Political Science 30: Political Inquiry
Section 4

Taylor Carlson
tncarlson@ucsd.edu

▸ Link to Stats Motivation of the Week

Learning Outcomes

By the end of section today, you should:

- Understand key concepts from lecture:
  - Parameters vs. Statistics
  - Sampling (bias, variability, margin of error)
  - Calculating measures of central tendency and dispersion
- Feel prepared for the midterm exam!

Warm Up

- What is the population of interest?
- What is the sample?
- What was Kirk's sampling method?
- Kirk clearly didn't take POLI 30. What would you suggest he do to improve his poll?

▸ Link to Video

Warm Up

- What is the population of interest?
  - All citizens of Stars Hollow
  - "A population is the entire group of cases about which you want information"
- What is the sample?
- What was Kirk's sampling method?
- Kirk clearly didn't take POLI 30. What would you suggest he do to improve his poll?

▸ Link to Video

Warm Up

- What is the population of interest?
  - All citizens of Stars Hollow
  - "A population is the entire group of cases about which you want information"
- What is the sample?
  - Kirk
  - "A sample is a subset of the population which is used to gain information about the whole population."
- What was Kirk's sampling method?
- Kirk clearly didn't take POLI 30. What would you suggest he do to improve his poll?

▸ Link to Video

Parameters vs. Statistics

- Parameter: number describing a population (usually unknown!)

Parameters vs. Statistics

- Parameter: number describing a population (usually unknown!)
  - In Kirk's poll, what was the parameter of interest? Was it known?

Parameters vs. Statistics

- Parameter: number describing a population (usually unknown!)
  - In Kirk's poll, what was the parameter of interest? Was it known?
- Statistic: a number describing a sample

Parameters vs. Statistics

- Parameter: number describing a population (usually unknown!)
  - In Kirk's poll, what was the parameter of interest? Was it known?
- Statistic: a number describing a sample
  - What was the statistic in Kirk's poll?

Parameters vs. Statistics

- Parameter: number describing a population (usually unknown!)
  - In Kirk's poll, what was the parameter of interest? Was it known?
- Statistic: a number describing a sample
  - What was the statistic in Kirk's poll?
- If our sample is representative of the population, sample statistics will closely approximate population parameters

Parameters vs. Statistics

- Parameter: number describing a population (usually unknown!)
  - In Kirk's poll, what was the parameter of interest? Was it known?
- Statistic: a number describing a sample
  - What was the statistic in Kirk's poll?
- If our sample is representative of the population, sample statistics will closely approximate population parameters
  - Was Kirk's poll representative?

Sampling: Bias and Variability

- Sampling Bias: consistent deviation of the sample statistic from the parameter (when your statistic is far away from the true population parameter)

Sampling: Bias and Variability

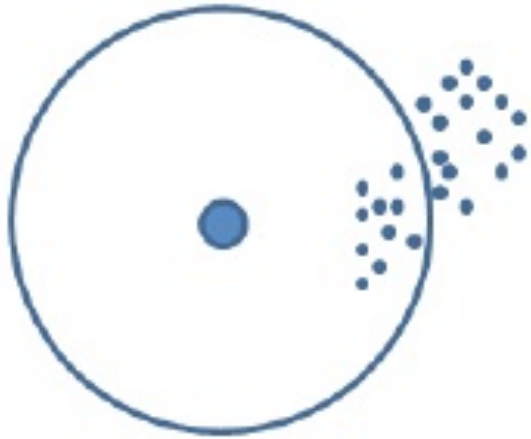
- Sampling Bias: consistent deviation of the sample statistic from the parameter (when your statistic is far away from the true population parameter)
- Sampling Variability: how far apart statistics are over many samples (how similar is the statistic from one sample to the statistic from another sample?)

Sampling: Bias? Variability?

Sampling: Bias? Variability?

Sampling: Bias? Variability?

High bias, Low variability



High bias, High variability

Low bias, High variability

Low bias, Low variability

So what? Why do we care?

Questions

What questions do you have?

Math!

I collected some data on the number of actors who have played a few popular superheroes over time. Note, I'm not an expert, this was my source: ▸ Link to Source

Superhero	Number of Actors
Batman	15
Superman	13
Wonderwoman	9
Daredevil	5
Spiderman	12
Catwoman	7

Math!

Find the mode, median, and mean of the number of actors who have portrayed these superheroes.

Superhero	Number of Actors
Batman	15
Superman	13
Wonderwoman	9
Daredevil	5
Spiderman	12
Catwoman	7
Captain America	11

What did you find?

- Mode?

What did you find?

- Mode? There isn't one!

What did you find?

- Mode? There isn't one!
- Median?

What did you find?

- Mode? There isn't one!
- Median?
  - 5, 7, 9, 11, 12, 13, 15 $\rightarrow$ **11**

What did you find?

- Mode? There isn't one!
- Median?
  - 5, 7, 9, 11, 12, 13, 15 $\rightarrow$ **11**
- Mean?

What did you find?

- Mode? There isn't one!
- Median?
  - 5, 7, 9, 11, 12, 13, 15 $\rightarrow$ **11**
- Mean?
  - $\frac{5+7+9+11+12+13+15}{7}$
  - $=\frac{72}{7}$
  - $=10.3$

How about a measure of dispersion: Variance

Note: This is my preferred method for calculating variance. There's a detailed example of this posted here: ▸ Link to Practice Example

$$Variance = \frac{\Sigma(X_i - \bar{X})^2}{N-1}$$

Calculating Variance

$$Variance = \frac{\Sigma(X_i - \bar{X})^2}{N-1}$$

Superhero	Number of Actors (X_i)
Batman	15
Superman	13
Wonderwoman	9
Daredevil	5
Spiderman	12
Catwoman	7
Captain America	11

Calculating Variance

$$Variance = \frac{\Sigma(X_i - \bar{X})^2}{N-1}$$

X_i	
15	
13	
9	
5	
12	
7	
11	

Calculating Variance

$$Variance = \frac{\Sigma(X_i - \bar{X})^2}{N-1}$$

X_i	$\bar{X}$
15	10.3
13	10.3
9	10.3
5	10.3
12	10.3
7	10.3
11	10.3

Calculating Variance

$$Variance = \frac{\Sigma(X_i - \bar{X})^2}{N-1}$$

X_i	$\bar{X}$	$X_i - \bar{X}$
15	10.3	15-10.3= 4.7
13	10.3	13-10.3= 2.7
9	10.3	9-10.3= -1.3
5	10.3	5-10.3= -5.3
12	10.3	12-10.3= 1.7
7	10.3	7-10.3= -3.3
11	10.3	11-10.3= 0.7

Calculating Variance

$$Variance = \frac{\Sigma(X_i - \bar{X})^2}{N-1}$$

X_i	$\bar{X}$	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
15	10.3	15-10.3= 4.7	4.7^2= 22.1
13	10.3	13-10.3= 2.7	2.7^2= 7.3
9	10.3	9-10.3= -1.3	$(-1.3)^2$= 1.7
5	10.3	5-10.3= -5.3	$(-5.3)^2$= 28.1
12	10.3	12-10.3= 1.7	1.7^2= 2.9
7	10.3	7-10.3= -3.3	$(-3.3)^2$= 10.9
11	10.3	11-10.3= 0.7	0.7^2= 0.5

Calculating Variance

$$Variance = \frac{\Sigma(X_i - \bar{X})^2}{N-1}$$

X_i	$\bar{X}$	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
15	10.3	15-10.3= 4.7	4.7^2= 22.1
13	10.3	13-10.3= 2.7	2.7^2= 7.3
9	10.3	9-10.3= -1.3	$(-1.3)^2$= 1.7
5	10.3	5-10.3= -5.3	$(-5.3)^2$= 28.1
12	10.3	12-10.3= 1.7	1.7^2= 2.9
7	10.3	7-10.3= -3.3	$(-3.3)^2$= 10.9
11	10.3	11-10.3= 0.7	0.7^2= 0.5
			73.5

$$22.1 + 7.3 + 1.7 + 28.1 + 2.9 + 10.9 + 0.5 = 73.5$$

Calculating Variance

$$Variance = \frac{\Sigma(X_i - \bar{X})^2}{N-1}$$
$$= \frac{73.5}{7-1}$$

Calculating Variance

$$Variance = \frac{\Sigma(X_i - \bar{X})^2}{N-1}$$

$$= \frac{73.5}{7-1}$$

$$Variance = \frac{\Sigma(X_i - \bar{X})^2}{N-1}$$

$$= \frac{73.5}{6}$$

Calculating Variance

$$Variance = \frac{\Sigma(X_i - \bar{X})^2}{N-1}$$
$$= \frac{73.5}{7-1}$$

$$Variance = \frac{\Sigma(X_i - \bar{X})^2}{N-1}$$
$$= \frac{73.5}{6}$$
$$= 12.3$$

Whew!


SUCCESS
Because you too can own this face of pure accomplishment