ISTA 116: Statistical Foundations for the Information Age

Dr. Colin Reimer Dawson

Last Revised August 28, 2011

1 Course Information

1.1 Course Description

Understanding uncertainty and variation in modern data: data summarization and description, rules of counting and basic probability, data visualization, graphical data summaries, working with large data sets, prediction of stochastic outputs from quantitative inputs. Operations with statistical computer packages such as R.

1.2 Prerequisites

MATH 109C or MATH 112 or placement beyond the level of College Algebra.

1.3 Course Goals

This course is intended as a first exposure to the nature of uncertainty and variation in data typical of the information age. Students will learn techniques for descriptive statistics, both quantitative and graphical, and will make use of these with real (often large) data sets in lab sessions and in homework. Several weeks are devoted to introducing basic concepts in probability, as a foundation for later study in inferential statistics. The statistical computer package R is used for illustration and exploration, as well as to demonstrate the central role of computing in modern data science.

1.4 Locations and Times

Lectures: MW 1:00-1:50
McClelland Park, Room 105

Lab: One 2 hr section per week

1.5 Required Texts

Verzani, J. (2005). *Using R for Introductory Statistics*. Boca Raton, FL: Chapman & Hall/CRC Press.

Owen, W. J. (2010). *The R Guide, ver. 2.5*. Department of Mathematics and Computer Science, University of Richmond, Richmond, VA. `http://cran.r-project.org/doc/contrib/Owen-TheRGuide.pdf` (available for free electronically)

1.6 Instructor Information

Dr. Colin Reimer Dawson
Office: Gould-Simpson 850
Email: `cdawson@email.arizona.edu`
Office Hours: W 2-4 and by appointment

1.7 Lab Instructors

Jeff Berry
Office: Douglass 304
`jjberry@email.arizona.edu`
Office Hours: Th 2:30-3:30 and F 11-1

Derek Green
Office: Gould-Simpson 830
`dtgreen@cs.arizona.edu`
Office Hours: M 10-12 and 2-3

Philip Sparks
Office: Gould-Simpson 228
`pspark@email.arizona.edu`
Office Hours: M 3-4 and Th 12-1

Anu Venkatesh
Office: Gould-Simpson 228
`anuvenk@email.arizona.edu`
Office Hours: Th 10-12

2 Policies and Grading

2.1 Grading

Grades are based on homework, quizzes, one midterm, and a comprehensive final exam. The grading scheme is as follows:

Component	Weight
Lab Assignments	400 pts
In-Class Quizzes	100 pts
Web Quizzes	50 pts
Midterm Exam	200 pts
Final Exam	200 pts
Exam Adjustment	50 pts
Total	1000 pts

Grade	Point Range
A	900-1000
B	800-899
C	700-799
D	650-699
E	0-649

Grade Disputes

Disputes about grades on a particular assignment or exam will be entertained for two weeks from the day the assignment is returned, or one week from the date of the final exam, whichever is sooner. These will be resolved by re-grading the entire assignment or exam. Note that this can result in a lower grade in the event that new mistakes are discovered. The final exam will be graded and made available for review by students within 48 hours of its completion, to allow time for any requested regrades.

"Curve" Policy

The instructor may decide at the end of the term to lower the thresholds stated above for all students. Any decision to do so will be based solely on aggregate class performance, and not on circumstances pertaining to any individual. Once any grade disputes have been resolved for individual assignments, and the final thresholds have been set for all students, the total number of points earned will automatically determine the final grade. **No negotiations about individual students' letter**

grades will be entertained once final grades are assigned, except as permitted by the policy stated above.

2.2 Homework Assignments

There will be 8 homework assignments, consisting of a mixture of paper-and-pencil and computational problems. Individual assignments are worth 50 points. Assignments are graded for correctness; however, genuine effort is nearly always worth partial credit.

Collaboration Policy

Students may discuss homework assignments with the instructor, TAs and each other, however, **each student must submit her or his own solutions**. Any fellow student collaborators must be identified along with the nature of the collaboration (e.g., "we worked together", "she helped me", "I helped him"), and any asymmetric collaboration should be limited to hints or explanation of concepts, and not detailed walkthroughs of particular solutions.

Late Policy

Assignments are due at the end of the first lab section after the announced due date, unless otherwise specified. Permission for an extension must be granted by the lab instructor in advance of the deadline in order to receive full credit for a late assignment. The lab instructors are under no obligation to grant such an extension.

Calculation of Final Homework Grade

Each student's lowest homework grade will be replaced by her/his highest homework grade, to produce a final score out of 400.

2.3 Quizzes

In-Class Quizzes

Several unnanounced quizzes will be given at the beginning of lectures throughout the semester, each consisting of one or two short and straightforward questions designed to demonstrate that students are paying attention in class and doing the required reading. Students who do so should find the quizzes easy.

The five highest quiz scores will be used to calculate the final quiz grade. **No make-up quizzes will be given for any reason**.

Web Quizzes

Web quizzes will occasionally be assigned via d2l. These are intended to provide practice with problems more like those that will appear on exams (and hence can be done without the computer). Students are encouraged to make at least their first attempt without any reference materials. These are graded for completion only.

2.4 Exams

One in-class exam will be given during the lecture on October 12$^{\text{th}}$. A comprehensive final exam will be given on December 12$^{\text{th}}$ from 10:30 A.M. to 12:30 P.M. The format of the exams will be mixed, and may include multiple choice, short verbal answer, and "math problems". Each exam is graded out of 200 pts. Each student's higher exam grade will be rescaled to 250 points.

Permitted Materials

Students will be allowed to use one double-sided $8\frac{1}{2} \times 11$" page of notes for the midterm, and two pages for the final. A simple, self-contained hand calculator may also be used. **No other electronic devices of any kind, including cell phones, iPods, laptops, etc., will be permitted during the exams**.

Make-up Policy

Exams may only be made up in case of a serious, unanticipated emergency. The student or an appointed proxy must contact the instructor in advance of the exam. Clearly worded documentation, from a doctor, dean, etc., demonstrating physical inability to take the exam at the appointed time, will be required to grant a make-up appointment. **No make-ups will be granted for personal reasons, including travel or personal hardship**. Any make-up exams that are granted must be scheduled for after the original time.

2.5   University Policies

Classroom Behavior

Students are expected to behave respectfully toward each other and to the instructor and TAs. Disrespectful behavior includes the use of cell phones or other electronic devices in the classroom during class hours.

The Arizona Board of Regents Student Code of Conduct is here: `http://dos.web.arizona.edu/uapolicies/scc5308abcd.html#sccphilosophy`

ABOR Policy 5-308, prohibits threats of physical harm to any member of the University community, including to oneself. See: `http://policy.web.arizona.edu/~policy/threaten.shtml`.

Special Needs and Accommodations

Students who need special accommodation or services should contact the

Disability Resources Center
1224 East Lowell Street, Tucson, AZ 85721
(520) 621-3268
FAX (520) 621-9423
email: `uadrc@email.arizona.edu`
web: `http://drc.arizona.edu/`.

You must register and request that the Center or DRC send official notification of your accommodations needs as soon as possible. Please plan to meet with the

instructor by appointment or during office hours to discuss accommodations and how the course requirements and activities may impact your ability to fully participate. The need for accommodations must be documented by the appropriate office.

Student Code of Academic Integrity

Students are encouraged to share intellectual views and discuss freely the principles and applications of course materials. However, graded work/exercises must be the product of independent effort unless otherwise instructed. Students are expected to adhere to the UA Code of Academic Integrity as described in the UA General Catalog. See: `http://dos.web.arizona.edu/uapolicies/`.

Confidentiality of Student Records

See `http://www.registrar.arizona.edu/ferpa/default.htm`

Subject to Change Statement

Information contained in this syllabus, other than the grade and absence policy, may be subject to change by the instructor, with advance notice.

3 Schedule of Lecture Topics and Readings

Topic	Approx. Dates	Readings	Add'l Lab Readings
Introduction	Aug. 22	Course Syllabus Ch. 1.1	Ch. 1.2-1.4 Owen, 1.1-1.8, 2.1
Univariate Categorical Data	Aug. 24	Ch. 2.1	Owen 3.3, 4.1
Univariate Numeric Data	Aug. 29, 31 Sept. 7	Ch. 2.2	Owen 5.1, 5.2
Characterizing Distributions	Sept. 12	Ch. 2.3	
Multivariate Categorical Data	Sept. 14, 19	Ch. 3.1, 4.1.1	Ch. 4.2
Grouped Numeric Data	Sept. 21	Ch. 3.2, 4.1.2	
Bivariate Numeric Data	Sept. 26, 28	Ch. 3.3, 4.1.3	
Simple Linear Regression	Oct. 3, 5	Ch. 3.4 to 3.4.3	
Catch-up and Review	Oct. 10	—	
MIDTERM EXAM	***Oct. 12***	—	
Intro to Probability	Oct. 17, 19	Notes	
Discrete Random Variables	Oct. 24, 26	Ch. 5.1.1	
The Bernoulli, Binomial and Discrete Uniform Distributions	Oct. 31	Ch. 5.2.2	Ch. 5.2.1 Owen 6.1, 6.3.1
Continuous Random Variables	Nov. 2, 7	Ch. 5.1.2	
The Normal Distribution	Nov. 9	Ch. 5.2.2	Owen 6.3.2
Random Samples	Nov. 14	Ch. 5.1.3	Owen 6.4
Sampling Distributions	Nov. 16, 21	Ch. 5.1.4	
The Central Limit Theorem	Nov. 23, 28	Ch. 5.3	
Simulation and Bootstrapping to Estimate Sampling Distributions	Nov. 30, Dec. 5	Notes	Ch. 6.1-6.3, 6.6
Catch-up and Review	Dec. 7	—	
FINAL EXAM	***Dec. 12***	—	

4 Schedule of Lab Topics

Dates	Topics	Assignment Due
Aug. 23-24	Intro and R Fundamentals	—
Aug. 30-31	R Fundamentals Cont'd, Tables, Plotting Bar Charts, Pie Charts	1
Sept. 6-7	Summary Statistics, Stem and Leaf Plots, Strip Charts, Histograms, Density Curves	—
Sept. 13-14	Visualizing Central Tendency and Variability, Transformations, Boxplots	2
Sept. 20-21	Contingency Tables, Margin Tables, Conditional Proportion Tables	—
Sept. 27-28	Overlaid Densities, Multiple Boxplots, Summary Statistics by Group, Scatterplots, Correlation	3*
Oct. 4-5	Linear Regression with lm() Prediction, Residual Plots	4*
Oct. 11-12	Midterm Review	—
Oct. 18-19	Probability Basics (no R this week)	—
Oct. 25-26	Sampling from a Population With and Without Replacement using sample()	5
Nov. 1-2	Computing Binomial Probabilities and Quantiles in R	—
Nov. 8-9	Normal Probabilities and Quantiles	6*
Nov. 15-16	Sampling from a Distribution	—
Nov. 22-23	Repeated Sampling, Visualizing Distributions of Sample Statistics	7
Nov. 29-30	Approximating Sampling Distributions with the CLT and with Simulation	—
Dec. 6-7	Bootstrapping, Catch Up	8

*Indicates a lab assignment with a Friday due date. These assignments are *not* due in lab, but are due in the d2l dropbox by 5:00 P.M. on the indicated day.

5   Lab Assignments and Web Quizzes

Number	Topics	Due Date
WQ 1	Course Policies	Aug. 26
Lab 1	R Fundamentals	Aug. 30-31
WQ 2	Central Tendency and Variability	Sept. 9
Lab 2	Univariate Data	Sept. 13-14
WQ 3	Contingency Tables	Sept. 23
WQ 4	Comparing Distributions	Sept. 28
Lab 3	Multivariate Data	Sept. 30*
Lab 4	Simple Linear Regression	Oct. 7*
WQ 5	Correlation and Regression	Oct. 10
Lab 5	Probability Intro	Oct. 25-26
WQ 6	Discrete Distributions	Nov. 4
Lab 6	Common Distributions	Nov. 11*
WQ 7	Sampling Schemes	Nov. 18
Lab 7	Sampling and Sampling Distributions	Nov. 22-23
WQ 8	Using the Central Limit Theorem	Dec. 2
Lab 8	Calculating Sample Probabilities	Dec. 6-7

*Indicates a lab assignment with a Friday due date. These assignments are *not* due in lab, but are due in the d2l dropbox by 5:00 P.M. on the indicated day.