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.

1. Střední hodnota
   (a) $E(a) = a$
   (b) $E(a + bX) = a + bE(X)$
   (c) $E(X - E(X)) = 0$
   (d) $E(\sum\limits_{i=1}^{n} X_i) = \sum\limits_{i=1}^{n} E(X_i)$
   (e) Jsou-li náhodné veličiny $X_1, \dots X_n$ stochasticky nezávislé, pak: $E(\prod\limits_{i=1}^{n} X_i) = \prod\limits_{i=1}^{n} E(X_i)$
2. Rozptyl
   (a) $D(a) = 0$
   (b) $D(a + bX) = b^2 D(X)$
   (c) $D(\sum\limits_{i=1}^{n} X_i) = \sum\limits_{i=1}^{n} D(X_i) + 2\sum\limits_{i=1}^{n-1} \sum\limits_{j=i+1}^{n} C(X_i, X_j)$. Jsou-li veličiny $X_1, \dots X_n$ nekorelované, pak platí: $D(\sum\limits_{i=1}^{n} X_i) = \sum\limits_{i=1}^{n} D(X_i)$
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) $C(\sum\limits_{i=1}^{n} X_i, \sum\limits_{j=1}^{m} Y_j) = \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} C(X_i, Y_j)$
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.

1. 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 $D(2X - 1)$.
   [$E(X) = 1$, $E(2X + 5) = 7$, $E(X^2) = 6$, $D(X) = 5$ a $D(2X - 1) = 20$]

2. 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) Poissonovo rozdělení: $X \sim Po(\lambda)$
   (d) geometrické rozdělení: $X \sim Ge(\theta)$
   (e) Pascalovo rozdělení: $X \sim Ps(k, \theta)$.

3. Náhodná veličina udává počet ok při hodu kostkou. Vypočtěte její rozptyl.
   [35/12]

4. Pravděpodobnost vyrobení zmetku na automatické lince je 0,1. Vyrobením zmetku se linka zastaví. Jaká je střední hodnota a rozptyl počtu výrobků do zastavení linky?
   $[E(X) = 10,\ D(X) = 90]$

5. 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) $E(2X + 3)$ $[a + 3]$
   (b) $E(3X^2 - 2X + 1)$ $[a^2 - a + 1]$
   (c) $D(2X + 3)$ $[\frac{a^2}{3}]$
   (d) $D(X^2 + 1)$ $[\frac{4a^4}{45}]$

6. Určete střední hodnotu a rozptyl náhodné veličiny X, pokud má
   (a) rovnoměrné spojité rozdělení: $X \sim Rs(a, b)$
   (b) exponenciální rozdělení: $X \sim Ex(\lambda)$
   (c) normální rozdělení: $X \sim N(\mu, \sigma^2)$
   (d) Cauchyovo rozdělení: $X \sim t(1)$.
   (e) Gamma rozdělení: $X \sim \Gamma(p, \lambda)$
   (f) Pearsonovo rozdělení: $X \sim \chi^2(n)$

7. 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 $D(Z) = 25$.
   $[a = 7]$

8. 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.
   $[E(Z) = 24]$

9. 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]

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