

2021

## STATISTICS — HONOURS

First Paper

(Group - A)

Full Marks : 50

*The figures in the margin indicate full marks.**Candidates are required to give their answers in their own words as far as practicable.**Notations and symbols are as usual.*Answer **any four** from **question nos. 1-8** (each carrying 5 marks) and **any two** from **question nos. 9-12** (each carrying 15 marks).

1. How can one classify data according to different scales of measurement. Explain with suitable examples. 5
2. What is a frequency curve? Give the broad categories under which frequency distributions may be put, indicating in each case the nature of the frequency curve. 5
3. When can stem and leaf display be useful? How can one construct such display? Mention one major drawback of such display. 5
4. What is Harmonic Mean? Explain with suitable example when one should use Harmonic Mean and when one should use Arithmetic Mean. 5
5. Suppose  $x$  is a variable with median  $M_i$ . If  $y = g(x)$  be a monotonically increasing or decreasing function of  $x$ , show that the median of  $y$  is  $g(M_i)$ . 5
6. Define correlation ratio. Explain : (i)  $e_{yx}^2 = 0$ , (ii)  $e_{yx}^2 = 1$ . 5
7. Explain the term independence and association for two attributes  $A$  and  $B$  in  $2 \times 2$  case. 5
8. 'Zero correlation does not imply independence.'— Explain. 5
9. (a) Suppose the given values of  $x$  are such that  $a \leq x_i \leq b$  for each  $i$ . Show that (i)  $a \leq \bar{x} \leq b$  and

$$(ii) 0 \leq s^2 \leq \frac{(b-a)^2}{4}.$$

Please Turn Over

- (b) Let  $x$  be a variable assuming the values  $i = 1, 2, \dots, k$  with frequencies  $f_i$  and let  $F_i'$  be the corresponding cumulative frequencies of the 'greater than' type, while  $F_i''$  are the cumulative totals of the 'greater than' type of these cumulative frequencies. If  $n$  be the total frequency and

$$T_1 = \frac{1}{n} \sum_{i=1}^k F_i', \quad T_2 = \frac{1}{n} \sum_{i=1}^k F_i''$$

show that  $S^2 = 2T_2 - T_1 - T_1^2$ . (1+6)+8

- 10.** (a) Consider any symmetrical frequency distribution for a discrete variable. Show that its central moments of odd orders must all be zero.
- (b) Show that (i)  $b_2 \geq 1$  and (ii)  $b_2 - b_1 - 1 \geq 0$ , where  $b_2$  and  $b_1$  have their usual meaning. Discuss, in detail, when  $b_2 = 1$  and  $b_2 - b_1 - 1 = 0$ .
- (c) Discuss Pearson's measure of skewness and show that Pearson's second measure of skewness lies between  $-3$  and  $3$ . 4+(2+4)+5
- 11.** (a) For the variables  $x$  and  $y$  such that  $s_x > 0$  and  $s_y > 0$ , show that  $r = \{\text{var}(x + y) - \text{var}(x - y)\} / 4s_x \cdot s_y$ .
- (b) Show that the correlation ratio  $e_{yx}$  is the simple correlation coefficient between  $y$  and the array mean of  $y$  corresponding to  $x$ .
- (c) Prove that,  $\text{corr}(y, Y) = |\text{corr}(y, x)|$ . Discuss  $r^2$  is called the coefficient of determination. 5+5+5
- 12.** (a) Discuss how Kendall's rank correlation coefficient had been developed. Also indicate how the formula can be adapted to the case of tied ranks.
- (b) Discuss in detail— Measures of association between two attributes  $A$  and  $B$  in  $2 \times 2$  case.
- (c) How can one estimate parameters while fitting a logit model? 7+5+3
-