Statistics for Decision Making - I

Full Marks: 40 Time: 2 hrs 30 minutes

Answer 1, 2 and 3 and any two from the rest.

1. i) The rural and urban population (in '0000) of a state over the years is given below. Draw a suitable diagram for the data: (4)

Year	Rural	Urban
1921	13994	2432
1931	14792	2897
1941	17197	4679
1951	20025	6282

ii) The following data gives a frequency distribution of the length of words in a short story. Give suitable diagrammatic representation for a) actual frequency distribution and b) cumulative frequency distribution. (3+3)=6

Length of word	2	3	4	5	6
Frequency	64	43	45	27	21

2. i) Compute the A.M., median and 10% trimmed mean for the following data:

33, 37, 2, 29, 86, 40, 34, 37, 28, 31. Give your comments.
If it is found later that 32 was wrongly recorded as 2, will your comment change? (5)

ii) Draw the boxplot for the corrected data of (5.i). (5)

3. The following scores are given by two judges in a drawing competition. Check how much they agree with each other in their judgement. (10)

Contestant No.	1	2	3	4	5	6	7	8	9	10	11	12
Judge I	33	32	48	40	39	27	27	35	30	20	24	16
Judge II	38	44	38	44	44	30	35	41	32	27	25	27

4. Prove the following:
   i) Mean deviation performs the best when taken about median
   ii) Write a short note on different measures of Skewness. (5+5)
5. i) If u=2x+3, v=3y-7 and the correlation coefficient between x and y is 0.86, what will be the correlation coefficient between u and v? Prove the results you use to get the correlation coefficient here. (4)
   ii) Derive the expression for the standard error in linear regression. How does it depend on the strength of linearity? (4+2)

6. i) Define explained and unexplained variation in the context of linear regression. (3)

ii) The correlation coefficient between two variables is 0.92. How much of the total variability is explained by the corresponding linear regression equation? Prove the result(s) you are using to solve this problem. (5)

iii) Consider the two equations $y = 6.93 + 0.587x$ and $x = 0.95+ 3.3390x$. Can you consider them as regression equations? Justify your answer. (2)