BIOSTATISTICS (BSC 490 & 420.27)

Fall 2019

Instructor: Steven Juliano	Office: 335 FSA	Phone: 438-2642
e-mail: sajulian@ilstu.edu	Web: https://about.illinoisstate.edu/sajulian/pages/default.aspx	
Lecture: MWF 9:00 – 9:50 AM FSA 129	**Laboratory:** Th 1:00 - 3:50 PM SLB 121	**Office Hours**: F 10:00-11:00AM, T 1:00-2:00 PM, & by Appointment

Teaching Assistant: Ian Rines	Office: 143 FSA	Office hours: T 1:30 – 3:00 PM **in SLB 121**, and by appointment
e-mail: igrines@ilstu.edu	Phone: 438-5438	

TEXTS: ***Experimental design and Data analysis for Biologists.*** 2002. GP Quinn & MJ Keough. 2002 Cambridge Univ. Press
Biostatistics Manual. 2018. S. A. Juliano [Files provided]

COURSE GOALS: This course is an introduction to applied statistics. The ideas and methods discussed will be those most relevant to biologists in general. You will acquire a working knowledge of basic statistical methods, and will be able to determine which procedures are most appropriate for a given circumstance. All of the statistical techniques relevant to biologists cannot be covered in one semester, however, once you have mastered the material in this course, you will be better equipped to understand and use more advanced statistical methods.

In the laboratory for this course you will gain experience in use of the *SAS 9.4* computer package for statistics. There are a number of good statistical packages available (e.g., *R, SPSS, JMP*), and some of you may already know how to use some of these. I will give examples and explain how to do things in *SAS*, and all of you will do the assignments using *SAS*. By learning enough about general aspects of statistical computation and interpretation, you will be able to generalize to other packages if you so choose.

GRADE: Although BSC 490 and 420.27 are nominally two different courses, in reality they are part of a single course. You will receive the same grade in both courses. Course grades will be determined as follows

Course Component		Percent of Final Grade
Exam I		17.5%
Exam II		17.5%
Cumulative Final Exam	(in class)	20.0%
	(take home)	10.0%
11 Homework Assignments		35.0%

Total Score	Yields Final Grade
≥85%	A
75 - 85%	B
65 - 75%	C
55 - 65%	D
<55%	F

Homework will involve using *SAS* to do analyses of statistical problems. Specific instructions on how to write up the report and what to include will be provided. In addition to homework assignments, I will give sets of study problems that will **not** be graded, but which will help you to learn the material. Parts of exams will be open book/open note/use the computer and will contain analysis questions. Some exams will include take-home sections, which will be similar to the homework assignments. Homework and take-home exams **will not be accepted late**. Turning in an incomplete homework assignment will produce a **much better** grade and learning experience than will turning in nothing at all. The 11 homework assignments, and **tentative** due dates are:

Assignment	Topics	Tentative Due Date (by 11:59 PM)
1	Summary Statistics	Thursday 29 August
2	Simulating data & Generating Random Numbers	Thursday 5 September
3	One and two sample tests	Thursday 12 September
4	One-way Fixed effects Analysis of Variance	Thursday 26 September
5	One-way Random effects Analysis of Variance	Thursday 3 October
6	Two-way Factorial Analysis of Variance	Thursday 17 October
7	Mixed model Analysis of Variance	Thursday 24 October
8	Two-stage nested Analysis of Variance	Thursday 7 November
9	Generalized Linear modesl for ANOVA	Thursday 14 November
10	Linear & Multiple Regression	Thursday 21 November
11	Analysis of Covariance	Thursday 5 December

COURSE OUTLINE & READINGS

Topic	Reading Assignment (Quinn & Keough 2002 *ED&DAFB*)
Introduction	pp. 1-7
Kinds of variables	
Frequency distributions	pp. 59-64
Random samples & populations	pp. 14-15
Descriptive Statistics	pp.14-17
Location and dispersion	
Relationships	pp. 72-75
Statistics vs. parameters	
Probability	pp. 7-13
Concepts	
Distributions	pp. 9-13
Normal distribution	pp. 17-19
Estimation	
Point vs. interval estimates	pp.19-25
t-distribution	
χ^2 distribution	
Hypothesis testing	pp. 3-5; 7
Null and alternative hypotheses	pp. 32-36
Assumptions	
Type 1 and type 2 errors	pp. 42-45
t-tests	pp. 33-39; 45
One tailed vs. two tailed tests	p. 37
Failure to meet assumptions	
Examples and consequences	
Transformations	pp. 64-68
Nonparametric tests	pp. 46-47
Randomization tests	pp. 45-46
Analysis of variance	
Assumptions and the model	pp. 171; 191-194
One way ANOVA	pp. 173-188; 190-191
Orthogonal contrasts	pp.196-199
Multiple comparisons	pp. 48-50; 199-201

EXAM I – <u>tentatively</u>: 19 September, during the lab period

Analysis of variance	
Fixed vs. random effects	pp. 176; 186-187; 188-191
Two way ANOVA	pp. 321-363
Factorial designs	pp. 221-241
Followup tests	pp. 251-261
Unbalanced designs	pp. 241-249
Nested designs	pp. 208-221

Topic	Reading Assignment (Quinn & Keough 2002 *ED&DAFB*)
Analysis of Variance	
Blocked designs	pp. 262-300
Failure of assumptions	pp. 249-250
Nonparametric analogs of ANOVA	
Assumptions	
Kruskal-Wallis test	pp. 195-196
Friedman's test	pp. 284-285
Follow up tests	
Generalized linear models	pp. 359-379
Experimental design	pp. 155-172
Randomization	
Replication	
Control	
Experimental units	
Complex designs; split plots	pp. 301-338

EXAM II – <u>tentatively</u>: 31 October during the lab period

Regression	
Assumptions	pp. 92-94
Reasons for doing regression	p. 78
Linear regression	pp. 77-94
Failure to meet assumptions	pp. 94-100; 104-106
Geometric mean regression	pp. 100-104
Comparing regression lines	pp. 90-92
Analysis of covariance	pp. 339-358
Polynomial regression	pp. 133-135
Multiple & stepwise regression	pp. 111-143; 153-154
Correlation	
Assumptions	pp. 75-76
Relationship to regression	pp. 72-76; 106
Partial correlation	
Nonparametric correlation	pp. 76-77
Frequency data	p. 380
Proportions	
Goodness of fit	pp. 381
χ^2 vs. likelihood ratio	pp.394-395
Contingency tables	pp. 381-393
Fisher's exact test	
Miscellaneous Methods	
Combining probabilities	pp. 50-51

CUMULATIVE FINAL – As scheduled by the University;
Take home part due: 11:59 PM Wed. 11 December

BIOSTATISTICS (BSC 490 & 420.27)

Fall 2019

LABORATORY SCHEDULE

Date	Laboratory Topics
22 August	Introduction to *SAS*; Data entry; Data manipulation; Summary Statistics
29 August	Probability: Generating & working with random numbers
5 September	One & two sample *t*-tests & Wilcoxon tests
12 September	One way ANOVA (fixed); Testing assumptions; Contrasts; Multiple comparisons; Nonparametric
19 September	**Exam I**
26 September	One way ANOVA (Random); Estimating variance components
3 October	Two Way factorial ANOVA; Interactions;
10 October	More Two Way ANOVA; Unbalanced designs; Least Squares Means for multiple comparisons; Fixed, random, and mixed model ANOVA in *SAS*
17 October	Mixed Model ANOVA: Comparing GLM, VARCOMP, MIXED
24 October	Two Stage Nested ANOVA; Estimating variance components
31 October	**Exam II** 
7 November	Generalized linear models for ANOVA: GENMOD, GLIMMIX, and comparison to GLM
14 November	Linear & Multiple regression; Residuals; Testing assumptions
21 November	Analysis of covariance; Test homogeneity of slopes; Estimate separate slopes  ***Steve in Costa Rica with Rainforest Ecology; Ian in charge***
28 November	*Thanksgiving break*
5 December	Loose ends/Review

Notes on *SAS*

It is **essential** that you read the assignments **before coming to lab.** This is particularly true for the first two weeks, when you will be learning about how to use *SAS*. Learning how to use *SAS* is **vital** to your success in this course, your sanity, and probably your success as a research student.

I no longer require a *SAS* manual as a reference. Instead I will show you how to look up the information you need in the very extensive documentation that is included with installed *SAS* software.