

1

G.P. codes 59199

Paper I

SOLUTION

Q.1]	(a)	Correct the following if necessary:	(10)
	(1)	Population is a finite subset of sample. True	(02)
	(2)	Data from a news paper is a primary data. True	(02)
	(3)	Temperature of a day represents a discrete random variable. False continuous variable.	(02)
	(4)	Kurtosis measures departure from symmetry. False Skewness measures departure from symmetry or Kurtosis measures the degree of peakedness of the curve about mode.	(02)
	(5)	Median is affected by extreme values. False Median is not affected by extreme values or mean is affected by extreme values.	(02)
Q.1]	(b)	Answer in one sentence:	(10)
	(1)	Distinguish between qualitative and quantitative data. Non-measurable and measurable.	(02)
	(2)	Explain briefly SRSWOR. Selection of units without replacing the selected unit. Eg. Distinct ranks	(02)
	(3)	State the formula for mode for grouped data, explaining the notations. Mode = $l_1 + \frac{(l_1 - l_2)(f_1 - f_0)}{(f_1 - f_0) + (f_1 - f_2)}$ where, l_1 : lower class-boundary of modal class, l_2 : upper class-boundary of modal class, f_0 : frequency of pre-modal class, f_1 : frequency of modal class, f_2 : frequency of post-modal class	(02)
	(4)	Explain the term negatively skewed distribution. If the density of observations is more for higher values of the variable than lower values, then the tail of the frequency distribution is elongated towards left side. Such a distribution is called negatively skewed distribution.	(02)
	(5)	Define r^{th} central moment for ungrouped and grouped data. $\mu_r = \frac{\sum (x - \bar{x})^r}{n}$ or $\frac{\sum f(x - \bar{x})^r}{N}$ where $N = \sum f$.	(02)
Q.2]		Attempt any TWO sub-questions:	(20)
	(1)	What is a Questionnaire? What are the requisites of a good Questionnaire? Questionnaire is a set of questions arranged in a logical order and systematic manner. (2) Requisites: 1. Covering letter explaining the importance of survey, 2. Easy and simple questions, 3. Arranged in a logical and systematic manner. 4. Multiple choice questions as far as possible, 5. Not too many or too less questions, 6. Related questions, 7. Unless required, no personal questions, 8. Questions to cross check answers, 9. No mathematical calculations unless required. (3)	(10)
	(2)	Explain the term scale of measurement. Write short note on (i) ordinal scale and (ii) Ratio scale of measurement.	(10)

(...2/-)

	<p>Scale of measurement is generally considered in terms of mathematical properties of characteristic. (2)</p> <p>(i) Ordinal scale: classifies observations on a characteristic into different classes. Allows to rank the observations but difference between adjacent ranks may not be equal. Eg. Socio-economic status of families. (4)</p> <p>(ii) Ratio scale: identifiable absolute zero point. One number can be justifiable as multiple of another. Differences and ratios are meaningful. Eg. Measures on physical dimensions such as height, weight etc. (4)</p>																	
(3)	<p>(i) State conditions for two attributes A and B to be completely associated and completely dissociated. State the type of association between attributes A and B for the following data: $(\alpha\beta) = 120, (\alpha) = 300, (\beta) = 200$ and $(B) = 300$. Completely associated: A always occurs with β and never with β (i.e. $(AB) = (A)$ or $(A\beta) = 0$) or B always occurs with A and never with α (i.e. $(AB) = (B)$ or $(\alpha B) = 0$). Completely dissociated: A never occurs with B (i.e. $(AB) = 0$) or α never occurs with β (i.e. $(\alpha\beta) = 0$). (2)</p> <table border="1" data-bbox="502 929 1029 1075"> <thead> <tr> <th></th> <th>B</th> <th>β</th> <th>Total</th> </tr> </thead> <tbody> <tr> <th>A</th> <td>120</td> <td>80</td> <td>200</td> </tr> <tr> <th>α</th> <td>180</td> <td>120</td> <td>300</td> </tr> <tr> <th>Total</th> <td>300</td> <td>200</td> <td>500</td> </tr> </tbody> </table> <p>$(A)(B)/N = (200*300)/500 = 120 = (AB)$ (2) A and B are independent. (1)</p> <p>(ii) In usual notations, prove that $Q = 2Y/(1+Y^2)$. (05)</p> $Q = \frac{(AB)(\alpha\beta) - (A\beta)(\alpha B)}{(AB)(\alpha\beta) + (A\beta)(\alpha B)}, Y = \frac{\sqrt{(AB)(\alpha\beta)} - \sqrt{(A\beta)(\alpha B)}}{\sqrt{(AB)(\alpha\beta)} + \sqrt{(A\beta)(\alpha B)}} \quad (2)$ <p><u>Proof:</u></p> $Y = \frac{1 - \frac{\sqrt{(A\beta)(\alpha B)}}{\sqrt{(AB)(\alpha\beta)}}}{1 + \frac{\sqrt{(A\beta)(\alpha B)}}{\sqrt{(AB)(\alpha\beta)}}}, K = \frac{(A\beta)(\alpha B)}{(AB)(\alpha\beta)}$ $= \frac{1 - \sqrt{K}}{1 + \sqrt{K}} \quad \therefore Y^2 = \frac{1 - 2\sqrt{K} + K}{1 + 2\sqrt{K} + K}$ $1 + Y^2 = \frac{2(1+K)}{(1+\sqrt{K})^2}, \quad 2Y = 2 \frac{1 - \sqrt{K}}{1 + \sqrt{K}}$ $\therefore \frac{2Y}{1+Y^2} = \frac{(1-\sqrt{K})(1+\sqrt{K})}{(1+K)} = \frac{1-K}{1+K} = Q. \quad (3)$		B	β	Total	A	120	80	200	α	180	120	300	Total	300	200	500	(05)
	B	β	Total															
A	120	80	200															
α	180	120	300															
Total	300	200	500															
(4)	(i) What is tabulation? What are its objective?	(05)																

	<p>Tabulation: a logical and systematic arrangement of statistical data in rows and columns. (2)</p> <p>Objectives: 1. To simplify complex data, 2. To clarify aim of survey, 3. To make comparisons, 4. To facilitate statistical analysis, 5. To facilitate future study, 6. To identify trend if any in data. (3)</p>	
	<p>(ii) Describe various parts of a statistical table. Title, table number, stubs and captions, body, source note, foot note.</p>	(05)
Q.3]	Attempt any TWO sub-questions:	(20)
(1)	<p>(A) Explain weighted and combined Arithmetic mean. $\bar{x}_w = \frac{\sum wx}{\sum w}$ and $\bar{x} = \frac{\sum nx}{\sum n}$ (2) + (3)</p> <p>(B) State merits and demerits of mode. Merits: 1. Easy to calculate and understand, 2. Not affected by extreme values, 3. Present in the data, 4. Can be calculated for open end class data, 5. Can be applied for both quantitative and qualitative data, 6. Can be located graphically. (3)</p> <p>Demerits: 1. Not rigidly defined, 2. Not based on all data, 3. not capable of further mathematical treatment, 4. It is insignificant for small data. (2)</p>	(05) (05)
(2)	<p>Write short note on (i) Histogram, (ii) Frequency polygon. (i) Histogram: adjacent bars, height = frequency and width = class width. Unequal class width, frequency density = freq/ width is considered. (6)</p> <p>(ii) Frequency polygon: X- axis- calss mark, Y-axis- frequency. Join points by straight line, extend end points to X-axis. (4)</p>	(10)
(3)	<p>(A) Write a note on Box and Whisker plot. Give procedure of construction.</p> <p>(B) Draw less than ogive for the following data: Weight (in Kgs.): 40-45 45-50 50-55 55-60 60-65 65-70 70-75 No. of Persons : 5 10 18 26 22 15 4 Hence, find median. <cf: 5 15 33 59 81 96 100 (1)</p> <p>Points plotted (UCB, <cf), upper class boundary, median. (2) + (2)</p>	(05) (05)
(4)	<p>For the following measures of central tendency, state the formula in each case by explaining clearly all the notations used:</p> <p>(i) Lower quartile, i^{th} decile and i^{th} percentile for grouped frequency distribution, $Q_1 = l_1 + \frac{(l_2 - l_1)(\frac{N}{4} - cf)}{f}$, $D_i = l_1 + \frac{(l_2 - l_1)(\frac{iN}{10} - cf)}{f}$</p> <p>(ii) Geometric mean and Harmonic mean for raw data. $P_1 = l_1 + \frac{(l_2 - l_1)(\frac{iN}{100} - cf)}{f}$, $G.M. = \sqrt[n]{x_1 x_2 \dots x_n}$ $H.M. = \frac{n}{\sum (\frac{1}{x})}$ (2) each</p>	(10)

4

4

Q.4]	Attempt any TWO sub-questions:																	
(1)	<p>Explain the concept of kurtosis. Discuss its various types. State various measures of kurtosis and their interpretations.</p> <p>Kurtosis: Measure of degree of peakedness. Platykurtic. Mesokurtic and Leptokurtic.</p> <p>Measures : measures based on moments β_2 and γ_2. Interpretations.</p>	(20) (10) (2) + (3) (5)																
(2)	<p>Define standard deviation and variance. Prove that standard deviation is invariant under shift of origin, but is affected by change in scale.</p> <p>S.D.: positive square root of a.m. of squares of the deviations of the observations from their a.m.</p> <p>Variance = (s.d.)²</p> <p>Give proof.</p>	(10) (2) (2) (6)																
(3)	<p>Define r^{th} raw moment about zero, μ'_r. Derive expressions for first four central moments in terms of raw moments about zero.</p> <p>$\mu'_r = \frac{\sum f x^r}{\sum f}$. $\mu_1 = 0$, $\mu_2 = \mu_2' - \mu_1'^2$, $\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3$</p> <p>$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4$ (1) + (1) + (2) + (2) + (2)</p>	(10)																
(4)	<p>(A) Define mean deviation from mean and quartile deviation for grouped data.</p> <p>M.D. from mean = $\frac{\sum f x - \bar{x} }{\sum f}$, Q.D. = $\frac{Q_3 - Q_1}{2}$ (2) + (3)</p> <p>(B) State relative measures of skewness.</p> <p>Karl Pearson's and Bowley's coefficient of skewness, measure based on moments</p>	(05) (05)																
Q.5]	Attempt any FOUR sub-questions:																	
(1)	<p>What is time series data and cross sectional data? Explain with illustration.</p> <p>Time series is chronological data. Cross sectional data is based on a number of characteristics over a time period. Example of each.</p>	(20) (05)																
(2)	<p>Write a note on Schedule.</p> <p>Schedule is a questionnaire to be filled by trained enumerators who collect information from respondents directly.</p>	(05)																
(3)	<p>If $(a\beta) = 165$, $(\alpha) = 510$, $(\beta) = 250$ and $(B) = 480$, find Yule's coefficient of association and comment.</p> <table border="1" data-bbox="287 1668 550 1792"> <tr> <td></td> <td>β</td> <td>B</td> <td>T</td> </tr> <tr> <td>A</td> <td>135</td> <td>85</td> <td>220</td> </tr> <tr> <td>a</td> <td>345</td> <td>165</td> <td>510</td> </tr> <tr> <td>T</td> <td>480</td> <td>250</td> <td>730</td> </tr> </table> <p>$Q = \frac{(AB)(\alpha\beta) - (A\beta)(\alpha B)}{(AB)(\alpha\beta) + (A\beta)(\alpha B)} = \frac{-7050}{57600} = -0.1224$ (2) + (1)</p> <p>-ve association</p>		β	B	T	A	135	85	220	a	345	165	510	T	480	250	730	(05)
	β	B	T															
A	135	85	220															
a	345	165	510															
T	480	250	730															
(4)	<p>Draw stem and leaf diagram for the following data:</p> <p>52, 61, 71, 86, 57, 63, 48, 70, 69, 62, 47, 45, 61, 77, 52, 62, 44, 58, 56, 40, 82, 61, 51, 50, 49, 75, 72, 80.</p>	(05)																

5

5

		4 0 4 5 7 8 9 5 0 1 2 2 6 7 8 6 1 1 1 2 2 3 9 7 0 1 2 5 7 8 0 2 6	
(5)	For the following distribution mode is 65. Find the value of 'a': Class-interval: 20-40 40-60 60-80 80-100 100-120 Frequency : 5 18 a 10 3	$\text{Mode} = l_1 + \frac{(l_2 - l_1)(f_1 - f_0)}{(f_1 - f_0) + (f_1 - f_2)} = 60 + \frac{20(a - 18)}{(a - 18) + (a - 10)} = 65 \Rightarrow a = 22$	(05)
(6)	Explain the concept of measures of dispersion. Measure of variation. Ex.		(05)
(7)	If the values of second and fourth central moments are 19.42 and 837.31 respectively, find β_2 and comment on the answer.	$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{837.31}{(19.42)^2} = 2.22 < 3. \text{ Dist is platykurtic.}$	(05)
