U3A - Maths

Term 2 / Week 5 21/05/21

Measures of Dispersion

Review:

- Arithmetic Mean (AM): $\bar{x} = \left(\frac{x_1 + x_2 + \cdots x_n}{n}\right)$
- Geometric Mean (GM): $= \sqrt[n]{x_1 x_2 \cdots x_n}$
- Harmonic Mean (H.M.): $n / \left(\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}\right)$
- Quadratic Mean (QM): $\sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}}$

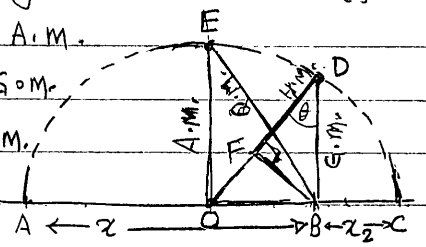
(Q.M. is the Root Mean Square value - when the mean value is zero, such as a sine wave)

- Pythagorean Means (for 2 values i.e. x_1 & x_2)

Q.M > A.M.

A.M. > G.M.

G.M. > H.M.



-2-

Home Work

- Show that the length 'DF' in the above Figure corresponds to Harmonic Mean.

Let Angle $O\hat{D}B = \theta$

Considering triangle ODB: $\cos\theta = \frac{BD}{OD}$

Considering triangle BFD: $\cos\theta = \frac{DF}{BD}$

$\therefore \cos\theta = \frac{DF}{BD} = \frac{BD}{OD}$ $\therefore DF = \frac{(BD)^2}{OD}$

we have $BD\ (= G.M.) = \sqrt[2]{x_1 x_2}$

$OD\ (= A.M.) = \left(\frac{x_1 + x_2}{2}\right)$

$\therefore DF = \frac{(\sqrt[2]{x_1 x_2})^2}{\left(\frac{x_1 + x_2}{2}\right)} = \left(\frac{2 x_1 x_2}{x_1 + x_2}\right)$

we have $H.M. = \frac{2}{(1/x_1 + 1/x_2)} = \left(\frac{2x_1 x_2}{x_1 + x_2}\right)$ ✓

-3-

Measures of Dispersion

- We have now defined "central" or "middle" value for a given "set of values" - we would like to know a "Measure of spread" for the given set of values.
- We invariably use "Arithmetic Mean" as it provides for mathematically simple and convenient equation for furthur statistical analysis.
- Let us now consider an Example

Ex: Given a set of values, namely, 5, 20, 25, 50, 75, 80, 151

-4-

we have Mean $\bar{x} = \frac{5 + 20 + \cdots + 151}{7} = \underline{58}$

Range as a Measure of Dispersion

we have Mean $= \underline{58}$

Range = Max. Val - Min. Val

$= 151 - 5 = \underline{146}$

- Range does provide a measure of dispersion.
- It uses only the extreme values and ignores all other values.
- It is good to have a measure which involves all the values
- Hence, the "Range" is rarely used in practice:
- Range is often specified as Range = [Min : Max] = [5 : 151]

Mean Deviation

- It is the Arithmetic Mean of the "deviations from the Mean value"

- Considering Ex. 1, we have,

Mean ($\bar{x}$) = 58

x: 5, 20, 25, 50, 75, 80, 151

$(x-\bar{x})$: −53, −38, −33, −8, 17, 22, 93

$$\text{Mean of Deviations} = \frac{\Sigma(x-\bar{x})}{N}$$

$$= \frac{-53-38-33-8+17+22+93}{7} = \frac{0}{7} = 0!$$

This does not help us!
The value is always zero!!

Alternative 1

- We can unceremoniously drop the "negative" signs.

- In other words, we can use the "absolute" or "modulus"

$$\text{Mean Deviation} = \frac{\Sigma(|x-\bar{x}|)}{N}$$

$$= \frac{53+38+33+8+17+22+93}{7} = \frac{264}{7}$$

$$= 37.174$$

- The above is often used by Economists and others.

- But true mathematicians do not like "arbitrarily" dropping the negative sign, as it is not a mathematical operation.

- True mathematicians prefer square of the deviations.

- Such a value is called the "Variance."

Alternative 2

Variance

Variance is defined as the mean of the square of deviations

$$\therefore \text{Variance } (s^2) = \frac{\Sigma(x-\bar{x})^2}{N}$$

- Considering Ex. 1, we have

$(x-\bar{x})^2$: $(-53)^2, (-38)^2, (-33)^2, (-8)^2, 17^2, 22^2, 93^2$

$$\therefore \underset{(s^2)}{\text{Variance}} = \frac{2809+1444+1089+64+289+484+8649}{7}$$

$$= 2118.29$$

- Variance has no practical significance, since it represents the "square of the values in the raw data"

- A more meaningful and practical value is the "square root" of the variance

- Hence we now define

$$\text{Standard Deviation } (s) = \sqrt{\frac{\Sigma(x-\bar{x})^2}{N}}$$

- Considering Ex. 1, we have

$$\text{Std. Dev. } (s) = \sqrt{(\text{variance})}$$

$$= \sqrt{2118.29}$$

$$= 46.03$$