

# WWS 600b: Introduction to Statistics

Junior Summer Institute 2011

Woodrow Wilson School

Princeton University

Lecture: M & Th. 9am-12pm Robertson 011

Precept: Th. 4 -5.30 pm

Instructor: Carlos Velasco

Office: TBD

Office hours: Thursday 2.00 - 4.00 pm or by appointment

Email: [cvelasco@princeton.edu](mailto:cvelasco@princeton.edu)

Preceptor: Miklos Bankuti

Office: TBD

Office hours: Tuesday 2.00 - 4.00 pm or by appointment

Email: [mbankuti@princeton.edu](mailto:mbankuti@princeton.edu)

## 1 Course Description

The purpose of this course is to introduce you to the basic empirical methods for the analysis of public policy. The main topics we will discuss over the next seven weeks include: descriptive statistics, probability theory, hypothesis testing, Bayesian inference, simple and multiple regression, and if there is time, maximum likelihood estimation. By the end of the summer you should feel comfortable analyzing empirical data, conducting hypotheses tests, interpreting regression results, and identifying the main problems in making statistical inferences.

## 2 Course Structure

This summer course has two major learning component modules: lectures and precepts. In the lecture we will go over essential concepts and examples illustrating their application. Precept time will be devoted to review some of the material covered in lecture and discussion of additional examples.

## 3 Grading

The grade for this course has the following components:

**Problem Sets:** 25%. There will be a total of four problem sets. Two of them will be due before the midterm; the other two before the final. We will not accept problem sets submitted after the deadline.

**Midterm Exam:** 35%. A take home midterm will be distributed at the end of the third week of the program and will be due back at the beginning of the following class. The exam will be available through Blackboard ([blackboard.princeton.edu](http://blackboard.princeton.edu)), and once you download it, you can work on it for no more than 24 hours.

**Final Exam:** 40%. A 90 minutes in-class final exam will be administered on the last session of the course.

## 4 Textbook

The main textbook for the course is:

- Freedman, David, Robert Pisani and Roger Purves. 2007. *Statistics*. Fourth edition. New York: W.W. Norton.

However, if you are interested and have the time, I also recommend that you take a look at:

- Huff, Darrell. 1953. *How to Lie with Statistics*. New York: W.W. Norton.
- Salsburg, David. 2001. *The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century*. New York: Holt Co.

## 5 Outline

- **Week 1**  
June 13: Introduction, describing one variable: density, central tendency, and dispersion  
*Required reading:* Preface and chapters 3-4 in Freedman et al.  
June 16: Describing relationships between two variables: correlation, causality, and experiments  
*Required reading:* Chapters 1-2 and 8-9 in Freedman et al.
- **Week 2**  
June 20: Probability theory and random variables  
*Required reading:* Chapters 13-15 in Freedman et al.  
June 23: Expectations, variances, covariances and conditional expectations. Problem Set #1 due in class.  
*Required reading:* Chapter 17 in Freedman et al.

- **Week 3**

June 27: Data collection and estimation theory

*Required reading:* Chapter 19 in Freedman et al.

June 30: Distribution of the sample mean, law of large numbers, and central limit theorem.

Problem Set #2 due in class. Midterm exam posted on **Blackboard** and due by Wednesday

July 5th at 5 pm in my mailbox.

*Required reading:* Chapters 5 and 18 in Freedman et al.

- **Week 4**

July 6 (Wednesday): Hypothesis testing

*Required reading:* Chapters 26 in Freedman et al.

July 7: Confidence intervals and comparing two populations.

*Required reading:* Chapters 21 and 27 in Freedman et al.

- **Week 5**

July 11: Bayesian inference and mechanics of simple linear regression. Problem Set #3 due in class.

*Required reading:* Chapters 10-12 in Freedman et al.

July 14: Trip to DC.

- **Week 6**

July 18: Assumptions of the classical linear regression model, inference, and omitted variable bias.

*Required reading:* TBA.

July 21: Mechanics of multiple regression; log, quadratic, categorical, dummy variables, and interactions.

*Required reading:* TBA.

- **Week 7**

July 25: Maximum likelihood estimation: Probit and logit models. Problem Set #4 due in class.

*Required reading:* TBA.

July 27: In-class final exam (Location: Robertson 011).