Last Updated: August 7, 2020

ADVANCED BIOMETRY
(BIOL 6090)

www.pelagicos.net/classes_advancedbiometry_fa20.htm

HAWAII PACIFIC UNIVERSITY	TIME: 9:00 – 10:15
MAKAPU'U CAMPUS (MKC)	DAYS: Monday & Wednesday
FALL SEMESTER, 2020	ROOM: OLC 103, Oceanic Institute

INSTRUCTOR

David Hyrenbach, Ph.D.
Phone: 808- 236-3563
Email: khyrenbach@hpu.edu
Office: EMSB, at Oceanic Institute (http://www.pelagicos.net/images/EMSB_Map.jpg)

OFFICE HOURS

Monday & Wednesday (EMSB, OI) 10:30 – 11:45
Monday & Wednesday (EMSB, OI) 14:00 – 15:00
… or by appointment

COURSE DESCRIPTION:

This hands-on workshop focuses on the analysis and interpretation of univariate data. Lectures and assignments emphasize the conceptual understanding and the practical use of statistical methods, with the goal of providing students with a tool-kit they will use in their thesis research and beyond. Topics covered include statistical distributions, ANOVA applications, regression and correlation, power analysis, and contingency / chi-square tests. Instruction to statistical theories and methods will be enhanced by discussing examples drawn from the scientific literature and conducting analyses using the R software.

RESOURCES:

Texts: We will use three texts:

Required: Focuses on the principles and mechanics of performing statistical analyses using the R software.

Discovering Statistics Using R, 2012. Andy Field, Jeremy Miles and Zoe Field. Sage Publications, Thousand Oaks (USA), ISBN: 9781446200469

Recommended: Focus on practical ways to summarize, analyze, and display data using the R software.

Handbook of Biological Statistics, 2014, (3rd edition). McDonald, J.H. 2014. Sparky House Publishing, Baltimore, Maryland. **(Available Online:** http://www.biostathandbook.com/). (**PDF version:** http://www.biostathandbook.com/HandbookBioStatThird.pdf

An R Companion for the Handbook of Biological Statistics, version 1.3.2. 2015. Mangiafico, S.S. (**Available Online:** http://www.rcompanion.org/rcompanion/). **(PDF version:** http://www.rcompanion.org/documents/RCompanionBioStatistics.pdf)

Web-site: You are encouraged to use the student resources for the required text, including example self-test answers, posted on the publisher's web-site https://studysites.sagepub.com/dsur/study/default.htm

Software: We will use two free-ware software packages available from the internet for PC and Mac.

- R: http://www.r-project.org/
- R studio: https://www.rstudio.com/products/rstudio/

STUDENT LEARNING OBJECTIVES:

- The main goal of this course is to provide students with the baseline of knowledge and the tools to perform and interpret statistical analyses of biological data. To this end, students will:

    - Develop the ability to state testable hypotheses and select the statistical tests to evaluate those hypotheses in a scientifically rigorous manner. Students will determine which analysis approaches are better suited to test the hypotheses and which underlying assumptions are required.
    - Learn to evaluate the assumptions specific to the various statistical parametric / non-parametric statistical tests and to determine what data transformations can be used to meet these assumptions. Given this assessment, students will determine which tests are appropriate for a given dataset.
    - Understand the different types of numerical data (categorical vs nominal, ordinal, discrete vs. continuous) and what options are available for their analysis.
    - Study the metrics and methods used for quantifying data distributions.
    - Evaluate the strengths and weaknesses of various numerical analysis, including the techniques for organizing, re-coding, and transforming data.
    - Explore how experimental design influences data analysis, including how to reduce the likelihood of type I (false positive) and type II (false negative) statistical errors, and statistical power.

- Another major focus of this course is to review the philosophical underpinnings and practice of the scientific method. Students will learn how experimental controls and sample size influence the findings of significance for a variety of analyses. Through the evaluation of power analysis and post-hoc tests, students will critically examine the reasons for non-significant outcomes.

- More specifically, students who successfully complete this course will be able to:

    - Understand the mechanics and theory of statistical testing, and the associated terminology: p-value, critical value, alpha, significance, type-I error, type-II error.
    - Understand the use of inferential and descriptive statistics
    - Understand the concept and function of the "null / alternate hypothesis" and be able to formulate hypotheses for their own research questions
    - Statistically analyze and interpret data using a portfolio of statistical tests
    - Identify the appropriate analytical procedures for a variety of data sets
    - Present data and statistical results in tabular and graphical format
    - Effectively analyze and present summary statistical (mean, median, mode, range, standard deviation, standard error, coefficient of variation, variance, confidence interval) by hand
    - Effectively analyze and graph data using the statistical software package R
    - Understand the limitations of the statistical procedures learned in class, and be able to critically evaluate the statistics used in the published papers and student projects

GRADING: Students will be assessed with quizzes, assignments, an individual project and class participation.

R Assignments	20%	(5 points each * 4 assignments)
Knowledge Exams	50%	(5 quizzes * 10 points each)
Individual Project	25%	(proposal, screening, analysis, presentation, write-up)
Class Participation	5%	(various class activities)
Total points: 100		(**Note:** I will scale all scores relative to the top score)

90 - 100%	A	I will use pluses (+) and minuses (-).
80 - 89%	B	
70 - 79%	C	Graduate students need at least a C- to pass MARS 6090.

NOTE: The homework sets are study guides for the quizzes and will not be graded.
Make sure you go over the homework and the quiz keys. They will be discussed in class.

1. **Quizzes: 50% (there will be 5 quizzes - 10 points each)**
   - Quizzes will be administered at the start of class and will last 30 or 60 minutes (in-person / online)
   - Quizzes will cover material from lectures, R assignments, and assigned readings.
   - Quizzes will be cumulative and may include questions from previous quizzes.
   - There will be no make-up quizzes, unless students have a valid (medical / family) reason.

2. **R Homework Sets: 20% (there will be 4 assignments - 5 points each)**
   - These assignments will allow you to practice implementing statistical tests using R.
   - The assignments will be discussed in class.

3. **Analysis Project: 25%**
   - Students will develop and complete an independent data analysis project.
   - Analyses may involve datasets from a graduate advisor, internship, thesis project, or the internet.
   - Students are encouraged to take on an ambitious project, using R tools beyond those discussed in class.
   - The project grade will reflect the complexity, rigor, and completeness of the analyses.
   - Students will turn in:
     - A proposal detailing the proposed questions and analyses (5 points)
     - A data screening report, evaluating data for normality and performing transformations (5 points)
     - A 15-minute presentation, outlining the project's goals, approach, and outcomes (5 points)
     - A write-up, summarizing the analysis methods and the results (10 points)

4. **Participation: 5%**
   Students will be evaluated on their effort and commitment to learning. This includes participating in online / in-person class meetings, readiness, and – if necessary - seeking extra help in class or office hours.

   Students will review the assignments and readings before coming to class. The instructor will assess student preparedness in class, using group activities and 5-minute papers at the beginning and the end of class. Students will be evaluated qualitatively (+ / -).

READING ASSIGNMENTS:

- To fully understand the theories, assumptions and output of the analytical procedures presented in this course it is essential to carefully read all assignments in the text books.
- You are advised to read the assigned material and watch the recorded lecture before coming to class.

DATASETS:

A primary goal of this course is to provide students with the opportunity to analyze real-world data. Ideally you should use data that you are somewhat familiar and excited about. Thus, students are encouraged to use their own data, collected as part of their thesis research, for the class project. This project is time consuming and demanding, but provides a great opportunity for making progress with the analysis of thesis data.

The dataset should involve at least one response variable and three or more driver variables, with a minimum of 30 data points. The only requirement concerning the type of data is that they should involve multiple biological and physical variables, measured concurrently in time and space. Students who do not have their own data set, will be able to use data sets provided by a thesis advisor, the class instructor, or downloaded from the internet.

OTHER CLASS POLICIES:

Attendance - Attendance is not mandatory, but participation in class activities counts towards your grade.

Please note that because of COVID-19, I will accommodate missed quizzes and class meetings.
If you are not feeling well, let me know by email and do not come to class. Good communication is key.

My goal is to meet with the entire class twice per week, but we may have to separate into two sections (monday / wednesday) for in-person meetings. Students will sign up for one of these sections and attend their day throughout the semester.

Recorded lectures will be posted online and in-person meetings will be used to go over homeworks, assignments, and other group activities.

Coming to class late - If you must come late or leave early discuss the need with me and try to make as small a disturbance as possible, by sitting close to the door.

Cell phones are **not allowed** in class. Turn them off before entering the room. You may leave the room to receive / make a call.

Laptops are allowed to take notes / view the lecture pdfs. This is a privilege which will be revoked if laptops are used for non-class activities (e.g., email / facebook). Students are encouraged to bring their laptops and work with R, during in-person class meetings.

OTHER UNIVERSITY POLICIES:

Academic Honesty:

Hawaii Pacific University provides a learning environment based upon academic excellence and integrity. In this course, it is expected that students will adhere to all university guidelines regarding academic honesty.

It is Hawaii Pacific University policy that any act of Academic Dishonesty will incur a penalty up to and including expulsion from the University. Any student who cheats on an academic exercise, lends unauthorized assistance to others, or who hands in a completed assignment that is not his or her work will be sanctioned. Any student who copies another student's assignment or uses their homework or test to guide their work, will receive a grade of "0" for that assignment. Depending on the severity of the case, a single event of academic dishonesty may result in either a zero on the assignment for everybody involved, an "FD" in the course, or expulsion from the University. The grade of "FD" represents an "F" for academic dishonesty and it will remain a permanent part of your academic record, and is not subject to HPU's normal retake policy.

Students are expected to comply with HPU's academic honesty policies and are furthermore required to internalize the principles of scientific integrity. In particular, it is academically dishonest to **plagiarize** (i.e., pass off someone else's intellectual work as your own). Directly quoting others, even with proper attribution of the source, is never done in scientific writing, so **there are no circumstances in which including someone else's writing in your paper will be acceptable.**

If you are unclear on what is and what is not plagiarism, please discuss it with me. For more information on what constitutes plagiarism, see the links in Campus Pipeline, under the Libraries folder. Purdue University's

Online Writing Lab provides excellent advice on how to avoid committing plagiarism: www.owl.english.purdue.edu/handouts/research/r_plagiar.html
Students are required to understand the reasons for this policy and internalize the principles of scientific integrity. To this end, all students will follow these guidelines:

- It is academically dishonest to try to pass off someone else's intellectual work as your own, or to help someone else to do so.
- All research results must be documented with laboratory notebooks, computer files of raw data.
- All work must be original.
- Data analysis and results must include clear and accurate descriptions of the methods; the results of others must be clearly identified, and the appropriate sources must be cited.
- If you use published materials (for facts, statements, images, data, etc.) in completing an assignment, credit must be given by citing the source – failure to properly cite sources is considered plagiarism.
- Directly quoting others is not acceptable in scientific writing. Thus, there are no circumstances in which including someone else's writing in your papers is permissible.

Special Needs Policy:

Under the Rehabilitation Act of 1973 (Section 504), the Americans with Disabilities Act, Title III (Public Accommodations) and Title V (Employment), and the Hawai'i Fair Employment Practice Law, Hawai'i Pacific University does not discriminate against individuals with disabilities. HPU will make reasonable accommodations in its policies, practices and procedures in order to: 1) allow students with disabilities to benefit from the services and facilities offered by the University and 2) employ otherwise qualified individuals with disabilities who are able to do essential tasks of specific jobs. HPU will accommodate known disabilities, unless to do so would impose an undue hardship. This is interpreted to mean significant difficulty (fundamentally altering the nature of the services and facilities provided by the University) or expense.

Note: Any student with a documented disability who would like to request accommodations should contact the University Disability Services Office (933-0816 (V), 933-3334 (TTY), Campus Center Room 311) as early in the semester as possible.

Counseling and Behavioral Health Services:

The Counseling and Behavioral Health Services (CBHS) department provides FREE and CONFIDENTIAL counseling services to current registered HPU students including the following counseling services: individual; couples; family; crisis Intervention; consultation services for students, parents, faculty and staff; grief and loss.

To schedule an appointment or for more information, please contact the Academic Advising & Behavioral Health office at either of the following: DOWNTOWN (808) 544-1198 or HAWAII LOA (808) 236-3578. The CBHS department can also be reached by email at: counseling@hpu.edu. They are unable to take appointments online or through email. All appointments must be scheduled by calling the numbers listed above. If this is an emergency, please call 911 or go to the nearest emergency room or hospital.

Last Updated: August 7, 2020

Class Schedule
(August 24 – December 2)

This is a tentative schedule, which may speed up / slow down, to keep up with the pace of student learning. Revised schedules will be posted periodically, throughout the semester.

NOTE: With 7 enrolled students, this schedule assumes the entire class will meet in person twice a week.

Date	***Topics (Recorded Lectures)***	***Reading Assignment***	***In Class Activity***
8/24	L1 - Introduction to the Course Three Statistical Frameworks	Field: Chapters 1.1 – 1.6 McDonald: Pp 1 – 13	Review Syllabus. Discuss Class Objectives.
8/26	L2 - The Scientific Method & Hypothesis Testing	Platt (1964) Quinn & Dunham (1983) Stephens et al. (2007)	Philosophical Discussion.
8/31	L3 - Describing & Summarizing Data	Field: Chapters 1.7 McDonald: Pp 1 – 13	Measurements & Data Types.
9/2	L4 - Sampling & Probability	Field: Chapter 2.1 – 2.3 McDonald: Pp 14 – 23 McDonald: Pp 131 – 133	Sampling & Probability.
9/7	**NO CLASS – HPU Holiday (Labor Day)**	Field: Chapter 3 Mangiafico: Pp 1 – 13	Install R and Run & RStudio.
9/9	L5 - Conditional Probability	Silver (2012) Pp. 232 - 261 Marcus & Davis (2013)	Conditional Probability.
9/14	L6 - Estimation: Central Tendency & Spread	Field: Chapters 2.3 – 2.6 McDonald: Pp 101 – 111 Mangiafico: Pp 78 – 87	Parameter Estimation.
9/16	L7 - Statistical Inference & Reliability *Quiz 1*	McDonald: Pp 112 – 121 Mangiafico: Pp 87 – 94 Silver (2012) Pp. 176 – 203	Confidence Intervals.
9/21	L8 - Estimation and Statistical Modelling	Field: Chapter 1.7, 2.1, 2.2 McDonald: Pp 16 - 28 Silver (2012) Pp. 142 – 175	Model Fitting and Hypothesis Testing
9/23	L9 - Probability Distributions	Field: Chapter 4 Limbert, et al. & (2001)	Probability Distributions
9/28	L10 - Parametric Statistics: Assumptions & Data Exploration	Field: Chapter 5.1 – 5.7 McDonald: Pp 112 – 121 Mangiafico: Pp 87 – 94	Data Exploration
9/30	L11 - Descriptive Statistics	Field: Chapter 3 and 4	Descriptive Statistics
10/5	L12 - Data Transformations	Field: Chapter 5.8 McDonald: Pp 141 – 145	Data Transformations
10/7	L13 - Dealing with lack of Normality *Quiz 2*	Field: Chapter 5	Data Analysis Flow
10/12	L14 - Statistical Modeling & Hypothesis Testing	Field: Chapter 1.7, 2.1, 2.2 McDonald: Pp 16 - 28 Silver (2012) Pp. 142 – 175	Model Fitting and Hypothesis Testing
10/14	L15 - Correlation	Field: Chapter 6 McDonald: Pp 191 – 214 Mangiafico: Pp 182 – 192	Correlation

Last Updated: August 7, 2020

Date	***Topics (Recorded Lectures)***	***Reading Assignment***	***In Class Activity***
10/19	L16 - Simple Linear Regression	Field: Chapter 7 McDonald: Pp 191 – 214 Mangiafico: Pp 182 - 192	Simple Regression
10/21	L17 - Multiple Regression	Field: Chapter 7 McDonald: Pp 231 – 238 Mangiafico: Pp 216 - 228	Multiple Regression
10/26	L18 - Information Theory	Burnham et al. 2011	AIC and AICs
10/28	L19 - Comparing two or more means *Quiz 3*	Field: Chapter 9 McDonald: Pp 121-130, 160-177 Mangiafico: Pp 94-100, 180-185	T tests
11/2	L20 - Analysis of Variance: Theory & Implementation	Field: Chapter 10 McDonald: Pp 45 - 156 Mangiafico: Pp 106 - 117	One-way ANOVA
11/4	L21 - Analysis of Covariance (ANCOVA)	Field: Chapter 11 McDonald: Pp 220-228 Mangiafico: Pp 206-215	ANCOVA
11/9	L22 - Multifactor ANOVA	Field: Chapter 13 McDonald: Pp 173-179 Mangiafico: Pp 143-160	Two-way ANOVA
11/11	**NO CLASS – HPU Holiday (Veteran's Day)**		
11/16	L23 - Repeated Measures ANOVA	Field: Chapter 14	RM ANOVA
11/18	L24 - Nested ANOVA *Quiz 4*	McDonald: Pp 165-172 Mangiafico: Pp 133-142	Nested ANOVA
11/23	L25 - Applications – Categorical Data (Contingency Tests)	Field: Chapter 18.1 – 18.7 McDonald: Pp 29-100 Mangiafico: Pp 14-77	Contingency Tests
11/25	L26 - Applications - Categorical Data (Logistic Regression)	Field: Chapter 8 McDonald: Pp 238-253 Mangiafico: Pp 228-255	Logistic Regression
11/30	L27 - Applications – Nonparametrics	Field: Chapter 15 McDonald: Pp 157-164, 186-189 Mangiafico: Pp 101-102, 118-128, 178-181	Non-parametric Tests
12/2	L28 - Applications – Statistical Power *Quiz 5*	Field: Chapter 2.6 Dayton (1998) The Economist (2013) Johnson (2002)	Statistical Power
12/7	Final Project Presentations (9:00- 11:15)		
12/11	Project Write-up Due by midnight (Via Email: khyrenba@gmail.com)		

OTHER IMPORTANT ADMINISTRATIVE DEADLINES:

Last Day to Register and to Drop A Class with 100% Tuition Refund (16-Week Classes)	Wednesday, September 2, 2020
Last Day to Drop (16-Week Classes) with 50% Tuition Refund for Dropped Credits – 50% Student Financial Responsibility for Dropped Credits	Wednesday, September 9, 2020
Last Day to Drop (16-Week Classes) without W Grade	Monday, September 21, 2020
Last Day to Drop (16-Week Classes) with 25% Tuition Refund for Dropped Credits – 75% Student Financial Responsibility for Dropped Credits	Monday, September 21, 2020
Last Day to Drop (16-Week Classes) with a W Grade	Monday, November 2, 2020