Úvod do pravděpodobnosti

9. října 2022

Definice(Základní pojmy). Dejinujme několik základních pojmů.

- **Elementární jev** ω je jev, jehož vnitřní strukturu už dále nerozlišujeme. Např. při hodu kostkou mohou nastat elemetnární jevy "padne 1", "padne 2", atd.
- **Základní množinu, výběrový prostor**, tj. množinu všech elementárních jevů ω, označme Ω.
- **Jev** A je přímo elemetární jev, nebo je libovolnou podmnožinou množiny všech elementárních jevů, tj. $A \subset \Omega$.
- Nechť $A = \emptyset$. Pak A nazveme jevem **nemožným**.
- Nechť $A = \Omega$. Pak A nazveme jevem **jistým**.
- Řekneme, že jev A **nastal**, pokud nastal elemetrární jev $\omega \in \Omega$ a navíc $\omega \in A$.

Definice(Operace s jevy). Nechť $A, B \subset \Omega$ jsou libovolné jevy.

- **Komplementární (opačný)** jev A^C k jevu A definujeme jako

$$\omega \in A^C \Leftrightarrow \omega \notin A,$$

  tj. jev A^C nastane právě tehdy, když nenastane jev A.
- Inkluzi $A \subset B$ definujeme následovně:

$$\omega \in A \Rightarrow \omega \in B,$$

  tj. nastane-li jev A, pak nastane i jev B.
- Řekneme, že jevy A, B jsou **totožné**, právě když platí:

$$A \subset B \wedge B \subset A,$$

  neboli jev A nastane právě tehdy, když nastane jev B.
- **Průnikem** jevů $A \cap B$ myslíme jev, při němž nastává jev A a zároveň jev B, tj.

$$(\omega \in A \cap B) \Leftrightarrow (\omega \in A \wedge \omega \in B).$$

  Je-li $A \cap B = \emptyset$, pak jsou A, B **disjunktní (neslučitelné)**.
- **Rozdílem** jevů A a B rozumíme

$$A - B = A \cap B^C.$$

- **Sjednocení** jevů $A \cup B$ je definováno jako:

$$\omega \in (A \cup B) \Leftrightarrow \omega \in A \vee \omega \in B,$$

  tj. jev $A \cup B$ nastane, nastane-li jev A nebo B. Jsou-li navíc A, B disjunktní, nazveme jejich sjednocení **direktním součinem** jevů A, B a značíme $A + B$. Sjednocení, resp. direktní součet více jevů budeme zapisovat následovně:

$$\bigcup_{j=1}^{+\infty} A_j, \text{ resp.} \sum_{j=1}^{+\infty} A_j.$$

Definice(σ-algebra). Buď Ω základní množina a mějme systém podmnožin $\mathcal{A} \subset \mathcal{P}(\Omega)$, který splňuje následující axiomy:

- $\Omega \in \mathcal{A}$,
- $(\forall A \in \mathcal{A})(A^C \in \mathcal{A})$,
- $(\forall j \in \mathbf{N})(A_j \in \mathcal{A}) \Rightarrow \bigcup_{j=1}^{+\infty} A_j \in \mathcal{A}$

Pak $\mathcal{A}$ nazýváme σ-algebrou jevů (množin) na množině Ω. Symbolem $\mathcal{P}(\Omega)$ značíme potenční množinu množiny Ω, tedy množinu

$$2^\Omega = \{B : B \subset \Omega\}$$

Definice(Axiomatická definice pravděpodobnosti). Mějme neprázdnou základní množinu Ω a na ní σ-algebru $\mathcal{A}$. Pak libovolnou funkci $\mathscr{P} : \mathcal{A} \to \mathbf{R}$ splňující tři Kolmogorovy axiomy

- $\mathscr{P}(\Omega) = 1$,
- $(\forall A \in \mathcal{A})(\mathscr{P}(A) \geqslant 0)$,
- $(\forall (A_j)_{j=1}^{+\infty} \in \mathcal{A},$ kde A_j jsou navzájem disjunktní$)\left(\mathscr{P}\left(\sum_{j=1}^{+\infty} A_j\right) = \sum_{j=1}^{+\infty} \mathscr{P}(A_j)\right)$,

nazveme **pravděpodobnostní mírou**. Pravděpodobnost je tedy množinová funkce, její definiční obor je soustava množin (množina množin) a její obor hodnot je interval $\langle 0, 1\rangle$.

Rozlišujeme tzv. **deterministické systémy**, systémy, které ze stejných počátečních podmínek dospějí do stále stejného stavu, např. klasická mechanika. Těmito systémy se zabývat nebudeme. Budeme se zabývat tzv. **stochastickými systémy**.

0.1 Klasifikace náhodných veličin

Náhodné veličiny můžeme rozlišit na **kvantitantivní**, tedy numerické, a **kvalitativní**, tedy slovní. Dále se rozlišují podle počtu prvků soustavy $\mathcal{A}$ na **diskrétní** a **spojité**.

0.1.1 Diskrétní náhodné veličiny

Pravděpodobnost je zde funkce $\mathscr{P} : \mathcal{A} \subset 2^\Omega \to \langle 0, 1\rangle$, kde $\mathrm{card}(\mathcal{A}) \in \mathbf{N}$. Máme tedy konečný počet výsledků (realizací) náhodné veličiny X.

Značení: Symbolem X budeme dále značit vybranou náhodnou veličinu (stochastickou i diskrétní). Náhodná veličina není náhodná a není to veličina. Jde o $\mathcal{A}$-měřitelnou funkci, viz definice z MAB4. Korektní definice zazní na přednáškách z MIPu, my se jí teď budeme snažit pro lepší pochopení vyhnout. Příkladem náhodné veličiny je hod kostkou, tlak pacienta, mzda zaměstnance a další. Symbolem x budeme značit realizaci (jednu) dané náhodné veličiny, tj. výsledek náhodného jevu. Pravděpodobnost, že náhodná veličina X nabude hodnoty právě x budeme značit symbolem $\mathscr{P}[X = x]$.

Pravděpodobnostní dotazy mohou být různého typu. Můžeme se ptát na pravděpodobnosti elementárních jevů, tedy např. $\mathscr{P}[X = 3] =?$, $\mathscr{P}[Y = 28 \text{ Kč}] =?$, $\mathscr{P}[Z = 13 \text{ min}] =?$. Můžeme ale mít i složitější dotaz, typu $\mathscr{P}[X \in \{1, 5, 6\}] = ?$, nebo dokonce $\mathscr{P}[Y \in \langle 4500, 5500\rangle] =?$. Základní pravděpodobnostní dotaz tedy bude

$$\mathscr{P}[X \in B] =?, \text{ kde } B \in \mathcal{A} \subset 2^\Omega, \text{ tj. } B \subset \Omega$$

Navíc dodejme, že pokud $\mathrm{card}(\Omega) \in \mathbf{N} \Rightarrow \mathrm{card}(B) \in \mathbf{N}$. V diskrétních případech navíc téměř vždy platí $\mathcal{A} = 2^\Omega$. Dále budeme často využívat následující zkrácený zápis, který nám bude víc připomínat práci s mírou:

$$\mathscr{P}[X \in A] = \mathscr{P}[A], \text{ kde } A \in \mathcal{A}.$$

Základní pravděpodobnostní popis probíhá pomocí:

- **pravděpodobnostní funkce**, v diskrétním případě jde o následující tabulku (bereme v potaz hod kostkou).

x	$\mathscr{P}[X = x]$
1	$\frac{1}{6}$
2	$\frac{1}{6}$
3	$\frac{1}{6}$
4	$\frac{1}{6}$
5	$\frac{1}{6}$
6	$\frac{1}{6}$

Potom např.:

$$\mathscr{P}[X \in \{2, 4, 6\}] = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}.$$

- **distribuční funkce** $F(x) = \mathscr{P}[X < x]$, která může být vyčíslena $\forall x \in \mathbf{R}$. Pokud uvážíme příklad s kostkou, pak např.:

$$F(-3) = \mathscr{P}[X < -3] = 0$$
$$F(3) = \mathscr{P}[X < 3] = \mathscr{P}[X \in \{1, 2\}] = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$
$$F(8) = \mathscr{P}[X < 8] = 1$$

Deskriptivní charakteristiky diskrétní náhodné veličiny X

- **Střední hodnota**, značíme ji $\mathrm{E}(X)$, vyjadřuje pravděpodobnostní těžiště, anglicky *expected value*.

$$\mathrm{E}(X) = \sum_{x \in \Omega} x \cdot \mathscr{P}[X = x]$$

Pro příklad s kostkou je tedy

$$\mathrm{E}(X) = \sum_{x=1}^{6} x \mathscr{P}[X = x] = 1\frac{1}{6} + 2\frac{1}{6} + 3\frac{1}{6} + \cdots + 6\frac{1}{6} = 3.5$$

- **Rozptyl** vyjadřuje míru rozptýlení hodnot kolem střední hodnoty, značíme $\mathrm{D}(X)$, nebo $\mathrm{VAR}(X)$, anglicky *variance*. Pokud $\mathrm{D}(X) = 0$, potom X je deterministický jev.

$$\mathrm{D}(X) = \mathrm{E}(X - \mathrm{E}(X))^2 = \sum_{x \in \Omega} (x - \mathrm{E}(X))^2 \cdot \mathscr{P}[X = x]$$

Označme $\lambda = \mathrm{E}(X) \in \mathbf{R}$ a zjednodušme vztah pro rozptyl.

$$\begin{aligned} \mathrm{D}(X) &= \sum_{x \in \Omega} (x^2 - 2x\lambda + \lambda^2) \cdot \mathscr{P}[X = x] = \sum_{x \in \Omega} x^2 \mathscr{P}[X = x] - 2\lambda \sum_{x \in \Omega} x \mathscr{P}[X = x] + \lambda^2 \mathscr{P}[X = x] = \\ &= \sum_{x \in \Omega} x^2 \mathscr{P}[X = x] - 2\lambda^2 + \lambda^2 = \sum_{x \in \Omega} x^2 \cdot \mathscr{P}[X = x] - \mathrm{E}^2(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) \end{aligned} \tag{0.1}$$

Dále bychom mohli definovat **směrodatnou odchylku**, anglicky *standard deviation* jako

$$\mathrm{SD}(X) = \sqrt{\mathrm{D}(X)}$$

- **K-tý moment** definujeme pro $k \in \mathbf{N}_0$ jako

$$\mathrm{E}(X^k) = \sum_{x \in \Omega} x^k \cdot \mathscr{P}[X = x].$$

Je zřejmé, že $\mathrm{E}(X^0) = \mathrm{E}(1) = 1$, což platí díky prvnímu axiomu z definice pravděpodobnostní míry. Dále $\mathrm{E}(X^1) = \mathrm{E}(X)$. Druhý moment $\mathrm{E}(X^2) = \sum_{x \in \Omega} x^2 \mathscr{P}[X = x]$ jsme již použili při úpravě vzorce pro rozptyl.

Rozdělení disktrérních náhodných veličin

- **Uniformní (rovnoměrné) rozdělení** je jednoparametrické rozdělení $\mathrm{Uni}(n)$, kde parametr n vyjadřuje počet elementárních jevů, tedy $n = \mathrm{card}(E)$. Pravděpodobnostní funkce vyjádřená tabulkou vypadaná následovně:

el. jevy	$\mathscr{P}[X = x]$
1	$\frac{1}{n}$
2	$\frac{1}{n}$
3	$\frac{1}{n}$
4	$\frac{1}{n}$
$\vdots$	$\vdots$
n	$\frac{1}{n}$

Ihned vidíme, že

$$\sum_{x=1}^{n} \mathscr{P}[X = x] = 1.$$

Skutečnost, že náhodná veličina má uniformní rozdělení s parametrem n zapisujeme jako $X \sim \mathrm{Uni}(n)$.

- **Alternativní (binomické) rozdělení** s parametrem $p \in (0,1)$, značíme $\mathrm{AB}(p)$. Jiný název je také Bernoulliho rozdělení. Pravděpodobnostní tabulka je

el. jevy	$\mathscr{P}[X=x]$
1	p
0	1-p

Opět platí, že $\sum_{x\in\Omega}\mathscr{P}[X=x]=1$, kde $\Omega=\{0,1\}$. Vypočteme střední hodnotu:

$$\mathrm{E}(X)=\sum_{x=0}^{1}x\mathscr{P}[X=x]=0\cdot(1-p)+1p=p,$$

dále rozptyl:

$$\mathrm{E}(X^2)=\sum_{x=0}^{1}x^2\mathscr{P}[X=x]=p$$
$$\mathrm{D}(X)=p-p^2=p(1-p)$$

- **Binomické rozdělení** $\mathrm{Bi}(n,p)$ vyjadřuje počet úspěchů v n opakovaných pokusech. Parametr p zde vyjadřuje pravděpodobnost úspěchu v každém jednotlivém pokusu. Pravděpodobnostní funkce má následující předpis:

$$\mathscr{P}[X=x]=\binom{n}{x}p^x(1-p)^{n-x},$$

kde x vyjadřuje počet úspěchů, $n-x$ počet neúspěchů v n pokusech a $1-p$ je prst neúspěchu. Výběrový prostor je zde $\Omega=\{0,1,2,\dots,n\}$. Vypočteme opět střední hodnotu:

$$\begin{aligned}\mathrm{E}(X)&=\sum_{x=0}^{n}x\cdot\binom{n}{x}p^x(1-p)^{n-x}=\sum_{x=1}^{n}x\cdot\binom{n}{x}p^x(1-p)^{n-x}=\sum_{x=1}^{n}\frac{n(n-1)!}{(x-1)!(n-x)!}p^x(1-p)^{n-x}=\\&=n\sum_{x=1}^{n}\frac{(n-1)!}{(x-1)!\,[n-1-(x-1)]!}p^x(1-p)^{n-1-(x-1)},\end{aligned}\tag{0.2}$$

nyní použijeme substituci $y=x-1$, tím se také změní meze sumy a vytkneme jedno n před celou sumu. Dostáváme tak

$$\begin{aligned}\mathrm{E}(X)&=n\sum_{x=1}^{n}\frac{(n-1)!}{(x-1)!\,[n-1-(x-1)]!}p^x(1-p)^{n-1-(x-1)}=np\sum_{y=0}^{n-1}\binom{n-1}{y}p^y(1-p)^{n-1-y}=\\&=np\,(p+(1-p))^{n-1}=np\end{aligned}\tag{0.3}$$

Analogickým způsobem čtenář jistě snadno ověří, že je v pořádku normalizační podmínka $\sum_{x=0}^{n}\mathscr{P}[X=x]=1$ a zároveň, že

$$\mathrm{D}(X)=np(1-p).$$

Vidíme tedy, že alternativní binomické rozdělení je speciálním připadem binomického, pro $n=1$.

- **Poissonovo rozdělení** $\mathrm{Po}(\lambda)$, kde $\lambda>0$, popisuje řídké jevy, nebo jevy, které se žádným způsobem neovlivňují. Výběrový prostor má nyní spočetně mnoho prvků, $\Omega=\mathbf{N}_0=\{0,1,2,3,\dots\}$. Předpis pravděpodobnostní funkce vypadá následovně:

$$\mathscr{P}[X=x]=\frac{\lambda^x}{x!}\mathrm{e}^{-\lambda}$$

Normovací podmínka je triviálně splněna, neboť

$$\sum_{x=0}^{\infty}\mathscr{P}[X=x]=\sum_{x=0}^{\infty}\frac{\lambda^x}{x!}\mathrm{e}^{-\lambda}=\mathrm{e}^{-\lambda}\mathrm{e}^{\lambda}=1,$$

kde byl využit Maclaurinův rozvoj funkce e^y. Analogicky spočtěme první a druhý moment náhodné veličiny X:

$$\mathrm{E}(X)=\sum_{x=1}^{\infty}\frac{\lambda^x}{x!}x\mathrm{e}^{-\lambda}=\mathrm{e}^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^x}{(x-1)!}=\mathrm{e}^{-\lambda}\sum_{y=0}^{\infty}\frac{\lambda^{y+1}}{y!}=\lambda\mathrm{e}^{-\lambda}\sum_{y=0}^{\infty}\frac{\lambda^y}{y!}=\lambda\mathrm{e}^{-\lambda}\mathrm{e}^{\lambda}=\lambda,\tag{0.4}$$

$$
\begin{aligned}
\mathrm{E}(X^2) &= \sum_{x=1}^{\infty} \frac{\lambda^x}{x!} x^2 \mathrm{e}^{-\lambda} = \mathrm{e}^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!} x = \mathrm{e}^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^{y+1}}{y!}(y+1) = \mathrm{e}^{-\lambda}\lambda \sum_{y=1}^{\infty} \frac{\lambda^y}{(y-1)!} + \mathrm{e}^{-\lambda}\lambda \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = \\
&= \lambda \mathrm{e}^{-\lambda} \sum_{z=0}^{\infty} \frac{\lambda^{z+1}}{z!} + \mathrm{e}^{-\lambda}\lambda \mathrm{e}^{\lambda} = \lambda^2 \mathrm{e}^{-\lambda} \sum_{z=0}^{\infty} \frac{\lambda^z}{z!} + \lambda = \lambda^2 + \lambda.
\end{aligned}
\tag{0.5}
$$

Nyní snadno spočteme rozptyl:

$$\mathrm{D}(x) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Pokud dále zadejifujeme **Index disperze** jako

$$\mathrm{Id}(X) = \frac{\mathrm{D}(X)}{\mathrm{E}(X)},$$

potom vidíme, že $\forall \mathrm{Po}(\lambda)$, platí:

$$\mathrm{Id}(X) = 1, \quad \forall \lambda \in \mathbf{R}.$$

Nyní uvedeme příklad na diskrétní náhodnou veličinu. Představme si, že jedeme tramvají v dnešní době covidu. Pravděpodobnost, že s námi jede nakažený člověk, je $p = 2\%$. Počet osob v tramvaji je $n \in \mathbf{N}$. Jaká je pravděpodobnost, že nejedeme s nakaženým člověkem?

Označíme si tedy jako náhodnou veličinu X setkání s k nakaženými osobami. Potom

$$X \sim \mathrm{Bi}(20, 0.02).$$

Chceme tedy $\mathscr{P}[X = 0]$, aplikujme tedy vzorec pro binomické rozdělení.

$$\mathscr{P}[X = 0] = \binom{20}{0} p^0 (1-p)^{n-0} = 0.98^{20}.$$

Promysleme si ještě jednou cestu k abstraktní pravděpodobnosti. Klasická funkce, se kterou jsme zvyklí pracovat, je zobrazení

$$f : \langle a, b \rangle \to \mathbf{R},$$

tzn. vyberu jedno $x \in \langle a, b \rangle$ a vyčíslím hodnotu $g(x)$. Pravděpodobnost je zobrazení

$$\mathscr{P} : \mathcal{A} \subset 2^{\Omega} \to \langle 0, 1 \rangle.$$

Vybereme tedy jednu množinu $A \in \mathcal{A}$ a vyčíslíme hodnotu pravděpodobnosti $\mathcal{P}[A]$. Pravděpodonost je tedy množinová funkce, tj. její definiční obor je $\mathrm{Dom}(\mathscr{P}) = \mathcal{A}$.

0.1.2 Spojité náhodné veličiny

V diskrétní pravděpodobnosti jsme pracovali s výběrovým prostorem Ω, kde

$$\mathrm{card}(\Omega) \leqslant \aleph_0.$$

Šlo tedy o konečně nebo spočetně mnoho elementárních jevů. Totiž obecně počet prvků množiny přirozených čísel $\mathbf{N}$ značíme

$$\mathrm{card}(\mathbf{N}) = \aleph_0.$$

Dále také platí, že $\mathrm{card}\mathbf{Z} = \mathrm{card}(\mathbf{N}) = \aleph_0$, tedy $\mathbf{Z} \sim \mathbf{N}$. Pro reálná čísla ale platí

$$\mathrm{card}(\mathbf{R}) > \mathrm{card}(\mathbf{N})$$

a označujeme

$$\mathrm{card}(\mathbf{R}) = \aleph_1.$$

Jiné označení pro $\mathrm{card}(\mathbf{R})$ může být také $\mathfrak{c}$ jako *kontinuum*.

O spojité pravděpodobnosti tedy mluvíme, pokud $\mathrm{card}(\Omega) = \mathfrak{c}$. Potom také $\mathrm{card}(\mathcal{A}) \geqslant \mathfrak{c}$. Příklad spojité náhodné veličiny X může být mzda zaměstnance. Potom

$$\mathscr{P}_X : \mathcal{A} \subset 2^{\Omega} \to \langle 0, 1 \rangle,$$

kde $\Omega = \langle 0, +\infty)$ a tedy $\mathrm{card}(\Omega) = \mathfrak{c}$. σ-algebra $\mathcal{A}$ může vypadat následovně:

$$\mathcal{A} = \{\emptyset, \langle a, b\rangle, \{a\} \cup \langle b, c\rangle, \langle 10, 6500\rangle \cap \mathbf{N}, \dots\}.$$

Základní pravděpodobnostní dotaz tedy bude opět $\mathscr{P}\left[X \in A\right] =?$, pro diskrétní veličiny bude $\mathrm{card}(A) \leqslant \aleph_0$, pro spojité potom $\mathrm{card}(A) = \mathfrak{c}$.

Základní charakteristiky spojité náhodné veličiny X

- **Distribuční funkci** jsme definovali již dříve pro diskrétní veličiny, zde tedy analogicky:

$$F(x) = \mathscr{P}\left[X < x\right] = \mathscr{P}\left[X \in (-\infty, x)\right] \quad \forall x \in \mathbf{R}$$

- **Hustota pravděpodobnosti** je zobrazení $f(x) : \mathbf{R} \to \mathbf{R}$. Je to spojitá alternativa pravděpodobnostní funkce pro diskrétní náhodné veličiny a $\forall x \in \mathbf{R}$ je definována následovně:

$$F(x) = \mathscr{P}\left[X < x\right] = \int_{-\infty}^{x} f(y) \,\mathrm{d}y.$$

  Nově tedy budeme skutečnost, že náhodná veličina X má hustotu pravděpodobnosti $g(x)$, neboli X je rozdělena podle $g(x)$, neboli má distribuci $g(x)$, značit symbolem

$$X \sim g(x).$$

  Nyní uvedeme vlastnosti hustoty pravděpodobnosti:

  – $f(x) \geqslant 0$ všude v $\mathbf{R}$.
  – $f(x)$ je integrabilní na $\mathbf{R}$, tj. $f(x) \in \mathcal{R}(\mathbf{R})$, kde $\mathcal{R}(M)$ je množina všech riemannovsky integrabilních funkcí na množině M, tj. $g(x) \in \mathcal{R}(\langle 0, 1\rangle) \Leftrightarrow \int_0^1 g(x) \,\mathrm{d}x \in \mathbf{R}$.
  – $\int_{\mathbf{R}} f(y) \,\mathrm{d}y \overset{!}{=} 1$, neboť jsme požadovali, aby $\mathscr{P}\left[X \in (-\infty, +\infty)\right] = 1$.
  – $f(x) \in \mathcal{PC}(\mathbf{R})$, tzn. hustota pravděpodobnosti má jen konečně mnoho bodů nespojitosti. Symbol $\mathcal{PC}$ znamená v angličtině *piecewise continuous*.

  Příkladem hustoty pravděpodobnosti může být funkce

$$g(x) = \Theta(x) C \frac{1}{\sqrt{x}} e^{-3x},$$

  kde $\Theta(x)$ je Heavisideova funkce a C je normovací konstanta taková, aby byla splněna podmínka $\int_{\mathbf{R}} g(x) \,\mathrm{d}x = 1$.

  Podobně zkoumejme vlastnosti distribuční funkce:

  – $F(x)$ je nezáporná.
  – $F(x) \leqslant 1$.
  – $F(x)$ je neklesající. Tato vlastnost vyplývá z Kolmogorových axiomů pravděpodobnosti, neboť:

$$x_1 < x_2 \Rightarrow F(x_1) = \mathscr{P}\left[X \in (-\infty, x_1)\right] \leqslant \mathscr{P}\left[X \in (-\infty, x_2)\right] = F(x_2)$$

  – $\lim_{x \to +\infty} F(x) = 1$, protože platí normovací podmínka $\mathscr{P}\left[X \in \mathbf{R}\right] = 1$ a tedy

$$\lim_{x \to +\infty} F(x) = \lim_{x \to +\infty} \mathscr{P}\left[X \in (-\infty, x)\right] = \mathscr{P}\left[X \in \mathbf{R}\right] = 1.$$

  – Jelikož opět platí, že $\mathscr{P}\left[X \in \emptyset\right] = 0$, pak analogicky předchozímu bodu dospějeme k tomu, že

$$\lim_{x \to -\infty} F(x) = 0$$

  – $F(x) \in \mathcal{PC}(\mathbf{R})$.

– Základní vztah mezi hustotou pravděpodobnosti a distribuční funkcí je

$$F'(x) = f(x).$$

Tento vztah snadno dokážeme. Vyjdeme z definičního vztahu

$$F(x) = \int_{-\infty}^{x} f(y)\ \mathtt{d}y$$

a tuto rovnost zderivujeme všude, kde to lze. Označme navíc $\frac{\mathtt{d}\tilde{F}}{\mathtt{d}y} = f(y)$. Dostáváme

$$F'(x) = \frac{\mathtt{d}}{\mathtt{d}x}\left[\tilde{F}(y)\right]_{-\infty}^{x},$$

$$F'(x) = \frac{\mathtt{d}}{\mathtt{d}x}\left[\tilde{F}(x) - \tilde{F}(-\infty)\right],$$

pokud nyní využijeme vlastnost distribuční funkce $F(-\infty) = 0$, obdržíme vztah

$$F'(x) = \frac{\mathtt{d}}{\mathtt{d}x}\tilde{F}(x),$$

$$F'(x) = \tilde{F}'(x),$$

dostali jsme opravdu $F'(x) = f(x)$. Tuto vlastnost můžeme uzavřít tak, že pokud je $F(x) \in \mathcal{C}^1(\mathbf{R})$, potom vztah $F'(x) = f(x)$ platí všude na $\mathbf{R}$.

Typický dotaz, který si budeme často pokládat, bude

$$\mathscr{P}\left[X \in \langle\alpha, \beta)\right].$$

Podívejme se tedy, jak tento dotaz obecně můžeme zodpovědět.

$$\mathscr{P}\left[X \in \langle\alpha, \beta)\right] = \mathscr{P}\left[X < \beta\right] - \mathscr{P}\left[X < \alpha\right] = F(\beta) - F(\alpha) = \int_{-\infty}^{\beta} f(y)\ \mathtt{d}y - \int_{-\infty}^{\alpha} f(y)\ \mathtt{d}y = \int_{\alpha}^{\beta} f(y)\ \mathtt{d}y$$

S jistotou můžeme říci, že

$$\forall A \in \mathcal{A} \quad \mathscr{P}\left[X \in A\right] = \int_A f(y)\ \mathtt{d}y.$$

Deskriptivní charakteristiky spojitých náhodných veličin

- **Střední hodnotu** náhodné veličiny $X \sim g(x)$ definujeme jako

$$\mathrm{E}(X) = \int_{\mathbf{R}} x g(x)\ \mathtt{d}x,$$

za předpokladu, že $xg(x) \in \mathcal{R}(\mathbf{R})$. (Uvažujeme-li Lebesgueův integrál, potom předpokládáme $xg(x) \in \mathcal{L}(\mathbf{R})$.)

Jako příklad uveďme náhodnou veličinu X, která představuje pravidelnou kostku a nechť $\Omega = \langle 0, 8)$. Jde tedy o rovnoměrně rozdělenou náhodnou veličinu, tedy $g(x) = A$, kde $A \in \mathbf{R}$. Z podmínky normalizace snadno zjistíme hodnotu A:

$$\int_{\Omega} g(x)\ \mathtt{d}x = \int_0^8 A\ \mathtt{d}x \stackrel{!}{=} 1 \Rightarrow A = \frac{1}{8}$$

Hustota pravděpodobnosti musí mít vždy definiční obor celé $\mathbf{R}$, takže její správný předpis je

$$g(x) = \begin{cases} \frac{1}{8}, \text{ pro } x \in \langle 0, 8) \\ 0, \text{ pro } x \notin \langle 0, 8). \end{cases}$$

To také znamená, že $\operatorname{supp}(g) = \langle 0, 8)$. Spočteme střední hodnotu takovéto náhodné veličiny:

$$\mathrm{E}(X) = \int_0^8 x g(x)\ \mathtt{d}x = \int_0^8 x\frac{1}{8}\ \mathtt{d}x = \frac{1}{8}\left[\frac{x^2}{2}\right]_0^8 = 4$$

- **Rozptyl** náhodné veličiny X definujeme jako

$$\mathrm{D}(X) = \mathrm{E}(X - \mathrm{E}(X))^2,$$

pokud $\mathrm{E}(X) \triangleq \lambda \in \mathbf{R}$. Pokud je hodnota $\mathrm{D}(X)$ malá, znamená to, že většina realizací X padne poblíž střední hodnoty λ. Pokud $\mathrm{D}(X) = 0$, potom všechny realizace X jsou rovny λ a X je v tomto případě deterministická náhodná veličina. Pro rozptyl spojitých veličin bude platit stejný vztah jako pro rozptyl diskrétních veličin:

$$\begin{aligned} \mathrm{D}(X) = \mathrm{E}(X-\lambda)^2 &= \int_{\mathbf{R}} (x-\lambda)^2 g(x)\ \mathtt{d}x = \int_{\mathbf{R}} x^2 g(x)\ \mathtt{d}x - 2\lambda \int_{\mathbf{R}} x g(x)\ \mathtt{d}x + \lambda^2 \int_{\mathbf{R}} g(x)\ \mathtt{d}x = \\ &= \int_{\mathbf{R}} x^2 g(x)\ \mathtt{d}x - \lambda^2 = \mathrm{E}(X^2) - \mathrm{E}(X). \end{aligned} \tag{0.6}$$

Pokud tedy existuje $\mathrm{E}(X^2)$, může být i $+\infty$, a $\mathrm{E}(X) \in \mathbf{R}$, potom $\mathrm{D}(X)$ existuje, případně je to opět $+\infty$.

- **K-tý moment** pro $k \in \mathbf{N}_0$ definujeme jako

$$\mathrm{E}(X^k) = \int_{\mathbf{R}} x^k g(x)\ \mathtt{d}x,$$

pokud $x^k g(x) \in \mathcal{R}(\mathbf{R})$. Vidíme, že $\mathrm{E}(X^0) = 1$ z normalizace, dále $\mathrm{E}(X^1) = \lambda$ a pro druhý moment platí $\mathrm{E}(X^2) = \mathrm{D}(X) + \lambda^2$.

Učebnicová rozdělení spojitých náhodných veličin

- **Uniformě rozdělená náhodná veličina** $X \sim \mathrm{U}(a, b)$, kde $-\infty < a < b < +\infty$ je taková náhodná veličina, jejíž hustota pravděpodobnosti je definována vztahem

$$g(x) = \begin{cases} \frac{1}{b-a}, & \text{pro } x \in \langle a, b) \\ 0, & \text{pro } x \notin \langle a, b). \end{cases}$$

Snadno zjistíme i tvar distribuční funkce. Pro $x \in \langle a, b)$ máme

$$F(x) = \mathscr{P}[X < x] = \int_{-\infty}^{x} g(y)\ \mathtt{d}y = \int_{a}^{x} g(y)\ \mathtt{d}y = \frac{1}{b-a}(x-a),$$

pro $x \in (-\infty, a)$, jelikož $g(y)$ je na tomto intervalu nulová,

$$F(x) = \int_{-\infty}^{a} g(y)\ \mathtt{d}y = 0$$

a nakonec pro $x \in \langle b, +\infty)$ dostáváme

$$F(x) = \int_{-\infty}^{x} g(y)\ \mathtt{d}y = \int_{a}^{x} g(y)\ \mathtt{d}y = \int_{a}^{b} g(y)\ \mathtt{d}y + \int_{b}^{x} g(y)\ \mathtt{d}y = 1,$$

kde jsme v předposlední rovnosti využili v prvním členu axiom normalizace, v druhém členu skutečnost, že $g(y)$ je mimo integral $\langle a, b)$ nulová. Pro $\forall A \subset \mathcal{A}$, kde $\mathcal{A} \subset 2^{\Omega}$, nám bude platit, že

$$\mathscr{P}[X \in A] = \int_{A} g(y)\ \mathtt{d}y.$$

Pokud ale $\mathcal{A} = 2^{\Omega}$, potom $\forall A \subset \Omega$ platí $\mathscr{P}[X \in A] = \int_A g(y)\ \mathtt{d}y$. Vezměme si množinu $B = \langle 0, 1) \cap \mathbf{Q}$. Potom

$$\mathscr{P}[X \in \langle 0, 1) \cap \mathbf{Q}] = \int_{B} \frac{1}{b-a}\ \mathtt{d}x$$

riemannovsky neexistuje! To je ale právě obtíž riemannova integálu. Jistě bychom chtěli znát pravděpodobnost, že náhodná veličina padne právě do množiny B, neboť se to zdá jako smysluplný dotaz. Abychom tento problém odstranili, zavádíme na přednáškách MAB4 právě již zmiňovaný Lebesgueův integrál. Množina B totiž bude lebesgueovsky měřitelná a bude platit, že

$$\mathscr{P}[X \in B] = 0.$$

- **Exponenciálně rozdělená náhodná veličina** $X \sim \mathrm{Exp}(\lambda)$, $\lambda > 0$, má hustotu pravděpodobnosti definovanou předpisem
$$g(x) = \Theta(x)\lambda \mathrm{e}^{-\lambda x}, \tag{0.7}$$
kde $\Theta(x)$ je Heavisideova funkce. Poznamenejme, že součin $\Theta(x)h(x)$ se klade za nulový, i pokud $h(x)$ není na $(-\infty, 0\rangle$ definována. Tedy např.
$$\Theta(x)\frac{1}{\sqrt{x}} = \begin{cases} \frac{1}{\sqrt{x}}, & \text{pro } x > 0 \\ 0, & \text{pro } x \leqslant 0 \end{cases}$$
Hustota pravděpodobsnoti zadaná vztahem (0.7) je již správně normovaná, neboť
$$\int_{\mathbf{R}} g(x)\ \mathrm{d}x = \int_{\mathbf{R}} \lambda\Theta(x)\mathrm{e}^{-\lambda x} = \int_0^{+\infty} \lambda \mathrm{e}^{-\lambda x} = \left[-\mathrm{e}^{-\lambda x}\right]_0^{+\infty} = 0.$$
Nyní odvodíme předpis pro distribuční funkci. Na intervalu $(-\infty, 0\rangle$ bude nulová. Pro $x > 0$ počítejme:
$$F(x) = \int_{-\infty}^{x} \Theta(x)\lambda \mathrm{e}^{-\lambda y}\ \mathrm{d}y = \int_0^x \lambda \mathrm{e}^{-\lambda y}\ \mathrm{d}y = \left[-\mathrm{e}^{-\lambda y}\right]_0^x = 1 - \mathrm{e}^{-\lambda x}.$$
Opravdu tedy $\lim_{x\to+\infty} F(x) = 1$ $\lim_{x\to-\infty} F(x) = 0$. Střední hodnota bude
$$\mathrm{E}(X) = \int_{\mathbf{R}} xg(x)\ \mathrm{d}x = \int_0^{+\infty} x\lambda \mathrm{e}^{-\lambda x}\ \mathrm{d}x = \left[-x\mathrm{e}^{-\lambda x}\right]_0^{+\infty} + \int_0^{+\infty} \mathrm{e}^{-\lambda x}\ \mathrm{d}x = \left[-\frac{1}{\lambda}\mathrm{e}^{-\lambda x}\right]_0^{+\infty} = \frac{1}{\lambda}$$
Analogickým způsobem, opět za použití metody per partes, spočteme, že $\mathrm{E}(X^2) = \frac{2}{\lambda^2}$ a $DX(X) = \frac{1}{\lambda^2}$

- **Erlangovsky rozdělená** náhodná veličina $X \sim \mathrm{Erl}(n, \lambda)$, kde pro parametry platí $\lambda > 0$ a $n \in \mathbf{N}$. Toto rozdělení popisuje časový interval mezi příchody $n+1$ událostí, např. příchod zákazníka. Hustota pravděpodobnosti je
$$g(x) = Ax^n \mathrm{e}^{-\lambda x}.$$
Vypočítejme normalizační konstantu A. Chceme, aby $\int_{\mathbf{R}} g(x)\ \mathrm{d}x \overset{!}{=} 1$:
$$\int_{\mathbf{R}} g(x)\ \mathrm{d}x = A\int_0^{+\infty} x^n \mathrm{e}^{-\lambda x}\ \mathrm{d}x = A\frac{n!}{\lambda^{n+1}} \overset{!}{=} 1,$$
takže $A = \frac{\lambda^{n+1}}{n!}$. Využili jsme známý vzorec
$$\int_0^{+\infty} x^n \mathrm{e}^{-\lambda x} = \frac{n!}{\lambda^{n+1}},$$
jehož platnost jistě čtenář sám snadno ověří. Hustota pravděpodobnosti má po znormování tvar
$$g(x) = \Theta(x)\frac{\lambda^{n+1}}{n!}x^n \mathrm{e}^{-\lambda x}.$$
Vypočteme nyní střední hodnotu:
$$\mathrm{E}(X) = \int_{\mathbf{R}} xg(x)\ \mathrm{d}x = \frac{\lambda^{n+1}}{n!}\int_0^{+\infty} x^{n+1}\mathrm{e}^{-\lambda x}\ \mathrm{d}x = \frac{\lambda^{n+1}}{n!}\frac{(n+1)!}{\lambda^{n+2}} = \frac{n+1}{\lambda}$$
O škálované náhodné veličině mluvíme právě tehdy, když
$$\mathrm{E}(X) \overset{!}{=} 1,$$
pro náhodnou veličinu $X \sim \mathrm{Erl}(n, \lambda)$ musí platit $\lambda \overset{!}{=} n+1$, škálovaná hustota pak bude mít předpis
$$g(x) = \Theta(x)\frac{(n+1)^{n+1}}{n!}x^n \mathrm{e}^{-(n+1)x}.$$
Analogicky vypočítáme rozptyl:
$$\begin{aligned} \mathrm{D}(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) &= \frac{\lambda^{n+1}}{n!}\int_0^{+\infty} x^{n+2}\mathrm{e}^{-\lambda x}\ \mathrm{d}x - \left(\frac{n+1}{\lambda}\right)^2 = \frac{(n+2)!}{\lambda^{n+3}}\frac{\lambda^{n+1}}{n!} - \frac{(n+1)^2}{\lambda^2} = \\ &= \frac{1}{\lambda^2}\left((n+2)-(n+1)\right)(n+1) = \frac{n+1}{\lambda^2} \end{aligned} \tag{0.8}$$
Ve škálované variantě $\mathrm{D}(X) = \frac{n+1}{(n+1)^2} = \frac{1}{n+1}$. Rozptyl klesá s n, maximální rozptyl tím pádem bude pro $n = 1$: $\mathrm{D}(X) = \frac{1}{2}$.

- **Gamma rozdělení** náhodné veličiny $X \sim \Gamma(\alpha, \lambda)$, s parametry $\alpha > -1$ a $\lambda > 0$, má hustotu pravděpodobnosti tvaru
$$g(x) = \Theta(x) A x^{\alpha} \mathrm{e}^{-\lambda x}.$$
Ihned vidíme, že pokud $\alpha \in \mathbf{N}$, pak $\Gamma(\alpha, \lambda) = \mathrm{Erl}(\alpha, \lambda)$, pro $\alpha = 0$ je $\Gamma(\alpha, \lambda) = \mathrm{Exp}(\lambda)$. Pokud $\alpha \in (0, +\infty) \backslash \mathbf{N}_0$, jde o tzv. ryzí Gamma rozdělení. Gamma funkce je definována jako
$$\Gamma(x) = \int_0^{+\infty} y^{x-1} \mathrm{e}^{-y} \, \mathrm{d}y, \text{ pro } x \in (0, +\infty),$$
tzn. $\mathrm{Dom}(\Gamma) = \mathbf{R}^+$. Připomeňme si nyní některé vlastnosti Gamma funkce. Triviálně platí:
$$\Gamma(1) = \int_0^{+\infty} \mathrm{e}^{-y} \, \mathrm{d}y = 1.$$
Dále, jelikož pro $n \in \mathbf{N}$ platí vzorec
$$\int_0^{+\infty} x^n \mathrm{e}^{-\lambda x} \, \mathrm{d}x = \frac{n!}{\lambda^{n+1}},$$
potom vidíme, že pokud vezmeme $\lambda = 1$, tak Gamma funkce je jakési spojité rozšíření faktoriálu, neboť
$$\Gamma(n) = \int_0^{+\infty} y^{n-1} \mathrm{e}^{-y} \, \mathrm{d}y = (n-1)!.$$
Pokud budeme uvažovat $x \in \mathbf{R}^+$, potom upravujme:
$$\Gamma(x+1) = \int_0^{+\infty} y^x \mathrm{e}^{-x} \, \mathrm{d}x = \left[-y^x \mathrm{e}^{-x}\right]_0^{+\infty} + x \int_0^{+\infty} y^{x-1} \mathrm{e}^{-y} \, \mathrm{d}y = x \int_0^{+\infty} y^{x-1} \mathrm{e}^{-y} \, \mathrm{d}y = x \cdot \Gamma(x),$$
Pro $x = n \in \mathbf{N}$ víme, že platí $\Gamma(n+1) = n!$, dostaneme za pomoci práce s faktoriály tentýž vztah:
$$\Gamma(n+1) = n! = n(n-1)! = n\Gamma(n).$$
Vezměme nyní hustotu Gamma rozdělení
$$g(x) = A\Theta(x) x^{\alpha} \mathrm{e}^{-\lambda x}$$
a určeme normovací konstantu A:
$$\int_{\mathbf{R}} g(x) \, \mathrm{d}x = A \int_0^{+\infty} x^{\alpha} \mathrm{e}^{-\lambda x} \, \mathrm{d}x = A \int_0^{+\infty} \left(\frac{y}{\lambda}\right)^{\alpha} \mathrm{e}^{-y} \frac{1}{\lambda} \, \mathrm{d}y = \frac{A}{\lambda^{\alpha+1}} \int_0^{+\infty} y^{\alpha} \mathrm{e}^{-y} \, \mathrm{d}y = \frac{A}{\lambda^{\alpha+1}} \Gamma(\alpha+1) \stackrel{!}{=} 1,$$
v druhé rovnosti byla použita substituce $y = \lambda x$, pro $\lambda > 0$. Získali jsme konstantu $A = \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)}$ a již správně normovaná hustota pravděpodobnosti má tvar
$$g(x) = \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} \Theta(x) x^{\alpha} \mathrm{e}^{-\lambda x}, \text{ kde } \alpha \in (-1, +\infty),\ \lambda \in \mathbf{R}^+.$$
Snadno vypočteme střední hodnotu náhodné veličiny X:
$$\mathrm{E}(X) = \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} \int_0^{+\infty} x^{\alpha+1} \mathrm{e}^{-\lambda x} \, \mathrm{d}x = \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} \frac{1}{\lambda^{\alpha+2}} \Gamma(\alpha+2) = \frac{1}{\lambda} \frac{\Gamma(\alpha+2)}{\Gamma(\alpha+1)} = \frac{1}{\lambda} \frac{(\alpha+1)\Gamma(\alpha+1)}{\Gamma(\alpha+1)} = \frac{\alpha+1}{\lambda},$$
pro splnění škálovací podmínky musí platit $\lambda = \alpha + 1$. Výpočet druhého momentu a rozptylu je zcela analogický, proto uvedeme pouze výsledek: $\mathrm{E}(X^2) = \frac{(\alpha+2)(\alpha+1)}{\lambda^2}$ a $\mathrm{D}(X) = \frac{\alpha+1}{\lambda^2}$.
Podívejme se nyní na distribuční funkci Erlangova a Gamma rozdělení. Nejprve zadefinujme speciální funkci, kterou budeme nazývat **Dolní neúplná Gamma funkce** $\gamma(s, x)$:
$$\gamma(s, x) = \int_0^x y^{s-1} \mathrm{e}^{-y} \, \mathrm{d}y. \tag{0.9}$$
Její hodnoty jsou tabelovány. Pokusme se ji upravit. Využijeme Maclaurinův rozvoj funkce $f(y) = \mathrm{e}^{-y}$:
$$\gamma(s, x) = \int_0^x y^{s-1} \mathrm{e}^{-y} \, \mathrm{d}y = \int_0^x y^{s-1} \sum_{n=0}^{\infty} \frac{y^n}{n!} (-1)^n \, \mathrm{d}y.$$

Jelikož tato řada stejnoměrně konverguje na každém uzavřeném intervalu $\langle 0, x\rangle$, na $\mathbf{R}+$ už ne, můžeme zaměnit pořadí sumy a integrálu:

$$\gamma(s,x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_0^x y^{n+s-1}\ \mathrm{d}y = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left[\frac{y^{n+s}}{n+s}\right]_0^x = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} x^{n+1},$$

Pokud m bude velké, můžeme dolní neúplnou Gamma funkci aproximovat konečnou řadou:

$$\gamma(s,x) = \sum_{n=0}^{m} \frac{(-1)^n}{n!} x^{n+1}.$$

Nyní se konečně podívejme, jak bude vypadat distribuční funkce:

$$\begin{aligned} F(x) = \mathscr{P}[X<x] = \int_{-\infty}^{x} g(y)\ \mathrm{d}y = \int_{-\infty}^{x} \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)}\Theta(y) y^{\alpha} \mathrm{e}^{-\lambda y}\ \mathrm{d}y = \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} \int_0^x y^{\alpha}\mathrm{e}^{-\lambda y}\ \mathrm{d}y = \\ = \frac{\lambda^{\alpha+1}}{\Gamma(\alpha+1)} \int_0^{\lambda x} \left(\frac{z}{\lambda}\right)^{\alpha} \mathrm{e}^{-z}\frac{1}{\lambda}\ \mathrm{d}z = \frac{1}{\Gamma(\alpha+1)} \int_0^{\lambda x} z^{\alpha}\mathrm{e}^{-z}\ \mathrm{d}z = \frac{\gamma(\alpha+1,\lambda x)}{\Gamma(\alpha+1)}. \end{aligned} \tag{0.10}$$

V první rovnosti na druhém řádku byla využita substituce $z = \lambda y$. Opět bude splněno, že $\lim_{x\to+\infty} F(x) = 1$ a $\lim_{x\to 0_+} F(x) = 0$.

- **Normálně (gaussovsky) rozdělená** náhodná veličina $X \sim \mathrm{N}(\mu,\sigma)$, kde $\mu \in \mathbf{R}$ a $\sigma > 0$. Hustota pravděpodobnosti je tzv. **Gaussova funkce**

$$g(x) = A\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Gaussovo rozdělení je symetrické kolem bodu $x = \mu$. Dále platí, že $\mu = \arg\max\ g(x)$ a $\max\ g(x) = A$. Připomeňme známý Gaussův integrál, který v následujících výpočtech mnohokrát použijeme:

$$\int_{\mathbf{R}} \mathrm{e}^{-ax^2}\ \mathrm{d}x = \sqrt{\frac{\pi}{a}}.$$

Nejprve spočítejme hodnotu konstanty A:

$$A\int_{\mathbf{R}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\ \mathrm{d}x = A\sqrt{2}\sigma \int_{\mathbf{R}} \mathrm{e}^{-y^2}\ \mathrm{d}y = A\sqrt{2}\sigma\sqrt{\pi} \stackrel{!}{=} 1,$$

využili jsme substituci $y = \frac{x-\mu}{\sqrt{2}\sigma}$. Hustota pravděpodobnosti má tedy tvar:

$$g(x) = \frac{1}{\sqrt{2pi\sigma^2}}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Nyní bude následovat důležitý výpočet střední hodnoty, který vede ke známému výsledku:

$$\mathrm{E}(X) = \int_{\mathbf{R}} xg(x)\ \mathrm{d}x = \frac{1}{\sqrt{2\pi}\sigma}\int_{\mathbf{R}} (x-\mu+\mu)\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\ \mathrm{d}x = \frac{1}{\sqrt{2\pi}\sigma}\int_{\mathbf{R}} (x-\mu)\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\ \mathrm{d}x + \frac{\mu}{\sqrt{2\pi}\sigma}\int_{\mathbf{R}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\ \mathrm{d}x = 0 + \mu\cdot 1,$$

v poslední rovnosti jsme využili vlastnosti μ-liché funkce

$$\int_{\mathbf{R}} f(x)\ \mathrm{d}x = 0.$$

Pro gaussovskou náhodnou veličinu tedy platí

$$\mathrm{E}(X) = \mu.$$

Pro výpočet rozptylu budeme potřebovat znát druhý moment:

$$\begin{aligned} \mathrm{E}(X^2) = \int_{\mathbf{R}} \frac{x^2}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\ \mathrm{d}x = \int_{\mathbf{R}} \frac{x^2 - 2x\mu + \mu^2 + 2x\mu - \mu^2}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\ \mathrm{d}x = \int_{\mathbf{R}} \frac{(x-\mu)^2}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\ \mathrm{d}x + \\ + 2\mu\int_{\mathbf{R}} \frac{x}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\ \mathrm{d}x - \mu^2\int_{\mathbf{R}} \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}\ \mathrm{d}x, \end{aligned} \tag{0.11}$$

ve druhém členu rozeznáváme střední hodnotu $\mathrm{E}(X) = \mu$ a ve třetím členu využijeme to, že hustota již je správně normovaná na jedničku. Dostáváme tedy

$$\mathrm{E}(X^2) = \int_{\mathbf{R}} \frac{(x-\mu)^2}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \,\mathrm{d}x + 2\mu^2 - \mu^2 = \mu^2 + \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} y^2 \mathrm{e}^{-y^2} 2\sigma^2 \sqrt{2} \,\mathrm{d}y = \mu^2 + \frac{2\sigma^2}{\sqrt{\pi}} \int_{\mathbf{R}} y^2 \mathrm{e}^{-y^2} \,\mathrm{d}y,$$

kde jsme využili substituci $y = \frac{x-\mu}{\sqrt{2}\sigma}$. Potřebujeme určit hodnotu Gaussova integrálu $\int_{\mathbf{R}} y^2 \mathrm{e}^{-y^2} \,\mathrm{d}y$, to provedeme záměnou integrálu a derivace. Doporučujeme čtenáři, aby si ověřil předpoklady nutné k této záměně. Víme, že

$$\int_{\mathbf{R}} \mathrm{e}^{-ax^2} \,\mathrm{d}x = \sqrt{\frac{\pi}{a}}.$$

Tento vztah nyní zderivujme podle parametru a. Dostáváme rovnost

$$-\int_{\mathbf{R}} x^2 \mathrm{e}^{-ax^2} \,\mathrm{d}x = \sqrt{\pi} \left(-\frac{1}{2}\right) a^{-\frac{3}{2}}.$$

Můžeme uzavřít výpočet druhého momentu:

$$\mathrm{E}(X^2) = \mu^2 + \frac{2\sigma^2}{\sqrt{\pi}} \int_{\mathbf{R}} y^2 \mathrm{e}^{-y^2} \,\mathrm{d}y = \mu^2 + \frac{1}{\sqrt{\pi}} \frac{1}{2} \sqrt{\pi} 2\sigma^2 = \mu^2 + \sigma^2.$$

Rozptyl gaussovsky rozdělené náhodné veličiny vychází jako

$$\mathrm{D}(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = \mu^2 + \sigma^2 - \mu^2 = \sigma^2.$$

Pokud je tedy σ^2 malé, hodnoty náhodné veličiny jsou velice málo odchýlené od střední hodnoty μ, pokud je σ^2 velké, potom jsou hodnoty náhodné veličiny rozptýlené po reálné ose. Věnujme se nyní distribuční funkci $F(x)$:

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(y-\mu)^2}{2\sigma^2}} \,\mathrm{d}y.$$

Pokud zadefinujeme tzv. **Chybovou funkci** $\mathrm{erf}(x)$, anglicky *Error function*, budeme moct distribuční funkci upravit. Definujme tedy:

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-z^2} \,\mathrm{d}z.$$

Najděme nyní vhodnou aproximaci chybové funkce. Použijeme Maclaurinův rozvoj fce e^{-z^2}:

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-z^2} \,\mathrm{d}z = \frac{2}{\sqrt{\pi}} \int_0^x \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{n!} = \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left[\frac{z^{2n+1}}{2n+1}\right]_0^x = \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)n!} x^{2n+1}$$

!!!!V ZÁPISCÍCH JE PODLE MĚ TATO ČÁST UDĚLANÁ ŠPATNĚ, PROTOŽE JE TAM POUŽITA MACLAURINOVA RADA PRO $\mathrm{e}^{-}Z$, TAKŽE TO NECHÁVÁM NA TOBĚ, str. 3, dokument !!!! Pomocí substituce $z = \frac{y-\mu}{\sqrt{2}\sigma}$ upravujme:

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(y-\mu)^2}{2\sigma^2}} \,\mathrm{d}y = \frac{1}{\pi} \int_{-\infty}^{\frac{x-\mu}{\sqrt{2}\sigma}} \mathrm{e}^{-z^2} \,\mathrm{d}z = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{0} \mathrm{e}^{-z^2} \,\mathrm{d}z + \frac{1}{\pi} \int_{0}^{\frac{x-\mu}{\sqrt{2}\sigma}} \mathrm{e}^{-z^2} \,\mathrm{d}z = \frac{1}{\pi}\frac{1}{2}\sqrt{\pi} + \frac{1}{\pi}\frac{\sqrt{\pi}}{2} \mathrm{erf}\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)$$

V poslední rovnosti je využito skutečnosti, že $\int_{-\infty}^{0} \mathrm{e}^{-z^2} \,\mathrm{d}z = \frac{1}{2} \int_{-\infty}^{+\infty} \mathrm{e}^{-z^2} \,\mathrm{d}z = \frac{1}{2}\sqrt{\pi}$. Celkem dostáváme

$$F(x) = \frac{1}{2} + \frac{1}{2} \mathrm{erf}\left(\frac{x-\mu}{\sqrt{2}\sigma}\right).$$

Hodnoty chybové funkce jsou tabelovány.

Další deskriptivní charakteristiky

- **Nosič hustoty pravděpodobnosti**, značíme supp, z anglického *support*, definujeme jako

$$\mathrm{supp}(g) = \{x \in \mathbf{R} : g(x) \neq 0\} = \{x \in \mathbf{R} : g(x) > 0\}.$$

- Hustotu $g(x)$ nazveme **s pozitivním nosičem**, právě tehdy, když
$$\text{supp}(g) \subset (0, +\infty),$$
tj. X nabývá pouze nezáporných hodnot. Uvažujme $a < 0$. Potom
$$F(a) = \mathscr{P}\left[X < a\right] = \int_{-\infty}^{a} g(x)\ \mathtt{d}x = \int_{-\infty}^{a} 0\ \mathtt{d}x = 0,$$
neboť $g(x)$ je nulová všude na $(-\infty, 0\rangle$.

- **Levé a pravé pomezí**, z angličtiny *ambit*, definujeme jako
$$\text{amb}_L = \inf \text{supp}(g),$$
$$\text{amb}_R = \sup \text{supp}(g).$$

- **Rozpětí nosiče**, anglicky *margin*, je definováno rozdílem
$$\text{marg}(g) = \text{amb}_R - \text{amb}_L.$$
Náhodná veličina vždy nabývá hodnot pouze z $\text{supp}(g)$, tj. pouze na množině délky $\text{marg}(g)$.

- Řekneme, že hustota pravděpodobnosti je **neposunutá**, jestliže
$$\text{amb}_L = 0.$$
Příkladem takových hustot, jsou třeba hustoty příslušející Exponenciálnímu, Erlangovu, nebo Gamma rozdělení.

- **Modus** náhodné veličiny definujeme jako
$$\arg\max g(x),$$
tj. jsou to body (nebo jeden bod), ve kterých $g(x)$ nabývá svého maxima.

- Pro spojitou a rostoucí distribuční funkci $F(x)$, **mediánem** rozumíme číslo $\tilde{x} \in \mathbf{R}$ takové, že 50 % hodnot je pod mediánem a 50 % hodnot je nad ním. Tedy
$$\tilde{x} = F^{-1}\left(\frac{1}{2}\right).$$
Obecně medián je kterékoliv číslo $\tilde{x} \in \mathbf{R}$, které splňuje
$$\mathscr{P}\left[X \leqslant \tilde{x}\right] \geqslant \frac{1}{2} \quad \wedge \quad \mathscr{P}\left[X \geqslant \tilde{x}\right] \leqslant \frac{1}{2}.$$

- Hustotu pravděpodobnosti nazýváme **unimodální**, právě tehdy, když
$$\exists c \in \mathbf{R} \text{ takové, že } f(x) \text{ je neklesající na } (-\infty, c) \text{ a zároveň nerostoucí na } (c, +\infty).$$

- Dále definujeme **kvantil**, nebo také α-kvantil $q(\alpha) \in \mathbf{R}$ jako
$$q(\alpha) = \inf\{x \in \mathbf{R} | F(x) \geqslant \alpha\}.$$
Tato podmínka jinými slovy znamená, že
$$\mathscr{P}\left[X \leqslant q(\alpha)\right] \geqslant \alpha \quad \wedge \quad \mathscr{P}\left[X \geqslant q(\alpha)\right] \geqslant 1 - \alpha.$$
To nám říká, že X nedosáhne hodnoty $q(\alpha)$ s pravděpodobsností $100 \cdot \alpha\ \%$ a X přesáhne hodnotu $q(\alpha)$ s pravděpodobností $100 \cdot (1 - \alpha)\ \%$. Z předchozích definic vidíme, že pro $\alpha = \frac{1}{2}$ se jedná o medián.

0.2 Problémy se zavedením pravděpodobnosti

Skutečnost, že Riemannův integrál je při zavádění pravděpodobnosti nevyhovující, a proto se buduje na přednáškách z MAB4 tzv. Lebesgueův integrál, jsme naznačili již dříve. Nyní se na tento problém podívejme obecněji. Uvažujme nyní náhodnou veličinu $X \sim \mathrm{U}(0,1)$ a spočtěme si pravděpodobnosti, že náhodná veličina X padne do různých množin.

$$
\begin{aligned}
\mathscr{P}\left[X < -3\right] &= \int_{-\infty}^{-3} g(x)\ \mathtt{d}x = \int_{-\infty}^{-3} 0\ \mathtt{d}x = 0,\\
\mathscr{P}\left[X < \frac{1}{2}\right] &= \int_{-\infty}^{\frac{1}{2}} g(x)\ \mathtt{d}x = \int_{-\infty}^{\frac{1}{2}} 1\ \mathtt{d}x = \frac{1}{2},\\
\mathscr{P}\left[X > \frac{1}{2}\right] &= \int_{\frac{1}{2}}^{+\infty} g(x)\ \mathtt{d}x = \int_{\frac{1}{2}}^{1} 1\ \mathtt{d}x = \frac{1}{2},\\
\mathscr{P}\left[X \in \mathbf{R}\right] &= \int_{\mathbf{R}} g(x)\ \mathtt{d}x = \int_{0}^{1} 1\ \mathtt{d}x = 1,\\
\mathscr{P}\left[X \in \emptyset\right] &= \int_{\emptyset} g(x)\ \mathtt{d}x = 0,
\end{aligned}
$$

v tomto případě výsledek vychází přímo z definice $\int_\emptyset f(x)\ \mathtt{d}x \overset{!}{=} 0$. Dále

$$
\mathscr{P}\left[X \in \langle 0,1)\right] = \int_0^1 1\ \mathtt{d}x = 1,
$$

problém ale nastane v následujícím případě:

$$
\mathscr{P}\left[X \in \langle 0,1) \cap \mathbf{Q}\right] = \int_{\langle 0,1)\cap\mathbf{Q}} 1\ \mathtt{d}x = \int_0^1 \mathcal{D}(x)\ \mathtt{d}x,
$$

kde $\mathcal{D}(x)$ představuje Dirichletovu funkce

$$
\mathcal{D}(x) = \begin{cases} 1, & \text{pro } x \in \mathbf{Q} \\ 0, & \text{pro } x \in \mathbf{R} \setminus \mathbf{Q}. \end{cases} \tag{0.12}
$$

Newtonův intergál můžeme rovnou opustit, protože $\mathcal{D}(x)$ je nespojitá. Zkusme spočítat $\int_0^1 \mathcal{D}(x)\ \mathtt{d}x$ riemannovsky. Uvažujme dělení, nyní ho označíme $\mathcal{A}$, intervalu $\langle 0,1)$ na intervaly $I_0 = \langle 0, x_0\rangle$, $I_1 = \langle x_0, x_1\rangle$, $\ldots$, $I_{n-1} = \langle x_{n-1}, 1\rangle$. V intervalu I_j $\forall j \in \{0, n-1\}$ označme

$$
v_j = \inf \mathcal{D}(x) \text{ a } V_j = \sup \mathcal{D}(x).
$$

Z předpisu Dirichletovy funkce (0.12) je jasné, že

$$
v_j = 0 \text{ a } V_j = 1 \quad \forall j \in \{0, n-1\}.
$$

Nyní vypočítáme dolní integrální součet:

$$
\mathrm{L}(\mathcal{A}, \mathcal{D}(x)) = \sum_{k=0}^{n-1} v_k (x_k - x_{k-1}) = 0,
$$

horní integrální součet:

$$
\mathrm{U}(\mathcal{A}, \mathcal{D}(x)) = \sum_{k=0}^{n-1} V_k (x_k - x_{k-1}) = 1,
$$

Dolní Riemannův integrál je:

$$
\mathrm{L}(\mathcal{D}(x)) = \sup_{\mathcal{A}} \mathrm{L}(\mathcal{A}, \mathcal{D}(x)) = 0
$$

a horní Riemannův integrál vychází jako:

$$
\mathrm{U}(\mathcal{D}(x)) = \inf_{\mathcal{A}} \mathrm{U}(\mathcal{A}, \mathcal{D}(x)) = 1.
$$

Protože ale
$$\mathrm{L}(\mathcal{D}(x)) \neq \mathrm{U}(\mathcal{D}(x)),$$
tak $\mathcal{D}(x) \notin \mathcal{R}(\langle 0,1\rangle)$, tj. $\mathcal{D}(x)$ není riemannovsky integrabilní a $\int_0^1 \mathcal{D}(x)\ \mathrm{d}x$ neexistuje. Pravděpodobnostní dotaz $\mathscr{P}[X \in \langle 0,1) \cap \mathbf{Q}]$ nemůže být zodpovězen!

Nyní uvažujme náhodnou veličinu X, která vyjadřuje náhodné losování čísel z intervalu $\langle 0,1)$ s rovnoměrným rozdělením. Budeme sledovat, kolik pokusů spadlo do $\langle 0,1) \cap \mathbf{Q}$ a kolik ne. Z teorie víme, že $\mathbf{Q} \sim \mathbf{N}$, tedy $\mathbf{Q}$ je spočetná množina. Z toho plyne, že čísel v $\langle 0,1) \cap \mathbf{Q}$ je spočetně mnoho. Oproti tomu, čísel v $\langle 0,1) \cap (\mathbf{Q} \cap \mathbf{R})$ je nespočetně mnoho, neboť $\langle 0,1) \cap (\mathbf{Q} \cap \mathbf{R}) \sim \mathbf{R}$. Můžeme tedy prohlásit, že iracionálních čísel v $\langle 0,1)$ je stejně jako všech čísel v $\langle 0,1)$, neboli skoro všechna čísla v $\langle 0,1)$ jsou iracionální. Pokud tedy půjde počet pokusů $M \to +\infty$, potom můžeme říct, že
$$\mathscr{P}[X \in \langle 0,1) \cap \mathbf{Q}] = 0 \text{ nebo } \mathscr{P}[X \in \langle 0,1) \cap (\mathbf{Q} \cup \mathbf{R})] = 1.$$
I přes tuto analýzu nám vyšlo, že dotaz $\mathscr{P}[X \in \langle 0,1) \cap \mathbf{Q}]$ nemůže být zodpovězen. Riemannův integrál je prostě nedostačující. V teorii Lebesgueova integrálu tento dotaz bude triviálně zodpovězen a dostaneme $\mathscr{P}[X \in \langle 0,1) \cap \mathbf{Q}] = 0$, neboť množina $\langle 0,1) \cap \mathbf{Q}$ má nulovou míru.

0.3 Matematické pozadí teorie pravděpodobnosti

Pravděpodobnost je zobrazení $\mathscr{P} : \mathcal{A} \to \langle 0,1\rangle$, definiční obor je množina množin $\mathrm{Dom}(\mathscr{P}) = \mathcal{A}$ a obor hodnot je $\mathrm{Ran}(\mathscr{P}) \subset \langle 0,1\rangle$. Pravděpodobnost $\mathscr{P}$ je tedy tzv. množinová funkce. $\mathcal{A}$ můžeme nazvat soustavou událostí. Dříve jsme už diskutovali, že základní pravděpodobností dotaz bude
$$\mathscr{P}[X \in B] =?, \text{ kde } B \in \mathcal{A} \subset 2^{\Omega}, \text{ tj. } B \subset \Omega.$$
Prvky Ω se nazývají elementární. Pokud $\mathrm{card}(\Omega) = n \in \mathbf{N}$, jde o diskrétní pravděpodobnost. V tomto případě zpravidla $\mathcal{A} = \{B : B \subset \Omega\} = 2^{\Omega}$. Pokud $\mathrm{card}(\Omega) = \aleph_0 \equiv \mathrm{card}(\mathbf{N})$, jde o tzv. spočetně diskrétní pravděpodobnost a také platí $\mathcal{A} = \{B : B \subset \Omega\} = 2^{\Omega}$. Nakonec pro $\mathrm{card}(\Omega) = \mathfrak{c} \equiv \mathrm{card}(\mathbf{R})$ mluvíme o spojité pravděpodobnosti. V posledním případě bude $\mathcal{A} \subset \{B : B \subset \Omega\} = 2^{\Omega}$.

Jak už bylo demonstrováno v minulé kapitole, pro všechny množiny $Y \in \mathcal{A} = 2^{\Omega}$, mluvíme o případě, kdy $\mathrm{card}(\Omega) = \mathfrak{c}$, nejde vyčíslit $\mathscr{P}[X \in Y]$. Například uvažujme náhodnou veličinu $X \sim \mathrm{U}(0,1)$. Pro $0 \leqslant a < b \leqslant 1$ počítáme
$$\mathscr{P}[X \in \langle a,b\rangle] = b - a.$$
Jak bylo dokázáno, tak $\mathscr{P}[X \in \mathbf{Q} \cap \langle 0,1)]$ nešlo vyčíslit. Řešení tohoto problému je, že do $\mathcal{A}$ se zahrnou pouze tzv. **měřitelné množiny** a proto pro spojité náhodné veličiny bereme $\mathcal{A} \subsetneqq 2^{\Omega}$.

Uvedeme některé elementární vlastnosti pravděpodobnosti $\mathscr{P}$:

1. $\Omega \in \mathcal{A} \ \wedge \ \mathcal{A} \subset 2^{\Omega}$,
2. $\emptyset \in \mathcal{A}$,
3. $\forall A \in \mathcal{A} : \mathscr{P}(A) \geqslant 0$,
4. $\mathscr{P}[\emptyset] = 0$,
5. $\mathscr{P}[\Omega] = 1$,
6. $\forall A, B \in \mathcal{A} : A \subset B \Rightarrow \mathscr{P}[A] \leqslant \mathscr{P}[B]$,
7. $\forall A, B \in \mathcal{A} : A \cap B = \emptyset \Rightarrow \mathscr{P}[A \cup B] = \mathscr{P}[A] + \mathscr{P}[B]$.

Míra je jakési zobecnění pojmu délka v jedné dimenzi, obsahu ve dvou dimenzích a objemu ve třech dimenzích. Uvažme nyní množinovou funkci
$$m(x) : \mathcal{A} \to \mathbf{R}^* \equiv \mathbf{R} \cup \{-\infty, +\infty\}.$$
Aby množinová funkce měla vlastnosti míry, tj. délky, obsahu, ..., musí splňovat následující axiomy:

1. axiom nulovosti $m(\emptyset) = 0$,
2. axiom prázdné množiny $\emptyset \in \mathcal{A}$,
3. axiom nezápornosti $\forall A \in \mathcal{A} : m(A) \geqslant 0$,

4. axiom monotonie $\forall A, B \in \mathcal{A} : A \subset B \Rightarrow m(A) \leqslant m(B)$,

5. axiom aditivity $\forall A, B \in \mathcal{A} : A \cap B = \emptyset \Rightarrow m(A \cup B) = m(A) + m(B)$.

Vidíme, že pokud se podíváme zpět na vlastnosti pravděpodobnosti, tak vlastnosti č. 2, 3, 4, 6 a 7 jsou právě zmíněné axiomy míry. Normalizační vlastnost 5, $\mathscr{P}[\Omega] = 1$, obecná míra mít nemusí. Pravděpodobnost je tedy konečná a normalizovaná míra.

Rozeberme nyní jaké příjemné vlastnosti bychom chtěli, aby splňovala soustava množin $\mathcal{A}$, se kterou pravděpodobnost pracuje. Existuje-li v $\mathcal{A}$ množina Ω tak, že

$$\forall A \in \mathcal{A} : A \subset \Omega, \tag{0.13}$$

pak množinu Ω nazveme prezidentem soustavy $\mathcal{A}$. Ne každá soustava ale prezidenta obsahuje. Např. pokud $\mathcal{A} = 2^{\{3,8,11\}}$, potom prezident $\Omega = \{3, 8, 11\}$. Pokud ale uvážíme soustavu polouzavřených intervalů

$$\mathcal{H}_1 = \{\emptyset; \langle\alpha, \beta\rangle : -\infty < \alpha < \beta < +\infty\}, \tag{0.14}$$

potom jediným kandidátem na prezidenta, který splňuje vlastnost (0.13), je interval $(-\infty, +\infty)$, který ale neleží v soustavě $\mathcal{H}_1$. Řekneme, že $\mathcal{A}$ je aditivní, právě tehdy, když

- $\forall A, B \in \mathcal{A} : A \cup B \in \mathcal{A}$.

Soustavu $\mathcal{A}$ nazveme okruhem, právě tehdy, když

- $\mathcal{A}$ je aditivní,
- $\forall A, B \in \mathcal{A} : A \setminus B \in \mathcal{A}$.

Řekneme, že $\mathcal{A}$ je algebra, právě tehdy, když

- $\mathcal{A}$ je okruh,
- v $\mathcal{A}$ existuje prezident Ω.

Soustava $\mathcal{A}$ je σ-algebrou, právě tehdy, když

- $\mathcal{A}$ je algebra,
- $\mathcal{A}$ je σ-aditivní, tj:

$$(A_i \in \mathcal{A})_{i=1}^{+\infty} \Rightarrow \bigcup_{i=1}^{+\infty} A_i \in \mathcal{A}.$$

V korektní definici Kolmogorovy pravděpodobnosti, kterou doktor Kůs zadefinuje na přednáškách z MIP, požadujeme právě, aby $\mathcal{A}$ byla σ-algebrou. Pro úplnost dodejme, že soustavu $\mathcal{A}$ nazveme polookruhem, právě tehdy, když

- $\forall A, B \in \mathcal{A} : A \cap B \in \mathcal{A}$,
- $\forall A, B \in \mathcal{A}$: buď $A \setminus B \in \mathcal{A}$, tedy každý okruh je i polookruhem, nebo stačí, že bude splněno:

$$\exists C_1, C_2, \ldots, C_m \in \mathcal{A} : A \setminus B = \biguplus_{i=1}^{m} C_i.$$

Příkladem polookruhu bez prezidenta je právě soustava $\mathcal{H}_1$, viz (0.14). Tato soustava je základní (výchozí) soustava při zobecňování pojmu délka, obsah a objem.

Z vlastností σ-algebry $\mathcal{A}$ tedy plyne, že pokud znám hodnotu $\mathcal{P}[X \in A]$ i hodnotu $\mathcal{P}[X \in B]$, potom musíme být schopni určit i hodnotu $\mathcal{P}[X \in A \cup B]$, stejně tak i $\mathcal{P}[X \in A \setminus B]$. Jelikož P je ve skutečnost jistý jev, potom $\mathcal{P}[X \in P] = 1$ a pokud známe posloupnost hodnot $\mathcal{P}[X \in A_1]$, $\mathcal{P}[X \in A_2]$, $\mathcal{P}[X \in A_3]$, $\ldots$, musíme být schopni určit také $\mathcal{P}\left[X \in \bigcup_{i=1}^{+\infty} A_i\right]$.

0.4 Lebesgueův integrál v teorii pravděpodobnosti

Uvažujeme opět pravděpodobnostní funkci

$$\mathscr{P} : \mathcal{A} \to \langle 0, +\infty \rangle,$$

která splňuje axiomy míry a pro prezidenta soustavy P platí $\mathscr{P} = 1$. Připomeňme si problém, který vznikal při zavedení pravděpodobnosti. Pro náhodnou veličinu $X \sim \mathrm{U}(0,1)$. Ukázali jsme, že

$$\mathscr{P}\left[\langle a, b)\right] = b - a.$$

Chtěli bychom, abychom mohli pro $\forall A \in \mathcal{A}$ počítat:

$$\mathscr{P}\left[X \in A\right] = \mathscr{P}\left[A\right] = \int_A g(x)\ \mathtt{d}x, \tag{0.15}$$

kde $g(x)$ je hustota pravděpodobnosti náhodné veličiny X. Problém jsme našli v případe, že $A = \langle \frac{1}{3}, 12) \cap \mathbf{Q}$. Pro naši uniformně rozdělenou náhodnou veličinu X ale integrál (0.15) neexistuje, neboť

$$\mathscr{P}\left[X \in \langle \frac{1}{3}, 12) \cap \mathbf{Q}\right] = \int_A g(x)\ \mathtt{d}x = \int_A 1\ \mathtt{d}x = \int_{\frac{1}{3}}^{\frac{1}{2}} \mathcal{D}(x)\ \mathtt{d}x$$

a poslední integrál z Dirichletovy funkce nelze riemannovsky vyčíslit. Nelze ale říct, že pravděpodobnost, že náhodná veličina padne do množiny $A = \langle \frac{1}{3}, 12) \cap \mathbf{Q}$ neexistuje, protože když budu losovat čísla z integralu $\langle 0, 1)$, budu si zapisovat výsledky, tak velmi rychle uvidím, že čím víc čísel tahám, tím je méně pravděpodobné, že si vytáhnu racionální číslo. Tedy z vlastní zkušenosti víme, že $\mathscr{P}\left[X \in \langle \frac{1}{3}, 12) \cap \mathbf{Q}\right]$ existuje. To tedy znamená, že někde na té cestě, kterou jsme se vydali, byla chyba v úvaze. Vše funguje správně, dokud nepoužijeme Riemannův integrál. Již dříve jsme si rozmysleli, že v intervalu $\langle 0, 1)$ jsou skoro všechna čísla iracionální. Tedy v řeči míry potom:

$$\mu\left(\langle 0, 1) \cap \mathbf{Q}\right) = 0,$$

a tedy máme jasnou intuici, že $\mathscr{P}\left[X \in \langle 0, 1 \cap \mathbf{Q}\right] = 0$. Konstrukci Lebesgueova integrálu pan proděkan podrobně projde na přednáškách z MAB4, proto jen rychle zrekapitulujeme hlavní kroky konstrukce. V první fázi konstruujeme integrál pro funkce ze Základního systému $\mathscr{Z}_\mu$, tedy pro tzv. schodovité funkce. Schodovitost funkce spočívá v tom, že

- $\mathrm{Dom}(f) = E$, kde E je prezidentem v $\mathcal{A}$,
- $\mathrm{Ran}(f) \subset \langle 0, +\infty)$,
- $\mathrm{Ran}(f) = \{d_1, d_2, \ldots, d_m\}$, tedy jde o konečnou množinu,
- $f(x)$ musí být finitní funkce, tedy $\mu(\mathrm{supp}(f)) < +\infty$,
- funkce $f(x)$ musí být μ-měřitelná (tuto vlastnost budeme diskutovat na MAB4, zde předpokládejme, že všechny funkce jsou μ-měřitelné).

Lebesgueův integrál pro funkce ze Základního systému definujeme jako

$$(\mathcal{L}) \int_E f(x)\ \mathtt{d}x = \sum_{i=1}^{m} d_i \cdot \mu(A_i),$$

kde $A_i = f^{-1}(d_i)$. Tento postup snadno aplikujeme na vyčíslení $(\mathcal{L}) \int_0^1 \mathcal{D}(x)\ \mathtt{d}x$. Totiž Dirichletova funkce $\mathcal{D}(x)$ je právě jedna ze schodovitých funkcí, neboť:

- $\mathrm{Dom}(\mathcal{D}) = \langle 0, 1)$, kde interval $\langle 0, 1)$ je prezidentem soustavy $\mathcal{A}$,
- $\mathrm{Ran}(\mathcal{D}) \subset \mathbf{R}_0^+$,
- $\mathrm{Ran}(\mathcal{D}) = \{0, 1\}$,
- funkce $\mathcal{D}$ je μ-měřitelná,
- podmínka finitnosti: $\mathrm{supp}(\mathcal{D}) = \mathbf{Q} \cup \langle 0, 1)$ a $\mu(\mathrm{supp}(\mathcal{D})) = 0$.

Proto

$$(\mathcal{L})\int_0^1 \mathcal{D}(x)\ \mathrm{d}x=\sum_{i=1}^{2} d_i\cdot\mu(A_i),$$

kde $d_1=0$, $d_2=1$, $A_1=(\mathbf{R}\setminus\mathbf{Q})\cap\langle 0,1)$ a $A_2=\mathbf{Q}\cap\langle 0,1)$. Výsledkem je

$$(\mathcal{L})\int_0^1 \mathcal{D}(x)\ \mathrm{d}x=\sum_{i=1}^{2} d_i\cdot\mu(A_i)=d_2\cdot\mu(A_2)=1\cdot 0=0.$$

Můžeme tedy tuto diskuzi uzavřít tím, že

$$(\mathcal{L})\int_U \mathcal{D}(x)\ \mathrm{d}x=0,$$

kde $U=W\cap\mathbf{Q}$ a W je libovolná množina $W\subset\mathbf{R}$, neboť $\mu(U)=0$.

Nyní jsme uvažovali, že $\mathcal{A}$ byla σ-algebra. Při konstrukci Legesgueova integrálu bereme speciálně $\mathcal{A}=\mathcal{M}$, kde $\mathcal{M}$ je tzv. Lebesgueova σ-algebra. V teorii pravděpodobnosti se používá výhradně soustava borelovských množin, neboli Borelova σ-algebra $\tilde{\mathcal{D}}$, kde $\tilde{\mathcal{D}}$ je nejmenší σ-okruh obsahující soustavu $\mathcal{D}=\{B\subset\mathbf{R}: B=B^\circ\}$. Na MAB4 bude tato problematika důkladně studována, zmiňme zde tedy pouze, že

$$\tilde{\mathcal{D}}\subset\mathcal{M}.$$

Na cvičeních bude dokázáno, že do soustavy $\tilde{\mathcal{D}}$ patří všechny otevřené množiny, uzavřené množiny, jednoprvkové množiny, konečně i spočetně prvkové množiny, polouzavřené intervaly, průniky všech zmíněných množin s $\mathbf{Q}$, atd.

Relevantní pravděpodobnosntní dotaz tedy bude pouze dotaz typu:

$$\mathscr{P}\left[X\in A\right]=(\mathcal{L})\int_A g(x)\ \mathrm{d}x,$$

kde $A\in\tilde{\mathcal{D}}$, množina A musí být borelovská. Chtěli bychom ale mít jistotu, že $(\mathcal{L})\int_A g(x)\ \mathrm{d}x$ vždy existuje! Pro schodovité funkce (i ve více dimenzích) je ale existence zřejmá.

Lebesgueův integrál nyní máme zavedený pouze pro schodovité funkce. Ve druhé fázi budeme Lebesgueův integrál konstruovat pro libovolné nezáporné funkce, které jsou měřitelné, $g(x): E\to\mathbf{R}$. Ke každé takovéto $g(x)$ existuje posloupnost funkcí

$$f_n(x)\in\mathcal{Z}_\mu: f_n(x)\nearrow g(x),$$

symbolem $f_n(x)\nearrow g(x)$ říkáme, že funkční posloupnost $f_n(x)$ konverguje zdola k funkci $g(x)$, tedy $\forall x\in E:$ $\lim_{n\to+\infty} f_n(x)=g(x)$, ale zároveň platí, že $f_n(c)$ je neklesající pro všechny $c\in E$. Díky tomu, že $f_n(x)\in\mathcal{Z}_\mu$, pak integrál zkonstruovaný v první fázi $(\mathcal{L})\int_E f_n(x)\ \mathrm{d}x$ jistě existuje, proto následující definice dává dobrý smysl:

$$(\mathcal{L})\int_E g(x)\ \mathrm{d}x=\lim_{n\to+\infty}(\mathcal{L})\int_E f_n(x)\ \mathrm{d}x.$$

Všechny nezáporné funkce tedy mají Lebesgueův integrál, protože jelikož $f_n(x)\nearrow g(x)$, potom určitě platí

$$\int_E f_n(x)\ \mathrm{d}x\leqslant\int_E f_{n+1}(x)\ \mathrm{d}x,$$

neboť $f_n(x)\leqslant f_{n+1}(x)$. Mám tedy číselnou posloupnost integrálu $\left((\mathcal{L})\int_E f_n(x)\ \mathrm{d}x\right)_{n=1}^{+\infty}$, která je rostoucí a limita rostoucí posloupnosti vždy existuje!

Ve třetí fázi definujeme Legesgueův integrál pro libovolné funkce. Tento krok ale v pravděpodobnosti dělat vůbec nemusíme, protože cílem bylo zodpovědět pravděpodobnostní dotaz

$$\mathscr{P}\left[X\in A\right]=(\mathcal{L})\int_A g(x)\ \mathrm{d}x,$$

kde $A\in\tilde{\mathcal{D}}$, tedy integruji hustoty pravděpodobnosti, ale to jsou vždy kladné funkce. Dokázali jsme tedy, že $(\mathcal{L})\int_A g(x)\ \mathrm{d}x$ existuje pro všechny množiny $A\in\tilde{\mathcal{D}}\subset\mathcal{M}\subset 2^E$, kde E je prezident, a pro úplně všechny funkce (hustoty) $g(x)$. Navíc pro pravděpodobnost z monotonie platí

$$(\mathcal{L})\int_A g(x)\ \mathrm{d}x\leqslant(\mathcal{L})\int_E g(x)\ \mathrm{d}x$$

a zároveň, jelikož E je prezident, musí platit axiom normalizace $\mathscr{P}[E]=(\mathcal{L})\int_E g(x)\ \mathrm{d}x=1$, pak všechny integrály $(\mathcal{L})\int_A g(x)\ \mathrm{d}x$ jsou dokonce konečné. Lebesgueův integrál je tedy ideální nástroj pro zodpovídání drtivé většiny pravděpodobnostních dotazů, protože všechny integrály z hustot pravděpodobnosti existují a jsou konečné. Jediné, co požadujeme je, aby ten pravděpodobnostní dotaz obsahovat množinu, která leží v množině lebegueovsky měřitelných množin (tedy v $\mathcal{M}$) a nebo zjednodušeně to lze udělat tak, že to A se bere ze soustavy borelovských množin $\tilde{\mathcal{D}}$, která je podsoustavou $\mathcal{M}$ a je nějak snázeji uchopitelná. Teorie pravděpodobnosti je vystavěná nad borelovskými množinami.