INDIAN STATISTICAL INSTITUTE
CHENNAI CENTRE
M.Stat First Year
2016-17 Semester II

Multivariate Analysis
End-Sem Examination

Total Marks 100. Date: 5 May 2017 Duration: 3 hours

Comments : This paper carries 110 *marks. Answer as much as you can. Maximum you can score is* 100

1. (a) Consider the classification problem of an observation $\mathbf{x}_0$ into either of normal populations Π_1 or Π_2 with common variance-covariance matrix Σ and μ_1 and μ_2 as their mean vectors respectively. Let $\delta = \mu_1 - \mu_2$ and K be defined by $\frac{1}{2}\delta'\Sigma^{-1}(\mu_1 + \mu_2)$. Show that $(\delta'\Sigma^{-1}\mathbf{x}_0 - K)$ can be written as $\frac{1}{2}(D_2^2 - D_1^2)$, where $D_i^2 = (\mathbf{x}_0 - \mu_i)'\Sigma^{-1}(\mathbf{x}_0 - \mu_i), \quad i = 1, 2$

   (b) Suppose, prior probabilities for Π_1 or Π_2 are respectively given by $q_1 = 0.6$, and $q_2 = 0.4$ and $\mu_1 = (2,4)', \quad \mu_2 = (6,8)'$, and $\Sigma = \begin{bmatrix} 4 & 3 \\ 3 & 9 \end{bmatrix}$
   Compute the Bayes error for $C(1|2)$ and $C(2|1) = c$ [15+10=25]

2. (a) Describe the $k-$factor model and how the related parameters are estimated in Principal Factor Analysis.

   (b) If the $k-$factor model holds, show that it is scale invariant but the factor loadings may not be unique.

   (c) Suppose that $X' = (X_1, X_2)$ has a bivariate multinomial distribution with $n = 1$, so that $X_1 = 1$ with probability p and $X_2 = 1 - X_1$. Find the principal components of X and their variances. [10+8+7=25]

3. (a) Let $\mathbf{y}_{ijk}$ be independently distributed p-variate Normal variable, where $i = 1, \ldots, g$, $j = 1, \ldots, b$ and $k = 1, \ldots, n$. Consider the fixed effect model

$$y_{ijk} = \mu + \tau_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}$$

$$\text{Where} \quad \sum_{i=1}^{g} \tau_i = 0, \quad \sum_{j=1}^{b} \beta_j = 0$$

$$\sum_{i=1}^{g} \gamma_{ij} = 0 = \sum_{j=1}^{b} \gamma_{ij} \quad \text{and}$$

$$\epsilon_{ijk} \sim N_p(\mathbf{0}, \Sigma)$$

Now define $\mathbf{z}_{ijk} = A\mathbf{y}_{ijk} + \mathbf{b}$, where $\mathbf{b}$ is a fixed p-vector and A is $p \times p$ non-singular matrix. Show that the test criterion to test equality of several multivariate means is invariant under the above transformation, alternatively $\Lambda_y^* = \Lambda_z^*$ where Wilk's Λ, is given by $\Lambda_{fac1}^* = \frac{|SSP_{res}|}{|SSP_{res} + SSP_{fac1}|}$, $\Lambda_{fac2}^* = \frac{|SSP_{res}|}{|SSP_{res} + SSP_{fac2}|}$, $\Lambda_{int}^* = \frac{|SSP_{res}|}{|SSP_{res} + SSP_{int}|}$.

(b) Three samples of sizes $n_1 = 3, n_2 = 2$ and $n_3 = 3$ have been obtained from three bivariate normal populations with same Variance-covariance matrix. Write down the MANOVA model and analyse the data to test equality of the three mean vectors. Describe and show your computations clearly to arrive at the test statistics.

Sample 1 $\begin{bmatrix} 9 \\ 3 \end{bmatrix} \begin{bmatrix} 6 \\ 2 \end{bmatrix} \begin{bmatrix} 9 \\ 7 \end{bmatrix}$

Sample 2 $\begin{bmatrix} 0 \\ 4 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

Sample 3 $\begin{bmatrix} 3 \\ 8 \end{bmatrix} \begin{bmatrix} 1 \\ 9 \end{bmatrix} \begin{bmatrix} 2 \\ 7 \end{bmatrix}$

Hint : Use the test statistics $\left(\frac{\sum n_i - g - 1}{g-1}\right)\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \sim F_{2(g-1), 2(\sum n_i - g - 1)}$. Where $p = 2$ and $g = 3$, $i = 1, 2, \cdots g$. $F_{4,8,0.05} = 3.83$ and $F_{4,8,0.01} = 7.00$ [20+15=35]

4. Let $\mathbf{X}_i, i = 1, 2, \ldots, n$ be a random sample of size n from $N_p(\mu\mathbf{1}, \Sigma)$ where μ is unknown but Σ is known, $\bar{\mathbf{x}}$ is the sample average and $\mathbf{1}' = (1, 1, \ldots, 1)$. Let $T_1 = \frac{\mathbf{1}'\bar{\mathbf{X}}}{\mathbf{1}'\mathbf{1}}$ and $T_2 = \frac{\mathbf{1}'\Sigma^{-1}\bar{\mathbf{X}}}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}$

   (a) Show that both T_1 and T_2 are unbiased for μ.

   (b) Obtain the variances of T_1 and T_2.

   (c) Show that $Var(T_2) \leq Var(T_1)$ [2+3+5=10]

5. Write short notes on

   (a) Clustering Methods and needs

   (b) Multi Dimensional Scaling

   (c) Canonical Correlation [5×3=15]