Probability & Statistics

Problem set №8. Week starting April 15$^{\text{th}}$

[Problems 1–2.] File climate.csv contains: latitude, longitude, annual precipitation (mm), average annual temperature (°C) and the altitude of voivodship cities.

1. Find regression line of temperature with respect to altitude.
2. Find regression line of temperature with respect to latitude and longitude. (Z depends on X and Y).
3. Random variable X has discrete uniform distribution, i.e.

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

Let r.vs. Y and Z be defined by

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

Find correlation coefficient ρ of variables Y i Z. (Answer: $\rho = 33/67$)

[Exercises 4–6] Random variables X_1, X_2, X_3 are independent and have the same continuous distribution. Cdf – $F(x)$, density – $f(x)$. Let $X_{(1)} = \min\{X_1, X_2, X_3\}$, $X_{(2)}$ – 2$^{\text{nd}}$-smallest value, $X_{(3)} = \max\{X_1, X_2, X_3\}$.

4. Prove that $f_{(2)}(x) = 6 \cdot F(x) \cdot (1 - F(x)) \cdot f(x)$.

[Exercises 5–6] Assume additionally that $X_k \sim \mathrm{U}[0, a]$, $k = 1, 2, 3$.

5. Let $Y_1 = \dfrac{X_1 + X_2 + X_3}{3}$, $Y_2 = X_{(2)}$, $Y_3 = \dfrac{X_{(1)} + X_{(3)}}{2}$. Prove that $\mathrm{E}(Y_k) = \dfrac{a}{2}$, $k = 1, 2, 3$.
   Hint: $\mathrm{E}(Y_1)$ from definition, $\mathrm{E}(Y_2)$ – integrating, $Y_3 = \dfrac{3Y_1 - Y_2}{2}$.
6. Check that $\mathrm{V}(Y_1) = \dfrac{a^2}{36}$, $\mathrm{V}(Y_2) = \dfrac{a^2}{20}$.
   Hint: Variance of independent variables' sum, $\mathrm{E}(Y_2^2)$ by integration.
7. **(2 p.)** Let (X, Y) denotes randomly chosen point on plane. Suppose the coordinates X and Y are independent and have distribution $N(0, 1)$. Cartesian coordinates (X, Y) are changed in polar coordinates (R, Θ), i.e. R and Θ are polar coordinates of point (X, Y). Prove that the density of variable (R, Θ) is given by

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

8. **(2 p.)** Random variable (X, Y) is the same like in the previous exercise. Let

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

(a) Check that density of the variable (D, Θ) satisfies identity $f(d, \Theta) = \frac{1}{2} \exp\left\{-\frac{d}{2}\right\} \frac{1}{2\pi}$, with $0 < d < \infty, \quad 0 < \Theta < 2\pi$.

(b) Check if variables D i Θ are independent.

(c) What is the distribution of the variable D?

9. Let independent random variables X, Y have distribution Gamma(b, p) and Gamma(b, q). Let $U = X + Y$ and $V = \frac{X}{X+Y}$. Prove that

(a) Variables U, V are independent.

(b) Variable $X + Y$ has distribution Gamma$(b, p + q)$.

(c) Random variable V has dostribution Beta(p, q), i.e. $f(x) = \frac{1}{B(p,q)} x^{p-1}(1-x)^{q-1}$, $x \in [0, 1]$.

Witold Karczewski