MACHAKOS UNIVERSITY

University Examinations 2021/2022 Academic Year

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS AND STATISTICS

THIRD/FOURTH YEAR FIRST SEMESTER EXAMINATION FOR

BACHELOR OF SCIENCE (MATHEMATICS AND COMPUTER SCIENCE)

BACHELOR OF SCIENCE (STATISTICS AND PROGRAMMING)

BACHELOR OF SCIENCE (ECONOMICS AND STATISTICS)

BACHELOR OF SCIENCE (ACTUARIAL SCIENCE)

BACHELOR OF SCIENCE (MATHEMATICS)

BACHELOR OF EDUCATION (SCIENCE)

SMA 360 MULTIVARIATE STATISTICAL METHODS I

DATE: 24/8/2022 **TIME: 8.30-10.30 AM**

INSTRUCTION:

Answer Question One and Any Other Two Questions

QUESTION ONE (30 MARKS)

a) Define the p-dimensional multivariate normal density. (2 marks)

b) State three properties of the multivariate normal distribution. (3 marks)

c) Suppose $\boldsymbol{X} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and given that $\boldsymbol{Y_1} = \boldsymbol{X_1} - \boldsymbol{X_2}$ and $\boldsymbol{Y_2} = \boldsymbol{X_1} - \boldsymbol{X_3}$**.** Determine

i. $E(\boldsymbol{Y})$ (2 marks)

ii. $Var(\boldsymbol{Y})$ (4 marks)

d) Given two random variables X and Y with joint and marginal probabilities as below, determine the variance covariance matrix.

	0	1	P(x)
-1	0.24	0.06	0.3
0	0.16	0.14	0.3
1	0.40	0.00	0.4
P(y)	0.8	0.2	1

(8 marks)

e) Given the covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} \mathbf{4} & \mathbf{1} & \mathbf{2} \\ \mathbf{1} & \mathbf{9} & \mathbf{-3} \\ \mathbf{2} & \mathbf{-3} & \mathbf{25} \end{bmatrix}$, determine the correlation matrix $\boldsymbol{\rho}$. (5 marks)

f) Determine the maximum likelihood estimates of the 2 x 1 mean vector $\boldsymbol{\mu}$ and the 2 x 2 covariance matrix $\boldsymbol{\Sigma}$ based on the random sample

$$\mathbf{X} = \begin{bmatrix} 3 & 6 \\ 4 & 4 \\ 5 & 7 \\ 4 & 7 \end{bmatrix}$$

from a bivariate normal population. (6 marks)

QUESTION TWO (20 MARKS)

a) Consider the data matrix $\boldsymbol{X} = \begin{bmatrix} 1 & 4 & 3 \\ 6 & 2 & 6 \\ 8 & 3 & 3 \end{bmatrix}$ and the linear combinations

$\boldsymbol{b}'\boldsymbol{X} = [1 \quad 1 \quad 1] \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$ and $\boldsymbol{c}'\boldsymbol{X} = [1 \quad 2 \quad -3] \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$. Evaluate the sample means, variances and covariance of $\boldsymbol{c}'\boldsymbol{X}$ and $\boldsymbol{b}'\boldsymbol{X}$. (10 marks)

b) Let $\boldsymbol{X} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\mu}' = [1 \quad -1 \quad 2]$ and $\boldsymbol{\Sigma} = \begin{bmatrix} 4 & 0 & -1 \\ 0 & 5 & 0 \\ -1 & 0 & 2 \end{bmatrix}$. Determine whether the following random variables are independent. Explain.

i. $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$

ii. $\boldsymbol{X}_1$ and $\boldsymbol{X}_3$

iii. $\boldsymbol{X}_2$ and $\boldsymbol{X}_3$

iv. $(\boldsymbol{X}_1, \boldsymbol{X}_3)$ and $\boldsymbol{X}_2$ (4 marks)

c) State the multivariate version of the central limit theorem. (3 marks)

d) State the law of large numbers and briefly explain its practical importance. (3 marks)

QUESTION THREE (20 MARKS)

a) Let $\mathbf{X}$ be $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with $\boldsymbol{\mu}' = [2 \quad -3 \quad 1]$ and $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}$. Determine the distribution of $3\boldsymbol{X}_1 - 2\boldsymbol{X}_2 + \boldsymbol{X}_3$. (5 marks)

b) Let $\boldsymbol{X} \sim \boldsymbol{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Suppose $\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{b}$ where $\mathbf{A}$ is a qxp matrix of constants and $\mathbf{b}$ is a non-zero $qx1$ vector of constants. Determine the distribution of $\mathbf{Y}$. (5 marks)

c) Given that $f(x_1, x_2)$ is the bivariate normal density, show that $f(x_1|x_2)$ is $N\left(\mu_1 + \frac{\sigma_{12}}{\sigma_{22}}(x_2 - \mu_2), \sigma_{11} - \frac{{\sigma_{12}}^2}{\sigma_{22}}\right)$. (10 marks)

QUESTION FOUR (20 MARKS)

a) Consider a bivariate normal distribution with $\mu_1 = 0$, $\mu_2 = 1$, $\sigma_{11} = 2$, $\sigma_{22} = 1$ and $\rho_{12} = 0.5$.

i. Write out the bivariate normal density. (8 marks)

ii. Write out the squared statistical distance expression as a function of x_1 and x_2. (2 marks)

b) Let $\boldsymbol{X_1}, \boldsymbol{X}_2, \boldsymbol{X}_3, \boldsymbol{X_4}, \boldsymbol{X_5}$ be independent and identically distributed random vectors with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

i. Determine the mean vector and covariance matrices for each of the two linear combinations of random vectors

$\frac{1}{5}\boldsymbol{X_1} + \frac{1}{5}\boldsymbol{X}_2 + \frac{1}{5}\boldsymbol{X}_3 + \frac{1}{5}\boldsymbol{X_4} + \frac{1}{5}\boldsymbol{X_5}$ and $\boldsymbol{X_1} - \boldsymbol{X}_2 + \boldsymbol{X}_3 - \boldsymbol{X_4} + \boldsymbol{X_5}$

in terms of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. (8 marks)

ii. Determine the covariance between the two linear combinations of random vectors. (2 marks)

QUESTION FIVE (20 MARKS)

a) Given the data

z_1	10	5	7	19	11	8
y	15	9	3	25	7	13

Fit the linear regression model $Y_j = \beta_0 + \beta_1 z_{jt} + \varepsilon_j,\ j = 1,2,\dots,6$. Specifically, calculate the least squares estimates $\widehat{\boldsymbol{\beta}}$, the fitted values $\widehat{\boldsymbol{y}}$, the residuals $\hat{\boldsymbol{\varepsilon}}$ and the residual sum of squares $\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}$. (10 marks)

b) Let $\boldsymbol{X_1}, \boldsymbol{X}_2, \dots, \boldsymbol{X_n}$ be a random sample from a joint distribution that has mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Prove that $\overline{\boldsymbol{X}}$ is an unbiased estimator of $\boldsymbol{\mu}$ and its covariance matrix is $\frac{1}{n}\boldsymbol{\Sigma}$. (10 marks)