Course Syllabus – CBB 400

Course Information

Course Number: CBB 400 WI24
Course Name: Introduction to Biostatistics
Term: WI 2024
Start Date: 01/02/2024
End Date: 03/22/2024
Credits: 3.0

Meeting Days / Times

Tuesdays and Thursdays, 2:00-3:30pm PT / 5:00-6:30pm ET
(See Calendar in Canvas for the most up-to-date schedule.)

Location

CA: Graduate Office Dining Room (Hazen Theory Building)
FL: C212

Course Managers

Role	Last Name	First Name	Email Address
Course Director	Waalen	Jill	jwaalen@scripps.edu
TA	Chen	Yuting	YutingChen@scripps.edu
TA	Shafiei Neyestanak	Mahdi	mshafiei@scripps.edu

Course Description

The focus of this course will be on basic concepts, principles, and practices in statistical analyses. Fundamentals of probability theory, distribution theory, and parameter estimation techniques will be discussed as the bases for understanding the application of statistical tests. The majority of the course will be dedicated to the application of statistical tests commonly encountered in scientific fields relevant to the students. This will include hypothesis testing and model building strategies, including assumption checking (e.g., normality, outliers, scale), graphical methods (e.g., scatter plots, box plots), and model diagnostics (e.g., serial correlation, normality). The lectures are given in sequence with each lecture building on the material covered in prior lectures.

The free on-line statistical package R will be used as an integral part of the lectures and coursework. No prior experience with R is required. Basic R programming will be learned as a practical complement to statistical concepts as they are covered. R code required to perform simulations and other procedures that demonstrate and apply these concepts will be provided for the students and discussed during regular class sessions. Optional Friday workshops will allow students to get individualize help with using the program. This course originates as an education and training course offered from the Scripps Research Translational Institute, which is dedicated to improving medicine and human health via biomedical discoveries. SRTI is a member of the NIH-funded Clinical and Translational Science Award (CTSA) Consortium.

Program Learning Outcomes

By the end of the program, students will have accomplished these objectives:
PLO1: Original Research – graduate students are expected to develop the skills critical for generating high-quality research output. This would include absorbing, recalling, and contextualizing scientific knowledge, evaluating scientific information and data, creating testable hypotheses and investigating hypotheses, mastering scientific tools and techniques, displaying ethical behavior, and receiving and giving feedback.
PLO2: Communication – graduate students are expected to demonstrate the oral, written, and media skills to effectively communicate the impact of a study or a body of work to the greater scientific community and to the public at large using a number of methods.
PLO3: Critical Thinking – graduate students are expected to develop a self-directed process to analyze information, form opinions or judgments, and use this process to improve the quality of their scientific thoughts, navigate problems, and make informed decisions.
PLO4: Intellectual Curiosity – graduate students are expected to acquire the capacity to build their intellectual curiosity and demonstrate problem solving approaches that serve their professional growth and ability to impact a field.
PLO5: Career and Professional Development – graduate students are expected to develop a variety of transferable skillsets throughout their graduate experience, including management and leadership, inclusiveness, resilience, scientific rigor, collaboration, accountability, time management, teamwork, networking, and career planning.

Course Learning Outcomes

Upon completion of this course students will be able to:
CLO1: Understand the concept of a random variable
CLO2: Be able to describe basic probability models
CLO3: Have a working knowledge of discrete and continuous distributions for random variables
CLO4: Implement basic parameter estimation techniques
CLO5: Implement basic hypothesis testing techniques
CLO6: Conduct analysis of contingency tables
CLO7: Set up the design and use of the analysis of variance (ANOVA)
CLO8: Implement regression models for many types of random variables
CLO9: Utilize statistical tools to analyze collected scientific data

CLO10: Utilize statistical software to assist in data analysis
CLO11: Understand statistical power and calculate sample sizes

Background Preparation (Prerequisites)

There are no prerequisites for this course. The course is intended to be an introduction to biostatistics for students not previously exposed to, or with limited exposure to, statistics or biostatistics. The course is intended to provide a basic understanding of biostatistics concepts and applied techniques. The course is suitable for laboratory biologists or clinical scientists who wish to learn biostatistics principles such as parameter estimation, hypothesis testing, and model building. Students are not expected to become experts in any of the topics covered during the course. Rather, students are expected to actively participate in lectures and discussions; ask questions and make comments on the topics being discussed; and think critically about the fundamentals of data analysis. Importantly, students should focus on these points in the context of their own research and their current and future data analyses as practicing scientists. Students are also expected to use the R statistical program (code provided) in the ways demonstrated in class to perform simulations that demonstrate concepts and to apply statistical techniques.

Course Materials

Required: *Open Intro Biostatistics*, 1st edition (available free online at https://www.openintro.org)

Course Format

This course is offered in a traditional classroom setting. Class sessions will involve a lecture on the topic listed as well as demonstrations using R statistical software. Students are expected to bring laptops to the lectures to be able to apply the R programming learned during class. Homework covering the theoretical and/or applied concepts will be assigned weekly. Most homework will require use of R. Alternative statistical software may be used only with prior approval from the Course Director. Please remember to turn off cell phones during class and avoid bringing other materials that may be disruptive.

Workshops

Need for Friday workshops will be determined during the course of the semester. All workshops will be optional.

Attendance Statement

Students are expected to attend all classes and to be on time. Students who are unable to attend class must seek permission for an excused absence from the Course Director. Unapproved absence or late attendance for 3 or more classes may result in a lower grade being

assigned or an incompletion recorded. If a student has to miss a class, s/he should arrange to get notes from a fellow student. Missed in-class assignments will not be available for make-up.

There will be a number of opportunities for class discussion, and thus students must come to class and be prepared to participate. Along these lines, if something is not clear, please ask. In general, students may feel free to interrupt the lecturer with questions, unless the lecturer explicitly states a preference for holding questions until the end. Students may also email the Course Director with questions or arrange an in-person meeting.

Scientific and Professional Ethics

The work you do in this course must be your own. Feel free to build on, react to, criticize, and analyze the ideas of others but, when you do, make it known whose ideas you are working with. You must explicitly acknowledge when your work builds on someone else's ideas, including ideas of classmates, professors, and authors you read. If you ever have questions about drawing the line between others' work and your own, ask the course professor who will give you clear guidance. Exams must be completed independently. Any collaboration on answers to exams, unless expressly permitted, may result in an automatic failing grade and possible expulsion from the Graduate Program.

Technology Requirements and Support

For issues related to Canvas, please contact the Graduate Office by email at: gradprgm@scripps.edu or by phone at: 858-784-8469.

Course Grading

Grading is in accordance with the academic policies of the Skaggs Graduate School. The breakdown of grading is as follows:

- Homework: 40%
- Midterm: 30%
- Final: 30%

Letter Grade	Percent	GPA	Description
A	93-100	4.00	Outstanding achievement. Student performance demonstrates full command of the course subject matter and evinces a high level of originality and/or creativity that far surpasses course expectations.
A-	90-92	3.67	Excellent achievement. Student performance demonstrates thorough knowledge of the course subject matter and exceeds course expectations by completing all requirements in a superior manner.

B+	87-89	3.33	Very good work. Student performance demonstrates above-average comprehension of the course subject matter and exceeds course expectations on all tasks as defined in the course syllabus. There is notable insight and originality.
B	83-86	3.00	Satisfactory work. Student performance meets designated course expectations and demonstrates understanding of the course subject matter at an acceptable level.
B-	80-82	2.67	Marginal work. Student performance demonstrates incomplete understanding of course subject matter. There is limited perception and originality.
C+	77-79	2.33	Unsatisfactory work. Student performance demonstrates incomplete and inadequate understanding of course subject matter. There is severely limited or no perception or originality. Course will not count toward degree.
C	73-76	2.00	Unsatisfactory work. Student performance demonstrates incomplete and inadequate understanding of course subject matter. There is severely limited or no perception or originality. Course will not count toward degree.
P	73-100	0.00	Satisfactory work. Student performance demonstrated complete and adequate understanding of course subject matter. Course will count toward degree.
F	0-72	0.00	Unacceptable work/Failure. Student performance is unacceptably low level of knowledge and understanding of course subject matter. Course will not count toward degree. Student may continue in program only with permission of the Dean.
I		0.00	Incomplete is assigned when work is of passing quality but is incomplete for a pre-approved reason. Once an incomplete grade is assigned, it remains on student's permanent record until a grade is awarded.
W		0.00	Withdrew from the course with Dean's permission beyond the second week of the term.

- All courses will be recorded and maintained in the student's permanent academic record; only courses that apply towards the degree will appear on the academic transcript. Non-credit or audited courses will not appear on the transcript.
- 4 core courses taken for a letter grade (pass = B- or higher for a core course)
- 2 elective courses taken pass/fail (pass = A, B, C for an elective)

Course Summary:

Date	Details
Mon Jan 1, 2024	No Class (New Year's Day day off)
Tue Jan 2, 2024	No Class
Thu Jan 4, 2024	Overview of Course; Introduction to R
Tue Jan 9, 2024	Descriptive Statistics
	Extra Credit Homework_1 (10 pts.)
	Extra Credit_Data Cleanup
Thu Jan 11, 2024	Binomial Distribution
Mon Jan 15, 2024	No Class (Martin Luther King Jr. Day)
Tue Jan 16, 2024	Other Discrete Distributions
Thu Jan 18, 2024	Normal Distribution
	Homework #1 Assigned
Tue Jan 23, 2024	No Class
Thu Jan 25, 2024	Central Limit Theorem/T-distribution
	Homework #1 Due
Tue Jan 30, 2024	Analysis of Variance (ANOVA)
	Homework #2 Assigned
Thu Feb 1, 2024	ANOVA; Non-parametric tests
Tue Feb 6, 2024	Power/Sample Size
	Homework #2
Thu Feb 8, 2024	Take Home Midterm Assigned
	Test of Proportions (Chi Square; Fisher's Exact)
Tue Feb 13, 2024	No Lecture/TA Time
Thu Feb 15, 2024	Take Home Midterm
	No Lecture
	Quiz Take Home Midterm Part 1
Mon Feb 19, 2024	No Class (President's Day)
Tue Feb 20, 2024	Linear Regression
Thu Feb 22, 2024	Homework #3 Assigned
	Multiple Linear Regression
Tue Feb 27, 2024	Logistic Regression
Thu Feb 29, 2024	Homework #3 Due
	Homework #3
	Multiple Logistic Regression
Tue Mar 5, 2024	Homework #4 Assigned
	Regression Modeling
Thu Mar 7, 2024	Survival Analysis
Tue Mar 12, 2024	Homework #4
	Bootstrapping/Review
	Homework #4 Due
	Take Home Final Exam Assigned
Tue Mar 19, 2024	Take Home Final Exam Due
	Take Home Final Exam