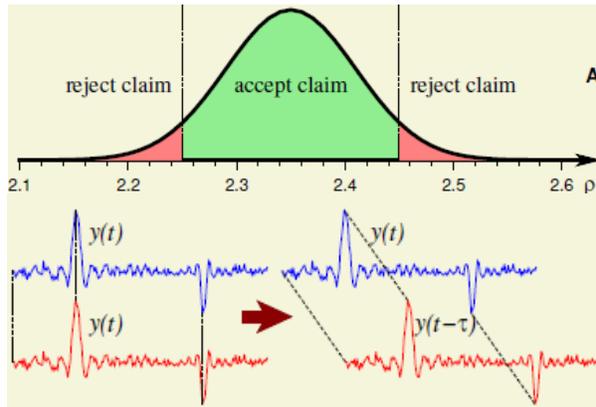# ERTH413/613: Introduction to Statistics and Data Analysis

- **Exploratory Data Analysis**
- **Error Propagation**
- **Probability Theory and Statistics**
- **Statistical Tests**
- **Curve Fitting and Regression**
- **Sequence Analysis**
- **Spectral Analysis**
- **Analysis of Directional Data**

Spring 2022, Instructed by Prof. Helen Janiszewski (hajanisz@hawaii.edu); Office: POST 614A.
Class Meetings: MWF 9:30-10:20 am, POST 703; **Office Hours:** TBD
Prerequisites: Math242 (2nd semester calculus), ERTH250 (scientific programming using Matlab/Python), or instructor consent
Textbook: Paul Wessels Lecture Notes. Recommended (optional) text: John C. Davis, Statistics and Data Analysis in Geology, 3rd Edition

OVERVIEW AND OBJECTIVES:
In this course students gain foundational understanding of the basic theory behind statistics, probability, and quantitative data analysis, as well as practical experience in working with real data sets using computer software. The course emphasizes solving problems and independent learning and inquiry. Students...

- Learn how to apply exploratory data analysis techniques to characterize their data or discover patterns within it
- Understand how to propagate errors in calculations of derived quantities
- Learn and apply concepts of samples, population, probability distributions, and the central limit theorem
- Doing formal hypothesis testing in interpreting data
- Gain an introduction to matrices, linear algebra, and least squares formalism for curve fitting and regression
- Explore various ways to examine sequential data
- Practice spectral analysis and apply concepts of aliasing and leakage
- Analyze directional data

Applications will emphasize geosciences but the course is *relevant to all fields of science*.

FORMAT AND WORKLOAD
You will be assigned readings, and problems to review prior to class meetings, and are expected to come to class with questions. Class time is an interactive learning environment and will be a combination of lecture, discussion, and working through examples and problem sets. Problem sets will be assigned approximately weekly and will involve using computer software to apply and practice using the techniques covered. The class will be taught using Matlab. You are welcome to use an alternative such as Python or R; however, the instructor is an expert in Matlab, has basic experience with Python, and has minimal experience with R. There will be a mid-term and a final exam.

STUDENTS TAKING ERTH613 will analyze a data set of their choice—ideally something related to their thesis/dissertation research—and gain professional practice presenting this work in writing and orally. This will include a ~6-10 page report (including figures, but not bibliography) and a 15-minute presentation at the end of class.

GRADING
Data analysis is a very hands-on activity and there will be weekly problem sets that require a mix of mathematical and computational manipulations. _Homework must be handed in at the assigned due date_, unless you have made prior arrangements with me. Otherwise, unexcused late homework will receive 10% less credit for each day it is late. _If you anticipate a conflict for exams, you must re-schedule the exam prior to the scheduled date._ The final grade for **ERTH413** will be a weighted average of grades for homework (60%), class engagement (10%), midterm (15%), and final exam (15%). For **ERTH 613** the grade is a weighted average of grades for homework (50%), class engagement (5%), midterm (15%), final exam (15%), and the project report (15%).

ATTENDANCE POLICY
Students who are enrolled in this course, but never attend will be flagged by the course instructor for non-participation before the last day to add/drop (for 100% tuition refund) deadline. Flagged students will be administratively dropped by the Office of the Registrar. Any changes to a student's enrollment status may affect financial aid eligibility and can result in the return of some of all of federal student financial aid.

EARTH SCIENCES PROGRAMMATIC STUDENT LEARNING OBJECTIVES
This course emphasizes three _student learning objectives_ for undergraduate and graduate students:
- Students can apply technical knowledge of computer applications and mathematics and physics to solving real-world problems in earth sciences
- Students use the scientific method to define, critically analyze, and solve a problem in earth science
- Students can communicate scientific knowledge in both oral presentations and in writing

WORKING COURSE SYLLABUS (TENTATIVE)

1. Exploring Data and Error Analysis
Week 1
Reading: Swan and Sandilands Handout, Wessel 1.1, 1.2, 2
1.1 Classification of data
1.2 Exploratory data analysis
2 Error Analysis (focus on section 2.3.4)

2. Basics of Probability and Statistics
Week 2
Reading: Wessel 3.1.1, 3.1.2
3.1 Probability Basics
  3.1.1 Permutations
  3.1.2 Combinations

Week 3
Reading: Wessel 3.1.3-4, 3.1.6-8, 3.2.1-3, 3.3.1
3.1 Probability Basics
   3.1.3 Discrete Probability
   3.1.3 Probability
   3.1.4 Some Rules of Probability
   3.1.6 Additional rules
   3.1.7 Conditional Probability
   3.1.8 Bayes Theorem
3.2 The M&Ms of Statistics
   3.2.1 Population and Samples
   3.2.2 Measure of central location (mean, median, mode)
   3.2.3 Measure of variation
3.3.1 Binomial Distribution, Hypergeometric Distribution

Week 4
Reading: Wessel 3.2.4-6, 3.3.1-2, 3.4.1
3.2 The M&Ms of Statistics
   3.2.4 Robust Estimation (MAD)
   3.2.5 Inference about the mean and Central Limits Theorem (Davis pages on Central Limit Theorem)
   3.2.6 Covariance and Correlation
3.3.1 and 3.4.1 Probability Distributions, Binomial and Normal Distributions
3.4.1 The Normal (Gaussian) Probability Density Function
3.4.1 and 3.3.2 Applications of the Normal Distribution & the Poisson's Distribution

3. Hypothesis Testing
Week 5
Reading: Wessel 3.5, 4.1, 4.2.2, 4.2.4-5
3.5. Estimating sample means and its confidence interval
4.1 Null Hypothesis
4.2. Parametric Tests (Students t, Chi-squared, F tests)
   4.2.2 One and two sample test of means
   4.2.4 Estimating the variance of a population
   4.2.5 Two sample test of F-test of variance

Week 6
Reading: Wessel 4.2.6-7
4.2 Parametric Tests (Students t, Chi-squared, F tests)
   4.2.6 Chi-squared test of a pdf
   4.2.7 test of linear correlation

Week 7
Reading: Wessel 4.3, 5.1-4
4.3 Parametric vs. Non-Parametric tests
   4.3.4: Spearman's Non Parametric test for correlation
5.1-5.2 Matrices: General concepts and definitions

5.3-5.4 Matrix Addition, Dot Product, and Matrix Multiplication
ERTH 613 students: Meet with Helen to discuss ideas for your class project this and next week.

4. Linear (Matrix) Algebra and Least Squares Regression
Week 8
Reading: Wessel 5.5, 5.7
5.5 Determinant of a Matrix
5.7 Matrix Division and the Inverse Matrix
Mid-Term

Week 9
Reading: Wessel 5.9
  5.9.1 Simple Regression and Curve Fitting
  5.9.3 General linear least-square regression

5. Regression
Week 10
Reading: Wessel 6.1
6.1 Line Fitting Revisited: Confidence Intervals on True Slope, Intercept, and Regression Line
ERTH 613 students: 1-page description of the data you would like to analyze for your project and what you'd like to do with your analysis

Week 11
Reading: Burnham & Anderson [2004], Hobbs & Hobbs (2015), "Bayesian Models: A Statistical Primer for Ecologists", Princeton Univ. Press.
Model Selection, Cross Validation

Week 12
Reading: Wessel 4.2.8, Draper & Smith excerpt
4.2.8 Analysis of Variance (ANOVA) of Linear Regression, One-way, Two-way ANOVA

6. Sequences and Time Series Analysis
Week 13
Reading: Wessel 7.5-6
7.5 Autocorrelation
7.6 Cross-correlation

7. Spectral Analysis
Week 14:
Reading: Wessel 8.1-3
8.1 Spectral Analysis: Basic Terminology
8.2 Spectral Analysis: Fitting the Fourier Series
8.3 The Periodogram or Discrete Power Spectrum
ERTH 613 students: 5 minute progress reports on projects.

Week 15
Wrap up and work on projects.

Week 16
Wrap up, review, in-class presentations.

Title IX:

The University of Hawai'i is committed to providing a learning, working and living environment that promotes personal integrity, civility, and mutual respect and is free of all forms of sex discrimination and gender-based violence, including sexual assault, sexual harassment, gender-based harassment, domestic violence, dating violence, and stalking. If you or someone you know is experiencing any of these, the University has staff and resources on your campus to support and assist you. Staff can also direct you to resources that are in the community. Here are some of your options:

As members of the University faculty, your instructors are required to immediately report any incident of potential sex discrimination or gender-based violence to the campus Title IX Coordinator. Although the Title IX Coordinator and your instructors cannot guarantee confidentiality, you will still have options about how your case will be handled. Our goal is to make sure you are aware of the range of options available to you and have access to the resources and support you need.

If you wish to remain ANONYMOUS, speak with someone CONFIDENTIALLY, or would like to receive information and support in a CONFIDENTIAL setting, use the **confidential resources available here**: http://www.manoa.hawaii.edu/titleix/resources.html#confidential

If you wish to directly REPORT an incident of sex discrimination or gender-based violence including sexual assault, sexual harassment, gender-based harassment, domestic violence, dating violence or stalking as well as receive information and support, contact: Dee UwonoTitle IX Coordinator (808) 956-2299 t9uhm@hawaii.edu.

Student Conduct and Academic Integrity:

University guidelines for acceptable student conduct are very specific and will be strictly followed. Please read the guidelines (http://www.studentaffairs.manoa.hawaii.edu/policies/conduct_code/) and contact me if you have any concerns. Cheating, of any form, will not be tolerated. Blind copying of intellectual material (text) from resources such as books, journals, and the internet is plagiarism and is illegal. You are encouraged to work together on your problem sets, but all work turned in for grading (including computer programs) must be yours, and yours alone. There will be no collaborations during exams.

Counseling & Student Development Center
Counseling and Student Development Center (CSDC) offers an array of services to meet counseling and testing needs among students, staff, and faculty of the UH Mānoa campus. Our multidisciplinary staff includes psychologists, psychiatrists, graduate-level therapists, and counselor trainees. It is our mission to uphold excellence in quality of care that is respectful to the socio-cultural diversity of our clientele. We offer walk-in, individual, and group counseling, as well as career and psychological assessments. When a client's needs could be best addressed by providers outside of the CSDC, we offer appropriate referrals in

the community. We also provide outreach events to increase awareness on issues relevant to healthy campus lifestyle. For more information, please visit the Counseling & Student Development Center web site at: http://www.manoa.hawaii.edu/counseling/.

Disability Access:

The Earth Sciences Department will make every effort to assist those with disability and related access needs. For confidential services, please contact the Office for Students with Disabilities (known as "Kokua") located in the Queen Lili'uokalani Center for Student Services (Room 013): KOKUA Program; 2600 Campus Road; Honolulu, Hawaii 96822. Voice: 956-7511; Email: kokua@hawaii.edu; URL: www.hawaii.edu/kokua

Department of Public Safety:

 (808)956-6911 (Emergency) / (808)956-8211 (Non-Emergency) http://manoa.hawaii.edu/dps/

Basic Needs:

Basic needs include food and housing, childcare, mental health, financial resources and transportation, among others. Student basic needs security is critical for ensuring strong academic performance, persistence and graduation and overall student well being. If you or someone you know are experiencing basic needs insecurity, please see the following resources: https://www.hawaii.edu/student-basicneeds/.