INDIAN STATISTICAL INSTITUTE, BANGALORE CENTRE
B.MATH - Second Year, Second Semester, 2008-09
Statistics - II, Midterm Examination, March 6, 2009

1. Suppose $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ are independent random samples, respectively, from $N(\mu_1, \sigma^2)$ and $N(\mu_2, \sigma^2)$, where $-\infty < \mu_1, \mu_2 < \infty$, $\sigma^2 > 0$.
(a) Find minimal sufficient statistics for (μ_1, μ_2, σ^2). Is it complete?
(b) Find the MLE and UMVUE of σ^2. [10]

2. For observations $Y_1, \ldots, Y_n$, consider the linear model

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, \ldots, n,$$

where x_i is the value of a co-variate corresponding to Y_i and ϵ_i are i.i.d. errors having the $N(0, \sigma^2)$ distribution. Here β_0, β_1 and $\sigma^2 > 0$ are unknown parameters and x_i are treated as known constants.
(a) Show that the distribution of $Y_1, \ldots, Y_n$ belongs to k-variate exponential family. Find k.
(b) Find minimal sufficient statistics for $(\beta_0, \beta_1, \sigma^2)$. Is it complete?
(c) Use properties of exponential family to set up equations and solve them to find MLE of $(\beta_0, \beta_1, \sigma^2)$. [10]

3. Consider a trial which ends up in 'Success' with probability p or 'Failure' with probability $1-p$, $0 < p < 1$. Let X denote the number of independent trials required to obtain the first 'Success'.
(a) Find the probability mass function of X.
(b) Find the Fisher information of p contained in X.
(c) Let $X_1, \ldots, X_n$ be a random sample from the distribution of X. Find the Cramer-Rao lower bound on the variance of an unbiased estimator of $\frac{1}{p}$ based on this random sample. [10]

4. Consider a random sample $X_1, X_2, \ldots, X_n$ from a population with density
$f(x|\theta) = \exp(-(x-\theta))$, for $x > \theta$ where $\theta > 0$ is an unknown parameter.
(a) Find the minimal sufficient statistics T for θ.
(b) Assuming that T is complete, find the UMVUE of θ.
(c) Show that $\sum_{i=1}^n (X_i - \bar{X})^2$ and T are independent random variables. [10]

5.(a) Let S and T be two statistics such that S has finite variance. Show that

$$\mathrm{Var}(S) = \mathrm{Var}(\mathrm{E}(S|T)) + \mathrm{E}(\mathrm{Var}(S|T)).$$

(b) Suppose $(X_1, X_2, \ldots, X_n)$ has probability distribution P_θ, $\theta \in \Theta$. Let $\delta(X_1, X_2, \ldots, X_n)$ be an estimator of θ with finite variance. Suppose that T is sufficient for θ, and let $\delta^*(T)$, defined by $\delta^*(t) = E(\delta(X_1, X_2, \ldots, X_n)|T = t)$, be the conditional expectation of $\delta(X_1, X_2, \ldots, X_n)$ given $T = t$. Then using (a), and without applying Jensen's Inequality, prove that

$$E(\delta^*(T) - \theta)^2 \leq E(\delta(X_1, X_2, \ldots, X_n) - \theta)^2,$$

with strict inequality unless $\delta = \delta^*$, or equivalently, δ is already a function of T. [10]