ŘEŠENÍ

Cvičení 1. Doba výpočtu (v sekundách) určité úlohy s náhodným vstupem je náhodná veličina s rozdělením s hustotou

$$f(x) = \begin{cases} \frac{2}{x^3} & \text{pro } x \geq 1, \\ 0 & x < 1. \end{cases}$$

(a) Určete pravděpodobnost toho, že výpočet skončí do 5 sekund. Vyznačte tuto pravděpodobnost v grafu hustoty f.

(b) Určete střední hodnotu doby výpočtu.

(c) Určete rozptyl doby výpočtu.

(d) Předpokládejme, že umíme generovat náhodnou veličinu U s rovnoměrným rozdělením na $[0, 1]$. Navrhněte, jak lze pomocí U generovat X.

Řešení. (a) Označme náhodnou veličinu $X :=$ doba do příjezdu vlaku. Ptáme se na $\mathbb{P}(X \leq 5)$. X je spojitá, a tedy

$$\mathbb{P}(X \leq 5) = \int_{-\infty}^{5} f(x)\mathrm{d}x = \int_{1}^{5} \frac{2}{x^3}\mathrm{d}x = 2\left[\frac{x^{-3+1}}{-3+1}\right]_{x=1}^{5} = \left[-\frac{1}{x^2}\right]_{x=1}^{5} = \frac{24}{25}$$

(b)

$$\mathbb{E}X = \int_{-\infty}^{\infty} xf(x)\mathrm{d}x = \int_{1}^{\infty} \frac{2}{x^2}\mathrm{d}x = 2\left[-\frac{1}{x}\right]_{x=1}^{\infty} = 2.$$

(c) Pro výpočet rozptylu si uvědomíme následující užitečný vztah

$$\begin{aligned} \mathrm{var}X &= \mathbb{E}(X - \mathbb{E}X)^2 \\ &= \mathbb{E}\left[X^2 - 2X(\mathbb{E}X) + (\mathbb{E}X)^2\right] \\ &= \mathbb{E}[X^2] - \mathbb{E}[2X(\mathbb{E}X)] + \mathbb{E}[(\mathbb{E}X)^2] \\ &= \mathbb{E}X^2 - 2(\mathbb{E}X)(\mathbb{E}X) + (\mathbb{E}X)^2 \\ &= \mathbb{E}X^2 - (\mathbb{E}X)^2 \end{aligned}$$

V první rovnosti jsme jen napsali definici, ve druhé roznásobili druhou mocninu, ve třetí využili linearitu střední hodnoty a ve čtvrté fakt, že $\mathbb{E}X$ je již nenáhodná konstanta, a tak jí můžeme vytknout před $\mathbb{E}$. Stačí nám již tedy dopočítat pouze $\mathbb{E}X^2$,

$$\mathbb{E}X^2 = \int_{-\infty}^{\infty} x^2 f(x)\mathrm{d}x = \int_{1}^{\infty} \frac{2}{x}\mathrm{d}x = 2\left[\log|x|\right]_{x=1}^{\infty} = 2(\lim_{x\to\infty} \log|x| - \log 1) = \infty.$$

A tedy $\mathrm{var}X = \mathbb{E}X^2 - (\mathrm{E}X)^2 = \infty$.

(d) Hledáme předpis $X = f(U)$, kde $f : \mathbb{R} \to \mathbb{R}$ je hledaná funkce. Vezmeme-li $f(t) = F_X^{-1}(t)$, kde F_X je distribuční funkce X, pak skutečně dostaneme rozdělení X, neboť

$$\mathbb{P}(X \leq x) = \mathbb{P}(F_X^{-1}(U) \leq x) = \mathbb{P}(U \leq F_X(x)) = F_U(F_X(x)) = F_X(x), \quad \forall x \in \mathbb{R}.$$

Vyžili jsme zde, že distribuční funkce U se chová jako identita na intervalu $[0, 1]$. Nejprve spočítáme, jak vypadá distribuční funkce X.

$$F_X(x) = \begin{cases} \int_1^x \frac{2}{t^3}\mathrm{d}t & \text{pro } x \geq 1, \\ 0 & x < 1. \end{cases}$$
$$= \begin{cases} 1 - \frac{1}{x^2} & \text{pro } x \geq 1, \\ 0 & x < 1. \end{cases}$$

Její inverzi, tj. *kvantilovou funkci* $F_X^{-1} : [0,1] \to \mathbb{R}$ najdeme jako řešení rovnice $y = F_X(x)$ pro x. Dostaneme, že rovnost platí pro

$$x \in \{\frac{1}{\sqrt{1-y}}, -\frac{1}{\sqrt{1-y}}\}.$$

Uvědomíme si, že náhodná veličina X nabývá pouze kladných hodnot, a tedy řešením je $f(t) = \frac{1}{\sqrt{1-t}}$. Dosazením generujeme X jako

$$X = \frac{1}{\sqrt{1-U}}.$$

Cvičení 2. Doba čekání na vlak je náhodná veličina X s exponenciálním rozdělením s hustotou

$$f(x) = \begin{cases} \frac{1}{5}e^{-x/5} & \text{pro } x \geq 0, \\ 0 & \text{pro } x < 0. \end{cases}$$

(a) Spočtěte střední hodnotu X.

(b) Spočtěte rozptyl X.

Řešení. (a)

$$\mathbb{E}X = \int_{-\infty}^{\infty} xf(x)\mathrm{d}x = \int_0^{\infty} x\frac{1}{5}e^{-x/5}\mathrm{d}x.$$

Nyní použijeme *per partes* pro $u = x, u' = 1, v' = \mathrm{e}^{-x/5}, v = -5\mathrm{e}^{-x/5}$. Dostaneme

$$\int_0^{\infty} x\frac{1}{5}e^{-x/5}\mathrm{d}x = \frac{1}{5}\left(\left[x(-5\mathrm{e}^{-x/5})\right]_{x=0}^{\infty} - \int_0^{\infty} -5\mathrm{e}^{-x/5}\mathrm{d}x\right) = \frac{1}{5}\left(0 - 25\left[\mathrm{e}^{-x/5}\right]_0^{\infty}\right) = 5.$$

Poznámka Všimněte si, že X má exponenciální rozdělení s parametrem $\frac{1}{5}$. Obecně pokud $X \sim Exp(\lambda)$, pak $\mathbb{E}X = \frac{1}{\lambda}$.

(b) Opět využijeme vztahu $\mathrm{var}X = \mathbb{E}X^2 - (\mathbb{E}X)^2$:

$$\mathbb{E}X^2 = \int_{-\infty}^{\infty} x^2 f(x)\mathrm{d}x = \int_0^{\infty} x^2\frac{1}{5}e^{-x/5}\mathrm{d}x.$$

Použijeme *per partes* pro $u = x^2, u' = 2x, v' = 1/5\mathrm{e}^{-x/5}, v = -\mathrm{e}^{-x/5}$. Dostaneme

$$\int_0^{\infty} x^2\frac{1}{5}\mathrm{e}^{-x/5}\mathrm{d}x = \left[x^2(-\mathrm{e}^{-x/5})\right]_{x=0}^{\infty} - \int_0^{\infty} 2x(-\mathrm{e}^{-x/5})\mathrm{d}x = 2\int_0^{\infty} x(-\mathrm{e}^{-x/5})\mathrm{d}x.$$

Zde využijeme znovu *per partes* s volbou jako v (a) a dostaneme výsledek 50. Rozptyl je tedy

$$\mathrm{var}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = 50 - 5^2 = 25.$$

Poznámka Pokud $X \sim Exp(\lambda)$, pak $\mathrm{var}X = \frac{1}{\lambda^2}$.

Cvičení 3. V kapse máte dvě padesátikoruny, jednu dvacetikorunu a jednu desetikorunu. Zloděj Vám z kapsy náhodně vybere dvě mince. Označme jako X náhodnou veličinu, která udává, o kolik peněz jste právě přišli.

(a) Určete rozdělení X a spočtěte Vaši očekávanou ztrátu.

(b) Určete rozptyl X.

(c) Zloděj si následně koupí kávu z automatu za 20 Kč a doma mu manželka zabaví čtyři pětiny z toho, co donese. Označme jako Y veličinu udávající částku, která zlodějovi po tom všem zůstane. Určete rozdělení a očekávanou hodnotu Y.

(d) Určete rozptyl veličiny Y. Jaký je vztah mezi momenty X a momenty Y?

Řešení. (a) Nejprve si můžeme uvědomit, na jakém prostoru vlastně náhodná veličina X působí. Označme

$$\Omega := \{\omega = \{m_1, m_2\};\ m_i \in \{10, 20, 50, 50\}\},$$
$$X(\omega) = X(\{m_1, m_2\}) = m_1 + m_2.$$

Takto definované Ω splňuje *klasickou definici pravděpodobnostního prostoru*, tedy všechny elementární jevy mají stejnou pravděpodobnost. Není to však jediná možná volba Ω, můžeme například uvažovat nerozlišitelnost padesátikorun, tj. $\Omega := \{\{m_1, m_2\};\ m_i \in \{10, 20, 50\}\}$. Potom ale elementární jevy nebudou mít stejnou pravděpodobnost. Která volba je lepší, to je otázka osobních preferencí.

Rozdělení X: X je diskrétní náhodná veličina a může nabývat hodnot z množiny $\{30, 60, 70, 100\}$. Rozdělení X je určeno pravděpodobnostmi p_1, p_2, p_3, p_4, kde

$$p_1 = \mathbb{P}(X = 30) = \mathbb{P}(\{10, 20\}) = \frac{1}{6},$$
$$p_2 = \mathbb{P}(X = 60) = \mathbb{P}(\{10, 50\}) = \frac{1}{3},$$
$$p_3 = \mathbb{P}(X = 70) = \mathbb{P}(\{20, 50\}) = \frac{1}{3},$$
$$p_4 = \mathbb{P}(X = 100) = \mathbb{P}(\{50, 50\}) = \frac{1}{6}.$$

Očekávaná (=střední) ztráta je z definice

$$\mathbb{E}X = \sum_{k=1}^{4} p_k \cdot x_k = \frac{1}{6} \cdot 30 + \frac{1}{3} \cdot 60 + \frac{1}{3} \cdot 70 + \frac{1}{6} \cdot 100 = 65$$

(b) Rozptyl ztráty je z definice

$$\mathrm{var} X = \sum_{k=1}^{4} p_k \cdot (x_k - \mathbb{E}X)^2 = \frac{1}{6} \cdot (30 - 65)^2 + \frac{1}{3} \cdot (60 - 65)^2 + \frac{1}{3} \cdot (70 - 65)^2 + \frac{1}{6} \cdot (100 - 65)^2 = 425$$

(c) Částka, která zlodějovi zůstane, je dána náhodou veličinou.

$$Y := (X - 20)\frac{1}{5}.$$

Náhodná veličina Y je taktéž diskrétní a nabývá hodnot $\{2, 8, 10, 16\}$. Její rozdělení je jednoznačně určeno pravděpodobnostmi $p_1^Y, p_2^Y, p_3^Y, p_4^Y$, kde

$$p_1^Y = \mathbb{P}(Y = 2) = \mathbb{P}(X = 30) = \frac{1}{6},$$
$$p_2^Y = \mathbb{P}(Y = 8) = \mathbb{P}(X = 60) = \frac{1}{3},$$
$$p_3^Y = \mathbb{P}(Y = 10) = \mathbb{P}(X = 70) = \frac{1}{3},$$
$$p_4^Y = \mathbb{P}(Y = 16) = \mathbb{P}(X = 10) = \frac{1}{6}.$$

Pro výpočet očekávané hodnoty Y máme dvě možnosti

- Z definice střední hodnoty:

$$\mathbb{E}Y = \sum_{k=1}^{4} p_k^Y \cdot y_k = \frac{1}{6}\cdot 2 + \frac{1}{3}\cdot 8 + \frac{1}{3}\cdot 10 + \frac{1}{6}\cdot 16 = 9$$

- Z linearity střední hodnoty

$$\mathbb{E}Y = \mathbb{E}[(X-20)\frac{1}{5}] = \frac{1}{5}\mathbb{E}X - \frac{1}{5}\cdot 20 = \frac{1}{5}\cdot 65 - 4 = 9$$

Pro výpočet rozptylu můžeme použít opět buď přímo definici nebo lépe vlastnost rozptylu

$$\mathrm{var}(a+bX) = b^2\mathrm{var}X, \quad \forall a,b\in\mathbb{R}, \forall X \text{náhodnou veličinu.}$$

Tedy,

$$\mathrm{var}Y = \mathrm{var}[(X-20)\frac{1}{5}] = \mathrm{var}[\frac{1}{5}X - 4] = \frac{1}{25}\mathrm{var}X = \frac{1}{25}\cdot 425 = 17.$$

Cvičení 4. Při přenosu binárního souboru se náhodně vybraný znak zkreslí s pravděpodobností $p \in (0,1)$ a jednotlivé znaky se zkreslují nezávisle na sobě. Náhodná veličina X udává počet zkreslených znaků v binární posloupnosti délky n.

(a) Připomeňte si, jaké rozdělení má X.

(b) Spočtěte střední hodnotu a rozptyl X pro $n = 1$.

(c) Spočtěte střední hodnotu a rozptyl X pro $n = 2$.

(d) Určete očekávaný počet zkreslených znaků v posloupnosti délky n.

(e) Spočtěte rozptyl veličiny X pro obecné n.

Řešení. (a) Označme

$$\Omega = \{\omega = (z_1, ..., z_n), z_i \in \{\text{zkresleno, nezkresleno}\}\},$$

$$X(\omega) = X((z_1,...,z_n)) = \sum_{i=1}^{n} \mathbf{1}\{z_i = \text{zkresleno}\} = \text{počet zkreslených znaků v posloupnosti } (z_1,...,z_n).$$

Náhodná veličina X je diskrétní a nabývá hodnot $\{0,...,n\}$. Její rozdělení je jednoznačně určeno pravděpodobnostmi $p_1,...,p_n$, kde

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

Připomeňme, že X má binomické rozdělení (značíme $X \sim Bi(n,p)$).

(b) Předpokládáme, že $n = 1$.

Poznámka: $Bi(1,p) = Alt(p)$.

Rozdělení X má tedy alternativní rozdělení a rozdělení X je určeno pravděpodobnostmi $p_0 = 1-p, p_1 = p$. Střední hodnota a rozptyl jsou z definice

$$\mathbb{E}X = 0\cdot(1-p) + 1\cdot p = p.$$

$$\mathrm{var}X = \mathbb{E}(X-\mathbb{E}X)^2 = (0-p)^2\cdot(1-p) + (1-p)^2\cdot p = p(1-p).$$

(c) Předpokládáme, že $n = 2$. Rozdělení X je určeno pravděpodobnostmi $p_0 = (1-p)^2, p_1 = 2p(1-p), p_2 = p^2$. Střední hodnota a je z definice

$$\mathbb{E}X = 0 \cdot (1-p)^2 + 1 \cdot 2p(1-p) + 2 \cdot p^2 = 2p.$$

Rozptyl zde spočítáme ze vztahu

$$\text{var}X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = 0^2 \cdot (1-p)^2 + 1^2 \cdot 2p(1-p) + 2^2 \cdot p^2 - 4p^2 = 2p(1-p).$$

(d) Přímý výpočet střední hodnoty a rozptylu jako v (b) a (c) se může velmi zkomplikovat uvažujeme-li obecné $n \in \mathbb{N}$. Proto využijeme následující, velmi důležitou vlastnost binomického rozdělení:

Poznámka: Je-li $X \sim Bi(1,p)$, pak X můžeme chápat jako součet n nezávislých pokusů, které mohou skončit úspěchem (1) nebo neúspěchem (0), tj. jako součet nezávislých alternativních náhodných veličin,

$$X = \sum_{i=1}^{n} Y_i,$$

kde $Y_1, ..., Y_n$ jsou nezávislé, stejně rozdělení s rozdělením $Alt(p)$.

Pak střední hodnotu spočítáme díky vlastnosti linearity jako

$$\mathbb{E}X = \mathbb{E}\sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} \mathbb{E}Y_i = \sum_{i=1}^{n} p = np.$$

Ve třetí rovnosti jsme využili výpočet v (b). Rozptyl pak z nezávislosti náhodných veličin $Y_1, ..., Y_n$ umíme spočítat jako

$$\text{var}X = \text{var}\sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} \text{var}Y_i = \sum_{i=1}^{n} p(1-p) = np(1-p).$$

Opět jsme zde využili znalost z (b).