

TEACHING THE TWO-WAY ANALYSIS OF VARIANCE USING TECHNOLOGY RESOURCES

Ramon Gomez
Florida International University
813 NW 133rd Court
Miami, FL 33182
gomezra@fiu.edu

1. Introduction

Educational researchers believe that advanced Statistics topics are now essential for a variety of college graduates, particularly psychology and biology majors. To address this issue many statistics departments offer a second undergraduate course including Analysis of Variance (ANOVA) models.

The intent of most statistical experiments is to determine the effect of one or more categorical variables (factors) on a quantitative variable (response). For example, an agricultural researcher can be interested in determining the effect of three types of irrigation levels and four types of fertilizers on the yield of the tomato crop in order to identify the optimal combination. The statistical technique used to analyze the significance of these effects is called Analysis of Variance. When two factors are involved in the experiment and the effect of one factor is not the same for every level of the other factor, the statistical model used is called two-way ANOVA with interaction. The students' understanding of this model is generally difficult for non-statistics majors because of the nature of the problem and complexity of computations.

The appropriate use of technology has been identified as a contributor to the students' motivation and understanding of statistics. Using technology can make college teaching of statistics more effective as it improves the quality of instruction, encourages students' active learning, and provides them with psychological incentives (Garfield, 1995; Higazi, 2002). Consistent with this statement, a committee created by the American Statistical Association produced in 2005 the Guidelines for Assessment and Instruction in Statistics Education (GAISE), which recommended the use of technology for introductory statistics courses at college level.

In this regard, PowerPoint has permeated all aspects of college teaching as a presentation technology resource. This success has been associated with the appropriate use of text, images, and graphics. Furthermore, the use of statistical software provides students with a tool that enhances their learning experience, by allowing them to engage the contents actively and analytically. Although the choice of statistical software generally depends on the context where students are expected to use their statistical skills and the resources available, SPSS (Statistical Package for Social Sciences) has been identified as one of the most commonly used packages at college level (Hulsizer & Woolf, 2009). The use of

PowerPoint and statistical software in undergraduate courses has been previously described in the literature as a facilitator of learning statistics (Lock, 2005; Gomez, 2010 & 2011)

This paper discusses the benefits of using PowerPoint and SPSS while teaching the two-way ANOVA with interaction as part of a second statistics course for undergraduate students at the university level. It summarizes the present author's experience at Florida International University (FIU) during recent years.

2. Method

2.1 Course description

The STA-3112, STA-3123, and STA-3194 are three credit-hours classes covering a range of topics: hypothesis testing based on one and two samples, analysis of variance models, regression analysis, chi-square tests for categorical data, and non-parametric statistics. They are second statistics courses having as a prerequisite a preceding class that includes descriptive statistics, probability and hypothesis testing based on a single sample.

The textbook for both STA-3112 & 3123 is "Statistics" by McClave and Sincich (2011, 2013), that emphasizes inference methods and stresses the development of statistical thinking. It includes many proposed exercises for which real data is utilized to illustrate statistical applications. These courses encompassed content from chapter 8 to 14 with chapter 10 devoted to the Analysis of Variance (ANOVA) models. In particular, the topic of Two-Way ANOVA with interaction is presented at the end of this chapter. The textbook for STA-3194 is "Biostatistics" by Wayne Daniel (2009) which requires mid-level mathematical prerequisites. It includes exercises focused on applications to the health sciences and introduces the Two-Way ANOVA with interaction in chapter 8.

The present author has taught the STA-3112 and STA-3194 courses each spring term between 2010 and 2012 integrating PowerPoint and the SPSS software. Typical enrollment for STA-3112 was 50 students seating in a classroom with a computer projection system. The STA-3194 course is offered to selected scholars enrolled in a special program (QBIC), with classes not exceeding 25 students seating in a computer lab.

2.2 Course design and organization

The traditional approach to teaching Statistics consists of using a board during lectures, a textbook as a reference, a calculator for computations and, more recently, supplementary material posted on a website. Two technology additions were integrated in our courses between 2010 and 2012: the daily use of PowerPoint for lectures as well as statistical software (SPSS) for data computations and analyses. This integration allowed for more class time to discuss statistical concepts and applications. Thus, a broader conceptual understanding of the material was promoted as well as active learning in the classroom.

Consequently, an interactive learning environment was generated where students had the opportunity to develop an increased rank of statistical literacy and reasoning.

The PowerPoint presentations, developed by the present author for this course were structured with the goal of increasing student participation during lectures in addition to satisfying the needs of scholars with a more visual oriented learning style. A course pack comprising the PowerPoint slides for all lectures was made available to the students at the beginning of the course, eliminating the hassle of frantic note-taking in class. The incorporation of SPSS involved the use of computer output to illustrate various topics allowing a more effective discussion of the statistical concepts and applications. The Instructor's style for lectures consisted of projecting slides from a presenter device and discussing their content with the students while moving around the classroom. This approach allowed a more direct interaction with learners.

2.3 In-class teaching approach for the two-way ANOVA with interaction

While teaching the topic of Two-Way ANOVA with interaction, the present author organized the material in five phases: a) Discussion of basic concepts b) Interpretation of the Interaction, c) Partition of the total variance and understanding of the ANOVA table, d) Hypothesis testing, and e) Complete discussion of an exercise.

A first example is use to describe the two-factor factorial design and introduce the basic definitions: response variable, factors, treatments, experimental units, and replicates. Then, the fundamental concept of this design, the interaction between factors, is discussed extensively using this example. The definitions of main effects and simple effects as well as the consequences for data analysis of this factor interaction are also discussed. PowerPoint is extremely helpful for these discussions by illustrating the numerical and graphical interpretation of the Interaction effect.

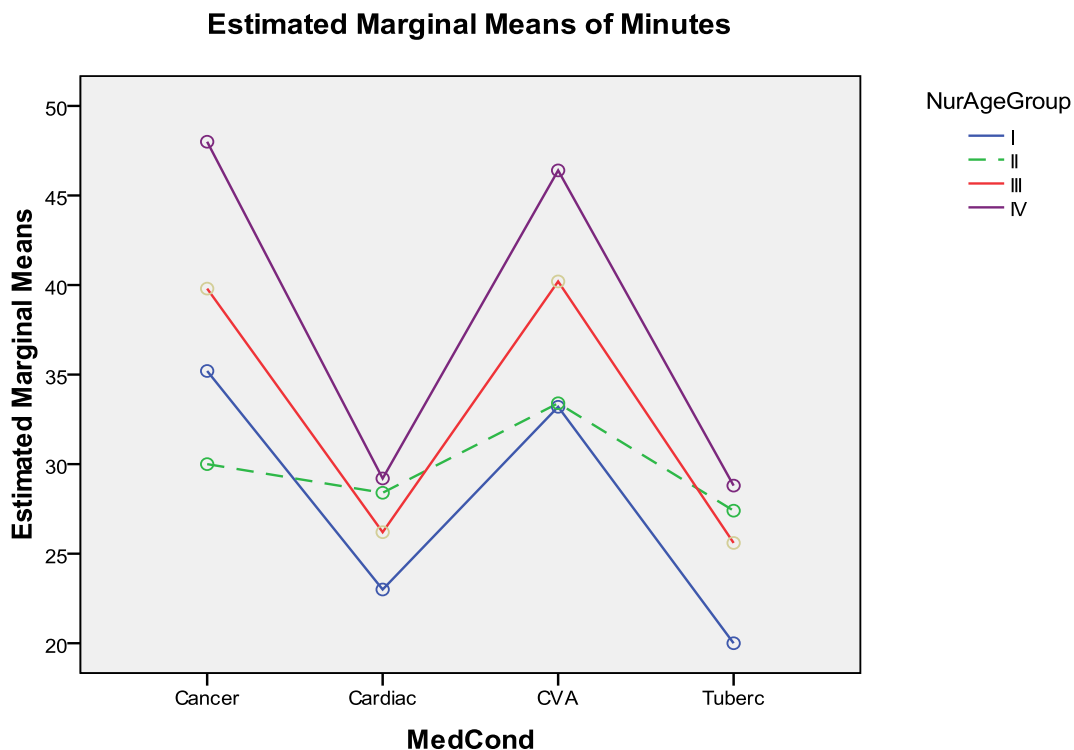
After that, a more technical stage begins with the partition of the total variance of the response variable, in general. This is a requirement for the introduction of the two-way ANOVA model with interaction that provides the statistical tool for data analysis. The endpoint of this phase is the associated ANOVA table that provides the F-test statistic for hypothesis testing. However, the calculation of the Sums of Squares components of the table from the raw data involves long and tedious computations using a hand calculator. The complexity of this data analysis is evidenced to the students, and consequently the need of statistical software warranted. Following this phase, the steps for the hypothesis testing procedure are discussed including the post-hoc multiple comparisons of means.

The final stage consists of a complete discussion of exercises. To illustrate this phase, a problem used in class is presented here. The exercise can be found in the textbook for STA-3194: Biostatistics by Wayne Daniel (2009). The problem describes a study whose objective is to determine the effect on the length of visit to patients for nurses according to the patient's Medical Condition (Factor A) and nurse's Age Group (Factor B). Four medical conditions are considered: cancer, cardiac, CVA, and tuberculosis. Four age

groups are defined: 20-29, 30-39, 40-49, and 50 or more years. Five nurses for each of the sixteen resulting combinations (treatments) were included in the study. Length of time for each of the eighty nurses was recorded.

The SPSS output for the Two-way ANOVA with Interaction procedure on these data, previously ran by the Instructor, is part of the PowerPoint presentation. The discussion is then initiated with the analysis of the ANOVA table from the SPSS output, that indicates a statistically significant Interaction between the two factors. This is graphically supported by the plot of treatments means (Figure 1), also obtained from the software output, showing that the effect of patients' medical condition for nurses' Age Group II (30-39 years old) is inconsistent with the other age groups.

Figure 1: Plot of treatment means (all age groups)



As a consequence of the statistically significant Interaction, the main effect of patient's Medical Condition is distorted. The combined means for each Medical Condition is not representative of all nurses' age groups. One-way ANOVAs may be conducted for the different Age Groups or, in this case, data from Group II may be separated while Groups I, III and IV are analyzed together.

The resulting SPSS output for the combined three groups confirmed a non-significant Interaction. The main effect of patient's Medical Condition is now clearly depicted with

nurses dedicating about 15 more minutes for Cancer and CVA patients compared to Cardiac and Tuberculosis patients. On the other hand, the One-way ANOVA for the nurses' Age Group II showed non-significant difference between medical conditions at the 1% significant level, with differences within 6 minutes (Table 1).

Table 1: Sample means by patient's Medical Condition

	Cancer	Cardiac	CVA	Tuberc.
Age Groups I, III, IV	41.0	26.1	39.9	24.8
Age Group II	30.0	28.4	33.4	27.4

All these analyses and discussions using technology resources contribute to an increased students' motivation and understanding that would be virtually impossible with a traditional teaching approach.

3. Results

The STA-3112 class taught by the present author during the spring of 2010-2012 years included three partial tests and a cumulative final exam. Students enrolled were mostly Biology majors. The second partial test comprises exclusively the analysis of variance models. This test always contains a problem related to the two-factor factorial design and two-way ANOVA model with interaction, having roughly a 30% of weight in the total score. Out of 130 students that took the test over the given period, a passing rate of 95% was achieved. Moreover, students' satisfaction was high demonstrated by the 91% of excellent/very good opinions about the overall quality of instruction, as assessed by the official university surveys. Also, the combined retention rate for the three groups was 98.5% during this period.

4. Discussion and Conclusions

The use of Power Point where text was presented in conjunction with tables, graphs and other pictorial representations assisted students, particularly those with a more visually oriented learning style. The course pack comprised of PowerPoint slides helped students to focus on class discussions by minimizing the note-taking process. Furthermore, the integration of computational technology provided an effective tool for this topic by generating more time for data analysis and conceptual understanding. The use of SPSS output for statistical graphs and tables led to a deeper problem comprehension. Instructor's mobility in the classroom, granted by the use of a presenter device, also facilitated communication with the students.

To assess the effectiveness of this technology based approach while teaching the Two-Way ANOVA with Interaction, the passing rates for STA-3112 (period 2010-2012) can be compared to the STA-3123 rates (period 2004-2009). Results for the STA-3194 should not be considered due the selected nature of the students and different class setting. The present author had taught the STA-3123 course using traditional resources and a more teacher centered approach during the summer term of the 2004-2009 years. Most students enrolled were Psychology majors and there were only two partial tests. The two-way ANOVA with interaction was included in the second test. The combined passing rate for this test was below 70%. Even though these outcomes are not direct evidence, considering the possible limitations of comparability due to different students' majors and assessment instruments, still the combined passing rate for STA-3112 students demonstrated to be overwhelmingly higher.

This discussion indicates that the use of technology resources in conjunction with an interactive approach provides a highly effective teaching-learning method for the Two-Way ANOVA model with Interaction. The following conclusions substantiate the improvements related to the quality of instruction and students' understanding:

- Students learned more effectively as demonstrated by the higher passing rate when compared to groups with a more traditional teaching method.
- Students' satisfaction was high as indicated by the 91% of excellent/very good opinions about the overall quality of instruction.
- Students' motivation was also high as indicated by the 98.5% retention rate.

References

- American Statistical Association. (2005). Guidelines for assessment and instruction in statistics education (GAISE): College Report. www.amstat.org/education/gaise.
- Daniel, W. W. (2009). Biostatistics: A foundation for analysis in the health sciences, 9th edition. Hoboken, NJ: John Wiley & Sons
- Garfield, J.B. (1995). How students learn statistics. *International Statistics Review*, 63, 25-34.
- Gomez, R. (2010). Using technology in large statistics classes. *Review of Higher Education and Self Learning*, Vol. 3, Issue 5, 109-113
- Gomez, R. (2010). Innovations in teaching undergraduate statistics courses for biology students. *Review of Higher Education and Self Learning*, Vol. 3, Issue 7, 8-13
- Gomez, R. (2011). Teaching a second statistics course for undergraduate psychology students. *Review of Higher Education and Self Learning*, Vol. 4, Issue 11, 1-6
- Higazi, S. M. (2002). Teaching statistics using technology. *ICOTS6 Proceedings*.
- Hulsizer, M. C. & Woolf, L. M. (2009). A guide to teaching statistics. Malden, MA: Wiley-Blackwell.

- Lock, R. (2005). Teaching a technology-enhanced course in a liberal arts environment. *Innovations in teaching statistics*, MAA Notes #65, 31-38.
- McClave, J. & Sincich, T. (2011). *Statistics*, 11th edition. Upper Saddle River, NJ: Prentice Hall.
- McClave, J. & Sincich, T. (2013). *Statistics*, 12th edition. Upper Saddle River, NJ: Prentice Hall.