

# STA130 Winter 2020

(Materials used in this presentation are provided by the U of T Statistical Sciences Department.

This presentation was prepared by Vivian Ngo.)

**[Github.com/vivianngo97/STA130-Winter-2020](https://github.com/vivianngo97/STA130-Winter-2020)**

**viv.ngo@mail.utoronto.ca**

# Agenda

- Vocabulary
- Ice breaker
- Group discussion
- Oral Presentations Preparation
- Oral Presentations
- Ticket out the door

# Vocabulary

- Type 1 and 2 error
- Comparing two population means/ proportions
- One- and Two-sample test
- Approximate permutation test
- *Refer to last week's list for hypothesis test vocabulary*

# Vocabulary: Type I error

- False positive!!!
- “false alarm”
- Probability of a false positive (type I error) is denoted as **alpha**
- Alpha is typically chosen to be 5%
- **Low tolerance** for this type of error!



# Vocabulary: Type II error

- False negative!
- “missed opportunity”



# Vocabulary: Type I and Type II error

		Reality	
		True	False
Measured or Perceived	True	Correct 😊	Type 1 error False Positive
	False	Type 2 error False Negative	Correct 😊

	Disease or Condition	No Disease or Condition
	A True Positive	B False Positive
	C False Negative	D True Negative
	Test Positive	Test Negative

# Vocabulary: A note about the errors

- In practice, we don't know if we committed one of these errors
- More tests -> more likely to have a type I error, but you won't know which test had it
- Future stats courses: learn ways to control the probabilities of getting these errors
- Which error do you think is more harmful/serious? Why?

# Real-world example: Somebody is being convicted of murder

- What is the null hypothesis?
- What is the alternative hypothesis?
- How could a type I error occur in this example?
- How could a type II error occur in this example?

Null hypothesis

Alternative  
hypothesis

Type 1 error

Type 2 error

# More examples...

- Do weight loss programs work?
- Let  $c$  = change in weight after using the weight loss program

Null hypothesis	They are not effective	$c = 0$
Alternative hypothesis	They are effective	$c \neq 0$ ( $c$ is not 0)
Type 1 error	Claiming that the weight loss program is effective even though it is not (false advertising)	
Type 2 error	Not being able to prove the effectiveness of the program even though it does work	

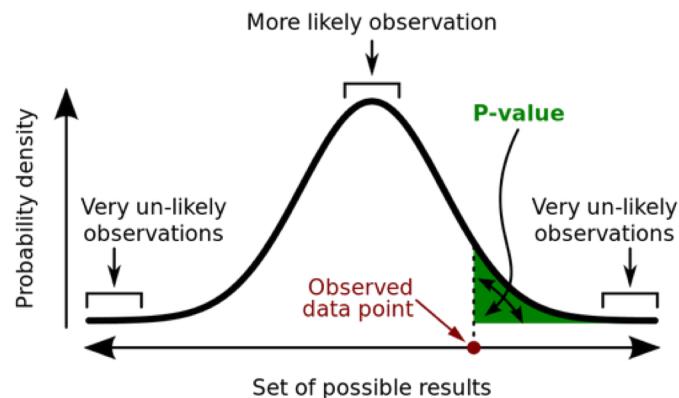
# More examples...

- Does smoking affect life expectancy?
- Let  $x$  = change in average life expectancy
- OR, let  $S_{mean}$  = average lifespan of smokers,  $N_{smean}$  = average lifespan of nonsmokers

Null hypothesis	Smoking does not affect life expectancy	$x = 0$ $S_{mean}=N_{smean}$
Alternative hypothesis	Smoking affects life expectancy	$x \neq 0$ $S_{mean} \neq N_{smean}$
Type 1 error	Claiming that smoking affects life expectancy when it does not	
Type 2 error	Not being able to show that smoking affects life expectancy	

# Back to p-values

- In question #2,
- Do you have stronger evidence against the null hypothesis of no difference in **median** anxiety scores for the two groups, or more evidence against the null hypothesis of no difference in **mean** anxiety scores?
- Hint: compare p-values



# Ice breaker!



# Group Discussions

- What makes a good presentation?



# Oral presentation tips

- Don't bring large distracting devices to present with
  - Perhaps use index cards or small notes
- Don't block visual aids
- Be mindful of your body language when other members are speaking
- Transition: "Next, my team member \_\_\_\_ will discuss..."
- Let us know when you are done: "That concludes our presentation. Thank you for listening..."
- Be clear and concise
- Don't trail for too long
- It's okay to take a break to gather your thoughts
- Make eye contact to everyone in the room
- Prepare visual aids beforehand
- Practice

# Oral Presentation Examples

- Tips for giving a great oral presentation: *Delivery*
  - 1. Be confident, make eye contact and avoid reading
  - 2. Avoid filler words – “ummm”, “like”, “you know”
  - 3. Speak slowly and it’s ok to pause (and breathe!)
  - 4. Remember to enunciate all the parts of each word
  - 5. Practice! Practice! Practice!

# Oral Presentations Preparation

- In groups of 3-4:
- Prepare a **5-minute presentation** summarizing one of the following research findings
- **Submit a draft outline**, one per group, on Quercus (for my reference)
- Your oral presentation, like a written summary, should include the following components:
  - 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.

# Oral Presentations!

- (a) A health survey asked 200 individuals aged 20-45 living in Toronto to report the number minutes they exercised last week. Researchers were interested in determining whether the average duration of exercise differed between people who consume alcohol and those who do not consume alcohol. Assume the researchers who conducted this study found that people who drank alcohol exercised, on average, 20 minutes per week. In contrast, people who did not drink alcohol exercised 40 minutes per week, on average. The researchers reported  $p=0.249$ .

# Oral Presentations!

- (b) A study was conducted to examine whether the sex of a baby is related to whether or not the baby's mother smoked while she was pregnant. The researchers used a birth registry of all children born in Ontario in 2018, which included approximately 130,000 births. The researchers found that 4% of mothers reported smoking during pregnancy and 52% of babies born to mothers who smoked were male. In contrast, 51% of babies born to mothers who did not smoke were male. The researchers reported a p-value of 0.50.

# Oral Presentations!

- (c) Based on results from a survey of graduates from the University of Toronto, we would like to compare the median salaries of graduates of statistics programs and graduates of computer science programs. 1,000 recent graduates who completed their Bachelor's degree in the last five years were included in the study; 80% of the respondents were female and 20% were male. Among statistics graduates, the median reported income was \$76,000. Among computer science grads, the median reported income was \$84,000. The researchers reported  $p=0.014$ .

# Oral Presentations!

- (d) A team of researchers were interested in understanding millennial's views regarding housing affordability in Toronto. The team interviewed 850 millennials currently living in Toronto. 84% reported that they felt housing prices were unaffordable in the city. Suppose the researchers were interested in testing whether this proportion was different from a study published last year, which found that 92% of millennials reported that housing costs were unaffordable. The researchers reported a p-value of 0.023.

# Oral Presentations!

- (e) Suppose a drug company was interested in testing a new weight-loss drug. They enrolled 20,000 participants and assigned 10,000 to take their new drug, SlimX, and 10,000 to take a placebo. The researchers found that over 2 months, participants who took SlimX lost, on average, 5 lbs. In comparison, the control group lost 4.5 lbs during the same time. The researchers reported a p-value of <0.0001.

# Ticket out the door

- Write down one aspect you thought you did well in this week's oral presentation and one aspect you think you could improve