**KINE 8970: Biostatistics I**

**Lectures**: TH 3:30-4:45pm in KINE 112

**FINAL EXAM:** None

**Instructor**: Keith Lohse, PhD

**|Office**: KINES 279  **|Office Hours**: Tuesday 1-3pm

**|Email**: lohse@auburn.edu (preferred) **|Phone**: 334-844-1982

**Goals of the course:**

1. To give students a detailed understanding of the fundamental processes of inferential statistics.
2. To provide an introduction to computerized data analysis and visualization using R.
3. To give students mastery of common statistical tests, their assumptions, and how to use these tests in research.
4. To provide students with knowledge, but also a digital library of scripts, data-sets, and readings that will allow them to take on progressively more challenging courses in statistics.

**Course Philosophy:**

I love to teach statistics—for several reasons. For instance, I absolutely believe that statistics is one of most important classes you can take, even if you do not go on to a career in research. Statistics are a set of tools—the best and often only ones we have—to learn about complex real world problems. Nearly everything you learn in your past, current, and future classes has used the types of statistical tools that we are going to explore. Also, statistics is an empowering class; it gives you the tools to understand experimental research, but also to do your ownexperimental research. Finally, statistics is challenging, it forces you to think analytically (reducing problems to their essential components), but also creatively (e.g., given a particular problem, how can we test the question we are interested in).

**Required Materials:**

***R:*** <http://www.r-project.org/>

***R Studio:*** <http://www.rstudio.com/products/rstudio/>

***Keith’s Biostatistics webpage:*** <https://sites.google.com/site/lohsekr/biostatistics>

**Grading and Assessment:**

As you will see, this course treats learning biostatistics like a game. "Biostatistics" is an unknown world that the student must explore. Hence, we pay homage to one of my favorite games by reminding students that, "***It's dangerous to go alone.***" If you are not familiar with the Legend of Zelda, in the beginning of the game the protagonist, Link, is given a sword to start his journey and told, “It’s dangerous to go alone! Take this.” and the old man gives him a sword.

Over the course of his adventures, Link not only learns how to use this sword more effectively, but learns to use a host of other weapons as well. Similarly, we start with learning some very basic statistics and how to implement them in the "base" package of R. As our skills grow, however, we will start to learn to use other packages to deal with different types of data and particularly difficult analyses.

***The course begins with all students having zero experience points (XP)***. There will be no tests on which you lose points during the course, but you can gain points in three different ways: (1) Earning ***experience*** by taking quizzes on the recorded lectures, (2) gaining ***skills and specialties*** by completing the in-lab assignments, and (3) completing ***quests*** (rhymes with tests) that require practical application of the knowledge you have gained in other areas of the course.

***To earn XP***, you can complete daily quizzes that will be given at the beginning of each class. You can also gain XP through completing lab assignments and completing quests.

The lab assignments are designed ***to help you gain skills***. These assignments are to be done in lab with the instructor. These skills will prepare you for the quests in the course and (hopefully) for the application of statistics in your own research. Sometimes lab assignments will be worth XP as well… but we never know when… so it’s important for you to attend class as often as possible (otherwise you might miss out on that sweet XP!).

***To complete a quest***, you will need to complete the practical module that goes with that quest with at least 80% accuracy. If you fail to complete a quest, or if you would simply like to score higher on the module, you are welcome to retake it. On your second attempt, you can earn up to 80% of the original xp (e.g., if you get 0% the first time and then 100% on the second attempt, you will finish with 80% on that module and beat the quest). There are four quests available in Biostatistics I and you need to complete all of the quests in order to pass the class.

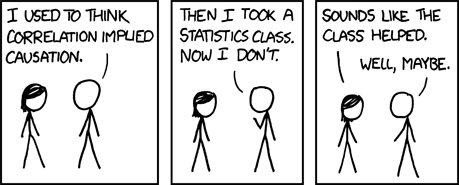
To start a quest (more information will be given in class), you can download the .pdf with all of the questions for that module and begin working through the questions in R. You can complete the module at home in the allotted time frame (usually over the weekend, Friday morning to Tuesday morning) using any of the materials from the course, ***but all students must work individually***.

**Accommodations**

Students who need accommodations are asked to electronically submit their approved accommodations through AU Access and to arrange a meeting during office hours to discuss your accommodations. If you have a conflict with my office hours, an alternate time can be arranged. To set up this meeting, please contact me by e-mail. If you have not established accommodations through the Office of Accessibility, but need accommodations, make an appointment with the Office of Accessibility, 1228 Haley Center, 844-2096 (V/TT).

**Fall 2016| Biostatistics I.**

|  |  |  |
| --- | --- | --- |
| **Date: 2016/** | **Online (before class)** | **In-Class Topics:** |
| 8/16 | Module 000  Module 001 | Syllabus and introduction to R. |
| 8/18 | Module 002 | Basic functions in R, importing data. |
|  |  |  |
| 8/23 | Module 010 | Introduction to statistical inference. |
| 8/25 | Module 011 | Descriptive statistics and basic data manipulation. |
|  |  |  |
| 8/30 | Module 012a  Module 012b | Describing data and visualization. |
| 9/1 | Module 013 | Effect size, variability, and sample size. |
|  |  | **Quest**: “***Goblins of the Shadowmarch***”  *Covers everything up to module 013. Due 9/6.* |
| 9/6 | Module 021  Module 022 | Probability and conditional probability. |
| 9/8 | Module 023a  Module 023b  Module 023c | Discrete probability distributions and expected value.  ***Binomial and Poisson distributions***. |
|  |  |  |
| 9/13 | Module 024  Module 025 | Continuous probability distributions, establishing probability in a continuous distribution. |
| 9/15 | Module 026  Module 027 | Linear combinations of random variables and sampling distributions of variables.  (Or, why the ***normal distribution*** is so cool!) |
|  |  |  |
| 9/20 | Module 030 | Measures of central tendency and measures of spread. The properties of estimators: unbiased, efficient, consistent. |
| 9/22 | Module 031a  Module 031b | Understanding the central limit theorem and the ***t-distribution***. |
|  |  |  |
| 9/27 | Module 032 | Deriving the standard error of mean, sampling distributions of other statistics. |
| 9/29 |  | Flex day for Review. |
|  |  | **Quest**: “***Spire of the Void***”  *Covers everything up to module 032. Due 10/4.* |
| 10/4 | Module 040 | Covariance and the correlation coefficient. |
| 10/6 | Module 041 | Simple regression and understanding residuals.  ***Data = Model + Error*** |
|  |  |  |
| 10/11 | Module 042  Module 043 | Introduction to the *general linear model (GLM).*  ***The F-distribution.*** |
| 10/13 | Module 044 | GLM and categorical predictors. |
|  |  |  |
| 10/18 | Module 045 | Two-level categorical analysis for independent groups. (i.e., ***student’s* t-test**) |
| 10/20 | Module 046 | Two-level categorical analysis for dependent groups.  (i.e., ***paired* t-test**)  ***Virtual meeting only, Dr. Lohse is out of town.*** |
|  |  |  |
| 10/25 | Module 047 | Regression diagnostics: normality of residuals, heteroscedasticity. |
| 10/27 | Module 048 | Regression diagnostics: leverage, outliers, and non-linear relationships. |
|  |  | **Quest**: “***The Dragonslayers’s Daughter.***”  *Covers everything up to module TBD. Due 11/1.* |
| 11/1 | Module 050 | Introduction to multi-variable GLM  (and how it is not the same as multivariate GLM).  ***Virtual meeting only, Dr. Lohse is out of town.*** |
| 11/3 | Module 051 | Multiple regression and comparison between models (ANOVA).  The Omnibus F and coefficients.  Tolerance, variance inflation, and collinearity.  ***Virtual meeting only, Dr. Lohse is out of town***. |
|  |  |  |
| 11/8 | Module 052 | Tolerance, variance inflation, and collinearity. |
| 11/10 | Module 053 | Multiple regression with categorical predictors (***ANOVA***). The difference between the Omnibus F and a contrast.  ***Virtual meeting only, Dr. Lohse is out of town.*** |
|  |  |  |
| 11/15 | Module 054 | Mixing continuous and categorical predictors (***ANCOVA***) |
| 11/17 | Module 055 | Can we really control for “baseline” differences? |
|  |  |  |
| 11/22 |  | Thanksgiving break, no class. |
| 11/24 |  |  |
|  |  |  |
| 11/29 | Module 056 | Interaction and non-linearity in multiple regression. |
| 12/1 | Module 056 | Flex day for review. |
|  |  | **Quest**: ***“Arcane Runes.*”**  *Covers everything up to module TBD. Due 12/6.* |
| 12/6 |  | Final exam period, no class. |
| 12/8 |  |  |



[https://xkcd.com/552/]