

TITLE: Improving Student Learning through Continuous Assessments in a Statistics Course—Pascal Deboeck, Psychology

OVERVIEW

A psychology professor modifies a graduate-level statistics course to enhance students' understanding, integration, and application of categorical data analysis in a term project.

BACKGROUND:

PSYC685/895 is a mid-level graduate course in categorical data analysis (statistics). The course is lecture-based and meets twice a week for 75 minutes each time. A third session held each week, taught by a graduate teaching assistant, consists of a one-hour lab session involving work with statistical software.

Since this course is cross-listed as an undergraduate course (for those completing a quantitative minor) and as a graduate course, students come from diverse backgrounds with large differences in statistical experiences and previous training. For example, I've had students from psychology (some with and some without a specific focus on statistical methodology), the School of Education, the School of Business, and a student from East Asian Studies. The course normally enrolls between seven and 11 students per semester. In a typical class there will be two or three students who require extensive additional help, as they may be unfamiliar with particular topics or are inexperienced with the use of statistical software.

The goal of this course is to introduce principles and analyses related to data with categorical outcomes. This course considers topics such as probability distributions with categorical data, contingency table analysis, the general linear model, logit models, and loglinear models. By the end of this course students are expected to:

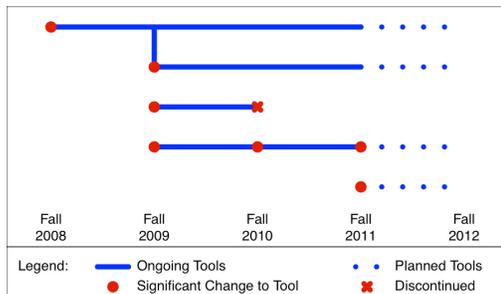
- learn to select methods appropriate for a question of interest;
- learn to apply categorical methods and interpret categorical analyses;
- demonstrate critical thinking about the application of categorical methods.

One course requirement is to submit a final project at the end of the semester. Students are expected to propose an analysis project based on their own data, draft a paper based on the analysis, review each other's projects, and subsequently revise the initial draft. This project is designed to be a practical exercise for the students, which requires them to demonstrate the ability to select methods appropriate for their question(s) of interest, learn to apply categorical methods, and interpret the results. The peer-review of other projects requires students to think critically about applications of categorical methods. However, in the first rendition of my class, I had a student whose project was lacking so many dimensions that it was difficult to ascertain how much had been learned in the class. Based on this experience I have made a series of ongoing changes to the class in order to scaffold, or support the development of, the skills required for the successful completion of the final project and future research projects.

Through these changes I aim to further reinforce the topics presented in class, and improve evaluation through changes to the class project and the addition of other measures of learning.

IMPLEMENTATION

Over the last few semesters, I have made a number of modifications to the term project and included additional assignments. The course now consists of several types of assignments that have been designed to scaffold students' understanding of categorical data analysis.



Final Project:

Each time I have taught this course, I have required students to apply their knowledge of categorical data analysis to any problem of their choosing. This project begins as a proposal (maximum 200 words) submitted to the instructor and GTA for comments. Students then present a draft of this paper to the rest of the class. The paper is subsequently revised based on feedback given by the instructor, the graduate teaching assistant, comments/questions made during the class presentation, and reviews written by two or three student reviewers.

In Fall 2008, when this project was first implemented, students generally performed as expected: Students were able to choose methods that were appropriate for their question(s) of interest, apply the method selected, and provide some interpretation of the results. However, one project left me entirely uncertain about how much knowledge or understanding the student was taking away from this class. This student's project was lacking on many dimensions: 1. the paper was poorly written and lacked detail; 2. there was improper use of statistical terms; 3. the student selected a very basic analysis and it was not clear from the paper that the analysis was correctly run; and 4. the student could not interpret the results from this basic analysis.

I realized that I not only needed to make my expectations clear to the students, but also include additional assignments to develop the skills to successfully complete this project. Accordingly, in Fall 2009, I provided a rubric to the students to indicate my expectations. This rubric included a list of requirements of a project that meets my expectations, one that is much above expectations, and one that is well below expectations. For example, a project that meets expectations would have "mostly accurate and clear use of statistical terms," while a project much above expectations would have "no misuse or ambiguity regarding statistical terms."

I also broke this project down into smaller assignments to allow students to get feedback at every step. This form of feedback came from three sources: the instructor, the teaching assistant, and peers. Students were given the rubric to use to evaluate their peers' projects. I find that the peer review process allows them to develop their understanding in applying the coursework to analyze and interpret different data sets with others who are knowledgeable about their topic.

With a small class it is difficult to gauge whether changes are working; however, outlining my expectations did appear to be a wise choice. The Fall 2009 class consisted of substantially fewer people specializing in statistical methodology, but their projects were at least as good as in the first year. Evaluation of student learning was more straightforward in the second year, but there were no students as extreme as the student described from the first year. I continued to use the same rubric for the final project in Fall 2010 and Fall 2011.

Online Wiki:

To allow students to share and consolidate information from the class, I added an online wiki in Fall 2009. Initially, students were asked to work in pairs. For each wiki, students had to write a short summary of a chapter (typically no more than a paragraph or two summarizing the key points), define important terms and formulas, and provide examples of how to analyze data using the ideas in the chapter. This project was poorly implemented in Fall 2009 (year 2). Most students waited until the end of the semester to submit the wiki assignments. While this still made it a useful tool to evaluate student learning, it made the wiki completely useless as a teaching tool. Students could not use the wiki to learn from each other, nor could they refer to the wiki as they were working on their class projects.

To use wikis as both a tool to evaluate student learning and as an additional teaching resource, starting the third year (Fall 2010) students were asked to improve upon the existing wikis, and also use the wikis as a resource. They were required to complete the wiki assignments one week following the completion of their assigned chapter in class. Students were encouraged to look over the wikis prior to the beginning of a new chapter; as the wikis were substantially shorter than the chapter, this was intended to briefly introduce concepts and terms from the chapter, thereby providing a scaffold for the class to build on. Students were also encouraged to look at the analysis examples, particularly as they planned their class project and were in the process of analyzing their own data.

Unfortunately, since students were adding new information to the previous wikis, students generally did not make meaningful contributions, and consequently wikis were not useful as a tool to evaluate student learning. On average, students' contributions were limited to adding code for new software rather than making changes that would help future classes, such as differing ways to understand concepts presented in class. The contributions in the second year likely did little to reinforce student learning of chapter material. The wiki assignments also served as a teaching resource, since students could use the wikis to enhance their understanding of course materials. I chose not to continue with the wiki assignment, as it did not seem likely that the wiki would become a growing document that would span over multiple semesters. If the document didn't grow over multiple semesters, there seemed to be little reason to use a wiki when other, simpler methods (e.g., written chapter summaries) would accomplish the same things with the

students. Accordingly, in Fall 2011 I discontinued the use of the wiki as an assignment but relied on it as a teaching resource.

Weekly Assignments

Another set of assignments consisted of problems completed each week during the lab session. The lab session included a 30 - 40 minute lecture given by the GTA, after which students completed a one-page assignment, typically consisting of three questions. Students had 10 - 20 minutes to complete this assignment. If they experienced difficulty, they had the option to turn the assignment in at a later date. The GTA graded these assignments and provided the instructor with a summary report that gave further indication as to how well the students were completing each assignment.

I added these assignments in the second year to provide more opportunities for me to give incremental feedback to students, as well to help reinforce students' learning by requiring additional practice of the skills they were acquiring. These assignments provided some measurement of student learning during the course of the semester. However, they were not a perfect indicator of student progression toward the class goals, as the assignments tended to reflect some combination of who was/was not diligent at completing the assignments and who was struggling with use of the statistical software, rather than identifying gaps between theoretical understanding and interpretation of data.

While in Fall 2009 the assignments were graded primarily for completion, in Fall 2010 the lab assignments carried minor penalties for poor completion. I made the assignments more challenging and designed the questions so that they could potentially tap into different aspects of student learning that were more aligned with my course goals – applying theoretical knowledge to analyze and interpret data sets, as well as demonstrating knowledge of statistical software. Using these assignments, I wanted to be able to identify different areas that students might have difficulties with, and help create a distribution of grades that would differentiate students who were having problems with the class material.

In Fall 2011, I made more improvements to the lab assignments. As Susan A. Ambrose (2010) highlights, "To develop mastery, students must acquire component skills, practice integrating them, and know when to apply what they have learned." While the early labs in 2011 were similar to those of the previous year, later labs further emphasized the acquisition of additional skills and practice integrating them. The final four labs in 2011 required students to not only practice component skills, such as calculating a particular statistic of interest, but to also integrate them into a 200-400 word summary which could serve as the results section of a paper. This provided additional practice for the final paper, and also allowed students to practice writing up results with four of the important methods examined in this course.

Survey

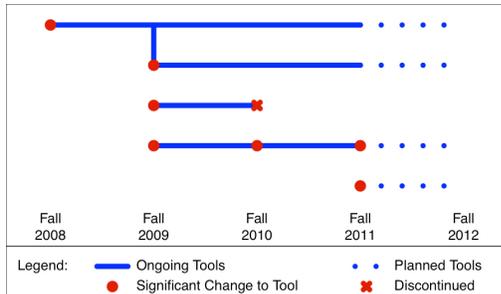
To get more feedback on student learning and examine the impact of this course on students' understanding of statistical concepts, I added a survey in Fall 2011. Prior to the beginning of the semester, I developed a list of important topics covered in this class. Once the semester started,

students indicated their expertise with each topic. More specifically, I asked them to indicate how much they thought that they could write on a particular topic: nothing, a page, a few pages, a half chapter or more. I administered the survey again at the end of the semester.

As this is an introductory course, my analyses focused less on the students who could already write several pages or more on a topic, and more on students who came into the class knowing a page or less about a topic—*did they feel that they could write more about a given topic having taken the class?* This survey would allow me to identify a handful of areas where the least learning occurred and focus my attention on altering those lectures.

Final Project
Grading Rubric
Wiki Assignment
Lab Assignment
Survey

STUDENT PERFORMANCE



There were several indications that the latest course modifications were successful. These indicators included student performance on the project and supporting assignments, student commentary on their peers' projects generated during the presentation session, and student feedback at the end of the semester.

Class project

After making these changes, I did not have any project that was below my expectations (as laid out in the grading rubric). The rubric was useful in evenly evaluating student learning, and I plan on using it again in upcoming semesters. Furthermore, breaking the project down into several components and getting continuous feedback resulted in mostly A-level work. Students selected methods that matched very well with their questions of interest, incorporated statistical principles and skills beyond the material presented in class, and included a well thought out discussion of statistical results.

Each of the final projects is organized in the following way:

- *Step 1: Student Categorical Proposal, including comments on proposal and retrospective reflection by the instructor*
- *Step 2: Student Paper Draft Excerpts, including comments and retrospective reflection by the instructor*
- *Step 3: Student Final Paper Excerpts (Key Changes), including retrospective reflection by instructor*

Online Wiki

The purpose of this assignment was to evaluate students' abilities to describe a topic learned in the class that was unlikely to be the same as the term project, while also providing students with another resource to help them learn about the class topics. The wikis were not as useful as they were in Fall 2009. Students were asked to add on to the existing wikis from the previous rendition of this course. Given that the wikis were very comprehensive to begin with, there wasn't much scope to make meaningful contributions. I discontinued the use of the wikis as an assignment in Fall 2011.

Lab Assignments

The initial lab assignments (prior to Fall 2011) did not give a good indication of student performance throughout the semester, as the assignments were more a reflection of programming expertise than class skills. The lab assignments are becoming increasingly useful, however, as was demonstrated in the most recent semester. While discussing student performance with my graduate teaching assistant, discussion turned to two students whom I was having difficulty reading in class. The questions students ask often provide me with a good impression as to how well they are grasping material. In the case of these two individuals, both were very quiet and showed few facial reactions, so it was unclear to me whether they were lost/confused or just had a more introverted personality. My graduate teaching assistant was able to give me a very clear impression based on the lab assignments, as one was completing the assignments very well and the other's assignments tended to not be completed as well. I am still assessing whether the change in the focus of the assignments—such that assignments increasingly require students to translate their results into words rather than just compute a statistic—has impacted the quality of students' learning.

Survey

The survey topics were divided into three sections: 1. topics that were related to background knowledge or critical tools/knowledge needed for the course; 2. the primary material taught by the course; and 3. material that was secondary or extra for this course. Students had to indicate if they could write a page more about the specific topics.

The background material pre-semester survey suggests varying degrees of experience with the topics. Some of the higher percentages, such as with statistical inferences, potentially suggest what topics could receive less attention. However, it appears that there is generally a need to introduce these topics. The post-semester results suggest students feel more familiar with the background material across all topics.

Across all topics, students believe they know much more about the topics following the course. However, the sections on non-independent data ("Models for a match pair of questions," "Generalized linear mixed models") suggest an area where students do not appear to be making the same gains as in other parts of the course.

The extra material survey indicates that students do gain some knowledge of material that was not of primary concern. However, there was no one particular topic that students seemed to especially know.

Given the amount of time used to construct this survey, administer it, and examine the data, the insights gained are well worth the investment. The results provide evidence to indicate that the background material is not redundant with other courses and:

1. most of the primary material is novel to most students,
2. most students feel they have made substantial gains in the primary material,
3. two related sections of the primary material might need to be better covered, and

4. some students are benefiting from the extra material despite the small amount of class time spent on these topics.

As this is a small course (seven to 11 students), and the makeup of the class varies dramatically each time it is taught, it will be interesting to see how consistent these values are for differing samples (e.g., more/fewer psychology students, quantitative students, undergraduates, etc.).

Detailed information on survey results can be found.

REFLECTIONS

The current generation of students experience a constant flood of statistics; news articles, reporters, politicians, scientists provide a bevy of claims frequently supported with statistics. Because of this, there is a greater burden on statisticians to try to ensure that the public is informed about how statistics work, what they can tell us, and their limitations. My courses, such as this one, are an opportunity to try to allow undergraduate and graduate students from a range of disciplines to think critically about the application of statistics to their own research, and in the process learn ways to think critically about the application of statistics by others.

This portfolio represents an ongoing experiment to better convey statistical methods. The course began with a design which I personally experienced in much of my graduate training, and that I observed coming to the University of Kansas; graduate level classes typically consisted of lecture, accompanied by a term project to assess student learning. In the first few years of teaching this course, I've explored ways to increase feedback throughout the semester—both for myself (that is, feedback about how students understand material), and for the students.

Future changes will continue to depend on the experiences with this class, and in teaching my other courses. In many of my statistical courses I am beginning to ask students to fill out a survey at the beginning and the end of the semester; this survey asks them to assess their knowledge on 10-15 major topics in the course. These surveys are starting to provide me more detailed feedback about topics with which students are typically already familiar prior to the beginning of class, and topics that show less growth. One of the next steps will be to create a correspondence between class assessments and this survey to see whether students can reasonably gauge their learning during the semester. Right now many of the changes are more incremental, although a recent and new course prep has allowed me to experiment with a very different style of lecturing. As with this course, depending on student outcomes and response, the experiment will inform changes made to all of my statistical courses.