

Pravdepodobnosť a štatistika pre informatikov

*Riešené príklady
časť 1: pravdepodobnosť*

Beáta Stehlíková

Fakulta matematiky, fyziky a informatiky
Univerzita Komenského v Bratislave

december 2025

1 Kombinatorická pravdepodobnosť, podmienená pravdepodobnosť

Príklad 1. Výrobok považujeme za chybný, ak nemá predpísaný rozmer alebo hmotnosť. Výrobkov, ktoré nemajú predpísaný rozmer, je 10 percent, ktoré nespĺňajú predpísanú hmotnosť, je 30 percent. Výrobkov bez chyby je 65 percent. Určte pravdepodobnosť, že náhodne vybraný výrobok nemá predpísanú hmotnosť, ale má predpísaný rozmer.

Riešenie. Označme:

- R – výrobok nemá predpísaný rozmer (komplement R^c potom znamená, že výrobok má predpísaný rozmer),
- H – výrobok nemá predpísanú hmotnosť (komplement H^c potom znamená, že výrobok má predpísanú hmotnosť).

Potom podľa zadania platí:

$$\mathbb{P}(R) = 0,10, \quad \mathbb{P}(H) = 0,30, \quad \mathbb{P}(R^c \cap H^c) = 0,65.$$

a hľadáme pravdepodobnosť $\mathbb{P}(H \cap R^c)$. Postupne budeme počítat (je užitočné nakresliť si Vennove diagramy):

- Výrobkov bez chyby je 65 percent, takže výrobkov s nejakou chybou je 35 percent. To znamená, že nemajú predpísaný rozmer alebo nemajú predpísanú hmotnosť. V našom značení to znamená, že $\mathbb{P}(R \cup H) = 0,35$.
- Podľa vzorca pre zjednotenie udalostí:

$$\mathbb{P}(R \cup H) = \mathbb{P}(R) + \mathbb{P}(H) - \mathbb{P}(R \cap H),$$

a teda pravdepodobnosť, že výrobok nemá ani predpísaný rozmer, ani predpísanú hmotnosť, je:

$$\mathbb{P}(R \cap H) = \mathbb{P}(R) + \mathbb{P}(H) - \mathbb{P}(R \cup H) = 0,10 + 0,30 - 0,35 = 0,05.$$

- Nakoniec použijeme rozklad:

$$\mathbb{P}(H) = \mathbb{P}(H \cap R) + \mathbb{P}(H \cap R^c),$$

z ktorého dostaneme

$$\mathbb{P}(H \cap R^c) = \mathbb{P}(H) - \mathbb{P}(H \cap R) = 0,30 - 0,05 = 0,25.$$

Pravdepodobnosť, že náhodne vybraný výrobok nemá predpísanú hmotnosť, ale má predpísaný rozmer, je 0,25.

Príklad 2. V antikvariáte sa cena knihy znižuje, ak má vytrhnutú aspoň jednu stranu, alebo sú jej stránky popísané. Kníh, ktoré majú vytrhnutú aspoň jednu stranu, je 20 percent. Kníh, ktoré sú popísané, je 30 percent a kníh bez chyby je 70 percent. Určte pravdepodobnosť, že náhodne vybraná kniha je popísaná, ale má všetky strany.

Riešenie. Označme:

- V – kniha má vytrhnutú aspoň jednu stranu (komplement V^c potom znamená, že kniha má všetky strany),
- P – kniha je popísaná (komplement P^c potom znamená, že kniha nie je popísaná).

Potom podľa zadania platí:

$$\mathbb{P}(V) = 0,20, \quad \mathbb{P}(P) = 0,30, \quad \mathbb{P}(V^c \cap P^c) = 0,70,$$

a hľadáme pravdepodobnosť $\mathbb{P}(P \cap V^c)$. Postupne budeme počítat:

- Kníh bez chyby je 70 percent, takže kníh s nejakou chybou je 30 percent. To znamená, že majú vytrhnutú aspoň jednu stranu alebo sú popísané. V našom značení to znamená, že

$$\mathbb{P}(V \cup P) = 0,30.$$

- Podľa vzorca pre zjednotenie udalostí:

$$\mathbb{P}(V \cup P) = \mathbb{P}(V) + \mathbb{P}(P) - \mathbb{P}(V \cap P),$$

a teda pravdepodobnosť, že kniha má vytrhnutú aspoň jednu stranu a zároveň je popísaná, je:

$$\mathbb{P}(V \cap P) = \mathbb{P}(V) + \mathbb{P}(P) - \mathbb{P}(V \cup P) = 0,20 + 0,30 - 0,30 = 0,20.$$

- Nakoniec použijeme rozklad:

$$\mathbb{P}(P) = \mathbb{P}(P \cap V) + \mathbb{P}(P \cap V^c),$$

z ktorého dostaneme

$$\mathbb{P}(P \cap V^c) = \mathbb{P}(P) - \mathbb{P}(P \cap V) = 0,30 - 0,20 = 0,10.$$

Pravdepodobnosť, že náhodne vybraná kniha je popísaná, ale má všetky strany, je 0,10.

Príklad 3. Hádzeme dvoma kockami. Vypočítajte pravdepodobnosť toho, že ak na prvej kocke padla dvojka, padne súčet väčší ako 6.

Riešenie. Označme:

- A – súčet bodov na dvoch kockách je väčší ako 6,
- B – na prvej kocke padla dvojka.

Hľadáme podmienenú pravdepodobnosť

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

Kocky sú pravidelné, takže $\mathbb{P}(B) = \frac{1}{6}$. Ak na prvej kocke padla dvojka, tak súčet je väčší ako 6, ak na druhej kocke je 5 alebo 6. Pravdepodobnosti $\mathbb{P}(A \cap B)$ teda

zodpovedajú dva priaznivé výsledky („prvá kocka 2, druhá kocka 5“ a „prvá kocka 2, druhá kocka 6“), pričom celkovo je 36 možností, ako môže vyzerateľ výsledok hádzania dvoma kockami. Preto

$$\mathbb{P}(A \cap B) = \mathbb{P}(\text{prvá kocka} = 2, \text{druhá kocka} \in \{5, 6\}) = \frac{2}{36} = \frac{1}{18}.$$

Potom:

$$\mathbb{P}(A | B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\frac{1}{18}}{\frac{1}{6}} = \frac{1}{3}.$$

Pravdepodobnosť, že súčet bude väčší ako 6, ak na prvej kocke padla dvojka, je $\frac{1}{3}$.

Príklad 4. Z výrobkov určitého druhu dosahuje 95 percent predpísanú kvalitu. V istom závode, ktorý vyrába 80 percent celkovej produkcie, však predpísanú kvalitu dosahuje 98 percent výrobkov. Bol náhodne vybraný výrobok a zistilo sa, že má predpísanú kvalitu. Aká je pravdepodobnosť, že bol vyrobený v spomínanom závode?

Riešenie. Označme:

- A – výrobok bol vyrobený v závode, ktorý vyrába 80 % produkcie,
- Q – výrobok má predpísanú kvalitu.

Podľa zadania platí:

$$\mathbb{P}(A) = 0,8, \quad \mathbb{P}(Q) = 0,95, \quad \mathbb{P}(Q | A) = 0,98.$$

Hľadáme podmienenú pravdepodobnosť, že výrobok s predpísanou kvalitou pochádza z uvedeného závodu, teda $\mathbb{P}(A | Q)$. Platí:

$$\mathbb{P}(A | Q) = \frac{\mathbb{P}(A \cap Q)}{\mathbb{P}(Q)} = \frac{\mathbb{P}(Q | A) \mathbb{P}(A)}{\mathbb{P}(Q)}.$$

Dosadíme:

$$\mathbb{P}(A | Q) = \frac{\mathbb{P}(Q | A) \mathbb{P}(A)}{\mathbb{P}(Q)} = \frac{0,98 \cdot 0,8}{0,95} = \frac{0,784}{0,95} \doteq 0,825.$$

Pravdepodobnosť, že náhodne vybraný výrobok s predpísanou kvalitou bol vyrobený v závode A, je približne 0,825.

Príklad 5. Traja športovci hádzu oštepom nezávisle jeden od druhého. Prvý prekoná hranicu 80 metrov s pravdepodobnosťou 0,8, druhý 0,7 a tretí 0,5. Každý z nich vykoná jeden hod. Aká je pravdepodobnosť, že aspoň raz bude prekonaná hranica 80 metrov?

Riešenie. Označme:

- A_1 - prvý športovec prekoná hranicu 80 m,
- A_2 - druhý športovec prekoná hranicu 80 m,

- A_3 - tretí športovec prekoná hranicu 80 m.

Podľa zadania platí:

$$\mathbb{P}(A_1) = 0,8, \quad \mathbb{P}(A_2) = 0,7, \quad \mathbb{P}(A_3) = 0,5.$$

Hľadáme pravdepodobnosť, že aspoň jeden športovec prekoná hranicu 80 m, teda $\mathbb{P}(A_1 \cup A_2 \cup A_3)$.

Je jednoduchšie použiť doplnkovú udalosť, že žiadny športovec neprekročí hranicu:

$$\mathbb{P}(A_1 \cup A_2 \cup A_3) = 1 - \mathbb{P}(A_1^c \cap A_2^c \cap A_3^c).$$

Keďže hody sú nezávislé, platí:

$$\begin{aligned} \mathbb{P}(A_1^c \cap A_2^c \cap A_3^c) &= \mathbb{P}(A_1^c) \mathbb{P}(A_2^c) \mathbb{P}(A_3^c) \\ &= (1 - 0,8)(1 - 0,7)(1 - 0,5) = 0,2 \cdot 0,3 \cdot 0,5 = 0,03. \end{aligned}$$

Teda

$$\mathbb{P}(A_1 \cup A_2 \cup A_3) = 1 - 0,03 = 0,97.$$

Pravdepodobnosť, že aspoň jeden športovec prekoná hranicu 80 m, je 0,97.

Príklad 6. Cez prenosný kanál sa prenášajú dva signály - 0 a 1. V dôsledku porúch sa jedno percento signálov zmení na opačné nezávisle od toho, či sa predchádzajúce signály preniesli správne alebo nie. Cez kanál bola vyslaná kombinácia 10110. Určte pravdepodobnosť toho, že

- celá kombinácia sa preniesie správne
- bude prijatá kombinácia 11110.

Riešenie. Označme:

- A_i - i -ty signál sa preniesol správne ($i = 1, \dots, 5$).

Podľa zadania platí $\mathbb{P}(A_i^c) = 0,01$ (a teda $\mathbb{P}(A_i) = 0,99$) a tieto udalosti sú nezávislé

- Pravdepodobnosť toho, že všetkých 5 signálov sa preniesie správne, je

$$\mathbb{P}(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5) = \mathbb{P}(A_1) \mathbb{P}(A_2) \mathbb{P}(A_3) \mathbb{P}(A_4) \mathbb{P}(A_5) = 0,99^5 \doteq 0,951.$$

Správny prenos celej kombinácie má pravdepodobnosť približne 0,951.

- To, že prijatá kombinácia je 11110, znamená:

- 1. signál: správne (A_1)
- 2. signál: chybný prenos (A_2^c)
- 3. signál: správne (A_3)
- 4. signál: správne (A_4)
- 5. signál: správne (A_5)

Potrebuje pravdepodobnosť:

$$\begin{aligned}\mathbb{P}(A_1 \cap A_2^c \cap A_3 \cap A_4 \cap A_5) &= \mathbb{P}(A_1)\mathbb{P}(A_2^c)\mathbb{P}(A_3)\mathbb{P}(A_4)\mathbb{P}(A_5) \\ &= 0,99^4 \cdot 0,01 \doteq 0,0096.\end{aligned}$$

Prijatie kombinácie 11110 pri vyslaní 10110 má pravdepodobnosť približne 0,0096.

Príklad 7. Zákazník náhodne vyberá obraz zo skupiny desiatich originálov a dvoch kópií. Radí sa s odborníkom, ktorý správne určí, či ide o originál alebo kópiu s pravdepodobnosťou $5/6$. Odborník tvrdí, že vybraný obraz je originál. Aká je pravdepodobnosť toho, že je to skutočne originál?

Riešenie. Označme:

- O - vybraný obraz je originál (a teda O^c - vybraný obraz je kópia),
- E - expert tvrdí, že obraz je originál.

Podľa zadania platí:

$$\mathbb{P}(O) = \frac{10}{12}, \quad \mathbb{P}(E | O) = \frac{5}{6},$$

a teda aj

$$\mathbb{P}(O^c) = \frac{2}{12}, \quad \mathbb{P}(E | O^c) = \frac{1}{6}$$

Hľadáme podmienenú pravdepodobnosť $\mathbb{P}(O | E)$. Z Bayesovho vzorca:

$$\mathbb{P}(O | E) = \frac{\mathbb{P}(E \cap O)}{\mathbb{P}(E)} = \frac{\mathbb{P}(E | O) \mathbb{P}(O)}{\mathbb{P}(E | O) \mathbb{P}(O) + \mathbb{P}(E | O^c) \mathbb{P}(O^c)}.$$

Dosadíme:

$$\mathbb{P}(O | E) = \frac{\frac{5}{6} \cdot \frac{10}{12}}{\frac{5}{6} \cdot \frac{10}{12} + \frac{1}{6} \cdot \frac{2}{12}} = \frac{50}{52} = \frac{25}{26} \doteq 0,962.$$

Pravdepodobnosť, že obraz je skutočne originál, ak odborník tvrdí, že je originál, je $\frac{25}{26}$, teda približne 0,962.

2 Diskrétné náhodné premenné

Príklad 1: Diskrétna náhodná premenná X nadobúda hodnoty z množiny $\{0,1,2,3,4\}$, pričom

$$\mathbb{P}(X = x) = c(x + 1)^2,$$

kde $c > 0$ je konštanta.

- (a) Určte hodnotu konštanty c .
- (b) Vypočítajte strednú hodnotu $\mathbb{E}(X)$.
- (c) Vypočítajte disperziu $\mathbb{D}(X)$.

Riešenie:

- (a) Musí platiť

$$\sum_{k=0}^4 \mathbb{P}(X = k) = 1.$$

Dosadíme pravdepodobnosti, ľavá strana potom je:

$$\sum_{k=0}^4 c(x + 1)^2 = c \sum_{k=0}^4 (x + 1)^2 = c(1^2 + 2^2 + 3^2 + 4^2 + 5^2) = 55c,$$

a teda $c = \frac{1}{55}$.

- (b) Stredná hodnota je daná vzťahom

$$\mathbb{E}(X) = \sum_{k=0}^4 k \mathbb{P}(X = k).$$

Dosadíme:

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=0}^4 k \frac{1}{55} (k + 1)^2 = \frac{1}{55} \sum_{k=0}^4 k(k + 1)^2 \\ &= \frac{1}{55} (0 \cdot 1 + 1 \cdot 4 + 2 \cdot 9 + 3 \cdot 16 + 4 \cdot 25) = \frac{170}{55} = \frac{34}{11}. \end{aligned}$$

- (c) Pri výpočte disperzie využijeme vzťah $\mathbb{D}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$, potrebujeme teda $\mathbb{E}(X^2)$. Keďže

$$\begin{aligned} \mathbb{E}(X^2) &= \sum_{k=0}^4 k^2 \mathbb{P}(X = k) = \frac{1}{55} \sum_{k=0}^4 k^2 (k + 1)^2 \\ &= \frac{1}{55} (0 \cdot 1 + 1 \cdot 4 + 4 \cdot 9 + 9 \cdot 16 + 16 \cdot 25) = \frac{584}{55}, \end{aligned}$$

dostaneme

$$\mathbb{D}(X) = \frac{584}{55} - \left(\frac{34}{11}\right)^2 = \frac{11692}{121}.$$

Príklad 2: Na stole je košík s ovocím, sú v ňom 4 jablká, 3 hrušky a 3 banány. Náhodne vyberieme dva kusy ovocia. Nech X je náhodná premenná, ktorá udáva počet jablák, ktoré sú v našom výbere.

- (a) Určte hodnoty, ktoré môže X nadobúdať a nájdite ich pravdepodobnosti.
- (b) Vypočítajte strednú hodnotu náhodnej premennej X .
- (c) Vypočítajte disperziu náhodnej premennej X .

Riešenie:

- (a) Náhodná premenná X nadobúda hodnoty z množiny $\{0,1,2\}$. Pri výpočte pravdepodobností využijeme, že celkový počet možností, ako vybrať dva kusy ovocia je $\binom{10}{2} = 45$. Ďalej:

- $X = 0$ (žiadne jablká) znamená, že vyberáme 2 kusy ovocia zo 6 (hrušky a banány). Počet možností, ako to spraviť, je $\binom{6}{2} = 15$, a teda

$$\mathbb{P}(X = 0) = \frac{15}{45} = \frac{1}{3}.$$

- $X = 1$ (jedno jablko): Vyberieme 1 jablko z 4 a 1 iné ovocie zo 6, počet možností je $\binom{4}{1}\binom{6}{1} = 24$, preto

$$\mathbb{P}(X = 1) = \frac{24}{45} = \frac{8}{15}.$$

- $X = 2$ (2 jablká): vyberieme 2 jablká zo 4, počet možností je $\binom{4}{2} = 6$, takže

$$\mathbb{P}(X = 2) = \frac{6}{45} = \frac{2}{15}.$$

- (b) Stredná hodnota je

$$\mathbb{E}(X) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{8}{15} + 2 \cdot \frac{2}{15} = \frac{12}{15} = \frac{4}{5}.$$

- (c) Najprv vypočítame $\mathbb{E}(X^2)$:

$$\mathbb{E}(X^2) = 0^2 \cdot \frac{1}{3} + 1^2 \cdot \frac{8}{15} + 2^2 \cdot \frac{2}{15} = \frac{16}{15}.$$

Disperzia potom je :

$$\mathbb{D}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{16}{15} - \left(\frac{4}{5}\right)^2 = \frac{16}{15} - \frac{16}{25} = \frac{32}{75}.$$

Príklad 3: Na športovom dni sa vyberá náhodne 3-členný tím zo 5 dievčat a 4 chlapcov. Nech X je náhodná premenná, ktorá udáva počet dievčat v tíme.

- (a) Určte hodnoty, ktoré môže X nadobúdať a nájdite ich pravdepodobnosti.
- (b) Vypočítajte strednú hodnotu náhodnej premennej X .
- (c) Vypočítajte disperziu náhodnej premennej X .

Riešenie:

- (a) Náhodná premenná X nadobúda hodnoty z množiny $\{0,1,2,3\}$. Pri výpočte pravdepodobností využijeme, že celkový počet možností, ako trojčlenný tím, je $\binom{9}{3} = 84$. Ďalej:

- $X = 0$ (žiadna dievča): vyberáme 3 chlapcov zo 4, počet možností je $\binom{4}{3} = 4$, preto

$$\mathbb{P}(X = 0) = \frac{4}{84} = \frac{1}{21}.$$

- $X = 1$ (jedno dievča): vyberieme 1 dievča z 5 a 2 chlapcov zo 4, počet možností je $\binom{5}{1}\binom{4}{2} = 5 \cdot 6 = 30$, teda

$$\mathbb{P}(X = 1) = \frac{30}{84} = \frac{5}{14}.$$

- $X = 2$ (dve dievčatá): vyberieme 2 dievčatá z 5 a 1 chlapca zo 4, počet možností je $\binom{5}{2}\binom{4}{1} = 10 \cdot 4 = 40$, teda

$$\mathbb{P}(X = 2) = \frac{40}{84} = \frac{10}{21}.$$

- $X = 3$ (tri dievčatá): vyberieme 3 dievčatá zo 5, počet možností je $\binom{5}{3} = 10$, teda

$$\mathbb{P}(X = 3) = \frac{10}{84} = \frac{5}{42}.$$

- (b) Stredná hodnota je

$$\mathbb{E}(X) = 0 \cdot \frac{1}{21} + 1 \cdot \frac{5}{14} + 2 \cdot \frac{10}{24} + 3 \cdot \frac{5}{42} = \frac{70}{42} = \frac{5}{3},$$

- (c) Najprv vypočítame $\mathbb{E}(X^2)$:

$$\mathbb{E}(X^2) = 0^2 \cdot \frac{1}{21} + 1^2 \cdot \frac{5}{14} + 2^2 \cdot \frac{10}{24} + 3^2 \cdot \frac{5}{42} = \frac{130}{42} = \frac{65}{21}.$$

Disperzia potom je :

$$\mathbb{D}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{65}{21} - \left(\frac{5}{3}\right)^2 = \frac{20}{63}.$$

Príklad 4: Na športovom dni sa vyberá náhodne 3-členný tím z 2 dievčat a 7 chlapcov. Nech X je náhodná premenná, ktorá udáva počet dievčat v tíme.

- (a) Určte hodnoty, ktoré môže X nadobúdať a nájdite ich pravdepodobnosti.
- (b) Vypočítajte strednú hodnotu náhodnej premennej X .
- (c) Vypočítajte disperziu náhodnej premennej X .

Riešenie:

- (a) Tím je, rovnako ako v predchádzajúcom príklade, trojčlenný, ale keďže sú len dve dievčatá, náhodná premenná X nenadobúda hodnoty z množiny $\{0,1,2,3\}$, ale iba $\{0,1,2\}$. Ďalej je výpočet analogický.

Pri výpočte pravdepodobností využijeme, že celkový počet možností, ako trojčlenný tím, je $\binom{9}{3} = 84$. Ďalej:

- $X = 0$ (žiadna dievča): vyberáme 3 chlapcov zo 7, počet možností je $\binom{7}{3} = 35$, preto

$$\mathbb{P}(X = 0) = \frac{35}{84}.$$

- $X = 1$ (jedno dievča): vyberieme 1 dievča z 2 a 2 chlapcov zo 7, počet možností je $\binom{2}{1} \binom{7}{2} = 2 \cdot 21 = 42$, teda

$$\mathbb{P}(X = 1) = \frac{42}{84} = \frac{1}{2}.$$

- $X = 2$ (dve dievčatá): vyberieme 2 dievčatá z 2 a 1 chlapca zo 7, počet možností je $\binom{2}{2} \binom{7}{1} = 1 \cdot 7 = 7$, teda

$$\mathbb{P}(X = 2) = \frac{7}{84} = \frac{1}{12}.$$

- (b) Stredná hodnota je

$$\mathbb{E}(X) = 0 \cdot \frac{35}{84} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{12} = \frac{4}{6} = \frac{2}{3}.$$

- (c) Najprv vypočítame $\mathbb{E}(X^2)$:

$$\mathbb{E}(X^2) = 0^2 \cdot \frac{35}{84} + 1^2 \cdot \frac{1}{2} + 2^2 \cdot \frac{1}{12} = \frac{5}{6}.$$

Disperzia potom je :

$$\mathbb{D}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \frac{5}{6} - \left(\frac{2}{3}\right)^2 = \frac{7}{18}.$$

Príklad 5: Predpokladajme, že v predchádzajúcich dvoch príkladoch nás zaujíma len stredná hodnota počtu dievčat. Teda riešime úlohy:

- (a) Na športovom dni sa vyberá náhodne 3-členný tím z 5 dievčat a 4 chlapcov. Nech X je náhodná premenná, ktorá udáva počet dievčat v tíme. Nájdite strednú hodnotu náhodnej premennej X .
- (b) Na športovom dni sa vyberá náhodne 3-členný tím z 2 dievčat a 7 chlapcov. Nech X je náhodná premenná, ktorá udáva počet dievčat v tíme. Nájdite strednú hodnotu náhodnej premennej X .

Okrem toho budeme riešiť úlohu:

- (c) Na športovom dni sa vyberá náhodne 10-členný tím z 15 dievčat a 15 chlapcov. Nech X je náhodná premenná, ktorá udáva počet dievčat v tíme. Nájdite strednú hodnotu náhodnej premennej X .

Riešenie: Je jasné, že v bode (c) by bol výpočet pomocou vyjadrenia pravdepodobností jednotlivých hodnôt zdlhavý. Pôjdeme na to preto inak – použijeme indikátory (náhodné premenné nadobúdajúce hodnoty 0 a 1). Výpočet každého z týchto príkladov bude veľmi podobný.

- (a) Definujeme náhodné premenné $X_1, \dots, X_5$ tak, že X_i sa bude rovnáť jednej, ak sa i -te dievča dostane do tímu a inak sa bude rovnáť nule:

$$I_i = \begin{cases} 1, & \text{ak } i\text{-te dievča je v tíme,} \\ 0, & \text{inak.} \end{cases}$$

Potom počet dievčat v tíme X sa dá vyjadriť ako $X = X_1 + \dots + X_5$, a teda

$$\mathbb{E}(X) = \mathbb{E}(X_1) + \dots + \mathbb{E}(X_5).$$

Máme:

$$\mathbb{E}(X_i) = 1 \cdot \mathbb{P}(X_i = 1) + 0 \cdot \mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1),$$

zostáva teda spočítať $\mathbb{P}(X_i = 1)$, teda že sa konkrétne dievča dostane do tímu. Všetkých možností, ako zostaviť trojčlenný tím z 9 účastníkov je $\binom{9}{3} = 84$. Možností, v ktorých je konkrétne sledované dievča v tíme, zodpovedajú tomu, že okrem nej ešte zo zvyšných 8 účastníkov vyberieme zostávajúcich dvoch členov tímu. Takýchto možností je $\binom{8}{2} = 28$. Teda pravdepodobnosť, že i -te dievča je v tíme, že $\frac{28}{84} = \frac{1}{3}$. Teda pre každé i je $\mathbb{E}(X_i) = \frac{1}{3}$ a

$$\mathbb{E}(X) = \mathbb{E}(X_1) + \dots + \mathbb{E}(X_5) = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = \frac{5}{3}.$$

- (b) Definujeme náhodné premenné X_1, X_2 tak, že X_i sa bude rovnáť jednej, ak sa i -te dievča dostane do tímu a inak sa bude rovnáť nule:

$$I_i = \begin{cases} 1, & \text{ak } i\text{-te dievča je v tíme,} \\ 0, & \text{inak.} \end{cases}$$

Potom počet dievčat v tíme X sa dá vyjadriť ako $X = X_1 + X_2$, a teda

$$\mathbb{E}(X) = \mathbb{E}(X_1) + \mathbb{E}(X_2).$$

Máme:

$$\mathbb{E}(X_i) = 1 \cdot \mathbb{P}(X_i = 1) + 0 \cdot \mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1),$$

zostáva teda spočítať $\mathbb{P}(X_i = 1)$, teda že sa konkrétne dievča dostane do tímu. Všetkých možností, ako zostaviť trojčlenný tím z 9 účastníkov je $\binom{9}{3} = 84$. Možností, v ktorých je konkrétne sledované dievča v tíme, zodpovedajú tomu, že okrem nej ešte zo zvyšných 8 účastníkov vyberieme zostávajúcich dvoch členov tímu. Takýchto možností je $\binom{8}{2} = 28$. Teda pravdepodobnosť, že i -te dievča je tíme, že $\frac{28}{84} = \frac{1}{3}$. Teda pre každé i je $\mathbb{E}(X_i) = \frac{1}{3}$ a

$$\mathbb{E}(X) = \mathbb{E}(X_1) + \mathbb{E}(X_2) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}.$$

- (c) Definujeme náhodné premenné $X_1, \dots, X_{15}$ tak, že X_i sa bude rovnáť jednej, ak sa i -te dievča dostane do tímu a inak sa bude rovnáť nule:

$$I_i = \begin{cases} 1, & \text{ak } i\text{-te dievča je v tíme,} \\ 0, & \text{inak.} \end{cases}$$

Potom počet dievčat v tíme X sa dá vyjadriť ako $X = X_1 + \dots + X_{15}$, a teda

$$\mathbb{E}(X) = \mathbb{E}(X_1) + \dots + \mathbb{E}(X_{15}).$$

Máme:

$$\mathbb{E}(X_i) = 1 \cdot \mathbb{P}(X_i = 1) + 0 \cdot \mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1),$$

zostáva teda spočítať $\mathbb{P}(X_i = 1)$, teda že sa konkrétne dievča dostane do tímu. Všetkých možností, ako zostaviť 10-členný tím z 30 účastníkov je $\binom{30}{10}$. Možností, v ktorých je konkrétne sledované dievča v tíme, zodpovedajú tomu, že okrem nej ešte zo zvyšných 29 účastníkov vyberieme zostávajúcich 9 členov tímu. Takýchto možností je $\binom{29}{9}$. Teda pravdepodobnosť, že i -te dievča je tíme, že $\frac{\binom{29}{9}}{\binom{30}{10}}$. Teda pre každé i je

$$\mathbb{E}(X_i) = \frac{\binom{29}{9}}{\binom{30}{10}} = \frac{\frac{29!}{9!20!}}{\frac{30!}{10!20!}} = \frac{10}{30} = \frac{1}{3}$$

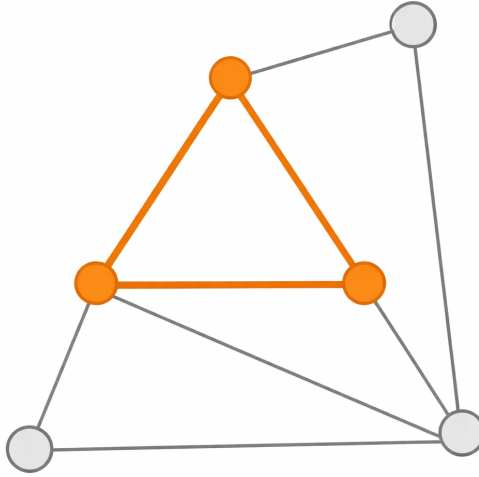
a

$$\mathbb{E}(X) = \mathbb{E}(X_1) + \dots + \mathbb{E}(X_{15}) = 15 \cdot \frac{1}{3} = 5$$

Príklad 6. Majme graf so 6 vrcholmi. Každá z potenciálnych hrán vznikne s pravdepodobnosťou 0,6 nezávisle od ostatných. Trojicu vrcholov nazveme trojuholníkom, ak sú každé dve z nich spojené hranou (na obr. 9 je ukážka grafu a jeden z trojuholníkov). Aká je stredná hodnota počtu trojuholníkov?

Riešenie. Pre každú trojicu vrcholov $\{i, j, k\}$ zavedieme indikátorovú premennú:

$$X_{ijk} = \begin{cases} 1, & \text{ak hrany } \{i, j\}, \{i, k\}, \{j, k\} \text{ v grafe existujú,} \\ 0, & \text{inak.} \end{cases}$$



Obr. 1: Ukážka grafu a jeden z trojuholníkov

Počet trojuholníkov v grafe môžeme zapísať ako súčet všetkých indikátorov:

$$X = \sum_{1 \leq i < j < k \leq 6} X_{ijk}.$$

a teda

$$\mathbb{E}(X) = \sum_{1 \leq i < j < k \leq 6} \mathbb{E}(X_{ijk}).$$

Pritom

$$\begin{aligned} \mathbb{E}(X_{ijk}) &= 0 \cdot \mathbb{P}(X_{ijk} = 0) + 1 \cdot \mathbb{P}(X_{ijk} = 1) = \mathbb{P}(X_{ijk} = 1) \\ &= \mathbb{P}(\text{hrany } \{i, j\}, \{i, k\}, \{j, k\} \text{ existujú}) = \\ &= \mathbb{P}(\text{hrana } \{i, j\} \text{ existuje}) \cdot \mathbb{P}(\text{hrana } \{i, k\} \text{ existuje}) \cdot \mathbb{P}(\text{hrana } \{j, k\} \text{ existuje}) \\ &= 0,6^3. \end{aligned}$$

Teda

$$\mathbb{E}(X) = \sum_{1 \leq i < j < k \leq 6} 0,6^3.$$

Počet trojíc vrcholov je: $\binom{6}{3} = 20$., takže suma má 20 sčítancov. To znamená, že

$$\mathbb{E}(X) = 20 \cdot 0,6^3 = 4,32.$$

Stredná hodnota počtu trojuholníkov je 4,32.

Príklad 7. Uvažujme abecedu s 26 písmenami (teda písmená bez diakritiky). Generujeme náhodnú postupnosť dĺžky 10^6 , kde každé písmeno sa vyberá nezávisle od ostatných s rovnakou pravdepodobnosťou pre každú možnosť. Nech X je náhodná premenná, ktorá udáva počet výskytov reťazca FMFI v tejto postupnosti. Vypočítajte jej strednú hodnotu.

Riešenie. Pre $i = 1, 2, \dots, 10^6 - 3$ (nakolko sledovaný reťazec má dĺžku 4, začať musí najneskôr na štvrtom mieste od konca) definujeme indikátory:

$$I_i = \begin{cases} 1, & \text{ak reťazec na pozíciách } i, i+1, i+2, i+3 \text{ je F, M, F, I,} \\ 0, & \text{inak.} \end{cases}$$

Počet výskytov reťazca môžeme zapísať ako súčet indikátorov:

$$X = \sum_{i=1}^{10^6-3} I_i.$$

Rovnako ako v predchádzajúcich príkladoch je $\mathbb{E}(I_i) = \mathbb{P}(I_i = 1)$. Dostaneme tak:

$$\mathbb{E}(X_n) = \sum_{i=1}^{10^6-3} \mathbb{E}(I_i) = \sum_{i=1}^{10^6-3} \mathbb{P}(I_i = 1) = \sum_{i=1}^{10^6-3} \left(\frac{1}{26}\right)^4 = (10^6 - 3) \cdot \frac{1}{26^4} = 2,188$$

Pre postupnosť dĺžky 1 000 000 je očakávaný počet výskytov reťazca FMFI približne 2,188.

Príklad 8. Šofér musí prejsť štyri križovatky. Na každej mu semafor povolí jazdu bez zastavenia s pravdepodobnosťou 0,5, nezávisle od predchádzajúcich križovatiek. Nech náhodná premenná X predstavuje počet križovatiek, ktoré šofér prejde pred tým, ako bude musieť prvýkrát zastať na červenu (resp. sa rovná 4, ak nebude stáť ani raz).

- Nájdite rozdelenie náhodnej premennej X .
- Nájdite distribučnú funkciu náhodnej premennej X a nakreslite jej graf.

Riešenie.

- Možné hodnoty náhodnej premennej X sú $\{0, 1, 2, 3, 4\}$. Ich pravdepodobnosti:

- $X = 0$ znamená, že šofér zastaví hneď na prvej križovatke.

$$\mathbb{P}(X = 0) = \mathbb{P}(\text{červená na 1. križovatke}) = \frac{1}{2}$$

- $X = 1$ znamená, že šofér prejde prvú križovatku na zelenú a potom zastaví na druhej na červenú. Z nezávislosti:

$$\mathbb{P}(X = 1) = \mathbb{P}(\text{zelená na 1. križovatke, červená na 2. križovatke}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

- $X = 2$ znamená, že šofér prejde prvú a druhú križovatku na zelenú, potom zastaví na tretej na červenú. Z nezávislosti:

$$\begin{aligned} \mathbb{P}(X = 2) &= \mathbb{P}(\text{zelená na 1. križovatke, zelená na 2., červená na 3}) \\ &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8} \end{aligned}$$

- $X = 3$ znamená, že šofér prejde prvú, druhú a tretiu križovatku na zelenú, potom zastaví na štvrtej na červenú. Z nezávislosti:

$$\begin{aligned}\mathbb{P}(X = 3) &= \mathbb{P}(\text{zelená na 1., zelená na 2., zelená na 3., červená na 4}) \\ &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{16}\end{aligned}$$

- $X = 4$ znamená, že šofér prejde všetky štyri križovatky na zelenú. Z nezávislosti:

$$\begin{aligned}\mathbb{P}(X = 4) &= \mathbb{P}(\text{zelená na 1., zelená na 2., zelená na 3., zelená na 4}) \\ &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{16}\end{aligned}$$

Rozdelenie náhodnej premennej X teda je:

x	0	1	2	3	4
$\mathbb{P}(X = x)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$

(b) Distribučná funkcia je definovaná ako $F(x) = \mathbb{P}(X \leq x)$, a teda

- pre $x < 0$ je

$$F(x) = 0$$

- pre $0 \leq x < 1$ je

$$F(x) = \mathbb{P}(X = 0) = \frac{1}{2}$$

- pre $1 \leq x < 2$ je

$$F(x) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$$

- pre $2 \leq x < 3$ je

$$F(x) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8}$$

- pre $3 \leq x < 4$ je

$$F(x) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) + \mathbb{P}(X = 3) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{15}{16}$$

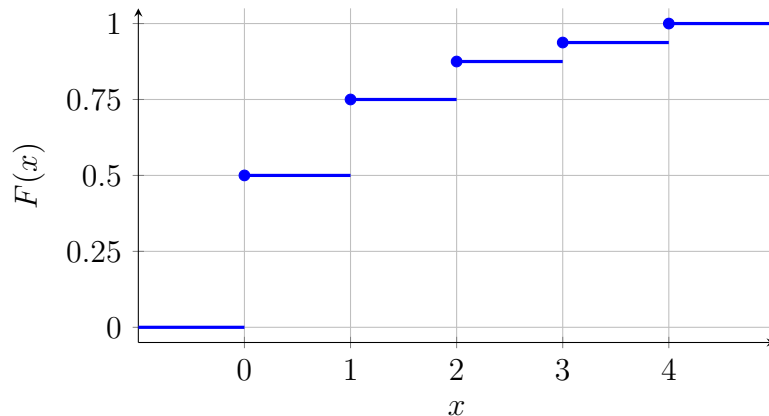
- pre $x \geq 4$ je

$$F(x) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) + \mathbb{P}(X = 3) + \mathbb{P}(X = 4) = 1$$

Celkovo:

$$F(x) = \begin{cases} 0 & \text{pre } x < 0, \\ \frac{1}{2} & \text{pre } 0 \leq x < 1, \\ \frac{3}{4} & \text{pre } 1 \leq x < 2, \\ \frac{7}{8} & \text{pre } 2 \leq x < 3, \\ \frac{15}{16} & \text{pre } 3 \leq x < 4, \\ 1 & \text{pre } x \geq 4. \end{cases}$$

Graf tejto funkcie je zobrazený na obrázku 2.



Obr. 2: Distribučná funkcia náhodnej premennej

Príklad 9. Podľa štatistík dopravného podniku sa v električke nachádza aspoň jeden čierny pasažier s pravdepodobnosťou 0,1. To, či sa v električke čierny pasažier nachádza, je nezávislé od ostatných električiek. Zistite, najmenej koľko električiek musí revízor skontrolovať, aby pravdepodobnosť, že nájde aspoň jedného čierneho pasažiera, bola aspoň 0,9.

Riešenie. Pravdepodobnosť, že v jednej električke nie je žiadny čierny pasažier, je $1 - 0,1 = 0,9$. Ak revízor skontroluje n električiek, pravdepodobnosť, že nenájde ani jedného čierneho pasažiera, je (vďaka nezávislosti) rovná $0,9^n$. To znamená, že pravdepodobnosť, že nájde aspoň jedného čierneho pasažiera, je $1 - 0,9^n$. Hľadáme najmenšie n , pre ktoré je

$$1 - 0,9^n \geq 0,9.$$

Riešime nerovnicu:

$$\begin{aligned} 0,9^n &\leq 0,1 \\ n \ln(0,9) &\leq \ln(0,1) \\ n &\geq \frac{\ln(0,1)}{\ln(0,9)} \doteq 21,86 \end{aligned}$$

Revízor musí skontrolovať aspoň 22 električiek.

Príklad 10. Piaty kamaráti, študenti, sa chystajú na výlet. Aby dostali zľavu pri návšteve hradu, potrebujú mať pri sebe preukaz študenti. Ten si každý z nich zabudne s pravdepodobnosťou 0,2, nezávisle od ostatných. Označme X náhodnú premennú, ktorá označuje, koľko si budú môcť vďaka preukazu študenta kúpiť zľavnenú vstupenku.

- Nájdite pravdepodobnosť, že zľavnenú vstupenku si kúpia všetci.
- Nájdite pravdepodobnosť, že zľavnenú vstupenku si kúpia najviac dvaja.
- Nájdite strednú hodnotu počtu kúpených zľavnených vstupeniek.

Riešenie. Počet študentov s preukazom je náhodná premenná s binomickým rozdelením, kde pravdepodobnosť úspechu p je pravdepodobnosť toho, že študent bude mať so sebou preukaz: $X \sim \text{Bin}(n = 5, p = 0.8)$. Teda

$$\mathbb{P}(X = k) = \binom{5}{k} 0.8^k 0.2^{5-k}, \quad k \in \{0, 1, 2, 3, 4, 5\}.$$

- (a) To, že zľavnenú vstupenku si kúpia všetci, znamená, že $X = 5$. Pravdepodobnosť tejto udalosti je

$$\mathbb{P}(X = 5) = \binom{5}{5} 0.8^5 0.2^0 \doteq 0,328$$

- (b) To, že zľavnenú vstupenku si kúpia najviac dvaja, znamená, že $X \leq 2$. Pravdepodobnosť tejto udalosti je

$$\begin{aligned} \mathbb{P}(X \leq 2) &= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) \\ &= \binom{5}{0} 0.8^0 0.2^5 + \binom{5}{1} 0.8^1 0.2^4 + \binom{5}{2} 0.8^2 0.2^3 \doteq 0,058. \end{aligned}$$

- (c) Stednú hodnotu počtu kúpených zľavnených vstupeniiek je stredná hodnota náhodnej premennej X . Stredná hodnota náhodnej premennej s binomickým rozdelením $\text{Bin}(n, p)$ je $n \cdot p$, čo je v našom prípade $5 \cdot 0,8 = 4$.

Príklad 11. Kniha bola vydaná v náklade 100 000 kusov. Pravdepodobnosť, že je kniha chybné zviazaná, je 0,000 1. Kvalita väzby určitej knihy je nezávislý od kvality ostatných. Pomocou aproximácie Poissonovým rozdelením vypočítajte pravdepodobnosť, že v celom náklade bude najviac 5 chybné zviazaných kníh.

Riešenie. Presné rozdelenie počtu chybné zviazaných kníh je binomické s parametrami $n = 100\,000$, $p = 0,000\,1$. Pretože n je veľké a p je malé, rozdelenie môžeme aproximovať Poissonovým rozdelením s parametrom $\lambda = n \cdot p$ (stredná hodnota), v našom prípade $\lambda = 100\,000 \cdot 0,000\,1 = 10$.

Pravdepodobnosti Poissonovho rozdelenia sú $\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$, a teda hľadaná pravdepodobnosť je

$$\begin{aligned} \mathbb{P}(X \leq 5) &= \sum_{k=0}^5 e^{-10} \frac{10^k}{k!} \\ &= e^{-10} \left(\frac{10^0}{0!} + \frac{10^1}{1!} + \frac{10^2}{2!} + \frac{10^3}{3!} + \frac{10^4}{4!} + \frac{10^5}{5!} \right) \doteq 0,066 \end{aligned}$$

Pravdepodobnosť, že v celom náklade bude najviac 5 chybné zviazaných kníh je približne 0,066.

Príklad 12. V lotérii sa predá veľa tiketov a pravdepodobnosť, že výhry prvej ceny je veľmi malá. V priemere 2 krát za rok niektorý hráč vyhrá prvú cenu. Aká je pravdepodobnosť, že počas roka nebude ani jeden takýto výherca?

Riešenie. Ide o zriedkavú udalosť (veľa pokusov, malá pravdepodobnosť úspechu), preto počet výhercov prvej ceny za rok modelujeme Poissonovým rozdelením. Máme zadanú strednú hodnotu.

Pravdepodobnosti Poissonovho rozdelenia sú $\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$, pričom stredná hodnota sa rovná λ . V našom prípade je $\lambda = 2$ a zaujíma nás $\mathbb{P}(X = 0)$:

$$\mathbb{P}(X = 0) = e^{-2} \frac{2^0}{0!} = e^{-2} \doteq 0,1353.$$

Pravdepodobnosť, že počas roka nebude ani jeden výherca prvej ceny je približne 0,1353.

3 Spojité náhodné premenné I. (všeobecné)

Príklad 1. Nech X je spojitá náhodná premenná s hustotou

$$f(x) = \frac{4}{81}x^3, \quad x \in (0, 3),$$

a $f(x) = 0$ mimo tohto intervalu.

- (a) Nájdite pravdepodobnosť, že $X \leq 2$.
- (b) Nájdite pravdepodobnosť, že $X < 2$.
- (c) Nájdite pravdepodobnosť, že $X \in (0, 1)$.
- (d) Nájdite pravdepodobnosť, že $X \in [0, 1]$.

Riešenie.

- (a) Pravdepodobnosť nájdeme pomocou integrálu:

$$\mathbb{P}(X \leq 2) = \int_{-\infty}^2 f(x) dx = \int_0^2 \frac{4}{81}x^3 dx = \frac{4}{81} \left[\frac{x^4}{4} \right]_0^2 = \frac{2^4}{81} = \frac{16}{81}.$$

- (b) Výpočet vedie k tonu istému integrálu ako v predchádzajúcom bode:

$$\mathbb{P}(X < 2) = \int_{-\infty}^2 f(x) dx = \int_0^2 \frac{4}{81}x^3 dx = \frac{16}{81}.$$

- (c)

$$\mathbb{P}(0 < X < 1) = \int_0^1 f(x) dx = \int_0^1 \frac{4}{81}x^3 dx = \frac{4}{81} \left[\frac{x^4}{4} \right]_0^1 = \frac{4}{81} \cdot \frac{1}{4} = \frac{1}{81}.$$

- (d) Výpočet vedie k tonu istému integrálu ako v predchádzajúcom bode:

$$\mathbb{P}(0 \leq X \leq 1) = \int_0^1 f(x) dx = \int_0^1 \frac{4}{81}x^3 dx = \frac{1}{81}.$$

Príklad 2. Nech Y je náhodná premenná, ktorej distribučná funkcia je na intervale $(-1, 1)$ daná predpisom

$$F(x) = \begin{cases} 0, & x \leq -1, \\ \frac{1}{4}(x+1)^2, & -1 < x < 1, \\ 1, & x \geq 1. \end{cases}$$

Vypočítajte pravdepodobnosti:

- (a) $\mathbb{P}(Y \leq 0)$,

- (b) $\mathbb{P}(Y \leq 2)$,
- (c) $\mathbb{P}(Y > 0)$,
- (d) $\mathbb{P}(Y \geq 0)$,
- (e) $\mathbb{P}(Y < \frac{1}{2})$,
- (f) $\mathbb{P}(Y \in (-\frac{1}{2}, \frac{1}{2}])$,
- (g) $\mathbb{P}(Y \in (-\frac{1}{2}, \frac{1}{2}))$.

Riešenie.

- (a) Podľa definície platí $F(a) = \mathbb{P}(Y \leq a)$, teda

$$\mathbb{P}(Y \leq 0) = F(0) = \frac{1}{4}.$$

- (b) Analogicky:

$$\mathbb{P}(Y \leq 2) = F(2) = 1$$

- (c) Udalosť $Y > 0$ je komplementom k udalosti: $Y \leq 0$, preto

$$\mathbb{P}(Y > 0) = 1 - \mathbb{P}(Y \leq 0) = 1 - F(0) = 1 - \frac{1}{4} = \frac{3}{4}.$$

- (d) Keďže máme spojitú náhodnú premennú, pravdepodobnosť, že nadobudne jednu konkrétnu hodnotu, je nulová. Preto

$$\mathbb{P}(Y \geq 0) = \mathbb{P}(Y > 0) + \mathbb{P}(Y = 0) = \mathbb{P}(Y > 0).$$

čo sme vypočítali v predchádzajúcom bode, a teda

$$\mathbb{P}(Y \geq 0) = \frac{3}{4}.$$

- (e)

$$\mathbb{P}(Y < \frac{1}{2}) = \mathbb{P}(Y \leq \frac{1}{2}) = F\left(\frac{1}{2}\right) = \frac{9}{16}.$$

- (f) Z definície distribučnej funkcie $F(a) = \mathbb{P}(Y \leq a)$ sme na prednáške odvodili, že

$$F(b) - F(a) = \mathbb{P}(Y \leq b) - \mathbb{P}(Y \leq a) = \mathbb{P}(a < Y \leq b) = \mathbb{P}(Y \in (a, b]),$$

a teda

$$\mathbb{P}(Y \in (-\frac{1}{2}, \frac{1}{2}]) = F\left(\frac{1}{2}\right) - F\left(-\frac{1}{2}\right) = \frac{9}{16} - \frac{1}{16} = \frac{1}{2}.$$

- (g) Keďže máme spojitú náhodnú premennú, pravdepodobnosť, že nadobudne jednu konkrétnu hodnotu, je nulová. Preto pravdepodobnosť, že nadobudne hodnotu z intervalov (a, b) , $[a, b]$, $(a, b]$, $[a, b)$ je rovnaká a rovná sa $F(b) - F(a)$. Teda

$$\mathbb{P}\left(Y \in \left(-\frac{1}{2}, \frac{1}{2}\right)\right) = F\left(\frac{1}{2}\right) - F\left(-\frac{1}{2}\right) = \frac{9}{16} - \frac{1}{16} = \frac{1}{2}.$$

Príklad 3. Čas (v hodinách) potrebný na doručenie balíka v meste závisí od dopravnej situácie a od miesta na trase kuriéra, ktorý balíky doručuje. Nech X je náhodná premenná označujúca čas doručenia balíka od okamihu jeho prevzatia kuriérom. Jej distribučná funkcia je daná predpisom:

$$F(x) = \begin{cases} 0, & x \leq 1, \\ \frac{(x-1)^2}{4}, & 1 < x \leq 2, \\ 1 - \frac{(5-x)^2}{12}, & 2 < x < 5, \\ 1, & x \geq 5. \end{cases}$$

- (a) Vypočítajte pravdepodobnosť, že doručenie potrvá najviac 2 hodiny.
- (b) Vypočítajte pravdepodobnosť, že doručenie potrvá viac ako 3 hodiny.
- (c) Vypočítajte pravdepodobnosť, že čas doručenia bude trvať viac ako 90 minút, ale menej ako 4 hodiny.
- (d) Odvoďte hustotu náhodnej premennej X .
- (e) Vypočítajte strednú hodnotu $\mathbb{E}(X)$.

Riešenie. Prvé tri zadania riešime ako v predchádzajúcom príklade:

- (a)

$$\mathbb{P}(X \leq 2) = F(2) = \frac{(2-1)^2}{4} = \frac{1}{4}.$$

- (b)

$$\mathbb{P}(X > 3) = 1 - F(3), \quad F(3) = 1 - \frac{(5-3)^2}{12} = 1 - \frac{4}{12} = \frac{2}{3}$$

- (c) 90 minút je 1,5 hodiny a vieme, že pravdepodobnosť nadobudnutia hodnôt medzi a a b je vždy $F(b) - F(a)$ bez ohľadu na to, či sa zahŕňajú aj krajné body. Preto:

$$\mathbb{P}(1,5 < X < 4) = F(4) - F(1,5) = \frac{11}{12} - \frac{1}{16} = \frac{41}{48}.$$

Pokračujeme ďalej:

(c) Hustota náhodnej premennej $f(x)$ je deriváciou distribučnej funkcie:

$$f(x) = F'(x) = \begin{cases} \frac{x-1}{2} & \text{pre } 1 < x \leq 2, \\ \frac{5-x}{6} & \text{pre } 2 < x < 5, \\ 0 & \text{inak.} \end{cases}$$

Na obr 3 je zobrazený graf distribučnej funkcie a hustoty

(d) Stredná hodnota $\mathbb{E}(X)$ spojitej náhodnej premennej sa počíta ako $\mathbb{E}(X) = \int_{-\infty}^{\infty} xf(x)dx$. V našom prípade:

$$\begin{aligned} \mathbb{E}(X) &= \int_{-\infty}^{\infty} xf(x)dx = \int_1^2 x \frac{x-1}{2} dx + \int_2^5 x \frac{5-x}{6} dx \\ &= \left[\frac{x^3}{6} - \frac{x^2}{4} \right]_1^2 + \left[\frac{5x^2}{12} - \frac{x^3}{18} \right]_2^5 \\ &= \left(\frac{8}{6} - \frac{4}{4} \right) - \left(\frac{1}{6} - \frac{1}{4} \right) + \left(\frac{125}{12} - \frac{125}{18} \right) - \left(\frac{20}{12} - \frac{8}{18} \right) = \frac{8}{3}. \end{aligned}$$

Stredná hodnota času doručenia je $\frac{8}{3}$ hodiny, teda 2 hodiny 40 minút

Príklad 4. Vypočítajte strednú hodnotu náhodnej premennej X z príkladu 1.

Riešenie. V tomto príklade sme mali zadanú hustotu

$$f(x) = \frac{4}{81}x^3, \quad x \in (0, 3),$$

a $f(x) = 0$ mimo tohto intervalu. Takže ju len dosadíme do vzťahu $\mathbb{E}(X) = \int_{-\infty}^{\infty} xf(x)dx$. Dostaneme:

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_0^3 x \frac{4}{81}x^3 dx = \frac{4}{81} \left[\frac{x^5}{5} \right]_0^3 = \frac{4 \cdot 3^5}{5 \cdot 81} = \frac{12}{5}.$$

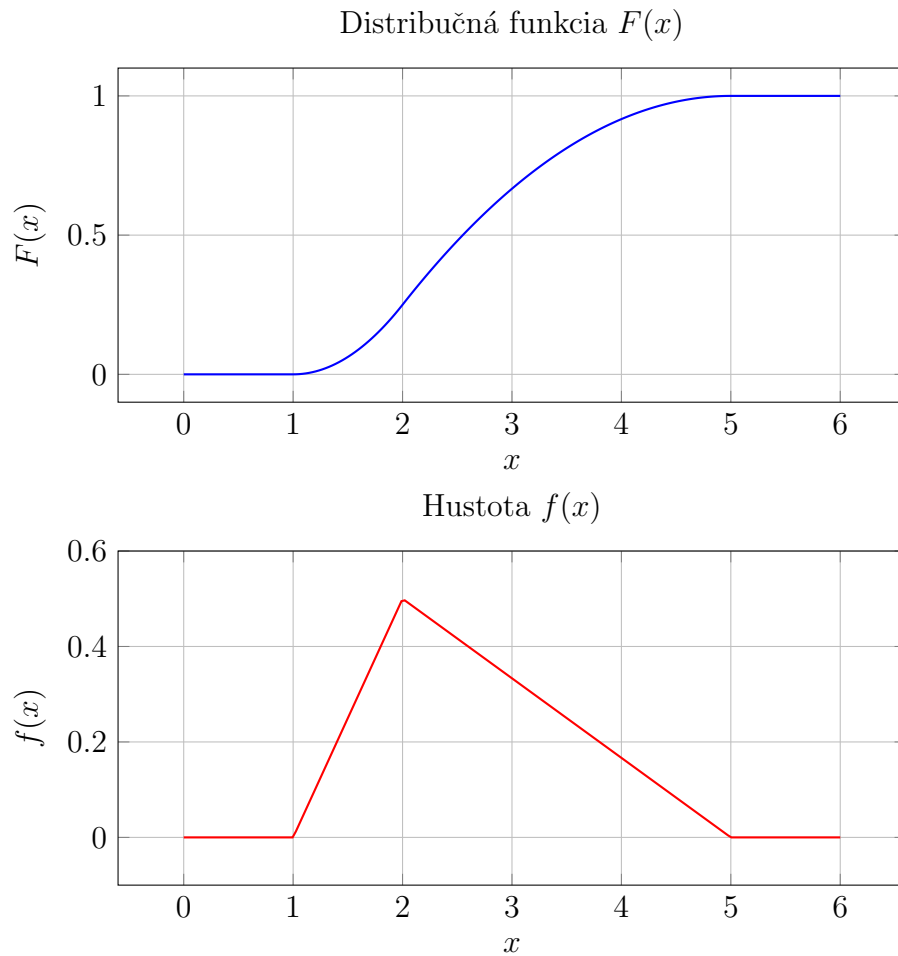
Príklad 5. Vypočítajte strednú hodnotu náhodnej premennej Y z príkladu 2.

Riešenie. V tomto príklade sme mali zadanú distribučnú funkciu

$$F(x) = \begin{cases} 0, & x \leq -1, \\ \frac{1}{4}(x+1)^2, & -1 < x < 1, \\ 1, & x \geq 1. \end{cases}$$

Z nej najskôr vypočítame hustotu a tú potom dosadíme do vzťahu $\mathbb{E}(Y) = \int_{-\infty}^{\infty} xf(x)dx$. Hustota je deriváciou distribučnej funkcie:

$$f(x) = F'(x) = \begin{cases} \frac{1}{2}(x+1), & -1 < x < 1, \\ 0, & \text{inak.} \end{cases}$$



Obr. 3: Distribučná funkcia a hustota

Stredná hodnota potom je

$$\begin{aligned}\mathbb{E}(Y) &= \int_{-\infty}^{\infty} x f(x) dx = \int_{-1}^1 x \frac{x+1}{2} dx = \frac{1}{2} \int_{-1}^1 (x^2 + x) dx \\ &= \frac{1}{2} \left[\frac{x^3}{3} + \frac{x^2}{2} \right]_{-1}^1 = \frac{1}{2} \left[\left(\frac{1}{3} + \frac{1}{2} \right) - \left(-\frac{1}{3} + \frac{1}{2} \right) \right] = \frac{1}{3}.\end{aligned}$$

Príklad 6. Vypočítajte disperziu náhodnej premennej Y z predchádzajúceho príkladu.

Riešenie. Disperzia je

$$\mathbb{D}(Y) = \mathbb{E}(Y^2) - (\mathbb{E}(Y))^2,$$

pričom sme v predchádzajúcom príklade vypočítali $\mathbb{E}(Y) = \frac{1}{3}$. Zostáva vypočítať $\mathbb{E}(Y^2)$, pričom využijeme hustotu odvodenú v predchádzajúcom príklade.

Vo všeobecnosti je $\mathbb{E}(Y^2) = \int_{-\infty}^{\infty} x^2 f(x) dx$. V našom prípade je to

$$\begin{aligned}\mathbb{E}(Y^2) &= \int_{-\infty}^{\infty} x^2 f(x) dx = \int_{-1}^1 x^2 \frac{x+1}{2} dx = \frac{1}{2} \left[\frac{x^4}{4} + \frac{x^3}{3} \right]_{-1}^1 \\ &= \frac{1}{2} \left[\left(\frac{1}{4} + \frac{1}{3} \right) - \left(\frac{1}{4} - \frac{1}{3} \right) \right] = \frac{1}{3}.\end{aligned}$$

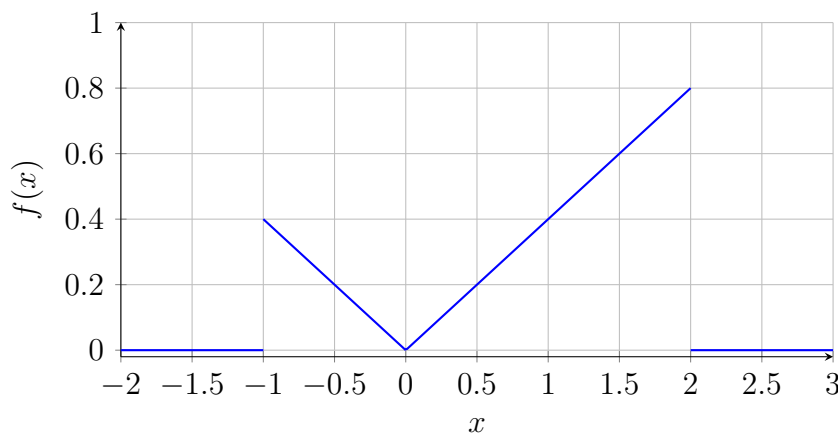
Disperzia teda je

$$\mathbb{D}(Y) = \mathbb{E}(Y^2) - (\mathbb{E}(Y))^2 = \frac{1}{3} - \left(\frac{1}{3} \right)^2 = \frac{2}{9}.$$

Príklad 7. Vypočítajte strednú hodnotu a disperziu náhodnej premennej X s hustotou

$$f(x) = \begin{cases} \frac{2}{5}|x| & \text{pre } -1 < x < 2, \\ 0 & \text{inak,} \end{cases}$$

ktorá je znázornená na obrázku 4.



Obr. 4: Graf hustoty náhodnej premennej X

Riešenie. Stredná hodnota je

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x) dx = \frac{2}{5} \int_{-1}^2 x|x| dx = \frac{2}{5} \int_{-1}^0 x \cdot (-x) dx + \frac{2}{5} \int_0^2 x \cdot x dx,$$

kde sme využili, že $|x| = -x$ pre $x \leq 0$ a $|x| = x$ pre $x \geq 0$. Vypočítame jednotlivé integrály:

$$\begin{aligned}\int_{-1}^0 x \cdot (-x) dx &= - \int_{-1}^0 x^2 dx = - \left[\frac{x^3}{3} \right]_{-1}^0 = -\frac{1}{3}, \\ \int_0^2 x \cdot x dx &= \left[\frac{x^3}{3} \right]_0^2 = \frac{8}{3},\end{aligned}$$

a teda

$$\mathbb{E}(X) = \frac{2}{5} \int_{-1}^0 x \cdot (-x) dx + \frac{2}{5} \int_0^2 x \cdot x dx = \frac{2}{5} \cdot \left(-\frac{1}{3}\right) + \frac{2}{5} \cdot \frac{8}{3} = \frac{14}{15}.$$

K výpočtu disperzie budeme potrebovať $\mathbb{E}(X^2)$:

$$\begin{aligned}\mathbb{E}(X^2) &= \int_{-\infty}^{\infty} x^2 f(x) dx = \frac{2}{5} \int_{-1}^2 x^2 |x| dx = \frac{2}{5} \int_{-1}^0 x^2 \cdot (-x) dx + \frac{2}{5} \int_0^2 x^2 \cdot x dx \\ &= -\frac{2}{5} \int_{-1}^0 x^3 dx + \frac{2}{5} \int_0^2 x^3 dx = -\frac{2}{5} \left[\frac{x^4}{4}\right]_{-1}^0 + \frac{2}{5} \left[\frac{x^4}{4}\right]_0^2 = \frac{2}{20} + \frac{32}{20} = \frac{17}{10}.\end{aligned}$$

Disperzia potom je

$$\mathbb{D}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \frac{17}{10} - \left(\frac{14}{15}\right)^2 = \frac{373}{450}.$$

Príklad 8. Hrana kocky má náhodnú dĺžku X centimerov, pričom hustota náhodnej premennej X je

$$f(x) = \frac{1}{2}x, \quad x \in (0, 2),$$

a $f(x) = 0$ mimo tohto intervalu. Kocky sú vyrobené z hliníka, ktorého hustota je $2,7 \text{ g/cm}^3$. Nájdite

- (a) strednú hodnotu a disperziu povrchu kocky,
- (b) strednú hodnotu a disperziu objemu kocky,
- (c) strednú hodnotu hmotnosti balenia, ktoré pozostáva z papierového vrečka s hmotnosťou 0,05 a piatich kociek,.

Riešenie. Stredná hodnota náhodnej premennej $g(X)$, ak X má hustotu f je $\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x)f(x) dx$. Napríklad sme počítali $\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx$. Analogicky sa dá vypočítať napr. $\mathbb{E}(X^3)$, $\mathbb{E}(\sqrt{X})$, $\mathbb{E}(e^X)$, atď.

- (a) Povrch kocky s hranou X je $S = 6X^2$.

$$\mathbb{E}(S) = \mathbb{E}(6X^2) = 6\mathbb{E}(X^2).$$

Vypočítame $\mathbb{E}[X^2]$:

$$\mathbb{E}(X^2) = \int_0^2 x^2 \cdot \frac{1}{2}x dx = \frac{1}{2} \int_0^2 x^3 dx = \frac{1}{2} \left[\frac{x^4}{4}\right]_0^2 = \frac{1}{2} \cdot \frac{16}{4} = 2.$$

Teda

$$\mathbb{E}(S) = 6\mathbb{E}(X^2) = 6 \cdot 2 = 12.$$

Disperzia je

$$\mathbb{D}(S) = \mathbb{D}(6X^2) = 36\mathbb{D}(X^2) = 36(\mathbb{E}(X^4) - (\mathbb{E}(X^2))^2).$$

Potrebujeeme $E(X^4)$:

$$E(X^4) = \int_0^2 x^4 \cdot \frac{1}{2}x \, dx = \frac{1}{2} \int_0^2 x^5 \, dx = \frac{1}{2} \left[\frac{x^6}{6} \right]_0^2 = \frac{1}{2} \cdot \frac{64}{6} = \frac{16}{3}.$$

Teda

$$\mathbb{D}(S) = 36(E(X^4) - (E(X^2))^2) = 36 \left(\frac{16}{3} - 4 \right) = 48.$$

- (b) Výpočet pre objem je veľmi podobný: Objem kocky je $V = X^3$ a stredná hodnota tohto objemu

$$\mathbb{E}(V) = \mathbb{E}(X^3) = \int_0^2 x^3 \cdot \frac{1}{2}x \, dx = \frac{1}{2} \int_0^2 x^4 \, dx = \frac{1}{2} \cdot \frac{32}{5} = \frac{16}{5}$$

Disperzia je

$$\mathbb{D}(V) = \mathbb{D}(X^3) = \mathbb{E}(X^6) - (\mathbb{E}(X^3))^2.$$

Spočítame $\mathbb{E}(X^6)$:

$$\mathbb{E}(X^6) = \int_0^2 x^6 \cdot \frac{1}{2}x \, dx = \frac{1}{2} \int_0^2 x^7 \, dx = \frac{1}{2} \left[\frac{x^8}{8} \right]_0^2 = \frac{1}{2} \cdot \frac{256}{8} = 16.$$

Teda

$$\mathbb{D}(V) = 16 - \left(\frac{16}{5} \right)^2 = \frac{144}{25}.$$

- (c) Objemy piatich kociek sú $V_1, \dots, V_5$ a celkový objem hliníka, z ktorého sú vyrobené, je $V_1 + \dots + V_5$ centimetrov kubických. Jeden centimeter kubický hliníka váži podľa zadania 2,7 gramov. Hmotnosť všetkých piatich kociek potom je $2,7(V_1 + \dots + V_5)$ gramov a spolu s vreckom je hmotnosť balenia

$$Y = 0,05 + 2,7(V_1 + \dots + V_5).$$

Stredná hodnota tejto hmotnosti je

$$\begin{aligned} \mathbb{E}(Y) &= \mathbb{E}(0,05 + 2,7(V_1 + \dots + V_5)) = 0,05 + 2,7(\mathbb{E}(V_1) + \dots + \mathbb{E}(V_5)) \\ &= 0,05 + 2,7 \cdot 5 \cdot \frac{16}{5} = 43,25. \end{aligned}$$

Stredná hodnota hmotnosti balenia je 43,25 gramov.

Príklad 9. Uvažujme hustotu náhodnej premennej X z predchádzajúceho príkladu:

$$f(x) = \frac{1}{2}x, \quad x \in (0, 2),$$

a $f(x) = 0$ mimo tohto intervalu. Odvoďte postup, ako generovať realizácie náhodnej premennej s týmto rozdelením, ak máme k dispozícii generátor náhodných čísel z rovnomerného rozdelenia na intervale $(0, 1)$

Riešenie. Univerzálny postup je taký, že generujeme realizácie náhodnej premennej U s rovnomerným rozdelením na $(0, 1)$ a dosadzujeme ich do inverznej funkcie k distribučnej funkcii nášho rozdelenia. Teda naše hodnoty X budú $X = F^{-1}(U)$. To znamená, že vyhovujú rovnosti $U = F(X)$, z ktorej pre dané U vyjadríme X .

V našom prípade máme zadanú hustotu (nie distribučnú funkciu), preto prvým krokom bude nájdenie distribučnej funkcie F . Vo všeobecnosti je $F(x) = \int_{-\infty}^x f(t)dt$, takže pre našu hustotu f dostaneme:

- pre $x \leq 0$: $F(x) = 0$

- pre $x \in (0, 2)$:

$$F(x) = \int_0^x \frac{1}{2}t dt = \frac{1}{2} \left[\frac{t^2}{2} \right]_0^x = \frac{x^2}{4}$$

- pre $x \geq 2$: $F(x) = 1$

Po vygenerovaní U z rovnomerného rozdelenia na $(0, 1)$ hľadáme také $X \in (0, 2)$, že $F(X) = U$. To znamená, že:

$$\begin{aligned} \frac{X^2}{4} &= U, \\ X^2 &= 4U. \end{aligned}$$

Tomuto vyhovujú dve hodnoty: $\pm\sqrt{4U}$, naša však má byť kladná, takže zoberieme

$$X = \sqrt{4U} = 2\sqrt{U}.$$

Generovanie teda prebieha tak, že vygenerujeme hodnotu U z rovnomerného rozdelenia na intervale $(0, 1)$ a potom zoberieme $2\sqrt{U}$.

```
import numpy as np

np.random.seed(42)          # kvoli reprodukovateľnosti
n = 10**6                   # počet simulácii
U = np.random.uniform(0, 1, n) # uniform
X = 2*np.sqrt(U)            # hrana kocky
```

Obr. 5: Generovanie náhodných čísel.

Poznámka. Pomocou tohto generátora si môžeme jednoducho skontrolovať výpočty strednej hodnoty a disperziu povrchu a objemu kocky z predchádzajúceho príkladu. Vygenerujeme hrany kocky práve odvodeným postupom, vypočítame príslušné povrchy a objemy a zo získaných hodnôt odhadneme strednú hodnotu (priemerom) a disperziu (výberovou disperziou). Výsledky sú na obrázku 6, vidíme, že sú veľmi blízke vypočítaným hodnotám.

```

# povrch kocky
S = 6*X**2
priemer_S = np.mean(S)
disperzia_S = np.var(S, ddof=1)
print("Povrch kocky")
print(f"Odhad strednej hodnoty: {priemer_S:.4f}")
print(f"Odhad disperzie: {disperzia_S:.4f}")

# objem kocky
V = X**3
priemer_V = np.mean(V)
disperzia_V = np.var(V, ddof=1)
print("\nObjem kocky")
print(f"Odhad strednej hodnoty: {priemer_V:.4f}")
print(f"Odhad disperzie: {disperzia_V:.4f}")

```

```

Povrch kocky
Odhad strednej hodnoty: 12.0080
Odhad disperzie: 47.9721

```

```

Objem kocky
Odhad strednej hodnoty: 3.2025
Odhad disperzie: 5.7596

```

Obr. 6: Odhad strednej hodnoty a disperzie.

Príklad 10. Odvodte hustotu náhodnej premennej vyjadrujúcej povrch kocky, ak je jej hrana náhodná s hustotou danou v predchádzajúcich dvoch príkladoch.

Riešenie. Máme hustotu náhodnej premennej X v tvare

$$f(x) = \frac{1}{2}x \quad \text{pre } x \in (0, 2),$$

a $f(x) = 0$ mimo tohto intervalu. Zaujímá nás hustota náhodnej premennej $S = 6X^2$. Táto náhodná premenná môže nadobúdať iba hodnoty z intervalu $(0, 6 \cdot 2^2)$, teda $(0, 24)$. Mimo tohto intervalu je teda hustota $f_S(x)$ nulová. Na intervale $(0, 24)$ najskôr nájdeme distribučnú funkciu, jej zderivovaním dostaneme hustotu.

Označme f hustotu a F distribučnú funkciu náhodnej premennej X ; f_S hustotu a F_S distribučnú funkciu náhodnej premennej S . Potom pre $x \in (0, 24)$:

$$F_S(x) = \mathbb{P}(S \leq x) = \mathbb{P}(6X^2 \leq x) = \mathbb{P}\left(X^2 \leq \frac{x}{6}\right) \stackrel{(1)}{=} \mathbb{P}\left(X \leq \sqrt{\frac{x}{6}}\right) \stackrel{(2)}{=} F\left(\sqrt{\frac{x}{6}}\right),$$

kde rovnosť (1) vyplýva z toho, že o náhodnej premennej X vieme, že nadobúda iba kladné hodnoty a rovnosť (2) vyplýva z definície distribučnej funkcie. Derivujeme

ako zloženú funkciu:

$$f_S(x) = F'_S(x) = F' \left(\sqrt{\frac{x}{6}} \right) \cdot \left(\sqrt{\frac{x}{6}} \right)'$$

Jednotlivé členy sú:

- $F' \left(\sqrt{\frac{x}{6}} \right) = f \left(\sqrt{\frac{x}{6}} \right)$, lebo derivácia distribučnej funkcie je hustota
- $\left(\sqrt{\frac{x}{6}} \right)' = \left(\frac{\sqrt{x}}{\sqrt{6}} \right)' = \frac{1}{2\sqrt{6}\sqrt{x}}$

Teda

$$f_S(x) = f \left(\sqrt{\frac{x}{6}} \right) \frac{1}{2\sqrt{6}\sqrt{x}},$$

kde napokon dosadíme hustotu f náhodnej premennej X (všimnime si, že vďaka tomu, že $x \in (0, 24)$, argument $\sqrt{\frac{x}{6}}$ je z intervalu $(0, 2)$, na ktorom je $f(x) = \frac{1}{2}x$):

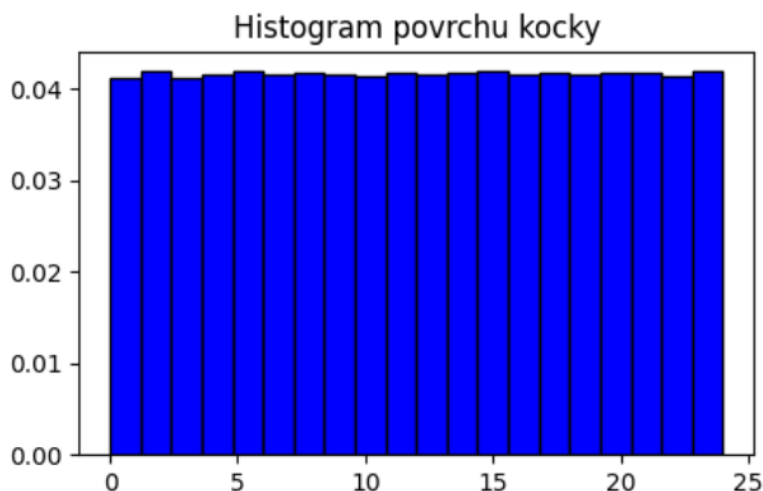
$$f_S(x) = \frac{1}{2} \sqrt{\frac{x}{6}} \frac{1}{2\sqrt{6}\sqrt{x}} = \frac{1}{24} \quad \text{pre } x \in (0, 24).$$

Mimo tohto intervalu je hustota nulová.

Poznámka. Na obrázku 7 zobrazíme na kontrolu histogram vygenerovaných hodnôt povrchu, vidíme, že zodpovedá vypočítanej hustote.

```
import matplotlib.pyplot as plt

plt.figure(figsize=(5, 3))
plt.hist(S, bins=20, density=True, color='blue', edgecolor='black')
plt.title('Histogram povrchu kocky')
plt.show()
```



Obr. 7: Histogram hodnôt povrchu kocky - na porovnanie s vypočítanou hustotou.

Príklad 11. Odvod'te hustotu náhodnej premennej vyjadrujúcej objem kocky, ak je jej hrana náhodná s hustotou danou v predchádzajúcich príkladoch.

Riešenie. Máme hustotu náhodnej premennej X v tvare

$$f(x) = \frac{1}{2}x \quad \text{pre } x \in (0, 2),$$

a $f(x) = 0$ mimo tohto intervalu. Zaujímá nás hustota náhodnej premennej $V = X^3$. Táto náhodná premenná môže nadobúdať iba hodnoty z intervalu $(0, 8)$. Mimo tohto intervalu je teda hustota $f_S(x)$ nulová. Na intervale $(0, 8)$ najskôr nájdeme distribučnú funkciu, jej zderivovaním dostaneme hustotu.

Označme f hustotu a F distribučnú funkciu náhodnej premennej X ; f_V hustotu a F_V distribučnú funkciu náhodnej premennej V . Potom pre $x \in (0, 8)$:

$$F_V(x) = \mathbb{P}(V \leq x) = \mathbb{P}(X^3 \leq x) = \mathbb{P}\left(X \leq x^{\frac{1}{3}}\right) \stackrel{(1)}{=} F\left(x^{\frac{1}{3}}\right),$$

kde rovnosť (1) vyplýva z definície distribučnej funkcie. Derivujeme ako zloženú funkciu:

$$f_V(x) = F'_V(x) = F'\left(x^{\frac{1}{3}}\right) \cdot \left(x^{\frac{1}{3}}\right)'$$

Jednotlivé členy sú:

- $F'\left(x^{\frac{1}{3}}\right) = f\left(x^{\frac{1}{3}}\right)$, lebo derivácia distribučnej funkcie je hustota
- $\left(x^{\frac{1}{3}}\right)' = \frac{1}{3}x^{-\frac{2}{3}}$

Teda

$$f_S(x) = f\left(x^{\frac{1}{3}}\right) \cdot \frac{1}{3}x^{-\frac{2}{3}},$$

kde napokon dosadíme hustotu f náhodnej premennej X (všimnime si, že vďaka tomu, že $x \in (0, 8)$, argument $x^{\frac{1}{3}}$ je z intervalu $(0, 2)$, na ktorom je $f(x) = \frac{1}{2}x$):

$$f_S(x) = \frac{1}{2}x^{\frac{1}{3}} \cdot \frac{1}{3}x^{-\frac{2}{3}} = \frac{1}{6}x^{-\frac{1}{3}} \quad \text{pre } x \in (0, 8).$$

Mimo tohto intervalu je hustota nulová.

Poznámka. Na obrázku 8 zobrazíme na kontrolu histogram vygenerovaných hodnôt objemu, do ktorého pridáme vypočítanú hustotu - vidíme, že sú podobné.

Príklad 12. Odvod'te postup generovania náhodných čísel z rozdelenia daného hustotou

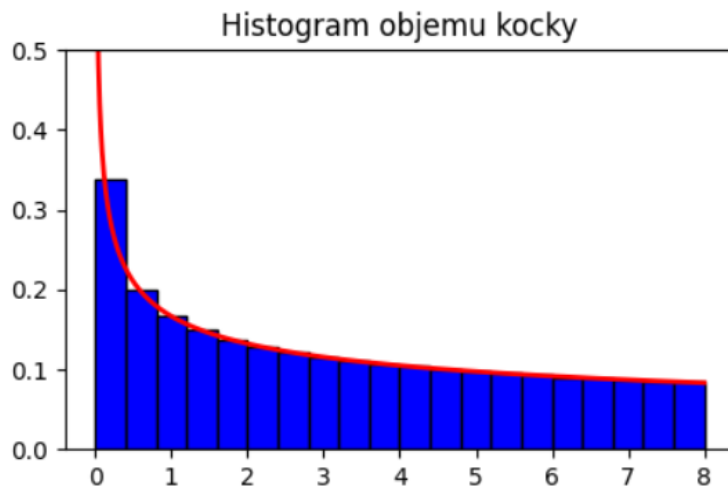
$$f(x) = \begin{cases} 3x^2, & \text{pre } x \in (0, 1), \\ 0, & \text{inak.} \end{cases}$$

ak máme k dispozícii generátor náhodných čísel z rovnomerného rozdelenia na intervale $(0, 1)$.

```
import matplotlib.pyplot as plt

plt.figure(figsize=(5, 3))
plt.hist(v, bins=20, density=True, color='blue', edgecolor='black')
plt.title('Histogram objemu kocky')

x = np.linspace(0.001, 8, 1000)
density = (1/6) * x**(-1/3)
plt.plot(x, density, linewidth=2, color='red')
plt.ylim([0, 0.5])
plt.show()
```



Obr. 8: Histogram hodnôt objemu kocky a porovnanie s vypočítanou hustotou (červená čiara).

Riešenie. Pripomeňme si univerzálny postup: generujeme realizácie náhodnej premennej U s rovnomerným rozdelením na $(0, 1)$ a dosadzujeme ich do inverznej funkcie k distribučnej funkcii nášho rozdelenia. Teda naše hodnoty X budú $X = F^{-1}(U)$. To znamená, že vyhovujú rovnosti $U = F(X)$, z ktorej pre dané U vyjadríme X .

V našom prípade máme zadanú hustotu (nie distribučnú funkciu), preto prvým krokom bude nájdenie distribučnej funkcie F . Vo všeobecnosti je $F(x) = \int_{-\infty}^x f(t)dt$, takže pre našu hustotu f dostaneme:

- pre $x \leq 0$: $F(x) = 0$
- pre $x \in (0, 1)$;

$$F(x) = \int_0^x 3t^2 dt = 3 \left[\frac{t^3}{3} \right]_0^x = x^3$$

- pre $x \geq 1$: $F(x) = 1$

Po vygenerovaní U z rovnomerného rozdelenia na $(0, 1)$ hľadáme také $X \in (0, 1)$, že $F(X) = U$. To znamená, že:

$$\begin{aligned} X^3 &= U, \\ X &= U^{\frac{1}{3}}. \end{aligned}$$

Generovanie teda prebieha tak, že vygenerujeme hodnotu U z rovnomerného rozdelenia na intervale $(0, 1)$ a potom zoberieme $U^{\frac{1}{3}}$.

Príklad 13. Odvodte postup generovania náhodných čísel z rozdelenia daného distribučnou funkciou

$$F(x) = \begin{cases} 0 & \text{pre } x < 0, \\ 1 - e^{-\left(\frac{x}{2}\right)^4} & \text{pre } x \geq 0. \end{cases}$$

ak máme k dispozícii generátor náhodných čísel z rovnomerného rozdelenia na intervale $(0, 1)$.

Riešenie. Tentokrát máme zadanú priamo distribučnú funkciu. Teda po vygenerovaní U z rovnomerného rozdelenia na $(0, 1)$ hľadáme také $X \geq 0$, že $F(X) = U$. To znamená, že:

$$\begin{aligned} 1 - e^{-\left(\frac{x}{2}\right)^4} &= U, \\ e^{-\left(\frac{x}{2}\right)^4} &= 1 - U, \\ -\left(\frac{x}{2}\right)^4 &= \ln(1 - U), \\ x^4 &= -2^4 \ln(1 - U). \end{aligned}$$

Keďže hľadáme $X \geq 0$, vyhovuje len $X = 2(-\ln(1 - U))^{\frac{1}{4}}$.

Generovanie teda prebieha tak, že vygenerujeme hodnotu U z rovnomerného rozdelenia na intervale $(0, 1)$ a potom zoberieme $X = 2(-\ln(1 - U))^{1/4}$.

4 Spojité náhodné premenné II. (konkrétne rozdelenia: rovnomerné, exponenciálne)

Príklad 1. Budovy v meste majú štvorcový pôdorys, pričom strana štvorca má rovnomerné rozdelenie na intervale $(6, 15)$.

- (a) Vypočítajte strednú hodnotu plochy budovy.
- (b) Nájdite hustotu plochy budovy.

Riešenie. Strana štvorca X má rovnomerného rozdelenie na intervale $(6, 15)$, teda jej hustota je

$$f(x) = \begin{cases} \frac{1}{9} & \text{pre } x \in (6, 15), \\ 0 & \text{inak.} \end{cases}$$

- (a) Plocha budovy je X^2 , teda jej stredná hodnota je

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_6^{15} x^2 \frac{1}{9} dx = \frac{1}{9} \left[\frac{x^3}{3} \right]_6^{15} = \frac{3159}{27} = 117.$$

Iný postup: Stredná hodnota náhodnej premennej s rovnomerným rozdelením na intervale je stred toho intervalu, teda $\mathbb{E}(X) = \frac{21}{2}$. Vo vzorcovníku nájdeme vzťah pre disperziu tohto rozdelenia: $\mathbb{D}(X) = \frac{(15-6)^2}{12} = \frac{81}{12} = \frac{27}{4}$. Teraz zo vzorca pre disperziu $\mathbb{D}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$ vyjadríme hľadanú hodnotu:

$$\mathbb{E}(X^2) = \mathbb{D}(X) + (\mathbb{E}(X))^2 = \frac{27}{4} + \left(\frac{21}{2}\right)^2 = \frac{468}{4} = 117$$

- (b) Plocha nadobúda hodnoty z intervalu $(36, 225)$, teda mimo tohto intervalu je jej hustota nulová. Na intervale $(36, 225)$ nájdeme distribučnú funkciu a jej derivovaním hustotu.

Označme F distribučnú funkciu strany štvorca X a f jej hustotu. Plochu označíme Y , jej distribučnú funkciu F_Y a hustotu f_Y . Pre $x \in (36, 225)$ je:

$$F_Y(x) = \mathbb{P}(Y \leq x) = \mathbb{P}(X^2 \leq x) \stackrel{(1)}{=} \mathbb{P}(X \leq \sqrt{x}) \stackrel{(2)}{=} F(\sqrt{x}).$$

Derivujeme:

$$f_Y(x) = F'_Y(x) \stackrel{(3)}{=} F'(\sqrt{x})(\sqrt{x})' \stackrel{(4)}{=} f(\sqrt{x}) \frac{1}{2\sqrt{x}} = \frac{1}{9 \cdot 2\sqrt{x}} = \frac{1}{18\sqrt{x}}$$

Vysvetlenie krokov:

- (1) Využili sme, že X aj x sú kladné, vtedy $X^2 \leq x \iff X \leq \sqrt{x}$
- (2) Definícia distribučnej funkcie.

- (3) Výsledný výraz z predchádzajúceho riadku derivujeme ako zloženú funkciu.
- (4) Derivácia distribučnej funkcie F je hustota f .

Teda hustota plochy domu je

$$f_Y(x) = \begin{cases} \frac{1}{18\sqrt{x}} & \text{pre } x \in (36, 225), \\ 0 & \text{inak.} \end{cases}$$

Všimnime si, že pomocou hustoty f_Y by sme integrovaním mohli nájsť strednú hodnotu $\mathbb{E}(Y)$ požadovanú v predchádzajúcom bode zadania. V prípade, že máme vypočítať len strednú hodnotu však máme aj kratšie postupy, ktoré sme videli v riešení časti (a).

Príklad 2. Výška škody má rovnomerné rozdelenie na intervale $(0, 10)$. Poistenie má spoluúčasť 2, čo znamená, že ak škoda túto hodnotu nedosiahne, poisťovňa nevyplatí nič a ak je škoda vyššia, poisťovňa vyplatí výšku škody zníženú o 2. Aká je stredná hodnota platby, ktorú poisťovňa vyplatí?

Riešenie. Nech X je výška škody, potom výplata poisťovne Y je pomocou X vyjadrená nasledovne:

$$Y = g(X) = \begin{cases} X - 2, & \text{pre } X \in (2, 10) \\ 0, & \text{inak.} \end{cases}$$

Náhodná premenná X má hustotu

$$f(x) = \frac{1}{10}, \quad 0 < x < 10.$$

Stredná hodnota platby poisťovne Y je

$$\mathbb{E}(Y) = \int_{-\infty}^{\infty} g(x)f(x) dx \stackrel{(1)}{=} \int_0^{10} g(x)\frac{1}{10} dx$$

kde sme v kroku (1) dosadili hustotu f . Teraz rozdelíme integrál podľa definície funkcie g :

$$\begin{aligned} \mathbb{E}(Y) &= \int_0^2 0 \cdot \frac{1}{10} dx + \int_2^{10} (x - 2) \cdot \frac{1}{10} dx = \frac{1}{10} \int_2^{10} (x - 2) dx \\ &= \frac{1}{10} \left[\frac{x^2}{2} - 2x \right]_2^{10} = \frac{32}{10} = 3,2. \end{aligned}$$

Stredná hodnota platby poisťovne je 3,2.

Príklad 3. Uvažujme situáciu z predchádzajúceho príkladu. Ako má poisťovňa nastaviť spoluúčasť, aby stredná hodnota platby, ktorú vyplatí, bola 3?

Riešenie. Budeme postupovať úplne analogicky ako v predchádzajúcom príklade, len výška spoluúčasti bude parameter $a \in (0, 10)$. Výsledná stredná hodnota vyplatenej sumy bude závisieť od a , a nakoniec zistíme, pri akej hodnote tohto parametra to bude požadovaná hodnota 3.

Nech X je výška škody, potom výplata poisťovne Y je pomocou X vyjadrená nasledovne:

$$Y = g(X) = \begin{cases} X - a, & \text{pre } X \in (a, 10) \\ 0, & \text{inak.} \end{cases}$$

Náhodná premenná X má hustotu

$$f(x) = \frac{1}{10}, \quad 0 < x < 10.$$

Stredná hodnota platby poisťovne Y je

$$\mathbb{E}(Y) = \int_{-\infty}^{\infty} g(x)f(x) dx \stackrel{(1)}{=} \int_0^{10} g(x)\frac{1}{10} dx$$

kde sme v kroku (1) dosadili hustotu f . Teraz rozdelíme integrál podľa definície funkcie g :

$$\begin{aligned} \mathbb{E}(Y) &= \int_0^a 0 \cdot \frac{1}{10} dx + \int_a^{10} (x - a) \cdot \frac{1}{10} dx = \frac{1}{10} \int_a^{10} (x - a) dx \\ &= \frac{1}{10} \left[\frac{x^2}{2} - ax \right]_a^{10} = \frac{1}{10} \left(50 - 10a + \frac{a^2}{2} \right) = 5 - a + \frac{a^2}{20}. \end{aligned}$$

Stredná hodnota platby poisťovne je $5 - a + \frac{a^2}{20}$. Riešime teda rovnicu

$$\begin{aligned} 5 - a + \frac{a^2}{20} &= 0, \\ \frac{a^2}{20} - a + 2 &= 0, \\ a^2 - 20a + 40 &= 0. \end{aligned}$$

Táto kvadratická rovnica má dva korene:

$$a_{1,2} = \frac{20 \pm \sqrt{240}}{2} = \frac{20 \pm \sqrt{16 \cdot 15}}{2} = \frac{20 \pm 4\sqrt{15}}{2} = 10 \pm 2\sqrt{15}$$

Naša hľadaná hodnota spoluúčasti však musí byť menšia ako 10, takže poisťovňa ju musí nastaviť na hodnotu

$$a = 10 - 2\sqrt{15} \doteq 2,254.$$

Príklad 4. Nech polomer gulôčky v centimetroch má rovnomerné rozdelenie na intervale $(1, 2)$. Odvodte distribučnú funkciu jej objemu. (Objem gule s polomerom R je $V = \frac{4}{3}\pi R^3$.)

Riešenie. Polomer R má rovnomerné rozdelenie na intervale $(1, 2)$, nech $V = \frac{4}{3}\pi R^3$ je jej objem. Objem nadobudne hodnoty z intervalu $(\frac{4}{3}\pi \cdot 8, \frac{4}{3}\pi \cdot 8)$, teda pre $x \leq \frac{4}{3}\pi$ je $F_V(x) = 0$ a pre $x \geq \frac{32}{3}\pi$ je $F_V(x) = 1$.

Pre $x \in (\frac{4}{3}\pi, \frac{32}{3}\pi)$ je:

$$F_V(x) = \mathbb{P}(V \leq x) = \mathbb{P}\left(\frac{4}{3}\pi R^3 \leq x\right) = \mathbb{P}\left(R \leq \sqrt[3]{\frac{3x}{4\pi}}\right) = F_R\left(\sqrt[3]{\frac{3x}{4\pi}}\right),$$

kde F_R je distribučná funkcia R . Všimnime si, že argument $\sqrt[3]{\frac{3x}{4\pi}}$ je presne z intervalu $(1, 2)$, teda z oboru hodnôt R .

O náhodnej premennej R vieme, že má rovnomerné rozdelenie na intervale $(1, 2)$, a teda má hustotu

$$f_R(x) = 1, \quad 1 < x < 2.$$

Nájdeme jej distribučnú funkciu $F_R(x) = \int_{-\infty}^x f_R(t)dt$ pre $x \in (1, 2)$:

$$F_R(x) = \int_{-\infty}^x f_R(t)dt = \int_1^x 1 dt = x - 1.$$

Teda na sledovanom intervale je

$$F_V(x) = F_R\left(\sqrt[3]{\frac{3x}{4\pi}}\right) = \sqrt[3]{\frac{3x}{4\pi}} - 1$$

a celkovo

$$F_V(x) = \begin{cases} 0 & \text{pre } x \leq \frac{4\pi}{3}, \\ \sqrt[3]{\frac{3x}{4\pi}} - 1 & \text{pre } \frac{4\pi}{3} < x < \frac{32\pi}{3}, \\ 1 & \text{pre } x \geq \frac{32\pi}{3}. \end{cases}$$

Príklad 5. Predpokladajme, že počet bodov z písomky je náhodná premenná s rovnomerným rozdelením na intervale¹ $(0, 20)$. Písomku písalo n študentov, ich výsledky môžeme považovať za nezávislé. Odvoďte hustotu:

- (a) minimálneho počtu bodov z písomky,
- (b) maximálneho počtu bodov z písomky.

Riešenie. Nech X_i pre $i = 1, 2, \dots, n$ je X_i $(0, 20)$. To znamená, že $\mathbb{P}(X_i \leq a) = \frac{a}{20}$

- (a) Nech $Y = \min(X_1, \dots, X_n)$. Hodnoty, ktoré môže nadobudnúť, sú z intervalu $(0, 20)$. Distribučná funkcia tohto minima je pre $y \in (0, 20)$ rovná:

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) \stackrel{(1)}{=} 1 - \mathbb{P}(Y > y) \stackrel{(2)}{=} 1 - \mathbb{P}(X_1 > y, \dots, X_n > y) \\ &\stackrel{(3)}{=} 1 - \mathbb{P}(X_1 > y) \dots \mathbb{P}(X_n > y) \stackrel{(4)}{=} 1 - \left(1 - \frac{y}{20}\right)^n. \end{aligned}$$

Vysvetlenia vyznačených úprav:

¹To znamená, že je možné získať ľubovoľný, aj neceločíselný počet bodov.

- (1) Pravdepodobnosť komplementu.
- (2) Nerovnosť $Y > y$ znamená, že minimum je väčšie ako y . To je ekvivalentné s tým, že všetky hodnoty sú väčšie ako y , teda $X_1 > y, \dots, X_n > y$.
- (3) X_i sú zo zadania nezávislé, takže $\mathbb{P}(X_1 > y, \dots, X_n > y) = \mathbb{P}(X_1 > y) \dots \mathbb{P}(X_n > y)$
- (4) Ako sme uviedli na začiatku, $\mathbb{P}(X_i \leq y) = \frac{y}{20}$. Tu potrebujeme komplement.

Hustotu teraz získame derivovaním tejto distribučnej funkcie. Pre $y \in (0, 20)$ je

$$f_Y(y) = F'_Y(y) = n \left(1 - \frac{y}{20}\right)^{n-1} \frac{1}{20}$$

a celkovo

$$f_Y(y) = \begin{cases} \frac{n}{20} \left(1 - \frac{y}{20}\right)^{n-1} & \text{pre } y \in (0, 20), \\ 0 & \text{inak.} \end{cases}$$

- (b) Nech $Z = \max(X_1, \dots, X_n)$. Hodnoty, ktoré môže nadobudnúť, sú z intervalu $(0, 20)$. Distribučná funkcia tohto maxima je pre $z \in (0, 20)$ rovná:

$$\begin{aligned} F_Z(z) &= \mathbb{P}(Z \leq z) \stackrel{(1)}{=} \mathbb{P}(X_1 \leq z, \dots, X_n \leq z) \\ &\stackrel{(2)}{=} \mathbb{P}(X_1 \leq z) \dots \mathbb{P}(X_n \leq z) \stackrel{(3)}{=} \left(\frac{z}{20}\right)^n \end{aligned} \quad (1)$$

Vysvetlenia vyznačených úprav:

- (1) Nerovnosť $Z \leq z$ znamená, že maximum je menšie alebo rovné z . To je ekvivalentné s tým, že všetky hodnoty sú menšie alebo rovné z , teda $X_1 \leq z, \dots, X_n \leq z$.
- (2) X_i sú zo zadania nezávislé, takže $\mathbb{P}(X_1 \leq z, \dots, X_n \leq z) = \mathbb{P}(X_1 \leq z) \dots \mathbb{P}(X_n \leq z)$
- (3) Ako sme uviedli na začiatku, $\mathbb{P}(X_i \leq z) = \frac{z}{20}$.

Hustotu teraz získame derivovaním tejto distribučnej funkcie. Pre $z \in (0, 20)$ je

$$f_Z(z) = F'_Z(z) = n \left(\frac{z}{20}\right)^{n-1} \cdot \frac{1}{20}$$

a celkovo

$$f_Z(z) = \begin{cases} \frac{n}{20} \left(\frac{z}{20}\right)^{n-1} & \text{pre } z \in (0, 20), \\ 0 & \text{inak.} \end{cases}$$

Príklad 6. Nájdite strednú hodnotu maximálneho počtu bodov z písomky z predchádzajúceho príkladu. Čomu sa rovná limita, keď počet študentov rastie do nekonečna?

Riešenie. Počítame integrál:

$$\begin{aligned}\mathbb{E}(Z) &= \int_{-\infty}^{\infty} z f_Z(z) dz = \int_0^{20} z \cdot \frac{n}{20} \left(\frac{z}{20}\right)^{n-1} dz = \frac{n}{20^n} \int_0^{20} z^n dz \\ &= \frac{n}{20^n} \left[\frac{z^{n+1}}{n+1} \right]_0^{20} = \frac{n}{20^n} \cdot \frac{20^{n+1}}{n+1} = \frac{20n}{n+1}.\end{aligned}\quad (2)$$

Ak $n \rightarrow \infty$, tak $\frac{n}{n+1} \rightarrow 1$, a teda

$$\lim_{n \rightarrow \infty} \mathbb{E}(Z) = \lim_{n \rightarrow \infty} \frac{20n}{n+1} = 20.$$

To má prirodzenú interpretáciu - ak písomku píše veľa študentov a výsledky majú rovnomerné rozdelenie, očakávaná hodnota najlepšej písomky sa približuje k maximálnemu možnému počtu bodov.

Príklad 7. Nech T je čas medzi príchodmi zákazníkov do predajne v minútach. Predpokladajme, že má exponenciálne rozdelenie so strednou hodnotou 5 minút. Práve prišiel zákazník.

- (a) Aká je pravdepodobnosť, že ďalší zákazník príde až po viac ako 10 minútach?
- (b) Aký je čas, do ktorého príde ďalší zákazník s pravdepodobnosťou 90 percent?
- (c) Prešlo 8 ďalších minút a zákazník ešte neprišiel. Aká je pravdepodobnosť, že naňho budeme čakať viac ako 10 ďalších minút?

Riešenie. Hustota exponenciálneho rozdelenia je $f_T(t) = \lambda e^{-\lambda t}$ pre $t > 0$ a jeho stredná hodnota je $\frac{1}{\lambda}$. Teda v našom prípade je $\lambda = \frac{1}{5} = 0,2$. Aby sme nemuseli integrovať hustotu v každej otázke, spravíme to na začiatku všeobecne a odvodíme si distribučnú funkciu. Pre $t > 0$ je

$$F(t) = \mathbb{P}(T \leq t) = \int_{-\infty}^t f(x) dx = \int_0^t \lambda e^{-\lambda x} dx = \lambda \left[\frac{e^{-\lambda x}}{-\lambda} \right]_0^t = 1 - e^{-\lambda t}.$$

V našom prípade teda

$$F(t) = 1 - e^{-0,2t} \text{ pre } t > 0.$$

- (a) To, že ďalší zákazník príde až po viac ako 10 minútach, znamená, že čas čakania je väčší ako 10:

$$\mathbb{P}(T > 10) = 1 - \mathbb{P}(T \leq 10) = 1 - F(10) = e^{-0,2 \cdot 10} = e^{-2} \doteq 0,135$$

Pravdepodobnosť, že ďalší zákazník príde až po viac ako 10 minútach, je 0,135.

- (b) Hľadáme t také, že $F_T(t) = 0,9$:

$$\begin{aligned}1 - e^{-0,2t} &= 0,9 \\ e^{-0,2t} &= 0,1 \\ t &= -\frac{\ln(0,1)}{0,2} \doteq 11,51\end{aligned}$$

S 90-percentnou pravdepodobnosťou príde ďalší zákazník do 11,51 minút.

- (c) Prešlo 8 ďalších minút a zákazník ešte neprišiel, teda $T > 8$. To, že naňho budeme čakať viac ako 10 ďalších minút, znamená, že medzi príchodmi zákazníkov bude viac ako 18 minút, teda $T > 18$. Hľadáme teda podmienenú pravdepodobnosť toho, že $T > 18$, ak vieme, že $T > 8$:

$$\begin{aligned}\mathbb{P}(T > 18 \mid T > 8) &\stackrel{(1)}{=} \frac{\mathbb{P}(T > 18, T > 8)}{\mathbb{P}(T > 8)} \stackrel{(2)}{=} \frac{\mathbb{P}(T > 18)}{\mathbb{P}(T > 8)} \stackrel{(3)}{=} \frac{1 - \mathbb{P}(T \leq 18)}{1 - \mathbb{P}(T \leq 8)} \\ &\stackrel{(4)}{=} \frac{1 - F(18)}{1 - F(8)} \stackrel{(5)}{=} \frac{e^{-0,2 \cdot 18}}{e^{-0,2 \cdot 8}} = e^{-0,2 \cdot 10} = e^{-2} \doteq 0,135\end{aligned}$$

Vysvetlenia:

- (1) definícia podmienenej pravdepodobnosti
- (2) to, že súčasne platí $T > 18$ aj $T > 8$ sa dá zjednodušiť na splnenie podmienky $T > 18$
- (3) komplementy
- (4) definícia distribučnej funkcie
- (5) dosadenie našej distribučnej funkcie

Pravdepodobnosť, že budeme čakať viac ako 10 ďalších minút, je 0,135.

Poznámka. Všimnime si, že v časti (c) vyšla rovnaká hodnota pravdepodobnosti ako v časti (a). Ide o charakteristickú vlastnosť exponenciálneho rozdelenia: ak čas čakania modelujeme týmto rozdelením, tak to, koľko sme čakali doteraz, nemá vplyv na pravdepodobnosť čakania v budúcnosti. Tejto vlastnosti sa hovorí „rozdelenie bez pamäte“.

Príklad 8. Predkladajme, prístroj obsahuje n častí, životnosť každej z nich má exponenciálne rozdelenie s rovnakou strednou hodnotou λ a tieto životnosti sú nezávislé. Prístroj sme ho práve zložili z nových funkčných častí. Nájdite hustotu doby fungovania prístroja v nasledujúcich dvoch prípadoch:

- (a) Prístroj prestane fungovať, ak prestane fungovať niektorá jeho časť (teda na jeho fungovanie sú potrebné všetky časti).
- (b) Prístroj funguje, ak funguje aspon jedna jeho časť.

Riešenie. Označme $X_1, \dots, X_n$ životnosti častí prístroja. V predchádzajúcom príklade sme odviedli distribučnú funkciu - pre $x > 0$ platí $\mathbb{P}(X_i \leq x) = 1 - e^{-\lambda x}$.

V prvom uvažovanom prípade je životnosť prístroja rovná $Y = \min(X_1, \dots, X_n)$, v druhom prípade $Z = \max(X_1, \dots, X_n)$. Máme teda nájsť hustoty týchto dvoch náhodných premenných. Oplatí sa porovnať postup riešenia s príkladom 5, kde sme počítali rozdelenia minima a maxima z náhodných premenných s rovnomerným rozdelením.

- (a) Máme $Y = \min(X_1, \dots, X_n)$. Hodnoty, ktoré môže nadobudnúť, sú kladné čísla. Distribučná funkcia tohto minima je pre $y > 0$ rovná:

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) \stackrel{(1)}{=} 1 - \mathbb{P}(Y > y) \stackrel{(2)}{=} 1 - \mathbb{P}(X_1 > y, \dots, X_n > y) \\ &\stackrel{(3)}{=} 1 - \mathbb{P}(X_1 > y) \dots \mathbb{P}(X_n > y) \stackrel{(4)}{=} 1 - (e^{-\lambda y})^n = 1 - e^{-n\lambda y}. \end{aligned}$$

Vysvetlenia vyznačených úprav:

- (1) Pravdepodobnosť komplementu.
- (2) Nerovnosť $Y > y$ znamená, že minimum je väčšie ako y . To je ekvivalentné s tým, že všetky hodnoty sú väčšie ako y , teda $X_1 > y, \dots, X_n > y$.
- (3) X_i sú zo zadania nezávislé, takže $\mathbb{P}(X_1 > y, \dots, X_n > y) = \mathbb{P}(X_1 > y) \dots \mathbb{P}(X_n > y)$.
- (4) Vieme, že $\mathbb{P}(X_i \leq y) = F(y) = 1 - e^{-\lambda y}$. Tu potrebujeme komplement.

Hustotu teraz získame derivovaním tejto distribučnej funkcie. Pre $y > 0$ je

$$f_Y(y) = F'_Y(y) = n\lambda e^{-n\lambda y}$$

a celkovo

$$f_Y(y) = \begin{cases} n\lambda e^{-n\lambda y} & \text{pre } y > 0, \\ 0 & \text{inak.} \end{cases}$$

Všimnime si, že nám vyšla hustota exponenciálneho rozdelenia s parametrom $n\lambda$.

- (b) Nech $Z = \max(X_1, \dots, X_n)$. Hodnoty, ktoré môže nadobudnúť, sú kladné čísla. Distribučná funkcia tohto maxima je pre $z > 0$ rovná:

$$\begin{aligned} F_Z(z) &= \mathbb{P}(Z \leq z) \stackrel{(1)}{=} \mathbb{P}(X_1 \leq z, \dots, X_n \leq z) \\ &\stackrel{(2)}{=} \mathbb{P}(X_1 \leq z) \dots \mathbb{P}(X_n \leq z) \stackrel{(3)}{=} (1 - e^{-\lambda z})^n \end{aligned} \quad (3)$$

Vysvetlenia vyznačených úprav:

- (1) Nerovnosť $Z \leq z$ znamená, že maximum je menšie alebo rovné z . To je ekvivalentné s tým, že všetky hodnoty sú menšie alebo rovné z , teda $X_1 \leq z, \dots, X_n \leq z$.
- (2) X_i sú zo zadania nezávislé, takže $\mathbb{P}(X_1 \leq z, \dots, X_n \leq z) = \mathbb{P}(X_1 \leq z) \dots \mathbb{P}(X_n \leq z)$.
- (3) Ako sme uviedli na začiatku, $\mathbb{P}(X_i \leq z) = 1 - e^{-\lambda z}$.

Hustotu teraz získame derivovaním tejto distribučnej funkcie. Pre $z > 0$ je

$$f_Z(z) = F'_Z(z) = n(1 - e^{-\lambda z})^{n-1} \cdot \lambda e^{-\lambda z}$$

a celkovo

$$f_Z(z) = \begin{cases} n(1 - e^{-\lambda z})^{n-1} \cdot \lambda e^{-\lambda z} & \text{pre } z > 0, \\ 0 & \text{inak.} \end{cases}$$

Príklad 9. Dokážte, že ak U má rovnomerné rozdelenie na intervale $(0, 1)$, tak $X = -\ln(U)$ má exponenciálne rozdelenie so strednou hodnotou 1.

Riešenie. Keďže U je z intervalu $(0, 1)$, logaritmus $\ln(U)$ je definovaný (nestane sa, že by sme mali robiť logaritmus z nuly alebo zo záporného čísla) a je záporný. Hodnota $-\ln(U)$ je teda kladná. Nájdeme jej distribučnú funkciu F_X a následne hustotu f_X pre $x > 0$:

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(-\ln(U) \leq x) = \mathbb{P}(\ln(U) \geq -x) = \mathbb{P}(U \geq e^{-x}) = 1 - \mathbb{P}(U < e^{-x}).$$

Pre náhodnú premennú U s rovnomerným rozdelením na $(0, 1)$ platí $\mathbb{P}(U < y) = y$ pre $y \in (0, 1)$, a teda

$$\mathbb{P}(U < e^{-x}) = e^{-x}.$$

Pre distribučnú funkciu F_X teda máme

$$F_X(x) = 1 - e^{-x} \quad \text{pre } x > 0.$$

Jej derivovaním dostaneme hustotu

$$f_X(x) = F'_X(x) = e^{-x} \quad \text{pre } x > 0,$$

čo je hustota exponenciálneho rozdelenia s parametrom λ . Stredná hodnota exponenciálneho rozdelenia je vo všeobecnosti $\frac{1}{\lambda}$, v našom prípade je to 1. To znamená, že f_X je hustota exponenciálneho rozdelenia so strednou hodnotou 1, čo sme mali dokázať.

Príklad 10. Budeme generovať náhodné čísla X s exponenciálnym rozdelením s parametrom λ a dosadzovať ich do distribučnej funkcie tohto rozdelenia. Teda vzniknú náhodné čísla $Y = 1 - e^{-\lambda X}$. Aké majú pravdepodobnostné rozdelenie

Riešenie. Realizácie náhodnej premennej X sú kladné, takže hodnoty $e^{-\lambda X}$ budú z intervalu $(0, 1)$ a tiež hodnoty $Y = 1 - e^{-\lambda X}$ budú z intervalu $(0, 1)$. Pre $y \in (0, 1)$ nájdeme distribučnú funkciu a hustotu náhodnej premenej Y (mimo tohto intervalu bude hustota nulová)

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) = \mathbb{P}(1 - e^{-\lambda X} \leq y) = \mathbb{P}(e^{-\lambda X} \geq 1 - y) = \mathbb{P}(-\lambda X \geq \ln(1 - y)) \\ &= \mathbb{P}\left(X \leq -\frac{1}{\lambda} \ln(1 - y)\right) = F_X\left(-\frac{1}{\lambda} \ln(1 - y)\right). \end{aligned} \quad (4)$$

Vieme, že pre $x > 0$ je $F_X(x) = 1 - e^{-\lambda x}$, a teda

$$F_Y(y) = 1 - e^{-\lambda \cdot \left(-\frac{1}{\lambda} \ln(1-y)\right)} = 1 - e^{\ln(1-y)} = 1 - (1 - y) = y.$$

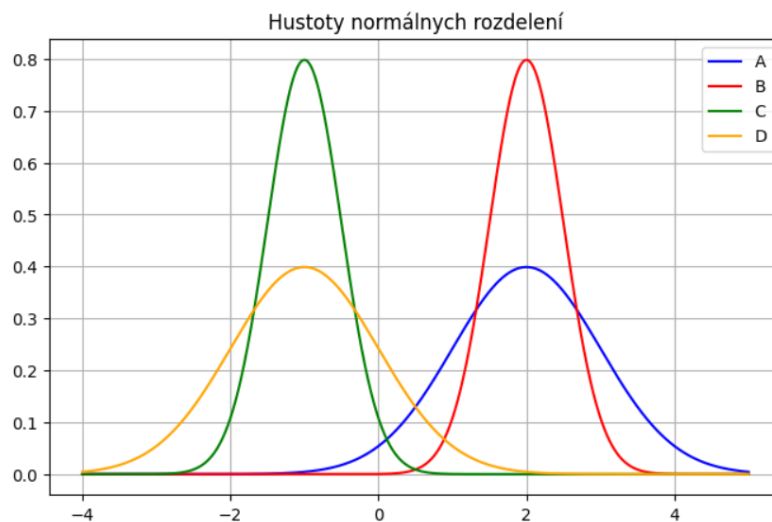
Hustota potom je

$$f_Y(y) = F'_Y(y) = 1 \quad \text{pre } y \in (0, 1),$$

čo je hustota rovnomerného rozdelenia na intervale $(0, 1)$. Uvedeným postupom teda generujeme náhodné čísla s rovnomerným rozdelením na intervale $(0, 1)$.

5 Spojité náhodné premenné III. (normálne rozdelenie)

Príklad 1. Na obrázku 9 sú hustoty nasledovných štyroch normálnych rozdelení: $\mathcal{N}(2, 1)$, $\mathcal{N}(-1, 1)$, $\mathcal{N}(2, \frac{1}{2})$, $\mathcal{N}(-1, \frac{1}{2})$. Priradiť rozdelenia ku grafom hustôt.



Obr. 9: Normálne rozdelenie

Riešenie. Normálne rozdelenie sa zapisuje v tvare $\mathcal{N}(\mu, \sigma^2)$, kde μ je stredná hodnota a σ^2 je disperzia. Graf hustoty je symetrický, takže μ je hodnota, okolo ktorej je hustota symetrická. Teda grafy A, B zodpovedajú $\mu = 2$ a grafy C, D zodpovedajú $\mu = -1$. Disperzia znamená rozptýlenosť, väčšia disperzia znamená väčšiu rozptýlenosť hodnôt. Takže A je $\mathcal{N}(2, 1)$, B je $\mathcal{N}(2, \frac{1}{2})$, C je $\mathcal{N}(-1, \frac{1}{2})$ a D je $\mathcal{N}(-1, 1)$.

Príklad 2. Nech X má normálne rozdelenie $\mathcal{N}(2, 1)$. Ktoré z nasledujúcich náhodných premenných majú tiež normálne rozdelenie? Určte ich strednú hodnotu a disperziu:

$$Y_1 = 1 + X^2, \quad Y_2 = -1 + 3X, \quad Y_3 = 1 - 2X, \quad Y_4 = e^{2X}.$$

Riešenie. Normálne rozdelenie sa zachováva pri transformácii $aX + b$, takže Y_2 a Y_3 majú normálne rozdelenie. Y_1 a Y_4 nevyhovujú už kvôli tomu, že môžu nadobúdať len kladné hodnoty.

Pre Y_2 dostaneme:

$$\mathbb{E}(Y_2) = \mathbb{E}(-1 + 3X) = -1 + 3 \cdot \mathbb{E}(X) = -1 + 3 \cdot 2 = 5,$$

$$\mathbb{D}(Y_2) = \mathbb{D}(-1 + 3X) = \mathbb{D}(3X) = 3^2 \cdot \mathbb{D}(X) = 3^2 \cdot 1 = 9,$$

teda $Y_2 \sim \mathcal{N}(5, 9)$.

Podobne pre Y_3 :

$$\mathbb{E}(Y_3) = \mathbb{E}(1 - 2X) = 1 - 2 \cdot \mathbb{E}(X) = 1 - 2 \cdot 2 = -3,$$

$$\mathbb{D}(Y_3) = \mathbb{D}(1 - 2X) = \mathbb{D}(-2X) = (-2)^2 \cdot \mathbb{D}(X) = (-2)^2 \cdot 1 = 4,$$

teda $Y_3 \sim \mathcal{N}(-3, 4)$.

Príkad 2. Vyjadrite nasledovné pravdepodobnosti pomocou distribučnej funkcie normalizovaného normálneho rozdelenia (teda $\mathcal{N}(0, 1)$), ktorú označíme Φ :

- (a) $\mathbb{P}(X > 3)$, ak $X \sim \mathcal{N}(1, 2)$
- (b) $\mathbb{P}(X \geq 2)$, ak $X \sim \mathcal{N}(0, 4)$
- (c) $\mathbb{P}(1 < X \leq 3)$, ak $X \sim \mathcal{N}(1, 2)$
- (d) $\mathbb{P}(|X - 1| \geq 2)$, ak $X \sim \mathcal{N}(2, 3)$

Riešenie. Výpočty sú založené na normalizácii: Ak X má normálne rozdelenie $\mathcal{N}(\mu, \sigma^2)$, tak

$$\frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

- (a) Máme $X \sim \mathcal{N}(1, 2)$, takže $\frac{X-1}{\sqrt{2}} \sim \mathcal{N}(0, 1)$. Preto:

$$\mathbb{P}(X > 3) = \mathbb{P}\left(\frac{X-1}{\sqrt{2}} > \frac{3-1}{\sqrt{2}}\right) = 1 - \mathbb{P}\left(\frac{X-1}{\sqrt{2}} \leq \frac{3-1}{\sqrt{2}}\right) = 1 - \Phi\left(\frac{2}{\sqrt{2}}\right).$$

- (b) Máme $X \sim \mathcal{N}(0, 4)$, takže $\frac{X-0}{\sqrt{4}} \sim \mathcal{N}(0, 1)$. Preto:

$$\mathbb{P}(X \geq 2) = \mathbb{P}\left(\frac{X-0}{2} \geq \frac{2-0}{2}\right) = 1 - \mathbb{P}\left(\frac{X-0}{2} < \frac{2-0}{2}\right) = 1 - \Phi(1),$$

kde sme navyše využili, že pre spojitú náhodnú premennú Y je $\mathbb{P}(Y < a) = \mathbb{P}(Y \leq a)$.

- (c) Máme $X \sim \mathcal{N}(1, 2)$, takže $\frac{X-1}{\sqrt{2}} \sim \mathcal{N}(0, 1)$. Preto:

$$\mathbb{P}(1 < X \leq 3) = \mathbb{P}\left(\frac{1-1}{\sqrt{2}} < \frac{X-1}{\sqrt{2}} \leq \frac{3-1}{\sqrt{2}}\right) = \Phi\left(\frac{2}{\sqrt{2}}\right) - \Phi(0) = \Phi\left(\frac{2}{\sqrt{2}}\right) - \frac{1}{2},$$

kde sme v poslednej rovnosti využili, že $\Phi(0) = \mathbb{P}(\mathcal{N}(0, 1) \leq 0) = \frac{1}{2}$

- (d) Máme $X \sim \mathcal{N}(2, 3)$, takže $\frac{X-2}{\sqrt{3}} \sim \mathcal{N}(0, 1)$. Preto:

$$\begin{aligned}\mathbb{P}(|X - 1| \geq 2) &= \mathbb{P}(X \leq -1) + \mathbb{P}(X \geq 3) \\ &= \mathbb{P}\left(\frac{X-2}{\sqrt{3}} \leq \frac{-1-2}{\sqrt{3}}\right) + \mathbb{P}\left(\frac{X-2}{\sqrt{3}} \geq \frac{3-2}{\sqrt{3}}\right) \\ &= \Phi\left(\frac{-3}{\sqrt{3}}\right) + 1 - \Phi\left(\frac{1}{\sqrt{3}}\right).\end{aligned}$$

Príklad 3. Nech náhodná premenná X má normálne rozdelenie $\mathcal{N}(\mu, \sigma^2)$. Dokážte, že pravdepodobnosť toho, že X sa od strednej hodnoty líši o menej ako k štandardných odchýliek, nezávisí od parametrov μ, σ^2 .

Riešenie. Nech náhodná premenná X má normálne rozdelenie $\mathcal{N}(\mu, \sigma^2)$. Budeme počítať pravdepodobnosť $\mathbb{P}(|X - \mu| < k\sigma)$, kde $k > 0$, kým sa nedostaneme k výrazu, ktorý neobsahuje parametre μ, σ^2 :

$$\mathbb{P}(|X - \mu| < k\sigma) = \mathbb{P}(-k\sigma < X - \mu < k\sigma)$$

Náhodná premenná $X - \mu$ ($0, \sigma^2$), takže pomocou normalizácie dostaneme

$$\mathbb{P}(|X - \mu| < k\sigma) = \mathbb{P}\left(-\frac{k\sigma}{\sigma} < \frac{X - \mu}{\sigma} < \frac{k\sigma}{\sigma}\right) = \mathbb{P}(-k < \frac{X - \mu}{\sigma} < k) = \Phi(k) - \Phi(-k).$$

Vidíme, že počítaná pravdepodobnosť závisí len od k , a nie od μ, σ^2 .

Poznámka. Konkrétne numerické hodnoty pre $k = 1, 2, 3$ sú všeobecne známe:

- V intervale $\mu \pm 1\sigma$ je približne 68 percent hodnôt.
- V intervale $\mu \pm 2\sigma$ je približne 95 percent hodnôt.
- V intervale $\mu \pm 3\sigma$ je približne 99,7 percent hodnôt.

Výpočet je na obrázku 10. Výsledok $\Phi(k) - \Phi(-k)$ je upravený pomocou rovnosti $\Phi(-k) = 1 - \Phi(k)$, ktorá vyplýva zo symetrie hustoty, vďaka ktorej je $\mathbb{P}(\mathcal{N}(0, 1) > x) = \mathbb{P}(\mathcal{N}(0, 1) < -x)$.

```
from scipy.stats import norm

k_list = [1, 2, 3]
pravd = [2*norm.cdf(k) - 1 for k in k_list]

for k, pravd in zip(k_list, pravd):
    print(f"P(|X- mu| <= {k}*sigma) = {pravd:.4f}")

P(|X- mu| <= 1*sigma) = 0.6827
P(|X- mu| <= 2*sigma) = 0.9545
P(|X- mu| <= 3*sigma) = 0.9973
```

Obr. 10: Výpočet podielu hodnôt v intervale pre normálne rozdelenie.

Príklad 4. Uvažujme dva výrobné procesy:

- X_1 je čas výroby súčiastky A, jeho rozdelenie je $X_1 \sim \mathcal{N}(10, 16)$
- X_2 je čas výroby súčiastky B, jeho rozdelenie je $X_2 \sim \mathcal{N}(12, 9)$

Predpokladáme, že X_1 a X_2 sú nezávislé. Pri výrobe súčiastky B, ktorá je finálnym produktom, je však potrebná súčiastka A. Výroba teda prebieha tak, že sa vyrobí súčiastka A a následne sa použije pri výrobe súčiastky B. Aká je pravdepodobnosť, že súčiastka B bude vyrobená o menej ako 25 hodín? Odpoveď vyjadrite pomocou distribučnej funkcie normalizovaného normálneho rozdelenia.

Riešenie. Celkový čas potrebný na výrobu súčiastky B je $X_1 + X_2$. Keďže X_1 a X_2 sú nezávislé s normálnym rozdelením, ich súčet má tiež normálne rozdelenie, a to s parametrami:

$$\mathbb{E}(X_1 + X_2) = \mathbb{E}(X_1) + \mathbb{E}(X_2) = 10 + 12 = 22,$$

$$\mathbb{D}(X_1 + X_2) \stackrel{(1)}{=} \mathbb{D}(X_1) + \mathbb{D}(X_2) = 16 + 9 = 25,$$

kde rovnosť (1) vyplýva z nezávislosti. Teda $X_1 + X_2 \sim \mathcal{N}(22, 25)$ a

$$\mathbb{P}(X_1 + X_2 < 25) = \mathbb{P}\left(\frac{X_1 + X_2 - 22}{\sqrt{25}} < \frac{25 - 22}{\sqrt{25}}\right) = \mathbb{P}\left(\mathcal{N}(0, 1) < \frac{3}{5}\right) = \Phi\left(\frac{3}{5}\right).$$

Príklad 5. Z historických dát sa vie, že škody, ktoré platila poisťovňa pri určitom type poistenia, majú normálne rozdelenie $X \sim \mathcal{N}(5000, 1000^2)$. Poisťovňa vybrala na analýzu náhodne 20 záznamov, kvôli náhodnosti výberu môžeme výšky škôd považovať za nezávislé. Aká je pravdepodobnosť, že priemerná škoda bude vyššia ako 5500? Odpoveď vyjadrite pomocou distribučnej funkcie normalizovaného normálneho rozdelenia.

Riešenie. Označme škody $X_1, \dots, X_{20}$ a $\bar{X}$ ich priemer. Potom priemer má tiež normálne rozdelenie a jeho parametre sú:

$$\begin{aligned} \mathbb{E}(\bar{X}) &= \mathbb{E}\left(\frac{1}{20}(X_1 + \dots + X_{20})\right) = \frac{1}{20}(\mathbb{E}(X_1) + \dots + \mathbb{E}(X_{20})) \\ &= \frac{1}{20}(20 \cdot 5000) = 5000, \end{aligned}$$

$$\begin{aligned} \mathbb{D}(\bar{X}) &= \mathbb{D}\left(\frac{1}{20}(X_1 + \dots + X_{20})\right) = \frac{1}{20^2}\mathbb{D}(X_1 + \dots + X_{20}) \\ &\stackrel{(1)}{=} \frac{1}{20^2}(\mathbb{D}(X_1) + \dots + \mathbb{D}(X_{20})) = \frac{1}{20^2}(20 \cdot 1000^2) = \frac{1000^2}{20}, \end{aligned}$$

kde rovnosť (1) vyplýva z nezávislosti. Teda $\bar{X} \sim \mathcal{N}(5000, \frac{1000^2}{20})$ a

$$\begin{aligned} \mathbb{P}(\bar{X} > 5500) &= \mathbb{P}\left(\frac{\bar{X} - 5000}{\sqrt{\frac{1000^2}{20}}} > \frac{5500 - 5000}{\sqrt{\frac{1000^2}{20}}}\right) = \mathbb{P}\left(\mathcal{N}(0, 1) > \frac{500}{\frac{1000}{\sqrt{20}}}\right) \\ &= \mathbb{P}\left(\mathcal{N}(0, 1) > \frac{\sqrt{20}}{2}\right) = 1 - \Phi\left(\frac{\sqrt{20}}{2}\right) = 1 - \Phi(\sqrt{5}). \end{aligned}$$

Príklad 6. Plastové vrečko na bonbóny má hmotnosť 0,05 gramov. Je v ňom 10 bonbónov. Hmotnosti bonbónov sú nezávislé, hmotnosť jedného bonbónu v gramoch

má normálne rozdelenie $\mathcal{N}(4, \frac{1}{8})$. Aká je pravdepodobnosť, že hmotnosť celého balenia je menšia ako 40 gramov? Odpoveď vyjadrite pomocou distribučnej funkcie normalizovaného normálneho rozdelenia.

Riešenie. Označme hmotnosti bonbónov $X_1, \dots, X_{10}$ a Y hmotnosť celého balenia, teda

$$Y = 0,05 + X_1 + \dots + X_{10}.$$

Potom Y má tiež normálne rozdelenie a jeho parametre sú:

$$\begin{aligned} \mathbb{E}(Y) &= \mathbb{E}(0,05 + X_1 + \dots + X_{10}) = 0,05 + \mathbb{E}(X_1) + \dots + \mathbb{E}(X_{10}) \\ &= 0,05 + 10 \cdot 4 = 40,05, \\ \mathbb{D}(Y) &= \mathbb{D}(0,05 + X_1 + \dots + X_{10}) = \mathbb{D}(X_1 + \dots + X_{10}) \\ &\stackrel{(1)}{=} \mathbb{D}(X_1) + \dots + \mathbb{D}(X_{10}) = 10 \cdot \frac{1}{8} = \frac{5}{4}, \end{aligned}$$

kde rovnosť (1) vyplýva z nezávislosti. Teda $Y \sim \mathcal{N}(40,05; \frac{5}{4})$ a

$$\begin{aligned} \mathbb{P}(Y < 40) &= \mathbb{P}\left(\frac{Y - 40,05}{\sqrt{\frac{5}{4}}} < \frac{40 - 40,05}{\sqrt{\frac{5}{4}}}\right) = \mathbb{P}\left(\mathcal{N}(0, 1) < -\frac{0,05}{\frac{\sqrt{5}}{2}}\right) \\ &= \mathbb{P}\left(\mathcal{N}(0, 1) < -\frac{0,1}{\sqrt{5}}\right) = \Phi\left(-\frac{0,1}{\sqrt{5}}\right). \end{aligned}$$

Príklad 7. Hmotnosti bonbónov sú nezávislé, hmotnosť jedného bonbónu v gramoch má rovnomerné rozdelenie na intervale $(9, 11)$. Aká je pravdepodobnosť, že 100 bonbónov bude mať celkovú hmotnosť menšiu ako 995 gramov? Použite centrálnu limitnú vetu a odpoveď vyjadrite pomocou distribučnej funkcie normalizovaného normálneho rozdelenia.

Riešenie. Označme hmotnosti bonbónov $X_1, \dots, X_{100}$ a

$$Y = X_1 + \dots + X_{100}.$$

Potom podľa centrálnej limitnej vety má súčet Y má približne normálne rozdelenie. Vypočítame ešte jeho strednú hodnotu a disperziu.

O jednotlivých bonbónoch vieme, že majú rovnomerné rozdelenie na intervale $(9, 11)$. To znamená, že $\mathbb{E}(X_i) = 10$ a (podľa vzorcovníka) $\mathbb{D}(X_i) = \frac{1}{3}$. Potom

$$\begin{aligned} \mathbb{E}(Y) &= \mathbb{E}(X_1 + \dots + X_{100}) = \mathbb{E}(X_1) + \dots + \mathbb{E}(X_{100}) = 100 \cdot 10 = 1000, \\ \mathbb{D}(Y) &= \mathbb{D}(X_1 + \dots + X_{100}) \stackrel{(1)}{=} \mathbb{D}(X_1) + \dots + \mathbb{D}(X_{100}) = 100 \cdot \frac{1}{3} = \frac{100}{3}, \end{aligned}$$

kde rovnosť (1) vyplýva z nezávislosti. Teda rozdelenie Y aproximujeme pomocou

normálneho rozdelenia $Y \sim \mathcal{N}(1000; \frac{100}{3})$ a

$$\begin{aligned}\mathbb{P}(Y < 995) &= \mathbb{P}\left(\frac{Y - 1000}{\sqrt{\frac{100}{3}}} < \frac{995 - 1000}{\sqrt{\frac{100}{3}}}\right) = \mathbb{P}\left(\mathcal{N}(0, 1) < -\frac{5}{\frac{10}{\sqrt{3}}}\right) \\ &= \mathbb{P}\left(\mathcal{N}(0, 1) < -\frac{\sqrt{3}}{2}\right) = \Phi\left(-\frac{\sqrt{3}}{2}\right).\end{aligned}$$

6 Náhodné vektory

Príklad 1. Janko Hraško si všímal počet správ, ktoré dostal počas jednotlivých hodín. Ak označíme

- X = počet správ na Whatsappe,
- Y = počet mailov,

tak združené rozdelenie náhodného vektora (X, Y) je dané nasledovnou tabuľkou:

	$Y = 0$	$Y = 1$	$Y = 2$
$X = 0$	0,12	0,08	0,05
$X = 1$	0,18	0,20	0,07
$X = 2$	0,10	0,12	0,08

- Nájdite marginálne rozdelenia jeho zložiek X a Y .
- Zistite, či sú počty správ na Whatsappe a počty notifikácií nezávislé náhodné premenné.
- Nájdite pravdepodobnosť aspoň jednej správy alebo mailu počas náhodne zvolenej hodiny.
- Vypočítajte podmienené rozdelenie počtu správ na Whatsappe, ak počas danej hodiny prišiel Jankovi Hraškovi jeden mail.
- Vypočítajte kovarianciu X a Y .

Riešenie.

- Marginálne rozdelenia zložiek X a Y nájdeme sčítaním pravdepodobností po riadkoch, resp. po stĺpcoch:

Rozdelenie X je

$$\mathbb{P}(X = 0) = 0,12 + 0,08 + 0,05 = 0,25,$$

$$\mathbb{P}(X = 1) = 0,18 + 0,20 + 0,07 = 0,45,$$

$$\mathbb{P}(X = 2) = 0,10 + 0,12 + 0,08 = 0,30.$$

a rozdelenie Y je

$$\mathbb{P}(Y = 0) = 0,12 + 0,18 + 0,10 = 0,40,$$

$$\mathbb{P}(Y = 1) = 0,08 + 0,20 + 0,12 = 0,40,$$

$$\mathbb{P}(Y = 2) = 0,05 + 0,07 + 0,08 = 0,20.$$

- X a Y nie sú nezávislé, lebo napríklad

$$\mathbb{P}(X = 0, Y = 0) = 0,12 \neq \mathbb{P}(X = 0) \cdot \mathbb{P}(Y = 0) = 0,25 \cdot 0,40 = 0,10.$$

(c) Aspoň jedna správa alebo mail počas náhodne zvolenej hodiny je komplement k udalosti $X = 0, Y = 0$, ktorá má podľa tabuľky pravdepodobnosť 0,12. Hľadaná pravdepodobnosť je teda 0,88.

(d) Hľadáme podmienené rozdelenie $X | Y = 1$. Z definície podmienenej pravdepodobnosti:

$$\mathbb{P}(X = x | Y = 1) = \frac{\mathbb{P}(X = x, Y = 1)}{\mathbb{P}(Y = 1)}$$

Teda pre náš náhodný vektor:

$$\mathbb{P}(X = 0 | Y = 1) = \frac{0,08}{0,40} = 0,20,$$

$$\mathbb{P}(X = 1 | Y = 1) = \frac{0,20}{0,40} = 0,50,$$

$$\mathbb{P}(X = 2 | Y = 1) = \frac{0,12}{0,40} = 0,30.$$

(e) Kovariancia X a Y sa vypočíta ako $\text{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$. Potrebujeme teda vypočítať tieto tri stredné hodnoty:

- Stredná hodnota X :

$$\mathbb{E}(X) = 0 \cdot 0,25 + 1 \cdot 0,45 + 2 \cdot 0,30 = 0 + 0,45 + 0,60 = 1,05$$

- Stredná hodnota Y :

$$\mathbb{E}(Y) = 0 \cdot 0,40 + 1 \cdot 0,40 + 2 \cdot 0,20 = 0 + 0,40 + 0,40 = 0,80$$

- Stredná hodnota súčinu $\mathbb{E}(XY)$:

$$\begin{aligned} \mathbb{E}(XY) &= 0 \cdot 0 \cdot 0,12 + 0 \cdot 1 \cdot 0,08 + 0 \cdot 2 \cdot 0,05 \\ &\quad + 1 \cdot 0 \cdot 0,18 + 1 \cdot 1 \cdot 0,20 + 1 \cdot 2 \cdot 0,07 \\ &\quad + 2 \cdot 0 \cdot 0,10 + 2 \cdot 1 \cdot 0,12 + 2 \cdot 2 \cdot 0,08 \end{aligned}$$

Nenulové sú v skutočnosti len štyri sčítance:

$$\mathbb{E}(XY) = 0,20 + 0,14 + 0 + 0,24 + 0,32 = 0,90.$$

Kovariancia teda je:

$$\text{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 0,90 - 1,05 \cdot 0,80 = 0,06.$$

Príklad 2. Nech X a Y sú nezávislé náhodné premenné s rovnomerným rozdelením na intervale $(0, 1)$. Nájdite hustotu náhodného vektora (m, M) , kde $m = \min(X, Y)$ a $M = \max(X, Y)$.

Riešenie. Je jasné, že (m, M) môže nadobúdať iba hodnoty spĺňajúce $0 < m \leq M < 1$. V tých bodoch roviny, ktoré tejto podmienke nevyhovujú, teda bude hustota

nulová. Pre ostatné najskôr nájdeme distribučnú funkciu a následne derivovaním nájdeme hustotu.

Počítajme teda najskôr distribučnú funkciu $F(a, b)$, ak $0 < a \leq b < 1$. Podľa definície je

$$F(a, b) = \mathbb{P}(m \leq a, M \leq b)$$

Minimum hodnôt X a Y má byť menšie alebo rovné a , to znamená, že aspoň jedna z hodnôt X, Y je menšia alebo rovná a . Súčasne maximum má byť menšie alebo rovné b , čo znamená, že žiadna z hodnôt nie je väčšia ako b . Hľadáme teda pravdepodobnosť toho, že žiadna z hodnôt nie je väčšia ako b , ale aspoň jedna je menšia alebo rovná a . To sa dá napísať ako rozdiel pravdepodobností:

$$\begin{aligned} F(a, b) &= \mathbb{P}(\{\text{žiadna z hodnôt } X, Y \text{ nie je väčšia ako } b\}) \\ &\quad - \mathbb{P}(\{\text{žiadna z hodnôt } X, Y \text{ nie je väčšia ako } b \text{ a obidve sú väčšie ako } a\}) \end{aligned}$$

Vypočítame tieto pravdepodobnosti:

$$\mathbb{P}(\{\text{žiadna z hodnôt } X, Y \text{ nie je väčšia ako } b\}) = \mathbb{P}(X \leq b, Y \leq b)$$

$$\stackrel{(1)}{=} \mathbb{P}(X \leq b) \cdot \mathbb{P}(Y \leq b) \stackrel{(2)}{=} b \cdot b = b^2$$

$$\mathbb{P}(\{\text{žiadna z hodnôt } X, Y \text{ nie je väčšia ako } b \text{ a obidve sú väčšie ako } a\})$$

$$= \mathbb{P}(X \in (a, b], Y \in (a, b]) \stackrel{(3)}{=} \mathbb{P}(X \in (a, b]) \cdot \mathbb{P}(Y \in (a, b]) \stackrel{(4)}{=} (b-a) \cdot (b-a) = (b-a)^2$$

Vysvetlenia úprav:

- (1) nezávislosť X a Y
- (2) X a Y majú rovnomerné rozdelenie na intervale $(0, 1)$
- (3) nezávislosť X a Y
- (4) X a Y majú rovnomerné rozdelenie na intervale $(0, 1)$

Teda

$$F(a, b) = b^2 - (b-a)^2 = 2ab - a^2$$

Hustota je derivácia distribučnej funkcie, postupne podľa jednotlivých premenných. V našom prípade je to podľa a a b . Po zderivovaní podľa a dostaneme $2b - a$. Ak toto zderivujeme podľa b , zostane konštanta 2. Teda

$$f(a, b) = \begin{cases} 2, & 0 < a \leq b < 1, \\ 0, & \text{inak.} \end{cases}$$

Poznámka. Všeobecnejší prípad s maximom a minimom n náhodných premenných bol na prednáške.

Príklad 3. Majme náhodné premenné:

- X = výška študenta v cm, s disperziou $\mathbb{D}(X) = 64$,

- Y = hmotnosť študenta v kg, s disperziou $\mathbb{D}(Y) = 49$.

Korelácia medzi výškou a hmotnosťou je $\rho_{X,Y} = 0,6$. Vypočítajte kovarianciu $\text{cov}(X, Y)$ a zostavte kovariančnú maticu náhodného vektora (X, Y) .

Riešenie. Definícia korelácie je $\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sqrt{\mathbb{D}(X)\mathbb{D}(Y)}}$, takže

$$\text{cov}(X, Y) = \rho_{X,Y} \sqrt{\mathbb{D}(X)\mathbb{D}(Y)} = 0,6 \cdot \sqrt{64 \cdot 49} = 0,6 \cdot 8 \cdot 7 = 33,6$$

a kovariančná matica je

$$\begin{pmatrix} \text{cov}(X, X) & \text{cov}(X, Y) \\ \text{cov}(X, Y) & \text{cov}(Y, Y) \end{pmatrix} = \begin{pmatrix} \mathbb{D}(X) & \text{cov}(X, Y) \\ \text{cov}(X, Y) & \mathbb{D}(Y) \end{pmatrix} = \begin{pmatrix} 64 & 33,6 \\ 33,6 & 49 \end{pmatrix}.$$

Príklad 4. K náhodným premenným X, Y z predchádzajúceho príklad pridajme ešte $\tilde{X}$, ktorá bude vyjadrovať výšku študenta v metroch. Odvodte, že korelácia $\tilde{X}$ a Y je rovnaká ako korelácia X a Y (čo znamená, že korelácia sa nezmení, ak výšku vyjadríme v metroch namiesto centimetrov).

Riešenie. Definícia korelácie je

$$\rho_{\tilde{X},Y} = \frac{\text{cov}(\tilde{X}, Y)}{\sqrt{\mathbb{D}(\tilde{X})\mathbb{D}(Y)}},$$

takže potrebujeme vypočítať $\text{cov}(\tilde{X}, Y)$ a $\mathbb{D}(\tilde{X})$. Keďže X je výška v centimetroch a $\tilde{X}$ je výška v metroch, platí $\tilde{X} = \frac{1}{100}X$. Preto:

$$\begin{aligned} \rho_{\tilde{X},Y} &= \frac{\text{cov}(\tilde{X}, Y)}{\sqrt{\mathbb{D}(\tilde{X})\mathbb{D}(Y)}} = \frac{\text{cov}(\frac{1}{100}X, Y)}{\sqrt{\mathbb{D}(\frac{1}{100}X)\mathbb{D}(Y)}} = \frac{\frac{1}{100}\text{cov}(X, Y)}{\sqrt{\frac{1}{100^2}\mathbb{D}(X)\mathbb{D}(Y)}} = \frac{\frac{1}{100}\text{cov}(X, Y)}{\frac{1}{100}\sqrt{\mathbb{D}(X)\mathbb{D}(Y)}} \\ &= \frac{\text{cov}(X, Y)}{\sqrt{\mathbb{D}(X)\mathbb{D}(Y)}} = \rho_{X,Y}, \end{aligned}$$

čo sme mali dokázať.

Príklad 5. Nech X je náhodná premenná s rozdelením:

$$\mathbb{P}(X = -1) = \frac{1}{4}, \quad \mathbb{P}(X = 0) = \frac{1}{2}, \quad \mathbb{P}(X = 1) = \frac{1}{4}.$$

Definujme náhodnú premennú

$$Y = X^2.$$

Dokážte, že X, Y sú nekorelované (teda že majú nulovú koreláciu).

Riešenie. Definícia korelácie je $\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sqrt{\mathbb{D}(X)\mathbb{D}(Y)}}$, takže potrebujeme ukázať, že $\text{cov}(X, Y) = 0$. Počítajme teda túto kovarianciu:

$$\text{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}(X^3) - \mathbb{E}(X)\mathbb{E}(X^2).$$

Z pravdepodobností zo zadania dostaneme:

$$\mathbb{E}(X) = (-1) \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} = 0$$

Náhodná premenná X^2 nadobúda hodnotu 0 s pravdepodobnosťou $\frac{1}{2}$ a hodnotu 1 s pravdepodobnosťou $\frac{1}{4}$. Jej stredná hodnota teda je

$$\mathbb{E}(X^2) = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} = \frac{1}{4}$$

Náhodná premenná X^3 má rovnaké rozdelenie ako X , takže

$$\mathbb{E}(X^3) = \mathbb{E}(X) = 0.$$

Kovariancia teda je

$$\text{cov}(X, Y) = \mathbb{E}(X^3) - \mathbb{E}(X)\mathbb{E}(X^2) = 0 - 0 \cdot \frac{1}{4} = 0,$$

čo sme potrebovali dokázať.

Poznámka. Máme teda ďalší príklad náhodných premenných, ktoré sú nekorelované, ale nie sú nezávislé. Opakovanie z prednášky: Ak sú náhodné premenné nezávislé, tak ich korelácia je nulová. Naopak to vo všeobecnosti neplatí (ako sme videli aj v tomto príklade), ale ak majú náhodné premenné normálne rozdelenie, tak z nekorelovanosti už nezávislosť vyplýva.

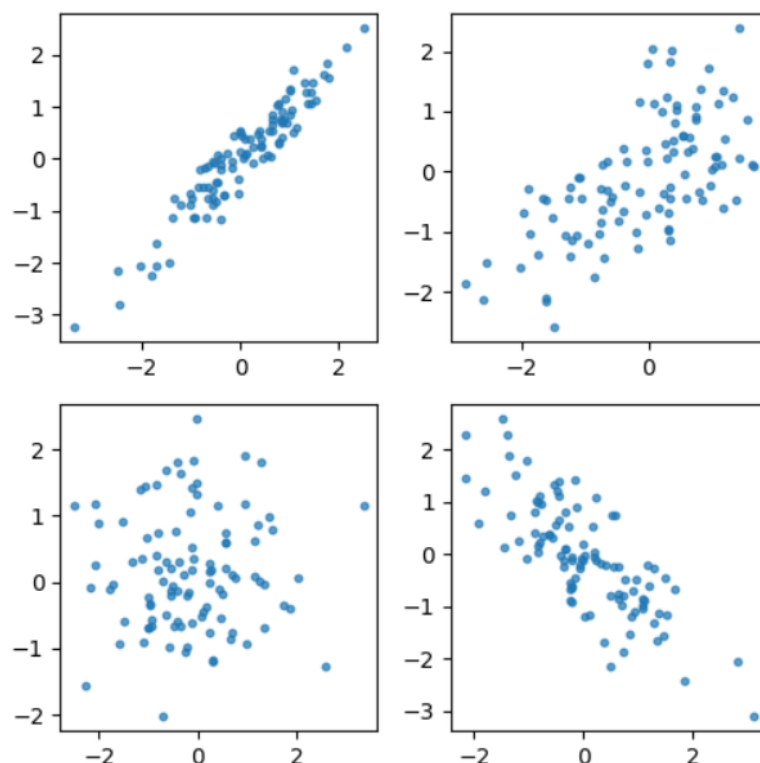
Príklad 6. Z dát bola vypočítaná výberová korelácia. Priradíte dátam na obrázku 11 hodnoty výberovej korelácie:

- (a) 0,94
- (b) -0,79
- (c) 0,84
- (d) 0,08

Riešenie Korelácia meria lineárnu závislosť. Kladné znamienko znamená rastúcu závislosť, záporné znamienko klesajúcu závislosť. Absolútna hodnota korelácie hovorí o sile závislosti - ak je blízka jednej, ide o veľmi silnú závislosť, ako klesá absolútna hodnota, tak klesá aj sila závislosti.

Teda priradenie bude:

- (a) 0,94 - vľavo hore
- (b) -0,79 - vpravo dolu
- (c) 0,84 - vpravo hore
- (d) 0,08 - vľavo dolu



Obr. 11: Korelácie

Príklad 7. Iba jedna sada dát na obrázku 12 bola vygenerovaná z dvojrozmerného normálneho rozdelenia. Ktorá?

Riešenie Na prednáške sme videli vygenerované dáta z rôznych dvojrozmerných normálnych rozdelení - typický obrázok vyzerá tak, že body vytvárajú približne elipsu (alebo v špeciálnom prípade kruh), pričom viac koncentrované sú okolo stredu. Tomuto zodpovedá graf vpravo hore.

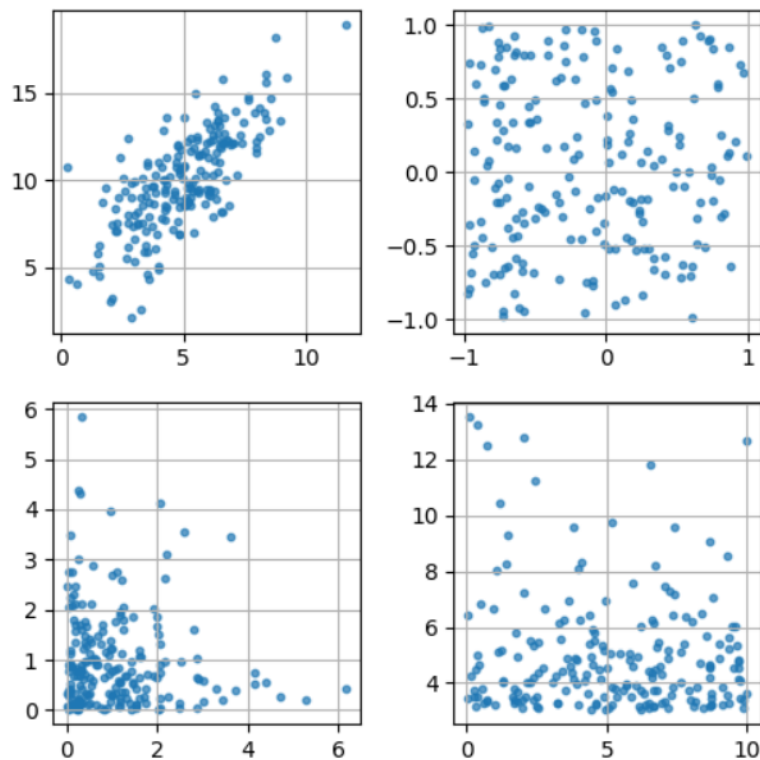
Príklad 8. Upravíme príklad 4 z časti o normálnom rozdelení - budeme uvažovať korelované náhodné premenné.

Uvažujme dva výrobné procesy:

- X_1 je čas výroby súčiastky A, jeho rozdelenie je $X_1 \sim \mathcal{N}(10, 16)$
- X_2 je čas výroby súčiastky B, jeho rozdelenie je $X_2 \sim \mathcal{N}(12, 9)$

Predpokladáme, že X_1 a X_2 sú korelované, čo súvisí s aktuálnym stavom stroja používaného pri výrobe súčiastok. Ich korelácia je 0,75. Pri výrobe súčiastky B, ktorá je finálnym produktom, je však potrebná súčiastka A. Výroba teda prebieha tak, že sa vyrobí súčiastka A a následne sa použije pri výrobe súčiastky B. Aká je pravdepodobnosť, že súčiastka B bude vyrobená o menej ako 25 hodín? Odpoveď vyjadrite pomocou distribučnej funkcie normalizovaného normálneho rozdelenia.

Riešenie. Celkový čas potrebný na výrobu súčiastky B je $X_1 + X_2$. Súčet $X_1 + X_2$



Obr. 12: Dvojrozmerné dáta

má tiež normálne rozdelenie, a to s parametrami:

$$\begin{aligned}\mathbb{E}(X_1 + X_2) &= \mathbb{E}(X_1) + \mathbb{E}(X_2) = 10 + 12 = 22, \\ \mathbb{D}(X_1 + X_2) &= \mathbb{D}(X_1) + \mathbb{D}(X_2) + 2\text{cov}(X_1, X_2).\end{aligned}$$

Potrebuje teda ešte vypočítať kovarianciu X_1 a X_2 : Podľa definície korelácie je

$$\text{cor}(X_1, X_2) = \frac{\text{cov}(X_1, X_2)}{\sqrt{\mathbb{D}(X_1)\mathbb{D}(X_2)}},$$

a teda

$$\text{cov}(X, Y) = \text{cor}(X, Y)\sqrt{\mathbb{D}(X)\mathbb{D}(Y)} = 0,75\sqrt{16 \cdot 9} = 9$$

Takže môžeme dopočítať disperziu:

$$\mathbb{D}(X_1 + X_2) = \mathbb{D}(X_1) + \mathbb{D}(X_2) + 2\text{cov}(X_1, X_2) = 16 + 9 + 2 \cdot 9 = 43$$

Teda $X_1 + X_2 \sim \mathcal{N}(22, 43)$ a

$$\mathbb{P}(X_1 + X_2 < 25) = \mathbb{P}\left(\frac{X_1 + X_2 - 22}{\sqrt{43}} < \frac{25 - 22}{\sqrt{43}}\right) = \mathbb{P}\left(\mathcal{N}(0, 1) < \frac{3}{\sqrt{43}}\right) = \Phi\left(\frac{3}{\sqrt{43}}\right).$$

Vzorcovník I. na skúšku (pravdepodobnosť)

Prehľad dôležitých diskretných rozdelení

Názov	Pravdepodobnosti	Hodnoty	Stredná hodnota	Disperzia
Bernoulliho	$P(X = 1) = p,$ $P(X = 0) = 1 - p$	0, 1	p	$p(1 - p)$
Binomické	$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$	0, 1, ..., n	np	$np(1 - p)$
Geometrické	$P(X = k) = (1 - p)^{k-1} p$	1, 2, ...	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Poissonovo	$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$	0, 1, 2, ...	λ	λ

Prehľad dôležitých spojitých rozdelení

Názov	Hustota	Stredná hodnota	Disperzia
Rovnomerné	$f(x) = \frac{1}{b-a}, x \in (a, b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponenciálne	$f(x) = \lambda e^{-\lambda x}, x > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Normálne	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in \mathbb{R}$	μ	σ^2

Základné pojmy a vlastnosti

- Distribučná funkcia: $F(x) = \mathbb{P}(X \leq x)$
- Hustota: $F(x) = \int_{-\infty}^x f(t)dt, f(x) = F'(x)$
- Stredná hodnota:

$$\mathbb{E}(X) = \sum_{i=1}^n x_i P(X = x_i), \mathbb{E}(X) = \sum_{i=1}^{\infty} x_i P(X = x_i), \mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x) dx$$

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f(x) dx$$

- Disperzia:

$$\mathbb{D}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$$

- Štandardná odchýlka: $\sqrt{\mathbb{D}(x)}$

- Vlastnosti:

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y), \quad \mathbb{E}(aX) = a\mathbb{E}(X)$$

$$\mathbb{D}(a + X) = \mathbb{D}(X), \quad \mathbb{D}(aX) = a^2\mathbb{D}(X)$$

Pre nezávislé náhodné premenné

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y), \quad \mathbb{D}(X + Y) = \mathbb{D}(X) + \mathbb{D}(Y)$$

- Nezávislosť diskrétnych náhodných premenných:

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \mathbb{P}(Y = y) \quad \forall x, y$$

- Kovariancia:

$$\text{cov}(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))).$$

Užitočný vzťah pre výpočet:

$$\text{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

Vlastnosti:

$$\text{cov}(aX + b, Y) = a \text{cov}(X, Y), \quad \text{cov}(X + Z, Y) = \text{cov}(X, Y) + \text{cov}(Z, Y)$$

- Korelácia:

$$\rho_{X,Y} = \text{cor}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\mathbb{D}(X)\mathbb{D}(Y)}},$$

- Disperzia súčtu náhodných premenných (môžu byť závislé):

$$\mathbb{D}(X + Y) = \mathbb{D}(X) + \mathbb{D}(Y) + 2 \text{cov}(X, Y)$$

- Distribučná funkcia náhodného vektora: $F(x_1, \dots, x_n) = \mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n)$

- Hustota spojitého náhodného vektora sa z distribučnej funkcie počíta postupným zderivovaním podľa jednotlivých premenných.

- Kovariančná matica dvojrozmerného náhodného vektora:

$$\begin{pmatrix} \text{cov}(X, X) & \text{cov}(X, Y) \\ \text{cov}(X, Y) & \text{cov}(Y, Y) \end{pmatrix} = \begin{pmatrix} \mathbb{D}(X) & \text{cov}(X, Y) \\ \text{cov}(X, Y) & \mathbb{D}(Y) \end{pmatrix}$$

Vo všeobecnosti je v i -tom riadku a j -tom stĺpci $\text{cov}(X_i, X_j)$

- Korelačná matica dvojrozmerného náhodného vektora:

$$\begin{pmatrix} \text{cor}(X, X) & \text{cor}(X, Y) \\ \text{cor}(X, Y) & \text{cor}(Y, Y) \end{pmatrix} = \begin{pmatrix} 1 & \text{cor}(X, Y) \\ \text{cor}(X, Y) & 1 \end{pmatrix}$$

Vo všeobecnosti je v i -tom riadku a j -tom stĺpci $\text{cor}(X_i, X_j)$