CEDARS-SINAI
Graduate School of Biomedical Sciences

Course Number: BMS510
Course Title: Biostatistics

Fall Trimester 2022
Day/Time of Class: Tuesday and Thursday 1pm- 2:30pm
Room: PDC Classroom
Credit Hours: 3

Primary Instructor:

Márcio Augusto Diniz

(marcio.diniz@cshs.org)

Co-instructor Name:

Michael Luu
(michael.luu@cshs.org)

Teaching Assistant:

Huyen (Hayley) Nguyen
(huyen.nguyen@cshs.org)

Office Hours:
Huyen – Monday 11am – 12pm – Davis Building
Marcio – Wednesday 11am – 12pm - Office G592 at Pacific Design Center
Michael – Friday 11am – 12pm – Microsoft Teams

Required Textbooks/Readings

1. P-value misuses – Nick and Inga

a) Head ML, Holman L, Lanfear R, Kahn AT, Jennions MD. The extent and consequences of p-hacking in science. PLoS Biol. 2015 Mar 13;13(3):e1002106.

b) Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, Altman DG. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. European journal of epidemiology. 2016 Apr;31(4):337-50.

c) Gelman A, Stern H. The difference between "significant" and "not significant" is not itself statistically significant. The American Statistician. 2006 Nov 1;60(4):328-31.

2. Reproducibility in Science – Na Jeong and Nimisha

a) Begley CG, Ioannidis JP. Reproducibility in science: improving the standard for basic and preclinical research. Circulation research. 2015 Jan 2;116(1):116-26.

b) Begley CG, Ellis LM. Raise standards for preclinical cancer research. Nature. 2012 Mar;483(7391):531-3.

c) Errington TM, Denis A, Perfito N, Iorns E, Nosek BA. Reproducibility in Cancer Biology: Challenges for assessing replicability in preclinical cancer biology. eLife. 2021 Dec 7;10:e67995.

3. Reporting in Pre-clinical sciences – Elena and Roberta

a) Percie du Sert N, Ahluwalia A, Alam S, Avey MT, Baker M, Browne WJ, Clark A, Cuthill IC, Dirnagl U, Emerson M, Garner P. Reporting animal research: Explanation and elaboration for the ARRIVE guidelines 2.0. PLoS biology. 2020 Jul 14;18(7):e3000411.

b) Avey MT, Moher D, Sullivan KJ, Fergusson D, Griffin G, Grimshaw JM, Hutton B, Lalu MM, Macleod M, Marshall J, Mei SH. The

devil is in the details: incomplete reporting in preclinical animal research. PLoS One. 2016 Nov 17;11(11):e0166733.

c) Serghiou S, Contopoulos-Ioannidis DG, Boyack KW, Riedel N, Wallach JD, Ioannidis JP. Assessment of transparency indicators across the biomedical literature: How open is open?. PLoS biology. 2021 Mar 1;19(3):e3001107.

4. Heterogeneity – Basia and Maya

a) Usui T, Macleod MR, McCann SK, Senior AM, Nakagawa S. Meta-analysis of variation suggests that embracing variability improves both replicability and generalizability in preclinical research. PLoS biology. 2021 May 19;19(5):e3001009.

b) Voelkl B, Vogt L, Sena ES, Würbel H. Reproducibility of preclinical animal research improves with heterogeneity of study samples. PLoS biology. 2018 Feb 22;16(2):e2003693.

c) Richter SH. Systematic heterogenization for better reproducibility in animal experimentation. Lab animal. 2017 Sep;46(9):343-9.

5. Sex as Biological Variable - Beyza

a) Garcia-Sifuentes Y, Maney DL. Reporting and misreporting of sex differences in the biological sciences. eLife 2021;10:e70817

b) Woitowich NC, Beery A, Woodruff T. Meta-research: a 10-year follow-up study of sex inclusion in the biological sciences. eLife. 2020 Jun 9;9:e56344.

c) Beery AK. Inclusion of females does not increase variability in rodent research studies. Current opinion in behavioral sciences. 2018 Oct 1;23:143-9.

Suggested Textbooks/Readings

1. Motulsky H. Intuitive biostatistics: a nonmathematical guide to statistical thinking. Oxford University Press, USA; 2013 Dec 13.
2. http://www.nature.com/collections/qghhqm/
3. http://www.ats.ucla.edu/stat/
4. http://r4ds.had.co.nz/

Students will be expected to read all assigned material before the class session in which it will be discussed. Readings supplementing the text will be placed on the class website. The reference readings are optional supplementary material to further enhance class discussion and provide more theory behind the topics we will discuss.

Policies

All students are required to read and understand the policies of Graduate Research Education: http://www.cedars-sinai.edu/Education/Graduate-Research-Education/Graduate-PhD-Program/Graduate-Program-Policy-Handbook.pdf.

These policies are meant to support the policies of Cedars-Sinai Medical Center which can be found here: http://cshsppmweb.csmc.edu/.

Due date policy: Assignments are turned in after their deadlines, there will be a reduction of 50% on their grades.

Plagiarism: It is characterized when an assignment is submitted by a student containing any work that is not produced by the same student and without acknowledging the sources. This is academic misconduct, and it will be reported to the Graduate Program.

Prerequisites: None.

Credit Hour Expectations

Students are expected to attend class for 1.5 hours, 1.5 times per week. It is the expectation that students will spend a minimum of two hours out of class per every in-class hour in preparation for this course.

Medical Library

The Medical Library is in South Tower, Plaza Level and is available for use from Monday-Friday from 8am-8pm. For 24-7 access, please contact Janet Hobbs (janet.hobbs@cshs.org). Electronic resources are available here: http://web.csmc.edu/research-and-education/the-medical-library/.

Course Learning Objectives

All students completing the Biostatistics Course are expected to have acquired the following knowledge and skills on:

1. How statistical methods play a fundamental role in reproducibility in science;
2. How to critically read statistical components of scientific papers;
3. How to manage and describe experimental data;
4. How to make decisions and interpret results of statistical analyses;
5. How to interact with statisticians.

Course Description

Our course will be developed as two parallel and interconnected activities:

1. Lectures on Statistical Concepts

Every living organism in the universe gathers information from the environment and makes decisions. Good decisions are rewarded with survival. Statistical science main objects of study are how to summarize, interpret information and make decisions. This places statistics in the core of all sciences.
The goal of this module is to introduce fundamental probability and statistical concepts used to design and analyze studies and experiments. You will learn how your opinion should change on the basis of data; how to estimate values of interest; what uncertainty is attached to these values; how to improve the efficiency of your investigations; and how your imperfect information should be used to make decisions for action. The preferred approach will be by presenting motivating case studies.

2. Hands on Statistical Programming with R

R is a language and environment for statistical computing and graphics. R provides a wide variety of statistical (e.g., classical statistical tests, Bayesian analysis, linear modelling) and graphical techniques, and is highly extensible. The S language is often the vehicle of choice for research in statistical methodology, and R provides an Open Source route to participation in that activity.
One of R's strengths is the ease with which well-designed publication-quality plots can be produced, including mathematical symbols and formulae where needed. Great care has been taken over the defaults for the minor design choices in graphics, but the user retains full control. R is available as Free Software under the terms of the Free Software Foundation's GNU General Public License in source code form. It compiles and runs on a wide variety of UNIX platforms and similar systems (including FreeBSD and Linux), Windows and MacOS.

Lecture Date	Lecture Subject	Homework Due Date (Weight)
Tuesday, August 30th	1. Introduction, Goals, Structure	
Thursday, September 1st	2. Datasets	
Tuesday, September 6th	3. R-basics	
Thursday, September 8th	4. Data manipulation: dplyr and tidyr	
Tuesday, September 13th	5. Data manipulation: dplyr and tidyr	
Thursday, September 15th	6. Descriptive Statistics	
Tuesday, September 20th	7. Descriptive Statistics	
Thursday, September 22nd	8. Descriptive Statistics: ggplot2	
Tuesday, September 27th	9. Descriptive Statistics: ggplot2	Datasets (10%)
Thursday, September 29th	10. Descriptive Statistics: ggplot2	
Tuesday, October 4th	11. Descriptive Statistics: ggplot2	
Thursday, October 6th	12. Probability	
Tuesday, October 11th	13. Probability	Descriptive tables (10%)
Thursday, October 13th	14. Estimation and Confidence Intervals	
Tuesday, October 18th	15. Estimation and Confidence Intervals	
Thursday, October 20th	16. Test of Hypothesis	
Tuesday, October 25th	17. Test of Hypothesis	Plots (20%)
Thursday, October 27th	18. Test of Hypothesis	
Tuesday, November 1st	19. Test of Hypothesis	
Thursday, November 3rd	20. Multiple Comparisons	
Tuesday, November 8th	21. Multiple Comparisons	
Thursday, November 10th	22. One-Way Analysis of Variance	
Tuesday, November 15th	23. One-Way Analysis of Variance	Test of Hypothesis (20%)
Thursday, November 17th	24. Two-Way Analysis of Variance	
Tuesday, November 22nd	Thanksgiving	
Thursday, November 24th	Thanksgiving	
Tuesday, November 29th	25. Two-Way Analysis of Variance	
Thursday, December 1st	26. P-values misuses	Paper discussion (40%)
Tuesday, December 6th	27. Reproducibility in Science	
Thursday, December 8th	28. Reporting in Pre-clinical Sciences	
Tuesday, December 13th	29. Heterogeneity in experiments	
Thursday, December 15th	30. Sex as Biological Variable	

Evaluation Criteria

To measure the intended outcomes, the student will have 6 assignments with different weights and deadlines showed at page 3.The final grade for the Biostatistics module will follow:

A-: [90%, 93%)	A: [93%, 96%)	A+: [96%, 100%]
B-: [70%, 77%)	B: [77%, 83%)	B+: [83%, 90%)
C-: [60%, 63%)	C: [63%, 67%)	C+: [67%, 70%)