KENYATTA UNIVERSITY

UNIVERSITY EXAMINATIONS 2011/2012

SECOND SEMESTER EXAMINATION FOR THE DEGREE OF BACHELOR OF SCIENCE

SMA 466: MULTIVARIATE METHODS

DATE: THURSDAY, 5TH APRIL 2012 **TIME: 8.00 A.M. – 10.00 A.M.**

INSTRUCTIONS

Attempt question one and any two other questions

Question one (30 marks)

a) Define the following terms
   i. Random vector
   ii. Data matrix
   iii. Total variance
   iv. Symmetric matrix
   v. Bayes' discriminant rule (5 marks)

b) Given the data matrix $X = \begin{bmatrix} 9 & 1 \\ 5 & 3 \\ 1 & 2 \end{bmatrix}$ obtain
   i. The mean vector
   ii. Sample covariance matrix
   iii. The generalized variance (6 marks)

c) Briefly explain the problems of :
   i. Principal component analysis
   ii. Discriminant analysis
   iii. Canonical analysis (9 marks)

d) Let $\underline{y}_{ij} = \underline{\mu} + \underline{t}_i + \underline{e}_{ij} \qquad i = 1, 2, \ldots, v \quad j = 1, 2, \ldots, n_i$

   Where $\underline{\mu}$ is grand mean vector, $\underline{t}_i$ is the vector of treatment effects, $\underline{e}_{ij}$ is the vector of random effects. It is assumed to be normally distributed ie $\underline{e}_{ij} \sim N(0, \Sigma)$ Σ *is unknown*
   Construct a test for the treatment effects.(10 marks)

Question Two (20 marks)

For a certain experiment with n = 42, it was found that

$$\underline{\overline{X}} = \begin{bmatrix} .564 \\ .603 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} .0144 & .0117 \\ .0117 & .0146 \end{bmatrix}$$

a) Construct a test for the null hypothesis $H_0 : \underline{\mu}' = [.55 \quad .60]$ at $\alpha = .05$ level of significance.(15 marks)

b) Write down the simultaneous confidence interval for $\underline{a}'\underline{\mu}$ assuming Σ is known and $\underline{a}$ is a non zero vector of constants. (5 marks)

Question Three (20 marks)

a. Define the vector of principal components of a normally distributed random vector with mean $\boldsymbol{\underline{\mu}}$ and variance covariance matrix $\boldsymbol{\Sigma}$.(3 marks)

b. Given the matrix $\Sigma = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix}$

i. Obtain the two principal components.(7 marks)

ii. What percentage is explained by the first principal component.(2 marks)

iii. If you wanted to discard one variable, which one would you discard?(8 marks)

Question Four (20 marks)

a. Define the maximum likelihood discriminant rule. (3marks)

b. Given $\underline{\overline{X}}_1 = (5.006, 3.428)'$ $\underline{\overline{X}}_2 = (5.936, 2.770)'$ and $n_1 = n_2 = 50$

$$S_1 = \begin{bmatrix} 0.1218 & 0.0972 \\ 0.0972 & 0.1408 \end{bmatrix}, \quad S_2 = \begin{bmatrix} 0.2611 & 0.0835 \\ 0.0835 & 0.0965 \end{bmatrix},$$

i. Assuming $\Pi_1 \sim N(\mu_1, \Sigma)$ *and* $\Pi_2 \sim N(\mu_2, \Sigma)$, derive the maximum likelihood discriminant rule. (14 marks)

ii. In which population would you allocate $\underline{\mathbf{x}}$ =(5.832, 3.252)' using the maximum likelihood discriminant rule? (3 marks)

Question Five (20 marks)

a) Write down the linear regression model in vector matrix notation explaining each term and giving any assumptions (3 marks)

b) Obtain the ordinary linear estimator of β and show that it is unbiased.(5 marks)

c) Given the following data

$$\underline{y}' = (15, 9, 3, 25, 7, 13), \qquad X = \begin{bmatrix} 2 & 10 \\ 3 & 5 \\ 3 & 7 \\ 6 & 19 \\ 7 & 11 \\ 9 & 18 \end{bmatrix}$$

i. Obtain the ordinary least squares estimator of β. (10 marks)

ii. Obtain also the value of y given the point 5,12(2marks)

****