Applying Advanced Econometric and Marketing Techniques to Analyze Various Issues in Computer Science Education

Lianen Ji,
Department of Computer Science and Technology
China University of Petroleum
Beijing, People's Republic of China

Jean Mayo and Ching-Kuang Shene
Department of Computer Science
Michigan Technological University
Houghton, MI USA

Issues

- My teaching evaluation is 3 out of 5, what is the problem or cause?
- 10% of CS students transferred, what is the problem or cause?
- Why students did not take my section? Am I too harsh? Was the timing wrong?
- Could small classes make my rating higher?
- Could grade inflation make my rating higher?
- Is teaching approach A better than B?
- Etc.

Commonly Used Techniques

- Mean, variance, confidence intervals, effect size, etc.
- Pre- and post- tests
- Control and treatment groups
- ANOVA
- Etc.

Problems:

- Even though a hypothesis test shows a significant effect, we do not know the cause!
- We may have difficulty to divide the students into control and treatment groups.
- Mean, variance, confidence interval, etc. do not tell a complete story.
- We need an advanced quantitative explanation rather than only using mean, variance, etc. and/or significant tests.
- We are all human being, can we avoid bias and very subjective judgment? How can we find this bias out?

Why the Modeling Approach?

1. Can build accurate and complex causal relation between the explanatory and dependent variables
2. Can determine a quantitative impact factor of each influential variable
3. Can have very sophisticated hypothesis testings to determine the importance of each variable
4. Can answer many more questions that the commonly used methods cannot.
5. Can handle various types of dependent variables (e.g., continuous, discrete, etc.)
6. Etc.

Example 1 (Aggregate Data):

Student average scores have ranges (e.g., in [0, 100]).
One cannot use regression for estimation.
For example, applying regression to the following yields incorrect α, β, etc.

$$\text{score} = \alpha + \beta\times(\text{working hours}) + \ldots$$

We should use the Tobit models for the limited dependent variable (score)

Example 2 (Aggregate Data):

Dependent variables may not be independent of each other and cannot be estimated separately.
For example, people think an instructor's rating may be related to student scores; but, student scores maybe in turn depend on instructor's rating:

$$\text{rating} = \alpha + \beta\times(\text{score}) + \ldots$$
$$\text{score} = \gamma + \delta\times(\text{rating}) + \ldots$$

We have simultaneous equations which can be estimated with 2SLS, 3SLS, FIML, etc. methods.

Example 3:

Dependent variables may be discrete events.
In this case, regression techniques do not apply.
For example, a student may choose to ***STAY*** in CS, ***LEAVE*** CS, and ***UNDECIDE*** (3 event) – retention.
Regression cannot model discrete events.
We may use multinomial or conditional logit models.

$$\text{prob}(\text{event } i \text{ occurs}) = \frac{\exp(X_i\beta)}{\sum_j \exp(X_j\beta)}$$

Example 4:

Sometimes events could be ordinal.
For example, the Likert scale is ordinal.
Additionally, events could be hierarchical.
For example, students may choose a course followed by a section of that course.
Or, students may choose an instructor followed by one of his/her course.
Could this hierarchical type of events affect instructor evaluation?
Could we model this hierarchical ordering of events to minimize evaluation bias?

Example 5:

Many factors could affect the rating of an instructor.
We intend to investigate and find the most influential factors that can truly describe student reactions on their choices (e.g., rating, retention, etc.)
This is exactly what conjoint analysis does in marketing analysis.

Example 6:

In a course evaluation, students may state their preference of an instructor (i.e., stated preference).
Could their opinion change AFTER completing the course with another evaluation (i.e., stated choice).
Could their opinion change again a few years AFTER they work in the field and understand the value of the knowledge they had learned in the class?
How do we address these issues so that course evaluation is no more perceptional and/or emotional?

Acknowledgements

This work is partially supported by a grant from the China Scholarship Council (CSC No.201306445011) and grants from the National Science Foundation (DUE-1140512, DUE-1245310 and IIS-1456763).