Statistical Inference

- hypothesis testing.
- Interpreting sampling distributions.
- interpreting p-values.
- One of the most important concept in statistical science.
- **Is something we observe in data meaningful, or could it simply be due to chance?**

P-value

- a p-value is the **probability** of obtaining an effect at least as extreme as the one in your sample data, **assuming the truth of the null hypothesis.**
- The probability of observing what you have observed (your test statistic) or more extreme.

- In class, you saw example on hypothesis testing for a proportion p
- Here, the example on student's height in UofT is little bit different, just mean to give you a different prospective of hypothesis testing and p-value.
- What is the difference?
- For the example: "the proportion of individuals who tilt to the right when they kiss." shown in class, each data point has one binary value, either x=1 (Yes, they tilt to the right) or x=0 (No, they don't.)
- For student's height in UofT example, each data point has one value, which is the height of the student.

1. State hypotheses

- Suppose, your friend tell you the average height of students in UofT is 165cm.
- Null hypothesis (H_0): $\mu_0 = 165\text{cm}$
- Alternative hypothesis (H_A): $\mu_0 \neq 165\text{cm}$

2. Test statistic

- Is a number, calculated from the data
- For UofT student height example, the test statistic we'll use is the average height of your sample of students.
- $\bar{X} = \frac{1}{n}\sum_i x_i$
- Parameter of the population?
- True average height of student: μ
- Do the Sampling and get: $\bar{X} = 172cm$ This is your test statistic.

3. Sampling distribution

- What is the sampling distribution assuming H_0 is true?
- Additional assumption about your population: we assume the height are normally distributed with mean 165cm, and some known variance σ^2
- $\bar{X} \sim N(\mu_0, \frac{\sigma^2}{n})$
- $H_0: \mu_0 = 165\text{cm}$ v.s $H_A: \mu_0 \neq 165\text{cm}$

- Difference between student's height example and the in-class proportion example:
- In-class example, you estimate a sampling distribution by
- **Simulation**: Randomly generate samples under the assumption that couples have no preference (i.e. they are equally likely to tilt their heads to the le or right.) For each sample, calculate the proportion who tilt their head to the right.
- Repeatedly sampling, draw a dot on the plot (which is your sampling distribution), Each dot represents the result for one simulation.
- While, here we assume the students' height are normally distributed.

4. Evaluate the evidence against H_0

- If H_0 is true, how unusual is our observed value?
- Normalize X and check standard normal table. $\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$
- You can compute $\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$, and after that, by checking the table, you get the p-value.

5. Make a conclusion

- Suppose, you calculate p-value = 0.11, what conclusion can you make?

P-value	Evidence
p-value > 0.10	no evidence against H_0
0.05 < p-value < 0.10	weak evidence against H_0
0.01 < p-value < 0.05	moderate evidence against H_0
0.001 < p-value < 0.01	strong evidence against H_0
p-value < 0.001	very strong evidence against H_0

- You can’t reject your friend’s claim!

Interpretation of p-value.

- 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.
- P-value is only meaningful in the sense of **statistically significance.**
- Statistical significance does not mean practical significance.
- It should only serve as a useful way to look at your data, but when you are trying make decision in practice, you must be very careful. Because sometimes, the conclusion you make based on the p-value, might be wrong, or doesn’t mean anything in practice.

- Example:
- a 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 (and get p=0.0001).
- While this is a statistically significant difference, it's likely not clinically meaningful. Statistically significant just means a result you obtain is unlikely due to chance!
- H_0: weight_loss_1 – weight_loss_2 < 0.1 (our drug is no difference than our competitor's drug)
- H_a: weight_loss_1 – weight_loss_2 > 0.1 (our drug is better!)

- If the above material doesn't make sense to you, don't worry too much, since this is a really important concept in stats, we may say that if you understand p-value and hypothesis testing, you understand this subject. (Will formally learn it in STA261)
- However, You should, at very least, remember the following:
- 1. You can never accept your hypotheses! You can only reject or fail to reject.
- 2. Evidence of statistically significance is either present or it's not. Never say that something is "almost" statistically significant.
- Often, we assess statistical significance based on a threshold of $p=0.05$. If $p>0.05$, then the chance of observing your outcome due to chance alone was greater than 5% (5 times in 100 or more).
- In this case, you would fail to reject the null hypothesis and would not accept the alternative hypothesis.

If we have time

- Some comment about the writing.
- Take 5 minutes to read the following writings, and think about the question:
- **1) is this a good or not so good example?**
- **2) what makes this a good/not so good example.**

- *Study skills and students' satisfaction with their performance positively affect their academic achievement. The current research was carried out to investigate the correlation of study skills with academic achievement among the medical and pharmacy students in 2013. This descriptive-analytical study was conducted on 148 students of basic medical sciences and pharmacy through convenience sampling. Data were collected by a valid and reliable questionnaire, consisting of two sections: Demographic information and questions about daily study hours, study skills in six domains, and students' satisfaction with study skills. Collected data sets were analyzed by SPSS-16 software. In total, 10.9% of students were reported to have favorable study skills. The minimum score was found for preparation for examination domain. Also, a significantly positive correlation was observed between students' study skills and their Grade Point Average (GPA) of previous term (P=0.001, r=0.269) and satisfaction with study skills (P=0.001, r=0.493). The findings indicated that students' study skills need to be improved. Given the significant relationship between study skills and GPA, as an index of academic achievement, and satisfaction, it is necessary to promote the students' study skills. These skills are suggested to be reinforced, with more emphasis on weaker domains.*

- *This study explored the pattern of video game usage and video game addiction among male college students and examined how video game addiction was related to expectations of college engagement, college grade point average (GPA), and on-campus drug and alcohol violations. Participants were 477 male, first year students at a liberal arts college. In the week before the start of classes, participants were given two surveys: one of expected college engagement, and the second of video game usage, including a measure of video game addiction. Results suggested that video game addiction is (a) negatively correlated with expected college engagement, (b) negatively correlated with college GPA, even when controlling for high school GPA, and (c) negatively correlated with drug and alcohol violations that occurred during the first year in college. Results are discussed in terms of implications for male students' engagement and success in college, and in terms of the construct validity of video game addiction.*

- **Notes about the 1st example: It clearly summarizes all components of a traditional abstract and makes it easy for readers to understand the focus of the research.**

- **Notes on the 2nd example: It doesn't specifically state why the problem is worth researching, though it is implied as the study focuses on addiction.**

- More examples with explanations for what makes them good/ poor can be found here:
- https://www.kibin.com/essay-writing-blog/10-good-abstract-examples/

When communicating about your work, it's important to include 4 critical items:

- The purpose. What is it that you're studying? Why should we care about the analytical work you've done?
- A summary of the methods you used. What did you do? Why did you do it this way?
- A summary of the results. We don't need to know everything you found – only the most critical things relating to your purpose! Remember, sometimes less is more!
- A conclusion. What is your take away message? Remember, a conclusion is not the place to present new findings.

- *Most importantly, your "story" should be clear, concise, cohesive, and complete! Like a real story, it should have a clear beginning, middle and end.*
- Complete — cover the important parts of the project, study, or analysis
- Concise — contain no excess wordiness or unnecessary information
- Clear — readable, well organized, and not too jargon-laden
- Cohesive — flows smoothly between the parts

Group discussion

- *For Question 1, what would you expect to happen your p-value if you used 10 simulations versus 10,000 simulations? Explain.*
- You may also discuss about any parts of the homework question, help each other if you cannot figure out how to solve some question or how to write code regarding those question.

In-class writing exercise

- Imagine you work as a business analyst for Coca Cola. Your CEO has asked you to deliver a summary of these new research findings (for question 1) by the end of the day. It's already 3:30 pm and it's a Friday!

  Remember, your CEO is a busy person. They only want the most important information and they don't want to read more than half a page of text. Use visuals to help get key points across, if you can. The CEO has only limited statistical background, so make sure everything is clear and makes sense! Remember to start with the purpose - your CEO is busy, they probably forgot what this was about! Also make sure to include a complete, but concise, summary of the methods, key results, and conclusion. You should also state how confident you in these research findings – remember, you're the statistics expert and the CEO is counting on *you* to summary this research!