4. cvičení z PST

14. října 2020

Důležitá poznámka: Náhodná veličina je funkce, která náhodnému výsledku (tj. elementárnímu jevu) přiřadí konkrétní hodnotu. Např. vybranému člověku přiřadí jeho tělesnou výšku. Jak je vidět, náhodná veličina vůbec není náhodná, co do hodnoty, kterou přiřazuje (ta je určena naprosto jasně). Náhodnost výstupu veličiny je dána náhodností jejího vstupu!

Abychom mohli s náhodnou veličinou X vůbec pracovat, potřebujeme umět určovat pravděpodobnosti toho, že hodnoty X budou např. v intervalu $\langle 1.5,\ 7.2) \subseteq \mathbb{R}$, což znamená, že množině $X^{-1}\langle 1.5,\ 7.2) = \{\omega \in \Omega \mid 1.5 \leq X(\omega) < 7.2\}$ musíme umět přiřadit její pravděpodobnost. To půjde jen tehdy, jestliže to bude množina měřitelná v daném systému jevů. Proto následující definice:

Definice: Náhodná veličina X v Kolmogorově modelu $(\Omega, \mathcal{A}, P)$ je takové zobrazení

$$X : \Omega \to \mathbb{R}$$

že vzor každého intervalu I v $\mathbb{R}$ je "přípustná" množina (neboli jev), tj. $X^{-1}(I) \in \mathcal{A}$. Tuto vlastnost stačí ověřit jen pro určité typy intervalů v $\mathbb{R}$:

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

4.1 (náhodná veličina v Kolmogorově modelu)
Zjistěte, zda $X : \Omega \to \mathbb{R}$ je náhodná veličina, pokud Kolmogorův model $(\Omega, \mathcal{A}, P)$ 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 $X(a) = 1$, $X(b) = 2$, $X(c) = 3$, a pokud není, navrhněte jiný (nekonstantní) předpis $Y : \Omega \to \mathbb{R}$, který veličinou bude.

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.

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).

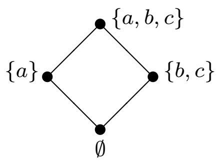
Jak tedy zvolit nějakou jinou (nekonstantní) náhodnou veličinu Y na Ω? Stačí např. položit

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

(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 $\mathcal{A}$:

$\mathcal{A}$:



Poznámka: Jak by mohla vypadat nějaká konkrétní fyzická realizace výše uvedené Kolmogorova modelu (zejména nás zajímá realizace výsledků Ω a jevů, tj. σ-algebry $\mathcal{A}$). Potřebujeme tudíž mít 3 výsledky a dva z nich nerozlišitelné:

- můžeme např. házetválečkem, který bude dopadat buď na hranu nebo na jednu z nerozlišitelných podstav (dlouhý válec bude padat spíše na hranu a plochý válec zase spíše na podstavy)
- nebo můžeme uvažovat urnu, která obsahuje koule dvou barev a kostky jedné barvy. Předměty pak vytahujeme potmě, takže umíme rozlišit jejich tvar ale ne barvu.

Důležitá poznámka (možná i nejdůležitější vůbec....): Co umožňuje náhodná veličina:

Náhodná veličina $X : \Omega \to \mathbb{R}$ umožňuje jednu podstatnou věc - zapomenout na původní prostor Ω všech výsledků (i s celou jeho strukturou jevů a jejich pravděpodobnostmi) a přitom stále umět vyčíslovat pravděpodobnosti pro X, které potřebujeme znát.

Veličina X totiž umí přenést a vytvořit rozdělení pravděpodobnosti tam, kde má své hodnoty - konkrétně vytvoří nový Kolmogorův model na reálné přímce $\mathbb{R}$. A to tak, že každý interval $I \subseteq \mathbb{R}$ bude mít prostě pravděpodobnost $P_X(I) := P\Big(X^{-1}(I)\Big)$.

Pravděpodobnosti P_X se říká rozdělení pravděpodobnosti veličiny X (na $\mathbb{R}$). Toto rozdělení můžeme úplně popsat pokud opět známe pravděpodobnosti jen některých speciálních intervalů: opět jsou to intervaly $(-\infty, t\rangle$ pro $t \in \mathbb{R}$. Jejich pravděpodobnosti pak přirozeně definují tzv. *distribuční funkci* $F_X : \mathbb{R} \to \langle 0, 1\rangle$ jako

$$F_X(t) := P(X \le t) \qquad \Big(= P\big(X^{-1}(-\infty, t\rangle\big)\Big)$$

Toto je jeden z nejdůležitějších pojmů v celém kurzu! Proto se jej snažte pochopit.

Definice: Veličina $X : \Omega \to \mathbb{R}$ má tzv. *diskrétní rozdělení* $\Leftrightarrow$ existuje **nejvýše spočetná** množina hodnot $A \subseteq \mathbb{R}$ taková, že $P(X \in A) = 1$.

(Veličina může nabývat i jiných hodnot, ale pak se tyto hodnoty objevují s nulovou pravděpodobností. To ovšem neznamená, že by nemohly nastat!)

4.2 (diskrétní veličina - hypergeometrické rozdělení)

V balíčku je 9 semen, z toho 3 už nevyklíčí. Náhodně vytáhneme 4 z nich a zasadíme. Náhodná veličina X je počet rostlin, které vyrostou. Určete:

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

(b) pravděpodobnostní funkci p_X a distribuční funkci F_X náhodné veličiny X.

(c) kolik průměrně vyroste rostlin.

(d) rozptyl $D(X)$ veličiny X.

(e) pravděpodobnost, že budeme mít alespoň dvě rostliny.

Řešení:
Máme k dispozici 6 dobrých a 3 špatná semena. Z nich vybíráme 4. Možné hodnoty X jsou tak $\{1,2,3,4\}$. Protože máme jen konečně mnoho hodnot, jde o veličinu s diskrétním rozdělením. Pravděpodobnostní funkce je

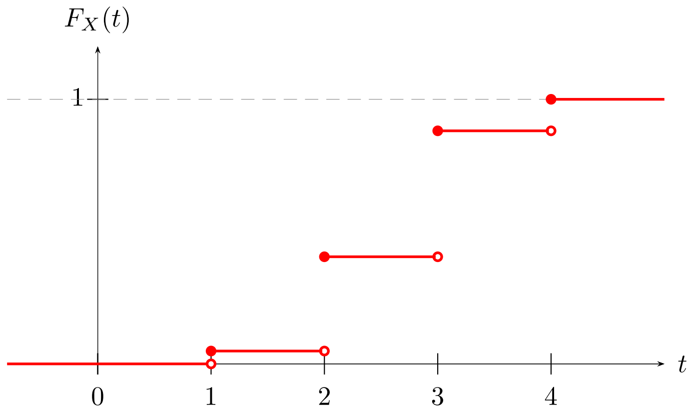
$$p_X(k) = P(X=k) = \begin{cases} \frac{\binom{6}{k}\cdot\binom{3}{4-k}}{\binom{9}{4}} & , k \in \{1,2,3,4\} \\ 0 & , \text{jinak.} \end{cases} .$$

Konkrétně:

k	1	2	3	4
$p_X(k)$	$\frac{\binom{6}{1}\cdot\binom{3}{3}}{\binom{9}{4}} = \frac{2}{42} \doteq 0.048$	$\frac{\binom{6}{2}\cdot\binom{3}{2}}{\binom{9}{4}} = \frac{15}{42} \doteq 0.357$	$\frac{\binom{6}{3}\cdot\binom{3}{1}}{\binom{9}{4}} = \frac{20}{42} \doteq 0.476$	$\frac{\binom{6}{4}\cdot\binom{3}{0}}{\binom{9}{4}} = \frac{5}{42} \doteq 0.119$

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

$$F_X(t) = \begin{cases} 0 & , t < 1 \\ p_X(0) = \frac{2}{42} \doteq 0.048 & , t \in \langle 1, 2) \\ p_X(0) + p_X(1) = \frac{17}{42} \doteq 0.405 & , t \in \langle 2, 3) \\ p_X(0) + p_X(1) + p_X(2) = \frac{37}{42} \doteq 0.881 & , t \in \langle 3, 4) \\ 1 & , t \ge 4. \end{cases}$$



Průměrný počet rostlin je střední hodnota naší diskrétní veličin X, což je

$$E(X) = \sum_{k \in \mathbb{R}} k \cdot p_X(k) = 1 \cdot \frac{2}{42} + 2 \cdot \frac{15}{42} + 3 \cdot \frac{20}{42} + 4 \cdot \frac{5}{42} = \frac{8}{3} \doteq 2.67 .$$

Střední hodnotu také můžeme spočítat na základě hypergeometrického rozdělení $\mathrm{Hyp}(n; K, N)$ jako

$$E(X) = n \cdot \frac{K}{N} = 4 \cdot \frac{6}{9} = \frac{8}{3} .$$

Rozptyl náhodné veličiny určuje, jak moc se hodnoty odchylují od střední hodnoty. Rozptyl je proto definován jako střední hodnota veličiny $\big(X - E(X)\big)^2$, která vyjadřuje kvadrát odchylky X od své střední hodnoty, tedy předpisem:

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

K tomu tedy potřebujeme ještě určit $E\big(X^2\big)$, což uděláme pomocí vzorce

$$E\big(h(X)\big) = \sum_{k\in\mathbb{R}} h(k)\cdot p_X(k),$$

kde $h:\mathbb{R}\to\mathbb{R}$ je (borelovsky) měřitelná funkce, např. spojitá. Takže máme

$$E(X^2) = \sum_{k\in\mathbb{R}} k^2\cdot p_X(k) = 1^2\cdot\tfrac{2}{42} + 2^2\cdot\tfrac{15}{42} + 3^2\cdot\tfrac{20}{42} + 4^2\cdot\tfrac{5}{42} = \tfrac{23}{3}\ .$$

Tedy

$$D(X) = E\big(X^2\big) - \Big(E\big(X\big)\Big)^2 = \tfrac{23}{3} - (\tfrac{8}{3})^2 = \tfrac{5}{9} \doteq 0.556\ .$$

Rozptyl také můžeme spočítat podle vzorce pro hypergeometrického rozdělení $\mathrm{Hyp}(n;K,N)$ jako

$$D(X) = n\cdot\frac{K}{N}\cdot\left(1-\frac{K}{N}\right)\cdot\frac{N-n}{N-1} = 4\cdot\frac{6}{9}\left(1-\frac{6}{9}\right)\cdot\frac{9-4}{9-1} = \frac{5}{9}\ .$$

A konečně, pravděpodobnost jevu

$$A = \text{“\textit{vybereme alespoň dvě dobrá semena}”}$$

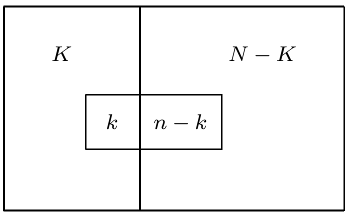
je

$$P(A) = P(X\geq 2) = 1 - P(X<2) = 1 - P(X=1) = 1 - \frac{1}{21} = \frac{20}{21} \doteq 0.952\ .$$

Poznámka: Náhodná veličiny X v tomto příkladu má tzv. *hypergeometrické* rozdělení

$$X \sim \mathrm{Hyp}(n;K,N)$$

kde N je počet prvků dané množiny (zde: počet semen), která obsahuje K prvků dané vlastnosti (zde: dobrá semena), a z níž vybíráme n prvků (kde $n\leq N$). Zajímá nás, jaká je pravděpodobnost, že z vybraných n prvků je právě k prvků uvedené vlastnosti.



V tomto případě je $n=4$, $K=6$ a $N=9$.

Pravděpodobnost je dána podílem příznivých možností ku všem. Příznivé jsou dány počtem způsobů jak vybrat k semen z K dobrých semen (které fyzicky rozlišujeme) násobeno počtem způsobů jak vybrat zbytek, tj. $n-k$ semen z $N-K$ špatných (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

$$1 = \max\{0, 4+6-9\} \leq k \leq \min\{4,6\} = 4\ .$$

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 $\mathrm{Hyp}(4;6,9)$ způsobem, který není těžké zobecnit pro libovolné hypergeometrické rozdělení:

$$E(X) = \sum_{k=1}^{4} k \cdot \underbrace{\frac{\binom{6}{k}\cdot\binom{3}{4-k}}{\binom{9}{4}}}_{=p_{\mathrm{Hyp}(4;6,9)}(k)} = \sum_{k=1}^{4} \frac{k \cdot \frac{6!}{k!(6-k)!}\cdot\binom{3}{4-k}}{\binom{9}{4}} = \sum_{k=1}^{4} \frac{6\cdot\frac{5!}{(k-1)!(5-(k-1))!}\cdot\binom{3}{4-k}}{\binom{9}{4}} =$$

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

$$= 6\cdot\frac{\frac{8!}{3!5!}}{\frac{9!}{4!5!}} = 4\cdot\frac{6}{9} = n\cdot\frac{K}{N}\ .$$

Limitně pro $N\to\infty$ a $K/N\to q$ a n pevně zvolené (tj. když poměr počtu dobrých semen K ku počtu všech semen 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 vybíráme n-krát z nekonečného množství s určeným podílem q dobrých semen.

Hypergeometrické rozdělení odpovídá výběru *bez* vracení, binomické zase výběru *s* vracením. V limitě se ovšem tato vlastnost u hypergeometrického vytratí, což je tím, že v obrovském množství semen N se už ztratí to, jestli tam těch pár semen n (pro $n << N$) vrátíme nebo ne.

Všimněme si ještě, že střední hodnoty vycházejí pro hypergeometrické i binomické rozdělení stejně v tomto smyslu:

$$E\Big(\mathrm{Hyp}(n;K,N)\Big) = n\cdot\frac{K}{N} = E\Big(\mathrm{Bi}(n,\tfrac{K}{N})\Big)$$

Rozptyly se pak už liší:

$$D\Big(\mathrm{Hyp}(n;K,N)\Big) = n\cdot\frac{K}{N}\cdot\Big(1-\frac{K}{N}\Big)\cdot\frac{N-n}{N-1}$$

$$D\Big(\mathrm{Bi}(n,\tfrac{K}{N})\Big) = n\cdot\frac{K}{N}\cdot\Big(1-\frac{K}{N}\Big)$$

4.3 (diskrétní veličina - binomické rozdělení)

Z urny, v níž jsou bílé a černé koule v poměru 9 : 1, vybíráme třikrát po jedné kouli, kterou vždy zase vracíme zpět. Náhodná veličina X je počet vybraných černých koulí. Určete:

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

(b) pravděpodobnostní funkci p_X a distribuční funkci F_X náhodné veličiny X.

(c) průměrný počet vybraných černých koulí.

(d) rozptyl $D(X)$ veličiny X.

(e) pravděpodobnost, že vytáhneme alespoň jednu černou kouli.

Řešení:

Budeme postupovat podobně jako v příkladu **4.2**.

Náhodná veličina X měří *počet úspěšných pokusů (vybrání černé koule) během n pokusů.* Zde je $n=3$ a pravděpodobnost každého (z nezávislých) pokusů je $p=0.1$.

Veličina X má tedy obor hodnot $\{0,1,2,3\}$ a (diskrétní) binomické rozdělení $\mathrm{Bi}(n,p)$. Pravděpodobnostní funkce proto je

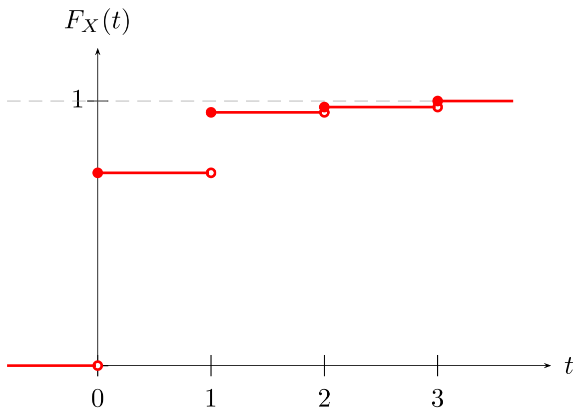
$$p_X(k) = P(X=k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & , k \in \{0,1,2,3\} \\ 0 & , \text{jinak.} \end{cases}$$

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

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



Průměrný počet vybraných černých koulí je střední hodnota naší diskrétní veličiny, tedy

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

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

Pro rozptyl $D(X) = E(X^2) - \Big(E(X)\Big)^2$ potřebujeme spočítat ještě

$$E(X^2) = \sum_{k \in \mathbb{R}} k^2 \cdot p_X(k) = 0^2 \cdot 0.729 + 1^2 \cdot 0.243 + 2^2 \cdot 0.027 + 3^2 \cdot 0.001 = 0.36$$

tedy

$$D(X) = E(X^2) - \Big(E(X)\Big)^2 = 0.36 - (0.3)^2 = 0.27$$

nebo podle vzorce pro binomické rozdělení $D(X) = np(1-p) = 3 \cdot 0.1 \cdot 0.9 = 0.27$.

A konečně, pravděpodobnost jevu

$$B = \text{“}\textit{vytáhneme alespoň jednu černou kouli}\text{”}$$

je

$$P(B) = P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) = 1 - 0.729 = 0.271\ .$$

Ještě si připomeneme, proč má veličina X vyjadřující počet úspěchů v n pokusech právě binomické rozdělení:

Jevy

$$A_i = \text{“\textit{úspěch v i-tém pokusu}”}$$

považujeme za nezávislé se stejnou pravděpodobností $P(A_i) = p$. Pak jev

$$B_k = \text{“\textit{nastane právě k úspěchů v n pokusech}”}$$

vyjádříme jako sjednocení disjunktních jevů (ty v největší závorce)

$$B_k = \bigcup_{\substack{K \subseteq \{1,\dots,n\} \\ |K| = k}} \left(\left(\bigcap_{i \in K} A_i \right) \cap \left(\bigcap_{j \in \{1,\dots,n\} \setminus K} \overline{A_j} \right) \right)$$

odkud máme ihned

$$P(X = k) = P(B_k) = \sum_{\substack{K \subseteq \{1,\dots,n\} \\ |K| = k}} \left(\left(\prod_{i \in K} P(A_i) \right) \cdot \left(\prod_{j \in \{1,\dots,n\} \setminus K} P(\overline{A_j}) \right) \right) =$$

$$= \sum_{\substack{K \subseteq \{1,\dots,n\} \\ |K| = k}} p^k \cdot (1-p)^{n-k} = \binom{n}{k} p^k \cdot (1-p)^{n-k}$$

4.4 (diskrétní veličina)

Do terče, nezávisle na sobě, střílí 3 střelci. První se trefí s pravděpodobností 70%, druhý s pravděpodobností 80% a třetí s pravděpodobností 90%. Náhodná veličina X je počet zásahů v terči. Určete:

(a) rozdělení pravděpodobnosti veličiny X.

(b) distribuční funkci F_X náhodné veličiny X.

(c) střední hodnotu $E(X)$ a rozptyl $D(X)$.

(d) pravděpodobnost, že v terči budou alespoň 2 zásahy.

Řešení:

Veličina X má obor hodnot $\{0, 1, 2, 3\}$ a tedy diskrétní rozdělení. Abychom si mohli odvodit pravděpodobnosti jednotlivých hodnot, označme si pro $i = 1, 2, 3$ jevy:

A_i = “*i-tý střelec se trefí*”,

které jsou nezávislé a mají pravděpodobnosti $P(A_1) = 0.7$, $P(A_1) = 0.8$ a $P(A_1) = 0.9$. Pravděpodobnostní funkce pak je

$$p_X(0) = P(X = 0) = P(\overline{A_1} \cap \overline{A_2} \cap \overline{A_3}) = (1 - 0.7) \cdot (1 - 0.8) \cdot (1 - 0.9) = 0.3 \cdot 0.2 \cdot 0.1 = \mathbf{0.006}$$

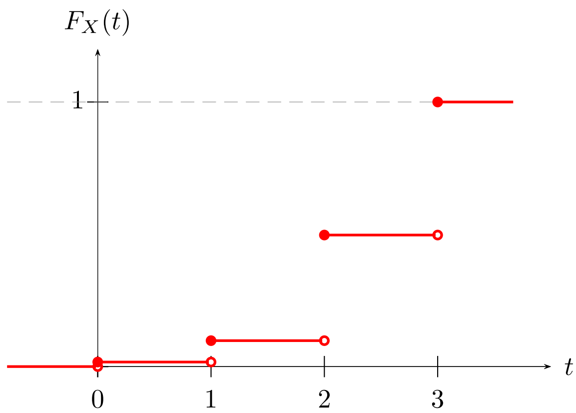
$$p_X(1) = P(X=1) = P\Big((A_1 \cap \overline{A_2} \cap \overline{A_3}) \ \cup \ (\overline{A_1} \cap A_2 \cap \overline{A_3}) \ \cup \ (\overline{A_1} \cap \overline{A_2} \cap A_3)\Big) =$$
$$= 0.7 \cdot 0.2 \cdot 0.1 + 0.3 \cdot 0.8 \cdot 0.1 + 0.3 \cdot 0.2 \cdot 0.9 = \mathbf{0.092}$$
$$p_X(2) = P(X=2) = P\Big((\overline{A_1} \cap A_2 \cap A_3) \ \cup \ (A_1 \cap \overline{A_2} \cap A_3) \ \cup \ (A_1 \cap A_2 \cap \overline{A_3})\Big) =$$
$$= 0.3 \cdot 0.8 \cdot 0.9 + 0.7 \cdot 0.2 \cdot 0.9 + 0.7 \cdot 0.8 \cdot 0.1 = \mathbf{0.398}$$
$$p_X(3) = P(X=3) = P(A_1 \cap A_2 \cap A_3) = 0.7 \cdot 0.8 \cdot 0.9 = 0.3 \cdot 0.2 \cdot 0.1 = \mathbf{0.504}$$

Konkrétně tedy:

k	0	1	2	3
$p_X(k)$	0.006	0.092	0.398	0.504

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

$$F_X(t) = \begin{cases} 0 & , t < 0 \\ p_X(0) = 0.006 & , t \in \langle 0, 1) \\ p_X(0) + p_X(1) = 0.098 & , t \in \langle 1, 2) \\ p_X(0) + p_X(1) + p_X(2) = 0.496 & , t \in \langle 2, 3) \\ 1 & , t \ge 3. \end{cases}$$



Průměrný počet zásahů v terči neboli střední hodnota naší diskrétní veličiny je

$$E(X) = \sum_{k \in \mathbb{R}} k \cdot p_X(k) = 0 \cdot 0.006 + 1 \cdot 0.092 + 2 \cdot 0.398 + 3 \cdot 0.504 = 2.4 \ .$$

Pro rozptyl $D(X) = E\big(X^2\big) - \Big(E\big(X\big)\Big)^2$ potřebujeme spočítat ještě

$$E(X^2) = \sum_{k \in \mathbb{R}} k^2 \cdot p_X(k) = 0^2 \cdot 0.006 + 1^2 \cdot 0.092 + 2^2 \cdot 0.398 + 3^2 \cdot 0.504 = 6.22$$

tedy

$$D(X) = E\big(X^2\big) - \Big(E\big(X\big)\Big)^2 = 6.22 - (2.4)^2 = 0.46 \ .$$

A konečně, pravděpodobnost jevu

$$B = \text{"}\textit{v terči budou alespoň 2 zásahy}\text{"}$$

je

$$P(B) = P(X \ge 2) = P(X=2) + P(X=3) = 0.398 + 0.504 = 0.902 \ .$$

4.5 (diskrétní veličina - binomické rozdělení)

Měli jsme tři stejné páry rukavic. Každá z nich se, nezávisle na ostatních, ztratila s pravděpodobností 0.1.

(a) Určete rozdělení veličin $X =$ *"počet zbylých levých rukavic"* a $Y =$ *"počet zbylých pravých rukavic"*.

(b) Jaká je pravděpodobnost, že pravých a levých rukavic je stejný počet?

(c) Jestliže víme, že pravých i levých rukavic zbylo stejně, jaká je pravděpodobnost, že se žádná rukavice neztratila?

Řešení:

(a) Veličina X označuje počet úspěchů (levá rukavice se neztratila s pravděpodobností $p = 0.9$) při $n = 3$ pokusech. Má tedy binomické rozdělení $\mathrm{Bi}(3, 0.9)$ (viz příklad **4.3**). Podobně to platí pro veličinu Y, která má tudíž také binomické rozdělení $\mathrm{Bi}(3, 0.9)$. Konkrétní hodnoty pravděpodobnostních funkcí

$$p_X(k) = p_Y(k) = \binom{n}{k} p^k \cdot (1-p)^{n-k} = \binom{3}{k} 0.9^k \cdot 0.1^{3-k} \text{ pro } k = 0, 1, 2, 3$$

udává tabulka

k	0	1	2	3
$p_X(k) = p_Y(k)$	0.001	0.027	0.243	0.729

(b) Obě diskrétní veličiny X a Y jsou tzv. *nezávislé*, což speciálně znamená, že

$$P(X = k \ \wedge \ Y = \ell) \ = \ P(X = k) \cdot P(Y = \ell) \ \text{ pro všechna } k, \ell \in \{0, 1, 2, 3\} \ .$$

To, že uvedené rovnosti platí, zdůvodníme pomocí jevů

$$A_i = \text{"}i\text{-tá levá rukavice se neztratila"}$$

$$B_j = \text{"}j\text{-tá pravá rukavice se neztratila"}$$

které jsou všechny dohromady nezávislé. Jev "$X = k$" pak vznikne pomocí průniků, sjednocení a doplňků pouze z jevů A_i, $i = 1, 2, 3$. Podobně jev "$Y = \ell$" vznikne pomocí průniků, sjednocení a doplňků pouze z jevů B_j, $j = 1, 2, 3$. Podle pravidel o manipulaci s nezávislými jevy tak musí být jevy "$X = k$" a "$Y = \ell$" nezávislé, tudíž výše uvedená rovnost platí.

S využitím této znalosti pak pro jev

$$C = \text{"pravých i levých rukavic zbyl stejný počet"}$$

máme

$$P(C) = P(X = Y) = \sum_{k=0}^{3} P(X = k \ \wedge \ Y = k) = \sum_{k=0}^{3} P(X = k) \cdot P(Y = k) =$$

$$= \sum_{k=0}^{3} \Big(p_X(k)\Big)^2 = 0.001^2 + 0.027^2 + 0.243^2 + 0.729^2 = 0.59122 \ .$$

(c) Podmíněná pravděpodobnost jevu, že se žádná rukavice neztratila (tj. $X = Y = 3$), pokud obou typu rukavic zbylo stejně (tj. $X = Y$), je

$$P(X = Y = 3 | X = Y) = \frac{P(X = Y = 3 \ \wedge \ X = Y)}{P(X = Y)} = \frac{P(X = Y = 3)}{P(X = Y)} =$$

$$= \frac{P(X = 3) \cdot P(Y = 3)}{P(X = Y)} \doteq \frac{0.729^2}{0.591} \doteq 0.899 \, .$$

4.6 (diskrétní veličina - hypergeometrické rozdělení)

V krabici jsou 4 nové a 3 použité tenisové míčky. Pro *první* hru se náhodně se vyberou 2 míčky, které se po skončení hry zase vrátí do krabice.

(a) Pro veličinu $X =$ *"počet nových míčků vybraných pro (první) hru"* určete její rozdělení a střední počet nových míčků mezi vybranými.

(b) Pro *druhou* hru se znovu náhodně vyberou opět 2 míčky. Určete pravděpodobnost toho, že všechny míčky, použité při této *druhé* hře, jsou nové.

Řešení:

(a) Ze $4+3=7$ míčků, z toho 4 nových, vybíráme 2 míčky. Veličina X má proto hodnoty $\{0,1,2\}$ a pravděpodobnostní funkci

$$p_X(k) = P(X=k) = \begin{cases} \frac{\binom{4}{k}\cdot\binom{3}{2-k}}{\binom{7}{2}} & , k \in \{0,1,2\} \\ 0 & , \text{jinak.} \end{cases}$$

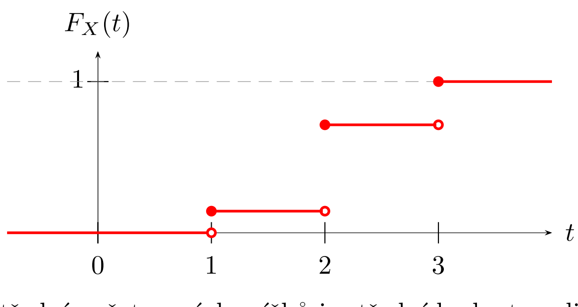
Je to tzv. hypergeometrické rozdělení (viz příklad **4.2**).

Konkrétně:

k	0	1	2
$p_X(k)$	$\frac{\binom{4}{0}\cdot\binom{3}{2}}{\binom{7}{2}} = \frac{1}{7} \doteq 0.143$	$\frac{\binom{4}{1}\cdot\binom{3}{1}}{\binom{7}{2}} = \frac{4}{7} \doteq 0.571$	$\frac{\binom{4}{2}\cdot\binom{3}{0}}{\binom{7}{2}} = \frac{2}{7} \doteq 0.286$

Distribuční funkce pro toto diskrétní rozdělení je

$$F_X(t) = P(X \le t) = \sum_{k \le t} p_X(k) = \begin{cases} 0 & , t < 0 \\ p_X(0) = \frac{1}{7} \doteq 0.143 & , t \in \langle 0,1) \\ p_X(0) + p_X(1) = \frac{5}{7} \doteq 0.714 & , t \in \langle 1,2) \\ 1 & , t \ge 2. \end{cases}$$



Střední počet nových míčků je střední hodnota veličiny X, tedy

$$E(X) = \sum_{k \in \mathbb{R}} k \cdot p_X(k) = 0 \cdot \frac{1}{7} + 1 \cdot \frac{4}{7} + 2 \cdot \frac{2}{7} = \frac{8}{7} \doteq 1.14\ .$$

Střední hodnotu také můžeme spočítat na základě hypergeometrického rozdělení $\mathrm{Hyp}(n; K, N) = \mathrm{Hyp}(2; 4, 7)$ jako

$$E(X) = n \cdot \frac{K}{N} = 2 \cdot \frac{4}{7} = \frac{8}{7} \; .$$

(b) Pro $i = 0, 1, 2$ si označme jevy:

A_i = *"pro první hru bylo vybráno i nových a* $2 - i$ *použitých míčku"* = *"X=i"*,
B = *"pro druhou hru byly vybrány* 2 *nové míčky"*.

Víme, že A_0, A_1 a A_2 je úplný disjunktní systém jevů. Hledáme pravděpodobnost $P(B)$. Z věty o úplné pravděpodobnosti máme

$$P(B) = \sum_{i=0}^{2} P(B|A_i) \cdot P(A_i)$$

Zbývá určit podmíněné pravděpodobnosti. Jestliže nastal jev A_i, je v krabici $4 - i$ nových míčků. Takže máme

$$P(B|A_i) = \frac{\binom{4-i}{2}}{\binom{7}{2}}$$

s hodnotami v tabulce

i	0	1	2
$P(B\|A_i)$	$\frac{\binom{4}{2}}{\binom{7}{2}} = \frac{6}{21}$	$\frac{\binom{3}{2}}{\binom{7}{2}} = \frac{3}{21}$	$\frac{\binom{2}{2}}{\binom{7}{2}} = \frac{1}{21}$

Výsledek tak je

$$P(B) = \sum_{i=0}^{2} P(B|A_i) \cdot P(A_i) = \sum_{i=0}^{2} P(B|A_i) \cdot P(X = i) = \tfrac{6}{21} \cdot \tfrac{1}{7} + \tfrac{3}{21} \cdot \tfrac{4}{7} + \tfrac{1}{21} \cdot \tfrac{2}{7} = \tfrac{20}{147} \doteq 0.136 \; .$$

Definice: Veličina $X : \Omega \to \mathbb{R}$ má tzv. *spojité rozdělení* $\Leftrightarrow$ její distribuční funkce F_X je spojitá.

Často je v tomto případě F_X tzv. *absolutně spojitá* $\Leftrightarrow$ ex. $f_X : \mathbb{R} \to \langle 0, +\infty)$, která je integrabilní a platí

$$F_X(t) = \int_{-\infty}^{t} f_X(u) \; \mathrm{d}u$$

pro každé $t \in \mathbb{R}$.

Funkce f_X se nazývá *hustotou pravděpodobnosti* veličiny X.

Odsud snadno máme, že pokud $I \subseteq \mathbb{R}$ je interval (nebo i nějaká množina poskládaná z intervalů), pak

$$P(X \in I) = \int_{I} f_X(u) \; \mathrm{d}u$$

tedy pravděpodobnost, že hodnoty veličiny X padnou do I, zjistíme prostě zintegrováním hustoty přes I (podobně zjišťujeme např. hmotnost nějaké křivky, když zintegrujeme hustotu (hmotnosti) přes danou křivku).

Poznamenejme ještě důležitou věc a sice, že hustota f_X NENÍ zdaleka určena jednoznačně jako funkce (např. její změnou v konečně mnoha bodech se nezmění příslušné integrály, takže i změněná funkce bude také hustotou). Pokud veličina X hustotu má, pak platí, že derivace funkce F_X, tj. $(F_X)'$ je hustotou veličiny X (tato derivace sice nemusí všude existovat, ale to není na překážku - tam, kde neexistuje, si prostě zvolíme libovolné nezáporné hodnoty).

Věta (o existenci hustoty):

Nechť F_X je spojitá a existují reálná čísla $a_0 < a_1 < \cdots < a_n$ taková, že na otevřených intervalech $(-\infty,\ a_1)$, $(a_1,\ a_2)$, $\ldots$, $(a_{n-1},\ a_n)$, $(a_n,\ +\infty)$ je F_X spojitě diferencovatelná. Pak veličina X má hustotu a funkce

$$f_X(t) = \begin{cases} (F_X)'(t) & \text{pokud tato derivace v bodě } t \text{ existuje a je konečná} \\ 0 & \text{jinak.} \end{cases}$$

je hustotou pravděpodobnosti veličiny X.

Poznámka: Stojí za to ještě poznamenat, že vlastnost absolutní spojitosti není samozřejmá pro spojité distribuční funkce - příkladem je tzv. *Cantorova funkce c*, také známá jako *"ďábelské schodiště"* (viz např. `https://en.wikipedia.org/wiki/Cantor_function`). Tato funkce (rozšířená z intervalu $\langle 0, 1\rangle$ přirozeně na celé $\mathbb{R}$) má nulovou derivaci c' až na (z hlediska integrálu) nepodstatnou množinu bodů. Přesto však Cantorova funkce není konstantní (a tudíž nelze získat integrováním své derivace - přislušné integrály $\int_{-\infty}^{t} c'(t)\ \mathrm{d}t = 0$ jsou všechny nulové). Tedy pro veličinu s touto distribuční funkcí neexistuje hustota pravděpodobnosti.

4.7 (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 hustotu pravděpodobnosti f_X.

(c) pravděpodobnost $P(X \geq \frac{1}{2})$.

(d) střední hodnotu $E(X)$ a rozptyl $D(X)$.

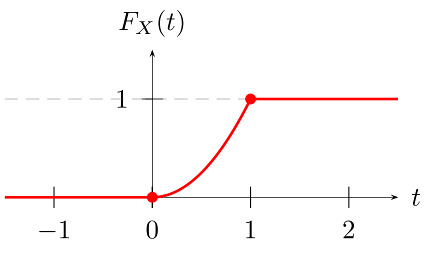
(e) 90% interval spolehlivosti se středem v $E(X)$, tedy takový interval $\langle E(X) - \varepsilon, E(X) + \varepsilon\rangle$, pro $\varepsilon > 0$, že $P\Big(|X - E(X)| \leq \varepsilon\Big) = 0.9$.

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

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



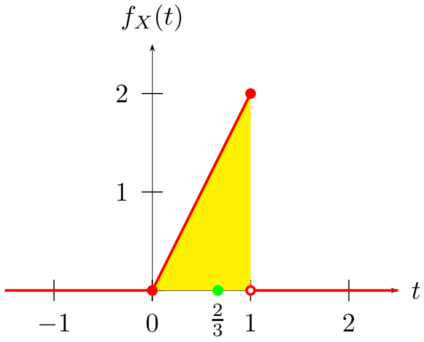
Veličina X má spojité rozdělení, protože distribuční funkce F_X je spojitá. Hustotu pravděpodobnosti (tj. funkci $f_X : \mathbb{R} \to \langle 0, +\infty)$, pro kterou je $F_X(t) = \int_{-\infty}^{t} f_X(u)\ du$ pro každé $t \in \mathbb{R}$) můžeme určit jako derivaci funkce F_X tam, kde derivace existuje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



Dále máme

$$P(X \geq \tfrac{1}{2}) = 1 - P(X < \tfrac{1}{2}) = 1 - P(X \leq \tfrac{1}{2}) = 1 - F_X(\tfrac{1}{2}) = 1 - (\tfrac{1}{2})^2 = \tfrac{3}{4}\ .$$

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.

Dále, pro rozptyl $D(X) = E(X^2) - \big(E(X)\big)^2$ potřebujeme ještě určit $E(X^2)$, což uděláme pomocí vzorce analogického tomu pro diskrétní rozdělení:

$$E\big(h(X)\big) = \int_{-\infty}^{\infty} h(t) \cdot f_X(t)\ \mathrm{d}t,$$

kde $h : \mathbb{R} \to \mathbb{R}$ je (borelovsky) měřitelná funkce, např. po částech spojitá. (Měřitelnost znamená, že h je "obyčejná" náhodná veličina ovšem vzhledem k σ-algebře borelovských množin, tedy $h^{-1}(I)$ je borelovská pro každý interval $I \subseteq \mathbb{R}$. No a k tomu stačí např. částečná spojitost.)

Tedy

$$E(X^2) = \int_{-\infty}^{\infty} t^2 \cdot f_X(t)\ \mathrm{d}t = \int_0^1 t^2 \cdot 2t\ \mathrm{d}t = \Big[\tfrac{2}{4} t^4\Big]_0^1 = \tfrac{1}{2}$$

a proto máme

$$D(X) = E(X^2) - \big(E(X)\big)^2 = \tfrac{1}{2} - (\tfrac{2}{3})^2 = \tfrac{1}{18}\ .$$

A konečně, hledáme $\varepsilon > 0$ takové, že

$$0.9 = P\Big(X \in \langle \tfrac{2}{3} - \varepsilon, \tfrac{2}{3} + \varepsilon \rangle\Big) = P\Big(X \leq \tfrac{2}{3} + \varepsilon\Big) - P\Big(X < \tfrac{2}{3} - \varepsilon\Big) = F_X\left(\tfrac{2}{3} + \varepsilon\right) - F_X\left(\tfrac{2}{3} - \varepsilon\right)\ .$$

Vzhledem k předpisu F_X teď máme dvě možnosti:

- $0 < \varepsilon \le \frac{1}{3}$:
  Pak je $\frac{2}{3}+\varepsilon, \frac{2}{3}-\varepsilon \in \langle 0,1\rangle$ a tedy
  $$0.9 = F_X\left(\tfrac{2}{3}+\varepsilon\right) - F_X\left(\tfrac{2}{3}-\varepsilon\right) = \left(\tfrac{2}{3}+\varepsilon\right)^2 - \left(\tfrac{2}{3}-\varepsilon\right)^2 = \tfrac{8}{3}\varepsilon$$
  odkud máme $\varepsilon = 0.9 \cdot \frac{3}{8} = 0.3375 \not\le \frac{1}{3}$, což je spor. Tato možnost tedy nedává řešení.
- $\frac{1}{3} < \varepsilon \le \frac{2}{3}$:
  Pak je $\frac{2}{3}+\varepsilon > 1$ a $\frac{2}{3}-\varepsilon \in \langle 0,1\rangle$ a tedy
  $$0.9 = F_X\left(\tfrac{2}{3}+\varepsilon\right) - F_X\left(\tfrac{2}{3}-\varepsilon\right) = 1 - \left(\tfrac{2}{3}-\varepsilon\right)^2 \quad \Rightarrow \quad \left(\tfrac{2}{3}-\varepsilon\right)^2 = 0.1$$
  takže $\varepsilon = \frac{2}{3} \pm \sqrt{0.1}$ a z omezujících podmínek pro ε je jediné řešení
  $$\varepsilon = \tfrac{2}{3} - \sqrt{0.1} \doteq 0.6496\ .$$

4.8 (spojitá náhodná veličina)

Bod se pohybuje v rovině po kružnici (se středem v počátku a poloměrem 1) stálou úhlovou rychlostí. Náhodná veličina X je průmět bodu na osu x. Pravděpodobnost nalezení X v intervalu $\langle a,b\rangle \subseteq \langle -1,1\rangle$ je dána jako
$$P\big(X \in \langle a,b\rangle\big) = \frac{\textit{“čas, co stráví průmět bodu v } \langle a,b\rangle\textit{”}}{\textit{“čas, co stráví průmět bodu v } \langle -1,1\rangle\textit{”}}$$

Určete:

(a) distribuční funkci F_X a hustotu pravděpodobnosti f_X.

(b) pravděpodobnost $P(\frac{1}{2} \le |X| < 3)$

(c) $\varepsilon > 0$ takové, že $P\big(X \in \langle -\varepsilon,\varepsilon\rangle\big) = \frac{1}{2}$.

(d) střední hodnotu $E(X)$ a rozptyl $D(X)$.

Řešení:
Pohyb bodu po dané kružnici je popsán v čase t jako $x(t) = \cos(\omega t)$ a $y(t) = \sin(\omega t)$, kde $\omega > 0$ je úhlová rychlost. Stačí uvažovat jen oběh bodu po půlkružnici a proto si pro t zvolíme časový interval např. $\langle 0, \frac{\pi}{\omega}\rangle$ (s rostoucím t nám pak souřadnice $x(t)$ bude klesat).

Pro interval $\langle a,b\rangle \subseteq \langle -1,1\rangle$ mějme teď takové (jednoznačně určené) $t_1, t_2 \in \langle 0, \frac{\pi}{\omega}\rangle$, $t_2 \le t_1$, že $a = \cos(\omega t_1)$ a $b = \cos(\omega t_2)$. Vzhledem k volbě časového intervalu máme

$t_1 = \frac{\arccos a}{\omega}$

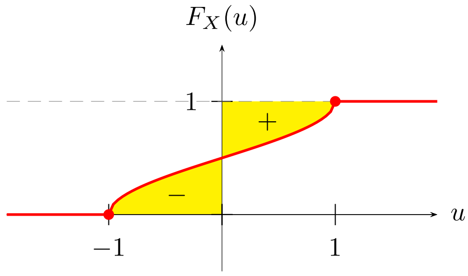
$t_2 = \frac{\arccos b}{\omega}$

Speciálně pro $a = -1$ a $b = 1$ je $t_1 = \frac{\pi}{\omega}$ a $t_2 = 0$.

Pak máme
$$P\big(X \in \langle a,b\rangle\big) = \frac{\textit{“čas, co stráví průmět bodu v } \langle a,b\rangle\textit{”}}{\textit{“čas, co stráví průmět bodu v } \langle -1,1\rangle\textit{”}} = \frac{t_1 - t_2}{\frac{\pi}{\omega} - 0} = \frac{\arccos(a) - \arccos(b)}{\pi}$$

Protože obor hodnot veličiny X je zřejmě $\langle -1, 1\rangle$, dostáváme distribuční funkci

$$F_X(u) = P(X \le u) = \begin{cases} 0 & , u \le -1 \\ P\big(X \in \langle -1, u\rangle\big) = \quad 1 - \frac{\arccos(u)}{\pi} & , u \in \langle -1, 1\rangle \\ 1 & , u \ge 1 \ . \end{cases}$$



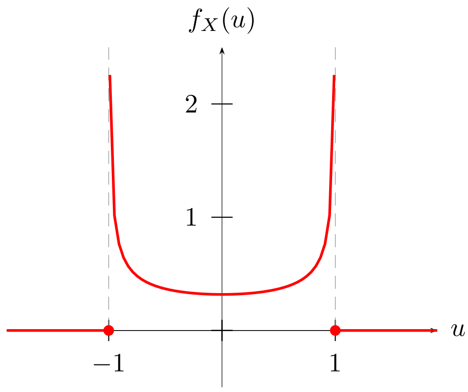
Hustotu f_X určíme z derivace F_X tam, kde tato derivace existuje (v jiných bodech si hodnotu f_X můžeme zvolit libovolně):

$$\frac{\mathrm{d}}{\mathrm{d}u} F_X(u) = \begin{cases} \frac{\mathrm{d}}{\mathrm{d}u}\left(1 - \frac{\arccos(u)}{\pi}\right) = \frac{1}{\pi}\cdot\frac{1}{\sqrt{1-u^2}} & , u \in (-1, 1) \\ 0 & , u < -1 \text{ nebo } u > 1 \ . \end{cases}$$

Takže funkce

$$f_X(u) = \begin{cases} \frac{1}{\pi}\cdot\frac{1}{\sqrt{1-u^2}} & , u \in (-1, 1) \\ 0 & , \text{jinak} \ . \end{cases}$$

je hustotou pravděpodobnosti veličiny X.



Jak je vidět, je f_X sudá funkce. Z toho např. vyplývá, že pro interval $\langle -b, -a\rangle \subseteq (-\infty, 0\rangle$ máme

$$P\big(X \in \langle -b, -a\rangle\big) = \int_{-b}^{-a} f_X(t)\ \mathrm{d}t = \left\{ \begin{smallmatrix} t=-u \\ dt=-du \\ u\in\langle a,b\rangle \end{smallmatrix} \right\} = \int_a^b f_X(u)\ \mathrm{d}u = P\big(X \in \langle a, b\rangle\big)$$

díky čemuž máme

$$P\left(\tfrac{1}{2} \le |X| < 3\right) = P\Big(X \in (-3, -\tfrac{1}{2}\rangle\Big) + P\Big(X \in \langle\tfrac{1}{2}, 3)\Big) = 2\cdot P\left(\tfrac{1}{2} \le X < 3\right) =$$

$$= 2\cdot\Big(P\left(X < 3\right) - P\left(X < \tfrac{1}{2}\right)\Big) = 2\cdot\Big(F_X(3) - F_X\left(\tfrac{1}{2}\right)\Big) = 2\cdot\Big(1 - 1 + \tfrac{1}{\pi}\cdot\arccos\left(\tfrac{1}{2}\right)\Big) = \tfrac{2}{\pi}\cdot\tfrac{\pi}{3} = \tfrac{2}{3} \ .$$

Podobně hledáme $0 < \varepsilon < 1$, aby

$$\tfrac{1}{2} = P\big(X \in \langle -\varepsilon, \varepsilon \rangle\big) = 2 \cdot P\big(0 \le X \le \varepsilon\big) = 2\big(F_X(\varepsilon) - F_X(0)\big) = \tfrac{2}{\pi}\big(\tfrac{\pi}{2} - \arccos(\varepsilon)\big)$$

tedy

$$\arccos(\varepsilon) = \tfrac{\pi}{4} \quad \Rightarrow \quad \varepsilon = \cos\tfrac{\pi}{4} = \tfrac{\sqrt{2}}{2} \ .$$

Střední hodnotu spočítáme pomocí hustoty pravděpodobnosti

$$E(X) = \int_{-\infty}^{\infty} \underbrace{t \cdot \underbrace{f_X(t)}_{\text{sudá funkce}}}_{\text{lichá funkce}} \ \mathrm{d}t = \frac{1}{\pi}\int_{-1}^{1} \frac{t}{\sqrt{1-t^2}} \ \mathrm{d}t = \frac{1}{\pi}\Big[-\sqrt{1-t^2} \Big]_{-1}^{1} = 0 \ .$$

Pro sudou hustotu tedy bude střední hodnota veličiny (pokud existuje!) vždy nulová. (POZOR: střední hodnota veličiny X obecně existovat nemusí!!). Existenci střední hodnoty (i bez počítání) v našem případě můžeme zjistit i z tvaru distribuční funkce F_X a sice pomocí tohoto vztahu:

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

jehož geometrická interpretace je, že střední hodnota je rovna součtu velikosti ploch s daným znaménkem v grafu F_X (viz žluté plochy na obrázku). Protože plochy jsou omezené, oba integrály existují, a protože obě plochy jsou stejně velké, je skutečně $E(X) = 0$.

Pro výpočet rozptylu $D(X)$ opět budeme potřebovat:

$$E(X^2) = \int_{-\infty}^{\infty} t^2 \cdot f_X(t) \ \mathrm{d}t = \frac{1}{\pi}\int_{-1}^{1} \frac{t^2}{\sqrt{1-t^2}} \ \mathrm{d}t = \left\{ \begin{smallmatrix} t=\cos(\alpha) \\ dt=-\sin(\alpha)d\alpha \\ \alpha\in\langle 0,\pi\rangle \end{smallmatrix} \right\} =$$

$$= \frac{1}{\pi}\int_{\pi}^{0} \frac{\cos^2(\alpha)}{\sin(\alpha)}\Big(-\sin(\alpha)\Big) \ \mathrm{d}\alpha = \frac{1}{\pi}\int_{0}^{\pi} \cos^2(\alpha) \ \mathrm{d}\alpha = \frac{1}{\pi} \cdot \frac{\pi}{2} = \frac{1}{2} \ .$$

Takže $D(X) = E(X^2) - \Big(E(X)\Big)^2 = \frac{1}{2}$.