GG413: Geological Data Analysis

Meetings: Tue/Thu 10:30-11:45, POST 702
Instructor/Office hours: Garrett Apuzen-Ito (gito@hawaii.edu) MWF 12:30-2:30, POST 810.
Prerequisites: Math242 (2nd semester calculus) GG250 (scientific programming using Matlab)
Textbook: Paul Wessel’s Lecture Notes. Recommended (optional) text: John C. Davis, Statistics and Data Analysis in Geology, 3rd Edition

OVERVIEW AND OBJECTIVES:
Quantitative skills are extremely important in the natural sciences. With the continued development of computer and internet technology, as well as advancements in data collection capabilities, the amount of data Earth scientists must process and interpret can be overwhelming. Being able to analyze data on a computer is a necessity and often a job requirement.

The main purpose of this course is to provide students with foundational understanding of the basic theory behind quantitative data analysis, and provide practical experience with real data sets using computer software (Matlab, Octave, or FreeMat). Students will learn the importance of knowing uncertainties, how they affect the significance of results, and how to assign confidence limits on computed solutions. Students will also...

• Learn how to apply exploratory data analysis techniques to characterize their data or discover structure 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
• Gain experience in doing formal hypothesis testing
• Be introduced to matrices, linear algebra, and least squares formalism for curve fitting and regression
• Explore various ways to examine sequential data
• Understand principals of spectral analysis and the key concepts of aliasing and leakage
• Be acquainted with statistical estimates and hypothesis testing relevant to directional data

Emphasis will be on techniques and data sets in the geosciences but the course is relevant to all fields of science.
STUDENT LEARNING OBJECTIVES
This course emphasizes two 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 geology and geophysics
- Students use the **scientific method** to solve a problem in earth science

FORMAT AND WORKLOAD
The class meets twice a week for lectures and for discussion of homework problems. You are encouraged to ask questions during class. Please be persistent: I tend to assume that *if there are no questions, then everyone must understand everything*. Homework will be assigned approximately weekly and will involve using Matlab to program and practice the techniques covered. There will be a mid-term and a final exam.

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 beginning of class on the 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 will be a weighted average of grades for homework (70%), mid term (15%), the final exam (15%).

EXAMPLE COURSE SYLLABUS

1. **Basic Statistical Concepts**
   **Week 1: Aug 27, 29**
   1.1 Classification of data
   1.2 Exploratory data analysis
   1.3 Error Analysis
   **Lecture Matlab scripts & example data**
   **HW #1** and **required datasets**

   **Week 2: Sept 3, 5 (HW #1 due)**
   1.4 Probability Basics
   **HW #2:** Probability

   **Week 3: Sept 10, 12 (HW #2 due)**
   1.5 The M&M’s of Statistics ([Davis pages on Central Limit Theorem](#))
   See **example script** for plotting the binomial and normal distributions.
   **HW3:** Statistics and Probability Distributions

2. **Hypothesis Testing**
   **Week 4: Sept 17, 19 (HW #3 due)**
   2.1 Null Hypothesis
   2.2. Parametric Tests (Student’s *t*, Chi-squared, *F* tests)
   **HW4:** Hypothesis Testing with Parametric Statistics
Week 5: Sept 24, 26 (HW #4 due)
2.2 Parametric Tests (Chi-squared goodness of fit test)
2.3 Non-Parametric Tests (sign test)
  
  **Hw5: Hypothesis Testing II:** see datasets “quakedays.d” and “rho.d”

Week 6: Oct. 1, 3 (HW #5 due)
2.3 Non-Parametric Tests (Mann-Whitney, Kolmogorov-Smirnov)
  
  **Hw6: Hypothesis Testing III,** see Matlab script “kolsmir.m”

3. Linear (Matrix) Algebra and Least Squares Inversion for Model Fitting
   
   Week 7: Oct. 8, 10 (HW #6 due)
3.1-3.9 Matrices and Matrix Math
3.10-3.11 Simple Regression and General Least Squares
  
  **Hw7: Least Squares Regression:** see dataset “hf.d”

Week 8: Oct. 15, 17
Review for Mid-term
  
  >>>> Midterm Thu. 10/17 (Covering material through HW #6) <<<<

Week 9: Oct. 22, 24 (HW #7 due)
3.11 General Least Squares
3.12 Weighted Least Squares
  
  Also **Draper & Smith excerpt**
  
  **Hw8: Lease Square Regression II:** see hawaii.dt, faultstep.d, and heaviside.m

4. Analysis of Variance (ANOVA)
   
   Week 10: Oct. 29, 31 (HW #8 due)
4.3 One-way ANOVA
4.4 Two-way ANOVA
  
  Also **Draper & Smith excerpt**
  
  **Hw9: ANOVA,** see Hw9_Prob1_Chromium.txt, Hw9_Prob2_NHawaiian.txt, and Hw9_Prob3_TensileStrength.txt

5. Sequences and Time Series Analysis
   
   Week 11: Nov. 5, 7 (HW #9 due)
5.1 Markov Chains
5.3 Series of Events
  
  **Hw10: Marcov Chains and Event Series.** See data file “aso.d”

Week 12: Nov. 12, 14 (HW #10 due)
5.8 Spectral Analysis
5.9 The “Periodogram”
  
  **Hw 11: Spectral Analysis.** See data file “honolulu_resampled.d”

Week 13: Nov. 19, 21 (HW #11 due)
5.9 The “Periodogram”
6.0 Analysis of Directional Data
Read Davis Hand out
Hw12: Analysis of Directional Data
See data files “Iceland_West.txt” and “Iceland_East.txt”, as well as Matlab script “polarhist.m”

Week 14: Nov. 26 (NO homework due this week)
6.0 Analysis of Directional Data
Happy Thanksgiving

Week 15: Dec. 3,5 (HW #12 due Thu 12/5)
Review for Final Exam

Week 16: Dec. 10,12
No Class

>>>> Final Exam is Thursday Dec 19, 9:45-11:45 a.m. <<<<