

**THE HELLER SCHOOL FOR SOCIAL POLICY AND MANAGEMENT
BRANDEIS UNIVERSITY**

HS246f
Statistics
Fall 2007 Module

Professor Stephen F. Fournier
(781) 736-3898

Meets: Monday 2:10 – 5:00
Room: TBD

Fournier@brandeis.edu

Class Website: LATTE.Brandeis.edu

Statistics for Managers: A Rationale

As a primary part of their work, managers constantly make decisions. Fundamentally, these decisions are all based on weighing the relevant information available. However, information is frequently derived from measures that are often quite difficult to make directly, thus introducing some degree of randomness. Even with numerous direct measures, many observed phenomena contain a random component (to a greater or lesser degree). In either case, randomness is important to managers because it means that the consequences of an action may not be predictable with certainty. Reducing randomness, understanding its size and potential impact in a management situation, is a valuable area of knowledge to which the study of statistics can contribute greatly.

Not all randomness is the same—it derives from myriad sources and can be accounted for and even reduced in various ways. Because some randomness reflects a lack of information, receiving more or better information can lessen that part of the randomness, an idea behind quality control. Although some phenomena are random, many random phenomena follow probability distributions that have systematic features. Different random processes have some well-studied probability distributions; statistical knowledge offers managers insight into these matters. Statistical models provide a way to estimate the true relationships we may theorize to exist; to inform the decisions we need to make. By incorporating the concepts of randomness, our models also include a relative degree of certainty related to results—a way for us to test if our results are ‘statistically significant’.

With randomness and uncertainty comes the very real possibility of error. We may determine that a result is ‘statistically significant’ and yet still be wrong. Even here, statistics offers a systematic way to examine how that error may come about. The models we will learn to develop and the measures we will learn to use can offer insight into how to treat the various sources of error, the probabilities of each of those errors, and ways to decide how to minimize the impacts of those errors.

Course Description: This seven-week module presents students with an introduction to the fundamentals of parametric statistics.¹ We will cover the essentials required for students to understand issues related to measurement and how to generate descriptive information and statistical analyses from these measurements. The intent of the course is to provide the

¹ We will touch upon a small number of non-parametric topics, but the primary focus will be as described.

background to understand the process of statistical model building and the course will culminate in learning how to generate multiple linear regression analyses. To this end, the focus will be primarily on understanding the importance of summary measures (most notably central tendency and dispersion), along with a study of fundamental statistical distributions (Z , t , χ^2 , F).

Although some initial work will be done by hand, computers are an integral part of this course and we will begin utilizing Excel immediately. I will introduce you to the basics of formulae and functions in Excel, along with a detailed explanation of the Data Analysis add-in. In addition, as part of this class we will also begin to use STATA as our statistics/data analysis package. There will also be a lab associated with this course for help in further understanding the software involved.

Management Discipline Skills and Competencies: A good manager must be able to understand measurement information provided and use that information in a variety of ways. Statistical model building is a primary tool in this process. This class will provide students with the ability to:

- Think critically about issues of measurement
 - What is being measured
 - Why is *this* being measured
 - How can I use this information to its best use
- Utilize information for both descriptive and analytical purposes
- Construct and generate causal models

MBA Program Distinctive Themes: This course is firmly located in the ‘Measuring and Managing Outcomes’ listing of the eight MBA themes. However, knowledge from this course forms a basis for work done in each of the other themes.

Course Requirements: Because this module meets only 7 times, it is crucial that students attend all sessions. There will be assigned problem sets (graded) for each session to be handed in before class begins (as solution sets will be posted and discussion of the work done will comprise the first part of the next session). Students will be expected to have prepared thoroughly for class (reading the assigned text and completing problem sets) and can assume that they will be called upon to contribute to in-class discussions. These problem sets are a crucial aspect of this course as they provide the information needed to assess how well the material is being both conveyed and, more importantly, understood. We will have both a mid-term (in-class) and a final (take home). My expectations are that by course end students will feel comfortable presenting summary statistical information (both descriptive and inferential) and be able to present analyses using linear regression techniques.

Course Reading: The course will rely primarily upon two online statistics texts: *The StatSoft Electronic Textbook*² and *The HyperStat Online Textbook*³. Selected chapters from other texts will be provided along with suggested titles for hardcopy textbooks. Two small primers (Statistics Primer, Regression Primer) and a few additional handouts also will be made available either in-class or online.

Provisions for Feedback: Feedback will be provided along a number of paths. The problem sets offer a rich weekly source on your progress, as well as in class interactions. The midterm

² URL: <http://www.statsoftinc.com/textbook/stathome.html>

³ URL: <http://davidmlane.com/hyperstat/>

and final will also provide strong indications of your progress. I will also offer two hours a week of office time at which I will be available if you have any problems/concerns about your work. As I spend much of my time online, e-mail questions are absolutely welcomed I will respond as quickly as possible.

Academic Integrity: Academic integrity is central to the mission of educational excellence at Brandeis University. Each student is expected to turn in work completed independently, except when assignments specifically authorize collaborative effort. It is not acceptable to use the words or ideas of any other person without proper acknowledgement of that source. This means that you must use footnotes and quotation marks to indicate the sources of any phrases, sentences, paragraphs or ideas found in published volumes, on the internet, or created by another student. Violations of university policies on academic integrity, described in Section 3 of *Rights and Responsibilities*, may result in failure in the course or on the assignment, and could end in suspension from the University. If you are in doubt about the instructions for any assignment in this course, you must ask for clarification.

Notice: If you have a documented disability on record at Brandeis University and require accommodations, please bring it to the instructor's attention prior to the second meeting of the class.

Course Schedule

Class Session.

Week	Topics Covered	Assignment and Lab Topics
1	Data and data measurement; types of variables. Distinction between descriptive and inferential statistics. Univariate statistics: summary measures; central tendency; dispersion.	Chapters 1-4 Lane text. StatSoft Chapters 'Basic Concepts' and 'Elementary Statistics.' Read Statistical Primer Handout. Problem set 1 (generate summary statistics by hand and in Excel)
2	Bivariate statistics: covariance and correlation. Concept of population vs. sample.	Statistical Primer Handout sections on covariance and correlation.
3	Probability and probability distributions: Discussion of summary measures: populations vs. samples. Introduction to concept of sampling distributions. Further work on variance. Generating Z scores and using z-table lookups. Introduction to the t-statistic and measures of significance. Concept of outliers. Introduction to the F-test and its use in comparing measures of variance.	StatSoft Chapters 5-8. Problem set 2 (utilize Z tables, generate confidence intervals, utilize t-tables). More Bivariate data (correlations); graphing in Excel. Generating statistical output in STATA and/or GRETL.
4	Re-examination of Z and t using Hypothesis test structure: duality of this approach vs. levels of significance. Significance of correlation. T-test for the difference	HyperStat Chapter 12 (Anova) Problem set 3 (t-test examples, review of univariate and bivariate tests). Doing t-tests in Excel and in

Week	Topics Covered	Assignment and Lab Topics
	between two means. Examination of correlation as introduction to causal modeling. Extension of t-test from two means to ANOVA for multiple means (further use of F-test).	STATA and/or GRETL. Comparison of output. Review of all material for test.
5	Introduction to Simple Linear Regression Analysis: further development of correlation concepts; discussion of 'least-squares' approach; assumptions of the model. Generating and interpreting output; understanding the x coefficient and its significance.	Simple Linear Regression Handout. StatSoft Chapter 'Linear Regression'. HyperStat Chapter 15.
6	Multiple Regression Analysis and review: extension from the simple model to multiple variable case.	Problem set 4 (multiple regression examples in Excel and in STATA and/or GRETL)
7	Model violations: implications and possible corrections. Limits to models. Review of all material for comprehensive take-home final.	Take home final.