Physics 252C Syllabus

Winter Quarter 2013

1 Welcome to Physics 252C

The aim of this course is to review some of the most commonly used data analysis techniques in high-energy nuclear and particle physics. We will review some of the fundamentals of probability, statistics with emphasis on the relevant topics which arise in data analysis in the high-energy physics. We will practice on how to apply the analysis techniques in commonly encountered practical problems in nuclear and particle physics, including parameter estimation, simulation, and advanced error analysis (both statistical and systematic).

As such, this course will have both a theoretical and practical aspects which will run in parallel. The goal will be on applying the theoretical concepts to actual calculations with data. Students will write their own computer programs to carry out calculations using supplied data sets.

2 Course Information and Contact

- Instructor: Prof. Manuel Calderón de la Barca Sánchez
- E-mail:`calderon@physics.ucdavis.edu`
- Office: Rm 397A
- Office Hours: Wed 11am-12:00pm
- TA: Christopher Flores
- E-mail:`chrflores@ucdavis.edu`
- Office: Rm 393
- Office Hours: Wed 4:30pm-5:30pm

3 Lecture Hours

Lectures are **Tuesday and Thursday, 9:00-10:20 am**. We will meet in **Phy 285**.

3.1 Text

G. Cowan, "Statistical Data Analysis", Oxford Science Publications, 1998.

Additional resources:

- L. Lyons, "Statistics for nuclear and particle physicists", Cambridge University Press, 1986.
- Particle Data Group, review of Probability, Statistics and Monte Carlo techniques.
- TRandom Pitfalls, Joel Henrich (University of Pennsylvania)
- The ROOT framework and the RooFit Package

- Minuit User Guide.
- Maximum Likelihood Primer, C. Blocker (Brandeis University).
- Unified approach to the classical statistical analysis of small signals, Gary Feldman and Robert Cousins, Phys. Rev. D. 57 3873 (1998).
- Presentation of search results: the CL_s technique, A. L. Read (University of Oslo).

4 252C Topics

We will follow the topics covered in the text, but we might skip some, emphasize some a bit more, and add some recent topics not covered in the text. Some of the things that we will plan to cover:

1. Probability and statistics fundamentals.
2. Probability distribution functions
3. Monte Carlo methods
4. Parameter estimation, Maximum Likelihood and least squares method
5. Statistical errors, confidence intervals and limits
6. Signal significance
7. Systematic uncertainties
8. Unfolding
9. Efficiency estimation
10. Multivariate analysis techniques

5 Grading

There will be weekly homework assignments including both theoretical problems and practical data analysis. The final course grade is based on class participation (20%), the homework (50%), and a final project of the student's choice which will be presented in class (15%) and for which a very short report is written (15%).

6 Final Project

The main requirement on the final project topic is that it involve some statistical calculation. It can be an extrapolation of the homework, a multivariate topic like boosted decision trees or neural networks (utilizing pre-existing packages if you want), developing a calculation such as the Kolmogorov-Smirnoff test or any number of things that we covered in class. If you are involved in research with some group and want to highlight the statistical treatment of your data, that would be even better.

Your project will be presented in a short, clear talk (15 minutes - I recommend no more than 10 slides) and a short paper. The presentations will be March 21 at 9am (barring conflicts with other finals) in Room 393 and the paper due by March 22.