27 December 2023 08:16 PM

① s.d is independent of change in origin but dependent of change in scale

i.e. if $y = a + bx$, then prove that $\boxed{\sigma_y = |b|\,\sigma_x}$.

$$\sigma_y^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}(\cancel{a} + bx_i - \cancel{a} - b\bar{x})^2$$

$$\boxed{\begin{aligned} y &= a + bx \\ \rightarrow \bar{y} &= a + b\bar{x} \end{aligned}}$$

$$\sigma_y^2 = \frac{b^2}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

$$\sigma_y^2 = b^2\sigma_x^2$$

$$\sigma_y = |b|\,\sigma_x \quad (\text{Proved})$$

② Show that the standard deviation is the min root mean square deviation.

Let $\Sigma(x_i - A)^2 = \Sigma\left[(x_i - \bar{x}) + (\bar{x} - A)\right]^2$

$$\Sigma(x_i - A)^2 = \Sigma(x_i - \bar{x})^2 + 2(\bar{x} - A)\underline{\Sigma(x_i - \bar{x})} + \underline{\Sigma(\bar{x} - A)^2}$$

$$= \Sigma(x_i - \bar{x})^2 + 0 + n(\bar{x} - A)^2$$

$$\ldots)^2 = (\Sigma(x_i - \bar{x})^2 + (\bar{x} - A)^2$$

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

min value of $(\bar{x}-A)^2 = 0$

$$\Rightarrow \bar{x} = A$$

$\Rightarrow \frac{1}{n}\Sigma(x_i-A)^2$ will be min at $\bar{x}=A$

$\Rightarrow \sqrt{\frac{1}{n}\Sigma(x_i-A)^2}$ will be min at $\bar{x}=A$

$\therefore$ $\boxed{\text{SD is less than RMSD}}$.

Q Find the mean and variance of the first 'n' natural numbers.

$x = 1, 2, 3, 4, \ldots, n$

$$\Sigma x = 1+2+\cdots+n = \frac{n(n+1)}{2}$$

$$\therefore \bar{x} = \frac{1}{n}\Sigma x = \frac{1}{\cancel{n}}\cdot\frac{\cancel{n}(n+1)}{2} = \frac{n+1}{2}$$

$$\text{Variance, } \sigma_x^2 = \frac{1}{n}\Sigma x_i^2 - \bar{x}^2$$

n

$x^2 : 1^2, 2^2, \dots, n^2$

$\Sigma x^2 = 1^2 + 2^2 + \dots + n^2 = \frac{n(2n+1)(n+1)}{6}$

$$\therefore \sigma_x^2 = \frac{1}{\cancel{n}} \times \frac{\cancel{n}(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2$$

$$= \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4}$$

$$= \frac{4(n+1)(2n+1) - 6(n+1)^2}{24}$$

$$= \frac{4(2n^2 + n + 2n + 1) - 6(n^2 + 1 + 2n)}{24}$$

$$= \frac{8n^2 + \cancel{12n} + 4 - 6n^2 - 6 - \cancel{12n}}{24}$$

$$= \frac{2n^2 - 2}{24} = \frac{n^2 - 1}{12}$$

$$\therefore s.d, \sigma_x = \sqrt{\frac{n^2-1}{12}} \quad \text{(ans)}$$

Q

	First set	Second set	Third set	Composite set

Q

	First set	Second Set	Third Set	Composite Set
No. of obs	$n_1 = 200$	$n_2 = 250$	$n_3 = 300$	$N = 750$
Mean	$\bar{x}_1 = 25$	$\bar{x}_2 = 10$	$\bar{x}_3 = 15$	$\bar{x} = ?$
Std Deviation	$\sigma_1 = 3$	$\sigma_2 = 4$	$\sigma_3 = 5$	$\sigma = ?$

Combined mean, $\bar{x} = \dfrac{n_1\bar{x}_1 + n_2\bar{x}_2 + n_3\bar{x}_3}{N} = 16$

Combined variance $\sigma^2 = \dfrac{(n_1\sigma_1^2 + n_2\sigma_2^2 + n_3\sigma_3^2) + n_1d_1^2 + n_2d_2^2 + n_3d_3^2}{N}$

where $d_i = \bar{x}_i - \bar{x}$

$d_1 = \bar{x}_1 - \bar{x} = 9$

$d_2 = \bar{x}_2 - \bar{x} = -6$

$d_3 = \bar{x}_3 - \bar{x} = -1$

$\sigma^2 = \dfrac{38800}{750}$

$\sigma^2 = 51.73$

$\sigma = \sqrt{51.73} = 7.19$ (ans).

$n_1 = 1$, $\bar{x}_1 = a-36$, $n_2 = 9$, $\bar{x}_2 = a$

Q In a batch of 10 children, The I.Q. of a dull boy is 36 below the average I.Q. of other children. Show that the standard deviation of I.Q for all the children cannot be less than 10.8.

If this s.d is actually 11.4, determine ... the s.d. will be when the

If this s.d is ...
what the s.d. will be when the
dull boy is left out.

'a' avg I Q of 9 children $= \bar{x}_2$

a - 36 ... of 1 dull boy $= \bar{x}_1$

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

$$\bar{x} = \frac{1 \times (a-36) + 9 \times a}{10}$$

$$\bar{x} = \frac{10a - 36}{10} = a - 3.6$$

$$d_1 = \bar{x}_1 - \bar{x} = (a-36) - (a-3.6)$$
$$= -32.4$$

$$d_2 = \bar{x}_2 - \bar{x} = a - (a - 3.6) = 3.6$$

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

group one has 1 obs $\therefore \sigma_1^2 = 0$.

$$\sigma^2 = \frac{0 + 9 \times \sigma_2^2 + (-32.4)^2 + 9(3.6)^2}{10}$$

σ^2 ... $9\sigma^2 + \ldots$ 1166.4 $\Rightarrow 10 \times 11.4^2 - 1166.4$

$$10\sigma^2 = 9\sigma_2^2 + \cancel{116.64}\ 1166.4 \Rightarrow \frac{10\times 11.4^2 - 1166.4}{9}$$

$$= 14.8$$

$$\boxed{\sigma^2 = 0.9\sigma_2^2 + 116.64}$$

$$\sigma_2^2 = 14.8$$

$$\sigma_2 = \sqrt{14.8}$$

$$= 3.85 \text{ (ans)}$$

$$\sigma^2 \geq 116.64$$

$$\sigma \geq \sqrt{116.64}$$

$$\sigma \geq 10.8 \qquad \text{(Proved)}.$$

$$\sigma = 11.4$$

$$\sigma^2 =$$

$$\sigma_2^2 = \frac{(11.4)^2 - 116.64}{0.9}$$

$$\sigma_2^2 = 14.8$$

Coefficient of variation, $C.V = \frac{sd}{\bar{x}} \times 100$

(1)

(2) coeff of M.D $= \frac{MD(\bar{x})}{\bar{x}} \times 100$

or $\frac{md(\text{median})}{\text{median}} \times 100$

(3) coeff of Q.D $= \frac{Q.D}{\text{med}} \times 100$

Q From some financial statistics, it is found that the monthly average electricity charges were Rs ~~[illegible]~~ 2460 ... Direct

Q From some financ... monthly average electricity charges were Rs ~~...~~ and s.d. Rs 120. The monthly average Direct wages was Rs 42,000 and s.d 1200. State which is the more variable?

	(A) Electricity	(B) Direct wage
Mean	2460	42 000
s.d	120	1200

$$CV_A = \frac{s.d_A}{\bar{x}_A} \times 100 = \frac{12\cancel{0}}{246\cancel{0}} \times 100 = \frac{1200}{246} = 4.87$$

$$CV_B = \frac{s.d_B}{\bar{x}_B} \times 100 = \frac{120\cancel{0}}{42\cancel{000}} \times 1\cancel{00} = \frac{120}{42} = 2.85$$

$CV_A > CV_B$

$\therefore$ Electricity charges are more variable than Direct wages.

$$Q.D = \frac{Q_3 - Q_1}{2}$$

$$CV \text{ of } QD = \frac{Q.D}{\text{Median}} = \frac{Q_3 - Q_1}{2Q_2}$$

Q Two variables A and B are given. ... coefficient of Q.D ... check which is more ...

Q - Two variables ... using coefficient of Q.D check which is more variable?

C of QD
(Q_1, Q_2, Q_3)

Series A

Mid value	Frequency
15	15
20	33
25	56
30	103
35	40
40	32
45	10

N

Series B (Q_1, Q_2, Q_3)
C Q.D B.

Mid-value	Frequency
100	340
150	492
200	890
250	1420
300	620
350	360
400	187
450	140

$Q_1 \rightarrow N/4$
$Q_2 \rightarrow N/2$
$Q_3 \rightarrow 3N/4$

H.W Try