



NORTHWESTERN  
UNIVERSITY

SCHOOL OF  
CONTINUING  
STUDIES

## MED\_INF 409: Biostatistics

### David A Stumpf, MD, PhD

d-stumpf@northwestern.edu

Phone: (847) 494-7589

Office Hours: By appointment only

### Course Description

This course is designed for the biomedical researcher and health care analyst. Topics include descriptive statistics; hypothesis development and testing; estimation; t-tests; chi-squared tests; analysis of variance; linear regression; correlation; nonparametric tests; survival analysis; odds ratio and attributable risk. Biomedical applications are discussed for each topic, as well as overall application of statistical methods in the informatics field.

### Text

D'Agostino, R.B., Sullivan, L.M., & Beiser, A.S. (2006). *Introductory applied biostatistics*. Belmont, CA: Brooks/Cole, Cengage Learning. [ISBN-13: 978-0534423995]

Statistics for the Terrified, Fifth Edition, John H. Kranzler, November, 2009, Pearson: Prentice Hall, ISBN-13:978-0-205-00406-5.

More Damned Lies and Statistics: How Numbers Confuse Public Issues by Joel Best. University of California Press. ISBN-13: **978-0520-23830-5**

### Software

SPSS – available for online student rental – no need to purchase

### Prerequisites

None.

### Learning Goals

The goals of this course are to:

- Discuss analytic approaches to health care problems.
- Summarize data in health care IT, clinical, research, and administrative analytics.
- Analyze health care data to help answer IT, clinical, research and administrative questions.
- Assess published research within the field of medical informatics.
- Formulate analytic solutions to medical informatics issues or problems within a healthcare context.

### Evaluation

The student's final grade will be determined as follows:

- Discussion Board participation = **35 points**
- Analytic Project Proposal = **25 points**
- Midterm Exam = **40points**
- Final Exam = **50points**

**Total Points: 150 pts.**

**Grading Scale**

93%–100%=A  
90%–92.9%=A-  
87%–89.9%=B+  
83%–86.9%=B  
80%–82.9%=B-  
77%–79.9%=C+  
73%–76.9%=C  
70%–72.9%=C-  
Below 70%=F

**Discussion Board Etiquette**

The purpose of the discussion boards is to allow students to freely exchange ideas. It is imperative to remain respectful of all viewpoints and positions and, when necessary, agree to respectfully disagree. While active and frequent participation is encouraged, cluttering a discussion board with inappropriate, irrelevant, or insignificant material will not earn additional points and may result in receiving less than full credit. Frequency is not unimportant, but content of the message is paramount. Please remember to cite all sources—when relevant—in order to avoid plagiarism.

All students are required to read, and respond where appropriate, on a regular basis to the material on the Discussion board. During the first sync session we will review requirements for posting content (if the class size is large, then posting requirements may be reduced).

**Proctored Assessment**

There is a proctored assessment requirement in this course. For additional information, please go to the Assignments section in Blackboard and scroll to the Proctored Exam Approval Application item.

**Attendance**

This course will not meet at a particular time each week. All course goals, session learning objectives, and assessments are supported through classroom elements that can be accessed at any time. To measure class participation (or attendance), your participation in threaded discussion boards is required, graded, and paramount to your success in this class. Please note that any scheduled synchronous or “live” meetings are considered supplemental and optional. While your attendance is highly encouraged, it is not required and you will not be graded on your attendance or participation.

**Late Work**

Work submitted late will have 10 points deducted for each day it is late. Assignments will not be accepted after two days and the student will receive zero points for the assignment.

**Learning Groups**

Learning groups are utilized in this course. More information about learning groups will be provided by the instructors via the Blackboard course site.

**Academic Integrity at Northwestern**

Students are required to comply with University regulations regarding academic integrity. If you are in doubt about what constitutes academic dishonesty, speak with your instructor or graduate coordinator before the assignment is due and/or examine the University Web site. Academic dishonesty includes, but is not limited to, cheating on an exam, obtaining an unfair advantage, and plagiarism (e.g., using material from readings without citing or copying another student's paper). Failure to maintain academic integrity will result in a grade sanction, possibly as severe as failing and being required to retake the course, and could lead to a suspension or expulsion from the program. Further penalties may apply. For more information, visit [www.scs.northwestern.edu/student/issues/academic\\_integrity.cfm](http://www.scs.northwestern.edu/student/issues/academic_integrity.cfm).

Plagiarism is one form of academic dishonesty. Students can familiarize themselves with the definition and examples of plagiarism, by visiting [www.northwestern.edu/uacc/plagiar.html](http://www.northwestern.edu/uacc/plagiar.html). A myriad of other sources can be found online.

Some assignments in this course may be required to be submitted through SafeAssign, a plagiarism detection and education tool. You can find an explanation of the tool at <http://wiki.safeassign.com/display/SAFE/How+Does+SafeAssign+Work>. In brief, SafeAssign compares the submitted assignment to millions of documents in large databases. It then generates a report showing the extent

to which text within a paper is similar to pre-existing sources. The user can see how or whether the flagged text is appropriately cited. SafeAssign also returns a percentage score, indicating the percentage of the submitted paper that is similar or identical to pre-existing sources. High scores are not necessarily bad, nor do they necessarily indicate plagiarism, since the score does not take into account how or whether material is cited. If a paper consisted of one long quote that was cited appropriately, it would score 100%. This would not be plagiarism, due to the appropriate citation. However, submitting one long quote would probably be a poor paper. Low scores are not necessarily good, nor do they necessarily indicate a lack of plagiarism. If a 50-page paper contained all original material, except for one short quote that was not cited, it might score around 1%. But, not citing a quotation is still plagiarism.

SafeAssign includes an option in which the student can submit a paper and see the resultant report before submitting a final copy to the instructor. This ideally will help students better understand and avoid plagiarism.

**Other Processes and Policies**

Please refer to your SCS student handbook at [www.scs.northwestern.edu/grad/information/handbook.cfm](http://www.scs.northwestern.edu/grad/information/handbook.cfm) for additional course and program processes and policies.

## Course Schedule

**Important Note:** Changes may occur to the syllabus at the instructor's discretion. When changes are made, students will be notified via an announcement in Blackboard.

### Session 1

#### Learning Objectives

After this session, the student will be able to:

- Explain use of statistics in medical informatics.
- Identify fundamental steps in analytics.
- Identify types of variables.
- Recognize levels of measurement.
- Distinguish between population and sample.
- Distinguish between parameter and statistic.
- Distinguish between various methods of sampling.
- Open a data file in SPSS to conduct basic operations.

#### Course Content

##### Textbook Reading

*D'Agostino*, Chapters 1 – 2

##### Online Reading

Review NIH website ([www.nih.gov](http://www.nih.gov)) for medical informatics and medical research

Research Fundamentals: Statistical Considerations in Research Design: A Simple Person's Approach  
Handout #1 – Levels of measurement

##### Self-Guided Practice

SPSS software basic package, current version

Review SPSS Tutorial in software:

- Help->tutorial->Introduction
- Helpful sections include: opening a data file, running an analysis, and viewing results. ...tutorial->reading data -> reading data from SPSS files, reading data from Excel spreadsheets

Print and Video SPSS Tutorials on Web: <http://calcnnet.mth.cmich.edu/org/spss/toc.htm>

Sample data file: pain\_medication.sav

Self-guided practice (not turned in); compare with self-guided practice answer key.

Files:

- Pain\_medication.sav
- Pain\_medication questions.doc
- Solutions to Questions for Pain medication.doc

#### Discussion Board

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students' comments is graded. For this session's discussion topic(s), visit the discussion board in Blackboard.

#### Assignments

None.

#### Sync Session

June 19 - First Sync Session 7-9pm CENTRAL

## **Session 2**

### **Learning Objectives**

After this session, the student will be able to:

- Distinguish between qualitative and quantitative variables.
- Distinguish between nominal, ordinal, interval, and ratio data use in SPSS.
- Distinguish between symmetrical, bimodal, and skewed distributions.
- Distinguish between positively and negatively skewed distributions.
- Define key terms in descriptive statistics.
- Identify outliers in a data set.
- Distinguish which form of data presentation is appropriate for different situations.
- Summarize data using appropriate forms of data presentation.
- Interpret descriptive statistics given SPSS output for a study.

### **Course Content**

#### **Textbook Reading**

*D'Agostino*, Chapter 2

#### **Online Reading**

Handout #1 – levels of measurement

#### **Multimedia**

Descriptive Statistics

#### **Self-Guided Practice**

SPSS software basic package, current version

Review SPSS tutorial:

- Help->tutorial→ examining summary statistics for individual variables.

Print and Video SPSS Tutorials on web: <http://calcnet.mth.cmich.edu/org/spss/toc.htm>

Self-guided practice (not turned in); compare with self-guided practice answer key

### **Discussion Board**

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students' comments is graded. For this session's discussion topic(s), visit the discussion board in Blackboard.

### **Assignments**

None.

### **Sync Session**

June 25 - Optional Open Office Session 7:30 pm Central

## **Session 3**

### **Learning Objectives**

After this session, the student will be able to:

- Define deterministic and random experiments.
- Define probability and its basic properties.
- Explain the probability distribution.
- Apply the rule of addition for probabilities for mutually exclusive events.
- Apply the rule of multiplication of probabilities for independent events.
- Calculate the probability of an event and its complement.
- State the properties of a binomial distribution.
- Compute probabilities using normal and binomial distributions.
- Interpret probabilities using normal and binomial distributions.
- Describe the relationship between the central limit theorem and the normal probability distribution.
- Differentiate between the normal distribution and the standard normal curve.
- Explain the Z score.
- Calculate the Z score given a value from a normal probability distribution (NPD) with a specified mean and standard deviation.
- Use the Z score table to determine the probability associated with any Z score.

### **Course Content**

#### **Textbook Reading**

*D'Agostino*, Chapters 3 - 5

#### **Multimedia**

Probability  
Binomial Distribution  
Central Limit Theorem  
Normal Probability Distribution

#### **Self-Guided Practice**

Self-guided practice (not turned in); compare with self-guided practice answer key

### **Discussion Board**

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students' comments is graded. For this session's discussion topic(s), visit the discussion board in Blackboard.

### **Assignments**

None.

### **Sync Session**

July 3 - Optional Open Office Session 7:30 pm Central

## **Session 4**

### **Learning Objectives**

After this session, the student will be able to:

- Explain the difference between parent distributions and sampling distributions.
- Describe the use of the central limit theorem to determine the properties of the probability distribution for a sampling distribution.
- Describe standard error of the mean and its relation to sample size.
- Compute the standard error of the mean.
- Interpret the standard error of the mean.
- Compute the confidence interval for a sample mean.
- Describe when the t-distribution should be used.
- Define the assumptions for one-sample procedures.
- Compute one-sample procedures.
- Interpret results obtained from one-sample procedures.
- State the null and alternative hypotheses for a general research hypothesis.
- Describe the five steps of hypothesis testing.
- Use the five steps of hypothesis testing for a one-sample test of a sample mean.
- Discuss use of one-sample procedures within medical informatics literature.

### **Course Content**

#### **Textbook Reading**

*D'Agostino*, Chapters 4 - 5

#### **Online Reading**

Confidence Intervals

#### **Self-Guided Practice**

Self-guided practice (not turned in); compare with self-guided practice answer key

### **Discussion Board**

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students' comments is graded. For this session's discussion topic(s), visit the discussion board in Blackboard.

### **Assignments**

The Practice Midterm is available for review.

### **Sync Session**

July 10 - Optional Open Office Session 7:30 pm Central

## **Session 5**

### **Learning Objectives**

After this session, the student will be able to:

- Define the t-test for an independent sample.
- Define the t-test for a dependent sample.
- State the assumptions for use of t-tests to analyze data.
- Describe the t-distribution.
- State the hypothesis for t-tests.
- Define type 1 and type 2 errors.
- Define power.
- Compute t.
- Determine the sample size needed for a two-sample procedure.
- Use the five steps of hypothesis testing for two sample tests of sample means.
- Discuss use of two-sample procedures within medical informatics literature.

### **Course Content**

#### **Textbook Reading**

*D'Agostino*, Chapter 6

#### **Self-Guided Practice**

Self-guided practice (not turned in); compare with self-guided practice answer key

Review SPSS Tutorial in software

Print and Video SPSS Tutorials on web: <http://calcnnet.mth.cmich.edu/org/spss/toc.htm>

### **Discussion Board**

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students' comments is graded. For this session's discussion topic(s), visit the discussion board in Blackboard.

### **Assignments**

For more information, click Assignments on the left navigation bar in Blackboard, and scroll to this assignment's item.

### **Sync Session**

July 17 - Second Sync Session 7-9 pm. "Take-home" Midterm Available July 18; due by Midnight, July 23



## **Session 6**

### **Learning Objectives**

After this session, the student will be able to:

- Differentiate between independent and dependent variables.
- Define ANOVA terminology.
- Describe the F distribution.
- Define the assumptions for ANOVA.
- Interpret ANOVA tables.
- Apply multiple comparison procedures in ANOVA.
- Use the five steps of hypothesis testing for ANOVA.
- Discuss use of ANOVA within medical informatics literature.

### **Course Content**

#### **Textbook Reading**

*D'Agostino*, Chapter 9

#### **Self-Guided Practice**

Self-guided practice (not turned in); compare with self-guided practice answer key

Review SPSS Tutorial in software

Print and Video SPSS Tutorials on web: <http://calcnet.mth.cmich.edu/org/spss/toc.htm>

### **Discussion Board**

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students' comments is graded. For this session's discussion topic(s), visit the discussion board in Blackboard.

### **Assignments**

None.

### **Sync Session**

July 23 - Optional Open Office Session 7:30 pm

## **Session 7**

### **Learning Objectives**

After this session, the student will be able to:

- Describe the chi-square distribution.
- Describe appropriate uses of the chi-square test.
- Define the assumptions for chi-square.
- Compute observed and expected values for a chi-square contingency table.
- Use the five steps of hypothesis testing for chi-square.
- Discuss use of chi-square within medical informatics literature.

### **Course Content**

#### **Textbook Reading**

*D'Agostino*, Chapter 7

#### **Online Reading**

Handout: Calculating Chi-square and Expected Frequencies

#### **Self-Guided Practice**

Self-guided practice (not turned in); compare with self-guided practice answer key

Review SPSS Tutorial in software

Print and Video SPSS Tutorials on web: <http://calcnet.mth.cmich.edu/org/spss/toc.htm>

### **Discussion Board**

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students' comments is graded. For this session's discussion topic(s), visit the discussion board in Blackboard.

### **Assignments**

None.

### **Sync Session**

July 31 - Optional Open Office Session 7:30 pm Central

## **Session 8**

### **Learning Objectives**

After this session, the student will be able to:

- Define key terms in correlation and regression.
- Describe independent ( $X$ ) and dependent ( $Y$ ) variables in correlation and regression.
- Define the assumptions for correlation and regression.
- Recognize linear and other models of regression.
- Compute  $r$ ,  $r^2$  and linear regression components.
- Interpret the nature and strength of association between two variables using a scatter gram,  $r$  and  $r^2$ .
- Interpret linear regression components.
- Use a regression equation to predict  $Y$ .
- Use the five steps of hypothesis testing for correlation and regression.
- Discuss use of correlation and linear regression within medical informatics literature.

### **Course Content**

#### **Textbook Reading**

*D'Agostino*, Chapter 10

#### **Multimedia**

Linear Regression

#### **Self-Guided Practice**

Self-guided practice (not turned in); compare with self-guided practice answer key

Review SPSS Tutorial in software

Print and Video SPSS Tutorials on web: <http://calcnet.mth.cmich.edu/org/spss/toc.htm>

### **Discussion Board**

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students' comments is graded. For this session's discussion topic(s), visit the discussion board in Blackboard.

### **Assignments**

None.

### **Sync Session**

Aug 7 - Optional Open Office Session 7:30 pm Central

## **Session 9**

### **Learning Objectives**

After this session, the student will be able to:

- Differentiate between parametric and nonparametric tests.
- Define the assumptions for various nonparametric tests.
- Identify the nonparametric test that would be used if the assumptions for the corresponding parametric test were not met.
- Use the five steps of hypothesis testing for nonparametric tests.
- Discuss use of nonparametric tests within medical informatics literature.
- Propose a research-based solution to a medical informatics problem within a healthcare context.

### **Course Content**

#### **Textbook Reading**

*D'Agostino*, Chapter 12

#### **Online Reading**

Handout: Nonparametric vs. Parametric Tests

#### **Self-Guided Practice**

Self-guided practice (not turned in); compare with self-guided practice answer key

Review SPSS Tutorial in software

Print and Video SPSS Tutorials on web: <http://calcnet.mth.cmich.edu/org/spss/toc.ht>

### **Discussion Board**

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students' comments is graded. For this session's discussion topic(s), visit the discussion board in Blackboard.

### **Assignments**

). For more information, click Assignments on the left navigation panel in Blackboard, and scroll to this assignment's item.

The Practice Final Exam is available for review.

### **Sync Session**

Aug 14 – Third and **Final Regular Sync Session 7-9pm**

## **Session 10**

### **Learning Objectives**

After this session, the student will be able to:

- No new learning objectives will be introduced in this Session.

### **Course Content**

None.

### **Discussion Board**

Each session you are required to participate in the session-specific discussion board forum. Your participation in both posting and responding to other students' comments is graded. For this session's discussion topic(s), visit the discussion board in Blackboard.

### **Assignments**

The Final Exam is due in Session 10 and must be submitted by 11:59pm Aug 26<sup>th</sup> to receive credit. For more information, click Assignments on the left navigation bar in Blackboard, and scroll to this assignment's item.

### **Sync Session**

None.