

Advanced Statistics I: Analysis of Variance and Regression ***Psychology 143 Spring, 2011***

Instructor

Alan Hartley
Office: Steele 118
Office Hours: 9:00-11:00 MW and by appt
E-mail: alan@scrippscollege.edu

Tutor

Zoe Ravich
By Appointment
Email: zravich4381@scrippscollege.edu

*Goals of the Course*¹

Research in the biological, behavioral, and social sciences employs two principal approaches: *experimental* designs in which the researcher manipulates or controls the conditions under which observations are made and *correlational* designs in which the researcher can gather information about the conditions along with the observation but doesn't control those conditions. Experimental data are commonly analyzed using *analysis of variance*; correlational data are explored and analyzed using *multiple regression*. The goal of this course is for you to gain substantial, practical skill at using these procedures to analyze data from designs that are commonly used in actual research projects. You will have extensive experience using computer-based statistical packages for data analysis (primarily SPSS, the *Statistical Package for the Social Sciences*).

Although we will concentrate on analysis of variance and regression, as class interests suggest and as time allows we will explore still more advanced topics: multivariate analysis of variance (MANOVA), logistic regression, and exploratory factor analysis. There will be a brief introduction to confirmatory factor analysis, path analysis and structural equation modeling (which are covered more extensively in Special Topics in Quantitative Methods: Structural Equation Modeling, PSY144, known affectionately as Advanced Advanced Statistics). You will learn, intuitively, what these statistics do, when they are appropriate, how to execute them in SPSS, and how to interpret the results.

In the end, you should be well prepared for your senior thesis, for graduate study in behavioral and social science, for basic research projects in graduate school, and for applied research projects in the "real world" (such as market research or public policy research). Individuals who have completed this course have frequently been successful with requests to be exempted from much or all of their graduate statistics requirements. The mathematical level of the course will not be high. The emphasis will be on intuitive understanding and skill at applying statistics. Nonetheless, we will make forays into nontrivial mathematics where that is important and I will often give you brief sketches of the mathematical underpinnings of

¹ **The Vice President for Academic Administration at Scripps College, Dean Amy Marcus-Newhall, has mandated that every course syllabus must contain "Learning Outcomes™" for the course. The outcomes of successfully completing this course will be that you will have mastered the two principal statistical tools for analysis of scientific, marketing, and public policy research, analysis of variance and multiple regression. You will know when their use is appropriate, how to carry them out, and how to interpret and present the results. You will also be substantially more likely to be exempted from graduate-level statistics requirements than if you did not take the course.**

the statistical methods you are learning.

Prerequisites

I expect that you are thoroughly familiar with the logic of statistical inference (hypothesis testing and the concept of significance). These topics won't be reviewed. I expect that you will have been introduced to one-variable (or one-way or single-factor) analysis of variance and to simple, two-variable correlation. We will review these topics, so you will have a chance to refresh what may have faded.

Organization of the Course

The most important aspect of this course is that it is relatively small. Because it is small, we can be very informal and can adjust our pace as we go. We can emphasize or deemphasize topics depending on your interests and needs. Evaluation can be based on in-class contributions, exercises, and projects rather than on formal examinations. Because the course is tailored to you, I will have to learn as we go which topics are hard for you, which are easy, how much time to devote to a topic, and what we can realistically accomplish. For this, I will rely on your assistance.

Evaluation

My specific goal for you is that you be able to pass an exam to be exempted from graduate statistics. If you are performing at that level, your grade in this class will be "A." If I feel that you are slipping below that standard, I will consult with you about how to overcome the problems and how to demonstrate your mastery (which may involve special assignments or a formal test). There will be homeworks and projects (larger homeworks). However, my evaluation of you is an amorphous subjective judgment based on your informal work (the questions you ask and answer, watching over your shoulder as you work, etc.) as well as formal work. For this reason I expect you will be in class every time.

Course Notes

I will post my lecture notes to Sakai in advance of each class (in the form of Word files). You are encouraged either to print the lecture notes and use them to take notes in class or to call them up in class, keep them in an open window, and add to them as you wish.

Texts

There are two assigned texts from the course. You are not required to have them. Many people do just fine relying on the course notes. It is often helpful to hear different approaches to the same material, so I will identify relevant readings in these texts as we go through the course.

Robert K. Leik, (1997). *Experimental Design and the Analysis of Variance*. Pine Forge Press. ISBN 978-080399006.

Paul D. Allison (1998). *Multiple Regression: A Primer*. Pine Forge Press. ISBN 9780761985334

Although there are two assigned texts, the books listed below can be very useful for additional reading on topics we take up and would be useful as references in the future.

James Jaccard, & Michael Becker. (1996). *Statistics for the Behavioral Sciences (4th ed.)*. Belmont, CA: Wadsworth. *(This has been the text for my course in Psychological Statistics, PSYC 103 for several years. If you do not own it, you may well be able to borrow a copy from someone who has taken the course. Earlier editions are just fine. Alternatively, you should be able to use the text from the statistics course you took. We will review the material in Chapters 11 and 12 (independent groups ANOVA and repeated measures ANOVA), and cover the material in Chapter 17(two-factor ANOVA).*

Jacob Cohen, & Patricia Cohen. (2003). *Applied Multiple Regression/ Correlation Analysis for the Behavioral Sciences*. Mahwah, NJ: Erlbaum. *(This is the latest edition of the book I turn to for reminders about or tutoring on any aspect of regression.)*

Lawrence Hamilton (1992). *Regression with graphics*. Belmont, CA: Brooks/Cole. *Really excellent coverage of regression.*

Timothy Keith. (2006). *Multiple regression and beyond*. Boston : Allyn and Bacon. *Excellent, understandable coverage of regression and other advanced topics.*

Roger Kirk. (1995). *Experimental Design*. Belmont, CA: Brooks/Cole. *(This is the best reference book there is for analysis of variance and the design of experiments. Any edition will do; I like the first one, 1968, best.)*

Barbara G. Tabachnick, Linda S. Fidell. (2001). *Computer-assisted research design and analysis*. Boston : Allyn and Bacon. *Excellent coverage of analysis of variance and related topics.*

Barbara G. Tabachnick, Linda S. Fidell. (2001). *Using multivariate statistics*. Boston : Allyn and Bacon. *Excellent, understandable coverage of regression and other advanced topics.*

Course Topics

Background

Analysis of Variance (ANOVA)

Review of single-factor ANOVA (Jaccard & Becker Ch. 12, 13)

Independent-groups ANOVA

Repeated-measures ANOVA

Comparisons among means

Tests for trend

Assumptions of ANOVA

Two-factor ANOVA (Jaccard & Becker, Ch. 17)

- Finding & interpreting interactions
- Completely-between-subjects designs
- Exploring interactions
 - Simple main effects
 - Treatment-contrast interactions
- Completely-within-subjects designs
- Split-plot designs

Multi-factor ANOVA

MANOVA: ANOVA with multiple dependent variables

Regression

Fundamentals

Review of bivariate correlation & regression

Multiple regression

- Doing & interpreting regression
- Exploring data with regression
- Looking for problems in the data: Regression Diagnostics
- Dummy Variables
- ANOVA is a special case of regression
- Analysis of Covariance (ANCOVA)

*Special topics in regression

Avoiding problems: Robust regression

Logistic regression

Non-linear relationships

**Factor Analysis*

Principal components analysis

Exploratory factor analysis

**Brief Overview of Structural Equation Modeling (covered in detail in Advanced Advanced Statistics PSY144):* Path Analysis, Confirmatory Factor Analysis, Structural Models

* *These topics will be covered as time allows and interest warrants.*