8. cvičení z PST

8. - 12. listopadu 2021

Připomenutí: Pro náhodný vektor (X,Y) definujeme *sdruženou distribuční funkci* $F_{(X,Y)} : \mathbb{R}^2 \to \langle 0,1\rangle$ jako

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

Je dobré si uvědomit, že náhodný vektor (X,Y) můžeme snadno sestavit z libovolných dvou náhodných veličin X a Y. Zatímco ale k počítání s veličinou X nám stačí znát jen její distribuční funkci F_X, k práci s vektorem nám NESTAČÍ znalost distribučních funkcí jeho složek! Potřebujeme totiž znát, jaký je vztah mezi veličinami X a Y, a ten je schovaný právě ve sdružené distribuční funkci.

Definice: Náhodný vektor (X,Y) má spojité rozdělení se *sdruženou hustotou pravděpodobnosti*

$$f_{X,Y} : \mathbb{R}^2 \to \langle 0,+\infty)$$

$\Leftrightarrow$ $f_{X,Y}$ je integrabilní funkce a pro každou "rozumnou" množinu $A \subseteq \mathbb{R}^2$ (tj. takovou, která se dá získat z intervalu v $\mathbb{R}^2$ pomocí sjednocování, průniku a doplňku) platí, že

$$P\big((X,Y) \in A\big) = \iint\limits_A f_{X,Y}(x,y)\ \mathrm{d}x\mathrm{d}y\ .$$

To nastává právě když

$$F_{X,Y}(a,b) = \int\limits_{-\infty}^{a} \int\limits_{-\infty}^{b} f_{X,Y}(x,y)\ \mathrm{d}x\mathrm{d}y$$

pro každé $a,b \in \mathbb{R}$.

Sdružená hustota $f_{X,Y}$ opět (jako u veličin) NENÍ zdaleka určena jednoznačně, co se týče její funkční hodnoty, ale pouze hodnotami integrálů z této funkce (např. její změnou v konečně mnoha bodech nebo na nějaké hladké křivce se nezmění příslušné integrály, takže i změněná funkce bude také hustotou). Přesněji, dvě nezáporné funkce $f_{X,Y}$ a $g_{X,Y}$ (s integrálem rovným jedné) jsou hustotami pro tutéž sdruženou distribuční funkci $F_{X,Y}$ právě když se rovnají *skoro všude* a zapisuje se to jako

$$f_{X,Y} = g_{X,Y} \quad \text{(s.v.)}\ .$$

(tj. mohou se lišit jen na takové množině $A \subseteq \mathbb{R}^2$, že $\iint\limits_A 1\ \mathrm{d}x\mathrm{d}y = 0$, tj. pokud A má nulový plošný obsah).

Věta: Nechť (X,Y) je náhodný vektor, který má sdruženou hustotu $f_{X,Y}$. Pak veličiny X a Y mají hustoty pravděpodobnosti

$$f_X(x) = \int\limits_{-\infty}^{\infty} f_{X,Y}(x,y)\ dy \quad \text{a} \quad f_Y(y) = \int\limits_{-\infty}^{\infty} f_{X,Y}(x,y)\ dx\ .$$

Následující podmínky jsou ekvivalentní:

- veličiny X a Y jsou nezávislé.
- $F_{(X,Y)}(a,b) = F_X(a) \cdot F_Y(b)$ pro všechna $(a,b) \in \mathbb{R}^2$.
- $f_{(X,Y)}(a,b) = f_X(a) \cdot f_Y(b)$ pro SKORO všechna $(a,b) \in \mathbb{R}^2$.

Postačující podmínka pro ZÁVISLOST veličin X a Y:
Nechť (X,Y) je náhodný vektor, který má sdruženou hustotu $f_{X,Y}$ a nechť f_X a f_Y jsou odpovídající hustoty pro X a pro Y. Položme

$$g(x,y) = f_X(x)\cdot f_Y(y) \quad \text{pro} \quad (x,y)\in\mathbb{R}^2\ .$$

Jestliže jsou obě funkce $f_{X,Y}$ a g spojité na otevřené množině $U\subseteq\mathbb{R}^2$ a alespoň pro jedno $(x_0,y_0)\in U$ je $f_{X,Y}(x_0,y_0)\neq f_X(x_0)\cdot f_Y(y_0)$, pak X a Y jsou ZÁVISLÉ veličiny.

8.1 (spojitý náhodný vektor)
Sdružená hustota náhodných veličin X a Y je

$$f_{(X,Y)}(x,y) = \begin{cases} ce^{-x-\frac{y}{2}}\,, & x>0, y>0\,,\\ 0\,, & \text{jinak}\,. \end{cases}$$

(a) Určete konstantu c.

(b) Jaká jsou marginální rozdělení veličin X a Y?

(c) Určete sdruženou distribuční funkci $F_{X,Y}$.

(d) Jsou veličiny X a Y nezávislé? Zdůvodněte.

(e) Jak vypadá jejich korelační matice?

(f) Určete pravděpodobnost $P(X>Y)$.

Řešení:

(a) Konstantu c určíme z podmínky

$$1 = \iint\limits_{\mathbb{R}^2} f_{X,Y}(x,y)\ dx\ dy = c\int_0^\infty\int_0^\infty e^{-x-\frac{y}{2}}\ dx\ dy = c\left(\int_0^\infty e^{-x}\ dx\right)\cdot\left(\int_0^\infty e^{-\frac{y}{2}}\ dy\right) =$$

$$= c\left[-e^{-x}\right]_{x=0}^{x=\infty}\cdot\left[-2e^{-\frac{y}{2}}\right]_{y=0}^{y=\infty} = 2c\ .$$

Tedy $c=\frac{1}{2}$.

(b) Marginální hustoty (tj. hustoty jednotlivých veličin X a Y) jsou

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\ dy = \begin{cases} \int_0^\infty \frac{1}{2}e^{-x-\frac{y}{2}}dy = e^{-x}\cdot\left[-e^{-\frac{y}{2}}\right]_0^\infty = e^{-x} & \text{pro } x>0,\\ \int_{-\infty}^{\infty} 0\ dy = 0 & \text{pro } x\le 0\ . \end{cases}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\ dx = \begin{cases} \int_0^\infty \frac{1}{2}e^{-x-\frac{y}{2}}dx = \frac{1}{2}e^{-\frac{y}{2}}\cdot\left[-e^{-x}\right]_0^\infty = \frac{1}{2}e^{-\frac{y}{2}} & \text{pro } y>0,\\ \int_{-\infty}^{\infty} 0\ dx = 0 & \text{pro } y\le 0\ . \end{cases}$$

Vidíme tedy, že obě rozdělení jsou exponenciální, konkrétně $X\sim\mathrm{Exp}(1)$ a $Y\sim\mathrm{Exp}(\frac{1}{2})$.

(c) Pro $a, b \geq 0$ máme

$$F_{X,Y}(a,b) = \int_{-\infty}^{a}\int_{-\infty}^{b} f_{X,Y}(x,y)\ \mathrm{d}y\mathrm{d}x = \int_{0}^{a}\int_{0}^{b} \tfrac{1}{2}e^{-x-\frac{y}{2}}\ \mathrm{d}y\mathrm{d}x = \left(\int_{0}^{a} e^{-x}\ dx\right)\cdot\left(\int_{0}^{b} \tfrac{1}{2}e^{-\frac{y}{2}}\ dy\right) =$$

$$= \left[-e^{-x}\right]_{x=0}^{x=a}\cdot\left[-e^{-\frac{y}{2}}\right]_{y=0}^{y=b} = (1-e^{-a})(1-e^{-\frac{b}{2}})$$

Pro ostatní případy (a,b) pak zřejmě bude $F_{X,Y}(a,b) = 0$ (protože budeme integrovat nulovou funkci).

(d) Složky X a Y jsou nezávislé právě tehdy, když

$$f_{X,Y}(x,y) = f_X(x)\cdot f_Y(y) \quad \text{pro skoro všechna} \quad (x,y)\in\mathbb{R}^2\ ,$$

což znamená, že množina bodů, kde uvedená rovnost neplatí má nulový plošný obsah.

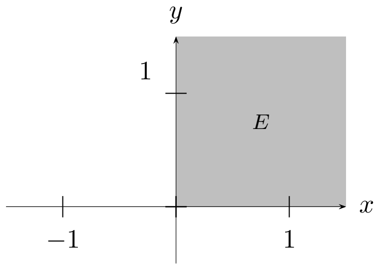
(Podmínce "skoro všude" se nelze vyhnout z toho důvodu, že hustoty nejsou jednoznačně definovány svými hodnotami, ale svými integrály.)

Jak je hned vidět, v našem případě je rovnost splněna dokonce všude, takže X a Y JSOU nezávislé.

(e) Z nezávislosti X, Y plyne okamžitě $\mathrm{cov}(X,Y) = 0$, tedy také $\mathrm{corr}(X,Y) = 0$ a korelační matice je tak

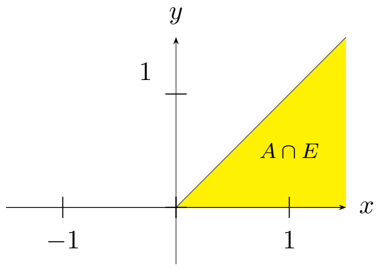
$$\mathrm{Corr}(X,Y) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

(f) Sdružená hustota $f_{(X,Y)}$ je nenulová na množině $E := (0,+\infty)^2$ (vyznačena šedě):



Jev "$X > Y$" je popsán jako $(X,Y) \in \underbrace{\{(x,y)\in\mathbb{R}^2 \mid x > y\}}_{A}$. Pravděpodobnost tohoto jevu tak dostaneme zintegrováním hustoty přes množinu A, neboli

$$P(X>Y) = P\Big((X,Y)\in A\Big) = \iint_{A} f_{X,Y}(x,y)\ dx\ dy = \iint_{A\cap E} f_{X,Y}(x,y)\ dx\ dy =$$



$$= \left\{A \cap E:\ 0 < y < x\right\} = \int_0^\infty \int_0^x \tfrac{1}{2} e^{-x-\frac{y}{2}}\ dy\ dx = \int_0^\infty e^{-x} \left[-e^{-\frac{y}{2}}\right]_{y=0}^{y=x}\ dx =$$

$$= \int_0^\infty e^{-x}(1 - e^{-\frac{x}{2}})\ dy = \int_0^\infty e^{-x} - e^{-\frac{3}{2}x}\ dx = \left[-e^{-x} + \tfrac{2}{3} e^{-\frac{3}{2}x}\right]_{x=0}^{x=\infty} = 1 - \tfrac{2}{3} = \tfrac{1}{3}\ .$$

8.2 (spojitý náhodný vektor)
Náhodný vektor (X, Y) má spojité rovnoměrné rozdělení v množině

$$E = \{(x, y) \in \mathbb{R}^2 \mid y \geq 0,\ x + y \leq 1,\ y - x \leq 1\}\ .$$

Určete:

(a) sdruženou hustotu $f_{X,Y}$.

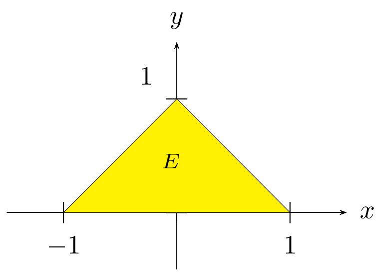
(b) marginální hustoty f_X, f_Y.

(c) zda jsou náhodné veličiny X a Y závislé či nezávislé.

(d) hodnotu $P(X + Y \geq \frac{1}{2})$.

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

Řešení:
Množina E je trojúhelník



(a) Spojité rovnoměrné rozdělení na E odpovídá modelu geometrické pravděpodobnosti na E (můžeme si pro větší názornost představovat E jako terč, do kterého se trefujeme všude se stejnou "intenzitou"). Sdružená hustota je tedy dána jako

$$f_{X,Y}(x, y) = \begin{cases} c & , (x, y) \in E \\ 0 & , \text{jinak,} \end{cases}$$

kde c je konstanta taková, aby integrál z hustoty byl roven 1, tedy

$$1 = \iint_{\mathbb{R}^2} f_{X,Y}(x, y)\ dx\ dy = \iint_E c\ dx\ dy = c \iint_E 1\ dx\ dy = c \cdot \text{obsah}(E) = c \cdot 1$$

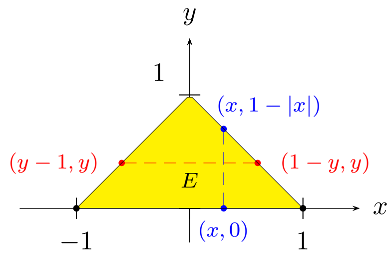
neboli

$$c = 1\ .$$

Integrál $\iint_E 1\ dx\ dy$ jsme mohli buď skutečně spočítat nebo prostě využít jeho geometrickou interpretaci, tj. že je to obsah trojúhelníka.

(b) Marginální rozdělení vektoru (X, Y) jsou rozdělení jeho jednotlivých složek, tj. veličin X a Y. Marginální hustoty f_X a f_Y jsou tedy hustoty náhodných veličin X a Y. V tomto případě je nejsnadněji získáme částečným zintegrováním sdružené hustoty, tedy integrací podél vhodného řezu množiny E, kterou si ještě vyjádříme jako

$$E: \quad -1 \leq x \leq 1 \ \ \& \ \ 0 \leq y \leq 1 - |x|$$



Pro $x \in \langle -1, 1\rangle$ tak máme

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\ dy = \int_0^{1-|x|} 1\ dy = 1 - |x|$$

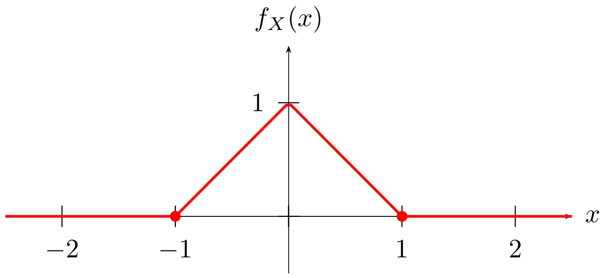
a pro $y \in \langle 0, 1\rangle$ máme

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\ dx = \int_{y-1}^{1-y} 1\ dx = 2(1-y)\ .$$

Celkově tedy

$$f_X(x) = \begin{cases} 1 - |x| & , \ x \in \langle -1, 1\rangle \\ 0 & , \ \text{jinak,} \end{cases}$$

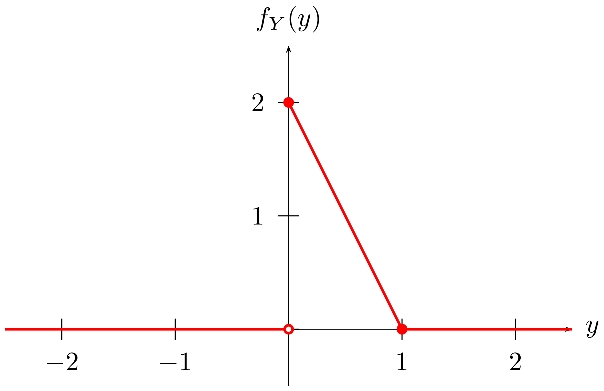
s grafem



a podobně

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

s grafem



(c) Veličiny X a Y jsou nezávislé právě když

$$F_{X,Y}(x,y) = F_X(x) \cdot F_Y(y) \text{ pro všechna } (x,y) \in \mathbb{R}^2 .$$

V případě existence sdružené hustoty je to ekvivalentní tomu, že

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y) \text{ pro SKORO všechna } (x,y) \in \mathbb{R}^2 .$$

(tj. množina bodů, kde rovnost NENASTÁVÁ, má nulový plošný obsah). Obě strany se mohou lišit např. v konečně mnoha bodech, nebo na nějaké křivce atd. Rovnost hustot *skoro všude* se velmi často opomíjí a nevyznačuje se. Nicméně to nic nemění na tom, že je potřeba tuhle věc mít na paměti.

Speciálně, pro nezávislé veličiny musí platit následující (pouze nutná podmínka!):

Nechť

$$S = \{(x,y) \in \mathbb{R}^2 \mid f_{X,Y}(x,y) \neq 0\}$$

a $\pi_1, \pi_2 : \mathbb{R}^2 \to \mathbb{R}$ jsou projekce na jednotlivé souřadné osy, tj. $\pi_1(x,y) = x$ a $\pi_2(x,y) = y$. Pokud jsou veličiny X a Y *nezávislé*, má množina

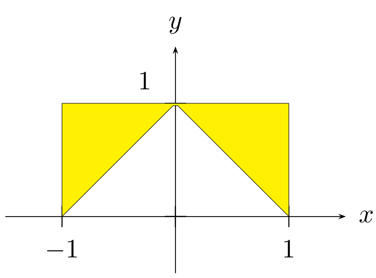
$$\Big(\pi_1(S) \times \pi_2(S)\Big) \setminus S$$

nulový obsah.

V našem konkretním případě $S = E$ a $\pi_1(E) = \langle -1, 1\rangle$ a $\pi_2(E) = \langle 0, 1\rangle$. Ovšem množina

$$\Big(\pi_1(E) \times \pi_2(E)\Big) \setminus E \;=\; \text{“obdélník”} \;\setminus\; \text{“trojúhelník”}$$

tj.



zřejmě nulový obsah NEMÁ. Veličiny X a Y tudíž NEJSOU nezávislé.

Naopak, pokud bychom hledali sdruženou hustotu vektoru (X', Y'), který by měl mít stejná marginální rozdělení jako vektor (X, Y), ale složky X' a Y' by byly nezávislé, stačilo by (podobně jako u diskrétních vektorů a jejich tabulek) položit

$$f_{X',Y'}(x,y) = f_X(x) \cdot f_Y(y) = \begin{cases} 2(1-|x|)(1-y), & (x,y) \in \langle -1,1\rangle \times \langle 0,1\rangle \\ 0, & \text{jinak.} \end{cases}$$

(d) Jev "$X + Y \geq \frac{1}{2}$" je množina $\Phi^{-1}(A)$, kde $\Phi = (X,Y) : \Omega \to \mathbb{R}^2$ je náš náhodný vektor a

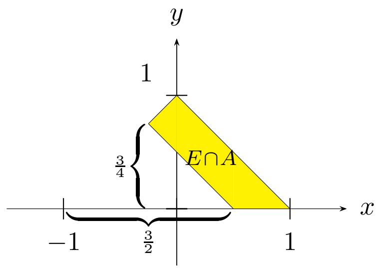
$$A = \{(x,y) \in \mathbb{R}^2 \mid x + y \geq \frac{1}{2}\}$$

tj. $\Phi^{-1}(A)$ je vzor množiny $A \subseteq \mathbb{R}^2$ při zobrazení Φ. Pravděpodobnost tohoto jevu tak dostaneme zintegrováním hustoty přes množinu A, neboli

$$P(X+Y \geq \tfrac{1}{2}) = \iint\limits_{A=\{(a,b)\in\mathbb{R}^2 \mid a+b \geq \frac{1}{2}\}} f_{X,Y}(x,y)\ dx\ dy = \iint\limits_{E\cap A} 1\ dx\ dy = \text{obsah}(E \cap A)\ .$$

Množina $E \cap A$ je tvaru

$$E \cap A : \quad \tfrac{1}{2} \leq x + y \leq 1 \quad \wedge \quad y - x \leq 1 \quad \wedge \quad y \geq 0$$



Velikost plochy $E \cap A$ spočítáme snadněji pomocí doplňkové plochy (tj. trojúhelníka s výškou $\frac{3}{4}$ a základnou $\frac{3}{2}$) do původního trojúhelníka E (s plochou 1):

$$P(X+Y \geq \tfrac{1}{2}) = 1 - \tfrac{1}{2} \cdot \tfrac{3}{4} \cdot \tfrac{3}{2} = \tfrac{7}{16}\ .$$

(e) Pro kovarianci máme

$$\varrho(X,Y) = \frac{E(XY) - E(X)\cdot E(Y)}{\sqrt{D(X)\cdot D(Y)}}\ .$$

Přitom je

$$E(XY) = \iint\limits_{\mathbb{R}^2} xy \cdot f_{X,Y}(x,y)\ dx\ dy = \iint\limits_{A} 2xy\ dx\ dy =$$

$$= \int\limits_0^1 \int\limits_{y-1}^{1-y} 2xy\ dx\ dy = \int\limits_0^1 y \underbrace{\left[x^2\right]_{x=y-1}^{x=1-y}}_{=0}\ dy = 0.$$

To, že tento integrál vyjde nula bylo vidět už na začátku z lichosti integrované funkce (tj. $xy \cdot f_{X,Y}(x,y)$) vzhledem k proměnné x (lichost této funkce je pochopitelně určena i tím, že množina A je symetrická podle osy y). Ze stejných důvodů bude nulový i následující integrál:

$$E(X) = \iint_{\mathbb{R}^2} x \cdot f_{X,Y}(x,y)\ dx\ dy = 0$$

což bylo vidět už z toho, že hustota f_X pro X byla sudá.
Takže $\varrho(X,Y) = 0$, ačkoliv veličiny X a Y nezávislé nejsou (viz (c)).

Poznamenejme ještě, že $E(Y) = \frac{1}{3}$, což snadno zjistíme integrováním pomocí hustoty f_Y a nebo z toho, že plocha pod hustotou je (pravoúhlý) trojúhelník, jehož x-ova složka těžiště je právě ve vzdálenosti $\frac{1}{3}$ od svislé odvěsny.

8.3 (spojitý náhodný vektor)
Sdružená hustota náhodných veličin X a Y je

$$f_{(X,Y)}(x,y) = \begin{cases} c(x+y)\,, & 0 \leq x \leq 1,\ 0 \leq y \leq 1\,, \\ 0\,, & \text{jinak}\,. \end{cases}$$

Určete:

(a) konstantu c.

(b) marginální hustoty f_X a f_Y.

(c) zda jsou veličiny X a Y nezávislé? Zdůvodněte.

(d) sdruženou distribuční funkci $F_{X,Y}$.

(e) střední hodnotu náhodného vektoru $E(X,Y)$ a kovarianci $\mathrm{cov}(X,Y)$.

Řešení:
(a) Konstantu c určíme z podmínky

$$1 = \iint_{\mathbb{R}^2} f_{X,Y}(x,y)\ dx\ dy = c\int_0^1\int_0^1 x+y\ dx\ dy = c\int_0^1 \left[\tfrac{x^2}{2} + xy\right]_{x=0}^{x=1} dy =$$

$$= c\int_0^1 \tfrac{1}{2} + y\ dy = c\left[\tfrac{y}{2} + \tfrac{y^2}{2}\right]_{y=0}^{y=1} = c\,.$$

Tedy $c = 1$.

(b) Marginální hustoty (tj. hustoty jednotlivých veličin X a Y) jsou

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\ dy = \begin{cases} \int_0^1 x+y\ dy = \left[xy + \frac{y^2}{2}\right]_{y=0}^{y=1} = x + \frac{1}{2} & ,\ 0 \leq x \leq 1, \\ \int_{-\infty}^{\infty} 0\ dy = 0 & ,\ \text{jinak}\,. \end{cases}$$

a symetricky

$$f_Y(y) = \begin{cases} y + \frac{1}{2} & , \ 0 \le y \le 1, \\ \\ 0 & , \ \text{jinak} \ . \end{cases}$$

(c) Porovnejme funkce $f_{X,Y}(x,y) = x + y$ a $f_X(x) \cdot f_Y(y) = (x + \frac{1}{2})(y + \frac{1}{2})$ na otevřené množině $U = (0,1)^2$, kde jsou obě spojité. Vidíme, že např. $f_{X,Y}(\frac{1}{3}, \frac{1}{3}) = \frac{2}{3} \neq (\frac{1}{3} + \frac{1}{2})^2 = f_X(\frac{1}{3}) \cdot f_Y(\frac{1}{3})$.

Podle kritéria o závislosti jsou tedy veličiny X a Y *závislé*.

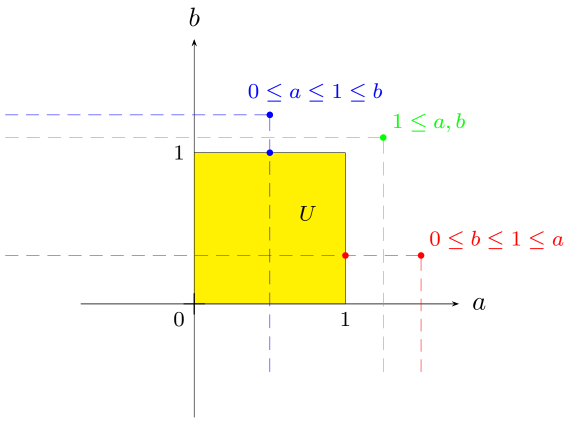
(d) Pro $a, b \in \langle 0, 1\rangle$ máme

$$F_{X,Y}(a,b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f_{X,Y}(x,y) \ \mathrm{d}y \ \mathrm{d}x = \int_{0}^{a} \int_{0}^{b} (x+y) \ \mathrm{d}y \ \mathrm{d}x = \int_{0}^{a} \left[xy + \frac{y^2}{2} \right]_{y=0}^{y=b} dx =$$

$$= \int_{0}^{a} xb + \frac{b^2}{2} \ dx = \left[\frac{x^2}{2} b + \frac{b^2}{2} x \right]_{x=0}^{x=a} = \frac{1}{2} ab(a+b) \ .$$

Pro případy $a < 0$ nebo $b < 0$ pak zřejmě bude $F_{X,Y}(a,b) = 0$ (protože budeme integrovat nulovou funkci).

Ostatní hodnoty $F_{X,Y}(a,b)$ už snadno určíme z obrázku (hustota $f_{X,Y}$ je nulová mimo žlutou oblast $\langle 0,1\rangle^2$):



Např. pro $0 \le b \le 1 \le a$ máme, že

$$F_{X,Y}(a,b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f_{X,Y}(x,y) \ \mathrm{d}y \ \mathrm{d}x = \int_{-\infty}^{1} \int_{-\infty}^{b} f_{X,Y}(x,y) \ \mathrm{d}y \ \mathrm{d}x = F_{X,Y}(1,b) \ .$$

Celkem máme

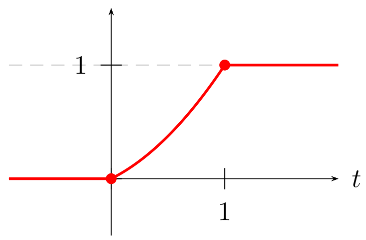
$$F_{X,Y}(a,b) = \begin{cases} 0 & , \ a \le 0 \text{ nebo } b \le 0 \\ \frac{1}{2} ab(a+b) & , \ 0 \le a, b \le 1 \\ F_{X,Y}(1,b) = \frac{1}{2}(b + b^2) & , \ 0 \le b \le 1 \le a \\ F_{X,Y}(a,1) = \frac{1}{2}(a + a^2) & , \ 0 \le a \le 1 \le b \\ 1 & , \ 1 \le a, b. \end{cases}$$

Připomeňme si ještě, jak se marginální distribuční funkce F_X a F_Y získávají jako limity sdružené distribuční funkce:

$$F_X(a) = P(X \leq a) = \lim_{b\to\infty} P(X \leq a, Y \leq b) = \lim_{b\to\infty} F_{X,Y}(a,b) = \begin{cases} 0 & , \ a \leq 0 \\ \frac{1}{2}(a+a^2) & , \ 0 \leq a \leq 1 \\ 1 & , \ 1 \leq a \end{cases}$$

$$F_Y(b) = P(Y \leq b) = \lim_{a\to\infty} P(X \leq a, Y \leq b) = \lim_{a\to\infty} F_{X,Y}(a,b) = \begin{cases} 0 & , \ b \leq 0 \\ \frac{1}{2}(b+b^2) & , \ 0 \leq b \leq 1 \\ 1 & , \ 1 \leq b \end{cases}$$

s grafy



(e) Střední hodnota náhodného vektoru (X,Y) popisuje jakou průměrnou hodnotu dvojic (x,y) můžeme při opakovaných měřeních očekávat. Střední hodnota je přirozeně definována jako

$$E(X,Y) := (E(X), E(Y)) \ .$$

Protože X a Y mají stejná rozdělení, tak máme

$$E(Y) = E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x) \ \mathrm{d}x = \int_0^1 x^2 + \tfrac{x}{2} \ \mathrm{d}x = \tfrac{1}{3} + \tfrac{1}{6} = \tfrac{1}{2}$$

takže $E(X,Y) = (\frac{1}{2}, \frac{1}{2})$.

Pro výpočet kovariance $\mathrm{cov}(X,Y) = E(XY) - E(X) \cdot E(Y)$ potřebujeme ještě hodnotu

$$E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \cdot f_{X,Y}(x,y) \ \mathrm{d}x \ \mathrm{d}y = \int_0^1 \int_0^1 xy(x+y) \ \mathrm{d}x \ \mathrm{d}y =$$

$$= \int_0^1 \int_0^1 x^2 y \ \mathrm{d}x \ \mathrm{d}y + \int_0^1 \int_0^1 xy^2 \ \mathrm{d}x \ \mathrm{d}y = \tfrac{1}{3} \cdot \tfrac{1}{2} + \tfrac{1}{2} \cdot \tfrac{1}{3} = \tfrac{1}{3}$$

takže $\mathrm{cov}(X,Y) = \frac{1}{3} - (\frac{1}{2})^2 = \frac{1}{12}$. Nenulovost kovariance také ukazuje na to, že X a Y jsou závislé.

Poznámky ke kovarianci a korelaci: Kovarianci náhodných veličin definujeme jako

$$\mathrm{cov}(X,Y) := E((X-EX)(Y-EY)) = \cdots = E(XY) - E(X) \cdot E(Y) \ .$$

Kovariance souvisí se skalárním součinem $X \bullet Y := E(X \cdot Y)$ vztahem $\mathrm{cov}(X,Y) = (X - EX) \bullet ((Y - EY)$.

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

$$\|X\| := \sqrt{\langle X|X\rangle} = \sqrt{E(X^2)}\,.$$

Všimněme, že pro X je $D(X) = \|X - E(X)\|^2$.

Korelaci mezi veličinami X a Y definujeme jako kosinus úhlu α mezi vektory $X - EX$ a $Y - EY$, tedy jako

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

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

Kovariance $\operatorname{cov}(\cdot,\cdot)$ má tyto vlastnosti (X, Y, Z jsou veličiny, $a, b, c, d \in \mathbb{R}$):

- je lineární v každé složce zvlášť (tj. je bilineární), tedy:

  $$\operatorname{cov}(aX + bY, Z) = a\cdot\operatorname{cov}(X,Z) + b\cdot\operatorname{cov}(Y,Z)$$
  $$\operatorname{cov}(Z, aX + bY) = a\cdot\operatorname{cov}(Z,X) + b\cdot\operatorname{cov}(Z,Y)$$

- symetrická, tj. $\operatorname{cov}(X,Y) = \operatorname{cov}(Y,X)$
- pozitivně semi-definitní, tj. $\operatorname{cov}(X,X) \geq 0$, kde navíc platí, že:
  $\operatorname{cov}(X,X) = 0 \;\Rightarrow\; \exists\alpha\in\mathbb{R}$, že $P(X=\alpha) = 1$ (neboli: X odpovídá konstantní veličině)
- $\operatorname{cov}(X+c, Y+d) = \operatorname{cov}(X,Y)$,
- $\operatorname{cov}(X,X) = D(X) = (\sigma_X)^2$.

Dále je $\varrho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)\cdot D(Y)}} = \frac{\operatorname{cov}(X,Y)}{\sigma_X\cdot\sigma_Y}$ a proto platí

$$\varrho(aX+c, bY+d) = \operatorname{sgn}(ab)\cdot\varrho(X,Y)$$

pro $a \neq 0$ a $b \neq 0$ a c, d libovolné konstanty (kde sgn je znaménková funkce).

Pro rozptyl $D(\cdot)$ díky tomu máme:

- $D(aX + c) = D(aX) = a^2\cdot D(X)$,
- $D(X + Y) = D(X) + D(Y) + 2\cdot\operatorname{cov}(X,Y)$.

Poznámka: Uvědomme si, že existuje několik stupňů "nezávislosti" veličin:

X a Y jsou nezávislé $\xRightarrow{\text{pokud cov ex.}}$ $\operatorname{cov}(X,Y) = 0$ (tj. $X - E(X)$ a $Y - E(Y)$ jsou kolmé) $\xRightarrow{\text{pokud } X,Y \text{ nejsou konst.}}$ X a Y jsou lineár. nezáv.

Konstantní veličina X spolu s jakoukoliv jinou veličinou Y vždy tvoří vzájemně nezávislé veličiny X a Y (tento případ je ale celkem nezajímavý).

8.4 (kovarianční a korelační matice)

Náhodný vektor (X, Y) má kovarianční matici.

$$\begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}$$

(a) Rozhodněte, zda jsou náhodné veličiny X a Y závislé či nezávislé.

(b) Napište korelační matici náhodného vektoru (X, Y). Určete $E(X + 2Y - 1)$, jestliže $E(X) = 2$ a $E(Y) = -1$.

(c) Napište kovarianční matici náhodného vektoru $(X + Y, -2Y + 1)$.

Řešení:

(a) Kovarianční matice pro (X,Y) představuje tyto parametry:

$$\begin{pmatrix} \mathrm{cov}(X,X) & \mathrm{cov}(X,Y) \\ \mathrm{cov}(Y,X) & \mathrm{cov}(Y,Y) \end{pmatrix} = \begin{pmatrix} D(X) & \mathrm{cov}(X,Y) \\ \mathrm{cov}(X,Y) & D(Y) \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}$$

Máme tedy $\mathrm{cov}(X,Y) = 2 \neq 0$ a proto veličiny X a Y musí být závislé. (Jestliže by kovariance byla nulová, nemohli bychom udělat o nezávislosti žádný závěr.)

(b) Korelační matice pro (X,Y):

$$\begin{pmatrix} \varrho(X,X) & \varrho(X,Y) \\ \varrho(Y,X) & \varrho(Y,Y) \end{pmatrix} = \begin{pmatrix} 1 & \frac{\mathrm{cov}(X,Y)}{\sqrt{D(X)\cdot D(Y)}} \\ \frac{\mathrm{cov}(X,Y)}{\sqrt{D(X)\cdot D(Y)}} & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & 1 \end{pmatrix}$$

protože máme $\varrho(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sqrt{D(X)\cdot D(Y)}} = \frac{2}{\sqrt{3\cdot 4}} = \frac{1}{\sqrt{3}}$.

Všimněme si, že kovarianční matice nám neudává střední hodnoty veličin X a Y. Jestliže $E(X) = 2$ a $E(Y) = -1$, pak $E(X + 2Y - 1) = E(X) + 2E(Y) - 1 = 2 + 2 \cdot (-1) - 1 = -1$.

(c) Pro kovarianční matici náhodného vektoru $(X + Y, -2Y + 1)$ potřebujeme:

$$D(X+Y) = D(X) + 2\mathrm{cov}(X,Y) + D(X) = 3 + 2\cdot 2 + 4 = 11$$

$$D(-2Y+1) = D(-2Y) = (-2)^2 \cdot D(Y) = 4 \cdot 4 = 16$$

$$\begin{aligned} &\mathrm{cov}(X+Y, -2Y+1) = \mathrm{cov}(X+Y, -2Y) = \mathrm{cov}(X, -2Y) + \mathrm{cov}(Y, -2Y) = \\ &= -2\mathrm{cov}(X,Y) - 2\mathrm{cov}(Y,Y) = (-2)\cdot 2 + (-2) \cdot 4 = -12 \end{aligned}$$

Hodnotu kovariance můžeme pomocí původní kovarianční matice také spočítat jako

$$\mathrm{cov}(X+Y, -2Y+1) = \mathrm{cov}(X+Y, -2Y) = (1,1)\begin{pmatrix} 3 & 2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} 0 \\ -2 \end{pmatrix} = -12$$

Kovarianční matice pro $(X + Y, -2Y + 1)$ tedy je:

$$\begin{pmatrix} 11 & -12 \\ -12 & 16 \end{pmatrix}$$

8.5 (kovariance, korelace)

Náhodný vektor (X,Y) má následující parametry:

$$E(X) = 10, \quad \sigma_X = 5, \quad E(Y) = 150, \quad \sigma_Y = 20, \quad \varrho(X,Y) = 0.5 \text{ (korelace)}.$$

(a) Stanovte střední hodnotu a rozptyl náhodných veličin

$$T = 2X + 3, \quad U = 200 - Y, \quad V = X + Y\ .$$

(b) Určete kovarianci $\mathrm{cov}(T,U)$ a rozhodněte, zda jsou náhodné veličiny T a U závislé či nezávislé.

Řešení:

(a) Střední hodnota je lineární zobrazeni ("veličina"$\mapsto$ "její střední hodnota") :

$$E(T) = E(2X+3) = 2E(X) + 3 = 2 \cdot 10 + 3 = 23$$

$$E(U) = E(200 - Y) = 200 - E(Y) = 200 - 150 = 50$$

$$E(V) = E(X + Y) = E(X) + E(Y) = 10 + 150 = 160$$

Rozptyl na druhou stranu NENÍ lineární zobrazení. (Je to tzv. kvadratická forma, tedy vzniká z bilineárního zobrazení - viz výše.):

$$D(T) = D(2X + 3) = D(2X) = 2^2 \cdot D(X) = 2^2 \cdot (\sigma_X)^2 = 4 \cdot 25 = 100$$

$$D(U) = D(200 - Y) = (-1)^2 \cdot D(Y) = (\sigma_Y)^2 = 400$$

$$\begin{aligned} D(V) &= D(X+Y) = \operatorname{cov}(X+Y, X+Y) = \\ &= \operatorname{cov}(X,X) + \operatorname{cov}(X,Y) + \operatorname{cov}(Y,X) + \operatorname{cov}(Y,Y) = \\ &= D(X) + 2 \cdot \operatorname{cov}(X,Y) + D(Y) = \\ &= (\sigma_X)^2 + 2 \cdot \sigma_X \cdot \sigma_Y \cdot \varrho(X,Y) + (\sigma_Y)^2 = \\ &= 25 + 2 \cdot 5 \cdot 20 \cdot 0.5 + 400 = 525 \ . \end{aligned}$$

(b) $\operatorname{cov}(T,U) = \operatorname{cov}(2X + 3, 200 - Y) = \operatorname{cov}(2X, -Y) =$

$$= 2 \cdot (-1) \cdot \operatorname{cov}(X,Y) = (-2) \cdot \sigma_X \cdot \sigma_Y \cdot \varrho(X,Y) = (-2) \cdot 5 \cdot 20 \cdot 0.5 = -100$$

Protože kovariance vyšla nenulová, můžeme hned říct, že U z T musí být *závislé.* V případě, že by kovariance byla nulová, bychom o nezávislosti nic usoudit nemohli!

8.6 (kovariance, kovarianční a korelační matice)
Pro náhodné veličiny X a Y platí, že

$$E(X) = 2, \quad E(Y) = -1, \quad D(X) = 3, \quad D(Y) = 4, \quad \operatorname{cov}(X,Y) = -2.$$

Pro náhodné veličiny

$$U = 3X + 4Y - 1 \quad \text{a} \quad V = -2X + 2Y + 3$$

určete

(a) koeficient kovariance $\operatorname{cov}(U,V)$ a střední hodnotu $E(U)$.

(b) rozptyl $D(X+Y)$.

(c) kovarianční a korelační matice náhodných vektoru (X,Y) a $(X,-2Y)$.

Řešení:
(a) Díky bilinearitě kovariance můžeme jednotlivé složky "roznásobit":

$$\operatorname{cov}(U,V) = \operatorname{cov}(3X + 4Y - 1,\ -2X + 2Y + 3) = \operatorname{cov}(3X + 4Y,\ -2X + 2Y) =$$

$$= 3 \cdot (-2) \cdot \operatorname{cov}(X,X) + 3 \cdot 2 \cdot \operatorname{cov}(X,Y) + 4 \cdot (-2) \cdot \operatorname{cov}(Y,X) + 4 \cdot 2 \cdot \operatorname{cov}(Y,Y) =$$

$$= (-6) \cdot D(X) + (-2) \cdot \operatorname{cov}(X,Y) + 8 \cdot D(Y) = (-6) \cdot 3 + (-2) \cdot (-2) + 8 \cdot 4 = 18 \ .$$

Jak je vidět, znalosti středních hodnot jsme zatím vůbec nepotřebovali!

Poznámka: Je dobré si všimnout, ze díky bilinearitě můžeme také používat přehlednější maticový zápis:

$$\operatorname{cov}(aX+bY, cX+dY) = (a,b)\begin{pmatrix}\operatorname{cov}(X,X) & \operatorname{cov}(X,Y)\\ \operatorname{cov}(Y,X) & \operatorname{cov}(Y,Y)\end{pmatrix}\begin{pmatrix}c\\ d\end{pmatrix}$$

V našem případě tedy

$$\operatorname{cov}(U,V) = (3,4)\begin{pmatrix}3 & -2\\ -2 & 4\end{pmatrix}\begin{pmatrix}-2\\ 2\end{pmatrix} = (3,4)\begin{pmatrix}-10\\ 12\end{pmatrix} = 18\ .$$

Střední hodnota:

$$E(U) = E(3X+4Y-1) = 3E(X)+4E(Y)-1 = 3\cdot 2 + 4\cdot(-1) - 1 = 1$$

(b) Využijeme vlastnosti kovariance:

$$D(X+Y) = \operatorname{cov}(X+Y, X+Y) = \operatorname{cov}(X,X) + \operatorname{cov}(X,Y) + \operatorname{cov}(Y,X) + \operatorname{cov}(Y,Y) =$$

$$= D(X) + 2\cdot\operatorname{cov}(X,Y) + D(Y) =$$

$$= 3 + 2\cdot(-2) + 4 = 3\ .$$

(c) Kovarianční matice pro (X,Y):

$$\begin{pmatrix}\operatorname{cov}(X,X) & \operatorname{cov}(X,Y)\\ \operatorname{cov}(Y,X) & \operatorname{cov}(Y,Y)\end{pmatrix} = \begin{pmatrix}D(X) & \operatorname{cov}(X,Y)\\ \operatorname{cov}(X,Y) & D(Y)\end{pmatrix} = \begin{pmatrix}3 & -2\\ -2 & 4\end{pmatrix}$$

Korelační matice pro (X,Y):

$$\begin{pmatrix}\varrho(X,X) & \varrho(X,Y)\\ \varrho(Y,X) & \varrho(Y,Y)\end{pmatrix} = \begin{pmatrix}1 & \frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)\cdot D(Y)}}\\ \frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)\cdot D(Y)}} & 1\end{pmatrix} = \begin{pmatrix}1 & -\frac{1}{\sqrt{3}}\\ -\frac{1}{\sqrt{3}} & 1\end{pmatrix}$$

protože máme $\varrho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sqrt{D(X)\cdot D(Y)}} = \frac{-2}{\sqrt{3\cdot 4}} = -\frac{1}{\sqrt{3}}$.

Podobně pro vektor $(X,-2Y)$ máme korelační matici:

$$\begin{pmatrix}D(X) & \operatorname{cov}(X,-2Y)\\ \operatorname{cov}(X,-2Y) & D(-2Y)\end{pmatrix} = \begin{pmatrix}D(X) & -2\cdot\operatorname{cov}(X,Y)\\ -2\cdot\operatorname{cov}(X,Y) & (-2)^2\cdot D(Y)\end{pmatrix} = \begin{pmatrix}3 & 4\\ 4 & 16\end{pmatrix}$$

a korelační matici:

$$\begin{pmatrix}1 & \frac{4}{\sqrt{3\cdot 16}}\\ \frac{4}{\sqrt{3\cdot 16}} & 1\end{pmatrix} = \begin{pmatrix}1 & \frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{3}} & 1\end{pmatrix}$$

což není překvapení, protože díky vlastnostem korelace (viz výše) máme, že:

$$\varrho(X,-2Y) = \operatorname{sgn}(-2)\cdot\varrho(X,Y) = -\varrho(X,Y)\ .$$

8.7 Délka hrany krychle je náhodná veličina $X \sim \mathrm{Ro}(1,2)$. Určete distribuční funkci náhodné veličiny Y popisující plochu povrchu této krychle.

(a) Určete distribuční funkci F_Y.

(b) Rozhodněte, zda jsou X a Y nezávislé a proč.

(c) Spočtěte $\operatorname{cov}(X,Y)$.

Řešení:
Máme veličiny

$X =$ *"délka hrany krychle"*

$Y =$ *"plocha povrchu krychle"*

takže $Y = 6 \cdot X^2$ a pro distribuční funkci náhodné veličiny Y dostáváme

$$F_Y(y) = P(Y \le y) = P(6X^2 \le y) = P(X^2 \le \tfrac{y}{6}) = \begin{cases} P\left(|X| \le \sqrt{\frac{y}{6}}\,\right) = F_X\left(\sqrt{\frac{y}{6}}\,\right) & , \ y \ge 0 \\ P(\emptyset) = 0 & , \ y < 0 \ . \end{cases}$$

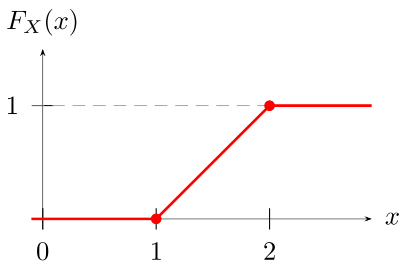
kde jsme využili toho, že obor hodnot pro X je $\langle 1, 2\rangle$, tedy $X \ge 0$ a speciálně tak platí, že $|X| = X$.

Teď si už si jen vyjádříme F_X a dosadíme:

Pro veličinu $X \sim \mathrm{Ro}(1,2)$ je její hustota $f_X(x) = \begin{cases} 1, & 1 \le x \le 2 \\ 0, & \text{jinak} \end{cases}$ a distribuční funkce

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt = \begin{cases} 0, & x < 1 \\ x - 1, & 1 \le x \le 2 \\ 1, & x > 2 \end{cases}$$

s grafem



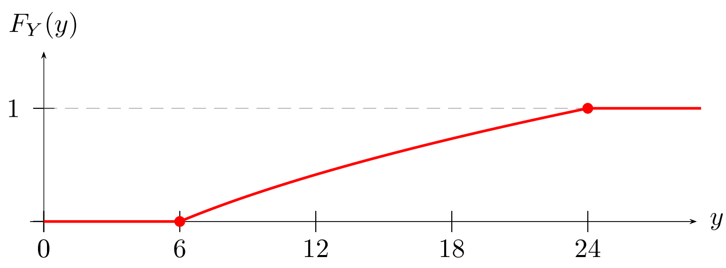
Do F_X (správně!) dosadíme $x = \sqrt{\frac{y}{6}}$ (pro $y \ge 0$) a přepíšeme podmínky pro y:

$$1 \le \sqrt{\frac{y}{6}} \le 2 \quad \Leftrightarrow \quad 1 \le \frac{y}{6} \le 4 \quad \Leftrightarrow \quad 6 \le y \le 24$$

(zbylé podmínky jsou podobné) a dostaneme tak

$$F_Y(y) = \begin{cases} 0, & y < 6 \\ \sqrt{\frac{y}{6}} - 1, & 6 \le y \le 24 \\ 1, & y > 24 \end{cases}$$

s grafem



(b) Jak se dá očekávat, pokud jedna veličina závisí svými hodnotami na druhé, nejspíš nezávislé nebudou. Výjimkou je jen jeden případ a celá situaci se dá popsat takto:

Věta: Nechť X a $h(X)$ jsou obě náhodné veličiny, kde $h : \mathbb{R} \to \mathbb{R}$ je borelovská funkce (např. spojitá). Pak X a $h(X)$ jsou nezávislé veličiny právě jen pokud

- $h(X)$ je konstantní veličina (přesněji: ex. $c \in \mathbb{R}$, že $P\big(h(X) = c\big) = 1$).

Podle distribuční funkce F_Y, není Y konstanta, tedy veličiny X a Y jsou závislé.
Jiným důvodem pro závislost je také to, že $\operatorname{cov}(X, Y) \neq 0$, jak spočítáme níže.

(c) Kovarianci vypočteme ze vztahu

$$\operatorname{cov}(X, Y) = \operatorname{cov}(X, 6X^2) = 6 \cdot \operatorname{cov}(X, X^2) = 6 \cdot \Big(E(X \cdot X^2) - E(X) \cdot E(X^2)\Big)$$

Stačí si tedy pro $n \geq 1$ zjistit

$$E(X^n) = \int_{-\infty}^{\infty} x^n \cdot f_X(x)\, dx = \int_{1}^{2} x^n\, dx = \left[\frac{x^{n+1}}{n+1}\right]_1^2 = \frac{2^n - 1}{n+1}$$

a dosazením dostaneme

$$\operatorname{cov}(X, Y) = 6 \cdot \left(\frac{7}{4} - \frac{1}{2} \cdot 1\right) = \frac{15}{2}\ .$$

8.8 (kovariance)

Nechť veličina X má rovnoměrné rozdělení na intervalu $(-1, 1)$ a $Y = X^2$.

(a) Určete distribuční funkci F_Y.

(b) Rozhodněte, zda jsou X a Y nezávislé a proč.

(c) Spočtěte $\operatorname{cov}(X, Y)$.

Řešení:

(a) Pro distribuční funkci F_Y náhodné veličiny Y dostáváme

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = \begin{cases} P\left(|X| \leq \sqrt{y}\ \right) = P\left(-\sqrt{y} \leq X \leq \sqrt{y}\ \right) = F_X\left(\sqrt{y}\ \right) - F_X\left(-\sqrt{y}\ \right) & ,\ y \geq 0 \\ P(\emptyset) = 0 & ,\ y < 0\ . \end{cases}$$

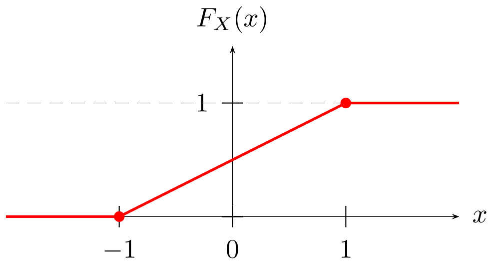
kde jsme využili toho, že X má spojité rozdělení.

Teď si už si jen vyjádříme F_X a dosadíme:

Pro veličinu $X \sim \mathrm{Ro}(-1, 1)$ je její hustota $f_X(x) = \begin{cases} \frac{1}{2}, & -1 \leq x \leq 1 \\ 0, & \text{jinak} \end{cases}$ a distribuční funkce

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt = \begin{cases} 0, & x < -1 \\ \frac{1}{2}(x+1), & -1 \leq x \leq 1 \\ 1, & x > 1 \end{cases}$$

s grafem



Pro dosazení do F_X teď uvážíme případy:

- $0 \le y \le 1$: pak je $-\sqrt{y}, \sqrt{y} \in \langle -1, 1 \rangle$ a tedy

$$F_Y(y) = \dots = F_X\left(\sqrt{y}\,\right) - F_X\left(-\sqrt{y}\,\right) = \tfrac{1}{2}(\sqrt{y}+1) - \tfrac{1}{2}(-\sqrt{y}+1) = \sqrt{y}$$

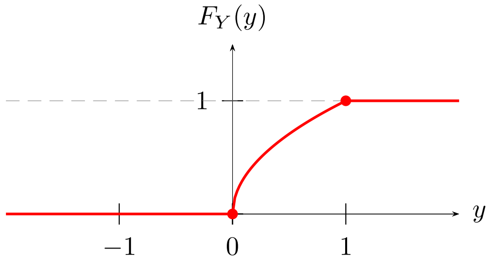
- $1 < y$: pak je $-\sqrt{y} < -1$ a $1 < \sqrt{y}$ a tedy

$$F_Y(y) = \dots = F_X\left(\sqrt{y}\,\right) - F_X\left(-\sqrt{y}\,\right) = 1 - 0 = 1$$

Celkově tedy je

$$F_Y(y) = \begin{cases} 0, & y < 0 \\ \sqrt{y}, & 0 \le y \le 1 \\ 1, & y > 1 \end{cases}$$

s grafem



(b) Jak se dá očekávat, pokud jedna veličina závisí svými hodnotami na druhé, nejspíš nezávislé nebudou. Výjimkou je jen jeden případ a celá situaci se dá popsat takto:

Věta: Nechť X a $h(X)$ jsou obě náhodné veličiny, kde $h : \mathbb{R} \to \mathbb{R}$ je borelovsky měřitelná funkce (např. spojitá). Pak X a $h(X)$ jsou nezávislé veličiny právě jen pokud

- $h(X)$ je skoro všude konstantní veličina, tj. ex. $c \in \mathbb{R}$, že $P\big(h(X) = c\big) = 1$.

V našem případě $Y = X^2$ taková není, což snadno ukážeme sporem:
Nechť $c \in \mathbb{R}$ je takové, že $P\big(X^2 = c\big) = 1$. Zřejmě musí být $c \ge 0$. Protože spojitá veličina X má nulovou pravděpodobnost nabytí dané hodnoty, dostáváme, že

$$1 = P\big(X^2 = c\big) = \underbrace{P(X = \sqrt{c})}_{0} + \underbrace{P(X = -\sqrt{c})}_{0} = 0$$

což je spor.

Takže veličiny X a Y jsou *závislé.*

(c) Kovarianci vypočteme ze vztahu

$$\operatorname{cov}(X, Y) = \operatorname{cov}(X, X^2) = E(X \cdot X^2) - E(X) \cdot E(X^2)$$

Hustota f_X je sudá funkce, takže

$$E(X^3) = \int_{-1}^{1} \underbrace{x^3 \cdot f_X(x)}_{\text{lichá funkce}} \, dx = 0$$

a podobně je $E(X) = 0$ a dosazením dostaneme, že $\operatorname{cov}(X, Y) = 0$. Máme tedy příklad toho, že veličiny X a Y mohou být závislé a přitom nekorelované.