ID : 14103

SUBJECT : STATISTICS

INSTRUCTOR : SIR RAZA

DEPARTMENT : MSMC (6^{TH} SEMESTER)

FINAL TERM PAPER

Saturday, June 27

Question No: 01

Find averages (A.M, G.M, H.M) of the following table (s) also justify their logical relationships.

a.

Number of children per family	Number of families
1	4
2	13
3	9
4	4
5	1

b.

marks	frequency
0 – 9	2
10 – 19	31
20 – 29	73
30 – 39	85
40 – 49	28

Solution:

Q1 (a)

Number of children per family x	Number of families f	$f.x$	$f.\log x$	f/x
1	4	4	$4\times\log 1 = 0$	$4/1 = 4$
2	13	26	$13\times\log 2 = 3.913$	$13/2 = 6.5$
3	9	27	$9\times\log 3 = 4.294$	$9/13 = 3$
4	4	16	$4\times\log 4 = 2.408$	$4/4 = 1$
5	1	5	$1\times\log 5 = 0.699$	$1/5 = 0.2$
	$\Sigma f = 31$	$\Sigma fx = 74$	$\Sigma f\log x = 11.314$	$\Sigma f/x = 14.7$

Arithematic Mean

$$A.M = \frac{\Sigma f x}{\Sigma f} = \frac{78}{31}$$

A.M = 2.[redacted]516

Geometric Mean

$$G.M = \text{Antilog}\left(\frac{\Sigma(f\log x)}{\Sigma f}\right)$$

$$G.M = \text{Antilog}\left(\frac{11.314}{31}\right)$$

$$G.M = \text{Antilog}\ (0.365)$$

$$G.M = 2.317$$

HORMONIC MEAN

$$H.M = \frac{\sum f}{\sum \left(f/x \right)}$$

$$H.M = \frac{31}{14.7}$$

$$H.M = 2.1088$$

Logic relationship of A.M, G.M, H.M

$$A.M \geq G.M \geq H.M$$

$$2.516 > 2.317 > 2.1088$$

logical Relationship of A.M, G.M & H.M Justified.

(b)

Class Interval	Frequency (F)	mid Point	f.m	f log m	F/m
0-9	2	4.5	2x4.5=9	2 log 4.5 = 1.306	2/4.5 = 0.444
10-19	31	14.5	31x14.5=449	31 log 14.5 = 36.002	31/14 = 2.138
20-29	73	24.5	73x24.5=1788.5	73 log 24.5 = 130.715	73/24.5 = 298
30-39	85	35.5	85x34.5 = 29325	85 log 34.5 = 46.154	85/34.5 = 2.464
40-49	28	44.5	28x44.5 = 1246	28 log 44.5 = 46.154	$\Sigma(F/m) = 8.655$
	$\Sigma F = 219$		$\Sigma f.m = 6425$	$\Sigma(F.\log m) = 315.586$	

Arthmatic Mean:

$$Am = \frac{\Sigma(fm)}{\Sigma F}$$

From the above data $\Sigma(fm) = \frac{6425}{219}$

A.M = 29.34

Gematric Mean

$$G.M = \text{Antilog}\left(\frac{\Sigma(f \times \log m)}{\Sigma f}\right)$$

$$G.M = \text{Anti log}\left(\frac{315.586}{219}\right)$$

$$G.M = \text{Anti log}\,(1.441)$$

$$G.M = 27.607$$

Hormonic Mean

$$H.M = \frac{\Sigma f}{\Sigma(f/m)}$$

$$H.M = \frac{219}{8.655}$$

$$H.M = \frac{219}{8.655}$$

$$= 25.303$$

logical relationship of A.M. GM, H.M

$$AM \geq GM \geq HM$$

Question No: 02

Find Median & Mode of the following tables

a.

Number of children per family	Number of families
1	4
2	13
3	9
4	4
5	1

b.

marks	frequency
0 – 9	2
10 – 19	31
20 – 29	73
30 – 39	85
40 – 49	28

Solution:

(5)

Q.2

(a)

x	f	c.f
1	4	4
2	13	17
3	9	26
4	4	30
5	1	31

$\Sigma f = 31$

$n = \Sigma f = 31$

n is odd.

Median is the central value

Position of central value $= \frac{n+1}{2} = \frac{31+1}{2}$

$= 16^{th}$ term.

From the tabel above it is clear that 16^{th} term is 2. Therfore,

[median = 2]

Mode:

Mode is the most frequent from The table. Thefore, it is clear that the highest frequency of 13 is 2.

Therfore;

Mode = 2

Q.2 (b)

Class Interval	F	Class Boundries	C.F	Mid-Point
0 - 9	2	-0.5 - 9.5	2	4.5
10 - 19	31	9.5 - 19.5	33	14.5
20 - 29	73 f_0	19.5 - 29.5	106	24.5
30 - 39	85 f_1	29.5 - 39.5	191	34.5
40 - 49	28 f_2	39.5 - 49.5	219	44.5

Median & mode lies there.

Median

$\left(\frac{n+1}{2}\right)^{th}$ term

$\left(\frac{219+1}{2}\right)^{th}$ term

$\left(\frac{220}{2}\right)^{th}$ term

110 th term

OR

$\frac{n}{2}$ th term

$\frac{219}{2}$ th term

109.5^{th} term

$L = 29.5$, $h = 10$

$f = 85$, $\Sigma f = 219$

$C = 106$

$$\text{Median} = L + \frac{h}{f}\left(\frac{\Sigma f}{2} - C\right)$$

$$\text{Median} = 29.5 + \frac{10}{85}\left(\frac{219}{2} - 106\right)$$

$$\text{Median} = 29.5 + 0.1176\,(35)$$

$$\text{Median} = 29.5 + 0.4118$$

$$\text{Median} \approx 30$$

$L_1 = 29.5$ $\qquad f_1 = 85$

$i = 10$ $\qquad f_2 = 28$

$f_0 = 73$

$$\text{Mode} = L_1 + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right) i$$

$$\text{Mode} = 29.5 + \left(\frac{85 - 73}{2(85 - 73 - 28)}\right) \times 10$$

$$\text{Mode} = 29 + \left(\frac{12}{69} \times 10\right)$$

$$\text{Mode} = 29 + 1.739$$

$$\text{Mode} = 30.74$$

Question No: 03

a. Find Semi Quartile Range & Semi Inter Quartile Range of Q2(a)
b. Find Variance and Co-efficient of variance of Q2(a)

Solution :

Q-3 (a)

Finding Semi Quratile Range of Q.2
& Semi Inter Quartile Range of Q.2

n	f	cf
1	4	4
2	13	17
3	9	26
4	4	30
5	1	31

$\Sigma f = 31$

Semi Q. Formula

$Q.D = \frac{Q_3 - Q_1}{2}$

Formula for Quartile Deviation

Formula for Inter Quratile diviation Range $= Q_3 - Q_1$

First Finding the quratile values

Finding Q_1 value

Formula,

$$Q_1 = \left(\frac{N+1}{4}\right)$$

$$= \frac{31+1}{4} = \frac{32}{4} = 8$$

$Q_1 = 8$, the value of 8 exists in the second class i-e 2.

So, the Quratile $1 = \boxed{2}$

Finding Q.3 Value,

$$Q_3 = 3\left(\frac{N+1}{4}\right) = 3\left(\frac{32}{4}\right) = 24^{th} \text{ value}$$

If we check the value of 24 in the cf table, the location tracks to

$Q_3 = 3$

Now Finding the semi Quartile range of Q_2.

Formula, $\frac{Q_3 - Q_1}{2} = \frac{3-1}{2} = \frac{\cancel{2}}{\cancel{2}}\,1$

The semi, Inter Quartile R. of Q_2
is, $\boxed{1}$

Finding semi Quaratile range of Q_2
calculated as,

$$\frac{Q_3 + Q_1}{2}$$

Mid or S.Q. Range $= \frac{3+1}{2} = \frac{\cancel{4}}{\cancel{2}} = 2$

The Interquratile range means
where the 50%. the data is
located in the table. i.e class (2)

Q-3 (b)

Finding the varience and coefficient of the varience

x	F
1	4
2	13
3	9
4	4
5	1

⇒ Meaure of dispersion

To find the variution First we 've to calculate the general value.

1- Range = Max − Min
Hence, the maximum value is 5 and the manimum is 1

$R = 5 - 1 \Rightarrow 4$

Calculating the varience so the data given

$$\text{St. Deviation} - \sigma = \Sigma f \left(\frac{x_i - x}{\Sigma F} \right)^2$$

$$\sigma = \sqrt{\frac{\Sigma F (x_i - \bar{x})^2}{\Sigma F}}$$

n	f	fx	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f(x_i - x)^2$
1	14	14	1-2.51	2.2801	14×2.2801 =31.921413×
2	13	26	2-2.51	0.2601	3.3813 9×
3	9	27	3-2.51	0.2401	2.1609 4×
4	4	16	4-2.51	2.2201	1× 8.88.4
5	1	5	5-2.51	6.2001	6.2001
	$\Sigma f = 31$	$\Sigma fx = 88$			$\Sigma f =$ 52.544

$$X_2 = \frac{\Sigma F n}{\Sigma F}$$

$$= \frac{88}{31} \Rightarrow 2.83 \qquad \delta = \frac{\Sigma f (x - \bar{x})^2}{\Sigma F}$$

Putting the values . $\delta = \sqrt{\dfrac{52.544}{31}}$

Variance = 1.694

To find the efficient we have to find the standard devision first.

$$S.D = \sqrt{\text{Variance}}$$

$$= \sqrt{\frac{52.5441}{31}} = \sqrt{1.694}$$

Standard deviation = 1.301

And finding co-effient of variation,

Formula,

$$C.V = \frac{\text{Standard deviation}}{\text{Mean}} \times 100$$

$$C.V = \frac{1.301}{2.516} \times 100$$

Coeffient of variation = 51.70 %

PTO....

Question No: 04

Write down the short notes on the followings:

- Range
- Quartile Range
- Semi Inter Quartile Range
- Variance
- Standard Deviation
- Coefficient of Variation

Solution :

Range:

In statistics, the range is a measure of spread: it's the difference between the highest value and the lowest value in a data set.

Range is pretty simple when it comes to statistics. I will further explain this with the help of an example.

For instance, we are given the following data

13, 19, 30, 15, 19, 32, 17, 28, 33, 19, 30, 34

Therefore to fine the range we subtract maximum value from the minimum value to find the range.

The largest value is denoted by "Xm", and smallest value by "Xo" (Xnot)

So in my set od data my maximum is 34, and my minimum value is 13. So when I subtract these two I get a value of 21.

Range= Xm – Xo de

34-13=21 (Range) Hence, 21 is the range for my data

Quartile Range:

Quartiles in statistics are values that divide your data into quarters. However, quartiles aren't shaped like pizza slices; Instead they divide your data into four segments according to where the numbers fall on the number line. The four quarters that divide a data set into quartiles are:

The lowest 25% of numbers.

The next lowest 25% of numbers (up to the median).

The second highest 25% of numbers (above the median).

The highest 25% of numbers.

Quartile rang by is a measure of statistical dispersion being equal to the difference between 75^{th} and 25^{th} percentile. It is also called mid spread or H-spread.

Semi-interquartile range

The semi-interquartile range is half of the difference between the upper quartile and the lower quartile.

The semi-interquartile range is a measure of spread or dispersion. It is computed as one half the difference between the 75^{th} percentile [often called (Q3)] and the 25^{th} percentile (Q1). The formula for semi-interquartile range is therefore: (Q3-Q1)/2.

Variance:

Variance in statistics is a measurement of the spread between numbers in a data set. That is, it measures how far each number in the set is from the mean and therefore from every other number in the set.
Variance is calculated by taking the differences between each number in the data set and the mean, then squaring the differences to make them positive, and finally dividing the sum of the squares by the number of values in the data set.

The Formula for Variance is (σ^2)

$$\text{Variance } \sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$$

Where

x_i = the i^{th} data point

$\bar{x}$ = the mean of all data points

n = The number of data points

Standard Division :

The standard deviation is a statistic that measures the dispersion of a dataset relative to its mean and is calculated as the square root of the variance. It is calculated as the square root of variance by determining the variation between each data point relative to the mean. If the data points are further from the mean, there is a higher deviation within the data set; thus, the more spread out the data, the higher the standard deviation.

Standard deviation is a statistical measurement in finance that, when applied to the annual rate of return of an investment, sheds light on the historical volatility of that investment. The greater the standard deviation of securities, the greater the variance between each price and the mean, which shows a larger price range. For example, a volatile stock has a high standard deviation, while the deviation of a stable blue-chip stock is usually rather low.

Formula for Standard Devision

$$\text{Standard Deviation} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

Coefficient of Variation:

The coefficient of variation (CV) is the ratio of the standard deviation to the mean. The higher the coefficient of variation, the greater the level of dispersion around the mean. It is generally expressed as a percentage. Without units, it allows for comparison between distributions of values whose scales of measurement are not comparable.

When we are presented with estimated values, the CV relates the standard deviation of the estimate to the value of this estimate. The lower the value of the coefficient of variation, the more precise the estimate.

End.