Cairo University

Institute of Statistical Studies and Research

Final Exam: StatM 606

Multivariate Statistical Analysis

Date: Saturday 6-6-2015

time: 3:00 - 6:00

Part I: Compulsory(10 marks)

Choose the correct answer from the following

1) Any linear combination of a multivariate normal random vector ***X*** is distributed according to:

A) multivariate Gamma distribution	B) Wishert distribution
C) multivariate Normal distribution	D) Hotling T^2 distribution

2) The covariance matrix of a multivariate distribution for a p-dimensional random vector have p variances and:

A) (p-1) covariances	B) p(p-1) covariances
C) 1/2 p(p-1) covariances	D) 1/4 p(p-1) covariances

3) The distribution of any subset of the multivariate normal random vector ***X*** is distributed according to:

A) multivariate Gamma distribution	B) Wishert distribution
C) multivariate Normal distribution	D) Fisher distribution

4) The covariance matrix of a multivariate distribution must be:

A) semi positive definite matrix	B) symmetric positive definite matrix
C) singular matrix	D) symmetric matrix

5) A set of random variables that are uncorrelated, are independently distributed iff they are distributed according to:

A) Hotling T^2 distribution	B) bivariate Normal distribution
C) normal distribution	D) Hotling T^2 distribution

Part II: Answer only three questions from the following

Question 1 (20 marks)

A) Consider a family that has four children and let X_1 represents the number of girls in the first two children and X_2 represents the number of girls in the last three children. Find

i) The joint distribution of X_1 and X_2.

ii) The conditional distribution of X_1 given that $X_2 = 2$.

B) Consider the bivariate distribution with joint probability density function

$$f(x_1, x_2) = \begin{cases} 2 & 0 < x_1 < x_1 < 1 \\ 0 & otherwise \end{cases}$$

i) Find the marginal distribution of both X_1 and X_2.

ii) Discuss the independence of X_1 and X_2.

Question 2 (20 marks)

A) Let $\boldsymbol{X}$ be a p-dimensional random vector with mean vector $\boldsymbol{\mu}$, variance-covariance matrix $\boldsymbol{\Sigma}$ and correlation matrix $\boldsymbol{\rho}$ where

$$\boldsymbol{X} = [X_1, X_2, \dots, X_p]', \ \boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_p]', \ \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{bmatrix} \text{ and}$$

$$\boldsymbol{\rho} = \begin{bmatrix} 1 & \cdots & \rho_{1p} \\ \vdots & \ddots & \vdots \\ \rho_{p1} & \cdots & 1 \end{bmatrix}.$$

i) Show that $E[(\boldsymbol{X} - \boldsymbol{\mu})(\boldsymbol{X} - \boldsymbol{\mu})'] = E[\boldsymbol{X}\boldsymbol{X}'] - \boldsymbol{\mu}\boldsymbol{\mu}'$

ii) Let **D** be a diagonal martix whose diagonal elements are the standard deviations of the random vector $\boldsymbol{X}$, show that $\boldsymbol{\Sigma} = \mathbf{D}\,\boldsymbol{\rho}\,\mathbf{D}$

iii) If $\boldsymbol{X} = [X_1, X_2, X_3]$, obtain the matrix **D**, the correlation matrix $\boldsymbol{\rho}$ and find

$$\mathrm{Var}(2X_1 - 5X_2 - X_3) \text{ given } \boldsymbol{\Sigma} = \begin{bmatrix} 25 & 2 & -3 \\ 2 & 4 & 1 \\ -3 & 1 & 9 \end{bmatrix}$$

B) Let **S** be the sample covariance matrix given by

$$\mathbf{S} = \begin{bmatrix} S_{11} & \cdots & S_{1p} \\ \vdots & \ddots & \vdots \\ S_{p1} & \cdots & S_{pp} \end{bmatrix}$$

where $S_{ij} = \dfrac{1}{n-1}\sum_{r=1}^{n}(X_{ri} - \overline{X}_i)(X_{rj} - \overline{X}_j)$, $\overline{X}_i = \dfrac{1}{n}\sum_{r=1}^{n} X_{ri}$ is the sample mean and n is the sample size. Show that the sample covariance matrix **S** is an unbiased estimator of the population covariance matrix $\boldsymbol{\Sigma}$.

Question 3 (20 marks)

A) Let $\boldsymbol{X}$ be a p-dimensional multivariate normal random vector with mean vector $\boldsymbol{\mu}$, and variance-covariance matrix $\boldsymbol{\Sigma}$ where

$$\boldsymbol{X} = [X_1, X_2, \dots, X_p]', \boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_p]', \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{bmatrix} \text{ and}$$

Assume that $\boldsymbol{X}$ is divided into two submatrices $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$, $\boldsymbol{\mu}$ is divided into $\boldsymbol{\mu}^{(1)}$ and $\boldsymbol{\mu}^{(2)}$ and the covariance matrix $\boldsymbol{\Sigma}$ is divided into $\boldsymbol{\Sigma}_{11} = \text{Cov}(\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(1)})$, $\boldsymbol{\Sigma}_{12} = \text{Cov}(\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)})$, $\boldsymbol{\Sigma}_{21} = \text{Cov}(\boldsymbol{X}^{(2)}, \boldsymbol{X}^{(1)})$ and $\boldsymbol{\Sigma}_{22} = \text{Cov}(\boldsymbol{X}^{(2)}, \boldsymbol{X}^{(2)})$. Show that the conditional distribution of $\boldsymbol{X}^{(1)}$ given $\boldsymbol{X}^{(2)} = \boldsymbol{x}^{(2)}$ is distributed according to the multivariate normal distribution with mean vector and covariance matrix given by $\text{E}[\boldsymbol{X}^{(1)}/\boldsymbol{X}^{(2)} = \boldsymbol{x}^{(2)}] = \boldsymbol{\mu}^{(1)} + \boldsymbol{\Sigma}_{12}\,\boldsymbol{\Sigma}_{22}^{-1}\,(\boldsymbol{x}^{(2)} - \boldsymbol{\mu}^{(1)})$, and $\text{Cov}[\boldsymbol{X}^{(1)}/\boldsymbol{X}^{(2)} = \boldsymbol{x}^{(2)}] = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$.

B) Given a 5-dimensional vector $\boldsymbol{X} = [X_1, X_2, \dots, X_5]'$ with mean vector

$$\boldsymbol{\mu} = [1, -2, 4, 0, 3]' \text{ and covariance } \boldsymbol{\Sigma} = \begin{bmatrix} 8 & 2 & 3 & 1 & 0 \\ 2 & 5 & -1 & 3 & 1 \\ 3 & -1 & 6 & -2 & -1 \\ 1 & 3 & -2 & 7 & 2 \\ 0 & 1 & -1 & -2 & 9 \end{bmatrix}.$$

Let $\boldsymbol{X}^{(1)} = [X_1, X_2, X_3]'$, $\boldsymbol{X}^{(2)} = [X_4, X_5]'$ and $\boldsymbol{X} = [\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(2)}]'$.

i) Find $\text{E}(\boldsymbol{X}^{(1)})$ and $\text{E}(\boldsymbol{X}^{(2)})$.

ii) Find $\text{Cov}(\boldsymbol{X}^{(1)}, \mathbf{X}^{(1)})$, $\text{Cov}(\boldsymbol{X}^{(1)}, \mathbf{X}^{(2)})$, $\text{Cov}(\boldsymbol{X}^{(2)}, \mathbf{X}^{(1)})$, and $\text{Cov}(\boldsymbol{X}^{(2)}, \mathbf{X}^{(2)})$.

Question 4 (20 marks)

A) Let $\boldsymbol{X}$ be a p-dimensional random vector distributed according to the multivariate normal distribution with mean vector $\boldsymbol{\mu}$, and variance-covariance matrix $\boldsymbol{\Sigma}$ where

$$\boldsymbol{X} = [X_1, X_2, \ldots, X_p]', \boldsymbol{\mu} = [\mu_1, \mu_2, \ldots, \mu_p]', \text{ and } \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{bmatrix}$$

Let $\boldsymbol{Y} = \mathbf{C}\,\boldsymbol{X}$ where $\mathbf{C}$ is a nonsingular squared matrix ($\mathbf{C} = [c_{ij}]$, i,j=1,2,...,p). Show that $\boldsymbol{Y}$ is distributed according to the multivariate normal distribution with mean vector $\mathbf{C}\,\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}\,\boldsymbol{\Sigma}\,\mathbf{C}'$.

B) Let $\boldsymbol{X} \backsim N_4(\mu, \Sigma)$, where $\boldsymbol{\mu} = [5,\ 5,\ 10,\ 10]'$ and $\boldsymbol{\Sigma} = \begin{bmatrix} 10 & 4 & 4 & 2 \\ 4 & 20 & 10 & 10 \\ 4 & 10 & 20 & 8 \\ 2 & 10 & 8 & 20 \end{bmatrix}$.

Find the marginal distribution of $\boldsymbol{X}^{(1)} = [X_1, X_4]'$ given $\boldsymbol{X}^{(2)} = [X_2, X_3]' = [5,5]'$.

Best Wishes and good luck

Dr. Mohamed Sabry