

Data analysis and error analysis

PHY 599 — Spring, 2010

<http://www.physics.nau.edu/~trilling/teaching/spring2010/phy599/>

Physical Sciences (Building 19) Room 218 TTh 9:35 a.m. – 10:50 a.m.

Instructor: Professor David E. Trilling

Course description:

This graduate course is intended to teach a set of critical data analysis and error analysis tools for use in any scientific or applied science application. We will cover a range of topics related to data analysis and error analysis. Chances are good that you will use a number of these tools in your future work, whether in physics or astronomy or another field, and in any of academia, research, or industry. We will not cover any of these topics in great detail. Instead, the goal is to expose you to a range of fundamental tools that are necessary for rigorous analysis of data and measurements. You would be amazed how many of your peers will not have this set of analytical tools. These skills will serve as a critical backbone for your future endeavors.

Course structure and approach:

It's a graduate course. I'm not going to load you up with problem sets and exams. In fact, there will be no exams. Instead, there will be a semester project in which you will present some new work that you have done, based on and extending from material that we have covered in the course. More on this later in the semester.

I will give you some exercises to do – more like practice than like homeworks.

Each student will have the opportunity (read: requirement) to teach one or more topic during the class. This is a good opportunity to learn something in detail, and to get practice teaching.

Some class meetings will be lectures. Some will be group work sessions. Some will be peer-taught. Towards the end of the semester we'll have computer time. And the last part of the semester will be presentations of your semester projects. Overall, I hope the class will be more seminar-like and less class-like. Don't be too formal, and for god's sake, don't just sit there quietly – discussion is encouraged.

Textbook and required materials:

Nothing required, but I strongly suggest that you acquire Bevington, third edition. Taylor's book is better written, but less rigorous; you might use that as a supplemental resource. Other resources may be useful as the course goes on.

You also need good, reliable access to a Unix/Linux computer. If you don't have that, we need to figure out how to get that for you.

The course web page:

The course web page is given at the very top of this syllabus. I'll post stuff there, including lecture notes, data sets, exercises, etc., throughout the course.

Office hours (how to find me and ask questions):

Just come find me. You know where my office is. Just don't come between 1:30–3:30 pm on Tuesdays and Thursdays, as I'm preparing for and teaching AST 183 during that time.

Grading and assignments:

Formally I will grade the exercises that I give you; your teaching presentations; and your final project. The final project will be worth 50% of your grade, and your teaching presentations and exercises will each be worth 25% of your grade.

A note about working together (statement on plagiarism and cheating):

Science works by sharing ideas. I encourage you to work together in this class. However, anything that has *only your name on it* should be written by you and you alone. Let me be clearer about the exercises: I encourage you to work together on these, by which I mean that you can discuss the best way to do the questions and compare answers. However, after working together, you should then write up your assignments by yourself. You should not have identical answers to anyone else in the class. If you do, you have cheated and perhaps plagiarized. This is absolutely not allowed in this class or this University, and I am very serious about this. Cheating pisses me off.

Here's the official statement: This course requires professional and ethical behavior. Plagiarism, or any form of cheating, violates this principle and will not be tolerated. The University regards acts of academic dishonesty as very serious offenses. Students charged with academic dishonesty are subject to the Arizona Board of Regents Code of Conduct and Procedures established by NAU.

Tentative list of topics to be covered (in some order)

- Introductory material (1 week)
 - Practicalities
 - What is data, what kinds of data sets are of interest
 - Practice data sets – yours and mine
 - Computer basics (Unix/Linux); how to do exercises for this class
- Error analysis (1–2 weeks)
 - Measurements; precision and accuracy
 - Sampling
 - Systematic, random, calibration errors
 - Error on mean, mean of errors, etc.
 - Propagation of errors, RSS, etc.
- Data analysis and properties of data (1–2 weeks)
 - Standard deviation, confidence intervals
 - Binomial, Gaussian, Poisson
- Data handling (3 weeks)
 - Significance
 - Best fit line, least squares, curve fitting
 - Correlations
 - χ^2 , likelihoods, f test
 - t test, K-S test
- Modeling (2 weeks)
 - Monte Carlo
 - Bayesian analysis
- Computer skills (3 weeks)

- Intro to programming – loops, variables, debugging, etc.
 - `awk` and `sed`
 - Python
 - Perl
 - MySQL
- Your projects (3 weeks)