Biometry (MTH207) Dr. Derek H. Ogle Winter, 2014

Description: Foundational principles of probability including methods of counting and specific probability distribution and density functions; linear models including one- and two-way ANOVA with post-hoc multiple comparisons and transformations, simple and multiple linear regression including transformations and indicator variables, and logistic regression; and basic principles of sample and experimental design. The theoretical constructs of all topics will be developed and applied to real-life or realistic situations in the life and natural sciences. Pre-requisite MTH107. 4 cr.

Importance: Statistics is central to scientific inquiry because hypotheses are generated from exploratory data analyses of existing or preliminary data and hypotheses are objectively tested with inferential statistical methods. As such, understanding statistics is a major component of understanding how to conduct and interpret scientific research. Understanding statistics includes knowing how to properly collect data ("design") and how to properly analyze data so that appropriate conclusions can be made ("analysis"). In this course you will learn (1) how to properly design a variety of studies that are commonly used in natural and social science fields and (2) how to properly analyze the data collected from those studies. Throughout the course you will focus on the generalities of major concepts so that your knowledge will extend beyond the specific situations examined in class. In addition, because the communication of results is critical to success as a researcher, you will also focus on correct report or manuscript writing.

Outcomes: At the end of this class you will be able to ...

1. explain why statistics is central to scientific inquiry (& your field of interest);
2. perform, including assumption checks, and interpret the results from one-way and two-way analysis of variances (ANOVA);
3. perform appropriate multiple comparisons following a significant analysis of variance result;
4. perform, including assumption checks, and interpret the results from simple linear (SLR) and indicator variable (IVR) regressions;
5. perform, including assumption checks, and interpret the results from simple logistic regression;
6. use variable transformations and interaction terms where appropriate;
7. compare and contrast the advantages and limitations of one-factor and two-factor experimental designs; and
8. write a concise, detailed, accurate, and interesting paper explaining the design and interpreting the results of a statistical analysis.

Web Page: **droglenc.wordpress.com/** then select the "Teaching" menu and "Biometry" submenu.
Location: CSE 236 (computer room) **Time:** MWF 1200-1250p, R 1200-1350
Contact: CSE239; 682-1300; dogle@northland.edu **Office Hours:** MWF 900-1000, or by appointment

Accommodations: Students in need of academic or medical accommodation should contact Kathleen Skoraczewski, Campus Counselor, ext. 1369, Ponzio Center Rm. 202B, kskoraczewski@northland.edu.

Conduct: Your conduct in class should revolve around the idea of being respectful of me and others in the class. I will ask you to leave class if you behave disrespectfully. Adhering to the following policies should be considered the minimum requirements to being respectful of others:

1) Arrive to class on time. If a situation arises where you will arrive late, then take the first available seat as quietly as possible.
2) Do not bring friends, children, pets, meals, or anything else to class that should not be there.
3) **Turn off and store out-of-sight cell phones and other electronic devices (personal laptops are ok).**
4) **Do not use the college or personal computers for other than assigned tasks.**
5) Do not engage in side discussions while I or others are speaking to the entire class.
6) Do not use disrespectful language when addressing others.

Textbook: I will provide a PDF document on the webpage as the text for this course. This document cannot be printed, but I will make arrangements on the first day of class for you to a hard copy (for the cost of copying) if you are interested.

Computing: Computers will be used extensively as you will be asked to type homework problems and perform statistical analyses using R (version 3 or newer) and RStudio, both of which can be obtained for free (see "R Resources" link on the class webpage). In addition, you must install the **`NCStats`** package for R (see "R Resources" link).

Grading: An overall percentage score will be computed from (with associated weights) homework assignments (20%), exams (40%), and a final portfolio (40%). Lowest possible final letter grades will be assigned by comparing your overall percentage score (rounded to a whole number) to the values shown below:

		A	92-100	**A-**	90-91
B+	87-89	**B**	82-86	**B-**	80-81
C+	77-79	**C**	70-76		
D+	67-69	**D**	60-66	**F**	0-59

Incomplete Grades: Under Northland College policy, an incomplete grade will be given ONLY under extreme circumstances beyond your control, such as a major illness. An incomplete grade will be given ONLY if you have successfully completed the entire course except for the final exam.

Attendance: For most students, achievement is strongly positively related to attendance. I will take attendance each day but attendance will not factor into your final grade with the following exception. At the end of the semester I may, at my discretion, decide to eliminate one or more of your homework or exam grades. Your attendance record will be one item considered when I make this decision.

Homework: Several homework assignments will be assigned per major course topic. Grading of each assignment will be described on the assignment sheet. Completed homework assignments will be due at the beginning of the class period on dates announced when the homework was assigned. Assignments will be accepted up to one day late but will be reduced by 50% of the worth of the assignment. Assignments not handed in at the beginning of the class period will be considered one day late.

Note that not all homework problems will be graded for content, some will be graded only for completeness (you will be responsible for checking content with on-line answer keys).

Each homework should include the following statement along with your signature

"I have neither given nor received unauthorized aid in completing this work, nor have I presented someone else's work as my own."

Unauthorized aid includes working "too much" with other students, either in the current class or not. In other words, you can discuss homework problems and course content with other students but if you find yourself always looking to someone else for help on how to do something or how to explain something then you are not substantively working independently. Start the homework early so that you can ask ME questions rather than relying on following exactly what a friend has done. If I suspect that you and another person have violated this rule then I will verbally warn each of you at the first such suggestion and then assign you no credit for each subsequent offending homework. If any homework is ***largely*** a verbatim copy or rephrasing of another person's work then ***both*** of you will receive no credit for that homework, even if it is the first such offense.

Exams: Exams will be given in class on **13-Feb**, **13-Mar**, and **17-Apr**. Make-up exams will be provided only if you have a fixed commitment of sufficient importance that was set before the beginning of the semester or a verifiable medical condition. You will not be allowed to make up an exam missed without my prior approval. The final exam will not be given early.

Portfolio: The portfolio project serves as the capstone experience for the course and allows you flexibility to demonstrate working knowledge of the major topics of the course. In general, the portfolio will consist of a thorough analysis of data using **four** (*cannot use any topic more than once*) of the following major topics in the course: (I) One-Way ANOVA (*must have more than two groups and cannot be a subset of your two-way ANOVA)*, (II) Two-Way ANOVA, (III) Simple Linear Regression (*cannot be a subset of your IVR*), (IV) Indicator Variable Regression, or (V) Logistic Regression. Alternatively, you can propose an alternative topic for my approval.

For each topic you must find or gather and appropriately analyze data. The data can originate from your original research, another class (but cannot be exactly like an assignment for another class), textbooks, journal articles (you may have to recreate the data from a graphic or table – see me for help with this), reputable internet sources, or any other reliable source. The data should be real, should be able to be analyzed with one of the five topics from this class, and should be interesting to you. ***Finding appropriate data may be the hardest part of the portfolio***. Thus, you must submit an initial proposal to me by **24-Feb** that briefly describes the data you will use, the types (i.e., quantitative or categorical) of response and explanatory variable(s), the question to be asked or hypotheses to be tested with the data, and which topic you will use to answer the question. The initial proposal is worth an ***all-or-nothing*** 10% of the overall grade for the portfolio. I strongly urge you to discuss your initial proposal with me before turning it in.

The final portfolio project will be a written research report detailing your results and conclusions for each topic. The report should contain the following sections:

- ***Introduction:*** describe any background information needed to understand the data, why the data are important to you or others, why you were interested in the data, and what the objectives of the study were (i.e., the research hypothesis).
- ***Methods:*** describe the methodology you (or the authors) used to collect and you used to analyze the data.
- ***Assumptions:*** describe all assumption checking and data transformations.
- ***Results:*** describe the results of your analysis.
- ***Discussion:*** describe what you found and how it relates to your reasoning from the introduction.
- ***Works cited:*** list of references (if any) used to support your project.

It is possible that all four analyses can be conducted from the same large data set. Thus, it is possible to have a single introduction and single methodology describing the data collection, but separate methods for analysis, assumption checking, results, and discussion. Alternatively, if different data sets were used for each topic then separate introductions, methods, assumption checking, results, and discussion will be required. The final portfolio should be typed, grammatically correct (proofread!), and refer to supporting graphics and tables. The final portfolio is due on **14-Apr**.