Indian Statistical Institute, Bangalore
B. Math (II), First semester 2016-2017
Backpaper Examination : Statistics (I)

Date: 29-12-2016 **Maximum Score 50** **Duration: 3 Hours**

1. Following is the data set of daily minimum temperature at a hill station recorded in $^\circ F$ during the month of April.

   77, 80, 82, 68, 65, 59, 61, 57, 50, 62, 61, 70, 69, 64, 67,

   70, 62, 65, 65, 73, 76, 87, 80, 82, 83, 79, 79, 71, 80, 77.

   (a) Make a stem and leaf plot of these data.

   (b) Find the sample mean $\overline{X}$.

   (c) Find $100p$-th percentile for $p = 0.2$ and 0.8.

   (d) Are they equidistant from the median M?

   (e) Draw the box plot and identify the outliers.

   (f) For a given trimming fraction α explain how would you obtain (no computations expected) the trimmed mean $\overline{X}_{\mathrm{T}}$ and the trimmed standard deviation S_{T}.

   (g) Decide on a trimming fraction just enough to eliminate all the outliers in the present case.

   (h) 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? $[3 + 2 + 2 + 2 + 4 + 3 + 1 + 3 = 20]$

2. Let $X_1, X_2, \cdots, X_n$ be a random sample from the distribution with *pdf* given by

$$f(x|\theta_1, \theta_2) = \frac{1}{\theta_2 - \theta_1} I_{(\theta_1, \theta_2)}(x)\,;\ \theta_1 < \theta_2 \in \mathbb{R}$$

   Obtain *method of moments (MOM) estimators and maximum likelihood estimators (MLE)* for θ_1 and θ_2. [10]

3. Let $X_1, X_2, \cdots, X_{2n}$ be a random sample from $Exp(\lambda)$. Let

   $Y_1 = 1(0)$ if $\sum_{i=1}^n X_i \geq \sum_{i=1}^n X_{n+i}\ (\sum_{i=1}^n X_i < \sum_{i=1}^n X_{n+i})$
   $Y_2 = 1(0)$ if $\sum_{i=1}^n X_i \leq \sum_{i=1}^n X_{n+i}\ (\sum_{i=1}^n X_i > \sum_{i=1}^n X_{n+i})$

   Find $\rho_{Y_1 Y_2}$. [8]

4. Let $X_1, X_2, \cdots, X_n$ be a random sample from $N(\mu, \sigma_0^2)$ where σ_0^2 is known. Derive the *likelihood ratio test* (*LRT*) for testing the hypotheses

$$H_0 : \mu = \mu_0 \ versus \ H_1 : \mu \neq \mu_0$$

   at level of significance α. Also obtain the *p value.* [10]

[PTO]

5. Let X be the number of defects in an electronic chip. A random sample of $n = 50$ chips was taken and all 50 chips were inspected. The number of defects in each of them was recorded. The results were as follows:

No of defects	0	1	2	3
Observed count	28	14	6	2

Does the hypothesis of the Poisson distribution as a model for these data seem appropriate? Take level of significance α to be 0.05. Also obtain the p *value.*

[12]