Professor: Corey Devin Anderson, Ph.D. (Evolution, Ecology, and Population Biology)

Preferred salutation: “Dr. Anderson”
E-mail: coreanderson@valdosta.edu

Face-to-face classes generally have the following:
a scheduled meeting place and
a scheduled time and day(s) of the week.

Days and time: Tues/Thurrs, 2:00 to 2:50 PM.

Days and time: Tues /Thurs, 3 to 4:50 PM.

Wednesday 08 Dec: 2:45 to 4:45 PM.

: Tuesday 9:30 AM to 11:30 AM

The lectures provide a survey of key concepts, topics, and tests in biostatistics; the labs are intended to reinforce the lecture material, as well a

Course overview

This is an upper division course on the statistical analysis of biological data ("biostatistics").

The catalog lists BIOL 1107, BIOL 1108, MATH 1112 or 1113, and MATH 1401 (old 2620) as prerequisites. I interpret this to mean that the course is upper division, that you have some basic math skills (precalculus or otherwise), and some sort of background in statistics. That said, I recognize that most students (even the graduate students) taking this course likely have a very rough understanding of statistics (and probably a fear of mathematics in general), even if they have had Math 1113 and Math 2620. By virtue of the course topic (biostatistics) you should expect this course to be quantitative (and computational), but I do not consider the math in this course to be advanced (so don't psyche yourself out!).

Statistics are

In the modern era, computers have facilitated the application of statistics by scientists and some methods (such as permutation testing) would be virtually impossible to implement without the aid of a computer. There are many statistical software packages available and, for many years, there was a trend towards programs with simple-to-use graphical user interfaces ("GUIs": Macintosh/Windows-style menus that you can point and click with a mouse…or, now, with your finger). Around the year 2000, many academics started to abandon proprietary statistical software with GUIs in favor of free, open-source statistical programming platforms. The most popular of these platforms is called "R".

R is a descendant of the statistical programming language "S" and has many advantages over proprietary statistical software packages that use graphical user interfaces. When you download R (for free), there are many functions that are built into the base distribution, but additional functionality can be contributed as packages of functions (developed by other R users) that can be downloaded and applied. Since R is open-source, anybody can see (or modify) the source code. Moreover, R is also a programming language, which means that you can write your own scripts (within which you can call functions from the base package or other contributed packages). This makes R extremely powerful because you are not limited to the choices on a menu; almost anything is possible.

The downside of R is that there is a steeper learning curve and, while anything is possible, it is not always obvious or easy to do certain things (such as custom graphics). On the flip side, once you learn the basics of R, some things are far easier (and faster) to do in R than with GUI-based software and, once the code is written, outputs can be reproduced much faster than what is possible with a GUI (especially if you have a lot of repetitive/batch procedures).

R is a statistical programming language and is intended for doing statistical analysis. Some functions that can be called in R may be written in other languages, such as "C". R is similar to "Python" (both are "interpreted" languages) and both R and Python are the most popular platforms for "Data Science." In the research sector and industry, there is a huge demand for people with R and Python programming skills.

In this course we will focus on statistical programming, which includes an introduction to some basic concepts in computer programming, such as how to write a "function" and a "loop". I recognize that most biology students do not have a strong background in computer science, so part of the goal (beyond teaching you the state-of-the-art in statistics) is to improve your computational skillset. This is a very important skill in the day and age of "big data." Many data sets (such as genomic data sets and large medical research studies) are simply too large for point and click interfaces; computer programming is now a required skill in most realms of biology, including medical science.

While some people have a better natural aptitude for computer programming, nobody is born knowing how to use R or how to write a loop. You get better at doing this sort of thing by practicing, a lot (I would suggest *every day*). Programming is a one-step forward, two-steps backwards process: solving one problem often leads to another…but if you are determined and keep battling, eventually you will succeed…and that's how you get better.

I have had students cry and moan about R, but then get into graduate school or get a job because they had training with R (and inevitably they thank me for it). I am not teaching you R because it is free, I am teaching you R because it is the most powerful platform for doing statistical analysis, period.

I use a rank-based (or “stack rank”) grading system; this means that you will be evaluated based on how well you perform (in terms of your point total) relative to other students in the class.

When possible, I like to use natural breaks in the point distribution to determine letter grades. For

Required text:

1) The analysis of Biological Data by Whitlock and Schluter (3rd edition); the publisher is W.H. Freeman (Macmillan Learning).

https://www.macmillanlearning.com/college/us/product/Analysis-of-Biological-Data/p/131922623X#:~:text=The%20Analysis%20of%20Biological%20Data,of%20statistics%20for%20biology%20students.&text=These%20include%20new%20calculation%20practice,medical%20and%20human%20health%20research.

This is an excellent introductory textbook, and most of the lecture material will be follow the topics in the book. I have chosen this book because it is easy to read (relative to most statistics texts), it has lots of practice problems, and it does an excellent job at explaining some of the more challenging concepts.

The 1 hour and 50 minute time slot for labs means that field based data collection is not feasible. Some semesters, we will do a class fishing trip (to compare types of bait and examine random effects of different fishermen). This will probably not happen in fall 2020 because of COVID-19 precautions.

Do NOT cheat on exams. You will receive a zero on the exam and will be reported to the Dean of Undergraduate Academic Affairs. I consider copying of problem sets/computer code to be cheating! Use must sign the agreement on code plagiarism to be formally matriculated into this course.

Some unit exam questions may require a calculator….so remember to bring one to the unit exams.

Unless you have special permission,
. Students who have cell phones out during exams will receive a zero on that exam. Any student caught photographing an exam will get an automatic "F", and will also be banned from retaking the course with Dr. Anderson.

I prefer that my lectures and labs not be recorded (especially without my consent), but if you feel as if you need to record a lecture, please place your recording device in the front of the classroom, so that I am aware that I am being recorded.

Students requiring classroom or testing accommodations because of documented disabilities should discuss their needs with the instructor at the beginning of the semester. Students not registered must contact the Access Office, Farber Hall, Phone; 245-2498. Website: http://www.valdosta.edu/access/
For some students, the presence of a medical condition places them at high risk for COVID-19. These students can use the online form to submit documentation of the condition to the Access Office to ensure confidentiality.

https://www.valdosta.edu/student/disability/forms/request-for-covid19-course-modification.php

The Access Office will then contact the advisor and department to indicate the receipt of documentation that supports the request for course substitutions or appropriate alternative assignments and virtual access to lectures.