Introduction to Confidence Intervals

Alison Kleffner

- Background information on Inference
- What is a Confidence Interval?
- How to calculate a Confidence Interval?
  - Simulation
  - Theory
- Drawing Conclusions

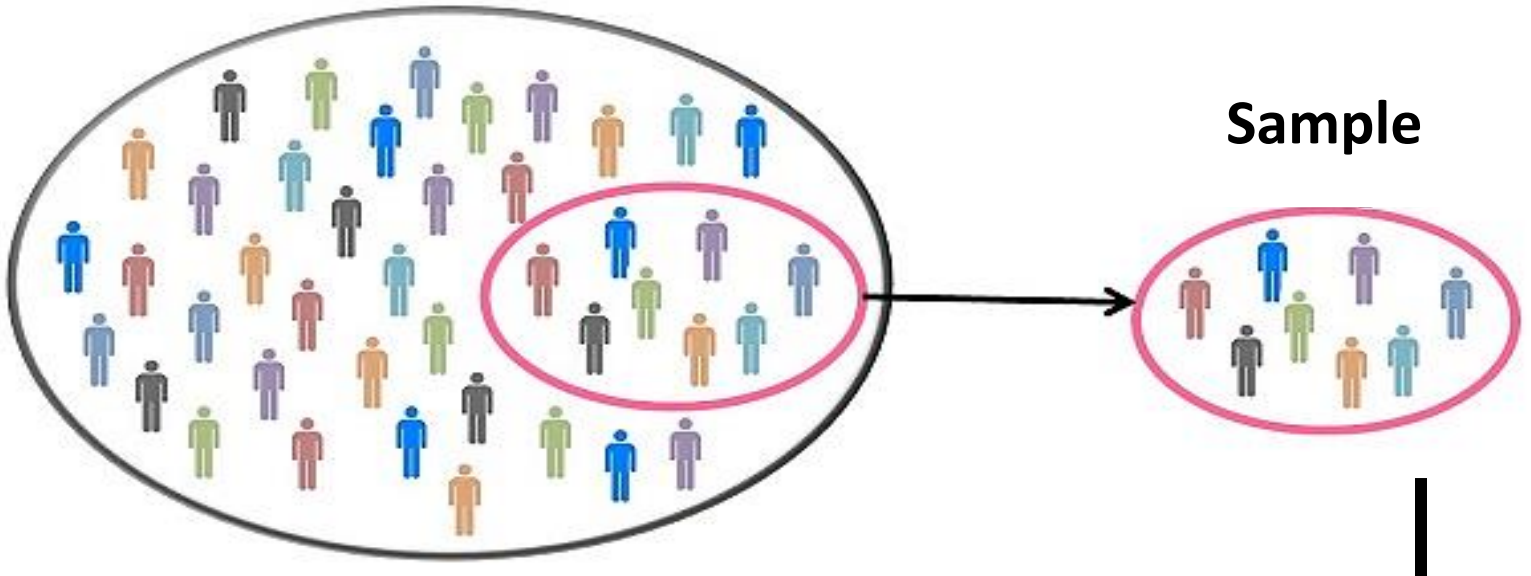
SCHEDULE

Background

- Generally, our goal in statistics is to gather information from a **sample** and then generalize the results from the sample to some larger **population** of interest.
- This is done by calculating **statistics** – numerical properties of the sample – and using them to estimate **parameters** – numerical properties of the population.
- This process is called **statistical inference**.
  - Is this value of the parameter plausible?

Overall Idea

We want to know about the population, and to do so, must work with the sample



Terminology

The whole purpose of statistics is to understand a set of information. This set of information is from a . . .

Population is the complete collection of ALL elements that are of interest for a given problem.

Parameter is the numerical measurement used to describe a population.

The population is often so big that obtaining all information about its elements is either difficult or impossible. So, we work with a more manageable set of data that we obtain from a . . .

Sample is sub-collection of elements drawn from a population.

Statistic is the numerical measurement used to describe a sample.

Example: Tire Story Falls Flat

A legendary story on college campuses concerns two students who miss a chemistry exam because of excessive partying but blame their absence on a flat tire. The professor allows them to take a make-up exam, and sends them to separate rooms to take it. The first question, worth 5 points is quite easy. The second question, worth 95 points, asks: Which tire was it?

- It has been conjectured that students select the front right more than would be expected with random chance.
- **Research Question –** *Do students pick which tire when flat in equal proportions?*
- To test this conjecture 32 students were asked, if they were in this situation, which tire would they say had gone flat. The results are as follows:

Left Front	Left Rear	Right Front	Right Rear
7	5	15	5

What is the population of interest?

What is the sample?

Example: Tire Story Falls Flat

What is the variable of interest?

Define parameter of interest.

Calculate the observed statistic.

Suppose students have no preference. How many in the sample would you expect to choose the right front tire?

Brief Overview of Hypothesis Testing

Summarize the **research question** and **parameter of interest**

Null hypothesis: what occurs due to random chance (written as H_0)

Alternate hypothesis: what the researcher is trying to support or what occurs if there is an effect (written as H_a or H_1)

We start by assuming null hypothesis is true and look for evidence to reject this idea

Alternative Hypothesis

One-sided test: testing for either less (<) than OR greater than (>) in the alternate hypothesis

Two-sided test: testing for “not equal to (≠)” in the alternate hypothesis

What are the hypotheses for the Tire Story Example?

Decision

Inference involves using the sample to decide one of the following two:

- The sample supports the research question – Reject H_0 in favor of H_a.
- The sample does *not* support the research question – Fail to reject H_0 in favor of H_a.

Must answer the following: How unusual would it be to a observe a result as extreme or more extreme by random chance alone (i.e. if the null is true)?

Decision commonly made using p-values or standardized statistics (z-test, t-test)

P-Values

$0 < p - value < 1$ (p-value is always between 0 and 1)

Compare p-value to significance level (denoted as α)

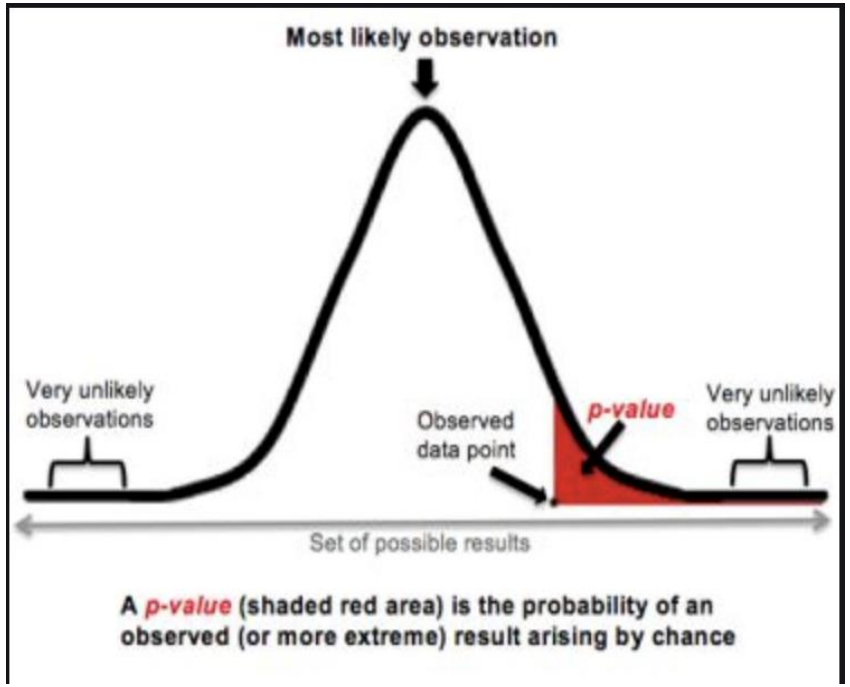
- α is probability of Type I error (incorrectly rejecting the null hypothesis)

Smaller p-value (p-value < α)

- Stronger evidence **against** the null hypothesis (reject)
- Support the alternative hypothesis

Larger p-value (p-value > α)

- Insufficient evidence against the null hypothesis (fail to reject)
  - We never Accept the Null Hypothesis!
- Cannot support the alternative hypothesis



New Goal: Estimation

What values of the parameter are plausible?

- Statisticians typically do this with a confidence interval?

Confidence Interval: A range of likely values for the parameter of interest

Who doesn't read books in America?

Who doesn't read books in America? | Pew Research Center

In a Pew Research Center a survey of U.S. adults conducted from Jan. 25-Feb. 8, 2021, asked the question: "Have you read a book in the past year, whether in print, electronic, or audio form?" Of the 1,502 US adults surveyed, 345 (23%) answered they have not read a book in the past year. They want to know if the proportion of US adult who have not read a book in the past year is different from 25%.

Ultimately, the Pew Research Center conducted this poll because they wanted to estimate the proportion of US adults in that (at the time of the survey) have not read a book in the past year. Keep this in mind as you answer the following questions.

Who doesn't read books in America

In a Pew Research Center a survey of U.S. adults conducted Jan. 25-Feb. 8, 2021, which asked the question: "Have you read a book in the past year, whether in print, electronic, or audio form?" Of the 1,502 US adults surveyed, 345 (23%) answered they have not read a book in the past year. They want to know if the proportion of US adult who have not read a book in the past year is different from 25%.

What is the population of interest?

What is the parameter of interest?

What is the sample?

What is the observed statistic of interest?

- **This is called a <u>point estimation</u>**

What if we took a different sample of 1502 US adults?

Who doesn't read books in America

In a Pew Research Center a survey of U.S. adults conducted Jan. 25-Feb. 8, 2021, which asked the question: "Have you read a book in the past year, whether in print, electronic, or audio form?" Of the 1,502 US adults surveyed, 345 (23%) answered they have not read a book in the past year. They want to know if the proportion of US adult who have not read a book in the past year is different from 25%.

Hypothesis

- Null: The long run proportion of US adults who have not read a book in the past year is equal to ____.
- Alternative: The long run proportion of US adults who have not read a book in the past year is different from ____.

Calculate your p-value using the **Theory Based Inference Applet**

- Click test of significance
- Make sure to use two sided
- Recall, n=1502 and $\hat{p} = 0.23$.

Based off your p-value, do you Reject (R) or Fail to Reject (FTR) at a significance level of 0.05?

When you are done, each group come up to fill out their section of the table.

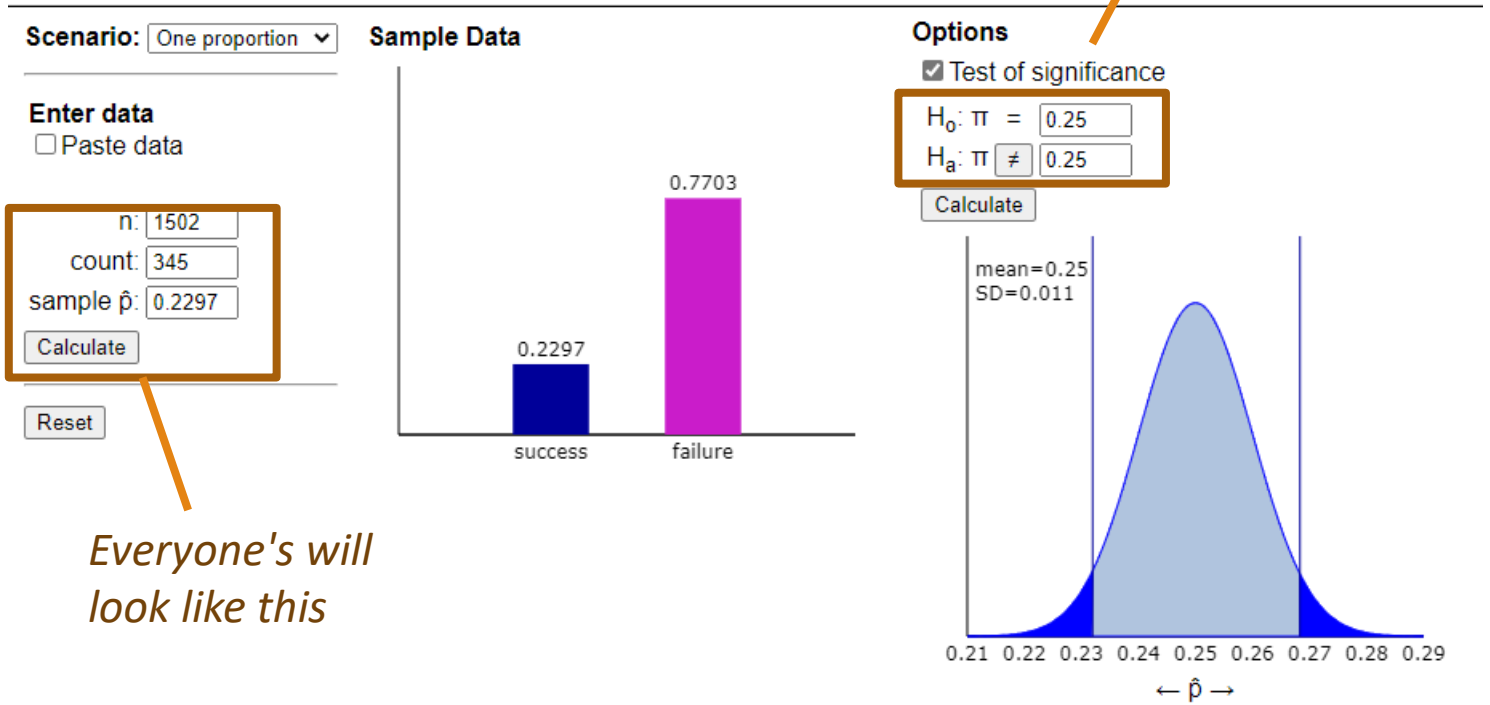
Theory Based Inference Applet

Calculate your p-value using the **Theory Based Inference**

- Click test of significance
- Make sure to use two sided
- Recall, n=1502 and $\hat{p} = 0.23$

Theory-Based Inference Applet (rossmanchance.com)

Each group will have a different value



Everyone's will look like this

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*
π	**0.18**	**0.19**	**0.2**	**0.21**	**0.22**	**0.23**	**0.24**	**0.25**	**0.26**	**0.27**	**0.28**
P-value											
R/FTR											

- What trends do we see?
- When we Fail to Reject are we saying the null or alternative is plausible?
- What is the interval of plausible values?
- *Interpretation:* We are 95% confident that the long-run proportion of US adult who have no read a book in the past year is between ____ and ____.

The General Form of a Confidence Interval

Notice, our confidence interval is centered around our observed statistic

Confidence interval = observed statistic ± margin of error

For this example, recall that the sample statistic is = 23%. The Pew Research Center reported the margin of error as follows:

> Sample: n=1,502 U.S. adults age 18 or older nationwide, including 1,202 cellphone interviews
> Interviewing dates: January 25, 2021 – February 8, 2021
> Margin of error: ± 2.9 percentage points for results based on Total [n=1,502]

Note that the confidence interval can be expressed as:

- 23% ± 2.9% (i.e., 20.1% ≤ π ≤ 25.9%).
- This agrees with what we saw under repeated sampling on the previous page.

Significance Levels (α)

- Significance Level + Confidence Level = 100%

$\alpha = 0.10$ relates to 90% confidence level
- Reject when p-value < 0.10
- Narrower than 95% confidence interval
- Rejecting null more often

$\alpha = 0.05$ relates to 95% confidence level
- Reject when p-value < 0.05
- Standard practice

$\alpha = 0.01$ relates to 99% confidence level
- Reject when p-value < 0.01
- Wider than 95% confidence interval
- Rejecting null less often

What if α=0.10?

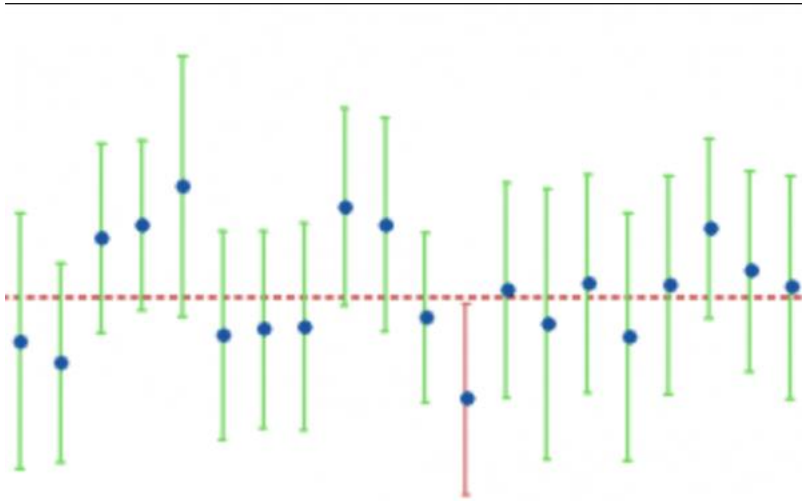
π	0.19	0.20	0.21	0.22	0.23	0.24	0.25	0.26	0.27
p-value	0.0001	0.004	0.0609	0.3645	0.9775	0.3497	0.0691	0.0074	0.0004
R or FTR									

Faster Ways to Calculate

Testing multiple parameter values can be very tedious

So, let’s use what we know about sampling distributions to help calculate this interval

What does confidence mean?



- What does “confidence” mean:
  - We expect 95% of all similarly constructed intervals to contain the true parameter value
- A 95% confidence interval indicates that 19 out of 20 samples (95%) from the same population will contain the population parameter.
- The percentage of confidence intervals that will contain the population parameter.

Interpretation

To interpret a confidence interval:

- With 95% confidence, the true parameter value falls within the confidence interval

Incorrect: There's a 95% chance that the true parameter values is within the confidence interval

- Why? The parameter value is fixed. "Chance" Implies that the parameter changes.

Example: With 95% confidence, the true proportion of US Adults that have not read a book in the past year is between 20.8% and 25.2%

Decision

If hypothesized value in interval – Fail to Reject

- Then the hypothesized value is a plausible value of the parameter

If hypothesized value not in interval – Reject

- Hypothesized value is not a plausible value of the parameter

Example:

- Hypothesized value: 0.25 -> within interval -> FTR
  - Likely value of the parameter
- We do not have evidence the true proportion of US Adults that have not read a book in the past year is different than 0.25.

Your Turn: Tire Story Falls Flat

A legendary story on college campuses concerns two students who miss a chemistry exam because of excessive partying but blame their absence on a flat tire. The professor allows them to take a make-up exam, and sends them to separate rooms to take it. The first question, worth 5 points is quite easy. The second question, worth 95 points, asks: Which tire was it?

- It has been conjectured that students select the front right more than would be expected with random chance.
- Remember $\hat{p} = \frac{15}{32}, n = 32$

Calculate the Confidence Interval using the formulas

Interpret the Confidence Interval

What decision can you make (R/FTR)? Why?

Review

A **Confidence Interval** is the range of likely value of the parameters.

Our confidence interval is centered around our observed statistic:

Confidence interval = observed statistic ± margin of error

A 95% confidence interval means we expect 95% of all similarly constructed intervals to contain the true parameter value