A word on Quiz 1 and Lab 1

Nicky Wakim

2024-02-05

Quiz 1

Overall

- Great job!
- Just a few things that I think are important to review before we move into more complicated models

Question 7

Which of the following statements is true about the value -0.834 in our regression table?

a. It is the estimate of the sample intercept

b. It is the estimate of the population intercept

c. It is the estimate of the sample slope

d. It is the estimate of the population slope

- Because -0.834 corresponds to the "Age" row of the regression table, this is the slope of our fitted regression line
- This means $\widehat{\beta}_1 = -0.834$
- $\widehat{\beta}_1$ is the estimate of the population slope
- -0.834 is just the realized value (the result of fitting the population model)
- We read $\widehat{\beta}_1 = -0.834$ as: the estimate of the population slope is -0.834
- While speaking, I might say "coefficient estimate." If I am saying estimate, then I mean the population estimate
  - Please stop and ask me if my language ever feels unclear

Question 9 and 10 (1/2)

- The following are required parts of the interpretation
  - Units of Y
  - Units of X
  - Mean/average/expected before Y when discussing intercept
  - Mean/average/expected before difference, increase, or decrease when discussing coefficient for continuous covariate $\hat{\beta}_1$
    - You can also have expected/average/mean before the Y, but not necessary
  - Confidence interval

Question 9 and 10 (2/2)

$\hat{Y} = 214 - 0.834(Age)$ ↙ 40 0

- **Intercept:** For someone 0 years old, the average peak exercise heart rate is 214.233 beats per minute (95% CI: 204.918, 223.548). ↳ 40
- **Slope:** For every one year increase in age, the peak exercise heart rate is expected to decrease by 0.834 beats per minute (95% CI: -0.982, -0.685).
  - OR: For every one year increase in age, the expected peak exercise heart rate decreases by 0.834 beats per minute (95% CI: -0.982, -0.685).

↳ $Age\_C = Age - 60$

More on coefficient interpretations

Population model: $Y = \beta_0 + \beta_1 X + \varepsilon$

[mean of Y|X]

$$E[Y|X] = \beta_0 + \beta_1 X$$

$E(Y|X) = Y - \varepsilon$

What is β_1 mean?

- Let's say we have $X = x_1$ and $X = x_2$
- The difference between x_1 and x_2 is 1: $x_1 - x_2 = 1$
- We don't have to know the actual values of the x's, just that their difference is 1
- Now, let's look at the expected values for each of those x's:

$$E[Y|x_1] = \beta_0 + \beta_1 x_1$$
$$E[Y|x_2] = \beta_0 + \beta_1 x_2$$

$E(Y|X)$: expected value of Y given X

- If we take the difference between the expected values, we get:

$$E[Y|x_1] - E[Y|x_2] = (\beta_0 + \beta_1 x_1) - (\beta_0 + \beta_1 x_2)$$
$$E[Y|x_1] - E[Y|x_2] = \beta_0 + \beta_1 x_1 - \beta_0 - \beta_1 x_2$$
$$E[Y|x_1] - E[Y|x_2] = \beta_1 x_1 - \beta_1 x_2$$
$$E[Y|x_1] - E[Y|x_2] = \beta_1 (x_1 - x_2)$$

$x_1 - x_2 = 1$

$$\beta_1 = \frac{E[Y|x_1] - E[Y|x_2]}{x_1 - x_2}$$
$$\beta_1 = \frac{E[Y|x_1] - E[Y|x_2]}{1}$$
$$\beta_1 = E[Y|x_1] - E[Y|x_2]$$

- So, we can consider β_1 as the difference in the expected Y for every 1 unit increase in X

Thinking of the expected value another way

- Or: we can look at β_1 another way:

$$\beta_1 = E[Y|x_1] - E[Y|x_2]$$

$$\beta_1 = E[(Y|x_1) - (Y|x_2)]$$

  - This would make β_1 the expected difference in Y for every 1 unit increase in X

Lab 1

Overall

- I really appreciated everyone's perspective!
- I definitely learned a few things while reading
- Biggest reason why points were lost: the research question was not focused enough
  - Asked for the IAT score (implicit measure) and one other variable

Limitations of the IAT

- Taking the test multiple times
  - A lot of us mentioned learning bias, which can definitely be true
    - Think about what direction that might bias our results
  - Problems with independence between observations
- Generalizability
  - Does it represent our population? When we just say “population,” is there an unsaid assumption on the population we are referring to?
  - Can we start to narrow the definition of our population to give context to our sample?

syllabus lab wording

Other notes

- Did not intend for us to get focused on the 3 social theories in the article
  - If it helps you contextualize, then go for it!
  - But make sure you are defining the social theories and connecting them to motivation for your research q
- Minor writing notes
  - While folks is a great, inclusive word to describe people, it is a little too informal in reports
    - Good alternative is “individuals”
  - Do not use “I” or “think” in report
    - Can use the royal “We”
- When we talk about our analysis, avoid how “individuals’” scores relate to their other measures.
  - Important to note that we are not making conclusions about the individual
  - We are using individual data to make conclusions about the population!

Moving forward

- Make sure you articulate the motivation for your research question
  - If you are interested in it, then there is likely some research discussing the relationship
  - Contextualize why this is a research question worth exploring
- If you want to review your intro, please come to me!
  - Revising early will be helpful for the report
  - I will grade each portion of the report expecting you make the needed changes
  - I did not make notes on all edits - tried to identify the bigger things
- Good sources for report help
  - Structuring research articles
  - Inclusive language practices
  - Guide on improving readability