

U.G. 1st Semester Examination - 2021**STATISTICS****[HONOURS]****Generic Elective Course (GE)****Course Code : STAT-H-GE-T-1****(Statistical Methods)**Full Marks : 50 (40+10) Time : $2\frac{1}{2}$ Hours*The figures in the right-hand margin indicate marks.**Candidates are required to give their answers in their own words as far as practicable.**Notations and symbols have their usual meaning.*

1. Answer any **five** questions: $2 \times 5 = 10$
- What are frequency type and non-frequency type data?
 - What is divided bar diagram?
 - Define Geometric mean and Harmonic mean.
 - How do you compare the variability of two sets of data with different means?
 - Distinguish between nominal and ordinal data.
 - For three variables, how do you interpret the partial correlation coefficient $r_{12.3}$?

[Turn over]

- What's the difference between correlation analysis and regression analysis?
 - When are two attributes said to be positively associated?
2. Answer any **two** questions: $5 \times 2 = 10$
- What is frequency polygon? How is it drawn? How can you obtain a frequency curve from such a polygon?
 - Compare standard deviation and mean deviation about mean as measures of dispersion.
 - For three variables, give a formula for the multiple correlation $r_{1.23}$ and discuss its properties.
 - Write down Yule's Coefficient of Association and discuss its Range. State its limitations.
3. Answer any **two** questions: $10 \times 2 = 20$
- Define Arithmetic Mean (AM), Geometric Mean (GM) and Harmonic Mean (HM). Show that $AM \geq GM \geq HM$.
 - In a frequency table, the upper boundary of each class interval has a constant ratio to the lower boundary. Show that the geometric

mean G may be expressed by the formula:

$$\log G = x_0 + \frac{c}{N} \sum_i f_i (i-1)$$

where x_0 is the logarithm of the mid-value of the first interval and c is the logarithm of the ratio between upper and lower boundaries.

- c) What is regression? Derive the least-square linear regression of y on x based on n paired observations on (x, y) . Also find the angle between the two regression lines.
- d) What is rank correlation? If d_i be the difference in the ranks of the i^{th} individual in two different characteristics, show that the maximum value of $\sum_{i=1}^n d_i^2$ is $\frac{1}{3}(n^3 - n)$. Hence or otherwise, show that rank correlation coefficient lies between -1 and +1.

[Internal Assessment : 10]
