

**Directorate of Assessments  
HOWARD UNIVERSITY COLLEGE OF DENTISTRY**

**RESEARCH SEMINAR\STATISTICS: AN INTRODUCTORY  
AND REVIEW COURSE**

**11714-HISP-327**

**SPRING SEMESTER 2012**

**GARNETT LEE HENLEY, MS CHS-III  
Assistant Professor and Director**

**Ms. Justine Keitt Teaching Assistant (TA)**

# Course Profile

## I. COURSE DESCRIPTION

Department	Directorate of Assessments and the Department of Diagnostic Services
Course Title	Pre-Doctoral Statistics Review and Research Seminar
Course Number	11714-HISP-327
Course Hours	3 Credit Hours
Class Location	Lecture Hall 2, Dixon Building
Semester Offered	Spring
Date & Time Offered	Friday Mornings, 0800 – 1000 Hours
Curricular Level	Pre-Doctoral
Instructor	Garnett Lee Henley, MS CHS-III
Office Location	Room 5D.7B (Dean's suite complex)
Office Telephone	202-806-0317 and <a href="mailto:ghenley@howard.edu">ghenley@howard.edu</a> . Ms. Keitt: 202-806-0239
Office Hours	Tuesdays 1200 – 1400 hours. and Thursdays 1200 – 1400 hours

## II. COURSE ORGANIZATION:

### COURSE OVERVIEW AND GOALS

This course is combined with Research Seminar and is designed to prepare Junior students with the tools and methods necessary to answer all statistics and epidemiology questions on Part II of the National Board. Upon completion of the course, students will understand the relationships that exist between data scales, data types and inferential testing. They will be able to calculate probability mathematics that are foundational to the understanding and application of biostatistics to medicine. They will be able to design a study and to select appropriate variables, research model designs, and tests. They will be able to interpret test results and to determine the level of significance or association that exists, based on probability and non-probability approaches. The Research Seminar part of the course is embedded in the homework assignments that require students to apply concepts learned in the course.

METHODOLOGY This course consists of 12 lectures, each 110 minutes long. Homework and reading assignments will be given each week and the student is expected to engage in scholarly discussion during each class.

CLASS SIZE: Up to 90

NUMBER OF TIMES ADMINISTERED PER ACADEMIC YEAR: Annually

## III. COURSE PREREQUISITES: NONE

## IV. TERMINAL BEHAVIORAL OBJECTIVES:

UPON COMPLETION OF THIS COURSE, THE STUDENT SHOULD BE ABLE TO:

Apply descriptive and inferential testing methods to appropriately analyze a data set

Apply appropriate methods to limit and measure statistical error.

Correctly interpret significance of outcomes from statistical testing.

Answer statistics related questions on Part II, National Boards

V. MEASUREMENT AND EVALUATION Homework 20%, Quizzes 30%, Midterm 25% and Final Exam 25%

VI. Course Failure Individual Evaluation. In accordance with CODA Standard 2-4, students that fail this course will be given a 2 week study period to review course information and then they must sit for a comprehensive examination that covers all of the concepts taught in the course. Failure on the comprehensive examination will result in a final grade of “F” for the course. Should the student pass the comprehensive examination, then the initial grade will be adjusted from an “F” to a “C”.

VII. WEB SITE FOR REFERENCE MATERIALS (Active as of 6pm, 15 Jan 2012). <http://www.garnetthenley.com>

VIII. LECTURE SCHEDULE

DATE	LECTURE TOPIC	READING ASSIGNMENT
Jan 13	Introduction/Overview	
Jan 13	Data Scales/Problem Statements/Research Design	From Whatever (by Henley), PowerPoint Video, Pagano p7-11
Jan 16	Martin Luther King's Birthday	
Jan 20	Introduction to Probability	From Whatever (by Henley), Pagano p125-135
Jan 27	Descriptive Statistics, Confidence Intervals	Pagano p11-52, 214-220
Feb 03	<b>Give Kids a Smile (No Class) This might change</b>	
Feb 10	Normal Distribution/Hypothesis Testing, Statistical Error	Pagano p52-55, 176-213, 232-257
Feb 17	Normal Distribution/Hypothesis Testing, Statistical Error	Pagano p176-213, 232-257
Feb 20	President's Day	
Feb 21-27	<b>Mid Term Week</b>	
Mar 02	Introduction to Sampling Theory	Pagano p514-525
Mar 09	t and f Distributions and Tests	Pagano p220-227, 259-317
Mar 12-16	<b>Spring Break</b>	
Mar 23**	t and f Distributions and Tests	Pagano p220-227, 259-317
Mar 30	Correlation and Regression	Pagano p398-435, 470-475
Apr 06	Chi Square Distribution and Tests	Pagano p136, 144-149, 342-360
Apr 13	Common Epidemiology Terms and Rates	Intro to Epidemiology (TBA)
Apr 20	Tie-Up Loose Ends, Review	

*Howard University is committed to providing an educational environment that is accessible to all students. In accordance with this commitment, students in need of accommodations due to a disability should contact the Office of the Dean for Special Student Services for verification and determination of reasonable accommodations as soon as possible after admission to the University, or at the beginning of each academic semester.*

## IX. Course Policies

A. Attendance and Promptness Students who do not come to class are still responsible for assignments and for understanding concepts taught in the class. Students will not be allowed to enter the classroom after 8:15 am.

B. Missed Examinations, Quizzes and Assignments There is no make-up for missed examinations, quizzes and assignments, unless the student presents a valid excuse from the Associate Dean for Student and Academic Affairs. Students with valid excuses will be not take the same exam or quiz that their classmates took, but will take an alternate examination that has the same level of difficulty. Student who do not have an official excuse will receive the grade of “F” for missed assignments, quizzes and examinations.

C. Academic integrity. Students will have unique examinations and therefore cannot cheat by sharing answers. Students caught cheating by using unauthorized visual or auditory or electronic devices will receive a grade of “F” and will be barred from taking the course for 2 years.

D. Incomplete Grades. The grade of “I” will be given to students who have legitimate excuses and were not able to meet course requirements. An incomplete grade will not be given if the student does not have a legitimate excuse for not meeting course requirements. The Associate Dean for Student and Academic Affairs is the only validation authority for Howard Dentistry academics. Otherwise, the letter grade of “F” will be given for incomplete work.

E. The syllabus will serve as the official guide and project line for course events. It will be modified only by joint consent of the course director and the class representative.

## X. Course Materials

A. **Note that pdf files will sometimes be posted on the web as substitute/supplemental reading material to support the primary textbook.**

B. Primary Textbook – “Principles of Biostatistics” Pagano and Gauvreau. Second Edition. Prooks/Cole Press

C. Principles of Epidemiology: An Introduction to Applied Epidemiology and Biostatistics 2<sup>nd</sup> Edition. CDC Epidemiology Program Office. (Provided by Prof. Henley)

D. From Whatever to Understanding and Applying Basic Statistics: Part I. by Garnett Henley

E. Winks Statistical Software. Texasoft Corporation. Part of book issue. <http://www.texasoft.com>

## COURSE OUTLINE

### Lecture 1

#### I. LECTURE TITLE: Data Scales/Problem Statements/Research Design

II. GOAL AND OBJECTIVES: 1) To introduce measures by which data can be classified into numeric systems; 2) To define common approaches to research modeling; and 3) To provide a foundational basis for critiquing the scientific literature. Upon completion of this lecture, students will be able to:

A. Given a set of data (quantitative and/or qualitative variables) and upon completion of this lecture, the student will be to identify the scale, data type, and variable type of each variable in the data set.

B. The student will know the two types of problem statements and how each is used.

C. The student will know how to identify independent and dependent variables and will be able to construct hypothesis statements.

D. The student will know the parts of a research article and understand how it should be evaluated.

#### III. TOPIC OUTLINE

A. Data Scales and Data Types

B. Problem Statements: Research Questions and Hypotheses

C. Variables and Variable Types

D. Relationship between Variable Types, Data Scales and Data Types

E. Experimental and Non-Experimental Models

F. Quantitative and Qualitative Modeling

G. Types of Journals

H. Parts of a Research Paper

I. How to Critique a Research Paper

#### IV. METHODS AND MATERIALS OF INSTRUCTION

A. LECTURE

B. HANDOUTS

C. DISCUSSION

V. READING ASSIGNMENT: Web-based at <http://www.garnetthenley.com>

## Lecture 2

### I. LECTURE TITLE: Introduction to Probability

II. GOAL AND OBJECTIVES: To introduce single event and multiple event probability as measurable statistical phenomena. Upon completion of the lecture, students will be able to:

- A. Using case examples, students will understand concepts of simple probability and conditional probability, in relation to mathematical models.
- B. Understand simplified rules that apply to multiplication and addition approaches
- C Understand probability as being a likelihood function and not an absolute
- D. Understand how bias affects probability distributions

### III. TOPIC OUTLINE

- A. Definition of Probability. Mathematical Rules
- C. Single Event Probabilities
- D. Multiple Event Probabilities
- E. Conditional Probabilities
- F. Probabilities from Cross-Tabulated Tables

### IV. METHODS AND MATERIALS OF INSTRUCTION

A. LECTURE

B. HANDOUTS

C. DISCUSSION

V. READING ASSIGNMENTS: Web-based at <http://www.garnetthenley.com>

## Lecture 3

### I. LECTURE TITLE: Introduction to Sampling Theory

II. GOAL AND OBJECTIVES: To provide the framework for understanding what sampling is, why it is necessary, and how it is performed. Upon completion of the lecture, students will be able to:

- A. Understand the concept of sampling in relation to the term population
- B. Understand the difference between probability-based and non-probability based models and when they are used
- C. Understand what sampling error is and how it is calculated
- D. Understand and calculate sample sizes for requirements that are proportional-based and mean difference-based.

### III. TOPIC OUTLINE

- A. Definition of Terms
- B. Sample relevance
- C. Probability and Non-Probability Approaches
- D. Sampling Error
- E. Calculation of Samples

### IV. METHODS AND MATERIALS OF INSTRUCTION

- A. LECTURE
- B. HANDOUTS
- C. DISCUSSION

V. READING ASSIGNMENT: Web-based at <http://www.garnetthenley.com>

## Lecture 4

### I. LECTURE TITLE: Common Epidemiology Terms and Methods

II. GOAL AND OBJECTIVES: To review some of the more common basic measurement methods associated epidemiology. Upon completion of the lecture, students will be able to:

- A. Identify study types used in epidemiology research
- B. Calculate and interpret measures of disease frequency and measures of association
- C. Understand meaning and uses of common epidemiology terms

### III. TOPIC OUTLINE

- A. Common terms used in epidemiology survey
- B. Epidemiology Models vs Bio-statistical Models
- C. Calculations related to epidemiology survey (Rates, Proportions, Indices)

### IV. METHODS AND MATERIALS OF INSTRUCTION

- A. LECTURE
- B. HANDOUTS
- C. DISCUSSION

V. READING ASSIGNMENT: Web-based at <http://www.garnetthenley.com>

## Lectures 5 and 6

### **I. LECTURE TITLE: Descriptive Statistics, Confidence Intervals, The Normal Distribution, Hypothesis Testing and Statistical Error**

II. GOAL AND OBJECTIVES: To introduce variance, error and normality as measurable statistical phenomena.

A. Given a set of data (quantitative and/or qualitative variables) and upon completion of this lecture, the student will understand how non-consecutiveness and variance in data influences measures of central tendency and measures of dispersion.

D. The student will know the characteristics of the normal distribution and will be able to use the normal distribution to make statistical estimations.

E. The student will understand the definitions of confidence interval, type I and type II errors,  $\alpha$ ,  $\beta$ ,  $p$ , power, acceptance region, and rejection region, and be able to apply the definitions to statistical error.

### III. TOPIC OUTLINE

A. MEASURES OF CENTRAL TENDENCY: Mean, Mode, Median

B. MEASURES OF DISPERSION: Variance, Standard Deviation, Coefficient of Variation, Standard Error of the Mean

C. CHARACTERISTICS OF THE NORMAL DISTRIBUTION

D. Z-SCORES AND APPLICATIONS OF THE NORMAL DISTRIBUTION

E. CONFIDENCE INTERVALS and STATISTICAL ERROR

F. DEFINITIONS: Significance,  $\alpha$ ,  $\beta$ ,  $p$ , power, acceptance region, rejection region

### IV. METHODS AND MATERIALS OF INSTRUCTION

A. LECTURE

B. HANDOUTS

C. DISCUSSION

V. READING ASSIGNMENTS: Found at: <http://www.garnetthenley.com/>

## Lecture 7

### **I. LECTURE TITLE: The t and F Distributions and Associated Tests**

II. GOAL AND OBJECTIVES: To study inferential testing approaches as mathematical assumptions that are based on probability.

A. Given a set of data (quantitative and qualitative) upon completion of this lecture, the student will be to select the appropriate inferential test to test the null hypothesis

B. The student will be able to decide when a parametric or non-parametric test is appropriate and will be able to perform the appropriate test using Winks software.

C. The student will apply the concept of “Homogeneity of Variance” in deciding which probability value presents the correct significance factor.

### III. TOPIC OUTLINE

A. CONCEPT OF INFERENTIAL TESTING

B. PARAMETRIC VS NON-PARAMETRIC TESTS

C. HYPOTHESIS TESTING

D. TESTS OF MEANS

E. TESTS OF VARIANCE (1 way and 2 way ANOVAs)

F. RELEVANCE OF EFFECT SIZE

### IV. METHODS AND MATERIALS OF INSTRUCTION

A. LECTURE

B. HANDOUTS

C. DISCUSSION

V. READING ASSIGNMENT: Web-based at <http://www.garnetthenley.com>

## Lecture 8

### I. LECTURE TITLE: Correlation Analysis, Linear and Logistic Regression

II. GOAL AND OBJECTIVES: To provide the framework for understanding how an equation can represent a relationship between predictor variables and their associated response. Upon completion of the lecture, students will be able to:

- A. Create a scatter plot and visually assess the nature of an association between two continuous variables.
- B. Interpret the calculated values of the correlation coefficient and the coefficient of determination, and understand the relationship between these two measures of association.
- C. Perform a simple linear regression and use the results to assess the magnitude and significance of the relationship between a continuous outcome variable and a continuous predictor variable for predicting values of the outcome variable.
- D. Understand why multiple regression techniques allow for the analysis of the relationship between an outcome and a predictor in the presence of confounding variables.
- E. Perform a multiple linear regression and use the results to assess the magnitude and significance of the relationship between a continuous outcome variable and multiple continuous and categorical predictor variables and for predicting values of the outcome variable.

### III. TOPIC OUTLINE

- A. Data Requirements for Regression Analysis (linear vs logistic)
- B. Relationship between Correlation Analysis, Linear and Logistic Regression
- C. Building Regression Models using Enter and Stepwise Procedures
- D. Interpreting the Results

## Lecture 9

### **I. LECTURE TITLE: The Chi-Square Distribution And Tests: Tests for Nominal/Ordinal Data**

II. GOAL AND OBJECTIVES: To introduce statistical methods of analysis for nominal and ordinal data, and to review means of assessing the reliability and validity of most testing methods and processes.

A. Given a set of data (quantitative and qualitative) upon completion of this lecture, the student will be to analyze data that is in the form of frequencies and percentages and is scaled as nominal or ordinal data.

B. The student will be able to determine when results from Fisher's exact test is preferred to the "Goodness of Fit" Chi-Square, and understand the dangers of using Yate's Correction Factor.

C. The student understand Odds Ratio and Relative Risk as embellishments to the Chi-Square test, and the student will be able to calculate the alternatives using Winks software.

D. The student will study tests of reliability and validity and will be able to apply that knowledge using Winks software.

### III. TOPIC OUTLINE

A. DATA REQUIREMENTS

B. INDEPENDENCE OF CATEGORICAL X AND Y

C. OBSERVED VS EXPECTED COUNTS

D. GOODNESS OF FIT

E. FISHER'S EXACT TEST AND YATE'S CORRECTION

F. USEFULNESS OF ODDS RATIO AND RELATIVE RISK

G. RELIABILITY AND VALIDITY

### IV. METHODS AND MATERIALS OF INSTRUCTION

A. LECTURE

B. HANDOUTS

C. DISCUSSION

V. READING ASSIGNMENT: Web-based at <http://www.garnetthenley.com>