

## 7. Náhodné veličiny a jejich charakteristiky

Příkladem náhodné veličiny je měření vzdálenosti - sčítá se zde nepřesnost měření, chyby přístrojů i samotného geodeta (viz chyby - systematické, hrubé a náhodné).

**Náhodná veličina** - mění své hodnoty v závislosti na náhodě. Značí se velkým písmenem. Existují dva typy:

- Diskrétní náhodná veličina - Může nabývat konečně nebo spočetně mnoha hodnot
- Spojitá náhodná veličina - Může nabývat libovolné hodnoty z určitého intervalu

### 7.1. Diskretní náhodná veličina

Pravděpodobnostní chování náhodné veličiny  $X$  je dáno rozdělením pravděpodobnosti náhodné veličiny  $X$ . V případě diskrétní náhodné veličiny je rozdělení dáno:

1. Výtčtem hodnot  $x$ , kterých může náhodná veličina  $X$  nabývat
2. Pravděpodobnostmi, s jakými může veličina těchto hodnot z množiny  $I$  nabývat, tj.  $P(X = x), x \in I$ .

*Příklad:* Uvažujme náhodnou veličinu  $X$ , která označuje počet líců ve třech po sobě jdoucích nezávislých hodech mincí. Jaké je její rozdělení?

$X$  může nabývat hodnot  $\{0, 1, 2, 3\}$  (v dalším textu L = líc, R = rub ;-)

**$X = 0$ :** Možnost RRR. Pravděpodobnost, že v prvním hoďu padne rub, je  $\frac{1}{2}$ , stejně jako pravděpodobnost, že padne v druhém a třetím (jevů jsou nezávislé). Proto je pravděpodobnost varianty RRR  $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$ , tzn.  $P(X = 0) = \frac{1}{8}$ .

**$X = 1$ :** Možnost LRR, RLR nebo RLL. Každý z těchto jevů má pravděpodobnost  $\frac{1}{8}$  (viz výše). Proto platí  $P(X = 1) = 3 \cdot \frac{1}{8} = \frac{3}{8}$ .

**$X = 2$ :** Padnou dva líce, tj. možnosti LLR, LRL nebo RLL. Každý z těchto jevů má pravděpodobnost  $\frac{1}{8}$  (viz výše). Proto platí  $P(X = 2) = 3 \cdot \frac{1}{8} = \frac{3}{8}$ .

**$X = 3$ :** Možnost LLL, pravděpodobnost stejná jako v prvním případě, tj.  $P(X = 3) = \frac{1}{8}$ .

$x$	0	1	2	3
$P(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Všimněte si, že součet všech pravděpodobností je jedna. Tato podmínka musí být splněna vždy, neboť různé hodnoty  $x \in I$  jsou jevy disjunkční a jejich sjednocení musí tvořit jev jistý  $\Omega$ , tedy platí:

$$\sum_{i \in I} P(X = i) = P(\Omega)$$

## 7.2. Charakteristiky diskrétní náhodné veličiny

**Střední hodnota náhodné veličiny** - charakterizuje polohu hodnot náhodné veličiny. Značí se  $E(X)$  a lze ji určit jako:

$$E(X) = \sum_{i \in I} x \cdot P(X = x),$$

kde  $I$  je množina hodnot, jichž náhodná veličina  $X$  nabývá.

*Příklad:* Určete střední hodnotu náhodné veličiny z předchozího příkladu.

Řešení:  $E(X) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{12}{8} = 1,5$

Nechť hodnota náhodné veličiny představuje polohu hmotného bodu na ose a hmotnost tohoto bodu odpovídá pravděpodobnosti. Pak velikost střední hodnoty odpovídá poloze těžiště od bodu nula.

Uvažujme nyní pokus, jehož výsledek je realizován jako náhodná veličina  $X$ . Jestliže pokus provedeme  $n$ -krát, získáme vektor  $(x_1, x_2, \dots, x_n) \cdot x_i$ , tedy počet líců v  $i$ -té trojici hodů. Pak

$$\bar{x} = \frac{\sum x_i}{n} \cong E(X),$$

tedy střední hodnota odpovídá dlouhodobému (sic!) aritmetickému průměru. Ale to platí jen pro velké hodnoty  $n$ !

Předložená úvaha je použitelná u hazardních her, kde uvažovanou náhodnou veličinou je zisk při jedné hře. Jestliže platí  $\sum x_i/n \cong E(X)$ , pak celkový zisk v  $n$  hrách se rovná přibližně  $n \cdot E(X)$  (tím si nejsem úplně jistej, buď pan docent kecá, nebo jsem si to špatně napsal...).

Na střední hodnotě záleží, zda je hra:

- Spravedlivá ...  $E(X) = 0$
- Výhodná ...  $E(X) > 0$
- Spravedlivá ...  $E(X) < 0$

*Příklad:* Hráč hází kostkou. Padne-li 6 ok, získá 6 Kč, pokud padne jiný počet ok, ztrácí korunu. Je hra výhodná? Řešení: Hráčův zisk má rozdělení:

$x$	-1	6
$P(X = x)$	$\frac{5}{6}$	$\frac{1}{6}$

$$E(X) = (-1) \cdot \frac{5}{6} + 6 \cdot \frac{1}{6} = \frac{1}{6} > 0$$

**Rozptyl** (též kolísavost nebo variabilita) vyjadřuje rozptýlenost hodnot. Značí se  $\text{Var}(X)$  nebo  $D(X)$  a platí:

$$\text{Var } X = \sum_{x \in I} [x - E(x)]^2 \cdot P(X = x) = \left[ \sum_{x \in I} x^2 \cdot P(X = x) \right] - E^2(x).$$

**Směrodatná odchylka** - definuje se jako druhá odmocnina z rozptylu, značí se  $SD(X)$  (z anglického standard deviation) nebo  $\sigma(X)$ .

$$SD(X) = \sqrt{\text{Var}(X)}$$

*Příklad:* Spočítejte rozptyl náhodné veličiny  $X$  označující počet líců ve třech po sobě jdoucích nezávislých hodech mincí (viz první příklad).

$$\text{Var}(X) = 0^2 \cdot \left(\frac{1}{8}\right) + 1^2 \cdot \left(\frac{3}{8}\right) + 2^2 \cdot \left(\frac{3}{8}\right) + 3^2 \cdot \left(\frac{1}{8}\right) - 1,5^2 = 0,75$$

Uvažujme nyní dvě hry: V první hře házíme mincí a získáme 1 Kč, jestliže padne líc a ztrácíme 1 Kč, jestliže padne rub. Ve druhé hře jsou pravidla stejná, hraje se ale o 10 Kč. Označme zisk v první hře jako  $X_1$  a zisk ve druhé hře  $X_2$ . Jaké je rozdělení náhodných veličin  $X_1$  a  $X_2$ ?

$x$	-1	1
$P(X_1 = x)$	$\frac{1}{2}$	$\frac{1}{2}$

$x$	-1	1
$P(X_2 = x)$	$\frac{1}{2}$	$\frac{1}{2}$

Obě hry jsou spravedlivé, ježto platí  $E(X_1) = 0$  a  $E(X_2) = 0$ . Sledujme nyní celkový zisk během  $n$  her, tj.  $\sum X_{1i}$ , resp.  $\sum X_{2i}$ , kde  $X_{1i}$  je zisk v  $i$ -tém opakování první hry a  $X_{2i}$  je zisk v  $i$ -tém opakování druhé hry.

$$\begin{aligned} \text{Var}(X_1) &= (-1)^2 \cdot \left(\frac{1}{2}\right) + 1^2 \cdot \left(\frac{1}{2}\right) = 1 \\ \text{Var}(X_2) &= (-10)^2 \cdot \left(\frac{1}{2}\right) + 10^2 \cdot \left(\frac{1}{2}\right) = 100 \end{aligned}$$

Z hodnot rozptylu je zřejmé, že v první hře můžeme dlouhodobě udržovat vyrovnané skóre. Protože však rozptyl hodnot ve druhé hře je velký, je to buď anebo - buď vyhraje velkou částku, nebo velkou částku prohrajeme, ale rozhodně neudržíme vyrovnané skóre.

### 7.3. Binomické rozdělení

Náhodná veličina  $X$  má binomické rozdělení  $\text{Bi}(n, p)$  s parametry  $n$  a  $p$ , jestliže nabývá hodnot  $x = 0, 1, 2, \dots, n$  s pravděpodobnostmi

$$P(X = x) = \binom{n}{x} \cdot p^x \cdot (1 - p)^{n-x}, \quad n \in \mathbb{N}, \quad p \in (0, 1)$$

Používá se tehdy, když určitý pokus má pouze dva možné výsledky - úspěch s pravděpodobností  $p$  a neúspěch s pravděpodobností  $(1 - p)$ . Počet úspěchů v sérii  $n$  po sobě jdoucích pokusů se řídí binomickým rozdělením. Jeho charakteristiky jsou:

**Střední hodnota:**  $E(X) = n \cdot p$

**Rozptyl:**  $\text{Var}(X) = n \cdot p \cdot (1 - p)$

Binomickému rozdělení s parametrem  $p = 1$  říkáme alternativní rozdělení a značíme jej  $A(X)$ . Rozdělení náhodné veličiny  $X$  řídicí se alternativní rozdělením  $A(X)$  je dáno:

$x$	0	1
$P(X = x)$	$1 - p$	$p$

Zřejmě zde tedy platí, že  $E(X) = p$  a  $\text{Var}(X) = p \cdot (1 - p)$  (ověřte!).

*Příklad:* Házáme třikrát po sobě mincí. Určete rozdělení náhodné veličiny  $X$  označující počet líců (tj. je to první příklad, jen jej neřešíme úvahou, ale použitím binomického rozdělení).

Úspěch = líc,  $p = \frac{1}{2}$ ,  $n = 3$ .

$$\text{Dosazení: } P(X = 0) = \binom{3}{0} \cdot \left(\frac{1}{2}\right)^0 \cdot \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$

$$P(X = 1) = \binom{3}{1} \cdot \left(\frac{1}{2}\right)^1 \cdot \left(\frac{1}{2}\right)^2 = \frac{3}{8}$$

$$P(X = 2) = \binom{3}{2} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^1 = \frac{3}{8}$$

$$P(X = 3) = \binom{3}{3} \cdot \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^0 = \frac{1}{8}$$

Všimněte si, že mocnina za kombinačním číslem představuje úspěšnou variantu a poslední závorka neúspěch. Celé schéma je podobné práci s binomickou větou - snižování jedné a zvyšování druhé mocniny.

#### 7.4. Poissonovo rozdělení

Náhodná veličina  $X$  má Poissonovo rozdělení s paramterem  $\lambda > 0$ , jestliže platí

$$P(X = x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$$

pro  $x = 0, 1, \dots$  a  $P(X = x) = 0$  všude jinde. Paramter  $\lambda$  se nazývá intenzita, neboť  $E(X) = \lambda$  a  $\text{Var}(X) = \lambda$ . Typickým příkladem veličiny, která má Poissonovo rozdělení je počet radioaktivních pulsů zaznamenaných Geiger-Müllerovým počítačem během nějakého časového intervalu  $(0, t)$ . Poissonovým procesem se modeluje např. počet příchozích nebo obsluhovaných zákazníků, počet vozidel, která projedou určitým úsekem za jednotku času a další.

*Příklad:* K holiči chodí “v průměru” čtyři zákazníci za hodinu. S jakou pravděpodobností přijde během půl hodiny alespoň jeden zákazník?

Náhodná veličina  $X$  = počet zákazníků, kteří přijdou během půl hodiny,  $\lambda = 2 \dots$  počet během půl hodiny.

$$P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-2} \cong 1 - 0,135 = 0,865$$