# 2021

## STATISTICS — HONOURS

**Paper : CC-1**

**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 of usual significance)***

1. Answer ***any five*** questions : 2×5

(a) Distinguish between cross-sectional data and time series data.

(b) When should one recommend the use of multiple bar diagram? Explain with an example.

(c) How should an average change when all values of the variables are increased or decreased in the same proportion?

(d) Point out one merit and one limitation of the range as a measure of dispersion.

(e) Two independent variables $x$ and $y$ have variances 3 and 4 respectively. Find the correlation coefficient between $(x+y)$ and $(x-y)$.

(f) Find the expression for the proportion of variation explained by the regression equation of $y$ on $x$ with respect to the total variation in the data.

(g) Define 'odds' and 'odds ratio' in a 2×2 contingency table.

(h) Write down the minimum and maximum values of the intra-class correlation coefficient when the covariate $x$ is observed in $k$ families, each consisting of $p$ members.

2. Answer ***any two*** questions : 5×2

(a) Why is a box plot said to be a five-measure diagram? How do you comment on skewness from a box plot?

(b) If the two regression equations of $y$ on $x$ and $x$ on $y$ are $a_1x + b_1y + c_1 = 0$ and $a_2x + b_2y + c_2 = 0$ respectively, show that $a_1b_2 \le a_2b_1$.

(c) Suppose that the variable $x$ takes positive values only and that the deviations $x_i - \bar{x}$ are small compared to $\bar{x}$. Show that $GM \simeq \bar{x}\left(1 - \frac{s^2}{2\bar{x}^2}\right)$.

3. Answer ***any three*** questions :

(a) (i) For a set of positive numbers, can the standard deviation be larger than mean? Can the IQR be higher than the median? Give an example to support your answer for each of the cases.

(ii) (A) Obtain the minimum value (if possible) of range of a variable $x$, if its s.d. is 4.

(B) Show that s.d. $\geq 5$ if the mean is 14, median is 9. 6+4

(b) Using Cauchy-Schwartz inequality or otherwise, prove that $b_2 - b_1 - 1 \geq 0$. Also, discuss the case where $b_2 - b_1 - 1 = 0$. Interpret the following cases :

(i) $b_1 = 1$ (ii) $b_2 = 0$ (iii) $b_1 = 1.5$ (iv) $b_2 = 3 + 1.5b_1$ 4+2+4

(c) (i) When are two attributes said to be independent?

(ii) What are the various methods of finding whether two attributes are associated? 2+8

(d) What is correlation ratio? Show that $0 \leq r^2 \leq e_{yx}^2 \leq 1$. Discuss the following cases :

(i) $e_{yx}^2 = r^2$, (ii) $e_{yx}^2 = e_{xy}^2 = r^2$, (iii) $e_{yx}^2 = 1$, (iv) $e_{yx}^2 = e_{xy}^2 = 1$. 2+4+4

(e) Discuss the rationale behind Kendall's $\tau$ coefficient for rank correlation. Also indicate how the formula can be adapted to the case of tied ranks. Show that Kendall's $\tau$ will lie between $-1$ and $+1$. Interpret the marginal cases. 2+3+3+2