Date: 15 March 2021

VIRTUAL COACHING CLASSES
ORGANISED BY BOS (ACADEMIC), ICAI

FOUNDATION LEVEL
PAPER 3: BUSINESS MATHEMATICS AND LOGICAL REASONING & STATISTICS

Faculty: CA Rashmi Lonikar, M.Sc., FCA, DISA

Sr. No.	Concept	Formula	Alternate Formula / Comment	
1.	Arithmetic Mean	$\frac{\sum_{i=1}^{n} x_i}{n}$	$\frac{\sum f_i x_i}{N}$	$\bar{x} = A + \frac{\sum f_i d_i}{N} \times C$
2.	Median	$M = l_1 + \left(\frac{\frac{N}{2} - N_l}{N_u - N_l} \right) \times C$		
3.	Quartiles	$Q = l_1 + \left(\frac{Np - N_l}{N_u - N_l} \right) \times C$		
4.	Mode	$\text{Mode} = l_1 + \left(\frac{f_0 - f_{-1}}{2f_0 - f_{-1} - f_1} \right) \times c$		
5.	Empirical Relation	Mean – Mode = 3(Mean – Median)		
6.	Geometric Mean	$G = (x_1 \times x_2 \times x_3 \ldots\ldots\ldots \times x_n)^{1/n}$	$G = (x_1^{f_1} \times x_2^{f_2} \times x_3^{f_3} \ldots\ldots\ldots \times x_n^{f_n})^{1/N}$	
7.	Harmonic Mean	$H = \frac{n}{\sum (1/x_i)}$	$H = \frac{N}{\sum \left[\frac{f_i}{x_i} \right]}$	

What is “Central Tendency”

- CT => Tendency of a data set to cluster around a particular value (usually central, middle)
- That central value is called “Measure of Central Tendency”
- Essentially, it is a number that is ***representative of a data set***
- Measures: Arithmetic Mean (AM), Median, Mode, Geometric Mean (GM), Harmonic Mean (HM)
- Why so many different measures?
  Ans: Not all data sets are same. Some data may be better represented by other measures
- Ex: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10: Almost all measures are good representations (AM=5.5, Me = 5, GM=4.5
- Ex: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10000: Here AM is 1004.5 but is this a good representation of the data? No. Only one value is high, others are single digits. So we should consider other measures

Criteria for Ideal Measure of CT

(i) It should be properly and unambiguously defined.

Opposite Ex: Let the central tendency measure be the value ***around the middle*** of the data set

(ii) It should be easy to comprehend.

(iii) It should be simple to compute.

Opposite Ex: Let the central tendency measure be ***5th root of product of data if data is equal to the square of the sum of price of similar commodity, if not, 6th root of a weighted average of any such similar data sets, raised to the power equal to the number of outliers in the aforementioned data sets***

(iv) It should be based on all the observations.

Opposite Ex: Let the central tendency measure be the ***mean of first 4 values, ignoring all others***

(v) It should have certain desirable mathematical properties.

(vi) It should be least affected by the presence of extreme observations.

Arithmetic Mean (AM) $\frac{\sum_{i=1}^{n} x_i}{n}$

- Defined as ***Sum of all observations divided by number of observations***
- Only data given: Simply add all data and divide by number of data points $\frac{\sum_{i=1}^{n} x_i}{n}$
- Frequency distribution given: Sum of "ClassMidPoint x Class Frequency" / total frequency $\frac{\sum f_i x_i}{N}$
- Assumed Mean method: When to use? 'Asked in question' or when data points are too large, causing calculator difficulties

$$\bar{x} = A + \frac{\sum f_i d_i}{N} \times C$$

Where $d = \frac{x_i - A}{c}$

Question Time!

- Find AM of 9 numbers 58, 62, 48, 53, 70, 52, 60, 84, 75 (Ans: 62.44)

- The weights of a group of BBA students of St. Xavier's College is as follows:

Weights (kg)	44-48	49-53	54-58	59-63	64-68	69-73
Number of students	3	4	5	7	9	8

Calculate AM (Ans: 61.42)

Properties of AM

- *1. If all the observations are constants say 'k' then AM is also 'k'*

- *2. The algebraic sum of deviations from the AM is always zero*

  Q: If there are three observations 15, 20, 25, then the sum of deviations from the AM is
  (a) 0 (b) 5 (c) -5 (d) None Of These

- *3. AM is affected by change of origin and/or scale*
  If we change origin say 'a' and change scale say 'b', such that the variable 'x' is changed to variable 'y' $y = a + bx$, then mean changes as $\bar{y} = a + b\bar{x}$
  Suppose I have a series 'x' 1, 2, 3, 4, 5 and $y = 3+2x$

  Q: If relation between x and y is given by $3y + 2x + 7 = 0$, find $\bar{y}$ if $\bar{x} = 15$ (Ans: -12.33)

- *4. COMBINED MEAN: If we have two groups with n1 and n2 observations and with means $\bar{x}1$, $\bar{x}2$ then their combined mean $\bar{x}$ is given by*

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$$

- **PROPERTIES**
- If x is the arithmetic mean of n observations $x_1, x_2, x_3, . . x_n$; then

---

- If p is added to all the items, then the AM increases by p.
  - the mean of the new observations is (x + p)

- If p is subtracted from all the items, then the AM decreases by p.
  - The mean of the new observations is (x - p).

- If every item is multiplied by a non –zero p, then the AM also gets multiplied by p.
  - the mean of the new observations is px.

- If every item is divided by p, then the AM also gets divided by p.
  - the mean of the new observations is (x/p).

$(x_1 - x) + (x_2 - x) + (x_3 - x) + ... + (x_n - x) = 0$. i.e summation of deviations from mean is zero

Let the numbers be 5, 10, 15, 20.

1. If the mean of 5 numbers 9, 8, 10, x, 12 is 15, find the value of x.

2. If the mean of five observations x, x + 4, x + 6, x + 8 and x + 12 is 16, find the value of x.

3. The mean of 40 numbers was found to be 38. Later on, it was detected that a number 56 was misread as 36. Find the correct mean of given numbers.

4. The average of x, y and z is 45. x is as much more than the average as y is less than the average. Find the value of z.
A. 45 years B. 35 years C. 60 years D. 15 years

5. The mean of eight numbers is 25. If five is subtracted from each number, what will be the new mean?

Properties of AM

- Q: The mean salary for a group of 40 workers is Rs 5200 per month and that for another group of 60 workers is Rs 6800 per month. Find the combined mean salary.

  Ans: Rs. 6160

- Q: If there are two groups containing 30 and 20 observations and having 50 and 60 arithmetic means, find the combined arithmetic mean.

  Ans : 54

Median

- Compared to AM, median is a ***positional measure*** i.e. its value depends on the size of data and the central position within said data
- It is defined as the ***middlemost value*** when the data is ***arranged in ascending or descending order***
- Example: Marks of students = 72, 85, 56, 80, 65, 52, 68

  72, 85, 56, 80, 65, 52, 68

Middlemost value = 80 so median = 80? **Careful!**

"First ARRANGE then Middle"

- Arranged data: 52, 56, 65, 68, 72, 80, 85 *Median = 68*
- When number of data points = even, take AM of middle two values
- Ex: 56, 82, 82, 96, 100, 106, 110, 120 *Median* = (96+10)/2 = *98*

Median

- For grouped frequency, formula for median:

$$M = l_1 + \left(\frac{\frac{N}{2} - N_l}{N_u - N_l} \right) \times C$$

Where,

l_1 = lower class boundary of the median class i.e. the class containing median.

N = total frequency.

N_l = less than cumulative frequency corresponding to l_1. (Pre median class)

N_u = less than cumulative frequency corresponding to l_2. (Post median class)

l_2 being the upper class boundary of the median class.

$C = l_2 - l_1$ = length of the median class.

Question Time!

- The weights of a group of BBA students of St. Xavier's College is as follows:

Weights (kg)	44-48	49-53	54-58	59-63	64-68	69-73
Number of students	3	4	5	7	9	8

Find median. (Ans: 62.79)

IMP: Write down the class boundaries and less than cumulative frequencies

Properties of Median

- 1. *If the variable 'x' is related to variable 'y' by the relation y = a + bx, then we can find median of y by the formula* $\boldsymbol{y_{me} = a + bx_{me}}$

  Q: If the median of x is 16 and x and y are related by 2x -5y -10 = 0. Find median of x. (Ans: 4.4)

- 2. *The sum of* ***absolute*** *deviations from median is minimum* $\Sigma |\ x - x_{me}\ |$

Partition Values / Quartiles / Fractiles

- Median divides the data set into two halves. What if we want 4 equal parts? What would those 3 numbers be, that would divide the set into 4 parts? **Quartiles**
- Similarly **Deciles** divide the set into 10 equal parts, and **Percentiles** divide the set into 100 equal parts

Term	Quartiles	Deciles	Percentiles
Divides data into	4 equal parts	10 equal parts	100 equal parts
Total number of such terms	3	9	99
Unclassified Data Formula (same for all!)	(n+1) x pth **term** Notice the **term**, do not just calculate (n+1)*p and tick option		
Grouped Data Formula (same for all!)	$Q = l_1 + \left(\frac{Np - N_l}{N_u - N_l}\right) \times C$		
Value of p	1/4, 2/4, 3/4	1/10, 2/10, … 9/10	1/100, 2/100, … 99/100

Question Time! (Part 1)

- The unclassified data for the weights for 10 persons - 82, 56, 65, 75, 75, 80, 82, 90, 120, 130
  Find Q1, D6 and P82

- Ans : Q1 = 62.75, D6 = 81.20, P82 = 120.20

- **CAREFUL – DON'T FORGET TO ARRANGE IN ASCENDING ORDER**

Question Time! (Part 2)

- The weights of a group of BBA students of St. Xavier's College is as follows:

Weights (kg)	44-48	49-53	54-58	59-63	64-68	69-73
Number of students	3	4	5	7	9	8

Calculate D7 and Q3. (Ans: 66.94, 67.94)

Imp: Write down the class boundaries and less than cumulative frequency!

Mode

- Defined as ***Value that occurs maximum number of times*** i.e. most common value

- Only data given: Simply count frequency of all unique data points and select the one with largest frequency

  Imp: More than one mode possible for given data
  Ex: 2, 4, 5, 5, 1, 1, 6, 7 Here both 1 and 5 occur two times, so both are modes. Such data is called **BIMODAL data**.

- Grouped frequency given:

$$\text{Mode} = l_1 + \left(\frac{f_0 - f_{-1}}{2f_0 - f_{-1} - f_1}\right) \times c$$

where,

l_1 = LCB of the modal class.
i.e. the class containing mode.
f_0 = frequency of the modal class
f_{-1} = frequency of the pre-modal class
f_1 = frequency of the post modal class
C = class length of the modal class

Question Time!

- The weights of a group of BBA students of St. Xavier's College is as follows:

Weights (kg)	44-48	49-53	54-58	59-63	64-68	69-73
Number of students	3	4	5	7	9	8

Find mode. (Ans: 66.83)

Find mean median mode for the following data

CLASS INTERVAL	FREQUENCY
13-15	3
16-18	11
19-21	14
22-24	7
25-27	5

20, 19.79, 19.4

Empirical Formula

- When it is difficult to calculate mode from grouped data, we may consider the empirical (i.e. experimentally obtained) formula:

$$\text{Mean} - \text{Mode} = 3(\text{Mean} - \text{Median})$$

- Note: This is an empirical formula, obtained by analysing large data sets. It is **NOT A THEOREM** so it may not always hold exactly. It is to be used only when asked in question and/or when data is such that direct calculation is not possible for all 3 terms

- Q: For a group of 200 students, the mean marks and median marks were found to be 55.60 and 52.40 respectively. Find the mode.
  Ans: 46

Revision Questions: AM, Mode, Median

- **For unclassified data:**

- 1. Find median of 5, 8, 6, 9, 11, 4 (Ans: 7)
- 2. Find the mode for 5, 8, 6, 4, 10, 15, 18, 10 (Ans: 10)
- 3. Find the first quartile for 15, 18, 10, 20, 23, 28, 12, 16 (Ans: 12.75)
- 4. Find the third decile for the numbers 15, 10, 20, 25, 18, 11, 9, 12 (Ans: 10.70)

- **For missing frequency**
- The following distribution has a mode of 44. Find its mean. (Missing f = 20, Mean = 48)

Marks	0-20	20-40	40-60	60-80	80-100
No. of Students	5	18	?	12	5

Geometric Mean

- Defined as **nth root of product of all 'n' observations**
- Ungrouped data given: **$GM = (x_1 \cdot x_2 \cdot x_3 \ldots \cdot x_n)^{(1/n)}$**
- Grouped data given: **$GM = \{(x_1)^{f1} \cdot (x_2)^{f2} \cdot (x_3)^{f3} \ldots \cdot (x_n)^{fn})^{(1/N)}$**
  Here, N = f1 + f2 + … + fn

Properties of GM

(i) Logarithm of G for a set of observations is the AM of the logarithm of the observations; i.e.

$$\log G = \frac{1}{r}\sum \log x \qquad \ldots\ldots\ldots\ldots\ldots\ldots(15.1.13)$$

(ii) if all the observations assumed by a variable are constants, say K > 0, then the GM of the observations is also K.

(iii) GM of the product of two variables is the product of their GM's i.e. if $z = xy$, then

$$\text{GM of } z = (\text{GM of } x) \times (\text{GM of } y) \qquad \ldots\ldots\ldots\ldots\ldots\ldots(15.1.14)$$

(iv) GM of the ratio of two variables is the ratio of the GM's of the two variables i.e. if $z = x/y$ then

$$\text{GM of } z = \frac{\text{GM of } x}{\text{GM of } y} \qquad \ldots\ldots\ldots\ldots\ldots\ldots(15.1.15)$$

Question Time!

- Q: Find the GM of 3, 6, 12 (Ans: 6)
- Q: Find the GM of 8, 24, 40 (Ans: $8\sqrt[3]{15}$)
- Q: If GM of X is 10 and GM of Y is 15, find GM of XY (Ans: 150)
- Q: If AM and GM for two numbers are 6.50 and 6 respectively, find the two numbers.
  (a) 6 and 7 (b) 9 and 4 (c) 10 and 3 (d) 8 and 5 (Ans: (b))

Harmonic Mean

- Defined as **reciprocal of AM of reciprocal of observations**
  Since we are dealing with reciprocal of observations, obviously they need to be non-zero (1/0 issue)
- Ungrouped data given: $$H = \frac{n}{\Sigma(1/x_i)}$$
- Grouped data given: $$H = \frac{N}{\Sigma\left[\frac{f_i}{x_i}\right]}$$

  Here, N = f1 + f2 + … + fn

Properties of HM

(i) If all the observations taken by a variable are constants, say k, then the HM of the observations is also k.

(ii) If there are two groups with n_1 and n_2 observations and H_1 and H_2 as respective HM's than the combined HM is given by

$$\frac{n_1 + n_2}{\frac{n_1}{H_1} + \frac{n_2}{H_2}} \qquad \text{..................(15.1.18)}$$

Question Time!

- Q: Find HM for 4, 6, 10 (Ans: 5.77)
- Q: Find HM for 2, 3, 5 (Ans: 2.91)
- Q: Two groups with 15 and 13 observations have harmonic means 75 and 65 respectively. Find combined harmonic mean. (Ans: 70)
- Q: An aeroplane flies from A to B at a speed of Rs 500 kph and returns at speed of 700 kph. Find the average speed. (Ans: 583.33 kph)

Revision Questions: AM, GM, HM

- Q: If AM and HM for two numbers are 5 and 3.2, find GM (Ans: 4)
- Q: If AM and GM for 10 observations are 15 and 15, find the value of HM (Ans: 15)
- Q: Compute AM,GM and HM for the numbers 6, 8, 12, 36 (Ans: 15.50, 12, 9.93)

THANK YOU