Discussion 3

Week of 02.02

UVA Batten

Agenda

- Announcements
- Conditional Expectation
- Reviewing Mean Comparison and Basic Regressions
- Understanding Bias

Announcements

- Quiz 2 due Monday, Feb. 2 at 9pm (Quiz 3 due Monday, Feb. 9 @ 9 pm)
- Homework 3 due Monday Feb. 9 at 9pm

- We are looking for SDAC notetakers in Lecture 002 (11:00am) and Discussion 103 (Tuesday). If you are interested, let us know.
- If you see a 🅿️ in the slides, it means there's a practice question.

Conditional probability is the likelihood of a certain outcome (x) given a precondition (y) that narrows the set of possible outcomes.

$$P(X|Y)$$

Conditional probability is the likelihood of a certain outcome (x) given a precondition (y) that narrows the set of possible outcomes.

P(UVA)

P(UVA | FRANK BATTEN SCHOOL of LEADERSHIP and PUBLIC POLICY)

P(FRANK BATTEN SCHOOL of LEADERSHIP and PUBLIC POLICY)

P(FRANK BATTEN SCHOOL of LEADERSHIP and PUBLIC POLICY | UVA)

Conditional expectation is the expected value (read: average) of a measurement. The notation is the same as conditional probability.

🅿️ *Decode the statistics included in this NYTimes article.*


The New York Times

OPINION

GUEST ESSAY

Steve Bannon Thinks Zohran Mamdani Is a Genius. It's Not a Feint.

$$\mathbb{E}[\text{Voted}_i | \text{Age 18 to 29} = 1, \text{Year}_t = 2021] = 0.111$$

$$\mathbb{E}[\text{Voted}_i | \text{Age 18 to 29} = 1, \text{Year}_t = 2025] = 0.413$$

$$\mathbb{E}[\text{Age 18 to 29}_i | \text{Voted}_i = 1, \text{Year}_t = 2021] = 0.089$$

$$\mathbb{E}[\text{Age 18 to 29}_i | \text{Voted}_i = 1, \text{Year}_t = 2025] = 0.166$$

🅿️ *Identify and write in conditional expectation two of the three statistics in this article.*

Catch-up contributions can put retirees way ahead

According to new data from Fidelity, just 8 percent of its clients who are 50 and over make use of the catch-up program. Vanguard found in its last "How America Saves" report that 16 percent contribute.

While those numbers sound really low, Vanguard senior research analyst Jean Young says there is a rosier picture in certain demographics. Among those 50+ who make more than $100,000 per year, the participation rate was 42 percent.

🅿️ Mean Comparison

“I work in a hospital 5 days a week, I’ve seen it all, and I’m tired of this situation, but to be honest I don’t think vaccines are the solution. At the end of the day, most of the people who end up hospitalized because of COVID-19 have received the vaccine. In fact, I think I heard that 75% of the people who are hospitalized were vaccinated.”

- What outcome is this worker noticing?
- What is the comparison that the worker should be making?
- What is the causal question/effect they are implying?

🅿️ Mean Comparison: Regression

In response to the previous statement, you gather vaccination and hospitalization data and compare hospitalization rates between those who have been vaccinated and those who have not. You find that 15% of vaccinated individuals who contract COVID-19 are hospitalized, whereas 23% of unvaccinated people who contract COVID-19 are hospitalized.

Using that information, find the values of α and β in the following regression:

$$\text{Hospitalized} = \alpha + \beta \text{Vaccinated} + \epsilon$$

🅿️ Mean Comparison: Regression

You find a study on the relationship between mortality (all causes) and the flu vaccine. The researchers use the following regression:

$$\text{Death}_i = \beta_0 + \beta_1 \text{Flu Vaccine}$$

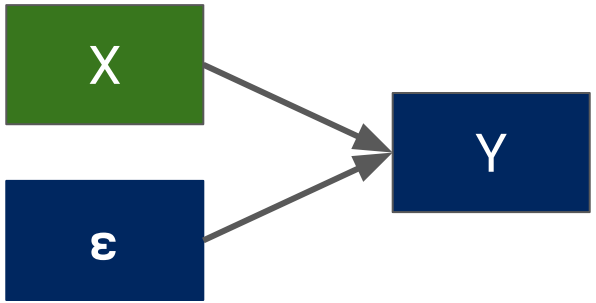
They present their results in the following table:

	Mortality Rate (per 100,000 individuals)
Vaccinated	4.1
Unvaccinated	5.8

What are the values of β_0 and β_1 in the above regression?

Why might we see the above results even if vaccination has no effect on all-cause mortality?

When examining **bias**, we want to envision the diagram below.



This equation also helps us understand mathematically how bias ($\beta_2\delta$) impacts our estimate of the causal effect. We will talk more about signing this bias next week.

If $Y_i = \alpha + \beta_1 X_{1i} + \epsilon_i$,

and $Y_i = \hat{\alpha} + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i}$,

then $\beta_1 = \hat{\beta}_1 + \hat{\beta}_2 \delta$.

🅿️ Understanding Bias

Studies have gone back and forth about the effect of the minimum wage on unemployment. You find data from your town showing that unemployment did not increase after it implemented a minimum wage increase.

- Why might a minimum wage increase have increased unemployment, even if a simple regression found no effect?
- Why might a minimum wage increase have decreased unemployment, even if a simple regression found no effect?

Additional Practice?

📝 **How to use conditional expectation language to detect BS comparisons** (For conditional expectation)

📝 **Connection between OLS and Differences in Averages** (For regressions and mean comparison)