6. T y p y   r o z d ě l e n í

Poznámka: V odst. 5.5 - 5.10 jsme uvedli příklady náhodných veličin a jejich distribučních funkcí. Poznali jsme, že se od sebe liší vlastnostmi distribučních funkcí. Uvedeme nyní typy rozdělení náhodné veličiny, jejich charakteristiky a vztahy pro výpočet pravděpodobností.

I. D i s k r é t n í   r o z d ě l e n í

6.1. Definice: Diskrétní rozdělení, pravděpodobnostní funkce. Říkáme, že náhodná veličina X má *diskrétní rozdělení,* jestliže nabývá pouze diskrétních hodnot. Funkce p, která je definována vztahem

$$p(x) = P(X = x), \quad x \in \boldsymbol{R},$$

se nazývá *pravděpodobnostní funkce.*

6.2. Věta: Vlastnosti pravděpodobnostní funkce. Jestliže má náhodná veličina X diskrétní rozdělení a p je její pravděpodobnostní funkce, pak platí:
a) Náhodná veličina X nabývá konečně nebo nejvýše spočetně mnoha hodnot. Ty tvoří konečnou nebo nekonečnou posloupnost
$M = \{x_i\} = \{x_1, x_2, \ldots\}$.
b) Je $0 \leq p(x) \leq 1, \quad x \in \boldsymbol{R}$.
c) $p(x) > 0 \Leftrightarrow x \in M$.
d) $\sum\limits_{x \in M} p(x) = 1$.
e) Distribuční funkce F je po úsecích konstatní. Body nespojiosti jsou pouze v bodech množiny M a pro $x_i \in M$ je
$F(x_i) - F(x_i-) = p(x_i) = P(X = x_i)$.

Pro distribuční funkci platí vztah:

$$F(x) = P(X \leq x) = \sum_{x_i \in M,\, x_i \leq x} p(x_i).$$

Poznámka: Pro řady z d) a e) používáme obvykle značení

$$\sum_{x\in \boldsymbol{R}} p(x) = \sum_{x\in M} p(x) \quad \text{a} \quad \sum_{x_i\in M,\, x_i\le x} p(x_i) = \sum_{z\le x} p(z),$$

kde sčítáme pouze kladné hodnoty argumentu.

Je zřejmé, že pravděpodobnostní funkce p je úplnou charakteristikou diskrétního rozdělení, která je jednodušší než distribuční funkce. Používáme ji proto k popisu náhodné veličiny častěji. Stačí tedy takovou náhodnou veličinou zadat posloupností $M = \{x_i\}$ hodnot, kterých náhodná veličina nabývá a jejich pravděpodobnostmi. Obvykle tak činíme pomocí tabulky.

x	x_1	x_2	$\dots$	x_k	$\dots$
$p(x)$	$p(x_1)$	$p(x_2)$	$\dots$	$p(x_k)$	$\dots$

6.3. Příklad: Zapišme do tabulky pravděpodobnostní funkce náhodných veličin z příkladů 5.5, 5.6 a 5.8.

a) Podle zadání je $X \in \{0,1\}$ a $P(X=0) = p(0) = 1-p$, $P(X=1) = p(1) = p$. Je tedy

x	0	1
$p(x)$	$1-p$	p

b) Ze zadání je $X \in \{0,1,2,\dots,n\}$ a $P(X=k) = p(k) = \binom{n}{k} p^k(1-p)^{n-k}$, $0 \le k \le n$. Tudíž

x	0	1	$\dots$	k	$\dots$	n
$p(x)$	$(1-p)^n$	$np(1-p)^{n-1}$	$\dots$	$\binom{n}{k} p^k(1-p)^{n-k}$	$\dots$	p^n

c) Ze zadání je $X \in \boldsymbol{N}$ a $P(X=k) = p(k) = p(1-p)^{k-1}$, $k = 0,1,2,\dots$. Tedy

x	1	2	$\dots$	k	$\dots$
$p(x)$	p	$p(1-p)$	$\dots$	$p(1-p)^{k-1}$	$\dots$

II. Spojité rozdělení

Poznámka: V příkladě 5.9 jsme uvedli příklad rozdělení, kdy byla distribuční funkce spojitá. V tomto případě je ale pravděpodobnost toho, že náhodná veličina nabývá jediné hodnoty rovna nule. Uvedeme vhodnější charakteristiku takových rozdělení než je distribuční funkce.

6.4. Definice: Spojité rozdělení. Říkáme, že náhodná veličina X má *spojité rozdělení,* jestliže existuje funkce f taková, že pro distribuční funkci F náhodné veličiny X platí

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t,\ \ x \in \boldsymbol{R}.$$

Funkci f nazýváme *hustotou rozdělení* náhodné veličiny X.

6.5. Věta: Vlastnosti hustoty. Reálná funkce $f : \boldsymbol{R} \to \boldsymbol{R}$ je hustotou rozdělení náhodné veličiny X, jestliže platí:

a) $f(x) \geq 0$ pro $x \in \boldsymbol{R}$; b) $\displaystyle\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1.$

Dále platí:

c) $F'(x) = f(x)$ pro skoro všechna $x \in \boldsymbol{R}$;

d) Pro $A \subset \boldsymbol{R}$ je $P(X \in A) = \int_A f(x)\,\mathrm{d}x$, speciálně je

$P(a < X < b) = \displaystyle\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a).$

III. Smíšené rozdělení

Poznámka: Distribuční funkce náhodné veličiny z příkladu 5.10 má distribuční funkci, která má jednak skoky v některých bodech a je spojitá a rostoucí v některých intervalech. Pro takové rozdělení se výpočet pravděpodobnosti provádí podle vzorců, které jsou sloučením vzorců z 6.2 a 6.5.

6.6. Definice: Smíšené rozdělení. Říkáme, že náhodná veličina X má *smíšené rozdělení,* jestliže je její distribuční funkce F nespojitá a pro její distribuční funkci platí:

$$F(x) = \sum_{t \leq x} (F(t) - F(t-)) + \int_{-\infty}^{x} F'(t)\,\mathrm{d}t,\ \ x \in \boldsymbol{R}.$$

Číselné charakteristiky náhodné veličiny.

Charakteristika polohy

Poznámka: Distribuční funkce nebo hustota či pravděpodobnostní funkce jsou úplným popisem rozdělení náhodné veličiny. V některých případech používáme k popisu jednodušších charakteristik, čísel, které v některých případech k popisu stačí.

6.7. Definice: Střední hodnota. Je-li X náhodná veličina, pak vážený průměr jejích hodnot podle pravděpodobnosti nazýváme *střední hodnotou* náhodné veličiny X a označujeme ji $E(X)$. Potom:

a) pro spojité rozdělení s hustotou f je $E(X) = \int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x$;

b) pro diskrétní rozdělení s pravděpodobnostní funkcí p je $E(X) = \sum\limits_{x \in \boldsymbol{R}} xp(x)$;

c) pro smíšené rozdělení s distribuční funkcí F je $E(X) = \sum\limits_{x \in \boldsymbol{R}} x[F(x) - F(x-)] + \int_{-\infty}^{\infty} xF'(x)\,\mathrm{d}x$,

pokud hodnoty ze vzorců existují.

6.8. Věta: Vlastnosti střední hodnoty. Pro střední hodnotu $E(X)$ náhodné veličiny X platí:

a) Je-li $X = a$, pak $E(X) = a$.

b) Je $E(\alpha X + \beta) = \alpha E(X) + \beta$.

c) Je-li $X \geq a > -\infty$, pak $E(X) \geq a$, je-li $X \leq b < \infty$, je $E(X) \leq b$, tedy pro $-\infty < a \leq X \leq b < \infty$ je $a \leq E(X) \leq b$.

Pro náhodné veličiny X a Y je $E(X+Y) = E(X) + E(Y)$.

Jsou-li nezávislé, pak i $E(XY) = E(X)E(Y)$.

Míra variability

6.9. Definice: Rozptyl a směrodatná odchylka. Je-li X náhodná veličina se střední hodnotou $E(X)$, pak hodnotu

$$D(X) = E([X - (E(X)]^2)$$

nazýváme *rozptylem* náhodné veličiny X. Hodnotu

$$\sigma(X) = \sqrt{D(X)}$$

nazýváme její *směrodatnou odchylkou.*

6.10. Věta: Vlastnosti rozptylu. Pro rozptyl náhodné veličiny X platí:

a) Je-li $X = a$, pak je $D(X) = 0$.
b) Pro náhodnou veličinu, která není konstantní je $D(X) > 0$.
c) Je $D(X) = E(X^2) - (E(X))^2$.
d) Je $D(\alpha X + \beta) = \alpha^2 D(X)$.
e) Jsou-li X a Y nezávislé náhodné veličiny, je $D(X + Y) = D(X) + D(Y)$.

6.11. Věta: Vzorec pro výpočet rozptylu. Rozptyl $D(X)$ náhodné veličiny X vypočteme podle vzorce:

a) má-li X spojité rozdělení s hustotou f, pak

$$D(X) = \int_{-\infty}^{\infty} (x - E(X))^2 f(x)\,\mathrm{d}x, \quad E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x;$$

b) má-li X diskrétní rozdělení s pravděpodobnostní funkcí p, pak

$$D(X) = \sum_{x \in \boldsymbol{R}} (x - E(X))^2 p(x), \quad E(X^2) = \sum_{x \in \boldsymbol{R}} x^2 p(x);$$

c) pro smíšené rozdělení s distribuční funkcí F je

$$D(X) = \sum_{x \in \boldsymbol{R}} (x - E(X))^2 [F(x) - F(x-)] + \int_{-\infty}^{\infty} (x - E(X))^2 F'(x)\,\mathrm{d}x,$$

$$E(X^2) = \sum_{x \in \boldsymbol{R}} x^2 [F(x) - F(x-)] + \int_{-\infty}^{\infty} x^2 F'(x)\,\mathrm{d}x,$$

pokud mají vzorce smysl.

Poznámka: V rozptylu sledujeme rozložení kvadrátů odchylek od střední hodnoty. Skutečné odchylky od střední hodnoty zachycuje směrodatná odchylka, která je v měřítku a v jednotkách v jakých jsou hodnoty náhodné veličiny X.