Biometry 560

**Instructor:**
Instructor: Jim Fordyce (4-2925) jfordyce@utk.edu
Office: 540 Hesler Hall
Office hours: Monday following class or by appointment
MWF 11:15-12:05 (488 HBB)

**TA:** Zach Marion zmarion@utk.edu (office hours: TBD)

**Communication:** Blackboard. Supplemental reading and assignments will be posted there.

**Text** (Recommended): Quinn, G.P. & M.J. Keough 2009. Experimental Design and Data Analysis for Biologist, Cambridge
Others: Crawley, M.J. 2012. The R Book. Wiley
Dalgaard, P. 2008. Introductory Statistics with R, Springer *
Maindonald, J. & Braun, W.J. 2006. Data Analysis and Graphics Using R, Cambridge
Gelman & Hill 2007. Data Analysis using regression and multilevel/hierarchical models

**Course Objectives:** The use of statistics is ubiquitous in the fields of ecology and evolution. An understanding of statistics is not only important for the design of experiments and analysis of one's own data, but also for the ability to critically read the literature (including literature that you might be asked to review). It is important for scientists to understand *what* various statistical approaches are doing and exactly what hypotheses are being tested, rather than simply having blind faith in the *All Mighty* $p < 0.05$. The goal of this course is to introduce basic statistical approaches in a way that we might understand what question is being addressed. No course (certainly not this one) can cover the whole body of statistical approaches used by researchers – it would take years and, at the end, you would find yourself writing a really thick book. Undoubtedly, at some point, you will have a data set that does not fit the approaches we will be covering. Thus, one objective is for you to feel comfortable with the language of statistics so that you can use books or consultants without too much confusion. Another objective of the course is to encourage (strongly) carefully thinking about the design of a study – from idea, to question, to hypothesis, to experimental design, to analysis, to (finally) interpretation of analysis. In a perfect world (and we all want a perfect world) all these steps should be done before the first datum is collected. One should never collect data and then ask, "How should I analyze this?" We will cover basic parametric and some non-parametric approaches, examine multivariate analysis, and discuss likelihood, Bayesian, and permutation approaches. Ultimately, we want to understand how to use statistics as a tool to understand and advance our research. Have fun – stats are fun.

**Assessment:**
Exercises(6) 50% - These will be given a week before they are due
Mid-term30% - October (middle – it's a take home)
Final 20% - December 9th (12:30-2:30)

The Mid-term and final will consist of analyses and description (including figures, etc. if necessary) of experimental data provided. Two documents, a pdf describing the analyses and your interpretation of the analyses (including figures) AND a document of annotated computer code should be emailed to me at <eebbiometry@gmail.com>.

**Software**:
We will be using the statistical programming language, R.
R is free at http://www.r-project.org/. R has become a standard tool in ecology and evolutionary biology (look through a recent issue of Evolution or Ecology and you'll notice R is commonly used). R also has great graphics abilities. We will like R.

Exercises will be assigned one week before they are due. They will largely consist of annotated computer code. The code should be emailed to me at <eebbiometry@gmail.com> by 11:15am the day they are due.

Tentative Schedule

| | | |
|---|---|---|
| August | | |
| 21 | Introduction | What's the question? |
| 23 | Data | Types of data / intro to JMP |
| 26 | Description of data | Distributions |
| 28 | Description of data | Distributions, summary statistics **** |
| 30 | Description of data | Distributions, summary statistics |
| | | |
| September | | |
| 2 | LABORDAY | |
| 4 | Hypothesis testing & experimental design **** | |
| 6 | Hypothesis testing & experimental design | |
| 9 | Correlation | Relationship between two variables |
| 11 | Correlation | Relationship between two variables |
| 13 | Regression | Linear regression |
| 16 | Regression | Linear regression ******* |
| 18 | Regression | Linear regression |
| 20 | Regression | multiple regression |
| 23 | Regression | multiple regression & model selection |
| 25 | Comparing 2 groups | t-test |
| 27 | ANOVA | One way |
| 30 | ANOVA | Multiple comparison (planned vs. unplanned & corrections) |
| | | |
| October | | |
| 2 | ANOVA | Multi-factor, Partitioning variance components |
| 4 | ANOVA | TBA |
| 7 | ANOVA | Mixed models (Fixed vs. Random effects) |
| 9 | ANOVA | Mixed models (Fixed vs. Random effects) |
| 11 | ANOVA | Mixed models (Fixed vs. Random effects) |
| 14 | ANOVA | Nested models |
| 16 | No class | Mid-term exam |
| 18 | No class | Mid-term exam |
| 21 | ANOVA | ANCOVA |
| 23 | Related samples | Paired t-test |
| 25 | Related samples | Repeated measures ANOVA |
| 28 | Multivariate | MANOVA |
| 30 | Multivariate | MANOVA |
| | | |
| November | | |
| 1 | Multivariate | MANOVA |
| 4 | Multivariate | Discriminant Function Analysis |
| 6 | Multivariate | Discriminant Function Analysis |
| 8 | Multivariate | Discriminate function analysis |
| 11 | Categorical data | odds ratio, Pearson's Chi-square, Fisher's exact test |
| 13 | Categorical data | odds ratio, Pearson's Chi-square, Fisher's exact test |
| 15 | Categorical data | odds ratio, Pearson's Chi-square, Fisher's exact test |
| 18 | Ordination | PCA,NMDS and others |
| 20 | Ordination | PCA,NMDS and others |
| 22 | Ordination | PCA,NMDS and others |
| 25 | | |
| 27 | DISCUSSION | |
| 30 | Thanksgiving | |
| | | |
| December | | |
| 2 | Discussion | |
| | | |
| 9 | Final (12:30-2:30) | |