

**EASTERN UNIVERSITY, SRI LANKA**  
**DEPARTMENT OF MATHEMATICS**  
**THIRD EXAMINATION IN SCIENCE -2008/2009**  
**SECOND SEMESTER (Sept./Oct., 2010)**  
**MT 308 - STATISTICS**  
**(PROPER & REPEAT)**

Answer all Questions

Time: Two hours

1. (a) The U.S department of interior releases figures on mineral production. Following are the 10 leading states in nonfuel mineral production in the United States.

States	Values(millions)
California	3350
Nevada	2800
Arisona	2550
Texas	2050
Florida	1900
Michigan	1670
Georgia	1660
Minnesota	1570
Utah	1420
Missouri	1320

- i. Calculate the mean, median and mode.
- ii. Calculate the range, interquartile range, sample variance and sample standard deviation.
- iii. Compute the Pearson's coefficient of skewness for these data.

- (b) The median and mode of the following wage distribution are known to be Rs33.50 and Rs34.00 respectively. Find the missing frequencies in the following table:

Wages(Rs)	0-10	10-20	20-30	30-40	40-50	50-60	60-70
Frequency	4	16	*	*	*	6	4

Total frequency = 230.

2. (a) In an arithmetic test given to a class consisting of  $n$  number of girls and  $n$  number of boys.  $n$  number of boys scored a mean mark of  $m_1$  with a standard deviation of  $s_1$  and the  $n$  number of girls scored a mean mark of  $m_2$  with a standard deviation of  $s_2$ . Show that the standard deviation  $s$  of the marks of all  $2n$  pupils is given by

$$s^2 = \frac{s_1^2 + s_2^2}{2} + \frac{(m_1 - m_2)^2}{4}.$$

- (b) Given the following results relating to two groups containing 20 and 30 observations. Calculate the coefficient of variation of all the 50 observations by combining both the groups.

	Group I	Group II
n	20	30
$\sum X$	45	55
$\sum X^2$	118	132

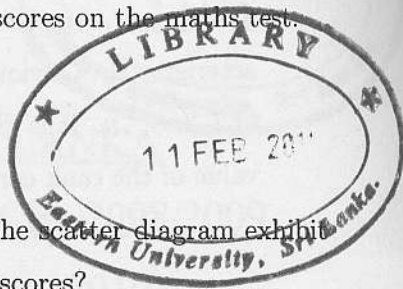
- (c) Number of employees for wages per employee and variance of the wage per employee for two factories are given below:

	Factory A	Factory B
n	50	100
Average (wage/employee)	120	85
Variance (wage/employee)	9	16

In which factory is there a greater variation in the distribution of wage/employee? Suppose that in factory A, the wage of an employee was wrongly entered as 120 instead of 100. What would be the correct variance for factory A?

3. Seven students were tested for stress before a mathematical test. The following table gives the stress scores of these students and their scores on the maths test.

stress score	6.5	4.0	2.5	7.2	8.1	3.4	5.5
test score	81	96	93	70	63	84	73



- (a) Construct a scatter diagram for these data. Does the scatter diagram exhibit a linear relationship between stress scores and test scores?
- (b) Calculate the mean of the stress and the mean of the test scores.
- (c) Without further calculation, draw the line of best fit.
- (d) Find the regression line of test scores on stress scores.
- (e) Construct the analysis of variance and test the hypothesis that the slope of the regression line is zero.
- (f) Predict the test scores of a student with a 7.5 stress score before a maths test, giving standard error for your prediction.
4. (a) Show that *Spearman's* rank correlation coefficients  $r_s$  is given by

$$r_s = 1 - 6 \frac{\sum_{i=1}^n d_i^2}{n(n^2 - 1)},$$

where  $n$  is the number of observation and  $d_i$  is the difference between ranks assigned to the  $i^{th}$  individual.

- (b) The following table shows the data on total costs in million rupees and output in million tons for a company over 10 time period.

Cost	4.39	2.38	2.86	2.77	4.04	3.64	1.93	1.65	3.10	4.66
Output	3.29	1.85	2.29	2.50	3.51	2.73	1.70	1.26	2.68	4.14

- Compute the *pearson's* correlation coefficient.
- Compute the *Spearman's* rank correlation coefficient.
- Comment your results on the basis of these two coefficients.

- (c) Let  $x_1, x_2, \dots, x_n$  be the ranks of  $n$  individuals according to a characteristic  $A$  and  $y_1, y_2, \dots, y_n$  be the ranks of the same individuals according to other characteristic  $B$ . Obviously,  $(x_1, x_2, \dots, x_n)$  and  $(y_1, y_2, \dots, y_n)$  are permutations of  $1, 2, \dots, n$ . It is given that  $x_i + y_i = 1 + n$ ,  $i = 1, 2, \dots, n$ . Show that the value of the rank correlation,  $r$ , between the two characteristic  $A$  and  $B$  is  $-1$ .