(3 Hours) [Total Marks: 60

N.B.

i. Attempt any **TWO** questions among question numbers 1, 2, 3, and any **TWO** questions among questions 4, 5, 6.
ii. Figures to the right indicate full marks.
iii. Simple non-programmable calculator is allowed.

1. (a) Let $\underline{X}$ be a $p \times 1$ random vector with $E(\underline{X}) = \underline{0}$ and $V(\underline{X}) = \Sigma$. Define multiple correlation coefficient and partial correlation coefficient. In usual notation prove that (09)

$$1 - \rho^2_{1.23\cdots p} = \left(1 - \rho^2_{12.34\ldots p}\right)\left(1 - \rho^2_{13.4\ldots p}\right)\left(1 - \rho^2_{14.5\ldots p}\right)\ldots\left(1 - \rho^2_{1p-1.p}\right)\left(1 - \rho^2_{1p}\right)$$

(b) Derive the Hotelling's T^2 statistic for testing $H_0 : \mu_1 = \mu_2 = \cdots = \mu_p$ based on a random sample of size n from $N_p(\mu, \Sigma)$ where Σ is unknown. (06)

2. Consider k independent p-variate normal populations given by $N_p(\underline{X}_\alpha, \underline{\mu}_\alpha, \Sigma_\alpha);\ \alpha = 1, 2, \ldots, k$. Derive the Likelihood Ratio Test to test the following two null hypotheses. (15)
(i) $H_a : \Sigma_1 = \Sigma_2 = \ldots = \Sigma_k$ (μ_α unknown : $\alpha = 1,2,\ldots, k.$)
(ii) $H_b : \mu_1 = \mu_2 = \ldots = \mu_k ; \Sigma_1 = \ldots \Sigma_k$
Give the corresponding asymptotic distributions.

3. (a) State the model, and hypothesis of a two-way multivariate analysis of variance (MANOVA) and explain every term in the model. Draw the MANOVA table and discuss the testing procedure. (11)

(b) Let $\underline{X} = (X_1, X_2, \ldots, X_p)'$ with $E(\underline{X}) = \underline{0}$ and $V(\underline{X}) = \Sigma$. Prove in usual notations that $\rho_{1.23\ldots p}$ is the maximum correlation between X_1 and any other linear function of $\underline{X}^*$ where $\underline{X}^* = (X_2, \ldots, X_p)'$ (04)

4. (a) Define population canonical correlations and population canonical variates. Give geometrical interpretation of the population canonical correlation analysis. (05)

(b) Let $\underline{Z}^{(1)} = V_{11}^{-1/2}\left(\underline{X}^{(1)} - \underline{\mu}^{(1)}\right)$ and $\underline{Z}^{(2)} = V_{22}^{-1/2}\left(\underline{X}^{(2)} - \underline{\mu}^{(2)}\right)$ be two vectors of standardized variables. If $\rho_1^*, \rho_2^*, \dots, \rho_p^*$ are the canonical correlations for the vectors $\underline{X}^{(1)}, \underline{X}^{(2)}$ and $(U_i, V_i) = \left(\underline{a}_i' \underline{X}^{(1)}, \underline{b}_i' \underline{X}^{(2)}\right),\ i = 1,2,..,p$ are the associated canonical variates, determine the canonical correlations and canonical variates for the $\underline{Z}^{(1)}, \underline{Z}^{(2)}$. (05)

(c) Show that if λ_i is an eigenvalue of $\Sigma_{11}^{-1/2}\,\Sigma_{12}\,\Sigma_{22}^{-1}\,\Sigma_{21}\,\Sigma_{11}^{-1/2}$ with associated eigenvector $\underline{e}_i, \lambda_i$ is also an eigenvalue of $\Sigma_{11}^{-1}\,\Sigma_{12}\,\Sigma_{22}^{-1}\,\Sigma_{21}$, with eigenvector $\Sigma_{11}^{-1/2}\,\underline{e}_i$, thereby giving an alternative calculation of canonical correlations and variates. (05)

5. (a) Let π_1 population be $N_p\left(\underline{\mu}_1, \Sigma\right)$ and π_2 population be $N_p\left(\underline{\mu}_2, \Sigma\right)$. Show that the estimated minimum expected cost of misclassification (ECM) rule is given by (05)

Allocate $\underline{x}_0$ to π_1 if $(\underline{\bar{x}}_1 - \underline{\bar{x}}_2)' s_{pooled}^{-1}\ \underline{x}_0 - \frac{1}{2}(\underline{\bar{x}}_1 - \underline{\bar{x}}_2)' s_{pooled}^{-1}\ (\bar{x}_1 + \bar{x}_2) \geq \ln \left[\left(\frac{c(1/2)}{c(2/1)}\right)\left(\frac{p_2}{p_1}\right)\right]$

Allocate $\underline{x}_0$ to π_2 otherwise

(b) For two populations show that $\frac{\left(\text{sum of squared distances from population s means to the overall mean}\right)}{(\text{variance of } y)}$ is proportional to $\frac{\left(\underline{l}'\underline{\delta}\right)^2}{\underline{l}'\Sigma\,\underline{l}}$ where $\underline{\delta} = \left(\underline{\mu}_1 - \underline{\mu}_2\right)$. Further, show that this ratio is maximized by the linear combination $\underline{l} = c\,\Sigma^{-1}\,\underline{\delta} = c\,\Sigma^{-1}\left(\underline{\mu}_1 - \underline{\mu}_2\right),\ c \neq 0.$ (06)

(c) Derive Fisher's discriminant function and the classification rule for two multivariate normal populations. (04)

6. (a) Show that canonical correlations are invariant under nonsingular linear transformations of the form $\underset{(p\times p)(p\times 1)}{C\,\underline{X}^{(1)}}$ and $\underset{(q\times q)(q\times 1)}{D\,\underline{X}^{(2)}}$ of the $\underline{X}^{(1)}, \underline{X}^{(2)}$ variables. (05)

(b) Show that the first canonical correlation is larger than the absolute value of any entry in ρ_{12}, in usual notations. (05)

(c) Let π_1 population be $N_p(\underline{\mu}_1, \Sigma)$ and π_2 population be $N_p(\underline{\mu}_2, \Sigma)$. Derive an expression for the optimum error rate when $p_1 = p_2 = \frac{1}{2}$. (05)

****