THE UNIVERSITY OF CHICAGO
Graduate School of Business

Business 424-01, Spring Quarter 1998, Mr. Ruey S. Tsay

Mid-term Exam

Notes:

1. Open book and notes. The exam time is 80 minutes.
2. Write your answers in a bluebook. Mark the solution clearly.

1. Let $\boldsymbol{Z} = (Z_1, Z_2, \cdots, Z_m)'$ be a m-dimensional Gaussian random variable with mean $\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_m)'$ and a positive definite covariance matrix $\boldsymbol{\Sigma} = (\sigma_{ij})$. Consider the following questions or statements. If the statement is true, briefly justify it such as citing a result from the textbook. If the statement is false, please explain why.

   (a) $(\boldsymbol{Z} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{Z} - \boldsymbol{\mu})$ has a (central) chi-square distribution with m degrees of freedom.

   (b) The conditional distribution of Z_1, given $Z_2 = z_2$ is normal with mean $\mu_1 + \frac{\sigma_{12}}{\sigma_{22}}(z_2 - \mu_2)$ and variance $\sigma_{11} + \frac{\sigma_{12}^2}{\sigma_2 2}$.

   (c) Consider the simple linear regression model

   $$Z_1 = \beta_0 + \beta_1 Z_2 + \epsilon.$$

   What is the expectation of the least squares estimate of β_1?

   (d) The conditional distribution of $3Z_1 - 2Z_2$ given $Z_3, \cdots, Z_m$ is normal.

   (e) Suppose that $\boldsymbol{Z}_1, \cdots, \boldsymbol{Z}_n$ are random samples from $\boldsymbol{Z}$. What is the distribution of the sample mean $\bar{\boldsymbol{Z}} = \sum_{i=1}^n \boldsymbol{Z}_i/n$?

2. Consider the multiple linear regression

   $$Y_i = \beta_0 + \beta_1 Z_{1i} + \cdots + \beta_r Z_{ri} + \epsilon_i, \quad i = 1, \cdots, n,$$

   where the $N \times (r+1)$ design matrix $\boldsymbol{Z}$ is of rank $(r+1)$. Assume that ϵ_i are *iid* $N(0, \sigma^2)$. Let $\hat{\beta}_j$ be the least squares estimator of β_j, $\hat{Y}_i = \hat{\beta}_0 - \sum_{j=1}^r \hat{\beta}_j Z_{ji}$ be the fitted value of Y_i, and $e_i = Y_i - \hat{Y}_i$ be the residual.

   (a) For any arbitrary real numbers $b_1, \cdots, b_r$, define $X_i = \sum_{j=1}^r b_j Z_{ji}$. Show that $\sum_{i=1}^n e_i X_i = 0$.

   (b) Show that $\sum_{i=1}^n (Y_i - \bar{Y})^2 = \sum_{i=1}^n (\hat{Y}_i - \bar{\hat{Y}})^2 + \sum_{i=1}^n e_i^2$.

3. Consider the data in Example 7.4 page 393 of the textbook that uses the model

$$Y_j = \beta_0 + \beta_1 Z_{j1} + \beta_2 Z_{j2} + \epsilon_j.$$

We have

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} 30.967 \\ 2.634 \\ 0.045 \end{bmatrix}, \quad (\boldsymbol{Z}'\boldsymbol{Z})^{-1} = \begin{bmatrix} 5.1523 & .2544 & -.1463 \\ .2544 & .0512 & -.0172 \\ -.1463 & -.0172 & .0067 \end{bmatrix}, \quad \hat{\sigma} = 3.4725$$

(a) What is the point prediction given $z_1 = 15.0$ and $z_2 = 60$?

(b) Construct a 95% prediction interval given $z_1 = 15.0$ and $z_2 = 60$. (Note: $t_{17}(0.025) = 2.11$.)

4. A researcher considered two indices measuring the severity of heart attacks. The values of these indices for $n = 40$ heart-attack patients arriving at a hospital emergency room produced the summary statistics

$$\bar{\boldsymbol{x}} = \begin{bmatrix} 46.1 \\ 50.4 \end{bmatrix}, \quad \boldsymbol{S} = \begin{bmatrix} 101.3 & 71.0 \\ 71.0 & 97.4 \end{bmatrix}$$

where $\bar{\boldsymbol{x}}$ and $\boldsymbol{S}$ are the sample mean and covariance matrix, respectively. The two indices are evaluated for each patient. Test for the equality of mean indices using the 5% significance level. Critical values of an F-distribution can be found from the textbook. You should give the test statistic, state the assumptions needed for its reference distribution, and provide your decision on the test.