Standard Deviation (S.D.)

It is a measure of dispersion or variation. It is universally used to show the scatter of individual measurements around the mean of all the measurements in a given distribution.

S.D. is the ~~cumulative~~ +ve square root of ~~[illegible]~~ variance or it is the square root of avg of the square deviation. (d)

" being measured from the mean of the observation

The idea of S.D. was given by Karl Pearson in 1983.

It is also known as root mean square deviation

Represented by the greek letter σ (sigma) or S or S.D.

$$\left\{\begin{array}{l} S.D. = \sigma = \sqrt{\text{variance}} \\ S.D. = \sigma \\ \text{Variance} = \sigma^2 \end{array}\right\}$$

For ungrouped data.

$$\sigma = \sqrt{\frac{\Sigma(x-\mu)^2}{n}} \quad (\text{if } n > 30)$$

$$\sigma = \sqrt{\frac{\Sigma(x-\mu)^2}{n-1}} \quad (\text{if } n \leq 30)$$

x = individual measurement
μ = mean
n = total no of observation
d = deviation

$$x - \mu = d.$$

$$\therefore \quad \sigma = \sqrt{\frac{\Sigma d^2}{n}} \quad (n > 30)$$

or

$$\sigma = \sqrt{\frac{\Sigma d^2}{n-1}} \quad (n \leq 30).$$

For grouped data :-

1. Discrete series.

$$\sigma = \sqrt{\frac{\Sigma f(x-\mu)^2}{\Sigma f}} \quad \text{or} \quad \sqrt{\frac{\Sigma f d^2}{\Sigma f}}$$

$(n > 30)$.

$$\sigma = \sqrt{\frac{\Sigma f(x-\mu)^2}{\Sigma f - 1}} \quad \text{or} \quad \sqrt{\frac{\Sigma f d^2}{\Sigma f - 1}}$$

$(\text{if } n \leq 30)$

$(\Sigma f = n)$

$$\sqrt{\frac{\Sigma f (x-\mu)^2}{\Sigma f}} \quad \text{or} \quad \sqrt{\frac{\Sigma f d^2}{\Sigma f}}$$

For continuous series

$$\sigma = \sqrt{\frac{\Sigma f(m-\mu)^2}{\Sigma f}} \therefore \sqrt{\frac{\Sigma fd^2}{\Sigma f}}$$

$(n > 30)$

$$\sigma = \sqrt{\frac{\Sigma f(m-\mu)^2}{\Sigma f - 1}} \text{ or } \sqrt{\frac{\Sigma fd^2}{\Sigma f - 1}}$$

$(n \leq 30)$

Merits of S.D.

1. Based on all the observations
2. Always calculated from airthmetic mean & not from median or mode.
3. Always a +ve quantity because by squaring the deviation -ve values become +ve.
4. Very imp statistical parameter & is based on algebric treatments

Demerits

1. Not easy to understand
2. difficult to calculate
3. gives more weightage to extreme values.

Importance of S.D.

1. It is most imp & useful parameter in case of normal frequency distribution because d is an indication of the spreading of measurement around

the centre of the distribution.

2. Used to remove all the errors of mean deviation.

3. useful in all conditions even if the arithmetic mean is in decimal.

4. It has uniformity in results because it is always obtained from arithmetic mean.

Q. Calculate standard deviation for 11, 16, 13, 20.

20
16
11
13
60

$$\text{Mean} = \frac{60}{4} = 15$$

x	$x - \mu$	$(x - \mu)^2$
11	$11 - 15 = -4$	16
16	$16 - 15 = 1$	1
13	$13 - 15 = -2$	4
20	$20 - 15 = 5$	25
		46

$$\sigma = \sqrt{\frac{\Sigma (x - \mu)^2}{n - 1}}$$

$$= \sqrt{\frac{46}{3}}$$

$$\sigma = \sqrt{15.33} = 3.915$$

Standard Error (S.E).

It is called as distribut^n of means. It is the diff b/n mean of populat^n & the sample.
S.E. is the ratio of S.D. of sample divided by square root of total no of observations.

$$S.E = \frac{S.D.}{\sqrt{n}} \quad \text{or} \quad \frac{\sigma^2}{n} \ (\text{Variance}).$$

If random sample of size 'n' are drawn from a normal population, the means of these sample will also show normal distribution.
The ʔ of means from a non - normal population will not be normal but will tend towards normality as 'n' increases in size.
Further the variance of the distribution of the means will decrease as 'n' incses

The variance of the population of all possible means of sample of size 'n' from a population with variance σ^2 is

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{S.D}{\sqrt{n}}$$

$\sigma_{\bar{x}}^2$ = Variance of means.

The value ($\sigma_{\bar{x}}$) is the standard deviat^n of the mean. The S.D. of mean is called S.E.

The concept was given by G.V. Yule in 1897.

The sample mean is often written together with it's S.E. as $(\bar{X} \pm S.E.)$

σ^2 = Variance

σ = S.D.

$\sigma_{\bar{x}}^2$ = Variance of Mean.

$\sigma_{\bar{x}}$ = S.D of mean or S.E.

Importance of S.E.

1. Very imp Parameter when the distributⁿ is normal or near normal.
2. It is a function of S.D. & sample size.
3. The greater the sample size smaller will be the S.E.

when sample size ↑ses to a very large no, S.E. becomes very small.

4. S.E. also used in hypothesis testing
5. It is of greatest use when the distribution of the estimate to w̄ they are attached is almost normal.

Q. Calculate the S.E. when the systolic B.P. in mm of Hg is given. Express the answer as X ± S.E.

151, 121, 125, 128, 134, 136, 138, 139, 141, 144, 145, 149.

$$\text{Mean} = \frac{\Sigma x}{n} = \frac{1651}{12} = 137.58$$

$$S.D = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$$

x	$x-\bar{x}$	$(x-\bar{x})^2$	
151	$151-137=14$	$(14)^2$	196
121	$121-137=-16$	$(-16)^2$	256
125	$125-137=-12$	$(-12)^2$	144
128	$128-137=-9$	$(-9)^2$	81
134	$134-137=-3$	$(-3)^2$	9
136	$136-137=-1$	$(-1)^2$	1
138	$138-137=1$	$(1)^2$	1
139	$139-137=2$	$(2)^2$	4
141	$141-137=4$	$(4)^2$	16
144	$144-137=7$	$(7)^2$	49
145	$145-137=8$	$(8)^2$	64
149	$149-137=12$	$(12)^2$	144
			965

151
121
125
128
134
136
138
139
141
144
145
149
~~1651~~
~~12~~ 137.58

137 − 121 = 16

151 − 137 = 14

137 − 125 = 12

16 × 16 = 196, 256

14 × 14 = 56, 196

$$\underline{S.D} = \sqrt{\frac{965}{11}} =$$

$$= \sqrt{87.72} = ~~9.17~~\ 9.36$$

$$\underline{S.E.} = \frac{S.D}{\sqrt{n}} = \frac{9.36}{\sqrt{12}} = \frac{9.36}{3.46}$$

$$= ~~2.65~~\ \boxed{2.70}$$