An Introduction to Statistics and Probability

3rd Edition

M. Nurul Islam

Mullick & Brothers

Chapter

2 Central Tendency

Central tendency:

If we consider a frequency distribution, it may be observed that there is a tendency among the values to cluster around a certain value.

The tendency to describe the observation with the central value is called central tendency.

Central tendency of distribution contributes a representative or typical value.

Measure of central tendency: Measure of location:

When we are interested in describing the position of a particular value relative to other values in the distribution, we use numerical indices known as measure of central tendency.

According Simpson and Kafka; "A measure of Central Tendency is a typical value around which other figures congregate."

A typical or representative value is found by measure of central tendency.

Types of measure of central tendency: There are five types of measure of central tendency:

1. Arithmetic Mean.
2. Geometric mean.
3. Harmonic mean.
4. Median.
5. Mode.

1. **Arithmetic Mean:** The arithmetic mean is the sum of a set of observations divided by the number of such observations. It is denoted by $\overline{X}$.

In ungrouped data,

Let $x_1, x_2, \ldots\ldots\ldots\ldots x_n$ be a set of n observations

Then the arithmetic mean,

$$\bar{x} = \frac{x_1 + x_2 + \ldots\ldots\ldots\ldots + x_n}{n}$$

$$= \frac{\sum_{i=1}^{n} x_i}{n}$$

In grouped data:

Let in frequency distribution, $f_1, f_2 \ldots\ldots\ldots\ldots f_n$ be the corresponding frequency of the $x_1, x_2 \ldots\ldots\ldots\ldots\ldots\ldots x_n$ of n observations.

Then the Arithmetic mean can be expressed as

$$\bar{x} = \frac{f_1 x_1 + f_2 x_2 + \ldots\ldots\ldots\ldots\ldots + f_n x_n}{f_1 + f_2 + \ldots\ldots\ldots\ldots\ldots + f_n}$$

$$= \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{k} f_i} = \frac{\sum_{i=1}^{k} f_i x_i}{n} \quad [\because f_1 + f_2 + \ldots\ldots + f_n = n]$$

2. **Geometric mean:** The geometric mean of n (non-zero)positive values $x_1, x_2, \ldots\ldots\ldots\ldots\ldots\ldots x_n$ is defined as the nth positive roots of their products.

It is denoted as G.

$$\therefore G = (x_1 \cdot x_2 \cdots\cdots\cdots x_n)^{1/n}$$

$$\Rightarrow \log G = \frac{1}{n} \sum_{i=1}^{n} \log x_i$$

$$\Rightarrow G = Anti\log\left[\frac{1}{n} \sum_{i=1}^{n} \log x_i\right]$$

In grouped data:

Let $f_1, f_2 \ldots\ldots\ldots\ldots f_n$ be the corresponding frequency of the x_1, x_2

................x_n of n observations. and $\sum f_i = N$

Then Geometric Mean,

$$G = \left(x_1^{f_1} \cdot x_2^{f_2} \,................\, x_n^{f_n}\right)^{1/N}$$

$$\Rightarrow \log G = \frac{1}{N}\sum_{i=1}^{n} f_i \log x_i$$

$$\Rightarrow G = Anti\log\left[\frac{1}{N}\sum_{i=1}^{n} f_i \log x_i\right]$$

In Combined Set,

Let, Geometric Mean, G_1 of n_1 number

Geometric Mean, G_2 of n_2 number

Geometric Mean, G_3 of n_3 number

And $n_1+n_2+n_3$ number Geometric mean, G

$$\log G = \frac{n_1 \log G_1 + n_2 \log G_2 + n_3 \log G_3}{n_1 + n_2 + n_3}$$

3. **Harmonic mean:** Harmonic mean is the reciprocal of the arithmetic mean of the reciprocal of the individual values. It is denoted by H.

In ungrouped /Simple Series:

Let x_1, x_2 ....x_n be a set of n observations,

Then Harmonic mean, $H = \dfrac{n}{\dfrac{1}{x_1} + \dfrac{1}{x_2} + \dfrac{1}{x_3} + ... + \dfrac{1}{x_n}}$

In grouped data:

Let in frequency distribution, f_1, f_2:............f_n be the corresponding frequency of x_1, x_2 ....x_n of n observations

Then Harmonic Mean

$$H = \frac{f_1 + f_2 + \dots\dots\dots f_n}{\frac{f_1}{x_1} + \frac{f_2}{x_2} + \frac{f_3}{x_3} + \dots\dots\dots\dots\dots + \frac{f_n}{x_n}}$$

$$= \frac{\sum_{i=1}^{n} f_i}{\sum_{i=1}^{n} \frac{f_i}{x_i}}$$

$$= \frac{n}{\sum_{i=1}^{n} \frac{f_i}{x_i}}$$

4. **Median:** The median may be defined as the middle most or central value of the variable when the values are arranged in order of magnitude or as the value such that greater and smaller values with equal frequency.

In case of a frequency curve, the median may be defined as that value of the variable which divides the area of the curve into two equal parts.

According to Connor:

The median is that value of the variable which devices the group in two equal parts, one part comprising all the values greater and the other all values less than the median.

In case of ungrouped data:

When the number of values n is odd;

The middle most value i.e. the $\frac{n+1}{2}$ th value in the arrangement will be the Median.

When n is even,

$$\text{The Median} = \frac{\frac{n}{2}th \text{ value} + \left(\frac{n}{2}+1\right)th \text{ value}}{2} \text{ in the arrangement}$$

In grouped data:

Median,

$$Me = L_1 + \frac{\frac{n}{2} - F_c}{f_m} \times h$$

Here,

L_1=Lower Limit of the median class

n=Total number.

F_c=Pre cumulative frequency of the median class

f_m= Frequency of the Median class.

h = Class interval of the Median class.

5. **Mode:** Mode is the value occurring most frequency in a series of items and around which the other items are distributed more densely.

According to Croxton & Cowden:

The Mode of a distribution is the value at the point around which the items tends to be more most heavily concentrated.

It may be regarded as the most typical of a series of values.

In case of ungrouped series: Mode can be determined by observations.

Let in a series, we have the number of 5, 7,4, 6, 5, 3, 4, 5, 6.

Here number 5 is the most frequently. (3 times)

$\therefore Mode = 5$

In case of grouped series: Mode can not be determined by observations or directly. Mode class is the class having the highest frequency.

Mode is determined by the following formula:

$$M_0 = L_0 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times c$$

$M_0 = Mode$

$L_0 = Lower$ Limit of the Modal class.

$\Delta_1 = Difference$ between frequency of Modal class and pre Modal class.

$\Delta_1 = Difference$ modal and post

C= Class interval of Modal class.

Criteria of an Ideal Measure of Central Tendency:

According to professor Yule, the characteristics of an ideal measure of central tendency are following:

1. It should be correctly and rigidly defined.
2. It should be readily comprehensive and easy to calculate.
3. It should be based on all the observations.
4. It should be suitable for further mathematical treatment.
5. It should be affected as little as possible by a sampling fluctuation.
6. It should not be unduly affected by a few extreme values.
7. It should be easy to understand.

Property of Arithmetic Mean:

1. **The sum of the deviations of the observations from their Mean is zero. i.e.** $\sum_{i=1}^{n}(x_i - \overline{x}) = 0$

Proof: Let x_1, x_2.....................x_n be a set of n observations of x.

Then Arithmetic Mean, $\overline{x} = \frac{x_1 + x_2 + \ldots\ldots\ldots\ldots + x_n}{n}$

$$= \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\Rightarrow n\overline{x} = \sum_{i=1}^{n} x_i \text{.........................}(1)$$

$\therefore$ Sum of the deviation of the observation from the A.M. is

$$(x_1 -)\overline{x} + (x_2 - \overline{x}) + \text{.........................} + (x_n - \overline{x})$$

$$= \sum_{i=1}^{n} (x_i - \overline{x})$$

$$= (x_1 + x_2 + \text{..........} + x_n) - (\overline{x} + \overline{x} + \text{..............} + \overline{x})$$

$$= \sum_{i=1}^{n} x_i - n\overline{x}$$

$$= n\overline{x} - n\overline{x} \qquad [\text{From equation (1)}]$$

$$= 0$$

2. **Sum of squares of deviations of the observations from A.M. is least.**

i.e. $\sum (x_i - \overline{X})^2 \leq \sum (x_i - a)^2$

Proof: Let $x_1, x_2, \text{..............} x_n$ be a set of n observations of variable x.

Then A.M. $\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$

a is any value and $a \neq \overline{x}$ $\therefore \sum (x_i - \overline{x}) = 0 \text{.............}(1)$

Then, $\sum(x_i - a)^2 = \sum(x_i - \overline{x} + \overline{x} - a)^2$

$= \sum(x_i - \overline{x})^2 + 2\sum(x_i - a)\overline{x}(-a) + \sum(\overline{x} - a)^2$

$= \sum(x_i - \overline{x})^2 + 0 + n(\overline{x} - a)^2$ [Equation(1)]

$= \sum(x_i - \overline{x})^2 + n(\overline{x} - a)^2$

$Since, a \neq \overline{x}, n(x_i - a)^2 > 0$

$\Rightarrow \sum(x_i - \overline{x})^2 \leq \sum(\overline{x} - a)^2$

3. **If a set of m observations $x_1, x_2 \ldots\ldots\ldots\ldots x_m$ and that of n observations $y_1, y_2, \ldots\ldots\ldots\ldots y_n$ have mean $\overline{x}$ and $\overline{y}$ respectively.**

Then the combined mean $\overline{z}$ of (m+n) observations of the sets is

$$\overline{z} = \frac{m\overline{x} + n\overline{y}}{m+n}$$

Proof: By definition,

$$\overline{x} = \frac{x_1 + x_2 + \ldots\ldots\ldots + x_m}{m} = \frac{\sum_{i=1}^{m} x_i}{m}$$

$$\Rightarrow \sum_{i=1}^{m} x_i = m\overline{x} \ldots\ldots\ldots\ldots (1)$$

Similarly,

$$\overline{y} = \frac{y_1 + y_2 + \ldots\ldots\ldots + y_n}{n} = \frac{\sum_{i=1}^{n} y_i}{n}$$

$$\Rightarrow \sum_{i=1}^{n} y_i = n\overline{y} \ldots\ldots\ldots\ldots (2)$$

Adding (1) & (2) $\Rightarrow \sum_{i=1}^{m} x_i + \sum_{i=1}^{n} y_i = m\overline{x} + n\overline{y}$

The combined set consists of the following (m+n) observations

$x_1.x_2..........x_m.y_1.y_2..............y_n$

If $\bar{z}$ is the mean of combined set,

Then,

$$\bar{z} = \frac{x_1.x_2.............x_m.y_1.y_2..................y_n}{m+n}$$

$$= \frac{\sum_{i=1}^{m} x_i + \sum_{i=1}^{n} y_i}{m} = \frac{m\bar{x} + n\bar{y}}{m+n}$$

4. A.M. depends on origin and scale of Measurement.

Proof: Let $x_1, x_2, ...x_n$ be the nth variable of variable x.

Then Arithmatic Mean, $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$

Now let us change the origin and scale of the variable x by a and c respectively and defined new variable u_i

Where,

$$u_i = \frac{x_i - a}{c} \qquad [i=1,2,3..................]$$

$$\Rightarrow x_i - a = cu_i$$

$$\Rightarrow x_i = a + cu_i$$

$$\Rightarrow \sum x_i = \sum a + \sum cu_i$$

$$\Rightarrow \frac{\sum x_i}{n} = \frac{\sum a}{n} + \frac{\sum cu_i}{n}$$

$$\Rightarrow \bar{x} = a + c\bar{u}$$

We see that, in the relation origin a and scale c is present, Hence A.M is not independent of origin and scale of measurement.

Theorem: For two non -zero positive values x_1 and x_2.

Prove that $AM \geq GM \geq HM$ where AM, GM and HM are respectively Arithmetic Mean, Geometric Mean and Harmonic

Mean.

Proof: By definition,

$$AM = \frac{x_1 + x_2}{2}, GM = \sqrt{x_1 \times x_2}, HM = \frac{2}{1/x_1 + 1/x_2}$$

Taking the difference of AM and GM

$$\therefore \text{ AM-GM} = \frac{x_1 + x_2}{2} - \sqrt{x_1 \times x_2}$$

$$= \frac{x_1 + x_2 - 2\sqrt{x_1 \times x_2}}{2}$$

$$= \frac{1}{2}\left(\sqrt{x_1} - \sqrt{x_2}\right)^2$$

$= \textit{Which}$ is always positive

$\therefore$ AM-GM ≥ 0

$\Rightarrow$ AM $\geq GM$................(1)

Again Taking the difference of GM and HM

$$GM - HM = \sqrt{x_1 \times x_2} - \frac{2}{1/x_1 + 1/x_2}$$

$$= \sqrt{x_1 \times x_2} - \frac{2x_1 x_2}{x_1 + x_2}$$

$$= \frac{\sqrt{x_1 \times x_2}}{x_1 + x_2}\left(x_1 + x_2 - 2\sqrt{x_1 \times x_2}\right)$$

$$= \frac{\sqrt{x_1 \times x_2}}{x_1 + x_2}\left(x_1 - x_2\right)^2$$

$= \textit{Which}$ is always positive

$\therefore$ GM $\geq$ HM...............(2)

From eq^n (1) & (2) $\Rightarrow$ AM $\geq$ GM $\geq$ HM.

Theorem: For a set of n- non zero positive values x_1 x_2..........x_n

Prove that. $AM \geq GM \geq HM$.

Proof: Let $x_1, x_2 \ldots\ldots\ldots x_n$ be a set of n observations of positive values.

$\therefore$ Arithmetic Mean, $(AM), A = \dfrac{x_1 + x_2 + \ldots\ldots\ldots x_n}{n}$

$= \dfrac{\sum_{i=1}^{n} x_i}{n}$ ....(1)

Geometric Mean, $G=(x_1 \cdot x_2 \ldots\ldots\ldots x_n)^{1/n}$ ....(2)

Harmonic Mean (HM), $H=\dfrac{n}{\dfrac{1}{x_1}+\dfrac{1}{x_2}+\ldots\ldots\ldots+\dfrac{1}{x_n}}$ ....(3)

Taking log in both sides of equation (2)

$\log G = \dfrac{1}{n}\log(x_1 \cdot x_2 \ldots\ldots\ldots x_n)$

$= 1/n(\log x_1 + \log x_2 + \ldots\ldots\ldots + \log x_n)$

$= 1/n \sum \log x_i$

$= 1/n \sum \log\left(\dfrac{\sum x_i}{n} - \dfrac{\sum x_i}{n} + x_i\right)$

$= 1/n \sum \log(A - A + x_i)$ [From equation (1)]

$= 1/n \sum \log A\left(1 - \dfrac{A - x_i}{A}\right)$

$= 1/n \sum \log A + 1/n \sum \log\left(1 - \dfrac{A - x_i}{A}\right)$

$= 1/n \sum \log A + 1/n \sum \left\{-\dfrac{A - x_i}{A} - \dfrac{1}{2}\left(\dfrac{A - x_i}{A}\right)^2 - \dfrac{1}{3}\left(\dfrac{A - x_i}{A}\right)^3 - \ldots\ldots\ldots\right\}$

$$= \frac{1}{n} n \log A - \frac{1}{n}\sum\left\{\frac{A-x_i}{A}+\frac{1}{2}\left(\frac{A-x_i}{A}\right)^2+\frac{1}{3}\left(\frac{A-x_i}{A}\right)^3+\dots\dots\right\}$$

$= \log A -$ a positive quantity

$\Rightarrow \log G \le \log A$

$\Rightarrow A \ge G$

$\Rightarrow$ A.M. $\ge$ G.M. ....(4)

Now using the relation A.M. $\ge$ G.M we have

$$\frac{x_1+x_2+\dots\dots+x_n}{n} \ge \left(x_1+x_2+\dots\dots+x_n\right)^{1/n}$$

Replacing $x_1, x_2, \dots\dots x_n$ by $\frac{1}{x_1}, \frac{1}{x_2}, \dots\dots \frac{1}{x_n}$

$$\frac{\frac{1}{x_1}+\frac{1}{x_2}+\dots\dots+\frac{1}{x_n}}{n} \ge \left(\frac{1}{x_1}\cdot\frac{1}{x_2}\dots\dots\frac{1}{x_n}\right)^{1/n}$$

$\Rightarrow \frac{1}{H} \ge \frac{1}{G}$

$\Rightarrow G \ge H$

$\Rightarrow GM \ge HM$ ....(5)

From eq^n (4) & (5)

$\Rightarrow AM \ge GM \ge HM.$

Theorem: For three non-zero positive values $AM \ge GM \ge HM$.

Let, x_1, x_2, x_3 be the three non zero positive values

$\therefore$ Arithmetic Mean, AM= $\frac{x_1+x_2+x_3}{3}$ ....(I)

Geometric Mean, GM= $(x_1 \cdot x_2 \cdot x_3)^{1/3}$ ....(II)

Harmonic Mean, HM= $\dfrac{3}{\dfrac{1}{x_1}+\dfrac{1}{x_2}+\dfrac{1}{x_3}}$ ....(III)

The squares of a real number can not be negative

$$\therefore \left(x_1^{1/3}-x_2^{1/3}\right)^2+\left(x_2^{1/3}-x_3^{1/3}\right)^2+\left(x_3^{1/3}-x_1^{1/3}\right)^2 \geq 0$$

$$\Rightarrow \frac{1}{2}\left(x_1^{1/3}+x_2^{1/3}+x_3^{1/3}\right)\left\{\left(x_1^{1/3}-x_2^{1/3}\right)^2+\left(x_2^{1/3}-x_3^{1/3}\right)^2+\left(x_3^{1/3}-x_1^{1/3}\right)^2\right\} \geq 0$$

$$\Rightarrow \left(x_1^{1/3}\right)^3+\left(x_2^{1/3}\right)^3+\left(x_3^{1/3}\right)^3-3x_1^{1/3}\cdot x_2^{1/3}\cdot x_3^{1/3} \geq 0$$

$$\Rightarrow x_1+x_2+x_3 \geq 3\left(x_1\cdot x_2\cdot x_3\right)^{1/3}$$

$$\Rightarrow \frac{x_1+x_2+x_3}{3} \geq \left(x_1\cdot x_2\cdot x_3\right)^{1/3}$$

$$\Rightarrow AM \geq GM \ldots\ldots\ldots\ldots(IV)$$

Now using the relation AM $\geq$ GM,

$$\therefore \frac{x_1+x_2+x_3}{3} \geq \left(x_1\cdot x_2\cdot x_3\right)^{1/3}$$

Replacing x_1, x_2, x_3 by $\dfrac{1}{x_1}, \dfrac{1}{x_2}, \dfrac{1}{x_3}$

$$\frac{\frac{1}{x_1}+\frac{1}{x_2}+\frac{1}{x_3}}{3} \geq \left(\frac{1}{x_1}\cdot\frac{1}{x_2}\cdot\frac{1}{x_3}\right)^{1/3}$$

$$\Rightarrow \frac{1}{HM} \geq \frac{1}{GM} \qquad \left[\text{From eq}^n \text{ (II) \& (III)}\right]$$

$$\Rightarrow GM \geq HM \ldots\ldots\ldots\ldots(V)$$

From eq^n (IV) & (V); AM $\geq GM \geq HM$.

Theorem: For four non-zero positive values, $AM \geq GM \geq HM$

Proof: Let, x_1, x_2, x_3, x_4 be four non-zero positive values.

$\therefore$ *Arithmetic* Mean, AM$=\frac{x_1+x_2+x_3+x_4}{4}$....(1)

$GM=(x_1x_2x_3x_4)^{1/4}$............(2)

$HM=\frac{4}{\frac{1}{x_1}+\frac{1}{x_2}+\frac{1}{x_3}+\frac{1}{x_4}}$...........(3)

The squares of a real number can not be negative

$(x_1-x_2)^2\geq 0$

$\Rightarrow (x_1+x_2)^2-4x_1x_2\geq 0$

$\Rightarrow (x_1+x_2)\geq 2\sqrt{x_1x_2}$

$\Rightarrow \frac{(x_1+x_2)}{2}\geq (x_1x_2)^{1/2}$.............(1)

Similarly, $\frac{(x_3+x_4)}{2}\geq (x_3x_4)^{1/2}$..........(2)

Again let, $y_1=\frac{x_1+x_2}{2}$ *and* $y_2=\frac{x_3+x_4}{2}$

From eq^n (1)

$\Rightarrow \frac{(y_1+y_2)}{2}\geq (y_1y_2)^{1/2}$

$\Rightarrow \frac{\left(\frac{x_1+x_2}{2}+\frac{x_3+x_4}{2}\right)}{2}\geq \left(\frac{x_1+x_2}{2}\right)^{1/2}\cdot\left(\frac{x_3+x_4}{2}\right)^{1/2}$

$\Rightarrow \frac{x_1+x_2+x_3+x_4}{4}\geq \left(\frac{x_1+x_2}{2}\right)^{1/2}\cdot\left(\frac{x_3+x_4}{2}\right)^{1/2}$............(3)

Taking root square both side of eq^n (1) & (2)

$$\Rightarrow \left(\frac{x_1+x_2}{2}\right)^{1/2} \geq (x_1x_2)^{1/4} \text{ and } \left(\frac{x_3+x_4}{2}\right)^{1/2} \geq (x_3x_4)^{1/4}$$

$$\therefore \left(\frac{x_1+x_2}{2}\right)^{1/2} \cdot \left(\frac{x_3+x_4}{2}\right)^{1/2} \geq (x_1x_2x_3x_4)^{1/4} \text{.......}(4)$$

$eqn(3) \& (4)$

$$\Rightarrow \frac{x_1+x_2+x_3+x_4}{4} \geq (x_1x_2x_3x_4)^{1/4}$$

$$\Rightarrow AM \geq GM \text{...........}(5)$$

Replacing x_1, x_2, x_3, x_4 by $\frac{1}{x_1}, \frac{1}{x_2}, \frac{1}{x_3}, \frac{1}{x_4}$

$$\therefore \frac{\frac{1}{x_1}+\frac{1}{x_2}+\frac{1}{x_3}+\frac{1}{x_4}}{4} \geq \left(\frac{1}{x_1}\cdot\frac{1}{x_2}\cdot\frac{1}{x_3}\cdot\frac{1}{x_4}\right)^{1/4}$$

$$\Rightarrow \frac{1}{HM} \geq \frac{1}{GM}$$

$$\Rightarrow GM \geq HM.$$

$$\therefore AM \geq GM \geq HM.$$

For two non-zero positive values, Prove that AM X HM=GM^2 $\Rightarrow GM = \sqrt{AM \times HM}$

Proof: Let x_1 and x_2 be the two positive values.

Arithmetic Mean, $\frac{x_1+x_2}{2}$

Geometric Mean, GM=$(x_1x_2)^{1/2}$

Harmonic Mean, HM=$\frac{2}{\frac{1}{x_1}+\frac{1}{x_1}}$

$$\therefore\ AM\times HM=\frac{x_1+x_2}{2}\times\frac{2}{\frac{1}{x_1}+\frac{1}{x_1}}$$

$$=\frac{x_1+x_2}{2}\frac{2x_1x_2}{x_1+x_2}$$

$$=x_1x_2$$

$$=\left(\sqrt{x_1x_2}\right)^2$$

$$=GM^2$$

$$AM\times HM=GM^2.$$

Theorem: A variable 'x' takes on n values which are in greater grogrssion obtain AM,GM & HM and show that AM X HM= GM²

Proof: Let the variable of x a, ar, ar^2, ...ar^{n-1} which are in geometric progression of n.

By definition, Arithmetic Mean, $AM=\frac{a+ar+ar^2+......+ar^{n-1}}{n}$

$$=\frac{a}{n}\left(1+r+r^2+.......r^{n-1}\right)$$

$$\therefore AM=\frac{a\left(1-r^n\right)}{n(1-r)}$$

Geometric Mean, $GM=\left(a.ar.ar^2................ar^{n-1}\right)^{1/n}$

$$=a^{n\times\frac{1}{n}}\left(1.r.r^2............r^{n-1}\right)^{1/n}$$

$$=a\left\{r^{1+2+..............+(n-1)}\right\}^{1/n}$$

$$=a\left(r\frac{n(n-1)}{2}\right)^{1/n}=ar^{\frac{(n-1)}{2}}$$

$$\Rightarrow GM^2=a^2r^{n-1}$$

Harmonic Mean, $HM=\dfrac{n}{\dfrac{1}{a}+\dfrac{1}{ar}+\dfrac{1}{ar^2}+\dots\dots\dots+\dfrac{1}{ar^{n-1}}}$

$$=\frac{n}{\frac{1}{a}\left(1+\frac{1}{r}+\frac{1}{r^2}+\dots\dots\dots+\frac{1}{r^{n-1}}\right)}$$

$$HM=\frac{na}{\dfrac{1-\left(\frac{1}{r}\right)^n}{1-\frac{1}{r}}}=\frac{na\times r^n(r-1)}{r(r^n-1)}$$

$$=\frac{na}{\dfrac{\frac{r^n-1}{r^n}}{\frac{r-1}{r}}}$$

$$AM\times HM=\frac{a(1-r^n)}{n(1-r)}\times\frac{nar^n(r-1)}{r(r^n-1)}$$

$$=a^2r^{n-1}$$

$$=GM^2$$

$$\therefore AM\times HM=GM^2.$$

What is the condition it will be true, AM=GM=HM

Let x_1 and x_2 be the two non-zero positive value.

Hence

Arithmetic Mean, $AM=\dfrac{x_1+x_2}{2}$

Geometric Mean, $GM=\sqrt{x_1x_2}$

Harmonic Mean, $HM=\dfrac{2}{\dfrac{1}{x_1}+\dfrac{1}{x_2}}$

Now it will be AM=GM,

$$When, \frac{x_1+x_2}{2}=\sqrt{x_1x_2}$$

$$\Rightarrow x_1+x_2=2\sqrt{x_1x_2}$$

$$\Rightarrow (x_1+x_2)^2=4x_1x_2$$

$$\Rightarrow (x_1+x_2)^2-4x_1x_2=0$$

$$\Rightarrow (x_1-x_2)^2=0$$

$$\Rightarrow x_1=x_2$$

It will be AM=GM when, $x_1=x_2$

Again it will be GM=HM

$$When, \sqrt{x_1x_2}=\frac{2}{\frac{1}{x_1}+\frac{1}{x_2}}$$

$$\Rightarrow \sqrt{x_1x_2}=\frac{2x_1x_2}{x_1+x_2}$$

$$\Rightarrow x_1+x_2=2\sqrt{x_1x_2}$$

$$\Rightarrow x_1+x_2-2\sqrt{x_1x_2}=0$$

$$\Rightarrow \left(\sqrt{x_1}-\sqrt{x_2}\right)^2=0$$

$$\Rightarrow \sqrt{x_1}=\sqrt{x_2}$$

$$\Rightarrow x_1=x_2$$

It will be AM=GM=HM When $x_1=x_2$

Median & Mode is preferable to AM:

1. When the distribution is highly asymmetrical.
2. When there is an open class interval at one or both ends of the distribution.
3. When it is difficult to measure the variable numerically.

Which criteria of Measure of Central Tendency are not satisfied by

(I) AM (II) Median. (III) Mode.

AM: It is least affected by extreme values.

Median:

1. It is not based on all observations.
2. It is not suitable for mathematical treatment.

Mode:

1. It is not based on all observations.
2. It is not suitable for mathematical treatment.
3. It is more affected by sampling fluctuations than Mean.

GM:

1. It is not easy to understand.
2. It is not easy to calculate.

HM:

1. It is not easy understand.
2. It is not easy to calculate.

Which measure is the best and why?

Arithmetic Mean is the best

Arithmetic Mean posses all characteristics of ideal measure of central tendency.

Such characteristics are not found in other measure of Central tendency.

The following criteria are not satisfied by

AM→It is less based on all the observations.

Median→

1. It is not based on all the observations
2. It is suitable for mathematical treatment.

Mode→

1. It is not based on all the observations

2. It is suitable for mathematical treatment.
3. It is affected by sampling fluctuations.

When A.M., G.M. H.M. Median and Mode cannot be calculated:

GM:

1. In case of being zero of one or more values in series.

HM:

1. In case of unknown of all the values of series.
2. In case of being zero of one or more values.

AM:

1. In case of multiplicative data.
2. Not determined of observations.
3. By graph.
4. Unknown of one or more values.

Median:

1. Increase of irregular distribution.

Mode:

1. In case of small distribution.

Graphical Method of determining Median:

The median can be determined with the help of graphs. Two methods of determining median graphically are available. The first one is with the help of cumulative frequency or ogive curve, and the second one gatton's method.

Determining median from Ogive curve:

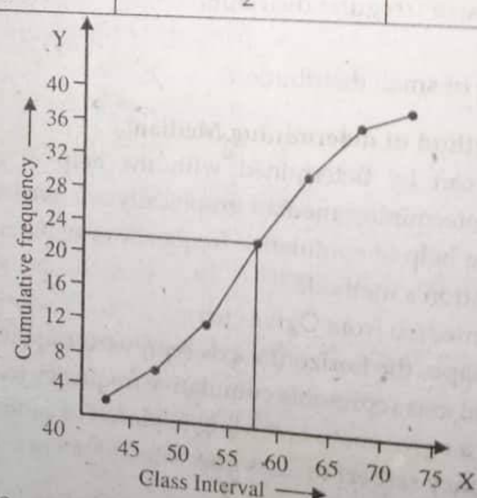
1. In a graph, the horizontal axis represents class interval and vertical axis represents cumulative frequency (c. f.). Draw a ogive curve according to placed of cumulative frequency respect to class interval.
2. The median value is determined and Median is $\frac{n}{2}$ th term.
3. The median value is marked on vertical axis and a straight line is drawn from the median point parallel to X-axis to

intersect the ogive.

4. Then the perpendicular line is drawn from the point of intersection to X-axis.
5. The point at which this perpendicular line touches the x-axis indicated the Median.

Example:

Class Interval	Frequency	C.F
40-45	2	2
45-50	4	6
50-55	6	12
55-60	10	22
60-65	8	30
65-70	6	36
70-75	2	38



Here $\frac{n}{2} = \frac{38}{2} = 19$; Vertical line intersect to the x-axis to the point of 58.5

∴ Median=58.5

Graphical Method of determining Mode:

In a frequency distribution, Mode can be determined by graphically /Histogram.

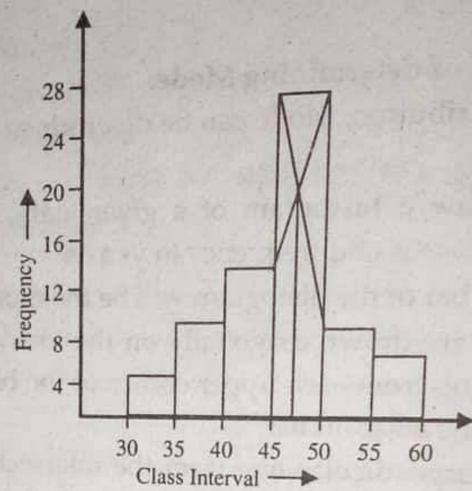
- At first draw a histogram of a given data in where class interval in x-axis and frequency in y- axis.
- The tallest bar of the histogram will be the modal value.
- Two lines are drawn diagonally on the inside of the modal class starting from each upper corner of the bar to the upper corner of the adjacent bar.
- Draw a perpendicular line from the intersection of the two diagonal lines to x-axis which gives the modal value.

Example:

Class Interval	Frequency
30-35	3
35-40	10
40-45	15
45-50	28
50-55	8
55-60	6

The perpendicular line from the intersection of the two diagonal lines to x-axis is 47

∴ Modal Value= 47



Arithmetic Mean:

Merits:

1. It is easy to understand
2. It is easy to calculate.
3. It is correctly and rigidly defined.
4. It is based on all observations.
5. It is suitable for mathematical treatment.
6. It is affected as little as possible by a sampling fluctuation.
7. Can be determined incase of being zero of grouped value.

Demerits:

1. It is affected by extreme values.
2. It is not easy to determine by observations as Median and Mode.
3. Not determined by graph as Median and Mode.

Geometric Mean:

Merits:

1. Correctly and rigidly defined.
2. Based on all observations.
3. Can be applied in mathematical treatment.
4. It is not affected by sampling fluctuations.

5. It is less affected by extreme values.

Demerits:

1. It is not easy to understand.
2. It is not easy to calculate.
3. Not determined in case being zero of one or two values of series.
4. Difficult to calculate because of logarithm.

Harmonic Mean:

Merits:

1. It is based on all the observations
2. It is suitable for mathematical treatment.
3. It is not less affected by extreme values than AM.

Demerits:

1. It is not easy to understand.
2. Not determined in case being zero of one or two values.

Mean:

Merits:

1. It is easy to understand and to calculate.
2. It is rigidly and correctly defined.
3. It is not affected by extreme values.
4. It is less affected by sampling fluctuations.
5. It is determined by graph.

Demerits:

1. It is not based on all observations.
2. It is not suitable for mathematical treatment.
3. To determine can be arranged to the order of magnitude.

Mode:

Merits:

1. Easy to understand and calculate.
2. It is not affected by extreme values.
3. Can be determined by graphs.

Demerits:

1. Not based on all observations.
2. Not suitable for mathematical treatment.
3. Sampling fluctuation is more than Mean.

Problem- 01:

AM and GM of two Numbers are 10 and 8 respectively. Find two numbers.

Let two numbers are a and b; a>b

$\frac{a+b}{2}=10 \qquad \Rightarrow a+b=20 \ldots\ldots\ldots\ldots(1)$

$\sqrt{a \times b}=8$ or $ab=64 \ldots\ldots\ldots\ldots\ldots(2)$

$$(a-b)^2=(a+b)^2-4ab$$
$$=20^2-4 \cdot 64$$
$$=144$$
$$\Rightarrow a-b=12 \ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$
$$(1)+(3) \Rightarrow$$
$$2a=32$$
$$or, a=16$$
$$(2) \Rightarrow b=\frac{64}{16}=4$$

$\therefore$ a=16 and b=4

Problem-02: Calculate the A.M. of the following data:

Marks	No. of student
0-10	12
10-20	18
20-30	27
30-40	20
40-50	17
50-60	6

Sol^n:

Marks	No. of students(f_i)	Mid value of the class (x_i)	f_ix_i
0-10	12	5	60
10-20	18	15	270
20-30	27	25	675
30-40	20	35	700
40-50	17	45	765
50-60	6	55	330
	$\sum f_i = 100$		$\sum f_i x_i = 2800$

$$\therefore A.M, \overline{x} = \frac{\sum_{i=1} f_i x_i}{\sum f_i} = \frac{2800}{100} = 28 \underline{Ans.}$$

Problem 03: Calculate G.M & H.M from the following data.

Class Interval	Frequency	Class Mid value (x_i)	log x_i	f_i logx_i
0-10	15	5	0.698	10.4845
10-20	20	15	1.1761	23.521
20-30	30	25	1.39794	41.9382
30-40	15	35	1.5441	23.1615
40-50	10	45	1.6532	16.5321
	N=90			$\sum f_i \log x_i =$ 115.6377

$$\therefore \text{GM=Anti log} \frac{\sum f_i \log x_i}{N}$$

$$= \text{Anti log} \frac{115.6377}{90} = \boxed{19} \quad \underline{\text{Ans.}}$$

HM:

Class interval	Mid value (x_i)	f_i	$\frac{f_i}{x_i}$
0-10	5	15	3
10-20	15	20	1.33
20-30	25	30	1.2
30-40	35	15	0.43
40-50	45	10	0.22
	N=90		$\sum\frac{f_i}{x_i}=6.18$

$$\therefore \text{HM}=\frac{n}{\sum\frac{f_i}{x_i}}=\frac{90}{6.18}=\boxed{14.56}\quad \underline{\underline{\text{Ans.}}}$$

Relationship among AM, HM, GM:

AM>GM>HM

Problem 04: AM and H.M of two numbers are given respectively. Calculate G.M.

Solution:

Given, AM=9, HM=4

We know,

$$\text{AM}\cdot\text{HM}=(\text{GM})^2$$

$$\Rightarrow 9\times 4=(\text{GM})^2$$

$$\Rightarrow \text{GM}=6\quad \underline{\underline{\text{Ans.}}}$$

Problem 05: AM and HM of two numbers are 5 and 1.8 respectively, Calculate two numbers:

Let two numbers are a and b; a>b

$$\frac{a+b}{2}=5$$

$$\Rightarrow a+b=10\text{..........}(I)$$

AM, $\dfrac{2}{\frac{1}{a}+\frac{1}{b}}=1.8\text{.............}(II)$

$$(II)\Rightarrow \frac{2}{\frac{1}{a}+\frac{1}{b}}=1.8$$

$$\Rightarrow \frac{2ab}{a+b}=1.8$$

$$\Rightarrow ab=\frac{1.8\times 10}{2}=9\text{..........}(III)$$

$$(a-b)^2=(a+b)^2-4ab$$

$$=10^2-4\times 9$$

$$=64$$

$$\Rightarrow a-b=8\text{.............}(III)$$

$$2a=18$$

$$\Rightarrow a=9$$

$$\therefore b=1$$

$$\therefore a=9 \text{ and } b=1$$

Problem 06: Calculate Mean, Median, and Mode from the data.

Class Interval	Frequencies
10-20	5
20-30	7
30-40	5
40-50	10
50-60	6

Solution:

Class Interval	Frequency (f_i)	Mid value (x_i)	$f_i x_i$	C.f.
10-20	5	15	75	5
20-30	7	25	175	12
30-40	12	35	420	24
40-50	10	45	450	34
50-60	6	55	330	40
	$\sum f_i = n = 40$		$\sum f_i x_i = 1450$	

$\therefore$ Mean, AM$=\dfrac{\sum f_i x_i}{\sum f_i = n} = \dfrac{1450}{40} = \boxed{36.25}$

$$Median, Me = L_1 + \frac{n/2 - f_c}{f_m} \times h$$

$\dfrac{n}{2} = \dfrac{40}{2} = 20$

$L_1 = 30$ [*Median* lies in the class 30-40]

$F_c = 12$

$f_m = 12$

$h = 10$

$\therefore Me = 30 + \dfrac{20-12}{12} \times 10$

$= 36.67$

Mode,

$M_0 = L_0 + \dfrac{\Delta_1}{\Delta_1 + \Delta_2} \times h$

$= 30 + \dfrac{5}{5+2} \times 10$

$= 37.14$

Mode lies in the class 30-40

$L_0 = 30$

$\Delta_1 = 12-7 = 5$

$\Delta_2 = 12-10 = 2$

$h = 10$

[Mode=3 Median-2 Mean]

Problem 07: In a examination of 675 candidates the examiner supplied the following informations:

Marks obtained	No. of Candidate
Less than 10%	7
Less than 20%	32
Less than 30%	95
Less than 40%	201
Less than 50%	381
Less than 60%	545
Less than 70%	631
Less than 80%	675

Sol^n:

Marks obtained	No. of Candidate	Mid value (x_i)	$f_i x_i$	C.f.
0-10	7	5	35	7
10-20	32	15	480	39
20-30	95	25	2375	134
30-40	201	35	7035	335
40-50	381	45	17145	716
50-60	545	55	29975	1261
60-70	631	65	41015	1892
70-80	675	75	50625	2567
	N = 2567		$\sum f_i x_i = 568685$	

$$\therefore \text{Mean, AM}, \bar{x} = \frac{\sum f_i x_i}{\sum f_i = n} = \frac{568685}{2567} = \boxed{221.53}$$

Problem 07: In a examination of 675 candidates the examiner supplied the following informations:

Marks obtained	No. of Candidate
Less than 10%	7
Less than 20%	32
Less than 30%	95
Less than 40%	201
Less than 50%	381
Less than 60%	545
Less than 70%	631
Less than 80%	675

Sol^n:

Marks obtained	No. of Candidate	Mid value (x_i)	$f_i x_i$	C.f.
0-10	7	5	35	7
10-20	32	15	480	39
20-30	95	25	2375	134
30-40	201	35	7035	335
40-50	381	45	17145	716
50-60	545	55	29975	1261
60-70	631	65	41015	1892
70-80	675	75	50625	2567
	N = 2567		$\sum f_i x_i = 568685$	

$$\therefore \text{Mean, AM}, \bar{x} = \frac{\sum f_i x_i}{\sum f_i = n} = \frac{568685}{2567} = \boxed{221.53}$$

...ch as p...
p_n are the base year and...
compute a quotient p_n/p_o, known as *price relative*
...tient by the value of the item obtained as a product
$p_n \times q_n$), where q_n is the current year quantity of the
weighted arithmetic average of the price relatives I_{0n} is
as

$$I_{0n} = \frac{\sum p_n q_n \left(\frac{p_n}{p_0} \right)}{\sum p_n q_n} \times 100 \qquad (3.9b)$$

where the sum extends over all commodities covered in the computation process. Thus an index number of 145 will imply that there has been an increase of 45 percent on the average in the cost of living over the period 1980-1990 for the family for which this index has been obtained.

3.7 SOME USEFUL PROPERTIES OF ARITHMETIC MEAN

Property 1: *The algebraic sum of the deviations of the observations from their mean is zero. Symbolically,*

$$\Sigma(x_i - \bar{x}) = 0 \qquad (3.10)$$

Proof. The expression (3.10) can be expanded as follows:

$$\Sigma(x_i - \bar{x}) = (x_1 - \bar{x}) + (x_2 - \bar{x}) + (x_3 - \bar{x}) + \dots\dots\dots + (x_n - \bar{x})$$

$= (x_1 + x_2 + x_3 + \ldots\ldots x_n) - (\bar{x} + \bar{x} + \bar{x} + \ldots\ldots + \bar{x})$
$= \sum x_i - n\bar{x} = n\bar{x} - n\bar{x} = 0.$

Hence the proof.

Example 3.6: Show that for the values 7, 9, 4, 8, 2,. $\sum(x_i - \bar{x}) = 0$

Solution: Here x_1=7, x_2=9, x_3=4, x_4=8, x_5=2 and $\bar{x}$ =6.

Hence,

$$\sum(x_i - \bar{x}) = (x_1 - \bar{x}) + (x_2 - \bar{x}) + (x_3 - \bar{x})(x_4 - \bar{x}) + (x_5 - \bar{x})$$
$$= (7-6)+(9-6)+(4-6)+(8-6)+(2-6)$$
$$=0$$

Property 2: *If a and b are constants such that $x = a \pm by$, where x and y are two variables assuming values x_1, x_2, ...,x_n and y_1, y_2....................,y_n respectively. Prove that $\bar{x} = a \pm b\bar{y}$*

Proof: Since $x = a \pm by$,

$$\sum x = \sum (a \pm by)$$
$$= (a \pm by_1) + (a \pm by_2) + \ldots + (a \pm by_n)$$
$$= (a + a + \ldots + a) \pm b(y_1 + y_2 + \ldots + y_n)$$
$$= na \pm b\sum y_i$$

Dividing both sides by n, we have

$\frac{\sum x_i}{n} = a \pm \frac{b\sum y_i}{n}$, from which we get

$$\bar{x} = a \pm b\bar{y}$$

Hence the proof

Example 3.7: Suppose X and Y are two variables connected by the equation $x = a + by$. Now, if the mean of y is known, the mean of x can be calculated using the *property* 2. For example, let y assume the values 2, 5, 6, 9,3, whose mean is $\bar{y} = 5$. Now if the relation between x and y is, x=2+4y then $\bar{x} = 2 + 4\bar{y}$ = 2+4×5=22. To verify this, let us find the values of x for the above values of y using the relation

$x=2+4y$. This gives the values of x as 10, 22, 26, 38, 14. The mean value of x is thus $(10+22+26+38+14)/5=22$. This completes the proof.

Property 3: *If* $u = x \pm y$, *prove that* $\overline{u} = \overline{x} \pm \overline{y}$

Proof: Since $u = x \pm y$,

$$\begin{aligned}\Sigma u_i &= \Sigma(x_i + y_i)\\ &= (x_1 \pm y_1) + (x_2 \pm y_2) + \ldots + (x_n \pm y_n)\\ &= (x_1 + x_2 + \ldots + x_n) \pm (y_1 + y_2 + \ldots + y_n)\\ &= \Sigma x_i \pm \Sigma y_i\end{aligned}$$

Dividing both sides by n, $\overline{u} = \overline{x} \pm \overline{y}$

More generally, we state the above result as saying that the mean of the sum of two or more variables is equal to the sum of their means.

Property 4: *If a set of m observations $x_1, x_2, \ldots, x_m$ and that of n observations $y_1, y_2, \ldots, y_n$ have means $\overline{x}$ and $\overline{y}$ respectively, then the combined mean $\overline{z}$ of $(m + n)$ observations of the two sets is*

$$\overline{z} = \frac{m\overline{x} + n\overline{y}}{m+n} \qquad (3.11)$$

Proof: By definition, the mean $\overline{x}$ of the first set is

$$\overline{x} = \frac{x_1 + x_2 + \ldots\ldots + x_m}{m} = \frac{\sum_i x_i}{m}$$

which gives

$$\sum_i x_i = m\overline{x} \qquad (3.11a)$$

Similarly, the mean $\overline{y}$ of the second set is

$$\overline{y} = \frac{y_1 + y_2 + \ldots\ldots + y_n}{n} = \frac{\sum_i y_i}{n}$$

from which

$$\sum_i y_i = n\overline{y} \qquad (3.11b)$$

Adding (3.11a) and (3.11b)

$$\sum_i x_i + \sum_i y_i = m\bar{x} + n\bar{y} \qquad (3.11c)$$

The combined set consists of the following $m + n$ observations

$x_1, x_2, \ldots\ldots\ldots\ldots x_m,\ y_1, y_2, \ldots\ldots\ldots\ldots\ldots y_n.$

Thus if $\bar{z}$ stands for the mean of the combined set,

$$\bar{z} = \frac{[x_1 + x_2 + \ldots.. + x_m] + [y_1 + y_2 + \ldots.. + y_n]}{m+n}$$

$$= \frac{\sum_i x_i + \sum_i y_i}{m+n} = \frac{m\bar{x} + n\bar{y}}{m+n} \qquad \text{[From3.1 lc]}$$

Hence the proof.

The result can be extended for k sets of observations, in which case the above expression assumes the form

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + \ldots\ldots + n_k\bar{x}_k}{n_1 + n_2 + \ldots\ldots.. + n_k} = \frac{\sum n_i\bar{x}_i}{\sum n_i} \qquad (3.11d)$$

where $\bar{x}_i$ (i=1, 2, ..k) is the mean of the ith set comprising n_i observations.

Example 3.8: If 20 students in a statistics course in one section receives an average score of 67, the 18 students in a second section receives 57, and the 12 students in the third section receives 62, then the overall score of these 20+18+12=50 students is computed using 3.11d as follows:

$$\bar{x} = \frac{20\times67 + 18\times57 + 12\times62}{20+18+12} = \frac{3110}{50} = 62.2$$

3.8 THE MEDIAN

Since an average is often referred to as a measure of central tendency, it should somehow reflect the *centre* of a series of observations. An arithmetic mean is very sensitive to the influence of a few large values in a collection of a predominantly small numbers; thus it sometimes fails to indicate the centre of a series of data. This is sometimes hence are not suitable for economic data. Harmonic mean can not be calculated when one or more of the observations are zero.

We will now prove an important relationship connecting arithmetic mean, geometric mean and harmonic mean in the form of a theorem.

Theorem 3.1: *For a set of n non-zero positive values* $x_1, x_2, \ldots x_n$, *prove that* $A \geq G \geq H$, where A, G and H are respectively the arithmetic mean, geometric mean and harmonic mean of the values.

Proof: By definition

$$A = \frac{\sum_{i=1}^{n} x_i}{n} \text{ and } G = (x_1\, x_2 \ldots x_n)^{\frac{1}{n}}$$

Taking logarithm of G

$$\log G = \frac{1}{n}\log(x_1 x_2 \ldots\ldots x_2) = \frac{1}{n}\sum_{i=1}^{n}\log x_i$$

$$= \frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{\sum_{i=1}^{n} x_i}{n} - \frac{\sum_{i=1}^{n} x_i}{n} + x_i\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\log(A - A + x_i)$$

$$= \frac{1}{n}\sum \log A\left(1 - \frac{A - x_i}{A}\right)$$

$$= \frac{1}{n}\sum \log A + \frac{1}{n}\sum \log\left(1 - \frac{A - x_i}{A}\right)$$

$$= \log A + \frac{1}{n}\sum\left\{-\frac{A - x_i}{A} - \frac{1}{2}\left(\frac{A - x_i}{A}\right)^2 - \frac{1}{3}\left(\frac{A - x_i}{A}\right)^3 - \ldots\right\} {}^{**}$$

** Note: The expansion is valid only when $(A - x_i)/A < 1$. This is true for almost all distributions excepting very skewed distributions.

$$= \log A - \frac{1}{n}\sum\left\{\frac{A-x_i}{A} + \frac{1}{2}\left(\frac{A-x_i}{A}\right)^2 + \frac{1}{3}\left(\frac{A-x_i}{A}\right)^3 + \ldots\ldots\right\}$$

$= \log A -$ a positive quantity.

Hence $A \geq G$

Using the relation $A \geq G$, we have

$$\frac{x_1 + x_2 + \ldots.. + x_n}{n} \geq (x_1\, x_2 \ldots\ldots.. x_n)^{\frac{1}{n}}$$

Replacing $x_1, x_2 .\ \ x_n$ by $1/x_1, 1/x_2, \ldots\ldots\ldots 1/x_n$ respectively

$$\frac{\frac{1}{x_1} + \frac{1}{x_2} + \ldots\ldots + \frac{1}{x_n}}{n} \geq \left(\frac{1}{x_1}\frac{1}{x_2}\ldots\ldots\frac{1}{x_n}\right)^{\frac{1}{n}}$$

That is

$$\frac{1}{H} \geq \frac{1}{G}$$

or

$$G \geq H$$

Combining (3.25) and (3.26), $A \geq G \geq H$

Hence the proof.

To establish the inequality $A \geq G \geq H$ for two values x_1 and x_2, consider first the difference $A-G$, where

$$A = \frac{x_1 + x_2}{2} \text{ and } G = \sqrt{x_1 x_2}$$

Taking the difference

$$A - G = \frac{x_1 + x_2}{2} - \sqrt{x_1 x_2}$$