OC

[144]

No of Printed pages: 03

SARDAR PATEL UNIVERSITY
M.Sc. 3rd Semester Examination
2012
Saturday, 1st December
2:30 p.m. to 5:30 p.m.
STATISTICS COURSE No. PS03CSTA02
(Multivariate Analysis)

Note: Figures to the right indicate full marks of the questions. **(Total Marks: 70)**

1 Wright the correct answer (each question carries one mark). **08**

(a) Let $\underline{X}$ have multivariate distribution. Then Σ is a variance-covariance matrix of $\underline{X}$ if and only if it is
(i) p. d. (ii) p. s. d. (iii) n. n. d. (iv) None

(b) Let $\underline{X} \sim N(\underline{\mu}, \Sigma)$, $\Sigma > 0$. Then the distribution of $(\underline{X}-\underline{\mu})'\Sigma^{-1}(\underline{X}-\underline{\mu})$ is
(i) $\chi_p^2(\underline{\mu}'\Sigma^{-1}\underline{\mu})$ (ii) $W_p(\Sigma,1)$ (iii) $W_p(\Sigma,n)$ (iv) χ_p^2

(c) If $W \sim W_p(I,n)$ then $|W|$ is distributed as
(i) χ_{np}^2 (ii) $W_p(I,1)$ (iii) χ_n^2 (iv) None

(d) Let $\underline{X} \sim N(\underline{0}, I)$. The characteristic function of $\underline{X}'\underline{X}$ is
(i) $|I_p - 2it\Theta\Sigma|^{-n/2}$ (ii) $|I_p - 2it\Theta|^{-n/2}$
(iii) $(1-2it)^{-n/2}$ (iv) none

(e) The Wilk's lamda distribution is a generalization of
(i) chi-square distribution (ii) t-distribution
(iii) F-distribution. (iv) beta distribution

(f) A multivariate regression model involves
(i) more than one explained variables (ii) more than one explanatory variables
(iii) more than one dummy variables (iv) none

(g) Test for equality of covariance matrices of k–independent normal population is also called
(i) test for homogeneity (ii) test for sphericity (iii) MANOVA (iv) none.

(h) In normal symmetric distribution all the variables
(i) have same means, same variances and covariances
(ii) are i.i.d. variables
(iii) are independent variables
(iv) none of the above

2 Answer any SEVEN of the following (each question carries two marks) **14**

(a) Let $\underline{X}^{\alpha} = (X_{\alpha 1},\ldots,X_{\alpha p})'$, $\alpha = 1,\ldots,N$, be independently distributed normal vectors with the same vector $\underline{\mu}$ and the same positive definite covariance matrix Σ, and let

$$V = \sum_{\alpha=1}^{N}(\underline{X}^{\alpha} - \overline{\underline{X}})(\underline{X}^{\alpha} - \overline{\underline{X}})' \text{ where } \overline{\underline{X}} = \frac{1}{N}\sum_{\alpha=1}^{N}\underline{X}^{\alpha} .$$

Show that $(\overline{\underline{X}}, V)$ are jointly complete sufficient for $(\underline{\mu}, \Sigma)$.

(b) Prove that V, defined in 2(a), is p. d. if and only if $N > p$.

(c) Let $\underline{X}_1,...,\underline{X}_N$ be a random sample of size N drawn from $N_p(\underline{\mu}, \Sigma)$, $\Sigma > 0$, and $N > p$. Let V be the corrected SP matrix partitioned as

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \begin{matrix} k \\ p-k \end{matrix}$$

k p-k

Obtain the distribution of $V_{11\bullet2} = V_{11} - V_{12}V_{22}^{-1}V_{21}$.

(d) If $X : p \times N$ is a data matrix from $N_p(\underline{0}, \Sigma)$, and if $C : N \times N$ is a symmetric real matrix then show that the distribution of XCX' is weighted sum of $W_p(\Sigma, 1)$ with weights are the eigenvalues of C matrix.

(e) Show that, in usual notations, $\frac{N-p}{p-1} \cdot \frac{\overline{R}^2}{1-\overline{R}^2} \sim F(p-1, N-p)$ when $\overline{\rho} = 0$.

(f) Obtain null-distribution of the sample correlation coefficient matrix $R = (r_{ij})$.

(g) Let $\underline{X}_1,...,\underline{X}_N$ be independently distributed as $N_4(\underline{\mu}, \Sigma)$, $\Sigma > 0$, $N > 4$. Derive the LRT for testing the hypothesis

$H_0 : \mu_1 - 2\mu_2 = \mu_2 - 2\mu_3 = \mu_3 - 2\mu_4$ against $H_1 \neq H_0$, where $\underline{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4)'$.

(h) Specifying the conditions, in usual notations, show that $B = \sum_{i=1}^{k} N_i(\overline{\underline{X}}_i - \overline{\underline{X}})(\overline{\underline{X}}_i - \overline{\underline{X}})'$ is distributed as $W_p(\Sigma, k-1)$.

(i) Write a brief note on a Growth Curve Model.

3 **(a)** If $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$ with rank of $\Sigma = k$, then obtain its density. **06**

(b) Show that $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$ with rank k if, and only if, **06**

$$\underline{X} = \underline{\mu} + B\underline{Y}, \qquad BB' = \Sigma$$

where B is p x k matrix of rank k and $\underline{Y} \sim N_k(\underline{0}, I)$.

Using above result define multivariate distribution.

OR

(b) Obtain the non-null distribution of the sample correlation coefficient matrix $R = (r_{ij})$. Hence find the distribution of R when p = 2. Also, obtain the non-null distribution of sample partial correlation coefficient between two variables when other variables are held fixed. **06**

4 **(a)** State and prove Bartlett's decomposition lemma. Using this derive the standard Wishart distribution. **06**

(b) Obtain the density of inverted Wishart distribution. Hence obtain the distributions of (i) σ^{pp}/v^{pp} and (ii) $\underline{h}'\Sigma^{-1}\underline{h}/\underline{h}'V^{-1}\underline{h}$, where $\underline{h}: p\times 1$ is an arbitrary vector. **06**

OR

(b) Derive distribution of Wilk's lamda statistic and discuss it application in testing of certain hypothesis in multivariate regression model. **06**

5 (a) Define the Hotelling's T^2 statistic. Show that it is invariant under a group of transformation (stated by you). Derive its density using one of the properties of Wishart distribution. **06**

(b) Describe Roy's Union-Intersection principle. Show that it is being used to test the hypothesis $H_0: \underline{\mu} = \underline{\mu}_0$ against $H_1: \underline{\mu} \neq \underline{\mu}_0$ when $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$. **06**

OR

(b) Derive the Hotelling's T^2 test for Randomized Block Design (RBD) when the underlying assumptions are not met. **06**

6 (a) Let $V_1,\ldots,V_k$ be independently distributed $W_p(I, n_i), i = 1,\ldots,k$, and let **06**

$V = \sum_{i=1}^{k} V_i$, $L_j = V^{1/2}V_jV^{1/2}$, $j = 1,\ldots,k-1$. Show that the joint probability

density function of $L_j, j = 1,\ldots,k-1$ is given by

$$C\prod_{i=1}^{k-1}|L_i|^{(n_i-p-1)/2}\left|I - \sum_{i=1}^{k-1}L_i\right|^{(n_k-p-1)/2}$$

where C is the normalizing constant.

(b) Derive the LRT test for homogeneity of k-independent p–variate normal populations. Also, derive its asymptotic distribution. **06**

OR

(b) Describe multivariate regression model with set of assumptions and estimate the parameters of the model. **06**

____