Assignment-2

Statistics - III

TRISHAN MONDAL

§ Problem1

Let $\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}_{p+p}\left(\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \begin{pmatrix} \Sigma_X & \Sigma_{XY} \\ \Sigma'_{XY} & \Sigma_Y \end{pmatrix}\right)$ and define: $U = X+Y, V = X-Y$. When is U independent of V?

Solution. At first of all notice that,

$$U = X + Y = \begin{pmatrix} I_p & I_p \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}$$

$$V = X - Y = \begin{pmatrix} I_p & -I_p \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}.$$

Define

$$A = \begin{pmatrix} I_p & I_p \\ I_p & -I_p \end{pmatrix}$$

Then, $(U,V)' = A(X,Y)'$, so

$$\begin{aligned}(U,V)' &\sim N_{p+p}\left(A\begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, A\begin{pmatrix} \Sigma_X & \Sigma_{XY} \\ \Sigma'_{XY} & \Sigma_Y \end{pmatrix}A^t\right) \\ &= N_{p+p}\left(\begin{pmatrix} \mu_X + \mu_Y \\ \mu_X - \mu_Y \end{pmatrix}, \begin{pmatrix} \Sigma_X + \Sigma^t_{XY} + \Sigma_{XY} + \Sigma_Y & \Sigma_X + \Sigma^t_{XY} - \Sigma_{XY} - \Sigma_Y \\ \Sigma_X - \Sigma^t_{XY} + \Sigma_{XY} - \Sigma_Y & \Sigma_X - \Sigma^t_{XY} - \Sigma_{XY} + \Sigma_Y \end{pmatrix}\right)\end{aligned}$$

Theorem : $X \sim N_p(\mu, \Sigma)$; Given $U = AX$ and $V = BX$, then U and Y are independent iff $\mathrm{Cov}(U,V) = A\Sigma B' = 0$

So for U and V to be independent we need

$$\Sigma_X + \Sigma^t_{XY} - \Sigma_{XY} - \Sigma_Y = 0_p \text{ and } \Sigma_X - \Sigma^t_{XY} + \Sigma_{XY} - \Sigma_Y = 0_p$$

Adding and subtracting the above equations implies that $\Sigma_X = \Sigma_Y$ and $\Sigma^t_{XY} = \Sigma_{XY}$. Clearly, if $\Sigma_X = \Sigma_Y$ and $\Sigma^t_{XY} = \Sigma_{XY}$ the above two equations hold, so we have found necessary and sufficient conditions for U and V to be independent. ■

§ Problem 2

Let Z_1, Z_2, Z_3 be i.i.d. $\mathcal{N}(0,1)$ and $0 < \rho < 1$. Define $X_1 = Z_1$, $X_2 = \rho Z_1 + \sqrt{1-\rho^2} Z_2$ and $X_3 = \rho X_2 + \sqrt{1-\rho^2} Z_3$. What is the joint distribution of $(X_1, X_2, X_3)'$?

Solution. Given, Z_1, Z_2, Z_3 i.i.d $\mathcal{N}(0,1)$. So, $(Z_1, Z_2, Z_3)'$ follows standard normal distribution of order 3. Notice that,

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ \rho & \sqrt{1-\rho^2} & 0 \\ \rho^2 & \rho\sqrt{1-\rho^2} & \sqrt{1-\rho^2} \end{pmatrix}}_{\text{Call this matrix } A} \begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \end{pmatrix}$$

Now, $\begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \end{pmatrix} \sim \mathcal{N}_3\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right)$ is the standard normal random vector. So,

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim \mathcal{N}\left(0, AI_3A'\right) = \mathcal{N}\left(0, AA'\right)$$

We will compute AA' as following,

$$\begin{aligned} AA' &= \begin{pmatrix} 1 & 0 & 0 \\ \rho & \sqrt{1-\rho^2} & 0 \\ \rho^2 & \rho\sqrt{1-\rho^2} & \sqrt{1-\rho^2} \end{pmatrix} \begin{pmatrix} 1 & \rho & \rho^2 \\ 0 & \sqrt{1-\rho^2} & \rho\sqrt{1-\rho^2} \\ 0 & 0 & \sqrt{1-\rho^2} \end{pmatrix} \\ &= \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix}. \end{aligned}$$

We can conclude that $(X_1, X_2, X_3)'$ follows a 3 -dimensional multivariate normal distribution whose mean is $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ and covariance matrix of it is, $\begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix}$.

$$\text{.i.e.} \quad \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix}\right)$$

■

§ Problem 3

Let Z_1, Z_2 be i.i.d. $N(0,1)$ and $0 < \rho < 1$. Define $X_1 = Z_1$, $X_2 = \rho Z_1 + \sqrt{1-\rho^2} Z_2$. Find X_3 such that

$$\operatorname{Cov}\left(X_1, X_2, X_3\right) = \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix}.$$

Solution. Introduce a new random variable Z_3. such that, Z_1, Z_2, Z_3 i.i.d $\mathcal{N}(0,1)$.

$$\begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, I_3\right)$$

Let, we can write X_3 as $\alpha Z_1 + \beta Z_2 + \gamma Z_3$. We want to determine α, β, γ explicitly.

$$\underset{\sim}{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ \rho & \sqrt{1-\rho^2} & 0 \\ \alpha & \beta & \gamma \end{pmatrix}}_{\text{call this matrix } A} \begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \end{pmatrix} = A\underset{\sim}{Z}$$

$$\Rightarrow \mathrm{Cov}(\underset{\sim}{X}) = \mathrm{Cov}(A\underset{\sim}{Z}) = A\,\mathrm{Cov}(\underset{\sim}{Z})A' = AA'$$

Now, $$AA' = \begin{pmatrix} 1 & 0 & 0 \\ \rho & \sqrt{1-\rho^2} & 0 \\ \alpha & \beta & \gamma \end{pmatrix} \begin{pmatrix} 1 & \rho & \alpha \\ 0 & \sqrt{1-\rho^2} & \beta \\ 0 & 0 & \gamma \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \rho & \alpha \\ \rho & 1 & \alpha\rho + \beta\sqrt{1-\rho^2} \\ \alpha & \alpha\rho + \beta\sqrt{1-\rho^2} & \alpha^2+\beta^2+\gamma^2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix}$$

By comparing entries of the matrices we get, $\alpha = \rho$ and, $\alpha\rho + \beta\sqrt{1-\rho^2} = \rho \Rightarrow \beta = \frac{\rho-\rho^2}{\sqrt{1-\rho^2}}$ and,

$$\begin{aligned}
&\alpha^2+\beta^2+\gamma^2 = 1 \\
&\Rightarrow \rho^2 + \rho^2\frac{(1-\rho)^2}{1-\rho^2} + \gamma^2 = 1 \\
&\Rightarrow \rho^2\left(1 + \frac{1-\rho}{1+\rho}\right) + \gamma^2 = 1 \\
&\Rightarrow \frac{2\rho^2}{1+p} + \gamma^2 = 1 \\
&\Rightarrow \gamma^2 = \frac{(1+\rho)-2\rho^2}{1+\rho} \\
&\Rightarrow \gamma = \sqrt{\frac{(1+\rho)-2\rho^2}{1+\rho}}
\end{aligned}$$

The term, $(1+\rho)-2\rho^2 = (1-\rho)(1+2\rho)$ is always positive as $0 < \rho < 1$. So, the required X_3 is,

$$X_3 = \rho Z_1 + \frac{\rho-\rho^2}{\sqrt{1-\rho^2}}Z_2 + \sqrt{\frac{(1+\rho)-2\rho^2}{1+\rho}}Z_3$$

** We have taken only the positive value of γ. ■

§ Problem 4

Let $\mathbf{Y} \sim \mathcal{N}_n\left(\theta, \sigma^2 I_n\right)$, and let $\mathbf{X} = A\mathbf{Y}, \mathbf{U} = B\mathbf{Y}$ and $\mathbf{V} = C\mathbf{Y}$, where A, B and C are all $r \times n$ matrices of rank $r < n$. If $\operatorname{Cov}(\mathbf{X}, \mathbf{U}) = 0$ and $\operatorname{Cov}(\mathbf{X}, \mathbf{V}) = 0$, show that $\mathbf{X}$ is independent of $\mathbf{U} + \mathbf{V}$

Solution. We should start observing that,

$$\begin{aligned}
&\operatorname{Cov}(\mathbf{X}, \mathbf{U}) = 0 \\
\Rightarrow &\operatorname{Cov}(A\mathbf{Y}, B\mathbf{Y}) = 0 \\
\Rightarrow &A\sigma^2 I_n B' = 0 \\
\Rightarrow &\sigma^2 AB' = 0 \cdots (1)
\end{aligned}$$

We can do the similar calculation for $\operatorname{Cov}(\mathbf{X}, \mathbf{V})$,

$$\begin{aligned}
&\operatorname{Cov}(\mathbf{X}, \mathbf{V}) = 0 \\
\Rightarrow &\operatorname{Cov}(A\mathbf{Y}, C\mathbf{Y}) = 0 \\
\Rightarrow &A\sigma^2 I_n C' = 0 \\
\Rightarrow &\sigma^2 AC' = 0 \cdots (2)
\end{aligned}$$

Theorem: If $\mathbf{X} \sim \mathcal{N}_p(\mu, \Sigma)$ then, for any $k \times p$ matrix A, $(A\mathbf{X}+b) \sim \mathcal{N}_p(A\mu+b, A\Sigma A)$.

Using the above theorem we can say that $\mathbf{U} + \mathbf{V} = (B + C)\mathbf{X}$ also follows multivariate normal. Now,

$$\begin{aligned}
&\operatorname{Cov}(\mathbf{X}, \mathbf{U} + \mathbf{V}) \\
&= \operatorname{Cov}(A\mathbf{Y}, (B + C)\mathbf{Y}) \\
&= A\sigma^2 I_n (B + C)' \\
&= \sigma^2 A(B + C)' \\
&= \sigma^2 A(B' + C') \\
&= 0 \text{ (by adding equation (1) and (2))}
\end{aligned}$$

From the theorem we used for ‘Problem1’ we can conclude that $\mathbf{X}$ is independent of $\mathbf{U} + \mathbf{V}$ ■

§ Problem 5

Let $Z \sim N(0, 1)$. Define

$$Y = \begin{cases} Z & \text{if } |Z| > c \\ -Z & \text{if } |Z| \leq c \end{cases}$$

Show that (Z, Y) has a joint distribution under which the marginal distributions are normal, but the joint distribution is not bivariate normal.

Solution. The joint distribution of (Y, Z) has marginal distribution as distributions of Y and Z respectively. Now we will look at the distribution of Y,

$$\begin{aligned}
F_Y(y) &= \mathbb{P}(Y \leqslant y) \\
&= \mathbb{P}(Z \leqslant y, |Z| > c) + \mathbb{P}(-Z \leqslant y, |Z| \leqslant c) \\
&= \mathbb{P}(Z \leq y, Z < -c) + \mathbb{P}(Z \leq y, -c \leq Z \leq c) + \mathbb{P}(Z < y, Z > c)
\end{aligned}$$

Now three seperate case will arise,

<u>Case 1</u> $y > c$ then,

$$\begin{aligned}
F_Y(y) &= \mathbb{P}(Z \leq y, Z < -c) + \mathbb{P}(Z \leq y, -c \leq Z \leq c) + \mathbb{P}(Z < y, Z > c) \\
&= \mathbb{P}(Z < -c) + \mathbb{P}(Z \leq c) - \mathbb{P}(Z < -c) + \mathbb{P}(Z \leq y) - \mathbb{P}(Z < c) \\
&= \mathbb{P}(Z \leq y) \\
&= F_Z(y)
\end{aligned}$$

<u>Case 2</u> $y < -c$ then,

$$\begin{aligned}
F_Y(y) &= \mathbb{P}(Z \leq y, Z < -c) + \mathbb{P}(Z \leq y, -c \leq Z \leq c) + \mathbb{P}(Z < y, Z > c) \\
&= \mathbb{P}(Z \leq y) + 0 + 0 \\
&= F_Z(y).
\end{aligned}$$

<u>Case 3</u> $-c < y < c$ then,

$$\begin{aligned}
F_Y(y) &= \mathbb{P}(Z \leq y, Z < -c) + \mathbb{P}(Z \leq y, -c \leq Z \leq c) + \mathbb{P}(Z < y, Z > c) \\
&= \mathbb{P}(Z < -c) + \mathbb{P}(Z \leqslant y) - \mathbb{P}(Z < -c) \\
&= \mathbb{P}(Z \leqslant y). \\
&= F_Z(y).
\end{aligned}$$

So,Y, Z has same distribution. And hence (Z, Y) has a joint distribution under which the marginal distributions are normal. Now notice that,

$$Y + Z = \begin{cases} 2Z & \omega \cdot p \quad \mathbb{P}(|Z| > c) \\ 0 & \omega \cdot p \quad \mathbb{P}(|Z| \leqslant c) \end{cases}$$

This is not univariate normal distribution. If $(Y, Z)'$ was distributed normally then for any α, $\alpha'(Y, Z)'$ should have been distributed normally. But this is not the case and hence the joint distribution of (U, V) is not bivariate normal. ■