

Základy teorie pravděpodobnosti

Náhodná veličina – Pravděpodobnost náhodné veličiny

Roman Biskup

(zapálený) statistik ve výslužbě, aktuálně analytik v praxi ;-)
roman.biskup(at)email.cz

12. února 2012



Obsah

Definice pojmů

Náhodná veličina

Pravděpodobnost náhodné veličiny

Popis náhodné veličiny

Diskrétní a spojitá náhodná veličina

Distribuční funkce a hustota pravděpodobnosti

Číselné charakteristiky náhodné veličiny

Medián a kvantil

Grafická vizualizace rozdělení náhodných veličin



Náhodná veličina

Náhodnou veličinou X budeme rozumět:

- ▶ (měřitelné) zobrazení, které náhodným jevům přiřazuje reálná čísla, tj.

$$X : \Omega \rightarrow \mathbb{R};$$

$$X(\omega) = x, \quad x \in \mathbb{R};$$

- ▶ respektive, proměnnou nabývající různých reálných hodnot.
- ▶ Dle počtu hodnot, jakých může náhodná veličina nabývat, ji lze rozdělit do dvou základních skupin a to náhodnou veličinu:
 - ℵ **diskrétní** – X nabývá nejvýše spočetného počtu hodnot (konečný a nekonečný počet);
 - ℵ **spojitou** – X nabývá nespočetného počtu hodnot $(\langle \bullet; \bullet \rangle, \mathbb{R})$.
- ▶ Zápisem $X(\omega) = x$ (nadále zkráceně $X = x$) budeme rozumět, že náhodná veličina X nabyla hodnoty x . Modeluje-li náhodná veličin náhodný pokus, pak to znamená, že jev ω nastal.
- ▶ Zápisem $P(X = x)$ budeme rozumět pravděpodobnost s jakou $X = x$.



Popis náhodné veličiny I

Náhodnou veličinu lze popsat (definovat):

- ▶ tabulkou hodnot náhodné veličiny X a příslušných pravděpodobností $P(X = x_i) = p_i, i = 1, 2, \dots, n$

x_i	1	2	3	4	5	6
p_i	1/6	1/6	1/6	2/6	1/12	1/12

- ▶ předpisem, který stanoví, jak vypočítat $P(X = x)$

$$P(X = x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}, \quad x = 0, 1, \dots;$$

- ▶ předpisem, který stanoví, jak vypočítat hustotu pravděpodobnosti $f(x)$ **Pozor:** $f(x) \neq P(X = x)$

$$f(x) = \frac{1}{b-a}, \quad x \in \langle a; b \rangle;$$



Popis náhodné veličiny II

- ▶ distribuční funkcí $F(x)$, definovanou jako: $F(x) = P(X \leq x)$

$$F(x) = \sum_{0 \leq i \leq x} e^{-\lambda} \cdot \frac{\lambda^i}{i!}, \quad i = 0, 1, \dots;$$

$$F(x) = \int_{-\infty}^x \frac{1}{b-a} dt, \quad x \in \mathbb{R};$$

- ▶ číselnými charakteristikami: EX , DX , ...

$$EX = \lambda, \quad DX = \lambda;$$

$$EX = \frac{a+b}{2}, \quad DX = \frac{b-a}{12}.$$



Diskrétní náhodná veličina

Nechť známe hodnoty x_i , $i = 1, 2, \dots, (n)$ náhodné veličiny X a k nim příslušné pravděpodobnosti $P(X = x_i) = p_i$, pak:

- ▶ hodnoty p_i pro $i = 1, 2, \dots, (n)$ označíme jako **pravděpodobnostní funkci**;

- ▶ $\sum_{i=1}^{n/\infty} p_i = 1$;

- ▶ $F(x) = \sum_{x_i \leq x} p_i$ – distribuční funkce je skokovitá – je konstantní mezi hodnotami x_i a právě v každé hodnotě x_i má skok velikosti p_i ;

- ▶ $EX = \sum_{i=1}^{n/\infty} x_i \cdot p_i$;

- ▶ $DX = \sum_{i=1}^{n/\infty} (x_i - EX)^2 \cdot p_i$.



Spojité náhodné veličiny

Nechť známe obor hodnot $x \in \mathbb{R}$ (respektive $x \in \langle \bullet; \bullet \rangle$) náhodné veličiny X a k ní příslušnou hustotu pravděpodobnosti $f(x)$, pak:

$$\blacktriangleright \int_{-\infty}^{\infty} f(x) dx = 1;$$

$$\blacktriangleright F(x) = \int_{-\infty}^x f(t) dt - \text{distribuční funkce je spojitá};$$

$$\blacktriangleright EX = \int_{-\infty}^{\infty} x \cdot f(x) dx;$$

$$\blacktriangleright DX = \int_{-\infty}^{\infty} (x - EX)^2 \cdot f(x) dx;$$

! $P(X = x_0) = 0$ – Pravděpodobnost, že **spojitá** náhodná veličina nabyde konkrétní hodnoty x_0 je **nulová** – **Paradox nulové pravděpodobnosti**.



Distribuční funkce

$$F(x) = P(X \leq x), F: \mathbb{R} \rightarrow \langle 0; 1 \rangle$$

Vlastnosti distribuční funkce:

1. $0 \leq F(x) \leq 1, \forall x \in \mathbb{R}$ – omezená funkce;
2. $\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow \infty} F(x) = 1$;
3. $F(a) \leq F(b), \forall a, b \in \mathbb{R}, a < b$ – neklesající funkce;
4. $\lim_{x \rightarrow x_0^+} F(x) = F(x_0), \forall x, x_0 \in \mathbb{R}$ – zprava spojitá;
5. $F'(x) = f(x), \forall x \in \mathbb{R}$, kde má F derivaci – její derivací je **hustota pravděpodobnosti**;
6. $P(a < X \leq b) = F(b) - F(a), \forall a, b, x \in \mathbb{R}, a \leq b$.



Hustota pravděpodobnosti

$$f(x) = \dots, f : \mathbb{R} \rightarrow \langle 0; \infty \rangle$$

$\leadsto$ slouží pro popis spojité náhodné veličiny

Vlastnosti hustoty pravděpodobnosti:

1. Po částech spojitá $\forall x \in \mathbb{R}$;
2. $0 \leq f(x)$, $\forall x \in \mathbb{R}$ – nezáporná funkce;
3. $\lim_{x \rightarrow -\infty} f(x) = 0$, $\lim_{x \rightarrow \infty} f(x) = 0$;
4. $F(x) = \int_{-\infty}^x f(t) dt$;
5. $P(a < X \leq b) = \int_a^b f(x) dx$, $\forall a, b, x \in \mathbb{R}$, $a \leq b$.



Střední hodnota – EX I

EX – numerický popis centrální tendence náhodné veličiny

- Vzorce

$$EX = \sum_{i=1}^{n/\infty} x_i \cdot p_i \quad \ddagger \quad EX = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

- Vlastnosti střední hodnoty

Nechť X , Y jsou libovolné náhodné veličiny, a a b reálná čísla pak:

- $Eb = b$ – střední hodnota konstanty je konstanta;
- $E(a \cdot X) = a \cdot EX$, pro $a \neq 0$;
- $E(X + b) = EX + b$;
- **$E(X + Y) = EX + EY$;**
- $X \geq 0 \Rightarrow EX \geq 0$ – střední hodnota z nezáporné náhodné veličiny je nezáporná;
- $E(X - EX) = 0$.
- Poznámky



Střední hodnota – EX II

- ▶ Nabývá-li náhodná veličina nekonečného počtu hodnot, pak nemusí střední hodnota vždy existovat.
- ▶ Výše zmíněné vlastnosti platí pouze tehdy, mají-li obě strany výrazů smysl.
- ▶ **Zvýrazněnou** vlastnost lze zobecnit na konečný počet náhodných veličin.



Rozptyl – DX I

DX – numerický popis variability náhodné veličiny

► Vzorce

$$DX = E(X - EX)^2 \quad \ddagger \quad DX = EX^2 - (EX)^2$$

$$DX = \sum_{i=1}^{n/\infty} (x_i - EX)^2 \cdot p_i \quad \ddagger \quad DX = \int_{-\infty}^{\infty} (x - EX)^2 \cdot f(x) dx$$

► Vlastnosti rozptylu

Nechť X , Y jsou libovolné náhodné veličiny, a a b reálná čísla pak:

- $Db = 0$ – rozptyl konstanty je nulový;
- $D(a \cdot X) = a^2 \cdot DX$, pro $a \neq 0$;
- $D(X + b) = DX$;
- Jsou-li X a Y nezávislé náhodné veličiny, pak $D(X + Y) = DX + DY$;
- $DX \geq 0$ – rozptyl je vždy nezáporný;
- Pro konečný rozptyl platí: $D(X/\sqrt{DX}) = 1$.

► Poznámky



Rozptyl – DX II

- ▶ Nabývá-li náhodná veličina nekonečného počtu hodnot, pak nemusí konečný rozptyl vždy existovat.
- ▶ Výše zmíněné vlastnosti platí pouze tehdy, mají-li obě strany výrazů smysl.
- ▶ **Zvýrazněnou** vlastnost lze zobecnit na konečný počet náhodných veličin.

$\sqrt{DX}$ $\sqrt{DX}$ – odmocnina z rozptylu se nazývá **směrodatná odchylka**



Medián a kvantil

$\tilde{x}$ Mediánem náhodné veličiny X je takové číslo $\tilde{x}$, pro které platí:

$$P(X \leq \tilde{x}) \geq 0,5 \quad \wedge \quad P(X \geq \tilde{x}) \geq 0,5.$$

$\tilde{x}_\alpha$ $\alpha \cdot 100\%$ (teoretickým) kvantilem náhodné veličiny X je takové číslo $\tilde{x}_\alpha$, $\alpha \in (0; 1)$, pro které platí:

$$P(X \leq \tilde{x}_\alpha) \geq \alpha \quad \wedge \quad P(X \geq \tilde{x}_\alpha) \geq 1 - \alpha.$$

► Poznámky

- Medián dělí obor náhodné veličiny na dvě zhruba stejně pravděpodobné části.
- Teoretický kvantil dělí obor náhodné veličiny na části, z nichž „nižší“ hodnoty teoreticky nastávají s pravděpodobností α a „vyšší“ s pravděpodobností $1 - \alpha$.
- Teoretický kvantil není vždy určen jednoznačně (diskrétní náhodná veličina).



Kvantil spojitě náhodné veličiny

- ▶ Na rozdíl od diskrétní náhodné veličiny lze kvantil $\tilde{x}_\alpha$ spojitě náhodné veličiny určit většinou jednoznačně (platí pro všechna námi probíraná rozdělení spojitě náhodné veličiny), tedy lze jednoznačně určit hodnotu $\tilde{x}_\alpha$ takovou, že:

$$F(\tilde{x}_\alpha) = P(X \leq \tilde{x}_\alpha) = \int_{-\infty}^{\tilde{x}_\alpha} f(x) dx = \alpha.$$

- ▶ Poznámky

- ▶ Kvantilovou funkci je možno chápat jako inverzní funkci k funkci distribuční (F^{-1}).
- ▶ Kvantil $\tilde{x}_\alpha$ je možno chápat jako hodnotu inverzní funkce k funkci distribuční v bodě α ($F^{-1}(\alpha)$), tedy
- ▶ $\alpha \cdot 100\%$ kvantil je hodnota inverzní distribuční funkce v bodě α :

$$\tilde{x}_\alpha = F^{-1}(\alpha).$$

- ▶ Kvantily některých základních rozdělení mají zavedené speciální značení (u_α , $t_\alpha(n)$, ...).



Grafická vizualizace rozdělení náhodných veličin

- ▶ Polygon četností – vizualizace diskrétního rozdělení
 - ▶ na vodorovnou osu se vynášejí hodnoty náhodné veličiny x_i
 - ▶ na svislou osu pravděpodobnost
 - ▶ nad jednotlivými hodnotami náhodné veličiny x_i jsou vynášeny hodnoty příslušné pravděpodobnosti p_i
 - ▶ jednotlivé vynesené pravděpodobnosti jsou navíc spojeny lomenou čarou
- ▶ Histogram četností – vizualizace diskrétního či spojitého rozdělení
 - ▶ podobné jako u polygonu četností
 - ▶ možnost shlukovat hodnoty do intervalů např. z důvodu srovnání se skutečností (viz třídění statistického znaku)
- ▶ Graf distribuční funkce
- ▶ Graf hustoty pravděpodobnosti

