# Course Title: Advanced Quantitative Research Methods

## Hilary Term 2017

## Lecturer/s: David Barron

# Course Aims and Objectives:

The course aims to give students a solid understanding of some more advanced quantitative methods that are in common use in management research. The emphasis is on understanding the situations in which each method is appropriate, on understanding the limitations of each method and on being able to interpret the results produced when these methods are implemented.

Students will also learn how to implement the methods using a fast-growing statistical analysis software system called R. This is an open-source package that can be freely downloaded from <http://cran.r-project.org>. Students should install it on their own laptops.

# Course Syllabus (by week/topic)

## Week One

### Dummy variables, interactions and non-linear effects in linear regression

### *This session will cover commonly used special types of explanatory variable: dummy variables (used to include categorical variables in regressions), non-linear effects (used when the effect of a variable depends on the value of that variable), and interactions (used when the effect of one variable depends on the value of a second variable).*

## Week Two

### Regression diagnostics and robust estimation

*The main purpose of this session is to examine various diagnostic techniques that can be used to check various aspects of a regression analysis, such as the presence of outliers, heteroskedasticity, multicollinearity, and sample selection bias. Methods for dealing with problems identified using these methods will also be discussed.*

## Week Three

### Maximum likelihood and logistic regression

*The principle of maximum likelihood estimation will be introduced. We will consider the application of this method to regression models for binary and ordinal outcome variables. We will also consider the interpretation of the results of such regressions.*

## Week Four

### Event models

*Special methods are required when the outcomes of interest are events (for example, the arrival of customers, or the occurrence of a merger). Sometimes we have counts of the number of events in a given period of time (eg, new firms founded in a year). In other situations we have the time at which events occurred (eg, the date on which a firm was liquidated). These two types of outcome require special regression methods, which we will discuss in this session.*

## Week Five

### Multilevel models

*Many data sets are hierarchical in nature (eg, employees within firms within industries). Special measures are needed to deal with such data, because individual observations cannot be assumed to be independent. In addition, we might be interested in modelling effects at different levels of analysis at the same time. We will consider how best to deal with such multilevel analyses.*

## Week Six

### Longitudinal data analysis

*Longitudinal data consist of repeated observations of a set of units (individual people, organizations, etc.). Such data, sometimes called panel data, can be very powerful. However, there are challenges involved in its analysis, in large part because it is unrealistic to treat successive observations of the same individual as being independent. Longitudinal data can be thought of as a special case of multilevel data, so we will consider extensions of the methods introduced last week to this class of data.*

## Week Seven

### Factor Analysis

*Factor analysis is a method for reducing a large number of variables to a smaller number of statistically independent variables, often used in conjunction with psychometric scales administered in surveys, but also with a wide range of other applications. We will consider exploratory factor analysis, principal components analysis, and factor rotation.*

## Week Eight

### Structural Equation Models

*Structural equation models can be seen as an extension of factor analysis, and include as a special case confirmatory factor analysis models. However, we will also look at their use in regression.*

# Course Readings (by week and/or general)

## General readings

UCLA Statistical Computing site: <http://www.ats.ucla.edu/stat/r/> I strongly recommend that all students look at the R starter kit on this site before the first session. It includes instructions for installing R on your computer; you should do this before coming to the first class.

Introduction to R: <http://cran.r-project.org/doc/manuals/R-intro.pdf>

Fox, John. 2011. An R Companion to Applied Regression. Sage.

Faraway, J. 2002. Practical Regression and ANOVA using R. <http://cran.r-project.org/doc/contrib/Faraway-PRA.pdf>

## Weeks One and Two

### Further topics in linear regression

* Agresti, Alan and Barbara Finlay. 1997. Statistical Methods for the Social Sciences. London: Prentice-Hall.
* Hair, J.F., R.E. Anderson, R.L. Tatham and W.C. Black. 1999. Multivariate Data Analysis with Readings. London: PrenticeHall. Chapters 2 and 3.
* White, Halbert. 1980. “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.” Econometrica 48: 817838.
* Berk, Richard A. 1983. “An Introduction to Sample Selection Bias in Sociological Data.” American Sociological Review 48: 386-398.

## Weeks Three and Four

### Maximum likelihood and regression

* Agresti, Alan and Barbara Finlay. 1997. Statistical Methods for the Social Sciences. London: Prentice-Hall. Chapter 15.
* Agresti, Alan. 1990. Categorical Data Analysis. London: PrenticeHall. Chapter 4.
* Maddala, G. S. 1983. Limited-Dependent and Qualitative Variables in Econometrics. Cambridge University Press. Chapter 2.
* Brostrom, G. 2012. Event History Analysis with R. Chapman & Hall.
* Mills, M. 2011. Introducing Survival and Event History Analysis. Sage.

## Week Five

### Multilevel models

* Gelman, Andrew and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models.* Cambridge: Cambridge University Press.

## Week Six

### Longitudinal data analysis

* Singer, Judith D. and John B. Willett. 2003. Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. Oxford University Press. Chapters 1-4

## Week Seven

### Factor Analysis.

* Agresti, Alan and Barbara Finlay. 1997. Statistical Methods for the Social Sciences. London: Prentice-Hall. Chapter 16.
* Hair, J.F., R.E. Anderson, R.L. Tatham and W.C. Black. 1999. Multivariate Data Analysis with Readings. London: PrenticeHall. Chapter 7

## Week Eight

### Structural Equation Models

* Fox, John. Structural Equation Models.http://socserv.mcmaster.ca/jfox/Books/Companion/appendix/Appendix-SEMs.pdf
* Long, J. Scott. Covariance Structure Models. Sage.
* Long, J. Scott. Confirmatory Factor Analysis. Sage
* The lavaan project. <http://lavaan.ugent.be>. This is an R package for estimating SEMs that I find easier to use than John Fox’s sem package.

# Assessment and Due Date

The Advanced Quantitative Research Method’s course will be assessed by the production of a report of **no more** than 6000 words on the following task:

Using any quantitative data and any statistical analysis software you chose, write a report in the format of a journal article (e.g., *Academy of Management Journal, Administrative Science Quarterly, American Sociological Review,* etc.). The article should involve the testing of hypotheses using some form of quantitative analysis that has been included in the course. In the article you should discuss the theory and relevant literature (though this part will likely be less thorough than would normally be acceptable in a journal article) and clearly state the hypotheses being tested as well as describing the statistical analysis. This should include a discussion of how the data were collected, key descriptive statistics of the variables used in the analysis, why the particular methods used were chosen, a clear explanation of the methods used, tests of whether any assumptions involved in the choice of method were met, a clear presentation of the results of the analyses carried out and your interpretations of these results.