INDIAN STATISTICAL INSTITUTE, BANGALORE CENTRE
B.MATH - Third Year, 2009-10
Statistics - III, Semesteral Examination, December 2, 2009
Marks are shown in square brackets. **Total Marks: 50**

1. 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 $(\mathbf{X}\hat{\beta}, RSS)$.
(b) Define the coefficient of determination and explain what it measures. [10]

2. Consider $Y_1, \ldots, Y_n$ i.i.d. $N(0, \sigma^2)$, $\sigma^2 > 0$, and let $X_i = \sum_{j=1}^{i} Y_j$ for $1 \le i \le n$.
(a) Find the covariance matrix of $\mathbf{X} = (X_1, X_2, \ldots, X_n)'$.
(b) Find the partial correlations $\rho_{12.3}$ and $\rho_{12.34}$ (between components of $\mathbf{X}$). [10]

3. A manufacturer of television sets is interested in the effect on tube conductivity of four different types of coating for colour picture tubes. The following conductivity data were obtained:

Coating type	Conductivity			
1	143	141	150	146
2	152	149	137	143
3	134	136	132	127
4	129	127	132	129

(a) Describe the methodology for determining whether conductivity for the different coating types significantly differ. Numerical computations are not needed.
(b) What is meant by a linear contrast in an experiment like this?
(c) What is the relation between the ANOVA null hypothesis and the hypotheses to check various linear contrasts? [10]

4. Consider the model:

$$y_{ij} = \mu + \alpha_i + \tau_j + \epsilon_{ij},$$

$1 \le i \le 4$, $j = 1, 2$, where ϵ_{ij} are i.i.d. $N(0, \sigma^2)$ and $0 = \tau_1 + \tau_2 = \sum_{i=1}^{4} \alpha_i$.
(a) Show that $\tau_1 - \tau_2$ and $\alpha_k - \alpha_l$, $1 \le k < l \le 4$ are estimable.
(b) Find the best linear unbiased estimators of the above mentioned linear contrasts.
(c) Find the variance of the estimators in (b) above and then provide an unbiased estimator for each of these variances. [10]

5. Describe the theory behind the normal probability plot (q-q plot). Why is it useful in linear regression? [10]