5. cvičení z PST

31. října 2018

5.1 (rozdělení veličiny vytvořené jako směs)

Program generuje čísla z intervalu $(0,1)$, která jsou rovnoměrně rozdělena. S pravděpodobnosti $c = \frac{1}{4}$ jsou ale někdy výstupem i čísla 0 a 1, a to v poměru $1:2$. Náhodná veličina Z udává hodnoty výstupu generátoru. Určete:

(a) distribuční funkci F_Z.

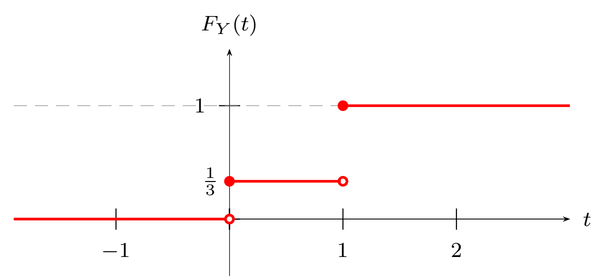
(b) pravděpodobnosti $P(0 \leq Z \leq \frac{1}{2})$ a $P(Z > \frac{1}{3})$.

(c) střední hodnotu $E(Z)$.

Řešení:
Náhodná veličina Z je směsí $\mathrm{Mix}_{1/4}(X,Y)$, kde veličina X je hodnota diskrétního výstupu (hodnoty $\{0,1\}$) a Y je hodnota spojitého výstupu (hodnoty z intervalu $(0,1)$).

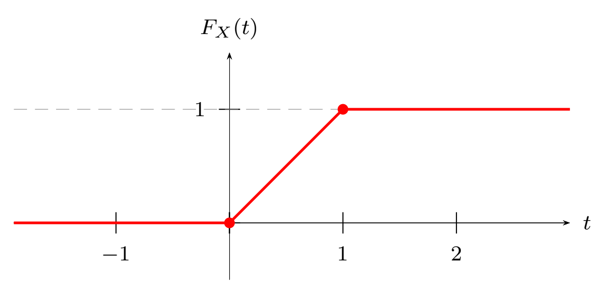
(a) Veličina X má alternativní rozdělení s pravděpodobnostní funkcí $p_X(0) = \frac{1}{3}$, $p_X(1) = \frac{2}{3}$ a distribuční funkcí

$$F_X(t) = \sum_{u \leq t} p_X(u) = \begin{cases} 0\,, & t < 0\,, \\ \frac{1}{3}\,, & 0 \leq t < 1\,, \\ 1\,, & t \geq 1\,. \end{cases}$$



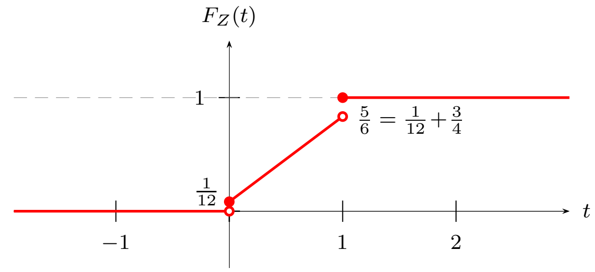
Veličina Y má spojité rozdělení s hustotou $f_Y(t) = \begin{cases} 1\,, & t \in (0,1)\,, \\ 0\,, & \text{jinak}\,, \end{cases}$ a distribuční funkcí

$$F_Y(t) = \int_{-\infty}^{t} f_X(u)\,\mathrm{d}u = \begin{cases} 0\,, & t < 0\,, \\ \int_0^t 1\,\mathrm{d}u = t\,, & 0 \leq t < 1\,, \\ 1\,, & t \geq 1\,. \end{cases}$$



Pro distribuční funkci veličiny $Z = \mathrm{Mix}_{1/4}(X, Y)$ pak platí

$$F_Z(t) = \tfrac{1}{4}F_X(t) + \tfrac{3}{4}F_Y(t) = \begin{cases} 0 & , t < 0\,, \\ \frac{1}{12} + \frac{3}{4}t & , 0 \le t < 1\,, \\ 1 & , t \ge 1\,. \end{cases}$$



(b) $P(0 \le Z \le \frac{1}{2}) = P(Z \le \frac{1}{2}) - P(Z < 0) = F_Z(\frac{1}{2}) - \lim\limits_{t \to 0_-} F_Z(t) = \frac{1}{12} + \frac{3}{4} \cdot \frac{1}{2} - 0 = \frac{11}{24}$

$P(Z > \frac{1}{3}) = 1 - P(Z \le \frac{1}{3}) = 1 - F_Z(\frac{1}{3}) = 1 - \left(\frac{1}{12} + \frac{1}{4}\right) = \frac{2}{3}$.

(c) Pro alternativní rozdělení veličiny X je $E(X) = 1 \cdot p_X(1) = \frac{2}{3}$ a pro rovnoměrné rozdělení veličiny Y na intervalu $(a, b) = (0, 1)$ je $E(Y) = \frac{a+b}{2} = \frac{1}{2}$. Takže pro $Z = \mathrm{Mix}_{1/4}(X, Y)$ je

$$E(Z) = \tfrac{1}{4}E(X) + \tfrac{3}{4}E(Y) = \tfrac{1}{4} \cdot \tfrac{2}{3} + \tfrac{3}{4} \cdot \tfrac{1}{2} = \tfrac{13}{24}\,.$$

5.2 (rozdělení veličiny vytvořené jako směs)

Náhodná veličina X má diskrétní rozdělení s pravděpodobnostní funkci $p_X(0) = 0.4$ a $p_X(1) = 0.6$. Náhodná veličina Y má spojité rovnoměrné rozdělení na intervalu $\langle -1, 2 \rangle$. Náhodná veličina Z je směsí $Z = \mathrm{Mix}_{1/4}(X, Y)$. Určete:

(a) distribuční funkci F_Z.

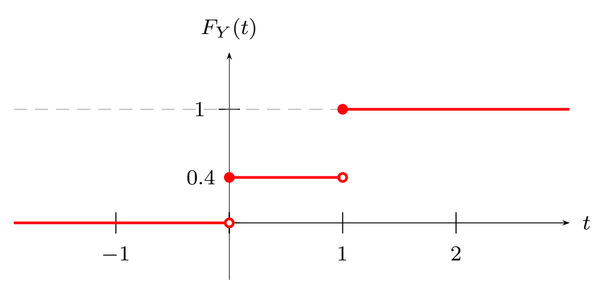
(b) pravděpodobnosti $P(0 \le Z \le 1)$ a $P(Z \ge 0.5)$.

(c) střední hodnotu $E(Z)$.

Řešení:

(a) Veličina X s alternativním rozdělením má distribuční funkci

$$F_X(t) = \sum_{u \le t} p_X(u) = \begin{cases} 0\,, & t < 0\,, \\ 0.4\,, & 0 \le t < 1\,, \\ 1\,, & t \ge 1\,. \end{cases}$$

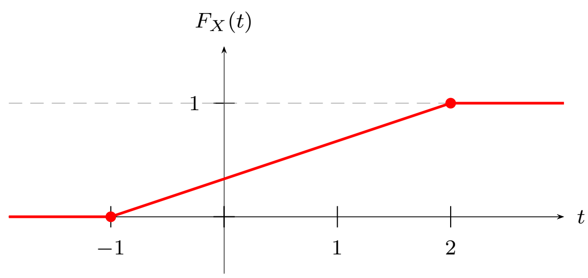


Veličina Y má spojité rovnoměrné rozdělení na $\langle -1, 2\rangle$ s hustotou

$$f_Y(t) = \begin{cases} \frac{1}{2-(-1)} = \frac{1}{3}\,, & t \in \langle -1, 2\rangle\,, \\ 0\,, & \text{jinak}\,, \end{cases}$$

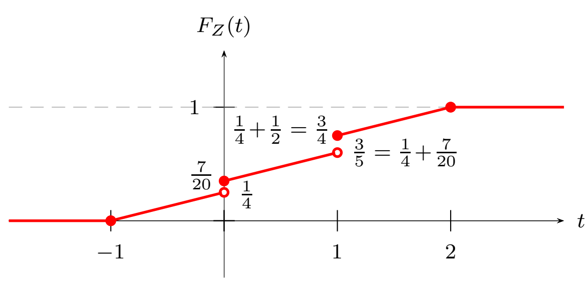
a distribuční funkcí

$$F_Y(t) = \int_{-\infty}^{t} f_X(u)\,\mathrm{d}u = \begin{cases} 0\,, & t < -1\,, \\ \int_{-1}^{t} \frac{1}{3}\,\mathrm{d}u = \frac{t+1}{3}\,, & -1 \le t < 2\,, \\ 1\,, & t \ge 2\,. \end{cases}$$



Pro distribuční funkci veličiny $Z = \mathrm{Mix}_{1/4}(X, Y)$ pak platí

$$F_Z(t) = \frac{1}{4}F_X(t) + \frac{3}{4}F_Y(t) = \begin{cases} 0 & , t < -1\,, \\ \frac{1}{4}\cdot 0 + \frac{3}{4}\cdot\frac{t+1}{3} = \frac{1}{4}t + \frac{1}{4} & , -1 \le t < 0\,, \\ \frac{1}{4}\cdot 0.4 + \frac{3}{4}\cdot\frac{t+1}{3} = \frac{1}{4}t + \frac{7}{20} & , 0 \le t < 1\,, \\ \frac{1}{4}\cdot 1 + \frac{3}{4}\cdot\frac{t+1}{3} = \frac{1}{4}t + \frac{1}{2} & , 1 \le t < 2\,, \\ 1 & , t \ge 2\,. \end{cases}$$



(b) $P(0 \le Z \le 1) = P(Z \le 1) - P(Z < 0) = F_Z(1) - \lim\limits_{t\to 0_-} F_Z(t) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$

$P(Z \ge 0.5) = 1 - P(Z < 0.5) = 1 - \lim\limits_{t\to 0.5_-} F_Z(t) = 1 - \left(\frac{1}{4}\cdot 0.5 + \frac{7}{20}\right) = \frac{21}{40}$.

(c) Pro alternativní rozdělení veličiny X je $E(X) = 1\cdot p_X(1) = 0.6 = \frac{3}{5}$ a pro rovnoměrné rozdělení veličiny Y na intervalu $(a, b) = (-1, 2)$ je $E(Y) = \frac{a+b}{2} = \frac{1}{2}$. Takže pro $Z = \mathrm{Mix}_{1/4}(X, Y)$ je

$$E(Z) = \frac{1}{4}E(X) + \frac{3}{4}E(Y) = \frac{1}{4}\cdot\frac{3}{5} + \frac{3}{4}\cdot\frac{1}{3} = \frac{2}{5}\,.$$

5.3 (rozklad na směs)

Náhodná veličina X má distribuční funkci

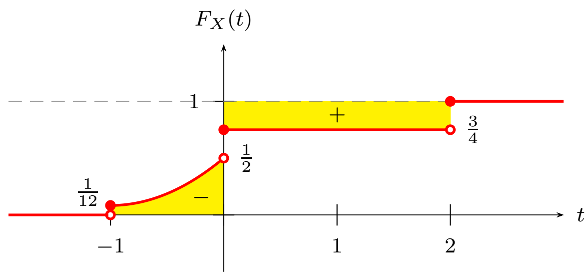
$$F_X(t) = \begin{cases} 0 & , t < -1 \\ \frac{5}{12}(t+1)^2 + \frac{1}{12} & , t \in \langle -1, 0) \\ \frac{3}{4} & , t \in \langle 0, 2) \\ 1 & , t \geq 2 \,. \end{cases}$$

(a) Určete $E(X)$.

(b) Najděte kvantil q_X.

(c) Vyjádřete X jako směs náhodných veličin U a V, z nichž U je diskrétní a V spojitá; popište a znázorněte jejich rozdělení.

Řešení:
(a) Znázorníme si graf F_X:



Střední hodnotu můžeme spočítat pomocí F_X jako

$$E(X) = -\int_{-\infty}^{0} F_X(t)\ \mathrm{d}t + \int_{0}^{\infty} (1 - F_X(t))\ \mathrm{d}t =$$

$$= -\int_{-1}^{0} \tfrac{5}{12}(t+1)^2 + \tfrac{1}{12}\ \mathrm{d}t + \int_{0}^{2} \tfrac{1}{4}\ \mathrm{d}t = -\tfrac{5}{36}\left[(t+1)^3\right]_{-1}^{0} - \tfrac{1}{12} + \tfrac{1}{2} = \tfrac{5}{18}\ .$$

Geometrický význam střední hodnoty $E(X)$ je rozdíl velikosti plochy nad grafem F_X v intervalu $(0, +\infty)$ a plochy pod grafem funkce F_X v intervalu $(-\infty, 0)$ (viz žlutá plocha se znaménky v obrázku).

(b) Kvantilová funkce q_X pro veličinu X je definována jako

$$q_X(\alpha) := \frac{1}{2}\Big(\sup\{t \in \mathbb{R} \mid F_X(t) \leq \alpha\} + \inf\{t \in \mathbb{R} \mid F_X(t) \geq \alpha\}\Big)$$

pro $\alpha \in (0,1)$. Tato definice vypadá poněkud složitě, ale je to jen kvůli případným bodům nespojitosti funkce F_X.

Poznámka: Někdy se používá i jiná definice, kde chceme, aby kvantilová funkce byla - stejně jako distribuční funkce - zprava spojitá. V tom případě je definice následující

$$\widetilde{q}_X(\alpha) := \sup\{t \in \mathbb{R} \mid F_X(t) \leq \alpha\} \quad \text{pro} \quad \alpha \in (0,1).$$

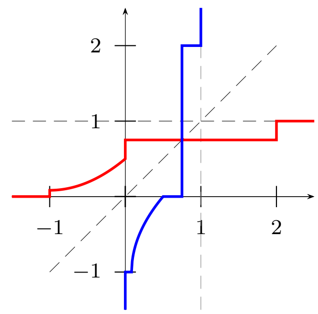
Vždy ale platí toto:

Věta: *Kvantilová funkce q_X je inverzní funkce k distribuční funkci F_X tam, kde F_X je na otevřeném intervalu $I \subseteq \mathbb{R}$ spojitá a ostře rostoucí.*

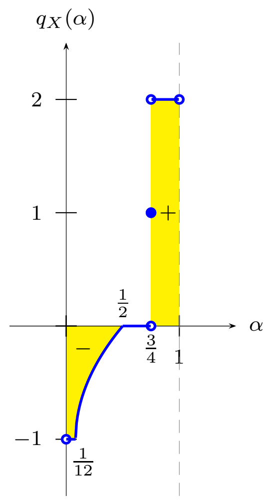
Obecněji pak platí následující: graf kvantilové funkce q_X získáme z grafu distribuční funkce F_X takto

- graf F_X doplníme na "souvislou čáru", tj. skoky funkce F_X nahradíme spojitou svislou úsečkou,
- tento útvar převrátíme podle osy 1. a 3. kvadrantu (tj. podle přímky "$x = y$"),
- tam, kde převrácený útvar není funkcí (tj. obsahuje svislé čáry) tyto úseky odstraníme a nahradíme jedinou hodnotou, a sice průměrem limit z práva a zleva (případné krajní úseky v bodech $\alpha = 0$ a $\alpha = 1$ odstraníme úplně, protože tam se kvantil nedefinuje).

V našem případě tedy graf F_X přejde na



a dostaneme graf q_X:



Kvantilovou funkci určíme také explicitně. Kvantil q_X je inverzní funkcí k ostře rostoucí funkci $F_X(t) = \frac{5}{12}(t+1)^2 + \frac{1}{12}$ pro $t \in (-1, 0)$. Tedy

$$\alpha = \frac{5}{12}(t+1)^2 + \frac{1}{12} \quad \Leftrightarrow \quad t = -1 + \sqrt{\frac{12}{5}\alpha - \frac{1}{5}}$$

a tudíž

$$q_Y(\alpha) = -1 + \sqrt{\frac{12}{5}\alpha - \frac{1}{5}} \quad \text{pro} \quad \alpha \in \left(\frac{1}{12}, \frac{1}{2}\right) \ .$$

Celkově:

$$q_X(\alpha) = \begin{cases} -1 & , \ \alpha \in (0, \frac{1}{12}) \\ -1 + \sqrt{\frac{12}{5}\alpha - \frac{1}{5}} & , \ \alpha \in \langle \frac{1}{12}, \frac{1}{2}) \\ 0 & , \ \alpha \in \langle \frac{1}{2}, \frac{3}{4}) \\ 1 & , \ \alpha = \frac{3}{4} \\ 2 & , \ \alpha \in (\frac{3}{4}, 1). \end{cases}$$

Jestliže nyní budeme uvažovat Kolmogorův model na intervalu $\Omega = (0, 1)$ s rovnoměrným rozdělením pravděpodobnosti, můžeme kvantilovou funkci $q_X : \Omega \to \mathbb{R}$ chápat jako jeden ze způsobů, jak si představit veličinu X, tj. q_X je nyní náhodná veličina se stejnou distribuční funkci jako má X (skutečně je $F_{q_X} = F_X$) - tj. q_X je jeden z modelů veličiny X.

Tato představa má tu výhodu, že na $\Omega = (0,1)$ můžeme "běžně"integrovat. Díky tomu, že interval $(0,1)$ má délku jedna, bude střední hodnota z q_X jednoduše integrál z této funkce (viz žluté plochy se znaménky na obrázku). Dostáváme tak jiný pohled na to, proč je $E(X) = \int_0^1 q_X(\alpha)\ d\alpha$.

(c) Připomeňme si, jak se hledá diskrétní a spojitá část veličiny $X : \Omega \to \mathbb{R}$. Diskrétní část U je zodpovědná za skoky v distribuční funkci F_X. Nejdříve si tedy určíme množinu všech bodů nespojitosti funkce F_X

$$A := \text{"}\textit{množina všech bodů nespojitosti funkce } F_X\text{"} = \{t \in \mathbb{R} \mid P(X = t) \neq 0\} = \{-1, 0, 2\}.$$

Diskrétní část U je zúžení X na množinu

$$\Omega_1 := X^{-1}(A) \subseteq \Omega,$$

tedy

$$U := X|_{\Omega_1},$$

a spojitá část je pak zúžení X na zbytek prostoru Ω, tj.

$$V := X|_{\overline{\Omega_1}}\ .$$

Pak máme

$$X = \mathrm{Mix}_c(U, V),$$

kde

$$c = P(\Omega_1) = P(X \in A) = \sum_{t \in A} P(X = t) = \text{"}\textit{součet všech skoků funkce } F_X\text{"} =$$
$$= P(X = -1) + P(X = 0) + P(X = 2)\ .$$

Skoky dostaneme pomocí vztahu

$$P(X = 0) = P(X \leq 0) - P(X < 0) = F_X(0) - \lim_{t \to 0_-} F_X(t) = \tfrac{3}{4} - \tfrac{1}{2} = \tfrac{1}{4}$$

a podobně

$$P(X = -1) = \tfrac{1}{12} - 0 = \tfrac{1}{12} \qquad \text{a} \qquad P(X = 2) = 1 - \tfrac{3}{4} = \tfrac{1}{4}.$$

Takže

$$c = P(X = -1) + P(X = 0) + P(X = 2) = \tfrac{1}{12} + \tfrac{1}{4} + \tfrac{1}{4} = \tfrac{7}{12}.$$

Pro distribuční funkce máme vztah

$$F_X(t) = cF_U(t) + (1 - c)F_V(t)$$

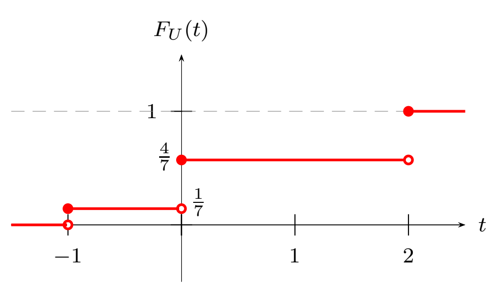
• **Popis diskrétní části U:**
Pro distribuční funkci F_U diskrétní veličiny U platí

$$cF_U(t) = \sum_{a \leq t} P(X = a) = \begin{cases} 0 & ,\ t < -1 \\ \frac{1}{12} & ,\ t \in \langle -1, 0) \\ \frac{1}{12} + \frac{1}{4} = \frac{1}{3} & ,\ t \in \langle 0, 2) \\ \frac{1}{12} + \frac{1}{4} + \frac{1}{4} = \frac{7}{12} & ,\ t \geq 2 \end{cases}$$

takže pro $\frac{1}{c} = \frac{12}{7}$ máme

$$F_U(t) = \begin{cases} 0 & ,\ t < -1 \\ \frac{12}{7} \cdot \frac{1}{12} = \frac{1}{7} & ,\ t \in \langle -1, 0) \\ \frac{12}{7} \cdot \frac{1}{3} = \frac{4}{7} & ,\ t \in \langle 0, 2) \\ \frac{12}{7} \cdot \frac{7}{12} = 1 & ,\ t \geq 2 \end{cases}$$

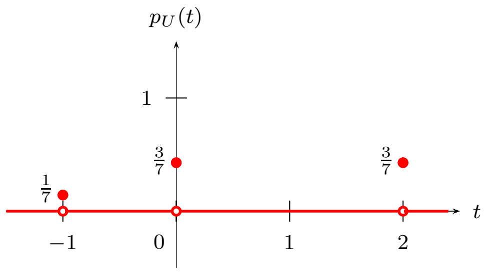
a graf F_U je:



Pro pravděpodobnostní funkci p_U diskrétní veličiny U platí

$$p_U(t) = \frac{1}{c} \cdot P(X = t) = \frac{12}{7} \cdot P(X = t) = \begin{cases} \frac{12}{7} \cdot \frac{1}{12} = \frac{1}{7} & , \ t = -1 \\ \frac{12}{7} \cdot \frac{1}{4} = \frac{3}{7} & , \ t = 0 \\ \frac{12}{7} \cdot \frac{1}{4} = \frac{3}{7} & , \ t = 2 \\ 0 & , \ \text{jinak} \end{cases}$$

s grafem:

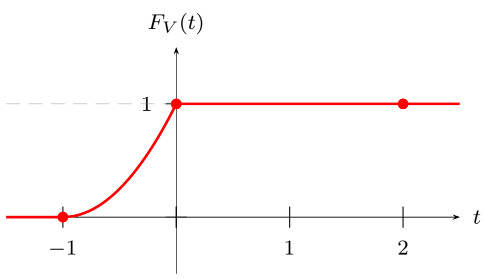


• Popis spojité části V:

Pro distribuční funkci spojité veličiny V pak ze vztahu $F_X = cF_U + (1-c)F_V$ dostáváme

$$F_V(t) = \frac{F_X(t) - cF_U(t)}{1-c} = \frac{12}{5}\Big(F_X(t) - cF_U(t)\Big) = \begin{cases} & 0 & , \ t < -1 \\ \frac{12}{5}\Big(\frac{5}{12}(t+1)^2 + \frac{1}{12} - \frac{1}{12}\Big) = & (t+1)^2 & , \ t \in \langle -1, 0) \\ \frac{12}{5}(\frac{3}{4} - \frac{1}{3}) = & 1 & , \ t \in \langle 0, 2) \\ \frac{12}{5}(1 - \frac{7}{12}) = & 1 & , \ t \geq 2 \end{cases}$$

Graf funkce F_V tedy dostaneme jednoduše tak, že části grafu F_X, které na sebe nenavazují, posuneme dolů tak, aby výsledek byl spojitý. Celý graf pak ve směru y natáhneme tak, aby v nekonečnu měl limitu rovnou 1:

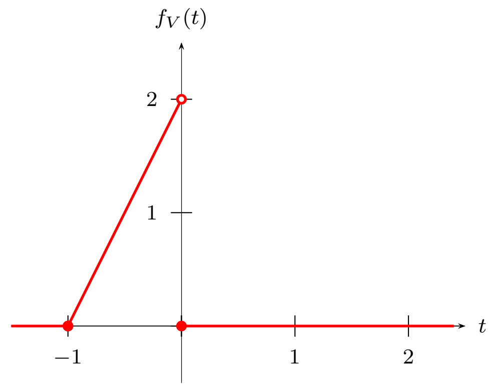


Hustotu f_V spojité veličiny V pak získáme derivací F_V pro body, kde derivace existuje. V ostatních (v tomto případě konečně mnoha bodech) na (nezáporných) hodnotách nezáleží. Takže můžeme psát toto:

$$f_V(t) = F_V'(t) = \Big(\frac{F_X - cF_U}{1-c}\Big)'(t) = \frac{1}{1-c} \cdot F_X'(t)$$

$$f_V(t) = \begin{cases} 2(t+1) & , \ t \in (-1, 0) \\ 0 & , \ \text{jinak.} \end{cases}$$

Graf f_V pak bude:



5.4 (rozklad na směs)

Náhodná veličina X má distribuční funkci

$$F_X(t) = \begin{cases} 0 & , t < 0 \\ \frac{1}{3}t + \frac{1}{2} & , t \in \langle 0, \frac{1}{2}) \\ \frac{2}{3} & , t \in \langle \frac{1}{2}, 1) \\ 1 & , t \geq 1\,. \end{cases}$$

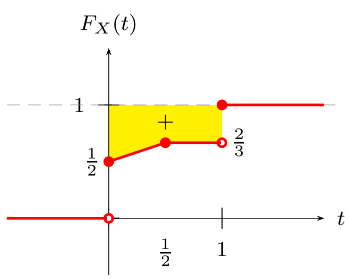
(a) Určete $E(X)$.

(b) Najděte kvantil q_X.

(c) Vyjádřete X jako směs náhodných veličin U a V, z nichž U je diskrétní a V spojitá; popište a znázorněte jejich rozdělení.

Řešení:

(a) Znázorníme si graf F_X:



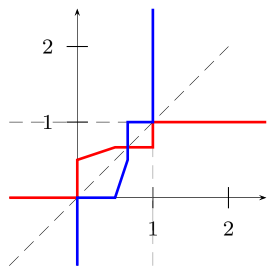
Střední hodnotu můžeme spočítat pomocí F_X jako

$$E(X) = -\int_{-\infty}^{0} F_X(t)\ \mathrm{d}t + \int_{0}^{\infty} (1 - F_X(t))\ \mathrm{d}t =$$

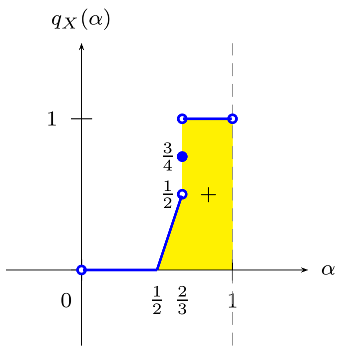
$$= \int_0^{1/2} \frac{1}{2} - \frac{1}{3}t\ \mathrm{d}t + \int_{1/2}^{1} \frac{1}{3}\ \mathrm{d}t = \frac{1}{4} - \frac{1}{6}\left[t^2\right]_0^{1/2} + \frac{1}{6} = \frac{3}{8}\ .$$

Geometrický význam střední hodnoty $E(X)$ je rozdíl velikosti plochy nad grafem F_X v intervalu $(0, +\infty)$ a plochy pod grafem funkce F_X v intervalu $(-\infty, 0)$ (viz žlutá plocha se znaménkem v obrázku).

(b) Kvantilová funkce q_X pro veličinu X je v jistém smyslu "inverzí"k F_X (podrobnosti viz předchozí příklad). V našem případě graf F_X přejde na



a dostaneme graf q_X:



Kvantilovou funkci určíme také explicitně. Kvantil q_X je inverzní funkcí k ostře rostoucí funkci $F_X(t) = \frac{1}{3}t + \frac{1}{2}$ pro $t \in (0, \frac{1}{2})$. Tedy

$$\alpha = \tfrac{1}{3}t + \tfrac{1}{2} \quad \Leftrightarrow \quad t = 3(\alpha - \tfrac{1}{2})$$

a tudíž

$$q_Y(\alpha) = 3(\alpha - \tfrac{1}{2}) \quad \text{pro} \quad \alpha \in \left(\tfrac{1}{2}, \tfrac{2}{3}\right) \ .$$

Celkově:

$$q_X(\alpha) = \begin{cases} 0 & , \ \alpha \in (0, \frac{1}{2}) \\ 3(\alpha - \frac{1}{2}) & , \ \alpha \in \langle \frac{1}{2}, \frac{2}{3}) \\ \frac{3}{4} & , \ \alpha = \frac{2}{3} \\ 1 & , \ \alpha \in (\frac{2}{3}, 1). \end{cases}$$

Střední hodnota $E(X)$ se dá spočítat také jednoduše jako $E(X) = \int_0^1 q_X(\alpha)\, d\alpha$ (viz žluté plochy se znaménky na obrázku).

(c) Připomeňme si, jak se hledá diskrétní a spojité části směsi $X = \mathrm{Mix}_c(U, V)$. Diskrétní část U je zodpovědná za skoky v distribuční funkci F_X. Pro koeficient c ve směsi platí

$$c = P(X \in \underbrace{\text{"\textit{množina všech bodu nespojitosti funkce} } F_X\text{"}}_{=\{0,1\}}) = P(X \in \{0,1\}) = P(X=0) + P(X=1) \ .$$

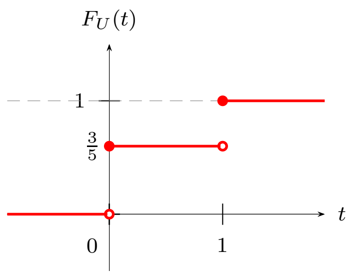
Hodnotu skoku dostaneme pomocí vztahu

$$P(X=0) = P(X \leq 0) - P(X < 0) = F_X(0) - \lim_{t \to 0_-} F_X(t) = \tfrac{1}{2} - 0 = \tfrac{1}{2}$$

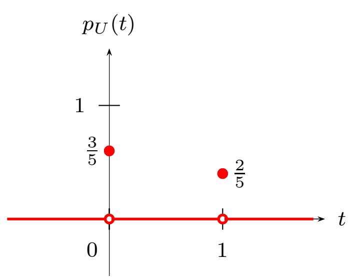
a podobně
$$P(X=1) = 1 - \tfrac{2}{3} = \tfrac{1}{3}\ .$$
Takže
$$c = P(X=0) + P(X=1) = \tfrac{1}{2} + \tfrac{1}{3} = \tfrac{5}{6}.$$
Pro distribuční funkce máme vztah
$$F_X(t) = cF_U(t) + (1-c)F_V(t)$$

- **Popis diskrétní části U:**

Pro distribuční funkci F_U diskrétní veličiny U platí
$$cF_U(t) = \sum_{a \le t} P(X=a) = \begin{cases} 0 & ,\ t<0 \\ \frac{1}{2} & ,\ t \in \langle 0,1) \\ \frac{1}{2}+\frac{1}{3} = \frac{5}{6} & ,\ t \ge 1 \end{cases}$$
takže pro $\frac{1}{c} = \frac{6}{5}$ máme
$$F_U(t) = \begin{cases} 0 & ,\ t<0 \\ \frac{6}{5}\cdot\frac{1}{2} = \frac{3}{5} & ,\ t \in \langle 0,1) \\ \frac{6}{5}\cdot\frac{5}{6} = 1 & ,\ t \ge 1 \end{cases}$$
a graf F_U je:



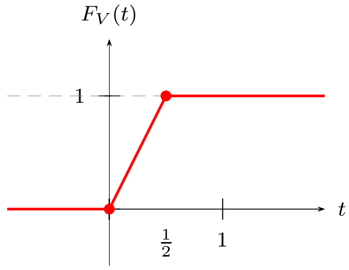
Pro pravděpodobnostní funkci p_U diskrétní veličiny U platí
$$p_U(t) = \tfrac{1}{c}\cdot P(X=t) = \tfrac{6}{5}\cdot P(X=t) = \begin{cases} \frac{6}{5}\cdot\frac{1}{2} = \frac{3}{5} & ,\ t=0 \\ \frac{6}{5}\cdot\frac{1}{3} = \frac{2}{5} & ,\ t=1 \\ 0 & ,\ \text{jinak} \end{cases}$$
s grafem:



- **Popis spojité části V:**

Pro distribuční funkci spojité veličiny V pak ze vztahu $F_X = cF_U + (1-c)F_V$ dostáváme
$$F_V(t) = \frac{F_X(t) - cF_U(t)}{1-c} = 6\Big(F_X(t) - cF_U(t)\Big) = \begin{cases} & 0 & ,\ t<0 \\ 6\Big(\frac{1}{3}t+\frac{1}{2}-\frac{1}{2}\Big) = & 2t & ,\ t\in\langle 0,\frac{1}{2}) \\ 6(\frac{2}{3}-\frac{1}{2}) = & 1 & ,\ t \in \langle \frac{1}{2},1) \\ 6(1-\frac{5}{6}) = & 1 & ,\ t \ge 1 \end{cases}$$

Graf funkce F_V tedy dostaneme jednoduše tak, že části grafu F_X, které na sebe nenavazují, posuneme dolů tak, aby výsledek byl spojitý. Celý graf pak ve směru y natáhneme tak, aby v nekonečnu měl limitu rovnou 1:

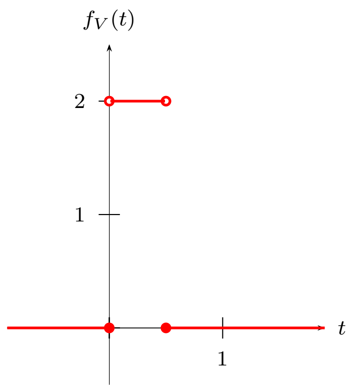


Hustotu f_V spojité veličiny V pak získáme derivací F_V pro body, kde derivace existuje. V ostatních (v tomto případě konečně mnoha bodech) na (nezáporných) hodnotách nezáleží. Takže můžeme psát toto:

$$f_V(t) = F_V'(t) = \left(\frac{F_X - cF_U}{1-c}\right)'(t) = \frac{1}{1-c} \cdot F_X'(t)$$

$$f_V(t) = \begin{cases} 2 & , t \in (0, \frac{1}{2}) \\ 0 & , \text{jinak.} \end{cases}$$

Graf f_V pak bude:



5.5 (geometrické rozdělení - střední hodnota, rozptyl)

Revizor najde v dané tramvaji alespoň jednoho černého pasažéra s pravděpodobností p. Náhodná veličina X je počet tramvají, které revizor projde předtím než najde černého pasažéra.

(a) Určete rozdělení veličiny X.

(b) Vypočtěte její střední hodnotu $E(X)$ a rozptyl $D(X)$.

(c) Jestliže je $p = 26\%$, kolik musí revizor zkontrolovat minimálně tramvají, aby s pravděpodobností alespoň 95% našel černého pasažéra?

Řešení:

(a) Veličina X je zřejmě diskrétní s oborem hodnot $\{0, 1, 2, \dots\}$. Dále si označme jevy

$$A_i = \text{ ''revizor v } i\text{-té tramvaji najde alespoň jednoho černého pasažéra''}$$

pro $i \in \mathbb{N}$. Jevy A_i jsou nezávislé a platí $P(A_i) = p$. Pak pro pravděpodobnostní funkci máme

$$p_X(0) = P(X = 0) = P(A_1) = p$$

a pro $k \in \mathbb{N}$ je

$$p_X(k) = P(X = k) = P(\overline{A_1} \cap \cdots \cap \overline{A_k} \cap A_{k+1}) = P(\overline{A_1}) \cdot \ldots \cdot P(\overline{A_k}) \cdot P(A_{k+1}) = p(1-p)^k$$

(a $p_X(t) = 0$ jinak). Veličina X má tzv. geometrické rozdělení (protože pravděpodobnostní funkce tvoří geometrickou posloupnost).

Distribuční funkce je po částech konstantní a pro $n \in \mathbb{N}_0$ máme

$$F_X(n) = \sum_{k \le n} p_X(k) = \sum_{k=0}^{n} p(1-p)^k = p \cdot \frac{1-(1-p)^{n+1}}{1-(1-p)} = 1-(1-p)^{n+1} \ .$$

(b) Pro výpočet $E(X)$ a $D(X)$ budeme potřebovat sčítat mocninné řady, kde využijeme toho, že uvnitř oboru jejich konvergence lze jejich derivaci získat tak, že je budeme derivovat člen po členu (tj. lze zaměňovat pořadí nekonečné sumace a derivace).

Ze vztahu $\sum_{k=0}^{\infty}(1-p)^k = \frac{1}{1-(1-p)} = \frac{1}{p}$ získáme derivací $\frac{\mathrm{d}}{\mathrm{d}p}$ vztah

$$\underbrace{\frac{\mathrm{d}}{\mathrm{d}p}\left(\frac{1}{p}\right)}_{=-\frac{1}{p^2}} = \frac{\mathrm{d}}{\mathrm{d}p}\left(\sum_{k=0}^{\infty}(1-p)^k\right) = \sum_{k=0}^{\infty}\frac{\mathrm{d}}{\mathrm{d}p}\left((1-p)^k\right) = -\sum_{k=0}^{\infty}k(1-p)^{k-1} = \tfrac{-1}{1-p}\sum_{k=0}^{\infty}k(1-p)^k$$

celkem tedy

$$\sum_{k=0}^{\infty}k(1-p)^k = \frac{1-p}{p^2}$$

což ihned použijeme k výpočtu střední hodnoty

$$E(X) = \sum_{k=0}^{\infty}kp_X(k) = \sum_{k=0}^{\infty}kp(1-p)^k = \frac{1-p}{p} \ .$$

Pro rozptyl máme

$$D(X) = E\Big(\big(X-E(X)\big)^2\Big) = E(X^2) - \Big(E(X)\Big)^2$$

a k výpočtu $E(X^2)$ budeme proto potřebovat ještě jednou použít derivaci na už získaný součet:

$$\underbrace{\frac{\mathrm{d}}{\mathrm{d}p}\left(\frac{1-p}{p^2}\right)}_{=-\frac{2}{p^3}+\frac{1}{p^2}=-\frac{2-p}{p^3}} = \frac{\mathrm{d}}{\mathrm{d}p}\left(\sum_{k=0}^{\infty}k(1-p)^k\right) = \sum_{k=0}^{\infty}\frac{\mathrm{d}}{\mathrm{d}p}\left(k(1-p)^k\right) = -\sum_{k=0}^{\infty}k^2(1-p)^{k-1} = \tfrac{-1}{1-p}\sum_{k=0}^{\infty}k^2(1-p)^k$$

celkem tedy

$$E(X^2) = \sum_{k=0}^{\infty}k^2p_X(k) = \sum_{k=0}^{\infty}k^2p(1-p)^k = \frac{(1-p)(2-p)}{p^2}$$

a tudíž pro rozptyl dostáváme výsledný tvar

$$D(X) = E(X^2) - \Big(E(X)\Big)^2 = \frac{(1-p)(2-p)}{p^2} - \left(\frac{1-p}{p}\right)^2 = \frac{1-p}{p^2}\big(2-p-(1-p)\big) = \frac{1-p}{p^2}$$

což si můžeme lépe zapamatovat jako

$$D(X) = \frac{1}{p} \cdot E(X) \ .$$

(c) Pro $n \in \mathbb{N}$ máme následující zřejmou rovnost jevů

”revizor v prvních n tramvajích najde alespoň jednoho černého pasažéra” = ”$X \le n-1$”

Chceme tedy znát nejmenší $n_0 \in \mathbb{N}$ takové, že $P\Big(X \le n_0 - 1\Big) \ge 0.95$. Takže

$$0.95 \le P\big(X \le n_0 - 1\big) = F_X(n_0 - 1) = 1 - (1-p)^{n_0}$$

neboli

$$0.05 \ge (1-p)^{n_0} = (1-0.26)^{n_0} = 0.74^{n_0}$$

$$\log 0.05 \ge n_0 \cdot \log 0.74$$

$$9.95 \doteq \frac{\log 0.05}{\log 0.74} \le n_0$$

Pozor, logaritmus je záporný pro hodnoty menší než 1! Revizor tedy musí projít alespoň $n_0 = 10$ tramvají.

5.6 (spojitá veličina - střední hodnota, rozptyl)

Nezáporná náhodná veličina X má hustotu

$$f_X(t) = \begin{cases} 0 & , \ t < 1 \\ \frac{c}{t^3} & , \ t \geq 1 \end{cases}$$

kde c je vhodná konstanta. Určete

(a) hodnotu c a distribuční funkci F_X,

(b) střední hodnotu $E(X)$ a rozptyl $D(X)$,

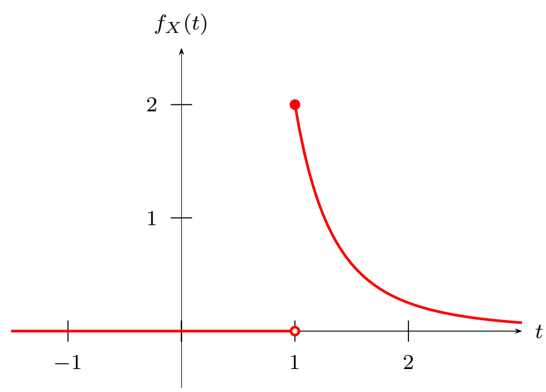
(c) pravděpodobnost $P(X > 2)$.

Řešení:

(a) Pro určení konstanty c potřebujeme, aby f byla nezáporná funkce a platilo, že

$$1 = \int_{-\infty}^{\infty} f_X(t) \ dt = c \int_1^{\infty} t^{-3} \ dt = c \Big[\tfrac{t^{-2}}{-2} \Big]_1^{\infty} = \tfrac{c}{2}$$

tedy $c = 2$. Graf f_X je



Distribuční funkce je rovna

$$F_X(t) = \int_{-\infty}^{t} f_X(u) \ du$$

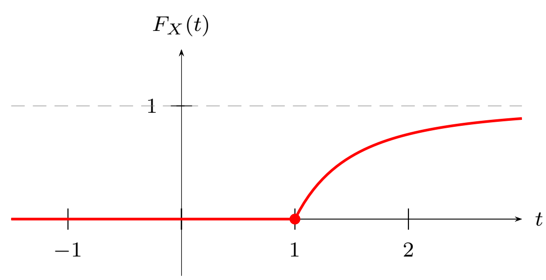
a pro $t \geq 1$ tedy máme

$$F_X(t) = \int_1^t 2u^{-3} \ du = \Big[-u^{-2} \Big]_1^t = 1 - \tfrac{1}{t^2}$$

celkově pak

$$F_X(t) = \begin{cases} 0 & , \ t \leq 1 \\ 1 - \frac{1}{t^2} & , \ t \geq 1 \end{cases}$$

s grafem



(b) Střední hodnotu a rozptyl určíme nejlépe z hustoty pravděpodobnosti:

$$E(X) = \int_{-\infty}^{\infty} t \cdot f_X(t)\ dt = \int_{1}^{\infty} \tfrac{2}{t^2}\ dt = \Big[-\tfrac{2}{t} \Big]_1^{\infty} = 2$$

Pro rozptyl $D(X) = E\Big(\big(X - E(X)\big)^2\Big) = E\big(X^2\big) - \Big(E(X)\Big)^2$ potřebujeme zjistit hodnotu $E(X^2)$. Pro každou borelovskou (tj. "rozumnou") funkci $h : \mathbb{R} \to \mathbb{R}$ máme, že $E(h(X)) = \int_{-\infty}^{\infty} h(t) \cdot f_X(t)\ dt$ pokud alespoň jeden z výrazů má smysl. Tedy

$$E(X^2) = \int_{-\infty}^{\infty} t^2 \cdot f_X(t)\ dt = \int_{1}^{\infty} \tfrac{2}{t}\ dt = \Big[2\ln(t)\Big]_1^{\infty} = +\infty$$

takže i rozptyl $D(X) = +\infty$.

(c)

$$P(X > 2) = \int_{2}^{\infty} f_X(t)\ dt = \int_{2}^{\infty} \tfrac{2}{t^3}\ dt = \Big[-\tfrac{1}{t^2} \Big]_2^{\infty} = \tfrac{1}{4}\ .$$