

University of Florida
College of Public Health & Health Professions Syllabus
PHC 6937 Survey of Advanced Biostatistical Methods for the Health Sciences (3 credits,
Summer A 2022)

Delivery Format: On-Campus
 Canvas: <http://elearning.ufl.edu>

Instructor Name: Jonathan Fischer
 Room Number: HPNP 1101 (M-F 2-3:15 PM)
 Phone Number: 352-294-5459
 Email Address: jfischer1@ufl.edu
 Office Hours: TBD
 Teaching Assistants: TBD
 Preferred Course Communications: Canvas messages

Prerequisites: *PHC6052 or permission of instructor*

PURPOSE AND OUTCOME

Course Overview

Uniquely blends the fundamentals of biostatistical inference with an introduction to advanced statistical techniques critical for the analysis of the growing compendium of health-related data. Topics span the analysis of high-dimensional, categorical, and longitudinal data from the health sciences. Applications utilize the statistical software packages R and SAS.

Course Description

A survey of biostatistical methods beyond basic testing and inference procedures. This course provides a one-semester overview of the advanced statistical procedures which are necessary for quantitative analysts and practitioners in the health sciences. Topics include complex regression models for independent, structured, and high-dimensional data, model/variable selection, and multiple testing, as well as strategies for imputing missing data, multivariate parametric and non-parametric methods with applications to health-related data. Students will have the option to use either the R or SAS software packages.

Relation to Program Outcomes

This is a biostatistics course designed for students in public health and the health sciences and is also appropriate for students in the biological sciences looking to expand their quantitative analysis capabilities. It will train students to rigorously analyze potentially complicated data, enabling the assessment of public health programs as well as the critical evaluation and/or production of original research in clinical and academic settings. Moreover, students will develop their quantitative communication and presentation skills.

Course Objectives and/or Goals

Upon completion of the course, students will be able to:

- 1) Apply statistical methods to categorical, count, or structured data that commonly arise in the health and biological sciences;
- 2) Quantify the impact of missing data and employ suitable methods in response, particularly for the censoring and missingness patterns that frequently occur in clinical and biological data;
- 3) Analyze patient survival data in epidemiological contexts;

- 4) Implement advanced dimension reduction and visualization tools to deal with high dimensional data, such as those arising in genomics;
- 5) Employ simple machine learning approaches for classification and clustering with applications in personalized medicine and case/control studies; and
- 6) Develop their own testable hypotheses and effectively communicate biostatistical techniques and results to an audience of clinicians or health workers, both via written reports and oral presentations.

Instructional Methods

Lectures will consist of the presentation of slides and demonstrations of data analysis using SAS and/or R; these resources, including recorded lectures, will be made available to students in Canvas. Examples make exclusive use of data from public health, clinical, biomedical, and genomic settings, and features unique to data from these sources are emphasized. Assessments will consist of homework assignments, two online quizzes, two take-home analysis assignments, and a final project (short written report and oral presentation). The course is split into three units corresponding to methods for independent, correlated, and multivariate data. These units are shaded orange, blue, and green, respectively, in the tabular schedule below. Assignments are shown in **bold** on the day on which they're due.

Topical Outline/Course Schedule

Here is a succinct summary of the purpose of each unit with examples of intended applications.

Unit 1—Advanced Methods for Independent Data: What are the optimal tools when our data are independent but not amenable to the simplest approaches? For instance, how should we best predict disease status or incidence, and how can we explicitly control for confounding effects when assessing the effectiveness of treatments?

Unit 2—Methods for Correlated Data: How do we handle data which are correlated, whether hierarchically or temporally? Examples include comparing measurements when some patients share the same hospital or physician, monitoring metrics of public health over time, and tracking patient disease survival times.

Unit 3—Multivariate Data Methods: How can we group or classify individuals based on clinical or genomic covariates? For example, are there subtypes within a certain disease?

Dates	Monday	Tuesday	Wednesday	Thursday	Friday
5/9-5/13	1 Review of EDA and Probability	2 Review of Testing and Inference	3 Multiple linear regression - Background	4 Multiple linear regression - Theory and Practice	5 Multiple linear regression - Modeling HW1 - Review
5/16-5/20	6 ANOVA – Standard design	7 ANOVA – Complex designs HW2 – Multiple Linear Reg.	8 Contingency tables – Odds ratios	9 Intro to Generalized linear models (GLM) HW3 - ANOVA	10 Logistic regression for binary outcomes

Dates	Monday	Tuesday	Wednesday	Thursday	Friday
5/23-5/27	11 Poisson/Negative binomial regression for count data	12 Linear mixed models (LMMs) - Background HW4 – Cont. Tables and Logistic Reg.	13 LMMs – Theory and Practice	14 LMMs & Quiz 1 Unit 1 Assignment: Independent Data	15 Modeling with LMMs
5/30-6/3	No Class Memorial Day	16 Generalized estimating equations - Background	17 GEEs/GLMMs	18 Modeling with GLMMs HW5 - LMM	19 Missing data methods – Handling incomplete clinical data
6/6-6/10	20 Survival analysis - Background	21 Survival analysis – Application to patient outcomes HW6 – GEE/GLMM	22 Clustering	23 PCA background	24 Clustering using PCA & Quiz 2 Unit 2 Assignment: Correlated Data
6/13-6/17	25 Classification for personalized clinical evaluation	26 Tree-based methods for regression and classification HW7 – PCA/Clustering	27 Wrap up/work on projects	28 Project presentations	29 Project presentations HW8 – Classification and Trees Project

Note: This schedule is tentative and subject to change. There are no required readings, though students may find it helpful to read the corresponding sections of the reference texts for each topic (see table below).

Course Materials and Technology

Texts: There are no required texts for this course. Below are several references.

- (RMB) Vittinghoff, E., Glidden, D.V., Shiboski, S.C., and C.E. McCulloch. *Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models*. Second Edition. Springer: New York, 2012.
- (AMS) Der, G. and B.S. Everitt. *Applied Medical Statistics Using SAS*. CRC Press: Boca Raton, 2013.
- (ISLA) James, G., Witten, D., Hastie, T., and R. Tibshirani. *An Introduction to Statistical Learning with Applications in R*. Springer: New York, 2013.

The following table lists the relevant chapters from the references for each topic. These are not required readings, but may be beneficial.

Topic	Day(s)	RMB	AMS	ISLA
Review	1-2	1-3	2	1-2
Multiple Linear Regression	3-5	4, 10	7, 8	3, 5, 6
ANOVA	6-7	3	6	
Contingency tables	8	3	4	
GLMs	9-11	5, 8	9, 10	4
LMMs	12-15	7	12, 13	
GLMMs	16-18	7	14	
Missing data	19	11	18	12

Topic	Day(s)	RMB	AMS	ISLA
Survival analysis	20-21	6	15, 16	11
Clustering	22-24			12
Classification	25-26			8

Software: Students will need to have access to R or SAS on their laptop. SAS can be used through UF Apps at <https://apps.ufl.edu/> or it can be purchased from the UF bookstore (See <https://software.ufl.edu/software-listings/sas-student-licensing.html> for SAS program purchase information and online documents). R and RStudio can be downloaded free from <https://www.r-project.org/> and <https://www.rstudio.com/products/rstudio/download/>.

For technical support for this class, please contact the UF Help Desk at:

- Learning-support@ufl.edu
- (352) 392-HELP - select option 2
- <https://lss.at.ufl.edu/help.shtml>

ACADEMIC REQUIREMENTS AND GRADING

Assignments

- Students are responsible for all course material, including reading required materials prior to class. The assessments will include homework, two exams, and a data analysis project. Please see the calendar above for tentative due dates and follow the course Canvas page for updates. Late submissions will receive zero credit unless arrangements were made in advance. Assignments must be submitted through Canvas.
- Eight **homework assignments** will be assigned and evaluated for both completion and correctness. They will consist of guided statistical analyses using software and provided data sets with follow-up questions regarding the generated output. These assignments will allow students to practice analyzing data using the methods presented in class along with the opportunity to interpret and explain their results.
- There are two take-home **quizzes** which will be completed on the dates indicated above. These quizzes will consist of multiple choice and short answer questions relating to the fundamental assumptions and applications of the methods learned in the corresponding unit.
- There are two take-home **analysis assignments**, and students will have three days to finish each of them. These assignments will consist of programming tasks with accompanying short answer questions requiring interpretation of results. These will test both the conceptual understanding and applied skills developed in lectures and homework assignments. When submitting these assignments on Canvas, both the answers to the questions and software outputs need to be submitted. All answers need to be clearly labeled in one file. All software output, along with SAS/R code, need to be in a separate file. Both assignments are due on a Friday and cover material through the Tuesday prior to the due date.
- Each student will complete a **final project** individually using one of the methods discussed in class. This project consists of the analysis of a data set of interest to the student, culminating in a written report (< 8 pages including figures) and short oral presentation (10-15 minutes). The report and presentation should walk through the scientific background and motivation before discussing the analysis, its conclusions, and any implications. These should be targeted towards a non-statistical audience in order to cultivate students' ability to communicate technical information to the public and practitioners who may or may not be statistically inclined. Projects will be assessed on the basis of accuracy and clarity, and additional guidance regarding these expectations, along with assistance selecting a suitable data set, will be provided early in the course. TurnItIn will be used to evaluate reports for originality.

Grading

Requirement	Percent of final grade	Notes
Homework	40%	8 homeworks, 5% each
Quizzes	20%	2 quizzes, 10% each
Data analysis assignments	20%	2 assignments, 10% each
Project	20%	Report and presentation, 10% each

Please be aware that a C- is not an acceptable grade for graduate students. The GPA for graduate students must be 3.0 based on 5000 level courses and above to graduate. A grade of C counts toward a graduate degree only if based on credits in courses numbered 5000 or higher that have been earned with a B+ or higher.

Percentage Earned	Letter Grade
[93,100]	A
[90,93)	A-
[87,90)	B+
[83-87)	B
[80-83)	B-
[77,80)	C+
[73,77)	C
[70,73)	C-
[67,70)	D+
[63,67)	D
[60,63)	D-
< 60	E

Letter Grade	Grade Points
A	4.0
A-	3.67
B+	3.33
B	3.0
B-	2.67
C+	2.33
C	2.0
C-	1.67
D+	1.33
D	1.0
D-	0.67
E	0.0
WF	0.0
I	0.0
NG	0.0
S-U	0.0

More information on UF grading policy may be found at <https://catalog.ufl.edu/graduate/regulations/>

Policy Related to Make up Exams or Other Work

Please notify me as soon as possible if unanticipated circumstances arise which interfere with your ability to complete an assignment on-time. Late work without prior notification and approval (unless there is an emergency) will receive a zero. Late work with prior approval (or due to an emergency) will be eligible for full credit. Any requests for make-ups due to technical issues must be accompanied by the ticket number received from e-learning support when the problem was reported to them. The ticket number will document the time and date of the problem. You must e-mail your instructor within 24 hours of the technical difficulty if you wish to request a make-up.

Policy Related to Required Class Attendance

We will follow the UF Attendance Policy. As this is a summer course, we will move at a fast pace, and it is critical that you keep up with the lectures. Please note all faculty are bound by the UF policy for excused absences. Excused absences must be consistent with university policies in the Graduate Catalog (<https://catalog.ufl.edu/graduate/regulations/>).

STUDENT EXPECTATIONS, ROLES, AND OPPORTUNITIES FOR INPUT

Expectations Regarding Course Behavior

Students are expected to show up for class prepared and on-time. Cell phones are to be silenced during class unless there is an emergency, in which case please inform the instructor.

Communication Guidelines

Please see the guidelines linked here: <http://biostat.ufl.edu/current-students/e-learning-resources/e-learning-basics/etiquette-online/>.

Academic Integrity

Students are expected to act in accordance with the University of Florida policy on academic integrity. As a student at the University of Florida, you have committed yourself to uphold the Honor Code, which includes the following pledge:

“We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honesty and integrity.”

You are expected to exhibit behavior consistent with this commitment to the UF academic community, and on all work submitted for credit at the University of Florida, the following pledge is either required or implied:

“On my honor, I have neither given nor received unauthorized aid in doing this assignment.”

It is your individual responsibility to know and comply with all university policies and procedures regarding academic integrity and the Student Honor Code. Violations of the Honor Code at the University of Florida will not be tolerated. Violations will be reported to the Dean of Students Office for consideration of disciplinary action. For additional information regarding Academic Integrity, please see Student Conduct and Honor Code or the Graduate Student Website for additional details:

<https://www.dso.ufl.edu/sccr/process/student-conduct-honor-code/>
<http://gradschool.ufl.edu/students/introduction.html>

Please remember cheating, lying, misrepresentation, or plagiarism in any form is unacceptable and inexcusable behavior.

Online Faculty Course Evaluation Process

Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at <https://gatorevals.aa.ufl.edu/students/>. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via <https://ufl.bluera.com/ufl/>. Summaries of course evaluation results are available to students at <https://gatorevals.aa.ufl.edu/public-results/>.

Policy Related to Guests Attending Class:

Only registered students are permitted to attend class. However, we recognize that students who are caretakers may face occasional unexpected challenges creating attendance barriers. Therefore, by exception, a department chair or his or her designee (e.g., instructors) may grant a student permission to bring a guest(s) for a total of two class sessions per semester. This is two sessions total across all courses. No further extensions will be granted. Students are responsible for course material regardless of attendance. For additional information, please review the Classroom Guests of Students policy in its entirety. Link to full policy: <http://facstaff.phhp.ufl.edu/services/resourceguide/getstarted.htm>

SUPPORT SERVICES

Accommodations for Students with Disabilities

Students with disabilities who experience learning barriers and would like to request academic accommodations should connect with the disability Resource Center. [Click here to get started with the Disability Resource Center](#). It is important for students to share their accommodation letter with their instructor and discuss their access needs, as early as possible in the semester.

Counseling and Student Health

Students sometimes experience stress from academic expectations and/or personal and interpersonal issues that may interfere with their academic performance. If you find yourself facing issues that have the potential to or are already negatively affecting your coursework, you are encouraged to talk with an instructor and/or seek help through University resources available to you.

- The Counseling and Wellness Center 352-392-1575 offers a variety of support services such as psychological assessment and intervention and assistance for math and test anxiety. Visit their web site for more information: <http://www.counseling.ufl.edu>. On line and in person assistance is available.
- You Matter We Care website: <http://www.umatter.ufl.edu/>. If you are feeling overwhelmed or stressed, you can reach out for help through the You Matter We Care website, which is staffed by Dean of Students and Counseling Center personnel.
- The Student Health Care Center at Shands is a satellite clinic of the main Student Health Care Center located on Fletcher Drive on campus. Student Health at Shands offers a variety of clinical services. The clinic is located on the second floor of the Dental Tower in the Health Science Center. For more information, contact the clinic at 392-0627 or check out the web site at: <https://shcc.ufl.edu/>
- Crisis intervention is always available 24/7 from:
Alachua County Crisis Center:
(352) 264-6789
<http://www.alachuacounty.us/DEPTS/CSS/CRISISCENTER/Pages/CrisisCenter.aspx>

Do not wait until you reach a crisis to come in and talk with us. We have helped many students through stressful situations impacting their academic performance. You are not alone so do not be afraid to ask for assistance.

Inclusive Learning Environment

Public health and health professions are based on the belief in human dignity and on respect for the individual. As we share our personal beliefs inside or outside of the classroom, it is always with the understanding that we value and respect diversity of background, experience, and opinion, where every individual feels valued. We believe in, and promote, openness and tolerance of differences in ethnicity and culture, and we respect differing personal, spiritual, religious and political values. We further believe that celebrating such diversity enriches the quality of the educational experiences we provide our students and enhances our own personal and professional relationships. We embrace The University of Florida's Non-Discrimination Policy, which reads, "The University shall actively promote equal opportunity policies and practices conforming to laws against discrimination. The University is committed to non-discrimination with respect to race, creed, color, religion, age, disability, sex, sexual orientation, gender identity and expression, marital status, national origin, political opinions or affiliations, genetic information and veteran status as protected under the Vietnam Era Veterans' Readjustment Assistance Act." If you have questions or concerns about your rights and responsibilities for inclusive learning environment, please see your instructor or refer to the Office of Multicultural & Diversity Affairs website: www.multicultural.ufl.edu