3 Md. Abdul Alim 4-08-2010 Session 2009-2010

# Chapter: Measures of Central Tendency

**Central Tendency:** In a representative sample, the value of a series of data have a tendency to cluster around a certain point usually at the center of the series is usually called central tendency and its numerical measures are called the measures of central location.

**Different Measures of Central Tendency:** The following are the important measures of central tendency which are generally used in business:

- Arithmetic mean
- Geometric mean
- Harmonic Mean
- Median
- Mode

**Arithmetic mean:** The sum of observations divided by the total number of observations.

**Calculation of Arithmetic Mean-Ungrouped Data:** For ungrouped data, arithmetic mean ~~may be~~ computed by applying any of the following methods:

- Direct method
- Short-cut method

**Direct method:** The arithmetic mean, often simply referred to as mean, is the total of the values of a set of observations divided by their total number of observations. Thus, if $X_1, X_2, X_3, ..... X_N$ represent the values of $N$ items or observations, the arithmetic mean denoted by $\bar{X}$ is defined as:

$$\bar{X} = \frac{X_1 + X_2 + ........ + X_N}{N} = \frac{\sum_{i=1}^{N} X_i}{N}$$

If the subscripts are dropped, the formula becomes:

$$\bar{X} = \frac{\sum X}{N}$$

**Example:** The monthly income (in Tk's) of 10 persons working in a firm is as follows:

14870 14930 15020 14460 14750 14920 15720 15160 14680 14890

Find average monthly income.

**Solution:** Let income be denoted by X. By using calculator, we get $\sum X = 149400$

$\bar{X} = \frac{\sum X}{N} = \frac{149400}{10} = 14940$. Hence, the average monthly income Tk.14940.

**Example:** Sun Com is studying the number of minutes used monthly by clients in a particular cell phone rate plan. Random sample of 12 clients showed the following number of minutes used last month.

| | | | | | |
|---|---|---|---|---|---|
| 90 | 77 | 94 | 89 | 119 | 112 |
| 91 | 110 | 92 | 100 | 113 | 83 |

What is the arithmetic mean number of minutes used?

**Solution:** Let the minute be denoted by X. By using calculator, we have $\sum X = 1170$.

$$\bar{X} = \frac{\sum X}{N} = \frac{1170}{12} = 97.5$$

The arithmetic mean number of minutes used last month by the sample of cell phone users is 97.5 minutes.

**Calculation of Arithmetic Mean-Grouped Data:** For grouped data, arithmetic mean is computed by applying any of the following methods:

- Direct method
- Short-cut method

**Direct Method:** When direct method is used

$\overline{X} = \frac{\sum fX}{N}$ where, $X$ = mid-point of various classes, $f$ = the frequency of each class and

$N$ = the total frequency

**Example.** The following are the figures of profits earned by 1400 companies during 1999-2000.

| Profits (Tk. lakhs) | No. of companies | Profits (Tk. lakhs) | No. of companies |
|---|---|---|---|
| 200-400 | 500 | 1000-1200 | 100 |
| 400-600 | 300 | 1200-1400 | 80 |
| 600-800 | 280 | 1400-1600 | 20 |
| 800-1000 | 120 | | |

Calculate the average profits for all companies.

**Solution: Calculation of average profits**

| Profits (Tk. lakhs) | Mid-point ($X$) | No. of companies ($f$) | $fX$ |
|---|---|---|---|
| 200-400 | 300 | 500 | 150000 |
| 400-600 | 500 | 300 | 150000 |
| 600-800 | 700 | 280 | 196000 |
| 800-1000 | 900 | 120 | 108000 |
| 1000-1200 | 1100 | 100 | 110000 |
| 1200-1400 | 1300 | 80 | 104000 |
| 1400-1600 | 1500 | 20 | 30000 |
| | | $N = 1400$ | $\sum fX = 848000$ |

We know that, $\overline{X} = \frac{\sum fX}{N}$. By using calculator we get, $\overline{X} = \frac{848000}{1400} = 605.71$. So the average profit is Tk. 605.71 lakhs.

**Properties of Arithmetic Mean**

**Property 1:** The sum of the deviations of set of observations from their arithmetic mean is zero i.e., $\sum_{i=1} f_i(X_i - \overline{X}) = 0$.

**Proof:** Let $\overline{X}$ be the mean of a set of observations $X_i$ with frequencies $f_i$ then

$\sum f_i(X_i - \overline{X}) = \sum_{i=1}^{n} f_i X_i - \overline{X}\sum_{i=1}^{n} f_i = n\overline{X} - n\overline{X} = 0$ when $\sum_{i=1}^{n} f_i = n$.

**Property 2:** The arithmetic mean of a set of $N$ constant observations $A$ is $A$.

**Proof:** Let us consider $N$ constant observations which are $A$ then

$\overline{X} = \frac{\sum A}{N} = \frac{NA}{N} = A$.



Md. Moyazzem Hossain
Lecturer, Dept. of Statistics, JU

**Property 3:** The sum of the squares of the deviations of a set of observations is minimum when the deviations are taken about the arithmetic mean i. e., $\sum_{i=1}^{n} f_i(X_i - A)^2 > \sum_{i=1}^{n} f_i(X_i - \bar{X})^2$ where, $A$ is any arbitrary constant.

2000

**Proof:** Let $\bar{X}$ be the mean of a set of observations $X_i$ with frequencies $f_i$ also let $A$ be any arbitrary value, we have,

$$\sum_{i=1}^{n} f_i(X_i - A)^2 = \sum_{i=1}^{n} f_i(X_i - \bar{X} + \bar{X} - A)^2$$

$$= \sum_{i=1}^{n} f_i(X_i - \bar{X})^2 + n(\bar{X} - A)^2 + 2(\bar{X} - A)\sum_{i=1}^{n} f_i(X_i - \bar{X})$$

$$= \sum_{i=1}^{n} f_i(X_i - \bar{X})^2 + n(\bar{X} - A)^2 \quad \text{Since } \sum_{i=1}^{n} f_i(X_i - \bar{X}) = 0$$

Also, $n(\bar{X} - A)^2$ is a positive quantity so that $\sum_{i=1}^{n} f_i(X_i - A)^2 > \sum_{i=1}^{n} f_i(X_i - \bar{X})^2$. Hence proved.

**Property 4:** If $\bar{X}_i; i = 1, 2, \ldots, k$ are the means of $k$ series of sizes $n_i; i = 1, 2, \ldots, k$ respectively, then the mean $\bar{X}$ of the composite series obtained by the formula,

$$\bar{X} = \frac{n_1\bar{X}_1 + n_2\bar{X}_2 + \ldots + n_k\bar{X}_k}{n_1 + n_2 + \ldots + n_k}$$

**Proof:** Let, $X_{11}, X_{12}, \ldots, X_{1n_1}$ be $n_1$ numbers in the first series, $X_{21}, X_{22}, \ldots, X_{2n_2}$ be $n_2$ numbers in the second series, .... $X_{k1}, X_{k2}, \ldots, X_{kn_k}$ be $n_k$ numbers in the $k^{th}$ series.

Now we have, $\bar{X}_1 = \dfrac{X_{11} + X_{12} + \ldots + X_{1n_1}}{n_1}$

$$\bar{X}_2 = \frac{X_{21} + X_{22} + \ldots + X_{2n_2}}{n_2}$$

..............................

$$\bar{X}_k = \frac{X_{k1} + X_{k2} + \ldots + X_{kn_k}}{n_k}$$

The arithmetic mean $\bar{X}$ of the composite series of size $n_1 + n_2 + \ldots + n_k$ is given by

$$\bar{X} = \frac{(X_{11} + X_{12} + \ldots + X_{1n_1}) + (X_{21} + X_{22} + \ldots + X_{2n_2}) + \ldots + (X_{k1} + X_{k2} + \ldots + X_{kn_k})}{n_1 + n_2 + \ldots + n_k}$$

$$\therefore \bar{X} = \frac{n_1\bar{X}_1 + n_2\bar{X}_2 + \ldots + n_k\bar{X}_k}{n_1 + n_2 + \ldots + n_k}$$

**Arithmetic mean for two or more related groups:** If we have the arithmetic mean and number of observations two or more than two related groups, we can compute combined average of these groups by applying the following formula.

$$\overline{X_{12}} = \frac{N_1\overline{X_1} + N_2\overline{X_2}}{N_1 + N_2}$$

Where, $\overline{X_{12}}$ = Combined mean of the two groups

$\overline{X_1}$ = Arithmetic mean of the first group

$\overline{X_2}$ = Arithmetic mean of the second group

$N_1$ = Number of observations in the first group

$N_2$ = Number of observations in the second group



Md. Moyazzem Hossain
Lecturer, Dept. of Statistics, JU

*** If we have to find out the combined mean of three series the above formula can be extended as follows:

$$\overline{X}_{123} = \frac{N_1\overline{X}_1 + N_2\overline{X}_2 + N_3\overline{X}_3}{N_1 + N_2 + N_3}$$

**Example:** There are two branches of a company employing 100 and 80 persons respectively. If arithmetic mean of the monthly salaries paid by two branches are Tk.1570 and Tk.1750 respectively, find the arithmetic mean of the salaries of the employees of the whole company.

**Solution:** We should compute the combined mean. The formula is

$$\overline{X}_{12} = \frac{N_1\overline{X}_1 + N_2\overline{X}_2}{N_1 + N_2}$$

Given that, $N_1=100$, $\overline{X}_1=1570$, $N_2=80$ and $\overline{X}_2=1750$

$$\overline{X}_{12} = \frac{(100\times1570+80\times1750)}{100+80} = \frac{297000}{180} = 1650$$

**Merits of Arithmetic Mean**

- It should be easy to understand.
- It should be easy to compute.
- It should be based on all the observations.
- It should be rigidly defined.
- It should be capable of further algebraic treatment.
- It should have sampling stability.

অযৌক্তিক

**Limitations of Arithmetic Mean:** Arithmetic mean is used unduly affected by the presence of extreme values. Also in open-end frequency distribution it is difficult to compute mean without making assumption regarding the size of the class interval of the open-end classes.

**Weighted Arithmetic Mean:** One of the limitations of the arithmetic mean is that it gives equal importance to all the observations. But there are cases where the relative importance of all the different observations is not the same. When that is so, we compute weighted arithmetic mean.

The terms 'weight' stands for the relative importance of the different observations. The formula for computing weighted arithmetic mean is

$$\overline{X}_w = \frac{\sum WX}{\sum W}$$

where, $\overline{X}_w$ = Represents the weighted arithmetic mean

X = The variable and W = Weights attached to the variable X.

**Uses of Weighted Arithmetic Mean:** Weighted mean is especially useful in the problems relating to the construction of index numbers and standardized birth and death rates.

**Example:** The cater Construction company pays it hourly employees \$16.50, \$17.50, or \$18.50 per hour. There are 26 hourly employees, 14 are paid at the \$16.50 rate, 10 at the \$17.50 rate and 2 at the \$18.50 rate. What is the mean hourly paid the 26 employees?

**Solution:** We know that, $\overline{X}_w = \frac{\sum WX}{\sum W}$

$$\overline{X}_w = \frac{14(\$16.50)+10(\$17.50)+2(\$18.50)}{14+10+2} = \frac{\$445.00}{26} = \$17.038$$

The weighted mean hourly wage is rounded to \$17.04

**Example:** Suppose that, the nearby Wendy's Restaurant sold medium, large and Biggie-sized soft drinks for \$0.90, \$1.25 and \$1.50 respectively of the last 10 drinks sold, 3 were medium, 4 were large and 3 were Biggie-sized. Find the mean price of the last 10 drinks sold.

**Solution:** We know that, $\overline{X}_w = \frac{\sum WX}{\sum W}$, $\overline{X}_w = \frac{3(\$0.90)+4(\$1.25)+3(\$1.50)}{3+4+3} = \frac{\$12.20}{10} = \$1.22$

## Geometric Mean

**For ungrouped data:** The geometric mean of a set of $N$ non-zero positive observations is the $N$ th root of their product. Let $X_1, X_2, X_3, ...., X_N$ be non-zero positive observations in a series of data.

Thus, the geometric mean $G.M = (X_1.X_2.X_3....X_N)^{\frac{1}{N}}$.

The calculation may sometime by simplified by taking logarithm, that is

$$Log\ G.M = \frac{1}{N}[\log X_1 + \log X_2 + ... + \log X_N] = \frac{\sum_{i=1}^{N} \log X_i}{N}$$

$$G.M = Anti\log\left(\frac{\sum_{i=1}^{N} \log X_i}{N}\right)$$

**Example:** Calculate the geometric mean for the following data:

6.5 169.0 11.0 112.5 14.2 75.0 35.5 215.0

**Solution:** We know that, $G.M = Anti\log\left(\frac{\sum_{i=1}^{N} \log X_i}{N}\right)$

Let us consider, the observations is denoted by $X$. Now by using calculator we get,

$\sum \log X = 13.043$ and $N = 8$. $\therefore G.M = Anti\log\left(\frac{13.043}{8}\right) = Anti\log\ (1.6304) = 42.697$

**Example:** Calculate the geometric mean of the following price relatives:

| Commodity | Price Relatives |
|---|---|
| Wheat | 207 |
| Rice | 198 |
| Pulses | 156 |
| Sugar | 124 |
| Salt | 107 |
| Oils | 196 |

**Solution:** We know that, $G.M = Anti\log\left(\frac{\sum_{i=1}^{N} \log X_i}{N}\right)$

Let us consider, the price relatives is denoted by $X$. Now by using calculator we get,

$\sum \log X = 13.2208$ and $N = 6$. $G.M = Anti\log\left(\frac{13.2208}{6}\right) = Anti\log\ (2.2035) = 159.77$



Md. Movazzem Hossain
Lecturer, Dept. of Statistics, JU

**For grouped data:** In grouped data for calculating geometric mean first we will find the mid-points and then apply the following formula: $G.M = Anti\log\left(\frac{\sum_{i=1}^{N} f_i \log X_i}{N}\right)$ where, $X$ = mid-point.

**Example:** Find out geometric mean from the following data:

| $X$ | 10 | 20 | 30 | 40 | 50 | 60 |
|---|---|---|---|---|---|---|
| $f$ | 12 | 15 | 25 | 10 | 6 | 2 |

**Solution:** We know that, $G.M = Anti\log\left(\frac{\sum_{i=1}^{N} f_i \log X_i}{N}\right)$. From the given data, by using calculator we get, $\sum f \log X = 98.214$ and $N = 70$. $G.M = Anti\log\left(\frac{98.214}{70}\right) = Anti\log\ (1.403) = 25.29$

**Example:** Calculate geometric mean for the following distribution.

| Weight (in lbs) | Frequency |
|---|---|
| 100-104 | 24 |
| 105-109 | 30 |
| 110-114 | 45 |
| 115-119 | 65 |
| 120-124 | 72 |

**Solution: Calculation of geometric mean**

| Weight (in lbs) | Midpoint ($x$) | Frequency |
|---|---|---|
| 100-104 | 102 | 24 |
| 105-109 | 107 | 30 |
| 110-114 | 112 | 45 |
| 115-119 | 117 | 65 |
| 120-124 | 122 | 72 |
| Total | | 236 |

We know that, $G.M = Anti\log\left(\frac{\sum_{i=1}^{N} f_i \log X_i}{N}\right)$

From the given data, by using calculator we get, $\sum f \log X = 485.95$ and $N = 236$.

$$G.M = Anti\log\left(\frac{485.95}{236}\right) = Anti\log\ (2.059) = 114.55$$

**Merits and Limitations of Geometric Mean:** Geometric mean is highly useful in averaging ratios and percentages and in determining rates of increase and decrease. It is also capable of algebraic manipulation. For example, if the geometric mean of two or more series and their number of observations are known, a combined geometric mean can easily be calculated.

However, compared to arithmetic mean, this average is more difficult to compute and interpret. Also geometric mean cannot be computed when there are both negative and positive values in a series or more observations are having zero value.



Md. Moyazzem Hossain
Lecturer, Dept. of Statistics, JU

**Harmonic Mean:** The harmonic mean is based on the reciprocal of the numbers averaged. It is defined as the reciprocal of the arithmetic mean of the reciprocal of the individuals' observations.

**For ungrouped data:** The harmonic mean, $H.M = \dfrac{N}{\left(\dfrac{1}{X_1}+\dfrac{1}{X_2}+\ldots+\dfrac{1}{X_N}\right)} = \dfrac{N}{\sum\left(\dfrac{1}{X}\right)}$

i.e. the harmonic mean of a set of $N$ non zero observations $X_1, X_2, \ldots, X_N$ in a series is the reciprocal of the arithmetic mean of the reciprocals.

**Example:** Calculate the Harmonic mean of the following series of monthly expenditure of a batch of students:

TK. 1250 1300 1750 1000 1450 1500 1550 1400 1500 1150

**Solution:** We know that, the harmonic mean, $H.M = \dfrac{N}{\left(\dfrac{1}{X_1}+\dfrac{1}{X_2}+\ldots+\dfrac{1}{X_N}\right)} = \dfrac{N}{\sum\left(\dfrac{1}{X}\right)}$

Let, monthly expenditure denoted by X. By using calculator we get,

| $x$ | $\frac{1}{x}$ |
|---|---|
| 1250 | 0.0008 |
| 1300 | 0.000769 |
| 1750 | 0.000571 |
| 1000 | 0.001 |
| 1450 | 0.00069 |
| 1500 | 0.000667 |
| 1550 | 0.000645 |
| 1400 | 0.000714 |
| 1500 | 0.000667 |
| 1150 | 0.00087 |
| Total | 0.007393 |

Here, $\sum\frac{1}{X} = 0.007393$ and $N = 10$ $\therefore H.M = \dfrac{10}{0.007393} = 1352.693$

**Example:** Calculate the Harmonic mean from the following figures:

9.7 0.0009 178.7 0.874 1238 0.012 89.9 78.4 0.989 0.008

**Solution:** We know that, the harmonic mean $H.M = \dfrac{N}{\left(\dfrac{1}{X_1}+\dfrac{1}{X_2}+\ldots+\dfrac{1}{X_N}\right)} = \dfrac{N}{\sum\left(\dfrac{1}{X}\right)}$

Let, the observations denoted by X. By using calculator we get, $\sum\dfrac{1}{X} = 1321.73$ and $N = 10$

$\therefore H.M = \dfrac{10}{1321.73} = 0.008.$

**For grouped data:** For grouped data,

$H.M = \dfrac{N}{\sum\left(f \times \dfrac{1}{X}\right)}$ where, $N = \sum f$ and $X$ may be considered as the mid values of the class intervals.



Md. Moazzem Hossain
Lecturer, Dept. of Statistics, IU

Example. Calculate harmonic mean from the following data:

| Marks | Frequency |
|---|---|
| 0-10 | 5 |
| 10-20 | 10 |
| 20-30 | 7 |
| 30-40 | 3 |
| 40-50 | 2 |

**Solution:** We know that, $H.M = \dfrac{N}{\sum\left(f \times \dfrac{1}{X}\right)}$

**Calculation of harmonic mean**

| Marks | Frequency (f) | Mid-points (x) | $f \times \dfrac{1}{X}$ |
|---|---|---|---|
| 0-10 | 5 | 5 | 1 |
| 10-20 | 10 | 15 | 0.667 |
| 20-30 | 7 | 25 | 0.280 |
| 30-40 | 3 | 35 | 0.086 |
| 40-50 | 2 | 45 | 0.044 |
| Total | 27 | | 2.077 |

We get, $\sum\left(f \times \dfrac{1}{X}\right) = 2.077$ and $N = \sum f = 27$ $\quad \therefore H.M = \dfrac{27}{2.077} = 13.00061 = 13$

**Theorem 1:** For two non zero positive observations $AH = G^2$, where $A$ is the arithmetic mean, $H$ is the harmonic mean and $G$ is the geometric mean.

**Proof:** Let the two observations be $X_1$ and $X_2$ then $A = \dfrac{X_1 + X_2}{2}$, $G = (X_1 X_2)^{\frac{1}{2}}$ and

$H = \dfrac{1}{\dfrac{1}{2}\left(\dfrac{1}{X_1} + \dfrac{1}{X_2}\right)} = \dfrac{2X_1X_2}{X_1 + X_2}$. Therefore, $AH = \dfrac{(X_1 + X_2)}{2} \times \dfrac{2X_1X_2}{(X_1 + X_2)} = X_1X_2 = G^2$.

Hence, proved.

**Theorem 2:** For $n$ non-zero positive observations,

Arithmetic mean ≥ Geometric mean ≥ Harmonic mean

**Proof:** Let $X_1, X_2, \ldots, X_n$ be $n$ non-zero positive observations. Also let $A, G$ and $H$ are as the Arithmetic mean, Geometric mean and Harmonic mean and $d_i = X_i - A$.

We know, $G = (X_1 X_2 \ldots X_n)^{\frac{1}{n}}$

Taking logarithm on the both sides, we have $Log\ G = \dfrac{1}{n}\sum_{i=1}^{n} \log X_i$

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

$$= \frac{1}{n} n \log A + \frac{1}{n}\sum_{n=1}^{n} \log\left(1 + \frac{d_i}{A}\right)$$

Expanding $\log\left(1 + \dfrac{d_i}{A}\right)$ in ascending power of $\dfrac{d_i}{A}$ and avoiding 3rd or more power we have,

$$\log\left(1 + \frac{d_i}{A}\right) = \frac{d_i}{A} - \frac{\left(\frac{d_i}{A}\right)^2}{2}$$

$$\log\left(1 + \frac{d_i}{A}\right) = \frac{d_i}{A} - \frac{\left(\frac{d_i}{A}\right)^2}{2} + \frac{\left(\frac{d}{A}\right)^3}{3} \ldots$$

Md. Mo... Hossain
Lecturer, Dept. of Statistics, II

Therefore, we get, $\log G = \log A + \frac{1}{n}\sum_{i=1}^{n}\frac{d_i}{A} - \frac{1}{n}\sum_{i=1}^{n}\frac{\left(\frac{d_i}{A}\right)^2}{2} = \log A + 0 -$ a positive quantity

$$\therefore \log G \le \log A \quad or \quad A \ge G \qquad (1)$$

Again, we know,

$$H = \frac{n}{\left(\frac{1}{X_1} + \frac{1}{X_2} + \dots + \frac{1}{X_n}\right)}$$

$$\frac{1}{H} = \frac{1}{n}\left(\frac{1}{X_1} + \frac{1}{X_2} + \dots + \frac{1}{X_n}\right) \ge \left(\frac{1}{X_1}.\frac{1}{X_2}.\dots.\frac{1}{X_n}\right)^{\frac{1}{n}} \ge \frac{1}{G}$$

$$\therefore G \ge H \qquad (2)$$

Combining equation 1 and 2 we have, $A \ge G \ge H$. Hence proved.

**Median:** The median is the measure of central tendency which appears in the "middle" of an ordered sequence of values. That is, half of the observations in a set of data are lower than it and half of the observations are greater than it. As distinct from the arithmetic mean which is calculated from the value of every observation in the series, the median is what is called a positional average. The term 'position' refers to the place of a value in a series. The place of the median in a series is such that an equal number of observations lie on either side of it.

For example, if the income of five persons is Tk.1000, 1200, 1500, 1600, 1800 then the median income would be Tk. 1500. Changing any or both of the first two values with any other numbers with value 1500 or less and/or changing any of the last two values to any other values with values of 1500 and more, would not affect the value of the median which would remain 1500.

**Calculation of Median –Ungrouped Data:** The median is defined as the middle most observation when the observations arranged in order of magnitude.

Arrange the data in ascending or descending order of magnitude. (Both arrangements would give the same answer)

আরোহন

**For ungrouped data**

- When $N$ (total number of observations) is odd, the middle most observation i.e. $\frac{(N+1)}{2}th$ observation will be the median in the series.
- When $N$ is even, the median will be the arithmetic mean of $\frac{N}{2}th$ and $\left(\frac{N}{2}+1\right)th$ observations in the series.

**Example:** From the following data of wages of 7 workers, compute the median wage:

Wages (in Tk.) 1600 1650 1580 1690 1660 1606 1640

**Solution: Calculation of median**

| Sl. No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Wages arranged in ascending order | 1580 | 1600 | 1606 | 1640 | 1650 | 1660 | 1690 |



Md. Moyazzem Hossain
Lecturer, Dept. of Statistics

Median = Size of $\frac{N+1}{2}th$ observation = $\frac{7+1}{2}th$ = $4^{th}$ observation

Value of $4^{th}$ observation is 1640. Hence median wage = Tk.1640

**Example:** The flowing table gives the monthly income of 12 families in a village.

| House No | Monthly income (Tk.) | House No | Monthly income (Tk.) |
|---|---|---|---|
| 1 | 587 | 7 | 805 |
| 2 | 693 | 8 | 907 |
| 3 | 595 | 9 | 763 |
| 4 | 780 | 10 | 865 |
| 5 | 840 | 11 | 768 |
| 6 | 760 | 12 | 894 |

Calculate the median income.

**Solution:** For calculating median the data have to arrange either in ascending or descending order. Here income has been arranged in ascending order.

| House No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Monthly income (Tk.) | 587 | 595 | 693 | 760 | 763 | 768 | 780 | 805 | 840 | 865 | 894 | 907 |

We know that, when $N$ is even, the median will be the arithmetic mean of $\frac{N}{2}$ th and $\left(\frac{N}{2}+1\right)$th observations in the series. We have, $N = 12$. Hence, $\frac{N}{2}$ th observation $= \frac{12}{2} = 6^{th}$ observation and $\left(\frac{N}{2}+1\right)$ observation $= (\frac{12}{2}+1)th = 7^{th}$ observation.

So that, Median $= \frac{(6\text{th observation}+7\text{th observation})}{2} = \frac{(768+780)}{2} = 774$

Hence, the median income = Tk.774.

**Calculation of Median –Grouped Data:** Apply the following formula for determining the exact value of median:

$$\text{Median} = L + \frac{\frac{N}{2} - p.c.f}{f} \times i$$

where $L$ = lower limit of median class i.e., the class in which the middle item in the distribution lies.

$p.c.f$ = preceding cumulative frequency to the median class

$f$ = frequency of the median class and $i$ = the class interval of the median class.

**Example:** Suppose 1500 workers are working in an industrial establishment. Their age is classified as follows:

| Age (yrs.) | No. of workers | Age (yrs.) | No. of workers |
|---|---|---|---|
| 18-22 | 120 | 38-42 | 184 |
| 22-26 | 125 | 42-46 | 162 |
| 26-30 | 280 | 46-50 | 86 |
| 30-34 | 260 | 50-54 | 75 |
| 34-38 | 155 | 54-58 | 53 |

Calculate the median age.

**Solution: Calculation of median age**

| Age (yrs.) | f | c. f |
|---|---|---|
| 18-22 | 120 | 120 |
| 22-26 | 125 | 245 |
| 26-30 | 280 | 525 |
| 30-34 | 260 | 785 |
| 34-38 | 155 | 940 |
| 38-42 | 184 | 1125 |
| 42-46 | 162 | 1286 |
| 46-50 | 86 | 1372 |
| 50-54 | 75 | 1447 |
| 54-58 | 53 | 1500 |

Median = Size of $\frac{N}{2}$ th observation = $\frac{1500}{2}$ = $750th$ observation.

Hence, the median lies in the class 30-34 26-30

$$\text{Median} = L + \frac{\frac{N}{2} - p.c.f}{f} \times i = 30 + \frac{750 - 525}{260} \times 4 = 30 + 3.46 = 33.46$$

Hence, the median age is 33.46 years.

**Merits and Limitations of Median:** The median is superior to arithmetic mean in certain respects. For example, it is especially in case of open-end distribution and also it is not influenced by the presence of extreme values. In fact, when extreme values are present in a series, the median is more satisfactory measure of central tendency than the mean.

However, since median is positional average, its value is not determined by each and every observation. Also median is not capable of algebraic treatment. For example, median cannot be used for determined the combined median of two or more groups.

**Mode:** Mode is the most typical or commonly value in a set of data. For example, if we take the values of six different observations as 5, 8, 10, 8, 5, 8 mode will be 8 as it has occurred maximum number of times. i.e 3 times. Graphically, it is the value on the X-axis below the peak or highest point. of the frequency curve as we can be seen from the flowing diagram.

**Calculation of mode –Ungrouped data:** For determining mode count the numbers of items the various values repeat themselves and the value which occurs the maximum number of times is the modal value.

**Example:** The following figures relate to the preferences with regard to size of screen in inches of T.V sets of 30 persons selected at random from a locality. Find the modal size of the T.V screen.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 12 | 20 | 12 | 24 | 27 | 20 | 12 | 20 | 27 | 24 |
| 24 | 20 | 12 | 20 | 24 | 27 | 24 | 24 | 20 | 24 |
| 24 | 20 | 24 | 24 | 12 | 24 | 20 | 27 | 24 | 24 |

**Solution: Calculation of modal size**

| Size in inches | Tally | Frequency |
|---|---|---|
| 12 | ~~IIII~~ | 5 |
| 20 | ~~IIII~~ III | 8 |
| 24 | IIII IIII III | 13 |
| 27 | IIII | 4 |

Since size 24 occurs the maximum number of items, therefore, the modal size of T.V screen is 24 inches.



Md. Monazzem Hossain
Lecturer, Dept. of Statistics

**Calculation of mode- Grouped data:** In case of grouped data the following formula is used for calculating mode:

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

where, L = Lower limit of the modal class

$\Delta_1$ = The difference between the frequency of the modal class and the frequency of the pre-modal class.

$\Delta_2$ = The difference between the frequency of the modal class and the frequency of the post-modal class.

$i$ = The size of the modal class.

**Example:** The following data relate to the sales of 100 companies:

| Sales (Tk. lakhs) | No. of companies |
|---|---|
| Below 60 | 12 |
| 60-62 | 18 |
| 62-64 | 25 |
| 64-66 | 30 |
| 66-68 | 10 |
| 68-70 | 3 |
| 70-72 | 2 |

Calculate the modal sales.

**Solution:** Since the maximum frequency 30 is in the class 64-66, therefore 64-66 is the modal class.

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

Here, L = 64, $\Delta_1 = (30\text{-}25) = 5$, $\Delta_2 = (30\text{-}10) = 20$, $i = 2$

Mode $= 64 + \frac{5}{5+20} \times 2 = 64 + \frac{10}{25} = 64.4$. Hence, modal sales are Tk.64.4 lakhs.

**Locating Mode Graphically:** In a frequency distribution the value of mode can also be determined graphically. The steps in calculation are

- Draw a histogram of the given data
- Draw two lines diagonally on the inside of the modal class bar, 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 the X-axis (horizontal scale) which gives us modal value.

**Example:** The daily profits in TK's of 100 shops are given as follows:

| Profits (in Tk's) | No. of shop |
|---|---|
| 0-100 | 12 |
| 100-200 | 18 |
| 200-300 | 27 |
| 300-400 | 20 |
| 400-500 | 17 |
| 500-600 | 6 |

Draw the histogram and then find the modal value. Check this value by direct calculation.

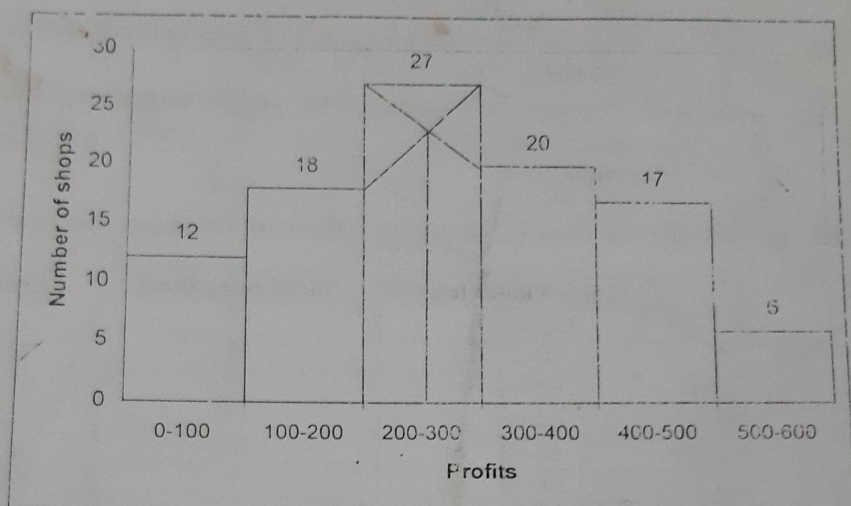


Md. Moyazzem Hossain
Lecturer, Dept. of Statistics, JU

**Solution: Direct calculation**

Mode lies in the class 200-300.

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i = 200 + \frac{9}{(9+7)} \times 100 = 256.25$$




From the above diagram, the modal value is also Tk. 256. Hence by both the methods we get the same value of mode.

## Merits and Limitations of Mode

### Merits

a) The mode is not affected by the extreme values.

b) It can be calculated from frequency distribution with open class.

c) Mode can be easily used to describe qualitative phenomenon.

### Limitations

a) It is not based on all observations.

b) It is difficult for algebraic treatments.

c) Mode is not clearly defined in case of bi-modal or multimodal distribution.

13