

Psychology 310P  
 Statistical Inference  
 Course Information Sheet, Fall, 2009

Instructor:	James H. Steiger, Professor Department of Psychology and Human Development 215A Hobbs Phone: 615-322-7060 Office Hours: F 10:10-11:00; MW (By Appointment) Email: james.h.steiger@vanderbilt.edu
Teaching Assistant:	Amy Weitlauf Email: <a href="mailto:amy.s.weitlauf@vanderbilt.edu">amy.s.weitlauf@vanderbilt.edu</a> Office: Hobbs 315B Office Hours: Thursday 12–2
Course Meeting Times:	MW 11:00-12:15 F 11:00-11:50 One of these meetings, on Friday, frequently will involve lab exercises in statistical computing. Generally this period will end enough before noon to allow students to attend the “Developmental Lunch Bunch” or other noontime seminar if they wish to do so.
Instructor Website:	<a href="http://www.statpower.net">http://www.statpower.net</a>
Textbook:	Glass & Hopkins, <i>Statistical Methods in Education and Psychology</i> , 3 <sup>rd</sup> edition ISBN: 0205142125  As part of an initiative I call “Operation Greenbook,” this text was not ordered for the bookstore this fall. Used copies are available from a variety of sources, including the students who took the course last year. New copies are available over the internet, and you may be able to “rent” a copy from another graduate student. The textbook is now very expensive, and not really worth it. Despite its considerable strengths, it is out of date in several respects.
Final Exam:	There are two scheduled administrations of the final exam. The primary exam is scheduled for December 17 at 9am. The “alternate exam” is scheduled for December 12 at 12 noon. From the university’s perspective, the alternate exam need not be given, but my policy is to give an alternate exam if the date is earlier than the primary exam. (This gives students the option of leaving earlier for the holiday.) <b>Do not schedule a plane flight that precludes your taking the exam during one of these periods.</b>
Grading:	Course grades will be scaled. Midterm Exam 30% Final Exam - 30% Lab Assignments, homework, class participation - 40%.
Missing Class:	Missing class for a legitimate excuse (such as illness, religious holiday) does <b><i>not relieve you of the responsibility for turning in homework assignments on time!</i></b> For example, if you miss a tutorial class for a religious holiday you must still email or deliver your homework assignment to the TA on or before the due date. This policy does not mean dispensations are not granted for legitimate reasons. However the onus is on you: If a personal situation impedes your academic progress, please inform me and I will make arrangements to accomodate the situation.

# *Outline of Statistical Topics*

## **Topics for review and diagnostic test (covers preparatory handouts and Glass and Hopkins, Chapters 1–6)**

The laws of linear transformation

Subscript and summation notation

Shape -- Tabulating and depicting data.

Frequency distributions and their properties

Cumulative Frequency

Relative Frequency

Cumulative Relative Frequency

“Axioms of Relative Frequency”

Skewness

Kurtosis

Percentiles

Central Tendency

The Mean

Estimating the mean from a grouped frequency distribution

Computing the mean for pooled groups

The Median

Variability

The Range

The Semi-Interquartile Range

The Variance and Standard Deviation

Computational formulas and their derivation

The normal distribution.

The standard score

Interval probability estimates with the normal curve

The bivariate normal distribution

Measures of relationship.

### **The linear transform**

Definition

Effect on measures of central tendency

Effect on measures of variability

### **Covariance and Correlation**

Effect of linear transforms on the covariance.

The product moment correlation coefficient.

Definition

Effect of linear transforms.

Computational formulas.

### **Linear regression**

General Algebraic Setup for Regression

Linear regression

Computational formulas

The “least squares” criterion

Solving for the optimal least squares regression coefficients

Using linear regression

Description

## Prediction

### Correlational Fallacies

- Inferring causality
- Confusing Perfect Correlation with Equivalence
- Problems with mixed distributions
- Problems with skewed marginals
- Restriction of range
- Shrinkage

## The laws of linear combination

- Mean of a linear combination
- Variance and covariance of linear combinations of variates
- A simple heuristic algorithm and its derivation

### The algebra of variances and covariances

- Deriving some classic results
- Orthogonality in Regression
- The “Rich get Richer” phenomenon

### Some Relationships between Geometry and Statistics

- Vectorial representation of numerical lists
- The geometry of variance and correlation
- Seeing a classic result through geometry

## Multiple Linear Regression – A Brief Introduction

- The fixed regressors multiple regression model
- Regression Effects
- Interactions
- Centering

## Conditional Distribution Theory

### Some basic definitions

- Bivariate distributions
- Marginal distributions
- Conditional distribution

### Bivariate normal distribution

#### Some simple consequences

- Regression toward the mean
- Conditional distribution calculations

## Inferential Statistics.

### Probability theory

#### Set theory

- Basic Definitions
- Set
- Subset
- Proper subset
- Null set
- Universal set

## Operations

- Union

- Intersection

- Complement

## Key Definitions

- Mutually Exclusive Sets

- Exhaustive Sets

- Set Partition

## Basic Theorems in Set Theory

Probabilistic experiments and sample spaces

Simple (elementary) and compound events

Probability axioms

Three basic theorems of probability theory

Conditional probability and independence

- A classic paradox

- Bayes's Theorem

Counting rules and combinatorics

- The General Path Rule

- Permutations

- Permutations with selection

- Combinations

Random variables and probability distributions

The binomial distribution

Expected value and variance of a random variable

- Effect of a linear transform.

Covariance and correlation for random variables.

- Linear combination rules revisited

The normal approximation to the binomial

## **Basic theoretical aspects of statistical estimation**

- Unbiasedness

- Efficiency

- Consistency

## **Standard confidence interval estimation logic**

- Inverting probability statements

- The critical distance approach – “Taking a stroll with Mr. Mu”

## **Hypothesis testing logic**

- Partitioning the parameter space

- Partitioning the sample space

- Test statistics

- Power and error rates

- Factors influencing power

## **Tests on means -- a general approach**

When population variances are known

- 1-sample, 2-sample, and k-sample tests on independent samples

- Dependent sample procedures

- Extensions for general asymptotic Z-statistics

- The central limit theorem

When population variances are unknown, independent samples

- Student's  $t$ -distribution

- Assumptions for the  $t$ -test

- Robustness

- 1-Sample  $t$

- 2-Sample  $t$

## *k*-Sample *t*

When population variances are unknown, dependent samples

Repeated measures

Matched samples

2-Sample, matched sample *t*- test

Generalized Matched Sample *t*

Assumptions

## **Confidence intervals on linear combinations of means**

Independent Samples

Dependent Samples

## **Additional Theoretical distributions for use in statistical inference**

The Chi-square distribution.

The F distribution.

Noncentral Chi-Square and F distributions

## **Tests on variances**

1-sample

2-sample independent

## **Tests on proportions**

1 and 2-sample independent and dependent procedures.

## **Tests on correlations**

Classical asymptotic theory -- The Pearson-Filon equations.

The Fisher (and related) normalizing transforms.

Independent and dependent sample tests on 1 or two correlations.

Generalized procedures for comparing elements of a correlation matrix.

Robustness problems in correlational significance tests.

## **The analysis of variance – 1-Factor Fixed and Random Effects.**

The models and their underlying statistical assumptions.

The meaning of "effect" in the analysis of variance.

Computational procedures.

Robustness properties of ANOVA.

Multiple comparison procedures.

Error rates (experimentwise, per experiment, per comparison) and the problems of multiple testing.

Planned orthogonal contrasts.

The Scheffe test.

The Tukey test.

Noncentrality Interval Estimation – Confidence Intervals on Effect Size

ANOVA as multiple regression