INDIAN STATISTICAL INSTITUTE, BANGALORE CENTRE
B.MATH - Third Year, 2017-18
Statistics - III, Semesteral Examination, November 20, 2017
Marks shown in square brackets. Time: 3 hours. Total Marks: 50

1. For $n \geq 4$, let $Z_j, 1 \leq j \leq n$ be i.i.d. $N(0, \sigma^2)$, $\sigma^2 > 0$. Define $X_1 = Z_1$, $X_2 = X_1 + Z_2$, $X_3 = X_1 + Z_3$ and $X_i = Z_i$ for $4 \leq i \leq n$.
Let $\mathbf{X} = (X_1, \ldots, X_n)'$.
(a) Find the probability distribution of $\mathbf{X}$.
(b) Find the partial correlation coefficient $\rho_{12.3}$ (between elements of $\mathbf{X}$).
(c) Without computation establish that here $\rho_{12.34} = \rho_{12.3}$. [13]

2. Consider the one-way model:

$$y_{ij} = \mu_i + \epsilon_{ij},\ 1 \leq j \leq n_i;\ 1 \leq i \leq k,$$

where ϵ_{ij} are i.i.d. $N(0, \sigma^2)$, $k \geq 4$ and $n_i > 1$ for all i.
(a) Show that $\mu_1 - \mu_2$ is estimable.
(b) What is the Bonferroni inequality used for multiple comparisons involving μ_i's?
(c) Construct a $100(1-\alpha)\%$ simultaneous confidence set for
$(\mu_1 - \mu_2, \mu_2 - \mu_3, \mu_3 - \mu_4)$. [12]

3. Let Y be a response variable and $X_1, \ldots, X_k$ be covariates. Also, let ρ_i denote the correlation coefficient between Y and X_i, and let R denote the multiple correlation coefficient between Y and $X_1, \ldots, X_k$.
(a) Show that $R \geq \max\{|r_i|, 1 \leq i \leq k\}$.
(b) What is the exact relationship between R and r_i's when $k = 1$? [10]

4. Consider the linear model:

$$y_i = \beta_1 x_i + \beta_2 z_i + \epsilon_i,\ 1 \leq i \leq n,$$

where ϵ_i are uncorrelated errors with mean 0 and common variance σ^2; also x_i's are not proportional to z_i's. Suppose one computes the least squares estimate of β_1 using the incorrect model,

$$y_i = \beta_1 x_i + \epsilon_i,$$

with the same assumptions on ϵ_i as given above. Compute the mean and variance of this estimate of β_1, and compare them with those of the *best linear unbiased* estimate of β_1, under the correct model. [15]