



MACHAKOS UNIVERSITY

University Examinations 2019/2020

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS, STATISTICS AND ACTUARIAL SCIENCE

SPECIAL/SUPPLEMENTARY EXAMINATION FOR

BACHELOR OF SCIENCE (MATHEMATICS AND COMPUTER SCIENCE)

BACHELOR OF SCIENCE (STATISTICS & PROGRAMMING)

BACHELOR OF SCIENCE (MATHEMATICS)

SMA 466: MULTIVARIATE STATISTICAL METHODS II

DATE: 21/1/2021

TIME: 8.30-10.30 AM

INSTRUCTIONS:

Answer question ONE and any other TWO questions

QUESTION ONE (30 MARKS)

- a) Briefly explain the following objectives of scientific investigations based on multivariate data identifying the area/topic under which it is covered.
- i. Data reduction. (2 marks)
 - ii. Data sorting and grouping. (2 marks)
 - iii. Hypothesis testing. (2 marks)
- b) Given that $\bar{X} \sim N_3(\mu, \Sigma)$ with

$$\mu = \begin{pmatrix} 24 \\ 41 \\ 47 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 36 & 12 & -6 \\ 12 & 16 & 2 \\ -6 & 2 & 4 \end{pmatrix}$$

Find;

- i. The distribution of $Y = \begin{pmatrix} X_1 + 2X_3 \\ X_1 - X_2 + X_3 \end{pmatrix}$ (4 marks)
 - ii. The distribution of X_1 given that $X_2 = 36$ and $X_3 = 49$. (4 marks)
 - iii. The regression function of X_3 on X_1 and X_2 . (4 marks)
 - iv. The correlation matrix for the data. (2 marks)
 - v. The partial correlation coefficient between X_1 and X_3 for fixed values of X_2 . (2 marks)
- c) Suppose that $X \sim N_p(\mu, \Sigma)$, find the Maximum Likelihood Estimator of μ . (3 marks)
 - d) Describe Wishart's distribution and state its role in Multivariate analysis. (5 marks)

QUESTION TWO (20 MARKS)

- a) Briefly explain the main idea behind discriminant analysis stating the two types or errors. (5 marks)
- b) Observations on three responses are collected for three treatments. The observation vectors $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ are as given by

Treatment	1	1	2	2	2	3	3	3
x_1	13	16	16	14	15	20	18	16
x_2	15	11	10	12	14	18	14	10
x_3	17	19	14	13	12	16	14	15

Find the matrix of sum of squares due to treatment, the matrix of residual sum of squares hence test at $\alpha = 0.05$ level of significance the hypothesis that the treatment have no effects.

(15 marks)

QUESTION THREE (20 MARKS)

- a) Discuss briefly hypothesis testing under Multivariate Analysis of Variance and identify three test statistics used in decision making. (5 marks)
- b) Given normal population $\Pi_1 : N_p(\mu_1, \Sigma)$, and $\Pi_2 : N_p(\mu_2, \Sigma)$. Suppose an observation x is to be classified into one of the two populations, derive a discriminant rule for classifying x . (5 marks)
- c) The following data was collected from two populations as labeled

Population	1	1	1	1	1	2	2	2
X_1	21	30	34	24	29	34	34	25
X_2	32	27	31	31	27	27	28	26
X_3	31	32	26	27	33	29	32	27

Given that the pooled covariance matrix is $S = \begin{pmatrix} 14 & -4 & -4 \\ -4 & 6 & -1 \\ -4 & -1 & 7 \end{pmatrix}$. Use an appropriate

discriminant rule to classify a new observation $x = \begin{pmatrix} 29.2 \\ 28.3 \\ 32.8 \end{pmatrix}$ into either population 1 or

population 2. (10 marks)

QUESTION FOUR (20 MARKS)

- a) The following is a data matrix of examination results in a number of units

X_1	26	35	31	42	36	43	38	41
X_2	40	34	45	39	37	42	38	30
X_2	37	41	43	39	41	35	39	34

Use it to answer the following questions

- i. Compute the Mean vector. (4 marks)

- ii. Find the covariance matrix. (6 marks)
- iii. Compute the correlation matrix. (6 marks)
- iv. Test at $\alpha = 0.1$ the hypothesis $H_0 : \mu_0 = \begin{pmatrix} 34 \\ 39 \\ 40.2 \end{pmatrix}$. (4 marks)

QUESTION FIVE (20 MARKS)

- a) Let $x_1, x_2, \dots, x_n$ be a random sample of size n taken on $\bar{X}$ and let $\bar{x}$ and S be the MLEs of μ and Σ respectively. Derive the test statistics for testing the hypothesis $H_0 : \mu = \mu_0$; vs $H_1 : \mu \neq \mu_0$. (7 marks)
- b) Let $X \sim N(\mu, \Sigma)$ be tri-variate normal random vector with sample variance-covariance matrix $S = \begin{pmatrix} 36 & 0 & -4.5 \\ 0 & 16 & 0 \\ -4.5 & 0 & 9 \end{pmatrix}$. Find the;
 - i. The first two principal component and the total variance explained by the components. (11 marks)
 - ii. Correlation between the first principal component and the second variable. (2 marks)