STA 130 Tutorial 10

David Veitch

University of Toronto

daveveitch.github.io

Mar 29, 2019

Agenda

1. Vocabulary
   - What is a Linear Model - Review
   - Observational Study
   - Establishing Causation
   - Confounding Variables
   - Kahoot
2. Presentations
3. Remaining Time - Work on Project

What is a Linear Model - Review

Example of linear model (linear in β's and hence a closed form solution to optimal β parameters): $y_i = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i}^2 + \epsilon_i$

Example of a nonlinear model: $y_i = \frac{\beta_0 + \beta_1 x_{1,i}}{1 + \beta_2 x_{2,i}}^{\beta_3} + \epsilon_i$

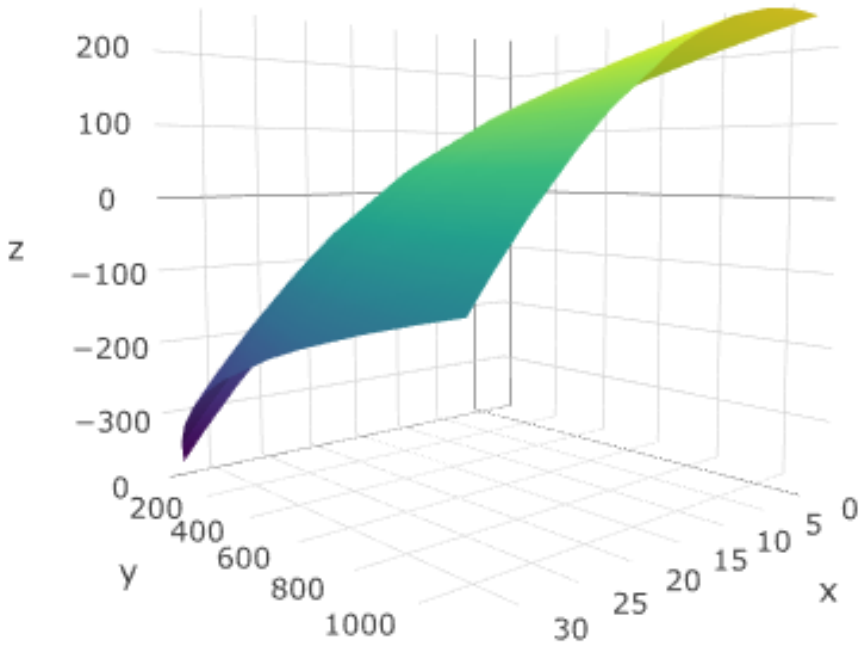


Figure: Chart of $y = -2.55 + 7.45x - .31y^2$

Observational Study

In an **observational study** varaiables are observed (measured and recorded) without manipulation of varibles or conditions by the researcher.

In an **experiment** variables and/or conditions are manipulated by the researcher and the impact on other variable(s) is measured and recorded.

Question: Come up with an example of an observational study, and an experiment, that could be conducted to study UofT students' academic success.

Establishing Causation

Data from an observational study we can only conclude _____ between variables not _____?

Establishing Causation

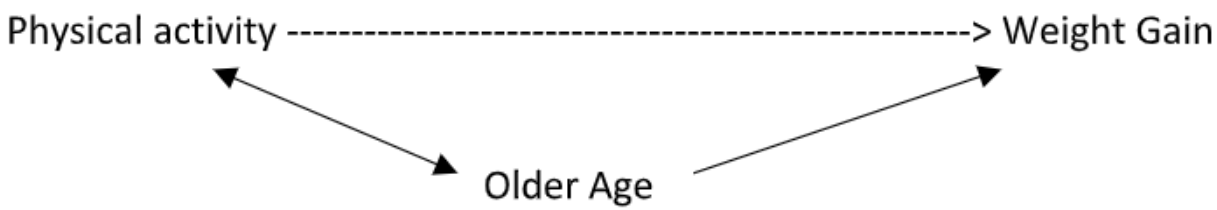
Data from an observational study we can only conclude **association** between variables not **causation**?

However in an experiment if there is a significant difference in the outcome betweeen two groups, we **may have** evidence that there is a **causal relationship** between treatment and outcome.

Confounding Variable

A variable that influences both the dependent variable and the independent variable causing a spurious association.

For example we may have a model where we explain weight gain by physical activity. However if we do not account for old age we may not be able to observe the true relationship between physical activity and weight gain.



Confounding Variable

Question: You are interested in investigating whether a Bachelor degree from UofT is associated with higher job earnings. What are some confounding variables you could think of, why are they confounders?

Kahoot

Presentation Outline

Presentation should include:

- Purpose of Presentation
- Methods Used
- Key Results
- Conclusion

Presentation Questions

- **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.
- **1b** What type of study did Hardy et al. conduct? What were the conclusions? Are there any limitations?
- **1c** Is age a confounder of this association? Why or why not?
- **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?

Remaining Time - Work on Project