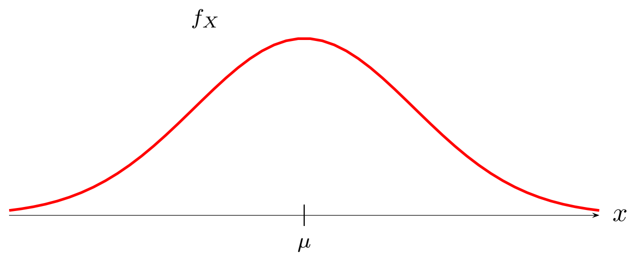
8. cvičení z PST

9. listopadu 2022

Poznámky k normálnímu rozdělení:
Veličina X má *normální* (neboli Gaussovo) rozdělení $N(\mu, \sigma^2)$ (kde $\mu \in \mathbb{R}$ a $\sigma > 0$), jestliže má hustotu

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{pro} \quad x \in \mathbb{R} \ .$$



Je to tedy spojité rozdělení, $E(X) = \mu$, $\text{var}(X) = \sigma^2$ a oborem hodnot veličiny X je celá reálná osa. Všimněme si ještě, že hustota f_X je symetrická vzhledem ke středu μ a proto platí $F_X(\mu) = \frac{1}{2}$.

Toto rozdělení je limitním rozdělením, které aproximuje součty nezávislých stejně (nebo podobně) rozdělených veličin (více později v Centrální limitní větě). Typicky se tedy objevuje u veličin, jejichž hodnoty jsou ovlivněny mnoha drobnými odchylkami (např. u chyb měření, výšky člověka apod.)

Značení: Pro náhodnou veličinu X s konečnou střední hodnotou a konečným rozptylem, položme

$$\text{norm}(X) := \frac{X - E(X)}{\sqrt{\text{var}(X)}} \ .$$

Speciálně tedy vidíme, že $E\big(\text{norm}(X)\big) = 0$ a $\text{var}\big(\text{norm}(X)\big) = 1$.

Platí: Pro takovouto veličinu X a konstanty $a > 0$ a $b \in \mathbb{R}$ je

$$\text{norm}(aX + b) = \text{norm}(X) \ .$$

Důležité vlastnosti normálního rozdělení:

- $X \sim N(\mu, \sigma^2) \ \Rightarrow \ \text{norm}(X) = \frac{X-\mu}{\sigma} \sim N(0,1)$ (je to tzv. normované normální rozdělení s hodnotami v tabulkách) dist. funkce pro $N(0,1)$ se značí Φ.

  V tomto případě pak máme $F_X(t) = P(X \le t) = P\Big(\underbrace{\frac{X-\mu}{\sigma}}_{=\text{norm}(X)} \le \frac{t-\mu}{\sigma}\Big) = \Phi\Big(\frac{t-\mu}{\sigma}\Big)$ pro všechna $t \in \mathbb{R}$.

- hustota $f_{N(0,1)}$ je sudá funkce $\ \Rightarrow \ \Phi(t) + \Phi(-t) = 1$ pro všechna $t \in \mathbb{R}$.

- Nechť $X_i \sim N(\mu_i, \sigma_i^2)$, pro $i = 1, 2$, jsou nezávislé. Pak $X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ (tj. speciálně součet nezávislých normálních rozdělení je zase normální.)

Pro vybraná čísla $t \ge 0$ se dají hodnoty Φ najít ve statistických tabulkách. Pro záporná čísla si pak pomůžeme vztahem $\Phi(-t) = 1 - \Phi(t)$.

Pro lepší představu o tom, jakou roli pro veličinu s normálním rozdělením $X \sim N(\mu, \sigma^2)$ hraje směrodatná odchylka σ se používá tzv.

pravidlo tří-sigma (`https://cs.wikipedia.org/wiki/Pravidlo_t\%C5\%99\%C3\%AD\_sigma`)

které je ovšem čistě jen technickou pomůckou:

Jestliže si budeme počítat pravděpodobnosti

$$P(|X-\mu| \leq k \cdot \sigma) = P\Big(\Big|\underbrace{\frac{X-\mu}{\sigma}}_{\sim N(0,1)}\Big| \leq k\Big) = \Phi(k) - \Phi(-k) = 2 \cdot \Phi(k) - 1 \quad \text{pro} \quad k = 1, 2, 3, \ldots$$

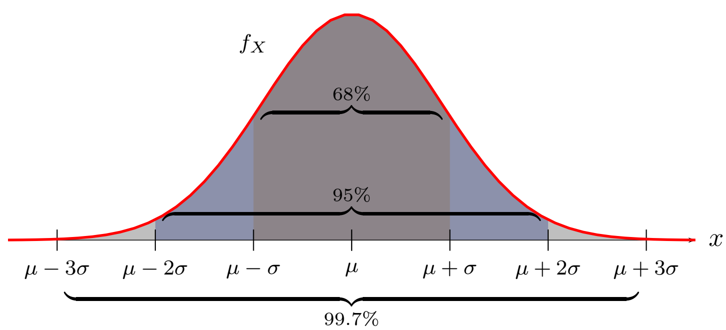
dostaneme postupně

$$P(|X-\mu| \leq \sigma) = 2 \cdot \Phi(1) - 1 \doteq 2 \cdot 0.8413 - 1 = 0.6826 \doteq 68\%$$

$$P(|X-\mu| \leq 2 \cdot \sigma) = 2 \cdot \Phi(2) - 1 \doteq 2 \cdot 0.9772 - 1 = 0.9544 \doteq 95\%$$

$$P(|X-\mu| \leq 3 \cdot \sigma) = 2 \cdot \Phi(3) - 1 \doteq 2 \cdot 0.99865 - 1 = 0.9973 \doteq 99.7\%$$

Pro vyšší hodnoty, tj. $k \geq 4$ už jsou pravděpodobnosti vpodstatě rovny 1, takže se v praxi příliš nepoužívají (záleží samozřejmě na zvolené přesnosti).



Příklad 8.1 *Výška dětí v 1. třídě je náhodná veličina* $X \sim N(130\,\text{cm},\ 36\,\text{cm}^2)$. *Jaká je pravděpodobnost, že náhodně vybrané dítě bude*

(a) větší než 136 *cm,*

(b) menší než 118 *cm,*

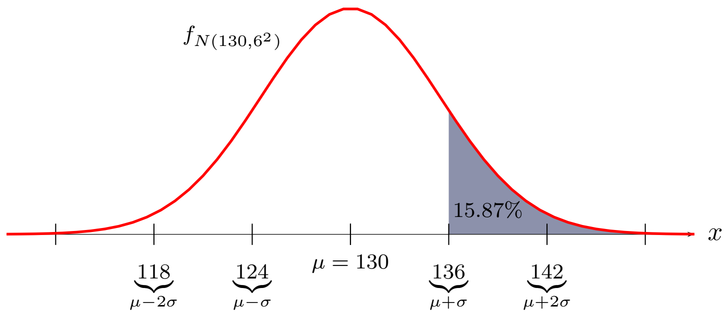
(c) mít výšku mezi 127 *a* 133 *cm?*

Řešení:
Nyní tedy máme $X \sim N(130,\ 36)$. Pro jednodušší zápis si ještě označme $Z := \text{norm}(X)$.

(a)

$$P(X > 136) = P\left(\underbrace{\frac{X-130}{\sqrt{36}}}_{Z} > \underbrace{\frac{136-130}{\sqrt{36}}}_{1}\right) = P(Z > 1) = 1 - P(Z \leq 1) =$$

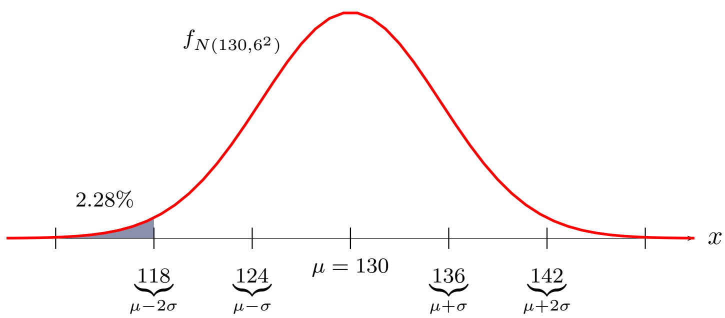
$$= 1 - \Phi(1) \doteq 1 - 0.8413 = 0.1587\ .$$



(b)

$$P(X < 118) = P\left(\frac{X-130}{\sqrt{36}} < \frac{118-130}{\sqrt{36}}\right) = P(Z < -2) = \Phi(-2) =$$

$$= 1 - \Phi(2) \doteq 1 - 0.9772 = 0.0228 \ .$$



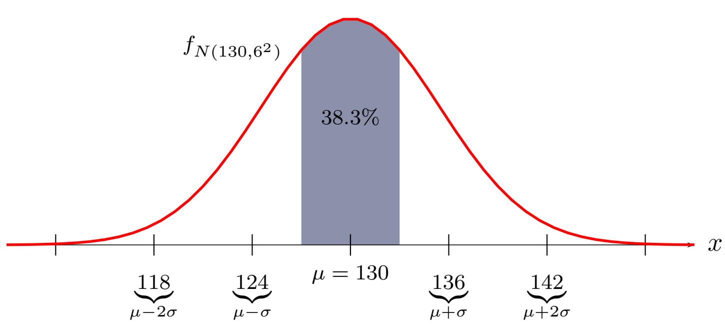
(c)

$$P(127 < X < 133) = P\left(\frac{127-130}{\sqrt{36}} < \frac{X-130}{\sqrt{36}} < \frac{133-130}{\sqrt{36}}\right) =$$

$$= P(-0.5 < Z < 0.5) = P(Z < 0.5) - P(Z \leq -0.5) =$$

$$= \Phi(0.5) - \Phi(-0.5) = \Phi(0.5) - \Big(1 - \Phi(0.5)\Big) =$$

$$= 2 \cdot \Phi(0.5) - 1 \doteq 2 \cdot 0.6915 - 1 = 0.383 \ .$$



Poznamenejme ještě, že hodnoty výšek, které nás zajímaly (tj. 136 cm, 118 cm atd.) se pohybují celkem blízko střední hodnoty $E(X) = 130$ cm, takže předpoklad o normálnosti rozdělení X byl přiměřený.

- -

Výpočty si můžeme i urychlit přímým vzorcem $F_X(t) = \Phi(\frac{t-\mu}{\sigma})$, kde $\mu = 130$ cm a $\sigma = 6$ cm:

(a) $P(X > 136) = 1 - F_X(136) = 1 - \Phi(\frac{136-130}{6}) = \cdots \doteq 0.1587$

(b) $P(X < 118) = F_X(118) = \Phi(\frac{118-130}{6}) = \cdots \doteq 0.0228$

(c) $P(127 < X < 133) = F_X(133) - F_X(127) = \Phi(\frac{133-130}{6}) - \Phi(\frac{127-130}{6}) = \cdots \doteq 0.383$

Příklad 8.2 *Oštěpařky Anna a Barbora mají střední hodnoty hodů po řadě* 67 *m a* 75 *m a směrodatné odchylky* 6 *m a* 3 *m. Předpokládejme nezávislá normální rozdělení. Odhadněte pravděpodobnost, že při jednom hodu hodí Anna dál.*

Řešení:
Náhodná veličina

$A =$“*délka hodu Anny*”

má rozdělení $N(67, 6^2)$ a veličina

$B =$“*délka hodu Barbory*”

má rozdělení $N(75, 3^2)$.

Zajímá nás $P(A > B) = P(A - B > 0)$. Protože veličiny A a B jsou nezávislé, tak veličina $Z := A - B$ má také normální rozdělení, a sice

$$Z \sim N(67 - 75, 6^2 + 3^2) = N(-8, 45)\ .$$

Takže

$$P(A > B) = P(Z > 0) = P\left(\underbrace{\frac{Z-(-8)}{\sqrt{45}}}_{\text{norm}(Z)} > \frac{0-(-8)}{\sqrt{45}}\right) = 1 - P\left(\text{norm}(Z) \leq \frac{8}{\sqrt{45}}\right) =$$

$$= 1 - \Phi\left(\tfrac{8}{\sqrt{45}}\right) \doteq 1 - \Phi(1.1926) \doteq 1 - 0.883 = 0.117\,.$$

POZOR! Zatímco střední hodnota je lineární zobrazení, tak rozptyl se chová jinak! Konkrétně je to takto:

Nechť X a Y jsou veličiny se střední hodnotou a konečným rozptylem. Pak

- $E(X \pm Y) = E(X) \pm E(Y)$
- $\text{var}(X \pm Y) = \text{var}(X) + \text{var}(Y) \pm 2 \cdot \text{cov}(X, Y)\ (\geq 0)$

Zde $\text{cov}(X, Y)$ je tzv. kovariance (viz poznámky níže). Speciálně, pokud X a Y jsou nezávislé, je $\text{cov}(X, Y) = 0$. Máme tedy:

- X a Y nezávislé $\Rightarrow$ $\text{var}(X \pm Y) = \text{var}(X) + \text{var}(Y)$

Tedy v tomto případě se rozptyly VŽDY sčítají!

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.

(Obdobně vznikne náhodný vektor s více složkami. My se teď zaměříme hlavně na dvousložkový případ.)

Náhodný vektor (X, Y) 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

$$P_{(X,Y)}(A) := P\Big(\underbrace{(X,Y)^{-1}(A)}_{\{\omega\in\Omega \mid (X(\omega),Y(\omega))\in A\}} \Big) .$$

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

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

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.

Opět si připomeňme, že 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 $a, b \in \mathbb{R}$.

Příklad 8.3 *Délka hrany krychle je náhodná veličina* $X \sim \mathrm{Ro}(1, 2)$. *Určete distribuční funkci náhodné veličiny* Y *popisující plochu povrchu této krychle.*

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

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

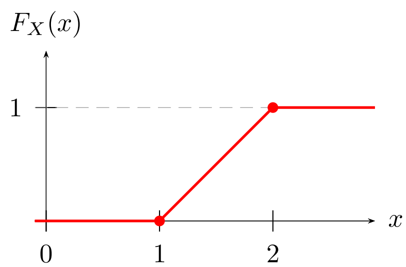
kde jsme využili toho, že obor hodnot pro X je $\langle 1, 2\rangle$, tedy $X \geq 0$ a speciálně tak platí, že $|X| = X$.

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



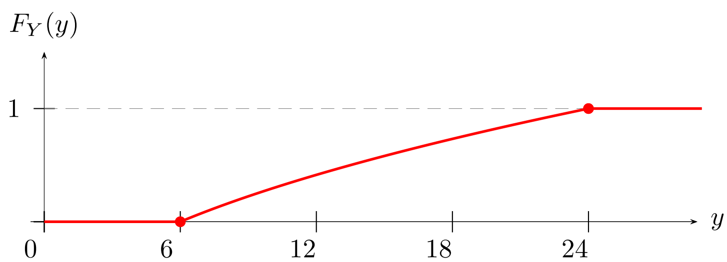
Do F_X (správně!) dosadíme $x = \sqrt{\frac{y}{6}}$ (pro $y \ge 0$) a přepíšeme podmínky pro y:

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

(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



Připomenutí: Kovariance pro X a Y je definována jako

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

Speciálně je $\mathrm{cov}(X,X) = \mathrm{var}(X)$.

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

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

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

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

- 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)
- $\operatorname{cov}(X + c, Y + d) = \operatorname{cov}(X, Y)$.

Platí : X a Y jsou nezávislé veliciny $\Longrightarrow$ $\operatorname{cov}(X, Y) = 0$.

(POZOR: Opačná implikace obecně neplatí!!)

Příklad 8.4 *Nechť* $X \sim \operatorname{Ro}(0, 2)$ *a* $Y = X^2 + 1$.

(a) Sestrojte distribuční funkci náhodné veličiny Y*.*

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

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

Řešení:

(a) Distribuční funkce F_Y se dá získat pomocí distribuční funkce F_X. Před výpočtem si ještě uvědomme, že $X \geq 0$ (protože její obor hodnot je $\langle 0, 2 \rangle$). Speciálně tedy $|X| = X$. Pro distribuční funkci náhodné veličiny Y máme

$$F_Y(y) = P(Y \leq y) = P(X^2 + 1 \leq y) =$$

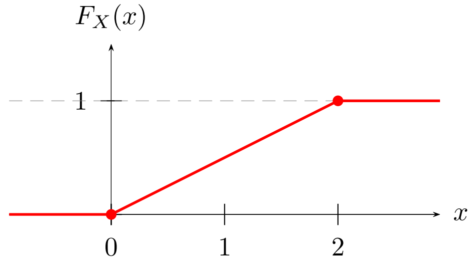
$$= P(X^2 \leq y - 1) = \begin{cases} P(\emptyset) = 0 & , \; y - 1 < 0 \\ P\left(|X| \leq \sqrt{y - 1}\,\right) = F_X\left(\sqrt{y - 1}\,\right) & , \; y - 1 \geq 0. \end{cases}$$

Teď už si stačí jen vyjádřit F_X a dosadit.

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

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

s grafem



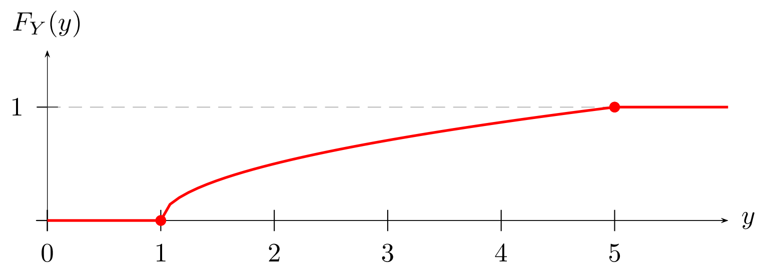
Do F_X (správně!) dosadíme $x = \sqrt{y-1}$ (pro $y - 1 \geq 0$) a přepíšeme podmínky pro y:

$$0 \leq \sqrt{y-1} \leq 2 \quad \Leftrightarrow \quad 0 \leq y-1 \leq 4 \quad \Leftrightarrow \quad 1 \leq y \leq 5$$

(zbylé podmínky jsou podobné) a dostaneme tak

$$F_Y(y) = \begin{cases} 0, & y < 1 \\ \frac{\sqrt{y-1}}{2}, & 1 \leq y \leq 5 \\ 1, & y > 5 \end{cases}$$

s grafem



(b) Kovarianci vypočteme ze vztahu

$$\text{cov}(X,Y) = \text{cov}(X, X^2+1) = \text{cov}(X, X^2) = E(X \cdot X^2) - E(X) \cdot E(X^2)$$

Stačí si tedy pro $n \geq 1$ zjistit

$$E(X^n) = \int_{-\infty}^{\infty} x^n \cdot f_X(x)\ dx = \int_0^2 \frac{x^n}{2}\ dx = \left[\frac{x^{n+1}}{2(n+1)}\right]_0^2 = \frac{2^n}{n+1}$$

a dosazením dostaneme

$$\text{cov}(X,Y) = 2 - 1 \cdot \frac{4}{3} = \frac{2}{3}\ .$$

(c) Protože $\text{cov}(X,Y) \neq 0$, jsou veličiny X a Y závislé. Toto zjištění ovšem můžeme udělat i bez výpočtu kovariance:

Veličiny X a Y jsou funkčně propojené, takže stačí najít podmínky, které naráz nemůžou splnit, ale jednotlivě, s nenulovými pravděpodobnostmi, ano. Z předchozích úprav už víme, že pro $1 < y < 5$ platí

$$Y \leq y \quad \Leftrightarrow \quad X^2 + 1 \leq y \quad \Leftrightarrow \quad X \leq \sqrt{y-1}$$

a pravděpodobnosti těchto jevů jsou (z tvaru F_X a F_Y) ostře mezi 0 a 1. Takže např. z volby $y = 2$ dostaneme, že

$$Y \leq 2 \quad \Leftrightarrow \quad X \leq 1$$

takže

$$P(\underbrace{Y \leq 2,\ X > 1}_{\emptyset}) = 0 \neq \underbrace{P\left(Y \leq 2\right)}_{F_Y(2)=0.5} \cdot \underbrace{P\left(X > 1\right)}_{1-F_X(1)=0.5}$$

z čehož plyne, že veličiny X a Y jsou závislé.

Poznámka: 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 $P\big(h(X) = c\big) = 1$).

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

$$\langle X|Y\rangle := E(X \cdot Y)$$

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

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

Mimo jiné si všimněme, že pro X je $\mathrm{var}(X) = \|X - E(X)\|^2$, takže platí

$$\|norm(X)\| = \left\| \frac{X - EX}{\sqrt{\mathrm{var}(X)}} \right\| = \frac{\|X - EX\|}{\sqrt{\mathrm{var}(X)}} = 1$$

neboli $norm(X)$ má délku skutečně znormovanou na hodnotu 1.

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

$$\mathrm{corr}(X,Y) := \frac{\langle X - EX \mid Y - EY\rangle}{\|X - EX\| \cdot \|Y - EY\|} = \cdots = \frac{E(XY) - E(X) \cdot E(Y)}{\sqrt{\mathrm{var}(X) \cdot \mathrm{var}(Y)}}\ .$$

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

A kromě toho máme:

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

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

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 pro *sdruženou distribuční funkci* máme

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

Věta: Nechť náhodný vektor (X,Y) má *diskrétní* rozdělení. Pak:

X a Y jsou nezávislé $\Leftrightarrow$ $P(X = i, Y = j) = P(X = i) \cdot P(Y = j)$ pro všechna $i, j \in \mathbb{R}$.

(Srovnejte to s obecně NEdiskrétním případem, kdy je nezávislost popsána jen nerovnostmi:

$P(X \le i, Y \le j) = P(X \le i) \cdot P(Y \le j)$ pro všechna $i, j \in \mathbb{R}$.)

Příklad 8.5 *Sdružené pravděpodobnosti náhodných veličin X a Y jsou dány následující tabulkou:*

	$X=0$	$X=1$	$X=2$
$Y=0$	1/4	1/8	0
$Y=1$	1/4	1/4	1/8

(a) *Jaká jsou jejich marginální rozdělení? Určete pravděpodobnost $P(X+Y>1)$.*

(b) *Určete střední hodnotu $E(X \cdot Y)$. Určete varianční a korelační matici.*

(c) *Jsou veličiny X a Y nezávislé? Zdůvodněte.*

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) U veličin X a Y předpokládáme obory hodnot určené tabulkou. Pak máme

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

a tedy

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

Marginální (česky: okrajová) rozdělení náhodného vektoru (X, Y) jsou rozdělení jeho jednotlivých složek, tedy veličin X a Y. Vektor (X, Y) má diskrétní rozdělení a obě veličiny X a Y budou proto mít také diskrétní rozdělení a pro jejich rozdělení platí:

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

Hodnoty pravděpodobností získáme tedy sečtením v řádcích (pro X) a sloupcích (pro Y) naší tabulky:

	$X=0$	$X=1$	$X=2$	$P(Y=j)$
$Y=0$	1/4	1/8	0	3/8
$Y=1$	1/4	1/4	1/8	5/8
$P(X=i)$	1/2	3/8	1/8	

Tedy

$$P(X = i) = \begin{cases} 1/2, & i = 0 \\ 3/8, & i = 1 \\ 1/8, & i = 2 \\ 0, & \text{jinak} \end{cases} \qquad \text{a} \qquad P(Y = j) = \begin{cases} 3/8, & j = 0 \\ 5/8, & j = 1 \\ 0, & \text{jinak.} \end{cases}$$

(b) Pro diskrétní vektor (X, Y) a borelovskou (tj. téměř každou, např. spojitou) funkci $h : \mathbb{R}^2 \to \mathbb{R}$ platí podobná věc jako jsme měli už u náhodné veličiny a sice:

$$E\big(h(X,Y)\big) = \sum_{(i,j)\in\mathbb{R}^2} h(i,j) \cdot P(X = i,\ Y = j)\ .$$

Odsud tedy snadno spočítáme střední hodnotu $E(XY)$:

$$\begin{aligned} E\big(X \cdot Y\big) &= \sum_{(i,j)\in\mathbb{R}^2} i \cdot j \cdot P(X = i,\ Y = j) = \\ &= 0 \cdot 0 \cdot \tfrac{1}{4} + 0 \cdot 1 \cdot \tfrac{1}{8} + 0 \cdot 2 \cdot 0 + 1 \cdot 0 \cdot \tfrac{1}{4} + 1 \cdot 1 \cdot \tfrac{1}{4} + 1 \cdot 2 \cdot \tfrac{1}{8} = \tfrac{1}{2} \end{aligned}$$

Kovarianci pak vypočteme ze vztahu

$$\mathrm{cov}(X,Y) = E(XY) - E(X) \cdot E(Y)$$

K tomu potřebujeme znát také

$$\begin{aligned} E(X) &= 0 \cdot \tfrac{1}{2} + 1 \cdot \tfrac{3}{8} + 2 \cdot \tfrac{1}{8} = \tfrac{5}{8} \\ E(Y) &= 0 \cdot \tfrac{3}{8} + 1 \cdot \tfrac{5}{8} = \tfrac{5}{8} \end{aligned}$$

Takže

$$\mathrm{cov}(X,Y) = E(XY) - E(X) \cdot E(Y) = \tfrac{1}{2} - \tfrac{5}{8} \cdot \tfrac{5}{8} = \tfrac{7}{64}.$$

Pro korelaci

$$\mathrm{corr}(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{var}(X)}\sqrt{\mathrm{var}(Y)}}$$

potřebujeme ještě znát rozptyly, takže si je dopočítáme:

$$\begin{aligned} E(X^2) &= 0^2 \cdot \tfrac{1}{2} + 1^2 \cdot \tfrac{3}{8} + 2^2 \cdot \tfrac{1}{8} = \tfrac{7}{8}, \\ E(Y^2) &= 0^2 \cdot \tfrac{3}{8} + 1^2 \cdot \tfrac{5}{8} = \tfrac{5}{8}, \\ \mathrm{var}(X) &= E(X^2) - \Big(E(X)\Big)^2 = \tfrac{7}{8} - (\tfrac{5}{8})^2 = \tfrac{31}{64}, \\ \mathrm{var}(Y) &= E(Y^2) - \Big(E(Y)\Big)^2 = \tfrac{5}{8} - (\tfrac{5}{8})^2 = \tfrac{15}{64}. \end{aligned}$$

Všimněme si ještě, že $Y \sim \mathrm{Alt}(\frac{5}{8})$, takže rozptyl jsme mohli spočítat jako $\mathrm{var}(Y) = \frac{5}{8} \cdot (1 - \frac{5}{8}) = \frac{15}{64}$. Korelace tedy je

$$\mathrm{corr}(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{var}(X)} \cdot \sqrt{\mathrm{var}(Y)}} = \frac{\frac{7}{64}}{\sqrt{\frac{31}{64}} \cdot \sqrt{\frac{15}{64}}} = \frac{7}{\sqrt{465}} \doteq 0.32462,$$

Varianční matice je tudíž

$$\mathrm{Cov}(X,Y) = \begin{pmatrix} \mathrm{cov}(X,X) & \mathrm{cov}(X,Y) \\ \mathrm{cov}(Y,X) & \mathrm{cov}(Y,Y) \end{pmatrix} = \begin{pmatrix} \mathrm{var}(X) & \mathrm{cov}(X,Y) \\ \mathrm{cov}(X,Y) & \mathrm{var}(Y) \end{pmatrix} = \begin{pmatrix} \frac{31}{64} & \frac{7}{64} \\ \frac{7}{64} & \frac{15}{64} \end{pmatrix}$$

a korelační matice je

$$\mathrm{Corr}(X,Y) = \begin{pmatrix} \mathrm{corr}(X,X) & \mathrm{corr}(X,Y) \\ \mathrm{corr}(Y,X) & \mathrm{corr}(Y,Y) \end{pmatrix} = \begin{pmatrix} 1 & \mathrm{corr}(X,Y) \\ \mathrm{corr}(X,Y) & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{7}{\sqrt{465}} \\ \frac{7}{\sqrt{465}} & 1 \end{pmatrix}.$$

(c) Protože nyní je např.

$$P(X = 2, Y = 0) = 0 \neq \tfrac{1}{8} \cdot \tfrac{3}{8} = P(X = 2) \cdot P(Y = 0).$$

jsou X a Y **závislé**.

Důvodem pro závislost X a Y je také fakt, že $\mathrm{cov}(X, Y) \neq 0$. Zdůrazněme ale, že pokud je kovariance nulová, nemůžeme (pouze na základě její znalosti) o nezávislosti obecně nic říct!

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).

Příklad 8.6 *Sdružená hustota náhodných veličin X a Y je*

$$f_{(X,Y)}(x, y) = \begin{cases} \frac{1}{2} e^{-x-\frac{y}{2}}\,, & x > 0, y > 0\,, \\ 0\,, & \textit{jinak}\,. \end{cases}$$

(a) Jaká jsou jejich marginální rozdělení?

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

(c) Určete varianční a korelační matici.

Řešení:

(a) 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 $Y \sim \mathrm{Exp}(\frac{1}{2})$.

(b) 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.

(c) Z nezávislosti X, Y plyne okamžitě $\mathrm{cov}(X,Y) = 0$, tedy také $\mathrm{corr}(X,Y) = 0$. Z $X \sim \mathrm{Exp}(1)$ a $Y \sim \mathrm{Exp}(\frac{1}{2})$ máme, že $\mathrm{var}(X) = 1^2 = 1$ a $\mathrm{var}(Y) = 2^2 = 4$. Tedy kovarianční matice

$$\mathrm{Cov}(X,Y) = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}.$$

je a korelační matice je tak

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