Department of Mathematics
Indian Institute of Technology Delhi

MAL 466 / MTL766 -- Multivariate Statistical Methods

Minor Examination

Time: 2 Hours **Date 01-05-2017** **Total Marks: 50**

YOU MAY ATTEMPT ALL <u>SEVEN</u> QUESTIONS. THE AWARDED SCORE WILL BE RESTRICTED TO 50.

Q1. Prove that the sample generalized variance computed from a $n \times p$ data matrix X is zero if and only if p deviation vectors defined on X are linearly dependent.

[5]

Q2. Define a 2-sample Hotelling T^2-statistic for testing equality of mean vectors of two Normal populations $N_p(\underset{\sim}{\mu_1}, \Sigma)$ and $N_p(\underset{\sim}{\mu_2}, \Sigma)$. Find the distribution of this statistic assuming the distribution of one sample T^2-statistic.

[6]

Q3. (a) Let $\underset{\sim}{X}$ have a $N_p(\underset{\sim}{\mu}, \Sigma)$ distribution. Prove that $(\underset{\sim}{X} - \underset{\sim}{\mu})^T \Sigma^{-1} (\underset{\sim}{X} - \underset{\sim}{\mu})$ has χ^2- distribution with p-degrees of freedom.

(b) Show that an approximate distribution of $n(\bar{\underset{\sim}{X}} - \underset{\sim}{\mu})^T \Sigma^{-1} (\bar{\underset{\sim}{X}} - \underset{\sim}{\mu})$ is χ^2- distribution with p-degrees of freedom for a large $n - p$, where $\bar{\underset{\sim}{X}}$ is the sample mean vector of a random sample of size n from any population having covariance matrix Σ.

[4 + 4 = 8]

Q4. Reaction times to visual stimuli were obtained from 20 young normal men under three conditions A, B and C. The sample mean reaction times in hundredths of seconds were 21.05, 21.65 and 28.95, respectively. The sample covariance matrix was

$$S = \begin{bmatrix} 2.2605 & 2.1763 & 1.6342 \\ 2.1763 & 2.6605 & 1.8237 \\ 1.6342 & 1.8237 & 2.4710 \end{bmatrix}$$

(a) Test at 5% level of significance the Null hypothesis of equal stimulus condition effects upon the differences B – A and C – B.

(b) Find 99% T^2-confidence intervals for the contrasts $\mu_B - \mu_A$ and $\mu_C - \mu_B$.

[6 + 5]

Q5. Consider thelinear regression model $\underset{\sim}{Y} = Z\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$, $E(\underset{\sim}{\varepsilon}) = 0$ and $Cov(\underset{\sim}{\varepsilon}) = \sigma^2 I$, where design matrix Z is of dimension $n \times (r + 1)$.

(a) Prove the decomposition

Total Sum of Squares = Sum of squares due to regression + Sum of squares due to residuals

and interpret the three sum of squares involved.

(b) If the vector $\underset{\sim}{\varepsilon}$ has $N_n(0, \sigma^2 I)$ distribution then prove that residual sum of squares follow $\sigma^2\chi^2$- distribution with $n - r - 1$ degrees of freedom.

adj A

[5+ 4]

Q6. The following is the sample correlation matrix of examination scores of six subjects for 220 students.

$$R = \begin{array}{c} \begin{array}{cccccc} \text{HUL} & \text{Stats} & \text{Phy} & \text{TOC} & \text{ADA} & \text{PDE} \end{array} \\ \begin{bmatrix} 1.0 & 0.439 & 0.410 & 0.288 & 0.329 & 0.248 \\ & 1.0 & 0.351 & 0.354 & 0.320 & 0.329 \\ & & 1.0 & 0.164 & 0.190 & 0.181 \\ & & & 1.0 & 0.595 & 0.470 \\ & & & & 1.0 & 0.464 \\ & & & & & 1.0 \end{bmatrix} \end{array}$$

A 2- Factor model fitted to the above data by using maximum likelihood method yielded the following estimated factor loadings:

Variable	Estimated Factor Loadings	
	F_1	F_2
HUL	0.553	0.429
Stats	0.568	0.288
Phy	0.392	0.450
TOC	0.740	- 0.273
ADA	0.724	- 0.211
PDE	0.595	- 0.132

(a) Obtain maximum likelihood estimates of

(i) The specific variances

(ii) The communalities

(iii) The proportion of variance explained by each factor.

(b) Obtain the rotated factor loadings obtained by factor rotation using orthogonal transformation giving a rigid clockwise rotation of 20°. Interpret the rotated factors F_1^* and F_2^*.

[5 +6]

Q7. The largest eigenvalue computed from the sample covariance matrix S obtained from a random sample of size 100 from Normal population $N_5(\mu, \Sigma)$ is $\hat{\lambda}_1 = 1.426$. Use this information to obtain an approximate 95% confidence interval for the variance of the first population principal component.

[4]

$\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{D} N_p(0, 2\Lambda^2\Sigma)$

$\Lambda = diag(\lambda_1, \lambda_2, \ldots$

$\leq (52 \; 3 + 1)\lambda$

$\sqrt{2}\; \lambda\; 3$