

PHYS 6810 Applied Statistics and Data Analysis in Physics

Spring Semester, 2016

3 Credits

Course Time: Mon/Wed 10:30-11:45am

Instructors: Profs. Alexander van der Horst, Michael Doering, Oleg Kargaltsev

Course Place: Corcoran Hall, room 209

Office Hours: Mon 3:30-4:30pm or by appointment

Office: Jan 11 – Mar 7: Samson Hall, room 312 (Doering) or room 213 (van der Horst)

Mar 21 – Apr 25: Samson Hall, room 214 (Kargaltsev)

Contact information (E-Mail): ajvanderhorst@gwu.edu

Prerequisites:

Programming experience and good working knowledge of *either* Matlab, Mathematica, Python, IDL, or R. *The default language for demonstration will be Mathematica (free for GW students). You will need to have a laptop for this class.*

Multivariable Calculus (MATH 2233 or equivalent), Linear Algebra (e.g., MATH 2184 or equivalent); calculus-based introductory physics (e.g., PHYS 1021, 1022, and 1023 or equivalent physics courses).

Web Site Resources:

http://www-library.desy.de/preparch/books/vstatmp_engl.pdf

<http://astrostatistics.psu.edu/>

Mathematica license for GW students: <https://seascf.seas.gwu.edu/wolfram-mathematica>

Additional resources can be found on the course page in the Blackboard:

<https://blackboard.gwu.edu/>

Synopsis

After taking this course, students will be able to explain the challenges and the best practices in statistical inference methods applied to physical science data, apply modern statistical methods, and create informative and appealing visualizations of the data and inferred statistically-sound trends, correlations and dependencies. Students will make extensive use of real data and numerical methods to illustrate abstract statistical concepts.

Students will complete projects involving hands-on analysis of various datasets as part of lab-like class activities (working in groups) and individual homework assignments. The overarching goal of this course is to help students to develop analytical and practical skills for physical (and other) data analysis and interpretation using solid statistical methods well recognized within the discipline.

Recommended reading materials for this course:

Bevington P. and Robinson D. K. “Data Reduction and Error Analysis for the Physical Sciences”, 2002, 3rd edition, ISBN-13: 978-0072472271

Bohm G. Zech, G. “Introduction to Statistics and Data Analysis for Physicists”,
http://www-library.desy.de/preparch/books/vstatmp_engl.pdf

John E. Freund's “Mathematical Statistics with Applications”, 2012, 8th edition, ISBN-13: 978-0321807090

Feigelson, E. Babu, J. “Modern Statistical Methods for Astronomy: With R Applications”, 2012, ISBN-13: 978-0521767279

Other materials provided by the instructors.

Table 1. Course Objectives

Content Objectives	Skills Objectives
<ol style="list-style-type: none">1. Be able to explain and discuss:<ul style="list-style-type: none">▪ the need and the nature of statistical data analysis▪ be able to identify, explain, and critically evaluate the statistical methods used in the research papers related to physical sciences2. Demonstrate deep understanding of fundamental statistical principles applicable to data analysis in physics and related disciplines3. Acquire knowledge of a broad spectrum of modern statistical software tools.	<ol style="list-style-type: none">1. Given a physical problem identify the statistical methods required to evaluate the impact of the uncertainties on the measurement or the choice of model2. Identify and use appropriate software tools to carry out statistical analysis of various problems3. Be able to apply statistical methods to a broad range of problems including those outside Physics (e.g., finance, signal processing)
<p>Assessments and tools:</p> <ul style="list-style-type: none">▪ Graded homeworks.▪ Reading Check-Up (RCU) Quizzes▪ Graded in-class projects assigned to groups▪ Graded Final Project	

Course Schedule. A preliminary schedule of topics for the course is shown in the table below. Various faculty members involved in the course as mentors and guest lecturers will enrich the course, and graduate students and post docs already involved in research will be invited to classes.

Table 2. Classroom Schedule and Activities

Class dates	Topics for Class Lectures and Activities
1/11 & 1/13	Uncertainties and errors in physical data. Real life examples. Laws of Probability.
1/20 & 1/25	Random variables, Expectation value, (co)variance, Practical applications.
1/27 & 2/1	Distributions and samples. Software tools.
2/3 & 2/8	Binomial, Poisson, and Normal Distributions. Central Limit Theorem. Student's t-distribution, F-distribution.
2/10 & 2/17	Tests of Hypotheses. F-test. Data fitting, chi-squared and its minimization, chi-squared test. Maximum Likelihood.
2/22 & 2/24	Best-fit parameter uncertainties, confidence intervals, covariance, and multidimensional fitting. Non-parametric statistics. Practical examples of fitting with real data.
2/29 & 3/2	Bootstrap and Monte Carlo methods. Practical applications.
3/7 & 3/9	Bayes' theorem. Bayesian Information Criterion. Practical applications.
3/21 & 3/23	Time series analysis. Fourier, Wavelets, and other methods. Practical applications.
3/28 & 3/30	Clustering and classification. Spatial statistics.
4/4 & 4/6	Supervised & unsupervised machine learning with applications to classification.
4/11 & 4/13	Practical applications of clustering and classification.
4/18 & 4/20	Data visualization techniques, linked view, dynamic updating. Practical applications.
4/25 & 4/27	Intro to High-Performance computing. Practical applications.

Course Overview

Course Format

This course will be a mix of lectures and lab-like activities. The activities can involve collaborative work in groups. It is critically important for these activities that you come to class having completed your reading assignments ahead of the class. It is expected that you will have read the recommended materials before each class. This way larger portion of the class time will be spent on practical applications of the methods and concepts described in the assignments.

Classroom Course Structure

Class time: You must read assigned textbook and other materials in order to get the most out of our in-class activities. Class time (outside the lecturing) will be spent on applying the theoretical methods and concepts described in the assigned reading, which will help you to obtain a better and deeper understanding of the underlying ideas and challenges. The class time is the best opportunity for asking questions about difficult concepts and anything you could not understand in the assigned reading. You are strongly encouraged to ask questions and initiate discussions in class.

Evaluation Criteria (how the course grade will be determined)

Reading Check-Up (RCU) Quizzes (10%): Multiple-choice quizzes used to check students' learning progress and preparation for class.

In class group projects (30%): A variety of short projects will require use of computational methods with your preferred software (choices are: Matlab, Mathematica, Python, IDL, or R). You must have your own copy of the software or be able to obtain it from the University. *You will need to have a laptop for this class.*

Individual Homework Assignments (35%): Regular homework assignments will include short computational projects to ensure good understanding of statistical methods and develop practical application skills.

Final Project (25%): A take-home project that will involve solving a data-based problem using statistical and computational methods.

Numerical course grades translate into letter grades using the following scale:

≥ 94.00	A	70.00 – 73.99	C
90.00 – 93.99	A-	66.00 – 69.99	C-
86.00 – 89.99	B+	61.00 – 65.99	D+
82.00 – 85.99	B	55.00 – 60.99	D
78.00 – 81.99	B-	50.00 – 54.99	D-
74.00-77.99	C+	≤ 49.99	F

Absences and Excuses: All requests to have an absence excused must follow standard University policy, that is, there must be a written note and supporting documentation to the instructor, explaining the absence. Verbal explanations cannot be accepted.

Academic Integrity

You must fully comply with the GW Code of Academic Integrity. It states: “Academic dishonesty is defined as cheating of any kind, including misrepresenting one's own work, taking credit for the work of others without crediting them and without appropriate authorization, and the fabrication of information.” For the remainder of the code, see: <http://www.gwu.edu/~ntegrity/code.html>

Support for Students Outside the Classroom

Disability Support Services (DSS)

Any student who may need an accommodation based on the potential impact of a disability should contact the Disability Support Services office at 202-994-8250 in the Marvin Center, Suite 242, to establish eligibility and to coordinate reasonable accommodations. For additional information please refer to: <http://gwired.gwu.edu/dss/>

Mental Health Services 202-994-5300

The Mental Health Services staff in the Colonial Health Center supports your academic and social success as you adjust to college life. Through individual and group counseling, crisis intervention, assessments, and referrals, and by partnering with the Active Minds student organization, GW has created a community of care. After-hours emergency care is also available 24/7 when you need support in a crisis. <http://counselingcenter.gwu.edu/>

Religious Holidays

The Faculty Senate has set guidelines pertaining to the observation of religious holidays. These have become university policy and are as follows:

1. that students notify faculty during the first week of the semester of their intention to be absent from class on their day(s) of religious observance.

2. that faculty continue to extend to these students the courtesy of absence without penalty on such occasions, including permission to make up examinations.
3. that faculty who intend to observe a religious holiday arrange at the beginning of the semester to re-schedule missed classes or to make other provisions for their course-related activities.
4. that, prior to each semester, the administration circulate to faculty a schedule of religious holidays most frequently observed by GW students
5. that student members of other religious groups are also entitled to the same courtesies and accommodations.
6. that the administration conveys this policy to students by including it in the Schedule of Classes and other places deemed appropriate.

Security

In the case of an emergency, if at all possible, the class should shelter in place. If the building that the class is in is affected, follow the evacuation procedures for the building. After evacuation, seek shelter at a predetermined rendezvous location.