ASSIGNMENT
SPRING-2020

Program: B.B.A / MMC
Course Title: BASIC STATISTICS / STATISTICS-I
Dated: 27th June ,2020
Instructor: Raza Ahmed Khan **Total marks: 50**

Important Instructions:

- **Assignment should be submitted within 6 – Hours.**
- **Submitted Document's format should be in word, pdf or in jpg.**
- **No Assignment will be accepted after due date mentioned above.**
- **Each question carries 12.5 Marks.**
- **Attempt All Questions**

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

a. Solution

Midpoint(x)	Freq.	Fx	logx	f logx	f/x
1	4	4	0	0	4
2	13	26	0.3010	3.913	6.5
3	9	27	0.4771	4.2939	3
4	4	16	0.6021	2.4084	1
5	1	5	0.6990	0.6990	0.2

$$Aritmatic\ Mean = \overline{X} = \frac{\sum fx}{\sum f}$$

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

$$= 2.52$$

$$Geometric\ Mean = G.M = anti\log\left(\frac{\sum f\log(x)}{\sum f}\right)$$

$$= anti\log\left(\frac{11.3143}{31}\right)$$

$$= anti\log(0.3650)$$

$$= 2.317$$

$$Harmonic\ Mean = H.M = anti\log\left(\frac{\sum f}{\sum \frac{f}{x}}\right)$$

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

$$= 2.11$$

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

$$2.52 \geq 2.317 \geq 2.11$$

b. Solution

classes	Freq.	Midpoint(x)	Fx	logx	f logx	f/x
0-9	2	4.5	9	0.6532	1.31	0.44
10-19	31	14.5	449.5	1.1614	36.00	2.138
20-29	73	24.5	1788.5	1.3892	101.41	2.980
30-39	85	34.5	2932.5	1.5378	130.71	2.464
40-49	28	44.5	1246	1.6484	46.15	0.629

$$Aritmatic\ Mean = \overline{X} = \frac{\sum fx}{\sum f}$$

$$= \frac{6425.5}{219}$$

$$= 29.34$$

$$Geometric\ Mean = G.M = anti\log\left(\frac{\sum f\log(x)}{\sum f}\right)$$

$$= anti\log\left(\frac{315.58}{219}\right)$$

$$= anti\log(1.44)$$

$$= 27.54$$

$$Harmonic\ Mean = H.M = anti\log\left(\frac{\sum f}{\sum \frac{f}{x}}\right)$$

$$= \frac{219}{8.651}$$

$$= 25.31$$

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

$$29.34 \geq 27.54 \geq 25.31$$

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

a. solution

X	Freq.	Cumulative Freq.
1	4	4
2	13	17
3	9	26
4	4	30
5	1	31

$$crossponding\ value\ of\ X\ at\left(\frac{n}{2}+1\right)^{th}\ group\ class$$

$$=\left(\frac{31}{2}+1\right)^{th}\ group\ class$$

$$=16.5^{th}\ group\ class$$

$$=17$$

$$crossponding\ value\ at\ X\ from\ 17\ is\ 2$$

$$Hence\ Median = 2$$

$$crossponding\ value\ of\ X\ at\ highest\ Freq. = 13$$

$$Hence\ Mode = 2$$

b. solution

classes	Freq.	Cumulative Freq.	Class boundaries
0-9	2	2	0-9.5
10-19	31	33	9.5-19.5
20-29	73	106	19.5-29.5
30-39	85	191	29.5-39.5
40-49	28	219	39.5-49.5

$$Median = l + \frac{h}{f}\left(\frac{n}{2} - C.F\right)$$

$$\frac{n^{th}}{2}\ group\ class = \frac{219^{th}}{2} = 109.5^{th}\ group\ class$$

$$l = lower\ class\ boundry$$

$$h = size, \qquad f = freq\ of\ group\ calss$$

$$C.F = cumulative\ \ freq\ of\ group\ calss$$

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

$$= 29.91$$

$$Mode = l + \frac{(f_m - f_1)}{(f_m - f_1) + (f_m - f_2)} \times h$$

$$= 29.5 + \frac{(85 - 73)}{(85 - 73) + (85 - 28)} \times 10$$

$$= 31.24$$

Question No: 03

a. Find Semi Quartile Range & Semi Inter Quartile Range of Q2(a)

classes	Freq.	Cumulative Freq.
1	4	4
2	13	17
3	9	26
4	4	30
5	1	31

$crossponding\ value\ of\ \ X\ \ at\ \left(\frac{n}{4}+1\right)^{th}\ group\ class$

$=\left(\frac{31}{4}+1\right)^{th}\ group\ class$

$=8.75^{th}\ group\ class$

$=17$

$crossponding\ value\ of\ \ X\ \ at\ 17\ is\ 2$

$Hence\ Q_1=2$

$crossponding\ value\ of\ \ X\ \ at\ \left(\frac{3n}{4}+1\right)^{th}\ group\ class$

$=\left(\frac{3\times 31}{4}+1\right)^{th}\ group\ class$

$=24.25^{th}\ group\ class$

$=26$

$crossponding\ value\ of\ \ X\ \ at\ 26\ is\ 3$

$Hence\ Q_3=3$

$S.Q.R=Q_3-Q_1$

$=3-2$

$=1$

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

$=\frac{3-2}{2}$

$=0.5$

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

$=\frac{3-2}{3+2}$

$=0.2$

b. Find Variance and Co-efficient of variance of Q2(a)

X	Freq.	x^2	$f\ x^2$
1	4	1	4
2	13	4	52
3	9	9	81
4	4	16	64
5	1	25	25

$$\text{var} = S^2 = \frac{\sum f\ X^2}{\sum f} - \left(\overline{X}\right)^2$$

$$S^2 = \frac{226}{31} - \left(2.52\right)^2$$

$$S^2 = 0.940$$

$$C.V. = \frac{\sqrt{S^2}}{\overline{X}} \times 100$$

$$C.V. = \frac{\sqrt{0.940}}{2.52} \times 100$$

$$C.V. = 38.47\%$$

Question No: 04

write down the short notes on the followings:

- **Range**

  Range is basically a difference between maximum and minimum value in a data. Mathematically it is written as

  $$Range = X_m - X_o$$

  It is the simple and easy method to measure the dispersion but it does not utilized all values in the data. So it is not consider as a good way to measure the dispersion. In a grouped, it is impossible to calculate range if data consist of open end classes.

- **Quartile Range**

  Quartiles distribute the whole data into four equal parts and denoted as Q1, Q2, and Q3 called lower quartile, median and upper quartile represents 25%, 50% and 75% of the data. Quartile range describes the range from lower to upper quartile and how much value lies in this range

- **Semi Inter Quartile Range**

  This is half of the difference between third and first quartile, denoted as

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

  Its main feature is that it consist the 50% of the data and not affected by extreme values. It is not widely in used because it does not utilized all values in the data, this measure only consist of 50% of the data and does not have any information about extreme values.

- **Variance**

  It defies the variation of the data. If it is calculated for population denoted as σ^2 and if calculated for sample denoted as S^2. Symbolically written as

  $$\text{var} = S^2 = \frac{\sum f\ X^2}{\sum f} - \left(\overline{X}\right)^2$$

- **Standard Deviation**

  Standard deviation defines how much data units are spread around their mean. Symbolically written as

$$S = \sqrt{\frac{\sum f\, X^2}{\sum f} - \left(\overline{X}\right)^2}$$

  Smaller value defines the data units are very near to their mean and highest values defines they are highly scattered around their mean.

- **Coefficient of Variation**

  It defines the variability in terms of percentage. Mathematically written as

$$C.V. = \frac{\sqrt{S^2}}{\overline{X}} \times 100$$

  It is used for comparison of variability between two or more data sets. Advantage is that it is unit less. Higher the CV defines more variability and lower the CV denies less variability

---