95-796, Statistics for IT Managers (Fall 2014)

Course Information

95-796, Statistics for IT Managers, is a 6 unit course, taught in Mini-1. There are four course sections (A,B,C,D). Classes begin the week of August 25th and end the week of October 6th, with final exams during the week of October 13th.

Class Schedule

Section A: Tuesdays and Thursdays, 9:00-10:20am, Hamburg Hall 1000 Section B: Tuesdays and Thursdays, 10:30-11:50am, Hamburg Hall 1004 Section C: Tuesdays and Thursdays, 3:00-4:20pm, Hamburg Hall 1000 Section D: Mondays and Wednesdays, 3:00-4:20pm, Hamburg Hall 1502 Review sessions (all sections): Fridays, 3:00-4:20pm, Doherty Hall 2210

Instructor

Daniel Neill (neill@cs.cmu.edu) *** Please follow our e-mail questions policy described below! ***
Office hours: Tuesdays and Thursdays, 4:30-5:15pm (or by appointment) Hamburg Hall 2105B, x8-3885

Faculty assistant: Natalia Pascal (npascal@andrew.cmu.edu), Hamburg Hall 2112

Teaching Assistants

Head TA: Sriram Somanchi (somanchi@cmu.edu) Office hours: Wednesdays 4:30-5:50pm and Fridays 1:30-2:50pm, Hamburg Hall 3030.

Mitun Jotheswari Balusamy (mjothesw@andrew.cmu.edu) Office hours: Thursdays noon-1:20pm, Hamburg Hall A117.

Alejandra Caro (acarorin@andrew.cmu.edu) Office hours: Thursdays 6:00-7:20pm, Hamburg Hall student lounge.

Yang Lu (yanglu2@andrew.cmu.edu)
Office hours: Tuesdays 6:00-7:20pm, Hamburg Hall A121.

Prem Viswanathan (prempv@cmu.edu)
Office hours: Wednesdays 10:30-11:50am, Hamburg Hall A115.

Chris Won (junheew@andrew.cmu.edu)
Office hours: Mondays 1:30-2:50pm, Hamburg Hall A115.

Yi Zhang (yzh@cmu.edu)
Office hours: Mondays 4:30-5:50pm and Thursdays 1:30-2:50pm, Hamburg Hall 3043.
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 IT management 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 IT management using the statistical software package Minitab.

Course Materials

*** It's fine to use the new 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 Blackboard, and Minitab software is available for your use.

Evaluation Method

Grades will be based entirely on six deliverables:

Three homework assignments: 30% (10% each)  
Two mini-projects: 30% (15% each)  
Final exam: 40%

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.

Mini-project 1 due Wednesday 9/10 (Section D) or Thursday 9/11 (Sections A,B,C).  
Homework 1 due Wednesday 9/17 (Section D) or Thursday 9/18 (Sections A,B,C).  
Homework 2 due Monday 9/29 (Section D) or Tuesday 9/30 (Sections A,B,C).  
Homework 3 due Monday 10/6 (Section D) or Tuesday 10/7 (Sections A,B,C).  
Mini-project 2 due Friday 10/10 at 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 online pricing and 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, BUT: you must turn in the assignment at the EARLIER of the due dates, and you must include the name, Andrew ID, and section (A, B, C, or D) of all team members. 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.

**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 IT management.
• 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 relating to online pricing and software cost estimation and describe the analysis results and conclusions.

The first two objectives will be assessed through all six deliverables (homeworks 1-3, mini-projects 1-2, and final exam). The third objective will be assessed through the two mini-projects.

**Grading Scale**

As noted above, I do not pre-specify the grading scale. Below is how the cutoffs came out for Fall 2012, 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 and that 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 the course instructor 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. 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 our head TA or the instructor. First contacts are assigned based on the first letter of your last name. For A-B, e-mail Chris (junheew@andrew.cmu.edu). For C, e-mail Alejandra (acarorin@andrew.cmu.edu). For D-G, e-mail Prem (prempv@cmu.edu). For H-L, e-mail Yi (yzh@cmu.edu). For M-R, e-mail Sriram (somanchi@cmu.edu). For S, e-mail Yiming (yimingzh@andrew.cmu.edu). For T-V, e-mail Lu (yanglu2@andrew.cmu.edu). For W-X, e-mail Mitun (mjothesw@andrew.cmu.edu). For Y-Z, e-mail Yiming (yimingzh@andrew.cmu.edu).

Course Outline

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

Lecture 1: Descriptive statistics
(Sections A,B,C: Tues. 8/26; Section D: Mon. 8/25) 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,C: Thurs. 8/28; Section D: Wed. 8/27) Basic probability (3.1-3.4) Conditional probability (3.5-3.6) Bayes' Theorem (3.8) Discrete random variables (4.1-4.3)

Review 1 (Yi)
(All sections: Fri. 8/29)
Minitab and example problems.

Monday 9/1 and Tuesday 9/2: NO CLASSES!

Lecture 3: Random variables
(Sections A,B,C: Thurs. 9/4; Section D: Wed. 9/3) Continuous random variables (4.5) The Uniform distribution (4.9) The Normal distribution (4.6)

Review 2 (Sriram)
(All sections: Fri. 9/5)
Minitab, example problems, and questions on mini-project 1.
Lecture 4: Normal distributions and sampling (Sections A,B,C: Tues. 9/9; Section D: Mon. 9/8) The Normal distribution, continued (4.6) Sampling distributions (4.10) Central Limit Theorem (4.11) Using Minitab for random variables and sampling

Module II: Hypothesis Testing and Inference (Chapters 5-7)

Lecture 5: Confidence intervals
(Sections A,B,C: Thurs. 9/11; Section D: Wed. 9/10) 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
Deliverable: Mini-Project 1 due at the beginning of class.

Review 3 (Alejandra)
(All sections: Fri. 9/12)
Minitab, example problems, and questions on HW 1.

Lecture 6: Hypothesis testing
(Sections A,B,C: Tues. 9/16; Section D: Mon. 9/15) 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)

Lecture 7: More hypothesis testing
(Sections A,B,C: Thurs. 9/18; Section D: Wed. 9/17) Type I and Type II errors (6.1) p-values (6.4) Using Minitab for 1-sample hypothesis testing
Deliverable: HW1 due at the beginning of class.

Review 4 (Sriram)
(All sections: Fri. 9/19)
Minitab, example problems, and questions on HW 2.

Lecture 8: Comparing two populations
(Sections A,B,C: Tues. 9/23; Section D: Mon. 9/22) 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,C: Thurs. 9/25; Section D: Wed. 9/24) Comparing two population means using paired differences (7.3) Using Minitab for paired differences

Review 5 (Prem)
(All sections: Fri. 9/26)
Minitab, example problems, and questions on HW 2.

Module III: Regression (Chapters 10-11)

Lecture 10: Simple regression
(Sections A,B,C: Tues. 9/30, Section D: Mon. 9/29) 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: HW2 due at the beginning of class.
Lecture 11: More simple regression
(Sections A,B,C: Thurs. 10/2; Section D: Wed. 10/1) Coefficients of correlation and determination (10.5)
Using the model for estimation and prediction (10.6) Using Minitab for simple regression

Review 6 (Mitun)
(All sections: Fri. 10/3)
Minitab, example problems, and questions on HW 3 and Mini-project 2.

Lecture 12: Multiple regression
(Sections A,B,C: Tues. 10/7; Section D: Mon. 10/6) 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) Making inferences using the model (11.3) Deliverable: HW3 due at the beginning of class.

Lecture 13: More multiple regression
(Sections A,B,C: Thurs. 10/9; Section D: Wed. 10/8) Using the model for estimation and prediction (11.4) Advanced topics in model building (11.5-11.7) Using Minitab for multiple regression

Review 7 (Sriram)
(All sections: Fri. 10/10)
Review for final exam.
Deliverable: Mini-Project 2 due at the beginning of the review session.

Final exam: date, time, and rooms TBD. It is likely that these will be held at your usual time and place on Thursday 10/16 (Sections A,B,C) and Wednesday 10/15 (Section D), but I will need to confirm.