Náhodná veličina

Nechť $(\Omega, \mathcal{A}, P)$ je pravděpodobnostní prostor. Zobrazení $X : \Omega \to \mathbb{R}$ se nazývá **náhodná veličina** (vzhledem k jevovému poli $\mathcal{A}$), právě když

$$\forall B \in \mathcal{B} : \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}, \tag{1}$$

tj. úplný vzor každé borelovské množiny je jevem.

Poznámka: Obraz $X(\omega)$ se nazývá **číselná realizace** náhodné veličiny X příslušná k možnému výsledku ω. Množinu $\{\omega \in \Omega : X(\omega) \in B\}$ zkráceně zapisujeme $\{X \in B\}$ nebo $(X \in B)$

Distribuční funkce

Funkce $F : \mathbb{R} \to \mathbb{R}$ definována vztahem

$$\forall x \in \mathbb{R} : F(x) = P(X \le x), \tag{2}$$

kde $P(X \le x)$ značí $P(\{\omega \in \Omega : X(\omega) \le x\})$, se nazývá **distribuční funkce** náhodné veličiny X (vzhledem k P).

Distribuční funkce má tyto vlastnosti:

1. je neklesající, tj. pro všechna $x_1 < x_2 : F(x_1) \le F(x_2)$;
2. je zprava spojitá, tj. pro všechna $x_0 \in \mathbb{R} : \lim\limits_{x \to x_0^+} F(x) = F(x_0)$;
3. $0 \le F(x) \le 1$;
4. Pro $x_0 \in \mathbb{R}$ libovolné, ale pevně zvolené je
$$P(X = x_0) = F(x_0) - \lim_{x \to x_0^-} F(x);$$
5. Pro $a, b \in \mathbb{R}, a < b$ je $P(a < X \le b) = F(b) - F(a)$.

Diskrétní náhodná veličina

Náhodná veličina X s distribuční funkcí $F(x)$ se nazývá **diskrétní**, jestliže existuje neprázdná, nejvýše spočetná podmnožina N reálných čísel a funkce $\pi : \mathbb{R} \to \mathbb{R}$ s těmito vlastnostmi:

1. $\pi(x) > 0$ pro $x \in N$
   $\pi(x) = 0$ pro $x \in \mathbb{R} - N$
2. $\sum\limits_{x \in \mathbb{R}} \pi(x) = \sum\limits_{x \in N} \pi(x) = 1$

a pro každé $x \in \mathbb{R}$ máme $F(x) = \sum\limits_{t \le x} \pi(t)$. Funkce $\pi(x)$ se nazývá **pravděpodobnostní funkce náhodné veličiny** X. Platí $\pi(x) = P(X = x)$.

1. Střelec střílí do terče až do prvního zásahu. Má v zásobě čtyři náboje. Pravděpodobnost zásahu je při každém výstřelu 0,6. Náhodná veličina X udává počet nespotřebovaných nábojů. Určete pravděpodobnostní a distribuční funkcí náhodné veličiny X a nakreslete grafy těchto funkcí.

2. Náhodná veličina X má pravděpodobnostní funkci

$$\pi(x) = P(X = x) = \begin{cases} \frac{3}{7} \cdot 0,7^x & x = 1, 2, \ldots \\ 0 & jinak \end{cases}$$

Jaká je pravděpodobnost, že tato náhodná veličina nabude hodnot:

(a) menších jak 3

(b) větších jak 4

(c) větších jak 1 a menších jak 4?

Spojitá náhodná veličina

Řekneme, že náhodná veličina X je (absolutně) **spojitá**, jestliže existuje nezáporná borelovská funkce f tak, že pro každé $x \in \mathbb{R}$ máme

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

Funkce f se nazývá **hustota** náhodné veličiny X a je určena jednoznačně až na borelovské množiny míry 0 a platí:

(a) $f(x) = \frac{\mathrm{d}F(x)}{\mathrm{d}x}$

(b) $\int\limits_{-\infty}^{\infty} f(t)\,\mathrm{d}t = 1$

3. Spojitá náhodná veličina má hustotu

$$f(x) = \begin{cases} ax & 0 \leq x \leq 1 \\ 0 & jinak \end{cases}$$

(a) Určete konstantu a.

(b) Vypočtěte pravděpodobnost, že x je větší jak $\frac{1}{3}$ a menší nebo rovno jak $\frac{2}{3}$.

(c) Určete distribuční funkci.

4. Určete $a \in \mathbb{R}$ tak, aby funkce

$$F(x) = \begin{cases} 0 & x < 0 \\ ax^2 & 0 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

byla distribuční funkcí.

Číselné charakteristiky náhodných veličin

Střední hodnota

Je-li dána diskrétní náhodá veličina X s pravděpodobnostní funkcí $\pi(x)$, pak číslo

$$E(X) = \sum_{x=-\infty}^{\infty} x \cdot \pi(x),$$

za předpokladu, že případná nekonečná řada absolutně konverguje, nazýváme střední hodnotou náhodné veličiny X.
Je-li náhodná veličina X spojitá s hustotou $f(x)$, pak číslo

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,\mathrm{d}x,$$

za předpokladu, že nevlastní Riemannův integrál absolutně konverguje, nazýváme její střední hodnotou.

Nechť $g(x)$ je borelovská funkce. Pak pro střední hodnotu náhodné veličiny $Y = g(X)$ platí:

$$E(Y) = \begin{cases} \sum\limits_{x\in\mathbb{R}} g(x)\pi(x) & \text{v diskrétním případě} \\ \int\limits_{-\infty}^{\infty} g(x)f(x)\,\mathrm{d}x & \text{ve spojitém případě} \end{cases}$$

pokud nekonečná řada, resp. Riemannův integrál, absolutně konvergují.

Rozptyl
Číslo

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

nazýváme rozptylem náhodné veličiny X za předpokladu, že všechny uvedené střední hodnoty existují.
Číslo $\sqrt{D(X)}$ nazýváme **směrodatnou odchylkou** náhodné veličiny X.

Kovariance
Číslo

$$C(X_1, X_2) = E[(X_1 - E(X_1))(X_2 - E(X_2))]$$

nazýváme kovariancí náhodných veličin X_1 a X_2 za předpokladu, že všechny uvedené střední hodnoty existují. Je-li $C(X_1, X_2) = 0$, pak řekneme, že náhodné veličiny X_1 a X_2 jsou **nekorelované**.

Korelace
Číslo

$$R(X_1, X_2) = E\left(\frac{X_1 - E(X_1)}{\sqrt{D(X_1)}} \cdot \frac{X_2 - E(X_2)}{\sqrt{D(X_2)}}\right),$$

$D(X_1)D(X_2) \neq 0$ nazveme korelací náhodných veličin X_1 a X_2 (a za předpokladu, že všechny střední hodnoty existují), $R(X_1, X_2) = 0$ jinak.

Vlastnosti:
Nechť a, a_1, a_2, b, b_1, b_2 jsou konstanty a $X_1, \dots X_n, Y_1, \dots Y_n$ náhodné veličiny definované na témže pravděpodobnostním prostoru.

(a) Střední hodnota

i. $E(a) = a$
ii. $E(a + bX) = a + bE(X)$
iii. $E(X - E(X)) = 0$
iv. $E(\sum\limits_{i=1}^{n} X_i) = \sum\limits_{i=1}^{n} E(X_i)$
v. Jsou-li náhodné veličiny $X_1, \dots X_n$ stochasticky nezávislé, pak: $E(\prod\limits_{i=1}^{n} X_i) = \prod\limits_{i=1}^{n} E(X_i)$

(b) Rozptyl

i. $D(a) = 0$
ii. $D(a + bX) = b^2 D(X)$
iii. $D(\sum\limits_{i=1}^{n} X_i) = \sum\limits_{i=1}^{n} D(X_i) + 2\sum\limits_{i=1}^{n-1}\sum\limits_{j=i+1}^{n} C(X_i, X_j)$. Jsou-li veličiny $X_1, \dots X_n$ nekorelované, pak platí: $D(\sum\limits_{i=1}^{n} X_i) = \sum\limits_{i=1}^{n} D(X_i)$

(c) Kovariance

i. $C(a_1, X_2) = C(X_1, a_2) = C(a_1, a_2) = 0$
ii. $C(a_1 + b_1X_1, a_2 + b_2X_2) = b_1b_2C(X_1, X_2)$
iii. $C(X, X) = D(X)$
iv. $C(X_1, X_2) = C(X_2, X_1)$
v. $C(X_1, X_2) = E(X_1X_2) - E(X_1)E(X_2)$
vi. $C(\sum\limits_{i=1}^{n} X_i, \sum\limits_{j=1}^{m} Y_j) = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} C(X_i, Y_j)$

(d) Korelace

i. $R(a_1, X_2) = R(X_1, a_2) = R(a_1, a_2) = 0$
ii. $R(a_1 + b_1X_1, a_2 + b_2X_2) = \operatorname{sgn}(b_1b_2)R(X_1, X_2)$
iii. $R(X_1, X_2) = R(X_2, X_1)$
iv. $R(X_1, X_2) = \frac{C(X_1,X_2)}{\sqrt{D(X_1)}\sqrt{D(X_2)}}$, $D(X_1)D(X_2) \neq 0$, $R(X_1, X_2)$ jinak.

5. Náhodná veličina X je dána pravděpodobnostní funkcí:

$$\pi(x) = \begin{cases} 1/3 & x = -2 \\ 1/2 & x = 3 \\ 1/6 & x = 1 \\ 0 & \text{jinak.} \end{cases}$$

Určete $E(X)$, $E(2X + 5)$, $E(X^2)$, $D(X)$ a $D(2X - 1)$.

6. Určete střední hodnotu a rozptyl náhodné veličiny X, pokud má

   (a) alternativní rozdělení: $X \sim A(\theta)$;

   (b) binomické rozdělení: $X \sim Bi(n, \theta)$;

   (c) rovnoměrně spojité rozdělení: $X \sim Rs(a, b)$.

7. Náhodná veličina udává počet ok při hodu kostkou. Vypočtěte její rozptyl.

8. Náhodná veličina X má konstatní hodnotu pravděpodobnosti v intervalu $(0, a)$, to znamená, že její hustota pravděpodobnosti má tvar:

$$f(x) = \begin{cases} \frac{1}{a} & \text{pro } 0 < x < a \\ 0 & \text{jinak} \end{cases}$$

   S použitím vlastností střední hodnoty a rozptylu určete

   (a) $E(2X + 3)$

   (b) $E(3X^2 - 2X + 1)$

   (c) $D(2X + 3)$

   (d) $D(X^2 + 1)$

9. Nekorelované náhodné veličiny X a Y mají rozptyly $D(X) = a$ a $D(Y) = 2$. Určete konstantu a, jestliže roztpyl náhodné veličiny $Z = 3Y - X$ je $D(Z) = 25$.

10. Náhodné veličiny X a Y jsou náhodné chyby na vstupu nějakého zařízení, mají charakteristiky $E(X) = -2$, $E(Y) = 4$, $D(X) = 4$ a $D(Y) = 9$. Koeficient korelace těchto chyb je $R(X, Y) = -0,5$. Chyba Z na výstupu závisí na chybách na vstupu následovně: $Z = 3X^2 + 2XY + Y^2 - 3$. Najděte střední hodnotu chyby na výstupu.

Domácí úkol: Znáte-li charakteristiky náhodných veličin X a Y, určete následující charakteristiky:

(a) $E(2X - Y + 4)$, [4]

(b) $D(2X - Y + 4)$, [21]

(c) $C(X + Y, X - Y)$, [3]

(d) $E[(X + Y)^2]$, [12]

(e) $E(3X^2 - 2Y)$. [11]

$E(X) = 1$, $E(Y) = 2$, $D(X) = 4$, $D(Y) = 1$, $C(X, Y) = -1$.