Multivariate Statistical Analysis Final Exam

January 12, 2008

2 pages, 10 problems, 100 points

1. Let components of $\mathbf{X}$ denote the responses to 3 kinds of treatments in a repeated measure. Assume that $\mathbf{X}$ is with a multivariate normal population of mean $\boldsymbol{\mu}$. From a certain experiment of 30 independent observations, we found that the sample average and the sample covariance matrix are $\bar{\mathbf{x}} = [-1 \quad 0 \quad 1]'$ and $\mathbf{S} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 8 & 2 \\ 0 & 2 & 5 \end{bmatrix}$, respectively. Suppose that the contrast matrix is $\mathbf{C} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$, test the hypothesis $H_0 : \mathbf{C}\boldsymbol{\mu} = 0$ at 10% significance level. (10%)

2. Two samples were taken from an experiment where two characteristics $\mathbf{X}_1$ and $\mathbf{X}_2$ were measured. Assume that samples 1 and 2 are taken from $N_2(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ and $N_2(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$, respectively. The sample covariance matrices and sample sizes of the two samples are $\mathbf{S}_1 = \begin{bmatrix} 6 & 1 \\ 1 & 4 \end{bmatrix}$, $n_1 = 16$ and $\mathbf{S}_2 = \begin{bmatrix} 10 & 3 \\ 3 & 6 \end{bmatrix}$, $n_2 = 16$, respectively. Use Box's test to Test the hypothesis $H_0 : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ .versus $H_1 : \boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2$ at 10% significance level. (10%)

3. Four subjects rate the performances of three voice synthesis systems, A, B, and C, in a 1 to 10 scale. The results are given as the following table

Subject\System	A	B	C
1	5	3	2
2	5	6	3
3	4	5	5
4	6	6	6

Test the hypothesis that all three systems are with equal performance at 5% significance level. (10%)

4. Consider the following independent samples. Population 1: $\begin{bmatrix} 6 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \end{bmatrix}, \begin{bmatrix} 7 \\ 3 \end{bmatrix}$; Population 2: $\begin{bmatrix} 8 \\ 3 \end{bmatrix}, \begin{bmatrix} 7 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \end{bmatrix}$; Population 3: $\begin{bmatrix} 4 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. Assume that all three

samples are from multivariate normal distributions with identical population covariance matrices. Are the mean vectors of those three populations the same at 10% significance level? (10%)

5. Calculate the least square estimates $\hat{\boldsymbol{\beta}}_{(1)}$, $\hat{\boldsymbol{\beta}}_{(2)}$ the residuals $\hat{\boldsymbol{\varepsilon}}_{(1)}$, $\hat{\boldsymbol{\varepsilon}}_{(2)}$ and the residual sum of squares for two straight line model $Y_1 = \beta_{01} + \beta_{11} z$, $Y_2 = \beta_{02} + \beta_{12} z$ fit to the data given in the table below. (10%)

z	-2	-1	0	1
y_1	2	1	3	4
y_2	8	6	5	2

6. Use the results of Problem 5 to construct a 90% confidence region for $E([Y_1, Y_2] \mid z = -0.5)$ and a 90% prediction region for a new observation $[Y_1, Y_2]$ when $z = -0.5$. (10%)

7. Establish the equation (7-50) in p. 404 of the text book. (See the hint for Exercise 7.5 in p. 420 of the textbook). (10%)

8. Suppose 200 observations of random variables X_1, X_2, X_3.are with the sample covariance matrix $\mathbf{S} = \begin{bmatrix} 12 & 6 & 0 \\ 6 & 7 & 0 \\ 0 & 0 & 5 \end{bmatrix}$. Find the first sample principal component Y_1, its proportion of total variance, and the 95% confidence interval for the first eigenvalue. (10%)

9. Assume an $m = 1$ orthogonal factor model. Estimate the loading matrix $\tilde{\mathbf{L}}$ and matrix of specific variance $\tilde{\boldsymbol{\Psi}}$ for the covariance matrix of Problem 8 using the principal component solution method. Compute also the residual matrix as well as the factor score at a sample $\mathbf{x} = [2 \quad 3 \quad 1]'$.using the unweighted least square procedure with sample mean $\bar{\mathbf{x}} = [1 \quad 0 \quad -1]'$. (10%)

10. Which part of this course is the most useful/interesting to you? Why? Give your suggestions to improve this course. (10%)