Indian Statistical Institute
Bangalore Centre
B.Math (Hons.) III Year 2010-2011
First Semester
Statistics III

Mid-semester Examination Date :28.09.11

Answer as many questions as possible. The maximum you can score is 60
All symbols have their usual meaning, unless stated otherwise. State clearly the results you use.

1. Consider the linear model

$$Y(n \times 1) = X(n \times p)\,\beta(p \times 1) + \varepsilon(n \times 1).$$

Here$E(\varepsilon) = 0$ and $Cov(\varepsilon) = \sigma^2 I_n$.

(a) Suppose l is in R^p. When is $l'\beta$ said to be estimable? Obtain the condition on l so that $l'\beta$ is estimable.

(b) Define error space and estimation space and obtain them in terms of the column space of X.

(c) Consider a vector a in the estimation space. Show that $a'Y$ is the BLUE of its expected value.

(d) While working with a linear model with three parameters $\beta_1, \beta_2, \beta_3$, one came came across the system of normal equations $X'X\beta = Z$, where $Z = (3, -2, -1)'$ and

$$X'X = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}.$$

Find two g-inverses G and H of $X'X$. Suppose $\hat{\beta} = GZ$ and $\tilde{\beta} = HZ$. Compute $\hat{\beta}_1 - \hat{\beta}_2$, $\tilde{\beta}_1 - \tilde{\beta}_2$, $\hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2$ and $\tilde{\beta}_0 + \tilde{\beta}_1 + \tilde{\beta}_2$. Explain the fact that the first two numbers are same while the last two are not.

[(1+1) + (2+2+3) + 3 + (2 x 2 + 2 + 3) = 21]

2. Consider the model of Q1. Assume that ε follows multivariate normal distribution.

(a) Derive the distribution of SS_E, the error sum of squares.

(b) Suppose $l'\beta$ is estimable.

(i) Show that $l'\hat{\beta}$ is independent of SS_E.

(ii) Show how you can find a 95% confidence interval for $l'\beta$.

(iii) How do you test the hypothesis $H_0 : l'\beta = \xi$?

(c) Consider a matrix $H(q \times p)$ $(q < p)$ of rank q such that $\rho(H) \subseteq \rho(X)$. Show that

(i) $\Sigma_H = Cov(H'\hat{\beta})$ is positive definite and

(ii) $(H'\hat{\beta})'(\Sigma_H)^{-1}H'\hat{\beta}$ is independent of SS_E.

(d) An experiment was conducted to find out whether all the four ovens in a bakery were operating at the same temperature. The average temperatures obtained and the Mean SS for Error are given below.

Temperatures : $O_1 : 232$, $O_2 : 244$, $O_3 : 229$, $O_4 : 251$.

$MSS_E = 634.3$.

Assuming that the temperature (in ^0C) of each oven is normally distributed find out whether the following pairs of ovens operate, on the average, at the same temperature. (i) O_1 and O_2, (ii) O_1 and O_3, (iii) O_2 and O_4. [Assume error d.f. =5]

[5+ (3+2+2) + (4+3) + 2x3 = 25]

3. (a) When is a vector of random variables said to follow multivariate normal distribution ?

(b) Suppose X follows $N_p(\mu, \Sigma)$, where Σ is p.d. Derive the density function of X.

(c) Suppose X of Q(b) is partitioned as $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$.

(i) Obtain a sufficient condition for X_1 and X_2 to be independent.

(ii) Derive the conditional density of X_1, given X_2.

[1 + 4 + (3+5) = 13]

4. Weighing design : You are asked to find the weights of $p(\geq 2)$ given balls using an ordinary balance. You can take n ($\geq p$) weighings but must use at least 2 balls in each weighings. [You can put each ball in anyone of the two pans].

(a) Assuming the weights of the balls to be $w_1, \cdots w_p$ and the weight you placed in the ith weighing to be $y_i, i = 1, \cdots n$, write an appropriate linear model.

(b) Take $p = 4$ and $n = 5$. Suggest an weighing design. Provide estimates (not necessary BLUE) of as many balls as you can from your design.

[4 + 4 = 8].