Přednáška 1.

- Náhodná veličina
- Pravděpodobnost
- Distribuční funkce
- Diskrétní náhodné veličiny
- Hustota pravděpodobnosti (spojité náhodné veličiny)
- Střední hodnota a rozptyl

Náhodná veličina

S každým **pokusem** nebo **hrou** je spojena množina všech možných výsledků. Označme symbolem Ω množinu všech možných, navzájem se vylučujících výsledků. Libovolný možný výsledek, označme $\omega \in \Omega$, nazveme *elementární jev*, Ω pak nazveme *základní pravděpodobnostní prostor*.

DEFINICE:

*Náhodná veličina **X** je zobrazení $X : \ \Omega \to R$ takové, že pro každé $x \in R$ platí*

$$X^{-1}\big((-\infty,x\rangle\big) = \{\omega \in \Omega \,|\, X(\omega) \leq x\} = [X \leq x] \subset \Omega.$$

Neboli: *Náhodná veličina **X** je reálná funkce definovaná na množině všech elementárních jevů (=pokusů experimentu), která každému jevu přiřadí reálné číslo.*

Množina čísel přiřazených elementárním jevům tvoří **obor hodnot** *náhodné veličiny*.

PŘÍKLAD: HOD 6 STĚNNOU KOSTKOU

- Pokus = hod kostkou, elementární jev $\omega \in \Omega = \{1, 2, 3, 4, 5, 6\}$. X je rovna hodnotě napsané na vrchní stěně kostky po dopadu. Obor náhodné veličiny je také $\{1, 2, 3, 4, 5, 6\}$.
- Pokus = hod 2 kostkami, pak $\omega \in \Omega = \{(i,j) | i = 1, \dots 6; j = 1, \dots 6; i \leq j\}$. X je rovna součtu padlých hodnot. Obor náhodné veličiny je $\{2, 3, \dots, 12\}$, jelikož např. $X(1,1) = 2$.

pravděpodobnost

KLASICKÁ DEFINICE PRAVDĚPODOBNOSTI:

- Uvažujme náhodný pokus, který může vykázat konečný počet „**n**“ vzájemně se vylučujících výsledků (např. tři hody kostkou)
- Předpokládáme, že všechny výsledky jsou stejně pravděpodobné (symetrie, homogenita)
- Jestliže „**m**“ z těchto výsledků má za následek realizaci jevu „**A**“ (např. padnou dvě šestky) a zbylých „**n-m**“ výsledků tuto realizaci vylučuje, potom pokládáme pravděpodobnost jevu „**A**“ rovnu

$$P(A) = \frac{m}{n} = \frac{number\ of\ outcomes\ in\ A}{total\ number\ of\ outcomes}$$

GEOMETRICKÁ DEFINICE PRAVDĚPODOBNOSTI:

- Zobecnění klasické definice pro nekonečně (nespočetně) mnoho možností
- Předpokládáme, že výsledek je stejně pravděpodobný v každém bodě objektu
- Pravděpodobnost jevu „$\mathbf{B}$“, neboli pravděpodobnost, že náhodná veličina má hodnotu v množině „$\mathbf{B}$“, vypočteme jako podíl velikosti (obsah, objem) množiny $\boldsymbol{S_B}$ všech bodů příznivých jevu „$\mathbf{B}$“ a velikosti celého objektu $\mathbf{S}$

$$P(B) = \frac{S_B}{S}$$

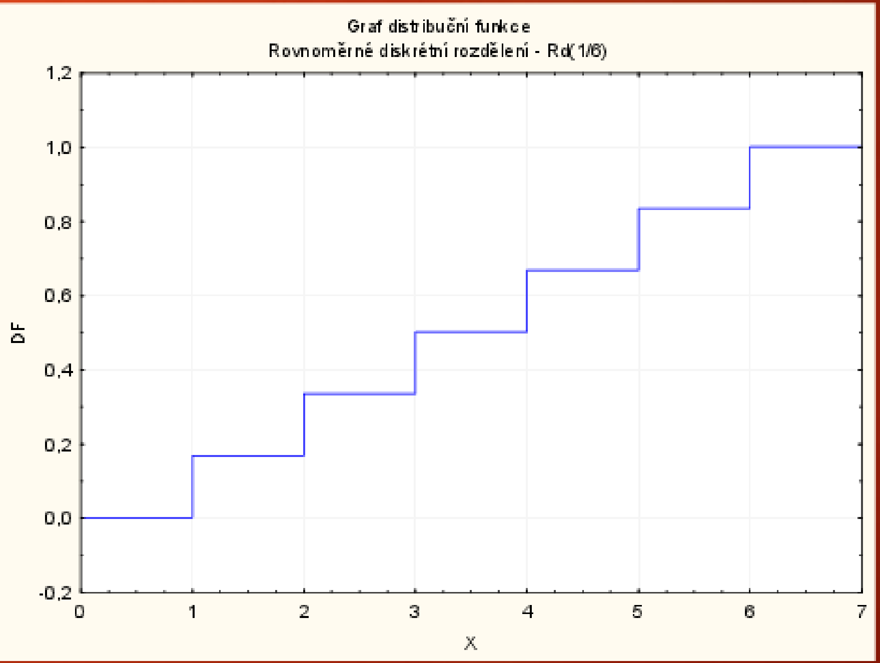
Distribuční funkce

DEFINICE:

Nechť $\boldsymbol{X}$ *je náhodná veličina. Funkci* $F_X : R \to \langle 0,1 \rangle$, *definovanou pro všechna* $x \in R$ *vztahem* $F_X(x) = P[X \leq x]$, *nazýváme distribuční funkcí náhodné veličiny* $\boldsymbol{X}$.

VLASTNOSTI:

- Obor hodnot distribuční funkce je $\langle 0,1 \rangle$.
- Distribuční funkce je neklesající.



Diskrétní náhodné veličiny
(příklady rozdělení pravděpodobnosti)

Veličina $\boldsymbol{X}$ *má diskrétní rozdělení, jestliže její obor hodnot* ***H*** *má spočetně mnoho prvků, tj.* $\sum_{x_k \in H} P[X = x_k] = 1$.

Funkci $P[X = x]$ nazveme pravděpodobnostní funkcí náhodné veličiny **X.**

ALTERNATIVNÍ (BERNOULLIHO) ROZDĚLENÍ:

Veličina $\boldsymbol{X}$ *má alternativní rozdělení s parametrem* $p \in (0,1)$*, jestliže platí* $P[X = 1] = p, \quad P[X = 0] = 1 - p$.

- Toto rozdělení popisuje náhodné jevy mající 2 možné výsledky (hod mincí, sudé vs liché číslo na kostce, pravda vs lež, hodnota bitu…), kde úspěch nastává s pravděpodobností $\mathbf{p}$.
- Zapisujeme, že náhodná veličina má Bernoulliho rozdělení jako $X \sim Be(p)$.

BINOMICKÉ ROZDĚLENÍ:

Uvažujme náhodnou veličinu $\mathbf{X}$ vyjadřující počet výskytů jevu $\mathbf{A}$ v $\mathbf{n}$ nezávislých pokusech. Jestliže pravděpodobnost jevu $\mathbf{A}$ je pokaždé rovna $\mathbf{p}$, pravděpodobnost, že jev $\mathbf{A}$ nastane z $\mathbf{n}$ pokusů právě $\mathbf{k}$-krát je $P[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$.

Řekneme, že náhodná veličina $\boldsymbol{X}$ *má binomické rozdělení (značíme* $X \sim Bi(n, p)$*) s parametry* $n \in N$ *a* $p \in (0,1)$*, jestliže* $P[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$.

Příklad: 4 hody mincí

Diskrétní náhodné veličiny
(příklady rozdělení pravděpodobnosti)

PASCALOVO/GEOMETRICKÉ ROZDĚLENÍ:

Veličina $\boldsymbol{X}$ *má Pascalovo rozdělení s parametrem* $p \in (0,1)$, *jestliže platí* $P[X = k] = p(1-p)^{k-1}$, *pro* $k = 1,2, \ldots$

Zapisujeme, že náhodná veličina má Pascalovo rozdělení jako $X \sim Pa(p)$.

Náhodná veličina $X \sim Pa(p)$ značí počet pokusů než nastane jev **A,** pokud jev nastává s pravděpodobností **p.**

POISSONOVO ROZDĚLENÍ:

Veličina $\boldsymbol{X}$ *má Poissonovo rozdělení s parametrem* $\lambda > 0$, *jestliže* $P[X = k] = e^{-\lambda}\frac{\lambda^k}{k!}$, *pro* $k = 0, 1, 2, \ldots$

Značíme $X \sim Po(\lambda)$.

Poissonovo rozdělení se využívá pro určení počtu událostí v časovém intervalu. Parametr λ je průměrný počet událostí za tento časový interval.

Pokud $X \sim Bi(n, p)$, pak pro velká **n** bude platit $n \cdot p = \lambda$, tedy Poissonovo rozdělení aproximuje Binomické.

Absolutně spojité náhodné veličiny
(definice Hustoty pravděpodobnosti)

DEFINICE:

*Náhodná veličina **X** má absolutně spojité rozdělení (ASR), jestliže existuje nezáporná reálná funkce $f_X(x)$ taková, že pro každé $x \in R$ platí $F_X(x) = \int_{-\infty}^{x} f_X(t)dt$.*

*Funkci $f_X(x)$ nazýváme **hustotou pravděpodobnosti** náhodné veličiny X.*

ZÁKLADNÍ VLASTNOSTI:

Pro X mající ASR platí:

- $f_X(x) = \frac{dF}{dx}(x)$,
- $P[X \in A] = \int_A f_X(x)\, dx$,
- $P[X = a] = 0$.

KAŽDÁ HUSTOTA PRAVDĚPODOBNOSTI DÁLE SPLŇUJE:

- $f_X(x) \geq 0$,
- $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$.

Příklady HUSTOT pro spojité veličiny

GAUSSOVO (NORMÁLNÍ) ROZDĚLENÍ:

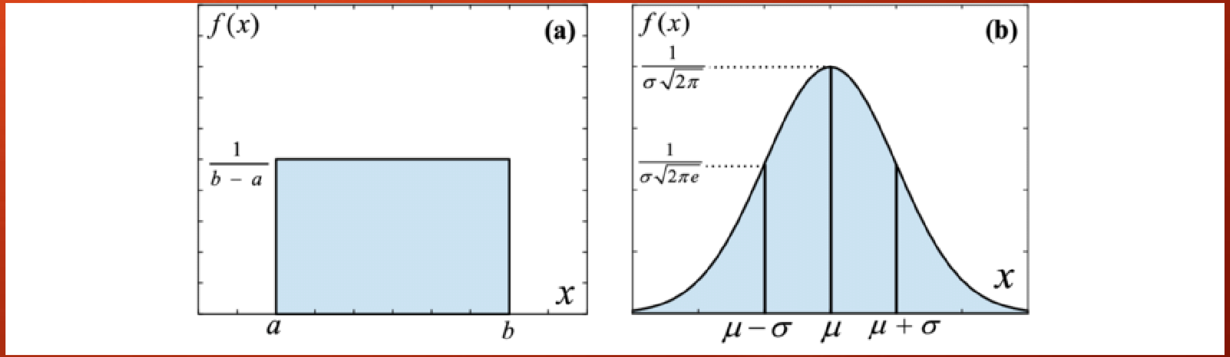
Řekneme, že náhodná veličina $\boldsymbol{X}$ *má normální (Gaussovo) rozdělení s parametry* $\mu \in R,\ \sigma^2 > 0$ *, jestliže její hustota pravděpodobnosti má tvar* $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ *.*

Značíme $X \sim N(\mu, \sigma^2)$ *. Distribuční funkci nelze vyjádřit analyticky pomocí základních funkcí.*
$F_X(x) = erf(x)$*; viz speciální funkce v Matlabu.*

UNIFORMNÍ (ROVNOMĚRNÉ) ROZDĚLENÍ:

Řekneme, že náhodná veličina $\boldsymbol{X}$ *má rovnoměrné rozdělení s parametry* $a < b,\ \ a, b \in R$*, jestliže její hustota má tvar* $f_X(x) = \frac{1}{b-a}$ *, pro* $x \in (a, b)$ *, a* $f_X(x) = 0$ *jinak.*

Značíme $\mathrm{X} \sim \mathrm{U}(\mathrm{a}, \mathrm{b})$.

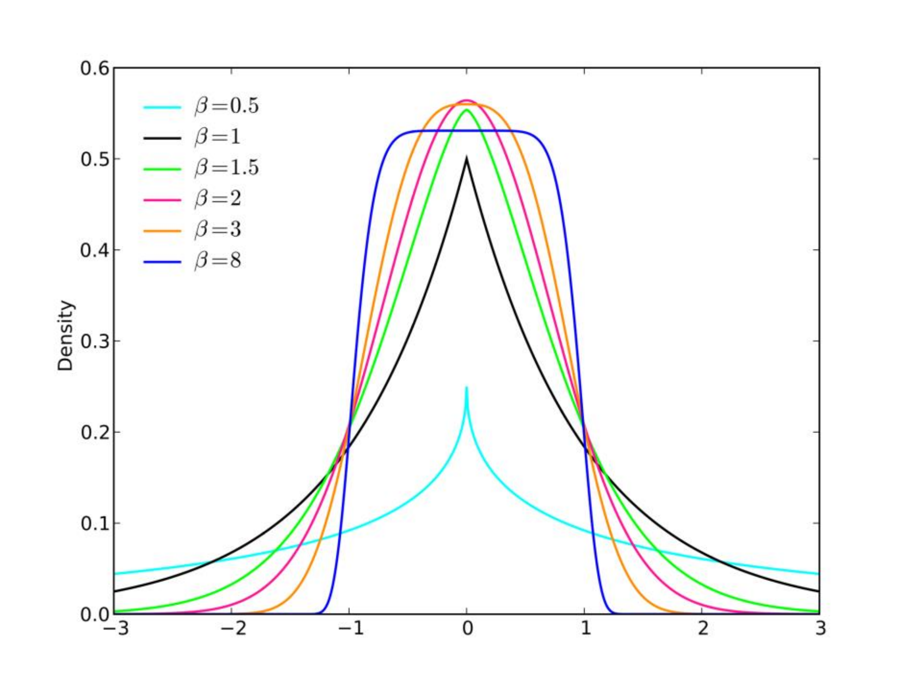


Příklady HUSTOT pro spojité veličiny

ZOBECNĚNÉ GAUSSOVO ROZDĚLENÍ (GGD):

Řekneme, že náhodná veličina $\boldsymbol{X}$ *má zobecněné normální (Gaussovo) rozdělení s parametry* $\mu \in R$, $\alpha, \beta > 0$ *, jestliže její hustota pravděpodobnosti má tvar* $f_X(x) = \frac{\beta}{2\alpha\Gamma\left(1/\beta\right)}\, e^{-\left(\frac{|x-\mu|}{\alpha}\right)^{\beta}}$.

Značíme $X \sim GGD(\mu, \alpha, \beta)$.



Střední hodnota a rozptyl

DISKRÉTNÍ NÁHODNÁ VELIČINA

ABSOLUTNĚ SPOJITÁ NÁHODNÁ VELIČINA

STŘEDNÍ HODNOTA

$$E[X] = \sum_i x_i \, P[X = x_i]$$

$$E[X] = \int_{-\infty}^{\infty} x \, p(x) \, dx$$

ROZPTYL

$$\mathrm{V}[X] = \sum_i (x_i - E[X])^2 \, P[X = x_i]$$

$$V[X] = \int_{-\infty}^{\infty} (x - E[X])^2 \, p(x) \, dx$$

Pro střední hodnoty diskrétních i ASR náhodných veličin **X,Y** a libovolné konstanty $a, c \in R$ platí:
$E[a + X + cY] = a + E[X] + cE[Y]$.

Sdružená hustota pravděpodobnosti

DEFINICE:

*Pravděpodobnost, že nastane jev „**A**“, a zároveň jev „**B**“, nazveme **sdružená pravděpodobnost** jevů A a B. Značíme* $P(A \cap B)$ *.*

Sdruženou hustotu pravděpodobnosti náhodných veličin **X** a **Y** pak značíme $\mathrm{p}(\mathrm{x}, \mathrm{y})$.

DEFINICE:

*Nechť jsou dány jevy **A**, **B** takové, že* $P(B) > 0$ *. **Podmíněnou pravděpodobností** jevu **A** za podmínky, že nastal jev **B**, nazveme číslo*

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Poznámka: Stejný vzorec platí i pro hustoty pravděpodobnosti.

$$p(x|y) = \frac{p(x, y)}{p(y)}$$

Statistická závislost a nezávislost

DEFINICE:

Řekneme, že náhodné veličiny $\boldsymbol{X}$ *a* $\boldsymbol{Y}$ *jsou statisticky* ***nezávislé****, jestliže platí* $p(x,y) = p(x)p(y)$ *.*

Hustoty $p(x), p(y)$ *nazýváme* ***marginální*** *hustoty pravděpodobnosti. Pro marginální hustoty platí*

$$p(x) = \int_{-\infty}^{\infty} p(x,y)dy$$

Nezávislé veličiny, které mají stejnou hustotu pravděpodobnosti (=stejné rozdělení / distribuční funkci) nazveme *nezávislé stejně rozdělené* a značíme *i.i.d. (*independent identically distributed*).*

PŘÍKLAD:

- Uvažujme hod kostkou a označme si 2 jevy **A** – padne sudé číslo a **B** - padne číslo nepřevyšující 2. Jsou jevy **A** a **B** nezávislé?
- Upravme si drobně zadání, a uvažujme jev **A** – padne sudé číslo a **B** - padne číslo nepřevyšující 3. Jsou jevy **A** a **B** i nyní nezávislé?

Cvičení 1

1. Dokažte, že rozdělení diskrétních náhodných veličin v přednášce splňují vlastnosti pravděpodobnostních rozložení. Bonus: Dokažte to i pro spojitá pravděpodobnostní rozložení.
2. Dokažte, že μ je střední hodnota Normálního rozdělení.
3. Matlab: Vykreslete křivku hustoty pravděpodobnosti náhodné veličiny $X\sim N(1,1)$ a vygenerujte N = 1000 i.i.d. samplů a zobrazte jejich histogram. Stejně tak pro $Y\sim U(0,3)$.
4. Babička mi v průměru posílá 5 SMS zpráv za rok. Jaká je pravděpodobnost, že mi jich příští rok pošle 6? A jaká, že maximálně 4? Ukažte na tomto příkladu, že Poissonovo rozdělení aproximuje Binomické pro velká n (n=365).
5. Matlab: Trefujte klávesu vždy na sudé vteřině a měřte odchylku, opakujte např. 50 krát. Pak vykreslete histogram odchylek a srovnejte s Normálním rozdělením.