Indian Statistical Institute
Bangalore Centre
B.Math (Hons.) III Year 2016-2017
First Semester
Statistics III

Mid-semester Examination Date :13.09.16

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 a random vector $X = (X_1, \cdots X_p)'$.

   (a) Find the 'best predictor' of X_1 among

   (i) **all** functions and (ii) **all linear** functions of

   $X_2, \cdots X_p$.

   (b) Denote 'the best linear predictor' of X_1 obtained in (a) (ii) by $X_{1.2\cdots p}$. Let $R_{1.2\cdots p} = X_1 - X_{1.2\cdots p}$.

   (i) Find variance of $X_{1.2\cdots p}$.

   (ii) Show that $R_{1.2\cdots p}$ is uncorrelated with every $X_j, j = 2, \cdots p$.

   [(3 + 4) + (2 + 3) = 12]

2. (a) When is a random vector $X = (X_1, \cdots X_p)'$ said to follow multivariate normal distribution ?

   (b) Suppose X follows $N_p(\mu, \Sigma)$. Find the distribution of $Y = B(k \times p) + b(k \times 1)$.

   (c) Consider X of Q(b). Partition X as $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ and μ and Σ accordingly.

   (i) Show that X_1 and X_2 are independent if and only if $\Sigma_{12} = 0$.

   (ii) Let $Y = X_1 + MX_2$. Assume that Σ_{22} is p.d. Show that Y is independent of X_2 if and only if $M = -\Sigma_{12}\Sigma_{22}^{-1}$.

   (iii) Assuming that Σ_{22} is p.d, find the conditional distribution of X_1, given $X_2 = t$.

   (d) Consider p random variables $X_1, X_2, \cdots X_p$. Fill in the blank in the following statement with justification.

   "When the joint distribution of $X_1, X_2, \cdots X_p$ is - -, 'the best predictor' of X_1, based on $X_2, \cdots X_p$ coincides with the best predictor among **all linear** functions of $X_2, \cdots X_p$".

   [1 + 3 + (3 + 3 + 5) + 5 = 20]

3. (a) Define generalized (g-)inverse of a matrix.

   (b)For an $m \times n$ matrix A show the following.

   (i) The column space of A is the same as that of AA'.

   (ii) $A'(AA')^-$ is a g-inverse of A. [1 + (3 + 3) =7]

4. 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 in terms of X matrix 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 β_1, β_2 and β_3, one 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 g-inverses G and H by filling up the blanks shown as * and computing constants c_1 and c_2.

$$G = c_1 \begin{bmatrix} 2 & 1 & * \\ 1 & 2 & * \\ * & * & * \end{bmatrix}.\ H = c_2 \begin{bmatrix} * & * & * \\ * & 2 & 1 \\ * & 1 & 2 \end{bmatrix}.$$

Suppose $\hat{\beta} = GZ$ and $\tilde{\beta} = HZ$. Compute $\hat{\beta}_1 - \hat{\beta}_2, \tilde{\beta}_1 - \tilde{\beta}_2, \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3, \tilde{\beta}_1 + \tilde{\beta}_2 + \tilde{\beta}_3$. Explain the fact that the first two numbers are same, but the last two are not.

(e) Consider a matrix $H(q \times p)(q < p)$ of rank q such that $\rho(H) \subset \rho(X)$. Find $Cov(H'\hat{\beta})$ and show that if is positive definite.

[(1+1) + (2+2+3) + 3 + (2 x 2 + 2 + 3) + (2 + 5) = 28]