# Role Play: A Useful Model for Engaging Students in Statistics Classes

Taeyong Park

Carnegie Mellon University in Qatar

# Engaging students in statistics classes

How can we better engage students in classes?

- Have you experienced any challenges in doing so?

# Engaging students in statistics classes

How can we better engage students in classes?

- Have you experienced any challenges in doing so?

In Business Statistics classes

- Varying levels of mathematical and computational skills.
- Some less motivated students: "Why should we learn probability, $p$-values, and confidence intervals?"

# Today's session: Overview

- Introduction to the role-play teaching model.
  - Theoretical model building for regression analysis.
  - Case studies from Harvard Business Publishing.
- A mock class.
  - You will participate as a student.
- Discussion and reflection.

# Background: My Regression Analysis course

Regression Analysis for Business Administration majors.

- Second and final course in the Business Statistics sequence.
- Topics: Linear, Logistic, Ordinal Logistic, Multinomial Logistic regression analyses from a statistical perspective.
- Assignments: Exams and a final research project.
- Software: R.

# The scope of today's session: Theoretical model building for regression analysis

The topic of model specification is ideal for a class seminar.

# The scope of today's session: Theoretical model building for regression analysis

The topic of model specification is ideal for a class seminar.

In explanatory research focused on "causal" relationships, regression model building should be guided by **theoretical considerations**:

1. Begin with a hypothesis about the relationship between two variables (outcome and explanatory).
2. Consider other variables that might also explain variation in the outcome variable.

# Theoretical model building for regression analysis

- It is critical to include all plausible explanatory variables to minimize omitted variable bias.

# Theoretical model building for regression analysis

- It is critical to include all plausible explanatory variables to minimize omitted variable bias.
- HOWEVER, we can never know *all* plausible explanatory variables; Also, the problem of overfitting.

  $\Rightarrow$ Different model choices are possible; what matters is a reasonable justification.

# Theoretical model building for regression analysis

- It is critical to include all plausible explanatory variables to minimize omitted variable bias.
- HOWEVER, we can never know *all* plausible explanatory variables; Also, the problem of overfitting.

  $\Rightarrow$ Different model choices are possible; what matters is a reasonable justification.
- It is important to teach students about this "subjective" nature of model building.

# Theoretical model building for regression analysis

- It is critical to include all plausible explanatory variables to minimize omitted variable bias.
- HOWEVER, we can never know *all* plausible explanatory variables; Also, the problem of overfitting.

  $\Rightarrow$ Different model choices are possible; what matters is a reasonable justification.
- It is important to teach students about this "subjective" nature of model building.

  $\Rightarrow$ I use case studies and role-play to help students understand this concept *in a more engaging way.*

# Case studies from Harvard Business Publishing

Click: `https://hbsp.harvard.edu/home/`

# Case studies from Harvard Business Publishing

- Access the case files here: https://osf.io/wf2pa/files/osfstorage
- **Please do not share the files with others**: Copyrighted materials.
  - The files were purchased specifically for today's session, with support from CMU-Qatar.
- Pricing (for reference):
  - $8.95 per person for non-degree-granting courses.
  - $4.95 per person for degree-granting courses.

# Example: A Dean's Dilemma

1. Assign roles.
   - Team Dean Iyer
   - Team Data Analytics 1
   - Team Data Analytics 2
   - Team Data Analytics 3
   - ⋮

# Example: A Dean's Dilemma

2. Encourage students to take on their assigned roles and act as those characters, with Team Dean leading the class and **minimizing the instructor intervention**. $\Rightarrow$ Students love this approach!

- Team Dean assigns each data team the task of developing a regression model.
- Each data team discusses and builds their model with substantive justification, then runs the model using R.
- After a set period, Team Dean invites each team to present their model specifications and results.
- Team Dean selects the model they wish to adopt and explains their decision-making process.

# Example: A Dean's Dilemma

3. Follow-up discussion.

# Example: A Dean's Dilemma

3. Follow-up discussion.
   - What factors should we consider when building a regression model?

# Example: A Dean's Dilemma

3. Follow-up discussion.
   - What factors should we consider when building a regression model?
   - Omitted variable bias vs. Overfitting.

# Example: A Dean's Dilemma

3. Follow-up discussion.
   - What factors should we consider when building a regression model?
   - Omitted variable bias vs. Overfitting.
   - If different models lead to different conclusions about our hypothesis, how should we interpret the results?

# Example: A Dean's Dilemma

3. Follow-up discussion.
   - What factors should we consider when building a regression model?
   - Omitted variable bias vs. Overfitting.
   - If different models lead to different conclusions about our hypothesis, how should we interpret the results?
   - Does a "correct" model truly exist? What does that mean in practice?

# Example: A Dean's Dilemma

3. Follow-up discussion.
   - What factors should we consider when building a regression model?
   - Omitted variable bias vs. Overfitting.
   - If different models lead to different conclusions about our hypothesis, how should we interpret the results?
   - Does a "correct" model truly exist? What does that mean in practice?
   - Why are robustness (or sensitivity) checks important in regression modeling?

# Advantages and challenges

Advantages:

# Advantages and challenges

Advantages:

- Students were actively engaged and enjoyed the activity.

# Advantages and challenges

Advantages:

- Students were actively engaged and enjoyed the activity.
- They appeared to appreciate seeing how statistical modeling applies in real-world contexts.

# Advantages and challenges

Advantages:

- Students were actively engaged and enjoyed the activity.
- They appeared to appreciate seeing how statistical modeling applies in real-world contexts.
- They developed a deeper understanding of the nature of real data analysis:
  - Multiple plausible models can exist for the same question.
  - The true (population) model is unknown.
  - Subjective judgment plays an important role.

# Advantages and challenges

Advantages:

- Students were actively engaged and enjoyed the activity.
- They appeared to appreciate seeing how statistical modeling applies in real-world contexts.
- They developed a deeper understanding of the nature of real data analysis:
  - Multiple plausible models can exist for the same question.
  - The true (population) model is unknown.
  - Subjective judgment plays an important role.

Challenges:

- Difficult to implement in large class settings.
- Some students remained less engaged despite the interactive format.

# Mock class

Case: GoodBelly

# Mock class

**Roles:**

- CFO
- Marketing Department
  - Caroline Dickerson 1
  - Caroline Dickerson 2
  - ...

**Instruction:** Please take on your assigned roles and act as those characters during the activity.

# Discussion and reflection

- Do you have any suggestions for improving the role-play method?
- How might you adapt or apply this method in your own classes?
- What other strategies have you used (or would like to try) to engage students, beyond role-play?