

Psych. 501 INTRODUCTORY STATISTICS

Instructor: R. Wilcox

Office Hours: TTh 8-9

Phone: 213-740-2258

e-mail: rwilcox@usc.edu

Text: Wilcox, R. R. (2011) *Modern Statistics for the Social and Behavioral Sciences: A Practical Introduction*. New York: Chapman & Hall/CRC press.

The primary goal is to provide a strong foundation for understanding and applying basic statistical techniques. A detailed outline of the material is given below. The basics include sampling distributions, expected values, hypothesis testing, the difference between parameters and statistics, least squares regression and correlation, and the basics of probability theory. No prior training in statistics is assumed. Classic methods are covered as well as some modern methods developed during the last half century that are aimed at dealing with skewed distributions, heavy-tailed distributions and heteroscedasticity.

This course covers chapters 1-9 in the text. Time permitting, portions of chapters 10 and 11 (ANOVA) will be covered.

There are three exams: two midterms and a final. Each test accounts for 30% of your grade. The computer lab is mandatory and counts for 10% of your grade. There is a possible 100 points for the course. A grade of A requires 92 or higher, 90-91 is A-, 88-89 is B+, 82-87 is a B, 80-81 a B-, 78-79 is a C+, etc.

Consequently, if you don't do the computer assignments, you can get an A- only if you get perfect scores on all of the exams.

A quiz is given after each chapter is covered in class. It is graded pass/fail. If you fail, you can hand in a take-home make-up quiz and get a pass. The quizzes can raise your grade but they can't lower it. You get a bonus point added to each exam if you do all of them prior to exam being given.

The first test will be on chapters 1-3 and will be given during the fifth week of the semester. So if you complete the quizzes prior to week 5, you get a bonus point on the exam. The second exam is on chapters 4-6 and will be given during the tenth week, and the final will cover chapters 7-9.

THE FINAL IS GIVEN ON THE DAY INDICATED BY THE SCHEDULE OF CLASSES. THERE ARE NO EXCEPTIONS.

Weekly layout:

week 1: Chapter 1 + Sections 2.1-2.4

week 2: Sections 2.5-2.10

week 3: Quiz, ch. 2, Sections 3.1-3.6 week 4: Sections 3.7-3.8, Quiz Ch 3.

week 5: Section 4.1-4.2, exam

week 6: Sections 4.3-4.9

week 7: Sections 4.10-4.13, 5.1-5.6 quiz ch 4

week 8: Sections 6.1-6.3, quiz ch 5

week 9: Sections 6.4-6.6, quiz ch 6

week 10: 7.1-7.2, exam

week 11: sections 7.3-7.6

week 12: 7.7-7.8, 8.1-8.3, quiz ch 7

week 13: 8.4-8.9

week 14: 8.10-8.13

week 15: 9.1-9.2, 10.1 and 12.1, review.

A MORE DETAILED OUTLINE OF THE MATERIAL:

1 INTRODUCTION

BROAD OVERVIEW

Samples versus Populations

Software

R Basics (The book illustrates how to apply standard methods)

2 NUMERICAL AND GRAPHICAL SUMMARIES OF DATA

Basic Notation

Measures of Location

The Sample Mean: Issues related to skewness and outliers

The Sample Median

Criticisms of the Median and Two General Strategies for Dealing with these Concerns

Measures of Variation or Scale

Sample Variance and Standard Deviation

The Interquartile Range

Robust measures of variation that have practical value.

Detecting Outliers

A method based on the mean and variance and why it is generally considered to be unsatisfactory.

Better Outlier Detection Rules (boxplot and MAD-median rules)

Measures of Location based on the strategy of removing outliers

Histograms and modern improvements

Stem-and-Leaf Displays

Skewness

Transforming Data

By modern standards, an ineffective method for dealing with skewness and outliers

Choosing a Measure of Location

3 PROBABILITY AND RELATED CONCEPTS

Basic Probability

Expected Values

Conditional Probability and Independence

Population Variance

The Binomial Probability Function

Continuous Variables and the Normal Curve

Computing Probabilities Associated with Normal Curves

Understanding the Effects of Non-normality:

Skewness and Heavy-Tailed Distributions

Pearson's Correlation and the Population Covariance

Computing the Population Covariance and Pearson's Correlation

Some Rules About Expected Values

Chi-Squared Distribution

4 SAMPLING DISTRIBUTIONS AND CONFIDENCE INTERVALS

Random Sampling

Sampling Distributions

Sampling Distribution of the Sample Mean

Computing Probabilities Associated with the Sample Mean

A Confidence Interval for the Population Mean Known Variance

Confidence Interval for the Population Mean Using Student's t

Relative Merits of Various Location Estimators

The Central Limit Theorem

Student's t , Non-normality and Common Misconceptions

Dealing with Skewness and Heavy-tailed Distributions

Transforming Data: An ineffective method for dealing with outliers

Confidence Interval for the Population Median

Estimating the Standard Error of the Sample Median

Concerns About Tied Values
Comments on Location Estimators that eliminate outliers
Confidence Intervals for the Probability of Success

5 HYPOTHESIS TESTING

The Basics of Hypothesis Testing
P-Value or Significance Level
Criticisms of Two-Sided Hypothesis Testing and P-Values
Tukey's strategy for dealing with this.
Power and Type II Errors
Testing Hypotheses about the Mean When the Variance Is Not Known
Controlling Power and Determining the Sample Size
Choosing the Sample Size Prior to Collecting Data
Judging the Sample Size, In Terms of Power, When Data Are Available
Practical Problems with Student's T Test and How They Might Be Addressed
Testing Hypotheses About the Population Median
Practical Reasons Why the Median Might Trim Too Many Observations and
What Might Be Done

6 REGRESSION AND CORRELATION

The Least Squares Principle
Confidence Intervals and Hypothesis Testing
Classic Inferential Techniques
Multiple Regression (Brief summary of basics)
Standardized Regression
Practical Concerns About Least Squares Regression and How
They Might Be Addressed
The Effect of Outliers on Least Squares Regression
Beware of Bad Leverage Points
Beware of Discarding Outliers Among the Y Values
(A Technically Sound Method is Covered in Ch 7)
Do Not Assume Homoscedasticity or that the Regression Line is Straight
Violating Assumptions When Testing Hypotheses
Dealing with Heteroscedasticity
Pearson's Correlation and the Coefficient of Determination
A Closer Look at Interpreting r
Testing the Hypothesis of a Zero Correlation
Dealing with Heteroscedasticity
When Is It Safe to Conclude that Two Variables Are Independent?
A Regression Method for Estimating the Median of Y Given X
Detecting Heteroscedasticity

7 BOOTSTRAP METHODS

Why They Are Important

Examples: Can remove outliers and test hypotheses in a technically sound manner.
Removing outliers and applying a standard method to the remaining data is disastrous.
Can handle tied values when comparing medians. Helps deal with heteroscedasticity.

The Percentile Bootstrap Method

Inferences About Measures of Location When Outliers Are
Empirically Determined and Eliminated

Bootstrap-t Method

Estimating Power When Testing Hypotheses

A Bootstrap Estimate of Standard Errors

Inferences about Pearson's Correlation: Dealing with Heteroscedasticity

Bootstrap Methods for Least Squares Regression

Detecting Associations Even When There Is Curvature

Quantile Regression

A Test for Homoscedasticity Using a Quantile Regression Approach

Regression: Which Predictors are Best?

Comparing Correlations

8 COMPARING TWO GROUPS

Student's T Test

Power Analysis and Sample Sizes

Relative Merits of Student's T

Welch's Heteroscedastic Method for Means

Why Testing Assumptions, to Justify a Standard Method, Generally Performs Poorly.

Criticisms of Methods that Test Hypotheses About Exact Equality.

Tukey's Strategy

Non-normality and Welch's Method

Three Modern Insights Regarding Methods for Comparing Means

Methods for Comparing Medians and Trimmed Means

Percentile Bootstrap Methods for Comparing Measures of
Location When Outliers Are Removed

Comparing Medians When There Are Tied Values

Some Guidelines on When To Use the Percentile Bootstrap Method

Bootstrap-t Methods for Comparing Means

Bootstrap-t Method When Comparing Trimmed Means

Estimating Power and Judging the Sample Sizes

Permutation Tests

Rank-Based and Nonparametric Methods

Wilcoxon-Mann-Whitney Test

Handling Tied Values and Heteroscedasticity

The Kolmogorov-Smirnov Test

Comparing All Quantiles Simultaneously
Graphical Methods for Comparing Groups
Error Bars
Plotting the Shift Function
Plotting the Distributions
Methods for Comparing Measures of Variation
Comparing Robust Measures of Variation
Measuring Effect Size
Comparing Correlations and Regression Slopes
Comparing Two Binomials
Making Decisions About Which Method To Use

9 COMPARING TWO DEPENDENT GROUPS

The Paired T Test

When Does the Paired T Test Perform Well?

Understanding When and Why It Can Perform Poorly.

Comparing Robust Measures of Location

Handling Missing Values

The Sign Test

Wilcoxon Signed Rank Test

Comparing Variances

Comparing Robust Measures of Scale

Comparing Quantiles

Plots for Dependent Groups

10 and 12. Basic ANOVA.

Statement on Academic Integrity:

USC seeks to maintain an optimal learning environment. General principles of academic honesty include the concept of respect for the intellectual property of others, the expectation that individual work will be submitted unless otherwise allowed by an instructor, and the obligations both to protect one's own academic work from misuse by others as well as to avoid using another's work as one's own. All students are expected to understand and abide by these principles. Scampus, the Student Guidebook, contains the Student Conduct Code in Section 11.00, while the recommended sanctions are located in Appendix A: <http://www.usc.edu/dept/publications/SCAMPUS/gov/>. Students will be referred to the Office of Student Judicial Affairs and Community Standards for further review, should there be any suspicion of academic dishonesty. The Review process can be found at: <http://www.usc.edu/student-affairs/SJACS/>.

Students who need to request accommodations based on a disability are required to register each semester with the Disability Services and Programs. In addition, a letter of verification to the

instructor from the Disability Services and Programs is needed for the semester you are enrolled in this course. If you have any questions concerning this procedure, please contact the course instructor and Disability Services and Programs at (213) 740-0776, STU 301.