Statistics I — Class Test 2 — November 18, 2024

Name: **Roll number:**

Write your name and roll number in capital letters.

Write your answers on the question paper. You may use separate sheets for calculations and rough work, but submit only the filled-in question paper.

Let $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ and $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$ be vectors representing paired numeric (X, Y) observations obtained in a study. For fixed but arbitrary $a, b, c, d \in \mathbb{R}$, define the n-vectors $\boldsymbol{U}$ and $\boldsymbol{V}$ as affine transformations of $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively, given by

$$\begin{array}{rcl} U_i &=& a + bX_i \\ V_i &=& c + dY_i \end{array} \quad \text{or, in matrix notation,} \quad \begin{pmatrix} U_i \\ V_i \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix} + \begin{pmatrix} b & 0 \\ 0 & d \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \end{pmatrix}$$

Define the sample variance $S^2(\boldsymbol{X})$ of $\boldsymbol{X}$, and the sample correlation coefficient $\operatorname{cor}(\boldsymbol{X}, \boldsymbol{Y})$ between $\boldsymbol{X}$ and $\boldsymbol{Y}$, by

$$S^2(\boldsymbol{X}) = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2 \text{ and } \operatorname{cor}(\boldsymbol{X}, \boldsymbol{Y}) = \frac{\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{S(\boldsymbol{X})\, S(\boldsymbol{Y})}.$$

Define the bivariate distance function $\Delta\left(\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}\right)$ by

$$\Delta\left(\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}\right) = \left\| \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \right\| = \sqrt{(x - \theta_1)^2 + (y - \theta_2)^2}.$$

Let $\hat{\mu}(\boldsymbol{X}, \boldsymbol{Y})$ be the value of (θ_1, θ_2) that minimises

$$\sum_{i=1}^{n} \Delta^2 \left(\begin{pmatrix} X_i \\ Y_i \end{pmatrix}, \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}\right).$$

Let $\hat{\nu}(\boldsymbol{X}, \boldsymbol{Y})$ be the value of (θ_1, θ_2) that minimises

$$\sum_{i=1}^{n} \Delta \left(\begin{pmatrix} X_i \\ Y_i \end{pmatrix}, \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}\right).$$

1. Prove that $\bar{U} = a + b\bar{X}$. [2]

2. Prove that $S(\boldsymbol{U}) = |b|\, S(\boldsymbol{X})$. [2]

---

3. Prove that $\mathrm{cor}(\boldsymbol{U}, \boldsymbol{V}) = \mathrm{cor}(\boldsymbol{X}, \boldsymbol{Y})$. [2]

---

4. Prove that [2]

$$\hat{\mu}(\boldsymbol{U}, \boldsymbol{V}) = \begin{pmatrix} a \\ c \end{pmatrix} + \begin{pmatrix} b & 0 \\ 0 & d \end{pmatrix} \hat{\mu}(\boldsymbol{X}, \boldsymbol{Y})$$

---

5. Prove that in general [2]

$$\hat{\nu}(\boldsymbol{U}, \boldsymbol{V}) \neq \begin{pmatrix} a \\ c \end{pmatrix} + \begin{pmatrix} b & 0 \\ 0 & d \end{pmatrix} \hat{\nu}(\boldsymbol{X}, \boldsymbol{Y})$$

Good luck!