PHY657

Statistics and Data Analysis in Physics

Syllabus

Instructor	Caleb Fink	Phone	TBA
Office	PHY 219	E-mail	cwfink@syr.edu
Office Hours	By Appointment		

Course Description:

This course provides the skills needed to analyze experimental and observational data without getting lost in abstract general principles. While these skills are critical to experimental physics and astrophysics, they are also relevant to numerical approaches in theoretical physics. Particular emphasis is devoted to modern machine learning techniques widely used to enhance experimenta sensitivity to rare phenomena, improved detection techniques and the solution of complex theoretical problems.

Additional Information:

Modern physics research requires knowledge of statistical methods and increasingly complex data analysis techniques.

A common theme in our experimental work is to compare data acquired with a specific physical model. Techniques to assess quantitatively the degree of data/experiment consistency will be examined. Statistics provides the natural mathematical framework for this comparison. The frequentist and Bayesian approaches will be introduced. In light of the fast-paced development of machine-learning-based algorithms, understanding Bayesian statistics is increasingly important and will be explored.

The world of scientific computing and, more generally, the world at large are deeply influenced by the rapid development of generative AI and large language models: we will explore different approaches to how these fast-evolving tools affect our work and society.

This course aims to provide the skills to analyze experimental and observational data without getting lost in abstract general principles. In addition, the development of critical skills to validate the outcome of a computational effort without threading the module as a black box will be emphasized. The vast influence of AI in our approach to our profession, and, more in general, in almost all aspects of our society, will be examined. Generative AI and large language models will be discussed.

Course Outline:
The course is organized around the modules below. In the course of the semester, we may adjust the module contents on the basis of the discussions and hot topics in specific areas of scientific inquiry. The detailed week work plan will be developed as we progress in the course, as the duration of each module will be tailored to the learning objective of the class. Lecture notes and planned activities will be posted on one week before the start of the module.

Unit	**Topic**
Module 1	Introduction – Fitting and model selection
Module 2	Probability distribution
Module 3	Linear Models for Regression
Module 4	Linear Models for Classification
Module 5	Neural Networks
Module 6	Graphical models
Module 7	Boosted decision trees
Module 8	Unsupervised learning and generative models

The module structure may be modified during the semester if additional modules are helpful to the learning outcomes.

Prerequisite / Co-requisite:

PHY607

Audience:

The course is intended for graduate students interested in applying statistical and machine learning techniques to data interpretation, model building, and analysis.

Credits:

3

Learning Objectives:

After taking this course, students will be able to:

1. Understand and assess experimental errors.
2. Evaluate the quality of a fitting algorithm or a specific hypothesis testing.
3. Use computational algorithms with critical reflection and not as “black boxes” without understanding their applicability or limitations.
4. Navigate complex software packages used in modern physics research and be able to apply them to current research topics.

Texts / Supplies – Recommended:

The main reference textbook is "Pattern Recognition and machine learning" by Christopher M. Bishop, it should be available at the link below

https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf

Additional resources will be added to PHY657 Google Drive

Several textbooks on the topic are useful references, and there is a vast body of reference material on the web. Examples of reference books are:

1. W.T. Eade, D. Drijard, F.E. James, M. Roos, B. Sadoulet "Statistical Methods in Experimental Physics."
2. Frederick James "Statistical Methods In Experimental Physics."
3. Louis Lyons "Statistics for nuclear and particle physics"
4. R.J. Barlow "Statistics A guide to the Use of Statistical Methods in the Physical Sciences
5. Probabilistic Machine Learning: An Introduction by Kevin P Murphy

Course Requirements and Expectations:

Each learning module includes assignments and topics for in-class discussion. For each module, there is some assigned coding to be done.

Please use Jupyter notebooks to develop your code. The outcomes of running the code with specific parameters must be documented and discussed in a short report submitted as a printed document before or on the date specified as the deadline included in the module reference notes. The code files must be uploaded by creating a branch with your team name and uploading the material to your branch. The report is individual, and you must submit only one notebook per team.

Methodology for in-class activities:

Each calls will start with a short discussion of the learning goals, with some mini-lecture or discussion session on the work done in class. A detailed schedule with the activities for each module will be uploaded to the class drive in advance of the session. During the first meeting, teams will be formed, and most of the work will be implemented with your partner. The report is where you examine your work and reflect on the technical issues and
the key takeaways from your work. Additional material to allow more in-depth exploration of each module will be posted on the GitHub repository for the class.

During the semester, we will establish connections to key research areas pursued in the department.

In class-etiquette

Please start your activity by discussing with your laboratory partner. The majority of the coding work is expected to happen in class.
Please do not:

1. Divide the coding work: each module is expected to be the outcome of a collaborative effort. During the in-class discussion you should be prepared to discuss all the joint work.
2. Refrain from multi-group discussion. If there is a topic that benefit from some input from me or from a more collective discussion, please bring it to my attention.
3. Be respectful of the different pace that different groups will adopt. This is a stron motivator for limiting large discussion session to planned times.

Relevant GitHub repository

https://github.com/cwfink1/phys657-spring2026#

Grading:

The final grade is based on the work that you submit during the semester [25% coding, 25% reports, 30% participation in the in-class discussion, 20% paper and presentation for the final in-class workshop on **AI: ethical, societal and technical issues**

Percentages are converted to letter grades based on the following Table:

92-100	A	72-75	C+
88-91	A-	68-71	C
84-87	B+	64-67	C-
80-83	B	60-63	D
76-79	B-	0-59	F

Each module will be graded based on the following:

1. Completeness of the work done
2. Clarity of the report
3. Clarity of your presentation in the in-class module discussion
4. To obtain full credit for one module, you must submit your work on time.

Academic Integrity:

As a pre-eminent and inclusive student-focused research institution, Syracuse University considers academic integrity at the forefront of learning, serving as a core value and guiding pillar of education. Syracuse University's Academic Integrity Policy provides students with the necessary guidelines to complete academic work with integrity throughout their studies. Students are required to uphold both course-specific and university-wide academic integrity expectations such as crediting your sources, doing your own work, communicating honestly, and supporting academic integrity. The full Syracuse University Academic Integrity Policy can be found by visiting class.syr/edu, selecting, "Academic Integrity," and "Expectations and Policy."

Upholding Academic Integrity includes the protection of faculty's intellectual property. Students should not upload, distribute, or share instructors' course materials, including presentations, assignments, exams, or other evaluative materials without permission. Using websites that charge fees or require uploading of course material (e.g., Chegg, Course Hero) to obtain exam solutions or assignments completed by others, which are then presented as your own violates academic integrity expectations in this course and may be classified as a Level 3 violation. All academic integrity expectations that apply to in-person assignments, quizzes, and exams also apply online.

Students found in violation of the policy are subject to grade sanctions determined by the course instructor and non-grade sanctions determined by the School or College where the course is offered. Students may not drop or withdraw from courses in which they face a suspected violation. Any established violation in this course may result in course failure regardless of violation level.

Artificial Intelligence Use

Syracuse's academic integrity expectations extend to the fast-growing realm of artificial intelligence (AI) and represent the importance of understanding, exploring, and evaluating emerging technologies.

Artificial Intelligence cases may not be submitted to the Center for Learning and Student Success with AI-detection results as the only evidence of an Academic Integrity Violation.

To view the policy in its entirety, please visit: Academic Integrity: Expectations and Policy.

ATTENDANCE POLICY (including Absence Notification)

Attendance in classes is expected in all courses at Syracuse University. It is a federal requirement that faculty promptly notify the university of students who do not attend or cease to attend any class. Faculty will use Early-Semester Progress Reports and Mid-Semester Progress Reports in Orange SUccess to alert the Registrar and Financial Aid Office on non-attendance. For more information visit: Information for Students: Non-attendance or Stopped Attending

If a student is unable to participate in-person or virtually for an extended period of time (48 hours or more), the student may request an absence notification from their home school/college Dean's Office or through Student Outreach and Support office. Instructors will be notified via the "Absence Notification" flag in Orange SUccess.

Barnes Center at the Arch (Health, Counseling, etc.) staff will not provide medical excuse notes for students. When Barnes Center staff determine it is medically necessary to remove a student from classes, they will coordinate with Student Outreach and Support case management staff to provide appropriate notification to faculty through Orange Success. For absences lasting less than 48 hours, students are encouraged to discuss academic arrangements directly with their faculty.

Additional information may be found at Student Outreach and Support: Absence Notifications.

a student's housing location/room.

DISCRIMINATION OR HARASSMENT

The University does not discriminate and prohibits harassment or discrimination related to any protected category including creed, ethnicity, citizenship, sexual orientation, national origin, sex, gender, pregnancy, disability, marital status, age, race, color, veteran status, military status, religion, sexual orientation, domestic violence status, genetic information, gender identity, gender expression or perceived gender.

Any complaint of discrimination or harassment related to any of these protected bases should be reported to Sheila Johnson-Willis, the University's Chief Equal Opportunity & Title IX Officer. She is responsible for coordinating compliance efforts under various laws including Titles VI, VII, IX and Section 504 of the Rehabilitation Act. She can be contacted at Equal Opportunity,

Inclusion, and Resolution Services, 005 Steele Hall, Syracuse University, Syracuse, NY 13244-1120; by email: titleix@syr.edu; or by telephone: 315-443-0211.

HEALTH & WELLNESS CONSIDERATIONS

Mental health and overall well-being are significant predictors of academic success. As such it is essential that during your college experience you develop the skills and resources effectively to navigate stress, anxiety, depression, and other mental health concerns. Please familiarize yourself with the range of resources the Barnes Center provides (https://ese.syr.edu/bewell/) and seek out support for mental health concerns as needed. Counseling services are available 24/7, 365 days, at 315-443-8000.

ORANGE ALERT

ORANGE ALERT, Syracuse University's crisis notification system, uses text messages, phone, and email alerts to provide rapid notification and instructions to members of the University community in the event of a crisis in progress. Crises could include an individual who is considered armed and dangerous, a hazardous materials incident, an explosion, or any other event in which there is an immediate threat of physical harm or death to campus community members. We recognize that faculty may consider activated cell phones as an interruption to their class. However, the Department of Public Safety recommends that faculty members leave their own cell phones on vibrate in order to receive text messages about a potential emergency situation. It is also recommended that faculty designate several class members to leave their cell phones on vibrate in order to receive notifications in the event of a critical incident. ORANGE ALERT contact information for students, faculty, and staff is drawn from the MySlice online information system; please keep your contact information current.

In the event of an emergency, please use one of the following numbers to reach us:

From any phone: 315.443.2224
From your cell phone: #78 (#SU)
Campus landline: 711

For more information on ORANGE ALERT, including how to update your contact information, visit the DPS website

Accommodations:

Syracuse University values diversity and inclusion; we are committed to a climate of mutual respect and full participation. There may be aspects of the instruction or design of this course that result in barriers to your inclusion and full participation in this course. I invite any student to contact me to discuss strategies and/or accommodations (academic adjustments) that may be essential to your success and to collaborate with the Center for Disability Resources (CDR) in this process.

If you would like to discuss disability accommodations or register with CDR, please visit Center for Disability Resources (https://disabilityservices.syr.edu/). Please call (315) 443-4498 or email disabilityresources@syr.edu for more detailed information.

Faith Tradition Observances:

Syracuse University's Religious Observances Policy (https://policies.syr.edu/policies/university-governance-ethics-integrity-and-legal-compliance/religious-observances-policy/) recognizes the diversity of faiths represented in the campus community and protects the rights of students, faculty, and staff to observe religious holy days according to their traditions. Under the policy, students are given an opportunity to make up any examination, study, or work requirements that may be missed due to a religious observance, provided they notify their instructors no later than the academic drop deadline. For observances occurring before the drop deadline, notification is required at least two academic days in advance. Students may enter their observances in MySlide under Student Services/Enrollment/My Religious Observances/Add a Notification.