Indian Statistical Institute, Bangalore
B. Math (II)
First semester 2008-2009
Mid-Semester Examination : Statistics (I)

Date: 23-09-2008 **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) Make a histogram of the frequency distribution.
   (c) Find the sample mean $\overline{x}$.
   (d) Find the sample standard deviation s.
   (e) Find the sample median m.
   (f) Find $100p$ percentiles for $p = 0.25$ and 0.75.
   (g) Find the first and third quartiles.
   (h) What proportion of the data lies within $\overline{x} \pm 3s$?
   (i) Draw the box plot and identify the outliers.
   (j) For the trimming fraction 0.1 obtain the trimmed mean $\overline{x}_T$.
   (k) Decide on trimming fraction just enough to eliminate the outliers and obtain the trimmed mean $\overline{x}_T$.
   (l) Also obtain the trimmed standard deviation s_T.

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

   (m) Between the box plot and the stem and leaf plot what do they tell us about the data set?

[06]

2. The following data were used in an experiment to determine the best way to cut sheets of insulating material so as to retain impact strength. The values are strengths in foot-pounds. Make a schematic diagram of box plots of two sets of specimens in the data and describe any differences in the distributions.

Lengthwise Specimens	1.15	0.84	0.88	0.91	0.86	0.88	0.92	0.87	0.93	0.95
Crosswise Specimens	0.89	0.69	0.46	0.85	0.73	0.67	0.78	0.77	0.80	0.79

$[10+4=14]$

3. Define 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]

4. In a large University a section of students is opposed to 'educational excursion' arguing that the expenditure is wasteful. A random sample of 100 students is drawn. Build a model to obtain the distribution of number of students opposed to 'educational excursion' in the random sample. What would be a good guess for the proportion of students opposed to 'educational excursion' in the entire University? If there were 20 students in the sample opposed to 'educational excursion' what would be the numerical value of your guess?

$[6 + 2 + 2 = 10]$

5. Assume a suitable model for the incoming calls to the EPBX, managed by Smt. Rathnamma. Suppose at time t_0 she needs to take a break and be away from the EPBX. If she decides to take a break for a period of 5 minutes, she would ideally not like to miss more than *two* calls and would be interested in knowing the chance of the same. Can you assist Smt. Rathnamma to find the relevant answer? If there is a call at time t_1 and T is the waiting time for the next call find $P[T > t]$. What is the distribution of T?

$[2 + 4 + 4 + 2 = 12]$