



MACHAKOS UNIVERSITY

University Examinations 2018/2019

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

FOURTH YEAR SECOND SEMESTER EXAMINATION FOR
BACHELOR OF SCIENCE (STATISTICS AND PROGRAMMING)

SMA 466: MULTIVARIATE ANALYSIS

DATE: 8/5/2019

TIME: 8:30 – 10:30 AM

INSTRUCTIONS:

Answer Question One and Any Other Two Question

QUESTION ONE (COMPULSORY) (30 MARKS)

- a) Let $\underline{X}$ be a $p \times 1$ random vector and $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ be a random sample of size n from the population.
- Describe
- i) $\underline{X}_r$ for some $r = 1, 2, \dots, n$
 - ii) The sample mean vector $\bar{\underline{X}}$
 - iii) The sample dispersion matrix S .
 - iv) The data matrix X (8 marks)
- b) Using the notations in part (a), show that
- i) $E(\bar{\underline{X}}) = \underline{\mu}$ and
 - ii) $E(S) = \frac{n-1}{n} \Sigma$ where $\underline{\mu}$ and Σ denote the population mean vector and variance-covariance vector respectively. (7 marks)
- c) Let $\underline{X}$ be a $p \times 1$ random vector and $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ be a random sample of size n from the population.
- Describe
- (i). $\underline{X}_r$ for some $r = 1, 2, \dots, n$

(ii). The sample mean vector $\bar{\underline{X}}$

(iii). The sample dispersion matrix $\underline{S}$.

(iv). The data matrix $\underline{X}$

(8 marks)

c) Suppose $\underline{X} = (x_1 \ x_2 \ x_3)^T$

$$\underline{\mu}_3 = (\mu_1, \mu_2, \mu_3)^T \text{ and } \underline{\Sigma} = \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & 0 \\ \rho^2 & 0 & 1 \end{pmatrix}$$

If $\underline{X}$ is $N_3[\underline{\mu}, \underline{\Sigma}]$

Show that $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ given x_3 has a multivariate normal distribution with mean $\underline{r}$ and

covariance matrix $\underline{V}$

where

$$\underline{r} = (\mu_1 + \rho^2(x_3 - \mu_3), \mu_2)^T$$

and

$$\underline{V} = \begin{pmatrix} 1 - \rho^4 & \rho \\ \rho & 1 \end{pmatrix} \quad (7 \text{ marks})$$

d) i) Describe the Wishart distribution,

$$\underline{W}_p = [\underline{\Sigma}, n]$$

ii) Let the random vector variable $\underline{Z}_{p \times 1}$ have the multivariate Normal distribution with mean $\underline{O}_{p \times 1}$, and covariance matrix $\underline{I}_{p \times p}$ where $\underline{I}_{p \times p}$ is a p by p identity matrix.

Determine the distribution of $\underline{Z}^T \underline{Z}$ and hence or otherwise show that

$$\underline{W}_p[\underline{I}_{p \times p}, P] = \chi^2_{(p)} \quad (8 \text{ marks})$$

QUESTION TWO (20 MARKS)

Consider a random vector $\underline{x}^1 = (x_1 \ x_2)$. Assume $\underline{x}$ has $N_2[\underline{\mu}, \underline{\Sigma}]$ distribution. A

sample of size $n = 25$ has $\overline{\underline{X}} = (185.72, 183.84)^1$

a) Suppose $\underline{\Sigma}$ is unknown and the sample dispersion matrix

$$S = \begin{pmatrix} 91.481 & 66.875 \\ 66.875 & 96.775 \end{pmatrix}$$

$$\text{Test } H_0: \underline{\Sigma} = \underline{\Sigma}_0 \quad \text{where } \underline{\Sigma}_0 = \begin{pmatrix} 100 & 0 \\ 0 & 100 \end{pmatrix}$$

Assume 5% level of significance.

(10 marks)

b) Suppose now $\underline{\Sigma} = \begin{pmatrix} 100 & 0 \\ 0 & 100 \end{pmatrix}$ while the mean vector $\underline{\mu}$ is unknown.

Test

$$H_0: \underline{\mu} = \underline{\mu}_0$$

where $\underline{\mu}_0 = (182 \ 182)^1$ at 5% level of significance.

(10 marks)

QUESTION THREE (20 MARKS)

Let the random vector $\underline{Y}$ be $N_4[\underline{\mu}, \underline{\Sigma}]$ where $\underline{\mu} = (2 \ 5 \ -2 \ 1)^1$, and

$$\underline{\Sigma} = \begin{pmatrix} 9 & 0 & 3 & 3 \\ 0 & 1 & -1 & 2 \\ 3 & -1 & 6 & -3 \\ 3 & 2 & -1 & 7 \end{pmatrix}$$

Let $\underline{Y} = (Y_1 \ Y_2 \ X_1 \ X_2)$ (6 marks)

a) Determine the distribution for $(Y_1 \ Y_2)^1 / (X_1 \ X_2)^1$

b) Hence or otherwise, calculate

i) $Cov(Y_1, Y_2)$

ii) $Cov((Y_1, Y_2) / (X_1, X_2)^1)$

iii) Comment on the results in b(i) and b(ii) (8 marks)

c) Find the partial correlation matrix

$$(Corr(Y_i, Y_j) / (X_1, X_2)^1) \quad (6 \text{ marks})$$

QUESTION FOUR (20 MARKS)

a) Show that if $\underline{Y}$ is $N_p[\underline{\mu}, \underline{\Sigma}]$ then an $r \times 1$ subvector of $\underline{Y}$ has an r -variate normal distribution with the same means, variances and covariances as in the original p -variate normal distribution. (6 marks)

b) Using part (a), or otherwise, Show that any individual variable Y_i in $\underline{Y}$ is distributed as $N_p[\underline{\mu}_i, \sigma_{ii}]$ where $\underline{\mu}_i = E(Y_i)$, $\sigma_{ii} = V(Y_i)$ (6 marks)

c) Show that if $\underline{Y}$ and $\underline{X}$ are jointly multivariate normal with $\underline{\Sigma}_{yx} \neq 0$, then the conditional distribution of $\underline{Y}$ given as $\underline{X}$, $f(\underline{Y}/\underline{X})$ is multivariate normal with mean vector and covariance matrix.

$$E(\underline{Y}/\underline{X}) = \underline{\mu}_y + \underline{\Sigma}_{yx} \underline{\Sigma}_{xx}^{-1} (\underline{X} - \underline{\mu}_x)$$

$$V(\underline{Y}/\underline{X}) = \underline{\Sigma}_{yy} - \underline{\Sigma}_{yx} \underline{\Sigma}_{xx}^{-1} \underline{\Sigma}_{yx}$$

respectively, given corresponding covariance matrices (8 marks)

QUESTION FIVE (20 MARKS)

a) Let $\underline{X}$ be a p -component random vector which has the multivariate distribution, $N_p[\underline{\mu}, \underline{\Sigma}]$. Show that for the sample mean vector $\bar{\underline{X}}$ and sample variance S ;

i) $\bar{\underline{X}} \sim N_p \left[\underline{\mu}, \frac{\underline{\Sigma}}{n} \right]$ (4 marks)

ii) $S \sim W_p \left[\frac{\underline{\Sigma}}{n}, n-1 \right]$ (4 marks)

iii) $\frac{n-p}{p} \left(\bar{\underline{X}} - \underline{\mu} \right)^t S^{-1} \left(\bar{\underline{X}} - \underline{\mu} \right) \sim F(p, n-1)$ (4 marks)

b) Let $\underline{Y}$ be a random vector with mean $\underline{\mu}$ and covariance matrix $\underline{\Sigma} = \mathbf{I}_{p \times p}$. Show that if $\underline{Y}$ is $N_p[\underline{\mu}, \underline{\Sigma}]$, then $E \{ (\underline{Y} - \underline{\mu})^t (\underline{Y} - \underline{\mu}) \} = p$ (8 marks)