MAASAI MARA UNIVERSITY

REGULAR UNIVERSITY EXAMINATIONS
2023/2024 ACADEMIC YEAR
SECOND YEAR FIRST SEMESTER

SCHOOL OF BUSINESS AND ECONOMICS
MASTER SCIENCE IN ECONOMICS AND STATISTICS

COURSE CODE: ECS 8202
COURSE TITLE: MULTIVARIATE ANALYSIS

DATE: 9/2/ 2024 **TIME: 1430-1730 HRS**

INSTRUCTIONS TO CANDIDATES

*Answer question **ONE (compulsory)** and any other **TWO** questions.*

QUESTION ONE (30 Marks)

a) Given that $X \sim N_3(\mu, \Sigma)$ with

$$\mu = \begin{pmatrix} 40 \\ 28 \\ 12 \end{pmatrix} \qquad \Sigma = \begin{pmatrix} 36 & -6 & -12/5 \\ -6 & 16 & 4 \\ -12/5 & 4 & 4 \end{pmatrix}$$

Find;

i. The distribution of $Y = \begin{pmatrix} X1-X3 \\ X1-X2-X3 \end{pmatrix}$ (3 marks)

ii. The correlation matrix for the data (3 marks)

iii. The distribution of X1 given that X2=23 and X3=14. (3 marks)

iv. The partial correlation coefficient between X1 and X2 for fixed values of X3 and that of X1and X3 for the fixed values of X2. (2 marks)

b) Use the data below for a bivariate normal distribution to test at $\alpha = 0.05$ level the hypothesis $H_0 = (17 \quad 15)'$ vs $H_1 \neq (17 \quad 15)'$.

$$X = \begin{pmatrix} 15 & 13 & 12 & 15 & 17 & 10 & 16 \\ 19 & 15 & 17 & 21 & 24 & 20 & 17 \end{pmatrix}$$ (5 marks)

c) Given that $\bar{X}_1 = (33 \quad 12 \quad 10)'$ comes from population I and

$\bar{X}_2 = (27 \quad 19 \quad 8)'$ Comes from population II and both populations have the same sample covariance matrix.

$$S = \begin{pmatrix} 20 & -4 & 15 \\ -4 & 16 & 0 \\ 15 & 0 & 4 \end{pmatrix}$$

Use discriminant rule to classify $X = (34 \quad 11 \quad 8)'$ (5 marks)

(c) For the matrix

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

i. Could A be a covariance matrix? Explain (2 marks)

ii. Obtain determinant of A and A^{-1} (3 marks)

iii. Compute the spectral decomposition of A (4 marks)

QUESTION TWO (20 Marks)

i. Observations on three responses are collected for two treatments as shown in the table below.

Treatment	1	1	1	2	2
x_1	12	13	8	11	10
x_2	19	18	14	11	15
x_3	8	9	7	14	12

find;

i) The matrix of sum of squares due to the treatment. (4 marks)

ii) The matrix of residual sum of squares (5 marks)

ii. Let random variables $\underline{x}' = [x_1, x_2, x_3]$ be distributed as $N_3(\underline{\mu}, \Sigma)$

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

Find the following

i. Correlation matrix of $\underline{x}$ (2 marks)

ii. The distribution of $z = 4x_1 - 6x_2 + x_3$ (3 marks)

iii. The distribution of $z = \begin{pmatrix} x_1 - x_2 + x_3 \\ 2x_1 + x_2 - x_3 \end{pmatrix}$ (3 marks)

iv. The Wilk's lambda statistic and use it to test the hypothesis that there is no treatment effects(use $\alpha = 0.05$) (3 marks)

QUESTION THREE (20 MARKS)

a) Describe briefly any two objectives of scientific investigation based on multivariate data (2 marks)

b) Briefly explain the idea behind MANOVA stating all the assumptions. (2 marks)

c) Distinguish between Factor Analysis and Canonical correlation Analysis (2 marks)

d) Consider data matrix for n=3 for a bivariate distribution

$$X = \begin{bmatrix} 6 & 10 & 8 \\ 9 & 6 & 3 \end{bmatrix}$$

$$\bar{X} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}$$

i. Evaluate the observed T^2 for $\underline{\mu}'_0 = [9 \quad 5]$. (6 marks)

ii. What is the sampling distribution of T^2 in this case? (6 marks)

QUESTION FOUR (20 MARKS)

a) Let X ~ N(μ, Σ) be a trivariate normal random vector. Suppose a certain sample gave

$$S=\begin{pmatrix} 64 & 0 & 9.6 \\ 0 & 16 & 0 \\ 9.6 & 0 & 36 \end{pmatrix}$$

find

i. The eigen values of this matrix (4 marks)

ii. The first two principal components (4 marks)

iii. The total variance explained by the two components. (4 marks)

b) Consider the covariance matrix

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

And the derived correlation matrix

$$\rho = \begin{bmatrix} 1 & 0.4 \\ 0.4 & 100 \end{bmatrix}$$

Determine the principal components for ∑ providing percentage of explained variability for each variate. (6 marks)

QUESTION FIVE (20 Marks)

a) Suppose

$$\sum \text{(covariance matrix)} = \begin{bmatrix} 4 & 1 & 2 \\ 1 & 9 & -3 \\ 2 & -3 & 25 \end{bmatrix}$$

Obtain standard deviation and population correlation matrix in the form of $v^{1/2}$ and ρ respectively. (5 marks)

b) Given the deviation vectors

$$e_1 = \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} \text{ and } e_2 = \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix}$$

Compute the sample variance-covariance matrix S_n and the sample correlation matrix γ using geometric concept. (3 marks)

c) The classic blue pullovers data given below is a data set consisting of 10 measurements of 4 variables. A textile shop manager is studying the sales of classic blue pullovers over 10 periods. He uses three different marketing methods and hopes to understand his sales as a fit of these variables using statistics. The variables measured are:

Sno.	**Sales**	**Price**	**Advert**	**Ass.Hours**
1	230	125	200	109
2	181	99	55	107
3	165	97	105	98
4	150	115	85	71
5	97	120	0	82
6	192	100	150	103
7	181	80	85	111
8	189	90	120	93
9	172	95	110	86
10	170	125	130	78

Note: $\mathbf{X_1}$**:** Number of sold pullovers

$\mathbf{X_2}$ **:** Price in (EUR)

$\mathbf{X_3}$ **:** Advertisement costs in local newspapers (in EUR)

$\mathbf{X_4}$ **:** Presence of a sales assistant (in hours per period)

i. Calculate the vector of the means for the four variables in the dataset (4 Marks)

ii. Calculate the sample covariance matrix (4 Marks)

iii. Calculate the sample correlation matrix (4 Marks)

/END/