Q4

Name = Muhammad Haider Khan

ID = 13702

6th semester

Statistics (final term Assignment/Exam)

Date = 27/06/2020

Sir Raza

Q1 Find Averages (A.M., G.M., H.M.) of the following tables: and also justify its logical Relation.

(a)

| number of Children per family | Number of family. |
|---|---|
| 1 | 4 |
| 2 | 13 |
| 3 | 9 |
| 4 | 4 |
| 5 | 1 |

| number of children per family (x) | number of family (f) | fx | f. log x | f/x |
|---|---|---|---|---|
| 1 | 4 | 4 | $4 \log 1 = 0$ | $4/1 = 4$ |
| 2 | 13 | 26 | $13 \log_2 = 3.913$ | $13/2 = 6.5$ |
| 3 | 9 | 27 | $9 \log 3 = 4.294$ | $9/3 = 3$ |
| 4 | 4 | 16 | $4 \log 4 = 2.408$ | $4/4 = 1$ |
| 5 | 1 | 5 | $1 \log 5 = 0.699$ | $1/5 = 0.2$ |
| $\Sigma f = 31$ | | $\Sigma fx = 78$ | $\Sigma f \log x = 11.314$ | $\Sigma f/x = 14.7$ |

Arthematic Mean:

$$A.M = \frac{\Sigma fx}{\Sigma f}$$

now as we know the value of $\Sigma f$ and $\Sigma fx$ in above table.

pulling value

$$Am = \frac{78}{31} = 2.516$$

Geomatric mean:-

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

putting value from the above table

$$G_m = \text{Antilog}\left(\frac{11.314}{31}\right)$$

$$G_m = \text{Antilog}(0.365)$$

$$G_m = 2.317$$

Hormonic mean

$$H_m = \frac{\Sigma f}{\Sigma(f/x)}$$

putting value from above table

$$H_m = \frac{31}{14.7}$$

$$H_m = 2.108$$

Logical Relationship:-

$$AM \geq GM \geq HM$$

$$2.516 > 2.317 > 2.108$$

Hence the logical Relation is justified.

(b)

| Class Interval mark | Frequency (f) | midpoint | f.m | flogm | f/m |
|---|---|---|---|---|---|
| 0-9 | 2 | 4.5 | 2×4.5 = 9 | 2log4.5 = 1.306 | 2/4.5 = 0.444 |
| 10-19 | 31 | 14.5 | 31×14.5 = 449 | 31log14.5 = 36.002 | 31/14.5 = 2.138 |
| 20-29 | 73 | 24.5 | 73×24.5 = 1788.5 | 73log24.5 = 101.409 | 73/24.5 = 2.98 |
| 30-39 | 85 | 34.5 | 85×34.5 = 2932.5 | 85log34.5 = 130.715 | 85/34.5 = 2.464 |
| 40-49 | 28 | 44.5 | 28×44.5 = 1246 | 28log44.5 = 46.154 | 28/44.5 = 0.629 |
| | Ef = 219 | | Efm = 6425 | E(flogm) = 315.586 | E(f/m) = 8.655 |

Arthematic mean:-

$$Am = \frac{E(fm)}{Ef}$$

as we know from the above data $E(fm) = 6425$ and $Ef = 219$ now putting the value in the arthematic formula we get.

$$Am = \frac{6425}{219} = 30.164$$

Geomatric mean:-

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

pulling value from abave table that has bun calculated.

$$Gm = \text{Antilog}\left(\frac{315.586}{219}\right)$$

$$Gm = \text{Antilog}(1.441)$$

$$\boxed{Gm = 27.607.}$$

Harmonic means:-

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

putting value

$$Hm = \frac{219}{8.655}$$

$$Hm = 25.3 .$$

Logical Relationship of Arithematic, Geomatic and Harmonic mean.

$$AM \geq GM \geq H.M$$

$30.164 > 27.6 > 25.3$

hence logical Relation is justified.

Q2 Find median and mode for the following table.

(a)

| number of children/family | number of family |
|---|---|
| 1 | 4 |
| 2 | 13 |
| 3 | 9 |
| 4 | 4 |
| 5 | 1 |

(b)

| mark | frequency |
|---|---|
| 0 - 9 | 2 |
| 10 - 19 | 31 |
| 20 - 29 | 73 |
| 30 - 39 | 85 |
| 40 - 49 | 28 |

(a) finding median for (a)

| x | f | c.f |
|---|---|---|
| 1 | 4 | 4 |
| 2 | 13 | 17 |
| 3 | 9 | 26 |
| 4 | 4 | 30 |
| 5 | 1 | (31) |

$\Sigma f = 31$

from the table we found that $\Sigma f = 31$ which is an odd number

now as we know that median is the central value

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

$\frac{31+1}{2} = \frac{32}{2} = 16$

(7)
(6)

from the above table we can see that the 16th terms is in 2 so technically our median = 2

Mode:

| number of children/fm | number of family. |
|---|---|
| 1 | 4 |
| 2 | 13 |
| 3 | 9 |
| 4 | 4 |
| 5 | 1 |

from the above table we see that mode is 2. because 2 is the most frequent value of the data.

Q2 (b)

(7)

| Class interval | f | class boundary | cf | 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 | 19.5 – 29.5 | 106 | 24.5 |
| 30 – 39 | 85 | 29.5 – 39.5 | 191 | 34.5 |
| 40 – 49 | 28 | 39.5 – 49.5 | 219 | 44.5 |
| | $n = \Sigma f = 219$ | | | |

→ Median lies here

median =

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

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

$\frac{220}{2}$

110th term

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

$L = 29.5 \quad h = 10 \quad f = 85 \quad \Sigma f = 219$

$C = 106$

putting values:

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

$29.5 + 0.1176\ (3.5)$

$29.5 + 0.4117$

$29.91 = \simeq 30$

(b)

mode

$L_1 = 29.5$ $f_1 = 85$

$i = 10$ $f_2 = 28$

$f_0 = 73$

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

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

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

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

$$\boxed{\text{mode} = 30.74}$$

Q3 find Quartile range and semi inter Quartile Range.

| x | f | cf |
|---|---|---|
| 1 | 4 | 4 |
| 2 | 13 | 17 |
| 3 | 9 | 26 |
| 4 | 4 | 30 |
| 5 | 1 | 31 |

$\Sigma f = 31$

$n = \Sigma f$

for $Q_1$

$$Q_1 = \frac{n+1}{4}$$

$$Q_1 = \frac{31+1}{4}$$

$$Q_1 = \frac{32}{4}$$

$$Q_1 = 8$$

8th term is 2 in table

$$Q_3 = 3\left(\frac{n+1}{4}\right)$$

$$Q_3 = 3\left(\frac{31+1}{4}\right)$$

$$Q_3 = 3\left(\frac{32}{4}\right)$$

$$Q_3 = 3(8)$$

$Q_3$ = 24th term lies in 3

$$Q_3 = 3$$

now to find Quartile range.

$$Q_r = Q_3 - Q_1$$

$$Q_r = 3 - 2$$

$$Q_r = 1$$

To find semi inter Quartile range we use formula.

$$S.I.Q.R = \frac{Q_3 - Q_1}{2}$$

$$S.I.Q.R = \frac{3-2}{2}$$

$$S.I.Q.R = \frac{1}{2}$$

(11)

(10)

Q3 b

| $x$ | $f$ | $fx$ | $x-\bar{x}$ | $(x-\bar{x})^2$ | $(x-\bar{x})^2$ |
|---|---|---|---|---|---|
| 1 | 4 | 4 | -1.5 | 2.25 | 9 |
| 2 | 13 | 26 | -2.5 | 6.25 | 81.25 |
| 3 | 9 | 27 | 0.5 | 6.25 | 2.25 |
| 4 | 4 | 16 | 1.5 | 2.25 | 9 |
| 5 | 1 | 5 | 2.5 | 6.25 | 6.25 |
| | $\Sigma f=31$ | $\Sigma fx=78$ | | | total 107.75 |

$$\text{mean} = \frac{\Sigma fx}{\Sigma f}$$

$$= \frac{78}{31}$$

$$= 2.5$$

now using the formula.

$$s^2 = \frac{\Sigma f (x-\bar{x})^2}{\Sigma f}$$

$$s^2 = \frac{107.75}{31}$$ → variance

$$\boxed{s^2 = 3.5 \ (3.475)}$$

Co-efficient of varience

$$C.V = \left[\frac{\text{varience}}{\bar{x}}\right] \times 100$$

$$CV = \frac{3.5}{2.516} \times 100$$

$$\boxed{CV = 140}$$

OR

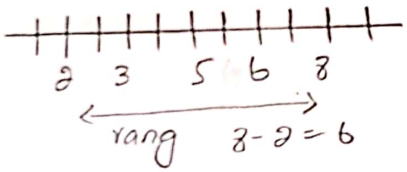
Q4

(a) **Range**:- Range is the difference between the lowest and the highest value.

For Example

{3, 5, 8, 2, 6} the lowest value is 2 and the highest value is 8 so the range is

$8 - 2 = 6$




**Range formula**:- The maximum value − the minimum value.

(b) **Quartile Range**:- Quartile range is a measure of statistical dispersion being equal to the difference between 75th and 25th percentile. Its also called midspread or H-spread.

Quartile Formula:- To find quartile you have to minus the lower quartile from the upper quartile

$$Q_2 = Q_3 - Q_1$$

For Example:- if we have total number of 100 and $Q_3 = 75$ and $Q_1 = 25$

$$Q_2 = 75 - 25$$

$$Q_2 = 50$$

---

(c) Semi Inter Quartile Range:- It is computed as one half of the difference between the 75th percentile and 25th percentile.

Formula:- highest percentile minus lowest percentile and then divided by 2.

$$(Q_3 - Q_1)/2$$

For Example:- we have $Q_3$ as 75 and $Q_1$ as 25 so

$$(Q_3 - Q_1)/2 \Rightarrow (75-25)/2 \Rightarrow 50/2 \Rightarrow 25$$

25 is Semi IQR

(d) Variance:- The variance is one of the measures of the dispersion that is measured of by how much the value in the dataset are likely to differ from the mean values. It is the average of the square of the deviation from the mean. Squaring the deviation making sure that the negative and positive deviation doesnot cancel eachother.

Formula:- The formula for the variance sample

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

where n is sample size and $\bar{x}$ is sample mean.

Now formula for variance of an entive population

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2$$

where N is the population size and $\mu$ is population mean

(e) Standard Deviation:- Standard deviation is the squar root of the variance. It is one of the measure of dispersion, That is the measure of by how much the value in the data set are likely to differ from the mean.

Standard deviation formula:-

For Sample:-

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

For entire population

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

(f) Co-efficient of Variation:- The statistical measure of dispersion of data point around the mean. The metric is completely comon to use compare the data dispersion between distinct data

Formula:-

$$\text{Coefficient of Variation} = \frac{\sigma}{\mu} \times 100\%$$

where $\sigma$ is standard deviation and $\mu$ is the mean.