U3A - Maths

Term 2 / Week 7 4-Jun-21

Sample Std. Dev. – Bessel Correction Factor

We have defined 3 equations for Standard deviation

1. Population Std. Dev (σ)

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

where μ – Population Mean
n – No. of values in Popln.

2. Sample Std. Dev. (S_N)

$$S_N = \sqrt{\frac{\sum (x-\bar{x})^2}{N}}$$

Where, $\bar{x}$ – Sample Mean
N – No. of values in Sample.

3. Sample Std. Dev. (S_{N-1})

$$S_{N-1} = \sqrt{\frac{\sum (x-\bar{x})^2}{(N-1)}}$$

where $\bar{x}$ – Sample Mean
N – No. of values in Sample

S_{N-1} – is the confusing one!

- One way of explaining this is to say that there are only (N-1) degrees of freedom since sum of deviations $\sum (x-\bar{x}) = 0$.
- But such a statement is true for population also! So it does not make sense!!

- The real use of S_{N-1} is to estimate "Standard deviation of a population" as a "Mean of Std. deviations obtained from the several Samples"

- In other words, S_{N-1} is only relevant when we are using "Sample Std. dev" to estimate "population std. dev".
- In such a case we need use a correction factor called the Bessel correction factor.

$$\therefore S_{N\,(corrected)} \equiv S_{N-1} = \sqrt{\frac{\sum (x-\bar{x})^2}{N} \times \left(\frac{N}{N-1}\right)}$$

Bessel Correction Factor.

- Hence, Bessel Correction Factor is

$$= \sqrt{\left(\frac{N}{N-1}\right)} \Rightarrow \text{depends on Sample size (N)}$$

- Let us explore the Bessel Correction factor.

Sample size (N)	$\sqrt{N/(N-1)}$	(%) Correction
2	1.41	41%
5	1.12	12%
10	1.05	5%
15	1.035	3.5%
20	1.026	2.6%
25	1.02	2.0%

- We can neglect the correction factor and use S_N instead of S_{N-1} for sample sizes above 25, since error is less than 2% !!
- We will now take an example to illustrate the use of S_{N-1} to estimate Popln. Std. Dev

Mean & Std Dev for "Sample" Data

Sample 1	Sample 2	Sample 3	Sample 4
51	47	43	47
47	47	43	39
45	41	49	47
45	49	37	43
43	47	45	43
41	41	45	53
47	45	51	51
43	39	39	47
43	49	45	41

(Above sample obtained using random numbers & f-Acc)
(Sample size = 9)

	Sample 1	Sample 2	Sample 3	Sample 4
Sample Mean	45.00	45.00	44.11	45.67
SD (N)	**2.83**	**3.53**	**4.12**	**4.32**
SD (N-1)	**3.00**	**3.74**	**4.37**	**4.58**

Mean of Sample Means = **44.94**

Mean of SD(N) **3.70**

Mean of SD(N-1) **3.92**

Mean & Std Dev for grouped "Population" data

x	f	f.x	x^2	f.x^2	f - Acc
37	2	74	1369	2738	2
39	4	156	1521	6084	6
41	10	410	1681	16810	16
43	15	645	1849	27735	31
45	19	855	2025	38475	50
47	19	893	2209	41971	69
49	15	735	2401	36015	84
51	10	510	2601	26010	94
53	4	212	2809	11236	98
55	2	110	3025	6050	100
Sum =>	100	4600		213124	

x => lengths of garden lizards
Mean = Sum(f.x)/sum(f)
Std Dev (Pop) = Sqrt[(Sum(f.x^2)/Sum(f)) - Mean^2)]

Pop Mean = **46.00** Pop Std Dev = **3.90**

"Mean of SD(N-1) values" - gives an estimate of the "Population Std Dev" (Almost perfect for this Example!!!)
"Mean of Sample Means" - gives an estimate of the "Population Mean"

Notes: 1. Bessel Correction Factor = Sqrt (N/(N-1)) = Sqrt (9/(9-1)) = 1.1
2. SD(N-1) = SD(N) x Bessel Correction Factor