

Interval/Ratio Variance

Sample 1	y-Mean (Residual)	(y-Mean) ²
12	-3.625	13.141
12	-3.625	13.141
14	-1.625	2.641
15	-0.625	0.391
17	1.375	1.891
18	2.375	5.641
18	2.375	5.641
19	3.375	11.391
Mean: 15.625	Total: 0	53.875

$$s^2 = \frac{\sum (y - \text{mean})^2}{n - 1}$$

$$s^2 = \frac{53.875}{7} = 7.696$$

$$s = \sqrt{S^2}$$

$$S = 2.774$$

23

General Rule Regarding Variance:

Under ordinary circumstances, variance as “s” or “cv” should not exceed 20-25% of value of mean, **unless the large cv denotes an effect.

**If so, must determine why
Example: (West, Haiti BP Study)

24

Understanding Variance, Standard Deviation, and the Coefficient of Variation

Sample 1	(Y-mean) ²	Sample 2	(Y-mean) ²
12	13.14	10	31.64
12	13.14	11	21.39
14	2.64	12	13.14
15	.391	14	2.64
17	1.89	15	.391
18	5.64	15	.391
18	5.64	16	.141
19	11.39	32	268.14
Mean: 15.625	Total: 53.86	Mean: 15.625	Total: 337.88

	Sample 1	Sample 2
S ²	7.696	48.267
S	2.774	6.948
CV	17.755	44.464

$$CV = \frac{s}{\bar{Y}} * 100$$

$$CV = \frac{2.774}{15.63} * 100 = 17.76$$

25

Nominal/Ordinal Data Variance

Mean = 1.44

Perio (Y)	Freq (F)	Σ FY	Y-Mean	Y-Mean ²	F(Y-Mean) ²
0	10	0	-1.44	2.074	20.74
1	6	6	-.44	.194	1.164
2	14	28	.56	2.12	29.68
3	5	15	1.56	2.434	12.17
Total	35	49		6.822	63.754

$$s^2 = \frac{\sum F (Y - \text{Mean})^2}{n - 1}$$

$$s^2 = \frac{63.754}{34} = 1.875$$

$$s = \sqrt{s^2} = \sqrt{1.875} = 1.369$$

26

Sheppard's Correction

- Grouping assumes that all observations in an interval are centered at "y" (the midpoint)
- Seldom is this the case
- Results in error in calculating variance and standard deviation
- Solution is to correct for the lack of even spread in an interval using Sheppard's Correction

$$s^2 (\text{corrected}) = s^2 - \frac{h^2}{12}$$

Where "h" is the size of the interval

27

A Few Words on SEM

- Is the standard deviation of sample mean's estimation of the population mean.
- Measures of how well a sample represents the population.
When the sample is representative, the standard error will be small.
- Division by the square root of the sample size is reflection of speed with which an increasing sample size gives an improved representation of the population.
- Allows an approximation of the confidence intervals where the true population mean can be found

$$SEM = \frac{s}{\sqrt{n}}$$

$$SEM = \sqrt{\frac{s^2}{n}}$$

28

SEM Example

Population Parameters								
N	Range	Minimum	Maximum	Sum	Mean	SEM	SD	
2000	81	18	99	118016	59.01	.537	24.011	

Descriptive Statistics of 10 Samples								
Samples	n	Range	Minimum	Maximum	Sum	Mean	SEM	SD
1	98	80	18	98	5942	60.63	2.424	23.998
2	109	81	18	99	5990	54.95	2.423	25.297
3	109	81	18	99	6473	59.39	2.294	23.949
4	108	80	18	98	6580	60.93	2.265	23.543
5	99	80	19	99	6069	61.30	2.247	22.355
6	102	79	19	98	6072	59.53	2.548	25.734
7	109	81	18	99	6265	57.48	2.481	25.904
8	106	77	21	98	6178	58.28	2.229	22.944
9	105	80	19	99	6560	62.48	2.433	24.930
10	114	80	19	99	6402	56.16	2.240	23.915

Grand Mean of the 10 Sample Means								
n	Range	Minimum	Maximum	Sum	Mean	SEM	SD	
10	7.53	54.95	62.48	591.13	59.1130	.75560	2.38943	

SEM and Confidence Interval of Mean

- Population parameters often not known (Mean and Standard Deviation)
- Sample provides an estimation
- True mean can be estimated in a range* as a measure of SEM and sample mean. This is a confidence interval
- Critical values for common confidence intervals below

Confidence Interval	Two-Tailed Critical Value
90%	1.64
95%	1.96
99%	2.58

Population Mean = Mean ± (Critical Value * SEM) 30

Estimating the Population Mean From Sample Data

- Involves calculating confidence interval of the mean (CI)
- CI is a range with upper and lower values
- We deduce that Population Mean is Within CI of Mean

Population Mean = Mean ± (Critical Value * SEM)

Upper Limit (UL) = Mean + Critical Value * SEM
 Lower Limit (LL) = Mean - Critical Value * SEM

31

A random sample of 300 Ferange males age 20 to 25 years old had a systolic pressure of 284 ± 15 mmHg emitting from their 3rd heart. What is the 3rd heart systolic pressure of the other males that age in the population at the 95% confidence interval?

Solution

Population Mean = Mean \pm (Critical Value * SEM)

Confidence Interval	Two-Tailed Critical Value	$SEM = \frac{s}{\sqrt{n}} = \frac{15}{\sqrt{300}} = 0.2887$
95%	1.96	

UL = Mean + Critical Value * SEM = $284 + (1.96 * 0.2887) = 284.57$

LL = Mean - Critical Value * SEM = $284 - (1.96 * 0.2887) = 283.43$

Conclusion: The 3rd heart systolic pressure for all men in that age group is somewhere between 283.43 and 284.57

32
