Indian Statistical Institute, Bangalore
B. Math (II)
First semester 2010-2011
Backpaper Examination : Statistics (I)

Date: 31-12-2010 **Maximum Score 80** **Duration: 3 Hours**

1. To establish a standard for parachute design, a researcher recorded the following fill times, in seconds, for 27 standard parachutes, obtained under controlled test conditions.

.59	.38	.47	.43	.44	.37	.43	.37	.27	.54	.39	.89	.48	.52
.51	.49	.38	.38	.23	.44	.40	.36	.33	.82	.51	.44	.37	

   (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 $\overline{X}_{\mathrm{T}}$.
   (j) Also obtain the trimmed standard deviation s_{T}.
   (k) If there are no outliers in these data only then explain how you would compute, trimmed mean $\overline{X}_{\mathrm{T}}$ and trimmed standard deviation s_{T} for given trimming fraction.
   (l) 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?

$[4 + 2 + 2 + 2 + 4 + 2 + 2 + 5 + 2 + 3 + 4 = 32]$

2. Define sample correlation coefficient r for the bivariate data $(x_1, y_1), (x_2, y_2), \cdots,$ (x_n, y_n). Show that r is invariant under the change of origin and scale. Show that $r^2 = 1$ if and only if there exist real numbers a and b, $b \neq 0$, such that $y_i = a + bx_i$; for $1 \leq i \leq n$.

$[2 + 4 + 4 = 10]$

$[PTO]$

3. Let $X_1,\ X_2, ..., X_n$ be a random sample from $Uniform(0, \theta)\,;\, \theta > 0$. Let $X_{(1)},\ X_{(2)}, ...,$ $X_{(n)}$ be the *order statistics*. Let further R be the range, $R = X_{(n)} - X_{(1)}$ and V be the midrange, $V = \frac{X_{(1)}+X_{(n)}}{2}$.

   (a) Obtain a *method of moments (mom)* estimator for the parameter θ.

   (b) Obtain *maximum likelihood estimator (mle)* for the parameter θ.

   (c) Obtain the joint density function of the range $R = X_{(n)} - X_{(1)}$ and midrange $V = \frac{X_{(1)}+X_{(n)}}{2}$.

   (d) Also obtain the marginal density functions of R and V.

$[3 + 5 + 6 + 4 = 18]$

4. Suppose you can draw a random sample from $U \sim uniform[0, 1]$. Explain how you would draw observations on a random variable W that has $F\,(2m, 2n)$ distribution, where m and n are positive integers.

[12]

5. The manufacturer of a certain type of automobile claims that under typical urban driving conditions the automobile will travel at least 20 *km* per litre of petrol. The owner of an automobile of this type notes the mileages that she obtained in her own urban driving conditions when she fills the tank with petrol on 9 different occasions. She finds that the results *km per litre*, on different occasions were as follows :

   $15.6, 18.6, 18.3, 20.1, 21.5, 18.4, 19.1, 20.4, 19.0.$

   (a) List carefully the assumptions you must make and formulate the problem of testing of hypotheses to ascertain the manufacturer's claim.

   (b) Carry out a test at 5% level of significance.

   (c) Find the *p-value* of the test.

   (d) Find 90% *confidence interval* for the expected distance travelled per litre of petrol.

$[3 + 8 + 2 + 3 = 16]$