INDIAN STATISTICAL INSTITUTE, BANGALORE CENTRE
B.MATH - Third Year, 2022-23
Statistics - III, Semesteral Examination, November 23, 2022

Time: $2\frac{1}{2}$ Hours **Total Marks: 50**

1. Suppose $\mathbf{Y} \sim N_p(\mu, \mathbf{\Sigma})$ where $\Sigma = \sigma^2(I_p + \rho \mathbf{1}\mathbf{1}')$, $0 < \rho < 1$.
(a) Show that $\sigma(I_p + \alpha \mathbf{1}\mathbf{1}') = \Sigma^{1/2}$ if $\alpha = (\sqrt{1+p\rho} - 1)/p$.
(b) Find the probability distribution of $\mathbf{Z} = \frac{1}{\sigma}\left(I_p - \frac{\alpha}{1+p\alpha}\mathbf{1}\mathbf{1}'\right)(\mathbf{Y} - \mu)$.
(c) Show that $\mathbf{Z}'\mathbf{Z} \sim \chi^2_r$. Find r.
(d) Find the partial correlation coefficient, $\rho_{13.2}$, between Y_1 and Y_3 given Y_2 $(\mathbf{Y} = (Y_1, Y_2, \ldots, Y_p)')$. [3+4+2+3]

2. Consider the model $\mathbf{Y} = \mathbf{X}\beta + \epsilon$, where $\mathbf{X}_{n\times p}$ has $\mathbf{1}$ as its first column and may not have full column rank; also $\epsilon \sim N_n(\mathbf{0}, \sigma^2 I_n)$. Let $\hat\beta = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}$ and $RSS = (\mathbf{Y} - \mathbf{X}\hat\beta)'(\mathbf{Y} - \mathbf{X}\hat\beta)$, where $(\mathbf{X}'\mathbf{X})^{-}$ is any generalized inverse of $(\mathbf{X}'\mathbf{X})$.
(a) Find the joint distribution of $(\mathbf{1}'\mathbf{Y}, RSS)$.
(b) Suppose $p = 2$. When do we have that $\hat\beta_0$ and $\hat\beta_1$ are independently distributed? [6+6]

3. Consider the one-way model:

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij},\ 1 \le j \le 10;\ 1 \le i \le 4,$$

where ϵ_{ij} are i.i.d. $N(0, \sigma^2)$, with the standard identifiability constraints on α_i.
(a) Show that $\alpha_1 - \alpha_2$ is estimable.
(b) What is the Bonferroni inequality used for multiple comparisons?
(c) Construct a $100(1-\alpha)\%$ simultaneous confidence set for $(\alpha_1 - \alpha_2, \alpha_2 - \alpha_3, \alpha_3 - \alpha_4)$. [3+3+6]

4. Consider testing the hypothesis $H_0 : \beta_1 = \beta_2 = \cdots = \beta_{p-1} = 0$ under the model $\mathbf{Y}_{n\times 1} = X_{n\times p}\beta_{p\times 1} + \epsilon_{n\times 1}$, where $\epsilon \sim N_n(\mathbf{0}, \sigma^2 I_n)$.
(a) Define the F-ratio statistic to test H_0. Find its expected value when H_0 is true, and again when H_0 is false. Explain why the expected value is larger when H_0 is false.
(b) Define the coefficient of determination, R^2. Find its probability distribution when H_0 is true, and show that it is a standard distribution. [9+5]