Political Science 30: Political Inquiry
Section 5

Taylor Carlson
tncarlson@ucsd.edu

▸ Link to Stats Motivation of the Week

"They've done studies, you know. 60% of the time, it works every time." — Brian, *Anchorman*

Learning Outcomes

By the end of section today, you should:

- Be able to calculate and interpret:
  - Margin of Error
  - 95% Confidence Interval
    - Proportion (nominal data)
    - Mean (interval/continuous data)
  - Explain the intuition behind a significance test
- Feel prepared to tackle Homework 2

Warm Up

- Name
- Reflect on the midterm. How do you feel it went for you? What strategies were effective? What will you do differently when studying for the final? Was there anything covered on the midterm that you still find confusing?

Announcements

- If you have questions on HW1, please come see me in office hours, especially if with Question 3 since you build on this for HW 2-4

Announcements

- If you have questions on HW1, please come see me in office hours, especially if with Question 3 since you build on this for HW 2-4
- See sample answer key for HW1 on TritonEd

Announcements

- If you have questions on HW1, please come see me in office hours, especially if with Question 3 since you build on this for HW 2-4
- See sample answer key for HW1 on TritonEd
- **HW 2 is due February 26th at the beginning of lecture**
  - Hard copy *and* Turn it In
  - Please staple!
  - You've already learned how to do all parts on this homework!

Announcements

- If you have questions on HW1, please come see me in office hours, especially if with Question 3 since you build on this for HW 2-4
- See sample answer key for HW1 on TritonEd
- **HW 2 is due February 26th at the beginning of lecture**
  - Hard copy *and* Turn it In
  - Please staple!
  - You've already learned how to do all parts on this homework!
- The course moves pretty quickly from here:

Announcements

- If you have questions on HW1, please come see me in office hours, especially if with Question 3 since you build on this for HW 2-4
- See sample answer key for HW1 on TritonEd
- **HW 2 is due February 26th at the beginning of lecture**
  - Hard copy *and* Turn it In
  - Please staple!
  - You've already learned how to do all parts on this homework!
- The course moves pretty quickly from here:
  - HW2 due Feb 26
  - HW3 due March 7
  - HW4 due March 14
  - Final exam March 23

My Advice:

- Stay on top of the readings, including the textbook
- Do the practice problems in the book each week (odds have answers in the back of the book)
- Do the practice problems I post online
- Come to lecture and section
- Come see us in office hours if you don't understand something
- Start the homework assignments early!

Quick Review: Normal Curves

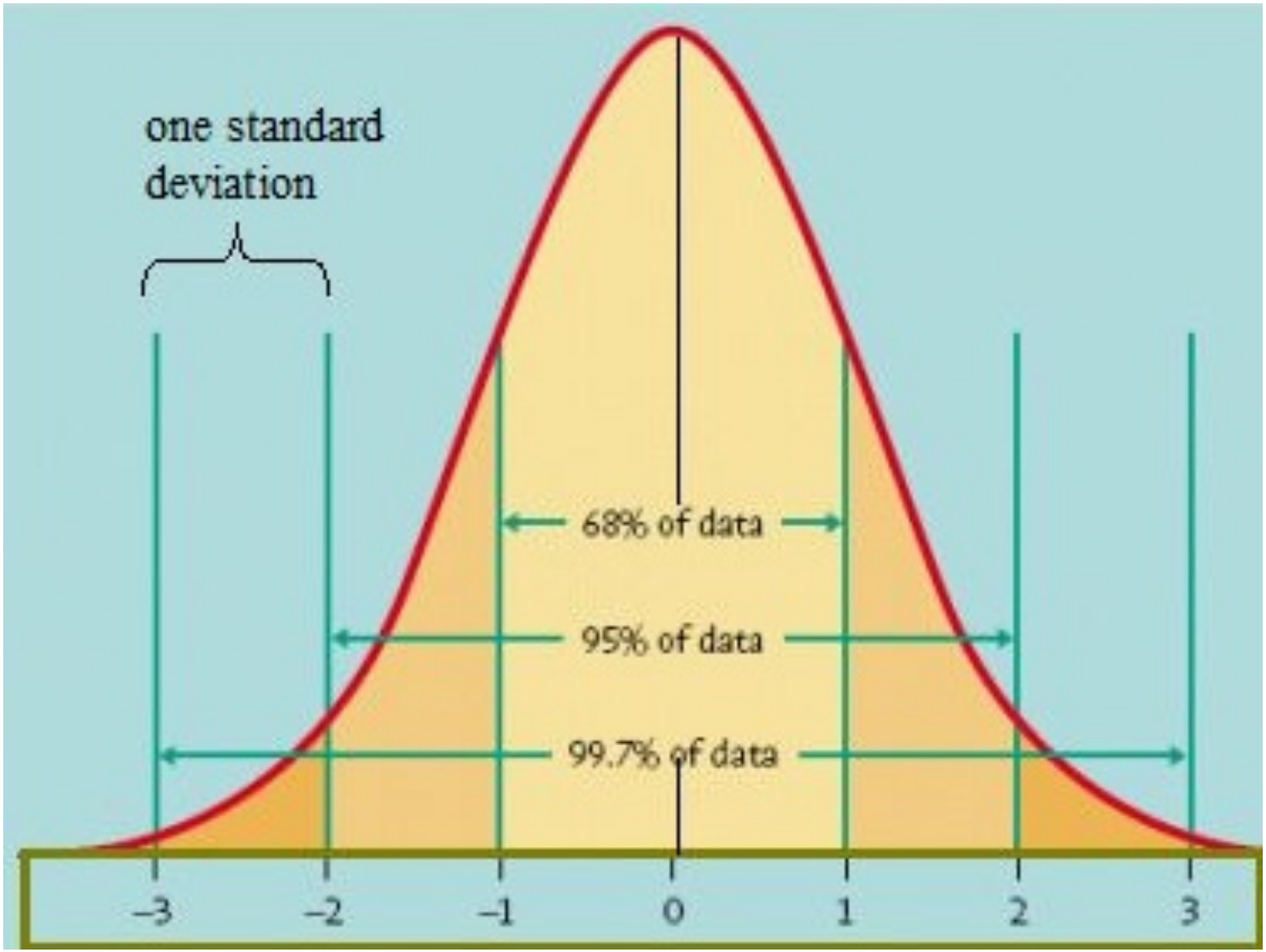
What do all normal curves have in common?

Quick Review: Normal Curves

What do all normal curves have in common?

- Symmetric
- Mean = Median
- Bell-shaped, with most of their density at the center

Quick Review: Normal Curves



Quick Review: Central Limit Theorem (CLT)

- If we draw lots and lots of samples and calculate the mean of each sample, these **sample means** will be **normally distributed** around the true population mean.

Quick Review: Central Limit Theorem (CLT)

- If we draw lots and lots of samples and calculate the mean of each sample, these **sample means** will be **normally distributed** around the true population mean.
- This means, that we can use the sample mean and the properties of the normal distribution to determine the **range** in which the **true population mean** resides.

Quick Review: Central Limit Theorem (CLT)

- If we draw lots and lots of samples and calculate the mean of each sample, these **sample means** will be **normally distributed** around the true population mean.
- This means, that we can use the sample mean and the properties of the normal distribution to determine the **range** in which the **true population mean** resides.
- That range is a ...

Confidence Interval

Confidence Interval

A confidence interval tells us the range of values in which the true population parameter resides at a given level of confidence (e.g. 95%, 99.7%).

Confidence Interval

Confidence Interval

What kind of data do we have?

Confidence Interval

What kind of data do we have?

- Nominal

Confidence Interval

What kind of data do we have?

- Nominal
  - Use **proportions**
  - $\hat{p} \pm 2 * \sqrt{\frac{(\hat{p})(1-\hat{p})}{N}}$

Confidence Interval

What kind of data do we have?

- Nominal
  - Use **proportions**
  - $\hat{p} \pm 2 * \sqrt{\frac{(\hat{p})(1-\hat{p})}{N}}$
- Interval / Continuous

Confidence Interval

What kind of data do we have?

- Nominal
  - Use **proportions**
  - $\hat{p} \pm 2 * \sqrt{\frac{(\hat{p})(1-\hat{p})}{N}}$
- Interval / Continuous
  - Use **means**
  - $\bar{X} \pm 2 * \frac{\hat{\sigma}}{\sqrt{N}}$

Confidence Interval of a Proportion (Nominal Data)

Awful Announcing recently conducted a poll on Twitter about whether individuals prefer the Winter or Summer Olympics. 2,353 individuals responded to the poll, and the results show that 36 percent prefer the Winter Olympics and 64 percent prefer the Summer Olympics. Assume that the respondents were a representative sample of Twitter Users.

Calculate the Margin of Error and the 95% Confidence Interval for the proportion of Twitter Users who prefer the Winter Olympics.

Confidence Interval of a Proportion (Nominal Data)

Calculate the Margin of Error:

Confidence Interval of a Proportion (Nominal Data)

Calculate the Margin of Error:

$$2 * \sqrt{\frac{(\hat{p})(1-\hat{p})}{N}}$$

Confidence Interval of a Proportion (Nominal Data)

Calculate the Margin of Error:

$$2 * \sqrt{\frac{(\hat{p})(1-\hat{p})}{N}}$$

$$2 * \sqrt{\frac{(.36)(1-.36)}{2353}}$$

Confidence Interval of a Proportion (Nominal Data)

Calculate the Margin of Error:

$$2 * \sqrt{\frac{(\hat{p})(1-\hat{p})}{N}}$$

$$2 * \sqrt{\frac{(.36)(1-.36)}{2353}}$$

$$2 * \sqrt{\frac{.23}{2353}}$$

Confidence Interval of a Proportion (Nominal Data)

Calculate the Margin of Error:

$$2 * \sqrt{\frac{(\hat{p})(1-\hat{p})}{N}}$$

$$2 * \sqrt{\frac{(.36)(1-.36)}{2353}}$$

$$2 * \sqrt{\frac{.23}{2353}}$$

$$2 * \sqrt{0.000097} \approx \mathbf{0.02}$$

Confidence Interval of a Proportion (Nominal Data)

Calculate the 95% Confidence Interval:

Confidence Interval of a Proportion (Nominal Data)

Calculate the 95% Confidence Interval:

$$\hat{p} \pm 2 * \sqrt{\frac{(\hat{p})(1-\hat{p})}{N}}$$

Confidence Interval of a Proportion (Nominal Data)

Calculate the 95% Confidence Interval:

$$\hat{p} \pm 2 * \sqrt{\frac{(\hat{p})(1-\hat{p})}{N}}$$

$$.36 \pm 2 * (0.01)$$

Confidence Interval of a Proportion (Nominal Data)

Calculate the 95% Confidence Interval:

$$\hat{p} \pm 2 * \sqrt{\frac{(\hat{p})(1-\hat{p})}{N}}$$

$$.36 \pm 2 * (0.01)$$

$$.36 \pm 0.02$$

Confidence Interval of a Proportion (Nominal Data)

Calculate the 95% Confidence Interval:

$$\hat{p} \pm 2 * \sqrt{\frac{(\hat{p})(1-\hat{p})}{N}}$$

$$.36 \pm 2 * (0.01)$$

$$.36 \pm 0.02$$

$$.34, \ .38$$

Confidence Interval of a Proportion (Nominal Data)

Calculate the 95% Confidence Interval:

$$\hat{p} \pm 2 * \sqrt{\frac{(\hat{p})(1-\hat{p})}{N}}$$

$$.36 \pm 2 * (0.01)$$

$$.36 \pm 0.02$$

$$.34, \ .38$$

Interpretation: We can conclude with 95% confidence that the true population proportion of Twitter users who prefer the Winter Olympics is between .34 and .38.

Confidence Interval of a Mean (Interval Data)

Chapter 6, Problem 8
We collect data from the FEC (Federal Election Committee) cataloguing all individual contributions to a given presidential primary campaign. The mean is $575 and the standard deviation is 250. If the sample size was 625 and we wanted to use a 95% confidence interval, what would we be able to confidently claim is the true mean contribution level for the population?

Confidence Interval Formula:

$$\bar{X} \pm 2 * \frac{\hat{\sigma}}{\sqrt{N}}$$

Confidence Interval of a Mean (Interval Data)

$$575 \pm 2 * \frac{250}{\sqrt{625}}$$
$$575 \pm 2 * \frac{250}{25}$$
$$575 \pm 2 * 10$$
$$575 \pm 20$$
$$555, 595$$

Interpretation: We can conclude with 95% confidence that the true population mean is between 555 and 595.

How might we turn these into significance tests?

Let's go back to the Olympics example.

- Expectation: The minority of Twitter users prefer the Winter Olympics.

How might we turn these into significance tests?

Let's go back to the Olympics example.

- Expectation: The minority of Twitter users prefer the Winter Olympics.
  - Null Hypothesis (H_0): the minority of Twitter users do not prefer the Winter Olympics

How might we turn these into significance tests?

Let's go back to the Olympics example.

- Expectation: The minority of Twitter users prefer the Winter Olympics.
  - Null Hypothesis (H_0): the minority of Twitter users do not prefer the Winter Olympics
    - $\hat{p}_{PrefersWinter} \geq 0.5$

How might we turn these into significance tests?

Let's go back to the Olympics example.

- Expectation: The minority of Twitter users prefer the Winter Olympics.
  - Null Hypothesis (H_0): the minority of Twitter users do not prefer the Winter Olympics
    - $\hat{p}_{PrefersWinter} \geq 0.5$
  - Alternative Hypothesis (H_1): the minority of Twitter users prefer the Winter Olympics

How might we turn these into significance tests?

Let's go back to the Olympics example.

- Expectation: The minority of Twitter users prefer the Winter Olympics.
  - Null Hypothesis (H_0): the minority of Twitter users do not prefer the Winter Olympics
    - $\hat{p}_{PrefersWinter} \geq 0.5$
  - Alternative Hypothesis (H_1): the minority of Twitter users prefer the Winter Olympics
    - $\hat{p}_{PrefersWinter} < 0.5$

How might we turn these into significance tests?

- Check the data – what does our observed proportion tell us?

How might we turn these into significance tests?

- Check the data – what does our observed proportion tell us?
  - 36% of Twitter users prefer the Winter Olympics, which is a minority.

How might we turn these into significance tests?

- Check the data – what does our observed proportion tell us?
  - 36% of Twitter users prefer the Winter Olympics, which is a minority.
- But, how sure are we that we're right?

How might we turn these into significance tests?

- Check the data – what does our observed proportion tell us?
  - 36% of Twitter users prefer the Winter Olympics, which is a minority.
- But, how sure are we that we're right?
- Recall the confidence interval we calculated earlier: (.34, .38)

How might we turn these into significance tests?

- Check the data – what does our observed proportion tell us?
  - 36% of Twitter users prefer the Winter Olympics, which is a minority.
- But, how sure are we that we're right?
- Recall the confidence interval we calculated earlier: (.34, .38)
- Recall that our null hypothesis is $\hat{p}_{PrefersWinter} \geq 0.5$

How might we turn these into significance tests?

- Check the data – what does our observed proportion tell us?
  - 36% of Twitter users prefer the Winter Olympics, which is a minority.
- But, how sure are we that we're right?
- Recall the confidence interval we calculated earlier: (.34, .38)
- Recall that our null hypothesis is $\hat{p}_{PrefersWinter} \geq 0.5$
- **Is 0.50 inside our confidence interval?**

How might we turn these into significance tests?

- Check the data – what does our observed proportion tell us?
  - 36% of Twitter users prefer the Winter Olympics, which is a minority.
- But, how sure are we that we're right?
- Recall the confidence interval we calculated earlier: (.34, .38)
- Recall that our null hypothesis is $\hat{p}_{PrefersWinter} \geq 0.5$
- **Is 0.50 inside our confidence interval?**
  - If YES $\rightarrow$ FAIL TO REJECT the null hypothesis
  - If NO $\rightarrow$ REJECT the null hypothesis