4. cvičení z PST

24. října 2018

Příklad **3.10**.

4.1 (náhodná veličina)
Zjistěte, zda $X : \Omega \to \mathbb{R}$ je náhodná veličina, pokud $(\Omega, \mathcal{A}, P)$ je

- $\Omega = \{a, b, c\}$,
- $\mathcal{A} = \big\{\emptyset, \{a\}, \{b, c\}, \{a, b, c\}\big\}$,
- $P(A) = \frac{1}{3}|A|$, $A \in \mathcal{A}$.

a platí-li, že

$X(a) = 1,\ X(b) = 2,\ X(c) = 3\ .$

Řešení:
Nejdříve bychom měli zkontrolovat, jestli máme opravdu Kolmogorův model:

Systém $\mathcal{A}$ je zřejmě uzavřen na sjednocení i na doplňky. Tedy $\mathcal{A}$ je σ-algebra, která je navíc podalgebrou $\exp(\{a, b, c\})$. Z tohoto důvodu budou splněny i požadavky na pravděpodobnost, protože ta má stejný předpis jako v jednom z předchozích příkladu, kde už to, jak víme, funguje. Tím spíše to musí platit i pro menší systém množin. Tedy $(\Omega, \mathcal{A}, P)$ je opravdu Kolmogorův model.

Definice: Náhodná veličina je takové zobrazení, že vzor každého intervalu v $\mathbb{R}$ je množina z $\mathcal{A}$. Tuto vlastnost stačí ověřit jen pro určité typy intervalů v $\mathbb{R}$:

$$X \text{ je náhodná veličina} \quad \Leftrightarrow \quad (\forall t \in \mathbb{R}) \ \ X^{-1}\Big((-\infty, t\rangle\Big) \in \mathcal{A}$$

Podíváme se, jak vypadají vzory všech potřebných intervalů:

$$X^{-1}\Big((-\infty, t\rangle\Big) = \{\omega \in \Omega \mid X(\omega) \leq t\} = \begin{cases} \emptyset & ,\ t \in (-\infty, 1) \\ \{a\} & ,\ t \in \langle 1, 2) \\ \{a, b\} & ,\ t \in \langle 2, 3) \\ \{a, b, c\} & ,\ t \in \langle 3, +\infty) \end{cases}$$

X tedy není náhodná veličina, protože např.

$$X^{-1}\Big((-\infty, 2.5\rangle\Big) = \{\omega \in \Omega \mid X(\omega) \leq 2.5\} = \{a, b\} \notin \mathcal{A}\ .$$

Problematické jsou hodnoty $t \in \langle 2, 3)$. K této situaci došlo proto, že zobrazení X oddělilo svými hodnotami prvky b a c, které jsou v rámci σ-algebry $\mathcal{A}$ nerozlišitelné (není už žádná menší množina z $\mathcal{A}$, která by obsahovala b a neobsahovala c nebo naopak).

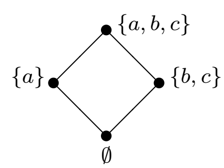
Jak tedy zvolit nějakou jinou (nekonstantní) náhodnou veličinu Y na Ω? Stačí např. položit

$$Y(a) = 1,\ Y(b) = Y(c) = 2\ .$$

(Problémového intervalu jsme se zbavili tak, že všem prvkům z množiny $\{\omega \in \Omega \mid X(\omega) \in \langle 2, 3)\}$ jsme "nastavili"nějakou stejnou společnou hodnotu.)

Můžeme si ještě pro názornost zakreslit Hasseův diagram pro $\mathcal{A}$:

$\mathcal{A}$:



4.2 (diskrétní veličina - binomické rozdělení)

Házíme třikrát mincí a náhodná veličina X má za hodnotu počet líců. Předpokládáme, že jsou hody nezávislé a pravděpodobnost, že padne líc v daném hodu, je $p = 0.1$. Určete:

(a) Typ rozdělení veličiny X a její obor hodnot.

(b) Pravděpodobnostní funkci p_X náhodné veličiny X.

(c) Distribuční funkci F_X náhodné veličiny X.

(d) Střední hodnotu $E(X)$ náhodné veličiny X.

Řešení:
Jde o diskrétní veličinu s binomickým rozdělením $\mathrm{Bi}(n, p)$, kde $n = 3$ a $p = 0.1$. Oborem hodnot je množina $\{0, 1, 2, 3\}$. Pravděpodobnostní funkce je

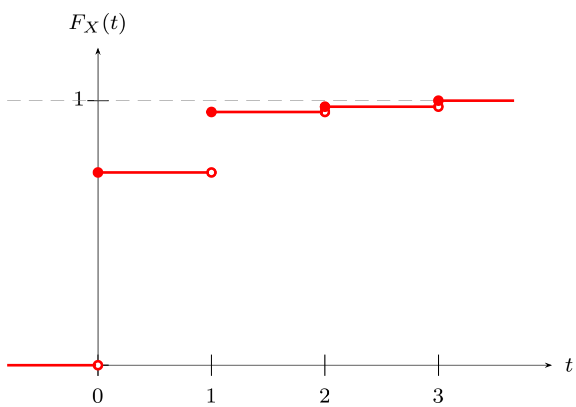
$$p_X(k) = P(X = k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & , k \in \{0, 1, 2, 3\} \\ 0 & , \text{jinak.} \end{cases}$$

Konkrétně:

k	0	1	2	3
$p_X(k)$	$(1-p)^3$	$3p(1-p)^2$	$3p^2(1-p)$	p^3
$p = 0.1$	0.729	0.243	0.027	0.001

Distribuční funkce $F_X(t) = P(X \leq t) = \sum_{k \leq t} p_X(k)$ pro diskrétní veličinu X má hodnoty

$$F_X(t) = \begin{cases} 0 & , t < 0 \\ p_X(0) = 0.729 & , t \in \langle 0, 1) \\ p_X(0) + p_X(1) = 0.972 & , t \in \langle 1, 2) \\ p_X(0) + p_X(1) + p_X(2) = 0.999 & , t \in \langle 2, 3) \\ 1 & , t \geq 3. \end{cases}$$



Střední hodnota pro diskrétní veličinu je

$$E(X) = \sum_{k \in \mathbb{R}} k \cdot p_X(k) = 0 \cdot 0.729 + 1 \cdot 0.243 + 2 \cdot 0.027 + 3 \cdot 0.001 = 0.3$$

nebo podle vzorce pro binomické rozdělení $E(X) = np = 3 \cdot 0.1 = 0.3$.

4.3 (diskrétní veličina - hypergeometrické rozdělení)

V osudí máme 3 bílé, 4 modré a 2 červené koule. Náhodně vytáhneme 4 z nich. Náhodná veličina X je určena počtem bílých koulí. Určete:

(a) Typ rozdělení veličiny X a její obor hodnot.

(b) Pravděpodobnostní funkci p_X náhodné veličiny X.

(c) Distribuční funkci F_X náhodné veličiny X.

(d) Střední hodnotu $E(X)$ náhodné veličiny X.

Řešení:
Jde o diskrétní veličinu s oborem hodnot $\{0,1,2,3\}$. Pravděpodobnostní funkce je

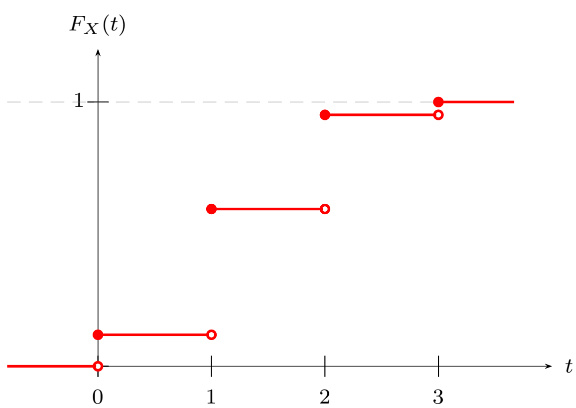
$$p_X(k) = P(X=k) = \begin{cases} \frac{\binom{3}{k}\cdot\binom{6}{4-k}}{\binom{9}{4}} & , k \in \{0,1,2,3\} \\ 0 & , \text{jinak.} \end{cases} .$$

Konkrétně:

k	0	1	2	3
$p_X(k)$	$\frac{\binom{3}{0}\cdot\binom{6}{4}}{\binom{9}{4}} = \frac{5}{42} \doteq 0.1190$	$\frac{\binom{3}{1}\cdot\binom{6}{3}}{\binom{9}{4}} = \frac{10}{21} \doteq 0.4762$	$\frac{\binom{3}{2}\cdot\binom{6}{2}}{\binom{9}{4}} = \frac{5}{14} \doteq 0.3571$	$\frac{\binom{3}{3}\cdot\binom{6}{1}}{\binom{9}{4}} = \frac{1}{21} \doteq 0.0476$

Distribuční funkce $F_X(t) = P(X \le t) = \sum_{k \le t} p_X(k)$ pro diskrétní veličinu X má hodnoty

$$F_X(t) = \begin{cases} 0 & , t < 0 \\ p_X(0) = \frac{5}{42} \doteq 0.1190 & , t \in \langle 0,1) \\ p_X(0) + p_X(1) = \frac{25}{42} \doteq 0.5952 & , t \in \langle 1,2) \\ p_X(0) + p_X(1) + p_X(2) = \frac{20}{21} \doteq 0.9524 & , t \in \langle 2,3) \\ 1 & , t \ge 3. \end{cases}$$



Střední hodnota pro diskrétní veličinu je

$$E(X) = \sum_{k\in\mathbb{R}} k \cdot p_X(k) = 0 \cdot \frac{5}{42} + 1 \cdot \frac{10}{21} + 2 \cdot \frac{5}{14} + 3 \cdot \frac{1}{21} = \frac{4}{3} \doteq 1.333 \ .$$

Střední hodnotu také můžeme spočítat na základě hypergeometrického rozdělení (viz poznámka níže).

Poznámka: Náhodná veličiny X v tomto příkladu má tzv. *hypergeometrické* rozdělení $\mathrm{Hyp}(N, K, n)$, kde N je počet prvků dané množiny (zde: počet koulí), která obsahuje K prvků dané vlastnosti (zde: bílé koule), a z níž vybíráme n prvků (kde $n \le N$). Zajímá nás, jaká je pravděpodobnost, že z vybraných n prvků je právě k prvků uvedené vlastnosti.

V tomto případě je $N = 9$, $K = 3$, $n = 4$.

Pravděpodobnost je dána podílem příznivých možností ku všem. Příznivé jsou dány počtem způsobů jak vybrat k koulí z K bílých koulí (které fyzicky rozlišujeme) násobeno počtem způsobů jak vybrat zbytek, tj. $n-k$ nebílých koulí z $N-K$ nebílých (které také fyzicky rozlišujeme). Celkem tedy

$$p_{\mathrm{Hyp}(N,K,n)}(k) = \frac{\binom{K}{k}\cdot\binom{N-K}{n-k}}{\binom{N}{n}}$$

přičemž obor hodnot pro k je

$$\max\{0, n+K-N\} \leq k \leq \min\{n, K\}$$

tedy

$$0 = \max\{0, 4+3-9\} \leq k \leq \min\{4, 3\} = 3\ .$$

Tuto podmínku dostaneme ihned z nerovností

$$k \leq K, \quad n-k \leq N-K,$$

$$k \leq n, \quad n-k \leq n\ .$$

Můžeme si ještě pro zajímavost spočítat střední hodnotu naší veličiny X s rozdělením $\mathrm{Hyp}(N=9, K=3, n=4)$ způsobem, který není těžké zobecnit pro libovolné hypergeometrické rozdělení:

$$E(X) = \sum_{k=0}^{3} k \cdot \underbrace{\frac{\binom{3}{k}\cdot\binom{6}{4-k}}{\binom{9}{4}}}_{=p_{\mathrm{Hyp}(9,3,4)}(k)} = \sum_{k=1}^{3} \frac{k \cdot \frac{3!}{k!(3-k)!} \cdot \binom{6}{4-k}}{\binom{9}{4}} = \sum_{k=1}^{3} \frac{3 \cdot \frac{2!}{(k-1)!(2-(k-1))!} \cdot \binom{6}{4-k}}{\binom{9}{4}} =$$

$$= \frac{3\cdot\binom{8}{3}}{\binom{9}{4}} \sum_{k=1}^{3} \frac{\binom{2}{k-1}\cdot\binom{6}{3-(k-1)}}{\binom{8}{3}} = \frac{3\cdot\binom{8}{3}}{\binom{9}{4}} \sum_{\ell=0}^{2} \underbrace{\frac{\binom{2}{\ell}\cdot\binom{6}{3-\ell}}{\binom{8}{3}}}_{=p_{\mathrm{Hyp}(8,2,3)}(\ell)} = \frac{3\cdot\binom{8}{3}}{\binom{9}{4}} \underbrace{\sum_{\ell=0}^{2} p_{\mathrm{Hyp}(8,2,3)}(\ell)}_{=1} = 3 \cdot \frac{\frac{8!}{3!5!}}{\frac{9!}{4!5!}} = 4 \cdot \frac{3}{9} = n \cdot \frac{K}{N}\ .$$

Limitně pro $N \to \infty$ a $K/N \to q$ a n pevně zvolené (tj. když poměr počtu bílých koulí K ku počtu všech koulí N se blíží ke konstantní hodnotě $q \in (0,1)$) se hypergeometrické rozdělení blíží binomickému rozdělení $\mathrm{Bi}(n,q)$. Toto binomické rozdělení odpovídá limitní situaci, kdy taháme koule z nekonečného množství s určeným podílem q bílých. Při n pokusech jsme předpokládali, ze koule NEVRACÍME (anebo je prostě vytáhneme všechny naráz). Jak ale vidíme, i přesto jsme dostali rozdělení pravděpodobnosti, které používáme, když opakovaně uskutečňujeme tentýž pokus vždy za STEJNÝCH podmínek (tj. opakovaně vytahujeme z osudí koule a VRACÍME je vždy zpátky). To je tím, že v obrovském množství koulí N se už v limitě ztratí to, jestli tam těch pár koulí n vrátíme nebo ne (protože $n << N$).

4.4 (spojitá náhodná veličina)

Házíme kuličku (o zanedbatelném průměru) na kruh o poloměru $r = 1$. Náhodná veličina X je vzdálenost dopadu kuličky od středu kruhu. Určete:

(a) Obor hodnot náhodné veličiny X.

(b) Distribuční funkci F_X a typ rozdělení náhodné veličiny X.

(c) Hustotu pravděpodobnosti f_X.

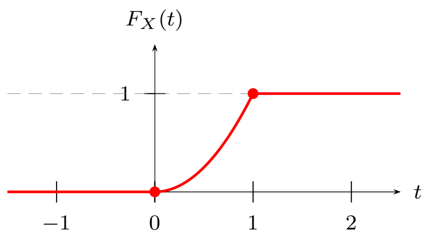
(d) Střední hodnotu $E(X)$.

Řešení:

Obor hodnot je interval $\langle 0, 1\rangle$. Množina všech možných výsledků Ω (tj. jevové pole) je kruh o poloměru 1. Předpokládáme Kolmogorův model daný geometrickou pravděpodobností.

Distribuční funkce je

$$F_X(t) = P(X \leq t) = \begin{cases} 0 & , t < 0 \\ \frac{\text{“obsah kruhu o poloměru } t\text{”}}{\text{“obsah kruhu o poloměru 1”}} = \frac{\pi t^2}{\pi 1^2} = t^2 & , t \in \langle 0, 1\rangle \\ 1 & , t > 1. \end{cases}$$



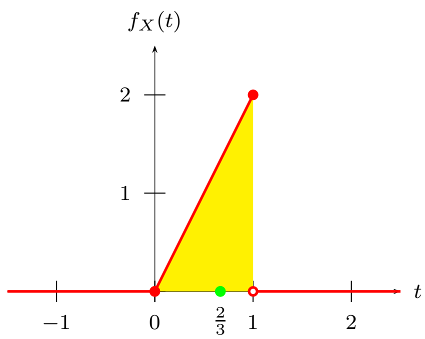
Veličina X je spojitá, protože distribuční funkce F_X je spojitá. Hustotu pravděpodobnosti (tj. funkci $f_X : \mathbb{R} \to \langle 0, +\infty)$, pro kterou je $F_X(t) = \int_{-\infty}^{t} f_X(u)\ \mathrm{d}u$ pro každé $t \in \mathbb{R}$) můžeme určit jako derivaci funkce F_X tam, kde derivace existuje (ve zbylých bodech si hustotu můžeme definovat libovolně nezáporně). Protože

$$\frac{\mathrm{d}}{\mathrm{d}t} F_X(t) = \begin{cases} 2t & , t \in (0,1) \\ 0 & , t < 0 \text{ nebo } t > 1 \ . \end{cases}$$

je funkce

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

hustotou pravděpodobnosti.



Střední hodnotu pro spojitou náhodnou veličinu s hustotou pravděpodobnosti můžeme spočítat jako

$$E(X) = \int_{-\infty}^{\infty} t \cdot f_X(t)\ \mathrm{d}t = \int_0^1 t \cdot 2t\ \mathrm{d}t = \tfrac{2}{3}\ .$$

Názornou interpretací střední hodnoty spojité veličiny X, která má hustotu, je, že $E(X)$ je vodorovná souřadnice (zelený bod) těžiště plochy, která je určena grafem hustoty f_X (žlutá plocha). V tomto případě speciálně musí platit, že $0 < E(X) < 1$, což je evidentně splněno. V případě, že by vypočtené hodnota $E(X)$ tyto nerovnosti nesplňovala, znamenalo by to, že jsme někde ve výpočtu udělali chybu.

4.5 (spojitá náhodná veličina - rovnoměrné rozdělení)

Nejmenší dílek na (analogových) stopkách má velikost 0.2 s. Čas se při určování zaokrouhluje tak, že k pozici ručičky se zvolí její nejbližší hodnota na stupnici. Nechť X představuje náhodnou veličinu danou rozdílem skutečného času a jeho zaokrouhlení. Určete:

(a) Obor hodnot náhodné veličiny X.

(b) Distribuční funkci F_X a typ rozdělení náhodné veličiny X.

(c) Hustotu pravděpodobnosti f_X.

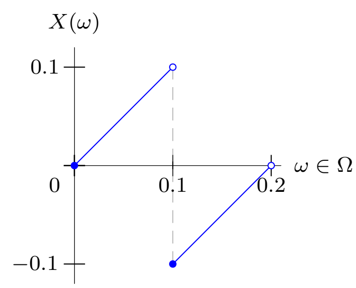
(d) Pravděpodobnost, že odchylka skutečného času od naměřené hodnoty bude menší než 0.05 s.

Řešení:

Obor hodnot náhodné veličiny X je interval $\langle -0.1,\ 0.1)$ v jednotkách "sekunda". Za množinu všech možných výsledků Ω (tj. jevové pole) si můžeme zvolit např. interval $\langle 0, 60)$ (v sekundách), ale z hlediska výpočtu bude jednodušší si zvolit $\Omega = \langle 0, 0.2)$. Předpokládáme Kolmogorův model daný geometrickou pravděpodobností na Ω. Pro $\omega \in \Omega$ je

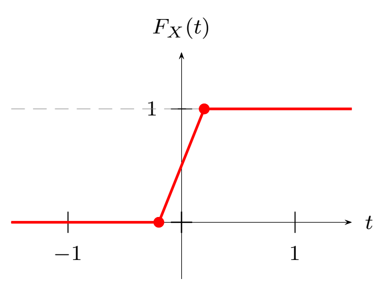
$$X(\omega) = \omega - \text{"nejbližší dílek k } \omega \text{ na stupnici"}.$$

Z grafu funkce $X : \Omega = \langle 0, 0.2) \to \mathbb{R}$



snadno odvodíme, že distribuční funkce je

$$F_X(t) = P(X \le t) = \begin{cases} 0 & , t < -0.1 \\ \frac{\text{"délka množiny } \{\omega \in \Omega \mid X(\omega) \le t\}\text{"}}{\text{"délka intervalu } \langle 0, 0.2)\text{"}} = \frac{t+0.1}{0.2} & , t \in \langle -0.1,\ 0.1\rangle \\ 1 & , t > 0.1. \end{cases}$$



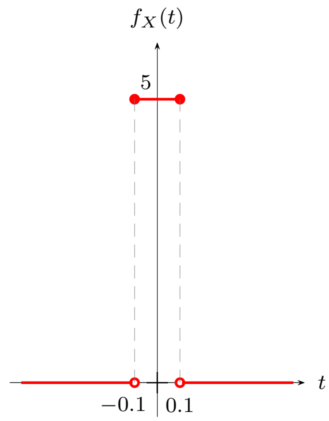
Veličina X je spojitá, protože distribuční funkce F_X je spojitá. Hustotu pravděpodobnosti (tj. funkci $f_X : \mathbb{R} \to \langle 0, +\infty)$, pro kterou je $F_X(t) = \int_{-\infty}^{t} f_X(u)\ du$ pro každé $t \in \mathbb{R}$) můžeme určit jako derivaci F_X tam, kde derivace existuje (ve zbylých bodech si hustotu můžeme definovat libovolně nezáporně). Protože

$$\frac{\mathrm{d}}{\mathrm{d}t} F_X(t) = \begin{cases} \frac{1}{0.2} & , t \in (-0.1,\ 0.1) \\ 0 & , t < -0.1 \text{ nebo } t > 0.1\ . \end{cases}$$

je funkce

$$f_X(t) = \begin{cases} \frac{1}{0.2} = 5 & , t \in \langle -0.1,\ 0.1\rangle \\ 0 & , \text{jinak}\ . \end{cases}$$

hustotou pravděpodobnosti.



Veličina X má tedy rovnoměrné rozdělení pravděpodobnosti na intervalu $\langle -0.1,\ 0.1\rangle$.
Pravděpodobnost, že odchylka skutečného času od naměřené hodnoty bude menší než 0.05 sekund, je

$$P(|X| < 0.05) = P(-0.05 < X < 0.05) = \int_{-0.05}^{0.05} f_X(t)\ \mathrm{d}t = \int_{-0.05}^{0.05} 5\ \mathrm{d}t = 0.5\ .$$