

Applied Biostatistics for Entomology and Ecology
(Statistical Techniques in Entomology)

ENT 535 – 3 Credits

Spring 2022

Course Description: As quantitative scientists we must responsibly conduct the scientific process from start to finish. However, most statistics courses focus solely on mathematical theory and leave students without the skills to appropriately frame questions in a reproducible manner. Through this course, students will 1) participate in discussions surrounding scientific inquiry, experimental design, open science and reproducible research, 2) gain hands-on experience with software use, data analysis and interpretation in R, and 3) follow through the process of scientific inquiry by learning to effectively communicate results using presentations and scientific writing.

Instructor:
Dr. Sara Hermann
(she/her/hers)
slh@psu.edu
Department of Entomology

Teaching Assistant:
Lillian Germeroth (Lilly)
(she/her/hers)
Lzg5389@psu.edu
Ecology MS Student

Class Time: Tuesday and Thursday 1:35-2:50

Class Location: Huck Life Sciences 007 (this room is spacious & will allow for social distancing!)

Office Hours: Office hours will be held on an as-needed basis, by appointment. If you desire to meet with the instructors and/or TA's, please send a message to both of us to arrange the time.

Prerequisites: None! However, there are some basic course requirements:

- 1) Math proficiency: Although there are no prerequisites, I do expect students to have basic math proficiency and hopefully some experience with basic statistical theory (even if you have to reach faaaar back for it!).
- 2) Class attendance: Be sure to attend class and please, be on time. Inability to be timely will affect your participation grade.
- 3) Don't be afraid!!!! Math and statistics tend to elicit a fear response in folk. Do not let that influence your experience in this class. No matter the stage, when you are in an career using statistics you will always be learning, troubleshooting and even experiencing failure (yes – even me!). We will work together to normalize this and find a way to work THROUGH it!

Learning Outcomes and Course Goals: As a budding scientist, it is crucial that you all are able to internalize the process of scientific inquiry from beginning to end. In this class, you will learn to efficiently and effectively design experiments, collect and manage data, analyze, visualize and interpret data in a reproducible and effective manner. Rather than theory, we will focus on practical skills that you, as a graduate student, will use. Importantly, we are going to build a foundation that will allow you to improve your applied statistical toolbox to the point that you can approach more complex statistical thinking and analysis in future endeavors. Comfort and confidence is key here – we can work towards that!

Through the course we will move progressively through the following framework:

Part 1: The Scientific Method and Statistical Thinking

How to frame a question, derive hypotheses and design an appropriate experiment

- Hypothesis testing – Biological versus Statistical Hypotheses
- What is the point of the study? Understanding Variables X and Y
- Manipulative versus Natural Experiments
- Snapshot versus Trajectory Studies
- Replication, Randomization, Independence and Confounding Factors

Part 2: Data – Collection, Cleaning and Organization

How to handle your data to make your work efficient and reproducible, identify common problems with data

- Best practices for spreadsheets and lab notebooks – Intro to Metadata
- Data storage – Use of public or private repositories
- Cleaning up messy data – errors, missing data, outliers, internal checks
- Responsible use of grey data or collaborative datasets
- Preparing data for analysis

Part 3: Analysis – Programming in R, Version Control, Statistical Tests

Learning the R language, practice, work through what different analysis can do in small group presentations and explore different statistical analyses with hands-on coding

- Intro to scripting in R and version control with R and Github
- Statistical Distributions
- Summary Statistics and Data Visualization
- Means, variance, standard error, standard deviation, confidence intervals
- T-test, paired and unpaired
- Chi-square test and non-parametric tests
- Intro to Regression analysis – Linear models and Generalized Linear Models
- ANOVA/ANCOVA

Part 4: Communication – Interpreting Results, Graphing and Presenting Data

Using technology to make our work accessible to others and efficiently convey results in scientific meetings and publications

- Plotting data and making publication-ready figures
- Visualization for outreach and public communication
- Creating effective extension bulletins with data
- Writing up statistics for a scientific paper
- Public speaking and preparing talks
- Social media as a means to disseminate ideas and results

Course Format: We will be meeting in-person this semester unless the University decides to move classes to remote format. If we have multiple COVID cases in the class (which I receive confidentially), I will move us to remote format for one week. In addition, I have two young children in school and daycare – if their classes should close due to COVID-19 or if they test positive for COVID-19, I will need to move to a remote format until the established quarantine window is complete. I will be completely transparent and quick to inform you all of any changes to the class format that will occur or any potential exposures that I may have had outside of the University contact tracing system that could affect you.

For more info on COVID-19 and University policies, practices and response to the pandemic, please visit: <https://virusinfo.psu.edu> or <https://ist.psu.edu/current/coronavirus>

Masking and COVID-19: Penn State University requires everyone to wear a face mask in all university buildings, including classrooms, regardless of vaccination status. ALL STUDENTS MUST wear a mask appropriately (i.e., covering both your mouth and nose) while you are indoors on campus. This is to protect your health and safety as well as the health and safety of your classmates, instructor, and the university community. Anyone attending class without a mask will be asked to put one on or leave. If you do not have a mask, we will provide one for you at the door. Instructors may end class if anyone present refuses to appropriately wear a mask for the duration of class. Students who refuse to wear masks appropriately will face disciplinary action for Code of Conduct violations. If you feel you cannot wear a mask during class, please speak with your academic advisor immediately about your options for altering your schedule. We ask you to please respect this policy for your health and well-being, for that of your peers, for us and for my children's safety.

Course Text: There is no formal course text, however, supplemental readings will be provided on an as needed basis. There are several fantastic online statistical resources as well as many useful textbooks. However, I find that there is no “one fits all” solution for the content I want to cover with you all. For any readings, I will supply scanned text or distribute electronic copies.

Course Materials: We will use computers for most class meetings. Therefore, class will be held in a traditional computer lab. In doing so, we will never exclude any student that does not have a laptop computer from participating and learning this material. However, if you do have a laptop computer that you would prefer to use, make sure it is fully charged prior to class to limit disruptions.

Attendance Policy: This course benefits from the rich discussions that student involvement provides. Therefore, we will keep track of attendance each class period. Religious observances

are not counted as absences, though observing students must inform the instructor in advance that they will not be present. Official university activities are excused absences if the student informs the instructor in advance and provides appropriate paperwork. Absence due to sickness does not require a doctor's note, but it is the student's obligation to inform the instructor promptly and to bring to the instructor's attention extended medical absences as soon as possible. It is the responsibility of the absent student to catch up on any missed material and do any make-up work required by the instructor. If absences occur in response to isolation or quarantine, please let your instructors know, we will keep your health status confidential.

Academic Integrity Statement: Academic integrity is the pursuit of scholarly activity in an open, honest and responsible manner. Academic integrity is a basic guiding principle for all academic activity at The Pennsylvania State University, and all members of the University community are expected to act in accordance with this principle. Consistent with this expectation, students should act with personal integrity, respect other students' dignity, rights and property, and help create and maintain an environment in which all can succeed through the fruits of their efforts. Academic integrity includes a commitment not to engage in or tolerate acts of falsification, misrepresentation or deception. Such acts of dishonesty violate the fundamental ethical principles of the University community and compromise the worth of work completed by others (see [Faculty Senate Policy 49-20](#) and [G-9 Procedures; Code of Conduct and Student Conduct Procedures](#)).

Academic Integrity Guidelines for the College of Agricultural Sciences can be found at <http://agsci.psu.edu/students/resources/academic-integrity>.

A lack of knowledge or understanding of the University's Academic Integrity policy and the types of actions it prohibits and/or requires does not excuse one from complying with the policy.

Disability Statement & Accommodations: Penn State welcomes students with disabilities into the University's educational programs. Every Penn State campus has an office for students with disabilities. Student Disability Resources (SDR) website provides [contact information for every Penn State campus](#). For further information, please visit [Student Disability Resources website](#).

In order to receive consideration for reasonable accommodations, you must contact the appropriate disability services office at the campus where you are officially enrolled, participate in an intake interview, and provide documentation: <http://equity.psu.edu/student-disability-resources/applying-for-services/documentation-guidelines>. If the documentation supports your request for reasonable accommodations, your campus's disability services office will provide you with an accommodation letter. Please share this letter with your instructors and discuss the

accommodations with them as early in your courses as possible. You must follow this process for every semester that you request accommodations.

I also stand in strong support of students who have children and understand that childcare can occasionally get cancelled due to events out of our control, including weather, child illness and now – COVID-19 quarantine. Don't hesitate to reach out to me directly to discuss options to attend class or receive missed materials if these situations arise.

Counseling and Psychological Services Statement:

Many students at Penn State, and in academia, face personal challenges or have psychological needs that may interfere with their academic progress, social development, or emotional wellbeing. The university offers a variety of confidential services to help you through difficult times, including individual and group counseling, crisis intervention, consultations, online chats, and mental health screenings. These services are provided by staff who welcome all students and embrace a philosophy respectful of clients' cultural and religious backgrounds, and sensitive to differences in race, ability, gender identity and sexual orientation.

- [Counseling and Psychological Services at University Park \(CAPS\)](#): 814-863-0395 (Monday to Friday 8:00 am to 5:00 pm)
- Penn State Crisis Line (24 hours/7 days/week): 1-877-229-6400
- Crisis Text Line (24 hours/7 days/week): Text "LIONS" to 741741
- Call Penn State Police at 814-863-1111 or call 911

*** I must say that I have been VERY disappointed to hear the experiences of some students with on-campus psychological services. Whether that be because of extremely long wait times or less than stellar interactions with psychologists. Please know that we also have long wait lists in the broader community, but there are options to get help and you shouldn't let one bad experience or a long wait affect your journey to get help. If you need SDR information or SDR accommodations, please contact them as soon as possible to avoid delaying your care ***

Educational Diversity & Equity/Report Bias Statement:

As your instructor, I strive to foster a diverse and inclusive environment for students, teaching assistants, other faculty, and staff across campus. Acts of intolerance, discrimination, or harassment due to age, ancestry, color, disability, gender, gender identity, national origin, race, religious belief, sexual orientation, or veteran status will not be tolerated and will be reported through Educational Equity via the [Report Bias webpage](#). I also encourage students who experience discrimination or harassment outside of the classroom to report it. In line with the

commitment of the university to support diversity in its myriad forms, in our class we will, "... embrace individual uniqueness, foster a culture of inclusion that supports both broad and specific diversity initiatives, leverage the educational and institutional benefits of diversity, and engage all individuals to help them thrive." - Penn State Equity & Diversity Website.

As your instructor, I promise to value your opinions and communicate in a respectful manner. I honor the uniqueness of you all as my colleagues and as my students. I promise to take ownership of my words and actions, apologize for my mistakes and to keep confidential any personal or professional discussions. Our work and community are strengthened by our rich variety of backgrounds and perspectives. I am committed to providing an atmosphere for learning that respects that diversity and encourages inclusion.

Exams

In this class there will be no traditional exams. In lieu of exams, we will have a semester long project where you work through some of your own data, or data found on the web. The project will culminate in a final presentation, a written report and peer grading.

Course Grading

<u>Class attendance, discussion, and participation</u>	20%
<u>Homework Assignments</u>	30%
<u>Group Presentations on Statistical Tests</u>	20%
<u>Final Project</u>	30%
<u>Presentation</u>	
<u>Written Report</u>	
<u>Project Analysis</u>	
Total	100%

Letter grades will be assigned based on the points that you earn. Class participation, attendance and other factors can be used to adjust your overall grade. In general, A = 93-100%; A- = 90-92%; B+ = 87-89%; B = 83-86%; B- = 80-82%; C+ = 77-79%; C = 70-77%; D = 60-69%; F = <60%.

- **Class Attendance, discussion, and participation:** attendance and participation will be crucial in this class. Once we start computer work, we will all be actively coding and/or following along with code in class. Active participation will ensure you succeed in the class!
- **Homework Assignments:** Homework will be distributed on the day they are listed in the syllabus and due the following class period. They are meant to serve as practice and confirmation

of your understanding of the material. They will be graded for completion and effort and feedback will be provided.

- **Final Project:** Students will be tasked to identify a question – either directly related to their graduate research or using a dataset from the web – and work through the hypotheses, relevant statistical tests and analysis of the data. Once analysis is completed, students will write up a mock manuscript (with special emphasis on methods, results and interpretation), make publication ready graphs of the results and present the work as a final project. Some class time will be dedicated to the projects and much feedback and guidance will be given from the instructors along the way.

***This syllabus is subject to change. Any changes to the syllabus will be distributed through electronic notification ***

Course Schedule and Assignments:

Week	Date	Section	Topic	Assignment (Red Assigned this date, black due this date)
1	11 Jan		Intro to the Class – meet your instructor and introduce yourselves	
	13 Jan		Continued Introductions	
2	18 Jan		Continued Introductions	
	20 Jan		The Origins of Statistics – issues in diversity, equity, and inclusion	
3	25 Jan		Statistical Thinking and Hypothesis Testing: How to formulate a good question	Reading: How to do Ecology HW: 1
	27 Jan		Experimental Design	Reading: How to do Ecology HW: 2
4	1 Feb		Managing and Curating Data, Cleaning and Organizing Data	Make a mock Data Sheet - HW: 3
	3 Feb		Intro to tools in Open Science and Reproducible Research	
5	8 Feb		Intro to Probability	Get R & RStudio & GitHub
	10 Feb		Probability Distributions, and a Hint of Transformation	
6	15 Feb		R Bootcamp #1 – getting comfortable with variables, vectors and functions	HW: 4
	17 Feb		R Bootcamp #2 – Getting data into R & R Markdown!	HW: 5
7	22 Feb		R Bootcamp #3 – Summary Statistics: Measures of Location and Spread – base plotting	HW: 6
	24 Feb		G-test; Chi-Square – Do Frequency Distributions Differ?	HW: 7
8	1 Mar		Paired T-test; Wilcoxon signed rank test – Do observations from two groups differ (Paired data)	HW: 8
	3 Mar		Unpaired T-test; Mann-Whitney U Test – Do observations from two groups differ (unpaired data)	HW: 9
9	8 Mar	Spring Break		
	10 Mar			

10	15 Mar		One-way ANOVA – Do observations from more than two groups differ	HW: 10
	17 Mar		Special Topics: Non-linear dose response curves	
11	22 Mar		Linear Regression – Is there a ‘Cause and Effect’ relationship between two variables?	
	24 Mar		ANCOVA – Cause and Effect, take 2	
12	29 Mar		Special Topics: Random Forest	
	31 Mar		Project Time	
13	5 Apr		Special Topics: PERMANOVA	
	7 Apr		Model Building and Model Selection	
14	12 Apr		How to Write Science	
	14 Apr		Advanced Data Visualization (ggplot): Presentation and Publication Ready Plots	
15	19 Apr		Project Time	
	21 Apr		Science Communication to Diverse Audiences	Peer-review of manuscript drafts
16	26 Apr		Student Presentations	Peer-grading 1
	28 Apr		Student Presentations	Peer-grading 2
17	3 May	Finals Week (Student Presentation Spillover)		
	5 May			