Seria 7. Zmienne losowe

Zadania statystyczne

Zad 1 *Znaleźć minimum funkcji* $\mathbf{E}(X-t)^2$*, gdzie* X *jest zmienną mającą wariancję.*

Bywa, zmienne losowe Xi Y są zależne, a nas interesuje stopień korelacji między nimi. Definiuje się kowariancję zmiennych losowych $\mathbf{Cov}(X,Y) = \mathbf{E}[(X-\mathbf{E}X)(Y-\mathbf{E}Y)]$. W szczególności $\mathbf{D}^2X = \mathbf{Cov}(X,X)$. W statystyce do badania korelacji zmiennych wykorzystuje się współczynnik korelacji

$$\rho(X,Y) = \frac{\mathbf{Cov}(X,Y)}{\sqrt{\mathbf{D}^2X}\sqrt{\mathbf{D}^2Y}}.$$

Można pokazać, że $|\rho| \leqslant 1$. Oraz $|\rho| = 1$ tylko w wypadku linowej zależności między zmiennymi X, Y.

Zad 2 *Pokaż, że* $\rho(X,Y) = \rho(Y,X)$*,* $\rho(aX+b,Y) = \mathrm{sgn} a \cdot \rho(X,Y)$*, o ile* $a \neq 0$.

Zad 3 *(Zagadnienie portfela inwestycyjnego) Zmienne losowe* X_0 *i* X_1 *mają dane* $\mathbf{E}X_0 = 0.05$, $\mathbf{E}X_1 = 0,07$, $\sqrt{\mathbf{D}^2X_0} = 0.02$, $\sqrt{\mathbf{D}^2X_1} = 0.03$, $\rho(X_0,X_1) = -0.5$*. Niech* $X_t = tX_0 + (1-t)X_1$, $t \in [0,1]$*. Dla jakiego* t *wariancja jest minimalna?*

Zad 4 *Niech* X *i* Y *będą zmiennymi losowymi o skończonej wariancji,* $\mathbf{D}^2X \neq 0$*. Wyznaczyć liczby* a *i* b *takie, że* $\mathbf{E}(Y-aX-b)^2$ *była minimalna. Prostą o równaniu*

$$y = \frac{\mathbf{Cov}(X,Y)}{\mathbf{D}^2X}(x-\mathbf{E}X) + \mathbf{E}Y.$$

nazywamy prostą liniowej regresji.

Zadania ze zmiennych niezależnych

Zad 5 *Przypuśćmy, że zmienne* $X_1, ..., X_n$ *są niezależne (mają wspólną dystrybuantę* F*). Znajdź rozkład zmiennych* $Y = \max\{X_1,...,X_n\}$*,* $Z = \max\{Y_1,...,Y_n\}$*. W szczególności rozwiąż zadanie w przypadku, gdy zmienne pochodzą z rozkładu wykładniczego* $\mathcal{E}(\lambda)$.

Zad 6 *Niech* $U = \min\{X,Y\}$*,* $V = \max\{X,Y\} - \min\{X,Y\}$*, gdzie* X, Y *niezależne wykładnicze z* $\mathcal{E}(1)$*. Pokaż, że zmienne* U, V *są niezależne.*

Zad 7 *Załóżmy, że zmienne niezależne* X*,* Y *pochodzą z rozkładu Poissona, odpowiednio* $\mathcal{P}(\mu)$*,* $\mathcal{P}(\nu)$*. Pokaż, że zmienna* $X+Y$ *ma rozkład* $X+Y$ *ma rozkład* $\mathcal{P}(\mu+\nu)$.

Istnieje ogólna metoda szukania rozkładów sum zmiennych losowych niezależnych.

Twierdzenie 1 *Jeśli* X_1*,* X_2 *są niezależne wówczas* $\mu_{X_1+X_2} = \mu_{X_1} * \mu_{X_2}$*, gdzie działanie* $* : \mathcal{M}\times\mathcal{M} \to \mathcal{M}$*, (*$\mathcal{M}$ *-miary probabilistyczne) nazywa się splotem i definiuje następująco*

$$\mu_{X_1} * \mu_{X_2}(B) = \int_{-\infty}^{\infty} \mu_{X_1}(B-y)\mu_{X_2}(dy).$$

W przypadku rozkładów ciągłych μ_{X_1}*,* μ_{X_2} *mają gęstości* f_1*,* f_2*. Twierdzenie daje, że*

$$\mu_{X_1+X_2} = \mu_{X_1} * \mu_{X_2}(B) = \int_{-\infty}^{\infty} \mu_{X_1}(B-y)\mu_{X_2}(dy) = \int_B \int_{-\infty}^{\infty} f_1(u-y)f_2(y)dydu = \int_B (f_1 * f_2)(u)du,$$

gdzie $f_1 * f_2(u) = \int_{-\infty}^{\infty} f_1(u-y)f_2(y)dy$*, jest gęstością rozkładu zmiennej* X_1+X_2.

Zad 8 *Niech* X_1, X_2 *będą zmiennymi niezależnymi z rozkładu jednostajnego* $\mathcal{U}[0,1]$*. Znajdź gęstość zmiennej* X_1+X_2.

Zad 9 *Niech zmienne $X_1, ..., X_n$ są niezależne i mają rozkłady wykładnicze z parametrem λ. Pokaż, że zmienna $X_1 + ... + X_n$ ma rozkład o gęstości $\frac{\lambda^n}{(n-1)!}e^{-\lambda x}1_{(0,\infty)}(x)$ (to znaczy $\gamma_{n,\lambda}$). .*

Niech $X = (X_1, ..., X_n)$ będzie wektorem gaussowskim, czyli zmienną losową o wartościach w $\mathbb{R}^n$ i gęstości

$$f(x) := \frac{\sqrt{\det Q}}{(\sqrt{2\pi})^n} e^{-\langle \mathbb{Q}(x-m),(x-m)\rangle}, \quad x = (x_1, ..., x_n) \in \mathbb{R}^n, \quad m = (\mathbf{E}X_1, ..., \mathbf{E}X_n), \quad Q = \mathbf{Cov}^{-1},$$

gdzie $\mathbf{Cov} = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ jest macierzą taką, że $a_{i,j} = \mathbf{Cov}(X_i, X_j)$. Wówczas zachodzi bardzo ważny fakt: dowolna kombinacja liniowa $\alpha_1 X_1 + ... + \alpha_n X_n + b$, gdzie $\alpha_1, ..., \alpha_n, b \in \mathbb{R}$ jest wektorem gaussowskim.

Zad 10 *Niech X ma rozkład gaussowski $\mathcal{N}(0, 1)$. Znajdź rozkład $aX + b$. Niech teraz $X_1, ..., X_n$ niezależne gaussowskie o rozkładach $\mathcal{N}(m_i, \sigma_i^2)$. Znajdź rozkład zmiennej $X_1 + ... + X_n$.*

Dla zmiennych niezależnych $X_1, ..., X_n$ zachodzi użyteczny wzór $\mathbf{E}X_1..., X_n = \mathbf{E}X_1 ... \mathbf{E}X_n$.

Zad 11 *Wykonujemy n doświadczeń Bernouliego z prawdopodobieństwem sukcesu $1 - \frac{1}{n}$. Policz wartość oczekiwaną zmiennej Y_n iloczynu liczby sukcesów. Ile wynosi $\lim_{n\to\infty} \mathbf{E}Y_n$?*

Nierówności probabilistyczne

Twierdzenie 2 *Zachodzą nierówności:*

- *(Nierówność Czebyszewa)* $\mathbf{P}(X \geqslant \epsilon) \leqslant \frac{\mathbf{E}X}{\epsilon}$;
- *(Nierówność Markowa)* $\mathbf{P}(|X| \geqslant \epsilon) \leqslant \frac{\mathbf{E}|X|^p}{\epsilon^p}$, $p > 0$;
- *(Nierówność Czebyszewa)* $\mathbf{P}(|X - \mathbf{E}X| \geqslant \epsilon) \leqslant \frac{\mathbf{D}^2 X}{\epsilon^2}$;
- *(Nierówność Czebyszewa)* $\mathbf{P}(X \geqslant \epsilon) \leqslant \frac{\mathbf{E}e^{\lambda x}}{e^{\lambda\epsilon}}$.

Zad 12 *Udowodnij regułę trzech. Niech $\sigma^2 = \mathbf{D}^2 X$, pokaż że $\mathbf{P}(|X - \mathbf{E}X| > 3\sigma) \leqslant \frac{1}{9}$.*

Zad 13 *Wykaż, że jeśli $X > 0$, $\mathbf{E}X > 0$, to $\frac{1}{\mathbf{E}X} \leqslant \mathbf{E}(\frac{1}{X})$.*

Zad 14 *Pokaż, że jeśli $0 < p < q$, to $(\mathbf{E}|X|^p)^{\frac{1}{p}} \leqslant (\mathbf{E}|X|^q)^{\frac{1}{q}}$.*

Zad 15 *Wykaż nierówność Schwarza. Jeśli $\mathbf{E}|X|^2 < \infty$, $\mathbf{E}|Y|^2 < \infty$, to $\mathbf{E}|XY| \leqslant \mathbf{E}|X|^2\mathbf{E}|Y|^2$.*