95-796, Statistics for IT Managers (Fall 2017)
90-777, Intermediate Statistics (Fall 2017)

Course Information

95-79 and 90-77 are 6 unit courses. There are five course sections (A,B,C,D,E). Classes begin the week of August 28th and end October 13th, with final exams October 16th and 17th.

Class Schedule

Section A: Mondays and Wednesdays, 3:00-4:20pm, Hamburg Hall 1002
Section B: Mondays and Wednesdays 4:30-5:50pm, Hamburg Hall 1002
Section C: Tuesdays and Thursdays, 10:30-11:50am, Hamburg Hall 1204
Section D: Tuesdays and Thursdays, 3:00-4:20pm, Hamburg Hall 1206
Section E: Tuesdays and Thursdays, 4:30-5:50pm, Hamburg Hall 1206
Review sessions (Sections A, B): Fridays, 9:00-10:20am, Hamburg Hall A301
Review sessions (Sections C,D, E): Fridays, 3:00-4:20pm, Hamburg Hall A301

Instructor

Daniel Nagin (dn03@andrew.cmu.edu)  *** Please follow our e-mail questions policy described below!
Office hours: Monday (11-12pm) and Wednesdays, 1:00-2:00pm. *** Feel free to stop by any of our office hours at the listed times. You do not need to make an appointment. ***
Hamburg Hall 2213, x8-8474

Teaching Assistants

Ciprian Domnisoru cdomniso@andrew.cmu.edu
Jennifer Hunter jlhunter@andrew.cmu.edu
Jacqueline Mauro <jmauro@andrew.cmu.edu>
Mallory Nobles <mnobles@andrew.cmu.edu>

TA Office Hours (No TA Hours Sunday, Sept. 3 or Monday, Sept. 4)

Sunday (3-5pm, HbH 1208)
Monday (6-8pm, HbH 1208)
Tuesday (12-2pm HbH 3034)
Wednesday (6-8pm, HbH 2009)

Course Description

This introductory course in data analysis and statistical inference requires no background in statistics. Its objective is to provide individuals who aspire to enter management or policy analysis positions with the basic statistical tools for analyzing and interpreting data. The course is divided into three distinct modules: descriptive statistics, statistical inference, and regression analysis. The emphasis of the classes on descriptive statistics is the calculation and interpretation of summary statistical measures for
describing raw data. The sessions on statistical inference are designed to provide you with the
background for executing and interpreting hypothesis tests and confidence intervals. The final
component of the course focuses on regression analysis, a widely used statistical methodology.
Throughout the course you will regularly analyze data relevant to management and policy analysis using
the statistical software package Minitab.

Course Objectives

Upon completion of this course, the student will be able to:

- Apply techniques for analyzing and interpreting data to real-world datasets relevant to
  management and policy analysis.
- Perform and interpret elementary statistical inferences (such as confidence intervals and
  hypothesis tests) both by hand and using the statistical software package Minitab.
- Analyze real data and describe the analysis results and conclusions.

Course Materials

*** You should be able to buy a used copy cheaply on Amazon, etc. It's fine to use the newer 12th
Edition, or the older 9th or 10th Edition instead, but the section numbers may be slightly different! ***

Also, lecture slides are posted here on Canvas, and Minitab software is available for your use.

Evaluation Method

Grades will be based entirely on six deliverables:
Four homework assignments: 40% (10% each)
One mini-project: 10%
Final exam: 50%

Due Dates

Reminder: all assignments are due at the beginning of your class period except where otherwise
noted. Assignments turned in more than five (5) minutes after class starts may be counted as "late" and
treated according to the late work policy below.

Homework 1 due Monday 9/11 (Sections A,B) or Tuesday 9/12 (Sections C,D,E).
Homework 2 due Monday 9/18 (Sections A,B) or Tuesday 10/19 (Sections C,D,E).
Homework 3 due Wednesday 9/27 (Sections A,B) or Thursday 9/28 (Sections C,D,E).
Homework 4 due Wednesday 10/4 (Sections A,B) or Thursday 10/5 (Sections C,D,E).
Mini-project due Friday 10/13 no later than 3:00pm (all sections). (Turn in at the beginning of Friday's
review session, or by e-mail.)

The mini-projects will be done in 2-person teams and will involve the analysis of real data relating
to software cost estimation. In your 2-3 page mini-project reports you will describe the results of your
analyses and your conclusions regarding the issues outlined in the assignment. Teams will be self-
selected, and both team members will receive the same grade.
Frequently asked question: can I work with someone from another section on the mini-project? **Answer:** YES, Please only submit a single copy of the project report.

Frequently asked question: do you curve our grades? **Answer:** Students’ numeric grades for deliverables, and their final numeric grades, will not be curved. However, the instructor will apply a curve when deciding on the final assignment of letter grades based on the numeric grades, in order to meet Heinz College standards for mean student GPA in core courses.

**Grading Scale**

As noted above, I do not pre-specify the grading scale. Below is how the cutoffs came out for a previous semester, but this year’s cutoffs will depend on the overall distribution of scores.

- A+ 96-100%
- B+ 80-83%
- C+ 68-71%
- A 88-95%
- B 76-79%
- C 64-67%
- A- 84-87%
- B- 72-75%
- C- 56-63%

**Cheating and Plagiarism Notice**

We encourage discussion among teams about the mini-projects and among individual students on homework assignments. However, the project hat is submitted for grading must be the work of the 2-person team alone. Similarly, completed homework assignments must be your work alone. Specifically, discussion of results that are identical or nearly identical across projects will be regarded as cheating. Also, your answers on the final exam must reflect your work alone. Sanctions for cheating include lowering your grade including failing the course. In egregious instances, the instructors may recommend the termination of your enrollment at CMU.

**Additional Course Policies**

**Late Work Policy:** You are expected to turn in all work on time (at the start of class on the due date). Because we understand that exceptional circumstances may arise, each student will be permitted to turn in one assignment (one homework or one mini-project) up to 48 hours late. Any other late assignments will not be accepted.

**Re-grade Policy:** Any requests for a re-grade must be submitted in writing to Professor Nagin or Head TA Ciprian Domnisoru within seven (7) days from when the graded assignment is returned. You must include a clear written explanation of why the regrade is necessary, stapled to your graded assignment. If a re-grade request is submitted, we may re-grade your entire assignment, which may either raise or lower your score. In general, we will only raise scores in cases where we have made an error in grading, and all decisions on re-grade requests are final.

**E-mail Questions Policy:** To balance the workload fairly among the teaching assistants, and to ensure a reasonable response time for questions received via e-mail, each student has been assigned one TA as a "first contact". This is the person that you should e-mail first with any questions that you might have regarding the course material, policies, etc. We will do our best to answer questions within 24 hours, or 48 hours on weekends. If you do not receive a response within this time, or if your first contact is unable to resolve your question, then you should feel free to e-mail the instructor or another TA. First contacts are assigned based on the first letter of your last name. For lasts names beginning with A-F e-mail Ciprian Domnisoru cdomniso@andrew.cmu.edu, for last names beginning with G-L e-mail Mallory
Nobles (mnobles@andrew.cmu.edu) for last name beginning with M-R e-mail For N-Q, e-mail Jacqueline Mauro jmauro@andrew.cmu.edu, and for last names beginning with S-Z email Jennifer Hunter jlhunter@andrew.cmu.edu.

Course Outline

Module I: Descriptive Statistics and Probability (Chapters 1-4)

Lecture 1: Descriptive statistics
(Sections A, B: Mon. 8/28; Sections C, D, E: Tue. 8/29)
Course overview
What is statistics (1.1-1.3)
Types of data (1.5)
Random sampling (1.6)
Histograms (2.2)
Measures of central tendency (2.4)
Measures of variability (2.5-2.6)
Box plots (2.8)
Using Minitab for descriptive statistics

Lecture 2: Probability
(Sections A, B: Wed. 8/30; Sections C, D, E: Thur. 8/31)
Basic probability (3.1-3.4)
Conditional probability (3.5-3.6)
Bayes' Theorem (3.8)
Discrete random variables (4.1-4.3)

FRIDAY SEPT. 1-LECTURE 2 CONTINUED + MINITAB TUTORIAL

NO CLASSES-MONDAY, 9/4 OR TUESDAY 9/5

Lecture 3: Random variables
(Sections A, B: Wed. 9/6; Sections C, D, E: Thur. 9/7)
Continuous random variables (4.5)
The Uniform distribution (4.9)
The Normal distribution (4.6)

Lecture 4: Normal distributions and sampling
(Sections A, B: Mon. 9/11; Sections C, D, E: Tue. 9/12)
The Normal distribution, continued (4.6)
Sampling distributions (4.10)
Central Limit Theorem (4.11)
Using Minitab for random variables and sampling

Deliverable: Homework 1 due at beginning of class
Module II: Hypothesis Testing and Inference (Chapters 5-7)

Lecture 5: Confidence intervals
(Sections A, B: Wed. 9/13; Sections C, D, E: Thur. 9/14)
Large-sample confidence intervals for the mean (5.2)
Small-sample confidence intervals for the mean (5.3)
Large-sample confidence intervals for the population proportion (5.4)
Determining the sample size (5.5)
Using Minitab for confidence intervals

Lecture 6: Hypothesis testing
(Sections A, B: Mon. 9/18; Sections C, D, E: Tue. 9/19)
Introduction to hypothesis testing (6.1-6.2)
Large-sample hypothesis tests for the mean (6.3)
Small-sample hypothesis tests for the mean (6.5)
Large-sample hypothesis tests for the population proportion (6.6)
Deliverable: Homework 2 due at beginning of class

Lecture 7: More hypothesis testing
(Sections A, B: Wed. 9/20; Sections C, D, E: Thur. 9/21)
Type I and Type II errors (6.1)
p-values (6.4)
Using Minitab for 1-sample hypothesis testing

Lecture 8: Comparing two populations
(Sections A, B: Mon. 9/25; Sections C, D, E: Tue. 9/26)
Comparing two population means (7.2)
Comparing two population proportions (7.4)
Using Minitab for 2-sample hypothesis testing

Lecture 9: Comparing two populations, continued
(Sections A, B: Wed. 9/27; Sections C, D, E: Thur. 9/28)
Comparing two population means using paired differences (7.3)
Using Minitab for paired differences
Deliverable: Homework 3 due at beginning of class

Module III: Regression (Chapters 10-11)

Lecture 10: Simple regression
(Sections A, B: Mon. 10/2; Sections C, D, E: Tue. 10/3)
Linear models (10.1)
Least squares linear regression (10.2)
Estimating the error of the model (10.3)
Making inferences using the model (10.4)
Deliverable: Homework 4 due at the beginning of class.

Lecture 11: More simple regression  
(Sections A, B: Wed. 10/4; Sections C, D, E: Thur. 10/5)  
Coefficients of correlation and determination (10.5)  
Using the model for estimation and prediction (10.6)  
Using Minitab for simple regression  
Deliverable: Homework 4 due at beginning of class

Lecture 12: Multiple regression  
(Sections A, B: Mon. 10/9; Sections C, D, E: Tue. 10/10)  
Multivariate linear models (11.1)  
Least squares linear regression (11.2)  
Estimating the error of the model (11.2)  
Multiple coefficient of determination (11.3)

Lecture 13: More multiple regression  
(Sections A, B: Wed. 10/11; Sections C, D, E: Thur. 10/12)  
Using the model for estimation and prediction (11.4)  
Advanced topics in model building (11.5-11.7)  
Using Minitab for multiple regression

Review (Fri. 10/13)  
Review for final exam.  
Deliverable: Mini-Project 1 due at the beginning of the review session.

Final Exam: (Sections A, B: Mon. 10/16; Sections C, D, E: Tue. 10/17)