Transformace náhodné veličiny

Borelovské funkce: Zobrazení $g : \mathbb{R}^n \to \mathbb{R}^m$ nazýváme **borelovskou funkcí**, právě když úplný vzor každé borelovské množiny je opět borelovská množina, tj.

$$\forall B \in \mathcal{B}^m; \{(x_1, \ldots, x_n) \in \mathbb{R}^n; (g_1(x_1, \ldots, x_n), \ldots, g_m(x_1, \ldots, x_n)) \in B\} \in \mathcal{B}^n$$

Pozn.: Jedná se zejména o funkce spojité.

Z dané náhodné veličiny $X : \Omega \to \mathbb{R}$ vytvoříme pomocí borelovské funkce $g : \mathbb{R} \to \mathbb{R}$ zobrazení $Y : \Omega \to \mathbb{R}$, dané takto:

$$\forall \omega \in \Omega; Y(\omega) = g(X(\omega)).$$

Toto zobrazení je opět náhodná veličina a nazývá se **transformovaná** náhodná veličina.

Náhodná veličina X má distribuční funkci $F(x)$ a pravděpodobnostní funkci $\pi(x)$ nebo hustotu $f(x)$. Náhodná veličina $Y = g(X)$ má distribuční funkci $F_*(y)$ a pravdepodobnostní funkci $\pi_*(y)$ nebo hustotu $f_*(y)$. Označíme τ inverzní funkci k funkci g, pak platí:

1. X je diskrétní náhodná veličina

$$\pi_*(y) = P(Y = y) = P(g(X) = y) = P(X = \tau(y)) = \pi(\tau(y))$$

2. X je spojitá náhodná veličina, g je rostoucí funkce a nechť existuje derivace funkce τ

$$F_*(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \leq \tau(y)) = F(\tau(y))$$
$$f_*(y) = \frac{\mathrm{d}F_*(y)}{\mathrm{d}y} = f(\tau(y))\frac{\mathrm{d}\tau(y)}{\mathrm{d}y}$$

3. X je spojitá náhodná veličina, g je klesající funkce a nechť existuje derivace funkce τ

$$F_*(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \geq \tau(y)) = 1 - P(X \leq \tau(y)) = 1 - F(\tau(y))$$
$$f_*(y) = \frac{\mathrm{d}F_*(y)}{\mathrm{d}y} = -f(\tau(y))\frac{\mathrm{d}\tau(y)}{\mathrm{d}y}$$

Celkem pro spojitou náhodnou veličinu X platí:

$$f_*(y) = \frac{\mathrm{d}F_*(y)}{\mathrm{d}y} = f(\tau(y))\left|\frac{\mathrm{d}\tau(y)}{\mathrm{d}y}\right|$$

1. Náhodná veličina X má normální rozdělení, tedy $X \sim N(\mu, \sigma^2)$, provedem transformaci $Y = a + bX$, kde $b \neq 0$. Určete $f_*(y)$.

2. Jaká je pravděpodobnost, že náhodná veličina X, která má rozdělení $N(20, 16)$, nabude hodnoty:

a) menší než 16

b) větší než 20

c) v mezích od 12 do 28

d) menší než 12 nebo větší než 28?

3. Náhodná veličina Y je funkcí náhodné veličiny X (tedy $Y = g(X)$). Určete, čemu se rovná hustota pravděpodobnosti $f_*(y)$ jestliže platí:

$$f(x) = \begin{cases} 2xe^{-x^2} & \text{pro } x > 0 \\ 0 & \text{jinak} \end{cases} \quad ; \qquad Y = X^2$$

4. Hledáme hustotu pravděpodobnosti náhodné veličiny $Y = X_1 + X_2$, pro $x_1, x_2 \in \mathbb{R}$, jestliže známe sdruženou hustotu pravděpodobnosti $f(x_1, x_2)$.

5. Hledáme pravděpodobnostní funkci náhodné veličiny Y, jestliže platí:

$$\pi(x) = \begin{cases} \frac{\lambda^x}{x!}e^{-\lambda} & \text{pro } x = 0, 1, \ldots \ \lambda > 0 \\ 0 & \text{jinak} \end{cases} \quad ; \qquad Y = 4X$$

Číselné charakteristiky náhodných veličin

Střední hodnota

Je-li dána diskrétní náhodná 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 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.

1. Střední hodnota
   (a) $E(a) = a$
   (b) $E(a + bX) = a + bE(X)$
   (c) $E(X - E(X)) = 0$
   (d) $E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(Xi)$
   (e) Jsou-li náhodné veličiny $X_1, \dots X_n$ stochasticky nezávislé, pak: $E(\prod_{i=1}^{n} X_i) = \prod_{i=1}^{n} E(X_i)$
2. Rozptyl
   (a) $D(a) = 0$
   (b) $D(a + bX) = b^2 D(X)$
   (c) $D(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} D(X_i) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} C(X_i, X_j)$. Jsou-li veličiny $X_1, \dots X_n$ nekorelované, pak platí: $D(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} D(X_i)$
3. Kovariance
   (a) $C(a_1, X_2) = C(X_1, a_2) = C(a_1, a_2) = 0$
   (b) $C(a_1 + b_1X_1, a_2 + b_2X_2) = b_1b_2C(X_1, X_2)$
   (c) $C(X, X) = D(X)$
   (d) $C(X_1, X_2) = C(X_2, X_1)$
   (e) $C(X_1, X_2) = E(X_1X_2) - E(X_1)E(X_2)$
   (f) $C(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} C(X_i, Y_j)$
4. Korelace
   (a) $R(a_1, X_2) = R(X_1, a_2) = R(a_1, a_2) = 0$
   (b) $R(a_1 + b_1X_1, a_2 + b_2X_2) = \mathrm{sgn}(b_1b_2)R(X_1, X_2)$
   (c) $R(X_1, X_2) = R(X_2, X_1)$
   (d) $R(X_1, X_2) = \frac{C(X_1,X_2)}{\sqrt{D(X_1)}\sqrt{D(X_2)}}$, $D(X_1)D(X_2) \neq 0$, $R(X_1, X_2)$ jinak.

6. 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 určíme

1. $E(2X + 3)$

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

3. $D(2X+3)$

4. $D(X^2+1)$

7. Určete střední hodnotu a rozptyl náhodné veličiny X, která má Poissonovo rozdělení s parametrem λ.

8. Určete střední hodnotu a rozptyl náhodné veličiny X, která má exponenciální rozdělení s parametrem λ.