

## STATISTICS: A WINDOW TO UNDERSTANDING DIVERSITY

Mahtash Esfandiari

Statistics Department, University of California, Los Angeles  
esfandia@stat.ucla.edu

*Students who enter UCLA College of Letters and Science since the Fall Quarter of 2015 are required to take a diversity course as a graduation requirement. Currently more than 200 courses qualify, but, only three are offered in the Physical Sciences Division; one of which entitled “Statistics: A Window to Understanding Diversity” is developed by the author. This experienced-based course allows the students to learn the theory of diversity and simultaneously use statistical methods to analyze extensive data collected by the author on diversity climate at UCLA. Based on student input, this course enabled them to critically think about diversity issues, develop a better understanding of diversity climate at UCLA, and adjust more successfully to our campus, and the diverse world in which we live.*

### INTRODUCTION

A diversity-focused course requirement has been a long-standing priority for UCLA. The Academic Senate’s Legislative Assembly — which includes faculty representatives from academic departments across campus — overwhelmingly approved the requirement. Based on this, students in the College of Letters and Science that approximately comprise 84 percent of the UCLA’s undergraduate student body will be required to take a diversity course that substantially addresses racial, ethnic, gender, socioeconomic, sexual orientation, religious or other types of diversity. The students need to pass this course with a minimum grade of C.

### BACKGROUND

The author has conducted extensive research along the reform movements in statistics education and GAISE guidelines for statistics education. These efforts include the role of active learning (Esfandiari, 2003), use of real data (Esfandiari, 2010), emphasis on diagnostic and formative evaluation through development of *test banks*, enhancement of statistical literacy through development and implementation of a *cyber-based online article bank* that emphasizes application of statistics to solving real world problems (Esfandiari, 2014), having teams of students work on projects that involve cleaning and analysis of large data sets, and helping *novice learners* see how an *expert* uses statistics in conducting scientific research (Esfandiari, 2012).

### AN OVERVIEW OF STATISTICS: “A WINDOW TO UNDERSTANDING DIVERSITY”

Statistics is all about analyzing variability and is defined as the “science of data”. All data exhibit variability. Diversity involves exploration of variability with respect to many dimensions such as race, ethnicity, gender, sexual orientation, socio-economic status, age, physical abilities, religious beliefs, political beliefs, or other ideologies. Thus, we could say that diversity is about examining and analyzing variability. This course is a revision of the upper division course that we offer at the Statistics Department entitled “Statistics for Social Sciences”. What makes this course unique is that it creates a rich context that helps physical science and social science majors see how applied statistical methods can be used to examine, analyze, and model variability of diversity issues such as gender roles, race, ethnicity, socio-economic status, sexual orientation, political view, exclusion, being first generation, micro-aggression, and relationship of campus climate to cognitive, affective, and moral development.

### THEORETICAL UNDERPINNINGS OF THE COURSE

The overall goals of this course include the cognitive, affective, and behavioral domains (Bloom, 1971). The cognitive goals relate to reading, discussion, analysis, and evaluation of diversity theory, and acquiring statistical knowledge for analysis of diversity data. The affective goals relate to helping students appreciate diversity, helping social science majors realize that statistics is not a branch of mathematics, and helping physical science majors realize the

importance of writing within context as a major component of applied statistics. The behavioral goal relates to helping the students adapt more successfully to our diverse campus and the diverse world in which we live. The theoretical framework of this course can be represented in figure one.

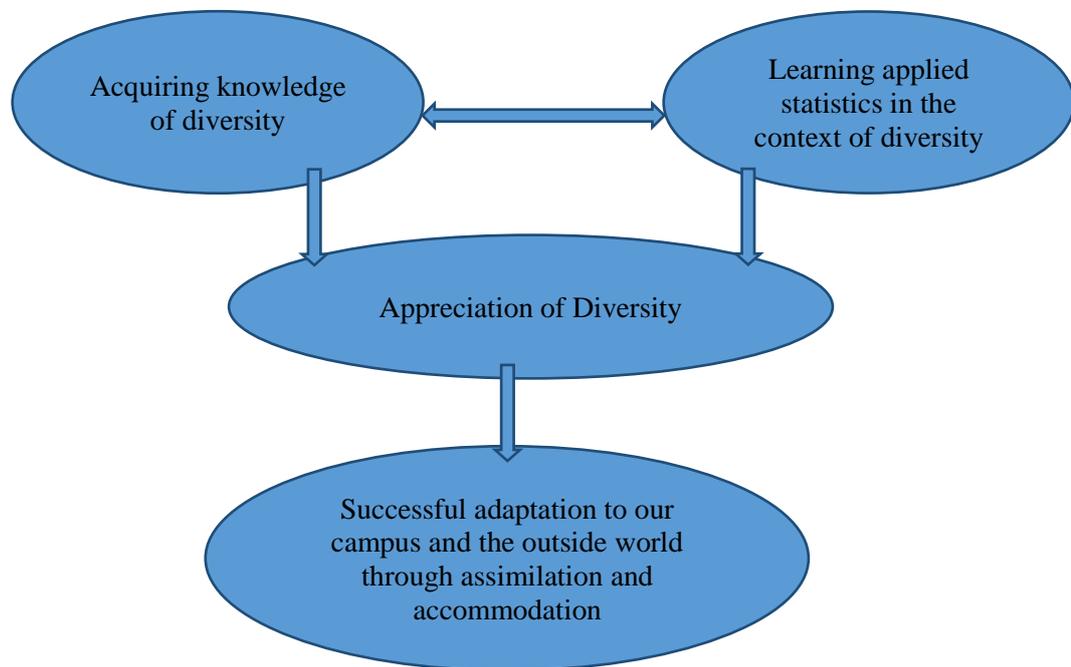


Figure 1. Schematic of the theoretical framework of the proposed course

Based on the theory of cognitive science (Piaget, 1978), successful *adaptation* to the environment, in this case our diverse campus and later to the outside world, happens through the active processes of *assimilation* and *accommodation*. *Assimilation* is the cognitive process of fitting new information into existing cognitive schemas, perceptions, and understanding. *Accommodation* happens when new information or experiences cause individuals to modify their schemas, perceptions and understandings.

Thus through learning the theory of diversity and *reflecting* upon it, along with using applied statistics to *make sense of diversity data*, the students take in the new information, *assimilation*, and modify their perceptions of diversity, *accommodation*, to develop appreciation of diversity and successfully adapt to the world around them.

#### INSTRUCTIONAL MATERIAL AND SOURCES OF DATA

A survey was designed by the author based on “UCLA Campus Climate Survey” and data was collected on a sample of 1000 undergraduates enrolled in a lower division statistics class that serves as a GE (General Education) requirement. After consultation with the relevant authorities, and removing identity information, the Office of Equity, Diversity, and Inclusion, provided the author with a similar data set on 5000 undergraduates at UCLA. Exploratory data analysis indicated that the two data sets are very similar in terms of sample characteristics and thus should be a good representative of UCLA undergraduates. These two data sets in addition to other large data sets related to diversity were used by the author in developing all of the instructional material in the course.

#### TEACHING STRATEGIES

The teaching methodology is simulation-based, with interactive use of R-studio during lecture, emphasis on active learning (Esfandiari, 2003), presentation of statistical concepts within context with emphasis on verbal and oral communications of finding to a statistical and non-statistical audience (Esfandiari, 2014), with heavy emphasis on group work.

The author has reached the objective of *active learning* and *student involvement* through development of a series of *review exercises* along with the answers to them for each lecture. Questions on the review exercises cover the whole cognitive domain including comprehension of statistical concepts, understanding simulations, explaining formulas, application of the right statistical method, analysis, oral, and verbal interpretation of outputs generated by R, as well as analysis, critique, and evaluation of tables and plots reported in the statistical section of refereed journal articles.

In each lecture, after teaching a statistical concept or strategy, the students start working on a review exercise that requires them to answer questions about the topic covered. They read the questions on their own, if need be, run the relevant analysis by using the given R commands, try to come up with a potential answer, discuss their answer with the student next to them, *pair and share*, and write the results within context. The answers to the review exercises have proven to help students with review and clarification of their misconceptions.

#### ACTIVITIES INVOLVED IN THE TEN WEEK QUARTER

During the first four weeks of the quarter: 1) students are divided into groups that are diverse with respect to gender, major, and ethnicity, 2) statistical concepts discussed in the pre-requisite statistics course are reviewed, 3) students read diversity literature, listen to lectures from scholars in field of diversity, write individual and group *reflection papers* on diversity topics, and complete statistics homeworks and labs in their groups.

In week six, students take a three-hour open-book, open-note, Open-Internet, midterm exam with access to R. This exam tests: 1) students' knowledge of what statistical methods they should use to answer different research questions, 2) their understanding of the mathematical underpinnings of the statistical methods taught, and 3) and their ability to use R in analyzing data, interpreting R outputs, writing statistical findings within context, and making sense of statistical tables and plots. Research has shown the author that this method of evaluation is very effective in terms of lowering student anxiety about statistics, improving their motivation and confidence about statistics, and allowing them to synthesize the material learned.

After the midterm, advanced statistical methods such as multiple linear regression and logistic regression are discussed. Each group is assigned a data set along with the relevant codebook. Students are required to use the assigned data sets for their final project. All group meets with instructor for half an hour in weeks eight, nine, and ten to discuss: 1) the questions they want to answer, relevant exploratory data analysis, statistical method they plan to use, analysis of the data, checking the relevant assumptions, and interpretation of results. Each group writes a final report and presents their results in the final's week. Typical topics covered in the projects include: 1) Analysis of UCLA students' perception of respect for diversity on our campus as a function of climate comfort, perception of prejudice and exclusion, whether the student is first generation, major, socio-economic status, ethnic background, and sexual orientation, 2) professor salary as a function of gender, ethnicity, and field of specialization, and 3) gender discrimination and how it affects appointing women to leadership positions and public office.

#### EVALUATION STRATEGIES

**Diagnostic Evaluation:** In the first week of the quarter, the students are given a number of non-graded online quizzes that can be tried multiple times and help with the review introductory statistics.

**Formative Evaluation:** Online quizzes, reflection papers, homeworks, labs, and review exercises are used to monitor student progress, and diagnose their misconceptions.

**Summative Evaluation:** Midterm exam, and final project are used as means of summative evaluation and ascertaining that the overall goals of the course have been reached.

#### SOME FINDINGS FROM AN END OF THE YEAR ANONYMOUS SURVEY

Of the 23 students who filled out an end-of-course anonymous survey (out of 25 students enrolled) ...

- 20 agreed that learning to analyze the provided data was effective or very effective for understanding diversity, and helping them *adjust more successfully to our diverse campus*.

- 16 said that the course was effective or very effective at helping them *adjust to the diverse world in which we live*.
- 23 mentioned that they *would recommend the course*.
- 23 said that they *learned more than they expected*, and they were not bored.
- 21 indicated that the guest speakers were effective or very effective in helping them *learn about the theory of diversity* and developing a more positive attitude toward diversity.
- 14 found the assigned readings and reflection papers to be effective or very effective in helping them learn about diversity and *developing a more positive attitude toward diversity*.
- 19 found *group projects* to be effective or very effective in helping them *play an active role in their learning*.
- 21 found *review exercises* to be effective or very effective in helping them *play an active role in their learning*.
- 20 found homeworks and review exercises to be effective or very effective with respect to training them *write statistical findings within context* and for a statistical and non-statistical audience.

To provide just one student comment that summarizes the benefits of this course from a student perspective: “The readings were interesting and allowed the stats majors to see how north campus majors are attempting to quantify diversity, while allowing non-stat majors to get a glimpse of how statistics are used to answer north campus questions. The guest speakers provided interesting and fresh perspectives, as well as demonstrating continual effort towards diversity inclusions.” North campus refers to students who major in social sciences.

#### REFERENCES

- Bloom, B., Hastings, J., & Maduas, M. (1971). *Handbook of formative and summative evaluation of student learning*. New York: Mc Graw Hill.
- Esfandiari, M., Nguyen, H. Yaglovskaia, Y. & Gould, R. (2010). Enhancing statistical literacy through open-ended questions that involve, context, data, and upper level thinking. Paper presented at *International Conference of Teaching Statistics, Yugoslavia*.
- Esfandiari, M., Sorenson, K., Zes, D., & Nichols, K. (2012). *Designing a Cyber-Based Article Bank to Enhance Statistics Education*. Presentation at UCLA Department of Statistics Seminar, USA.
- Esfandiari, M., Sorenson, K., Zes, D., & Nichols, K. (2014). *Enhancing Statistical Literacy through Analysis of Scientific Journal Articles*. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA*. Voorburg, The Netherlands: International Statistical Institute.
- Esfandiari, M. (2003). *Application of a Generative Model of Teaching and Technical Writing to Teaching Applied Statistics*. Presentation at the IX Seminar on Applied Statistics, Brazil.
- Piaget, J., & Brown, T. (1978). *The equilibrium of cognitive structures: The central problem of intellectual development*. Chicago: University of Chicago Press.