MIND MAP

Statistics

1. Arithmetic Mean (A.M.)

For Ungrouped Distribution

If $x_1, x_2 \dots x_n$ are 'n' values of variate then their AM ($\bar{x}$) is $\bar{x} = \dfrac{x_1 + x_2 + \dots\dots + x_n}{n}$

$$\boxed{\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \text{ OR } \sum_{i=1}^{n} x_i = n\bar{x}}$$

For Ungrouped Freq. Distribution

Values: $x_1, x_2, \dots x_n$

Frequencies: $f_1, f_2, \dots f_n$

AM of this frequency distribution is

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + \dots\dots + f_nx_n}{(f_1 + f_2 + \dots\dots + f_n)}$$

$$\boxed{\bar{x} = \frac{\sum_{i=1}^{n} f_ix_i}{\sum_{i=1}^{n} f_i} \text{ OR } \bar{x} = \frac{\sum_{i=1}^{n} f_ix_i}{N}}$$

Where, $N = \sum_{i=1}^{n} f_i$

For Grouped Frequency Distribution

1. Find the middle values of each group (or class)
2. Change it into Ungrouped Frequency Distribution.
3. Then use the same formula used for Ungrouped Frequency Distribution.

AM by Short Method (By change of origin)

In this method we take deviation (d_i) of each variate (x_i) from any central value (a).

Let, $x_i - a = d_i \Rightarrow x_i = a + d_i$

Then, $\boxed{\bar{x} = \dfrac{\sum f_i(a + d_i)}{N} \text{ OR } \bar{x} = a + \dfrac{\sum f_i d_i}{N}}$

Where, a is assumed mean (any central value of variate x_i)

Step Deviation Method

(By change of origin and scale)

In this method, if each value of deviation (d_i) is divided by any common number say, (h)

Then, let $\dfrac{d_i}{h} = u_i \Rightarrow d_i = hu_i$

Then, $\boxed{\bar{x} = a + \dfrac{\sum f_i u_i}{N} \times h}$

2. Properties of AM

(1) The sum of deviation of variate from their AM is always zero

$$\boxed{\sum (x_i - \bar{x}) = 0 \text{ OR } \sum f_i(x_i - \bar{x}) = 0}$$

(2) If the AM of the observations $x_1, x_2, \dots\dots, x_n$ is $\bar{x}$.

Then, AM of $x_1 \pm \lambda, x_2 \pm \lambda, \dots\dots, x_n \pm \lambda$ is

$\boxed{\bar{x} \pm \lambda}$

(3) If the AM of the observations $x_1, x_2, \dots\dots, x_n$ is $\bar{x}$.

Then, AM of $\lambda x_1, \lambda x_2, \dots\dots, \lambda x_n$ is $\boxed{\lambda\bar{x}}$

(4) If the AM of the observations $x_1, x_2, \dots\dots, x_n$ is $\bar{x}$.

Then, AM of $ax_1 \pm b, ax_2 \pm b, \dots\dots\dots, ax_n \pm b$ is $\boxed{a\bar{x} \pm b}$

(Where a, b are constants)

(5) The sum of square of deviation of variate from their AM is minimum.

i.e. $\sum (x_i - \bar{x})^2 \le \sum (x_i - a)^2$

(where a is any central value)

(6) AM is not affected by any change in assumed mean.

MIND MAP

Statistics

3. Weighted Mean (W.M.)

Let the 'n' values be : $x_1, x_2, ... x_n$

& Their Weightages are : $w_1, w_2, ... w_n$

Then, WM of this distribution is

$$W.M. = \frac{w_1x_1 + w_2x_2 + ... + w_nx_n}{w_1 + w_2 + ... + w_n}$$

$$= \sum_{i=1}^{n} w_i x_i \Big/ \sum_{i=1}^{n} w_i$$

4. Combined Mean (C.M.)

If, $(\bar{x}_1, \bar{x}_2, ..., \bar{x}_k)$ are mean of observations of 'K' series which have $(n_1, n_2, ..., n_k)$ terms respectively then their combined mean is –

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + ... + n_K\bar{x}_K}{n_1 + n_2 + ... + n_k}$$

5. Geometric Mean (GM)

For Ungrouped Distribution

If $x_1, x_2, ... x_n$ are n positive values then their GM is

$$GM = (x_1 \times x_2 \times x_3 \times \cdots \times x_n)^{1/n}$$

"Thus, the n^{th} root of the product of n +ve values" is known as GM of these values.

For Ungrouped Distribution

$$GM = \text{antilog}\left[\frac{1}{n}\sum_{i=1}^{n} \log x_i\right]$$

For Frequency Distribution

Let 'n' values be : $x_1, x_2, ... x_n$

Their frequencies : $f_1, f_2, ... f_n$

GM of this freq. distribution is

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

Here, $N = \sum f_i$

6. Harmonic Mean (HM)

For Ungrouped Distribution

If $x_1, x_2 .... x_n$ are 'n' non-zero values then their HM is

$$\frac{1}{\frac{1}{n}\left(\frac{1}{x_1} + \frac{1}{x_2} + ... + \frac{1}{x_n}\right)}$$

$$= \frac{n}{\left(\frac{1}{x_1} + \frac{1}{x_2} + ... + \frac{1}{x_n}\right)}$$

For Ungrouped Distribution

$$HM = \frac{1}{\frac{1}{n}\left(\frac{1}{x_1} + \frac{1}{x_2} + ... + \frac{1}{x_n}\right)}$$

Thus, the reciprocal of the mean of the reciprocal of given values is known as Harmonic Mean of these values.

For Frequency Distribution

Let 'n' values be : $x_1, x_2, ... x_n$

Their frequencies : $f_1, f_2, ... f_n$

Then, HM of this frequency distribution is

$$HM = \frac{1}{\frac{1}{N}\left(f_1 \times \frac{1}{x_1} + f_2 \times \frac{1}{x_2} + ... + f_n\frac{1}{x_n}\right)}$$

$$HM = \frac{N}{\left(\frac{f_1}{x_1} + \frac{f_2}{x_2} + ... + \frac{f_n}{x_n}\right)}$$

7. Median

- Median is the middle value of a series when the values of the series are arranged in ascending or descending order
- It divides an arranged series into two equal parts.

MIND MAP

Statistics

7. Median

For Ungrouped Distribution

- Let 'n' be the number of terms in a series.
- Firstly arrange the series in ascending or descending order. Then, in this arranged series,

$$\text{Median} = \begin{cases} \left(\dfrac{n+1}{2}\right) \text{th term} & (n \text{ is odd}) \\ \dfrac{\left(\frac{n}{2}\right) \text{th term} + \left(\frac{n}{2}+1\right) \text{th term}}{2} & (n \text{ is even}) \end{cases}$$

For Ungrouped Frequency Distribution

1. Firstly arrange the values of variate x_i, in ascending order with their frequency.
2. Prepare cumulative frequency (c.f) column and find the value of N.

$$\text{Median} = \begin{cases} \left(\dfrac{N+1}{2}\right) \text{th term} & (N \text{ is odd}) \\ \dfrac{\left(\frac{N}{2}\right) \text{th term} + \left(\frac{N}{2}+1\right) \text{th term}}{2} & (N \text{ is even}) \end{cases}$$

For Grouped Frequency Distribution

1. Prepare the cummulative frequency (c.f.) column and find the value of $N/2$
2. Find the class which contain the value of c.f. which is equal or just greater to $N/2$. This is median class.

$$\text{Median} = \ell + \frac{\left(\frac{N}{2} - F\right)}{f} \times h$$

Where, ℓ – lower limit of median class.
F – c.f. of the class preceding median class.
f – frequency of median class.
h – Class interval of median class.

8. Mode

Mode is the value of variate which have maximum frequency.

For Ungrouped Distribution

The value of variate which is repeated max. no. of times is known as its mode.

For Ungrouped Frequency Distribution

The value of variate which have maximum frequency.

For Grouped Frequency Distribution

Find the class which have max frequency. This is the Modal Class.

$$\text{Mode} = \ell + \frac{(f_0 - f_1)}{(2f_0 - f_1 - f_2)} \times h$$

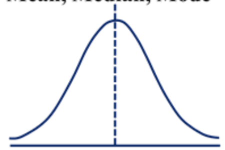
Where, ℓ – lower limit of modal class.
f_0 – frequency of modal class.
f_1 – frequency of the class preceding modal class.
f_2 – frequency of the class succeeding modal class.
h – class interval of modal class.

9. Relation b/w Mean, Median and Mode

Symmetric and Asymmetric (Skewed) Distribution

A distribution is called Symmetric, when the value of Mean, Median & Mode of the distribution are same (equal).

Mean, Median, Mode



Symmetrical Distribution

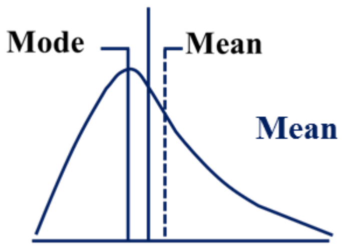
MIND MAP

Statistics

9. Relation b/w Mean, Median and Mode

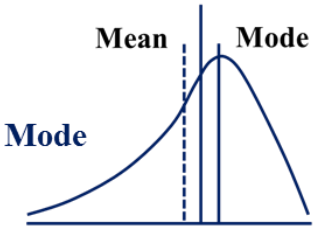
Symmetric & Asymmetric (Skewed) Distribution

A distribution which is Not Symmetric is called Skewed (or asymmetric) distribution.



Positive Skewed

Mean > Median > Mode



Negative Skewed

Mean < Median < Mode

Note :

For a Skewed Distribution –

Mean – Mode = 3 (Mean – Median)

$$\boxed{\text{Mode} = 3\text{Median} - 2\text{Mean}}$$

10. Range

The range of a distribution is the difference of Highest Value (H) and Least Value (L) of the distribution.

$$\boxed{\text{Range} = H - L}$$

The Range of the Grouped Distribution is the difference of Upper Limit of Max Class (H) and Lower Limit of minimum class (L).

Coefficent of Range =

$$\frac{\text{Diff. of extreme values}}{\text{Sum of extreme values}} = \frac{H - L}{H + L}$$

11. Mean Deviation (M.D.)

The mean of absolute value of deviation of each variate from any central value is called Mean Deviation about that central value (mean, median, mode).

Let A be any central value of a distribution then M.D. about A is defined as

$$\text{M. D.} = \frac{\sum_{i=1}^{n} |x_i - A|}{n}$$ (For Ungrouped Distribution)

(For Frequency Distribution) $$\text{M. D.} = \frac{\sum_{i=1}^{n} f_i |x_i - A|}{N}, \quad N = \sum f_i$$

12. Variance and Standard Deviation

Variance :

The Mean of square of deviation of each variate from their mean is called Variance. ❖ It is denoted by σ^2

Standard Deviation :

❖ The positive square root of the variance is called Standard Deviation.

❖ It is denoted by σ.

$$\boxed{\text{S. D.} = +\sqrt{\text{Variance}}}$$

Formulae for Variance

For Ungrouped Distribution :

(1) $\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n} = \frac{\sum x_i^2}{n} - \frac{2\bar{x}\sum x_i}{n} + \frac{\sum \bar{x}^2}{n}$

(2) $\sigma^2 = \frac{\sum x_i^2}{n} - \bar{x}^2 = \frac{\sum x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2$

(3) $\sigma^2 = \frac{\sum d_i^2}{n} - \left(\frac{\sum d_i}{n}\right)^2$; $[d_i = x_i - a]$

(4) $\sigma^2 = h^2\left[\frac{\sum u_i^2}{n} - \left(\frac{\sum u_i}{n}\right)^2\right]$; $\left[u_i = \frac{d_i}{h}\right]$

For Frequency Distribution : (1) $\sigma^2 = \frac{\sum f_i (x_i - \bar{x})^2}{N}$

MIND MAP

Statistics

Formulae for Variance

For Frequency Distribution :

(1) $\sigma^2 = \dfrac{\sum f_i(x_i - \bar{x})^2}{N}$

(2) $\sigma^2 = \dfrac{\sum f_i x_i^2}{N} - \bar{x}^2 = \dfrac{\sum f_i x_i^2}{N} - \left(\dfrac{\sum f_i x_i}{N}\right)^2$

(3) $\sigma^2 = \dfrac{\sum f_i d_i^2}{N} - \left(\dfrac{\sum f_i d_i}{N}\right)^2$

(4) $\sigma^2 = h^2\left[\dfrac{\sum f_i u_i^2}{N} - \left(\dfrac{\sum f_i u_i}{N}\right)^2\right]$

Note:

(1) $\sigma_x^2 = \sigma_d^2 = h^2 \times \sigma_u^2 = \sigma^2$

(2) The variance of first 'n' Natural No.s is $\sigma^2 = \dfrac{n^2 - 1}{12}$

(3) The variance of A.P. $a, a + d, a + 2d, \ldots\ldots, a + 2nd$ is

$$\sigma^2 = \frac{n(n+1)}{3}d^2$$

Properties of Variance

Property – 1

Let the distribution be : $x_1, x_2, \ldots\ldots, x_n$

And if Variance is : σ^2

Then,

Var. $(x_1 \pm \lambda, x_2 \pm \lambda, \ldots\ldots\ldots, x_n \pm \lambda)$:

$\boxed{\sigma^2}$

Var. $(\lambda x_1, \lambda x_2, \ldots\ldots\ldots\ldots, \lambda x_n)$:

$\boxed{\lambda^2\sigma^2}$

Var. $(ax_1 \pm b, ax_2 \pm b, \ldots., ax_n \pm b)$:

$\boxed{a^2\sigma^2}$

Where λ, a, b are constants.

Thus, the Variance and S.D. does not depend on change of origin (add & subtract) but they depend on change of scale (Multiplication & Division).

Property – 2

We have,

Distribution 1

(Mean : $\bar{x}_1$, Variance : σ_1^2,

No. of Terms : n_1)

Distribution 2 –

(Mean : $\bar{x}_2$, Variance : σ_2^2,

No. of Terms : n_2)

And their Combined Mean : $\bar{x}$

Then, their combined variance σ^2 is

$$\boxed{\sigma^2 = \frac{n_1(\sigma_1^2 + d_1^2) + n_2(\sigma_2^2 + d_2^2)}{n_1 + n_2}}$$

Where,

$d_1 = \bar{x}_1 - \bar{x}$

$d_2 = \bar{x}_2 - \bar{x}$

$$\sigma^2 = \frac{n_1\sigma_1^2 + n_2\sigma_2^2}{n_1 + n_2} + \frac{n_1 d_1^2 + n_2 d_2^2}{n_1 + n_2}$$

$$= \frac{n_1\sigma_1^2 + n_2\sigma_2^2}{n_1 + n_2} + \frac{n_1 n_2}{(n_1 + n_2)^2}(\bar{x}_1 - \bar{x}_2)^2$$