STA130 T0209

Week 10: Confounding and Ethics

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

This presentation was prepared by Vivian Ngo.)

Agenda

- Reminder: final project
  - Emails with group number
  - Know your evaluation schedule! (for evaluating other groups) – on Quercus
- Material, vocabulary, homework discussion
- Group work and presentations
- Remaining time: final poster project

Material and Vocabulary Review

- Observational study
- Confounding
- Confounded
- Association
- Causation
- Experiment
- Randomized (controlled) trial
- Control group
- Ethics/ ethical issue/ ethical consideration
- Consent
- Information
- Comprehension
- Voluntariness
- Transparent

Material and Vocabulary Review

- **Observation study** vs **experiment**
  - researcher simply observes natural processes VS researcher imposes certain conditions in order to investigate the effect
- E.g. Interested in investigating the relationship between physical activity and weight gain among adults
  - Observational study:
    - identify a group of people, hopefully representative of the population you're interested in. For example, using a health survey that has sampled all of the Ontario population. Investigate whether less physically active people were more likely to gain weight than less active people (e.g. over 5, 10, ?? years). Could also look at weight gain difference.
  - Experiment:
    - identify a group of people, hopefully representative of the population you're interested in. Assign at random some people to be physically active and others to not be physically active (called the control group). Investigate whether the physically active group were less/more likely to gain weight (over X years) than the less physically active. Could also look at weight gain difference.

Material and Vocabulary Review

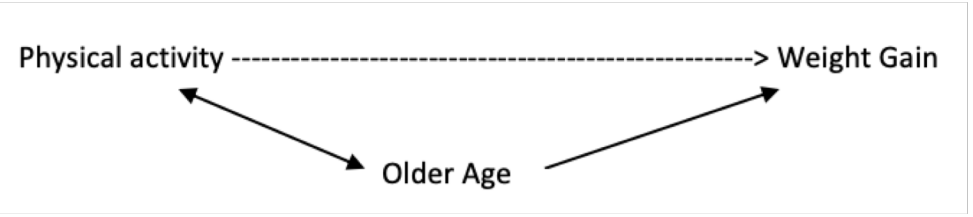
- **Association** vs **causation**:
  - Association does not mean causation!
  - Association -> two things are related.
  - Causation means that one thing *causes* the other thing to occur.
  - Observational studies only look at association
  - RCTs help investigate causation
  - However**, causation is a very strong statement** to make and several factors must be considered, such as the Bradford-Hill Criteria discussed in class. You must be *very* careful when making causal statements.

Material and Vocabulary Review

- **Confounding**
  - A variable that influences both the dependent variable and independent variable, causing a spurious association
  - A variable that makes it seem like a relationship between X and Y exists, but in reality, it is only (or mostly) because of the confounder.

Material and Vocabulary Review

- **Confounding**
- In randomized controlled trials, confounding should theoretically not exist because groups are assigned at random.
  - This means that all covariates should be balanced and both groups should be more or less exactly the same – except one is assigned to receive the treatment and the other is not.
- In observational studies, this is not the case.
  - There are MANY reasons why people engage in different behaviors and these may be related to both your exposure and outcome of interest.



Material and Vocabulary Review

- **Confounding**
- E.g. suppose you were interested in investigating whether a Bachelor degree from certain universities (e.g. UofT, Oxford, McMaster, UWaterloo, Harvard, Ryerson, etc.) were associated with higher job earnings. Are there any confounders you can think of, why may this may be a confounder?

Material and Vocabulary Review

- **Confounding**
- E.g. suppose you were interested in investigating whether a Bachelor degree from certain universities, e.g. UofT, Oxford, McMaster, UWaterloo, Harvard, Ryerson, etc., were associated with higher job earnings. Are there any confounders you can think of, why may this may be a confounder?
  - Programs
  - Student's personalities

Material and Vocabulary Review

- **Ethics**
  - Research involving humans and animals requires ethics board approval. Such work must reach ethical and scientific standards.
  - E.g. We could not conduct an experiment where we assign to people to smoke or not smoke. Why?
  - E.g. Homework #4: A data scientist compiled data from several public sources (voter registration, political contributions, tax records) that were used to predict sexual orientation of individuals in a community.
    - What ethical considerations arise that should guide use of such data sets?

Oral presentations

THE 4 C'S: Calm; Confident; Clear; Concise

Tips for giving a great oral presentation: Content

1. What is the main message you want to get across?
2. Create an (organized) outline of your presentation
3. Define terms early
4. Make clear transitions between parts of your presentation
5. Make your data/ figures meaningful
6. Summarize

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!

Group presentation

- Prepare a 5 minute oral presentation based on the following topics (next slide)
- When not presenting:
  - One person from each group evaluate other students, upload rubric to Quercus
  - Write down any questions you have
  - Presentation rubric on Quercus and Github
- **You should state the purpose of your presentation, briefly explain the methods used, key results and include an appropriate conclusion**

Group presentation

- **Group 1. Consider question 1a, specifically 1a iv).**
- Interpret the p-value of this test to compare the mean improvement for Lumosity versus crossword puzzles. How does it compare to the p-value estimated using the randomization test earlier in this question? Is this surprising? Why or why not? Make sure to explain the methods you used.
- **Group 2. Consider question 1b.**
- What type of study did Hardy et al. conduct? What were the conclusions? Are there any limitations?

- **Group 3. Consider question 1c.**
- Is age a confounder of this association? Why or why not?

Group 4. Consider Question 1d and 2.

- What ethical considerations did Hardy et al. make in their study? Why were these steps necessary? Consider the Statistical Society of Canada (SSC) Code of Ethical Statistical Practice, what practices should you consider while completing your poster project?

	4 (Excellent)	3 (Good)	2 (Adequate)	1 (Poor)
Context	The context and connection to the problem are clear.	Some context was provided and all variables/concepts were mentioned. Some aspects were not clear.	Very little context was provided and only some variables/ concepts were mentioned.	No context and mentioning of any variables/ concepts covering in this week's materials.
Structure	Well organized, follows a logical structure.	The organization follows some logical structure.	Some structure but difficult to follow.	There is no structure, very difficult to follow.
Conclusion	There is a clear central idea and the conclusion is correct.	A central idea or conclusion is present. The conclusion might be incorrect.	The central idea or conclusion is weak and not supported.	The central idea or conclusion is missing. Incorrect conclusion.
Transitions	The progression is logical. Effective use of transitions.	The progression is controlled. The use of transitions is mostly meaningful.	Minor disruptions in flow and weak transitions.	Weak progression and lack of transitions.
Vocabulary	Good use of statistical terms and appropriate choice of words.	Use of statistical terms and phrases mostly correct, demonstrates understanding of concepts.	Some use of statistical terms/ phrases and some understanding of concepts demonstrated.	Inaccurate or incorrect use of statistical terms or phrases and a lack of understanding statistical concepts.
Presentation Skills	Regular eye contact with all parts of the audience. The audience was engaged. The presenter held the audience's attention. Appropriate speaking volume & body language. Good pace.	Somewhat regular eye contact or eye contact with some of the audience The audience was mostly engaged. The presenter mostly spoke at a suitable volume. Spoke too quietly at times. Some fidgeting. Going too fast/slow.	Focused on only one or two members of the audience. Sporadic eye contact. The audience was not engaged. Speaker could be heard by only some of the audience. Body language was distracting.	Minimal (or no) eye contact. The audience was never engaged. The presenter did not speak clearly. Presenter was very difficult to hear.
Preparedness/ Participation	Extremely prepared and rehearsed. The presenter was confident.	Mostly prepared but some dependence on or reading off of notes. The presenter seemed fairly confident.	The presenter was not well prepared. The presenter did not seem confident.	Evident lack of preparation/rehearsal. Complete dependence on notes.

Last tutorial!