MACHAKOS UNIVERSITY

University Examinations for 2024/2025 Academic Year

SCHOOL OF PURE AND APPLIED SCIENCES
DEPARTMENT OF MATHEMATICS, STATISTICS AND ACTUARIAL SCIENCES

FOURTH YEAR FIRST SEMESTER EXAMINATION FOR DEGREE IN
BAC BACHELOR OF SCIENCE (STATISTICS AND COMPUTER PROGRAMMING)
BACHELOR OF SCIENCE (MATHEMATICS AND COMPUTER SCIENCE)
BACHELOR OF SCIENCE (ACTUARIAL SCIENCE)
BACHELOR OF EDUCATION (SCIENCE)

SMA 466: MULTIVARIATE METHODS II

DATE: 13TH JANUARY, 2025 **TIME: 8.30 AM – 10.30 AM**

INSTRUCTIONS:
*Answer question **ONE (Compulsory)** and **any other TWO** questions*

QUESTION ONE (30 MARKS), *COMPULSORY*

a) State two main objectives of discriminant analysis. (2 marks)

b) State two assumptions of the discriminant analysis for two groups. (2 marks)

c) Briefly explain the difference between principal component analysis and discriminant analysis. (4 marks)

d) Compute the decomposition of the following matrix. (5 marks)

$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$$

e) Consider the covariance matrix

$$\Sigma = \begin{bmatrix} 1 & 4 \\ 4 & 100 \end{bmatrix}$$

Determine the correlation matrix, ρ (3 marks)

f) Given the vectors $\mathbf{x}^T = (4, 1, 2)$ and $\mathbf{y}^T = (2, 1,1)$, determine;

(i) the length of x, (2 marks)

(ii) the length of y and (3 marks)

g) Compute the eigenvalues and eigenvectors of the matrix $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ (5 marks)

h) Using the result of (g) above, compute the normalized eigenvectors. (4 marks)

QUESTION TWO (20 MARKS)

a) Consider the bivariate normal population with $\mu_1 = 1,\ \mu_2 = 3,\ \sigma_{11}^2 = 2,\ \sigma_{22}^2 = 1,\ \rho_{12} = 0.8$

i. Write out the bivariate normal density. (6 marks)

ii. Write out the squared statistical distance expression $(x-\mu)'\sum^{-1}(x-\mu)$ as a function of x_1 and x_2. (4 marks)

b) Let $\mathbf{X}$ be $N_3(\boldsymbol{\mu}, \Sigma)$ with $\boldsymbol{\mu}' = (2, 3, 1)$ and

$$\sum = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

i. Find the distribution of $3x_1 - 3x_2 + x_3$ (4 marks)

ii. If $\mathbf{X}' = (3, 2, 1)$ the conditional distribution of $X_1 = x_1$ given $\boldsymbol{X}_2 = (x_2, x_3)$ (6 marks)

QUESTION THREE (20 MARKS)

a) State three (3) factors to consider when deciding on the number of principal components to retain. (3 marks)

b) Give three (3) general comparisons between of the principal components obtained from the correlation matrix and those obtained from S. (3 marks)

c) Data on $x_1 = sales$ and $x_2 = profits$ for the 10 largest companies in the world were obtained and the data summarized as follows

$$\bar{X} = \begin{bmatrix} 155.60 \\ 14.70 \end{bmatrix},\ S = \begin{bmatrix} 7476.45 & 303.62 \\ 303.62 & 26.19 \end{bmatrix}$$

i. Determine that sample principal components and their variances for these data. (10 marks)

ii. Find the proportion of the total sample variance explained by $\hat{y}_2$. (4 marks)

QUESTION FOUR (20 MARKS)

(a) State two (2) properties of the T^2 test. (2 marks)

(b) Evaluate T^2, for testing $H_0: \boldsymbol{\mu}' = (7, 11)$, using the data

$$X = \begin{bmatrix} 2 & 12 \\ 8 & 9 \\ 6 & 9 \\ 8 & 10 \end{bmatrix}$$ (8 marks)

(c) Specify the distribution of T^2 for the case of (a) above. (3 marks)

(d) Using (a) and (b) above test H_0 at $\alpha = 0.05$ level. What is your conclusion? (4 marks)

(e) Suppose Σ is known.

i. What would be the best replacement of T^2? (1 marks)

ii. What is the distribution of the replacement of T^2? (2 marks)

QUESTION FIVE (20 MARKS)

(a) Briefly explain the following terms as used in canonical correlation analysis.

i. Canonical variates. (3 marks)

ii. Canonical correlations. (3 marks)

(b) Differentiate between multivariate analysis and canonical correlation analysis. (2 marks)

(c) Let $Z = \begin{pmatrix} X \\ Y \end{pmatrix}$ be a 5x1 real vector random variable where X is 3x1 and Y is 2x1. Let the covariance matrix of Z be $\sum$ where

$$Z = \begin{bmatrix} 3 & 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}, \Sigma_{11} = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

$$\Sigma_{22} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \Sigma_{21} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \Sigma_{12} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$$

Construct the pairs of canonical variables. (12 marks)