OSNOVE TEORIJE VERJETNOSTI

Zapiski s predavanj

Martin Raič

Kazalo

1.

Verjetnostni prostori

1.1 Osnovni pojmi verjetnosti

Pojem *verjetnost* (angl. *probability*) v praksi navadno interpretiramo kot predvideni (ocenjeni, pričakovani) *delež* poskusov, pri katerih se zgodi določen dogodek.

Primer: stavek:

'Verjetnost, da boste pri statistiki dobili oceno 10, je 9%'

si navadno razlagamo kot:

'Ocenjujemo, da bo 9% takih, kot ste vi, pri statistiki dobilo oceno 10.'

Take ocene dostikrat temeljijo na podatkih iz preteklosti, včasih pa tudi na simetriji ali fizikalnih dejstvih.

Primer: če je kovanec simetričen, domnevamo, da je *pošten*, kar pomeni, da cifra in grb padeta z enako verjetnostjo, t. j. 1/2. Niso pa vsi kovanci vedno pošteni, še zlasti ne, če kovanec zavrtimo in čakamo, da se prevrne. Bolje je, če kovanec vržemo v zrak, tako da se velikokrat zavrti.

Zapišimo zdaj to malo splošneje. *Poskus* je realizacija natančno določenih pogojev, pri katerih opazujemo enega ali več *dogodkov*. Po realizaciji poskusa mora biti za vsak dogodek jasno, ali se je *zgodil* ali ne. To ponazorimo z diagramom:

$$\text{poskus} \dashrightarrow \text{dogodek}$$

pri čemer pikčasta puščica pomeni, da je to, na kar kaže, znano šele po realizaciji poskusa.

Pri prejšnjih dveh primerih je bil poskus opravljen izpit (v tiste, ki odnehajo, se tu ne bomo spuščali) in met kovanca. Pri kovancu je smiselno obravnavati dogodka 'pade grb' in 'pade cifra', pri izpitu pa je dogodkov, ki jih je smiselno obravnavati, več, npr. 'dobimo 10', 'dobimo 6', 'dobimo več kot 7' itd.

Ali se je dani dogodek zgodil ali ne, izvemo šele po realizaciji poskusa. Navadno pa želimo kaj izračunati, še preden izvedemo poskuse. Zato je smiselno vzpostaviti pojem, na podlagi katerega lahko *a priori* (že pred izvedbo poskusa) opredelimo, ali se je posamezen dogodek zgodil ali ne. To vlogo igra pojem *izida*: če poznamo izid, moramo za vsak dogodek vedeti, ali se je zgodil ali ne. Po realizaciji poskusa pa moramo vedeti tudi, kateri izid se je zgodil (kako se je izšel poskus). To prikažemo z diagramom:

$$\text{poskus} \dashrightarrow \text{izid} \longrightarrow \text{dogodek}$$

pri čemer pikčasta puščica spet pomeni, da je to, na kar kaže, znano šele po realizaciji poskusa, neprekinjena puščica pa, da je znano *a priori*, že vnaprej.

Teorija verjetnosti se gradi od izidov naprej: vsi možni izidi sestavljajo množico, dogodki pa so njene podmnožice. Množico izidov bomo označevali z Ω.

Primer: pri metu kovanca je množica izidov $\Omega = \{c, g\}$, možni dogodki pa so štirje: $\emptyset = \{\}$, $\{c\}$, $\{g\}$ in $\{c, g\} = \Omega$. Prvi dogodek je *nemogoč*, drugi pa *gotov.*

Primer: pri metu standardne kocke je smiselno postaviti:

$$\Omega = \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\}\,.$$

Možnih dogodkov je $2^6 = 64$ (za vsak izid imamo dve možnosti – ali je v izbranem dogodku ali pa ga ni notri). Oglejmo si naslednjih šest:

$$\begin{aligned}
A &= \{\text{pade natanko šest pik }\} = \{⚅\}\,,\\
B &= \{\text{pade najmanj pet pik}\} = \{⚄, ⚅\}\,,\\
C &= \{\text{pade liho število pik}\} = \{⚀, ⚂, ⚄\}\,,\\
D &= \{\text{pade manj kot šest pik}\} = \{⚀, ⚁, ⚂, ⚃, ⚄\}\,,\\
E &= \{\text{pade največ šest pik}\} = \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\} = \Omega\\
D &= \{\text{pade več kot šest pik}\} = \{\} = \emptyset\,.
\end{aligned} \tag{1.1.1}$$

Dogodek F je *nemogoč*, dogodek E pa je *gotov.*

Pripomnimo naj, da za dogodke ne moremo vedno proglasiti vseh podmnožic prostora izidov, temveč le določeno družino podmnožic, ki mora imeti posebne lastnosti. Več o tem v nadaljevanju.

V teoriji je *verjetnost* preslikava, ki vsakemu dogodku A priredi število $\mathbb{P}(A)$ iz intervala $[0, 1]$, njegovo verjetnost. Omenili smo, da verjetnost navadno interpretiramo kot oceno za delež poskusov. V teoriji zahtevamo, da ima verjetnost določene lastnosti, ki jih ima tudi delež. To so tako imenovani *aksiomi Kolmogorova*, ki jih bomo predstavili v nadaljevanju.

Primer: pri poštenem kovancu so verjetnosti vseh možnih dogodkov enake:

$$\mathbb{P}(\emptyset) = 0\,, \quad \mathbb{P}(\{c\}) = \frac{1}{2}\,, \quad \mathbb{P}(\{g\}) = \frac{1}{2}\,, \quad \mathbb{P}(\{c, g\}) = 1\,.$$

Primer verjetnosti pri nepoštenem (pristranskem) kovancu:

$$\mathbb{P}(\emptyset) = 0\,, \quad \mathbb{P}(\{c\}) = 0{\cdot}7\,, \quad \mathbb{P}(\{g\}) = 0{\cdot}3\,, \quad \mathbb{P}(\{c, g\}) = 1\,.$$

1.2 Diskretna verjetnost

Omenili smo že, da je v teoriji verjetnost nekaj, kar se obnaša podobno kot delež. No, če verjetnost izida (dogodka) kar *definiramo* kot delež poskusov, ki se izidejo na predpisani način (pri katerih se zgodi dani dogodek), velja:

- Verjetnost vsakega izida je število iz intervala $[0, 1]$.
- Verjetnost vsakega dogodka je vsota verjetnosti izidov, ki jih dani dogodek zajema.
- Vsota verjetnosti vseh izidov je enaka 1.

Če torej v tem primeru poznamo verjetnosti posameznih izidov, poznamo tudi verjetnosti dogodkov. To nam da navdih za naslednjo splošnejšo, bolj teoretično definicijo verjetnosti.

Verjetnostna funkcija (angl. *probability mass function, pmf*) na danem prostoru izidov Ω je predpis, ki vsakemu izidu $\omega \in \Omega$ priredi število iz intervala $[0, 1]$, ki ga bomo navadno označevali s $p(\omega)$ in mu rekli *verjetnost izida.* Vsota verjetnosti vseh izidov mora biti 1:

$$\sum_{\omega \in \Omega} p(\omega) = 1\,.$$

Če je $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, kjer so vsi izidi $\omega_1, \ldots, \omega_n$ različni, zgornji zapis pomeni:

$$p(\omega_1) + p(\omega_2) + \cdots + p(\omega_n) = 1\,.$$

Primer: pri poštenem kovancu je $p(c) = p(g) = \frac{1}{2}$, pri prejšnjem primeru pristranskega kovanca pa je $p(c) = 0{\cdot}7$ in $p(g) = 0{\cdot}3$.

Verjetnostno funkcijo lahko zapišemo z *verjetnostno shemo*, ki je zapis oblike:

$$\begin{pmatrix} \omega_1 & \omega_2 & \cdots & \omega_n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix}\,,$$

kjer so $\omega_1, \omega_2, \ldots, \omega_n$ vsi možni izidi, t. j. $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$. Verjetnostna shema je v *kanonični obliki*, če so vse vrednosti $\omega_1, \omega_2, \ldots, \omega_n$ različne. Tako zapisana verjetnostna shema definira verjetnostno funkcijo po predpisu:

$$p(\omega_1) = p_1, \ p(\omega_2) = p_2, \ \ldots, \ p(\omega_n) = p_n \, .$$

Verjetnostna shema poštenega kovanca: $\begin{pmatrix} c & g \\ 1/2 & 1/2 \end{pmatrix}$.

Primer verjetnostne sheme pristranskega kovanca: $\begin{pmatrix} c & g \\ 0\text{˙}7 & 0\text{˙}3 \end{pmatrix}$.

Primer: $\begin{pmatrix} c & g \\ 0\text{˙}5 & 0\text{˙}3 \end{pmatrix}$, $\begin{pmatrix} c & g \\ 0\text{˙}8 & 0\text{˙}3 \end{pmatrix}$ in $\begin{pmatrix} c & g \\ 1\text{˙}2 & -0\text{˙}2 \end{pmatrix}$ niso verjetnostne sheme: pri prvi je vsota verjetnosti manjša od 1, pri drugi večja od 1, pri tretji pa je sicer enaka 1, a verjetnosti niso iz intervala $[0, 1]$.

Včasih ima smisel, da se vrednosti v verjetnostni shemi ponavljajo. V tem primeru vrednost verjetnostne funkcije dobimo s seštevanjem:

$$p(\omega) = \sum_{i;\omega_i = \omega} p_i \, .$$

in tako lahko vsako verjetnostno shemo zapišemo v kanonični obliki.

Primer: verjetnostna shema:

$$\begin{pmatrix} a & b & a & c & b \\ 0\text{˙}1 & 0\text{˙}2 & 0\text{˙}3 & 0\text{˙}1 & 0\text{˙}3 \end{pmatrix}$$

je ekvivalentna verjetnostni shemi v kanonični obliki:

$$\begin{pmatrix} a & b & c \\ 0\text{˙}4 & 0\text{˙}5 & 0\text{˙}1 \end{pmatrix} .$$

Diskretni verjetnostni prostor je množica izidov Ω skupaj z verjetnostno funkcijo. Verjetnost dogodka A z verjetnostno funkcijo definiramo kot:

$$\mathbb{P}(A) := \sum_{\omega \in A} p(\omega) \, .$$

Za *enostavne* dogodke (t. j. tiste, ki vsebujejo en sam izid) velja:

$$\mathbb{P}(\{\omega\}) = p(\omega) \, .$$

Če je verjetnostna funkcija podana z verjetnostno shemo (v kanonični ali nekanonični obliki), velja:

$$\mathbb{P}(A) = \sum_{i;\omega_i \in A} p_i \, .$$

Torej za vsako verjetnostno shemo velja $p_1+p_2+\cdots+p_n=1$. Velja tudi, da vsaka shema zgornje oblike, v kateri je $0 \le p_i \le 1$ za vse i in še $p_1+p_2+\cdots+p_n=1$, predstavlja verjetnostno shemo neke verjetnostne funkcije.

Primer: poštena kocka ima verjetnostno shemo:

$$\begin{pmatrix} \text{⚀} & \text{⚁} & \text{⚂} & \text{⚃} & \text{⚄} & \text{⚅} \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}.$$

in verjetnosti dogodkov iz (1.1.1) so enake:

$$\mathbb{P}(A)=\frac{1}{6}, \quad \mathbb{P}(B)=\frac{2}{6}=\frac{1}{3}, \quad \mathbb{P}(C)=\frac{3}{6}=\frac{1}{2}, \quad \mathbb{P}(D)=\frac{5}{6}, \quad \mathbb{P}(E)=1, \quad \mathbb{P}(F)=0.$$

Pri pošteni kocki so vsi izidi enako verjetni. V splošnem, če verjetnostni prostor vsebuje n enako verjetnih izidov, je $p(\omega)=1/n$ za vse ω. Verjetnosti dogodkov se potem izražajo kot razmerja:

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

Pri pošteni kocki lahko rečemo tudi, da je število pik izbrano *na slepo* iz množice $\{1,2,3,4,5,6\}$. Splošno, *slepa* izbira (angl. *by chance, uniformly at random*) pomeni, da so vse možnosti, ki jih lahko izberemo, enako verjetne.

Pri slepi izbiri je torej verjetnost dogodka kar delež izidov, ki so zajeti v tem dogodku. To spomninja na motivacijo za verjetnost iz podrazdelka 1.1, kjer je verjetnost ocena za delež *poskusov*, pri katerih se zgodi dani dogodek. No, če verjetnost dejansko definiramo kar kot delež poskusov, to pomeni, da smo privzeli, da je poskus izbran na slepo.

Seveda pa ni nujno, da so vsi izidi enako verjetni.

Primer: pri nepošteni kocki z verjetnostno shemo:

$$\begin{pmatrix} \text{⚀} & \text{⚁} & \text{⚂} & \text{⚃} & \text{⚄} & \text{⚅} \\ 0{\cdot}05 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}35 \end{pmatrix} \tag{1.2.1}$$

so verjetnosti prej omenjenih dogodkov enake:

$$\begin{gathered} \mathbb{P}(A)=p(\text{⚅})=0{\cdot}35, \qquad \mathbb{P}(B)=p(\text{⚄})+p(\text{⚅})=0{\cdot}5, \\ \mathbb{P}(C)=p(\text{⚀})+p(\text{⚂})+p(\text{⚄})=0{\cdot}35, \\ \mathbb{P}(D)=p(\text{⚀})+p(\text{⚁})+p(\text{⚂})+p(\text{⚃})+p(\text{⚄})=0{\cdot}65, \\ \mathbb{P}(E)=1, \qquad \mathbb{P}(F)=0. \end{gathered}$$

Posvetimo se zdaj še malo slepi izbiri. Pomembna primera sta slepo *vlečenje* z vračanjem in brez vračanja. Slepo vlečenje r elementov z vračanjem iz dane množice pomeni, da so vse urejene r-terice $(a_1,a_2,\ldots,a_r)$ izvlečenih elementov enako verjetne (ločimo torej različne vrstne rede vlečenja). Slepo vlečenje brez vračanja pa pomeni, da so enako verjetne vse tiste r-terice, ki imajo vse elemente različne.

Primer: Iz množice kart $\{\text{as}, \text{kralj}, \text{dama}, \text{fant}\}$ na slepo izvlečemo dve karti. Kolikšna je verjetnost, da bo med njima vsaj ena dama?

Odgovor je odvisen od tega, ali vlečemo z vračanjem ali brez. Če vlečemo z vračanjem, verjetnostni prostor sestavlja naslednjih 16 enako verjetnih izidov:

AA	AK	AQ	AJ
KA	KK	KQ	KJ
QA	QK	QQ	QJ
JA	JK	JQ	JJ

kjer prva črka pomeni prvo, druga pa drugi izvlečeno karto, in kjer se držimo standardnih kratic A za asa, K za kralja, Q za damo in J za fanta. Vsaj eno damo izvlečemo pri natanko pri natanko 7 izidih. Verjetnost tega dogodka je torej $7/16 = 0^{\cdot}4375$.

Če pa vlečemo brez vračanja, verjetnostni prostor sestavlja le naslednjih 12 enako verjetnih izidov:

	AK	AQ	AJ
KA		KQ	KJ
QA	QK		QJ
JA	JK	JQ	

in vsaj eno damo izvlečemo pri natanko 6 izidih (in v tem primeru je to isto kot reči, da izvlečemo natanko eno damo). Verjetnost tega dogodka je zdaj torej $6/12 = 1/2 = 0^{\cdot}5$.

Slepo vlečenje je pomembno v statistiki, kjer mu pravimo *enostavno slučajno vzorčenje* s *ponavljanjem* ali brez, naboru izvlečenih elementov pa *vzorec*. Na statistični množici velikosti n obstaja n^r možnih vzorcev s ponavljanjem in:

$$n(n-1)(n-2)\cdots(n-r+1)$$

vzorcev brez ponavljanja.

Pri statističnem sklepanju vrstni red vlečenja navadno ni pomemben (podobno kot tudi ni bil pomemben pri dogodku, da je vsaj ena izvlečena karta dama). Če pri vzorčenju brez ponavljanja zanemarimo vrstni red, se ohrani, da so vsi vzorci enako verjetni (kar pa ni res pri vzorčenju s ponavljanjem). Vzorcev brez ponavljanja brez informacije o vrstnem redu pa je:

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

Izrazu na desni pravimo *binomski simbol*. Le-ta v verjetnosti igra zelo pomembno vlogo.

Primer izračuna binomskega koeficienta: $\displaystyle\binom{10}{3} = \frac{10\cdot 9\cdot 8}{1\cdot 2\cdot 3} = 120.$

Velja pa tudi $\displaystyle\binom{10}{7} = 120.$

Nasploh velja $\binom{n}{r} = \binom{n}{n-r}$. Čisto vseeno je namreč, ali povemo, katerih r elementov smo vzeli v vzorec ali pa katerih $n - r$ nismo vzeli.

1.3 Relacije in operacije na dogodkih

- Pravimo, da je A *način* dogodka B, če vsi izidi, ki so v dogodku A, pripadajo tudi dogodku B. Pišemo $A \subseteq B$ ali tudi $B \supseteq A$. Posebej velja, da je vsak dogodek tudi način samega sebe: $A \subseteq A$.
- *Unija* (tudi *vsota*) dogodkov A in B je dogodek, ki ga sestavljajo tisti izidi, ki so v A *ali* v B. Unijo označujemo z $A \cup B$, v starejši literaturi tudi z $A + B$.
- *Presek* (tudi *produkt*) dogodkov A in B je dogodek, ki ga sestavljajo tisti izidi, ki so v A *in hkrati* v B. Presek označujemo z $A \cap B$, v starejši literaturi tudi z AB.
- *Razlika* dogodkov A in B je dogodek, ki ga sestavljajo tisti izidi, ki so v A, ne pa tudi v B. Razliko bomo označevali z $A \setminus B$, možna pa je tudi oznaka $A - B$.
- Dogodka A in B sta *nezdružljiva*, če je $A \cap B = \emptyset$.
- Dogodka A in B sta si *nasprotna*, če sta nezdružljiva in je $A \cup B = \Omega$. Pravimo tudi, da je dogodek A nasproten dogodku B oziroma B nasproten dogodku A.
- Dogodek, ki je (v danem prostoru izidov) nasproten danemu dogodku A, je en sam in mu pravimo tudi *komplement* dogodka A. Označevali ga bomo z $\bar{A}$, možna pa je tudi oznaka A^{c}.

Primer: Če spet vzamemo standardno kocko in dogodke (1.1.1), velja:

- A je način dogodka B in C je način dogodka D: $A \subseteq B$, $C \subseteq D$.
- $B \cup C = \{$⚀, ⚂, ⚄, ⚅$\}$.
- $B \cap C = \{$⚄$\}$.
- $B \setminus C = \{$⚅$\} = A$.
- $C \setminus B = \{$⚀, ⚂$\}$.
- Dogodka A in C sta nezdružljiva, prav tako dogodka A in D.
- Dogodka A in D sta si nasprotna: $D = \bar{A}$, $A = \bar{D}$.

Nasploh veljajo naslednje zveze:

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

Zadnjima dvema zvezama pravimo *de Morganova zakona*.

Verjetnost, definirana s pomočjo verjetnostne funkcije, ima naslednje lastnosti:

- Za vsak dogodek A je $0 \le \mathbb{P}(A) \le 1$.
- $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$.
- Če je A način dogodka B, je $\mathbb{P}(A) \le \mathbb{P}(B)$. Primer: pri dogodkih (1.1.1) na nepošteni kocki (1.2.1) je $A \subseteq B$ in $\mathbb{P}(A) = 0{\cdot}35 \le 0{\cdot}5 = \mathbb{P}(B)$.
- $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$. Primer: pri dogodkih (1.1.1) na nepošteni kocki (1.2.1) je $\mathbb{P}(D) = \mathbb{P}(\bar{A}) = 0{\cdot}65 = 1 - \mathbb{P}(A)$.
- Če sta dogodka A in B nezdružljiva, velja $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$. Primer: pri dogodkih (1.1.1) na nepošteni kocki je $\mathbb{P}(A) = 0{\cdot}35$, $\mathbb{P}(C) = 0{\cdot}35$ in $\mathbb{P}(A \cup C) = 0{\cdot}7 = \mathbb{P}(A) + \mathbb{P}(C)$.

Nasploh pa verjetnost unije ni vsota verjetnosti.

Primer: pri dogodkih (1.1.1) na nepošteni kocki (1.2.1) velja:

$$\begin{aligned} \mathbb{P}(B) &= 0{\cdot}15 + 0{\cdot}35 = 0{\cdot}5\,, \\ \mathbb{P}(C) &= 0{\cdot}05 + 0{\cdot}15 + 0{\cdot}15 = 0{\cdot}35\,, \\ \mathbb{P}(B \cup C) &= 0{\cdot}05 + 0{\cdot}15 + 0{\cdot}15 + 0{\cdot}35 = 0{\cdot}7\,, \end{aligned}$$

medtem ko je $\mathbb{P}(B) + \mathbb{P}(C) = 0.85$. Ko namreč seštejemo $\mathbb{P}(B) + \mathbb{P}(C)$, verjetnost izida ⚄, ki tvori presek $B \cap C$, štejemo dvakrat. Če želimo dobiti $\mathbb{P}(B \cup C)$, ga moramo nazaj odšteti.

Tako dobimo formulo za verjetnost unije v splošnem:

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

1.4 Aksiomi Kolmogorova

Včasih želimo gledati tudi verjetnost, ki se je ne da opisati z verjetnostno funkcijo ali pa bi bilo to nenaravno. V tem primeru definiramo verjetnost *neposredno* na dogodkih. Žal pa se izkaže, da tega ne moremo vedno storiti na vseh podmnožicah prostora izidov Ω, temveč le na določeni družini podmnožic. Pojem dogodka bomo omejili na izbrano družino, za katero bomo zahtevali, da je σ-*algebra*.

Družina podmnožic $\mathscr{F}$ dane množice Ω je σ-*algebra*, če velja:

- $\emptyset \in \mathscr{F}$.
- Če je $A \in \mathscr{F}$, je tudi $\bar{A} \in \mathscr{F}$.
- Za poljubno zaporedje dogodkov $A_1, A_2, A_3, \ldots$ mora biti tudi $A_1 \cup A_2 \cup A_3 \cup \cdots \in F$. Slednji dogodek sestavljajo vsi izidi, ki so v vsaj enem izmed dogodkov $A_1, A_2, A_3, \ldots$.

Če je $\mathscr{F}$ σ-algebra, mora veljati tudi:

- $\Omega \in \mathscr{F}$, saj je $\Omega = \bar{\emptyset}$.
- Za poljubna $A, B \in \mathscr{F}$ mora biti $A \cup B \in F$, saj je $A \cup B = A \cup B \cup \emptyset \cup \emptyset \cup \cdots$.
- Za poljubna $A, B \in \mathscr{F}$ mora biti $A \cap B \in F$, saj je $A \cap B = \overline{\bar{A} \cup \bar{B}}$.

Verjetnost (ali *verjetnostna mera*) na σ-algebri $\mathscr{F}$, ki jo sestavljajo določene podmnožice množice izidov Ω (dogodki), je predpis, ki vsakemu dogodku A priredi število, ki ga označimo s $\mathbb{P}(A)$ in mu pravimo *verjetnost* dogodka A. Pri tem mora veljati:

- $0 \le \mathbb{P}(A) \le 1$ za vsak dogodek A.
- $\mathbb{P}(\emptyset) = 0, \quad \mathbb{P}(\Omega) = 1$.
- Če sta dogodka A in B nezdružljiva, mora biti $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.
- Za poljubno zaporedje dogodkov $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$ mora veljati:

$$\mathbb{P}(A_1 \cup A_2 \cup A_3 \cup \cdots) = \lim_{n\to\infty} \mathbb{P}(A_n)\,.$$

Zgornjim pravilom pravimo *aksiomi Kolmogorova.* Prvim trem pravilom bomo rekli *osnovni aksiomi verjetnosti.*

Množico izidov skupaj z verjetnostno mero imenujemo *verjetnostni prostor.*

Vsaka verjetnost, definirana s pomočjo verjetnostne funkcije, izpolnjuje aksiome Kolmogorova, pri čemer so lahko dogodki kar vse podmnožice prostora izidov Ω. Diskretni verjetnostni prostori so torej poseben primer verjetnostnih prostorov, definiranih z aksiomi Kolmogorova. Zato je na mestu naslednja definicija.

Verjetnostna mera je *diskretna*, če izvira iz verjetnostne funkcije. Natančneje, to je tedaj, ko so dogodki kar vse podmnožice prostora izidov Ω in ko obstaja taka funkcija $p\colon\ \Omega\ \to [0,1]$, da za vsak dogodek A velja:

$$\mathbb{P}(A) = \sum_{\omega \in A} p(\omega)\,.$$

Verjetnostna funkcija p je natančno določena, saj velja $p(\omega) = \mathbb{P}(\{\omega\})$. Zato lahko ekvivalentno definiramo, da je verjetnostna mera diskretna, če so dogodki kar vse podmnožice prostora izidov Ω in če za vsak dogodek A velja:

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\})\,.$$

Diskretne verjetnostne mere so torej tiste, pri katerih je verjetnost določena že na enostavnih dogodkih. Velja tudi, da je verjetnostna mera $\mathbb{P}$ iskkretna natenko tedaj, ko je $p(\omega) := \mathbb{P}(\{\omega\})$ verjetnostna funkcija, t. j. ko je $\sum_{\omega\in\Omega} p(\omega) = 1$.

Verjetnost, definirana s pomočjo aksiomov Kolmogorova, ima vse lastnosti, ki smo jih nanizali za diskretne verjetnostne mere:

- $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$: to velja zato, ker sta dogodka A in $\bar{A}$ nezdružljiva, njuna unija pa je gotov dogodek Ω, ki ima verjetnost 1.
- Če je $A \subseteq B$, je $\mathbb{P}(A) \le \mathbb{P}(B)$: to velja zato, ker je v tem primeru $B = A \cup (B \setminus A)$ in kar sta dogodka A in $B \setminus A$ nezdružljiva, mora veljati $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \ge \mathbb{P}(A)$.
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$: to velja zato, ker je $A \cup B$ unija nezdružljivih dogodkov A in $B \setminus A$, B pa je unija nezdružljivih dogodkov $B \setminus A$ in $A \cap B$. Sledi:

  $$\begin{aligned} \mathbb{P}(A \cap B) &= \mathbb{P}(A) + \mathbb{P}(B \setminus A) \\ \mathbb{P}(B) &= \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B) \end{aligned}$$

  Če enačbi odštejemo in uredimo, dobimo prej omenjeno zvezo.
- Če so $A_1, A_2, \ldots, A_n$ paroma nezdružljivi (t. j. poljubna dva z različnima indeksoma sta nezdružljiva), velja:

  $$\mathbb{P}(A_1 \cup A_2 \cup \cdots \cup A_n) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \cdots + \mathbb{P}(A_n)\,.$$

  Trditev velja tudi za neskončno zaporedje paroma nezdružljivih dogodkov:

  $$\mathbb{P}(A_1 \cup A_2 \cup \cdots) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \cdots\,,$$

  kjer je vrednost izraza na desni, ki mu pravimo *vrsta*, definirana kot limita delnih vsot. Zgornja formula je (ob privzetju ostalih) ekvivalentna zadnjima dvema aksiomoma Kolmogorova.

Iz zadnje lastnosti sledi, da je vsaka verjetnostna mera na končnem ali števno neskončnem verjetnostnem prostoru diskretna.

Primer nediskretne verjetnostne mere. Avtobus vozi na 10 minut. Pridemo na slepo na postajo. Kolikšna je verjetnost, da bomo čakali več kot 7 minut?

Za prostor izidov tu postavimo interval $(0, 10]$. Posamezen izid naj recimo predstavlja čas, ki je minil od odhoda zadnjega avtobusa, na katerega nismo mogli priti (izid 10 pomeni, da je ob istem času kot mi prišel avtobus in nas je voznik ravno še spustil noter). Dogodek, da čakamo več kot 7 minut, ustreza intervalu $(0, 3]$. Torej je verjetnost tega dogodka enaka 30%.

Kako smo dobili zadnji rezultat? Tako, da smo *dolžino* intervala $(0, 3]$ delili z dolžino celotnega intervala $(0, 10]$. Če je torej dogodek $I \subseteq (0, 10]$ interval, je njegova verjetnost enaka:

$$\mathbb{P}(I) = \frac{\text{dolžina}(I)}{10} \,.$$

Interval je lahko odprt, zaprt ali polodprt. Lahko je tudi izrojen, torej ena sama točka. To ustreza vprašanju, kolikšna je verjetnost, da bomo čakali *natanko* 7 minut (niti delčka sekunde več niti delčka sekunde manj). Ker je dolžina ena same točke enaka nič, je tudi verjetnost dogodka, da bomo čakali natanko 7 minut, enaka nič (takim dogodkom pravimo *skoraj nemogoči*). Od tod sledi tudi, da ta verjetnostna mera ni diskretna, saj je pripadajoča funkcija p povsod enaka nič in zato ne more biti verjetnostna funkcija.

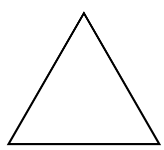
Prejšnji primer je primer *slepe izbire*, definirane splošnejše. Za slepo izbiro iz končne množice smo že definirali, da je to takrat, ko so vse možnosti enako verjetne. Z drugimi besedami, če Ω končna množica, je izid izbran na slepo, če za vsak dogodek A velja $\mathbb{P}(A) = \dfrac{\sharp(A)}{\sharp(\Omega)}$.

Slepa izbira iz intervala na realni osi, ravninskega lika ali prostorskega telesa pa pomeni, da je:

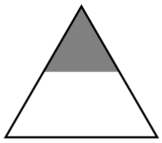
$$\mathbb{P}(A) = \frac{\text{mera}(A)}{\text{mera}(\Omega)} \,,$$

kjer je mera dolžina (za intervale na realni osi), ploščina (za ravninske like) oziroma prostornina (za prostorska telesa).

Primer: na slepo izberemo točko iz trikotnika v ravnini z vodoravno osnovnico:



in zanima nas verjetnost, da je izbrana točka na več kot polovici višine trikotnika. Dogodek A je prikazan kot osenčeno območje:



Dogodek A je trikotnik s polovično osnovnico in polovično višino trikotnika, ki ponazarja cel verjetnostni prostor Ω, torej je njegova ploščina četrtina ploščine trikotnika Ω. Sledi $\mathbb{P}(A) = 1/4$.

2.

Pogojna verjetnost

2.1 Definicija pogojne verjetnosti

Pogojna verjetnost dogodka A glede na dogodek B, $\mathbb{P}(A \mid B)$, je verjetnost, če vemo, da se je zgodil B. Privzemimo za hip, da je verjetnost dogodka definirana kar kot delež poskusov, pri katerih se zgodi dani dogodek, torej: t. j.

$$\mathbb{P}(A) = \frac{\text{število poskusov, pri katerih se zgodi } A}{\text{število vseh poskusov}},$$

je smiselno definirati:

$$\mathbb{P}(A \mid B) = \frac{\text{število poskusov, pri katerih se zgodi } A\text{, med tistimi, pri katerih se zgodi } B}{\text{število poskusov, pri katerih se zgodi } B}$$

Velja:

$$\begin{aligned}
\mathbb{P}(A \mid B) &= \frac{\text{število poskusov, pri katerih se zgodita } A \text{ in } B}{\text{število poskusov, pri katerih se zgodi } B} = \\
&= \frac{\dfrac{\text{število poskusov, pri katerih se zgodita } A \text{ in } B}{\text{število vseh poskusov}}}{\dfrac{\text{število poskusov, pri katerih se zgodi } B}{\text{število vseh poskusov}}} = \\
&= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.
\end{aligned}$$

Slednje vzamemo za definicijo pogojne verjetnosti v splošnem, ko verjetnost ni nujno delež poskusov:

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

Omenili smo že, da definicija verjetnosti z deležem poskusov pomeni slepo izbiro poskusa. No, če gre za slepo izbiro, t. j. če so vsi izidi enako verjetni, to pomeni:

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

Primer: Vzemimo pošteno kocko in dogodka iz (1.1.1):

$$
\begin{aligned}
B &= \{\text{pade najmanj pet pik}\} = \{\text{⚄}, \text{⚅}\}\,, \\
D &= \{\text{pade manj kot šest pik}\} = \{\text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}\}\,,
\end{aligned}
$$

Tedaj velja $\mathbb{P}(B \mid D) = 1/5 = 0{\cdot}2$.

Če pa kocka ni poštena, temveč sledi porazdelitveni shemi (1.2.1):

$$
\begin{pmatrix} \text{⚀} & \text{⚁} & \text{⚂} & \text{⚃} & \text{⚄} & \text{⚅} \\ 0{\cdot}05 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}35 \end{pmatrix}\,,
$$

velja $\mathbb{P}(B \mid D) = \dfrac{0{\cdot}15}{0{\cdot}05 + 0{\cdot}15 + 0{\cdot}15 + 0{\cdot}15 + 0{\cdot}15} = \dfrac{0{\cdot}15}{0{\cdot}65} \doteq 0{\cdot}231$.

Pri neodvisnih dogodkih se pogojna verjetnost ujema z brezpogojno: če sta A in B neodvisna, velja:

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

2.2 Relejni poskusi

Dostikrat poznamo določene pogojne verjetnosti, želeli pa bi izračunati brezpogojno verjetnost. To se dogaja pri *relejnih poskusih*, ki potekajo v več fazah. Recimo, da se v prvi fazi lahko zgodi dogodek A, v drugi pa dogodek B. Tedaj navadno po naravi stvari poznamo $\mathbb{P}(A)$ in $\mathbb{P}(B \mid A)$, brezpogojno verjetnost $\mathbb{P}(A \cap B)$ pa izračunamo po formuli:

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

Primer: Žena pošlje moža na trg po glavo solate, ki jo prodajata dve branjevki, Francka in Micka. Verjetnost, da mož kupi solato pri Francki, je 40%, da kupi pri Micki, pa 60%. Francka ima 10%, Micka pa 20% nagnitih glav solate. Privzamemo, da Francka in Micka izbirata solato na slepo.

Tukaj gre za dvofazni relejni poskus: v prvi fazi mož izbira branjevko, v drugi pa branjevka izbira solato.

Žena si zastavlja naslednja vprašanja:

1. Najprej jo zanima verjetnost, da bo mož prinesel domov nagnito glavo solate.
2. Nato gre z mislimi še malo dlje in se vpraša po verjetnosti, da bo šel k Micki in tam kupil nagnito glavo solate.
3. Končno se mož vrne domov – z nagnito glavo solate. Sedaj se žena vpraša po *pogojni* verjetnosti, da je solato kupil pri Micki.

Zaenkrat znamo dogovoriti na drugo vprašanje: če je G dogodek, da mož prinese domov nagnito gnilo solate, M pa dogodek, da kupi solato pri Micki, vemo:

$$\mathbb{P}(M) = 0\text{˙}6\,, \quad \mathbb{P}(G \mid M) = 0\text{˙}2\,.$$

Odgovor na iskano vprašanje je:

$$\mathbb{P}(M \cap G) = \mathbb{P}(M)\,\mathbb{P}(G \mid M) = 0\text{˙}6 \cdot 0\text{˙}2 = 0\text{˙}12\,.$$

Za odgovor na prvo in tretje vprašanje pa potrebujemo še malo teorije.

Dogodki $H_1, H_2, \ldots, H_n$ tvorijo *popoln sistem dogodkov* ali tudi *razčlenitev* (angl. *partition*) prostora izidov, če so paroma nezdružljivi, njihova unija pa je gotov dogodek. Z drugimi besedami, vsak izid pripada natanko enemu izmed dogodkov $H_1, H_2, \cdots H_n$.

Če je $H_1, H_2, \ldots, H_n$ popoln sistem dogodkov in A poljuben dogodek, so tudi dogodki $H_1 \cap A, H_2 \cap A, \ldots, H_n \cap A$ paroma nezdružljivi, njihova unija pa je dogodek A. Sledi:

$$\mathbb{P}(A) = \mathbb{P}(H_1 \cap A) + \mathbb{P}(H_2 \cap A) + \cdots + \mathbb{P}(H_n \cap A)$$

oziroma:

$$\mathbb{P}(A) = \mathbb{P}(H_1)\,\mathbb{P}(A \mid H_1) + \mathbb{P}(H_2)\,\mathbb{P}(A \mid H_2) + \cdots + \mathbb{P}(H_n)\,\mathbb{P}(A \mid H_n)\,.$$

Zgornji formuli pravimo *izrek o polni verjetnosti* (angl. *law of total probability*). Uporaben je pri dvofaznih relejnih poskusih. Dogodki $H_1, H_2, \ldots, H_n$ predstavljajo vse možnosti za prvo fazo poskusa, zato jim pravimo tudi *hipoteze*, dogodek A pa se nanaša na drugo fazo poskusa.

Sedaj lahko odgovorimo na prvo ženino vprašanje. Če z F označimo še dogodek, da mož kupi solato pri Francki, dogodka F in M sestavljata popoln sistem dogodkov. Velja:

$$\mathbb{P}(F) = 0\text{˙}4\,, \quad \mathbb{P}(G \mid M) = 0\text{˙}1\,.$$

Po izreku o polni verjetnosti velja:

$$\mathbb{P}(G) = \mathbb{P}(F)\,\mathbb{P}(G \mid F) + \mathbb{P}(M)\,\mathbb{P}(G \mid M) = 0\text{˙}4 \cdot 0\text{˙}1 + 0\text{˙}6 \cdot 0\text{˙}2 = 0\text{˙}16\,.$$

Preostane še tretje vprašanje. Tu skušamo iz druge faze poskusa, ki je *opazljiva* (angl. *observable*) – žena opazi nagnito glavo solate, rekonstruirati prvo fazo, ki ni opazljiva – žena ne ve, pri kateri branjevki je mož kupil solato. V splošnem kontekstu nas torej zanimajo pogojne verjetnosti neopazljivih hipotez glede na opaženo drugo fazo poskusa, t. j. $\mathbb{P}(H_i \mid A)$. Po definiciji pogojne verjetnosti je:

$$\mathbb{P}(H_i \mid A) = \frac{\mathbb{P}(H_i \cap H)}{\mathbb{P}(A)} = \frac{\mathbb{P}(H_i)\,\mathbb{P}(A \mid H_i)}{\mathbb{P}(A)}\,.$$

Če v imenovalec vstavimo izrek o polni verjetnosti, dobimo *Bayesovo formulo*:

$$\mathbb{P}(H_i \mid A) = \frac{\mathbb{P}(H_i)\,\mathbb{P}(A \mid H_i)}{\mathbb{P}(H_1)\,\mathbb{P}(A \mid H_1) + \mathbb{P}(H_2)\,\mathbb{P}(A \mid H_2) + \cdots + \mathbb{P}(H_n)\,\mathbb{P}(A \mid H_n)}\,.$$

Bayesova formula torej potrebuje brezpogojne verjetnosti dogodkov H_i, ki jim pravimo tudi *apriorne* verjetnosti (apriorne so zato, ker se nanašajo na čas pred realizacijo druge faze poskusa), da pa nam pogojne verjetnosti, ki jim pravimo tudi *aposteriorne* verjetnosti (ker se nanašajo na čas po realizaciji druge faze poskusa).

Odgovor na tretje vprašanje pri našem primeru je torej:

$$\mathbb{P}(M \mid G) = \frac{\mathbb{P}(M)\,\mathbb{P}(G \mid M)}{\mathbb{P}(F)\,\mathbb{P}(F \mid M) + \mathbb{P}(M)\,\mathbb{P}(G \mid M)} = \frac{0\text{˙}6 \cdot 0\text{˙}2}{0\text{˙}4 \cdot 0\text{˙}1 + 0\text{˙}6 \cdot 0\text{˙}2} = 0\text{˙}75\,.$$

Dogodek, da mož kupi solato pri Micki, ima torej apriorno verjetnost 0˙6, njegova aposteriorna verjetnost glede na dano opažanje pa je 0˙75.

Na Bayesovi formuli temelji cela veja statistike, ki ji pravimo *Bayesova statistika*. Ta se razlikuje od *klasične* inferenčne statistike po tem, da potrebuje apriorne verjetnosti za tisto, kar nas zanima. Klasična statistika ne privzema nikakršnih apriornih verjetnosti, zato pa so tudi njeni rezultati šibkejše narave.

3.

Slučajne spremenljivke

3.1 Osnovni pojmi

Ideja slučajne spremenljivke je, da je to število ali kakšna druga vrednost, ki ni natančno določena, je pa v zvezi z njo smiselno gledati verjetnost. Pač pa postane natančno določena, ko enkrat realiziramo poskus. Z drugimi besedami, če poznamo izid poskusa, poznamo vrednosti vseh slučajnih spremenljivk. Tako je smiselno postaviti naslednjo definicijo:

Slučajna spremenljivka z vrednostmi v množici M je predpis, ki vsakemu izidu priredi določeno vrednost v M, torej preslikava iz Ω v M.

Slučajne spremenljivke označujemo tako kot statistične, torej navadno z velikimi tiskanimi črkami. Statistična spremenljivka postane slučajna spremenljivka, če slučajno (npr. na slepo) izberemo en element iz statistične množice.

Tako kot statistične tudi slučajne spremenljivke ločimo po tipu (torej po strukturi, ki je definirana na množici M) na imenske, urejenostne, intervalske in razmernostne, poleg tega pa imamo še dihotomne slučajne spremenljivke.

Tudi veliko drugih pojmov iz opisne statistike, torej definirane za statistične spremenljivke, je možno definirati tudi za slučajne spremenljivke. Osnovno vodilo pri tem je, naj pojma sovpadata, če je slučajna spremenljivka dobljena iz statistične s slepo izbiro enote.

Primer. *Briškula* se igra s 40 kartami, ki jih tvorijo vse možne kombinacije 4 *barv* (špada – meč, kopa – pokal, bašton – palica in denar) in 10 *moči* (2, 4, 5, 6, 7, fant – F, kaval – C, kralj – R, trojka – 3 in as – A). Moč določa, katera karta pobere, če gre za isto barvo. Poleg tega ima vsaka karta glede na moč tudi svojo vrednost in vrednosti se na koncu seštejejo.

Komplet kart je tako statistična množica, prejšnji pojmi pa statistične spremenljivke na njej:

- *Barva* je imenska spremenljivka (vsaj dokler ni znan adut – briškula).

- *Moč* je urejenostna spremenljivka: $2 < 4 < 5 < 6 < 7 < F < C < R < 3 < A$).
- *Vrednost* je razmernostna spremenljivka. Enaka je 0 za *liše* (moči 2, 4, 5, 6, 7), 2 za fante, 3 za kavale, 4 za kralje, 10 za trojke in 11 za ase.

Lahko pa s kupa slučajno izberemo eno karto (ali pa gledamo aduta spodaj). V tem primeru postane množica kart verjetnostni prostor, barva, moč in vrednost pa slučajne spremenljivke.

Če barvo označimo z B, moč z M in vrednost z V ter gledamo izid–karto ω, ki naj bo špadni fant (fant pri mečih), velja:

$$B(\omega) = \text{špada}\,, \quad M(\omega) = \text{fant}\,, \quad V(\omega) = 2\,.$$

Porazdelitev slučajne spremenljivke pove, katere vrednosti zavzame s kolikšnimi verjetnosti. V splošnem porazdelitev opišemo tako, da za vse množice A iz določene σ-algebre povemo $\mathbb{P}(X \in A)$, t. j. verjetnost, da X pripada množici A. Toda če je X definirana na diskretnem verjetnostnem prostoru, je dovolj povedati *točkaste verjetnosti* $\mathbb{P}(X \in x)$ za vse možne vrednosti x, saj velja:

$$\mathbb{P}(X \in A) = \sum_{x \in A} \mathbb{P}(X = x)\,.$$

Za slučajno spremenljivko, za katero velja zgornja formula (in torej lahko njeno porazdelitev opišemo s točkastimi verjetnostmi), pravimo, da je *diskretna* oz. diskretno porazdeljena. Pravimo tudi, da je njena porazdelitev diskretna.

Porazdelitev diskretne slučajne spremenljivke lahko opišemo s *porazdelitveno shemo*, ki odgovarja verjetnostni shemi. Če X lahko zavzame vrednosti $a_1, a_2, \ldots, a_k$, ki so vse različne, in če je $\mathbb{P}(X = a_1) = p_1$, $\mathbb{P}(X = a_2) = p_2$, $\ldots$, $\mathbb{P}(X = a_k) = p_k$, pišemo:

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

Seveda velja $p_1 + p_2 + \cdots p_k = 1$.

Če vrednosti $a_1, \ldots, a_k$ niso vse različne, moramo seštevati. V vsakem primeru velja $\mathbb{P}(X \in A) = \sum_{i; a_i \in A} p_i$.

Primer: Če na slepo izberemo karto pri briškuli, so porazdelitve njene barve, moči in vrednosti podane s shemami:

$$B \sim \begin{pmatrix} \text{špada} & \text{bašton} & \text{kopa} & \text{denar} \\ \frac{10}{40} & \frac{10}{40} & \frac{10}{40} & \frac{10}{40} \end{pmatrix} = \begin{pmatrix} \text{špada} & \text{bašton} & \text{kopa} & \text{denar} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}\,,$$

$$M \sim \begin{pmatrix} 2 & 4 & 5 & 6 & 7 & \text{fant} & \text{kaval} & \text{kralj} & \text{trojka} & \text{as} \\ \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} \end{pmatrix}\,,$$

$$V \sim \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 2 & 3 & 4 & 10 & 11 \\ \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} \end{pmatrix} = \begin{pmatrix} 0 & 2 & 3 & 4 & 10 & 11 \\ \frac{1}{2} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} \end{pmatrix}$$

Verjetnost dogodka, da bo vrednost karte manjša of 4, je enaka:

$$\mathbb{P}(V < 4) = \mathbb{P}(V = 0) + \mathbb{P}(V = 2) + \mathbb{P}(V = 3) = \frac{1}{2} + \frac{1}{10} + \frac{1}{10} = \frac{7}{10}\,.$$

Verjetnost dogodka, da bo vrednost karte strogo med 3 in 7, pa je enaka:

$$\mathbb{P}(3 < V < 7) = \mathbb{P}(V = 4) = \frac{1}{10}\,.$$

Diskretna porazdelitev je *enakomerna* na množici $\{a_1, a_2, \ldots, a_k\}$, če so vse vrednosti $a_1, a_2, \ldots, a_k$ enako verjetne, druge vrednosti pa niso zavzete. To pa lahko povemo tudi drugače: *slepa izbira* stvari, ki je definirana kot vrednost slučajne spremenljivke, iz dane množice pomeni, da je porazdelitev te slučajne spremenljivke enakomerna na tej množici.

Pri prej omenjenih slučajnih spremenljivkah pri briškuli sta tako barva B in moč M porazdeljeni enakomerno, vrednost V pa ne. Slepa izbira karte pri briškuli torej pomeni slepo izbiro barve iz množice $\{\text{špada}, \text{bašton}, \text{kopa}, \text{denar}\}$, in slepo izbiro moči iz množice $\{2, 4, 5, 6, 7, \text{fant}, \text{kaval}, \text{kralj}, \text{trojka}, \text{as}\}$, ne pomeni pa slepe izbire vrednosti.

Modus diskretne slučajne spremenljivke je tista njena vrednost, kjer je njena verjetnostna funkcija največja. Modusov je lahko tudi več.

Primer: porazdelitev s shemo:

$$\begin{pmatrix} 0 & 2 & 3 & 4 & 10 & 11 \\ \frac{1}{2} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} & \frac{1}{10} \end{pmatrix}$$

ima modus 0. Porazdelitev s shemo:

$$\begin{pmatrix} 0 & 1 & 2 & 3 \\ 0\text{'}125 & 0\text{'}375 & 0\text{'}375 & 0\text{'}125 \end{pmatrix}$$

ima dva modusa (1 in 2). Pri enakomerni porazdelitvi pa so kar vse vrednosti modusi.

3.2 Navzkrižne porazdelitve

Podobno kot smo pri opisni statistiki gledali povezanost med dvema statističnima spremenljivkama, moramo tudi v verjetnosti dostikrat hkrati gledati dve ali več slučajnih spremenljivk. Če sta na istem verjetnostnem prostoru definirani dve slučajni spremenljivki X in Y, je njuna *navzkrižna* ali *skupna porazdelitev* (angl. *joint distribution*) opišemo z navzkrižnimi verjetnostmi $\mathbb{P}(X \in A, Y \in B)$, pri diskretnih slučajnih spremenljivkah pa je dovolj povedati navzkrižne točkaste verjetnosti $\mathbb{P}(X = a, Y = b)$.

Primer: iz dobro premešanega kupa 10 kart za briškulo ene same barve na slepo vzamemo dve karti. Označimo z V_1 vrednost prve, z V_2 pa druge izvlečene karte. Če jemljemo z vračanjem, je možnih 100 enako verjetnih izidov in navzkrižno porazdelitev slučajnih spremenljivk V_1 in V_2 lahko predstavimo s tabelo:

	$V_2 = 0$	$V_2 = 2$	$V_2 = 3$	$V_2 = 4$	$V_2 = 10$	$V_2 = 11$
$V_1 = 0$	$\frac{1}{4}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$
$V_1 = 2$	$\frac{1}{20}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$
$V_1 = 3$	$\frac{1}{20}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$
$V_1 = 4$	$\frac{1}{20}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$
$V_1 = 10$	$\frac{1}{20}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$
$V_1 = 11$	$\frac{1}{20}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$

Izračunajmo še verjetnost dogodka, da je skupna vrednost obeh kart manjša od 4. Velja:

$$\begin{aligned}\mathbb{P}(V_1 + V_2 < 4) &= \mathbb{P}(V_1 = 0, V_2 = 0) + \mathbb{P}(V_1 = 0, V_1 = 2) + \mathbb{P}(V_1 = 0, V_2 = 3) + \\ &\quad + \mathbb{P}(V_1 = 2, V_2 = 0) + \mathbb{P}(V_1 = 3, V_2 = 0) = \\ &= \frac{9}{20} = 0{\cdot}45\,.\end{aligned}$$

Če pa jemljemo brez vračanja, imamo 90 enako verjetnih izidov in navzkrižno porazdelitev predstavlja tabela:

	$V_2 = 0$	$V_2 = 2$	$V_2 = 3$	$V_2 = 4$	$V_2 = 10$	$V_2 = 11$
$V_1 = 0$	$\frac{2}{9}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$
$V_1 = 2$	$\frac{1}{18}$	0	$\frac{1}{90}$	$\frac{1}{90}$	$\frac{1}{90}$	$\frac{1}{90}$
$V_1 = 3$	$\frac{1}{18}$	$\frac{1}{90}$	0	$\frac{1}{90}$	$\frac{1}{90}$	$\frac{1}{90}$
$V_1 = 4$	$\frac{1}{18}$	$\frac{1}{90}$	$\frac{1}{90}$	0	$\frac{1}{90}$	$\frac{1}{90}$
$V_1 = 10$	$\frac{1}{18}$	$\frac{1}{90}$	$\frac{1}{90}$	$\frac{1}{90}$	0	$\frac{1}{90}$
$V_1 = 11$	$\frac{1}{18}$	$\frac{1}{90}$	$\frac{1}{90}$	$\frac{1}{90}$	$\frac{1}{90}$	0

in velja $\mathbb{P}(V_1 + V_2 < 4) = \dfrac{4}{9} \doteq 0{\cdot}444$.

Če imamo podano navzkrižno porazdelitev slučajnih spremenljivk X in Y, poznamo tudi porazdelitvi posamičnih slučajnih spremenljivk. Pravimo jima *robni porazdelitvi*, to pa zato, ker ju v diskretnem primeru dobimo s seštevanjem navzkrižnih verjetnosti, te pa navadno zapišemo na rob tabele. Če lahko Y zavzame vrednosti $b_1, b_2, \ldots, b_m$ (vse različne), velja:

$$\mathbb{P}(X = a_i) = \mathbb{P}(X = a_i, Y = b_1) + \mathbb{P}(X = a_i, Y = b_2) + \cdots + \mathbb{P}(X = a_i, Y = b_n)\,.$$

Z navzkrižnimi porazdelitvami je povezana še ena različica slepe izbire. Če izbiramo dve stvari, ki sta definirani kot vrednosti dveh slučajnih spremenljivk X in Y, pri čemer lahko X zavzame vrednosti $\{a_1, \ldots, a_m\}$ (vse različne), Y pa vrednosti $\{b_1, \ldots, b_n\}$ (prav tako vse različne), pravimo, da X in Y izberemo *na slepo in neodvisno*, če so vsi dogodki:

$$\begin{array}{cccc}
\{X = a_1, Y = b_1\}, & \{X = a_1, Y = b_2\}, & \cdots & \{X = a_1, Y = b_n\}, \\
\{X = a_2, Y = b_1\}, & \{X = a_2, Y = b_2\}, & \cdots & \{X = a_2, Y = b_n\}, \\
\vdots & \vdots & \ddots & \vdots \\
\{X = a_m, Y = b_1\}, & \{X = a_m, Y = b_2\}, & \cdots & \{X = a_m, Y = b_n\},
\end{array}$$

enako verjetni, torej če ima vsak izmed njih verjetnost $\dfrac{1}{mn}$.

Podobno definiramo tudi slepo in neodvisno izbiro treh ali več stvari.

Poseben primer slepe in neodvisne izbire je vlečenje s ponavljanjem: če izvlečemo dva elementa, slučajna spremenljivka X predstavlja prvi, slučajna spremenljivka Y pa drugi izvlečeni element. Tedaj imata X in Y isto zalogo vrednosti. Ni pa to nujno.

Primer: neodvisno vržemo pošten kovanec in pošteno standardno kocko, kar pomeni, da na slepo in neodvisno izberemo grb oz. cifro na kovancu in eno izmed šestih strani kocke. Če X predstavlja stran kovanca, Y pa stran kocke, dobimo naslednjo navzkrižno porazdelitev:

	$Y =$ ⚀	$Y =$ ⚁	$Y =$ ⚂	$Y =$ ⚃	$Y =$ ⚄	$Y =$ ⚅
$X =$ cifra	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$
$X =$ grb	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$

Recimo zdaj, da dobimo za cifro na kovancu dva evra in za vsako piko na kocki en evro. Dogodek, da skupaj dobimo vsaj 6 evrov, lahko tedaj zapišemo v obliki:

$\{X = \text{cifra}, Y =$ ⚃$\} \cup \{X = \text{cifra}, Y =$ ⚄$\} \cup \{X = \text{cifra}, Y =$ ⚅$\} \cup \{X = \text{grb}, Y =$ ⚅$\}$

in njegova verjetnost je enaka $\dfrac{4}{12} = \dfrac{1}{3}$.

Primer: če na slepo izvlečemo karto iz kupa pri briškuli, to pomeni tudi, da na slepo in neodvisno izberemo njeno moč in barvo. Naj bo A dogodek, da izvlečemo figuro (kralja, kavala ali fanta) B pa dogodek, da izvlečemo denar ali kopo. Tedaj velja:

$$\begin{aligned}
\mathbb{P}(A) &= \frac{12}{40} = \frac{3}{10} = 0{\cdot}3 \\
\mathbb{P}(B) &= \frac{20}{40} = \frac{2}{4} = \frac{1}{2} = 0{\cdot}5 \\
\mathbb{P}(A \cap B) &= \frac{6}{40} = \frac{3}{20} = 0{\cdot}15\,.
\end{aligned}$$

Opazimo, da je $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$.

Opažanja iz zgornjega primera lahko posplošimo na naslednji dve pomembni ugotovitvi:

- Če dve ali več slučajnih spremenljivk izbiramo na slepo in neodvisno, je tudi posamezna slučajna spremenljivka izbrana na slepo (z drugimi besedami, ima enakomerno porazdelitev).

- Če sta X in Y izbrani na slepo in neodvisno ter je A dogodek, ki je enolično določen z X (pišemo tudi $A \in \sigma(X)$), B pa je dogodek, ki je enolično določen z Y (pišemo tudi $B \in \sigma(Y)$), velja $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$.

Slednje opažanje pa vodi v pojem neodvisnosti dogodkov. To je posplošitev primera, ko dogodka izhajata iz slepe in neodvisne izbire tako kot zgoraj.

Dogodka A in B sta *neodvisna*, če velja $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$.

Neodvisnost je navadno nekaj, iz česar izhajamo, kar privzamemo a priori. Včasih pa dobimo neodvisnost tudi kar tako.

Primer: pri pošteni kocki sta naslednja dogodka iz (1.1.1):

$$\begin{aligned} B &= \{\text{pade najmanj pet pik}\} = \{\text{⚄}, \text{⚅}\}\,, \\ C &= \{\text{pade liho število pik}\} = \{\text{⚀}, \text{⚂}, \text{⚄}\}\,, \end{aligned}$$

neodvisna. Velja namreč $B \cap C = \{\text{⚄}\}$ in:

$$\mathbb{P}(B) = \frac{1}{3}\,, \quad \mathbb{P}(C) = \frac{1}{2}\,, \quad \mathbb{P}(B \cap C) = \frac{1}{6} = \mathbb{P}(B)\,\mathbb{P}(C)\,.$$

Neodvisnost lahko izrazimo s pomočjo pogojne verjetnosti: če je $\mathbb{P}(B) > 0$, sta A in B neodvisna natanko tedaj, ko je $\mathbb{P}(A \mid B) = \mathbb{P}(A)$. To lahko interpretiramo kot 'A je neodvisen od B'.

Pri neodvisnosti je pomembno naslednje dejstvo: *če sta dogodka A in B neodvisna, to velja tudi za dogodka A in $\bar{B}$.* To lahko tudi dokažemo:

- Ker sta $A \cap B$ in $A \cap \bar{B} = A \setminus B$ nezdružljiva z unijo A, velja
$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap \bar{B})$.

- $\mathbb{P}(A \cap \bar{B}) = \mathbb{P}(A) - \mathbb{P}(A \cap B) = \mathbb{P}(A) - \mathbb{P}(A)\,\mathbb{P}(B) = \mathbb{P}(A)\big(1 - \mathbb{P}(B)\big) =$
$= \mathbb{P}(A)\,\mathbb{P}(\bar{B})$.

To pomeni, da je neodvisnost dogodkov A in B ekvivalentna neodvisnosti katerih koli dogodkov $\tilde{A}$ in $\tilde{B}$, kjer je $\tilde{A} = A$ ali $\tilde{A} = \bar{A}$ in $\tilde{B} = B$ ali $\tilde{B} = \bar{B}$.

Neodvisnost dogodkov motivira neodvisnost slučajnih spremenljivk.

Slučajni spremenljivki X in Y sta *neodvisni*, če sta kateri koli dogodek, ki je natančno določen z X, in kateri koli dogodek, ki je natančno določen z Y, neodvisna.

Za diskretni slučajni spremenljivki X in Y se da dokazati, da sta neodvisni natanko tedaj, ko za poljubna relevantna a in b velja $\mathbb{P}(X = a, Y = b) = \mathbb{P}(X = a)\,\mathbb{P}(Y = b)$.

Za enakomerno porazdeljene slučajne spremenljivke (slepo izbiro) pa se da dokazati naslednje: X in Y sta izbrani na slepo in neodvisno natanko tedaj, ko sta X in Y vsaka zase izbrani na slepo in ko sta neodvisni. Tako formulacija 'na slepo in neodvisno' dobi smisel.

Primer: pri slepem vlečenju s ponavljanjem sta prvi in drugi izvlečeni element neodvisna. To pa ne velja za vlečenje brez ponavljanja.

Primer: če na slepo izvlečemo karto iz kupa pri briškuli, sta tudi njena *vrednost* in barva neodvisni, čeprav vrednost ni porazdeljena enakomerno.

Primer: slučajni spremenljivki $X \sim \begin{pmatrix} 1 & 2 \\ 0{\cdot}4 & 0{\cdot}6 \end{pmatrix}$ in $Y \sim \begin{pmatrix} 1 & 2 & 3 \\ 0{\cdot}2 & 0{\cdot}5 & 0{\cdot}2 \end{pmatrix}$ sta neodvisni natanko tedaj, ko je njuna navzkrižna porazdelitev podana s tabelo:

	$Y = 1$	$Y = 2$	$Y = 3$	
$X = 1$	0·12	0·2	0·08	0·4
$X = 2$	0·18	0·3	0·12	0·6
	0·3	0·5	0·2	1

Pojem neodvisnosti lahko posplošimo tudi na več dogodkov in slučajnih spremenljivk, a pri dogodkih moramo biti nekoliko previdni. Izkaže se, da ima smisel definirati, da so dogodki $A_1, A_2, \ldots, A_n$ neodvisni, če veljajo vse zveze oblike:

$$\mathbb{P}(\tilde{A}_1 \cap \tilde{A}_2 \cap \cdots \tilde{A}_n) = \mathbb{P}(\tilde{A}_1)\,\mathbb{P}(\tilde{A}_2) \cdots \mathbb{P}(\tilde{A}_n)\,,$$

kjer za vsak i lahko postavimo bodisi $\tilde{A}_i = A_i$ bodisi $\tilde{A}_i = \bar{A}_i$. V tem primeru se namreč neodvisnost ohrani, če določene dogodke izločimo. Če so recimo A, B in C neodvisni, sta tudi A in B neodvisna.

Primer: vržemo dva poštena in neodvisna kovanca in si oglejmo naslednje dogodke:

$$\begin{aligned} A &:= \{\text{prvič pade cifra}\} \\ B &:= \{\text{drugič pade cifra}\} \\ C &:= \{\text{prvič in drugič pade enako}\} \end{aligned}$$

Za verjetnostni prostor lahko tedaj postavimo $\Omega = \{cc, cg, gc, gg\}$ in vsi štirje izidi so enako verjetni. Velja:

$$\begin{gathered} A = \{cc, cg\}\,, \quad B = \{cc, gc\}\,, \quad C = \{cc, gg\}\,, \\ A \cap B = A \cap C = B \cap C = A \cap B \cap C = \{cc\} \end{gathered}$$

in še:

$$\begin{gathered} \mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{2}\,, \\ \mathbb{P}(A \cap B) = \mathbb{P}(A \cap C) = \mathbb{P}(B \cap C) = \mathbb{P}(A \cap B \cap C) = \frac{1}{4}\,, \end{gathered}$$

od koder dobimo, da sta dogodka A in B neodvisna, prav tako tudi A in C ter B in C. Dogodki A, B in C pa so odvisni.

Primer: na slepo izberemo en element iz množice $\{A, K, Q, J, 10, 9, 8, 7\}$, ki naj bo kar verjetnostni prostor. Definirajmo naslednje dogodke:

$$F := \{A, K, Q, J\}\,,$$
$$G := \{A, 9, 8, 7\}\,,$$
$$H := \{A, K, Q, 10\}\,.$$

Dogodki F, G in H niso neodvisni, čeprav velja $\mathbb{P}(F \cap G \cap H) = \mathbb{P}(F)\,\mathbb{P}(G)\,\mathbb{P}(H)$.

Pri slučajnih spremenljivkah pa je posplošitev bolj premočrtna: slučajne spremenljivke $X_1, X_2, \ldots, X_n$ so neodvisne, če za vse možne nabore dogodkov $A_1, A_2, \ldots, A_n$, kjer je dogodek A_i enolično določen z X_i, velja:

$$\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathbb{P}(A_1)\,\mathbb{P}(A_2) \cdots \mathbb{P}(A_n)\,.$$

Pri diskretnih slučajnih spremenljivkah pa je dovolj, da se množijo točkaste verjetnosti:

$$\mathbb{P}(X_1 = a_1, X_2 = a_2, \ldots, X_n = a_n) = \mathbb{P}(X_1 = a_1)\,\mathbb{P}(X_2 = a_2) \cdots \mathbb{P}(X_n = a_n)\,.$$

3.3 Bernoullijeva zaporedja poskusov

Bernoullijevo zaporedje poskusov je zaporedje slučajnih poskusov, pri čemer pri vsakem gledamo, ali *uspe* ali *ne uspe*. Pri tem morajo biti poskusi neodvisni (t. j. dogodki, da posamezen poskus uspe, morajo biti neodvisni) in vsak poskus mora uspeti z enako verjetnostjo, ki jo bomo označevali s p.

Primer. mečemo pošteno kocko, meti so neodvisni. Če je poskus met kocke, uspešen poskus pa je met, pri katerem pade šestica, dobimo Bernoullijevo zaporedje poskusov.

Recimo, da pod temi pogoji petkrat vržemo pošteno kocko.

- Verjetnost, da šestica pade v prvem metu, je $\dfrac{1}{6} \doteq 0{\cdot}167$.
- Verjetnost, da šestica pade v petem metu, je $\dfrac{1}{6} \doteq 0{\cdot}167$.
- Verjetnost, da šestica pade v prvem in petem metu, je $\dfrac{1}{6} \cdot \dfrac{1}{6} = \dfrac{1}{36} \doteq 0{\cdot}0278$.
- Verjetnost, da šestica pade vsaj enkrat v petih metih, ni $\dfrac{5}{6}$. Verjetnosti dogodkov, da v posameznem metu pade šestica, namreč ne smemo kar sešteti, ker ti dogodki niso paroma nezdružljivi. Pač pa si lahko pomagamo z nasprotnim dogodkom, ki je

dogodek, da šestica ne pade v nobenem metu. Njegova verjetnost je $\left(\frac{5}{6}\right)^5 \doteq 0{\cdot}402$.

Verjetnost dogodka, da šestica pade vsaj enkrat, pa je potemtakem $1 - \left(\frac{5}{6}\right)^5 \doteq 0{\cdot}598$.

- Verjetnost dogodka, da šestica pade v prvem in samo v prvem metu, je $\frac{1}{6} \cdot \left(\frac{5}{6}\right)^4 \doteq 0{\cdot}0804$.
- Verjetnost dogodka, da šestica pade natanko enkrat, je vsota verjetnosti dogodkov, da pade v posameznem metu, sicer pa ne. Te verjetnosti lahko seštevamo, ker so dogodki paroma nezdružljivi. Torej je verjetnost danega dogodka enaka $5 \cdot \frac{1}{6} \cdot \left(\frac{5}{6}\right)^4 \doteq 0{\cdot}402$.
- Verjetnost dogodka, da šestica pade natanko dvakrat, je vsota verjetnosti, da pade natanko v i-tem in j-tem metu, po vseh možnih neurejenih parih $\{i, j\}$, kjer sta i in j različna, t. j.:

$$\{1,2\}\,,\ \{1,3\}\,,\ \{1,4\}\,,\ \{1,5\}\,,\ \{2,3\}\,,\ \{2,4\}\,,\ \{2,5\}\,,\ \{3,4\}\,,\ \{3,5\}\,,\ \{4,5\}\,.$$

  Teh neurejenih parov je toliko, kot je vseh možnih izbir dveh elementov izmed petih, torej $\binom{5}{2} = 10$. Za vsak par $\{i, j\}$ je verjetnost, da pade šestica natanko v i-tem in j-tem metu, enaka $\left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^3 \doteq 0{\cdot}0161$. Iskana verjetnost je torej enaka $10 \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^3 \doteq 0{\cdot}161$.
- Verjetnost dogodka, da šestica pade manj kot trikrat, je:

$$\left(\frac{5}{6}\right)^5 + 5 \cdot \frac{1}{6} \cdot \left(\frac{5}{6}\right)^4 + 10 \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^3 \doteq 0{\cdot}965\,.$$

Primer: če kocko vržemo 10-krat in želimo izračunati verjetnost dogodka, da šestica pade natanko trikrat, moramo prešteti vse možne izbire treh metov, v katerih pade šestica. Teh izbir je $\binom{10}{3} = 120$, iskana verjetnost pa je enaka:

$$\binom{10}{3} \cdot \left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^7 \doteq 0{\cdot}155\,.$$

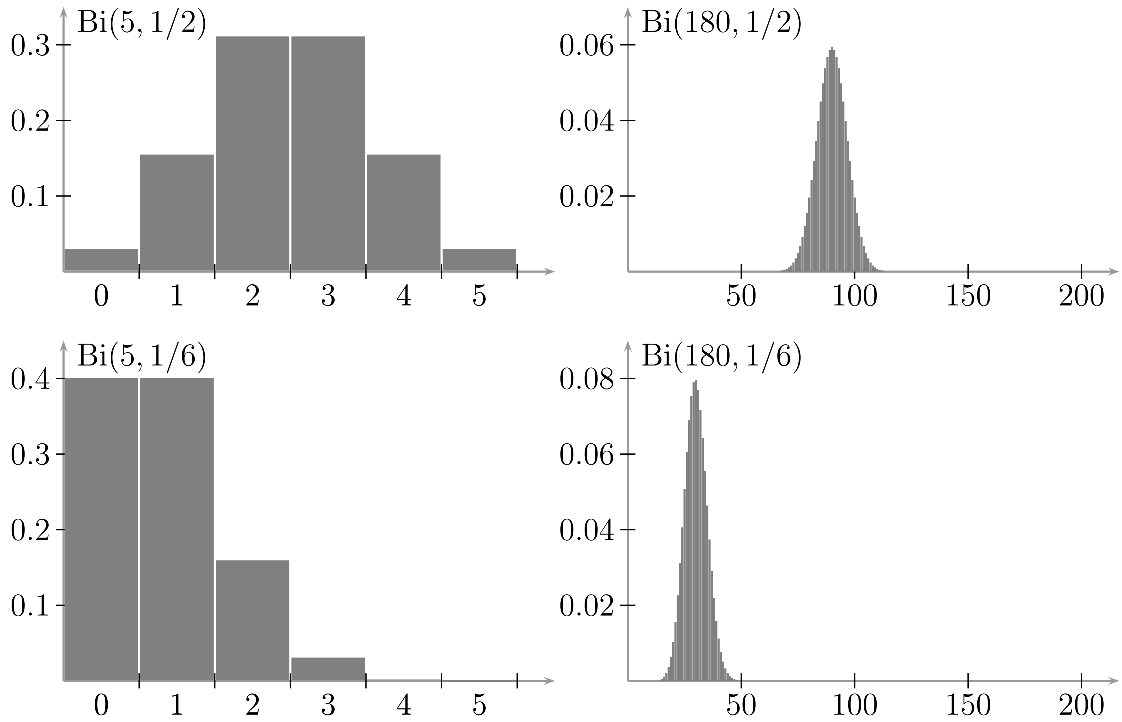
V splošnem lahko vsakemu Bernoullijevemu zaporedju n poskusov priredimo slučajno spremenljivko S, ki pove število uspelih poskusov. Za $k = 0, 1, \ldots, n$ velja:

$$\mathbb{P}(S = k) = \binom{n}{k} p^k (1-p)^{n-k}\,,$$

kjer je p verjetnost, da posamezen poskus uspe. Zgornji obrazec imenujemo *Bernoullijeva formula*, porazdelitvi tako definirane slučajne spremenljivke S pa pravimo *binomska porazdelitev*. Pišemo $S \sim \mathrm{Bi}(n,p)$. Ta zapis je po definiciji ekvivalenten veljavnosti zgornje formule za vse $k = 0, 1, \ldots, n$.

Če je torej $S \sim \mathrm{Bi}(5, 1/6)$, je npr. $\mathbb{P}(S = 2) \doteq 0{\cdot}161$, $\mathbb{P}(S < 3) \doteq 0{\cdot}965$, $\mathbb{P}(S = 5/2) = 0$ in $\mathbb{P}(S < 6) = 1$.

Če pa bi 5-krat neodvisno vrgli pošten kovanec in gledali število grbov, bi imelo le-to porazdelitev $\mathrm{Bi}(5, 1/2)$. Nekaj histogramov:



Za velike n je verjetnosti pri binomski porazdelitvi $\mathrm{Bi}(n,p)$ brez sodobnih pripomočkov težko računati. Zato so že zgodaj iznašli približne obrazce. Najširše uporabni sta Laplaceovi približni formuli.

Naj bo $S \sim \mathrm{Bi}(n,p)$. *Laplaceova lokalna formula* pravi:

$$\mathbb{P}(S = k) \approx \frac{1}{\sigma\sqrt{2\pi}} e^{-(k-np)^2/(2\sigma^2)} \,,$$

kjer je $\sigma = \sqrt{np(1-p)}$. Produkt np je *pričakovana vrednost*, σ pa je *standardni odklon* binomske porazdelitve. Natančnejšo definicijo bomo navedli v podrazdelku 3.6, v grobem pa lahko rečemo, da pri slučajni spremenljivki S lahko pričakujemo vrednosti blizu np, in sicer z odmiki reda velikosti σ. Slednja vrednost je tudi glavno merilo za natančnost obrazca: večji kot je σ, natančnejši je obrazec. Razumno natančnost dosežemo pri $\sigma \geq 5$. Toda to velja za napako glede na največjo točkasto verjetnost. Če pa natančnost merimo z relativno napako, mora biti še absolutna razlika $|k - np|$ majhna v primerjavi s $\sigma^{4/3}$ (razumno natančnost dosežemo pri $|k - np| \leq \sigma^{4/3}$).

Primer: 180-krat vržemo pošteno kocko, meti so neodvisni. Če z S označimo število šestic, je $S \sim \mathrm{Bi}(180, 1/6)$. Velja $\sigma = \sqrt{180 \cdot \frac{1}{6} \cdot \frac{5}{6}} = 5$. Nekaj primerov približnih izračunov:

- $\mathbb{P}(S = 30) \approx \frac{1}{5\sqrt{2\pi}}\, e^{0} \doteq 0{\cdot}07979$. Točen rezultat: $0{\cdot}07956$.
- $\mathbb{P}(S = 35) \approx \frac{1}{5\sqrt{2\pi}}\, e^{-1/2} \doteq 0{\cdot}04839$. Točen rezultat: $0{\cdot}04640$.
- $\mathbb{P}(S = 25) \approx \frac{1}{5\sqrt{2\pi}}\, e^{-1/2} \doteq 0{\cdot}04839$. Točen rezultat: $0{\cdot}05072$.
- $\mathbb{P}(S = 20) \approx \frac{1}{5\sqrt{2\pi}}\, e^{-2} \doteq 0{\cdot}01080$. Točen rezultat: $0{\cdot}01079$.
- $\mathbb{P}(S = 15) \approx \frac{1}{5\sqrt{2\pi}}\, e^{-9/2} \doteq 0{\cdot}00089$. Točen rezultat: $0{\cdot}00052$.

Pri zadnjem primeru je absolutna napaka še vedno majhna v primerjavi z maksimalno verjetnostjo $\mathbb{P}(S = 30)$, relativna napaka pa je velika: velja $|k - np| = 15$, medtem ko je $\sigma^{4/3} \doteq 8{\cdot}55$.

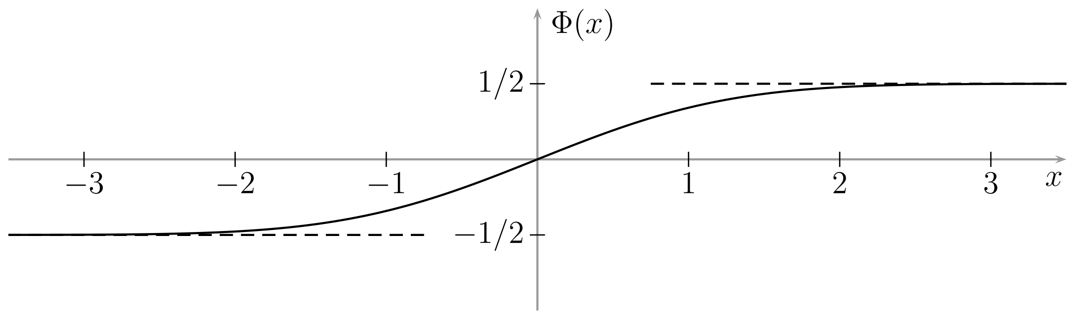
Laplaceova integralska formula pa pravi:

$$\mathbb{P}(a < S < b) \approx \mathbb{P}(a \le S \le b) \approx \Phi\left(\frac{b - np}{\sigma}\right) - \Phi\left(\frac{a - np}{\sigma}\right),$$

kjer je Φ *Gaussov verjetnostni integral*:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2}\,\mathrm{d}t\,.$$

To je *specialna* funkcija: je pomembna, ne da pa se izraziti z elementarnimi funkcijami. Njene vrednosti lahko odčitavamo iz tabel. Graf:



Na grafu so razvidne naslednje lastnosti funkcije Φ:

- $\Phi(0) = 0$;
- $\lim_{x\to\infty} \Phi(x) = \frac{1}{2}$;
- $\lim_{x\to-\infty} \Phi(x) = -\frac{1}{2}$;
- $\Phi(-x) = -\Phi(x)$ (lihost).

Pripomnimo naj, da definicije funkcije Φ v literaturi zelo variirajo, zato se mora bralec vselej prepričati, kako je ta funkcija v posameznem viru definirana.

Tudi za natančnost Laplaceove integralske formule je merilo standardni odklon σ: večji kot je, natančnejša je. Razumno natančnost spet dosežemo pri $\sigma \geq 5$. Pomembno pa je tudi, da večjo natančnost dosežemo, če sta a in b *na polovici* med dvema zaporednima celima številoma. To je skupaj s pogojem, da je vsaj ena od absolutnih razlik $|a - np|$ in $|b - np|$ majhna v primerjavi s $\sigma^{4/3}$, tudi jamstvo za majhno relativno napako (razumno natančnost spet dosežemo pri $|a - np| \leq \sigma^{4/3}$ ali $|b - np| \leq \sigma^{4/3}$).

Primer: če je spet $S \sim \mathrm{Bi}(180, 1/6)$, gremo lahko računat:

$$\mathbb{P}(25 \leq S \leq 35) \approx \Phi(1) - \Phi(-1) = 2\Phi(1) \doteq 0\text{'}6827\,.$$

Ker je $\{25 \leq S \leq 35\} = \{24 < S < 36\}$, lahko isto verjetnost približno računamo tudi kot:

$$\mathbb{P}(24 \leq S \leq 36) \approx \Phi(1\text{'}2) - \Phi(-1\text{'}2) = 2\Phi(1\text{'}2) \doteq 0\text{'}7699\,.$$

Najnatančnejša aproksimacija pa pride, če jo računamo kot:

$$\mathbb{P}(24\text{'}5 < S < 35\text{'}5) \approx \Phi(1\text{'}1) - \Phi(-1\text{'}1) = 2\Phi(1\text{'}1) \doteq 0\text{'}7287\,.$$

Točen rezultat: 0'7292.

Še en primer izračuna:

$$\mathbb{P}(S < 20) = \mathbb{P}(-\infty < S < 19\text{'}5) \approx \Phi(-2\text{'}1) - \Phi(-\infty) = \frac{1}{2} - \Phi(2\text{'}1) \doteq 0\text{'}0179\,.$$

Točen rezultat: 0'0142.

Za spodnjo mejo smo tukaj postavili $-\infty$. To je najenostavneje, čeprav bi se bralcu morda zdela naravnejša spodnja meja 0. Toda razlika je zanemarljiva, saj bi se $\Phi(\infty) = 0\text{'}5$ zamenjal s $\Phi(6) \doteq 0\text{'}499999999013$.

Laplaceova integralska formula pove, da je binomska porazdelitev za vslike σ približno *normalna*. Normalna ali tudi *Gaussova* porazdelitev s sredino μ in standardnim odklonom σ, kjer je $\mu \in \mathbb{R}$ in $\sigma > 0$, je določena z lastnostjo, da slučajna spremenljivka X, ki ima to porazdelitev, zadošča:

$$\mathbb{P}(a < X < b) = \mathbb{P}(a \leq X \leq b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)\,.$$

Pišemo $X \sim \mathrm{N}(\mu, \sigma)$. Tako lahko Laplaceovo integralsko formulo zapišemo kot približno ujemanje porazdelitev:

$$\mathrm{Bi}(n, p) \approx \mathrm{N}\big(np, \sqrt{np(1-p)}\big)\,.$$

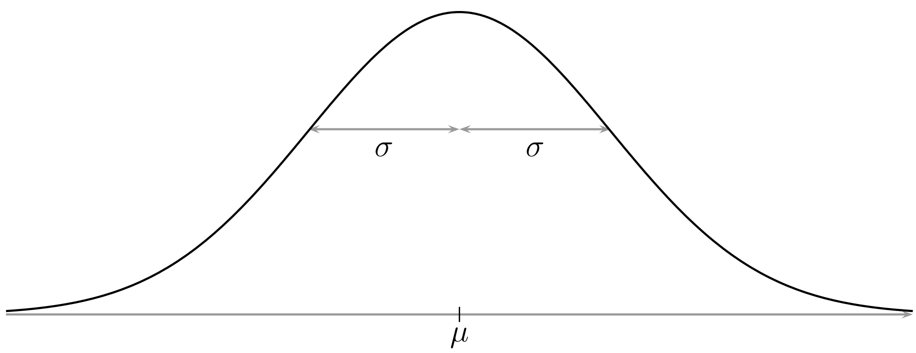
Za $\sigma = 0$ dobimo *degenerirano* normalno porazdelitev: če je $X \sim \mathrm{N}(\mu, 0)$, to pomeni, da je $X = \mu$ z verjetnostjo 1.

Standardna normalna porazdelitev pa je porazdelitev $\mathrm{N}(0, 1)$. Za $X \sim \mathrm{N}(0, 1)$ torej velja $\mathbb{P}(a < X < b) = \mathbb{P}(a \le X \le b) = \Phi(b) - \Phi(a)$.

Primer: če je $X \sim \mathrm{N}(50, 20)$, je:

- $\mathbb{P}(35 < X < 80) = \mathbb{P}(35 \le X < 80) = \mathbb{P}(25 < X \le 80) = \mathbb{P}(35 \le X \le 80) = \Phi(1\dot{}5) + \Phi(0\dot{}75) \doteq 0\dot{}7066$;
- $\mathbb{P}(X < 0) = \frac{1}{2} - \Phi(2\dot{}5) \doteq 0\dot{}0062$.

Histogram porazdelitve celoštevilske slučajne spremenljivke, katere porazdelitev je približno normalna, ima obliko *Gaussove krivulje*:



Slučajna spremenljivka, porazdeljena normalno $\mathrm{N}(\mu, \sigma)$, se:

- z verjetnostjo približno 68% nahaja med $\mu - \sigma$ do $\mu + \sigma$;
- z verjetnostjo približno 95% nahaja med $\mu - 2\sigma$ do $\mu + 2\sigma$;
- z verjetnostjo približno 99˙7% nahaja med $\mu - 3\sigma$ do $\mu + 3\sigma$.

3.4 Funkcije slučajnih spremenljivk

Iz slučajne spremenljivke X in funkcije f dobimo novo slučajno spremenljivko $f(X)$, ki izidu ω priredi vrednost $f(X(\omega))$. Če ima X diskretno porazdelitev, podano s shemo:

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_k \\ p_1 & p_2 & \cdots & p_k \end{pmatrix},$$

je tudi $f(X)$ diskretna, njeno porazdelitev pa določa shema:

$$\begin{pmatrix} f(a_1) & f(a_2) & \cdots & f(a_k) \\ p_1 & p_2 & \cdots & p_k \end{pmatrix}.$$

Toda tudi če so vrednosti $a_1, \ldots, a_k$ vse različne, vrednosti $f(a_1), \ldots, f(a_k)$ niso nujno različne. Če hočemo torej dobiti točkaste verjetnosti (verjetnostno funkcijo), moramo v splošnem seštevati.

Primer. Naj bo:

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 0{\cdot}05 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}35 \end{pmatrix}$$

in $Y = (X-2)^2$. Tedaj sicer lahko pišemo:

$$X \sim \begin{pmatrix} 4 & 1 & 0 & 1 & 4 & 9 \\ 0{\cdot}05 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}35 \end{pmatrix},$$

toda verjetnost $\mathbb{P}(Y = 4)$ ni enaka niti $0{\cdot}05$ niti $0{\cdot}15$, temveč $0{\cdot}05 + 0{\cdot}15 = 0{\cdot}2$. Porazdelitvena shema, iz katere lahko točkaste verjetnosti neposredno razberemo, je:

$$X \sim \begin{pmatrix} 0 & 1 & 4 & 9 \\ 0{\cdot}05 & 0{\cdot}3 & 0{\cdot}2 & 0{\cdot}35 \end{pmatrix}.$$

Lahko gledamo tudi funkcije več slučajnih spremenljivk.

Primer. Slučajni spremenljivki X in Y naj imata navzkrižno porazdelitev:

	$Y=1$	$Y=2$	$Y=3$	
$X=1$	0	0.25	0.15	0.4
$X=2$	0.25	0.25	0.1	0.6
	0.25	0.5	0.25	

Tedaj je njuna vsota $X + Y$ funkcija obeh slučajnih spremenljivk in velja:

$$X + Y \sim \begin{pmatrix} 3 & 4 & 5 \\ 0{\cdot}5 & 0{\cdot}4 & 0{\cdot}1 \end{pmatrix}.$$

Če pa vzamemo neodvisni slučajni spremenljivki X' in Y' z enakima porazdelitvama kot X in Y, t. j.:

	$Y'=1$	$Y'=2$	$Y'=3$	
$X'=1$	0.1	0.2	0.1	0.4
$X'=2$	0.15	0.3	0.15	0.6
	0.25	0.5	0.25	

je $X'+Y' \sim \begin{pmatrix} 2 & 3 & 4 & 5 \\ 0\text{˙}1 & 0\text{˙}35 & 0\text{˙}4 & 0\text{˙}15 \end{pmatrix}$, čeprav sta robni porazdelitvi v obeh primerih enaki.

Nauk: za izračun porazdelitve funkcije dveh ali več slučajnih spremenljivk potrebujemo *navzkrižno* porazdelitev.

3.5 Karakteristike urejenostnih slučajnih spremenljivk

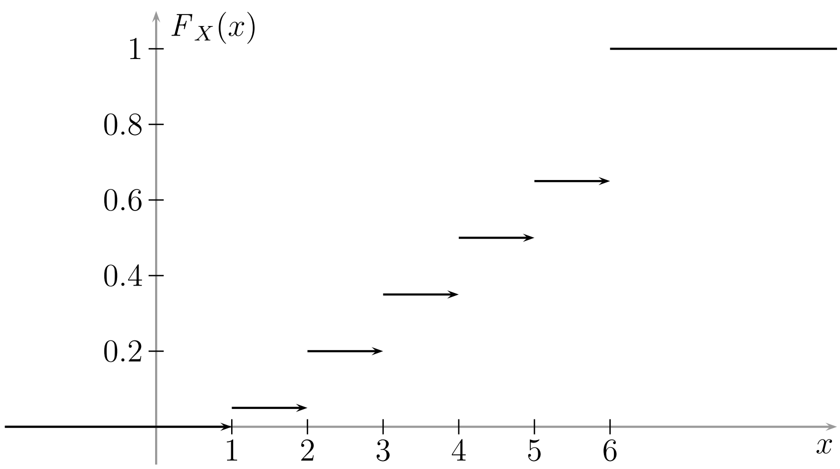
Eden temeljnih pojmov pri urejenostnih slučajnih spremenljivkah je *kumulativna porazdelitvena funkcija*:

$$F_X(x) = \mathbb{P}(X \le x)\,.$$

Primer: za $X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 0\text{˙}05 & 0\text{˙}15 & 0\text{˙}15 & 0\text{˙}15 & 0\text{˙}15 & 0\text{˙}35 \end{pmatrix}$ je npr.:

$$\begin{aligned} F_X(2) &= \mathbb{P}(X=1) + \mathbb{P}(X=2) = 0\text{˙}2\,, \\ F_X(1\text{˙}9) &= \mathbb{P}(X=1) = 0\text{˙}05\,, \\ F_X(0) &= 0\,, \\ F_X(6) &= 1\,, \\ F_X(1000) &= 1\,. \end{aligned}$$

Graf:



Če urejenostna slučajna spremenljivka X zavzame vrednosti v množici, ki se da urejenostno vložiti v realna števila, je s kumulativno porazdelitveno funkcijo njena porazdelitev natančno določena. Velja namreč:

$$\mathbb{P}(X < x) = F_X(x^-) = \lim_{y \uparrow x} F_X(y)$$

ter za $a \le b$ še:

$$\begin{aligned}
\mathbb{P}(a \le X \le b) &= \mathbb{P}(X \le b) - \mathbb{P}(X < a) = F_X(b) - F_X(a^-)\,, \\
\mathbb{P}(a < X \le b) &= \mathbb{P}(X \le b) - \mathbb{P}(X \le a) = F_X(b) - F_X(a)\,, \\
\mathbb{P}(a \le X < b) &= \mathbb{P}(X < b) - \mathbb{P}(X < a) = F_X(b^-) - F_X(a^-)\,, \\
\mathbb{P}(a < X < b) &= \mathbb{P}(X < b) - \mathbb{P}(X \le a) = F_X(b^-) - F_X(a)\,.
\end{aligned}$$

Velja tudi:

$$\mathbb{P}(X = x) = \mathbb{P}(X \le x) - \mathbb{P}(X < x) = F_X(x) - F(x^-)\,.$$

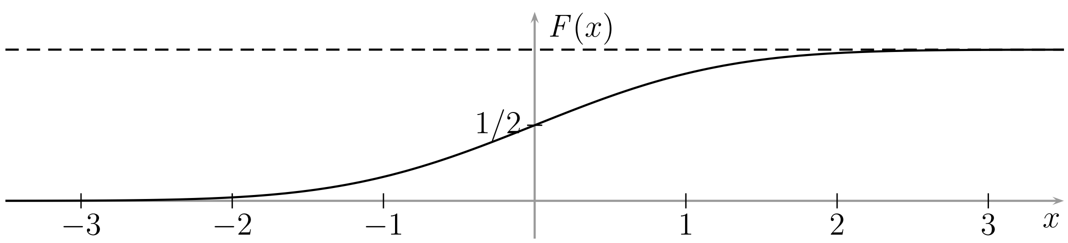
Točkasta verjetnost je torej enaka *skoku* kumulativne porazdelitvene funkcije v dani točki.

Kumulativna porazdelitvena funkcija F ima naslednje lastnosti:

- Je naraščajoča, a ne nujno strogo: če je $a \le b$, je $F(a) \le F(b)$.
- $\lim_{x \to -\infty} F(x) = 0$ in $\lim_{x \to \infty} F(x) = 1$.
- Je z desne zvezna: $\lim_{y \downarrow x} F(y) = F(x^+) = F(x)$.

Velja tudi, da je vsaka funkcija z zgornjimi lastnostmi kumulativna porazdelitvena funkcija neke slučajne spremenljivke.

Kumulativna porazdelitvena funkcija slučajne spremenljivke, ki zavzame kvečjemu končno mnogo vrednosti, je stopničasta. To pa ni res za vsako kumulativno porazdelitveno funkcijo. Kumulativna porazdelitvena funkcija standardne normalne porazdelitve je tako $F(x) = \frac{1}{2} + \Phi(x)$. Graf:



1−

Kvantil slučajne spremenljivke za verjetnost p je vrednost q, za katero velja:

$$\mathbb{P}(X < q) \le p \quad \text{in} \quad \mathbb{P}(X \le q) \ge p\,.$$

To je posplošitev kvantila spremenljivke na statistični množici: slednji je kvantil porazdelitve slučajne spremenljivke, ki jo dobimo tako, da enoto na populaciji izberemo na slepo.

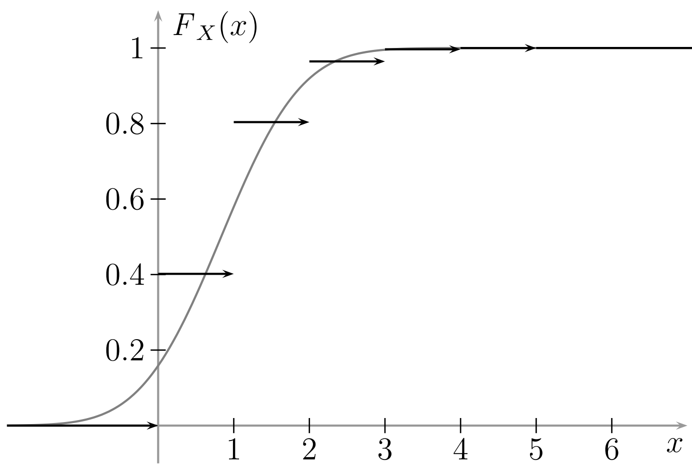
Kvantile lahko grafično določimo iz grafa kumulativne porazdelitvene funkcije: v skokih graf dopolnimo z ustreznimi navpičnimi daljicami in kvantil za verjetnost p je vsaka abscisa točke na dobljeni črti, ki ima ordinato enako p. Gledamo torej presečišča dopolnjenega grafa z ustrezno vodoravnico.

Primer. Če je $X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 0\text{'}05 & 0\text{'}15 & 0\text{'}15 & 0\text{'}15 & 0\text{'}15 & 0\text{'}35 \end{pmatrix}$, je kvantil slučajne spremenljivke X za verjetnost $0\text{'}3$ natančno določen, in sicer je enak 3:

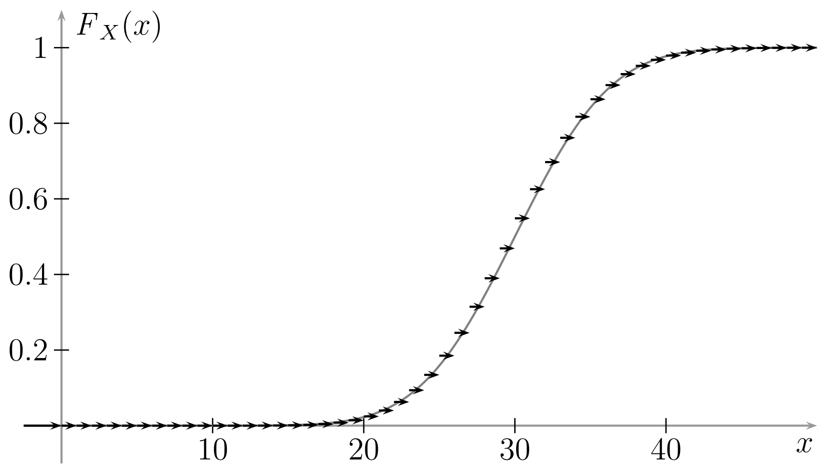
- Velja $\mathbb{P}(X < 3) = 0\text{'}2 \le 0\text{'}3$ in $\mathbb{P}(X \le 3) = 0\text{'}35 \ge 0\text{'}3$.
- Brž ko je $x < 3$, je $\mathbb{P}(X \le x) \le 0\text{'}2 < 0\text{'}3$, zato x ni ustrezni kvantil.
- Brž ko je $x > 3$, je $\mathbb{P}(X < x) \ge 0\text{'}35 > 0\text{'}3$, zato x ni ustrezni kvantil.

Podobno je tudi prvi kvartil lahko enak le 3. Mediana pa ni natančno določena, lahko je kar koli iz intervala $[4, 5]$.

Primer: kumulativna porazdelitvena funkcija števila šestic pri 5 metih poštene kocke skupaj s kumulativno porazdelitveno funkcijo ustrezne aproksimativne normalne porazdelitve:



Primer: kumulativna porazdelitvena funkcija števila šestic pri 180 metih poštene kocke skupaj s kumulativno porazdelitveno funkcijo ustrezne aproksimativne normalne porazdelitve:



3.6 Karakteristike intervalskih slučajnih spremenljivk

Eden temeljnih pojmov pri intervalskih slučajnih spremenljivkah je *matematično upanje* ali tudi *pričakovana vrednost*. Le-to sicer ni vselej definirano, lahko pa ga izračunamo za vsako slučajno spremenljivko, ki zavzame le končno mnogo vrednosti. Matematično upanje slučajne spremenljivke:

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

je definirano kot:

$$\mathbb{E}(X) := a_1 p_1 + a_2 p_2 + \cdots + a_k p_k \,.$$

Pri tem ni nujno, da so vse vrednosti $a_1, a_2, \ldots, a_k$ različne. Izkaže se, da je matematično upanje neodvisno od sheme, s katero podamo porazdelitev.

Matematično upanje je posplošitev aritmetične sredine: aritmetična sredina spremenljivke, definirane na določeni statistični množici, je matematično slučajne spremenljivke, ki nastane, če enoto iz množice izberemo na slepo.

Matematično upanje je mera centralne tendence: v grobem lahko rečemo, da lahko pri realizacijah slučajne spremenljivke X pričakujemo vrednosti blizu $\mathbb{E}(X)$.

Primer: za $X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 0\text{'}1 & 0\text{'}15 & 0\text{'}15 & 0\text{'}15 & 0\text{'}15 & 0\text{'}3 \end{pmatrix}$ je:

$$\mathbb{E}(X) = 1 \cdot 0\text{'}1 + 2 \cdot 0\text{'}15 + 3 \cdot 0\text{'}15 + 4 \cdot 0\text{'}15 + 5 \cdot 0\text{'}15 + 6 \cdot 0\text{'}3 = 4 \,.$$

Neposredno iz porazdelitvene sheme se da izračunati tudi matematično upanje poljubne funkcije slučajne spremenljivke. Velja namreč:

$$\mathbb{E}\big[f(X)\big] = f(a_1) p_1 + f(a_2) p_2 + \cdots + f(a_k) p_k \,.$$

Primer. Naj bo spet $X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 0\text{'}1 & 0\text{'}15 & 0\text{'}15 & 0\text{'}15 & 0\text{'}15 & 0\text{'}3 \end{pmatrix}$ in $Y = (X-4)^2$. Tedaj lahko $\mathbb{E}(Y)$ izračunamo na dva načina. Lahko upoštevamo, da je
$Y \sim \begin{pmatrix} 0 & 1 & 4 & 9 \\ 0\text{'}15 & 0\text{'}3 & 0\text{'}45 & 0\text{'}1 \end{pmatrix}$ in zato:

$$\mathbb{E}(Y) = 0 \cdot 0\text{'}15 + 1 \cdot 0\text{'}3 + 4 \cdot 0\text{'}45 + 9 \cdot 0\text{'}1 = 3$$

ali pa:

$$\begin{aligned} \mathbb{E}(Y) &= (1-4)^2 \cdot 0\text{'}1 + (2-4)^2 \cdot 0\text{'}15 + (3-4)^2 \cdot 0\text{'}15 + (4-4)^2 \cdot 0\text{'}15 + \\ &\qquad + (5-4)^2 \cdot 0\text{'}15 + (6-4)^2 \cdot 0\text{'}3 = \\ &= 3\,. \end{aligned}$$

Matematično upanje lahko izmenjamo z linearno funkcijo. Velja namreč:

$$\mathbb{E}(kX + n) = k\,\mathbb{E}(X) + n\,.$$

To pa ne velja kar za vsako funkcijo: v splošnem je $\mathbb{E}\big[f(X)\big] \neq f\big(\mathbb{E}(X)\big)$.

Primer. Naj bo spet $X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 0\text{'}1 & 0\text{'}15 & 0\text{'}15 & 0\text{'}15 & 0\text{'}15 & 0\text{'}3 \end{pmatrix}$ in $Z = 10X + 1000$. Tedaj lahko $\mathbb{E}(Z)$ spet izračunamo na dva načina. Lahko upoštevamo, da je
$Z \sim \begin{pmatrix} 1010 & 1020 & 1030 & 1040 & 1050 & 1060 \\ 0\text{'}1 & 0\text{'}15 & 0\text{'}15 & 0\text{'}15 & 0\text{'}15 & 0\text{'}3 \end{pmatrix}$ in zato:

$$\mathbb{E}(Z) = 1010 \cdot 0\text{'}1 + 1020 \cdot 0\text{'}15 + 1030 \cdot 0\text{'}15 + 1040 \cdot 0\text{'}15 + 1050 \cdot 0\text{'}15 + 1060 \cdot 0\text{'}3 = 1040$$

ali pa preprosto izračunamo $\mathbb{E}(Z) = 10 \cdot 4 + 1000 = 1040$.

Če pa postavimo $Y = (X-4)^2$, smo v prejšnjem primeru izračunali, da je $\mathbb{E}(Y) = 3$, medtem ko je $(4-4)^2 = 0$.

Izmenljivost matematičnega upanja je temelj za *u-metodo*, ki jo poznamo že iz računanja aritmetične sredine: za poljuben $u \in \mathbb{R}$ velja:

$$\mathbb{E}(X) = u + \mathbb{E}(X - u)\,.$$

Primer: za $X \sim \begin{pmatrix} 490 & 500 & 520 \\ 0\text{'}1 & 0\text{'}3 & 0\text{'}6 \end{pmatrix}$ lahko izračunamo:

$$\mathbb{E}(X) = 500 + (-10) \cdot 0\text{'}1 + 20 \cdot 0\text{'}6 = 511\,.$$

Matematično upanje je tudi *aditivno*: za slučajni spremenljivki velja:

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)\,.$$

Matematično upanje vsote je torej določeno že z robnima porazdelitvama, čeprav to ne velja za samo porazdelitev vsote.

Primer. Slučajni spremenljivki X in Y naj imata navzkrižno porazdelitev:

	$Y=1$	$Y=2$	$Y=3$	
$X=1$	0	0.25	0.15	0.4
$X=2$	0.25	0.25	0.1	0.6
	0.25	0.5	0.25	

Iz robnih porazdelitev dobimo:

$$\mathbb{E}(X) = 1 \cdot 0{.}4 + 2 \cdot 0{.}6 = 1{.}6\,,$$
$$\mathbb{E}(Y) = 1 \cdot 0{.}25 + 2 \cdot 0{.}5 + 3 \cdot 0{.}25 = 2\,.$$

Nadalje je $X + Y \sim \begin{pmatrix} 3 & 4 & 5 \\ 0{.}5 & 0{.}4 & 0{.}1 \end{pmatrix}$ in posledično:

$$\mathbb{E}(X+Y) = 3 \cdot 0{.}5 + 4 \cdot 0{.}4 + 5 \cdot 0{.}1 = 3{.}6 = 1{.}6 + 2\,.$$

Če pa vzamemo neodvisni slučajni spremenljivki X' in Y' z enakima porazdelitvama kot X in Y, t. j.:

	$Y'=1$	$Y'=2$	$Y'=3$	
$X'=1$	0.1	0.2	0.1	0.4
$X'=2$	0.15	0.3	0.15	0.6
	0.25	0.5	0.25	

je seveda prav tako $\mathbb{E}(X') = 1{.}6$ in $\mathbb{E}(Y') = 2$. Vsota je sicer drugače porazdeljena: $X' + Y' \sim \begin{pmatrix} 2 & 3 & 4 & 5 \\ 0{.}1 & 0{.}35 & 0{.}4 & 0{.}15 \end{pmatrix}$, toda spet je:

$$\mathbb{E}(X'+Y') = 2 \cdot 0{.}1 + 3 \cdot 0{.}35 + 4 \cdot 0{.}4 + 5 \cdot 0{.}15 = 3{.}6 = 1{.}6 + 2\,.$$

Zaenkrat smo matematično upanje definirali le za slučajne spremenljivke, ki zavzamejo končno mnogo vrednosti. Da se ga smiselno posplošiti na veliko slučajnih spremenljivk, ne pa na vse. Omenimo naj, da za $X \sim \mathrm{N}(\mu, \sigma)$ velja $\mathbb{E}(X) = \mu$.

Za intervalske spremenljivke na statističnih množicah smo definirali tudi varianco. Varianca slučajne spremenljivke je definirana kot:

$$\mathrm{var}(X) = \mathbb{E}\Big[\big(X - \mathbb{E}(X)\big)^2\Big]\,.$$

Izkaže pa se, da velja tudi:

$$\mathrm{var}(X) = \mathbb{E}(X^2) - \big(\mathbb{E}(X)\big)^2\,,$$

kar često pride prav pri računanju peš. Uporaba te formule za računalnik pa je lahko nevarna, ker je formula lahko *numerično nestabilna*: majhne napake pri računanju vodijo v velike napake pri končnem rezultatu. To je zato, ker se lahko zgodi, da se dve veliki količini odštejeta v majhno količino.

Primer. Varianco slučajne spremenljivke $X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 0{\cdot}1 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}3 \end{pmatrix}$ lahko izračunamo na vsaj dva načina. Prvič, ker vemo, da je $\mathbb{E}(X) = 4$, je $\operatorname{var}(X) = \mathbb{E}\big[(X - 4)^2\big] = 3$. Drugič, lahko izračunamo:

$$\mathbb{E}(X^2) = 1 \cdot 0{\cdot}1 + 4 \cdot 0{\cdot}15 + 9 \cdot 0{\cdot}15 + 16 \cdot 0{\cdot}15 + 25 \cdot 0{\cdot}15 + 36 \cdot 0{\cdot}3 = 19$$

in posledično $\operatorname{var}(X) = 19 - 16 = 3$.

Varianca je mera razpršenosti: v grobem pove, za koliko se realizirane vrednosti slučajne spremenljivke X lahko razlikujejo. Natančneje, je mera za *kvadrate* teh razlik. Za varianco velja:

$$\operatorname{var}(aX + b) = a^2 \operatorname{var}(X) + b\,.$$

Primer: za $X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 0{\cdot}1 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}15 & 0{\cdot}3 \end{pmatrix}$ in $Z = 10X + 1000$ velja $\operatorname{var}(Z) = 100 \operatorname{var}(X) = 300$. Opazimo še, da bi se, če bi šli računat $\operatorname{var}(Z) = \mathbb{E}(Z^2) - \big(\mathbb{E}(Z)\big)^2$, odšteli dve količini z vrednostjo približno milijon.

Ker varianca ni neposredno primerljiva z vrednostmi slučajne spremenljivke, je spet ugodno uvesti *standardni odklon*:

$$\sigma(X) = \sqrt{\operatorname{var}(X)}\,.$$

Velja $\sigma(aX + b) = |a|\sigma(X)$. Za slučajni spremenljivki iz prejšnjega primera je tako $\sigma(X) = \sqrt{3} \doteq 1{\cdot}732$ in $\sigma(Z) = \sqrt{300} = 10\sqrt{3} \doteq 17{\cdot}32$.

Standardni odklon slučajne spremenljivke, porazdeljene normalno $\mathrm{N}(\mu, \sigma)$, je enak σ. Tako je uporaba znaka σ za drugi parameter pri normalni porazdelitvi in za standardni odklon konsistentna.

Neobčutljivost variance za prištetje deterministične količine vodi v izpeljavo u-metode za varianco:

$$\operatorname{var}(X) = \mathbb{E}\big[(X - u)^2\big] - \big(\mathbb{E}(X - u)\big)^2\,.$$

Primer. Naj bo $Z \sim \begin{pmatrix} 3 & 4 & 5 \\ 0{\cdot}5 & 0{\cdot}4 & 0{\cdot}1 \end{pmatrix}$. Če izberemo $u = 4$, dobimo:

$$\begin{aligned} \mathbb{E}(Z - u) &= (-1) \cdot 0{\cdot}5 + 1 \cdot 0{\cdot}1 = -0{\cdot}4 \\ \mathbb{E}\big[(Z - u)^2\big] &= (-1)^2 \cdot 0{\cdot}5 + 1^2 \cdot 0{\cdot}1 = 0{\cdot}6 \\ \operatorname{var}(Z) &= 0{\cdot}6 - 0{\cdot}16 = 0{\cdot}44\,. \end{aligned}$$

Varianca je tudi aditivna, vendar ne v tolikšni splošnosti kot matematično upanje. Enakost:

$$\operatorname{var}(X + Y) = \operatorname{var}(X) + \operatorname{var}(Y)$$

velja, če sta X in Y neodvisni, v splošnem pa ne. Tako recimo račun:

$$\operatorname{var}(2X) = \operatorname{var}(X + X) = \operatorname{var}(X) + \operatorname{var}(X) = 2\operatorname{var}(X)$$

ni pravilen, ker ne velja druga enakost: X in X namreč nista neodvisni, brž ko X ni skoraj gotovo konstantna. Pravilno je $\operatorname{var}(2X) = 4\operatorname{var}(X)$.

Primer. Spet vzemimo slučajni spremenljivki X in Y z navzkrižno porazdelitvijo:

	$Y = 1$	$Y = 2$	$Y = 3$	
$X = 1$	0	0.25	0.15	0.4
$X = 2$	0.25	0.25	0.1	0.6
	0.25	0.5	0.25	

pri kateri pride $X+Y \sim \begin{pmatrix} 3 & 4 & 5 \\ 0\text{˙}5 & 0\text{˙}4 & 0\text{˙}1 \end{pmatrix}$. Velja $\operatorname{var}(X) = 0\text{˙}24$, $\operatorname{var}(Y) = 0\text{˙}5$ in $\operatorname{var}(X + Y) = 0\text{˙}44 \neq \operatorname{var}(X) + \operatorname{var}(Y)$.

$$\mathbb{E}(X + Y) = 3 \cdot 0\text{˙}5 + 4 \cdot 0\text{˙}4 + 5 \cdot 0\text{˙}1 = 3\text{˙}6 = 1\text{˙}6 + 2\,.$$

Če pa vzamemo neodvisni slučajni spremenljivki X' in Y' z enakima porazdelitvama kot X in Y, t. j.:

	$Y' = 1$	$Y' = 2$	$Y' = 3$	
$X' = 1$	0.1	0.2	0.1	0.4
$X' = 2$	0.15	0.3	0.15	0.6
	0.25	0.5	0.25	

je seveda prav tako $\operatorname{var}(X') = 0\text{˙}24$ in $\operatorname{var}(Y') = 0.5$ in iz $X'+Y' \sim \begin{pmatrix} 2 & 3 & 4 & 5 \\ 0\text{˙}1 & 0\text{˙}35 & 0\text{˙}4 & 0\text{˙}15 \end{pmatrix}$ dobimo $\operatorname{var}(X'+Y') = 0\text{˙}74$. Tega pa ne bi bilo treba računati, saj sta X' in Y' neodvisni, zato je $\operatorname{var}(X' + Y') = 0\text{˙}24 + 0\text{˙}5 = 0\text{˙}74$.

Primer. Bernoullijevemu zaporedju slučajnih poskusov, pri katerih vsak uspe z verjetnostjo p, lahko priredimo zaporedje slučajnih spremenljivk $X_1, X_2, \ldots$ na naslednji način:

$$X_i = \begin{cases} 1 & ;\ i\text{-ti poskus uspe} \\ 0 & ;\ i\text{-ti poskus ne uspe} \end{cases}.$$

Tedaj so te slučajne spremenljivke neodvisne, vsota prvih n, $S = X_1 + X_2 + \cdots + X_n$, pa je natančno število uspelih med prvimi n poskusi, torej je $S \sim \mathrm{Bi}(n, p)$. Ni težko izračunati, da je $\mathbb{E}(X_i) = p$ in $\operatorname{var}(X_i) = p(1 - p)$. Sledi $\mathbb{E}(S) = np$ in $\operatorname{var}(S) = np(1 - p)$.

Sklep: za *vsako* slučajno spremenljivko $S \sim \mathrm{Bi}(n, p)$ je $\mathbb{E}(S) = np$ in $\operatorname{var}(S) = np(1 - p)$. To je zato, ker sta matematično upanje in varianca odvisna le od porazdelitve sučajne spremenljivke, ne pa tudi od njene konstrukcije.

Zdaj lahko Laplaceovo integralsko formulo pogledamo še drugače: če gre n proti neskončno, p pa ostane stran od 0 in 1, je binomska porazdelitev približno normalna z ustreznim matematičnim upanjem in varianco.

Kovarianca slučajnih spremenljivk X in Y je definirana kot:

$$K(X,Y) = \mathbb{E}\Big[\big(X - \mathbb{E}(X)\big)\big(Y - \mathbb{E}(Y)\big)\Big] .$$

Tudi kovarianca je posplošitev kovariance spremenljivk na statistični množici. Velja tudi:

$$K(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y)$$

in celo:

$$K(X,Y) = \mathbb{E}\big[(X-u)(Y-v)\big] - \mathbb{E}(X-u)\,\mathbb{E}(Y-v) .$$

Primer. Za slučajni spremenljivki X in Y z navzkrižno porazdelitvijo:

	$Y=1$	$Y=2$	$Y=3$	
$X=1$	0	0.25	0.15	0.4
$X=2$	0.25	0.25	0.1	0.6
	0.25	0.5	0.25	

velja:

$$\mathbb{E}(XY) = 1\cdot 1\cdot 0 + 1\cdot 2\cdot 0\text{'}25 + 1\cdot 3\cdot 0\text{'}15 + 2\cdot 1\cdot 0\text{'}25 + 2\cdot 2\cdot 0\text{'}25 + 2\cdot 3\cdot 0\text{'}1 = 3\text{'}05 ,$$

torej:

$$K(X,Y) = 3\text{'}05 - 1\text{'}6\cdot 2 = -0\text{'}15 .$$

Za slučajni spremenljivki s kovarianco nič pravimo, da sta *nekorelirani.* Slučajni spremenljivki X in Y sta torej nekorelirani natanko tedaj, ko je $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$. Poljubni neodvisni slučajni spremenljivki sta nekorelirani, obratno pa ne velja nujno.

Primer: pri navzkrižni porazdelitvi:

	$Y=-1$	$Y=0$	$Y=1$	
$X=-1$	0.1	0	0.1	0.2
$X=0$	0	0.6	0	0.6
$X=1$	0.1	0	0.1	0.2
	0.2	0.6	0.2	

sta X in Y nekorelirani, saj je $\mathbb{E}(X) = \mathbb{E}(Y) = \mathbb{E}(XY) = 0$. Nista pa neodvisni, saj je npr. $\mathbb{P}(X=-1, Y=0) = 0$, medtem ko je $\mathbb{P}(X=-1)\,\mathbb{P}(Y=0) \doteq 0\text{'}12$.

A že če sta X in Y nekorelirani, velja $\operatorname{var}(X+Y) = \operatorname{var}(X) + \operatorname{var}(Y)$. V splošnem namreč velja formula:

$$\operatorname{var}(X+Y) = \operatorname{var}(X) + 2\,K(X,Y) + \operatorname{var}(Y)\,.$$

Korelacijski koeficient med slučajnima spremenljivkama X in Y je tako kot za spremenljivke na statističnih množicah definiran s predpisom:

$$r(X,Y) = \frac{K(X,Y)}{\sigma(X)\,\sigma(Y)}\,.$$

Tudi pri slučajnih spremenljivkah velja $-1 \le r(X,Y) \le 1$ in korelacijski koeficient lahko kvalitativno opredeljujemo na enak način kot pri statističnih spremenljivkah.

Primer. Za slučajni spremenljivki X in Y z navzkrižno porazdelitvijo:

	$Y=1$	$Y=2$	$Y=3$	
$X=1$	0	0.25	0.15	0.4
$X=2$	0.25	0.25	0.1	0.6
	0.25	0.5	0.25	

velja:

$$r(X,Y) = \frac{-0\dot{}15}{\sqrt{0\dot{}24}\,\sqrt{0\dot{}5}} \doteq -0\dot{}433\,.$$

Gre torej za zmerno negativno korelacijo.

3.7 Centralni limitni izrek

Pri binomski porazdelitvi $\mathrm{Bi}(n,p)$ smo videli, da je, ko je n velik, p pa dovolj stran od 0 ali 1 blizu normalni (Gaussovi) z ustreznim matematičnim upanjem in varianco. Videli smo tudi, da se da binomska porazdelitev zapisati kot vsota neodvisnih slučajnih spremenljivk. To velja tudi za splošnejše vsote.

Centralni limitni izrek (ohlapna formulacija): slučajna spremenljivka S, ki je vsota veliko neodvisnih slučajnih spremenljivk z dovolj lepimi porazdelitvami, od katerih nobena posebej ne izstopa, je porazdeljena približno normalno $\mathrm{N}(\mu,\sigma)$, kjer je $\mu = \mathbb{E}(S)$ in $\sigma^2 = \operatorname{var}(S)$. Velja torej:

$$\mathbb{P}(a < S < b) \approx \mathbb{P}(a \le S \le b) \approx \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

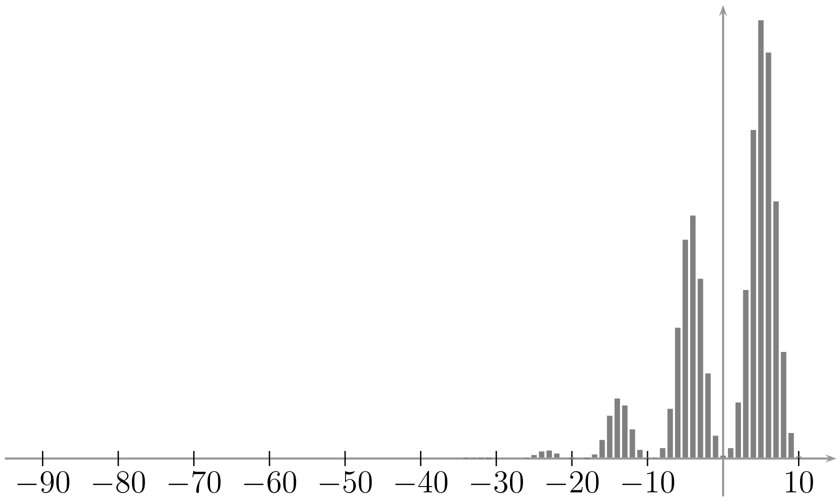
Z drugimi besedami, porazdelitev količine, ki nastopi kot rezultat veliko majhnih neodvisnih slučajnih vplivov, ki se med seboj seštevajo, sledi Gaussovi krivulji.

Primer. Loterija izda srečko, ki stane 1 evro. Srečka z verjetnostjo 5% zadene 10 evrov, z verjetnostjo 45% zadene 1 evro (v tem primeru torej kupec dobi povrnjen vložek), z verjetnostjo 50% pa ne zadene nič. Zanima nas dobiček loterije, če proda določeno število neodvisnih srečk (to vsaj približno velja, če igralec srečko napiše sam).

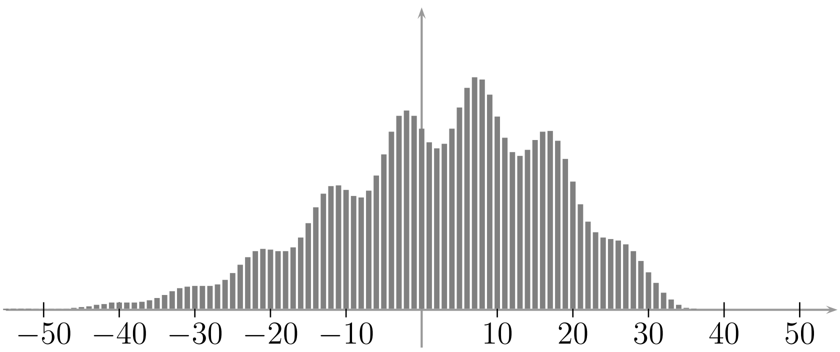
Pri n prodanih srečkah lahko dobiček zapišemo kot vsoto $S = X_1 + X_2 + \cdots + X_n$, kjer so $X_1, \ldots, X_n$ neodvisne s porazdelitvijo:

$$\begin{pmatrix} -9 & 0 & 1 \\ 0\text{'}05 & 0\text{'}45 & 0\text{'}5 \end{pmatrix} .$$

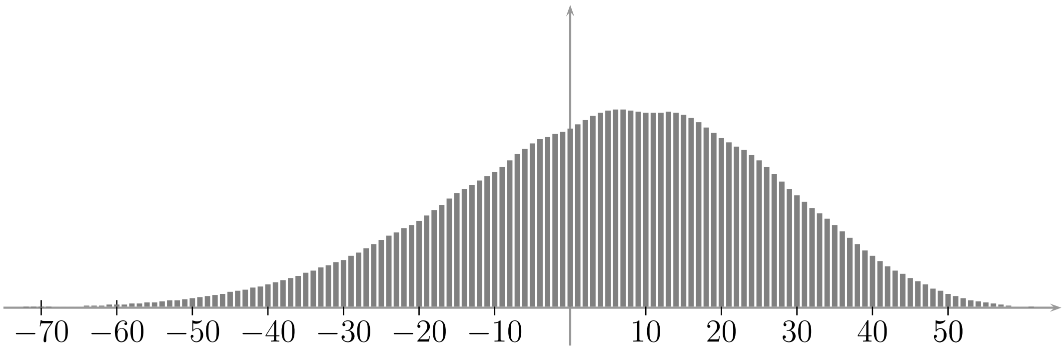
Če loterija proda 10 srečk, porazdelitev dobička S prikazuje naslednji histogram:



Če loterija proda 50 srečk, je histogram naslednji:



Pri 100 prodanih srečkah pa je naslednji:



Izračunajmo za ta primer približno verjetnost, da ima loterija izgubo! Velja:

$$\mathbb{E}(X_i) = 0\text{˙}05\,, \quad \mathrm{var}(X_i) = 4\text{˙}5475\,, \quad \sigma(X_i) \doteq 2\text{˙}1325\,,$$

torej:

$$\mathbb{E}(S) = 5\,, \quad \mathrm{var}(S) = 454\text{˙}75\,, \quad \sigma(S) \doteq 21\text{˙}325\,,$$

od koder dobimo:

$$\mathbb{P}(S < 0) = \mathbb{P}(-\infty < S < -0\text{˙}5) \approx \Phi\left(\frac{-0\text{˙}5 - 5}{21\text{˙}325}\right) - \Phi(-\infty) \doteq \frac{1}{2} - \Phi(0\text{˙}2579) \doteq 0\text{˙}3982\,.$$

Točen rezultat: 0˙376765.

Lahko pa se vprašamo tudi, najmanj koliko srečk mora loterija prodati, če naj bo verjetnost, da ima izgubo, največ 5%. Če loterija proda n srečk, je $\mathbb{E}(S) = 0\text{˙}05\,n$ in $\sigma(S) = \sqrt{4\text{˙}5475\,n}$, torej je:

$$\mathbb{P}(S < 0) \approx \Phi\left(\frac{-0\text{˙}5 - 0\text{˙}05\,n}{\sqrt{4\text{˙}5475\,n}}\right) + \frac{1}{2}\,.$$

Izraz v ulomku pa je za nadaljnjo obravnavo nekoliko zapleten, zato si privoščimo, da zanemarimo tudi člen 0˙5. Dobimo:

$$\mathbb{P}(S < 0) \approx \frac{1}{2} - \Phi\left(0\text{˙}05\sqrt{\frac{n}{4\text{˙}5475}}\right)\,.$$

Rešiti moramo torej neenačbo:

$$\frac{1}{2} - \Phi\left(0\text{˙}05\sqrt{\frac{n}{4\text{˙}5475}}\right) \le 0\text{˙}05$$

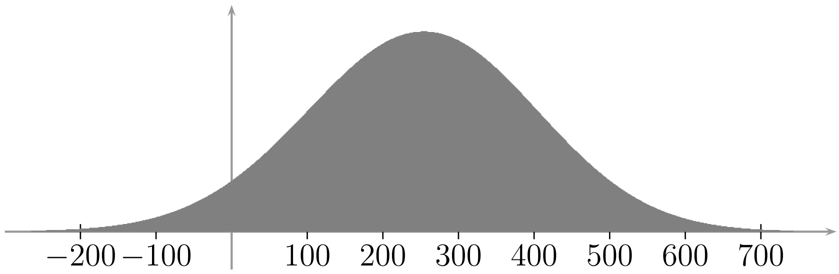
oziroma:

$$\Phi\left(0\text{˙}05\sqrt{\frac{n}{4\text{˙}5475}}\right) \ge 0\text{˙}45\,.$$

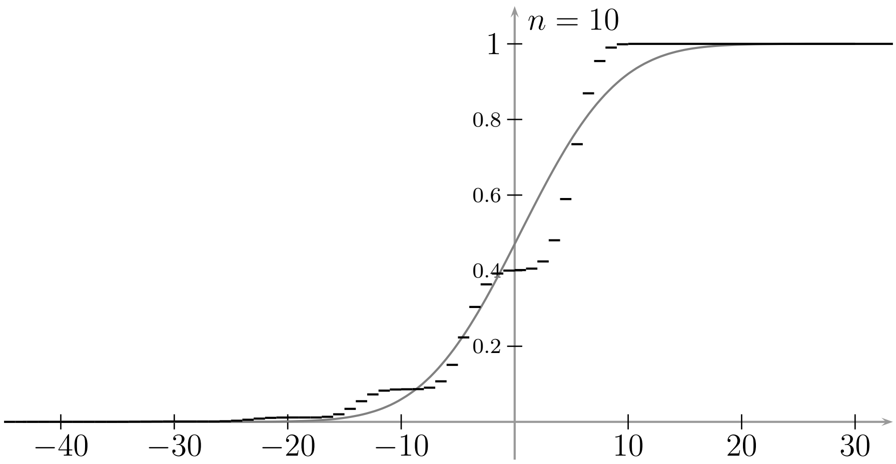
Ker je $\Phi(1\text{˙}645) \doteq 0\text{˙}45$, je pogoj za število srečk približno:

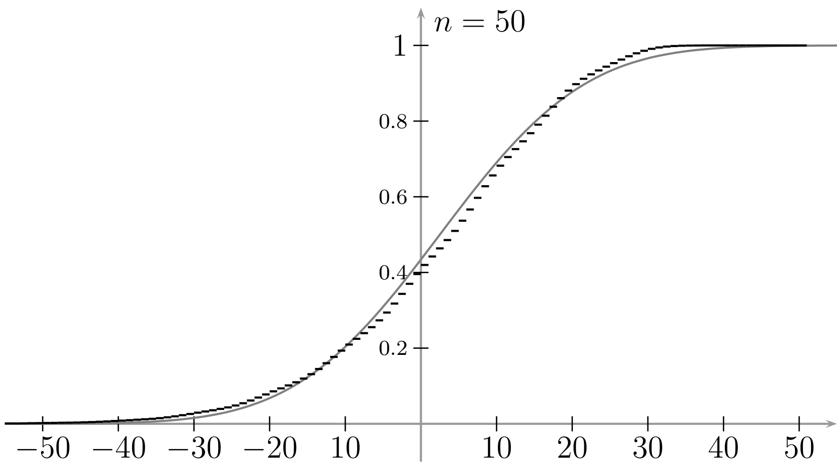
$$0\text{˙}05\sqrt{\frac{n}{4\text{˙}5475}} \ge 1\text{˙}645\,,$$

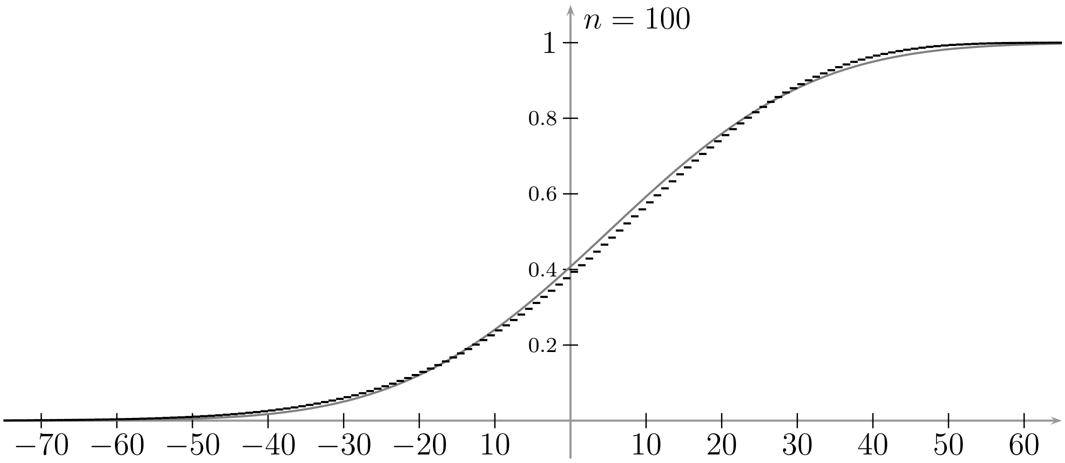
kar velja za $n \ge 4923$ (zaokrožili smo navzgor). Točna verjetnost izgube pri 4923 prodanih srečkah pa je nekoliko večja, 0˙05117. Pri 5000 prodanih srečkah pa je verjetnost izgube 0˙04987. Histogram:

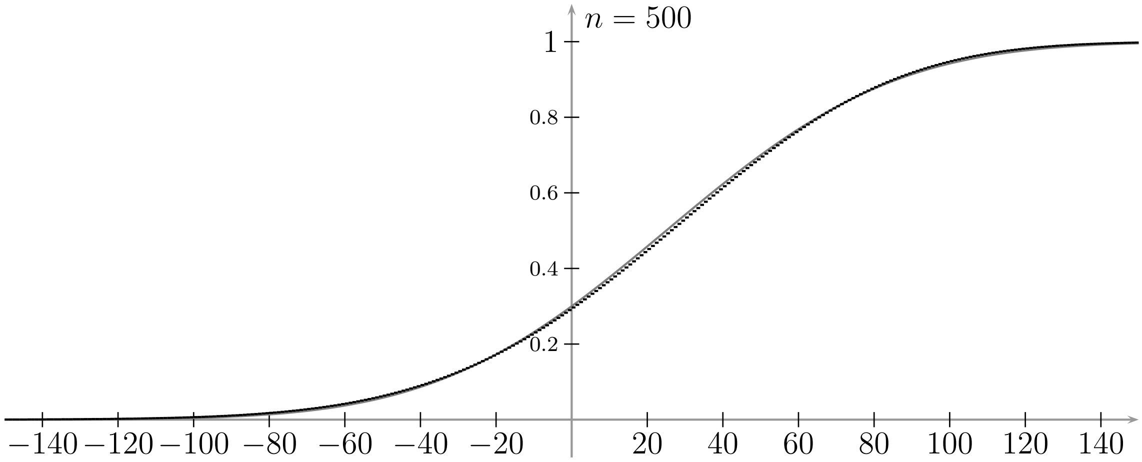


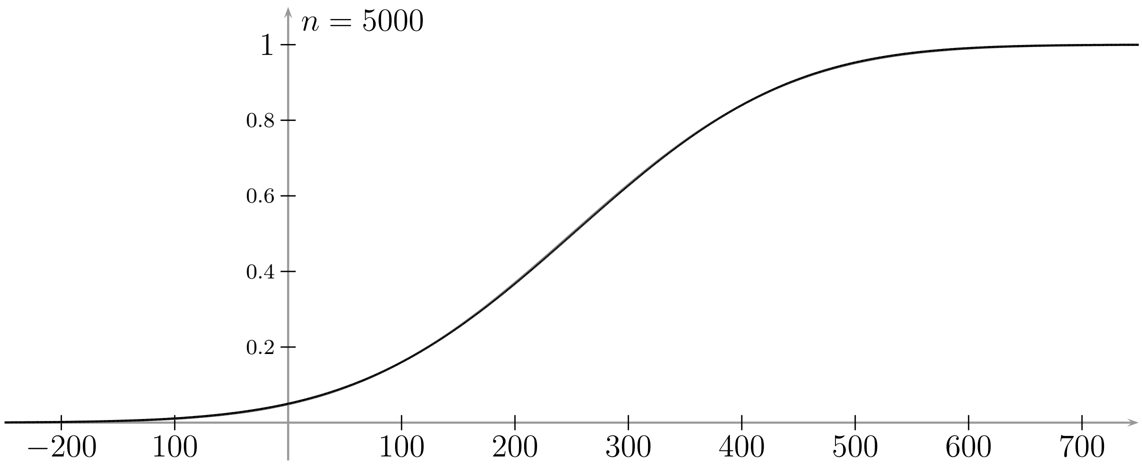
Prikažimo še grafe kumulativnih porazdelitvenih funkcij dobička prodanih srečk skupaj s pripadajočimi normalnimi slučajnimi spremenljivkami:



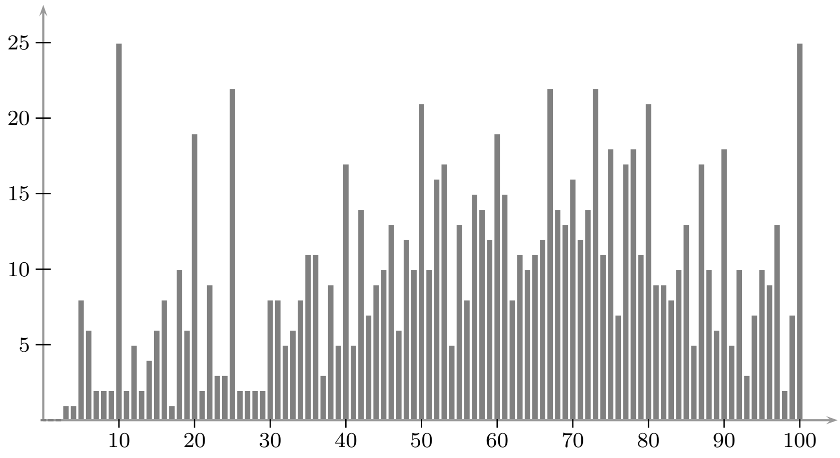




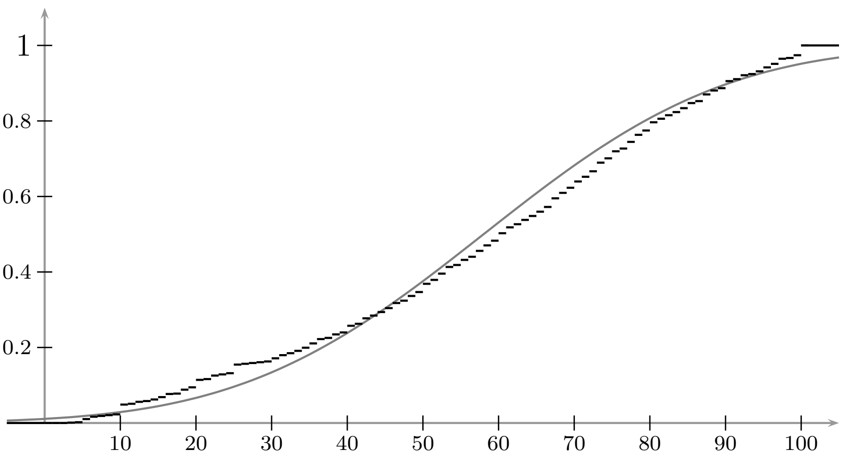




Veliko podatkov pa ne pomeni vedno (približno) normalne porazdelitve. Spodaj so prikazani rezultati 963 drugih kolokvijev iz predmeta Verjetnost in statistika za študente računalništva na Univerzi v Ljubljani v letih od 1997 do 2010:



Relativne kumulativne frekvence skupaj z normalno aproksimacijo ($\mu \doteq 58{\cdot}02, \sigma \doteq 25{\cdot}31$):



Literatura

[1] R. Jamnik: Verjetnostni račun. ZOTKS, Ljubljana, 1987.

[2] R. Jamnik: Verjetnostni račun in statistika. DMFA, Ljubljana, 1995.

[3] M. Hladnik: Verjetnost in statistika, zapiski s predavanj. FRI, Ljubljana, 2002.

[4] J. A. Čibej: Matematika: kombinatorika, verjetnostni račun, statistika. DZS, Ljubljana, 1994.

[5] A. Jurišić: Verjetnostni račun in statistika. Dosegljivo na: http://lkrv.fri.uni-lj.si/~ajurisic/stat10/