UNIVERSITY OF NORTH BENGAL

B.Sc. Honours 1st Semester Examination, 2020

GE1-STATISTICS

Full Marks: 40

ASSIGNMENT

The figures in the margin indicate full marks.
All symbols are of usual significance.

1. Answer the following: $2 \times 4 = 8$

   (a) Define the terms: Frequency, Relative Frequency, Cumulative Frequency (of less than and more than types), and Frequency Density.

   (b) The numbers 3.2, 5.8, 7.9 and 4.5 have frequencies x, $(x+2)$, $(x-3)$ and $(x+6)$ respectively. If the arithmetic mean is 4.876 find the value of x.

   (c) If $y_i = x_i - c$, $(i = 1, 2, ..., n)$ where c is a constant. Prove that $\bar{x} = c + \bar{y}$.

   (d) If the first quartile is 142 and semi-interquartile range is 18, what is the third quartile?

2. Answer any *four* from *six* questions: $8 \times 4 = 32$

   (a) (i) Find the first, the second and the third central moments of the frequency distribution of expenditure (Rs. Per month) given below: 4

Expenditure	3-6	6-9	9-12	12-15	15-18	18-21	21-24	Total
No. of Families	28	292	389	212	59	18	2	1000

        (ii) Fit a straight line to the following data and estimate the most probable yield of rice for 40 inches of water: 4

Water x (inches)	12	18	24	30	36	42	48
Yield y (tons)	5.27	5.68	6.25	7.21	8.02	8.71	8.42

   (b) (i) Prove that the co-efficient of correlation is the geometric mean of the co-efficients of regression.   3

        (ii) Explain why the standard deviation is regarded as superior to the other measures of dispersion. What is its chief defect?   5

   (c) (i) Prove that for a given set of observations the sum of the squares of derivations is the minimum, when deviations are taken from the arithmetic mean.   2

(ii) The mean and S.D of a group of 25 observations were found to be 30 and 3 respectively. After the calculations were made, it was found that two of the observations were incorrect, which were recorded as 29 and 31. Find the mean and S.D if the incorrect observations were excluded.   6

(d) (i) Show that the angle between two regression lines y on x and x on y will be   5

$$\tan^{-1}\left[\frac{s_x s_y (1-r^2)}{(s_x^2 + s_y^2)\,|r|}\right]$$

where s_x, s_y and r be the standard deviation of x, standard deviation of y and correlation coefficient respectively.

(ii) Define Histogram. How is it constructed?   3

(e) (i) Show that the combined standard deviation of two distributions Pooled together is given by the expression: 4

$$NS^2 = n_1 s_1^2 + n_2 s_2^2 + \frac{n_1 n_2}{N}(\bar{x}_1 - \bar{x}_2)^2$$

where the symbols have their usual meanings.

(ii) Deduce Spearman's formula for rank correlation co-efficient?   4

(f) (i) What do you mean by skewness?   1

(ii) Prove that $\beta_2 \geq 1$,   3

                    $\beta_2 - \beta_1 - 1 \geq 0$

(iii) Prove that $AM \geq GM \geq HM$ and state the case when both of them are equal.   4

——×——