



ECON 210f: Fundamentals of Statistical Analysis

Brandeis International Business School

Fall Module I 2012

Monday and Wednesday 5:00 – 6:20

Location: TBA

Stephen Fournier

Hours: TBA

781-736-3898 Office: Heller Room 154

Fournier@brandeis.edu

Overview

This module introduces statistical thinking and fundamental analytical methods to first-year IBS graduate students with little or no prior statistics training. The intent of the module is to develop a set of tools to enable critical analysis of statistical evidence to inform business decisions. Topics include the development and use of both descriptive and inferential statistics, continuous probability distributions, estimation, tests of hypotheses, and regression modeling. We will use Stata as our statistical software for the computational tasks in the course.

Required Text

We will use Stock and Watson's Introduction to Econometrics for both the ECON210 and ECON211 modules. We will also take advantage of two on-line statistics textbooks for most of our readings: David M. Lane's online [Hyperstat](#) as well as [Statsoft's Electronic Textbook](#). Articles and additional readings will also be supplied on-line.

Rationale for the Course

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.

Learning Goals: A good manager must be able to understand measurement information provided and use that information in a variety of ways. Statistical analysis and model building are primary tools in this process. Upon completion of this class students will be able to:

- Identify, evaluate, modify and manipulate data
- Utilize this data for descriptive purposes by generating, analyzing and presenting a variety of univariate statistics (sum, count, minimum, maximum, mean, median, variance, standard deviation, etc.)
- Utilize this data for inferential purposes by generating, analyzing and presenting a variety of bivariate statistics (correlations, t-tests, f-tests, chi-squares, etc.)
- Utilize this data for inferential purposes by generating, analyzing and presenting simple and multiple regression models

Prerequisites

There are no formal prerequisites for this module. Students should be familiar with Excel and basic algebra. *If you have had more than one prior course in statistics, this course may be too elementary for you. See me if you are in doubt.* The target audience for this course is those students with little or no prior background, and their needs will have the highest priority in the class.

Keeping Informed

We’ll make regular use of LATTE and a course mailing list (registering in the course automatically adds you to both lists). All lecture notes, handouts, links and other supporting materials will be available via LATTE, and any late-breaking news will reach you via the mailing list. Please check your Brandeis email regularly to keep apprised of important course-related announcements.

Participation & Contributions

IBS is truly an international community of learners, each with something to contribute to the enterprise. Each student in the class should regard participation as a chance to contribute to our joint efforts and helping fellow students to learn. Moreover, because this module aims to build both understanding and effective communication skills, class participation is important. Statistics is not a spectator sport; you will learn by *doing* rather than by watching. Participation can take many forms, and each student is expected to contribute actively, freely, and effectively to the classroom experience by raising questions, demonstrating preparedness and proficiency in the analysis of problems and cases, and explaining the implications of particular analyses in context. To this end, *class attendance is required, and students should use name cards.*

Written Assignments

We will have weekly problem sets, a take home midterm (to be done individually), and the final will consist of a group project (see below). All written material is due before the beginning of the next class unless otherwise specified. In your written work, the *clarity and correctness of your*

explanations is at least as important as the numerical correctness of your analysis. All assignments are due before the start of class in electronic format via LATTE. If you are absent, your paper should arrive electronically before class that day.

Final Project

Small teams will produce an original regression-based project. Instead of an in-class final exam, each group will present its projects to the class on the date reserved for a final exam. Assignment details will be distributed during the second week of the module. Your final grade in the course will be computed using these weights:

| | |
|---|------|
| Participation | 15% |
| Homework sets (3) | 25% |
| Take home Midterm | 35% |
| Evaluation Project (incl. presentation) | 25% |
| TOTAL | 100% |

Your course grade will reflect your performance on assignments relative to your classmates. Typically, grades range from A to B- for most students.

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.

Disabilities

If you are a student with a documented disability on record at Brandeis and wish to have a reasonable accommodation made for you in this class, please see me immediately.

Study Groups

Working with one or two partners is an excellent way to gain deep understanding of this subject. I encourage small groups to work on assignments, with a few caveats:

- Be sure that you are neither carrying nor being carried by the group; each member of the group is entitled to learn.
- For problem sets, you may work alone or with as many as 3 partners, but each person should submit their own problem set.
- Each group member retains the right to “go it alone.” Joining a group is not a marriage, and you can leave a group at any time.

Course Outline

Note: This module meets only thirteen times for about one-and-one-half hours per session; your attendance and involvement are crucial.

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 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). | HyperStat Chapter 12 (Anova) Problem set 3 (t-test examples, review of univariate and bivariate tests). Doing t-tests in Excel and in 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. |