INDIAN STATISTICAL INSTITUTE, BANGALORE CENTRE
B.MATH - Third Year, 2009-10
Statistics - III, Midterm Examination, September 14, 2009

Total Marks: 50   **Time: 3 hours**

1. In the model $\mathbf{Y} = \mathbf{X}\beta + \epsilon$, where $\epsilon \sim N_n(\mathbf{0}, \sigma^2 I_n)$, show that $\mathbf{a}'\beta$ is estimable if and only if $\mathbf{a}'(\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{X} = \mathbf{a}'$.   [**10**]

2. Consider the model $\mathbf{Y} = X\beta + \epsilon$, where $\epsilon \sim N_n(\mathbf{0}, \sigma^2\Gamma)$, where Γ is a known positive definite matrix, not necessarily equal to I_n and $X_{n \times p}$ has full column rank. Least squares is not applied when ϵ_i are correlated or have unequal variance. Let $\Gamma = TT'$ and define $\mathbf{Z} = T^{-1}\mathbf{Y}$.
(a) What is the probability distribution of $\mathbf{Z}$?
(b) Write down a linear model for $\mathbf{Z}$.
(c) What is the least squares estimate of β from this new model?
(d) Find $E(\hat{\beta})$ and $Cov(\hat{\beta})$.   [**20**]

3. Consider the following data on water flow (in 100 Cubic Feet per Second) at Two Points (Libby and Newgate) in January During 1931–43 in Kootenai River. A dam was being planned on the river at Newgate, B.C., where it crossed the Canadian border. The question was how the flow at Newgate could be estimated from that at Libby.

Year	1931	32	33	34	35	36	37	38	39	40	41	42	43
Newgate(y)	19.7	18.0	26.1	44.9	26.1	19.9	15.7	27.6	24.9	23.4	23.1	31.3	23.8
Libby(x)	27.1	20.9	33.4	77.6	37.0	21.6	17.6	35.1	32.6	26.0	27.6	38.7	27.8

The following summary measures are available.

$\bar{x} = 32.53846,$ $\bar{y} = 24.96154,$
$\sum(x_i - \bar{x})^2 = 2693.15,$ $\sum(y_i - \bar{y})^2 = 642.07,$ $\sum(x_i - \bar{x})(y_i - \bar{y}) = 1278.59.$

Fit the linear model $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, ϵ_i i.i.d $N(0, \sigma^2)$ and
(a) give the fitted least squares line;
(b) give an estimate for the error variance σ^2;
(c) provide the ANOVA table;
(d) discuss goodness-of-fit using F-ratio and R^2.   [**20**]