Kasarani Campus
Off Thika Road
P. O. Box 49274, 00101
NAIROBI
Westlands Campus
Pamstech House
Woodvale Grove
Tel. 4442212
Fax: 4444175

KIRIRI WOMENS' UNIVERSITY OF SCIENCE AND TECHNOLOGY
UNIVERSITY EXAMINATIONS, 2024/2025 ACADEMIC YEAR
FOURTH YEAR, FIRST SEMESTER EXAMINATION
FOR THE DEGREE OF BACHELOR OF SCIENCE IN MATHEMATICS

KMA 2410 MULTIVARIATE STATISTICAL METHODS I

Date: 13TH AUGUST, 2024
Time: 11:30 AM – 1:30 PM

INSTRUCTIONS TO CANDIDATES
ANSWER QUESTION ONE (COMPULSORY) AND ANY OTHER TWO QUESTIONS

QUESTION ONE: COMPULSORY (30 MARKS)

a) Suppose the data below is a random sample from a bivariate normal distribution

$$\underline{X}^T = \begin{pmatrix} 3 & 4 & 5 & 4 \\ 6 & 4 & 7 & 7 \end{pmatrix}^T$$

Obtain:

i. The maximum likelihood estimate of the mean vector **(3 Marks)**

ii. The maximum likelihood estimate of the covariance matrix **(4 Marks)**

b) Suppose $Y \sim N_3(\underline{\mu}, \sum)$ where $\mu = \begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix}$ and $\sum = \begin{bmatrix} 4 & -3 & 0 \\ -3 & 6 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

Examine the independence between (Y_1, Y_3) and Y_2 **(5 Marks)**

c) Let $\underline{X}$ be a 3-variate random vector such that $\Sigma = \begin{bmatrix} 5 & 2 & 3 \\ 2 & 3 & 0 \\ 3 & 0 & 2 \end{bmatrix}$

Obtain the correlation matrix of the random vector $\underline{Y} = (Y_1, Y_2, Y_3)'$ where

$Y_1 = x_1 + x_2, \quad Y_2 = x_1 + x_3, \quad Y_3 = x_2 + x_3$ **(6 Marks)**

d) Explain the following concepts and their significance as used in multivariate analysis

i. Central Limit Theorem **(2 Marks)**

ii. Law of Large Numbers **(2 Marks)**

e) Define $Y = a'\underline{X}$, $a \neq 0$, where a is a vector of constants. Use the characteristic function technique to show that $Y \sim N_p(a'\underline{\mu}, a'\Sigma a)$ **(4 Marks)**

f) The variates $\underline{X}' = (X_1, X_2, X_3)$ and $\underline{Y}' = (Y_1, Y_2, Y_3)$ are distributed independently according to the trivariate normal populations with respective parameters

$$\mu_1 = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}, \quad \sum{}_1 = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 4 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \quad \mu_2 = \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix}, \quad \sum{}_2 = \begin{bmatrix} 4 & 2 & 0 \\ 2 & 4 & 2 \\ 0 & 2 & 4 \end{bmatrix},$$

Determine the distribution of $Z = 3\underline{X}' - \underline{Y}'$ **(4 marks)**

QUESTION TWO: (20 MARKS)

a) Consider a p-dimensional random vector $\underline{X}$ and let A be a matrix of constants compatible with $\underline{X}$ and $\underline{b}$ be a vector of constants of the same size as $A\underline{X}$

i. Given that $X = (X_1, X_2, X_3)'$ with mean vector $\underline{\mu}' = (1, -1, 3)'$ and $\Sigma = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & 3 \\ 0 & 3 & 10 \end{bmatrix}$ Let $Y_1 = X_1 - X_2 + X_3$, $Y_2 = X_1 + X_2 - 3X_3$ and $Y_3 = 2X_1 - X_2 - X_3$. Find the dispersion matrix of $Y = (Y_1, Y_2, Y_3)'$ **(5 Marks)**

ii. Comment on the independence of Y_1 and Y_3 using correlation matrix **(5 Marks)**

b) Consider the following sample data matrix with three variables

$$X = \begin{pmatrix} 9 & 12 & 3 \\ 2 & 8 & 4 \\ 6 & 6 & 0 \\ 5 & 4 & 2 \\ 8 & 10 & 1 \end{pmatrix}$$

Obtain:

i. Sample covariance matrix **(6 Marks)**

ii. Use the obtained sample covariance matrix in (i) above to find variance of $Z = \begin{pmatrix} X_1 - X_2 + X_3 \\ 2X_1 + X_2 - X_3 \end{pmatrix}$ **(4 Marks)**

QUESTION THREE: (20 MARKS)

a) Consider the random vector $\underline{X} = (X_1, X_2, X_3)'$ with a pdf given by

$$f(\underline{x}) = \begin{cases} kx_1x_2x_3, & 0 < x_1, x_2 < 1, \quad 0 < x_3 < 3 \\ 0, & otherwise \end{cases}$$

i. Find the value of k **(2 Marks)**

ii. Obtain the dispersion matrix for this distribution **(6 Marks)**

b) Given that $X = (X_1, X_2, X_3)'$ with mean vector $\underline{\mu}' = (1, -1, 2)'$ and $\Sigma = \begin{bmatrix} 7 & 3 & 2 \\ 3 & 4 & 1 \\ 2 & 1 & 2 \end{bmatrix}$.

Partition $\underline{X}$ as $X_1 = X_1$ and $\underline{X}_2 = (X_2, X_3)'$. Obtain the regression coefficients for the regression line of X_1 on X_2 *and* X_3 **(7 Marks)**

c) Let $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$, where Σ is symmetric positive definite. Given that the moment generating function of $\underline{X}$ is given by $M_{\underline{X}}(\underline{t}) = \exp\left\{\underline{t}'\underline{\mu} + \frac{1}{2}\underline{t}'\Sigma\underline{t}\right\}$, show that $E(\underline{X}) = \underline{\mu}$ and $Var(\underline{X}) = \Sigma$ **(5 Marks)**

QUESTION FOUR: (20 MARKS)

a) The table below shows laboratory results of three characteristics of soil chemical contents (measured in milliequivalents per 100 g) from some 10 different locations. The variables are

Y_1= available soil calcium,
Y_2 = exchangeable soil calcium,
Y_3 = turnip green calcium.

Location	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$\mathbf{Y_1}$	35	35	40	10	6	20	35	35	35	30
$\mathbf{Y_2}$	3.5	4.9	30.0	2.8	2.7	2.8	4.6	10.9	8.0	1.6
$\mathbf{Y_3}$	2.80	2.70	4.38	3.21	2.73	2.81	2.88	2.90	3.28	3.20

Assuming normal distribution, use the provided data to obtain the following:

i. Mean vector **(3 Marks)**
ii. Dispersion matrix **(5 Marks)**
iii. Conditional distribution of Y_3 given $(Y_1, Y_2) = (7.0, 3.0)'$ **(5 Marks)**

b) Let the joint density function of two random variables X and Y be given by

$$f(x,y) = \begin{cases} k(x+4y), & 0 \le x \le 2, 0 \le y \le 1 \\ 0, & elsewhere \end{cases}$$

i. Find the value of the constant *k* **(2 Marks)**
ii. Obtain the marginal densities of X and Y **(5 Marks)**

QUESTION FIVE: (20 MARKS)

a) Let $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_n$ be ***n*** independent observation vectors from a multivariate normal population with mean vector $\underline{\mu}$ and covariance matrix Σ. Define the sample mean vector

$\overline{\underline{X}} = \frac{1}{n}\sum_{r=1}^{n}\underline{X}_r$ and sample covariance matrix as $S = ((s_{ij}))$, $s_{ij} = \frac{1}{n}\sum_{r=1}^{n}(X_{ri} - \overline{X}_i)(X_{rj} - \overline{X}_j)$

i. Derive the distribution of $\overline{\underline{X}}$ (Hint: show that $\overline{\underline{X}} \sim N(\underline{\mu}, \Sigma)$) **(5 Marks)**
ii. Show that the sample covariance matrix S is biased for Σ **(4 Marks)**
iii. Hence or otherwise, obtain the unbiased estimator for Σ **(2 Marks)**

b) Given that $X = (X_1, X_2, X_3)'$ with mean vector $\underline{\mu}' = (1, -1, 2)'$ and $\Sigma = \begin{bmatrix} 7 & 3 & 2 \\ 3 & 4 & 1 \\ 2 & 1 & 2 \end{bmatrix}$.

Partition $\underline{X}$ as X_2 and $\underline{X}_2 = (X_1, X_3)'$. Hence or otherwise, find

i. $E(X_2 / \underline{X}_2)$ **(4 Marks)**
ii. $Var(X_2 / \underline{X}_2)$ **(5 Marks)**