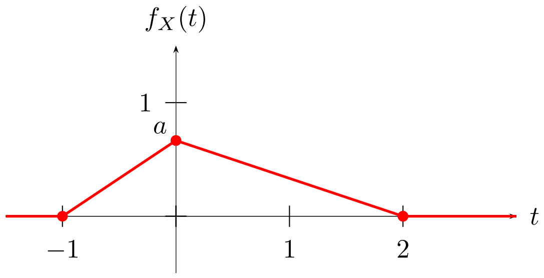
7. cvičení z PST

1. -5. listopadu 2021

7.1 Náhodná veličina X má hustotu danou grafem



kde a je vhodná konstanta. Určete:

(a) hodnotu c a distribuční funkci F_X,

(b) střední hodnotu $E(X)$, rozptyl $D(X)$ a $E\left(|X|^3\right)$,

(c) určete rozdělení veličin $Y = -2X + 1$ a $W = h(X)$, kde h je “ořezávací” funkce na interval $\langle 0, 2\rangle$:

$$h(x) = \begin{cases} 0, & x \leq 0, \\ x, & 0 < x < 2, \\ 2, & x \geq 2. \end{cases}$$

Řešení:

(a) Nezáporná integrabilní funkce $f_X : \mathbb{R} \to \langle 0, +\infty)$ je hustotou pravděpodobnosti právě když platí, že $\int_{-\infty}^{\infty} f_X(t)\, dt = 1$.

$$1 = \int_{-\infty}^{\infty} f_X(t)\, dt = \text{"}\textit{obsah trojúhelníku pod grafem } f_X\text{"} = \frac{3a}{2}\,.$$

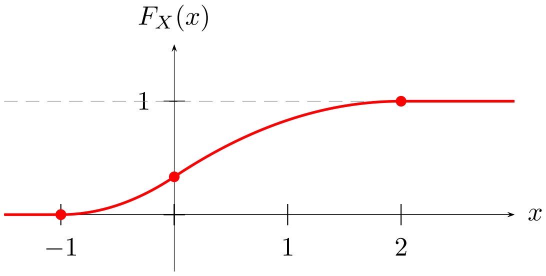
Tedy $a = \frac{2}{3}$. A hustota f_X má tedy předpis

$$f_X(t) = \begin{cases} 0 & , t \leq -1 \\ \frac{2}{3}(t+1) & , t \in \langle -1, 0) \\ -\frac{1}{3}(t-2) & , t \in \langle 0, 2) \\ 0 & , t \geq 2 \end{cases}$$

Distribuční funkce F_X je tudíž rovna

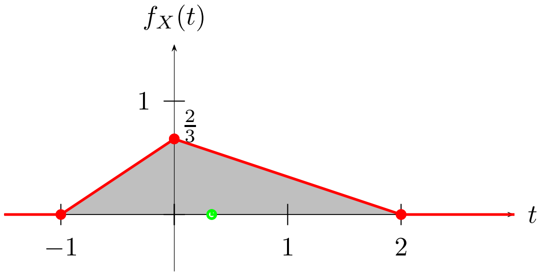
$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt = \begin{cases} \int\limits_{-\infty}^{x} 0\ dt = \mathbf{0} & , x \leq -1 \\ 0 + \int\limits_{-1}^{x} \frac{2}{3}(t+1)\ dt = \left[\frac{1}{3}(t+1)^2\right]_{-1}^{x} = \mathbf{\frac{1}{3}(x+1)^2} & , x \in \langle -1, 0) \\ \int\limits_{-1}^{0} f_X(t)dt + \int\limits_{0}^{x} \frac{1}{3}(2-t)dt = \frac{1}{3} + \left[-\frac{1}{6}(2-t)^2\right]_0^x = \mathbf{1 - \frac{1}{6}(x-2)^2} & , x \in \langle 0, 2) \\ \mathbf{1} & , x \geq 2 \end{cases}$$

s grafem



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

A to buď jako vodorovnou souřadnici (zelený bod) těžiště (šedé) plochy (trojúhelníku) pod grafem hustoty:



$$E(X) = \tfrac{2}{3} \cdot \tfrac{-1+2}{2} = \tfrac{1}{3}$$

anebo výpočtem

$$E(X) = \int_{-\infty}^{\infty} t \cdot f_X(t)\ dt = \int\limits_{-1}^{0} \tfrac{2}{3}t(t+1)\ dt + \int\limits_{0}^{2} \tfrac{1}{3}t(2-t)dt =$$

$$= \tfrac{2}{3}\left[\tfrac{t^3}{3} + \tfrac{t^2}{2}\right]_{-1}^{0} + \tfrac{1}{3}\left[t^2 - \tfrac{t^3}{3}\right]_{0}^{2} = -\tfrac{2}{3}(-\tfrac{1}{3} + \tfrac{1}{2}) + \tfrac{1}{3}(4 - \tfrac{8}{3}) = -\tfrac{1}{9} + \tfrac{4}{9} = \boldsymbol{\tfrac{1}{3}}$$

Pro rozptyl spočítáme ještě

$$E(X^2) = \int_{-\infty}^{\infty} t^2 \cdot f_X(t)\ dt = \int\limits_{-1}^{0} \tfrac{2}{3}t^2(t+1)\ dt + \int\limits_{0}^{2} \tfrac{1}{3}t^2(2-t)dt =$$

$$= \tfrac{2}{3}\left[\tfrac{t^4}{4} + \tfrac{t^3}{3}\right]_{-1}^{0} + \tfrac{1}{3}\left[\tfrac{2t^3}{3} - \tfrac{t^4}{4}\right]_{0}^{2} = -\tfrac{2}{3}(\tfrac{1}{4} - \tfrac{1}{3}) + \tfrac{1}{3}(\tfrac{16}{3} - 4) = -\tfrac{1}{18} + \tfrac{2}{9} = \boldsymbol{\tfrac{1}{6}}$$

Tedy $D(X) = E(X^2) - \Big(E(X)\Big)^2 = \frac{1}{6} - (\frac{1}{3})^2 = \boldsymbol{\frac{1}{18}}$.

Pro obecnou (borelovskou, tj. např. po částech spojitou funkci) $h : \mathbb{R} \to \mathbb{R}$ platí, že

$$E(h(X)) = \int_{-\infty}^{\infty} h(t) \cdot f_X(t)\ dt$$

tedy v našem případě to je

$$E\left(|X|^3\right) = \int_{-\infty}^{\infty} |t|^3 \cdot f_X(t)\ dt = \int_{-1}^{0} -\tfrac{2}{3}t^3(t+1)\ dt + \int_{0}^{2} \tfrac{1}{3}t^3(2-t)dt =$$

$$= -\tfrac{2}{3}\left[\tfrac{t^5}{5}+\tfrac{t^4}{4}\right]_{-1}^{0} + \tfrac{1}{3}\left[\tfrac{t^4}{2}-\tfrac{t^5}{5}\right]_{0}^{2} = \tfrac{2}{3}(-\tfrac{1}{5}+\tfrac{1}{4}) + \tfrac{1}{3}(8-\tfrac{32}{5}) = \tfrac{1}{30}+\tfrac{8}{15} = \mathbf{\tfrac{17}{30}}\ .$$

(c) Určíme distribuční funkci transformované veličiny $Y = g(X)$, kde $g(x) = -2x+1$:

$$F_Y(y) = P(Y \le y) = P(-2X+1 \le y) = P(X \ge \tfrac{1}{2} - \tfrac{y}{2}) =$$

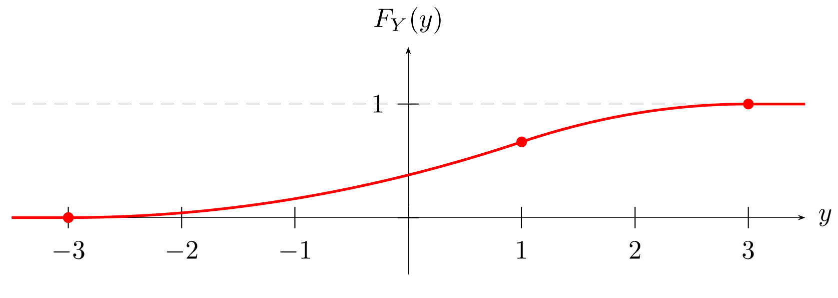
$$= 1 - P(X < \tfrac{1}{2} - \tfrac{y}{2}) = 1 - F_X(\tfrac{1}{2} - \tfrac{y}{2})$$

kde jsme využili spojitosti funkce F_X. Nyní už stačí dosadit $x = g^{-1}(y) = \frac{1}{2} - \frac{y}{2}$ do předpisu pro F_X a využít toho, že díky prostotě klesající funkce f platí, že

$$x \in \langle a, b\rangle \ \Leftrightarrow\ y = g(x) \in \langle g(b), g(a)\rangle$$

$$F_Y(y) = 1 - F_X(\tfrac{1}{2}-\tfrac{y}{2}) = \begin{cases} 1 - 0 = \mathbf{1} & , y \ge g(-1) = 3 \\ 1 - \tfrac{1}{3}(\tfrac{1}{2}-\tfrac{y}{2}+1)^2 = \mathbf{1 - \tfrac{1}{12}(y-3)^2} & , y \in \langle g(0), g(-1)\rangle = \langle 1, 3\rangle \\ 1 - \left(1 - \tfrac{1}{6}(\tfrac{1}{2}-\tfrac{y}{2}-2)^2\right) = \mathbf{\tfrac{1}{24}(y+3)^2} & , y \in \langle g(2), g(0)\rangle = \langle -3, 1\rangle \\ 1 - 1 = \mathbf{0} & , y \le g(2) = -3 \end{cases}$$

s grafem

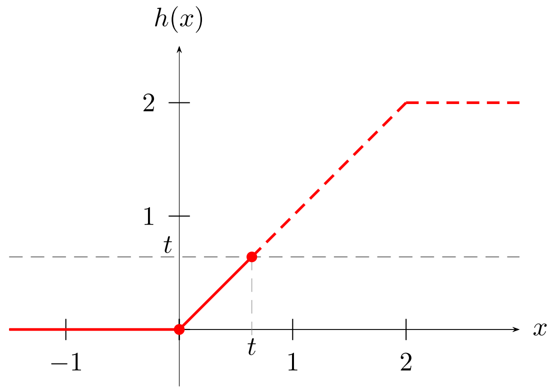


- -

Pro distribuční funkci veličiny $W = h(X)$ máme

$$F_W(t) = P(W \le t) = P(h(X) \le t) = P\Big(h(X) \in (-\infty, t\rangle\Big) = P\Big(X \in h^{-1}(-\infty, t\rangle\Big)$$

Podíváme se proto, jak vypadají množiny $h^{-1}(-\infty, t\rangle$ (tj. vzor intervalu $(-\infty, t\rangle$ při zobrazení h). To snadno uvidíme z grafu funkce h:



Když "uřízneme"z grafu všechno, co je nad danou hladinou t a zbytek promítneme na vodorovnou osu, dostaneme právě hledanou množinu. Tedy:

$$h^{-1}(-\infty, t\rangle = \begin{cases} \emptyset & , \ t < 0 \\ (-\infty, t\rangle & , \ 0 \leq t < 2 \\ \mathbb{R} & , \ t \geq 2. \end{cases}$$

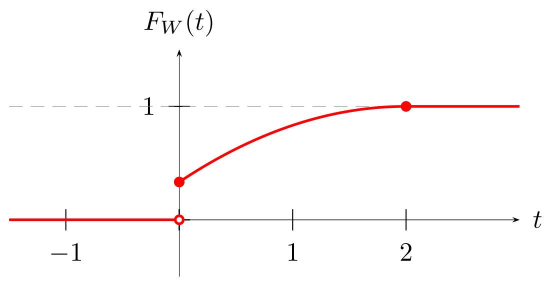
Takže můžeme psát:

$$F_W(t) = P\Big(X \in h^{-1}(-\infty, t\rangle\Big) = \begin{cases} P(X \in \emptyset) = & 0 & , \ t < 0 \\ P\Big(X \in (-\infty, t\rangle\Big) = & F_X(t) & , \ 0 \leq t < 2 \\ P(X \in \mathbb{R}) = & 1 & , \ t \geq 2. \end{cases}$$

Funkci F_W teď získáme výše spočítaným oříznutím na interval $\langle 0, 2)$:

$$F_W(t) = \begin{cases} 0 & , \ t < 0 \\ 1 - \frac{1}{6}(x-2)^2 & , \ t \in \langle 0, \, 2) \\ 1 & , \ t \geq 2 \end{cases}$$

a její graf je:



Veličina W má tedy smíšené rozdělení.

Poznámka: Jak "fyzicky"realizovat ořezávací funkci na interval $\langle 0, 2\rangle$: Jestliže uvážíme např. veličinu z přechozího cvičení X =*"kolik naprší milimetrů srážek v daný den"*, pak funkce h bude odpovídat použití odměrného válce o výšce 2 (např. decimetry), protože takovýto válec umí změřit jen hodnoty vodního sloupce od 0 do 2 decimetrů. Hodnoty nad 2 dm, kdy válec přeteče, pak už od sebe rozeznat neumíme. Tedy

$h(X) =$ *"kolik zjistime milimetrů srážek v daný den v odměrnám válci (výšky 2 dm)"*.

7.2 Nezáporná náhodná veličina X má hustotu

$$f_X(t) = \begin{cases} c \cdot t^2 & , 0 \leq t \leq 2 \\ 0 & , \text{jinak.} \end{cases}$$

kde c je vhodná konstanta. Určete:

(a) hodnotu c a distribuční funkci F_X,

(b) střední hodnotu $E(X)$ a rozptyl $D(X)$ a $E\left(\sqrt{|X|}\right)$,

(c) určete rozdělení veličin $Y = -2X + 1$ a $W = h(X)$, kde h je "ořezávací" funkce na interval $\langle -1, 1 \rangle$:

$$h(x) = \begin{cases} -1, & x \leq -1, \\ x, & -1 < x < 1, \\ 1, & x \geq 1. \end{cases}$$

Řešení:

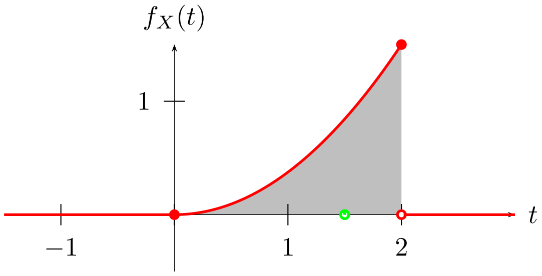
(a) Nezáporná integrabilní funkce $f : \mathbb{R} \to \langle 0, +\infty)$ je hustotou pravděpodobnosti právě když platí, že $\int_{-\infty}^{\infty} f_X(t)\, dt = 1$.

$$1 = \int_{-\infty}^{\infty} f_X(t)\, dt = \int_0^2 ct^2\, dt = c\left[\tfrac{t^3}{3}\right]_0^2 = \frac{8c}{3}\,.$$

Tedy $c = \frac{3}{8}$. A hustota f_X má tedy předpis

$$f_X(t) = \begin{cases} \frac{3}{8} \cdot t^2 & , 0 \leq t \leq 2 \\ 0 & , \text{jinak.} \end{cases}$$

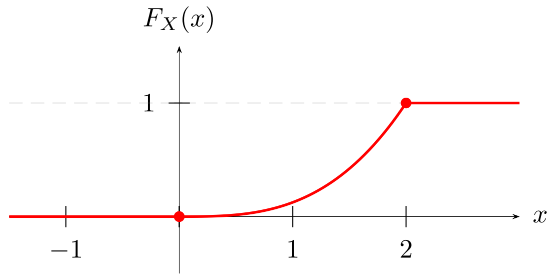
s grafem



Distribuční funkce F_X je tudíž rovna

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt = \begin{cases} \int\limits_{-\infty}^{x} 0\, dt = \mathbf{0} & , x \leq 0 \\ 0 + \int\limits_{0}^{x} \frac{3}{8}t^2\, dt = \left[\frac{t^3}{8}\right]_0^x = \mathbf{\frac{x^3}{8}} & , x \in \langle 0, 2 \rangle \\ \mathbf{1} & , x \geq 2 \end{cases}$$

s grafem



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

$$E(X) = \int_{-\infty}^{\infty} t \cdot f_X(t)\; dt = \int_0^2 \tfrac{3}{8} t^3 \; dt = \tfrac{3}{8} \left[\tfrac{t^4}{4}\right]_0^2 = \mathbf{\tfrac{3}{2}}$$

(Geometrický význam $E(X)$ je vodorovná souřadnice těžiště (šedé) plochy pod grafem f_X. Tato souřadnice je vyznačena jako zelený bod.)

Pro rozptyl spočítáme ještě

$$E(X^2) = \int_{-\infty}^{\infty} t^2 \cdot f_X(t)\; dt = \int_0^2 \tfrac{3}{8} t^4 \; dt = \tfrac{3}{8} \left[\tfrac{t^5}{5}\right]_0^2 = \mathbf{\tfrac{12}{5}}$$

Tedy $D(X) = E(X^2) - \Big(E(X)\Big)^2 = \frac{12}{5} - (\frac{3}{2})^2 = \mathbf{\frac{3}{20}}$.

a nakonec

$$E\left(\sqrt{|X|}\right) = \int_{-\infty}^{\infty} \sqrt{|t|} \cdot f_X(t)\; dt = \int_0^2 \tfrac{3}{8} \underbrace{\sqrt{t} \cdot t^2}_{t^{5/2}} \; dt = \tfrac{3}{8} \left[\tfrac{t^{7/2}}{\frac{7}{2}}\right]_0^2 = \mathbf{\tfrac{6}{7}\sqrt{2}}\ .$$

(c) Určíme distribuční funkci transformované veličiny $Y = g(X)$, kde $g(x) = -2x + 1$:

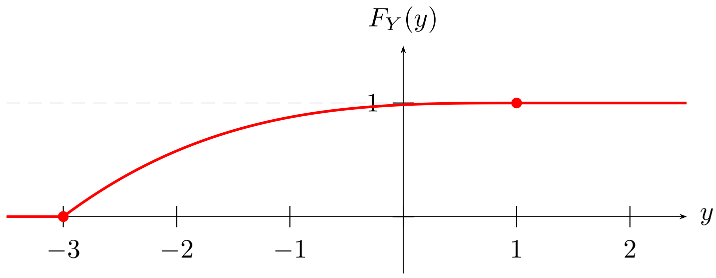
$$\begin{aligned} F_Y(y) = P(Y \le y) = P(-2X + 1 \le y) = P(X \ge \tfrac{1}{2} - \tfrac{y}{2}) = \\ = 1 - P(X < \tfrac{1}{2} - \tfrac{y}{2}) = 1 - F_X(\tfrac{1}{2} - \tfrac{y}{2}) \end{aligned}$$

kde jsme využili spojitosti funkce F_X. Nyní už stačí dosadit $x = g^{-1}(y) = \frac{1}{2} - \frac{y}{2}$ do předpisu pro F_X a využít toho, že díky prostotě klesající funkce g platí, že

$$x \in \langle a, b\rangle \;\; \Leftrightarrow \;\; y = g(x) \in \langle g(b), g(a)\rangle$$

$$F_Y(y) = 1 - F_X(\tfrac{1}{2} - \tfrac{y}{2}) = \begin{cases} 1 - 0 = \mathbf{1} & , y \ge g(0) = 1 \\ 1 - \frac{1}{8}(\frac{1}{2} - \frac{y}{2})^3 = \mathbf{1 + \frac{1}{64}(y-1)^3} & , y \in \langle g(2), g(0)\rangle = \langle -3, 1\rangle \\ 1 - 1 = \mathbf{0} & , y \le g(2) = -3 \end{cases}$$

s grafem

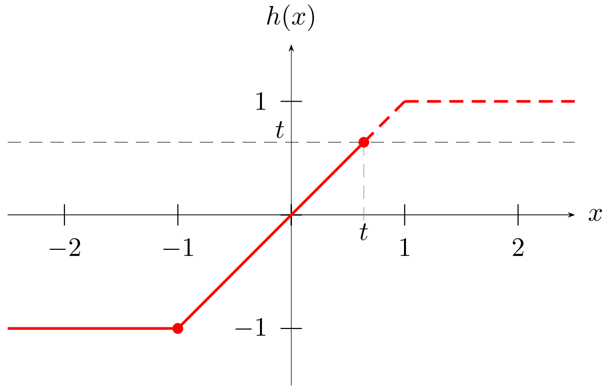


- -

Pro distribuční funkci veličiny $W = h(X)$ máme

$$F_W(t) = P(W \leq t) = P(h(X) \leq t) = P\Big(h(X) \in (-\infty, t\rangle\Big) = P\Big(X \in h^{-1}(-\infty, t\rangle\Big)$$

Podíváme se proto, jak vypadají množiny $h^{-1}(-\infty, t\rangle$. To snadno uvidíme z grafu funkce h:



Když "uřízneme"z grafu všechno, co je nad danou hladinou t a zbytek promítneme na vodorovnou osu, dostaneme právě hledanou množinu. Tedy:

$$h^{-1}(-\infty, t\rangle = \begin{cases} \emptyset & , \ t < -1 \\ (-\infty, t\rangle & , \ -1 \leq t < 1 \\ \mathbb{R} & , \ t \geq 1. \end{cases}$$

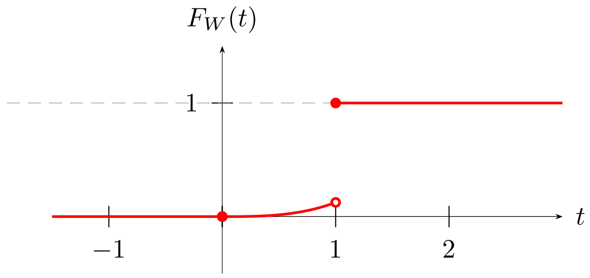
Takže můžeme psát:

$$F_W(t) = P\Big(X \in h^{-1}(-\infty, t\rangle\Big) = \begin{cases} P(X \in \emptyset) = & 0 & , \ t < -1 \\ P\Big(X \in (-\infty, t\rangle\Big) = & F_X(t) & , \ -1 \leq t < 1 \\ P(X \in \mathbb{R}) = & 1 & , \ t \geq 1. \end{cases}$$

Funkci F_W teď získáme výše spočítaným oříznutím na interval $\langle -1, 1)$:

$$F_W(t) = \begin{cases} 0 & , \ t < 0 \\ \frac{t^3}{8} & , \ t \in \langle 0,\ 1) \\ 1 & , \ t \geq 1 \end{cases}$$

a její graf je:



Veličina W má tedy smíšené rozdělení.

Poznámka: Jak "fyzicky"realizovat ořezávací funkci na interval $\langle -1, 1\rangle$: Jestliže uvážíme např. veličinu $X =$*"velikost a směr protékajícího elektrického proudu"*, pak funkce h bude odpovídat použití (analogového, tj. ručičkového) voltmetru o rozsahu ± 1 Ampera. Proudy o absolutní velikosti větší než 1 A, už přístroj nezměří, protože se ručička zastaví o příslušnou zarážku na stupnici. Tedy $h(X) =$*"proud, co naměříme daným ampérmetrem"*.

7.3 Nezáporná náhodná veličina X má hustotu

$$f_X(t) = \begin{cases} c \cdot \frac{1}{t^4} & , 1 \leq t \\ 0 & , \text{jinak.} \end{cases}$$

kde c je vhodná konstanta. Určete:

(a) hodnotu c a distribuční funkci F_X a pravděpodobnost $P\left((X-2)^2 + 1 \in \langle 0, 5\rangle\right)$,

(b) střední hodnotu $E(X)$, rozptyl $D(X)$ a $E\left(\ln|X|\right)$,

(c) určete rozdělení veličin $Y = -2X + 1$ a $W = h(X)$, kde h je "ořezávací" funkce na interval $\langle 1, 2\rangle$:

$$h(x) = \begin{cases} 1, & x \leq 1, \\ x, & 1 < x < 2, \\ 2, & x \geq 2. \end{cases}$$

Řešení:

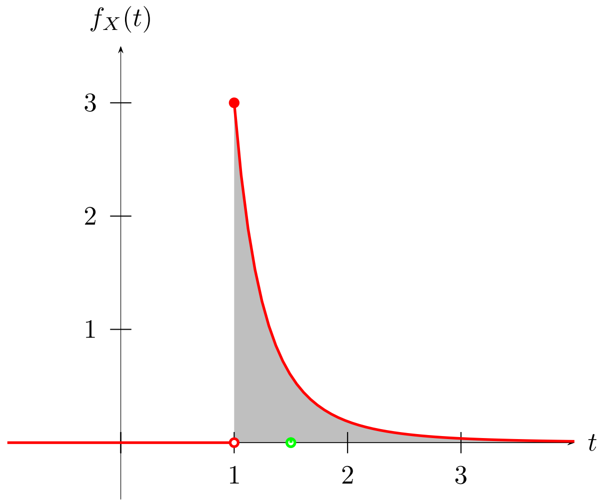
(a) Nezáporná integrabilní funkce $f : \mathbb{R} \to \langle 0, +\infty)$ je hustotou pravděpodobnosti právě když platí, že $\int_{-\infty}^{\infty} f_X(t)\, dt = 1$.

$$1 = \int_{-\infty}^{\infty} f_X(t)\, dt = \int_1^{\infty} \frac{c}{t^4}\, dt = c\Big[-\tfrac{1}{3t^3}\Big]_1^{\infty} = \frac{c}{3}\,.$$

Tedy $c = 3$. A hustota f_X má tedy předpis

$$f_X(t) = \begin{cases} \frac{3}{t^4} & , 1 \leq t \\ 0 & , \text{jinak.} \end{cases}$$

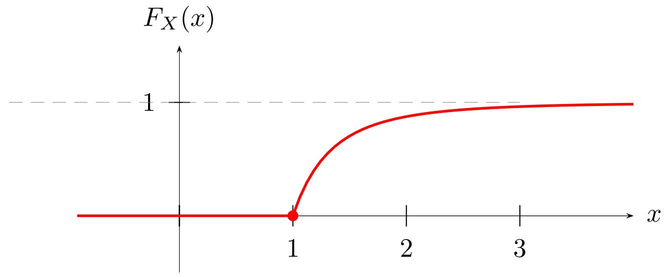
s grafem



Distribuční funkce F_X je tudíž rovna

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\ dt = \begin{cases} \int\limits_{-\infty}^{x} 0\ dt = \mathbf{0} & , x \leq 1 \\ 0 + \int\limits_{0}^{x} \frac{3}{t^4}\ dt = \left[-\frac{1}{t^3}\right]_0^x = \mathbf{1} - \frac{\mathbf{1}}{\mathbf{x^3}} & , x \geq 1 \end{cases}$$

s grafem



Pro určení pravděpodobnosti máme:

$$(X-2)^2 + 1 \in \langle 0, 5\rangle \ \Leftrightarrow\ 0 \leq (X-2)^2 + 1 \leq 5 \ \Leftrightarrow\ -1 \leq (X-2)^2 \leq 4 \ \Leftrightarrow$$
$$\Leftrightarrow\ (X-2)^2 \leq 4 \ \Leftrightarrow\ |X-2| \leq \sqrt{4} = 2 \ \Leftrightarrow\ 0 \leq X \leq 4$$

tedy

$$P\left((X-2)^2 + 1 \in \langle 0, 5\rangle\right) = P\left(0 \leq X \leq 4\right) = F_X(4) - F_X(0) = 1 - \tfrac{1}{4^3} - 0 = \tfrac{63}{64}\ .$$

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

$$E(X) = \int_{-\infty}^{\infty} t \cdot f_X(t)\ dt = \int\limits_{1}^{\infty} \tfrac{3}{t^3}\ dt = 3\left[-\tfrac{1}{2t^2}\right]_1^{\infty} = \tfrac{\mathbf{3}}{\mathbf{2}}$$

(Geometrický význam $E(X)$ je vodorovná souřadnice těžiště (šedé) plochy pod grafem f_X. Tato souřadnice je vyznačena jako zelený bod.)

Pro rozptyl spočítáme ještě

$$E(X^2) = \int_{-\infty}^{\infty} t^2 \cdot f_X(t)\ dt = \int\limits_1^{\infty} \tfrac{3}{t^2}\ dt = 3\Big[-\tfrac{1}{t}\Big]_1^{\infty} = \mathbf{3}$$

Tedy $D(X) = E(X^2) - \Big(E(X)\Big)^2 = 3 - (\frac{3}{2})^2 = \mathbf{\frac{3}{4}}$.
a nakonec

$$E\left(\ln|X|\right) = \int_{-\infty}^{\infty} \ln|t| \cdot f_X(t)\ dt = \int\limits_1^{\infty} \ln t \cdot \tfrac{3}{t^4}\ dt = \{\text{per partes}\} =$$

$$= \underbrace{\Big[\ln t \cdot (-\tfrac{1}{t^3})\Big]_1^{\infty}}_{=0} + \int\limits_1^{\infty} \tfrac{1}{t} \cdot \tfrac{1}{t^3}\ dt = \Big[-\tfrac{1}{3t^3}\Big]_1^{\infty} = \mathbf{\tfrac{1}{3}}\ .$$

(c) Určíme distribuční funkci transformované veličiny $Y = g(X)$, kde $g(x) = -2x + 1$:

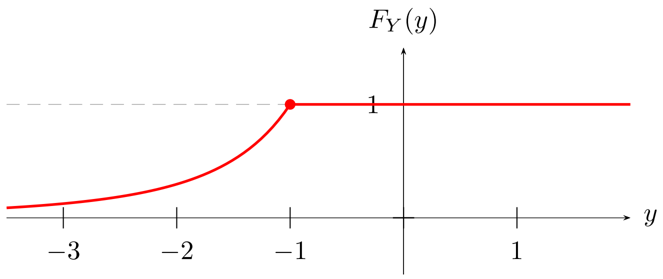
$$F_Y(y) = P(Y \le y) = P(-2X + 1 \le y) = P(X \ge \tfrac{1}{2} - \tfrac{y}{2}) =$$

$$= 1 - P(X < \tfrac{1}{2} - \tfrac{y}{2}) = 1 - F_X(\tfrac{1}{2} - \tfrac{y}{2})$$

kde jsme využili spojitosti funkce F_X.

$$F_Y(y) = 1 - F_X(\tfrac{1}{2} - \tfrac{y}{2}) = \begin{cases} 1 - 0 = \mathbf{1} & , \frac{1}{2} - \frac{y}{2} \le 1 \ \Leftrightarrow\ y \ge -1 \\ 1 - \left(1 - \frac{1}{\left(\frac{1}{2} - \frac{y}{2}\right)^3}\right) = \mathbf{\frac{8}{(1-y)^3}} & , \frac{1}{2} - \frac{y}{2} \ge 1 \ \Leftrightarrow\ y \le -1 \end{cases}$$

s grafem

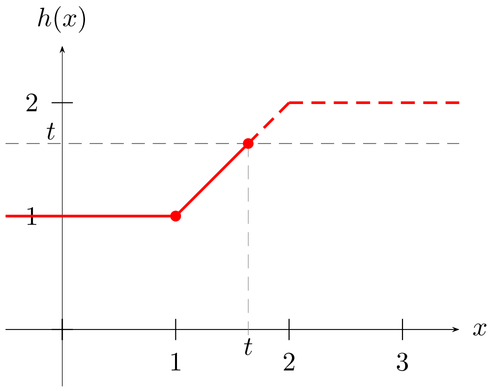


- -

Pro distribuční funkci veličiny $W = h(X)$ máme

$$F_W(t) = P(W \le t) = P(h(X) \le t) = P\Big(h(X) \in (-\infty, t\rangle\Big) = P\Big(X \in h^{-1}(-\infty, t\rangle\Big)$$

Podíváme se proto, jak vypadají množiny $h^{-1}(-\infty, t\rangle$. To snadno uvidíme z grafu funkce h:



Když "uřízneme"z grafu všechno, co je nad danou hladinou t a zbytek promítneme na vodorovnou osu, dostaneme právě hledanou množinu. Tedy:

$$h^{-1}(-\infty,t\rangle = \begin{cases} \emptyset & , \ t<1 \\ (-\infty,t\rangle & , \ 1 \le t < 2 \\ \mathbb{R} & , \ t \ge 2. \end{cases}$$

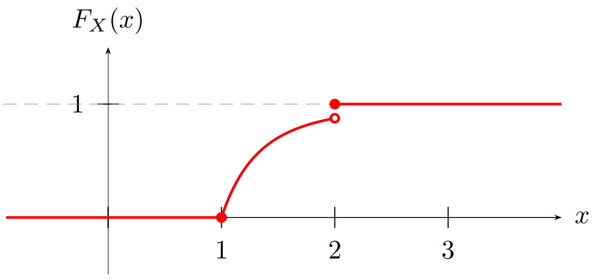
Takže můžeme psát:

$$F_W(t) = P\Big(X \in h^{-1}(-\infty,t\rangle\Big) = \begin{cases} P(X \in \emptyset) = & 0 & , \ t<1 \\ P\Big(X \in (-\infty,t\rangle\Big) = & F_X(t) & , \ 1 \le t < 2 \\ P(X \in \mathbb{R}) = & 1 & , \ t \ge 2. \end{cases}$$

Funkci F_W teď získáme výše spočítaným oříznutím na interval $\langle 1,2)$:

$$F_W(t) = \begin{cases} 0 & , t<1 \\ 1-\frac{1}{x^3} & , t \in \langle 1,\,2) \\ 1 & , t \ge 2 \end{cases}$$

a její graf je:



Veličina W má tedy smíšené rozdělení.

Poznámka: Jak "fyzicky"realizovat ořezávací funkci na interval $\langle 1,2\rangle$: Jestliže uvážíme např. veličinu X =*"výška člověka"*, pak funkce h bude odpovídat např. použití okna ve stěně, jehož spodní okraj je ve výšce 1 m a horní okraj v 2 m, a kterým pozorujeme postavu daného člověka, jenž stojí u stěny v místě okna. Pokud je menší než 1 m, nedokážeme určit jeho výšku (a tyto případy pro nás splývají) a pokud je větší než 2 m, pak také nevíme, kolik přesně měří (a tyto případy opět splývají). Tedy $h(X)$ =*"výška člověka určená na základě jeho obrazu v okně"*.

Připomenutí: Jestliže máme dvě náhodné veličiny $X, Y : \Omega \to \mathbb{R}$, pak zobrazení

$$(X, Y) : \Omega \to \mathbb{R}^2$$

nazýváme náhodný (dvousložkový) vektor.

Tedy náhodnému výsledku ω (tj. elementárnímu jevu) přiřadíme dvojici hodnot $(x, y) \in \mathbb{R}^2$. Např. vybranému člověku z množiny lidí Ω přiřadíme jeho tělesnou výšku a hmotnost.

Náhodný vektor (X, Y) opět umí přenést a vytvořit rozdělení pravděpodobnosti na $\mathbb{R}^2$ - a to tak, že každá "rozumná" množina $A \subseteq \mathbb{R}^2$ (např. otevřená množina nebo interval atd.) bude mít prostě pravděpodobnost

$$P_{(X,Y)}(A) := P\Big((X, Y)^{-1}(A)\Big) \ .$$

Rozdělení této pravděpodobnosti $P_{(X,Y)}$ na $\mathbb{R}^2$ můžeme opět úplně popsat pokud známe pravděpodobnosti jen některých speciálních intervalů a ty nám definují tzv. *sdruženou distribuční funkci* $F_{(X,Y)} : \mathbb{R}^2 \to \langle 0, 1\rangle$ jako

$$F_{(X,Y)}(a, b) := P(X \le a, Y \le b)$$

Pojem nezávislosti pro jevy umožňuje přirozeně definovat nezávislost veličin:

Definice: Veličiny X a Y jsou nezávislé $\Leftrightarrow$ $P(X \in I, Y \in J) = P(X \in I) \cdot P(Y \in J)$ pro libovolné intervaly $I, J \subseteq \mathbb{R}$.

Věta: X a Y jsou nezávislé $\Leftrightarrow$ $F_{(X,Y)}(a, b) = F_X(a) \cdot F_Y(b)$ pro všechna $a, b \in \mathbb{R}$.

Definice: Náhodný vektor (X, Y) má *diskrétní rozdělení* $\Leftrightarrow$ existuje $A \subseteq \mathbb{R}^2$, která je konečná nebo spočetná a taková, že $P((X, Y) \in A) = 1$. (tedy vektor má nejvýše spočetně mnoho "zajímavých" hodnot)

V tomto případě pak rozdělení vektoru (X, Y) popisuje *sdružená pravděpodobnostní funkce* $p_{(X,Y)} : \mathbb{R}^2 \to \langle 0, +\infty)$ definovaná jako

$$p_{(X,Y)}(a, b) := P(X = a, Y = b)$$

a platí

$$F_{(X,Y)}(a, b) = P(X \le a, Y \le b) = \sum_{\substack{u \le a \\ t \le b}} p_{(X,Y)}(u, t) \ .$$

Věta: Nechť náhodný vektor (X, Y) má diskrétní rozdělení. Pak jsou veličiny X a Y nezávislé $\Leftrightarrow$

$$p_{X,Y}(i, j) = p_X(i) \cdot p_Y(j) \quad \text{pro všechna} \quad (i, j) \in \mathbb{R}^2 \ .$$

Poznámky ke kovarianci a korelaci: Náhodné veličiny (jako funkce na pravděpodobnostním prostoru Ω) tvoří přirozeně (reálný) vektorový prostor (kde ještě navíc dvě veličiny budeme pokládat za totožné, pokud se rovnají s pravděpodobností 1). Na vektorovém *pod*prostoru veličin s konečnou střední hodnotou a konečným rozptylem pak můžeme přirozeným způsobem zavést skalární součin jako

$$X \bullet Y := E(X \cdot Y)$$

Díky němu můžeme přirozeně zavést *normu* $\|X\|$ (neboli "délku"vektoru X) jako

$$\|X\| := \sqrt{X \bullet X} = \sqrt{E(X^2)} \ .$$

Mimo jiné si všimněme, že pro X je $D(X) = \|X - E(X)\|^2$.

Skalární součin nám dále umožňuje měřit také úhel mezi dvěma vektory. Pro veličiny X a Y je užitečné znát, jestli jejich výchylky vůči středním hodnotám (tj. veličiny $X - EX$ a $Y - EY$) mají podobné chování (tj. jestli korelují). Zavádíme proto korelaci mezi veličinami X a Y jako kosinus úhlu α mezi vektory $X - EX$ a $Y - EY$, tedy

$$\varrho(X,Y) := \frac{(X - EX) \bullet (Y - EY)}{\|X - EX\| \cdot \|Y - EY\|} = \dots = \frac{E(XY) - E(X) \cdot E(Y)}{\sqrt{D(X) \cdot D(Y)}} .$$

(Veličiny $X - EX$ a $Y - EY$ mají nulovou střední hodnotu).

7.4 (náhodný vektor - diskrétní)

Diskrétní náhodný vektor (X,Y) má sdruženou pravděpodobnostní funkci $p_{X,Y}$ dánu tabulkou:

$p_{X,Y}(x,y)$: y / x	0	1	2
0	1/6	1/9	1/9
1	1/9	2/9	0
2	1/6	0	1/9

(a) Stanovte pravděpodobnost $P(\frac{1}{2} \leq X \leq 3 \;\&\; Y \geq 1)$.

(b) Určete rozdělení veličiny $Z = X - Y$.

(c) Určete marginální rozdělení veličin X a Y.

(d) Zjistěte, zda X a Y jsou nezávislé. Pokud ne, popište rozdělení náhodného vektoru (X', Y') se stejnými marginálními rozděleními, jehož složky jsou nezávislé veličiny.

(e) Vypočtěte korelaci $\varrho(X,Y)$.

Řešení:

Na začátku bychom si měli pro pořádek ještě ověřit, že součet všech pravděpodobností v tabulce je $= 1$ (pokud by byl např. < 1, pak nemáme úplnou informaci o rozdělení a nemůžeme dál pokračovat).

(a) Zřejmě

$$P\left(\frac{1}{2} \leq X \leq 3 \;\&\; Y \geq 1\right) = P\Big((X,Y) \in \{(1,1),(1,2),(2,1),(2,2)\}\Big) = \frac{2}{9} + 0 + 0 + \frac{1}{9} = \frac{1}{3} .$$

(b) U veličin X a Y předpokládáme obory hodnot určené tabulkou (ostatní hodnoty mohou být také případně nabyty, ale s nulovou pravděpodobností). Hodnoty z veličiny $Z = X - Y$ si pro přehlednost zapíšeme také do tabulky

$z(x,y) = x - y$: y / x	0	1	2
0	0	-1	-2
1	1	0	-1
2	2	1	0

a pravděpodobnostní funkce p_Z vznikne nasčítáním pravděpodobností pro jednotlivé případy

$$p_Z(z) = P(X - Y = z) = \sum_{\substack{x-y=z \\ x,y \in \mathbb{R}}} \underbrace{P(X = x,\; Y = y)}_{p_{X,Y}(x,y)} = \begin{cases} \frac{1}{9} & , z = -2 \\ \frac{1}{9} + 0 = \frac{1}{9} & , z = -1 \\ \frac{1}{6} + \frac{2}{9} + \frac{1}{9} = \frac{1}{2} & , z = 0 \\ \frac{1}{9} + 0 = \frac{1}{9} & , z = 1 \\ \frac{1}{6} & , z = 2 \\ 0 & , \text{jinak.} \end{cases}$$

(c) Marginální pravděpodobnostní funkce p_X a p_Y (tj. pravděpodobnostní funkce jednotlivých složek vektoru) jsou

$$p_X(i) = P(X = i) = P(X = i, Y \in \mathbb{R}) = \sum_{j \in \mathbb{R}} P(X = i, Y = j) = \sum_{j \in \mathbb{R}} p_{X,Y}(i, j)$$

$$p_Y(j) = P(Y = j) = \cdots = \sum_{i \in \mathbb{R}} p_{X,Y}(i, j)$$

a jejich hodnoty tudíž získáme sečtením pravděpodobností v řádcích (p_X) a sloupcích (p_Y) naší tabulky:

$X \backslash Y$	0	1	2	p_X
0	1/6	1/9	1/9	7/18
1	1/9	2/9	0	1/3
2	1/6	0	1/9	5/18
p_Y	4/9	1/3	2/9	

(d) Veličiny X a Y jsou nezávislé právě když

$$F_{X,Y}(i, j) = F_X(i) \cdot F_Y(j) \quad \text{pro všechna} \quad (i, j) \in \mathbb{R}^2$$

což je v případě existence sdružené hustoty ekvivalentní podmínce

$$p_{X,Y}(i, j) = p_X(i) \cdot p_Y(j) \quad \text{pro všechna} \quad (i, j) \in \mathbb{R}^2 \ .$$

Protože v našem případě např. $p_{X,Y}(0, 0) = \frac{1}{18} \neq \frac{1}{3} \cdot \frac{1}{3} = p_X(0) \cdot p_Y(0)$, tak X a Y jsou závislé.

Nechť (X', Y') je nyní náhodný vektor s nezávislými složkami takovými, že $p_{X'} = p_X$ a $p_{Y'} = p_Y$. Pak pro jeho pravděpodobnostní funkci máme $p_{X',Y'}(i, j) = p_{X'}(i) \cdot p_{Y'}(j)$ a můžeme ji tak popsat následující tabulkou:

$X' \backslash Y'$	0	1	2	$p_{X'}$
0	$\frac{4}{9} \cdot \frac{7}{18} = \frac{14}{81}$	$\frac{1}{3} \cdot \frac{7}{18} = \frac{7}{54}$	$\frac{2}{9} \cdot \frac{7}{18} = \frac{7}{81}$	7/18
1	$\frac{4}{9} \cdot \frac{1}{3} = \frac{4}{27}$	$\frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$	$\frac{2}{9} \cdot \frac{1}{3} = \frac{2}{27}$	1/3
2	$\frac{4}{9} \cdot \frac{5}{18} = \frac{10}{81}$	$\frac{1}{3} \cdot \frac{5}{18} = \frac{5}{54}$	$\frac{2}{9} \cdot \frac{5}{18} = \frac{10}{81}$	5/18
$p_{Y'}$	4/9	1/3	2/9	

(e) Korelace je dána jako

$$\varrho(X, Y) = \frac{cov(X, Y)}{\sqrt{D(X)} \cdot \sqrt{D(Y)}} = \frac{E\Big(\big(X - EX\big) \cdot \big(Y - EY\big)\Big)}{\sqrt{E\Big(\big(X - EX\big)^2\Big)} \cdot \sqrt{E\Big(\big(Y - EY\big)^2\Big)}} =$$

$$= \frac{E(XY) - E(X) \cdot E(Y)}{\sqrt{E(X^2) - E(X)^2} \cdot \sqrt{E(Y^2) - E(Y)^2}} \ .$$

Spočítáme jednotlivé střední hodnoty (můžeme si pomoci i součinem matic):

$$E(X) = 0 \cdot \frac{7}{18} + 1 \cdot \frac{1}{3} + 2 \cdot \frac{5}{18} = \frac{8}{9}$$

$$E(X^2) = 0^2 \cdot \frac{7}{18} + 1^2 \cdot \frac{1}{3} + 2^2 \cdot \frac{5}{18} = \frac{13}{9}$$

$$E(Y) = 0 \cdot \frac{4}{9} + 1 \cdot \frac{1}{3} + 2 \cdot \frac{2}{9} = \frac{7}{9}$$

$$E(Y^2) = 0^2 \cdot \frac{4}{9} + 1^2 \cdot \frac{1}{3} + 2^2 \cdot \frac{2}{9} = \frac{11}{9}$$

$$E(XY) = \sum_{(i,j)\in\mathbb{R}^2} ij \cdot p_{X,Y}(i,j) = (0,1,2) \cdot \begin{pmatrix} 1/6 & 1/9 & 1/9 \\ 1/9 & 2/9 & 0 \\ 1/6 & 0 & 1/9 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} =$$

$$= (1,2) \cdot \begin{pmatrix} 2/9 & 0 \\ 0 & 1/9 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{2}{3}$$

Korelace tedy je

$$\varrho(X,Y) = \frac{E(XY) - E(X) \cdot E(Y)}{\sqrt{E(X^2) - E(X)^2} \cdot \sqrt{E(Y^2) - E(Y)^2}} = \frac{\frac{2}{3} - \frac{8}{9} \cdot \frac{7}{9}}{\sqrt{\frac{13}{9} - (\frac{8}{9})^2} \cdot \sqrt{\frac{11}{9} - (\frac{7}{9})^2}} =$$

$$= -\frac{1}{5}\sqrt{\frac{2}{53}} \doteq -0.0389 \, .$$

Úhel $\alpha \in \langle 0, \pi \rangle$ mezi našimi náhodnými veličinami $X - EX$ a $Y - EY$ je tedy

$$\arccos\big(\varrho(X,Y)\big) \doteq \arccos(-0.0389) \doteq 92,23^\circ \, .$$

7.5 (náhodný vektor - diskrétní)

Diskrétní náhodný vektor (X,Y) má sdruženou pravděpodobnostní funkci $p_{X,Y}$ dánu tabulkou:

$p_{X,Y}(x,y)$: y / x	0	1	2
-1	$1/8$	0	$1/8$
0	0	$1/4$	$1/4$
1	$1/8$	$1/8$	0

Určete:

(a) pravděpodobnost $P(X + Y > 1)$.

(b) rozdělení veličiny $Z = X^2 \cdot Y$.

(c) marginální pravděpodobnostní funkce p_X a p_Y náhodných veličin X a Y.

(d) zda jsou náhodné veličiny X a Y závislé či nezávislé. Pokud nejsou, popište rozdělení náhodného vektoru (X', Y') se stejnými marginálními rozděleními, jehož složky jsou nezávislé veličiny.

(e) koeficient korelace $\varrho(X,Y)$.

Řešení:
Postup je analogický jako v **8.5**.

(a) Máme

$$X \cdot Y > 1 \quad \Leftrightarrow \quad (X, Y) \in \{(0,2),(1,1),(1,2)\}$$

a tedy

$$P\,(X \cdot Y > 1) = P\Big((X,Y) \in \{(0,2),(1,1),(1,2)\}\Big) = \frac{1}{4} + \frac{1}{8} + 0 = \frac{3}{8}\,.$$

(b) U veličin X a Y předpokládáme obory hodnot určené tabulkou (ostatní hodnoty mohou být také případně nabyty, ale s nulovou pravděpodobností). Hodnoty z veličiny $Z = X^2 \cdot Y$ si pro přehlednost zapíšeme také do tabulky

$z(x,y) = x^2 \cdot y$ y / x	0	1	2
-1	0	1	2
0	0	0	0
1	0	1	2

a pravděpodobnostní funkce p_Z vznikne nasčítáním pravděpodobností pro jednotlivé případy

$$p_Z(z) = P(X^2 \cdot Y = z) = \sum_{\substack{x^2 \cdot y = z \\ x,y \in \mathbb{R}}} \underbrace{P(X = x,\ Y = y)}_{p_{X,Y}(x,y)} = \begin{cases} \frac{1}{8} + 0 + \frac{1}{8} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4} & , z = 0 \\ 0 + \frac{1}{8} = \frac{1}{8} & , z = 1 \\ \frac{1}{8} + 0 = \frac{1}{8} & , z = 2 \\ 0 & , \text{jinak.} \end{cases}$$

(c) Marginální pravděpodobnostní funkce p_X a p_Y (tj. pravděpodobnostní funkce jednotlivých složek vektoru) získáme pro jednotlivé hodnoty sečtením pravděpodobností v řádcích (p_X) a sloupcích (p_Y) naší tabulky:

X \ Y	0	1	2	p_X
-1	1/8	0	1/8	1/4
0	0	1/4	1/4	1/2
1	1/8	1/8	0	1/4
p_Y	1/4	3/8	3/8	

(d) V tabulce se vyskytla nulová pravděpodobnost, konkrétně

$$p_{X,Y}(4,2) = 0 \neq \frac{1}{4} \cdot \frac{3}{8} = p_X(4) \cdot p_Y(2)$$

takže X a Y jsou závislé.

Nechť (X', Y') je nyní náhodný vektor s nezávislými složkami takovými, že $p_{X'} = p_X$ a $p_{Y'} = p_Y$. Pak pro jeho pravděpodobnostní funkci máme $p_{X',Y'}(i,j) = p_{X'}(i) \cdot p_{Y'}(j)$ a můžeme ji tak popsat následující tabulkou:

$X' \backslash Y'$	0	1	2	$p_{X'}$
-1	$\frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16}$	$\frac{3}{8} \cdot \frac{1}{4} = \frac{3}{32}$	$\frac{3}{8} \cdot \frac{1}{4} = \frac{3}{32}$	1/4
0	$\frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}$	$\frac{3}{8} \cdot \frac{1}{2} = \frac{3}{16}$	$\frac{3}{8} \cdot \frac{1}{2} = \frac{3}{16}$	1/2
1	$\frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16}$	$\frac{3}{8} \cdot \frac{1}{4} = \frac{3}{32}$	$\frac{3}{8} \cdot \frac{1}{4} = \frac{3}{32}$	1/4
$p_{Y'}$	1/4	3/8	3/8	

(e) Spočítáme korelaci X a Y:

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

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

$$E(Y) = 0 \cdot \frac{1}{4} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} = \frac{9}{8}$$

$$E(Y^2) = 0^2 \cdot \frac{1}{4} + 1^2 \cdot \frac{3}{8} + 2^2 \cdot \frac{3}{8} = \frac{15}{8}$$

$$E(XY) = \sum_{(i,j) \in \mathbb{R}^2} ij \cdot p_{X,Y}(i,j) = (-1) \cdot 2 \cdot \frac{1}{8} + 1 \cdot 1 \cdot \frac{1}{8} = -\frac{1}{8}$$

Korelace tedy je

$$\varrho(X,Y) = \frac{E(XY) - E(X) \cdot E(Y)}{\sqrt{E(X^2) - E(X)^2} \cdot \sqrt{E(Y^2) - E(Y)^2}} = \frac{-\frac{1}{8} - 0 \cdot \frac{9}{8}}{\sqrt{\frac{1}{2} - 0^2} \cdot \sqrt{\frac{15}{8} - (\frac{9}{8})^2}} =$$

$$= -\sqrt{\frac{2}{39}} \doteq -0.2264 \doteq \arccos(103,09^\circ) \ .$$

Úhel mezi náhodnými veličinami $X - E(X)$ a $Y - E(Y)$ je pak $103,09^\circ$.