



STA130
TUT0110
Week4

TA: Gloria
shiyi.hou@mail.utoronto.ca

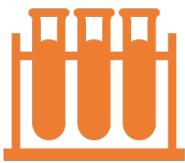
Reminder:

- Tutorial and OH are two separate thing.
- OH are Tues 2-5pm, Wed 12-3pm and Thurs 11am – 2pm @ HS390.

Addressing last week's feedback

- For the midterm/exam: need to know the meaning of the code, don't need to write the code from scratch
- To download the tutorial slides, :
<https://gloriahou1.github.io/STA130winter2020/>
- Time on writing activity
- Future Tutorials Structure
- Applications of Statistics/Data Science

Agenda



Vocabularies:
hypothesis testing



Group
discussion



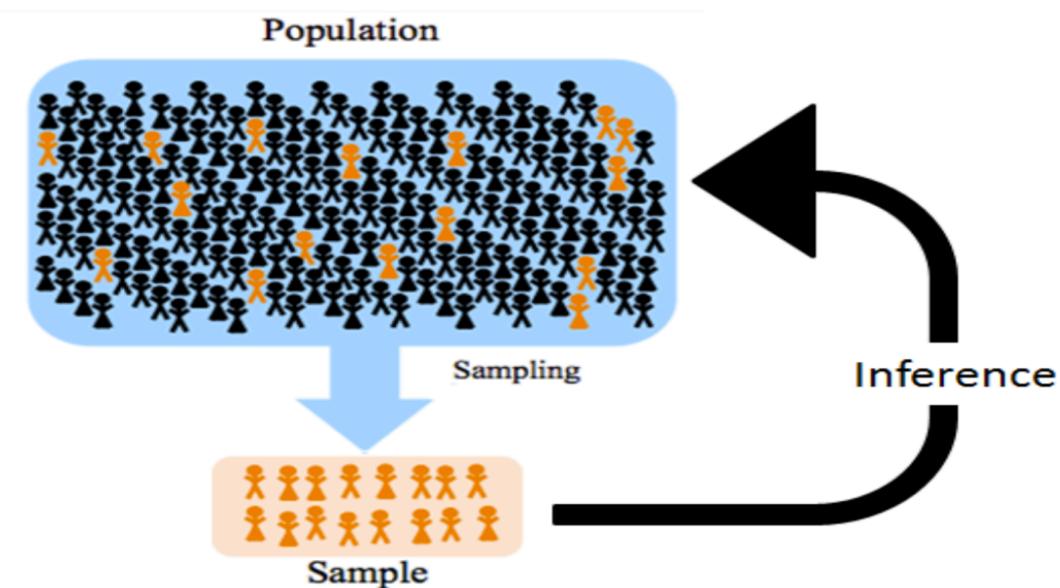
How to give oral
presentations



Group
presentation

Vocabularies

- random
- generalize
- inference



Vocabularies

- hypothesis testing
- null hypothesis
- alternative hypothesis
- assumption

Null hypothesis (H_0): The Wheel of Destiny spinner is fair

$$H_0 : p_{red} = 0.5$$

Alternative hypothesis (H_A or H_a or H_1): The Wheel of Destiny spinner is not fair

$$H_A : p_{red} \neq 0.5$$

```
pvalue <- sim %>%
  filter(p_red >= 0.64 | p_red <= 0.36) %>%
  summarise(p_value = n() / repetitions)
```

Vocabularies

- simulation
- simulation statistic
- sample
- population
- loop

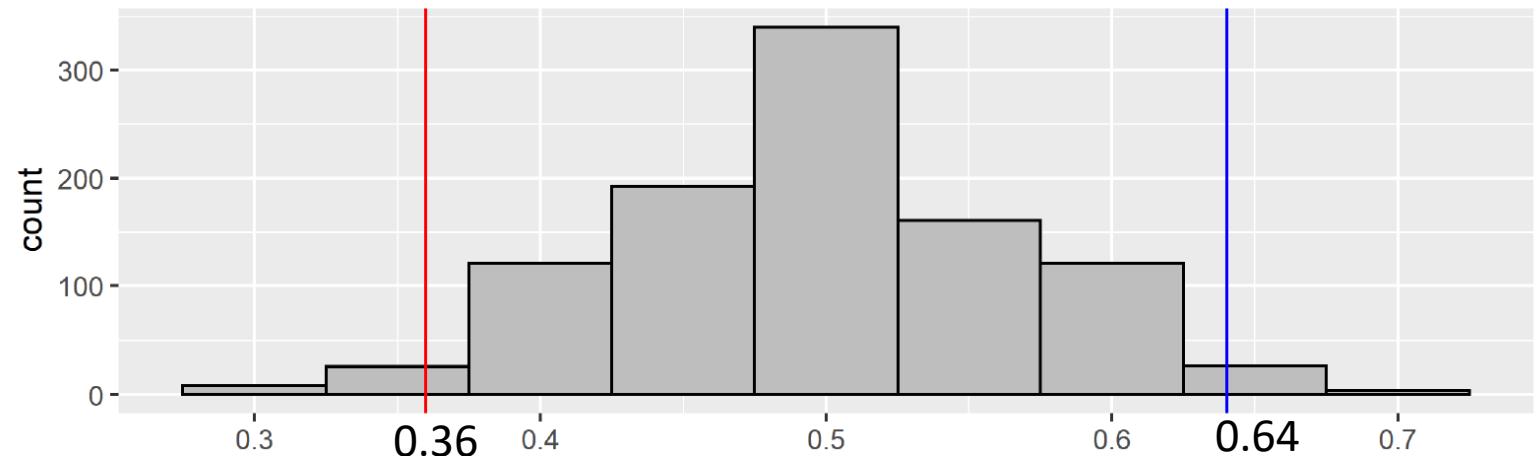
```
coin <- c("heads", "tails")
flips <- sample(coin,
                 size = 50,
                 prob = c(0.5, 0.5),
                 replace = TRUE)
flips

for (i in 1:repetitions){
  new_sim <- sample(c("red", "black"),
                    size = n_observations,
                    prob = c(0.5, 0.5),
                    replace = TRUE)
  sim_p <- sum(new_sim == "red") / n_observations

  simulated_stats[i] <- sim_p; # add new value to vector of results
}
```

Vocabularies

- statistic
- test statistic
- parameter
- observed value
- sampling distribution
- meaningful difference
- “at least as extreme”/ “as extreme or more extreme”



Vocabularies

- probability
- p-value
- statistically significant
- significance level
- strength of evidence

e.g. strong, moderate, weak, or no evidence

P-value	Evidence
$p\text{-value} > 0.10$	no evidence against H_0
$0.05 < p\text{-value} < 0.10$	weak evidence against H_0
$0.01 < p\text{-value} < 0.05$	moderate evidence against H_0
$0.001 < p\text{-value} < 0.01$	strong evidence against H_0
$p\text{-value} < 0.001$	very strong evidence against H_0

Interpret p-values

- In technical terms, a p-value is the probability of obtaining an effect ⁽¹⁾at least as extreme as ⁽²⁾the one in your sample data, ⁽³⁾assuming the truth of the null hypothesis.
- Important notes:
 - P-value can't be zero
 - Never accept the null-hypothesis.
 - Statistical significance level is predefined.
 - Evidence of statistical significance is either present or not present. (no "almost")

What does p-value mean?

- For example, suppose that a vaccine effectiveness study produced a p-value of 0.04. This p-value indicates that if the vaccine had no effect, you'd obtain the observed difference or more in 4% of studies due to random sampling error.
- Critically, p-values address only one question: **how likely** are your data, assuming a true null hypothesis?
- It does not measure support for the alternative hypothesis.

What doesn't p-value mean?

- Statistical significance does not mean practical significance.
 - Getting a low p-value and conclude a difference is statistically significant, doesn't mean the difference will automatically be important.
 - For example, a large clinical trial investigating a new weight loss drug found that people who took their drug loss 0.1 pounds more over the course of a year compared to those who took their competitor's drug ($p=0.0001$). While this is a statistically significant difference, it's likely not clinically meaningful
- Statistically significant just means a result is unlikely due to chance alone!



Group Discussion

- For Question 1, what would you expect to happen your p-value if you used 10 simulations versus 10,000 simulations? Explain.



Group Discussion

- Approximately 10% of the general population is left-handed. Suppose that the university is conducting a study to see if this percentage is the same among their students. This would help inform classroom renovations to ensure sufficient left-handed (and right-handed) seating. Suppose 500 students are randomly selected and asked whether or not they are left-handed. Suppose that 63 of these 500 students respond that they are left-handed. Say you used R to estimate the sampling distribution of the test statistic under the assumption that the prevalence of left-handedness among University of Toronto students matches the general population and you computed the p-value of the above hypothesis test based on this sampling distribution.
- Which of the following statements is/are valid description of the P-value you computed.



Group Discussion

- i. The probability that the proportion of U of T students who are left-handed matches the general population.
- ii. The probability that the proportion of U of T students who are left-handed does not match the general population.
- iii. The probability of obtaining a number of left-handed students in a sample of 500 students at least as extreme as the result in this study.
- iv. The probability of obtaining a number of left-handed students in a sample of 500 students at least as extreme as the result in this study, if the prevalence of left-handedness among all U of T students matches the general population.

What makes a good presentation?



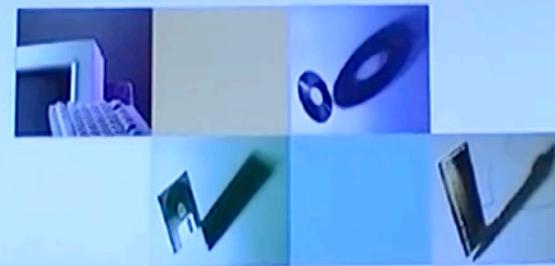
Advanced Presentation Skills



A Total Success Course
Presented by Warren Wint

What did you like or dislike about the presentation?

Advanced Presentation Skills



A Total Success Course
Presented by Warren Wint

What makes this a good presentation?

Advice for delivering a good presentation

THE 4 C'S:

- Calm
- Confident
- Clear
- Concise

Advice on content



What is the main message you want to get across?



Create an (organized) outline of your presentation



Define terms early



Make clear transitions between parts of your presentation



Make your data/figures meaningful



Summarize

Advice on delivery

 Be confident, make eye contact and avoid reading

 Avoid filler words – “ummm”, “like”, “you know”

 Speak slowly and it's ok to pause (and breathe!)

 Remember to enunciate all the parts of each word

 Practice! Practice! Practice!

Oral presentation

- Form a group of 4 people: prepare for a group presentation.
 - Each group will have 20 minutes to prepare
 - Each group will present for 5 minutes
 - Each group member needs to speak
 - Submit an outline for your presentation on Quercus

Prepare for your presentation

- Contextualize the problem
- Summarize the methods. E.g. State hypotheses; define the test statistic; etc.
- Summarize their findings
- Conclusion
- Limitations (optional, but good practice). E.g. sample size, study design issues, etc.



Presentation topic

- summarizing the one of their major research findings from Question 2.
 - You can pretend you've been asked by the Chair of the Dept of Statistical Sciences to present their work at the next faculty meeting.