Rachunek prawdopodobieństwa i statystyka

Lista zadań nr 8. Tydzień rozpoczynający się 29. kwietnia

Zadania

[Do zadań 1–2] W pliku klimat.csv znajduje się: szerokość i długość geograficzna, roczna suma opadów (mm), średnia temperatura roczna (°C) i wysokość nad poziomem morza miast wojewódzkich.

1. Wyznaczyć prostą regresji temperatury względem wysokości npm.
2. Wyznaczyć prostą regresji temperatury względem długości i szerokości. (Z zależy od X i od Y).
3. Zmienna losowa X ma dyskretny rozkład jednostajny

$$P(X = i) = \frac{1}{100}, \quad i \in \{1, 2, \dots, 99, 100\}.$$

Zmienne losowe Y oraz Z określone są następująco

$$Y = \begin{cases} 1, & 2|X \vee 3|X, \\ 0, & \text{wpw}, \end{cases} \qquad Z = \begin{cases} 1, & 3|X, \\ 0, & \text{wpw}. \end{cases}$$

Znaleźć wartość współczynnika korelacji ρ zmiennych Y i Z. (Odp.: $\rho = {}^{33}/_{67}$)

[Do zadań 4–6] Zakładamy, że zmienne X_1, X_2, X_3 są niezależne i mają ten sam ciągły rozkład o dystrybuancie $F(x)$ i gęstości $f(x)$. Tworzymy nowe zmienne losowe, mianowicie: $X_{(1)} = \min\{X_1, X_2, X_3\}$, $X_{(2)}$ to druga co do wielkości wartość, $X_{(3)} = \max\{X_1, X_2, X_3\}$.

4. Udowodnić, że $f_{(2)}(x) = 6 \cdot F(x) \cdot (1 - F(x)) \cdot f(x)$.

[Do zadań 5–6] Dodatkowo zakładamy, że $X_k \sim \mathrm{U}[0, a]$, $k = 1, 2, 3$.

5. Niech $Y_1 = \dfrac{X_1 + X_2 + X_3}{3}$, $Y_2 = X_{(2)}$, $Y_3 = \dfrac{X_{(1)} + X_{(3)}}{2}$. Udowodnić, że wartości oczekiwane są takie same: $\mathrm{E}(Y_1) = \mathrm{E}(Y_2) = \mathrm{E}(Y_3) = \dfrac{a}{2}$.

Wsk.: $\mathrm{E}(Y_1)$ z własności wartości oczekiwanej, $\mathrm{E}(Y_2)$ – całkowanie, $Y_3 = \dfrac{3Y_1 - Y_2}{2}$.

6. Wykazać, że : $\mathrm{V}(Y_1) = \dfrac{a^2}{36}$, $\mathrm{V}(Y_2) = \dfrac{a^2}{20}$.

Wsk.: Wariancja sumy niezależnych zmiennych losowych, $\mathrm{E}(Y_2^2)$ poprzez całkowanie.

7. **(2 p.)** Niech (X, Y) oznacza wybrany losowo punkt na płaszczyźnie. Załóżmy, że współrzędne X i Y są niezależne i podlegają rozkładowi $N(0, 1)$. Od zmiennej (X, Y) przechodzimy do zmiennej (R, Θ), gdzie R i Θ są współrzędnymi biegunowymi punktu (X, Y). Wykazać, że gęstość zmiennej (R, Θ) określona jest wzorem

$$g(r, \Theta) = \frac{1}{2\pi} r \cdot \exp\left\{-\frac{r^2}{2}\right\}, \quad \text{gdzie } 0 < \Theta < 2\pi, \ 0 < r < \infty.$$

8. **(2 p.)** Znaczenie zmiennej (X, Y) niech będzie takie, jak w poprzednim zadaniu. Niech

$$D = R^2 = X^2 + Y^2, \quad \Theta = \tan^{-1} \frac{Y}{X}.$$

(a) Udowodnić, że gęstość zmiennej (D, Θ) to: $f(d,\Theta) = \dfrac{1}{2}\,\exp\left\{-\dfrac{d}{2}\right\}\,\dfrac{1}{2\pi}$,

gdzie $0 < d < \infty, \quad 0 < \Theta < 2\pi$.

(b) Sprawdzić czy zmienne D i Θ są niezależne.

(c) Jaki rozkład ma zmienna D?

9. Załóżmy, że niezależne zmienne losowe X, Y mają rozkłady, odpowiednio, Gamma(b,p) i Gamma(b,q). Niech $U = X + Y$ oraz $V = \dfrac{X}{X+Y}$. Wykazać, że

(a) Zmienne U i V są niezależne.

(b) $X + Y$ ma rozkład Gamma$(b, p+q)$.

(c) Zmienna V ma rozkład Beta(p,q), tzn. $f(x) = \dfrac{1}{B(p,q)} x^{p-1}(1-x)^{q-1}$, $x \in [0,1]$.

Witold Karczewski