Educational Psychology 711-002: Applied Regression Analysis
University of Wisconsin–Madison / Department of Educational Psychology
Fall 2014

Lecture: TuTh 9:30-10:45 AM Room: 123 Noland
Instructor: Peter M. Steiner
Office: 1061 Educational Sciences, (608 262 0842)
e-mail: psteiner@wisc.edu
Office Hours: TuTh 11:00-12:00 AM, and by appointment

Objectives: This course provides an introduction to basic regression techniques and serves as a solid foundation for more advanced methods like hierarchical linear modeling (HLM) or structural equation modeling (SEM). The general objective is to make you familiar with the theoretical foundations of regression analysis as well as its application to real datasets. During the course you will learn how to specify and interpret regression models using data provided by the instructor and your own data. All the analyses will be done in R which is a free language and environment for statistical computing and graphics (http://www.r-project.org/). R will also be used for demonstrating computational aspects related to regression techniques (e.g., nonparametric regression). By the end of the course you should be able to (i) understand the main concepts of regression, (ii) analyse your own data and draw meaningful conclusions with regard to your research questions, and (iii) critically read and evaluate social science publications that use regression models as analytic tool.

Prerequisites: Introductory statistics classes like Ed Psych 760, 761, or equivalent. Familiarity with a major statistical software program (e.g., R, Stata, SPSS/PASW or SAS) or a programming language is highly desirable.

Requirements & Grading:

	Requirement	max. points	weight for final grade
1.	14 weekly assignments	14×4 pts	70%
2.	Paper	4 pts	30% (i.e., points will be multiplied by 6)

Thus, the maximum number of points you can achieve is 80 points ($14 \times 4 + 6 \times 4$):

A: 70 pts or more AB: 60-69 pts B: 50-59 pts BC: 40-49 pts
C: 30-39 pts D: 20-29 pts F: 19 pts or less

Books & Lecture Notes:
Required Textbooks:

Fox, J. (2008). *Applied Regression Analysis and Generalized Linear Models* (2nd edition). Thousand Oaks, CA: Sage. Online resources (appendix, data, etc.): http://socserv.socsci.mcmaster.ca/jfox/Books/Applied-Regression-2E/index.html

Fox, J., & Weisberg, S. (2011). *An R Companion to Applied Regression* (2nd edition). Thousand Oaks, CA: Sage. Online resources (appendices, scripts, etc.): http://socserv.socsci.mcmaster.ca/jfox/Books/Companion/index.html

Lecture notes will be available online but they are *not* substituting the required textbooks. Note that you are required to read selected chapters of both books *before* each class (see below).

Software:
Throughout the course, we use the free software package R (http://www.r-project.org/).

Course Schedule: Topics, Readings & Assignments

Week 1 Sep 2 & 4	Introduction to the course and Intro to R *Readings*: ARA, Chapter 1; CAR, Chapters 1-2
Week 2 Sep 9 & 11	Intro to R; Review of basic statistical concepts *Readings*: ARA, Chapter 3; CAR, Chapters 1-2 & 3.1-3.3
Week 3 Sep 16 & 18	Intro to R; Basic statistical concepts & simple nonparametric regression *Readings*: ARA, Chapter 2; CAR, Section 3.1-3.3, 7.1
Week 4 Sep 23 & 25	Simple nonparametric regression *Readings*: ARA, Chapter 2
Week 5 Sep 30 & Oct 2	Linear regression *Readings*: ARA, Chapter 5; CAR, Sections 4.1-4.2, 4.8-4.9
Week 6 Oct 7 & 9	Linear regression & Statistical inference *Readings*: ARA, Chapter 6; CAR, Sections 4.1-4.4, 4.8-4.9
Week 7 Oct 14 & 16	Linear regression & Statistical inference *Readings*: ARA, Chapter 6; CAR, Sections 4.1-4.4, 4.8-4.9
Week 8 Oct 21 & 23	Regression with categorical independent variables & interactions *Readings*: ARA, Chapter 7; CAR, Sections 4.2.3, 4.6 ***Project proposal due (Oct 23)***
Week 9 Oct 28 & 30	Regression with interactions; Review *Readings*: ARA, Chapter 7; CAR, Sections 4.2.3, 4.6
Week 10 Nov 4 & 6	Review: Intro to matrix algebra & linear regression in matrix form *Readings*: ARA, Chapter 9 & Online Appendix B; CAR, Sections 2.3, 8.2.1, 8.2.2
Week 11 Nov 11 & 13	Review: Linear regression in matrix form; Model diagnostics I *Readings*: ARA, Chapter 11 & 12; CAR, Chapter 6
Week 12 Nov 18 & 20	Model diagnostics II *Readings*: ARA, Chapter 12 & 13; CAR, Chapter 6
Week 13 Nov 25	Generalized linear model (logit model) *Readings*: ARA, Chapter 14.1; CAR, Sections 5.1-5.4
Week 14 Dec 2 & 4	Generalized linear model (logit model) *Readings*: ARA, Chapter 14.1; CAR, Sections 5.1-5.4
Week 15 Dec 9 & 11	Generalized linear model (logit & probit model) *Readings*: ARA, Chapter 14.1; CAR Sections 5.1-5.4 ***Paper due (Dec 11)***

Indicated chapters & sections of ARA and CAR need to be read *before* each class!
ARA: Fox, J. (2008). *Applied Regression Analysis and Generalized Linear Models*
CAR: Fox, J., & Weisberg, S. (2011). *An R Companion to Applied Regression*

Weekly Assignment and Paper

Weekly Assignments (Week 1 to 14): The weekly assignments consist of focused exercises (e.g. writing and applying an R-function or running a regression model and interpreting the R-output). These assignments are handed out on Thursday and are due on Tuesday of the following week. Weekly assignments are strictly limited to a two page maximum!

Paper: The main assignment consists of writing a paper using data from a study you work on or a dataset that is of interest to you. If you do not have your own dataset, I will provide you with one. You will need to hand in a brief proposal describing the problem and the dataset (max. of two pages). Ideally, the dataset should meet the following requirements:

- at least one continuous (e.g., achievements scores or income)
- at least 3 continuous and 3 categorical predictors (independent variables)
- there is no strict sample size requirement, but more than 100 observations are desirable.

In the *project proposal* (max. of two pages) you should briefly describe the substantive problem along with a few key references. Since regression analysis is the required method you should describe the dataset in enough detail such that I can assess whether the data are suitable for our purposes. In an appendix to the proposal you should give descriptive statistics for the key variables (the independent variables and predictors as mentioned above).
The *paper* then should analyze the formulated research question using appropriate regression techniques. You should begin by providing descriptive statistics for the most important variables entering your regression model. Tables and plots should be provided if they are useful in describing the data. Then you should try to specify a useful model using substantive theory and model diagnostics. The final model might require including higher order terms (e.g., quadratic terms or interactions) or transforming variables (e.g., log transformations). The paper should be written in a style consistent with the major publication outlet in your field (e.g., APA formatting style). In particular, it should cover the following sections: Introduction & research question, data and methods, results, and conclusions & discussion (plus references and appendix). More technical descriptions of the regression model (e.g., diagnostics) and R syntax can be included in an appendix. The maximum number of pages is limited to 10 pages (excluding the appendix). The paper should be written as coherently as possible, as if you were submitting it for publication.

Weekly assignments and the paper will be scored with 0, 1, 2, 3, or 4 pts:

4 = Excellent understanding of the assignment. All of the required elements are present and correctly interpreted. The paper demonstrates excellent depth of understanding with respect to the link between theory, results, and interpretation. Write-up is in an appropriate style with minor typographical errors.
3 = Good understanding of the assignment. All of the required elements are present and correctly interpreted. Minor lack/problems in understanding or in the depth of the analysis. Write-up is in an appropriate style with minor typographical errors.
2 = Adequate understanding of the assignment. All the required components are present but without sufficient depth of understanding with respect to the interpretation of results. Write-up is in appropriate style but there are some grammatical and typographical errors.
1 = Inadequate understanding of the statistical procedures required for analyzing the problem. Some required elements of the assignments are missing, such as tests of assumptions or interpretation of important results. Severe problems in understanding the link between theory, results, and interpretation. Write-up may be in the appropriate style but there are numerous typographical and grammatical errors.
0 = No understanding of the nature of the assignment. Incorrect interpretations of results. Write-up is not in an appropriate style and there are numerous typographical, logical, and grammatical errors. Also applies if the assignment is not handed in timely.

R Resources

- R can be downloaded from http://www.r-project.org/.
- Online contributed documentations (contain lots of introductory papers/texts on R) are available from http://cran.r-project.org/other-docs.html .
- For a guide on the programming style in R see http://google-styleguide.googlecode.com/svn/trunk/google-r-style.html .
- You may also use an editor that uses syntax highlighting and provides other nice features:
    - RStudio: http://rstudio.org/
    - Tinn-R: http://sourceforge.net/projects/tinn-r/ (install also the TinnR R-package)
- For a short interactive online-introduction see http://tryr.codeschool.com/ .

R Textbooks

- Dalgaard P. (2008). *Introductory Statistics with R*. Springer.
- Everitt, B.S., & Hothorn, T. (2006). *A Handbook of Statistical Analyses Using R*. Chapman & Hall/CRC.
- Spector, P. (2008). *Data Manipulation with R*. Springer. [A useful book for manipulating data in R.]

Other Excellent Textbooks on Applied Regression

- Cohen, P., Cohen, J., West, S. G., & Aiken, L. S. (2003). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (3rd edition). Mahwah: Lawrence Erlbaum Associates.
- Wooldridge, J. M. (2009). *Introductory Econometrics: A Modern Approach* (4th edition). Mason, OH: South-Western Cengage Learning.
- Weisberg, S. (2005). *Applied Linear Regression* (3rd edition). Hoboken NJ: John Wiley.