Indian Statistical Institute, Bangalore
B. Math (II)
First semester 2011-2012
Mid-Semester Examination : Statistics (I)
Date: 21-09-2011 **Maximum Score 70** **Duration: 3 Hours**

1. Blood sugar levels in a random sample of 31 persons were recorded in *mg/dl (milligram per decilitre).* The data are as follows

080	120	110	082	070	143	092	095	095	090	078
090	075	068	220	145	100	135	120	075	165	
121	075	135	071	079	120	145	108	124	077.	

   (a) Make a stem and leaf plot of these data.
   (b) Find the sample mean $\overline{X}$.
   (c) Find the sample standard deviation s.
   (d) Find the sample median M.
   (e) Find $100p$-th percentiles for $p = 0.25$ and 0.75.
   (f) Find the first and third quartiles.
   (g) What proportion of the data lies within $\overline{X} \pm 3s$?
   (h) Draw the box plot and identify the outliers.
   (i) Decide on trimming fraction just enough to eliminate the outliers and obtain the trimmed mean $\overline{X}_{\text{T}}$.
   (j) Also obtain the trimmed standard deviation s_{T}.
   (k) Between the box plot and the stem and leaf plot what do they tell us about the above data set? In very general terms what can you say about the population from which the data arrived?

$[4 + 2 + 2 + 2 + 4 + 2 + 2 + 5 + 3 + 4 + 4 = 34]$

2. A social activist works towards the welfare of petrol bunk workers. She collects donations from the customers arriving at petrol bunks to fuel their cars. She stops as soon as she gets contributions from 20 customers. **Build** a probability model for X, the number of customers she has to request for donations till she gets exactly 20 donors. Find the expected number of customers she would have to request for donations till she gets 20 donors on a randomly chosen day.

$[6 + 4 = 10]$

[PTO]

3. Suppose that in a Poisson process the average rate at which radioactive particles strike a certain target is 2 particles per minute. What is the probability that 5 or more particles will strike the target in a given time interval of 2 minute duration. If there is a strike at time t_1 and T is the waiting time thereafter for the next 10 strikes then find $P[T > t]$. Hence show that the distribution of T is *Gamma* and identify its parameters. Recall that the *probability density function (pdf)* of $Gamma(\alpha, \lambda)$ distribution is given by

$$f(x; \alpha, \lambda) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} e^{-\lambda x} x^{a-1} I_{(0,\infty)}(x); \ \alpha > 0, \lambda > 0; \ -\infty < x < \infty.$$

$[3 + 6 + 3 = 12]$

4. Students of a central university in town B travelled to town A in central India to participate in a cultural *fest*. From there they proceeded to their respective home towns. A typical student had to travel anywhere between 0 to 1000 km to reach home from A. It may be assumed that the distance from A to home town in thousand km is uniformly distributed over $(0, 1)$. Suppose n randomly chosen students had to travel distances $X_1, X_2, \cdots, X_n$ from B to their respective home towns *via* A (in thousand km). Obtain *method of moments* (MOM) *estimator* as well as *maximum likelihood estimator* (MLE) for θ, the actual distance in km between the towns A and B. *Comment* on the estimators.

$[2 + 4 + 4 + 3 = 13]$

5. A production unit produces batteries for cell phones. The efficiency X of a battery is defined suitably, albeit technically, on the scale of 0 to 1. It is found that X has $\beta(a, b)$ distribution $a > 0$, $b > 0$. Recall that the *probability density function (pdf)* of $\beta(a, b)$ distribution is given by

$$f(x; a, b) = \frac{1}{\beta(a, b)} x^{a-1} (1 - x)^{b-1} I_{(0,1)}(x); \ a > 0, b > 0; \ -\infty < x < \infty.$$

For n randomly chosen batteries data are available on $Y_1, Y_2, \cdots, Y_n$ but not on $X_1, X_2, \cdots, X_n$; where $Y_i = \frac{X_i}{1 - X_i}$, $1 \leq i \leq n$. Find *method of moments* (MOM) *estimator* for a and b.

[10]