Measures of variability

- Measures of central tendency: typical or average value for a sample or population
- Can be misleading
- How are values distributed

Measures of variability

- How much variability is there, in a sample, or in a population?
- Crude measure:

  Range = highest value minus lowest value
- More sophisticated measure of variability:

  Standard deviation

Standard deviation

- Average of the differences between each individual score and mean of all scores?
- How far is average score from the mean?
- How scattered are the data?

Standard deviation: ungrouped data

$$s = \sqrt{\frac{\sum (Y - \overline{Y})^2}{n}}$$

Check precedence order for calculation!

Standard deviation: grouped data

$$s = \sqrt{\frac{\sum f(Y - \overline{Y})^2}{n}}$$

Check precedence order for calculation!

Standard deviation: example

- Find mean: $\overline{Y} = \frac{\sum fY}{n}$

= 1663/76 = 21.9

Standard deviation example:
scores on a midterm

interval	midpt.	f	fY	$Y - \overline{Y}$	$(Y - \overline{Y})^2$	$f(Y - \overline{Y})^2$
1-5	3	x 0	= 0	-18.9	357.21	0
6-10	8	x 1	= 8	-13.9	193.21	193.21
11-15	13	x 6	= 78	-8.9	79.21	475.26
16-20	18	x 19	= 342	-3.9	15.21	288.99
21-25	23	x 33	= 759	1.1	1.21	39.93
26-30	28	x 17	= 476	6.1	37.21	632.57
		76	1663			1629.96

Standard deviation:
grouped data

$$s = \sqrt{\frac{\sum f(Y - \overline{Y})^2}{n}}$$

Check precedence order for calculation!

Standard deviation: example

- Divide sum by n (1629.96/76) = 21.4
- Take square root
- Answer: s = 4.6

Standard deviation: calculating formula

$$s = \sqrt{\frac{\sum fY^2}{n} - \left(\sum fY / n\right)^2}$$

Mean of the squares

Square of the means

Standard deviations:
calculating formula

Y	Y^2	f	fY^2	fY
3	9	0	0	0
8	64	1	64	8
13	169	6	1014	78
18	324	19	6156	342
23	529	33	17,457	759
28	784	17	13,328	476
		76	38,019	1663

Standard deviation:
calculating formula

$$s = \sqrt{(38{,}019/76) - (1663/76)^2}$$

$$s = \sqrt{500.25 - (21.88)^2}$$

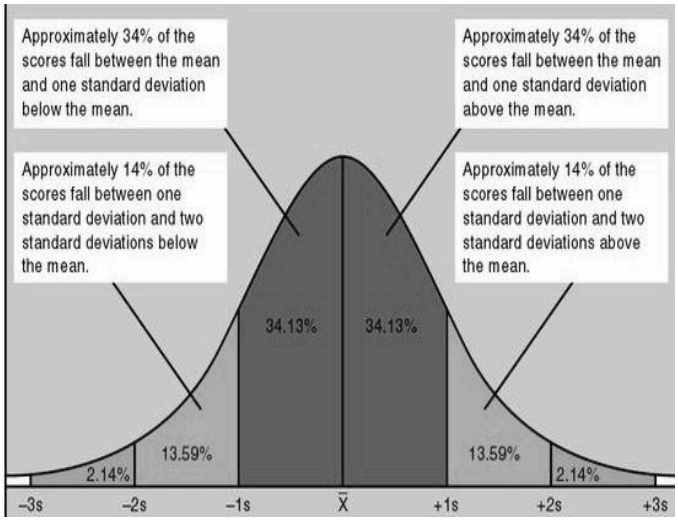
$$s = \sqrt{21.5156}$$

$$s = 4.6$$

Standard deviation

- If the population is normally distributed, you can estimate the percentage of the population within 1, 2, and 3 standard deviations of the mean:
- 68% of cases lie +/- 1 s.d. of mean
- 95% of cases lie +/- 2 s.d.'s of mean
- 99.9% of cases lie +/- 3 s.d.'s of mean

Standard deviation



$\overline{Y} = 21.9$