BINDURA UNIVERSITY OF SCIENCE EDUCATION SFM214

BACHELOR OF STATISTICS AND FINANCIAL MATHEMATICS

MULTIVARIATE METHODS

Time : 3 hours OCT 2024

Candidates should attempt ALL questions in section A and at most TWO questions in section B.

SECTION A (40 marks)

Candidates may attempt ALL questions being careful to number them A1 to A5

A1. The data matrix for variables X_1 and X_2 respectively is given as:

$$\mathbb{X} = \begin{bmatrix} 60 & 14 \\ 50 & 15 \\ 45 & 14 \\ 55 & 13 \\ 40 & 10 \\ 20 & 12 \end{bmatrix}$$

Find the

(a) mean vector matrix, $\bar{\mathbb{X}}$ [2]

(b) sample variance-covariance matrix, $\sum$ [4]

(c) correlation matrix, $\mathbb{R}$. [2]

A2. Consider a random vector $\mathbb{X}' = [X_1, X_2, X_3, X_4]$ with mean vector $\mu'_X = [3, 2, -2, 0]$ and variance-covariance matrix

$$\sum_X = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Let

$$A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 1 & -3 \end{bmatrix}$$

(a) Find E(AX), the mean of AX. [2]

(b) Find Cov(AX), the variances and covariances of AX. [3]

(c) Determine from (ii) above pairs of linear combinations that have zero covariances. [3]

A3. (a) Let X be a normally distributed random vector with

$$\mu = \begin{pmatrix} -3 \\ 1 \\ -2 \end{pmatrix} \text{ and } \sum = \begin{bmatrix} 4 & 0 & -1 \\ 0 & 5 & 0 \\ -1 & 0 & 2 \end{bmatrix}$$

Determine whether the following are independent or not and justify your answers.

(i) X_1 and X_3 [2]

(ii) $X_1 + X_3$ and $X_1 - 2X_2$ [3]

(iii) X_1 and $X_1 + 3X_2 - 2X_3$. [3]

(b) Let X be $\mathbb{N}_3(\mu, \sum)$ with $\mu = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$ and $\sum = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 3 & 3 \end{pmatrix}$

Find the distribution of $3X_1 - 2X_2 + X_3$. [4]

A4. Let $A = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}$. Compute the spectral decomposition of A, [6]

A5. Suppose

$$\sum = \begin{bmatrix} 4 & 1 & 2 \\ 1 & 9 & -3 \\ 2 & -3 & 25 \end{bmatrix}$$

is the variance-covarince matrix for random variables X_1, X_2 and X_3. Find the correlation matrix $\mathbb{R}$ for $\sum$. [6]

SECTION B (60 marks)

Candidates may attempt TWO questions being careful to number them B6 to B8.

B6. (a) To study the per capita consumption of chicken in Zimbabwe, you are given data below, where Y denotes per capita consumption of chickens (Kg) and X denotes real disposable income per capita ($).

Y	36	72	48	51	80	40	55	72	39	47
X	240	450	250	320	450	250	330	430	240	320

(i) Construct a scatter plot of the data and the marginal dot diagrams. [4]

(ii) Infer the sign of the sample covariance S_{12} from the scatter plot. [2]

(iii) Display the sample mean vector matrix $\bar{\mathbb{X}}$, the sample variance-covariance matrix $\mathbb{S}_n$ and the sample correlation matrix $\mathbb{R}$. [13]

(b) Consider a bivariate normal population with mean $\mu_1 = 0$, $\mu_2 = 2$, $\sigma_{11} = 2$, $\sigma_{22} = 1$, and $\rho_{12} = 5$.

(i) Find the bivariate normal density function for the above normal population. [6]

(ii) Write out the squared generalised distance expression $(x-\mu)'\sum^{-1}(x-\mu)$ as a function of x_1 and x_2. [2]

(iii) Determine and sketch the constant density-contour that contains 50% of the probability for (ii) above. [3]

B7. (a) You are given the random vector $\mathbb{X}' = [X_1, X_2, ...X_4]$ with mean vector $\mu_X' = [4, 3, 2, 1]$ and variance-covariance matrix

$$\sum = \begin{bmatrix} 3 & 0 & 2 & 2 \\ 0 & 1 & 1 & 0 \\ 2 & 1 & 9 & -2 \\ 2 & 0 & -2 & 4 \end{bmatrix}$$

Partition X as

$$X = \begin{bmatrix} X_1 \\ X_2 \\ .... \\ X_3 \\ X_4 \end{bmatrix} = \begin{bmatrix} X^{(1)} \\ .... \\ X^{(2)} \end{bmatrix}$$

Let $A = \begin{bmatrix} 1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -2 \\ 2 & -1 \end{bmatrix}$ and consider the linear combinations $AX^{(1)}$ and $BX^{(2)}$. Find

(i) $E(AX^{(1)})$. [3]

(ii) $E(BX^{(2)})$. [3]

(iii) $Cov(BX^{(2)})$. [3]

(iv) $Cov(AX^{(1)})$. [3]

(v) $Cov(AX^{(1)}, BX^{(2)})$ [6]

(b) Suppose that $n_1 = 11$ and $n_2 = 12$ observations are sampled from two different bivariate normal distributions that have a common covariance matrix $\sum$ and possibly different mean vectors μ_1 and μ_2. The sample mean vectors and pooled covariance matrix are:

$$\bar{X}_1 = \begin{pmatrix} -1 \\ -2 \end{pmatrix}, \quad \bar{X}_2 = \begin{pmatrix} 4 \\ 5 \end{pmatrix}, \quad \mathbb{S} = \begin{pmatrix} 13 & 6 \\ 6 & 22 \end{pmatrix}$$

Use the Hotelling two sample T^2-statistic to test for a difference in the population mean vectors. Use $\alpha = 0.05$. [12]

B8. (a) Consider a sample of 10 observations recorded on two varibles X_1 and X_2. Suppose that the mean vector matrix for the sample is $\bar{\mathbb{X}} = \begin{pmatrix} 192 \\ 278.4 \end{pmatrix}$ and the sample variance-covariance matrix is $\sum = \begin{pmatrix} 121.78 & 76.3 \\ 76.3 & 92.9 \end{pmatrix}$.

Calculate the 95% simultaneous confidence interval for the mean μ. [12]

(b) An Agronomist interested in curbing desertification planted 8 drought resistant trees of the same type at different depths and measured the extra height they gained and the extra leaf span they gained after a few weeks. The expert wants to investigate the effect of the different planting depth on the growth of the trees. The data in appropriate units are recorded in the table below.

Tree	1	2	3	4	5	6	7	8
Depth (cm)	5	5	5	8	8	12	12	12
Height (cm)	9	6	9	0	2	3	1	2
Span (cm)	3	2	7	4	0	8	9	7

(i) Breakdown the observations into $X_{ij} = \bar{X} + (\bar{X}_i - \bar{X}) + (X_{ij} - \bar{X}_i)$. [10]

(ii) Construct a one-way MANOVA table for the data. [8]

END OF QUESTION PAPER