1 nalazi se na prvom mjestu, pa trebamo ispermutirati preostalih pet brojeva što možemo napraviti na $5! = 120$ načina.

(c) Ovdje koristimo pravilo komplementa. Naime, permutacija koje završavaju brojem 5 ima $5!$, pa permutacija koje ne završavaju brojem 5 ima $6! - 5! = 600$.

(d) Permutacija koje počinju s 36 ima $4! = 24$ jer je pozicija brojeva 3 i 6 određena, dok preostala četiri broja treba ispermutirati.

□

Primjer 1.17. *Na koliko načina možemo na policu rasporediti m knjiga iz matematike i n knjiga iz informatike, tako da*

(a) knjige iz matematike budu na prvih m mjesta,

(b) sve knjige iz matematike budu jedna do druge?

Rješenje:

(a) Knjige iz matematike možemo rasporediti na $m!$ načina (one su na početku police jedna do druge). Neovisno o tome, knjige iz informatike možemo rasporediti na $n!$ načina. One su također jedna do druge, na kraju police. Po pravilu uzastopnog prebrojavanja, ukupan broj mogućih rasporeda je $m! \cdot n!$.

(b) Knjige iz matematike se moraju nalaziti jedna do druge pa ih možemo shvatiti kao jedan 'blok'. Taj blok trebamo permutirati s n knjiga iz informatike. Broj permutacija tih $n+1$ elemenata je $(n+1)!$. Međutim knjige iz matematike unutar bloka možemo permutirati na $m!$ načina. Stoga je, prema pravilu uzastopnog prebrojavanja, broj traženih rasporeda jednak $(n+1)! \cdot m!$. □

Primjer 1.18. *Delegaciju studenata Visoke škole za primijenjeno računarstvo čini šest studenata tako da sa svake od tri godine postoje po dva predstavnika koji su različitih spolova. Na koliko načina možemo te predstavnike sjesti za ravni stol, jednog do drugog, tako da*

(a) mogu sjesti po svojoj volji,

(b) predstavnici iste godine sjede jedan do drugog,

(c) jedna do druge ne smiju sjediti dvije osobe istog spola?

Rješenje:

(a) Svaka permutacija određuje jedan raspored. Broj različitih rasporeda je $6! = 720$.

(b) Rasporedimo najprije parove s iste godine. Tri su para, pa za njihov raspored imamo $3! = 6$ mogućnosti. Unutar svaka dva mjesta određena za jedan par, dva su moguća rasporeda. Zato je ukupan broj svih mogućih rasporeda $6 \cdot (2!)^3 = 6 \cdot 8 = 48$.

(c) Sve muškarce možemo rasporediti na $3! = 6$ načina, baš kao i sve žene. Ukupan broj rasporeda na kojima je prva osoba muškarac (gledano, recimo, s lijeve strane) je $6 \cdot 6$. Jednako toliko ima rasporeda u kojima je prva osoba žena. Ukupan broj svih rasporeda je $2 \cdot 6 \cdot 6 = 72$.

□

Permutacije s ponavljanjem

Želimo li izračunati broj permutacija od n elemenata među kojima ima i jednakih, njihov će broj biti očito manji. Naime, neke od permutacija ispisanih na prethodno opisan način, nećemo više moći razlikovati i njihov će se ukupni broj smanjiti. Pogledajmo sljedeći primjer.

Primjer 1.19.

Slova riječi $TATA$ možemo permutirati na sljedećih šest načina: $AATT$, $ATAT$, $ATTA$, $TAAT$, $TATA$, $TTAA$.
Pretpostavimo za trenutak da možemo razlikovati pojedina slova A i T. Razlikovanjem pojedinih slova A i T iz svake od ovih permutacija dobili bismo $2! \cdot 2! = 4$ nove permutacije. Na primjer, permutaciji $AATT$ bi odgovarale permutacije $A_1A_2T_1T_2$, $A_1A_2T_2T_1$, $A_2A_1T_1T_2$, $A_2A_1T_2T_1$.
Na taj bismo način dobili ukupno $P_4 = 24$ permutacije od četiri različita elementa. Zato za broj permutacija P slova u riječi $TATA$ vrijedi:

$$2! \cdot 2! \cdot P = P_4 \Longrightarrow P = \frac{4!}{2! \cdot 2!} = 6.$$

□

Poopćavajući ovo razmatranje, dolazimo do sljedećeg zaključka:

Permutacije s ponavljanjem
Neka u nizu $s_1, s_2, \ldots, s_n$ postoji prva skupina od k_1 identičnih elemenata, druga skupina od k_2 identičnih elemenata, …, $r-$ta skupina od k_r identičnih elemenata, $k_1 + k_2 + \ldots + k_r = n$. Bilo koji razmještaj elemenata takvog niza nazivamo permutacijom s ponavljanjem. Njihov ukupni broj označavamo s $P_n^{k_1,k_2,\ldots,k_r}$ i vrijedi

$$P_n^{k_1,k_2,\ldots,k_r} = \frac{n!}{k_1! \cdot k_2! \cdots k_r!}$$

Primjer 1.20. *Koliko se različitih riječi može napisati od slova riječi MATEMATIKA?*

Rješenje: Ovdje je riječ o nizu slova A,A,A,E, I, K, M, M, T, T. Zato je

$$N = P_{10}^{3,1,1,1,2,2} = \frac{10!}{3!1!1!1!2!2!} = 151200.$$

Po dogovoru, u ovakvim primjerima ne pišemo broj 1 niti u oznaci niti u razlomcima:

$$N = P_{10}^{3,2,2} = \frac{10!}{3!2!2!}.$$

□

1.3 Kombinacije

U Primjeru 1.9, gdje smo računali broj odigranih partija na šahovskom turniru, poredak šahista nije bio važan. Zato smo ukupan broj uređenih parova partija morali podijeliti s 2, jer je svaka partija bila uračunata dvaput. U mnogim problemima prebrojavanja poredak izabranih elemenata nije bitan. Na primjer, u igri LOTO 7 od 39 nije važno kojim se redom izvlači prvih 7 brojeva, već samo koji su to brojevi. Na koliko se načina može izvući 7 brojeva od 39? Općenitije, pitamo se: Na koliko se načina može izvući k elemenata iz skupa S od n elemenata, ne pazeći na njihov poredak? Označimo taj broj s C_n^k. Svaki izbor k različitih elemenata skupa $S = \{a_1, a_2, \dots, a_n\}$ određuje jedan njegov podskup koji ima k elemenata. Pokazali smo kako niz nula i jedinica duljine n određuje neki podskup skupa S.

Broj C_n^k jednak je:

- broju načina na koji iz skupa od n elemenata možemo izvući k elemenata, ne pazeći na njihov poredak,
- broju različitih podskupova s k elemenata uzetih iz skupa od n elemenata.

Odredimo taj broj. Izbor jednog podskupa koji ima k elemenata određen je nizom nula i jedinica duljine n, ali takvih da u njemu postoji točno k jedinica.

Ilustrirajmo izbor podskupova koji imaju dva elementa na skupu $S = \{a_1, a_2, a_3, a_4\}$. Na koliko načina možemo odabrati dva njegova elementa?
Ispišimo niz nula i jedinica i njemu odgovarajući izbor elemenata ovog skupa:

1100	a_1, a_2
1010	a_1, a_3
1001	a_1, a_4
0110	a_2, a_3
0101	a_2, a_4
0011	a_3, a_4

Broj svih načina jednak je broju svih permutacija niza 1100 kojih ima

$$P_4^{2,2} = \frac{4!}{2!2!} = 6.$$

Vidimo da je ukupan broj načina jednak broju permutacija u nizu od n nula i jedinica, u kojem ima k jedinica i $n-k$ nula:

$$C_n^k = P_n^{k,n-k} = \frac{n!}{k!(n-k)!}.$$

Taj broj nazivamo binomni koeficijent, označavamo ga s $\binom{n}{k}$, te čitamo 'n povrh k'. Dakle, vrijedi

$$C_n^k = \binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}.$$

Prisjetimo se kako za binomne koeficijente vrijede sljedeće relacije:

$$\binom{n}{0} = 1, \ \binom{n}{1} = n, \ \binom{n}{2} = \frac{n(n-1)}{2}, \ \binom{n}{3} = \frac{n(n-1)(n-2)}{6}, \ \binom{n}{n} = 1.$$

Također, binomni koeficijenti posjeduju svojstvo simetrije, odnosno $\binom{n}{k} = \binom{n}{n-k}$, za $k \leq n$. Tako je, primjerice, $\binom{11}{9} = \binom{11}{11-9} = \binom{11}{2} = \frac{11 \cdot 10}{2} = 55$.

Kombinacije
Svaki podskup od k različitih elemenata skupa S nazivamo kombinacijom u skupu S. Broj različitih kombinacija je $C_n^k = \frac{n!}{k!(n-k)!} = \binom{n}{k}$.

Primjer 1.21. *Na koliko se načina u igri LOTO može izvući 7 brojeva od 39 zadanih?*

Rješenje: Iz skupa od 39 brojeva moramo odabrati 7 brojeva. To se može napraviti na

$$\binom{39}{7} = \frac{39 \cdot 38 \cdot 37 \cdot 36 \cdot 35 \cdot 34 \cdot 33}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7} = 15380937$$

načina. □

Primjer 1.22. *Imamo n predmeta u jednoj vreći i m predmeta u drugoj vreći. Iz prve se vreće vadi r predmeta, a iz druge s predmeta. Izvađene predmete nanizujemo jedan do drugog. Koliko takvih nizova možemo dobiti?*

Rješenje: Iz prve vreće, od n predmeta moramo odabrati r predmeta, a to možemo napraviti na $\binom{n}{r}$ načina. Iz druge vreće, trebamo od m predmeta odabrati s predmeta, što možemo napraviti na $\binom{m}{s}$ načina. Dakle, ukupno smo izvadili $r+s$ predmeta koje permutiramo na $(r+s)!$ načina. Konačno, prema pravilu produkta, traženi broj nizova jednak je

$$\binom{n}{r}\binom{m}{s}(r+s)!.$$

□

Primjer 1.23. *Od 7 žena i 4 muškarca treba izabrati delegaciju. Na koliko se načina može izabrati delegacija tako da se ona sastoji od*

a) petero ljudi,

b) petero ljudi, i to 3 žene i 2 muškarca,

c) bilo kojeg broja ljudi ali mora biti jednako muškaraca i žena?

Rješenje:

a) Imamo ukupno $7+4=11$ ljudi pa petero možemo odabrati na

$$\binom{11}{5} = \frac{11\cdot 10\cdot 9\cdot 8\cdot 7}{1\cdot 2\cdot 3\cdot 4\cdot 5} = 462 \text{ načina}$$

b) Tri žene možemo odabrati na $\binom{7}{3} = \frac{7\cdot 6\cdot 5}{1\cdot 2\cdot 3} = 35$ načina. Dva muškarca možemo odabrati na $\binom{4}{2} = \frac{4\cdot 3}{2} = 6$ načina, pa prema pravilu uzastopnog prebrojavanja dobivamo da je broj traženih delegacija jednak $\binom{7}{3}\binom{4}{2} = 35\cdot 6 = 210$.

c) Mogućnosti su da se izabere po jedna osoba svakog spola, po dvije, tri i četiri. Primjenom pravila umnoška i zbroja dobivamo da je traženi broj jednak

$$\binom{7}{1}\binom{4}{1} + \binom{7}{2}\binom{4}{2} + \binom{7}{3}\binom{4}{3} + \binom{7}{4}\binom{4}{4} = 329.$$

□

Primjer 1.24. *Koliko je različitih mogućnosti podjele* 32 *karte na* 4 *igrača tako da se svakom dijeli odjednom po* 8 *karata?*

Rješenje: Prvom igraču možemo podijeliti 8 karata na $\binom{32}{8}$ načina. Sada je preostalo $32-8=24$ karte te ih drugom igraču možemo podijeliti na $\binom{24}{8}$ načina. Trećem igraču karte možemo podijeliti na $\binom{16}{8}$ načina, te četvrtom igraču na $\binom{8}{8}$ načina. Dakle, ukupan broj podjela karata je

$$\binom{32}{8}\binom{24}{8}\binom{16}{8}\binom{8}{8} = 99561092450391000.$$

□

Primjer 1.25. *Na koliko se načina iz snopa od* 32 *karte može odabrati* 8 *karata tako da među njima bude*

a) točno tri kralja i tri asa,

b) barem jedan as?

Rješenje:

a) Kako imamo po četiri kralja i asa njih možemo odabrati (i jedne i druge) na $\binom{4}{3} = \binom{4}{1} = 4$ načina. Nakon što smo odabrali kraljeve i aseve, od preostalih $32-4-4=24$ karata odabiremo dvije na $\binom{24}{2} = \frac{24\cdot 23}{2} = 276$ načina. Konačno, prema pravilu uzastopnog prebrojavanja broj traženih odabira je

$$\binom{4}{3}\binom{4}{3}\binom{24}{2} = 4\cdot 4\cdot 276 = 4416.$$

b) Ovdje koristimo pravilo komplementa. Odredimo koliko ima odabira u kojima nema niti jednog asa. Kako imamo 4 asa imamo $32 - 4 = 28$ karata koje nisu asevi. Stoga imamo ukupno $\binom{28}{8}$ odabira po 8 karata u kojima nema niti jednog asa. Zato je traženi broj odabira koje sadrže barem jednog asa jednak

$$\binom{32}{8} - \binom{28}{8} = 7410195.$$

□

1.4 Zadatci za ponavljanje

Zadatak 1.1. Koliko ima šesteroznamenkastih brojeva sastavljenih od znamenki $\{1, 2, 3, 4, 5, 6\}$ koji su parni? Pri tome je svaka znamenka upotrijebljena točno jednom.

Rješenje: Kako tražimo parne brojeve, zadnja znamenka pripada skupu $\{2, 4, 6\}$ pa je možemo odabrati na 3 načina. Preostalih 5 znamenaka trebamo rasporediti na 5 mjesta, što možemo napraviti na $5!$ načina. Dakle, traženih brojeva ima $3 \cdot 5! = 3 \cdot 120 = 360$. □

Zadatak 1.2. Na policu treba složiti 4 knjige iz matematike, 3 iz fizike, 3 iz kemije i 2 iz biologije, tako da knjige iz iste struke budu jedna do druge. Na koliko je načina to moguće napraviti?

Rješenje: Sve knjige iz iste struke moraju biti jedna do druge, pa svaku struku možemo shvatiti kao jedan blok. Dakle, imamo četiri bloka koje možemo ispermutirati na $4!$ načina. Nakon toga, permutiramo knjige unutar svakog pojedinog bloka, što redom možemo napraviti na $4!, 3!, 3!, 2!$ načina. Konačno, prema pravilu umnoška, traženi broj rasporeda je

$$4! \cdot 4! \cdot 3! \cdot 3! \cdot 2! = 24 \cdot 24 \cdot 6 \cdot 6 \cdot 2 = 41472.$$

□

Zadatak 1.3. U kutiji je 6 bijelih i 5 crnih kuglica. Na koliko načina možemo odabrati 4 kuglice tako da budu:

(a) bilo koje boje,

(b) 2 bijele i 2 crne,

(c) sve iste boje?

Rješenje:

(a) Iz skupa od 11 kuglica, 4 kuglice možemo odabrati na $\binom{11}{4} = 330$ načina.

(b) Bijele kuglice odabiremo na $\binom{6}{2}$ načina, a crne na $\binom{5}{2}$ načina. Zato je broj odabira jednak $\binom{6}{2}\binom{5}{2} = 150$.

(c) Kako su sve kuglice iste boje, odabiremo ili 4 bijele ili 4 crne kuglice. Bijele kuglice možemo odabrati na $\binom{6}{4}$ načina, a crne na $\binom{5}{4}$ načina. Ti odabiri su disjunktni, pa je prema pravilu zbroja traženi broj jednak $\binom{6}{4} + \binom{5}{4} = 20$.

□

Zadatak 1.4. Hokejašku momčad čine 2 vratara, 7 braniča i 10 napadača. Na koliko načina se može odabrati početna postava koju čine 1 vratar, 2 braniča i 3 napadača?

Rješenje: Vratara, braniče i napadače možemo redom odabrati na $\binom{2}{1}$, $\binom{7}{2}$ i $\binom{10}{3}$ načina, pa početnu momčad možemo odabrati na $\binom{2}{1}\binom{7}{2}\binom{10}{3} = 5040$ načina. □

Zadatak 1.5. Koliko ima sedmeroznamenkastih brojeva

(a) kojima su sve znamenke neparne,

(b) koji imaju dvije znamenke 1, tri znamenke 2 i dvije znamenke 3?

Rješenje:

(a) Svaku znamenku odabiremo iz skupa $\{1, 3, 5, 7, 9\}$, dakle, na 5 načina. Kako se znamenke smiju ponavljati, traženih sedmeroznamenkastih brojeva ima $5^7 = 78125$.

(b) Trebamo odrediti koliko se sedmeroznamenkastih brojeva može zapisati pomoću znamenki $1, 1, 2, 2, 2, 3, 3$. Prema formuli za broj permutacija s ponavljanjem, taj je broj jednak

$$\frac{7!}{2!3!2!} = 210.$$

□

Zadatak 1.6. Koliko ima peteroznamenkastih brojeva koji imaju

(a) barem dvije znamenke 5,

(b) točno dvije znamenke 5?

Rješenje:

(a) Ovdje ćemo koristiti pravilo komplementa. Od broja svih peteroznamenkastih brojeva oduzet ćemo one koji imaju najviše jednu peticu. Dakle, trebamo odrediti koliko ima peteroznamenkastih brojeva koji ne sadrže peticu te koliko je onih koji sadrže točno jednu peticu.

Peteroznamenkastih brojeva koji nemaju peticu ima $8 \cdot 9^4 = 52488$, zato jer na prvom mjestu ne smije biti nula i petica, a na preostala četiri mjesta ne smije biti petica.

Pogledajmo koliko ima peteroznamenkastih brojeva s jednom peticom. Tu razlikujemo dva slučaja, ovisno o tome da li se petica nalazi ili ne nalazi na prvom mjestu. Ako je petica na prvom mjestu, svako od preostalih mjesta možemo popuniti na 9 načina, pa takvih brojeva ima $9^4 = 6561$. Ako se petica ne nalazi na prvom mjestu, onda je možemo rasporediti na 4 načina. Nakon što smo rasporedili peticu, prvo mjesto možemo popuniti na 8 načina jer tu, osim petice, ne smije biti i nula, a svako od preostala tri mjesta na 9 načina, pa takvih brojeva ima $4 \cdot 8 \cdot 9^3 = 23328$. Dakle, peteroznamenkastih brojeva s jednom peticom ima $23328 + 6561 = 29889$.

Konačno, svih peteroznamenkastih brojeva ima $9 \cdot 10^4 = 90000$, pa onih koji imaju barem dvije petice ima

$$90000 - (52488 + 29889) = 7623.$$

(b) Moramo razlikovati dva slučaja, ovisno o tome da li se petica nalazi na početku ili ne. Ako se petica ne nalazi na prvom mjestu, onda dvije petice na preostala mjesta možemo rasporediti na $\binom{4}{2}$ načina. Sada, prvo mjesto možemo popuniti na 8 načina (jer tu osim petice ne smije biti ni nula), a preostala dva mjesta na 9 načina. Dakle, peteroznamenkastih brojeva s dvije petice, koje nisu na prvom mjestu, ima $\binom{4}{2} \cdot 8 \cdot 9^2 = 3888$.

Ako je jedna petica na prvom mjestu, onda drugu možemo rasporediti na 4 načina te preostala tri mjesta popuniti na po 9 načina. Prema tome, takvih brojeva ima $4 \cdot 9^3 = 2916$.

Konačno, peteroznamenkastih brojeva s točno dvije petice ima

$$3888 + 2916 = 6804.$$

□

Zadatak 1.7. Dokažite kombinatorno da za svaki prirodan broj n vrijedi identitet

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n-1} + \binom{n}{n} = 2^n.$$

Rješenje: Promotrimo neki $n-$člani skup S. Znamo da on ima 2^n podskupova, uključujući i prazan skup. S druge strane, sve podskupove od S možemo rastaviti na prazan skup, jednočlane podskupove, dvočlane podskupove, ..., $n-$člane podskupove. Prazan skup je jedan, odnosno $\binom{n}{0}$, jednočlanih ima $\binom{n}{1}$, dvočlanih $\binom{n}{2}$, itd. Dakle, lijeva strana identiteta predstavlja zbroj broja jednočlanih, dvočlanih, ..., $n-$članih podskupova od S te praznog skupa, što je broj svih podskupova od S. Time je identitet dokazan. □

2 Vjerojatnost

Kao i kod mnogih drugih grana matematike, razvoj teorije vjerojatnosti bio je potaknut različitim praktičnim problemima. Obično se za početak teorije vjerojatnosti uzima 1654. godina kada se slavnim francuskim matematičarima Pascalu i Fermatu obratio ugledni građanin de Méré s pitanjem da li se preporučljivo kladiti da će se u 24 uzastopna bacanja dviju kocki pojaviti dvostruka šestica. Iza takvog i sličnih pitanja zapravo su se krili matematički modeli za rješavanje korisnih problema praktične matematike.

Jedna od osnovnih poteškoća u razvoju teorije vjerojatnosti bila je definicija vjerojatnosti, pojam koji je intuitivno bio vrlo jasan, a ipak matematički neuhvatljiv. Traženje takve definicije trajalo je gotovo tri stoljeća i obilježeno je mnogim poteškoćama. Problem je konačno riješio ruski matematičar A. Kolmogorov 1933. godine aksiomatskim pristupom teoriji vjerojatnosti. No prije nego što dođemo do te definicije, potrebno je uvesti pojam događaja i algebre događaja.

2.1 Događaji i algebra događaja

Pojam događaja možemo jednostavno dovesti u vezu sa **slučajnim pokusom**. Tako nazivamo svaki pokus čiji ishod nije unaprijed određen. Taj ishod ovisi o nekim nepredvidivim okolnostima i stoga je slučajan. Bacimo li igraću kocku, ona će pasti na jedan od šest brojeva, na koji – unaprijed ne možemo znati. Vrijeme ispravnog rada žarulje nitko ne može unaprijed predvidjeti.

Ishod slučajnog pokusa zovemo **elementarni događaj** i označavamo slovom ω. Takvih ishoda može biti konačno mnogo, ali jednako tako i beskonačno mnogo. Biranje na sreću jedne točke unutar jediničnog intervala $[0,1]$ ima kao mogući ishod beskonačno mnogo elementarnih događaja.

Skup svih mogućih ishoda, odnosno svih elementarnih događaja, označavamo slovom Ω. Po volji odabrane događaje vezane uz taj pokus označavat ćemo velikim slovima latinične abecede: A, B, C .... Oni se sastoje od određenog broja elementarnih događaja. Stoga su to podskupovi skupa Ω. Pogledajmo sada neke najjednostavnije primjere slučajnih pokusa te njihovih elementarnih događaja.

Primjer 2.1. *Bacamo jednu igraću kocku čije su strane označene brojevima od* 1 *do* 6*. Odredite elementarne događaje, skup* Ω *te događaje*

$$\begin{aligned} A &= \{\textit{pao je neparni broj}\} \\ B &= \{\textit{pao je broj veći od 3}\} \\ C &= \{\textit{pao je parni broj manji od 6}\} \\ D &= \{\textit{pao je prost broj}\}. \end{aligned}$$

Rješenje: Elementarni događaji su brojevi na koje kocka može pasti:

$$\omega_1 = 1,\ \omega_2 = 2,\ \omega_3 = 3,\ \omega_4 = 4,\ \omega_5 = 5,\ \omega_6 = 6.$$

Stoga je skup svih elementarnih događaja $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\} = \{1, 2, 3, 4, 5, 6\}$.
Zadani događaji su redom

$$\begin{aligned}
A &= \{\text{pao je neparni broj}\} = \{1,3,5\} \\
B &= \{\text{pao je broj veći od 3}\} = \{4,5,6\} \\
C &= \{\text{pao je parni broj manji od 6}\} = \{2,4\} \\
D &= \{\text{pao je prost broj}\} = \{2,3,5\}.
\end{aligned}$$

□

U pokusu koji ima samo konačno mnogo ishoda događaj je bilo koji podskup od Ω. U prethodnom primjeru skup Ω ima šest elemenata, pa on ima $2^6 = 64$ različitih podskupova. Dakle, toliko će biti i različitih događaja u prethodnom pokusu. Među njima ima: $\binom{6}{1} = 6$ jednočlanih (elementarnih) događaja, $\binom{6}{2} = 15$ događaja od po dva elementarna, $\binom{6}{3} = 20$ događaja s tri elementarna, $\binom{6}{4} = 15$ događaja s četiri elementarna i $\binom{6}{5} = 6$ događaja s pet elementarnih. Naravno da je sve te događaje teško riječima opisati na prihvatljiv način, poput opisa u prethodnom primjeru.
Ovom popisu događaja nedostaju još dva: već spomenuti događaj Ω koji sadrži sve elementarne događaje, te jedan događaj koji ne sadrži niti jedan elementarni. Skup Ω i sam je događaj, on se ostvaruje pri svakom ishodu pokusa. Nazivamo ga stoga **sigurni događaj**. Njegova suprotnost je nemoguć događaj, koji se pri realizaciji pokusa nikad ne može ostvariti. Označavamo ga simbolom $\emptyset$. Na primjer, u prethodnom primjeru $E = \{\text{pao je broj veći od 7}\}$ predstavlja nemoguć događaj, odnosno $\emptyset$. S druge strane, $F = \{\text{pao je broj manji od 7}\}$ predstavlja siguran događaj, odnosno Ω.
Pogledajmo sada još jedan primjer.

Primjer 2.2. *Košarkaš izvodi slobodno bacanje tri puta. Za svako slobodno bacanje bilježimo je li postignut koš* $(+)$ *ili ne* $(-)$. *Odredite* Ω, *elementarne događaje te događaje*

$$\begin{aligned}
A &= \{\textit{košarkaš je postigao dva koša}\}, \\
B &= \{\textit{košarkaš je pogodio koš u trećem bacanju}\}, \\
C &= \{\textit{košarkaš je barem jednom pogodio i barem jednom promašio koš}\}, \\
D &= \{\textit{košarkaš je pogodio koš dva puta zaredom}\}.
\end{aligned}$$

Rješenje: Elementarnih događaja ima osam. To su

$$\begin{array}{llll}
\omega_1 = ---, & \omega_2 = --+, & \omega_3 = -+-, & \omega_4 = +--, \\
\omega_5 = -++, & \omega_6 = +-+, & \omega_7 = ++-, & \omega_8 = +++.
\end{array}$$

Pri tome, poredak nabrajanja nije važan. Ovdje smo, kratkoće radi, s $-\ -\ +$ označili uređenu trojku $(-,-,+)$ i slično za ostale elementarne događaje. Događaji A, B, C i D su redom

$$\begin{aligned}
A &= \{\text{košarkaš je postigao dva koša}\} = \{\omega_5, \omega_6, \omega_7\}, \\
B &= \{\text{košarkaš je pogodio koš u trećem bacanju}\} = \{\omega_2, \omega_5, \omega_6, \omega_8\}, \\
C &= \{\text{košarkaš je barem jednom pogodio i barem jednom promašio koš}\} = \{\omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7\}, \\
D &= \{\text{košarkaš je pogodio koš dva puta zaredom}\} = \{\omega_5, \omega_7, \omega_8\}.
\end{aligned}$$

□

Uspoređivanje događaja

Kažemo da događaj A **povlači** događaj B ako realizacija događaja A povlači realizaciju događaja B. To znači da B sadrži sve elementarne događaje koji ulaze u događaj A. Pišemo $A \subset B$ ili $A \Longrightarrow B$.

Primjer 2.3. *Istovremeno bacamo pet novčića. Označimo događaje*

$$\begin{aligned} A &= \{\textit{najviše jedan novčić pokazuje glavu}\}, \\ B &= \{\textit{barem tri novčića pokazuju pismo}\}. \end{aligned}$$

Da li događaj A povlači događaj B? Da li događaj B povlači događaj A?

Rješenje: U događaj A ulaze svi elementarni događaji kod kojih najviše jedan novčić pokazuje glavu. To znači da je broj pisama u takvim elementarnim događajima jednak četiri ili pet. Očito, svi ti elementarni događaji pripadaju skupu B, zato jer su u B elementarni događaji koji imaju barem tri pisma. Dakle, $A \Longrightarrow B$.

Obrat ne vrijedi jer u B imamo elementarne događaje s točno tri pisma, a oni ne pripadaju događaju A. Prema tome, događaj B ne povlači događaj A. □

Primjer 2.4. *Istovremeno bacamo dvije kocke. Označimo događaje*

$$\begin{aligned} A &= \{\textit{zbroj brojeva na kockama jednak je } 10\}, \\ B &= \{\textit{oba broja veća su od } 3\}. \end{aligned}$$

Dokažite da događaj A povlači događaj B. Da li događaj B povlači događaj A?

Rješenje: Događaj A sastoji se od elementarnih događaja $(4,6)$, $(5,5)$ i $(6,4)$ jer samo na ta tri načina možemo postići zbroj 10. Pri tome prva komponenta u uređenom paru označava broj na prvoj, a druga komponenta broj na drugoj kocki. Dobivena tri elementarna događaja imaju komponente veće od 3, pa oni sigurno pripadaju događaju B. Prema tome, $A \Longrightarrow B$.

Događaj B očito ne povlači događaj A zato jer, na primjer, događaj B sadrži elementarni događaj $(4,5)$ koji nije u A. □

Ukoliko vrijedi $A \subset B$ i $B \subset A$, onda kažemo da su A i B **ekvivalentni** ili **jednaki** i pišemo $A = B$. Ekvivalentni događaji sastoje se od istih elementarnih događaja.

Suprotnost ovoj situaciji je ona u kojoj A i B nemaju zajedničkih elementarnih događaja. Događaji A i B su **disjunktni**, ako se istovremeno ne mogu ostvariti i jedan i drugi. Kažemo još da se A i B **međusobno isključuju**.

Primjer 2.5. *Svaki od četiri strijelca gađa istu metu jedanput. Promotrimo događaje*

$$\begin{aligned} A &= \{\textit{tri strijelca su pogodila metu}\}, \\ B &= \{\textit{jedan strijelac je promašio metu}\}, \\ C &= \{\textit{dva strijelca su pogodila, a dva promašila metu}\}. \end{aligned}$$

Koji od ovih događaja povlače neki drugi, koji su ekvivalentni, a koji se međusobno isključuju?

Rješenje: Očito vrijedi $A \Longrightarrow B$ i $B \Longrightarrow A$. Prema tome, događaji A i B su ekvivalentni, tj. $A = B$. Nadalje događaji A i C te B i C su disjunktni, odnosno vrijedi $A \cap C = \emptyset$ i $B \cap C = \emptyset$. □

Algebra događaja

Familiju svih događaja koji se pojavljuju u nekom pokusu označavat ćemo s $\mathcal{F}$ i zvati **algebra događaja**. Riječ algebra ovdje sugerira da ćemo na događajima moći činiti određene operacije nalik na algebarske.

Neka su A i B događaji. Pomoću njih možemo načiniti nove događaje:

Unija i presjek događaja
Događaj koji se ostvaruje ako se ostvario barem jedan od događaja A, B naziva se unija ili zbroj (suma) događaja i označava s $A \cup B$, $A + B$, A ili B.
Događaj koji se ostvaruje ako su se ostvarila oba događaja A i B naziva se presjek ili umnožak (produkt) događaja i označava s $A \cap B$, AB, A i B.

Primjer 2.6. *Bacamo jednu kocku. Istaknimo događaje*

$$\begin{aligned} A &= \{\textit{pao je neparan broj}\} \\ B &= \{\textit{pao je prost broj}\}. \end{aligned}$$

Odredite događaje $A \cup B$ i $A \cap B$.

Rješenje: Imamo da je

$$\begin{aligned} A \cup B &= \{\text{pao je broj koji je neparan ili prost}\} = \{1, 2, 3, 5\} \\ A \cap B &= \{\text{pao je broj koji je i neparan i prost}\} = \{3, 5\}. \end{aligned}$$

□

Operacije unije i presjeka mogu se definirati i za nekoliko događaja. Unija n događaja je događaj

$$A = A_1 \cup A_2 \cup \ldots \cup A_n$$

koji se ostvaruje ako se ostvario barem jedan od događaja $A_1, A_2, \ldots, A_n$.
Presjek n događaja je događaj

$$A = A_1 \cap A_2 \cap \ldots \cap A_n$$

koji se ostvaruje ako se ostvario svaki od događaja $A_1, A_2, \ldots, A_n$.

Razlika događaja. Komplement događaja
Događaj koji se ostvaruje ako se ostvari događaj A, a da se ne ostvari događaj B, nazivamo razlika događaja A i B i označavamo s $A \backslash B$, $A - B$.
Događaj $\Omega \backslash A$ nazivamo komplementom ili suprotnim događajem događaja A. On se ostvaruje ako i samo ako se A nije ostvario. Označavamo ga s $\overline{A}$ ili s A^c.

Očito, operacija komplementiranja je involutorna, odnosno vrijedi $\overline{\overline{A}} = A$. Nadalje, važna formula koja razliku događaja izražava pomoću presjeka glasi:

$$A \backslash B = A \cap \overline{B}.$$

Uvjerite se grafički, odnosno pomoću Venn-Eulerovih dijagrama, u ispravnost prethodne formule.

Spomenimo još na kraju ove točke relacije koje povezuju operacije komplementiranja, unije i presjeka događaja:

$$\begin{aligned}\overline{A \cup B} &= \overline{A} \cap \overline{B},\\ \overline{A \cap B} &= \overline{A} \cup \overline{B}.\end{aligned}$$

Te formule nazivamo **de Morganovi zakoni**. Uvjerite se u njihovu ispravnost pomoću Venn-Eulerovih dijagrama!

2.2 Aksiomatska definicija vjerojatnosti

Sada možemo precizno iskazati aksiomatsku definiciju vjerojatnosti Kolmogorova iz 1933. godine, koju smo najavili na početku ovog poglavlja. To je funkcija koja svakom događaju pridružuje realan broj.

Vjerojatnost
Vjerojatnost je preslikavanje $P : \mathcal{F} \mapsto [0, 1]$ definirano na algebri događaja $\mathcal{F}$, koje ima svojstva
1) $P(\Omega) = 1$, $P(\emptyset) = 0$ (**normiranost**),
2) ako je $A \subset B$, onda vrijedi $P(A) \leq P(B)$ (**monotonost**),
3) ako su A i B disjunktni događaji, onda je $P(A \cup B) = P(A) + P(B)$ (**aditivnost**).
Broj $P(A)$ nazivamo vjerojatnost događaja A.

Svojstva vjerojatnosti

Izvedimo neka dodatna svojstva vjerojatnosti koja neposredno slijede iz navedene definicije. Neka je A po volji odabran događaj, a $\overline{A}$ njegov komplement. Onda vrijedi $A \cup \overline{A} = \Omega$ i pritom su A i $\overline{A}$ disjunktni. Iskoristimo li sada svojstva normiranosti i aditivnosti, imamo da je

$$1 = P(\Omega) = P(A \cup \overline{A}) = P(A) + P(\overline{A}),$$

odakle je $P(\overline{A}) = 1 - P(A)$. Time smo dobili formulu za vjerojatnost komplementa događaja.

Vjerojatnost komplementa
Za svaki događaj A vrijedi $P(\overline{A}) = 1 - P(A)$.

Iz definicije vidimo da je vjerojatnost unije dvaju disjunktnih događaja jednaka zbroju vjerojatnosti tih događaja. Pokažimo sada kako se računa vjerojatnost unije u slučaju kada skupovi nisu disjunktni.

Vjerojatnost unije
Za bilo koja dva događaja vrijedi $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Da dokažemo ovo svojstvo, događaj $A \cup B$ prikazat ćemo kao uniju dvaju disjunktnih događaja:

$$A \cup B = A \cup (B \cap \overline{A}).$$

Uvjerite se u ispravnost prethodne relacije pomoću Venn-Eulerovih dijagrama! Slično tome, B možemo rastaviti ovako:

$$B = (A \cap B) \cup (B \cap \overline{A})$$

i ponovo su događaji s desne strane jednakosti disjunktni.
Primijenimo li svojstvo aditivnosti vjerojatnosti na prethodne dvije relacije, dobivamo

$$\begin{aligned} P(A \cup B) &= P(A) + P(B \cap \overline{A}) \\ P(B) &= P(A \cap B) + P(B \cap \overline{A}). \end{aligned}$$

Konačno, oduzimanjem prethodnih dviju jednakosti dobivamo traženu formulu:

$$P(A \cup B) - P(B) = P(A) - P(A \cap B).$$

□

Primjer 2.7. *Neka su A i B događaji takvi da je $P(A) = \frac{1}{3}$, $P(B) = \frac{1}{4}$ i $P(A \cup B) = \frac{1}{2}$. Izračunajte vjerojatnosti $P(A \cap B)$, $P(\overline{A})$, $P(\overline{B})$, $P(\overline{A} \cap \overline{B})$, $P(\overline{A} \cup \overline{B})$, $P(A \cap \overline{B})$, $P(\overline{A} \cap B)$.*

Rješenje: Iz formule za vjerojatnost unije dobivamo da je

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = \frac{1}{3} + \frac{1}{4} - \frac{1}{2} = \frac{4+3-6}{12} = \frac{1}{12}.$$

Zbog vjerojatnosti komplementa vrijedi

$$P(\overline{A}) = 1 - P(A) = \frac{2}{3} \quad \text{i} \quad P(\overline{B}) = 1 - P(B) = \frac{3}{4}.$$

Kod računanja sljedećih dviju vjerojatnosti koristimo de Morganove zakone:

$$\begin{aligned} P(\overline{A} \cap \overline{B}) &= P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - \frac{1}{2} = \frac{1}{2}, \\ P(\overline{A} \cup \overline{B}) &= P(\overline{A \cap B}) = 1 - P(A \cap B) = 1 - \frac{1}{12} = \frac{11}{12}. \end{aligned}$$

Lagano vidimo, pomoću Venn-Eulerovih dijagrama, kako su događaji $A \cap B$ i $A \cap \overline{B}$ disjunktni te da je $(A \cap B) \cup (A \cap \overline{B}) = A$. Zbog toga je

$$P(A \cap \overline{B}) = P(A) - P(A \cap B) = \frac{1}{3} - \frac{1}{12} = \frac{1}{4}.$$

Slično je i

$$P(\overline{A} \cap B) = P(B) - P(A \cap B) = \frac{1}{4} - \frac{1}{12} = \frac{1}{6}.$$

□

2.3 Konačni vjerojatnosni prostor

U ovoj točki ćemo aksiomatsku definiciju vjerojatnosti primijeniti na vjerojatnosni prostor Ω koji se sastoji od konačno mnogo elementarnih događaja. Takav prostor nazivamo **konačni vjerojatnosni prostor**. Označimo njegove elemente s $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$. Događaj u ovakvom prostoru je svaki podskup od Ω. Vjerojatnost bilo kojeg događaja moći ćemo odrediti ako znamo vjerojatnosti elementarnih događaja, tj. ako poznajemo brojeve

$$p_1 = P(\{\omega_1\}), \quad p_2 = P(\{\omega_2\}), \quad \ldots, \quad p_n = P(\{\omega_n\}).$$

Očito je $p_1 > 0$, $p_2 > 0$, ..., $p_n > 0$. Nadalje, kako je $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$, a elementarni događaji su međusobno disjunktni, zbog aditivnosti i normiranosti vjerojatnosti vrijedi relacija $P(\{\omega_1\}) + P(\{\omega_2\}) + \cdots + P(\{\omega_n\}) = P(\Omega) = 1$.
Prema tome, pozitivni brojevi $p_1, p_2, \dots, p_n$ imaju svojstvo

$$p_1 + p_2 + \cdots + p_n = 1.$$

Neka je $A \in \mathcal{F}$ bilo koji događaj. On se sastoji od nekoliko elementarnih događaja, recimo m:

$$A = \{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_m}\}.$$

Vjerojatnost događaja A računamo tako da zbrojimo vjerojatnosti tih elementarnih događaja:

$$P(A) = p_{i_1} + p_{i_2} + \cdots p_{i_m}.$$

Prethodna razmatranja promotrit ćemo na nekim uobičajenim modelima konačnih vjerojatnosnih prostora.

Bacanje simetričnog novčića ili kocke

Kod bacanja simetričnog novčića imamo dva elementarna događaja ω_1 = P, ω_2 = G. Pod simetričnim novčićem podrazumijevamo ispravan novčić kod kojeg je način bacanja uobičajen, pa je prirodno pretpostaviti da su vjerojatnosti pojavljivanja obaju događaja jednake:

$$p_1 = P(\{\omega_1\}) = \frac{1}{2}, \quad p_2 = P(\{\omega_2\}) = \frac{1}{2}.$$

Slično, za simetričnu kocku prirodno je uzeti $p_i = P(\{\omega_i\}) = \frac{1}{6}$, za svaku od šest mogućnosti na koje kocka može pasti.

Primjer 2.8. *Bacamo simetričnu kocku. Odredite vjerojatnosti sljedećih događaja:*

$$\begin{aligned} A &= \{\textit{pao je neparan broj}\} \\ B &= \{\textit{pao je broj manji od } 5\}. \end{aligned}$$

Rješenje: Označimo li $\Omega = \{1, 2, 3, 4, 5, 6\}$, onda imamo da je

$$\begin{aligned} P(A) &= P(\{1,3,5\}) = 3 \cdot \frac{1}{6} = \frac{1}{2} \\ P(B) &= P(\{1,2,3,4\}) = 4 \cdot \frac{1}{6} = \frac{2}{3}. \end{aligned}$$

□

Ponekad ćemo izostavljati riječ simetričan te pretpostavljati da su novčić i kocka simetrični, ukoliko nije drukčije naglašeno.

Bacanje nekoliko novčića ili kocki

Kod bacanja dvaju novčića imamo četiri elementarna događaja:

	1. novčić	2. novčić
ω_1	P	P
ω_2	P	G
ω_3	G	P
ω_4	G	G

Kako bismo lakše razlikovali elementarne događaje ω_2 i ω_3, možemo zamisliti da bacamo dva različita novčića ili da jedan novčić bacamo dva puta te da razlikujemo bacanja. Sva četiri elementarna događaja su jednako vjerojatna, tj. vrijedi

$$p_1 = P(\{\omega_1\}) = \frac{1}{4}, \quad p_2 = P(\{\omega_2\}) = \frac{1}{4}, \quad p_3 = P(\{\omega_3\}) = \frac{1}{4}, \quad p_4 = P(\{\omega_4\}) = \frac{1}{4}.$$

Slično, kod bacanja dviju kocki imamo 36 elementarnih događaja. Da bismo razlikovali događaje poput $(3,4)$ i $(4,3)$, možemo zamisliti da su kocke obojene različitim bojama ili da umjesto dvije kocke istovremeno, bacamo jednu kocku dva puta tako da znamo što je rezultat prvog, a što rezultat drugog bacanja. Ako su kocke simetrične, prirodno je pretpostaviti da su svi elementarni događaji jednako vjerojatni te da vjerojatnost njihovog pojavljivanja iznosi $1/36$. Dakako, prethodna razmatranja lagano možemo proširiti na bilo koji broj novčića ili kocki.

Primjer 2.9. *Bacamo dvije simetrične kocke. Kolika je vjerojatnost da je zbroj brojeva na kockama jednak* 8*?*

Rješenje: Vjerojatnosni prostor Ω sastoji se od 36 jednako vjerojatnih elementarnih događaja. Traženi događaj A sastoji se od 5 elementarnih događaja, tj. $A = \{(2,6),(3,5),(4,4),(5,3),(6,2)\}$. Zbog toga je

$$P(A) = 5 \cdot \frac{1}{36} = \frac{5}{36}.$$

□

Uočimo, u primjerima koje smo do sada promatrali, svi elementarni događaji bili su jednako vjerojatni. Dakako, jednako vjerojatni elementarni događaji javljaju se i u drugim modelima vjerojatnosnih prostora.

Primjer 2.10. *Slučajno odabiremo broj iz skupa* $\{1,2,3,\ldots,89,90\}$*. Kolika je vjerojatnost da je odabrani broj djeljiv s* 3*? Kolika je vjerojatnost da je odabrani broj djeljiv s* 5*? Kolika je vjerojatnost da je odabrani broj djeljiv s* 3 *ili s* 5*?*

Rješenje: Označimo s A, B, C događaje:

$$\begin{aligned}
A &= \{\text{odabran je broj djeljiv s } 3\},\\
B &= \{\text{odabran je broj djeljiv s } 5\},\\
C &= \{\text{odabran je broj djeljiv s } 3 \text{ ili s } 5\}.
\end{aligned}$$

Kako je svaki treći broj djeljiv s 3, u skupu $\{1,2,3,\ldots,89,90\}$ ima 30 brojeva djeljivih s 3. Vjerojatnost odabira svakog od njih jednaka je $\frac{1}{90}$, pa je vjerojatnost događaja A jednaka

$$P(A) = \frac{30}{90} = \frac{1}{3}.$$

Slično, u skupu $\{1,2,3,\ldots,89,90\}$ ima $\frac{90}{5} = 18$ brojeva djeljivih s 5, pa je

$$P(B) = \frac{18}{90} = \frac{1}{5}.$$

Treći događaj C je unija prvih dvaju. Prvi dojam da je broj povoljnih ishoda $48 = 30 + 18$ je pogrešan, jer događaji A i B nisu disjunktni. Njihov presjek čine brojevi djeljivi s 3 i s 5. Očito, to su brojevi djeljivi s 15, odnosno $15, 30, 45, 60, 75, 90$. Zato je

$$P(C) = \frac{48-6}{90} = \frac{42}{90} = \frac{7}{15}.$$

Primijetimo da je ovdje $P(A\cap B) = \dfrac{6}{90} = \dfrac{1}{15}$, pa isti rezultat dobivamo iz formule za vjerojatnost unije:

$$P(C) = P(A\cup B) = P(A) + P(B) - P(A\cap B) = \frac{1}{3} + \frac{1}{5} - \frac{1}{15} = \frac{5+3-1}{15} = \frac{7}{15}.$$

□

U sljedećoj točki ponovno ćemo se vratiti na prostore s jednako vjerojatnim elementarnim događajima, a u nastavku ćemo proučiti nekoliko primjera vjerojatnosnih prostora kod kojih elementarni događaji nisu jednako vjerojatni.

Modeli s elementarnim događajima koji nisu jednako vjerojatni

Pretpostavimo da je novčić načinjen od dva različita materijala. Tada on nije simetričan pa, na primjer, vjerojatnost pojavljivanja pisma može biti veća od vjerojatnosti pojavljivanja glave.

Nadalje, slučajnom pokusu u kojemu bacamo dva novčića možemo pridružiti i vjerojatnosni prostor koji se sastoji od tri elementarna događaja:

$$\begin{aligned}
\omega_1 &= \{\text{pala su dva pisma}\},\\
\omega_2 &= \{\text{palo je jedno pismo i jedna glava}\},\\
\omega_3 &= \{\text{pale su dvije glave}\}
\end{aligned}$$

Ovaj pristup je također ispravan. Međutim, vjerojatnosti ovih elementarnih događaja nisu jednake, nego mora biti

$$P(\{\omega_1\}) = \frac{1}{4}, \quad P(\{\omega_2\}) = \frac{1}{2}, \quad P(\{\omega_3\}) = \frac{1}{4}.$$

Promotrimo sada još neke vjerojatnosne prostore kod kojih elementarni događaji nisu jednako vjerojatni.

Primjer 2.11. *U košari se nalazi jedna jabuka, dvije kruške, dvije naranče i jedna mandarina. Ivo slučajno odabire jednu voćku iz košare. Odredite vjerojatnost sljedećih događaja:*

$$\begin{aligned}
A &= \{\textit{nije odabrana kruška}\},\\
B &= \{\textit{odabrana je kruška, naranča ili mandarina}\},\\
C &= \{\textit{odabran je limun}\}.
\end{aligned}$$

Pretpostavljamo da je vjerojatnost odabira bilo koje voćke jednako vjerojatna.

Rješenje: Zadani slučajni pokus ima četiri moguća ishoda: $\omega_1 = J$, $\omega_2 = K$, $\omega_3 = N$ i $\omega_4 = M$. Ako pretpostavimo da je vjerojatnost odabira bilo koje voćke jednako vjerojatna, onda je razumno pridijeliti ovim elementarnim događajima vjerojatnosti

$$p_1 = P(\{\omega_1\}) = \frac{1}{6},\ p_2 = P(\{\omega_2\}) = \frac{1}{3},\ p_3 = P(\{\omega_3\}) = \frac{1}{3},\ p_4 = P(\{\omega_4\}) = \frac{1}{6}.$$

Događaju A odgovaraju sljedeći elementarni događaji: $A = \{\omega_1, \omega_3, \omega_4\}$, te je

$$P(A) = p_1 + p_3 + p_4 = \frac{1}{6} + \frac{1}{3} + \frac{1}{6} = \frac{2}{3}.$$

Mogli smo računati pomoću suprotnog događaja: $\overline{A} = \{\omega_2\}$:

$$P(A) = 1 - P(\overline{A}) = 1 - p_2 = 1 - \frac{1}{3} = \frac{2}{3}.$$

Događaju B odgovaraju elementarni događaji: $B = \{\omega_2, \omega_3, \omega_4\}$, te je

$$P(B) = p_2 + p_3 + p_4 = \frac{1}{3} + \frac{1}{3} + \frac{1}{6} = \frac{5}{6}.$$

Događaj C za ovu košaru nije moguć zato jer se u njoj ne nalazi niti jedan limun. Zbog toga je $P(C) = 0$. □

Primjer 2.12. *Dva prijatelja, Mirko i Slavko igraju stolni tenis. Slavko je nešto bolji stolnotenisač i u pravilu, od pet mečeva on pobjeđuje Mirka u tri meča, pa možemo pretpostaviti da je vjerojatnost pobjede Slavka u svakom pojedinom meču jednaka* 0.6. *Mirko i Slavko su odlučili igrati mečeve zaredom sve dok Mirko ne pobijedi, ali će odigrati najviše pet mečeva. Opišite vjerojatnosni prostor te izračunajte vjerojatnosti sljedećih događaja*

$$\begin{aligned} A &= \{\textit{Mirko će pobijediti Slavka u prvom ili drugom meču}\}, \\ B &= \{\textit{Mirko će pobijediti Slavka nakon drugog meča}\}, \\ C &= \{\textit{Mirko neće pobijediti Slavka}\}. \end{aligned}$$

Rješenje: Skup Ω sastoji se od 6 elementarnih događaja. Kako su pobjede Mirka i Slavka disjunktni događaji, vjerojatnost pobjede Mirka u pojedinom meču iznosi 0.4, uz pretpostavku da je vjerojatnost pobjede Slavka jednaka 0.6. Stoga je prirodno svakom od 6 elementarnih događaja pridružiti sljedeće vjerojatnosti:

$$\begin{array}{ll} \omega_1 = \text{M} & p_1 = 0.4 \\ \omega_2 = \text{SM} & p_2 = 0.6 \cdot 0.4 = 0.24 \\ \omega_3 = \text{SSM} & p_3 = 0.6 \cdot 0.6 \cdot 0.4 = 0.144 \\ \omega_4 = \text{SSSM} & p_4 = 0.6 \cdot 0.6 \cdot 0.6 \cdot 0.4 = 0.0864 \\ \omega_5 = \text{SSSSM} & p_5 = 0.6 \cdot 0.6 \cdot 0.6 \cdot 0.6 \cdot 0.4 = 0.05184 \\ \omega_6 = \text{SSSSS} & p_5 = 0.6 \cdot 0.6 \cdot 0.6 \cdot 0.6 \cdot 0.6 = 0.07776. \end{array}$$

Očito vrijedi $A = \{\omega_1, \omega_2\}$, $B = \{\omega_3, \omega_4, \omega_5\}$, $C = \{\omega_6\}$. Stoga ćemo vjerojatnosti događaja A, B i C dobiti tako da zbrojimo vjerojatnosti pripadnih elementarnih događaja:

$$\begin{aligned} P(A) &= p_1 + p_2 = 0.4 + 0.24 = 0.64, \\ P(B) &= p_3 + p_4 + p_5 = 0.144 + 0.0864 + 0.05184 = 0.28224, \\ P(C) &= p_6 = 0.07776. \end{aligned}$$

□

2.4 Klasični vjerojatnosni prostor

U prethodnoj točki promatrali smo slučajne pokuse bacanja simetričnog novčića ili kocke. Kod tih pokusa svi su elementarni događaji bili jednako vjerojatni. Takvi su još, na primjer, slučajni pokusi izvlačenja karte iz snopa ili izvlačenja broja u lutriji. U ovoj točki promatrat ćemo slučajne pokuse s jednako vjerojatnim elementarnim događajima.
Neka je $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ skup svih elementarnih događaja i $p_1, p_2, \ldots, p_n$ pripadne vjerojatnosti. Kako su svi ti brojevi jednaki, a njihov je zbroj 1, vrijedi

$$p_i = P(\{\omega_i\}) = \frac{1}{n}, \quad i = 1, 2, \ldots, n.$$

Ovakav vjerojatnosni prostor nazivamo klasični vjerojatnosni prostor jer se problemi iz kojih je potekla teorija vjerojatnosti mogu opisati ovim modelom.

Neka je $A \subset \Omega$ bilo koji događaj. Da bismo izračunali vjerojatnost događaja A, nije nam više potrebno znati koje elementarne događaje A sadrži, nego samo njihov broj. Naime, ako A sadrži m elementarnih događaja, $A = \{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_m}\}$, tada je

$$P(A) = p_{i_1} + p_{i_2} + \cdots + p_{i_m} = m \cdot \frac{1}{n} = \frac{m}{n}.$$

Ovu formulu možemo interpretirati na sljedeći način: Svaki elementarni događaj nazovimo mogućim ishodom (svi su jednako vjerojatni). Tako je

$$n = \text{broj svih mogućih ishoda.}$$

Elementarne događaje koji su sadržani u A nazovimo povoljnima za događaj A:

$$m = \text{broj svih povoljnih ishoda.}$$

Klasična vjerojatnost
U klasičnom vjerojatnosnom prostoru vjerojatnost događaja računa se formulom:

$$P(A) = \frac{m}{n} = \frac{\text{broj povoljnih ishoda}}{\text{broj mogućih ishoda}}$$

Računanje vjerojatnosti u klasičnom vjerojatnosnom prostoru povezano je s prebrojavanjem elemenata konačnih skupova, čime se bavi kombinatorika. Stoga je za rješavanje složenijih zadataka nužno poznavanje temeljnih pojmova iz kombinatorike, što smo naučili u prvom poglavlju.

Primjer 2.13. *Bacamo dvije simetrične kocke. Odredite skup elementarnih događaja. Koliko ima ukupno događaja u ovom pokusu? Nadalje, odredite koji je od sljedećih dvaju događaja vjerojatniji:*

$$\begin{aligned} A &= \{\textit{zbroj brojeva na obje kocke je najviše } 4\} \\ B &= \{\textit{zbroj brojeva na obje kocke je } 8\}. \end{aligned}$$

Rješenje: Skup elementarnih događaja sastoji se od uređenih parova:

$$\Omega = \{(\omega_i, \omega_j) : \omega_i, \omega_j \in \{1, 2, 3, 4, 5, 6\}\}.$$

Broj svih mogućih događaja je $|\Omega| = 36$. Nadalje, vrijedi $\mathcal{F} = \mathcal{P}(\Omega)$, pa je broj svih mogućih događaja $|\mathcal{F}| = 2^{36}$. Zbog simetrije vrijedi

$$P\{(\omega_i, \omega_j)\} = \frac{1}{36}.$$

Odredimo sada sve elementarne događaje od kojih su sastavljeni događaji A i B. Kod događaja A je zbroj na kockama jednak 2,3 ili 4, pa je $A = \{(1,1), (1,2), (2,1), (1,3), (2,2), (3,1)\}$. Dakle,

$$P(A) = \frac{6}{36} = \frac{1}{6}.$$

S druge strane, $B = \{(2,6), (3,5), (4,4), (5,3), (6,2)\}$, pa je

$$P(B) = \frac{5}{36}.$$

Konačno, kako je

$$P(A) = \frac{1}{6} > \frac{5}{36} = P(B),$$

slijedi da je događaj A vjerojatniji od događaja B. □

Primjer 2.14. *Istovremeno bacamo četiri igraće kocke. Izračunajte vjerojatnost da*

(a) se pojavila barem jedna šestica,

(b) su se pojavile točno dvije šestice,

(c) sve kocke pokazuju različite brojeve,

(d) dvije kocke pokazuju šestice, a dvije jedinice.

Rješenje: Kako svaka kocka može pasti na 6 brojeva, broj svih elementarnih, odnosno povoljnih događaja je $6^4 = 1296$.

(a) Ovdje koristimo formulu za vjerojatnost komplementa zato jer je lakše prebrojiti sva bacanja u kojima se nije pojavila niti jedna šestica. Broj takvih bacanja je $5^4 = 625$ jer tu imamo po pet mogućnosti za svaku kocku. Prema tome, vjerojatnost da se pojavila jedna šestica iznosi

$$1 - \frac{5^4}{6^4} = 1 - \frac{625}{1296} = \frac{671}{1296} \approx 0.52.$$

(b) U ovom dijelu zadatka povoljne mogućnosti su nizovi duljine 4 koji sadrže točno dvije šestice. Takvih nizova ima $\binom{4}{2} \cdot 5^2 = 150$, zato jer dva mjesta na kojima su šestice možemo odabrati na $\binom{4}{2}$ načina, a svako od preostalih dvaju mjesta popuniti na 5 načina. Tražena vjerojatnost je

$$\frac{\binom{4}{2} \cdot 5^2}{6^4} = \frac{150}{1296} \approx 0.12.$$

(c) Povoljne mogućnosti su nizovi duljine 4 s različitim znamenkama iz skupa $\{1, 2, 3, 4, 5, 6\}$. Njih ima $6 \cdot 5 \cdot 4 \cdot 3 = 360$, zato jer prvo mjesto možemo popuniti na 6 načina, drugo na 5, treće na 4 i četvrto na preostala 3 načina. Stoga je tražena vjerojatnost jednaka

$$\frac{6 \cdot 5 \cdot 4 \cdot 3}{6^4} = \frac{60}{216} \approx 0.28.$$

(d) Ovdje su povoljne mogućnosti nizovi duljine 4 sastavljeni od točno dvije jedinice i dvije šestice. Iz prethodnog poglavlja znamo da su to permutacije s ponavljanjem, pa je njihov broj jednak $\frac{4!}{2!2!} = 6$. Konačno, vjerojatnost iznosi

$$\frac{4!}{2! \cdot 2! \cdot 6^4} = \frac{6}{1296} \approx 0.005.$$

$\square$

Uočimo kako smo u prethodnom zadatku rezultate zaokruživali na dva ili tri decimalna mjesta. U zadatcima koje ćemo rješavati izbjegavat ćemo oznaku za približnu vrijednost i naprosto pisati znak jednakosti. Nadalje, ukoliko budemo imali velike brojeve, često ćemo rezultat ostavljati u obliku razlomaka koji sadrže binomne koeficijente, faktorijele itd.

Primjer 2.15. *Kolika je vjerojatnost da će igrač koji je zaokružio jednu kombinaciju u igri LOTO* 7 *od* 39 *pogoditi* *(a) svih sedam brojeva,*

(b) tri broja,

(c) barem dva broja?

Rješenje:

(a) Različitih kombinacija ima $\binom{39}{7}$. Povoljna je samo jedna kombinacija zato jer je $\binom{7}{7} = 1$. Prema tome, vjerojatnost pogotka svih sedam brojeva iznosi

$$p_7 = \frac{1}{\binom{39}{7}} = \frac{1}{15380937} = 6.5 \cdot 10^{-8}.$$

(b) Tri broja između sedam izvučenih možemo odabrati na $\binom{7}{3}$ načina, a četiri broja iz skupa od 32 broja koji nisu izvučeni na $\binom{32}{4}$ načina. Zato je vjerojatnost dobitka od tri pogotka jednaka

$$p_3 = \frac{\binom{7}{3}\binom{32}{4}}{\binom{39}{7}} = \frac{1258600}{15380937} = 0.082.$$

(c) Ako igrač pogodi barem 2 broja znači da je pogodio 2, 3, 4, 5, 6 ili 7 brojeva. Stoga nam je ovdje lakše računati pomoću komplementarnog događaja. Ako je događaj $A = \{\text{igrač je pogodio barem dva broja}\}$, onda je $\overline{A} = \{\text{igrač je pogodio jedan ili nijedan broj}\}$. Slično kao pod b) vjerojatnost pogotka jednog broja je

$$p_1 = \frac{\binom{7}{1}\binom{32}{6}}{\binom{39}{7}} = \frac{906192}{15380937} = 5.89 \cdot 10^{-2},$$

a niti jednog broja

$$p_0 = \frac{\binom{7}{0}\binom{32}{7}}{\binom{39}{7}} = 0.219.$$

Zbog toga je $P(\overline{A}) = 5.89 \cdot 10^{-2} + 0.219 = 0.278$, pa je vjerojatnost pogotka barem dvaju brojeva $P(A) = 1 - P(\overline{A}) = 1 - 0.278 = 0.722$.

□

Kako bismo što bolje razumjeli model iz prethodnog primjera, pogledajmo još jedan primjer s kartama.

Primjer 2.16. *Iz snopa od* 52 *karte izvlačimo odjednom* 5 *karata. Kolika je vjerojatnost da među izvučenim kartama ima*

(a) 4 *asa,*

(b) 3 *asa i* 2 *kralja,*

(c) 1 *as,* 1 *kralj,* 1 *dama,* 1 *dečko i* 1 *desetka,*

(d) barem jedan as?

Rješenje: Broj mogućih ishoda, odnosno broj načina na koje možemo izvući 5 karata je $n = \binom{52}{5}$.

(a) Četiri asa možemo odabrati na $\binom{4}{4}$ načina, a petu kartu od preostalih 48 karata na $\binom{48}{1}$ načina, odnosno $m = \binom{4}{4}\binom{48}{1}$. Zato je

$$p = \frac{m}{n} = \frac{\binom{4}{4}\binom{48}{1}}{\binom{52}{5}} = 1.84 \cdot 10^{-5}.$$

(b) Kako imamo po četiri asa i kralja, broj povoljnih mogućnosti je $m = \binom{4}{3}\binom{4}{2}$, pa je

$$p = \frac{m}{n} = \frac{\binom{4}{3}\binom{4}{2}}{\binom{52}{5}} = 9.23 \cdot 10^{-6}.$$

(c) Ima ukupno po četiri asa, kralja, dame, dečka i desetke, pa je broj povoljnih mogućnosti $m = \binom{4}{1}^5$. Zato je

$$p = \frac{m}{n} = \frac{\binom{4}{1}^5}{\binom{52}{5}} = 3.94 \cdot 10^{-4}.$$

(d) Ovdje vjerojatnost računamo preko suprotnog događaja. Broj izvlačenja u kojima nema niti jednog asa jednak je $\binom{48}{5}$, jer odabiremo iz skupa karata koje nisu asevi. Zato je tražena vjerojatnost jednaka

$$p = 1 - \frac{\binom{48}{5}}{\binom{52}{5}} = 0.34.$$

□

Primjer 2.17. *U skupini od deset ljudi, šest ih ima plave oči, a četiri smeđe oči. Biramo na sreću tri osobe. Odredite vjerojatnosti sljedećih događaja:*

$$\begin{aligned} A &= \{\textit{sve tri osobe imaju plave oči}\} \\ B &= \{\textit{sve tri osobe imaju istu boju očiju}\} \\ C &= \{\textit{dvije osobe imaju plave, a jedna smeđe oči}\} \end{aligned}$$

Rješenje: Tri osobe iz skupa od 10 ljudi možemo odabrati na $n = \binom{10}{3}$ načina. Povoljnih za događaj A je $m = \binom{6}{3}$, jer se na toliko načina mogu odabrati tri osobe iz skupa od šest osoba s plavim očima. Zato je

$$P(A) = \frac{m}{n} = \frac{\binom{6}{3}}{\binom{10}{3}} = \frac{20}{120} = \frac{1}{6}.$$

Broj povoljnih ishoda za događaj B je

$$m = \binom{6}{3} + \binom{4}{3}$$

jer boja očiju odabranih osoba može biti plava ili smeđa. Zato je

$$P(B) = \frac{m}{n} = \frac{\binom{6}{3} + \binom{4}{3}}{\binom{10}{3}} = \frac{24}{120} = \frac{1}{5}.$$

Odredimo broj povoljnih ishoda za događaj C. Dvije osobe s plavim očima možemo odabrati na $\binom{6}{2}$ načina, a jednu sa smeđim na $\binom{4}{1}$ načina:

$$m = \binom{6}{2}\binom{4}{1} = 60 \implies P(C) = \frac{60}{120} = \frac{1}{2}.$$

□

Primjer 2.18. *Kolika je vjerojatnost da u skupini od* 5 *osoba barem dvije slave rođendan istog dana?*

Rješenje: Iskoristit ćemo komplementarni događaj jer nam je lakše odrediti vjerojatnost da svih pet osoba slavi rođendan u različitim danima.

Broj svih mogućosti za rođendane jednak je $n = 365^5$, zato jer svaka osoba može imati rođendan bilo kojeg od 365 dana u godini. Nadalje, broj mogućnosti tako da svaka osoba ima rođendan različitog dana u godini jednak je $m = 365 \cdot 364 \cdot 363 \cdot 362 \cdot 361$, zato jer dan za prvu osobu možemo odabrati na 365 načina, za drugu osobu na 364 načina, itd.

Konačno, prema formuli za suprotnu vjerojatnost imamo da je

$$p = 1 - \frac{m}{n} = 1 - \frac{365 \cdot 364 \cdot 363 \cdot 362 \cdot 361}{365^5} = 0.027.$$

□

Primjer 2.19. *Lokalni autobus Omiš-Split polazi s autobusnog kolodvora u Omišu i ima stajališta na* 10 *stanica, uključujući i posljednju u Splitu. U Omišu se u autobus ukrcalo* 9 *putnika. Izračunajte vjerojatnosti da*

(a) svih 9 *putnika izađe na različitim stanicama,*

(b) jedan od odabranih putnika (Ivan) izađe na posljednjoj stanici,

(c) točno 3 *putnika izađu na posljednjoj stanici.*

Pretpostavljamo da svaki od 9 *putnika s jednakom vjerojatnošcu izlazi na svakoj od* 10 *stanica te ne promatramo ostale putnike koji ulaze u autobus.*

Rješenje: Broj mogućih ishoda jednak je $n = 10^9$ zato jer svaki putnik može izaći na bilo kojoj od 10 stanica.

(a) Povoljnih mogućnosti ima $m = 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 = 10!$, zato jer prvi putnik može izaći na 10 načina, drugi na 9, treći na 8, ..., te posljednji na preostala 2 načina. Zbog toga je

$$p = \frac{m}{n} = \frac{10!}{10^9} = 0.004.$$

(b) Ivanova stanica je određena (to je posljednja), preostalih osam je slobodno. Stoga je $m = 10^8$, pa je

$$p = \frac{m}{n} = \frac{10^8}{10^9} = 0.1.$$

(c) Očito, 3 putnika koji izlaze na posljednjoj stanici možemo odabrati na na $\binom{9}{3}$ načina. Preostalih 6 putnika može izaći na bilo kojoj od preostalih 9 stanica. Zato je $m = \binom{9}{3} \cdot 9^6$, pa tražena vjerojatnost iznosi

$$p = \frac{m}{n} = \frac{\binom{9}{3} \cdot 9^6}{10^9} = 0.045.$$

□

2.5 Neki primjeri beskonačnih vjerojatnosnih prostora

U prethodnim točkama smo promatrali različite vjerojatnosne modele i naučili mnogo o svojstvima vjerojatnosti. Zajedničko svojstvo u svim primjerima koje smo promatrali je bilo to što je skup svih elementarnih događaja bio konačan.

Međutim još uvijek nismo u stanju odgovoriti na mnoga pitanja vezana uz vrlo jednostavne modele. Na primjer, bacamo igraću kocku sve dok se ne pojavi šestica. Kolika je vjerojatnost da se šestica nikad neće pojaviti? Ili, slučajno odabiremo točku unutar jediničnog kvadrata. Kolika je vjerojatnost da ćemo odabrati njegovo središte? Zajedničko je svojstvo u oba ova pokusa to što je skup Ω elementarnih događaja beskonačan.

U prvom slučaju, broj bacanja u kojem se šestica može pojaviti je bilo koji prirodni broj, pa elementarnih događaja ima prebrojivo mnogo. Prisjetimo se, prebrojiv skup je skup koji ima beskonačno mnogo elemenata, te čije elemente možemo poredati u niz. Takvi su, na primjer, skup prirodnih, cijelih i racionalnih brojeva.

U drugom slučaju, elementarni događaj je izbor bilo koje točke unutar kvadrata. Tih događaja ima neprebrojivo mnogo. Prisjetimo se, bilo koji interval ili čitav skup realnih brojeva su primjeri neprebrojivih skupova. Njihove elemente ne možemo poredati u niz.

Čini se da je vjerojatnost događaja u oba spomenuta primjera jednaka nuli. No kako uopće možemo računati vjerojatnost u takvom prostoru? Za razliku od konačnog, u beskonačnom prostoru algebra događaja i vjerojatnost moraju zadovoljavati neka dodatna svojstva.

Prisjetimo se, u konačnom vjerojatnosnom prostoru izvodili smo operacije unije, presjeka i razlike s konačno mnogo događaja. U beskonačnom prostoru moramo izvoditi operacije s beskonačno mnogo događaja. Zbog toga ćemo zahtijevati da unija prebrojivo mnogo događaja također bude događaj. Takva algebra događaja naziva se $\sigma-$algebra. Također, funkcija vjerojatnosti u takvom prostoru mora zadovoljavati uvjet prebrojive aditivnosti. Ovdje ćemo se ograničiti na prebrojive vjerojatnosne prostore $\Omega = \{\omega_1, \omega_2, \ldots\}$ čiji će svi podskupovi biti događaji, te će navedeni uvjeti biti zadovoljeni.

Pogledajmo nekoliko primjera beskonačnih prebrojivih vjerojatnosnih prostora.

Primjer 2.20. *U kutiji se nalaze tri bijele i tri crne kuglice. Slučajno izvlačimo kuglice, jednu za drugom, sve dok ne izvučemo bijelu kuglicu. Opišite vjerojatnosni prostor, odredite elementarne događaje i pripadne vjerojatnosti u svakom od sljedeća dva načina izvlačenja:*

(a) nakon izvlačenja kuglica se vraća u kutiju

(b) izvučena kuglica ne vraća se natrag.

Rješenje:

(a) Vjerojatnosni prostor je beskonačan. Elementarni događaji i pripadne vjerojatnosti su:

$$\begin{array}{ll}
\omega_1 = \mathrm{B} & P(\omega_1) = 1/2, \\
\omega_2 = \mathrm{CB} & P(\omega_2) = 1/2 \cdot 1/2, \\
\omega_3 = \mathrm{CCB} & P(\omega_3) = 1/2 \cdot 1/2 \cdot 1/2, \\
\omega_4 = \mathrm{CCCB} & P(\omega_4) = 1/2 \cdot 1/2 \cdot 1/2 \cdot 1/2, \\
\vdots & \\
\omega_n = \mathrm{C} \cdots \mathrm{CB} & P(\omega_n) = (1/2)^n . \\
\vdots &
\end{array}$$

(b) U ovom je slučaju vjerojatnosni prostor konačan. Sastoji se od četiri elementarna događaja. Pripadne vjerojatnosti su redom jednake

$$\begin{array}{lll}
\omega_1 = \mathrm{B} & P(\omega_1) = 1/2 & = 1/2, \\
\omega_2 = \mathrm{CB} & P(\omega_2) = 1/2 \cdot 3/5 & = 3/10, \\
\omega_3 = \mathrm{CCB} & P(\omega_3) = 1/2 \cdot 2/5 \cdot 3/4 & = 3/20, \\
\omega_4 = \mathrm{CCCB} & P(\omega_4) = 1/2 \cdot 2/5 \cdot 1/4 \cdot 3/3 & = 1/20.
\end{array}$$

□

Primjer 2.21. *Simetričnu kocku bacamo sve dok se ne pojavi šestica.*

(a) Odredite elementarne događaje i njihove vjerojatnosti.

(b) Odredite vjerojatnost da se šestica pojavila u prva tri bacanja.

(c) Odredite vjerojatnost da se šestica nije pojavila u prva tri bacanja.

(d) Odredite vjerojatnost da se šestica uopće nije pojavila.

Rješenje: (a) Vjerojatnosni prostor je beskonačan. Elementarni događaji i pripadne vjerojatnosti su:

$$\begin{array}{ll}
\omega_1 = 6 & P(\omega_1) = 1/6, \\
\omega_2 = \overline{6}6 & P(\omega_2) = 5/6 \cdot 1/6, \\
\omega_3 = \overline{66}6 & P(\omega_3) = 5/6 \cdot 5/6 \cdot 1/6, \\
\omega_4 = \overline{666}6 & P(\omega_4) = 5/6 \cdot 5/6 \cdot 5/6 \cdot 1/6, \\
\vdots & \\
\omega_n = \overline{6 \cdots 6}6 & P(\omega_n) = (5/6)^{n-1} \cdot 1/6. \\
\vdots &
\end{array}$$

(b) Neka je $A = \{$šestica se pojavila u prva tri bacanja$\}$. Tada je $A = \{\omega_1, \omega_2, \omega_3\}$, pa je

$$P(A) = P(\omega_1) + P(\omega_2) + P(\omega_3) = \frac{1}{6} + \frac{5}{36} + \frac{25}{216} = \frac{91}{216}.$$

(c) Ako se šestica nije pojavila u prva tri bacanja, onda tražimo vjerojatnost događaja $\overline{A} = \{\omega_4, \omega_5, \omega_6, \ldots\}$. Prema pravilu komplementa vrijedi

$$P(\overline{A}) = 1 - P(A) = 1 - \frac{91}{216} = \frac{125}{216}.$$

Isti rezultat mogli smo dobiti i direktno, odnosno

$$P(\overline{A}) = \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} = \frac{125}{216}.$$

Razjasnimo zašto je i ovaj način rješavanja ispravan! Naime, kod događaja $\overline{A}$ važno je samo to da se šestica nije pojavila u prva tri bacanja, a što je bilo poslije, to za ovaj događaj nije važno.

(d) Neka je $B = \{$šestica se uopće nije pojavila$\}$. Pokažimo da je vjerojatnost događaja B jednaka nuli. U tu svrhu neka je $B_n = \{$šestica nije pala u prvih n bacanja$\}$. Iz (c) dijela zadatka vidimo da je $P(B_n) = (5/6)^n$. Očito je $B \subseteq B_n$ za svaki prirodan broj n, pa je

$$P(B) \leq \left(\frac{5}{6}\right)^n, \quad n \in \mathbf{N}.$$

Pustimo li sada da n teži u beskonačnost, vidimo da izraz $(5/6)^n$ teži k nuli, pa je $P(B) = 0$. $\square$

Prelaskom na limes, u prethodnom primjeru smo koristili takozvano svojstvo neprekinutosti vjerojatnosti. Pokazuje se da je to svojstvo ekvivalentno prebrojivoj aditivnosti vjerojatnosti. Dakle, u slučaju kada vjerojatnost zadovoljava uvjet prebrojive aditivnosti, možemo koristiti limes, što će i biti slučaj u problemima koje promatramo.

Primjer 2.22. *Simetričnu kocku bacamo sve dok se ne pojavi šestica. Odredite vjerojatnost da se šestica pojavi u neparnom bacanju.*

Rješenje: Ovdje imamo isti vjerojatnosni prostor kao i u prethodnom primjeru. Uz oznake iz tog primjera trebamo odrediti vjerojatnost događaja $C = \{\omega_1, \omega_3, \omega_5, \ldots\}$. Kako se događaj C sastoji od beskonačno mnogo elementarnih događaja, trebat ćemo izračunati beskonačnu sumu:

$$\begin{aligned} P(C) &= P(\omega_1) + P(\omega_3) + P(\omega_5) + \cdots \\ &= \frac{1}{6} + \frac{1}{6}\left(\frac{5}{6}\right)^2 + \frac{1}{6}\left(\frac{5}{6}\right)^4 + \cdots \\ &= \frac{1}{6} + \frac{1}{6}\left(\frac{25}{36}\right)^1 + \frac{1}{6}\left(\frac{25}{36}\right)^2 + \cdots \end{aligned}$$

U prethodnoj sumi prepoznajemo beskonačni konvergentni geometrijski red. Prisjetimo se formule za sumu tog reda:

$$a + aq + aq^2 + aq^3 + \cdots + aq^n + \cdots = \frac{a}{1-q}, \quad \text{gdje je} \quad |q| < 1.$$

U našem je slučaju $a = \dfrac{1}{6}$ i $q = \dfrac{25}{36}$, pa je

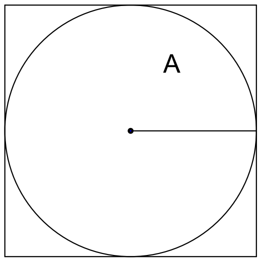
$$P(C) = \frac{1}{6} \cdot \frac{1}{1 - \frac{25}{36}} = \frac{6}{11}.$$

$\square$

2.6 Geometrijska vjerojatnost

U prethodnoj točki vidjeli smo primjere beskonačnih prebrojivih vjerojatnosnih prostora. Ovdje ćemo promatrati neke primjere prostora koji nisu prebrojivi.

Promotrimo pokus u kojem biramo na slučajan način točku unutar kvadrata Ω sa stranicom duljine 2 cm. Istaknimo jedan podskup tog kvadrata. Neka je A krug upisan u kvadrat Ω.



Biramo li točku unutar kvadrata, možemo se upitati kolika je vjerojatnost da će ta točka biti odabrana unutar kruga. U ovdje opisanom pokusu prirodno je traženoj vjerojatnosti pridružiti omjer površina kruga i kvadrata. Kako je polumjer upisanog kruga jednak polovini duljine stranice kvadrata, odnosno 1 cm, površina kruga iznosi $1^2\pi = \pi$ cm^2. Nadalje, površina kvadrata je $2^2 = 4$ cm^2, pa promatranom pokusu možemo pridružiti događaj A s pripadnom vjerojatnošću

$$P(A) = P\{\text{točka je pala u krug } A\} = \frac{\pi}{4} = 0.785.$$

Upitamo li se kolika je vjerojatnost da slučajno odabrana točka unutar kvadrata ne upadne u krug, trebat ćemo površinu dijela kvadrata koji je van kruga podijeliti s površinom samog kvadrata. Očito, površina dijela kvadrata koji je van kruga iznosi $(4-\pi)$ cm^2, pa možemo promatrati događaj $\overline{A}$ s pripadnom vjerojatnošću

$$P(\overline{A}) = P\{\text{točka nije pala u krug } A\} = \frac{4-\pi}{4} = 0.215.$$

Dakle, vjerojatnosti smo dobili promatrajući omjer površina podskupova i čitavog kvadrata. Takva situacija vrijedi i općenitije, pa pogledajmo stoga definiciju geometrijske vjerojatnosti.

Geometrijska vjerojatnost
Pretpostavimo da slučajno odabiremo točku unutar ograničenog skupa Ω u prostoru. Tada je vjerojatnost da točka padne unutar nekog podskupa $A \subseteq \Omega$ jednaka

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

gdje m označava mjeru odgovarajućeg skupa. Pri tome je mjera duljina, površina ili obujam danog skupa točaka. Ovako definiranu vjerojatnost nazivamo geometrijska vjerojatnost.

Prethodnom formulom je doista definirana vjerojatnost. Provjerimo jesu li ispunjena svojstva $1^\circ - 3^\circ$ iz definicije vjerojatnosti.
1°. U geometrijskoj vjerojatnosti nemoguć događaj je izbor točke unutar praznog skupa. Mjera praznog skupa je 0, pa je

$$P(\emptyset) = \frac{m(\emptyset)}{m(\Omega)} = 0.$$

Nadalje vjerojatnost odabira točke unutar čitavog skupa Ω je siguran događaj pa je

$$P(\Omega) = \frac{m(\Omega)}{m(\Omega)} = 1.$$

2°. Ako su A i B podskupovi od Ω takvi da je $A \subset B$, onda je $m(A) \leq m(B)$. Zato je

$$P(A) = \frac{m(A)}{m(\Omega)} \leq \frac{m(B)}{m(\Omega)} = P(B).$$

3°. Ako su A i B disjunktni podskupovi od Ω, onda je mjera njihove unije jednaka zbroju mjera pojedinih skupova. Zato je vjerojatnost da točka bude izabrana unutar jednog od podskupova jednaka:

$$P(A \cup B) = \frac{m(A \cup B)}{m(\Omega)} = \frac{m(A)}{m(\Omega)} + \frac{m(B)}{m(\Omega)} = P(A) + P(B).$$

U primjerima koji slijede trebat ćemo samostalno modelirati problemske zadatke pomoću geometrijske vjerojatnosti.

Primjer 2.23. *Unutar intervala $[0, 1]$ biraju se na sreću dva broja x i y. Odredite vjerojatnost*

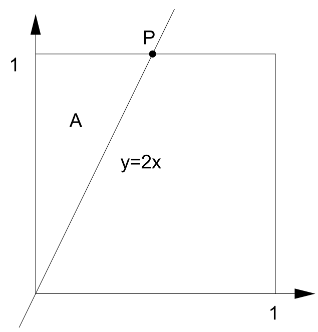
(a) da broj y bude veći od dvostrukog broja x,

(b) da zbroj brojeva x i y bude manji od $\frac{5}{4}$,

(c) da brojevi x i y budu jednaki.

Rješenje: Izbor dvaju brojeva x i y unutar intervala $[0, 1]$ odgovara izboru jedne točke (x, y) unutar jediničnog kvadrata $[0, 1] \times [0, 1]$. Označimo taj kvadrat s Ω. On predstavlja skup elementarnih događaja. Da bismo odredili tražene vjerojatnosti, moramo izračunati površinu podskupova od Ω koji odgovaraju tim događajima. Površina kvadrata je očito $m(\Omega) = 1^2 = 1$.

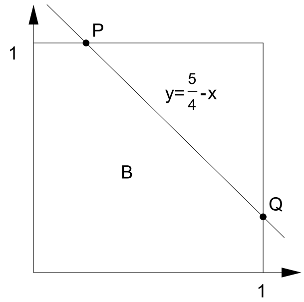
(a) Označimo s A traženi događaj. Tom događaju odgovara skup svih točaka jediničnog kvadrata za koje je $y > 2x$. To je skup svih točaka iznad pravca $y = 2x$ koje se nalaze unutar kvadrata Ω. Prema tome, skup A je pravokutni trokut kojemu je jedna stranica jednaka 1. Odredimo duljinu druge stranice. Neka je P točka presjeka pravca $y = 2x$ i kvadrata Ω, kao na slici. Kako je $y-$koordinata točke P jednaka 1, iz jednadžbe $y = 2x$ dobivamo da je $x-$koordinata te točke jednaka $\frac{1}{2}$. Zbog toga je površina trokuta jednaka $m(A) = \frac{1}{2} \cdot 1 \cdot \frac{1}{2} = \frac{1}{4}$.



Konačno, vjerojatnost događaja A je jednaka

$$P(A) = \frac{m(A)}{m(\Omega)} = \frac{\frac{1}{4}}{1} = \frac{1}{4}.$$

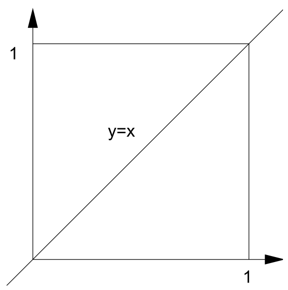
(b) Neka je B traženi događaj. Događaj B predstavlja skup svih točaka kvadrata Ω koje se nalaze ispod pravca $y = \frac{5}{4} - x$.



Odredimo sada koordinate sjecišta pravca $y = \frac{5}{4} - x$ i kvadrata Ω. Kako je $y-$koordinata točke P jednaka 1, iz jednadžbe pravca imamo da je njena $x-$koordinata jednaka $x = \frac{5}{4} - y = \frac{5}{4} - 1 = \frac{1}{4}$. Na isti način dobivamo koordinate točke Q, tj. $Q = \left(1, \frac{1}{4}\right)$. Zbog toga, površinu lika B računamo tako da od površine kvadrata oduzmemo površinu jednakokračnog pravokutnog trokuta s katetama duljine $\frac{3}{4}$. Imamo:

$$P(B) = \frac{m(B)}{m(\Omega)} = \frac{1 - \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{3}{4}}{1} = \frac{23}{32}.$$

(c) Skup svih točaka (x, y) za koje je $x = y$ je dijagonala kvadrata.



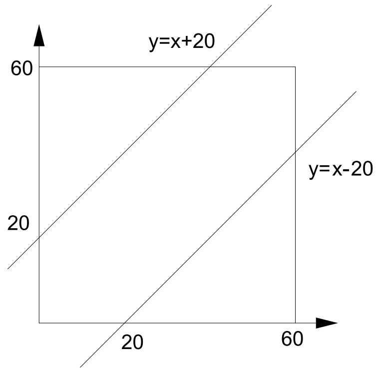
Površina takvog skupa je 0. Zato je i pripadna vjerojatnost jednaka nuli: birajući na sreću dva broja unutar intervala $[0, 1]$ vjerojatnost da ta dva broja budu jednaka jest 0. To ne znači da je ovaj događaj nemoguć, jer u vjerojatnosnom modelu s neprebrojivo mnogo ishoda (kao što je geometrijski model vjerojatnosti) razlikuju se pojmovi nemogućeg događaja i događaja s vjerojatnošću nula! Kao posljedica ovog, događaji $\{x < y\}$ i $\{x \leq y\}$ imaju jednaku vjerojatnost. Isto vrijedi i za slične skupove.

□

Primjer 2.24. *Maja i Ana izlaze navečer neovisno jedna o drugoj, u na sreću odabranom trenutku između* 21 *i* 22 *sata. Po dolasku u omiljeni kafić zadržavaju se na tom mjestu* 20 *minuta, ali najkasnije do* 22 *sata, kad odlaze u disko. Kolika je vjerojatnost da će se Maja i Ana sresti u kafiću?*

Rješenje: Neka je x trenutak dolaska Maje, a y trenutak dolaska Ane u kafić. Brojevi x i y su na sreću odabrani brojevi u intervalu $[0, 60]$ (vrijeme brojimo u minutama). Maja i Ana će se sresti ako je $|y - x| < 20$. Kako iz $|y - x| < 20$ slijedi da je $-20 < y - x < 20$ imamo

$$P\{|y - x| < 20\} = P\{-20 < y - x < 20\} = P\{x - 20 < y < x + 20\}.$$



Dakle, nacrtamo li u koordinatnom sustavu kvadrat i pravce $y = x - 20$ i $y = x + 20$ vidimo kako moramo izračunati površinu između paralelnih pravaca $y = x - 20$ i $y = x + 20$ koja se nalazi unutar kvadrata. Tu površinu računamo tako da od površine kvadrata oduzmemo površine dvaju jednakokračnih pravokutnih trokuta s katetama duljine 40 (vidi sliku). Dakle,

$$P\{|y - x| < 20\} = \frac{m(A)}{m(\Omega)} = \frac{60^2 - 40^2}{40^2} = 0.56.$$

□

Primjer 2.25. *Slučajno odabiremo dva broja unutar intervala* $[0, 1]$. *Kolika je vjerojatnost da oni zadovoljavaju sustav nejednadžbi*

$$\begin{aligned} x^2 + y^2 &\leq 2x \\ x^2 + y^2 &\leq 2y \text{ ?} \end{aligned}$$

Rješenje: Kao što smo već vidjeli, odabir dvaju brojeva unutar intervala $[0, 1]$ ekvivalentan je odabiru točke unutar jediničnog kvadrata

$$\Omega = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1\}.$$

Uočimo kako su danim nejednadžbama opisana područja unutar dviju kružnica. Naime, transformiranjem nejednadžbe $x^2 + y^2 \leq 2x$ i nadopunjavanjem na potpun kvadrat imamo redom

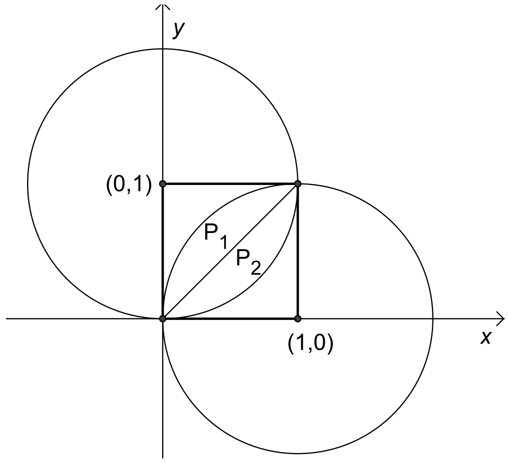
$$x^2 - 2x + y^2 \leq 0 \quad \Longrightarrow \quad x^2 - 2x + 1 + y^2 \leq 1 \quad \Longrightarrow (x - 1)^2 + y^2 \leq 1,$$

pa tu nejednadžbu zadovoljavaju sve točke koje leže unutar kružnice sa središtem u točki $(1, 0)$ i polumjerom 1. Slično, nejednadžba $x^2 + y^2 \leq 2y$ se svodi na oblik

$$x^2 + y^2 - 2y \leq 0 \quad \Longrightarrow \quad x^2 + y^2 - 2y + 1 \leq 1 \quad \Longrightarrow x^2 + (y - 1)^2 \leq 1,$$

pa je to područje unutar kružnice sa središtem u točki $(0, 1)$ i polumjerom 1.

Prema tome, dva broja x, y iz intervala $[0, 1]$ zadovoljavaju zadani sustav nejednadžbi ako leže unutar područja omeđenog kružnicama $(x - 1)^2 + y^2 = 1$ i $x^2 + (y - 1)^2 = 1$, kao na slici.



Uočimo kako se to područje sastoji od dva jednaka dijela P_1 i P_2. Površinu svakog od njih možemo dobiti kao razliku površine četvrtine kruga i polovine kvadrata, odnosno

$$m(P_1) = \frac{1^2\pi}{4} - \frac{1^2}{2} = \frac{\pi}{4} - \frac{1}{2}.$$

Označimo li s A traženi događaj, slijedi da je

$$m(A) = m(P_1) + m(P_2) = 2m(P_1) = \frac{\pi}{2} - 1,$$

odnosno

$$P(A) = \frac{m(A)}{M(\Omega)} = \frac{\frac{\pi}{2} - 1}{1} = \frac{\pi}{2} - 1 = 0.57.$$

□

2.7 Zadatci za ponavljanje

Zadatak 2.1. Neka su A i B događaji takvi da je $P(\overline{A}) = 0.6$, $P(\overline{B}) = 0.5$ i $P(\overline{A} \cup \overline{B}) = 0.8$. Izračunajte vjerojatnosti sljedećih događaja:

(a) $P(A \cap B)$,

(b) $P(A \cup B)$,

(c) $P(\overline{A} \cap B)$.

Rješenje:

(a) Iz formule za vjerojatnost komplementa i de Morganove formule imamo redom

$$P(A \cap B) = 1 - P(\overline{A \cap B}) = 1 - P(\overline{A} \cup \overline{B}) = 1 - 0.8 = 0.2.$$

(b) Kako je $P(A) = 1 - P(\overline{A}) = 1 - 0.6 = 0.4$, $P(B) = 1 - P(\overline{B}) = 1 - 0.5 = 0.5$, korištenjem formule za vjerojatnost unije dobivamo:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.4 + 0.5 - 0.2 = 0.7.$$

(c) Događaji $A \cap B$ i $\overline{A} \cap B$ su disjunktni, a njihova unija je događaj B. Zbog toga vrijedi

$$P(\overline{A} \cap B) = P(B) - P(A \cap B) = 0.5 - 0.2 = 0.3.$$

□

Zadatak 2.2. Tri žene i četiri muškarca stoje u redu na šalteru u banci. Kolika je vjerojatnost da su

(a) žene na prva tri mjesta u tom redu,

(b) žene jedna do druge u tom redu?

Rješenje:

(a) Broj povoljnih mogućnosti je $n = 7!$, jer permutiramo svih 7 ljudi. Kako se žene nalaze na prva tri mjesta, njih permutiramo na 3! načina. Muškarce permutiramo na preostala četiri mjesta, pa je broj povoljnih mogućnosti $m = 3! \cdot 4!$. Dakle, tražena vjerojatnost iznosi

$$p = \frac{3! \cdot 4!}{7!} = \frac{1}{35} = 0.029.$$

(b) Kako se žene nalaze jedna do druge, možemo ih shvatiti kao jedan blok. Dakle, permutiramo 5 objekata (4 muškarca i blok žena), tako da permutiramo i unutar jednog od objekata. Prema tome, broj povoljnih mogućnosti je $m = 5! \cdot 3!$, pa vjerojatnost iznosi

$$p = \frac{5! \cdot 3!}{7!} = \frac{1}{7} = 0.143.$$

□

Zadatak 2.3. Bačeno je 12 kocaka. Kolika je vjerojatnost da će se svaki od brojeva $1, 2, 3, 4, 5, 6$ pojaviti točno dvaput?

Rješenje: Rezultat pokusa je niz od 12 brojeva odabranih iz skupa $\{1, 2, 3, 4, 5, 6\}$. Na svakoj kocki može se pojaviti bilo koji od tih brojeva, pa je broj svih mogućih ishoda jednak $n = 6^{12}$.

S druge strane, povoljnih događaja ima isto koliko i različitih permutacija (s ponavljanjem) niza $\{1, 1, 2, 2, 3, 3, 4, 4, 5, 5, 6, 6\}$ jer svaka takva permutacija daje povoljan ishod bacanja:

$$m = P_{12}^{2,2,2,2,2,2} = \frac{12!}{(2!)^6}.$$

Tako dobivamo

$$p = \frac{m}{n} = \frac{12!}{(2!)^6 \cdot 6^{12}} = 0.00344.$$

□

Zadatak 2.4. Avion prevozi veliki kontejner s monitorima. U kontejneru je 800 monitora od kojih je 15 neispravno. Pri slijetanju aviona u zračnu luku inspekcija slučajnim odabirom uzima 10 monitora na pregled. Kolika je vjerojatnost da su među odabranima

(a) točno dva neispravna,

(b) barem tri neispravna?

Rješenje:

(a) Broj svih mogućnosti jednak je $n = \binom{800}{10}$, jer se od 800 monitora bira njih 10. Nadalje, od 15 neispravnih, 2 monitora odabiremo na $\binom{15}{2}$ načina, od preostalih 785 njih 8 biramo na $\binom{785}{8}$ načina. Zato je broj povoljnih mogućnosti jednak $m = \binom{785}{8}\binom{15}{2}$. Prema tome, tražena vjerojatnost je

$$p = \frac{m}{n} = \frac{\binom{785}{8}\binom{15}{2}}{\binom{800}{10}}.$$

(b) Traženu vjerojatnost odredit ćemo pomoću suprotnog događaja. Suprotan događaj predstavljaju odabiri u kojima su najviše dva monitora neispravna. Dakle, od svih odabira trebamo oduzeti one u kojima su 0,1 ili 2 neispravna monitora. Zbog toga vjerojatnost iznosi

$$1 - \frac{\binom{785}{10}}{\binom{800}{10}} - \frac{\binom{785}{9}\binom{15}{1}}{\binom{800}{10}} - \frac{\binom{785}{8}\binom{15}{2}}{\binom{800}{10}}.$$

Dakako, zbog velikih brojeva rješenje ćemo ostaviti u ovom obliku.

□

Zadatak 2.5. U urni se nalaze četiri bijele i tri crne kuglice. Izvlačimo na sreću po jednu kuglicu, bez vraćanja, sve dok ne izvučemo crnu kuglicu. Kolika je vjerojatnost da će pokus završiti u prva četiri izvlačenja?

Rješenje: U ovom slučaju je vjerojatnosni prostor konačan, zato jer će crna kuglica biti izvučena u najviše pet izvlačenja. Odredimo stoga sve elementarne događaje i pripadne vjerojatnosti.

$$\begin{array}{lll}
\omega_1 = \mathrm{C} & P(\omega_1) = 3/7 & = 3/7, \\
\omega_2 = \mathrm{BC} & P(\omega_2) = 4/7 \cdot 3/6 & = 2/7, \\
\omega_3 = \mathrm{BBC} & P(\omega_3) = 4/7 \cdot 3/6 \cdot 3/5 & = 6/35, \\
\omega_4 = \mathrm{BBBC} & P(\omega_4) = 4/7 \cdot 3/6 \cdot 2/5 \cdot 3/4 & = 3/35, \\
\omega_5 = \mathrm{BBBBC} & P(\omega_5) = 4/7 \cdot 3/6 \cdot 2/5 \cdot 1/4 \cdot 3/3 & = 1/35.
\end{array}$$

Kako tražimo vjerojatnost da pokus završi u prva četiri izvlačenja, moramo odrediti vjerojatnost događaja $A = \{\omega_1, \omega_2, \omega_3, \omega_4\}$. Konačno, kako su događaji A i $\{\omega_5\}$ komplementarni imamo da je

$$P(A) = 1 - P(\omega_5) = 1 - \frac{1}{35} = \frac{34}{35} = 0.971.$$

□

Zadatak 2.6. Bacamo igraću kocku sve dok se ne pojavi petica ili šestica. Kolika je vjerojatnost da pokus neće završiti u prva tri bacanja kocke?

Rješenje: Ovdje imamo beskonačan vjerojatnosni prostor. Neka je $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n, \ldots\}$, pri čemu ω_n označava elementarni događaj u kojemu je petica ili šestica postignuta u $n-$tom bacanju. Ako je $A = \{\omega_1, \omega_2, \omega_3\}$, onda je traženi događaj jednak komplementu od A, tj. tražimo $P(\overline{A})$.

Vjerojatnost pojave petice ili šestice u svakom pojedinom bacanju jednaka je $\frac{2}{6} = \frac{1}{3}$, pa je

$$\begin{aligned} P(\omega_1) &= \frac{1}{3} \\ P(\omega_2) &= \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9} \\ P(\omega_3) &= \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{4}{27}. \end{aligned}$$

Prema tome,

$$P(A) = P(\omega_1) + P(\omega_2) + P(\omega_3) = \frac{1}{3} + \frac{2}{9} + \frac{4}{27} = \frac{19}{27},$$

pa je vjerojatnost događaja $\overline{A}$ jednaka

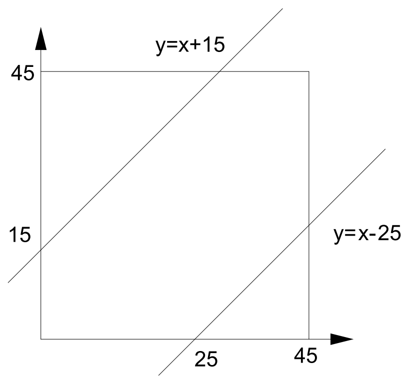
$$P(\overline{A}) = 1 - P(A) = 1 - \frac{19}{27} = \frac{8}{27} = 0.296.$$

□

Zadatak 2.7. Dva automobila, nezavisno jedan od drugog, slučajno dolaze u autopraonicu između 10:00 i 10:45. Prvom automobilu je za pranje potrebno 15 minuta, a drugom 25 minuta. Kolika je vjerojatnost da niti jedan od njih neće morati čekati na pranje?

Rješenje: Neka je x trenutak dolaska prvog, a y trenutak dolaska drugog automobila u autopraonicu. Brojevi x i y su na sreću odabrani brojevi u intervalu $[0, 45]$ (vrijeme brojimo u minutama). Automobili neće morati čekati na pranje ako je $y - x \geq 15$ (prvo je došao prvi automobil) ili $x - y \geq 25$ (prvo je došao drugi automobil). Prema tome, tražimo vjerojatnost događaja

$$P\{y - x \geq 15 \text{ ili } x - y \geq 25\} = P\{y \geq x + 15 \text{ ili } y \leq x - 25\}.$$



Prema tome, nacrtamo li u koordinatnom sustavu kvadrat i pravce $y = x + 15$ i $y = x - 25$ vidimo kako su povoljni događaji one točke kvadrata koje su iznad prvog ili ispod drugog pravca. Područje iznad pravca $y = x + 15$ je jednakokračni pravokutni trokut s katetom duljine 30, a područje ispod pravca $y = x - 25$ je jednakokračni pravokutni trokut s katetom duljine 20 (vidi sliku). Konačno, vjerojatnost da automobili neće čekati na pranje, iznosi

$$P\{y \geq x + 15 \text{ ili } y \leq x - 25\} = \frac{\frac{1}{2} \cdot 30^2 + \frac{1}{2} \cdot 20^2}{45^2} = \frac{650}{2025} = \frac{26}{81} = 0.321.$$

□

3 Uvjetna vjerojatnost

3.1 Uvjetna vjerojatnost

Pretpostavimo da urna sadrži trideset kuglica označenih brojevima $1, 2, 3, \ldots 29, 30$. Vjerojatnost izvlačenja kuglice s brojem 1 očito je jednaka $1/30$. Nakon izvlačenja, mi sa sigurnošću znamo je li se taj događaj ostvario ili ne. Pretpostavimo međutim da je netko pogledao na izvučenu kuglicu i rekao nam: izvučena je kuglica s neparnim brojem. Kolika je sada vjerojatnost da je izvučena kuglica s brojem 1? Očito, ta se vjerojatnost promijenila. Kako imamo petnaest mogućnosti $1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29$, ta je vjerojatnost jednaka $1/15$.

Promotrimo još jedan primjer pridružen gornjem slučajnom pokusu s kuglicama. Neka su A i B događaji

$$\begin{aligned} A &= \{\text{izvučena je kuglica s neparnim brojem}\} \\ B &= \{\text{izvučena je kuglica s brojem djeljivim s 3}\}. \end{aligned}$$

Od 30 elementarnih događaja, događajima A i B pripadaju sljedeći:

$$\begin{aligned} A &= \{1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29\} \\ B &= \{3, 6, 9, 12, 15, 18, 21, 24, 27, 30\}. \end{aligned}$$

Zato je

$$P(A) = \frac{15}{30}, \qquad P(B) = \frac{10}{30}.$$

Kolika je vjerojatnost događaja A, ako je poznato da se realizirao događaj B? U tom je slučaju dovoljno samo promotriti elementarne događaje koji sačinjavaju B (jer samo neki od njih dolaze u obzir) i među njima tražiti one povoljne za događaj A – time tražimo elementarne događaje za presjek $A \cap B$ tih događaja. Ova vjerojatnost ovisi o događaju B te je nazivamo uvjetna vjerojatnost i označavamo je simbolom P_B. Uvjetnu vjerojatnost $P_B(A)$ događaja A čitamo: vjerojatnost od A uz uvjet B. Iz prethodnih razmatranja imamo da je

$$P_B(A) = \frac{5}{10},$$

zato jer je $A \cap B = \{3, 9, 15, 21, 27\}$.

Definicija uvjetne vjerojatnosti

Kako bismo dobili općenitu formulu za uvjetnu vjerojatnost, primijetimo da brojnik 5 označava broj elementarnih događaja koji su povoljni i za događaj A i za događaj B. Već smo vidjeli da je

$A \cap B = \{3, 9, 15, 21, 27\}$. Zbog toga, prethodnu vjerojatnost možemo pisati u sljedećem obliku:

$$P_B(A) = \frac{5}{10} = \frac{\frac{5}{30}}{\frac{10}{30}} = \frac{P(A \cap B)}{P(B)}.$$

Zbog toga je opravdana sljedeća definicija:

Uvjetna vjerojatnost
Neka je $B \in \mathcal{F}$ događaj pozitivne vjerojatnosti: $P(B) > 0$. Uvjetna vjerojatnost uz uvjet B je funkcija $P_B : \mathcal{F} \mapsto [0,1]$ definirana formulom

$$P_B(A) := \frac{P(A \cap B)}{P(B)}, \qquad \forall A \in \mathcal{F}.$$

Lako se uvjeriti da je prethodnom formulom doista definirana funkcija vjerojatnosti, tj. da zadovoljava aksiome vjerojatnosti iz točke 2.2. Uobičajeno je da se uvjetna vjerojatnost P_B označava i formulom $P(\cdot|B)$, dakle, za vjerojatnost događaja A uz uvjet B pisat ćemo $P(A|B)$ umjesto $P_B(A)$.

Primjer 3.1. *Iz snopa od* 52 *karte slučajno je izvučena jedna karta. Kolika je vjerojatnost da je ta karta*

(a) kralj, ako je poznato da je izvučena karta pik boje,

(b) pik boje, ako je poznato da je izvučena karta kralj?

Rješenje: Neka je $A = \{\text{izvučen je kralj}\}$, $B = \{\text{izvučena je karta pik boje}\}$. Očito, u (a) dijelu zadatka tražimo $P(A|B)$, a u (b) dijelu $P(B|A)$. Nadalje, presjek događaja A i B je $A \cap B = \{\text{izvučen je kralj pik boje}\}$. Sada, kako je

$$P(A) = \frac{4}{52}, \quad P(B) = \frac{13}{52}, \qquad P(A \cap B) = \frac{1}{52},$$

imamo:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{52}}{\frac{13}{52}} = \frac{1}{13}, \qquad P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{52}}{\frac{4}{52}} = \frac{1}{4}.$$

□

Primjer 3.2. *U jednoj kutiji je* 8 *bijelih i* 10 *crnih kuglica. Slučajno odabiremo* 3 *kuglice. Kolika je vjerojatnost da su sve izvučene kuglice bijele, ako je poznato da su sve izvučene kuglice iste boje?*

Rješenje: Neka je $A = \{\text{izvučene kuglice su bijele}\}$, $B = \{\text{izvučene kuglice su iste boje}\}$. Očito je $A \cap B = \{\text{izvučene kuglice su bijele}\}$, a mi tražimo uvjetnu vjerojatnost $P(A|B)$. Sada je

$$P(A \cap B) = \frac{\binom{8}{3}}{\binom{18}{3}} \qquad \text{i} \qquad P(B) = \frac{\binom{8}{3} + \binom{10}{3}}{\binom{18}{3}}.$$

Pri tome druga jednakost vrijedi zato što su povoljne mogućnosti odabir triju crvenih ili triju bijelih kuglica. Konačno, dobivamo da je

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\binom{8}{3}}{\binom{8}{3} + \binom{10}{3}} = \frac{7}{22} = 0.3182.$$

□

Primjer 3.3. *Neka su A i B događaji takvi da je $P(A) = \frac{13}{30}$, $P(B) = \frac{7}{15}$, $P(A \cup B) = \frac{11}{15}$. Odredite uvjetne vjerojatnosti $P(A|B)$ i $P(\overline{A}|B)$.*

Rješenje: Iz zadanih vjerojatnosti prvo računamo vjerojatnost presjeka:

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = \frac{13}{30} + \frac{7}{15} - \frac{11}{15} = \frac{5}{30} = \frac{1}{6}.$$

Zatim računamo uvjetnu vjerojatnost po definiciji:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{6}}{\frac{7}{15}} = \frac{5}{14}.$$

Budući da je uvjetna vjerojatnost i sama vjerojatnost, vrijedi pravilo komplementa:

$$P(\overline{A}|B) = 1 - P(A|B) = 1 - \frac{5}{14} = \frac{9}{14}.$$

□

Primjer 3.4. *Kolika je vjerojatnost da slučajno odabrana permutacija skupa $\{1,2,3,4,5\}$ počinje parnom znamenkom, ako je poznato da završava neparnom znamenkom?*

Rješenje: Neka je

$$\begin{aligned} A &= \{\text{permutacija počinje parnom znamenkom}\} \\ B &= \{\text{permutacija završava neparnom znamenkom}\}. \end{aligned}$$

Očito je $A \cap B = \{$permutacija počinje parnom te završava neparnom znamenkom$\}$, a mi tražimo uvjetnu vjerojatnost $P(A|B)$.
Broj permutacija skupa $\{1,2,3,4,5\}$, koje završavaju neparnom znamenkom jednak je $3 \cdot 4!$, zato jer kraj možemo odabrati na tri načina (znamenke $1,3,5$), a ostatak permutiramo na $4!$ načina. Zbog toga je

$$P(B) = \frac{3 \cdot 4!}{5!} = \frac{3}{5}$$

Slično, broj permutacija koje počinju parnim brojem te završavaju neparnim, jednak je $2 \cdot 3 \cdot 3!$, zato jer početak možemo odabrati na 2 načina, kraj na 3 načina, te ostatak ispermutirati na $3!$ načina. Stoga je

$$P(A \cap B) = \frac{2 \cdot 3 \cdot 3!}{5!} = \frac{3}{10}.$$

Konačno, iz formule za uvjetnu vjerojatnost imamo

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{3}{10}}{\frac{3}{5}} = \frac{1}{2}.$$

□

U mnogim primjerima lakše računamo uvjetnu vjerojatnost nego vjerojatnost presjeka događaja. Stoga se formula za uvjetnu vjerojatnost često koristi u računanju presjeka, odnosno umnoška, dvaju događaja:

Vjerojatnost presjeka (umnoška)
Vjerojatnost presjeka dvaju događaja računa se pomoću formule $P(A \cap B) = P(B)P(A|B)$.

Ako zamijenimo uloge događaja A i B (koji oboje imaju pozitivnu vjerojatnost), dobivamo sličnu formulu

$$P(A \cap B) = P(A)P(B|A).$$

Primjer 3.5. *U košari se nalazi* 5 *jabuka i* 10 *krušaka. Izvlačimo dvije voćke, jednu za drugom, bez vraćanja. Kolika je vjerojatnost da su izvučene dvije jabuke?*

Rješenje: Voćke izvlačimo jednu po jednu. Neka su A i B događaji

$$\begin{aligned} A &= \{\text{prva voćka je jabuka}\} \\ B &= \{\text{druga voćka je jabuka}\}. \end{aligned}$$

Tada je $A \cap B$ događaj čiju vjerojatnost tražimo. Očito je:

$$P(A) = \frac{5}{15} = \frac{1}{3}.$$

Nakon što izvučemo prvu voćku, u košari je preostalo četrnaest voćki, od kojih su četiri jabuke. Stoga je

$$P(B|A) = \frac{4}{14} = \frac{2}{7}.$$

Konačno, imamo

$$P(A \cap B) = P(A)P(B|A) = \frac{1}{3} \cdot \frac{2}{7} = \frac{2}{21}.$$

Uočimo kako smo traženu vjerojatnost mogli naći i klasičnim pristupom:

$$P(A \cap B) = \frac{\binom{5}{2}}{\binom{15}{2}} = \frac{2}{21}.$$

□

Na sličan način možemo računati i vjerojatnost presjeka više događaja. Na primjer,

$$P(A \cap B \cap C) = P(A) \cdot P(B|A) \cdot P(C|A \cap B).$$

Moguće su i druge kombinacije događaja s desne strane.

Primjer 3.6. *Student polaže usmeni ispit iz kolegija Vjerojatnost i statistika. Zna odgovore na* 35 *pitanja od mogućih* 50. *Ako se tri pitanja izvlače jedno za drugim, bez vraćanja, odredite vjerojatnost da će student znati odgovoriti na sva tri pitanja.*

Rješenje: Neka je A traženi događaj i neka je $A_i = \{\text{student zna odgovor na } i\text{-to pitanje}\}$, $i = 1, 2, 3$. Tada je $A = A_1 \cap A_2 \cap A_3$ i računamo vjerojatnost po formuli

$$P(A) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2).$$

Pojedine vjerojatnosti su

$P(A_1) = \dfrac{35}{50}$, zato jer student zna odgovor na 35 pitanja,

$P(A_2|A_1) = \dfrac{34}{49}$, nakon što je izvučeno pitanje na koje student zna odgovor, ostalo je 49 pitanja od kojih student zna odgovor na njih 34,

$P(A_3|A_1 \cap A_2) = \dfrac{33}{48}$.

Dakle, $P(A) = \dfrac{35}{50} \cdot \dfrac{34}{49} \cdot \dfrac{33}{48} = 0.334$. Do istog rezultata mogli smo doći klasičnim putem:

$$P(A) = \frac{\binom{35}{3}}{\binom{50}{3}}.$$

□

3.2 Nezavisni događaji

U Primjeru 3.5 smo iz košare izvlačili voćke, jednu za drugom bez vraćanja. Dakako, nakon svake izvučene voćke vjerojatnost odabira pojedine vrste voća se mijenjala. Proučimo sljedeći primjer koji je usko u vezi sa spomenutim primjerom.

Primjer 3.7. *U košari se nalazi* 5 *jabuka i* 10 *krušaka. Izvlačimo dvije voćke, jednu za drugom. Kolika je vjerojatnost da će druga izvučena voćka biti jabuka, ako je prva izvučena voćka bila jabuka? Kolika je ta vjerojatnost ako je prva voćka bila kruška? Izračunajte obje ove vjerojatnosti u sljedeća dva slučaja:*

(a) prva se voćka nakon izvlačenja ne vraća u košaru,

(b) prva se voćka nakon izvlačenja vraća u košaru.

Rješenje: Označimo s A i B događaje

$$\begin{aligned} A &= \{\text{prva voćka je jabuka}\} \\ B &= \{\text{druga voćka je jabuka}\}. \end{aligned}$$

Tražimo uvjetne vjerojatnosti $P(B|A)$ i $P(B|\overline{A})$. Očito je

$$P(A) = \frac{5}{15} = \frac{1}{3}, \qquad P(\overline{A}) = \frac{10}{15} = \frac{2}{3}.$$

(a) Nakon izvlačenja prve voćke, u košari imamo jednu voćku manje. Zato je

$$P(B|A) = \frac{4}{14} = \frac{2}{7}, \qquad P(B|\overline{A}) = \frac{5}{14}.$$

(b) Ako voćku nakon izvlačenja vratimo u košaru, prije izvlačenja druge voćke imat ćemo istu situaciju: 5 jabuka i 10 krušaka, bez obzira je li se ostvario događaj A ili nije:

$$P(B|A) = \frac{5}{15} = \frac{1}{3}, \qquad P(B|\overline{A}) = \frac{5}{15} = \frac{1}{3}.$$

Kažemo da *realizacija događaja A ne utječe na vjerojatnost realizacije događaja B.* □

Neka događaji A i B imaju pozitivnu vjerojatnost. Neka je $P(B|A) = P(B)$, tj. vjerojatnost događaja B ne mijenja se nakon što nam je poznato da se realizirao događaj A. Tada kažemo da su A i B **nezavisni događaji**.

U tom slučaju vrijedi

$$P(A \cap B) = P(A)P(B|A) = P(A)P(B).$$

Ako je ispunjena prethodna jednakost, onda za uvjetnu vjerojatnost vrijede sljedeće dvije relacije:

$$\begin{aligned} P(B|A) &= \frac{P(A \cap B)}{P(A)} = \frac{P(A)P(B)}{P(A)} = P(B), \\ P(A|B) &= \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A). \end{aligned}$$

Definicija i kriterij nezavisnosti događaja
Za događaje A i B kažemo da su nezavisni, ako vrijedi bilo koja od jednakosti: $P(A|B) = P(A)$ ili $P(B|A) = P(B)$.
Nužan i dovoljan uvjet za nezavisnost jest relacija $P(A \cap B) = P(A)P(B)$.

Primjer 3.8. *Jesu li događaji A i B nezavisni ako je $P(A) = 0.2$, $P(B) = 0.5$ i $P(A \cup B) = 0.6$?*

Rješenje: Izračunajmo prvo vjerojatnost presjeka događaja A i B:

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.2 + 0.5 - 0.6 = 0.1.$$

Nadalje, imamo

$$P(A)P(B) = 0.2 \cdot 0.5 = 0.1 = P(A \cap B),$$

pa su događaji A i B nezavisni. □

Primjer 3.9. *Bacamo dvije kocke. Kolika je vjerojatnost da na prvoj kocki bude prost, a na drugoj složen broj?*

Rješenje: Rezultat na jednoj kocki nezavisan je od rezultata na drugoj kocki. Vjerojatnost pojave prostog broja na prvoj kocki je $\frac{1}{2}$, zato jer su mogući prosti brojevi $2, 3, 5$. Vjerojatnost pojave složenog broja na drugoj kocki je $\frac{1}{3}$ zato jer su mogući složeni brojevi 4 i 6. Traženi događaj je presjek ovih dvaju (nezavisnih) događaja, pa je njegova vjerojatnost jednaka $\frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$. □

Primjer 3.10. *Neka su A i B nezavisni događaji. Dokažite da su tada događaji $\overline{A}$ i B također nezavisni.*

Rješenje: Iz pravila komplementa i nezavisnosti događaja A i B imamo:

$$P(\overline{A})P(B) = (1 - P(A))P(B) = P(B) - P(A)P(B) = P(B) - P(A \cap B).$$

Nadalje, iz Venn-Eulerovih dijagrama lagano uočavamo kako su događaji $A \cap B$ i $\overline{A} \cap B$ disjunktni, te je njihova unija jednaka događaju B. Zbog toga je

$$P(A \cap B) + P(\overline{A} \cap B) = P(B),$$

odakle je $P(B) - P(A \cap B) = P(\overline{A} \cap B)$. Kombinirajući ovu relaciju s gornjom relacijom slijedi da je

$$P(\overline{A})P(B) = P(\overline{A} \cap B),$$

odakle slijedi da su događaji $\overline{A}$ i B nezavisni. □

Nezavisnost skupine događaja definira se na složeniji način.

Nezavisnost događaja
Događaji $A_1, A_2, \ldots, A_n$ su nezavisni ako za svaki k, $2 \leq k \leq n$ i svaki izbor $A_{i_1}, A_{i_2}, \ldots A_{i_k}$ nekolicine tih događaja vrijedi

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) P(A_{i_2}) \cdots P(A_{i_k}).$$

Primjer 3.11. *Kada su tri događaja A, B, C nezavisna?*

Rješenje: U skladu s gornjom definicijom mora vrijediti:

$$\begin{aligned} P(A \cap B) &= P(A)P(B), \\ P(B \cap C) &= P(B)P(C), \\ P(C \cap A) &= P(C)P(A), \\ P(A \cap B \cap C) &= P(A)P(B)P(C). \end{aligned}$$

Prema tome, događaji A, B i C su nezavisni ako i samo ako su ispunjene prethodne četiri jednakosti. $\square$

Prema tome, tri događaja ne moraju biti nezavisna ako samo zadovoljavaju jednakost

$$P(A \cap B \cap C) = P(A)P(B)P(C).$$

Nezavisnost triju događaja znači da će sve uvjetne vjerojatnosti u kojima se ti događaji javljaju biti jednake bezuvjetnima. Tako je, na primjer,

$$P(A|B \cap C) = \frac{P(A \cap B \cap C)}{P(B \cap C)} = \frac{P(A)P(B)P(C)}{P(B)P(C)} = P(A).$$

Slično vrijedi i za druge moguće kombinacije uvjetnih vjerojatnosti.

Primjer 3.12. *Bacamo dvije kocke. Dani su događaji:*

$$\begin{aligned} A &= \{\textit{na prvoj kocki je pao broj manji ili jednak } 3\ \}, \\ B &= \{\textit{na drugoj kocki je pao broj veći od } 2\}, \\ C &= \{\textit{zbroj dobivenih brojeva je } 7\}. \end{aligned}$$

Da li su događaji A, B i C nezavisni?

Rješenje: Ovaj pokus možemo opisati klasičnim modelom jer je svaki od 36 elementarnih događaja jednako vjerojatan. Imamo redom:

$$P(A) = \frac{3 \cdot 6}{36} = \frac{1}{2}, \qquad P(B) = \frac{6 \cdot 4}{36} = \frac{2}{3}, \qquad P(C) = \frac{6}{36} = \frac{1}{6}.$$

Sada, za presjeke dvaju događaja vrijedi

$$\begin{aligned} P(A \cap B) &= \frac{3 \cdot 4}{36} = \frac{1}{3} = P(A)P(B), \\ P(B \cap C) &= \frac{4}{36} = \frac{1}{9} = P(B)P(C), \\ P(C \cap A) &= \frac{3}{36} = \frac{1}{12} = P(C)P(A). \end{aligned}$$

Međutim, događaj $A \cap B \cap C$ sadrži samo tri elementarna događaja, te vrijedi

$$P(A \cap B \cap C) = \frac{3}{36} = \frac{1}{12} \neq \frac{1}{18} = P(A)P(B)P(C).$$

Prema tome, događaji A, B i C nisu nezavisni, iako je svaki par događaja međusobno nezavisan. $\square$

Primjer 3.13. *Tri strijelca gađaju metu nezavisno jedan od drugog. Vjerojatnost pogotka prvog je* 0.85*, drugog* 0.91*, a trećeg* 0.73*. Odredite vjerojatnosti sljedećih događaja:*

(a) sva tri strijelca su pogodila metu,

(b) barem jedan strijelac je pogodio metu,

(c) točno jedan strijelac je pogodio metu.

Rješenje: Neka je

$$\begin{aligned} A &= \{\text{prvi strijelac je pogodio metu}\}, \\ B &= \{\text{drugi strijelac je pogodio metu}\}, \\ C &= \{\text{treći strijelac je pogodio metu}\}. \end{aligned}$$

(a) Kako su događaji A, B i C nezavisni, imamo da je

$$P(A \cap B \cap C) = P(A)P(B)P(C) = 0.85 \cdot 0.91 \cdot 0.73 = 0.565.$$

(b) Lakše nam je izračunati vjerojatnost suprotnog događaja, odnosno vjerojatnost da niti jedan strijelac nije pogodio metu. Očito, komplementi događaja A, B i C su nezavisni pa je

$$P(A \cup B \cup C) = 1 - P(\overline{A} \cap \overline{B} \cap \overline{C}) = 1 - P(\overline{A})P(\overline{B})P(\overline{C}) = 1 - 0.15 \cdot 0.09 \cdot 0.27 = 0.996.$$

(c) Ovdje imamo uniju triju međusobno isključivih događaja, odnosno događaja kod kojih jedan strijelac pogađa, a preostala dva promašuju metu. Očito, događaji $A, \overline{B}$ i $\overline{C}$ su nezavisni. Slično vrijedi i za preostale kombinacije, pa je vjerojatnost pogotka točno jednog strijelca jednaka

$$\begin{aligned} & P\left((A \cap \overline{B} \cap \overline{C}) \cup (\overline{A} \cap B \cap \overline{C}) \cup (\overline{A} \cap \overline{B} \cap C)\right) \\ & = P(A \cap \overline{B} \cap \overline{C}) + P(\overline{A} \cap B \cap \overline{C}) + P(\overline{A} \cap \overline{B} \cap C) \\ & = P(A)P(\overline{B})P(\overline{C}) + P(\overline{A})P(B)P(\overline{C}) + P(\overline{A})P(\overline{B})P(C) \\ & = 0.85 \cdot 0.09 \cdot 0.27 + 0.15 \cdot 0.91 \cdot 0.27 + 0.15 \cdot 0.09 \cdot 0.73 = 0.067. \end{aligned}$$

□

3.3 Formula potpune vjerojatnosti

Pri računanju vjerojatnosti ponekad moramo sve moguće ishode podijeliti u različite skupine. Pogledajmo sljedeći primjer.

Primjer 3.14. *Prvi stroj daje* 4% *neispravnih proizvoda, a drugi* 5% *neispravnih proizvoda. Uzima se* 70 *proizvoda proizvedenih na prvom stroju i* 30 *proizvedenih na drugom stroju te sprema u skladište. Iz skladišta se slučajno bira jedan proizvod. Kolika je vjerojatnost da je proizvod ispravan?*

Rješenje: Odaberemo na sreću jedan proizvod u skladištu. Dvije su mogućnosti:

$$\begin{aligned} H_1 &= \{\text{proizvod je napravljen na prvom stroju}\}, \\ H_2 &= \{\text{proizvod je napravljen na drugom stroju}\}. \end{aligned}$$

Vjerojatnosti odabira proizvoda s prvog, odnosno drugog stroja redom su jednake

$$P(H_1) = \frac{70}{100} = 0.7, \qquad P(H_2) = \frac{30}{100} = 0.3.$$

Neka je A traženi događaj:

$$A = \{\text{odabran je ispravan proizvod}\}.$$

Događaj A možemo rastaviti na dva disjunktna događaja, tj. vrijedi $A = (A \cap H_1) \cup (A \cap H_2)$. Pri tome je:

$$\begin{aligned} A \cap H_1 &= \{\text{odabrani ispravni proizvod napravljen je na prvom stroju}\}, \\ A \cap H_2 &= \{\text{odabrani ispravni proizvod napravljen je na drugom stroju}\}. \end{aligned}$$

Zato je

$$P(A) = P(A \cap H_1) + P(A \cap H_2).$$

Vjerojatnosti presjeka događaja računamo na poznati način, pomoću uvjetne vjerojatnosti:

$$P(A \cap H_1) = P(H_1) \cdot P(A|H_1).$$

Vjerojatnost da je proizvod ispravan, ako je poznato da je napravljen na prvom stroju je, prema podatcima

$$P(A|H_1) = 0.96 \quad \Longrightarrow \quad P(A \cap H_1) = 0.7 \cdot 0.96 = 0.672.$$

Analogno tome vrijedi

$$P(A|H_2) = 0.95 \quad \Longrightarrow \quad P(A \cap H_2) = 0.3 \cdot 0.95 = 0.285.$$

Sada dobivamo

$$P(A) = P(A \cap H_1) + P(A \cap H_2) = 0.672 + 0.285 = 0.957.$$

□

Poopćimo razmatranje iz prethodnog primjera na slučaj kada se može pojaviti više različitih mogućnosti. Pretpostavimo da skup elementarnih događaja možemo rastaviti u n međusobno disjunktnih događaja:

$$\Omega = H_1 \cup H_2 \cup \ldots \cup H_n,$$

pri čemu su događaji H_i, H_j disjunktni za $i \neq j$ i vrijedi $P(H_i) > 0$ za svaki i. Ovakav rastav nazivamo **particija vjerojatnosnog prostora**. Kažemo još da familija $H_1, H_2, \ldots, H_n$ čini **potpun sustav događaja**.

Neka je $A \subset \Omega$ bilo koji događaj. Familijom $H_1, H_2, \ldots, H_n$ i on je rastavljen na događaje:

$$A = (A \cap H_1) \cup (A \cap H_2) \cup \ldots \cup (A \cap H_n).$$

Kako su događaji $A \cap H_i$ međusobno disjunktni, vrijedi:

$$\begin{aligned} P(A) &= P(A \cap H_1) + P(A \cap H_2) + \ldots + P(A \cap H_n) \\ &= P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + \ldots + P(H_n)P(A|H_n). \end{aligned}$$

Formula potpune vjerojatnosti
Neka je $\{H_1, H_2, \ldots, H_n\}$ potpun sustav događaja. Za svaki događaj $A \subset \Omega$ vrijedi

$$P(A) = \sum_{i=1}^{n} P(H_i)P(A|H_i).$$

Događaje $H_1, H_2, \ldots, H_n$ nazivamo **hipotezama**. Tijekom realizacije nekog pokusa ostvaruje se točno jedna hipoteza.

Primjer 3.15. *Ivan u lijevom džepu ima dvije kovanice od* 2 *kune i četiri kovanice od* 5 *kuna, a u desnom džepu tri kovanice od* 2 *kune i četiri od* 5 *kuna. Ivan slučajno odabire jednu kovanicu iz lijevog džepa i prebacuje je u desni džep. Kolika je vjerojatnost da nakon toga on iz desnog džepa izvuče kovanicu od* 5 *kuna?*

Rješenje: Vjerojatnost izbora kovanice od 5 kuna ovisi o tome koja je kovanica prebačena iz lijevog u desni džep. Postavimo sljedeće hipoteze:

$$\begin{aligned} H_1 &= \{\text{iz lijevog džepa je prebačena kovanica od 2 kune}\}, \quad P(H_1) = \frac{1}{3} \\ H_2 &= \{\text{iz lijevog džepa je prebačena kovanica od 5 kuna}\}, \quad P(H_2) = \frac{2}{3}. \end{aligned}$$

Označimo s A događaj čiju vjerojatnost tražimo: iz desnog džepa je izvučena kovanica od 5 kuna. Ako se ostvari prva hipoteza, tada se u desnom džepu nalaze po četiri kovanice od po 2 i 5 kuna. Zato je

$$P(A|H_1) = \frac{1}{2}.$$

Ako se ostvari druga hipoteza, tada se u desnom džepu nalaze tri kovanice od 2 kune i pet kovanica od 5 kuna. Zato je

$$P(A|H_2) = \frac{5}{8}.$$

Prema formuli potpune vjerojatnosti vrijedi

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) = \frac{1}{3} \cdot \frac{1}{2} + \frac{2}{3} \cdot \frac{5}{8} = \frac{7}{12}.$$

□

Primjer 3.16. *Kako nije bio potpuno siguran u dijagnozu, liječnik je pacijentu propisao tri različita lijeka koja mora redovito piti. Vjerojatnosti da pacijent pozitivno reagira na svaki od triju propisanih lijekova ne ovise jedna o drugoj i iznose* 0.7. *Ako pacijent pozitivno reagira na jedan lijek, vjerojatnost da će on ozdraviti iznosi* 0.5*, ako pozitivno reagira na dva lijeka, ozdravit će s vjerojatnošću* 0.7*, a ako pozitivno reagira na sva tri lijeka, ozdravit će s vjerojatnošću* 0.9*. Nađi vjerojatnost da pacijent ozdravi.*

Rješenje: Označimo događaj i hipoteze:

$$\begin{aligned} A &= \{\text{pacijent je ozdravio}\}, \\ H_i &= \{\text{pacijent pozitivno reagira na } i \text{ lijekova}\}, \quad i = 0, 1, 2, 3. \end{aligned}$$

Tada je

$$P(H_0) = 0.3^3, \quad P(H_1) = \binom{3}{1} \cdot 0.7 \cdot 0.3^2, \quad P(H_2) = \binom{3}{2} \cdot 0.7^2 \cdot 0.3, \quad P(H_3) = 0.7^3.$$

Prema uvjetima zadatka vrijedi:

$$P(A|H_0) = 0, \qquad P(A|H_1) = 0.5, \qquad P(A|H_2) = 0.7, \qquad P(A|H_3) = 0.9.$$

Konačno, po formuli uvjetne vjerojatnosti imamo

$$P(A) = \sum_{i=0}^{3} P(H_i)P(A|H_i) = 0.712.$$

□

3.4 Bayesova formula

Iz poznatih relacija
$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$$
možemo napisati
$$P(B|A) = \frac{P(B)P(A|B)}{P(A)}.$$
Ovu formulu koristimo uglavnom onda kada je događaj B jedna od hipoteza $H_1, H_2, \ldots, H_n$, na koje je razbijen skup Ω:
$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{P(A)}.$$
Pritom se vjerojatnost $P(A)$ računa uglavnom pomoću formule potpune vjerojatnosti. Tako dobivamo **Bayesovu formulu**.

Bayesova formula
Ako su $H_1, H_2, \ldots, H_n$ hipoteze te A događaj, onda vrijedi:
$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{\sum_{j=1}^{n} P(H_j)P(A|H_j)}.$$

Bayesovu formulu koristimo pri računanju aposteriornih vjerojatnosti pojedinih hipoteza. Prije početka pokusa svaka hipoteza ima svoju vjerojatnost realizacije $P(H_i)$. Nakon realizacije pokusa, ako znamo koji se elementarni događaj ostvario, tada je nestala neizvjesnost: ostvarila se samo jedna od mogućih hipoteza $H_1, H_2, \ldots, H_n$, dok za sve ostale znamo sa sigurnošću da se nisu ostvarile.

Pretpostavimo međutim da nam nije poznato koji se elementarni događaj ostvario, već umjesto toga znamo da se ostvario događaj $A \subset \Omega$. U tom slučaju ne znamo točno koja je od hipoteza $H_1, H_2, \ldots, H_n$ nastupila, ali dodatna informacija o realizaciji događaja A mijenja **apriorne vjerojatnosti** pojedinih hipoteza. Pomoću Bayesove formule računamo uvjetne vjerojatnosti $P(H_1|A), P(H_2|A), \ldots, P(H_n|A)$, koje nazivamo aposteriornim vjerojatnostima pojedinih hipoteza. Uočimo još kako je zbroj vjerojatnosti svih aposteriornih hipoteza jednak 1.

Primjer 3.17. *Bacamo kocku. Neka su H_1 i H_2 hipoteze*
$$\begin{aligned} H_1 &= \{\textit{kocka pokazuje parni broj}\}, \\ H_2 &= \{\textit{kocka pokazuje neparni broj}\}, \end{aligned}$$
te neka je A događaj
$$A = \{\textit{kocka pokazuje broj manji od } 4\}.$$
Odredite vjerojatnosti aposteriornih hipoteza $P(H_1|A)$ i $P(H_2|A)$.

Rješenje: Prije bacanja kocke, apriorne vjerojatnosti pojedinih hipoteza su
$$P(H_1) = \frac{1}{2}, \qquad P(H_2) = \frac{1}{2}.$$
Lagano računamo i uvjetne vjerojatnosti događaja A:
$$P(A|H_1) = \frac{1}{3}, \qquad P(A|H_2) = \frac{2}{3}.$$

Izračunajmo sada aposteriorne vjerojatnosti pojedinih hipoteza. Ako se ostvario događaj $A = \{1, 2, 3\}$, onda očito hipoteza H_2 postaje vjerojatnija od hipoteze H_1, budući da događaj A sadrži dva neparna i samo jedan paran broj.

Konačno, prema Bayesovoj formuli dobivamo nove, aposteriorne vjerojatnosti:

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

□

Primjer 3.18. *U vreći se nalazi 5 novčića među kojima je jedan lažni jer ima pismo s obje strane. Slučajno odabrani novčić baca se 3 puta i svaki se put pojavi pismo. Kolika je vjerojatnost da je odabran lažni novčić?*

Rješenje: Slučajni pokus je bacanje novčića tri puta. Dakle, ishode dijelimo prema tome je li bacan lažni ili pravi novčić. Prema tome, imamo sljedeće hipoteze:

$$\begin{aligned} H_1 &= \{\text{odabran je lažni novčić}\}, \quad P(H_1) = \frac{1}{5} \\ H_2 &= \{\text{odabran je pravi novčić}\}, \quad P(H_2) = \frac{4}{5}. \end{aligned}$$

No, ishod pokusa je poznat – tri pisma. Zanima nas vjerojatnost hipoteze H_1, nakon što znamo da se ostvario događaj $A = \{\text{pala su tri pisma}\}$. Dakle, trebamo izračunati $P(H_1|A)$.
Izračunajmo prvo uvjetne vjerojatnosti. Očito je

$$P(A|H_1) = 1, \qquad P(A|H_2) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}.$$

Sada, prema Bayesovoj formuli imamo:

$$P(H_1|A) = \frac{P(H_1)P(A|H_1)}{P(H_1)P(A|H_1) + P(H_2)P(A|H_2)} = \frac{\frac{1}{5}\cdot 1}{\frac{1}{5}\cdot 1 + \frac{4}{5}\cdot\frac{1}{8}} = \frac{2}{3}.$$

□

Primjer 3.19. *U prvoj posudi nalaze se* 3 *jestive i* 3 *otrovne gljive, a u drugoj* 4 *jestive i* 2 *otrovne. Iz prve posude na sreću odabiremo dvije gljive i prebacimo ih u drugu. Zatim iz druge posude izvlačimo jednu gljivu.*

(a) Kolika je vjerojatnost da izvučena gljiva iz druge posude bude jestiva?

(b) Odredite vjerojatnost da su obje gljive prebačene iz prve posude bile jestive, ako je iz druge izvučena jestiva gljiva.

Rješenje: (a) Nazovimo traženi događaj

$$A = \{\text{gljiva izvučena iz druge posude je jestiva}\}.$$

Pri prebacivanju dviju gljiva u drugu posudu postoje tri mogućnosti:

$$\begin{aligned}
H_0 &= \{\text{niti jedna prebačena gljiva nije jestiva}\}, & P(H_0) &= \frac{\binom{3}{2}}{\binom{6}{2}} = \frac{1}{5},\\
H_1 &= \{\text{jedna prebačena gljiva je jestiva}\}, & P(H_1) &= \frac{\binom{3}{1}\binom{3}{1}}{\binom{6}{2}} = \frac{3}{5},\\
H_2 &= \{\text{obje prebačene gljive su jestive}\}, & P(H_2) &= \frac{\binom{3}{2}}{\binom{6}{2}} = \frac{1}{5}.
\end{aligned}$$

Vrijedi

$$P(A|H_0) = \frac{1}{2}, \qquad P(A|H_1) = \frac{5}{8}, \qquad P(A|H_2) = \frac{3}{4}.$$

Po formuli potpune vjerojatnosti dobivamo

$$P(A) = \sum_{i=0}^{2} P(H_i)P(A|H_i) = \frac{1}{5}\cdot\frac{1}{2} + \frac{3}{5}\cdot\frac{5}{8} + \frac{1}{5}\cdot\frac{3}{4} = \frac{5}{8}.$$

(b) Uočimo kako smo dio ovog zadatka već riješili u (a) dijelu. Trebamo izračunati $P(H_2|A)$. Stoga, iz Bayesove formule imamo:

$$P(H_2|A) = \frac{P(H_2)P(A|H_2)}{\sum_{i=0}^{2} P(H_i)P(A|H_i)} = \frac{\frac{1}{5}\cdot\frac{3}{4}}{\frac{5}{8}} = \frac{6}{25}.$$

□

Primjer 3.20. *U jednoj od anketa provedenih prije predsjedničkih izbora,* 41% *građana je podržalo kandidata X.Y.,* 37% *građana ga nije podržalo, a* 22% *građana nije imalo izgrađeno mišljenje o kandidatu. Od onih koji podržavaju kandidata X.Y,* 62% *je s visokim obrazovanjem, od onih koji ga ne podržavaju* 40% *je s visokim obrazovanjem, a od neodlučnih je* 15% *s visokim obrazovanjem. Novinar političkog tjednika slučajnim odabirom upita jednog od anketiranih, visoko obrazovanog građanina da li podržava kandidata X.Y., na što mu on nije htio odgovoriti. Kolika je vjerojatnost da taj građanin ne podržava kandidata X.Y.?*

Rješenje: Stavimo $A = \{\text{građanin je visoko obrazovan}\}$. Imamo sljedeće hipoteze

$$\begin{aligned}
H_1 &= \{\text{građanin podržava predsjedničkog kandidata}\}, & P(H_1) = 0.41,\\
H_2 &= \{\text{građanin ne podržava predsjedničkog kandidata}\}, & P(H_2) = 0.37,\\
H_3 &= \{\text{građanin nema izgrađeno mišljenje o kandidatu}\}, & P(H_3) = 0.22.
\end{aligned}$$

Nadalje, poznate su nam i uvjetne vjerojatnosti:

$$P(A|H_1) = 0.62, \quad P(A|H_2) = 0.40, \quad P(A|H_3) = 0.15.$$

Sada računamo vjerojatnost da građanin ne podržava predsjedničkog kandidata, ako znamo da je visoko obrazovan, odnosno $P(H_2|A)$:

$$P(H_2|A) = \frac{P(H_2)P(A|H_2)}{\sum_{i=1}^{3} P(H_i)P(A|H_i)} = \frac{0.37\cdot 0.4}{0.41\cdot 0.62 + 0.37\cdot 0.4 + 0.22\cdot 0.15} = 0.34.$$

□

3.5 Zadatci za ponavljanje

Zadatak 3.1. Neka su A i B događaji takvi da je $P(A) = \dfrac{1}{4}$, $P(B) = \dfrac{2}{3}$ i $P(A \cap B) = \dfrac{1}{12}$. Koliko je $P(\overline{B}|\overline{A})$?

Rješenje: Iskoristimo li formulu za uvjetnu vjerojatnost i de Morganovu formulu za komplement unije, imamo da je

$$P(\overline{B}|\overline{A}) = \frac{P(\overline{B} \cap \overline{A})}{P(\overline{A})} = \frac{P(\overline{A \cup B})}{P(\overline{A})} = \frac{1 - P(A \cup B)}{1 - P(A)}.$$

Iz formule za vjerojatnost unije imamo da je

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{4} + \frac{2}{3} - \frac{1}{12} = \frac{5}{6},$$

pa uvrštavanjem u prethodnu formulu dobivamo

$$P(\overline{B}|\overline{A}) = \frac{1 - \frac{5}{6}}{1 - \frac{1}{4}} = \frac{2}{9}.$$

□

Zadatak 3.2. Pokus se sastoji u bacanju dviju kocaka. Promatramo događaje

$$\begin{aligned} A &= \{\text{pojavila se barem jedna četvorka}\}, \\ B &= \{\text{pojavila se barem jedna dvojka}\}, \\ C &= \{\text{pojavio se jedan paran i jedan neparan broj}\}. \end{aligned}$$

(a) Odredite uvjetnu vjerojatnost $P(B|C)$

(b) Ispitajte da li su događaji A i B nezavisni.

Rješenje:

(a) Iz formule za uvjetnu vjerojatnost imamo da je

$$P(B|C) = \frac{P(B \cap C)}{P(C)},$$

pa moramo izračunati vjerojatnosti događaja $B \cap C$ i C. Događaj $B \cap C$ predstavljaju svi ishodi koji sadrže jednu dvojku i jedan neparan broj, pa je

$$P(B \cap C) = \frac{1 \cdot 3 + 3 \cdot 1}{36} = \frac{1}{6}.$$

Slično je

$$P(C) = \frac{3 \cdot 3 + 3 \cdot 3}{36} = \frac{1}{2},$$

pa je

$$P(B|C) = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}.$$

(b) Događaji A i B su nezavisni ako vrijedi jednakost $P(A \cap B) = P(A) \cdot P(B)$, pa moramo izračunati vjerojatnosti pripadnih događaja. Primjenom pravila komplementa imamo da je

$$P(A) = P(B) = 1 - \left(\frac{5}{6}\right)^2 = \frac{11}{36}.$$

Nadalje, događaj $A \cap B$ sadrži samo dva elementarna događaja (to su $(2,4)$ i $(4,2)$), pa je

$$P(A \cap B) = \frac{2}{36} = \frac{1}{18}.$$

Konačno, događaji A i B su zavisni zato jer je

$$P(A \cap B) = \frac{1}{18} \neq \frac{121}{1296} = \frac{11}{36} \cdot \frac{11}{36} = P(A) \cdot P(B).$$

□

Zadatak 3.3. Bacamo četiri igraće kocke. Zadani su događaji

$$\begin{aligned} A &= \{\text{pala su četiri različita broja}\}, \\ B &= \{\text{pala je barem jedna jedinica}\}. \end{aligned}$$

Izračunajte uvjetne vjerojatnosti $P(A|B)$ i $P(B|A)$ te ispitajte da li su događaji A i B nezavisni.

Rješenje: Trebamo izračunati vjerojatnosti događaja A, B i $A \cap B$. Lagano dobivamo

$$\begin{aligned} P(A) &= \frac{6 \cdot 5 \cdot 4 \cdot 3}{6^4} = \frac{5}{18}, \\ P(B) &= 1 - \left(\frac{5}{6}\right)^4 = \frac{671}{1296}, \end{aligned}$$

pri čemu smo kod događaja B koristili pravilo komplementa. Nadalje, događaj $A \cap B$ sastoji se od uređenih četvorki s različitim brojevima i točno jednom jedinicom. Kako jedinicu možemo odabrati na jednoj od četiriju kocaka, imamo da je

$$P(A \cap B) = \frac{\binom{4}{1} \cdot 5 \cdot 4 \cdot 3}{6^4} = \frac{5}{27}.$$

Zato su tražene uvjetne vjerojatnosti redom jednake

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{5}{27}}{\frac{671}{1296}} = \frac{240}{671}$$

i

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{5}{27}}{\frac{5}{18}} = \frac{2}{3}.$$

Prema tome, događaji A i B su zavisni zato jer je

$$P(A \cap B) = \frac{5}{27} \neq \frac{5}{18} \cdot \frac{671}{1296} = P(A)P(B).$$

□

Zadatak 3.4. Vjerojatnosti pogotka u metu za svakog od tri strijelca su redom 0.8, 0.6, 0.5. Kolika je vjerojatnost da će meta biti pogođena točno dvaput?

Rješenje: Označimo događaje:

$$\begin{aligned} A &= \{\text{prvi strijelac je pogodio metu}\}, \\ B &= \{\text{drugi strijelac je pogodio metu}\}, \\ C &= \{\text{treći strijelac je pogodio metu}\}. \end{aligned}$$

Ovdje imamo uniju triju međusobno disjunktnih događaja, odnosno događaja kod kojih dva strijelca pogađaju, a jedan promašuje metu. Kako su događaji A, B, C nezavisni, slijedi da su i događaji $\overline{A}, B$, C nezavisni. Slično vrijedi i za preostale kombinacije, pa je vjerojatnost pogotka točno dvaju strijelaca jednaka

$$\begin{aligned} &P\left((\overline{A}\cap B\cap C)\cup(A\cap\overline{B}\cap C)\cup(A\cap B\cap\overline{C})\right) \\ &= P(\overline{A}\cap B\cap C)+P(A\cap\overline{B}\cap C)+P(A\cap B\cap\overline{C}) \\ &= P(\overline{A})P(B)P(C)+P(A)P(\overline{B})P(C)+P(A)P(B)P(\overline{C}) \\ &= 0.2\cdot 0.6\cdot 0.5+0.8\cdot 0.4\cdot 0.5+0.8\cdot 0.6\cdot 0.5 = 0.46. \end{aligned}$$

□

Zadatak 3.5. Gospodin Šporkaćun ima u ponedjeljak ujutro u ladici 10 pari čarapa od kojih su 4 para prljavih. Oblači jedan par čarapa i navečer ga vraća u ladicu.

(a) Kolika je vjerojatnost da će u utorak obući čiste čarape?

(b) Kolika je vjerojatnost da će u srijedu obući čiste čarape?

(c) Ako je gospodin Šporkaćun u srijedu obukao čiste čarape, kolika je vjerojatnost da je u ponedjeljak i utorak obukao čiste čarape?

Rješenje:

(a) Vjerojatnost da gospodin Šporkaćun u utorak obuće čiste čarape ovisi o tome kakve je čarape obukao u ponedjeljak. Zato postavljamo sljedeće hipoteze i računamo njihove vjerojatnosti:

$$\begin{aligned} H_1 &= \{\text{ponedjeljak čiste}\}, \quad P(H_1)=\frac{6}{10}, \\ H_2 &= \{\text{ponedjeljak prljave}\}, \quad P(H_2)=\frac{4}{10}. \end{aligned}$$

Označimo s A traženi događaj. Ako je gospodin Šporkaćun u ponedjeljak obukao čiste čarape, onda je $P(A|H_1)=\frac{5}{10}$, a ako je obukao prljave, onda je $P(A|H_2)=\frac{6}{10}$. Uvrstimo li dobivene vjerojatnosti u formulu za potpunu vjerojatnost dobivamo

$$P(A)=P(H_1)\cdot P(A|H_1)+P(H_2)\cdot P(A|H_2)=\frac{6}{10}\cdot\frac{5}{10}+\frac{4}{10}\cdot\frac{6}{10}=0.54.$$

(b) U ovom slučaju tražena vjerojatnost ovisi o tome kakve će čarape gospodin Šporkaćun obući u ponedjeljak i utorak. Zbog toga imamo četiri hipoteze:

$$\begin{aligned}
H_1 &= \{\text{ponedjeljak i utorak čiste}\}, \quad P(H_1) = \frac{6}{10} \cdot \frac{5}{10} \\
H_2 &= \{\text{ponedjeljak čiste, utorak prljave}\}, \quad P(H_2) = \frac{6}{10} \cdot \frac{5}{10} \\
H_3 &= \{\text{ponedjeljak prljave, utorak čiste}\}, \quad P(H_3) = \frac{4}{10} \cdot \frac{6}{10} \\
H_4 &= \{\text{ponedjeljak i utorak prljave}\}, \quad P(H_4) = \frac{4}{10} \cdot \frac{4}{10}.
\end{aligned}$$

Ukoliko s B označimo traženi događaj, onda su uvjetne vjerojatnosti redom jednake:

$$P(B|H_1) = \frac{4}{10}, \quad P(B|H_2) = \frac{5}{10}, \quad P(B|H_3) = \frac{5}{10}, \quad P(B|H_4) = \frac{6}{10}.$$

Konačno, uvrstimo li dobivene podatke u formulu za potpunu vjerojatnost imamo da je

$$P(B) = \sum_{i=1}^{4} P(H_i)P(B|H_i) = 0.486.$$

(c) Uz oznake kao u (b) dijelu zadatka tražimo vjerojatnost $P(H_1|B)$. Prema Bayesovoj formuli imamo da je

$$P(H_1|B) = \frac{P(H_1)P(B|H_1)}{\sum_{i=1}^{4} P(H_i)P(B|H_i)} = \frac{0.3 \cdot 0.4}{0.486} = 0.247.$$

□

Zadatak 3.6. Utvrđeno je da 34% ljudi ima krvnu grupu `O`, 37% ljudi grupu `A`, 21% grupu `B` i 8% krvnu grupu `AB`. Kod transfuzije, osim svoje krvne grupe, osoba krvne grupe `AB` može primiti bilo koju grupu, osoba krvne grupe `A` ili `B` može primiti krvnu grupu `O`, a osoba krvne grupe `O` ne može ništa primiti osim svoje krvne grupe `O`.

(a) Odredite vjerojatnost da slučajno odabrana osoba može primiti krv od slučajno odabrane osobe.

(b) Ukoliko je osoba uspješno primila krv, kolika je vjerojatnost da ona ima krvnu grupu `AB`?

Rješenje:

(a) Neka je A traženi događaj, tj. $A = \{\text{slučajno odabrana osoba je uspješno primila krv}\}$. Imamo četiri moguće hipoteze:

$$\begin{aligned}
H_1 &= \{\text{slučajno odabrana osoba ima krvnu grupu } \texttt{O}\}, \\
H_2 &= \{\text{slučajno odabrana osoba ima krvnu grupu } \texttt{A}\}, \\
H_3 &= \{\text{slučajno odabrana osoba ima krvnu grupu } \texttt{B}\}, \\
H_4 &= \{\text{slučajno odabrana osoba ima krvnu grupu } \texttt{AB}\}.
\end{aligned}$$

Prema uvjetu zadatka je $P(H_1) = 0.34$, $P(H_2) = 0.37$, $P(H_3) = 0.21$ i $P(H_4) = 0.08$. Kako osoba krvne grupe `O` može primiti samo svoju krvnu grupu imamo da je $P(A|H_1) = 0.34$. Slično, imamo da je $P(A|H_2) = 0.34 + 0.37 = 0.71$, $P(A|H_3) = 0.34 + 0.21 = 0.55$ i $P(A|H_4) = 0.34 + 0.37 + 0.21 + 0.08 = 1$. Zbog toga je

$$P(A) = \sum_{i=1}^{4} P(H_i)P(A|H_i) = 0.574.$$

(b) Uz oznake kao u (a) dijelu zadatka tražimo vjerojatnost $P(H_4|A)$. Prema Bayesovoj formuli dobivamo

$$P(H_4|A) = \frac{P(H_4)P(A|H_4)}{\sum_{i=1}^{4} P(H_i)P(A|H_i)} = \frac{0.08 \cdot 1}{0.574} = 0.139.$$

$\square$

Zadatak 3.7. Ako su A, B i C nezavisni događaji, onda su A i $B \cup C$ također nezavisni. Dokažite!

Rješenje: Trebamo pokazati da vrijedi relacija $P(A \cap (B \cup C)) = P(A)P(B \cup C)$. Krenimo od lijeve strane jednakosti. Iskoristimo li distributivnost presjeka i unije događaja te formulu za vjerojatnost unije, imamo da je

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

Sada, kako su događaji A, B, C nezavisni, vrijedi $P(A \cap B) = P(A)P(B)$, $P(A \cap C) = P(A)P(C)$ i $P(A \cap B \cap C) = P(A)P(B)P(C)$, pa je

$$\begin{aligned} P(A \cap (B \cup C)) &= P(A)P(B) + P(A)P(C) - P(A)P(B)P(C) \\ &= P(A)\left[P(B) + P(C) - P(B)P(C)\right] \\ &= P(A)P(B \cup C), \end{aligned}$$

što je i trebalo dokazati. $\square$

4 Diskretne slučajne varijable

4.1 Diskretne slučajne varijable

U vezi sa slučajnim pokusima najčešće provodimo nekakva mjerenja, tj. svakom rezultatu slučajnog pokusa pridružujemo neki realan broj. Prema tome, važno je promatrati realne funkcije na vjerojatnosnom prostoru Ω, koje ćemo zvati **slučajne varijable**. Promotrimo, na primjer, slučajni pokus bacanja kocke. Tu je prirodno svakom elementarnom događaju pridružiti broj na koji je kocka pala. Time je definirana slučajna varijabla $\Omega \mapsto \{1, 2, 3, 4, 5, 6\}$.

Uz jedan slučajni pokus može biti povezano i više slučajnih varijabli. Na primjer, ako bacamo dvije kocke, onda se kao slučajne varijable pridružene tom pokusu mogu uzeti zbroj brojeva na kockama, apsolutna vrijednost razlike brojeva na kockama, veći od brojeva na kockama, razlika kvadrata brojeva na kockama itd.

Očito, u slučaju diskretnog vjerojatnosnog prostora slučajna varijabla može poprimiti najviše prebrojivo mnogo različitih vrijednosti. No, prirodno je pretpostaviti da neka slučajna varijabla može poprimiti sve realne vrijednosti ili vrijednosti iz nekog intervala. Prema tome, razlikujemo dvije vrste slučajnih varijabli: diskretne i neprekinute slučajne varijable. U ovom poglavlju proučavat ćemo diskretne slučajne varijable.

Neka je $S = \{x_1, x_2, \ldots\}$ podskup skupa prirodnih ili cijelih brojeva (konačan ili beskonačno prebrojiv). Promatrat ćemo slučajne varijable koje svakom elementarnom događaju pridružuju neku vrijednost iz skupa S. Neka je X preslikavanje sa skupa Ω u skup S. Naravno, uz to je preslikavanje prirodno postaviti pitanje "kolika je vjerojatnost da slučajna varijabla poprimi neku vrijednost x_i iz skupa S". Zbog toga skup

$$A_i := \{\omega \in \Omega : X(\omega) = x_i\}$$

mora biti događaj, odnosno element $\sigma-$algebre $\mathcal{F}$ svih događaja. Ako je taj uvjet ispunjen, za preslikavanje X ćemo reći da je slučajna varijabla.

Slučajna varijabla
Preslikavanje $X : \Omega \mapsto S$ je **diskretna slučajna varijabla** ako je za svaki $x_i \in S$ skup $A_i := \{\omega \in \Omega : X(\omega) = x_i\}$ događaj. Označimo $p_i := P(A_i) = P(X = x_i)$.
Za ove brojeve je $p_i > 0$ i $\sum p_i = 1$. **Zakon razdiobe** slučajne varijable X sastoji se od područja vrijednosti koje ona poprima i odgovarajućih vjerojatnosti. Pišemo

$$X \sim \begin{pmatrix} x_1 & x_2 & x_3 & \cdots \\ p_1 & p_2 & p_3 & \cdots \end{pmatrix}.$$

Pogledajmo sada u konkretnim slučajevima neke primjere slučajnih varijabli.

Primjer 4.1. *Bacamo tri igraće kocke. Neka slučajna varijabla X bilježi broj šestica. Odredite razdiobu te slučajne varijable.*

Rješenje: Vjerojatnosni prostor sastoji se od osam elementarnih događaja. Ispišimo ih, te naznačimo odgovarajuće vrijednosti slučajne varijable X:

$$\begin{array}{ll}
\omega_1 = \overline{6}\,\overline{6}\,\overline{6}, & X(\omega_1) = 0 \\
\omega_2 = \overline{6}\,\overline{6}\,6, & X(\omega_2) = 1 \\
\omega_3 = \overline{6}\,6\,\overline{6}, & X(\omega_3) = 1 \\
\omega_4 = \overline{6}\,6\,6, & X(\omega_4) = 2 \\
\omega_5 = 6\,\overline{6}\,\overline{6}, & X(\omega_5) = 1 \\
\omega_6 = 6\,\overline{6}\,6, & X(\omega_6) = 2 \\
\omega_7 = 6\,6\,\overline{6}, & X(\omega_7) = 2 \\
\omega_8 = 6\,6\,6, & X(\omega_8) = 3.
\end{array}$$

Pri tome 6 označava pojavljivanje šestice, a $\overline{6}$ komplementarni događaj. Vidimo da X poprima vrijednosti u skupu $\{0, 1, 2, 3\}$, a vjerojatnosti su

$$\begin{aligned}
p_1 &= P(X=0) = P(\omega_1) = \frac{125}{216}, \\
p_2 &= P(X=1) = P(\{\omega_2, \omega_3, \omega_5\}) = 3 \cdot \frac{25}{216} = \frac{75}{216}, \\
p_3 &= P(X=2) = P(\{\omega_4, \omega_6, \omega_7\}) = 3 \cdot \frac{5}{216} = \frac{15}{216}, \\
p_4 &= P(X=3) = P(\omega_8) = \frac{1}{216}.
\end{aligned}$$

Dakle, zakon razdiobe slučajne varijable X je $X \sim \begin{pmatrix} 0 & 1 & 2 & 3 \\ \frac{125}{216} & \frac{75}{216} & \frac{15}{216} & \frac{1}{216} \end{pmatrix}$.
Uočimo kao je zbroj svih vjerojatnosti u razdiobi jednak 1. □

Primjer 4.2. *U kutiji se nalazi* 7 *kuglica od kojih su* 3 *bijele. Slučajno odabiremo* 3 *kuglice te neka slučajna varijabla bilježi broj izvučenih bijelih kuglica. Odredite razdiobu slučajne varijable X.*

Rješenje: Slučajna varijabla X može poprimiti vrijednosti iz skupa $\{0, 1, 2, 3\}$, ovisno o tome koliko je bijelih kuglica izvučeno. Sada imamo redom

$$\begin{aligned}
p_1 &= P(X=0) = \frac{\binom{4}{3}}{\binom{7}{3}} = \frac{4}{35}, \\
p_2 &= P(X=1) = \frac{\binom{3}{1}\binom{4}{2}}{\binom{7}{3}} = \frac{18}{35}, \\
p_3 &= P(X=2) = \frac{\binom{3}{2}\binom{4}{1}}{\binom{7}{3}} = \frac{12}{35}, \\
p_4 &= P(X=3) = \frac{\binom{3}{3}}{\binom{7}{3}} = \frac{1}{35}.
\end{aligned}$$

Prema tome, zakon razdiobe slučajne varijable X je $X \sim \begin{pmatrix} 0 & 1 & 2 & 3 \\ \frac{4}{35} & \frac{18}{35} & \frac{12}{35} & \frac{1}{35} \end{pmatrix}$. □

Na razdiobu iz prethodnog primjera ponovo ćemo se vratiti u točki 4.3.

Nezavisne slučajne varijable

Promotrimo jednostavan pokus u kojemu bacamo dvije kocke. Neka X bilježi broj na prvoj kocki, a Y broj na drugoj kocki. Prirodno je pretpostaviti da dobiveni brojevi na kockama ne ovise jedan o drugome. Tako na primjer, vrijedi

$$P(X = 4, Y = 5) = P(\{4, 5\}) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = P(X = 4) \cdot P(Y = 5).$$

Slično je

$$P(X \leq 3, Y \geq 3) = \frac{12}{36} = \frac{3}{6} \cdot \frac{4}{6} = P(X \leq 3) \cdot P(Y \geq 3)$$

ili

$$P(X = 3, 3 \leq Y \leq 5) = \frac{3}{36} = \frac{1}{6} \cdot \frac{3}{6} = P(X = 3) \cdot P(3 \leq Y \leq 5).$$

Navedeni primjeri upućuju nas na prirodnu definiciju nezavisnosti slučajnih varijabli.

Nezavisne slučajne varijable - definicija
Slučajne varijable $X, Y : \Omega \mapsto S$ su **nezavisne** ako za sve $x_i, y_j \in S$ vrijedi

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j).$$

Iz prethodne definicije lako se izvodi važno svojstvo nezavisnih slučajnih varijabli.

Nezavisne slučajne varijable - osnovno svojstvo
Ako su $X, Y : \Omega \mapsto S$ nezavisne slučajne varijable, onda za sve $A, B \subset S$ vrijedi relacija

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

Definicija nezavisnosti proširuje se i na skup od konačno ili prebrojivo mnogo slučajnih varijabli:

Nezavisnost niza slučajnih varijabli
Slučajne varijable $X_1, X_2, \ldots, X_n$ definirane na istom vjerojatnosnom prostoru su nezavisne, ako za sve $A_1, A_2, \ldots, A_n \subset S$ vrijedi

$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n) = P(X_1 \in A_1)P(X_2 \in A_2) \cdots P(X_n \in A_n).$$

Primjer 4.3. *Bacamo novčić sve dok se ne pojavi pismo. Neka slučajna varijabla X označava potreban broj bacanja. Odredite zakon razdiobe slučajne varijable X.*

Rješenje: Označimo sa X_i slučajne varijable: rezultate $i-$tog bacanja. Preciznije,

$$X_i \sim \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

pri čemu slučajna varijabla X_i poprima vrijednost 1 ako se pojavilo pismo, te vrijednost 0 inače. Očito, slučajne varijable X_i su nezavisne i jednako distribuirane. Sada imamo:

$$\begin{aligned} P(X = n) &= P(X_1 = 0, X_2 = 0, \ldots, X_{n-1} = 0, X_n = 1) \\ &= P(X_1 = 0) \cdot P(X_2 = 0) \cdots P(X_{n-1} = 0) \cdot P(X_n = 1) \\ &= \left(\frac{1}{2}\right)^{n-1} \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^{n}, \quad n \geq 1. \end{aligned}$$

Dakle, zakon razdiobe slučajne varijable X glasi

$$X \sim \begin{pmatrix} 1 & 2 & \dots & n & \dots \\ \frac{1}{2} & \frac{1}{4} & \dots & (\frac{1}{2})^n & \dots \end{pmatrix}.$$

□

Funkcije diskretnih slučajnih varijabli

Neka je X diskretna slučajna varijabla s poznatim zakonom razdiobe, $\phi : \mathbf{R} \mapsto \mathbf{R}$ zadana funkcija i $Y = \phi(X)$. Ako je

$$X \sim \begin{pmatrix} x_1 & x_2 & \dots \\ p_1 & p_2 & \dots \end{pmatrix}$$

zakon razdiobe varijable X, tada je

$$Y \sim \begin{pmatrix} \phi(x_1) & \phi(x_2) & \dots \\ p_1 & p_2 & \dots \end{pmatrix}$$

zakon razdiobe varijable Y. Njega dovodimo na reducirani oblik

$$Y \sim \begin{pmatrix} y_1 & y_2 & \dots \\ q_1 & q_2 & \dots \end{pmatrix},$$

gdje su $y_1, y_2, \dots$ sve različite vrijednosti iz skupa $\{\phi(x_1), \phi(x_2), \dots\}$, pri čemu se odgovarajuće vjerojatnosti zbrajaju. Pogledajmo konkretan primjer:

Primjer 4.4. *Slučajna varijabla X ima zakon razdiobe*

$$X \sim \begin{pmatrix} -3 & -2 & 2 & 3 \\ 0.5 & 0.1 & 0.2 & 0.2 \end{pmatrix}$$

Odredite zakon razdiobe varijable $Y = X^4$.

Rješenje:

$$Y \sim \begin{pmatrix} 81 & 16 & 16 & 81 \\ 0.5 & 0.1 & 0.2 & 0.2 \end{pmatrix} = \begin{pmatrix} 16 & 81 \\ 0.3 & 0.7 \end{pmatrix}$$

□

Primjer 4.5. *Nezavisne slučajne varijable X_1 i X_2 imaju isti zakon razdiobe*

$$X_1, X_2 \sim \begin{pmatrix} -1 & 0 & 1 \\ 0.2 & 0.7 & 0.1 \end{pmatrix}.$$

Odredite zakon razdiobe slučajnih varijabli $Y = X_1 + X_2$ i $Z = X_1 X_2$.

Rješenje: Slučajna varijabla Y poprima vrijednosti u skupu $\{-2, -1, 0, 1, 2\}$ s vjerojatnostima

$$\begin{aligned}
P(Y = -2) &= P(X_1 = -1, X_2 = -1) = 0.2 \cdot 0.2 = 0.04, \\
P(Y = -1) &= P(X_1 = -1, X_2 = 0) + P(X_1 = 0, X_2 = -1) = 2 \cdot 0.2 \cdot 0.7 = 0.28, \\
P(Y = 0) &= P(X_1 = -1, X_2 = 1) + P(X_1 = 0, X_2 = 0) + P(X_1 = 1, X_2 = -1) \\
&= 2 \cdot 0.2 \cdot 0.1 + 0.7 \cdot 0.7 = 0.53, \\
P(Y = 1) &= P(X_1 = 0, X_2 = 1) + P(X_1 = 1, X_2 = 0) = 2 \cdot 0.7 \cdot 0.1 = 0.14, \\
P(Y = 2) &= P(X_1 = 1, X_2 = 1) = 0.1 \cdot 0.1 = 0.01.
\end{aligned}$$

Dakle, razdioba slučajne varijable Y je

$$Y \sim \begin{pmatrix} -2 & -1 & 0 & 1 & 2 \\ 0.04 & 0.28 & 0.53 & 0.14 & 0.01 \end{pmatrix}.$$

Slučajna varijabla Z poprima vrijednosti u skupu $\{-1, 0, 1\}$ s vjerojatnostima

$$\begin{aligned} P(Z = -1) &= P(X_1 = -1, X_2 = 1) + P(X_1 = 1, X_2 = -1) = 2 \cdot 0.2 \cdot 0.1 = 0.04, \\ P(Z = 0) &= P(X_1 = 0) + P(X_1 \neq 0, X_2 = 0) = 0.7 + 0.3 \cdot 0.7 = 0.91, \\ P(Z = 1) &= P(X_1 = -1, X_2 = -1) + P(X_1 = 1, X_2 = 1) = 0.2^2 + 0.1^2 = 0.05. \end{aligned}$$

Prema tome, razdioba od Z je

$$Z \sim \begin{pmatrix} -1 & 0 & 1 \\ 0.04 & 0.91 & 0.05 \end{pmatrix}.$$

□

4.2 Numeričke karakteristike diskretnih slučajnih varijabli

Očekivanje slučajne varijable

Već smo prije istaknuli kako kod slučajnih pokusa najčešće vršimo nekakva mjerenja. Preciznije, slučajnim varijablama pridružujemo određene **numeričke karakteristike**, pomoću kojih opisujemo njihova svojstva. Najvažnija karakteristika je **očekivanje**.

Očekivanje slučajne varijable
Neka slučajna varijabla X ima zakon razdiobe:

$$X \sim \begin{pmatrix} x_1 & x_2 & x_3 & \dots \\ p_1 & p_2 & p_3 & \dots \end{pmatrix}.$$

Očekivanje slučajne varijable X definirano je kao suma

$$E(X) = \sum_i x_i p_i.$$

Često se očekivanje slučajne varijable označava i simbolima $\overline{x}$ ili m_X.

Na primjer, ako bacamo igraću kocku, a slučajna varijabla X bilježi broj na kocki, onda je njena razdioba

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

a pripadno očekivanje

$$E(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{7}{2}.$$

Vidimo da očekivanje slučajne varijable ne mora biti jednako nekoj od njenih mogućih vrijednosti. Osim toga, očekivanje ne mora biti blisko niti realizaciji s najvećom vjerojatnosti. Na primjer, za slučajnu varijablu

$$Y \sim \begin{pmatrix} 1 & 2 & 3 & 100 \\ 0.7 & 0.1 & 0.1 & 0.1 \end{pmatrix}$$

vrijedi

$$E(Y) = 1 \cdot 0.7 + 2 \cdot 0.1 + 3 \cdot 0.1 + 100 \cdot 0.1 = 11.2.$$

Prema tome, očekivanje možemo shvatiti kao srednju vrijednost, odnosno težinsku aritmetičku sredinu gdje su težine jednake vjerojatnostima pojedinih realizacija slučajne varijable.

Ako slučajna varijabla ima prebrojivo mnogo realizacija, onda njeno očekivanje ne mora postojati. Na primjer, za slučajnu varijablu

$$Z \sim \begin{pmatrix} 1 & 3 & 9 & \dots & 3^{n-1} & \dots \\ \frac{2}{3} & \frac{2}{9} & \frac{2}{27} & \dots & \frac{2}{3^n} & \dots \end{pmatrix}$$

vrijedi

$$E(Z) = 1 \cdot \frac{2}{3} + 3 \cdot \frac{2}{9} + 9 \cdot \frac{2}{27} + \ldots + 3^{n-1} \cdot \frac{2}{3^n} + \ldots = \frac{2}{3} + \frac{2}{3} + \frac{2}{3} + \ldots = +\infty.$$

Kako pripadni red divergira, slučajna varijabla Z nema očekivanja.

Primjer 4.6. *U kutiji se nalazi pet kuglica od kojih su dvije bijele. Izvlačimo na sreću po jednu kuglicu, bez vraćanja, sve dok ne izvučemo bijelu kuglicu. Neka X označava pokušaj u kojemu je izvučena bijela kuglica. Izračunajte očekivanje slučajne varijable X.*

Rješenje: Najprije trebamo naći razdiobu slučajne varijable X. Očito, slučajna varijabla X poprima vrijednosti iz skupa $\{1, 2, 3, 4\}$, jer će najkasnije u četvrtom pokušaju biti izvučena bijela kuglica. Sada imamo:

$$\begin{aligned} P(X = 1) &= \frac{2}{5}, \\ P(X = 2) &= \frac{3}{5} \cdot \frac{2}{4} = \frac{3}{10}, \\ P(X = 3) &= \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} = \frac{1}{5}, \\ P(X = 4) &= \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{1}{3} \cdot \frac{2}{2} = \frac{1}{10}, \end{aligned}$$

pa je zakon razdiobe

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

Zbog toga je očekivanje jednako

$$E(X) = 1 \cdot \frac{2}{5} + 2 \cdot \frac{3}{10} + 3 \cdot \frac{1}{5} + 4 \cdot \frac{1}{10} = 1.$$

□

Lagano se pokazuju sljedeća dva važna svojstva očekivanja:

Svojstva očekivanja

Teorem 4.1 *Neka su X i Y slučajne varijable definirane na istom vjerojatnosnom prostoru. Očekivanje ima svojstvo linearnosti, za sve realne brojeve α i β vrijedi*

$$E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y).$$

Ako su slučajne varijable X i Y nezavisne, onda vrijedi

$$E(XY) = E(X)E(Y).$$

U prethodnoj točki smo tražili razdiobu slučajne varijable $Y = \phi(X)$ koja je funkcija slučajne varijable X s poznatom razdiobom. Kako ćemo odrediti očekivanje od Y? Jedna je mogućnost da odredimo razdiobu od Y i zatim primijenimo formulu za očekivanje na varijablu Y.

Primjer 4.7. *Zadana je slučajna varijabla* $X \sim \begin{pmatrix} -3 & -1 & 0 & 1 & 3 \\ \frac{1}{3} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{3} \end{pmatrix}$. *Izračunajte* $E(Y)$ *ako je* $Y = X^2$.

Rješenje: Odredimo razdiobu od Y:
$Y = X^2 \sim \begin{pmatrix} (-3)^2 & (-1)^2 & 0^2 & 1^2 & 3^2 \\ \frac{1}{3} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{3} \end{pmatrix} \sim \begin{pmatrix} 9 & 1 & 0 & 1 & 9 \\ \frac{1}{3} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{3} \end{pmatrix} \sim \begin{pmatrix} 0 & 1 & 9 \\ \frac{1}{9} & \frac{2}{9} & \frac{2}{3} \end{pmatrix}$.
Zato je

$$E(Y) = 0 \cdot \frac{1}{9} + 1 \cdot \frac{2}{9} + 9 \cdot \frac{2}{3} = \frac{56}{9}.$$

□

Uočimo kako je u ovom primjeru $E(X) = 0$, pa je $E(X^2) \neq E(X)^2$. Dakle, svojstvo umnoška očekivanja iz Teorema 4.1. ne vrijedi za bilo koje slučajne varijable.

Druga mogućnost za računanje očekivanja funkcije slučajne varijable je upotreba formule

$$E(\phi(X)) = \sum_i \phi(x_i) p_i.$$

Do nje dolazimo tako da u gornjem postupku ne svodimo zakon razdiobe varijable Y na reducirani oblik, nego očekivanje računamo iz nesređenog oblika. Iz

$$Y = X^2 \sim \begin{pmatrix} 9 & 1 & 0 & 1 & 9 \\ \frac{1}{3} & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} & \frac{1}{3} \end{pmatrix}$$

dobivamo

$$E(Y) = 9 \cdot \frac{1}{3} + 1 \cdot \frac{1}{9} + 0 \cdot \frac{1}{9} + 1 \cdot \frac{1}{9} + 9 \cdot \frac{1}{3} = \frac{56}{9}.$$

Dakle, ukoliko nas zanima samo očekivanje funkcije slučajne varijable, a ne njena razdioba, slučajnu varijablu nećemo svoditi na reducirani oblik, nego ćemo direktno računati očekivanje.

U nastavku ćemo promatrati funkciju oblika $\phi(x) = (x - a)^2$ koja će nam dati još jednu važnu numeričku karakteristiku slučajne varijable.

Disperzija slučajne varijable

Disperzija slučajne varijable je numerička karakteristika koja, u određenom smislu, mjeri odstupanje od srednje vrijednosti, odnosno očekivanja. Navedimo precizno definiciju te veličine.

Disperzija slučajne varijable
Disperzija (rasipanje, varijanca) slučajne varijable X definira se formulom

$$D(X) = E\left[(X - E(X))^2\right].$$

Formula iz prethodne definicije nepraktična je za računanje pa ćemo je transformirati. Označimo li $E(X) = m_X$, zbog linearnosti očekivanja vrijedi relacija

$$\begin{aligned} E\left[(X - m_X)^2\right] &= E\left[X^2 - 2Xm_X + m_X^2\right] \\ &= E(X^2) - 2m_X E(X) + E(m_X^2) \\ &= E(X^2) - m_X^2, \end{aligned}$$

pri čemu smo koristili činjenicu da je očekivanje konstantne slučajne varijable jednako toj konstanti.

Računanje disperzije
Disperziju slučajne varijable X računamo po formuli

$$D(X) = E(X^2) - m_X^2 = \sum_i x_i^2 p_i - \left(\sum_i x_i p_i \right)^2 .$$

Disperzija slučajne varijable ima neka važna svojstva, koja ćemo iskazati u sljedećem teoremu.

Svojstva disperzije
Teorem 4.2 *Za slučajnu varijablu X i realni broj α vrijedi*

$$D(\alpha X) = \alpha^2 D(X).$$

Ako su slučajne varijable X i Y nezavisne, onda vrijedi

$$D(X+Y) = D(X) + D(Y).$$

Dokaz: Iskoristimo li svojstva očekivanja, za prvu formulu dobivamo

$$D(\alpha X) = E\left[(\alpha X)^2\right] - E\left[(\alpha X)\right]^2 = E(\alpha^2 X^2) - \left[\alpha E(X)\right]^2 = \alpha^2 E(X^2) - \alpha^2 \left[E(X)\right]^2 = \alpha^2 D(X).$$

Nadalje, ako su X i Y nezavisne, onda vrijedi

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

□

Svojstva očekivanja i disperzije moraju se dobro usvojiti. Zbog toga ćemo ponajprije razmotriti sljedeći primjer.

Primjer 4.8. *Nezavisne slučajne varijable X i Y imaju identičnu razdiobu s očekivanjem 2 i disperzijom 5. Kolika je disperzija slučajne varijable $2X + 3Y$? Koliko je očekivanje, a kolika disperzija slučajne varijable $3X - 3Y$?*

Rješenje: U prvom slučaju vrijedi

$$D(2X + 3Y) = 2^2 \cdot D(X) + 3^2 \cdot D(Y) = 13 \cdot 5 = 65.$$

U drugom slučaju je

$$E(3X - 3Y) = 3E(X) - 3E(Y) = 6 - 6 = 0,$$
$$D(3X - 3Y) = 3^2 D(X) + (-3)^2 D(Y) = 9D(X) + 9D(Y) = 90.$$

□

Uočimo kako je disperzija slučajne varijable uvijek pozitivna. Ta činjenica slijedi direktno iz formule

$$D(X) = E\left[(X - m_X)^2\right] = \sum_i (x_i - m_X)^2 p_i,$$

zato jer su svi pribrojnici u sumi nenegativni. Pogledajmo kada će disperzija biti jednaka nuli. Naime, iz prethodne formule zaključujemo da je to moguće samo ako je $x_i = m_X$ za svaki i, a to znači da se sve realizacije slučajne varijable X podudaraju. Drugim riječima, tada X nije slučajna, nego uvijek poprima istu vrijednost.

Veličinu $\sigma_X = \sqrt{D(X)}$ nazivamo **standardna devijacija (odstupanje)** varijable X.

Primjer 4.9. *Zadane su nezavisne slučajne varijable*

$$X \sim \begin{pmatrix} -10 & -5 & -2 \\ 0.3 & 0.2 & 0.5 \end{pmatrix}, \qquad Y \sim \begin{pmatrix} 5 & 10 \\ 0.4 & 0.6 \end{pmatrix}.$$

Izračunajte očekivanje slučajne varijable XY *i disperziju slučajne varijable* $2X - 3Y$.

Rješenje: Kako su X i Y nezavisne slučajne varijable, iz svojstava očekivanja i disperzije vidimo da ne trebamo tražiti razdiobe slučajnih varijabli XY i $2X - 3Y$. Dovoljno je izračunati očekivanje i disperziju varijabli X i Y. Imamo redom:

$$\begin{aligned} &E(X) = -10 \cdot 0.3 - 5 \cdot 0.2 - 2 \cdot 0.5 = -5, \\ &E(X^2) = 100 \cdot 0.3 + 25 \cdot 0.2 + 4 \cdot 0.5 = 37, \\ &D(X) = E(X^2) - E(X)^2 = 37 - (-5)^2 = 12, \\ &E(Y) = 5 \cdot 0.4 + 10 \cdot 0.6 = 8, \\ &E(Y^2) = 25 \cdot 0.4 + 100 \cdot 0.6 = 70, \\ &D(Y) = E(Y^2) - E(Y)^2 = 70 - 8^2 = 6. \end{aligned}$$

Sada, zbog nezavisnosti slučajnih varijabli, dobivamo da je

$$E(XY) = E(X)E(Y) = (-5) \cdot 8 = -40$$

i

$$D(2X - 3Y) = 4D(X) + 9D(Y) = 102.$$

□

Primjer 4.10. *Slučajno odabiremo dva broja iz skupa* $\{1, 2, 3, 4, 5, 6, 7\}$*, pri čemu isti broj možemo odabrati dva puta. Neka je* X *maksimalna vrijednost od ta dva, na sreću odabrana broja. Odredite očekivanje od* X.

Rješenje: Vjerojatnosni prostor sastoji se od 49 jednako vjerojatnih elementarnih događaja. Razdiobu slučajne varijable lagano ćemo odrediti služeći se sljedećom tablicom:

max	1	2	3	4	5	6	7
1	1	2	3	4	5	6	7
2	2	2	3	4	5	6	7
3	3	3	3	4	5	6	7
4	4	4	4	4	5	6	7
5	5	5	5	5	5	6	6
6	6	6	6	6	6	6	7
7	7	7	7	7	7	7	7

Iz tablice vidimo da slučajna varijabla X poprima vrijednost 1 na jednom elementarnom događaju, vrijednost 2 na tri elementarna događaja, itd. Zbog toga, njena razdioba glasi

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \frac{1}{49} & \frac{3}{49} & \frac{5}{49} & \frac{7}{49} & \frac{9}{49} & \frac{11}{49} & \frac{13}{49} \end{pmatrix}$$

Prema tome, traženo očekivanje je

$$E(X) = 1 \cdot \frac{1}{49} + 2 \cdot \frac{3}{49} + 3 \cdot \frac{5}{49} + 4 \cdot \frac{7}{49} + 5 \cdot \frac{9}{49} + 6 \cdot \frac{11}{49} + 7 \cdot \frac{13}{49} = \frac{36}{7}$$

□

Primjer 4.11. *Bacamo dva simetrična tetraedra kojima su na stranama napisani brojevi od* 1 *do* 4*. Neka slučajna varijabla* X *označava umnožak brojeva na donjim stranama tetraedara na koje su oni pali. Odredite očekivanje i disperziju slučajne varijable* X*.*

Rješenje: Vjerojatnosni prostor sastoji se od 16 jednako vjerojatnih elementarnih događaja. Odredimo vrijednost slučajne varijable X na tim događajima:

X	1	2	3	4
1	1	2	3	4
2	2	4	6	8
3	3	6	9	12
4	4	8	12	16

Iz tablice vidimo da je razdioba slučajne varijable X dana s

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 6 & 8 & 9 & 12 & 16 \\ \frac{1}{16} & \frac{2}{16} & \frac{2}{16} & \frac{3}{16} & \frac{2}{16} & \frac{2}{16} & \frac{1}{16} & \frac{2}{16} & \frac{1}{16} \end{pmatrix}.$$

Sada računamo

$$\begin{aligned} E(X) &= 1 \cdot \frac{1}{16} + 2 \cdot \frac{2}{16} + \cdots + 16 \cdot \frac{1}{16} = 6.25, \\ D(X) &= 1 \cdot \frac{1}{16} + 4 \cdot \frac{2}{36} + \cdots + 256 \cdot \frac{1}{16} - 6.25^2 = 56.25 - 6.25^2 = 17.1875. \end{aligned}$$

Uočimo kako smo očekivanje ove slučajne varijable mogli dobiti jednostavnije. U tu svrhu, definirajmo nezavisne slučajne varijable:

$$\begin{aligned} X_1 &= \{\text{broj na prvom tetraedru}\}, \\ X_2 &= \{\text{broj na drugom tetraedru}\}. \end{aligned}$$

Očito su X_1 i X_2 identički distribuirane, ali i nezavisne slučajne varijable, s razdiobom

$$X_1, X_2 \sim \begin{pmatrix} 1 & 2 & 3 & 4 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$

Stoga su i očekivanja tih dviju slučajnih varijabli jednaka:

$$E(X_1) = E(X_2) = 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{4} = 2.5.$$

Konačno, kako je $X = X_1 X_2$, slijedi da je

$$E(X) = E(X_1 X_2) = E(X_1)E(X_2) = 2.5^2 = 6.25,$$

što smo dobili i prije. □

4.3 Primjeri diskretnih razdioba

U ovoj točki proučavamo razdiobe nekih važnih diskretnih slučajnih varijabli. Na početku ćemo kratko spomenuti dvije razdiobe s kojima smo se već prije susreli.

Jednolika razdioba

Kažemo da diskretna slučajna varijabla ima **jednoliku** ili **uniformnu razdiobu** ako ona poprima konačno mnogo vrijednosti s jednakim vjerojatnostima. Preciznije, ako slučajna varijabla X s vrijednostima u skupu $\{1, 2, \ldots, n\}$ ima jednoliku razdiobu, onda je ona dana sa

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

S ovom razdiobom smo se već susreli, na primjer kod bacanja kocke. U tom slučaju, slučajna varijabla poprima vrijednosti iz skupa $\{1, 2, 3, 4, 5, 6\}$ s vjerojatnostima $\frac{1}{6}$.

Primjer 4.12. *U kutiji se nalazi* 10 *kuglica od kojih je samo jedna bijela. Izvlačimo na sreću jednu po jednu kuglicu iz kutije, bez vraćanja. Neka X označava pokušaj u kojem je izvučena bijela kuglica. Odredite razdiobu i očekivanje od X.*

Rješenje: Slučajna varijabla X poprima vrijednosti iz skupa $\{1, 2, 3, \ldots, 10\}$, ovisno o tome u kojem je pokušaju izvučena bijela kuglica. Izračunajmo pripadne vjerojatnosti:

$$\begin{aligned}
P(X = 1) &= \frac{1}{10}, \\
P(X = 2) &= \frac{9}{10} \cdot \frac{1}{9} = \frac{1}{10}, \\
P(X = 3) &= \frac{9}{10} \cdot \frac{8}{9} \cdot \frac{1}{8} = \frac{1}{10}, \\
&\vdots \\
P(X = 10) &= \frac{9}{10} \cdot \frac{8}{9} \cdot \frac{7}{8} \cdot \frac{6}{7} \cdot \frac{5}{6} \cdot \frac{4}{5} \cdot \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{1}{2} \cdot \frac{1}{1} = \frac{1}{10}.
\end{aligned}$$

Prema tome, X ima jednoliku razdiobu

$$X \sim \begin{pmatrix} 1 & 2 & \cdots & 10 \\ \frac{1}{10} & \frac{1}{10} & \cdots & \frac{1}{10} \end{pmatrix},$$

pa je traženo očekivanje jednako

$$E(X) = \frac{1 + 2 + \cdots + 10}{10} = 5.5.$$

□

Hipergeometrijska razdioba

Neka su m i n prirodni brojevi takvi da je $n \leq m$ te neka je $r \in \{0, 1, 2, \ldots, m\}$. Kažemo da slučajna varijabla X ima **hipergeometrijsku razdiobu** ako ona poprima vrijednosti u skupu $\{0, 1, 2, \ldots, n\}$ s vjerojatnostima

$$P(X = k) = \frac{\binom{r}{k}\binom{m-r}{n-k}}{\binom{m}{n}}, \quad k \in \{0, 1, 2, \ldots, n\}.$$

Ovu razdiobu smo već promatrali u Primjeru 4.2, gdje smo iz skupa od 7 kuglica, među kojima su 3 bijele, odabirali 3 kuglice. Tada slučajna varijabla X koja bilježi broj bijelih kuglica ima hipergeometrijsku razdiobu s parametrima $m = 7$, $n = r = 3$. Pogledajmo općeniti model u kojemu se pojavljuje hipergeometrijska razdioba.

Primjer 4.13. *U kutiji se nalazi m proizvoda od kojih je r oštećenih. Slučajno odabiremo n proizvoda te neka slučajna varijabla X bilježi broj oštećenih proizvoda među odabranima. Pokažite da X ima hipergeometrijsku razdiobu.*

Rješenje: Izračunajmo vjerojatnost da među odabranih n proizvoda ima točno k oštećenih, $k \in \{0, 1, 2, \ldots, n\}$. Očito vrijedi

$$P(X = k) = \frac{\binom{r}{k}\binom{m-r}{n-k}}{\binom{m}{n}},$$

pa X ima hipergeometrijsku razdiobu. □

U nastavku ćemo detaljnije proučavati svojstva triju važnih diskretnih razdioba. Prva od njih je geometrijska razdioba.

Geometrijska razdioba

Neka je pri izvođenju nekog pokusa vjerojatnost realizacije događaja A jednaka p. Ponavljamo taj pokus u nepromijenjenim uvjetima do prve realizacije tog događaja. Neka slučajna varijabla X mjeri broj pokusa u kojem se realizirao događaj A. Onda kažemo da X ima **geometrijsku razdiobu** s parametrom p i pišemo $X \sim \mathcal{G}(p)$.

Uočimo kako smo se s opisanim modelom već sreli u točki 2.5, gdje smo proučavali beskonačne prebrojive vjerojatnosne prostore.

Primjer 4.14. *Odredite razdiobu slučajne varijable $X \sim \mathcal{G}(p)$.*

Rješenje: Ispišimo elementarne događaje, vrijednost slučajne varijable i pripadne vjerojatnosti u ovom pokusu:

$$\begin{array}{lll}
\omega_1 = A, & P(\omega_1) = p, & X(\omega_1) = 1, \\
\omega_2 = \overline{A}A, & P(\omega_2) = qp, & X(\omega_2) = 2, \\
\omega_3 = \overline{A}\,\overline{A}A, & P(\omega_3) = q^2p, & X(\omega_3) = 3, \\
\vdots & & \\
\omega_n = \overline{A}\cdots\overline{A}A, & P(\omega_n) = q^{n-1}p, & X(\omega_n) = n, \\
\vdots & &
\end{array}$$

Pri tome je $q = 1 - p$. Dakle, zakon razdiobe glasi

$$X \sim \begin{pmatrix} 1 & 2 & 3 & \cdots & n & \cdots \\ p & qp & q^2p & \cdots & q^{n-1}p & \cdots \end{pmatrix}.$$

Uočimo kako je, prema formuli za sumu geometrijskog reda,

$$p + qp + q^2p + \cdots + q^{n-1}p + \cdots = p\left(1 + q + q^2 + \cdots\right) = p \cdot \frac{1}{1-q} = \frac{p}{p} = 1.$$

□

Ponovimo još jednom. Kod geometrijske razdiobe je

$$p_k = P(X = k) = p(1-p)^{k-1}, \qquad k \in \{1, 2, 3, \ldots\},$$

jer se u prvih $k - 1$ pokusa događaj A nije realizirao, a pojavio se u $k-$tom pokusu. Primijetimo da vrijedi (uz uobičajenu oznaku $q = 1 - p$)

$$P(X > k) = (1-p)^k = q^k,$$

jer se tada događaj A nije ostvario u prvih k pokusa.

Odredimo očekivanje geometrijske razdiobe s parametrom p. Dakle, trebamo odrediti sumu beskonačnog reda

$$E(X) = \sum_{k=1}^{\infty} k \cdot pq^{k-1} = p \sum_{k=1}^{\infty} kq^{k-1}.$$

Da bismo odredili tu sumu, krenut ćemo od formule za sumu beskonačnog geometrijskog reda:

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}, \qquad |x| < 1.$$

Deriviranjem prethodnog izraza po varijabli x, dobivamo formulu

$$\sum_{k=1}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2}, \qquad |x| < 1.$$

Uvrstimo li u prethodnu formulu $x = q$, imamo da je $\sum_{k=1}^{\infty} kq^{k-1} = \frac{1}{(1-q)^2} = \frac{1}{p^2}$. Zbog toga je očekivanje geometrijske razdiobe s parametrom p jednako

$$E(X) = p \sum_{k=1}^{\infty} kq^{k-1} = p \cdot \frac{1}{p^2} = \frac{1}{p}.$$

Na sličan način se pokazuje da za disperziju geometrijske razdiobe vrijedi $D(X) = \frac{q}{p^2}$.

Geometrijska razdioba – očekivanje i disperzija
Očekivanje i disperzija slučajne varijable $X \sim \mathcal{G}(p)$ dani su sljedećim formulama:

$$E(X) = \frac{1}{p}, \qquad D(X) = \frac{q}{p^2}.$$

Uočimo kako je očekivanje geometrijske razdiobe jednako recipročnoj vrijednosti parametra – vjerojatnosti p. Taj je rezultat u skladu s iskustvom. Na primjer, pri bacanju kocke, vjerojatnost pojavljivanja šestice jednaka je $p = 1/6$. Broj bacanja kocke do pojave šestice mjeri slučajna varijabla koja ima geometrijsku razdiobu. Očekivani broj ponavljanja je $E(X) = 1/p = 6$.

Primjer 4.15. *Kolika je vjerojatnost da će se šestica pojaviti u prva tri bacanja kocke?*

Rješenje: Odgovor na ovo pitanje daje zakon razdiobe slučajne varijable:

$$P(X \leq 3) = 1 - P(X > 3) = 1 - \left(\frac{5}{6}\right)^3 = 0.421.$$

□

Ukoliko se šestica nije pojavila u prva tri bacanja kocke, vjerojatnost da se ona pojavi u iduća tri bacanja ostaje ista. To je važno svojstvo koje karakterizira geometrijsku razdiobu. Iskažimo precizno to svojstvo.

Odsustvo pamćenja – karakterizacija geometrijske razdiobe
Teorem 4.3 *Slučajna varijabla X koja poprima vrijednosti u skupu $\{1, 2, 3, \dots\}$ ima geometrijsku razdiobu onda i samo onda ako za sve $k, m \geq 1$ vrijedi:*

$$P(X = k + m | X > k) = P(X = m).$$

Ovaj teorem ukazuje na činjenicu da kod geometrijske razdiobe nije važno što je bilo prije. Na primjer, ako se šestica nije pojavila u prvih pet bacanja, očekivani broj (novih) bacanja do njezine pojave opet je jednak 6. Kažemo da geometrijska razdioba nema pamćenje.

Primjer 4.16. *Slučajni pokus sastoji se od istovremenog bacanja simetričnog novčića i kocke. Pokus ponavljamo sve dok se ne pojavi pismo ili šestica. Slučajna varijabla bilježi redni broj pokusa u kojem se ostvario taj događaj. Dokažite da slučajna varijabla X ima geometrijsku razdiobu, te joj odredite parametar.*

Rješenje: Neka je X_1 slučajna varijabla koja bilježi redni broj pokusa u kojem se prvi put pojavilo pismo. Tada X_1 ima geometrijsku razdiobu i vrijedi $X_1 \sim \mathcal{G}(\frac{1}{2})$. Slično, ako je X_2 slučajna varijabla koja bilježi redni broj pokusa u kojem se pojavila šestica, onda je $X_2 \sim \mathcal{G}(\frac{1}{6})$.

Uočimo kako se slučajna varijabla X koja registrira prvo pojavljivanje bilo kojeg od ovih dvaju događaja može zapisati formulom $X = \min\{X_1, X_2\}$.

Sada, zbog nezavisnosti varijabli X_1 i X_2 vrijedi

$$\begin{aligned} P(X > k) &= P(\min\{X_1, X_2\} > k) = P(X_1 > k, X_2 > k) = P(X_1 > k)P(X_2 > k) \\ &= \left(1 - \frac{1}{2}\right)^k \left(1 - \frac{1}{6}\right)^k = \left(\frac{5}{12}\right)^k, \quad k = 1, 2, 3, \ldots \end{aligned}$$

Nadalje, očito vrijedi $P(X > k - 1) = P(X = k) + P(X > k)$, odakle dobivamo da je

$$\begin{aligned} P(X = k) &= P(X > k - 1) - P(X > k) = \left(\frac{5}{12}\right)^{k-1} - \left(\frac{5}{12}\right)^k = \left(1 - \frac{5}{12}\right)\left(\frac{5}{12}\right)^{k-1} \\ &= \frac{7}{12} \cdot \left(\frac{5}{12}\right)^{k-1}. \end{aligned}$$

Konačno, iz prethodne relacije zaključujemo da slučajna varijabla X ima geometrijsku razdiobu s parametrom $p = \dfrac{7}{12}$. □

U prethodnom primjeru vidjeli smo slučajni pokus s dva nezavisna obilježja. Dakako, takva razmatranja se mogu proširiti i na više pokusa s nezavisnim obilježjima.

Binomna razdioba

Najvažnija diskretna razdioba je **binomna razdioba**. Neka je p vjerojatnost realizacije događaja A pri izvođenju nekog pokusa. Pretpostavimo da isti pokus ponavljamo n puta. Neka slučajna varijabla X mjeri broj pojavljivanja događaja A. Onda kažemo da X ima binomnu razdiobu s parametrima n i p te pišemo $X \sim \mathcal{B}(n, p)$.

Odredimo razdiobu te slučajne varijable. X poprima vrijednosti u skupu $\{0, 1, 2, \ldots, n\}$. Odredimo vjerojatnost $p_k = P(X = k)$. Ako se realizirao događaj $\{X = k\}$, to znači da se u n pokusa A ostvario točno k puta, a $n - k$ puta se nije ostvario. Broj različitih mogućnosti za odabir pokusa u kojima se A ostvario je $\binom{n}{k}$. Zato je

$$p_k = P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

Uobičajeno je označiti $q := 1 - p$.

Binomna razdioba – definicija
Kažemo da slučajna varijabla X ima binomnu razdiobu s parametrima n i p i pišemo $X \sim \mathcal{B}(n,p)$, ako X poprima vrijednosti unutar skupa $\{0,1,2,\ldots,n\}$ s vjerojatnostima

$$p_k = P(X=k) = \binom{n}{k} p^k q^{n-k}.$$

Primjer 4.17. *Simetrični novčić bacamo* 9 *puta. Označimo s A i B događaje*

$$\begin{aligned} A &= \{\textit{pismo se pojavilo } 4 \textit{ puta}\} \\ B &= \{\textit{pismo se pojavilo } 5 \textit{ puta}\}. \end{aligned}$$

Da li je vjerojatniji događaj A, događaj B ili su događaji jednako vjerojatni?

Rješenje: Broj dobivenih pisama u zadanom slučajnom pokusu ravna se po binomnoj razdiobi $\mathcal{B}(9,\frac{1}{2})$. Zbog toga vrijedi

$$P(A) = \binom{9}{4} \cdot \frac{1}{2^4} \cdot \frac{1}{2^5} = \frac{63}{256},$$
$$P(B) = \binom{9}{5} \cdot \frac{1}{2^5} \cdot \frac{1}{2^4} = \frac{63}{256}.$$

Prema tome, događaji A i B su jednako vjerojatni. To je u skladu sa svojstvom simetrije binomnih koeficijenata. □

Primjer 4.18. *Pokus se sastoji u bacanju triju kocki. Izračunajte vjerojatnost da se u* 6 *nezavisnih pokusa* 3 *puta pojavi točno* 1 *šestica.*

Rješenje: Neka je A događaj

$$A = \{\text{pri bacanju triju kocki pojavila se točno 1 šestica}\}.$$

Broj pojavljivanja šestice je binomna slučajna varijabla $\mathcal{B}(3,\frac{1}{6})$, budući da se bacaju tri kocke, a vjerojatnost pojavljivanja šestice iznosi $\dfrac{1}{6}$. Zato je vjerojatnost događaja A jednaka

$$p = P(A) = \binom{3}{1}\frac{1}{6} \cdot \left(\frac{5}{6}\right)^2 = \frac{25}{72}.$$

Broj realizacija događaja A pri ponavljanju 6 pokusa je binomna slučajna varijabla $\mathcal{B}(6,p)$. Vjerojatnost da se on pojavi točno 3 puta iznosi

$$\binom{6}{3} p^3 q^3 = \binom{6}{3} \left(\frac{25}{72}\right)^3 \left(\frac{47}{72}\right)^3 = 0.233.$$

□

Binomna razdioba ima mnoga važna svojstva, jedno od njih je **stabilnost**: Ako su $X_1 \sim \mathcal{B}(n_1,p)$ i $X_2 \sim \mathcal{B}(n_2,p)$ nezavisne binomne slučajne varijable, onda je $X_1 + X_2 \sim \mathcal{B}(n_1+n_2,p)$. Pokažimo to svojstvo. Koristeći činjenicu da su X_1 i X_2 nezavisne binomne slučajne varijable, imamo redom

$$\begin{aligned}
P(X_1 + X_2 = k) &= \sum_{i=0}^{k} P(X_1 = i, X_2 = k - i) = \sum_{i=0}^{k} P(X_1 = i) P(X_2 = k - i) \\
&= \sum_{i=0}^{k} \binom{n_1}{i} p^i q^{n_1 - i} \binom{n_2}{k-i} p^{k-i} q^{n_2 - k + i} = p^k q^{n_1 + n_2 - k} \sum_{i=0}^{k} \binom{n_1}{i} \binom{n_2}{k-i} \\
&= \binom{n_1 + n_2}{k} p^k q^{n_1 + n_2 - k},
\end{aligned}$$

pri čemu posljednji znak jednakosti vrijedi zbog takozvane *Vandermondeove konvolucije* za binomne koeficijente.

Dobiveni rezultat je prirodan. Riječ je o tome da smo promatrali isti pokus, s time da smo ga podijelili u dva dijela: u prvom smo pokus ponovili n_1 puta, a u drugom n_2 puta. Broj realizacija događaja A u cijelom pokusu jednak je zbroju tih realizacija u pojedinim dijelovima.

Očekivanje i disperzija binomne razdiobe

U nastavku želimo odrediti očekivanje i disperziju binomne slučajne varijable. Međutim, prije toga ćemo promotriti najjednostavniji primjer binomne slučajne varijable. **Bernoullijeva** ili **indikatorska** slučajna varijabla je slučajna varijabla koja poprima samo dvije vrijednosti: 1 s vjerojatnošću p i 0 s vjerojatnošću $q = 1 - p$. Ona bilježi realizaciju događaja A u jednom pokusu.

Ako su $X_1, X_2, \ldots, X_n$ Bernoullijeve nezavisne varijable s istim parametrom p, tada je njihov zbroj $X_1 + X_2 + \cdots + X_n$ binomna slučajna varijabla $X \sim \mathcal{B}(n, p)$. Ova tvrdnja slijedi zbog svojstva stabilnosti binomnih slučajnih varijabli.

Na temelju toga možemo lagano odrediti očekivanje i disperziju binomne slučajne varijable. Naime, za indikatorsku varijablu vrijedi

$$E(X_i) = 0 \cdot q + 1 \cdot p = p, \quad D(X_i) = 0^2 \cdot q + 1^2 \cdot p - p^2 = p(1-p) = pq, \quad i = 1, 2, \ldots, n,$$

pa je, zbog nezavisnosti

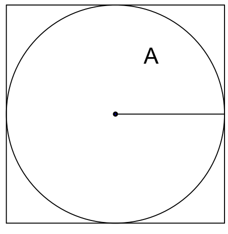
$$\begin{aligned}
E(X) &= E(X_1) + E(X_2) + \cdots + E(X_n) = np, \\
D(X) &= D(X_1) + D(X_2) + \cdots + D(X_n) = npq.
\end{aligned}$$

Binomna razdioba – očekivanje i disperzija
Očekivanje i disperzija binomne slučajne varijable $\mathcal{B}(n, p)$ dani su sljedećim formulama:

$$E(X) = np, \qquad D(X) = npq$$

Primjer 4.19. *U kvadrat je upisan krug. Izračunajte vjerojatnost da će se od šest na sreću odabranih točaka unutar kvadrata barem dvije naći unutar kruga.*

Rješenje: Promotrimo događaj $A = \{$slučajno odabrana točka leži unutar kruga$\}$



Neka je a duljina stranice kvadrata. Polumjer upisanog kruga je $\frac{a}{2}$, pa je površina upisanog kruga $\left(\frac{a}{2}\right)^2 \pi = \frac{a^2\pi}{4}$. Zato je vjerojatnost događaja A jednaka

$$p = \frac{\frac{1}{4}a^2\pi}{a^2} = \frac{\pi}{4}.$$

Broj realizacija događaja A pri ponavljanju 6 pokusa je binomna slučajna varijabla $X \sim \mathcal{B}(6, p)$. Trebamo odrediti $P(X \geq 2)$. Stoga, zbog svojstva komplementa imamo

$$\begin{aligned} P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - \binom{6}{0}\left(\frac{\pi}{4}\right)^0\left(1 - \frac{\pi}{4}\right)^6 - \binom{6}{1}\left(\frac{\pi}{4}\right)^1\left(1 - \frac{\pi}{4}\right)^5 \\ &= 1 - 9.9 \cdot 10^{-5} - 2.2 \cdot 10^{-3} = 0.998. \end{aligned}$$

□

Poissonova razdioba

Poissonovu razdiobu dobivamo kao granični slučaj binomne, u slučaju kada broj pokusa neograničeno raste. Ulogu vjerojatnosti p pojavljivanja događaja preuzima *intenzitet* λ pojavljivanja događaja.

Promotrimo sljedeći primjer. Organizatori neke nagradne igre pakiraju srećke u kutije. Među srećkama je $n = 120$ dobitnih koje se slučajno raspoređuju u $m = 30$ kutija. Neka je X slučajna varijabla: broj dobitnih srećki unutar jedne kutije. Kakva je razdioba slučajne varijable X?

Pretpostavljamo da se svaka dobitna srećka, neovisno jedna o drugoj, s jednakom vjerojatnošću može naći unutar bilo koje kutije. Vjerojatnost da se dobitna srećka nađe unutar odabrane kutije je $p = 1/m = 1/30$. Broj dobitnih srećki unutar te kutije je binomna slučajna varijabla s parametrima $n = 120$ i $p = 1/m = 1/30$.

Primijetimo da je očekivani broj dobitnih srećki unutar jedne kutije jednak $E(X) = np = n/m$. Označimo tu veličinu s λ. Ona označava *intenzitet* pojavljivanja dobitnih srećki unutar neke kutije. Uočimo kako je u ovom primjeru $\lambda = 4$.

Model sličan ovome pojavljuje se pri promatranju broja poziva koji će stići na neku telefonsku centralu u nekoj jedinici vremena. Ako za $m = 60$ minuta na centralu stigne *u prosjeku* $n = 360$ poziva, tada je broj poziva unutar jedne minute – baš kao i u prošlom primjeru sa dobitnim srećkama – binomna razdioba s parametrima $n = 360$, $p = 1/60$. Primijetimo da je očekivani broj poziva $\lambda = 360/60 = 6$.

Razlika između ova dva primjera jest u tome što je broj dobitnih srećki bio unaprijed poznat, i ograničen odozgo. Ukupan broj poziva u drugom primjeru nije poznat, nego je dan kao *statistička veličina*. Dakako, razumno je pretpostaviti da taj broj nije ograničen odozgo.

Veza između binomne i Poissonove razdiobe

Pomoću sljedećeg teorema lakše ćemo približno računati vjerojatnosti kod binomne razdiobe.

Aproksimacija binomne razdiobe
Teorem 4.4 *Neka je n velik, a p malen. Označimo $\lambda = np$. Tada vrijedi sljedeća aproksimacija:*

$$\binom{n}{k} p^k q^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda}.$$

Dobivena aproksimacija upravo služi za definiciju Poissonove razdiobe.

Poissonova razdioba – definicija i numeričke karakteristike
Kažemo da slučajna varijabla X ima Poissonovu razdiobu s parametrom $\lambda > 0$ i pišemo $X \sim \mathcal{P}(\lambda)$ ako ona poprima vrijednosti unutar skupa $\{0, 1, 2, \ldots\}$ s vjerojatnostima

$$p_k = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Za očekivanje i disperziju ove razdiobe vrijedi $E(X) = D(X) = \lambda$.

Pokažimo da je očekivanje Poissonove razdiobe jednako parametru razdiobe. Imamo redom

$$E(X) = \sum_{k=0}^{\infty} k p_k = \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = \lambda,$$

pri čemu smo koristili formulu

$$e^x = \sum_{j=0}^{\infty} \frac{x^j}{j!},$$

za razvoj eksponencijalne funkcije u Taylorov red. Tu formulu svakako treba zapamtiti jer će nam biti od koristi prilikom rješavanja zadataka. Na sličan način se dobiva i formula za disperziju.

Primjer 4.20. *U poznati shopping centar tijekom jednog sata ušlo je* 360 *ljudi. Odredite vjerojatnost*

(a) da tijekom jedne minute nitko nije ušao u shopping centar,

(b) da je barem troje ljudi ušlo u shopping centar.

Rješenje: Neka je slučajna varijabla X broj poziva u jednoj (bilo kojoj) minuti. To je Poissonova varijabla s intenzitetom λ koji je jednak očekivanoj vrijednosti $\lambda = \frac{360}{60} = 6$. Dalje imamo:

$$\begin{aligned} P(X = 0) &= \frac{6^0}{0!} e^{-6} = e^{-6} = 0.002, \\ P(X \geq 3) &= 1 - P(X = 0) - P(X = 1) - P(X = 2) \\ &= 1 - e^{-6} - 6e^{-6} - 18e^{-6} = 1 - 25e^{-6} \\ &= 0.938. \end{aligned}$$

□

Poissonova razdioba, kao i binomna, ima svojstvo **stabilnosti**: Ako su $X_1 \sim \mathcal{P}(\lambda_1)$ i $X_2 \sim \mathcal{P}(\lambda_2)$ nezavisne Poissonove slučajne varijable, onda je $X_1 + X_2 \sim \mathcal{P}(\lambda_1 + \lambda_2)$. Pokažimo to svojstvo. Koristeći činjenicu da su X_1 i X_2 nezavisne Poissonove slučajne varijable, imamo redom

$$\begin{aligned} P(X_1 + X_2 = k) &= \sum_{i=0}^{k} P(X_1 = i, X_2 = k - i) = \sum_{i=0}^{k} P(X_1 = i) P(X_2 = k - i) \\ &= \sum_{i=0}^{k} \frac{{\lambda_1}^i}{i!} e^{-\lambda_1} \frac{{\lambda_2}^{k-i}}{(k-i)!} e^{-\lambda_2} = \frac{e^{-(\lambda_1 + \lambda_2)}}{k!} \sum_{i=0}^{k} \frac{k!}{i!(k-i)!} {\lambda_1}^i {\lambda_2}^{k-i} \\ & \frac{e^{-(\lambda_1 + \lambda_2)}}{k!} \sum_{i=0}^{k} \binom{k}{i} {\lambda_1}^i {\lambda_2}^{k-i} = \frac{(\lambda_1 + \lambda_2)^k}{k!} e^{-(\lambda_1 + \lambda_2)}, \end{aligned}$$

pri čemu smo iskoristili formulu za binomni teorem.

Aproksimacija binomne razdiobe Poissonovom

Na osnovu Teorema 4.4, za veliko n i maleno p, binomna razdioba $\mathcal{B}(n,p)$ može se aproksimirati Poissonovom razdiobom $\mathcal{P}(np)$.

Primjer 4.21. *Tvornica žarulja isporučuje veliku količinu žarulja, među kojima je* 0.8% *škarta. Žarulje se pakiraju u kutije od po* 100 *komada. Koliki će postotak kutija biti bez ijednog škarta, a koliki s dva ili više škartova?*

Rješenje: Broj škartnih proizvoda u jednoj kutiji je slučajna varijabla X distribuirana po binomnom zakonu $\mathcal{B}(100, 0.008)$. Zbog toga je

$$\begin{aligned} P(X=0) &= \binom{100}{0} 0.008^0 \cdot 0.992^{100} = 0.4479, \\ P(X \geq 2) &= 1 - P(X=0) - P(X=1) \\ &= 1 - 0.992^{100} - \binom{100}{1} 0.008 \cdot 0.992^{99} = 0.1909. \end{aligned}$$

Možemo aproksimirati $X \approx \mathcal{P}(0.8)$. Jednostavniji račun daje

$$\begin{aligned} P(X=0) &= e^{-0.8} = 0.4493, \\ P(X \geq 2) &= 1 - e^{-0.8} - 0.8e^{-0.8} = 0.1912. \end{aligned}$$

Uočimo kako se rezultati dobiveni na dva načina podudaraju u dvije decimale. Dakako, povećanjem uzorka, odnosno broja n, točnost postaje sve veća. □

Primjer 4.22. *Neki tvornički stroj sastoji se od* 2000 *dijelova. Vjerojatnost kvara svakog pojedinog dijela u toku radnog tjedna iznosi* 0.001. *Ukoliko se pojedini dio pokvari, vjerojatnost da stroj prestane s radom iznosi* 0.03. *Odredite vjerojatnost da stroj prestane raditi u toku radnog tjedna.*

Rješenje: Neka je slučajna varijabla X broj pokvarenih dijelova stroja. Tada je $X \sim \mathcal{B}(2000, 0.001)$. Označimo

$$\begin{aligned} H_k &= \{\text{pokvarilo se } k \text{ dijelova stroja}\}, \\ A &= \{\text{stroj je prestao s radom}\}. \end{aligned}$$

Vrijedi

$$P(A) = \sum_{k=1}^{2000} P(H_k) P(A|H_k).$$

Uočimo, vjerojatnost da stroj nastavi s radom, ukoliko je pokvareno k dijelova, iznosi 0.97^k. Zbog toga je

$$P(A|H_k) = 1 - P(\overline{A}|H_k) = 1 - 0.97^k.$$

Nadalje, vrijedi

$$P(H_k) = P(X=k) = \binom{2000}{k} 0.001^k \cdot 0.999^{2000-k},$$

što je nepraktično za daljnji račun. Zato aproksimiramo binomnu razdiobu $\mathcal{B}(2000, 0.001)$ s Poissonovom razdiobom $\mathcal{P}(2)$. Uz tu aproksimaciju vrijedi

$$P(H_k) = \frac{2^k}{k!} e^{-2},$$

pa je

$$P(A) = \sum_{k=1}^{2000} (1 - 0.97^k) \frac{2^k}{k!} e^{-2} = e^{-2} \left(\sum_{k=1}^{2000} \frac{2^k}{k!} - \sum_{k=1}^{2000} \frac{1.94^k}{k!} \right).$$

Sada, kako je

$$e^x = \sum_{j=0}^{\infty} \frac{x^j}{j!},$$

posljednje dvije sume možemo aproksimirati pomoću eksponencijalne funkcije. Uočimo kako pri tome u objema sumama nedostaje nulti član. Zbog toga je

$$P(A) = e^{-2} \left[(e^2 - 1) - (e^{1.94} - 1) \right] = 1 - e^{-0.06} = 0.058.$$

$\square$

4.4 Zadatci za ponavljanje

Zadatak 4.1. Neka su X i Y nezavisne slučajne varijable distribuirane po zakonu

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

Odredite disperziju slučajne varijable $Z = \dfrac{X}{Y}$.

Rješenje: Lagano provjeravamo kako slučajna varijabla $Z = X/Y$ poprima vrijednosti iz skupa $\{1/2, 1, 2\}$. Stoga, zbog nezavisnosti varijabli X i Y imamo:

$$\begin{aligned}
P(Z = 1/2) &= P(X = 1, Y = 2) = \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}, \\
P(Z = 1) &= P(X = 1, Y = 1) + P(X = 2, Y = 2) = \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{2}, \\
P(Z = 2) &= P(X = 2, Y = 1) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}.
\end{aligned}$$

Prema tome, razdioba slučajne varijable Z je

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

pa je očekivanje jednako

$$E(X) = \frac{1}{2} \cdot \frac{1}{3} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{6} = 1.$$

Konačno, disperzija iznosi

$$D(X) = E(X^2) - (E(X))^2 = \frac{1}{4} \cdot \frac{1}{3} + 1 \cdot \frac{1}{2} + 4 \cdot \frac{1}{6} - 1^2 = \frac{1}{4}.$$

$\square$

Zadatak 4.2. Zaposlenik u servisu računalne opreme dobije na testiranje 4 ispravne i 3 neispravne matične ploče. Ukoliko zaposlenik testira jednu po jednu matičnu ploču (bez ponavljanja), koliki je očekivani broj pokušaja dok ne naiđe na neispravnu ploču?

Rješenje: Neka je X slučajna varijabla koja bilježi redni broj izvlačenja u kojemu je izvučena neispravna matična ploča. Očito, X poprima vrijednosti iz skupa $\{1, 2, 3, 4, 5\}$. Odredimo njenu razdiobu. Imamo da je

$$\begin{aligned}
P(X=1) &= \frac{3}{7}, \\
P(X=2) &= \frac{4}{7}\cdot\frac{3}{6} = \frac{2}{7}, \\
P(X=3) &= \frac{4}{7}\cdot\frac{3}{6}\cdot\frac{3}{5} = \frac{6}{35}, \\
P(X=4) &= \frac{4}{7}\cdot\frac{3}{6}\cdot\frac{2}{5}\cdot\frac{3}{4} = \frac{3}{35}, \\
P(X=5) &= \frac{4}{7}\cdot\frac{3}{6}\cdot\frac{2}{5}\cdot\frac{1}{4}\cdot\frac{3}{3} = \frac{1}{35}.
\end{aligned}$$

Dakle, razdioba slučajne varijable X je

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ \frac{3}{7} & \frac{2}{7} & \frac{6}{35} & \frac{3}{35} & \frac{1}{35} \end{pmatrix},$$

pa je traženo očekivanje

$$E(X) = 1\cdot\frac{3}{7} + 2\cdot\frac{2}{7} + 3\cdot\frac{6}{35} + 4\cdot\frac{3}{35} + 5\cdot\frac{1}{35} = \frac{70}{35} = 2.$$

□

Zadatak 4.3. Igrač baca jednu kocku. Ako se pojavi broj 1, baca kocku ponovo, ali samo još jedanput. Slučajna varijabla X jednaka je rezultatu prvog bacanja, odnosno zbroju dvaju bacanja ukoliko je kocka bačena dvaput. Izračunajte očekivanje i disperziju slučajne varijable X.

Rješenje: Vrijednosti slučajne varijable X pripadaju skupu $\{2, 3, 4, 5, 6, 7\}$. Očito, svaka od vrijednosti $2, 3, 4, 5, 6$ može se dobiti kao rezultat prvog bacanja te kao zbroj prvog i drugog bacanja. S druge strane, vrijednost 7 se može dobiti samo kao zbroj prvog i drugog bacanja. Zbog toga dobivamo sljedeće vjerojatnosti:

$$\begin{aligned}
P(X=2) &= \frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} = \frac{7}{36}, \\
P(X=3) &= \frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} = \frac{7}{36}, \\
P(X=4) &= \frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} = \frac{7}{36}, \\
P(X=5) &= \frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} = \frac{7}{36}, \\
P(X=6) &= \frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} = \frac{7}{36}, \\
P(X=7) &= \frac{1}{6}\cdot\frac{1}{6} = \frac{1}{36}.
\end{aligned}$$

Prema tome, razdioba slučajne varijable X je

$$X \sim \begin{pmatrix} 2 & 3 & 4 & 5 & 6 & 7 \\ \frac{7}{36} & \frac{7}{36} & \frac{7}{36} & \frac{7}{36} & \frac{7}{36} & \frac{1}{36} \end{pmatrix}.$$

Sada je pripadno očekivanje jednako

$$E(X) = (2+3+4+5+6)\cdot\frac{7}{36} + 7\cdot\frac{1}{36} = \frac{49}{12} = 4.08,$$

a disperzija

$$D(X) = (2^2+3^2+4^2+5^2+6^2)\cdot\frac{7}{36} + 7^2\cdot\frac{1}{36} - 4.08^2 = 2.215.$$

□

Zadatak 4.4. Bacamo dvije simetrične kocke. Slučajna varijabla X mjeri apsolutnu vrijednost razlike brojeva na kockama. Odredite očekivanje i disperziju slučajne varijable X.

Rješenje: Vjerojatnosni prostor sastoji se od 36 jednako vjerojatnih elementarnih događaja. Odredimo vrijednosti varijable X na tim događajima.

X	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

Sada, iz gornje tablice lagano dobivamo razdiobu zadane slučajne varijable:

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

Očekivanje slučajne varijable X je

$$E(X) = 0\cdot\frac{6}{36} + 1\cdot\frac{10}{36} + 2\cdot\frac{8}{36} + 3\cdot\frac{6}{36} + 4\cdot\frac{4}{36} + 5\cdot\frac{2}{36} = \frac{70}{36}.$$

Konačno, disperzija slučajne varijable X je

$$D(X) = 0\cdot\frac{6}{36} + 1\cdot\frac{10}{36} + 4\cdot\frac{8}{36} + 9\cdot\frac{6}{36} + 16\cdot\frac{4}{36} + 25\cdot\frac{2}{36} - \left(\frac{70}{36}\right)^2 = 2.052.$$

□

Zadatak 4.5. Bacamo četiri kocke. Kolika je vjerojatnost događaja A i B ako je

$$\begin{aligned} A &= \{\text{šestica se pojavila barem tri puta}\} \\ B &= \{\text{šestica se pojavila najviše tri puta}\}. \end{aligned}$$

Rješenje: Broj šestica je binomna slučajna varijabla $X \sim \mathcal{B}(4, \frac{1}{6})$. Zbog toga je

$$P(A) = P(X=3) + P(X=4) = \binom{4}{3}\left(\frac{1}{6}\right)^3\cdot\frac{5}{6} + \binom{4}{4}\left(\frac{1}{6}\right)^4 = \frac{7}{432} = 0.016.$$

Vjerojatnost događaja B najlakše računamo pomoću suprotne vjerojatnosti:

$$P(B) = 1 - P(X=4) = 1 - \binom{4}{4}\left(\frac{1}{6}\right)^4 = \frac{1295}{1296} = 0.999.$$

□

Zadatak 4.6. Neka je X binomna slučajna varijabla s očekivanjem $E(X) = 1.5$ i disperzijom $D(X) = 1.125$. Izračunajte vjerojatnost događaja $\{X = 4\}$.

Rješenje: Znamo da za binomnu razdiobu $X \sim \mathcal{B}(n, p)$ vrijedi $E(X) = np$ i $D(X) = npq$, pa ćemo nepoznate parametre n i p izračunati iz očekivanja i disperzije. Iz formule za disperziju dobivamo da je

$$q = \frac{D(X)}{np} = \frac{D(X)}{E(X)} = \frac{1.125}{1.5} = 0.75.$$

Zbog toga je $p = 1 - q = 1 - 0.75 = 0.25$, pa iz formule za očekivanje dobivamo i parametar n:

$$n = \frac{E(X)}{p} = \frac{1.5}{0.25} = 6.$$

Prema tome, riječ je o binomnoj razdiobi $X \sim \mathcal{B}(6, 0.25)$, pa je vjerojatnost traženog događaja jednaka

$$P(X = 4) = \binom{6}{4} 0.25^4 0.75^2 = 0.033.$$

□

Zadatak 4.7. Na nekom graničnom prijelazu u Hrvatsku uđe prosječno 210 automobila tijekom jednog sata. Odredite vjerojatnost da na tom graničnom prijelazu tijekom jedne minute u Hrvatsku uđu barem 3 automobila.

Rješenje: Ovo je Poissonova slučajna varijabla s parametrom $\lambda = \dfrac{210}{60} = 3.5$. Zato je tražena vjerojatnost jednaka

$$\begin{aligned} P(X \geq 3) &= 1 - P(X = 0) - P(X = 1) - P(X = 2) \\ &= 1 - e^{-3.5}\left(1 + \frac{3.5}{1!} + \frac{3.5^2}{2!}\right) \\ &= 1 - 10.625 \cdot e^{-3.5} = 0.679. \end{aligned}$$

□

Zadatak 4.8. Slučajna varijabla X ima Poissonovu razdiobu. Ako je $P(X = 2) = P(X = 4)$, izračunajte očekivanje $E(X)$ i vjerojatnost $P(X \geq 3)$.

Rješenje: Očekivanje Poissonove razdiobe jednako je parametru λ kojeg ćemo izračunati iz uvjeta $P(X = 2) = P(X = 3)$. Iz zadanog uvjeta slijedi da je

$$\frac{\lambda^2}{2!} e^{-\lambda} = \frac{\lambda^3}{3!} e^{-\lambda}.$$

Dobivenu jednakost možemo podijeliti s $e^{-\lambda}$, pa sređivanjem dobivamo jednadžbu $3\lambda^2 - \lambda^3 = 0$, odnosno

$$\lambda^2(3 - \lambda) = 0.$$

Rješenja te jednadžbe su $\lambda = 0$ i $\lambda = 3$. Dakako, rješenje $\lambda = 0$ nema smisla pa je parametar razdiobe $\lambda = 3$. Stoga je i očekivanje $E(X)$ jednako 3.

Nadalje, tražena vjerojatnost je

$$\begin{aligned} P(X \geq 4) &= 1 - P(X=0) - P(X=1) - P(X=2) - P(X=3) \\ &= 1 - e^{-3}\left(1 + \frac{3}{1!} + \frac{3^2}{2!} + \frac{3^3}{3!}\right) \\ &= 1 - 13 \cdot e^{-3} = 0.353. \end{aligned}$$

□

Zadatak 4.9. Neka su X i Y nezavisne i jednako distribuirane slučajne varijable s očekivanjem μ i disperzijom σ^2. Izračunajte očekivanje slučajne varijable $(X-Y)^2$.

Rješenje: Iz svojstava očekivanja imamo da je

$$\begin{aligned} E\left((X-Y)^2\right) &= E\left(X^2 - 2XY + Y^2\right) = E\left(X^2\right) - 2E(XY) + E\left(Y^2\right) \\ &= E\left(X^2\right) - 2E(X)E(Y) + E\left(Y^2\right). \end{aligned}$$

Sada, kako su X i Y jednako distribuirane, slijedi da je $E(X) = E(Y)$ i $E\left(X^2\right) = E\left(Y^2\right)$, odakle dobivamo

$$E\left((X-Y)^2\right) = 2E\left(X^2\right) - 2\left(E(X)\right)^2 = 2\left[E\left(X^2\right) - \left(E(X)\right)^2\right].$$

Konačno, kako je $D(X) = E\left(X^2\right) - \left(E(X)\right)^2 = \sigma^2$, slijedi da je $E\left((X-Y)^2\right) = 2\sigma^2$. Uočite kako dobiveni rezultat ne ovisi o očekivanju μ, nego samo o disperziji σ^2. □

5 Neprekinute slučajne varijable

5.1 Slučajne varijable i razdiobe

U prethodnom poglavlju proučavali smo diskretne slučajne varijable. Svaka takva slučajna varijabla bila je jednoznačno određena svojom razdiobom. Nadalje, pomoću te razdiobe računali smo numeričke karakteristike slučajne varijable, kao što su očekivanje i disperzija.

Ovdje promatramo slučajne varijable kojima je skup vrijednosti neki interval u skupu realnih brojeva. Takve slučajne varijable mogu poprimiti svaku vrijednost unutar tog intervala pa im ne možemo pridružiti tablicu razdiobe kao u diskretnom slučaju. Osim toga, takve slučajne varijable poprimaju neprebrojivo mnogo vrijednosti, pa je vrijednost svake od tih realizacija jednaka nuli.

Prilikom proučavanja neprekinutih slučajnih varijabli koristit ćemo se aparatom matematičke analize. Nizove brojeva koji definiraju razdiobu zamijenit ćemo realnim funkcijama, a umjesto sumiranja koristit ćemo tehnike integralnog i diferencijalnog računa.

Slučajna varijabla

Kao što smo već napomenuli, slučajnim varijablama koje poprimaju neprebrojivo mnogo vrijednosti ne možemo pridružiti tablicu razdiobe, nego ćemo im pridružiti odgovarajuću funkciju. No, prije toga ćemo definirati slučajnu varijablu. Ta se definicija neznatno razlikuje od definicije diskretne slučajne varijable, a uključuje i tu vrstu slučajnih varijabli.

Slučajna varijabla i funkcija razdiobe
Neka je $(\Omega, \mathcal{F}, P)$ vjerojatnosni prostor. Preslikavanje $X : \Omega \to \mathbf{R}$ nazivamo **slučajna varijabla** ako je za svaki $x \in \mathbf{R}$ skup $A_x = \{\omega \in \Omega : X(\omega) < x\}$ događaj, odnosno element algebre $\mathcal{F}$.
Skup $\{\omega \in \Omega : X(\omega) < x\}$ označavat ćemo kraće sa $\{X < x\}$.
Funkcija razdiobe slučajne varijable X je funkcija $F : \mathbf{R} \to [0, 1]$ definirana formulom

$$F(x) = P(\{X < x\}).$$

Funkcija razdiobe i njezina derivacija, ukoliko postoji, bit će najvažniji pojmovi vezani uz slučajnu varijablu. Naime, slučajna varijabla je jednoznačno određena pripadnom funkcijom razdiobe.

Kao što smo već napomenuli, ima smisla promatrati funkciju razdiobe i za diskretne slučajne varijable. Pogledajmo sljedeći primjer.

Primjer 5.1. *Zadana je slučajna varijabla* $X \sim \begin{pmatrix} -1 & 0 & 1 \\ \frac{5}{24} & \frac{1}{2} & \frac{7}{24} \end{pmatrix}$. *Odredite funkciju razdiobe varijable X i nacrtajte njezin graf.*

Rješenje: Ako je $x \leq -1$, tada događaj $\{X < x\}$ ima vjerojatnost 0 te je $F(x) = 0$ za takve x. Za

$-1 < x \leq 0$ vrijedi

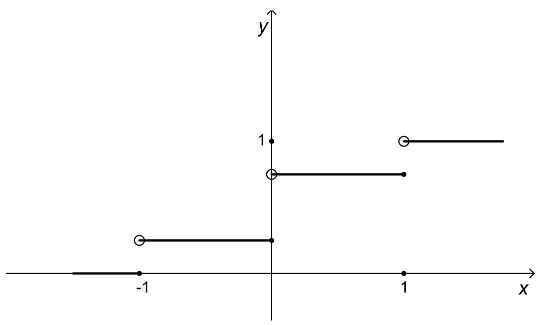
$$P(X < x) = P(X = -1) = \frac{5}{24}.$$

Za $0 < x \leq 1$ vrijedi

$$P(X < x) = P(X = -1) + P(X = 0) = \frac{5}{24} + \frac{1}{2} = \frac{17}{24},$$

itd. Tako dobivamo $F(x) = \begin{cases} 0, & x \leq -1, \\ \frac{5}{24}, & -1 < x \leq 0, \\ \frac{17}{24}, & 0 < x \leq 1, \\ 1, & 1 < x. \end{cases}$

Graf funkcije razdiobe je prikazan na donjoj slici:



$\square$

Očito, funkcija razdiobe diskretne slučajne varijable sa zakonom

$$X \sim \begin{pmatrix} x_1 & x_2 & x_3 & \ldots \\ p_1 & p_2 & p_3 & \ldots \end{pmatrix}$$

je stepenasta funkcija sa skokovima u točkama $x_1, x_2, \ldots$. Iznosi skokova su vjerojatnosti $p_1, p_2, \ldots$. Dakle, funkcija razdiobe je lijevo od najmanje vrijednosti jednaka nuli, a desno od najveće, jedan. Uočimo kako je ta stepenasta funkcija neopadajuća te je neprekinuta slijeva u točkama prekida. Ta svojstva vrijede i za bilo koju funkciju razdiobe, a slijede lagano iz definicije funkcije razdiobe.

Važna svojstva općenite funkcije razdiobe sadržana su u sljedećem teoremu, kojeg navodimo bez dokaza.

Svojstva funkcije razdiobe

Teorem 5.1 *Neka je F funkcija razdiobe slučajne varijable X. Ona posjeduje svojstva:*

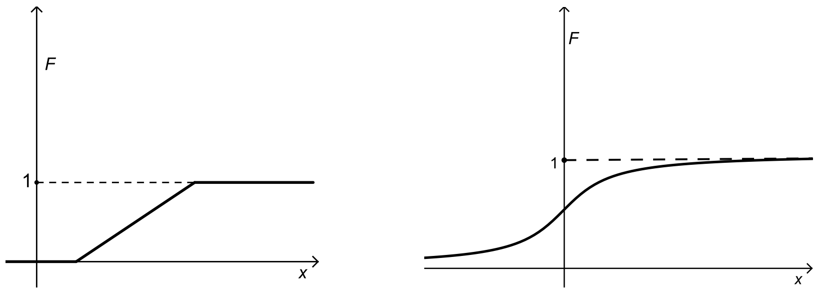
1° $P(\{x_1 \leq X < x_2\}) = F(x_2) - F(x_1)$,

2° *F je neopadajuća:* $x_1 < x_2 \Longrightarrow F(x_1) \leq F(x_2)$,

3° $\lim\limits_{x \to -\infty} F(x) = 0$, $\lim\limits_{x \to \infty} F(x) = 1$,

4° *F je neprekinuta slijeva:* $F(x - 0) := \lim\limits_{\varepsilon \downarrow 0} F(x - \varepsilon) = F(x)$, $\forall x \in \mathbf{R}$.

Promotrimo sada svojstva iz Teorema 5.1 na primjerima grafova nekih funkcija razdioba.



Kao što vidimo iz grafova, vrijednosti obiju funkcija razdioba su u minus beskonačnosti ili jednake ili teže k nuli, dok su u plus beskonačnosti ili jednake ili teže k jedinici. Također, obje funkcije razdiobe su neopadajuće. Dakako, funkcija razdiobe može imati prekide, ali prema prethodnom teoremu ona mora biti neprekinuta slijeva, što smo vidjeli u Primjeru 5.1.

Zadržimo se još malo na diskretnoj razdiobi sa zakonom

$$X \sim \begin{pmatrix} x_1 & x_2 & x_3 & \dots \\ p_1 & p_2 & p_3 & \dots \end{pmatrix}.$$

Toj razdiobi možemo pridružiti takozvanu **funkciju gustoće**, odnosno funkciju $f : \mathbf{R} \to \mathbf{R}$, definiranu formulom

$$f(x) = P(X = x) = \begin{cases} 0, & x \neq x_i, \\ p_i, & x = x_i. \end{cases}$$

Iz Primjera 5.1 vidimo da u diskretnom slučaju vrijedi relacija

$$F(x) = \sum_{x_i < x} p_i = \sum_{y < x} f(y),$$

gdje je F funkcija razdiobe diskretne varijable X. Htjeli bismo naći analogon prethodne relacije kod neprebrojivih slučajnih varijabli. U nastavku ćemo vidjeti kako se prethodna relacija može proširiti u slučaju kada postoji gustoća takve slučajne varijable, pri čemu će ulogu sume će preuzeti integral. To će nas i dovesti do pojma neprekinutih slučajnih varijabli koje ćemo proučavati u ovom poglavlju.

Neprekinute slučajne varijable

Krenimo odmah s definicijom neprekinute ili kontinuirane slučajne varijable i definicijom gustoće.

Neprekinute slučajne varijable. Gustoća razdiobe
Za slučajnu varijablu X kažemo da je **neprekinuta** ako postoji nenegativna funkcija $f : \mathbf{R} \to \mathbf{R}$ takva da vrijedi

$$F(x) = \int_{-\infty}^{x} f(t)dt.$$

Funkcija f naziva se **gustoća razdiobe vjerojatnosti** slučajne varijable X. Ona nije nužno neprekinuta, no u točkama neprekinutosti od f vrijedi

$$f(x) = F'(x).$$

Očito, funkcija razdiobe neprekinute slučajne varijable je i sama neprekinuta, jer je to funkcija gornje granice integrala. Zbog toga je $P(X = x) = F(x + 0) - F(x) = 0$ za svaki $x \in \mathbf{R}$. Iz te relacije zaključujemo da su događaji $\{x_1 < X < x_2\}$, $\{x_1 \leq X < x_2\}$, $\{x_1 < X \leq x_2\}$,

$\{x_1 \le X \le x_2\}$ jednako vjerojatni. Prema Teoremu 5.1 njihova se vjerojatnost računa pomoću formule

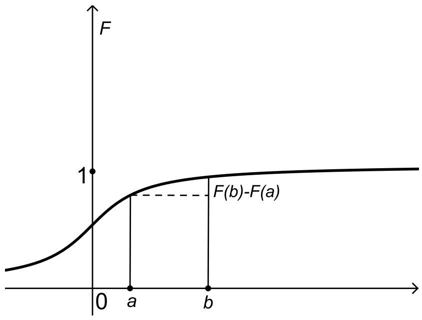
$$P(x_1 < X < x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(t)dt.$$

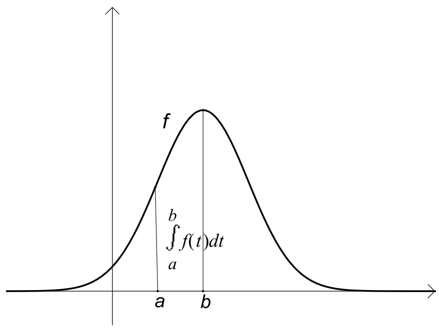
Također, iz Teorema 5.1 slijedi da je funkcija gustoće pozitivna funkcija s integralom

$$\int_{-\infty}^{\infty} f(t)dt = 1.$$

Prisjetimo li se matematičke analize, prethodna relacija kazuje kako je površina ispod grafa funkcije gustoće uvijek jednaka jedan.

Na donjoj slici vidimo primjer razdiobe neprekinute slučajne varijable i pripadne gustoće.





Do sada smo naučili da je funkcija razdiobe neopadajuća funkcija s vrijednostima unutar intervala $[0,1]$. Zato ćemo neke formule zapisivati u skraćenom obliku. Na primjer, umjesto zapisa s vitičastim zagradama,

$$F(x) = \begin{cases} 0, & x \le 0, \\ 3x, & 0 < x < \frac{1}{3}, \\ 1, & 1 \le x, \end{cases}$$

pisat ćemo kratko

$$F(x) = 3x, \quad 0 < x < \frac{1}{3},$$

jer je tada nužno $F(x) = 0$ za $x \le 0$ i $F(x) = 1$ za $x \ge 1/3$.

Također, ako gustoću razdiobe definiramo nekom formulom za $x \in [a,b]$, tada smatramo da je van tog intervala po definiciji jednaka nuli. Konačno, ako je funkcija F ili f definirana nekom formulom bez naznake područja definicije, onda će to redovito biti čitav skup realnih brojeva.

Jednolika razdioba

Sada ćemo opisati jednu važnu neprekinutu razdiobu. Prisjetimo se slučaja kojeg smo imali kada je područje vrijednosti slučajne varijable $S = \{x_1, x_2, \ldots, x_n\}$, bio konačan skup. Slučajna varijabla X, koja je poprimala vrijednosti unutar S s jednakim vjerojatnostima, opisivala je pokus biranja na sreću elemenata skupa S. Razdioba takve varijable dana je s

$$X \sim \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{pmatrix},$$

odnosno X je jednolika razdioba $X \sim \mathcal{U}(n)$.

Pretpostavimo sada da je skup S interval $[a,b]$. Točaka unutar tog intervala ima beskonačno (neprebrojivo) mnogo. Dakle, toj slučajnoj varijabli ne možemo pridružiti razdiobu, pa je razumno pretpostaviti da je gustoća te slučajne varijable konstantna funkcija na intervalu $[a,b]$ te jednaka nuli van tog intervala. Tu konstantu možemo naći iz činjenice da je površina ispod funkcije gustoće jednaka jedan. Dakle, ako je $f(x) = K$, $x \in [a,b]$, onda iz

$$1 = \int_a^b K dx = K \int_a^b dx = K(b-a),$$

dobivamo da je $K = \frac{1}{b-a}$. Sada lagano možemo izračunati i pripadnu funkciju razdiobe. Naime, vrijedi

$$F(x) = \int_a^x \frac{1}{b-a} dt = \frac{1}{b-a} \int_a^x dt = \frac{1}{b-a} \cdot t\Big|_a^x = \frac{x-a}{b-a}.$$

Time je u potpunosti opisana jednolika razdioba na intervalu $[a,b]$.

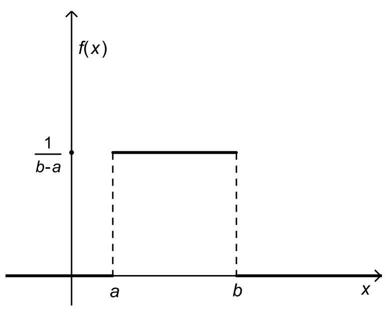
Jednolika razdioba, funkcija razdiobe i gustoća

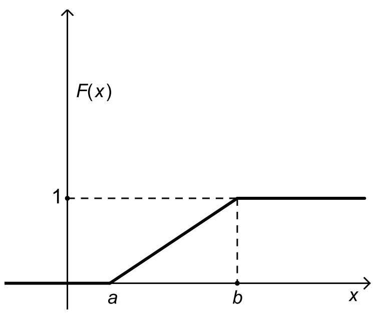
Za slučajnu varijablu X kažemo da je **jednoliko** (**uniformno**) distribuirana na intervalu $[a,b]$ ako je zadana funkcijom razdiobe odnosno funkcijom gustoće:

$$F(x) = \frac{x-a}{b-a}, \quad a \leq x \leq b,$$
$$f(x) = \frac{1}{b-a}, \quad a \leq x \leq b.$$

Pišemo $X \sim \mathcal{U}(a,b)$.

Grafovi funkcije gustoće i pripadne funkcije razdiobe izgledaju ovako:





Uočimo kako jednolika razdioba predstavlja afinu funkciju na intervalu $[a,b]$. Gustoća jednolike razdiobe je konstantna na tom intervalu. Odatle i ime razdiobi.

Promotrimo sada jedan složeniji primjer.

Primjer 5.2. *Slučajno odabiremo dva broja unutar intervala* $[0,1]$. *Definirajmo slučajnu varijablu* Z *kao aritmetičku sredinu odabranih brojeva. Odredite zakon razdiobe i funkciju gustoće slučajne varijable* Z.

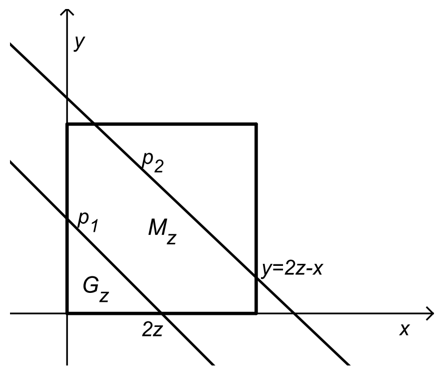
Rješenje: Izbor dvaju brojeva x i y unutar intervala $[0,1]$ ekvivalentan je izboru jedne točke (x,y) unutar kvadrata $[0,1] \times [0,1]$. Vrijednost koju slučajna varijabla Z poprima jednaka je $Z = \dfrac{x+y}{2}$, pa ona poprima vrijednosti iz intervala $[0,1]$. Očito vrijedi ekvivalencija

$$\frac{x+y}{2} < z \Longleftrightarrow y < 2z - x,$$

pa je

$$F(z) = P(Z < z) = P(y < 2z - x).$$

No moramo razlikovati dva slučaja, ovisno o tome da li je $z \leq \frac{1}{2}$ ili $z \geq \frac{1}{2}$ (vidi sliku).



U prvom slučaju, područje ispod pravca p_1 je jednakokračan pravokutan trokut G_z sa stranicom duljine $2z$. Zato je

$$F(z) = P(Z < z) = m(G_z) = \frac{4z^2}{2} = 2z^2, \quad 0 \le z \le \frac{1}{2}.$$

U drugom slučaju, područje ispod pravca p_2 je peterokut M_z kojeg dobivamo ako od kvadrata odstranimo njegov dio iznad pravca p_2. Zato je

$$F(z) = P(Z < z) = m(M_z) = 1 - \frac{(2-2z)^2}{2} = -2z^2 + 4z - 1, \quad \frac{1}{2} \le z \le 1.$$

Prema tome, funkcija razdiobe slučajne varijable Z je

$$F(z) = \begin{cases} 2z^2, & 0 \le z \le \frac{1}{2}, \\ -2z^2 + 4z - 1, & \frac{1}{2} \le z \le 1. \end{cases}$$

Funkciju gustoće nalazimo deriviranjem funkcije razdiobe na svakoj komponenti:

$$f(z) = \begin{cases} 4z, & 0 \le z \le \frac{1}{2}, \\ -4z + 4, & \frac{1}{2} \le z \le 1. \end{cases}$$

□

Nezavisnost slučajnih varijabli

Pojam nezavisnosti slučajnih varijabli upoznali smo za varijable diskretnog tipa. Potpuno iste definicije vrijede i za općenite slučajne varijable.

Definicija nezavisnosti
Kažemo da su slučajne varijable X i Y **nezavisne**, ukoliko za sve intervale A, B iz skupa $\mathbf{R}$ vrijedi

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B).$$

Očekivanje i disperzija

Očekivanje i disperziju definiramo pomoću gustoće slučajne varijable. Naime, u usporedbi s diskretnim slučajem ulogu vjerojatnosti preuzima gustoća, a sumu zamijenjujemo integralom.

Dakle, ako je X neprekinuta slučajna varijabla s gustoćom f, njezino očekivanje definiramo na način

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx.$$

Ako ovaj nepravi integral ne konvergira, očekivanje ne postoji.

Označimo $\overline{x} = E(X)$. Disperzija $D(X)$ slučajne varijable X računa se uz pomoć formula:

$$D(X) = E\left[(X-\overline{x})^2\right] = E(X^2) - \overline{x}^2.$$

Prema tome,

$$D(X) = \int_{-\infty}^{\infty} (x-\overline{x})^2 f(x) dx = \int_{-\infty}^{\infty} x^2 f(x) dx - \overline{x}^2.$$

Osnovna svojstva očekivanja i disperzije vrijede kao i kod diskretnih slučajnih varijabli.

Svojstva očekivanja i disperzije

Za sve slučajne varijable X, Y i realne brojeve α, β vrijedi svojstvo linearnosti očekivanja:

$$E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y).$$

Za disperziju vrijedi

$$D(\alpha X) = \alpha^2 D(X).$$

Ako su X i Y nezavisne, onda vrijede relacije

$$E(XY) = E(X)E(Y) \text{ i } D(X+Y) = D(X) + D(Y).$$

Promotrimo sada nekoliko primjera u kojima ćemo računati očekivanje i disperziju slučajne varijable.

Primjer 5.3. *Izračunajte očekivanje i disperziju jednolike razdiobe na intervalu* $[0,1]$.

Rješenje: Kako je gustoća promatrane razdiobe jednaka $f(x) = \frac{1}{1-0} = 1$, $0 \le x \le 1$, imamo redom

$$\begin{aligned}
E(X) &= \int_0^1 x dx = \left.\frac{x^2}{2}\right|_0^1 = \frac{1}{2}, \\
D(X) &= \int_0^1 x^2 dx - \left(\frac{1}{2}\right)^2 = \left.\frac{x^3}{3}\right|_0^1 - \frac{1}{4} = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.
\end{aligned}$$

□

Primjer 5.4. *Zadana je funkcija*

$$f(x) = \begin{cases} 1 - Cx & , 0 \le x \le 2, \\ 0 & , \textit{inače} \end{cases},$$

gdje je C neka realna konstanta.

(a) Odredite konstantu C tako da f bude gustoća razdiobe slučajne varijable X.

(b) Odredite funkciju razdiobe $F(x)$.

(c) Izračunajte vjerojatnost događaja $\{1 < X < 2\}$.

(d) Izračunajte očekivanje $E(X)$.

Rješenje: (a) Konstantu C određujemo iz uvjeta $\displaystyle\int_0^2 f(x)dx = 1$. Naime, kako je

$$\int_0^2 (1 - Cx)dx = \left(x - \frac{Cx^2}{2}\right)\Bigg|_0^2 = 2 - 2C,$$

konstanta C zadovoljava jednadžbu $2 - 2C = 1$, čije je rješenje $C = \dfrac{1}{2}$.
(b) Funkcija razdiobe je

$$F(x) = \int_0^x \left(1 - \frac{t}{2}\right) dt = \left(t - \frac{t^2}{4}\right)\Bigg|_0^x = x - \frac{x^2}{4}, \quad 0 \leq x \leq 2.$$

(c) Vjerojatnost događaja $\{1 < X < 2\}$ računamo pomoću funkcije razdiobe:

$$P(1 < X < 2) = F(2) - F(1) = 1 - \frac{3}{4} = \frac{1}{4}.$$

Vjerojatnost događaja također možemo računati pomoću funkcije gustoće:

$$P(1 < X < 2) = \int_1^2 \left(1 - \frac{x}{2}\right) dx = \left(x - \frac{x^2}{4}\right)\Bigg|_1^2 = \frac{1}{4}.$$

(d) Očekivanje slučajne varijable X je

$$E(X) = \int_0^2 x\left(1 - \frac{x}{2}\right) dx = \int_0^2 \left(x - \frac{x^2}{2}\right) dx = \left(\frac{x^2}{2} - \frac{x^3}{6}\right)\Bigg|_0^2 = 2 - \frac{4}{3} = \frac{2}{3}.$$

□

Primjer 5.5. *Izračunajte očekivanje i disperziju slučajne varijable iz Primjera* 5.2.

Rješenje: Očekivanje i disperziju računamo pomoću funkcije gustoće koju smo odredili u Primjeru 5.2:

$$f(z) = \begin{cases} 4z, & 0 \leq z \leq \frac{1}{2}, \\ -4z + 4, & \frac{1}{2} \leq z \leq 1. \end{cases}$$

Imamo da je

$$\begin{aligned} E(Z) &= \int_0^{\frac{1}{2}} z \cdot 4z dz + \int_{\frac{1}{2}}^1 z(4 - 4z)dz = 4\int_0^{\frac{1}{2}} z^2 dz + 4\int_{\frac{1}{2}}^1 (z - z^2)dz \\ &= \frac{4z^3}{3}\Bigg|_0^{\frac{1}{2}} + 4\left(\frac{z^2}{2} - \frac{z^3}{3}\right)\Bigg|_{\frac{1}{2}}^1 = \frac{1}{6} + \frac{1}{3} = \frac{1}{2}. \end{aligned}$$

Dalje je

$$\begin{aligned} E(Z^2) &= \int_0^{\frac{1}{2}} z^2 \cdot 4z dz + \int_{\frac{1}{2}}^1 z^2(4 - 4z)dz = 4\int_0^{\frac{1}{2}} z^3 dz + 4\int_{\frac{1}{2}}^1 (z^2 - z^3)dz \\ &= z^4\Bigg|_0^{\frac{1}{2}} + 4\left(\frac{z^3}{3} - \frac{z^4}{4}\right)\Bigg|_{\frac{1}{2}}^1 = \frac{1}{16} + \frac{11}{48} = \frac{7}{24}. \end{aligned}$$

Konačno, disperzija slučajne varijable Z je

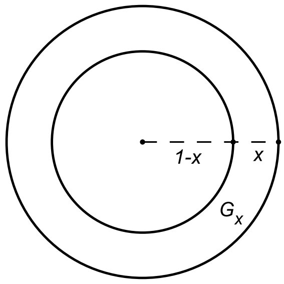
$$D(Z) = E(Z^2) - (E(Z))^2 = \frac{7}{24} - \left(\frac{1}{2}\right)^2 = \frac{1}{24}.$$

□

Primjer 5.6. *Biramo na sreću točku unutar kruga polumjera 1. Neka je vrijednost slučajne varijable X udaljenost te točke od ruba kruga. Odredite razdiobu i očekivanje slučajne varijable X.*

Rješenje: Odredimo najprije funkciju razdiobe:

$$F(x) = P(X < x) = P(T \in G_x) = \frac{m(G_x)}{m(\Omega)} = \frac{1^2\pi - (1-x)^2\pi}{1^2\pi} = 1-(1-x)^2 = 2x-x^2, \quad 0 \le x \le 1.$$



Pri tome Ω označava krug polumjera 1, a G_x kružni vijenac širine x.

Nadalje, gustoća razdiobe X je

$$f(x) = F'(x) = 2 - 2x, \quad 0 \le x \le 1,$$

pa je očekivanje

$$E(X) = \int_0^1 x(2-2x)dx = \int_0^1 (2x - 2x^2)dx = \left(x^2 - \frac{2x^3}{3}\right)\Bigg|_0^1 = 1 - \frac{2}{3} = \frac{1}{3}.$$

□

5.2 Eksponencijalna razdioba

Eksponencijalna razdioba pojavljuje se kod problema vezanih uz vrijeme ispravnog rada nekog uređaja čija se svojstva ne mijenjaju u vremenu. Na primjer, to može biti vrijeme ispravnog rada žarulje, vrijeme do prvog poziva u telefonskoj centrali, vrijeme do ulova prve ribe, vrijeme do pojave neke nesreće i općenito vrijeme do pojave nekog događaja čija je vjerojatnost pojavljivanja u svakom kratkom intervalu jednake duljine jednaka.

Eksponencijalna razdioba, definicija

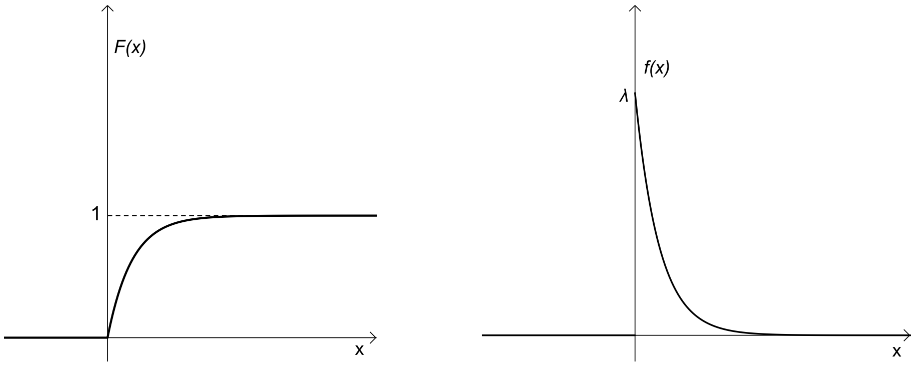
Kažemo da slučajna varijabla X ima **eksponencijalnu razdiobu** s parametrom $\lambda > 0$ ako ona poprima pozitivne vrijednosti s gustoćom

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0.$$

Pišemo $X \sim \mathcal{E}(\lambda)$. Njezina funkcija razdiobe je

$$F(x) = 1 - e^{-\lambda x}, \quad x > 0.$$

Slika prikazuje funkciju razdiobe i gustoće eksponencijalne razdiobe.



Primjer 5.7. *Vrijeme X ispravnog rada računala je slučajna varijabla s eksponencijalnom razdiobom. Vjerojatnost ispravnog rada računala tijekom jedne godine iznosi* 0.9. *Kolika je vjerojatnost da će od* 15 *računala u računarskom praktikumu njih barem* 13 *raditi ispravno tijekom* 2 *godine?*

Rješenje: Slučajna varijabla X ponaša se po eksponencijalnom zakonu, odnosno $X \sim \mathcal{E}(\lambda)$, pa je $F(x) = 1 - e^{-\lambda x}$, gdje je λ parametar te razdiobe.

Iz uvjeta zadatka vidimo da je $P(X > 1) = 0.9$. No, s druge strane imamo da je

$$P(X > 1) = 1 - P(X < 1) = 1 - F(1) = 1 - \left(1 - e^{-\lambda}\right) = e^{-\lambda} = 0.9.$$

Iz prethodne jednakosti možemo izračunati parametar λ, ali u ovom zadatku to neće biti potrebno. Naime, mi trebamo naći vjerojatnost da računalo radi ispravno dvije godine. Imamo da je

$$P(X > 2) = 1 - P(X < 2) = 1 - F(2) = 1 - \left(1 - e^{-2\lambda}\right) = e^{-2\lambda} = \left(e^{-\lambda}\right)^2 = 0.81.$$

Sada trebamo odrediti vjerojatnost da od 15 računala, njih barem 13 radi ispravno tijekom 2 godine. To je binomna razdioba $Y \sim \mathcal{B}(15, 0.81)$. Konačno, dobivamo

$$P(Y \geq 13) = \binom{15}{13} 0.81^{13} \cdot 0.19^2 + \binom{15}{14} 0.81^{14} \cdot 0.19 + \binom{15}{15} 0.81^{15} = 0.436.$$

□

Očekivanje i disperzija eksponencijalne razdiobe

Odredimo sada očekivanje i disperziju eksponencijalne razdiobe s parametrom λ. Kako za gustoću vrijedi $f(x) = \lambda e^{-\lambda x}$, $x > 0$, primjenom formule za parcijalnu integraciju imamo redom

$$E(X) = \int_0^\infty x f(x) dx = \lambda \int_0^\infty x e^{-\lambda x} dx = \left(-x e^{-\lambda x}\right)\Big|_0^\infty + \int_0^\infty e^{-\lambda x} dx = 0 + \frac{e^{-\lambda x}}{-\lambda}\Big|_0^\infty = \frac{1}{\lambda}.$$

Na sličan način dobivamo da je $E(X^2) = \dfrac{2}{\lambda^2}$, pa je

$$D(X) = E(X^2) - [E(X)]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Iz prethodnog izvoda vidimo da je očekivanje eksponencijalne razdiobe recipročna vrijednost njezinog parametra: što je parametar razdiobe veći, to je manje njezino očekivanje.

Primjer 5.8. *Tvornički stroj se tijekom godine dana pokvario četiri puta. Kolika je vjerojatnost da će prvi mjesec sljedeće godine raditi ispravno?*

Rješenje: Neka slučajna varijabla X mjeri vrijeme do prvog kvara stroja. Tada X ima eksponencijalnu razdiobu. Iz podataka koje imamo zaključujemo da je očekivano vrijeme do kvara jednako $\frac{12}{4} = 3$. To znači da je $\frac{1}{\lambda} = 3$, odnosno $\lambda = \frac{1}{3}$. Konačno, iz definicije funkcije razdiobe dobivamo da je

$$P(X > 1) = 1 - P(X < 1) = 1 - F(1) = 1 - \left(1 - e^{-\frac{1}{3}}\right) = e^{-\frac{1}{3}} = 0.717.$$

□

Primjer 5.9. *Vrijeme do prvog poziva u nekoj telefonskoj centrali je slučajna varijabla distribuirana po eksponencijalnom zakonu s očekivanjem 3 minute. Kolika je vjerojatnost da će se prvi poziv ostvariti tijekom*

(a) prve minute,

(b) prve dvije minute,

(c) druge minute, ako je poznato da nije bilo poziva tijekom prve minute,

(d) treće i četvrte minute, ako je poznato da nije bilo poziva tijekom prve dvije minute?

Rješenje: Kako je očekivanje jednako 3, slijedi da je $\lambda = \frac{1}{3}$, tj. $X \sim \mathcal{E}\left(\frac{1}{3}\right)$. Sada imamo redom:

(a) $P(X < 1) = F(1) = 1 - e^{-\frac{1}{3}} = 0.283$,

(b) $P(X < 2) = F(2) = 1 - e^{-\frac{2}{3}} = 0.486$.

(c) Ovdje imamo uvjetnu vjerojatnost, odnosno

$$\begin{aligned} P(1 < X < 2 \,|X > 1) &= \frac{P(1 < X < 2, X > 1)}{P(X > 1)} = \frac{P(1 < X < 2)}{P(X > 1)} \\ &= \frac{F(2) - F(1)}{1 - F(1)} = \frac{\left(1 - e^{-\frac{2}{3}}\right) - \left(1 - e^{-\frac{1}{3}}\right)}{e^{-\frac{1}{3}}} \\ &= \frac{e^{-\frac{1}{3}} - e^{-\frac{2}{3}}}{e^{-\frac{1}{3}}} = 1 - e^{-\frac{1}{3}} = 0.283. \end{aligned}$$

(d) Slično kao i u (c) dijelu zadatka imamo

$$\begin{aligned} P(2 < X < 4 \,|X > 2) &= \frac{P(2 < X < 4, X > 2)}{P(X > 2)} = \frac{P(2 < X < 4)}{P(X > 2)} \\ &= \frac{F(4) - F(2)}{1 - F(2)} = \frac{\left(1 - e^{-\frac{4}{3}}\right) - \left(1 - e^{-\frac{2}{3}}\right)}{e^{-\frac{2}{3}}} \\ &= \frac{e^{-\frac{2}{3}} - e^{-\frac{4}{3}}}{e^{-\frac{2}{3}}} = 1 - e^{-\frac{2}{3}} = 0.486. \end{aligned}$$

□

Uočimo kako se u prethodnom primjeru prva i treća vjerojatnost podudaraju. Isto vrijedi i za preostale dvije vjerojatnosti. To svojstvo vrijedi i općenitije, a karakterizira eksponencijalnu razdiobu.

Odsustvo pamćenja

Svojstvo iz prethodnog primjera može se poopćiti. Za sve $x, t > 0$ vrijedi

$$P(X < x + t \mid X > t) = P(X < x).$$

Prethodnu relaciju možemo interpretirati na sljedeći način: eksponencijalna razdioba nema pamćenja. Na primjer, ako je očekivano vrijeme do prvog poziva u telefonskoj centrali 3 minute, i ako u prve dvije minute nije bilo niti jednog poziva, tada očekivano vrijeme do prvog poziva ostaje i dalje tri minute, bez obzira na proteklo vrijeme. Kao što smo već napomenuli, ovo svojstvo karakterizira eksponencijalnu razdiobu.

Karakterizacija eksponencijalne razdiobe
Teorem 5.2 *Neka za slučajnu varijablu X, koja poprima samo pozitivne vrijednosti, za sve pozitivne x i t vrijedi*

$$P(X < x + t \mid X > t) = P(X < x).$$

Tada X ima eksponencijalnu razdiobu.

5.3 Normalna razdioba

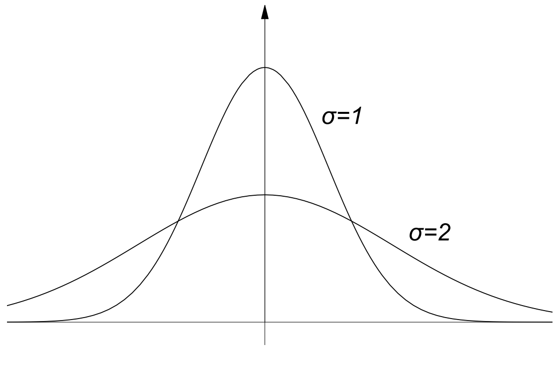
U ovoj točki upoznat ćemo najvažniju neprekinutu razdiobu, odnosno normalnu razdiobu. Naime, ta razdioba se najčešće pojavljuje u primjenama. Razlog za to je činjenica da takvu razdiobu ima slučajna varijabla koja je dobivena kao zbroj velikog broja međusobno nezavisnih slučajnih varijabli. Normalnu razdiobu precizno definiramo pomoću njezine funkcije gustoće.

Normalna razdioba, definicija
Slučajna varijabla X ima **normalnu razdiobu** s parametrima $a \in \mathbf{R}$ i $\sigma^2 > 0$ ako je X neprekinuta slučajna varijabla s gustoćom

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right).$$

Pišemo $X \sim \mathcal{N}(a, \sigma^2)$.

Graf ove funkcije je zvonolika krivulja koju nazivamo **Gaussova krivulja**. Zbog toga se normalna razdioba često naziva i Gaussova razdioba. Na donjoj slici prikazan je graf gustoće normalne razdiobe za parametre $a = 0$ te $\sigma = 1$ i $\sigma = 2$.



Promotrimo li formulu kojom je zadana funkcija gustoće, vidimo da integral te funkcije nije elementaran. Prema tome, kod normalne razdiobe ne postoji eksplicitna formula za funkciju razdiobe kao što je bio slučaj kod jednolike ili eksponencijalne razdiobe. Zato ćemo morati računati s približnim vrijednostima te će nam od posebnog interesa biti takozvana jedinična normalna razdioba.

Jedinična normalna razdioba

Odaberemo li parametre $a = 0$ i $\sigma = 1$, dobivamo slučajnu varijablu $\mathcal{N}(0,1)$ koju nazivamo jedinična normalna razdioba. Neka je ϕ gustoća te razdiobe, odnosno

$$\phi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}.$$

Pripadna funkcija razdiobe Φ definirana je formulom

$$\Phi(u) = \int_{-\infty}^{u} \phi(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{1}{2}t^2} dt.$$

Kao što smo već uočili, posljednji integral nije elementaran te se ne može eksplicitno izraziti pomoću elementarnih funkcija. Zbog toga ćemo funkciju Φ izraziti pomoću jedne malo jednostavnije funkcije čije su vrijednosti tabelirane.

Uočimo ponajprije kako je gustoća ϕ parna funkcija. Zbog toga vrijede sljedeća svojstva:

$$\int_{-\infty}^{0} \phi(t)dt = \frac{1}{2} \int_{-\infty}^{\infty} \phi(t)dt = \frac{1}{2},$$

$$\int_{0}^{u} \phi(t)dt = \frac{1}{2} \int_{-u}^{u} \phi(t)dt.$$

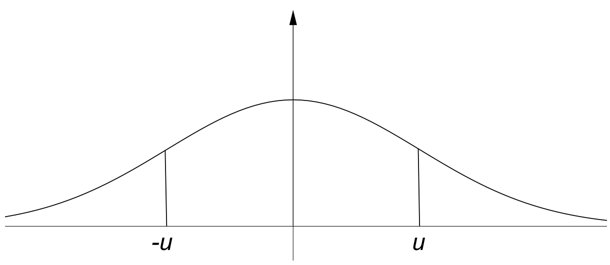
Sada možemo pisati

$$\Phi(u) = \int_{-\infty}^{u} \phi(t)dt = \int_{-\infty}^{0} \phi(t)dt + \int_{0}^{u} \phi(t)dt = \frac{1}{2} + \frac{1}{2} \int_{-u}^{u} \phi(t)dt = \frac{1}{2}\left[1 + \Phi^*(u)\right].$$

Prema tome, funkcija razdiobe Φ može se prikazati pomoću funkcije

$$\Phi^*(u) = \frac{1}{\sqrt{2\pi}} \int_{-u}^{u} e^{-\frac{1}{2}t^2} dt.$$

Očito, vrijednost funkcije Φ^* jednaka je površini ispod grafa funkcije nad odgovarajućim simetričnim intervalom.



Integral kojim je definirana funkcija Φ^* također nije elementaran, pa ćemo koristiti tabelirane vrijednosti te funkcije. Iz matematičke analize znamo da zamjenom granica integral mijenja predznak.

Zbog toga je funkcija Φ^* neparna, pa je dovoljno znati njezine vrijednosti za pozitivne vrijednosti od u. Tablica s vrijednostima funkcije Φ^* nalazi se na kraju ove knjige.

No što ako uzmemo normalnu razdiobu različitu od jedinične? Dakako, ne možemo tabelirati vrijednosti za sve moguće normalne razdiobe. Upravo nam je zbog toga važna jedinična normalna razdioba. Naime, svaku normalnu razdiobu različitu od jedinične, svodit ćemo na jediničnu razdiobu.

Lagano se pokazuje da bilo koju normalnu razdiobu linearnom transformacijom svodimo na jediničnu normalnu razdiobu. Preciznije, vrijedi sljedeća veza između jedinične i općenite normalne razdiobe.

Veza jedinične i opće normalne razdiobe
Jedinična i opća normalna razdioba mogu se dobiti jedna iz druge linearnom transformacijom:
$$X \sim \mathcal{N}(0,1) \implies a + \sigma X \sim \mathcal{N}(a,\sigma^2),$$
$$X \sim \mathcal{N}(a,\sigma^2) \implies \frac{X-a}{\sigma} \sim \mathcal{N}(0,1).$$

Očekivanje i disperzija: parametri normalne razdiobe

Izračunajmo očekivanje i disperziju jedinične normalne razdiobe $X \sim \mathcal{N}(0,1)$. Tada je $a = 0$, $\sigma = 1$ te je gustoća
$$\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}.$$
Prema tome, za očekivanje jedinične normalne razdiobe vrijedi
$$E(X) = \int_{-\infty}^{\infty} t\phi(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} te^{-\frac{1}{2}t^2} dt = 0,$$
zato jer je podintegralna funkcija neparna.

Disperziju jedinične normalne razdiobe računamo primjenom formule za parcijalnu integraciju. Imamo redom
$$\begin{aligned} D(X) &= E(X^2) - [E(X)]^2 = E(X^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-\frac{1}{2}t^2} dt \\ &= \left(\frac{-te^{-\frac{1}{2}t^2}}{\sqrt{2\pi}} \right) \Bigg|_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}t^2} dt = 0 + \int_{-\infty}^{\infty} \phi(t) dt = 1, \end{aligned}$$
pri čemu smo iskoristili činjenicu da je integral gustoće jednak jedinici.

Očekivanje i disperziju općenite normalne razdiobe odredit ćemo iz relacije koja povezuje jediničnu i općenitu razdiobu. Neka je $X \sim \mathcal{N}(a,\sigma^2)$. Tada je $\frac{X-a}{\sigma}$ jedinična normalna razdioba. Stoga je
$$E\left(\frac{X-a}{\sigma}\right) = 0.$$
Iz prethodne jednakosti i svojstava očekivanja dobivamo da je
$$\frac{1}{\sigma}E(X) - E\left(\frac{a}{\sigma}\right) = \frac{1}{\sigma}E(X) - \frac{a}{\sigma} = 0,$$
odakle je $E(X) = a$.

Slično, kako je
$$D\left(\frac{X-a}{\sigma}\right) = 1,$$

iz svojstava disperzije imamo da je

$$\frac{1}{\sigma^2}D(X) + D\left(\frac{a}{\sigma}\right) = \frac{1}{\sigma^2}D(X) + 0 = 1,$$

pa je $D(X) = \sigma^2$.

Time smo odredili značenje parametara normalne razdiobe: parametar a jednak je očekivanju, a parametar σ^2 disperziji normalne razdiobe. Korijen iz disperzije je σ, koji se naziva standardno odstupanje ili standardna devijacija normalne razdiobe.

Računanje vjerojatnosti normalne razdiobe

Neka je X jedinična normalna razdioba. U nastavku ćemo vidjeti kako se računaju vjerojatnosti da slučajna varijabla poprimi vrijednost unutar nekog intervala. Naime, iz Teorema 5.1 i definicije funkcije Φ^* lagano dobivamo sljedeće pravilo.

Računanje vjerojatnosti za jediničnu normalnu razdiobu

Za jediničnu normalnu razdiobu X vrijedi

$$P(u_1 < X < u_2) = \Phi(u_2) - \Phi(u_1) = \frac{1}{2}\left[\Phi^*(u_2) - \Phi^*(u_1)\right].$$

Posebno, u slučaju simetričnog intervala vrijedi

$$P(|X| < u) = \frac{1}{2}\left[\Phi^*(u) - \Phi^*(-u)\right] = \Phi^*(u).$$

Primjer 5.10. *Neka je X jedinična normalna varijabla. Odredite vjerojatnost događaja*
(a) $0 < X < 2$ *(b)* $-2 < X < 1$ *(c)* $-2 < X < 2$ *(d)* $-2 < X < -1$ *(e)* $X < 2$ *(f)* $X > 2$

Rješenje:

$$\begin{aligned}
&P(0 < X < 2) = \frac{1}{2}\left[\Phi^*(2) - \Phi^*(0)\right] = \frac{1}{2}\Phi^*(2) = 0.477,\\
&P(-2 < X < 1) = \frac{1}{2}\left[\Phi^*(1) - \Phi^*(-2)\right] = \frac{1}{2}\left[\Phi^*(1) + \Phi^*(2)\right] = 0.819,\\
&P(-2 < X < 2) = \Phi^*(2) = 0.954,\\
&P(-2 < X < -1) = \frac{1}{2}\left[\Phi^*(-1) - \Phi^*(-2)\right] = \frac{1}{2}\left[\Phi^*(2) - \Phi^*(1)\right] = 0.136,\\
&P(X < 2) = \Phi(2) = \frac{1}{2} + \frac{1}{2}\Phi^*(2) = 0.977,\\
&P(X > 2) = 1 - \Phi(2) = \frac{1}{2} - \frac{1}{2}\Phi^*(2) = 0.023.
\end{aligned}$$

□

U prethodnom primjeru naučili smo računati vjerojatnosti jedinične normalne razdiobe. Vjerojatnost općenite normalne razdiobe računamo svođenjem na jediničnu normalnu razdiobu. Neka je $X \sim \mathcal{N}(a, \sigma^2)$. Tada je $\dfrac{X-a}{\sigma} \sim \mathcal{N}(0, 1)$ i funkciju razdiobe F varijable X možemo izraziti uz pomoć funkcije Φ:

$$F(x) = \Phi(u) = \frac{1}{2}\left[1 + \Phi^*(u)\right], \quad u = \frac{x-a}{\sigma}.$$

Tada vrijedi sljedeće pravilo:

$$P(x_1 < X < x_2) = P\left(\frac{x_1 - a}{\sigma} < \frac{X - a}{\sigma} < \frac{x_2 - a}{\sigma}\right) = \Phi\left(\frac{x_2 - a}{\sigma}\right) - \Phi\left(\frac{x_1 - a}{\sigma}\right)$$
$$= \frac{1}{2}\left[\Phi^*\left(\frac{x_2 - a}{\sigma}\right) - \Phi^*\left(\frac{x_1 - a}{\sigma}\right)\right].$$

Primjer 5.11. *Neka je* $X \sim \mathcal{N}(2, 9)$. *Izračunajte vjerojatnosti*
(a) $P(-1 < X < 5)$ *(b)* $P(5 < X < 8)$ *(c)* $P(X < 3.5)$.

Rješenje: Ako X ima općenitu normalnu razdiobu, tada ćemo sa $\widetilde{X}$ označavati pripadnu jediničnu normalnu slučajnu varijablu, odnosno $\widetilde{X} = \dfrac{X - a}{\sigma}$. Tada imamo redom

$$P(-1 < X < 5) = P\left(\frac{-1 - 2}{3} < \frac{X - 2}{3} < \frac{5 - 2}{3}\right) = P(-1 < \widetilde{X} < 1) = \Phi^*(1) = 0.683,$$

$$P(5 < X < 8) = P\left(\frac{5 - 2}{3} < \frac{X - 2}{3} < \frac{8 - 2}{3}\right) = P(1 < \widetilde{X} < 2) = \frac{1}{2}\left[\Phi^*(2) - \Phi^*(1)\right] = 0.136,$$

$$P(X < 3.5) = P\left(\frac{X - 2}{3} < \frac{3.5 - 2}{3}\right) = P(\widetilde{X} < 0.5) = \frac{1}{2} + \frac{1}{2}\Phi^*(0.5) = 0.692.$$

□

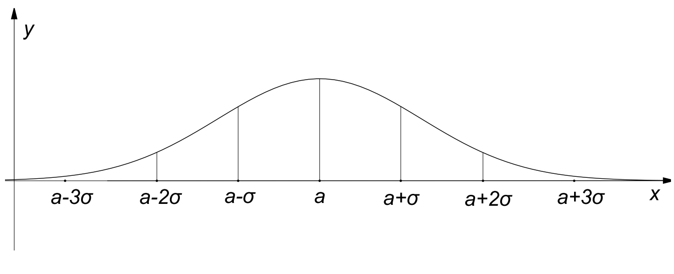
Promotrimo li tablicu s vrijednostima funkcije Φ^*, vidimo da ona sadrži samo vrijednosti za argumente od 0 do 4. Očito za argumente veće od 4 vrijednost te funkcije je približno jednaka 1. S druge strane, funkcija gustoće poprima najveću vrijednost u točki očekivanja te s obje strane pada prema nuli. Dakako, brzina tog pada ovisi o disperziji. Sljedeće pravilo detaljnije objašnjava te činjenice.

Pravilo tri sigma

Neka je $X \sim \mathcal{N}(a, \sigma^2)$ općenita normalna razdioba. Izračunajmo vjerojatnosti

$$P(|X - a| < k\sigma),$$

za parametre $k = 1, 2, 3$.



Svođenjem na jediničnu razdiobu dobivamo

$$P(|X - a| < k\sigma) = P(-k\sigma < X - a < k\sigma) = P(-k < \widetilde{X} < k) = \Phi^*(k).$$

Sada, iz tablica čitamo $\Phi^*(1) = 0.6827$, $\Phi^*(2) = 0.9545$, $\Phi^*(3) = 0.9973$. Prema tome, normalna varijabla s vjerojatnošću 99.73%, tj. gotovo sigurno, poprima vrijednosti unutar intervala $(a - 3\sigma, a + 3\sigma)$. To svojstvo naziva se pravilo tri sigma. Dakako, normalna slučajna varijabla također može poprimiti vrijednosti i izvan tog intervala, ali je vjerojatnost takvog događaja približno jednaka nuli.

Primjer 5.12. *Prosječan vijek trajanja automobilske gume, izražen u prijeđenim kilometrima, je normalna slučajna varijabla s očekivanjem* 35000 *km i odstupanjem* 5000 *km.*

(a) Kolika je vjerojatnost da guma neće puknuti do prijeđenih 41000 *km?*

(b) Kolika je vjerojatnost da će guma puknuti između 32000 *km i* 39000 *km?*

Rješenje: Neka je X zadana slučajna varijabla, odnosno $X \sim \mathcal{N}(35000, 5000^2)$. Sada redom imamo da je

$$P(X > 41000) = P\left(\frac{X - 35000}{5000} > \frac{41000 - 35000}{5000}\right) = P\left(\widetilde{X} > 1.2\right) = \frac{1}{2} - \frac{1}{2}\Phi^*(1.2) = 0.115,$$

$$\begin{aligned} P(32000 < X < 39000) &= P\left(\frac{32000 - 35000}{5000} < \frac{X - 35000}{5000} < \frac{39000 - 35000}{5000}\right) \\ &= \frac{1}{2}\left[\Phi^*\left(\frac{4}{5}\right) - \Phi^*\left(-\frac{3}{5}\right)\right] = \frac{1}{2}\left[\Phi^*(0.8) + \Phi^*(0.6)\right] = 0.514. \end{aligned}$$

□

Primjer 5.13. *Vrijeme koje student provede na putu od kuće do fakulteta je slučajna varijabla približno distribuirana po normalnom zakonu s očekivanjem* 50 *minuta. Student kreće iz kuće u* 07 : 20 *da bi stigao na predavanje koje počinje u* 08 : 15*. Ako je vjerojatnost da će stići na fakultet u vremenskom intervalu od* 08 : 05 *do* 08 : 15 *jednaka* 0.383*, kolika je vjerojatnost da će kasniti na predavanje više od* 5 *minuta?*

Rješenje: Neka je X zadana slučajna varijabla, odnosno $X \sim \mathcal{N}(50, \sigma^2)$. Iz zadanih uvjeta izračunat ćemo disperziju tj. odstupanje razdiobe. Ako student stiže na fakultet u vremenskom intervalu od 08 : 05 do 08 : 15, onda njegov put do fakulteta traje od 45 do 55 minuta. Iz uvjeta zadatka imamo da je $P(45 < X < 55) = 0.383$. Svođenjem na jediničnu razdiobu dobivamo da je

$$P\left(\frac{45 - 50}{\sigma} < \frac{X - 50}{\sigma} < \frac{55 - 50}{\sigma}\right) = P\left(-\frac{5}{\sigma} < \widetilde{X} < \frac{5}{\sigma}\right) = \Phi^*\left(\frac{5}{\sigma}\right) = 0.383.$$

Sada, iz tablica normalne razdiobe imamo da je

$$\frac{5}{\sigma} = \frac{1}{2},$$

odakle je $\sigma = 10$. Konačno, student će kasniti na predavanje više od 5 minuta ako njegov put traje više od 60 minuta, pa je

$$P(X > 60) = P\left(\frac{X - 50}{10} > \frac{60 - 50}{10}\right) = P\left(\widetilde{X} > 1\right) = \frac{1}{2} - \frac{1}{2}\Phi^*(1) = 0.159.$$

□

Stabilnost normalne razdiobe

Prisjetimo se, Poissonova razdioba posjeduje svojstvo stabilnosti, odnosno zbroj nekoliko Poissonovih slučajnih varijabli je opet Poissonova slučajna varijabla. Isto vrijedi i za binomnu razdiobu.

Normalna razdioba ima pojačano svojstvo stabilnosti. Naime, ne samo da je zbroj normalnih varijabli opet normalna varijabla, nego je i linearna kombinacija normalnih varijabli normalna varijabla. O tom važnom svojstvu govori sljedeći teorem.

Stabilnost normalne razdiobe
Teorem 5.3 *Neka su X_1 i X_2 nezavisne slučajne varijable s normalnim razdiobama $X_1 \sim \mathcal{N}(a_1, \sigma_1^2)$, $X_2 \sim \mathcal{N}(a_2, \sigma_2^2)$ i α_1, α_2 bilo koji realni brojevi. Tada vrijedi*

$$\alpha_1 X_1 + \alpha_2 X_2 \sim \mathcal{N}\left(\alpha_1 a_1 + \alpha_2 a_2, \alpha_1^2 \sigma_1^2 + \alpha_2^2 \sigma_2^2\right)$$

Može se pokazati da navedeno svojstvo karakterizira normalnu razdiobu, odnosno da je normalna razdioba jedina koja posjeduje pojačano svojstvo stabilnosti.

Primjer 5.14. *Međusobno nezavisne slučajne varijable X, Y, Z podvrgavaju se normalnim razdiobama redom $X \sim \mathcal{N}(1,1)$, $Y \sim \mathcal{N}(2,2)$, $Z \sim \mathcal{N}(3,3)$. Izračunajte vjerojatnost događaja $\{X + 3Z > 2Y\}$.*

Rješenje: Neka je $W = X - 2Y + 3Z$. Tražimo vjerojatnost $P(W > 0)$. Prema teoremu o stabilnosti normalne razdiobe, W je normalna slučajna varijabla s parametrima

$$\begin{aligned} a &= 1 \cdot 1 - 2 \cdot 2 + 3 \cdot 3 = 6, \\ \sigma^2 &= 1 \cdot 1 + 4 \cdot 2 + 9 \cdot 3 = 36. \end{aligned}$$

Dalje imamo

$$P(W > 0) = P\left(\frac{W-6}{6} > \frac{-6}{6}\right) = P(\widetilde{W} > -1) = \frac{1}{2} - \frac{1}{2}\Phi^*(-1) = \frac{1}{2} + \frac{1}{2}\Phi^*(1) = 0.841.$$

□

Primjer 5.15. *Masa limuna podvrgava se normalnom zakonu s parametrima $a_1 = 15$ dkg i $\sigma_1 = 2$ dkg, a masa naranče normalnom zakonu s parametrima $a_2 = 25$ dkg i $\sigma_2 = 3$ dkg. U jednu vrećicu pakiraju se po tri limuna i tri naranče. Odredite vjerojatnost da se masa tako načinjenog paketa kreće između 115 dkg i 130 dkg. Zanemarite masu vrećice.*

Rješenje: Neka su X_L i X_N slučajne varijable koje opisuju masu limuna, odnosno naranče. Prema uvjetima zadatka je

$$X_L \sim \mathcal{N}(15, 2^2), \qquad X_N \sim \mathcal{N}(25, 3^2).$$

Označimo s Y masu paketa. On sadrži tri limuna i tri naranče, pa bismo mogli napisati $Y = 3X_L + 3X_N$. Međutim, taj zapis nije ispravan. Naime, mi ne računamo trostruku masu limuna, nego zbroj triju masa koje su nezavisne jedna od druge, a imaju istu razdiobu. Zato je ispravno napisati

$$Y = X_L' + X_L'' + X_L''' + X_N' + X_N'' + X_N'''.$$

Pri tome su X_L', X_L'' i X_L''' nezavisne kopije slučajne varijable X_L, odnosno, nezavisne slučajne varijable koje imaju istu razdiobu kao i X_L. Isto vrijedi i za slučajne varijable X_N', X_N'' i X_N'''.

Prema teoremu 5.3, Y je normalna slučajna varijabla s razdiobom

$$Y \sim \mathcal{N}(15 + 15 + 15 + 25 + 25 + 25, 4 + 4 + 4 + 9 + 9 + 9) = \mathcal{N}(120, 39).$$

Trebamo izračunati vjerojatnost $P(115 < Y < 130)$:

$$P\left(\frac{115-120}{\sqrt{39}} < \frac{Y-120}{\sqrt{39}} < \frac{130-120}{\sqrt{39}}\right) = P(-0.801 < \widetilde{Y} < 1.601)$$
$$= \frac{1}{2}\left[\Phi^*(1.601) + \Phi^*(0.801)\right] = \frac{1}{2}\left[0.893 + 0.577\right] = 0.735.$$

□

Aproksimacija binomne razdiobe normalnom

Jedno od važnih svojstava normalne razdiobe jest i to da pomoću nje možemo aproksimirati binomnu razdiobu. Potreba za tom aproksimacijom javlja se zbog teškoće računanja binomnih koeficijenata za velike brojeve.

Naime, neka je $X \sim \mathcal{B}(n, p)$ binomna slučajna varijabla. Ukoliko je n dovoljno velik, razdioba ove varijable nalikuje funkciji gustoće normalne varijable $Y \sim \mathcal{N}(np, npq)$. Za fiksni n kvaliteta aproksimacije je bolja što je p bliži $\frac{1}{2}$. Navedimo sada rezultat pomoću kojeg približno računamo vjerojatnost da binomna slučajna varijabla poprimi određenu vrijednost.

Lokalni teorem Moivre-Laplacea
Teorem 5.4 *Vjerojatnosti realizacija binomne slučajne varijable $\mathcal{B}(n, p)$ mogu se aproksimirati pomoću funkcije gustoće f normalne varijable $\mathcal{N}(pq, npq)$:*

$$P(X = m) \approx P(m - \frac{1}{2} < Y < m + \frac{1}{2}) \approx f(m) = \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(m-np)^2}{2npq}}.$$

Primjer 5.16. *Tvornica žarulja isporučuje veliku seriju žarulja. Vjerojatnost da slučajno odabrana žarulja iz velike serije bude neispravna iznosi* 0.01*. Kolika je vjerojatnost da će među* 2000 *slučajno odabranih žarulja njih* 20 *biti neispravno?*

Rješenje: Broj neispravnih žarulja je slučajna varijabla s razdiobom $X \sim \mathcal{B}(2000, 0.01)$. Stoga je

$$P(X = 20) = \binom{2000}{20} \cdot 0.01^{20} \cdot 0.99^{1980}.$$

Uočimo kako je teško izračunati vrijednost prethodnog izraza pomoću džepnog računala. Zbog toga ćemo približnu vrijednost te vjerojatnosti izračunati pomoću lokalnog Moivre-Laplaceovog teorema. Imamo

$$P(X = 20) \approx \frac{1}{\sqrt{2\pi \cdot 2000 \cdot 0.01 \cdot 0.99}} e^{-\frac{(20-2000\cdot 0.01)^2}{2\cdot 2000\cdot 0.01\cdot 0.99}} = \frac{1}{\sqrt{39.6\pi}} = 0.089.$$

□

Centralni granični teorem – Teorem Moivre-Laplacea

Kao što smo već vidjeli, pomoću lokalnog teorema Moivre-Laplacea približno računamo vjerojatnost da binomna varijabla poprimi određenu vrijednost. Postavlja se pitanje kako odrediti vjerojatnost da takva varijabla poprimi neku od vrijednosti iz diskretnog intervala. Odgovor na to pitanje daje centralni granični teorem, jedan od najvažnijih teorema u teoriji vjerojatnosti. Mi ćemo ovdje iskazati najvažniji posebni slučaj tog teorema, koji se odnosi na binomne slučajne varijable.

Neka X ima binomnu razdiobu, $X \sim \mathcal{B}(n, p)$. Očekivanje ove varijable je np, a disperzija npq. Zato varijabla $\dfrac{X - np}{\sqrt{npq}}$ ima očekivanje 0 i disperziju 1.

Teorem Moivre-Laplacea

Teorem 5.5 *Neka je* $X \sim \mathcal{B}(n,p)$*. Za veliki* n *razdioba slučajne varijable* $\frac{X-np}{\sqrt{npq}}$ *može se aproksimirati jediničnom normalnom razdiobom*

$$P\left(x_1 < \frac{\mathcal{B}(n,p)-np}{\sqrt{npq}} < x_2\right) \approx \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{1}{2}u^2} du.$$

Pišemo $\mathcal{B}(n,p) \approx \mathcal{N}(np, npq)$*, za dovoljno veliki* n*.*

Aproksimacija je dobra već za relativno male vrijednosti broja n, npr. za $n \geq 10$. Nadalje, aproksimacija je sve bolja (i točnija za male vrijednosti od n) što je vrijednost parametra p bliža $\frac{1}{2}$.

Primjer 5.17. *Izračunajte vjerojatnost da se u* 18000 *bacanja ispravne kocke broj šestica nalazi između* 2975 *i* 3050.

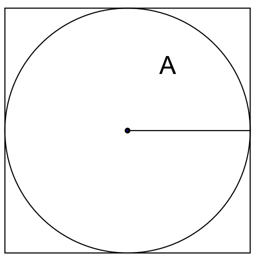
Rješenje: Označimo sa X broj pojavljenih šestica. Tada je $X \sim \mathcal{B}\big(18000, \frac{1}{6}\big)$. Međutim, vjerojatnost događaja $\{2975 < X < 3050\}$ ne može se jednostavno izračunati, zbog teškog računanja binomnih koeficijenata i velikog broja pribrojnika. Stoga X aproksimiramo normalnom razdiobom $\mathcal{N}(3000, 2500)$. Zbog velikog broja $n = 10000$ aproksimacija će biti jako dobra.

$$\begin{aligned} P(2975 < X < 3050) &= P\left(\frac{2975-3000}{\sqrt{2500}} < \frac{X-3000}{\sqrt{2500}} < \frac{3050-3000}{\sqrt{2500}}\right) \\ &= P(-0.5 < \widetilde{X} < 1) = \frac{1}{2}\left[\Phi^*(1) + \Phi^*(0.5)\right] \\ &= \frac{1}{2}\left[0.683 + 0.383\right] = 0.533. \end{aligned}$$

□

Primjer 5.18. *Slučajno odabiremo* 1000 *točaka unutar kvadrata. Kolika je vjerojatnost da će više od* 800 *točaka biti unutar kruga upisanog tom kvadratu?*

Rješenje: Neka je A krug upisan u kvadrat te neka je a njegov polumjer. Tada je duljina stranice kvadrata jednaka $2a$.



Kako je površina kruga jednaka $a^2\pi$, a površina kvadrata $(2a)^2 = 4a^2$, vjerojatnost da slučajno odabrana točka pripada krugu iznosi

$$p = \frac{a^2\pi}{4a^2} = \frac{\pi}{4} = 0.785.$$

Broj točaka unutar kruga je binomna slučajna varijabla $X \sim \mathcal{B}(1000, 0.785)$. Ponovno ćemo, služeći se centralnim graničnim teoremom, aproksimirati X normalnom razdiobom s parametrima

$a = 1000 \cdot 0.785 = 785$ i $\sigma^2 = 1000 \cdot 0.785 \cdot 0.215 = 168.775$, odnosno $X \sim \mathcal{N}(785, 168.775)$. Tražimo vjerojatnost događaja $\{X > 800\}$:

$$\begin{aligned} P(X > 800) &= P(\frac{X - 785}{\sqrt{168.775}} > \frac{800 - 785}{\sqrt{168.775}}) = \frac{1}{2}\left[1 - \Phi^*\left(\frac{15}{\sqrt{168.775}}\right)\right] \\ &= \frac{1}{2}\left[1 - \Phi^*(1.155)\right] = 0.124. \end{aligned}$$

□

U slučaju kada broj ponavljanja pokusa u binomnoj razdiobi nije velik, moramo napraviti korekciju granica prilikom aproksimacije normalnom razdiobom, kako bismo dobili što bolju aproksimaciju. Preciznije, ako binomna slučajna varijabla poprima vrijednosti unutar nekog diskretnog intervala onda donju granicu tog intervala treba umanjiti za 0.5, a gornju povećati za 0.5. Ta korekcija se primjenjuje za relativno male vrijednosti od n. Naime, događaju $\{X = k\}$ za diskretnu varijablu X odgovara događaj $\{k - 0.5 < X < k + 0.5\}$ kod aproksimacije neprekinutom varijablom.

O tome govori sljedeći primjer.

Primjer 5.19. *Simetrični novčić bacamo 80 puta. Neka slučajna varijabla X bilježi broj pisama. Izračunajte vjerojatnost događaja $\{35 \leq X \leq 50\}$.*

Rješenje: Broj pojavljivanja pisama u 80 bacanja novčića je slučajna varijabla $X \sim \mathcal{B}(80, 0.5)$ koju možemo aproksimirati normalnom $\mathcal{N}(40, 20)$. Kako je $n = 80$ relativno malen broj, napravit ćemo korekciju granica zadanog intervala. Dakle, donju granicu intervala smanjit ćemo za 0.5, a gornju povećati za 0.5. Stoga imamo

$$\begin{aligned} P(35 \leq X \leq 50) &= P\left(\frac{34.5 - 40}{\sqrt{20}} < \frac{X - 40}{\sqrt{20}} < \frac{50.5 - 40}{\sqrt{20}}\right) \\ &= P\left(-\frac{5.5}{\sqrt{20}} < \widetilde{X} < \frac{10.5}{\sqrt{20}}\right) \\ &= P\left(-1.23 < \widetilde{X} < 2.35\right) \\ &= \frac{1}{2}\left[\Phi^*(2.35) + \Phi^*(1.23)\right] = 0.881. \end{aligned}$$

□

5.4 Zadatci za ponavljanje

Zadatak 5.1. Zadana je funkcija

$$f(x) = \begin{cases} \alpha(x^2 + 2x), & 0 \leq x \leq 1 \\ 0, & \text{inače,} \end{cases}$$

gdje je α neka realna konstanta.

(a) Odredite konstantu α tako da f bude gustoća razdiobe slučajne varijable X.

(b) Odredite funkciju razdiobe $F(x)$.

(c) Izračunajte vjerojatnost događaja $\{0.2 < X < 0.4\}$.

(d) Izračunajte očekivanje $E(X)$.

Rješenje:

(a) Konstantu α određujemo iz uvjeta $\int_0^1 f(x)dx = 1$. Imamo da je

$$\int_0^1 \alpha(x^2+2x)dx = \alpha\left(\frac{x^3}{3}+2\frac{x^2}{2}\right)\Big|_0^1 = \frac{\alpha}{3}+\alpha = 1,$$

odakle je $4\alpha = 3$, odnosno $\alpha = \frac{3}{4}$.

(b) Funkcija razdiobe je

$$F(x) = \int_0^x \frac{3}{4}(t^2+2t)dt = \frac{3}{4}\left(\frac{t^3}{3}+2\frac{t^2}{2}\right)\Big|_0^x = \frac{x^3+3x^2}{4}, \qquad 0 \le x \le 1.$$

(c) Vjerojatnost događaja $\{0.2 < X < 0.4\}$ računamo pomoću funkcije razdiobe:

$$P(0.2 < X < 0.4) = F(0.4) - F(0.2) = 0.136 - 0.032 = 0.104.$$

(d) Očekivanje slučajne varijable X je

$$E(X) = \int_0^1 \frac{3}{4}x\left(x^2+2x\right)dx = \int_0^1 \left(\frac{3}{4}x^3+\frac{3}{2}x^2\right)dx = \left(\frac{3}{16}x^4+\frac{x^3}{2}\right)\Big|_0^1 = \frac{11}{16} = 0.6875.$$

□

Zadatak 5.2. Slučajna varijabla X dana je funkcijom razdiobe

$$F(x) = \frac{x^3+x}{2}, \qquad 0 \le x \le 1.$$

Izračunajte očekivanje i disperziju slučajne varijable X.

Rješenje: Kako bismo odredili očekivanje i disperziju, prvo trebamo odrediti gustoću slučajne varijable X. Deriviranjem dobivamo

$$f(x) = F'(x) = \frac{3x^2+1}{2}, \qquad 0 \le x \le 1.$$

Očekivanje slučajne varijable X je

$$E(X) = \int_0^1 x \cdot \frac{3x^2+1}{2}dx = \frac{1}{2}\int_0^1 (3x^3+x)dx = \frac{1}{2}\left(\frac{3}{4}x^4+\frac{1}{2}x^2\right)\Big|_0^1 = \frac{5}{8}.$$

Nadalje,

$$E(X^2) = \int_0^1 x^2 \cdot \frac{3x^2+1}{2}dx = \frac{1}{2}\int_0^1 (3x^4+x^2)dx = \frac{1}{2}\left(\frac{3}{5}x^5+\frac{1}{3}x^3\right)\Big|_0^1 = \frac{7}{15},$$

pa je disperzija jednaka

$$D(X) = E(X^2) - (E(X))^2 = \frac{7}{15} - \left(\frac{5}{8}\right)^2 = \frac{73}{960}.$$

□

Zadatak 5.3. Slučajno odabiremo točku na dužini duljine 1. Slučajna varijabla X mjeri udaljenost odabrane točke od polovišta dužine. Izračunajte očekivanje slučajne varijable X.

Rješenje: Odabir točke na dužini duljine 1 ekvivalentan je odabiru broja y unutar intervala $\Omega = [0,1]$. Očito, slučajna varijabla koja mjeri udaljenost od polovišta određena je formulom

$$X = \left| y - \frac{1}{2} \right|.$$

Odredimo razdiobu slučajne varijable X. Slučajna varijabla X poprima vrijednosti u intervalu $\left[0, \frac{1}{2}\right]$ i vrijedi

$$F(x) = P(X < x) = P\left(\left|y - \frac{1}{2}\right| < x\right) = P\left(-x < y - \frac{1}{2} < x\right) = P\left(\frac{1}{2} - x < y < \frac{1}{2} + x\right).$$

Sada, kako je duljina intervala $G_x = [\frac{1}{2} - x, \frac{1}{2} + x]$ jednaka $2x$, slijedi da je

$$F(x) = \frac{m(G_x)}{m(\Omega)} = \frac{2x}{1} = 2x, \quad 0 \leq x \leq \frac{1}{2}.$$

Konačno, deriviranjem dobivamo funkciju gustoće

$$f(x) = F'(x) = 2, \quad 0 \leq x \leq \frac{1}{2},$$

pa je traženo očekivanje jednako

$$E(X) = \int_0^{\frac{1}{2}} 2x dx = x^2 \Big|_0^{\frac{1}{2}} = \frac{1}{4}.$$

□

Zadatak 5.4. Roko lovi ribu. Vrijeme do ulova ribe je slučajna varijabla X s gustoćom

$$f(x) = \frac{1}{3} e^{-\frac{x}{3}}, \quad x \geq 0.$$

Ako Roko nije ulovio ribu tijekom prva tri sata, kolika je vjerojatnost da je on ulovi tijekom četvrtog sata?

Rješenje: Iz formule za gustoću prepoznajemo da je riječ o eksponencijalnoj razdiobi s parametrom $\lambda = \frac{1}{3}$, tj. $X \sim \mathcal{E}\left(\frac{1}{3}\right)$. Znamo da je pripadna funkcija razdiobe dana formulom $F(x) = 1 - e^{-\frac{x}{3}}$, $x \geq 0$, pa imamo:

$$\begin{aligned}
P(3 < X < 4 \,|X > 3) &= \frac{P(3 < X < 4, X > 3)}{P(X > 3)} = \frac{P(3 < X < 4)}{P(X > 3)} \\
&= \frac{F(4) - F(3)}{1 - F(3)} = \frac{\left(1 - e^{-\frac{4}{3}}\right) - \left(1 - e^{-1}\right)}{e^{-1}} \\
&= \frac{e^{-1} - e^{-\frac{4}{3}}}{e^{-1}} = 1 - e^{-\frac{1}{3}} = 0.283.
\end{aligned}$$

Uočimo kako smo zadatak mogli riješiti i jednostavnije, uzmemo li u obzir svojstvo odsustva pamćenja eksponencijalne razdiobe. Naime, vrijedi

$$P(3 < X < 4 \,|X > 3) = P(X < 1) = F(1) = 1 - e^{-\frac{1}{3}} = 0.283.$$

□

Zadatak 5.5. Slučajna varijabla X ima normalnu razdiobu s očekivanjem $a = 2$ i vrijedi $P(X < 4) = 0.6915$. Izračunajte vjerojatnost događaja $\{-2 < X < 5\}$.

Rješenje: Neka je $X \sim \mathcal{N}(2, \sigma^2)$. Iz zadanih uvjeta izračunat ćemo disperziju razdiobe. Imamo redom

$$P\left(\frac{X-2}{\sigma} < \frac{4-2}{\sigma}\right) = P\left(\widetilde{X} < \frac{2}{\sigma}\right) = \frac{1}{2} + \frac{1}{2}\Phi^*\left(\frac{2}{\sigma}\right) = 0.6915,$$

odakle je

$$\Phi^*\left(\frac{2}{\sigma}\right) = 0.383.$$

Iz tablica normalne razdiobe imamo da je $\frac{2}{\sigma} = \frac{1}{2}$, odakle je $\sigma = 4$. Sada lagano dobivamo

$$P(-2 < X < 5) = P\left(\frac{-2-2}{4} < \frac{X-2}{4} < \frac{5-2}{4}\right) = P\left(-1 < \widetilde{X} < \frac{3}{4}\right)$$
$$= \frac{1}{2}\left[\Phi^*(0.75) + \Phi^*(1)\right] = \frac{1}{2}(0.683 + 0.547) = 0.615.$$

□

Zadatak 5.6. Izračunajte vjerojatnost da u 3600 bacanja ispravnog novčića broj pisama bude veći od broja glava.

Rješenje: Označimo sa X broj pojavljenih pisama. Tada je $X \sim \mathcal{B}\big(3600, \frac{1}{2}\big)$. Tu binomnu razdiobu aproksimiramo normalnom razdiobom s parametrima $a = 3600 \cdot \frac{1}{2} = 1800$ i $\sigma^2 = 3600 \cdot \frac{1}{2} \cdot \frac{1}{2} = 900$, tj $X \sim \mathcal{N}(1800, 900)$. Zbog toga je tražena vjerojatnost jednaka

$$P(X > 1800) = P\left(\frac{X - 1800}{30} > \frac{1800-1800}{30}\right) = P\left(\widetilde{X} > 0\right) = \frac{1}{2}\left[1 - \Phi^*(0)\right] = 0.5,$$

što je logičan rezultat. □

Zadatak 5.7. Masa proizvoda X je normalna slučajna varijabla s očekivanjem 6 i odstupanjem 0.6, a masa proizvoda Y je normalna slučajna varijabla s očekivanjem 5 i odstupanjem 0.3. Ako na lijevi krak vage stavimo 4 proizvoda tipa X, a na desni krak iste vage 5 proizvoda tipa Y, kolika je vjerojatnost da će lijevi krak pretegnuti?

Rješenje: Prema uvjetima zadatka vrijedi $X \sim \mathcal{N}(6, 0.6^2)$ i $Y = \mathcal{N}(5, 0.3^2)$. Na lijevom kraku vage nalaze se četiri proizvoda X. Masa svakog od njih je distribuirana po zakonu $X \sim \mathcal{N}(6, 0.6^2)$ i te slučajne varijable su nezavisne. Prema tome, masa na lijevom kraku vage opisana je slučajnom varijablom

$$X_1 + X_2 + X_3 + X_4,$$

pri čemu sve četiri slučajne varijable imaju isti zakon razdiobe kao i slučajna varijabla X.

Slično, masa na desnom kraku vage opisana je slučajnom varijablom

$$Y_1 + Y_2 + Y_3 + Y_4 + Y_5,$$

pri čemu svih pet slučajnih varijabli imaju isti zakon razdiobe kao i slučajna varijabla Y.

Mi tražimo vjerojatnost događaja

$$\{X_1 + X_2 + X_3 + X_4 > Y_1 + Y_2 + Y_3 + Y_4 + Y_5\},$$

odnosno $P(Z > 0)$, gdje je Z slučajna varijabla

$$Z = X_1 + X_2 + X_3 + X_4 - Y_1 - Y_2 - Y_3 - Y_4 - Y_5.$$

Prema teoremu o stabilnosti, Z je normalna slučajna varijabla s parametrima

$$\begin{aligned} a &= 6+6+6+6-5-5-5-5-5 = -1 \\ \sigma^2 &= 0.36+0.36+0.36+0.36+0.09+0.09+0.09+0.09+0.09 = 1.89. \end{aligned}$$

Zbog toga je tražena vjerojatnost jednaka

$$P(Z > 0) = P\left(\frac{Z+1}{\sqrt{1.89}} > \frac{0+1}{\sqrt{1.89}}\right) = P(\widetilde{Z} > 0.727) = \frac{1}{2}\left[1 - \Phi^*(0.727)\right] = 0.234.$$

$\square$

6 Opisna statistika

Kocka je bačena 1000 puta, pri čemu se jedinica pojavila 350 puta. Da li je kocka simetrična te da li je način bacanja bio ispravan? Tijekom radnog tjedna student bilježi vrijeme koje provede na putu od kuće do fakulteta pri čemu dobiva sljedeće podatke: 45, 49, 53, 41, 50 minuta. Ako je poznato da je to vrijeme distribuirano po normalnom zakonu, kako možemo odrediti parametre te razdiobe?

Na ova i slična pitanja, odgovor daje **matematička statistika**. U ovom poglavlju upoznat ćemo se s osnovnim pojmovima **opisne** ili **deskriptivne statistike**. Opisna ili deskriptivna statistika je grana statistike koja se bavi predočavanjem i opisivanjem glavnih karakteristika sakupljenih podataka (tablice, grafikoni, histogrami, srednje vrijednosti).

Prilikom opažanja ili eksperimentiranja, pažnja istraživača redovito je usmjerena na jednu ili više veličina. Promatramo li samo jednu veličinu, u oznaci X, onda je rezultat jednog mjerenja realni broj x. Višestrukim ponavljanjem mjerenja veličine X dobivamo konačan niz brojeva $x_1, x_2, \ldots, x_n$ kao rezultat od n ponovljenih mjerenja veličine X. Taj niz je **realizacija** veličine X. Veličina X obično se naziva **statističko obilježje**, a dobiveni niz $x_1, x_2, \ldots, x_n$ predstavlja **statističke podatke** o promatranom statističkom obilježju X.

6.1 Grafički prikaz podataka

U ovoj točki upoznat ćemo se s osnovnim grafičkim prikazima statističkih podataka. Promotrimo ponajprije sljedeći primjer.

Primjer 6.1. *U nekoj srednjoj školi ima* 20 *razrednih odjela. Na kraju polugodišta zabilježeni su sljedeći podatci o broju negativnih ocjena iz matematike u pojedinom razrednom odjelu:*

$$\begin{array}{cccccccccc} 4 & 5 & 3 & 5 & 1 & 3 & 1 & 6 & 6 & 2 \\ 3 & 3 & 4 & 6 & 4 & 3 & 1 & 4 & 1 & 4 \end{array}.$$

Na ovom primjeru objasnit ćemo pojmove koje smo definirali u uvodu ovog poglavlja te ćemo definirati još neke pojmove koji će nam biti važni prilikom grafičkog predočavanja podataka.

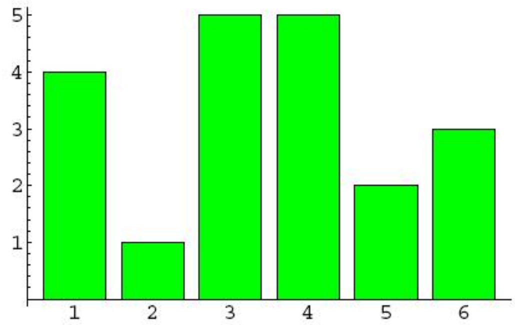
- statističko obilježje X predstavlja broj negativnih ocjena iz matematike u pojedinom razrednom odjelu
- X poprima vrijednosti iz skupa $\{1, 2, 3, 4, 5, 6\}$. Taj skup obično označavamo s $\operatorname{Im} X$, tj. $\operatorname{Im} X = \{1, 2, 3, 4, 5, 6\}$
- u ovom primjeru je $\operatorname{Im} X$ diskretan, tj. konačan skup, pa kažemo da je X **diskretno obilježje**
- obilježje može biti **numeričko** ili **nenumeričko**
- nenumeričko obilježje nazivamo i **kategorija**, npr. spol, boja, vrsta ...

- svakom elementu $a_i \in \operatorname{Im} X$ možemo pridružiti broj f_i tj. **frekvenciju pojavljivanja elementa** a_i u nizu podataka
- broj $f_{r_i} = \frac{f_i}{n}$ naziva se relativna frekvencija elementa a_i. Broj n predstavlja broj ponavljanja pokusa, odnosno broj mjerenja podataka. U ovom primjeru je $n = 20$.

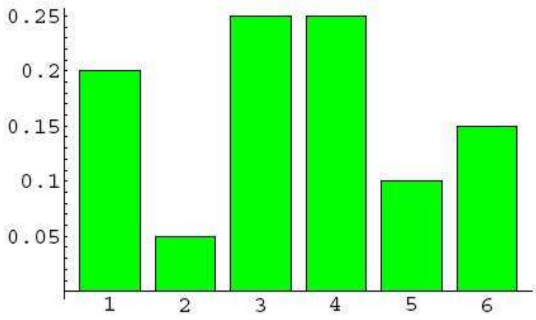
Prikažimo podatke iz Primjera 6.1 u **tablici frekvencija**.

a_i	f_i	f_{r_i}
1	4	$4/20 = 0.2$
2	1	$1/20 = 0.05$
3	5	$5/20 = 0.25$
4	5	$5/20 = 0.25$
5	2	$2/20 = 0.1$
6	3	$3/20 = 0.15$
Σ	20	1.00

Podatke također možemo prikazivati i pomoću **stupčastog dijagrama** (**bar chart**).



Uočimo kako je na prethodnoj slici visina svakog stupca jednaka odgovarajućoj frekvenciji. Stupčasti dijagram može se crtati i tako da visina svakog stupca bude jednaka odgovarajućoj relativnoj frekvenciji. Tada je ukupna površina svih stupaca jednaka 1, što je često bolje zbog usporedbe za različite vrijednosti n:



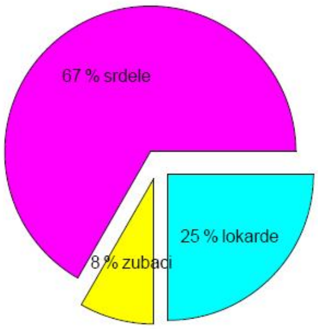
Ukoliko statističko obilježje može poprimiti malo različitih vrijednosti, onda je zgodno nacrtati **strukturni krug** (**pie chart**). Pogledajmo sljedeći primjer.

Primjer 6.2. *Ribar Roko ulovio je* 288 *riba od kojih je* 71 *lokarda,* 24 *zubatca i* 193 *srdele. Nacrtajte strukturni krug za dane podatke.*

Rješenje: Odredimo prvo tablicu frekvencija:

vrsta ribe	f_i	f_{r_i}	%
lokarda	71	$71/288 = 0.2465$	24.65%
zubatac	24	$24/288 = 0.0833$	8.33%
srdela	193	193/288=0.6702	67.02%
Σ	288	1.00	100%

Sada, strukturni krug izgleda ovako:

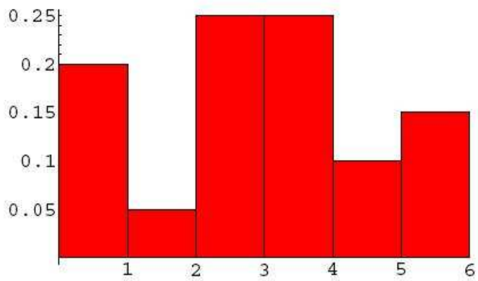


□

Slično kao i stupčasti dijagram, možemo nacrtati i **histogram**. Kod histograma, svaki stupac ima širinu 1 te se svaka dva stupca dodiruju. Visine stupaca jednake su odgovarajućim frekvencijama, odnosno relativnim frekvencijama. Ukoliko je visina svakog stupca jednaka odgovarajućim relativnim frekvencijama, površina svih stupaca je očito jednaka 1. Dakako, crtanje histograma nema smisla za nenumeričke vrijednosti.

Primjer 6.3. *Nacrtajte histogram za podatke iz Primjera* 6.1.

Rješenje:



□

U sljedećem primjeru promatrat ćemo neprekinuto statističko obilježje, odnosno obilježje koje može poprimiti vrijednosti iz nekog intervala realnih brojeva.

Primjer 6.4. *Na sistematskom pregledu studenata Visoke škole za primijenjeno računarstvo mjerena je visina (u metrima) 30 studenata. Dobiveni su sljedeći podatci:*

1.86	1.74	1.68	1.59	1.76	1.65	1.79	1.83	1.79	1.69
1.75	1.72	1.86	1.70	1.74	1.76	1.78	1.60	1.76	1.69
1.82	1.79	1.72	1.75	1.72	1.85	1.88	1.78	1.72	1.80

Nacrtajte histogram za dane podatke.

Rješenje: Kao što smo već napomenuli, mjereno statističko obilježje X je neprekinuto, odnosno poprima vrijednosti iz nekog intervala.

Dobivene podatke trebamo rasporediti u razrede. Prvo trebamo odrediti $x_{\max}$ i $x_{\min}$. Iz tablice vidimo da je $x_{\min} = 1.59$ i $x_{\max} = 1.88$. Sada trebamo odabrati adekvatan broj razreda, on je okvirno jednak vrijednosti $\sqrt{n}$. Kako je u našem slučaju $n = 30$, odabrat ćemo prvi veći cijeli broj, a to je $k = 6$.

Nadalje, određujemo zajedničku širinu razreda c. Imamo da je

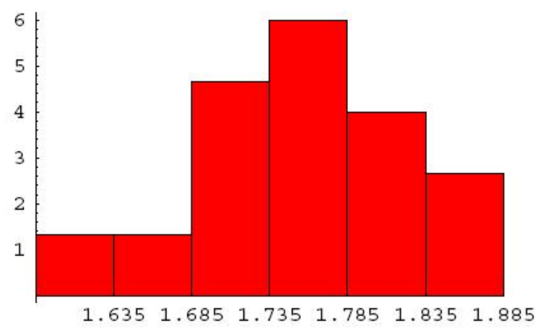
$$\frac{x_{\max} - x_{\min}}{k} = \frac{1.88 - 1.59}{6} = 0.0483.$$

Kako smo podatke mjerili na dvije decimale, tako ćemo i širinu razreda zaokružiti (uvijek na više) na dvije decimale. Prema tome, širina pojedinog razreda je $c = 0.05$.

Konačno, odredimo razrede $I_1, I_2, I_3, I_4, I_5, I_6$. Pri tome $I_1 \cup I_2 \cup I_3 \cup I_4 \cup I_5 \cup I_6$ treba obuhvaćati sve podatke, te je kraj jednog intervala jednak početku sljedećeg intervala. Prema tome imamo ovakvu situaciju:

razredi	f_i	$f_{r_i} = f_i/n$	f_{r_i}/c
$I_1 = [1.585, 1.635]$	2	0.067	1.34
$I_2 = [1.635, 1.685]$	2	0.067	1.34
$I_3 = [1.685, 1.735]$	7	0.233	4.66
$I_4 = [1.735, 1.785]$	9	0.3	6
$I_5 = [1.785, 1.835]$	6	0.2	4
$I_6 = [1.835, 1.885]$	4	0.133	2.66
Σ	30	1	20

Sada možemo nacrtati traženi histogram podataka. Sada širina stupca više nije proizvoljna, ona je jednaka širini razreda, tj. $c = 0.05$. Prema tome, da bi površina svih stupaca bila jednaka 1, na ordinati odabiremo vrijednosti $\frac{f_{r_i}}{c}$, a ne f_{r_i}, zato jer je $20 \cdot c = 20 \cdot 0.05 = 1$. Na kraju, histogram izgleda ovako:



□

6.2 Srednje vrijednosti uzorka

U ovoj točki upoznat ćemo neke pojmove kojima opisujemo srednje vrijednosti nekog uzorka. Preciznije, upoznat ćemo veličine kao što su **aritmetička sredina**, **medijan** i **mod**.

Aritmetička sredina

Neka je X numeričko statističko obilježje, te neka su $x_1, x_2, \ldots, x_n$ realizacije varijable X, odnosno statistički podatci. Aritmetička sredina podataka $x_1, x_2, \ldots, x_n$ je broj

$$\overline{x} = \frac{1}{n}\left(x_1 + x_2 + \ldots + x_n\right) = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Ako se u nizu $x_1, x_2, \ldots, x_n$ pojavljuju samo brojevi $a_1, a_2, \ldots, a_k$ s frekvencijama $f_1, f_2, \ldots, f_k$, onda je

$$\overline{x} = \frac{1}{n}\left(f_1 a_1 + f_2 a_2 + \ldots + f_k a_k\right) = \frac{1}{n}\sum_{j=1}^{k} f_j a_j.$$

Uočimo još da je $\{a_1, a_2, \ldots, a_k\} \subseteq \operatorname{Im} X$.

Primjer 6.5. *Kontrolor uzima uzorke od po* 30 *proizvoda i svaki put zapiše broj defektnih proizvoda u uzorku. Nakon* 20 *pregledanih takvih uzoraka dobiveni su sljedeći podatci:*

$$\begin{array}{cccccccccc} 0 & 0 & 1 & 1 & 0 & 0 & 3 & 1 & 0 & 2 \\ 0 & 1 & 0 & 0 & 4 & 0 & 0 & 3 & 2 & 0 \end{array}.$$

Odredite aritmetičku sredinu broja defektnih proizvoda s obzirom na zadane podatke.

Rješenje: Ovdje je $X =$"broj defektnih proizvoda u uzorku od 30 proizvoda". Stoga je $\operatorname{Im} X = \{0, 1, 2, 3, \ldots, 30\}$. Frekvencije pojavljivanja defektnih proizvoda u kontroliranim uzorcima prikazane su u sljedećoj tablici:

i	f_i
0	11
1	4
2	2
3	2
4	1
$5 - 30$	0
Σ	20

Prema tome, tražena aritmetička sredina jednaka je

$$\overline{x} = \frac{11 \cdot 0 + 4 \cdot 1 + 2 \cdot 2 + 2 \cdot 3 + 1 \cdot 4}{20} = \frac{18}{20} = 0.9.$$

$\square$

Medijan

Pojam medijana također promatramo samo za numeričke varijable. Neka je X numerička varijabla. Medijan je vrijednost od X za koju vrijedi da je 50% podataka manje od ili jednako toj vrijednosti i 50% podataka je veće od ili jednako njoj.

Kako bismo odredili medijan m niza $x_1, x_2, \ldots, x_n$, podatke je potrebno poredati po veličini:

$$x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}.$$

Sada, u ovisnosti o tome da li je broj podataka n paran ili neparan, slijedi da je medijan jednak

$$\begin{aligned} m &= x_{(k)}, & n = 2k-1, k \in \mathbf{N} \\ m &= \frac{1}{2}\left(x_{(k)} + x_{(k+1)}\right), & n = 2k, k \in \mathbf{N}. \end{aligned}$$

Primjer 6.6. *Mjeri se vrijeme izvođenja neke radne operacije. Podatci dobiveni u* 20 *nezavisnih mjerenja su (u sekundama):*

24	28	22	26	24	27	26	25	26	23
30	26	29	25	27	24	26	25	24	27

.

Odredite medijan za dana mjerenja.

Rješenje: Sortirajmo prvo podatke po veličini:

$$22, 23, 24, 24, 24, 24, 25, 25, 25, 26, 26, 26, 26, 26, 27, 27, 27, 28, 29, 30.$$

Sada, kako je $n = 20 = 2 \cdot 10$, slijedi da je

$$m = \frac{1}{2}\left(x_{(10)} + x_{(11)}\right) = \frac{1}{2}(26 + 26) = 26.$$

□

Formulu za medijan možemo napisati u sažetijem obliku, odnosno da ne razmatramo posebno parni i neparni slučaj. Pri tome ćemo iskoristiti formulu

$$x_{\left(\frac{p}{q}\right)} = x_{(k)} + \frac{r}{q}\left(x_{(k+1)} - x_{(k)}\right),$$

gdje su p i q cijeli brojevi i $p = k \cdot q + r$, $0 \leq r < q$, odnosno $\frac{p}{q} = k + \frac{r}{q}$. Drugim riječima, k je cjelobrojni količnik, a r je ostatak pri dijeljenju broja p s brojem q.

Uz takvu notaciju medijan m niza $x_1, x_2, \ldots, x_n$ jednak je

$$m = x_{\left(\frac{n+1}{2}\right)},$$

pa prethodni primjer možemo računati i ovako:

$$m = x_{\left(\frac{21}{2}\right)} = x_{\left(10+\frac{1}{2}\right)} = x_{(10)} + \frac{1}{2}\left(x_{(11)} - x_{(10)}\right) = 26 + \frac{1}{2}(26 - 26) = 26.$$

Navedenu formulu često ćemo koristiti u idućoj točki.

Mod

Mod je ona vrijednost statističkog obilježja koja se u uzorku javlja s najvećom frekvencijom. Koristan je kod statističkih obilježja koja nisu numerička, odnosno gdje ne možemo računati aritmetičku sredinu. Uzorak u kojemu postoje dvije vrijednosti s najvećom frekvencijom naziva se bimodalni uzorak, dok se uzorak sa samo jednim modom naziva unimodalni uzorak. Ako svi podatci imaju istu frekvenciju pojavljivanja u uzorku, tada uzorak nema mod.

Primjer 6.7. *Odredite mod za podatke iz Primjera* 6.1 *i Primjera* 6.6.

Rješenje: U Primjeru 6.1 imamo dvije vrijednosti: mod $= 3$ i mod $= 4$. Prema tome, to je bimodalan uzorak. Nadalje, u Primjeru 6.6 je mod $= 26$, tj. imamo unimodalan uzorak. □

6.3 Mjere raspršenja

U prošloj točki smo definirali veličine kojima opisujemo srednje vrijednosti nekog uzorka. Ovdje ćemo definirati veličine pomoću kojih opisujemo koliko je neki uzorak raspršen, odnosno disperziran. To su **raspon**, **interkvartil**, **varijanca** i **standardna devijacija**.

Raspon podataka

Neka je $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$ uređeni niz podataka. Broj

$$d = x_{(n)} - x_{(1)}$$

naziva se raspon uzorka.

Primjer 6.8. *Odredite raspon uzorka iz Primjera* 6.6.

Rješenje: Određivanjem najveće i najmanje vrijednosti dobivamo da je raspon uzorka jednak

$$d = 30 - 22 = 8.$$

□

Interkvartil

Kako bismo definirali interkvartil, potrebne su nam definicije gornjeg i donjeg kvartila.

Donji kvartil q_L je ona vrijednost uzorka za koju vrijedi da je 25% svih podataka manje ili jednako od nje i 75% svih podataka veće ili jednako od nje. Donji kvartil računamo po formuli

$$q_L = x_{\left(\frac{n+1}{4}\right)}.$$

Gornji kvartil q_U je ona vrijednost uzorka za koju vrijedi da je 75% svih podataka manje ili jednako od nje i 25% svih podataka veće ili jednako od nje. Gornji kvartil računamo po formuli

$$q_U = x_{\left(\frac{3(n+1)}{4}\right)}.$$

Interkvartil d_q jednak je razlici gornjeg i donjeg kvartila:

$$d_q = q_U - q_L.$$

Primjer 6.9. *Odredite interkvartil za podatke iz Primjera* 6.4.

Rješenje: Podatke prvo trebamo poredati po veličini:

$$1.59, 1.60, 1.65, 1.68, 1.69, 1.69, 1.70, 1.72, 1.72, 1.72, 1.72, 1.74, 1.74, 1.75, 1.75,$$
$$1.76, 1.76, 1.76, 1.78, 1.78, 1.79, 1.79, 1.79, 1.80, 1.82, 1.83, 1.85, 1.86, 1.86, 1.88.$$

Sada računamo donji i gornji kvartil. Imamo da je

$$\begin{aligned} q_L &= x_{\left(\frac{n+1}{4}\right)} = x_{\left(\frac{30+1}{4}\right)} = x_{\left(7+\frac{3}{4}\right)} \\ &= x_{(7)} + \frac{3}{4}\left(x_{(8)} - x_{(7)}\right) \\ &= 1.70 + \frac{3}{4}\left(1.72 - 1.70\right) = 1.715 \approx 1.72 \end{aligned}$$

i

$$\begin{aligned} q_U &= x_{\left(\frac{3(n+1)}{4}\right)} = x_{\left(\frac{93}{4}\right)} = x_{\left(23+\frac{1}{4}\right)} \\ &= x_{(23)} + \frac{1}{4}\left(x_{(24)} - x_{(23)}\right) \\ &= 1.79 + \frac{1}{4}(1.80 - 1.79) = 1.7925 \approx 1.79. \end{aligned}$$

Prema tome, interkvartil je jednak

$$d_q = q_U - q_L = 1.79 - 1.72 = 0.07.$$

□

Uređenu petorku $\left(x_{(1)}, q_L, m, q_U, x_{(n)}\right)$ nazivamo **karakteristična petorka uzorka**. Pomoću nje crtamo tzv. "box and whisker" dijagram, odnosno **dijagram pravokutnika**. Pri formiranju dijagrama pravokutnika "outlieri" su sve vrijednosti koje su od donjeg ili gornjeg kvartila udaljene za više od $\frac{3}{2}d_q$. "Brkovi" su najmanja i najveća vrijednost koje nisu outlieri. Outlieri se posebno naznačavaju na dijagramu pravokutnika.

Primjer 6.10. *Na nekom fakultetu je odabran uzorak od* 40 *studenata i izmjerene su im visine:*

140	188	175	176	177	168	162	181
183	187	187	162	184	161	180	169
195	171	170	199	181	169	189	191
172	182	183	178	180	165	185	205
183	187	188	182	163	179	178	188

(a) Odredite karakterističnu petorku uzorka.

(b) Nacrtajte dijagram pravokutnika.

Rješenje:
(a) Uredimo podatke:

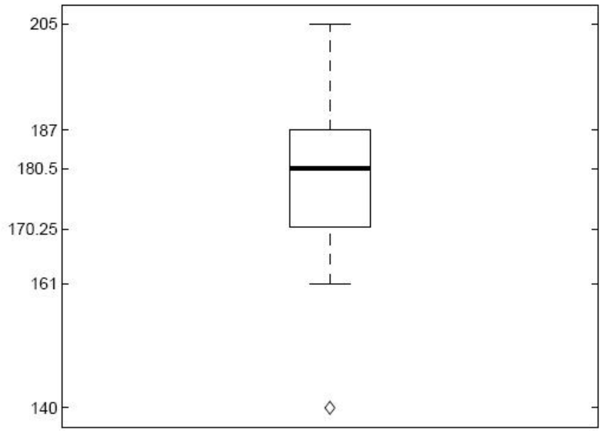
140	161	162	162	163	165	168	169
169	170	171	172	175	176	177	178
178	179	180	180	181	181	182	182
183	183	183	184	185	187	187	187
188	188	188	189	191	195	199	205

Tada je
$n = 40$,
$x_{(1)} = 140$,
$q_L = x_{\left(\frac{40+1}{4}\right)} = x_{\left(10+\frac{1}{4}\right)} = x_{(10)} + \frac{1}{4}\left(x_{(11)} - x_{(10)}\right) = 170 + \frac{1}{4}(171 - 170) = 170.25$,
$m = x_{\left(\frac{40+1}{2}\right)} = x_{\left(20+\frac{1}{2}\right)} = x_{(20)} + \frac{1}{2}\left(x_{(21)} - x_{(20)}\right) = 180 + \frac{1}{2}(181 - 180) = 180.5$,
$q_U = x_{\left(\frac{3(40+1)}{4}\right)} = x_{\left(30+\frac{3}{4}\right)} = x_{(30)} + \frac{3}{4}\left(x_{(31)} - x_{(30)}\right) = 187 + \frac{3}{4}(187 - 187) = 187$,
$x_{(40)} = 205$.
Prema tome, karakteristična petorka je $(140, 170.25, 180.5, 187, 205)$.

(b) Interkvartil iznosi $d_q = q_U - q_L = 187 - 170.25 = 16.75$ pa je

$$q_L - \frac{3}{2} d_q = 145.125 \quad \text{i} \quad q_U + \frac{3}{2} d_q = 212.125.$$

Sada lagano crtamo dijagram pravokutnika.



□

Napomenimo još da medijan, te gornji i donji kvartil spadaju u mjere lokacije. Tu još spadaju decili, percentili i općenito kvantili, ali ovdje nećemo uvoditi te veličine.

Uzoračka disperzija i standardno odstupanje

Uzoračka disperzija (varijanca) s^2 niza $x_1, x_2, \ldots, x_n$ dana je formulom

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2 ,$$

pri čemu je $\overline{x}$ aritmetička sredina uzorka $x_1, x_2, \ldots, x_n$.

Uzoračko standardno odstupanje s (devijacija) jednako je drugom korijenu uzoračke disperzije, odnosno

$$s = +\sqrt{s^2}.$$

Formula za uzoračku disperziju je nepraktična za računanje pa ćemo je prikazati u pogodnijem obliku. Naime, vrijedi

$$\begin{aligned} s^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i^2-2x_i\overline{x}+\overline{x}^2\right) \\ &= \frac{1}{n-1}\left(\sum_{i=1}^{n}x_i^2-2\overline{x}\sum_{i=1}^{n}x_i+\sum_{i=1}^{n}\overline{x}^2\right) \\ &= \frac{1}{n-1}\left(\sum_{i=1}^{n}x_i^2-2n\overline{x}^2+n\overline{x}^2\right) \\ &= \frac{1}{n-1}\left(\sum_{i=1}^{n}x_i^2-n\overline{x}^2\right). \end{aligned}$$

Ovaj oblik formule je puno praktičniji za računanje.

Nadalje, ako se u uzorku $x_1, x_2, \ldots, x_n$ vrijednosti $a_1, a_2, \ldots, a_k$ pojavljuju s frekvencijama $f_1, f_2, \ldots, f_k$, onda vrijedi

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{k}(a_i-\overline{x})^2\cdot f_i = \frac{1}{n-1}\left(\sum_{i=1}^{k}f_i\cdot a_i^2-n\overline{x}^2\right).$$

Primjer 6.11. *Izračunajte disperziju i standardno odstupanje za podatke iz Primjera* 6.6.

Rješenje: Napravimo prvo frekvencijsku tablicu u koju ćemo unijeti vrijednosti $f_i a_i$ i $f_i a_i^2$.

a_i	f_i	$f_i a_i$	$f_i a_i^2$
22	1	22	484
23	1	23	529
24	4	96	2304
25	3	75	1875
26	5	130	3380
27	3	81	2187
28	1	28	784
29	1	29	841
30	1	30	900
Σ	20	514	13284

Sada je

$$\overline{x} = \frac{514}{20} = 25.7,$$

$$s^2 = \frac{1}{19}\left(13284 - 20\cdot 25.7^2\right) = 3.91$$

i

$$s = \sqrt{3.91} = 1.98.$$

□

6.4 Mjere oblika

Slično kao što se definira uzoračka disperzija, može se definirati i uzorački centralni moment reda k, $k \in \mathbf{N}$:

$$\mu_k = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^k .$$

Uočimo da je centralni moment reda 1 uvijek jednak nuli. Naime, vrijedi

$$\mu_1 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x}) = \frac{1}{n-1} \left(\sum_{i=1}^{n} x_i - n\overline{x} \right) = \frac{1}{n-1} (n\overline{x} - n\overline{x}) = 0.$$

Centralni moment reda 2 je upravo uzoračka disperzija. Značajan je i centralni moment reda 3, odnosno

$$\mu_3 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^3 .$$

U tu svrhu pogledajmo sljedeći primjer.

Primjer 6.12. *Zadan je uzorak* $0.2, 1.5, 2.5, 4.5, 5.5, 6.8$. *Odredite centralni moment reda* 3 *za taj uzorak.*

Rješenje: Aritmetička sredina zadanog uzorka jednaka je

$$\overline{x} = \frac{0.2 + 1.5 + 2.5 + 4.5 + 5.5 + 6.8}{6} = \frac{21}{6} = 3.5.$$

Sada je centralni moment reda 3 jednak

$$\begin{aligned} \mu_3 &= \frac{1}{5} \left((0.2 - 3.5)^3 + (1.5 - 3.5)^3 + (2.5 - 3.5)^3 + (4.5 - 3.5)^3 + (5.5 - 3.5)^3 + (6.8 - 3.5)^3 \right) \\ &= \frac{1}{5} (-35.937 - 8 - 1 + 1 + 8 + 35.937) = 0. \end{aligned}$$

□

Iz prethodnog primjera zaključujemo kako je kod uzorka simetričnog s obzirom na njegovu aritmetičku sredinu, centralni moment reda 3 uvijek jednak nuli. Stoga se pomoću te veličine definira takozvani **koeficijent asimetrije** uzorka.

Koeficijent asimetrije uzorka definiran je formulom

$$\alpha_3 = \frac{\mu_3}{s^3} = \frac{1}{n-1} \sum_{i=1}^{n} \left(\frac{x_i - \overline{x}}{s} \right)^3 .$$

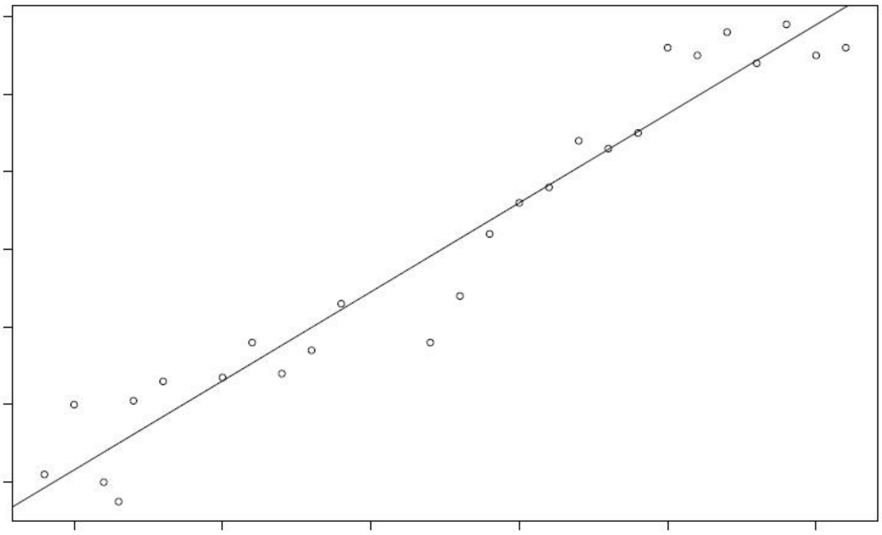
Dakako, ako se u uzorku $x_1, x_2, \ldots, x_n$ vrijednosti $a_1, a_2, \ldots, a_k$ pojavljuju s frekvencijama $f_1, f_2, \ldots, f_k$, onda je

$$\alpha_3 = \frac{1}{n-1} \sum_{i=1}^{k} f_i \cdot \left(\frac{a_i - \overline{x}}{s} \right)^3$$

Ako je $\alpha_3 = 0$ kažemo da je uzorak simetričan. Nadalje, ako je $\alpha_3 > 0$ kažemo da je uzorak pozitivno asimetričan, dok je za $\alpha_3 < 0$ uzorak negativno asimetričan.

6.5 Linearna regresija

Pretpostavimo da imamo n parova podataka (x_i, y_i), $i = 1, 2, \ldots, n$. Želimo odrediti vezu između nezavisne varijable x i zavisne varijable y. Nadalje, pretpostavimo da je ta veza linearna, odnosno da je graf pripadajuće funkcije pravac $y = \alpha x + \beta$. Drugim riječima, mi tražimo pravac koji najbolje aproksimira dane parove točaka, kao na slici.



Pomoću takozvane metode najmanjih kvadrata dobiva se da je procjenitelj pravac

$$y = \hat{\alpha} x + \hat{\beta},$$

gdje je

$$\hat{\alpha} = \frac{s_{xy}}{s_x^2}, \quad \hat{\beta} = \overline{y} - \hat{\alpha}\overline{x},$$

$$s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2\right),$$

$$s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \overline{y})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} y_i^2 - n\overline{y}^2\right),$$

i

$$s_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i y_i - n\overline{x}\,\overline{y}\right).$$

Primjer 6.13. *Šestero studenata upitano je koliko su se sati pripremali za završni ispit. Njihovi odgovori na to pitanje uspoređeni su s bodovima na ispitu (maksimalni broj bodova je* 30*). Dobiveni su ovi rezultati:*

sati učenja	0.50	1.25	1.50	2.00	2.25	3.50
bodovi	19	23	25	26	28	30

(a) Procijenite pravac linearne regresije za ove podatke.

(b) Procijenite koliko student ima bodova na ispitu ako se za njega pripremao jedan sat.

(c) Procijenite koliko se za ispit pripremao student koji ima 24 boda na ispitu.

Rješenje: (a) Ovdje je $n = 6$, pa za pripadne aritmetičke sredine $\overline{x}$ i $\overline{y}$ imamo da je

$$\overline{x} = \frac{1}{6}(0.50 + 1.25 + 1.50 + 2.00 + 2.25 + 3.50) = 1.833$$

i

$$\overline{y} = \frac{1}{6}(19 + 23 + 25 + 26 + 28 + 30) = 25.167.$$

Nadalje,

$$\sum_{i=1}^{6} x_i^2 = 0.50^2 + 1.25^2 + 1.50^2 + 2.00^2 + 2.25^2 + 3.50^2 = 25.375,$$

pa je

$$s_x^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2\right) = \frac{1}{5}\left(25.375 - 6 \cdot 1.833^2\right) = 1.043.$$

Dalje imamo

$$\begin{aligned}\sum_{i=1}^{6} x_i y_i &= 0.50 \cdot 19 + 1.25 \cdot 23 + 1.50 \cdot 25 + 2.00 \cdot 26 + 2.25 \cdot 28 + 3.50 \cdot 30 \\ &= 295.750,\end{aligned}$$

odakle je

$$s_{xy} = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i y_i - n\overline{x}\,\overline{y}\right) = \frac{1}{5}(295.750 - 6 \cdot 1.833 \cdot 25.167) = 3.793.$$

Konačno, koeficijenti $\hat{\alpha}$ i $\hat{\beta}$ su jednaki

$$\hat{\alpha} = \frac{s_{xy}}{s_x^2} = \frac{3.793}{1.043} = 3.637$$

i

$$\hat{\beta} = \overline{y} - \hat{\alpha}\overline{x} = 25.167 - 3.637 \cdot 1.833 = 18.5,$$

pa je traženi pravac linearne regresije

$$y = \hat{\alpha}x + \hat{\beta} = 3.637x + 18.5.$$

(b) Student koji se pripremao 1 sat, po ovoj metodi ima približno

$$y(1) = 3.637 \cdot 1 + 18.5 = 22.137 \approx 22$$

boda na ispitu.
(c) Ovdje trebamo naći x takav da je

$$3.637x + 18.5 = 24.$$

Rješenje te jednadžbe je $x \approx 1.50$, pa se student za ispit pripremao oko sat i pol. $\square$

Usko u vezi s linearnom regresijom je **Pearsonov koeficijent korelacije**. Pearsonov koeficijent korelacije r varijabli x_i i y_i, $i = 1, 2, \ldots, n$, definiran je formulom

$$r = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_i - \overline{x}}{s_x}\cdot\frac{y_i-\overline{y}}{s_y}\right) = \frac{s_{xy}}{s_x\cdot s_y}.$$

Može se pokazati da Pearsonov koeficijent korelacije uvijek poprima vrijednosti unutar intervala $[-1, 1]$, odnosno uvijek je $-1 \leq r \leq 1$.

Ako je $r = 0$ onda nema korelacije između varijabli x_i i y_i, $i = 1, 2, \ldots, n$. Ako je $r > 0$ onda je korelacija pozitivna. To znači da ako varijabla x raste, onda u pravilu raste i varijabla y. S druge strane, ako je $r < 0$ korelacija je negativna pa ako x raste, onda u pravilu varijabla y pada.

Primjer 6.14. *Odredite Pearsonov koeficijent korelacije za podatke iz Primjera* 6.13.

Rješenje: U prethodnom primjeru smo dobili da je $s_{xy} = 3.793$ i $s_x^2 = 1.043$, odakle je $s_x = 1.021$. Trebamo još izračunati s_y. Slično kao i u prethodnom primjeru dobivamo

$$\sum_{i=1}^{6} y_i^2 = 19^2 + 23^2 + 25^2 + 26^2 + 28^2 + 30^2 = 3875,$$

odakle je

$$s_y^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} y_i^2 - n\overline{y}^2\right) = \frac{1}{5}\left(3875 - 6\cdot 25.167^2\right) = 14.947,$$

odnosno $s_y = 3.866$. Konačno, uvrštavanjem dobivenih podataka dobivamo da je

$$r = \frac{s_{xy}}{s_x\cdot s_y} = \frac{3.793}{1.021\cdot 3.866} = 0.96,$$

što je vrlo visoka korelacija. □

6.6 Zadatci za ponavljanje

Zadatak 6.1. Tijekom 20 dana zabilježeni su sljedeći podatci o broju prometnih nesreća na širem području nekog grada:

$$\begin{array}{cccccccccc} 3 & 0 & 3 & 5 & 0 & 4 & 0 & 6 & 1 & 3 \\ 3 & 4 & 0 & 3 & 1 & 4 & 5 & 0 & 10 & 1 \end{array}.$$

(a) Prikažite zadane podatke pomoću frekvencijske tablice.

(b) Odredite aritmetičku sredinu uzorka.

(c) Odredite disperziju i standardno odstupanje uzorka.

(d) Odredite mod uzorka.

(e) Odredite medijan uzorka.

(f) Odredite raspon uzorka.

(g) Odredite donji i gornji kvartil, te interkvartil uzorka.

(h) Odredite koeficijent asimetrije uzorka.

Rješenje:

(a) Tablica frekvencija izgleda ovako:

a_i	f_i	f_{r_i}
0	5	$5/20 = 0.25$
1	3	$3/20 = 0.15$
3	5	$5/20 = 0.25$
4	3	$3/20 = 0.15$
5	2	$2/20 = 0.1$
6	1	$1/20 = 0.05$
10	1	$1/20 = 0.05$
Σ	20	1.00

(b) Kako je $n = 20$, aritmetička sredina iznosi

$$\overline{x} = \frac{5\cdot 0 + 3\cdot 1 + 5\cdot 3 + 3\cdot 4 + 2\cdot 5 + 1\cdot 6 + 1\cdot 10}{20} = \frac{56}{20} = 2.8.$$

(c) Izračunajmo prvo disperziju:

$$s^2 = \frac{1}{19}\left(5\cdot 0^2 + 3\cdot 1^2 + 5\cdot 3^2 + 3\cdot 4^2 + 2\cdot 5^2 + 1\cdot 6^2 + 1\cdot 10^2 - 20\cdot 2.8^2\right) = 6.59.$$

Zbog toga je standardno odstupanje jednako $s = \sqrt{s^2} = \sqrt{6.59} = 2.567$.

(d) U ovom uzorku imamo dvije vrijednosti s najvećom frekvencijom, tj. mod $= 0$ i mod $= 3$.

(e) Da bismo odredili medijan, prvo ćemo podatke poredati po veličini:

$$0, 0, 0, 0, 0, 1, 1, 1, 3, 3, 3, 3, 3, 4, 4, 4, 5, 5, 6, 10.$$

Sada je medijan jednak

$$m = x_{\left(\frac{n+1}{2}\right)} = x_{\left(\frac{21}{2}\right)} = x_{\left(10+\frac{1}{2}\right)} = x_{(10)} + \frac{1}{2}\left(x_{(11)} - x_{(10)}\right) = 3 + \frac{1}{2}(3-3) = 3.$$

(f) Raspon uzorka je $d = x_{(20)} - x_{(1)} = 10 - 0 = 10$.

(g) Donji i gornji kvartil su redom jednaki

$$q_L = x_{\left(\frac{n+1}{4}\right)} = x_{\left(\frac{20+1}{4}\right)} = x_{\left(5+\frac{1}{4}\right)} = x_{(5)} + \frac{1}{4}\left(x_{(6)} - x_{(5)}\right) = 0 + \frac{1}{4}(1-0) = 0.25,$$

$$q_U = x_{\left(\frac{3(n+1)}{4}\right)} = x_{\left(\frac{3(20+1)}{4}\right)} = x_{\left(15+\frac{3}{4}\right)} = x_{(15)} + \frac{3}{4}\left(x_{(16)} - x_{(15)}\right) = 4 + \frac{3}{4}(4-4) = 4,$$

pa je interkvartil jednak

$$d_q = q_U - q_L = 4 - 0.25 = 3.75.$$

(h) Formulu za koeficijent asimetrije možemo zapisati u obliku

$$\alpha_3 = \frac{1}{(n-1)s^3}\sum_{i=1}^{n} f_i(a_i - \overline{x})^3,$$

pa je

$$\alpha_3 = \frac{1}{19\cdot 2.567^3}\Big[5\cdot(0-2.8)^3 + 3\cdot(1-2.8)^3 + 5\cdot(3-2.8)^3 + 3\cdot(4-2.8)^3 + 2\cdot(5-2.8)^3 + 1\cdot(6-2.8)^3 + 1\cdot(10-2.8)^3\Big] = 0.95.$$

□

Zadatak 6.2. Profesorica povijesti je tijekom školskog sata ispitala 6 učenika. Nakon što ih je ocijenila, upitala je svakog od učenika koliko je sati jučer proveo pred računalom igrajući igrice. U priloženoj tablici x_i označava dobivenu ocjenu, a y_i broj sati provedenih pred računalom:

x_i	5	5	4	4	3	2
y_i	3	1	2	3	4	5

(a) Odredite pravac linearne regresije za dane podatke.

(b) Odredite Pearsonov koeficijent korelacije za dane podatke. Da li je korelacija između varijabli pozitivna ili negativna?

Rješenje: (a) Imamo 6 parova podataka, pa za pripadne aritmetičke sredine $\overline{x}$ i $\overline{y}$ imamo da je

$$\overline{x} = \frac{1}{6}(5+5+4+4+3+2) = 3.833$$

i

$$\overline{y} = \frac{1}{6}(3+1+2+3+4+5) = 3.$$

Nadalje,

$$\sum_{i=1}^{6} x_i^2 = 5^2+5^2+4^2+4^2+3^2+2^2 = 95,$$

pa je

$$s_x^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2\right) = \frac{1}{5}\left(95 - 6\cdot 3.833^2\right) = 1.37.$$

Dalje imamo

$$\sum_{i=1}^{6} x_i y_i = 5\cdot 3+5\cdot 1+4\cdot 2+4\cdot 3+3\cdot 4+2\cdot 5 = 62,$$

odakle je

$$s_{xy} = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i y_i - n\overline{x}\,\overline{y}\right) = \frac{1}{5}(62 - 6\cdot 3.833\cdot 3) = -1.399.$$

Konačno, koeficijenti $\hat{\alpha}$ i $\hat{\beta}$ su jednaki

$$\hat{\alpha} = \frac{s_{xy}}{s_x^2} = -\frac{1.399}{1.37} = -1.021$$

i

$$\hat{\beta} = \overline{y} - \hat{\alpha}\overline{x} = 3 + 1.021\cdot 3.833 = 6.913,$$

pa je traženi pravac linearne regresije

$$y = \hat{\alpha}x + \hat{\beta} = -1.021x + 6.913.$$

(b) U prvom dijelu zadatka dobili smo da je $s_{xy} = -1.399$ i $s_x^2 = 1.37$, odakle je $s_x = 1.17$. Trebamo još izračunati s_y. Slično kao i u prethodnom primjeru imamo

$$\sum_{i=1}^{6} y_i^2 = 3^2+1^2+2^2+3^2+4^2+5^2 = 64,$$

odakle je

$$s_y^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} y_i^2 - n\overline{y}^2\right) = \frac{1}{5}\left(64 - 6 \cdot 3^2\right) = 2,$$

odnosno $s_y = 1.414$. Konačno, uvrštavanjem dobivenih podataka dobivamo da je

$$r = \frac{s_{xy}}{s_x \cdot s_y} = -\frac{1.399}{1.17 \cdot 1.414} = -0.846,$$

pa su varijable negativno korelirane. $\square$