

MED INF 409 Syllabus

Dr. Richard Gershon
Vice-Chair, Research Associate Professor
Department of Medical Social Sciences, Northwestern University

BIOSTATICS AND MEDICAL INFORMATICS **SUMMER 2010**

Contact Information:

Phone: 1-312-503-3453

E-mail: gershon@northwestern.edu

Office hours: (by appointment only)

Course Description: The biostatistics course is designed for informatics specialists who would like to understand the research methods and statistics underlying biomedical research. Topics include: descriptive statistics, hypothesis testing, estimation, t-tests, chi-squared tests, analysis of variance, linear regression, correlation, nonparametric tests and item response theory. Biomedical applications are discussed for each topic, as well as overall application of statistical methods in the informatics field. This course is the biostats requirement for students in the distance learning MMI program. On-campus students take PUB HLTH 302.

Text:

Statistics for the Terrified, Fifth Edition, John H. Kranzler, 2010, Pearson: Prentice Hall.

NOTE: Hopefully this is the (new) version received by the bookstore. If necessary purchase the Fourth Edition. I don't particularly care which stats text that you actually use, but former students in this course did recommend this one.

Statistics at Square One. 11th Edition. TDV Swinscow and MJ Campbell. London: BMJ Books, 2002, ISBN: 978-1-4051-9100-5.

More Damned Lies and Statistics: How Numbers Confuse Public Issues by Joel Best. University of California Press.

Software:

OPTIONAL: SPSS Version 18 also known as PASW Version 18.

This is the product from which I will demonstrate the practical application of statistics and which you will need in order to complete some of your assignments. You may definitely use an earlier version or even another stats software package if you are stats savvy (although if you are very stats savvy you should consider seeking departmental permission to place out of this course).

There are online vendors who will "lease" graduate students this software for 6 months for less than \$50. I will post at least one such source on the Blackboard Announcements the first week of class.

You can wait until after the first class session to determine whether or not you should lease this software.

Learning Goals: The goals of this course are to:

- 1) become a critical consumer of the basic statistics utilized in medical informatics research.
- 2) understand how to recognize the indicators of statistical significance
- 3) be able to distinguish between unbiased and biased research models
- 4) be able to conduct simple statical analyses

Evaluation: Your final grade for the course will be based on the following activities:

- RAT 1 - RAT 7 = (4% each) 28%
- Statistical Practice 1-5 (4% each) 20%
- Participation in class discussions and online discussions 22%
- Final Exam = 30%

Grading Scale: While I reserve the right to apply a curve at the end of the quarter, I have utilized the following grading scale for the last few quarters. A curve could be applied when new exam content is added/subtracted and where that content results in a significant change in student performance (typically this would be the case when an exam is too difficult, necessitating a curve to improve student grades, but could also be the case if new exams are added which are too easy, and the grading made more difficult).

93-100 A

90-93 A-

87-90 B+

83-87 B

80-83 B-

77-80 C+

73-77 C

70-73 C-

Online Activities:

- Discussion Board:
You can start your own threads on any topic related to the session. You should check the Board to read what your classmates and instructor will have posted and respond to those posts. Students will receive credit for posing valuable questions and/or answering the questions posed by others.
- Synchronous Sessions:
These are weekly meetings that you'll have with your instructor and classmates and, possibly, a guest speaker. These sessions will be recorded and accessible if you have to miss the live meeting for any reason.

Please set aside time on your calendar now for the synchronous sessions. If you must miss a session, please let the instructor know in advance, and if you know that you will miss more than one session, please let the instructor know as soon as possible. Regular synchronous course attendance is mandatory for this section (a portion of the grade is based upon class participation).

- **Class Participation:**
... is a fundamental component of the class. Each week at least half the class will be called upon to participate in a discussion based upon the Best reading. As long as you are prepared to participate in that discussion, and are present in class to do so, you will receive full credit for participation. Should you miss more than 2 class sessions for any reason, or are not prepared to participate, you will be downgraded accordingly.
- **Readings:**
The Assigned readings must be read before the synchronous portion of the session in order to fully participate in the session.
- **Quizzes:**
Each week there will be a short, true-false and/or multiple choice quiz called a Readiness Assessment Test (RAT) covering basic concepts from your readings given during the course combined with material presented in the lecture- one at the end of each unit of material.

You must individually complete each RAT in the timeframe starting 9 am central on the Tuesday after the synchronous session and ending 6 days later, at 10 pm central time on the Sunday following the synch session.

I will review the RATs and identify issues which should be reviewed during the next week's session. Scores from the RATs WILL count toward your final grade, and should be indicative of how you are likely to perform on the final examination.

- **Statistical Practice**
During the quarter there will be five statistical practice exercises. Each exercise will be explained at the end of a class session and will be due one week later. For each practice session a hypothesis will be posted along with a dataset. Your assignment will be to analyze the dataset using SPSS (or any other statistics package if your choosing). Your submission will be a one paragraph summary of the results including a statement as to whether or not the results of the hypothesis test that you conducted were significant, and what conclusion can be drawn. You should cut and paste a copy of your SPSS output file into the second page of your submission. Statistical practice submissions will be graded as Pass (5 points) or Fail (0 points). Failed submissions can be re-written within one week for full credit less one point (4 points if submitted within one week following the original submission). Similarly, late submissions which ultimately result in a "Pass" will be reduced by one point per week.
- **Homework:**
Prior to the start of each class, you must complete the assigned reading and complete the previous Readiness Assessment Test individually. Five times during the course you will also have a statistical practice exercise which must be submitted prior to the following class to obtain full credit.
- **Final Exam:**

This course requires that students take the Final Exam at an approved testing center. More details on the exam will come later in the quarter, but each student is responsible for choosing a

testing center to be approved by SCS. Complete the following assignments before the first session:

1. Click on the proctor approval application form link (on the Assignment page under the Session 1 tab). You do not need to save this PDF anywhere or send it to your instructor as an attachment. Simply fill it out, then click "Submit by Email," found on the lower right-hand corner of the form. This part of the assignment is necessary for Northwestern to ensure all testing centers are prepared for the appropriate MMI students.
2. Let your instructor know you've submitted your testing center information by going through the "Submit Work" link, then the Session 1 Assignments folder. Click on the "View/Complete Assignment" link below and writing "Done" in the Comments box. Then click "Submit" at the bottom of the page. This part of the assignment is to ensure you receive credit for this assignment through Blackboard.

Discussion Board Etiquette: The purpose of Discussion Boards is to allow students to freely exchange ideas and participation is highly encouraged. It is important that we always remain respectful of one another's viewpoints and positions and, when necessary, agree to disagree, respectfully. 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. Although frequency is not unimportant, content of the message is paramount. Please remember to cite all sources – when relevant – in order to avoid plagiarism.

Attendance: This course is mostly asynchronous, meaning that we will not meet at a particular time each week. Even though we will not be meeting face-to-face in a physical classroom, attendance at the synchronous sessions and participation on Discussion Boards are required and paramount to your success in this class.

Late Work: Late work will be accepted with certain conditions. Advance notice is not necessary. Late assignments will be penalized 1 point per week regardless of the reason. Any assignments not submitted within 1 week of the last regular class day will be marked as 0.

A final exam taken after the final exam period will require advanced notice, may be subject to departmental approval, and will definitely result in an incomplete or failing grade for the quarter.

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., taking material from readings without citation 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:

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 the site <http://www.northwestern.edu/uacc/plagiar.html>. Myriad other sources can be found online, as well.

There is currently no plan for essay based assignments for this course. If this changes, these new assignments may be required to be submitted through SafeAssign, a plagiarism detection and education

tool. You can find an explanation of the tool [here](#). In brief, SafeAssign compares the submitted assignment to millions of documents in very large databases. It then generates a report showing the extent to which text within a paper is very similar or identical to pre-existing sources. The user can then see how or whether the flagged text is cited appropriately, if at all. 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 doesn't take into account how or whether material is cited. (If a paper consisted of just one long quote that was cited appropriately, the score would be 100%. This wouldn't be plagiarism, due to the appropriate citation. However, just submitting one long quote would probably be a pretty bad paper.) Low scores are not necessarily good, nor do they necessarily indicate a lack of plagiarism. (If a 50-page paper had all original material, except for one short quote that was not cited, the score might be around 1%. But, not citing a quotation would still be plagiarism.)

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

Other Processes and Policies: Please refer to your SCS student handbook at <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.

Monday, June 14 – Saturday, June 19, 2010: Blackboard Upgrade & Outage (no access to Blackboard)

Due to the short nature of the summer quarter and since we will not have class on the independence day holiday (celebrated Monday, July 5), care must be taken to keep up with our "rapid" coverage of the course material.

Session 1 - June 21 Introduction to Biostatistics

This session will cover:

- Orientation to the course
- Basic vocabulary of biostatistics
- Introduction to biostatistics
- Frequency distributions and percentiles
- Central tendency & variability
- Z-scores & normal distributions
- Overview of inferential statistics
- Probability

Learning goals for this session:

- Define frequency distributions and percentiles
- Define central tendency and variability
- Define z-scores
- Define normal distributions
- Define inferential statistics
- Define probability

Readings	Assignments	Synchronous Meeting
<p>Assigned Readings:</p> <ul style="list-style-type: none"> • Kranzler chapters 1, 2, 4, 5, 6, and 7 • Campbell chapter 1-4 	<ul style="list-style-type: none"> • Introduce yourself in response to the first question in the Discussion Board and respond to the other questions in the Discussion Board • Complete and submit a Proctor Approval Application form (see under "Final Exam," above for more information) <p>Readiness Assessment Test (RAT) #1</p>	<p>Synchronous Session: Join the online discussion with your instructor and classmates.</p> <p>Consult the Announcements page for instructions on accessing the session.</p> <p>The link persists through the entire quarter: please bookmark it so you can access the synch sessions directly from your browser.</p>

Session 2 – June 28
Introduction to Biostatistics II

This session will cover:

- Sampling distributions
- Logic of hypothesis testing
- Power
- Logic of parameter estimation

Learning goals for this session:

- Define sampling distributions
- Define logic of hypothesis testing
- Define the logic of parameter estimation
- Define power

<u>Readings</u>	<u>Assignments</u>	<u>Synchronous Meeting</u>
Assigned Readings: <ul style="list-style-type: none"> • Kranzler chapter 10 • Campbell chapter 5 • Best chapter 1 	Discussion Board <ul style="list-style-type: none"> • Respond to the questions or begin your own forum topics in the Discussion Board Readiness Assessment Test (RAT) #2	Synchronous Session: Join the online discussion with your instructor and classmates. For dates and times, see the schedule on the <i>Course Resources page</i> . Consult the Announcements page for instructions on accessing the session.

July 5
Observation of Independence Day

NO CLASS

Session 3 – July 12 The Single Sample t Statistic

This session will cover:

- Using SPSS (PAW)
- The single sample t statistic

Learning goals for this session:

- Define the single sample t statistic
- Recognize when the procedures are used
- Recognize the pre-requisites
- Construct a null hypothesis
- Construct an alternative hypothesis
- Interpret the results of the procedures
- Understand the basics of using SPSS (PAW)
- Understand how to use SPSS to conduct a single sample t test

Readings	Assignments	Synchronous Meeting
<p>Assigned Readings:</p> <ul style="list-style-type: none"> • Kranzler chapter 3 • Campbell chapter 6 (tentative) • Best chapter 2 	<p>Discussion Board</p> <ul style="list-style-type: none"> • Respond to the questions or begin your own forum topics in the Discussion Board <p>Readiness Assessment Test (RAT) #3</p> <p>Statistics Practice - t-test</p>	<p>Synchronous Session:</p> <p>Join the online discussion with your instructor and classmates.</p> <p>For dates and times, see the schedule on the <i>Course Resources page</i>.</p> <p>Consult the Announcements page for instructions on accessing the session.</p>

Session 4 – July 19
t-test for Independent Samples and Dependent Sample t

This session will cover:

- t-test for independent samples
- dependent sample t

Learning goals for this session:

- Define the t-test for independent samples
- Define the dependent sample t
- Recognize when the procedures are used
- Recognize the pre-requisites
- Construct a null hypothesis
- Construct an alternative hypothesis
- Interpret the results of the procedures
- Understand how to conduct analyses for each of these tests in SPSS
- Understand how to interpret the results for each of these tests in an SPSS output

<u>Readings</u>	<u>Assignments</u>	<u>Synchronous Meeting</u>
Assigned Readings: <ul style="list-style-type: none"> • Kranzler chapter 11 • Campbell chapter 7 • Best chapter 3 	Discussion Board <ul style="list-style-type: none"> • Respond to the questions or begin your own forum topics in the Discussion Board Readiness Assessment Test (RAT) #4 Statistics Practice: Independent t-test	Synchronous Session: Join the online discussion with your instructor and classmates. For dates and times, see the schedule on the <i>Course Resources page</i> . Consult the Announcements page for instructions on accessing the session.

Session 5 – July 26
Comparing Population Variances and Analysis of Variance (ANOVA)

This session will cover:

- Comparing population variances
- Analysis of variance

Learning goals for this session:

- Define population variances
- Define analysis of variance
- Recognize when the procedures are used
- Recognize the pre-requisites
- Construct a null hypothesis
- Construct an alternative hypothesis
- Interpret the results of the procedures
- Understand how to conduct analyses for each of these tests in SPSS
- Understand how to interpret the results for each of these tests in an SPSS output

<u>Readings</u>	<u>Assignments</u>	<u>Synchronous Meeting</u>
Assigned Readings: <ul style="list-style-type: none"> • Kranzler chapter 12 • Best chapter 4 	Discussion Board <ul style="list-style-type: none"> • Respond to the questions or begin your own forum topics in the Discussion Board Readiness Assessment Test (RAT) #5 Statistics Practice: ANOVA	Synchronous Session: Join the online discussion with your instructor and classmates. For dates and times, see the schedule on the <i>Course Resources page</i> . Consult the Announcements page for instructions on accessing the session.

Session 6 – August 2
Regression and Correlation

This session will cover:

- Regression & correlation

Learning goals for this session:

- Define regression
- Define correlation
- Recognize when the procedures are used
- Recognize the pre-requisites
- Construct a null hypothesis
- Construct an alternative hypothesis
- Interpret the results of the procedures
- Understand how to conduct analyses for each of these tests in SPSS
- Understand how to interpret the results for each of these tests in an SPSS output

<u>Readings</u>	<u>Assignments</u>	<u>Synchronous Meeting</u>
<p>Assigned Readings:</p> <ul style="list-style-type: none"> • Kranzler chapters 8 and 9 • Campbell chapter 11 • Best chapter 5 	<p>Discussion Board</p> <ul style="list-style-type: none"> • Respond to the questions or begin your own forum topics in the Discussion Board <p>Readiness Assessment Test (RAT) #6</p> <p>Statistics Practice: Correlation</p>	<p>Synchronous Session: Join the online discussion with your instructor and classmates.</p> <p>For dates and times, see the schedule on the <i>Course Resources page</i>.</p> <p>Consult the Announcements page for instructions on accessing the session.</p>

Session 7 – August 9
Chi-Square

This session will cover:

- Chi-Square

Learning goals for this session:

- Define Chi-square
- Recognize when the procedure is used
- Recognize the pre-requisites
- Construct a null hypothesis
- Construct an alternative hypothesis
- Interpret the results of the procedure
- Understand how to conduct analyses for each of these tests in SPSS
- Understand how to interpret the results for each of these tests in an SPSS output

<u>Readings</u>	<u>Assignments</u>	<u>Synchronous Meeting</u>
Assigned Readings: <ul style="list-style-type: none"> • Kranzler chapter 13 • Campbell chapter 8 • Best chapter 6 	Discussion Board <ul style="list-style-type: none"> • Respond to the questions or begin your own forum topics in the Discussion Board Readiness Assessment Test (RAT) #7 Statistics Practice: Chi-Square	Synchronous Session: Join the online discussion with your instructor and classmates. For dates and times, see the schedule on the <i>Course Resources page</i> . Consult the Announcements page for instructions on accessing the session.

Session 8 – August 16 Special Topics

This session will cover:

- Factor Analysis
- Item Response Theory
- Computer Adaptive Testing
- Final Examination Review

Learning goals for this session:

- Define factor analysis
- Define and understand how to use Item Response Theory
- Understand how and when to use the results of the procedures
- Final preparation: Recognize when to use what statistic

Readings	Assignments	Synchronous Meeting
Assigned Readings: <ul style="list-style-type: none">• Best chapter 7• Campbell chapters 9-13• Chapter 8 from <i>Introduction to Rasch Measurement</i>, Gershon -- posted on Blackboard	Discussion Board <ul style="list-style-type: none">• Respond to the questions or begin your own forum topics in the Discussion Board	Synchronous Session: <p>Join the online discussion with your instructor and classmates.</p> <p>For dates and times, see the schedule on the <i>Course Resources page</i>.</p> <p>Consult the Announcements page for instructions on accessing the session.</p>

The Final Exam will be posted at 12:01 am Central August 17 is due by 11:55 pm CST August 23.

August 31 - Grades will be posted to Caesar (Please note that it can take the registrar 1-2 days to make grades visible!)