Essential Mathematics and Statistics for Scientists

Carnegie Mellon University
School of Computer Science

Computational Biology Department

02-680 – Fall 2023

Revised: September 5, 2023
(change since start of classes marked in orange)

Class Location and Times: Posner Hall (POS) 153,
Tuesday & Thursday 3:30pm–4:50pm (Lecture),
Friday 5:00pm-6:20pm (Recitation)

Course Description: This course rigorously introduces fundamental topics in mathematics and statistics to first-year master's students as preparation for more advanced computational coursework. Topics are sampled from information theory, graph theory, proof techniques, phylogenetics, combinatorics, set theory, linear algebra, neural networks, probability distributions and densities, multivariate probability distributions, maximum likelihood estimation, statistical inference, hypothesis testing, Bayesian inference, and stochastic processes. Students completing this course will obtain a broad skillset of mathematical techniques and statistical inference as well as a deep understanding of mathematical proof. They will have the quantitative foundation to immediately step into an introductory master's level machine learning or automation course. This background will also serve students well in advanced courses that apply concepts in machine learning to scientific datasets, such as 02-710 (Computational Genomics) or 02-750 (Automation of Biological Research). The course grade will be computed as the result of homework assignments, midterm tests, and class participation.

Prerequisite: High school mathematics (including single- and multi-variable calculus) and ability to reason mathematically.

Course Goals: Students completing this course will have a broad overview of mathematical techniques, as well as a deep understanding of proof techniques and statistical tests. They will have the quantitative foundation to immediately step into a machine learning or automation course aimed at computational biology or other science MS students. This background will also serve them well in advanced courses that apply concepts in machine learning to scientific datasets, such as 02-710 (Computational Genomics).

Contents

1 Instructional Staff

Instructors

Dr. Dan DeBlasio	email:	deblasio@cmu.edu
	office:	GHC 7707
	office hours:	Tuesday 11:00am-12:00pm
	appointments:	calendly.deblasiolab.org

Dr. Jose Lugo-Martinez	email:	jlugomar@andrew.cmu.edu
	office:	GHC 7411
	office hours:	Friday 12:00-1:00pm
	appointments:	by email

Teaching Assistants

| **Peiran Jiang** | email: | peiran@cmu.edu |
| | office hours: | 10:00-11:00am Thursdays GHC 7404 (/Zoom, see Piazza) |

| **Vratin Srivastava** | email: | vratins@andrew.cmu.edu |
| | office hours: | 5:00-6:00pm Mondays GHC 7404 |

2 Logistics

Course Topic Schedule: Can be found on Google Sheets[1] and is subject to change.

Topics covered this semester:

- Sets and functions
- Linear algebra
- Matrices and vectors
- Probabilities
- Probability distributions
- Probability densities
- Multivariate distributions
- Maximum likelihood estimation
- Statistical inference
- Hypothesis testing

Textbook: There are no required textbooks for this course. Below are textbooks that can be good references:

- **Mathematics for Machine Learning** by Deisenroth, Faisal, Ong
- **All of Statistics: A Concise Course in Statistical Inference** by Larry Wasserman
- **Probability and Statistics** by Morris DeGroot and Mark Schervish
- **Linear Algebra Done Right** by Sheldon Axler

Communication platforms:

- **Canvas Homepage.** The course homepage will be hosted on Canvas. Canvas will be used for attendance and as a central repository for grades. You should be automatically enrolled at `https://canvas.cmu.edu/courses/36680`.
- **Discussion Forum.** An online forum is provided on Piazza as an area for discussion and questions. The forum will be moderated by the course staff who will respond to questions, but students are encouraged to help each other via discussion. However, assignment specifics should not be discussed — any hints will be provided by the teaching staff. You can find the class on Piazza at `https://piazza.com/cmu/fall2023/02680`.

3 Grading

Grades are communicated to students in a timely manner. It is the students' responsibility to keep track of their grades by compiling the grades they receive. Your semester grade will be based on a combination of homework assignments, weekly quizzes, class participation, mid-term assessment, and a final exam.

The approximate percentages are as follows:

---

[1]`https://docs.google.com/spreadsheets/d/1VhhVtQ6OiOiUol-YNR7yw1RdQCwQw4jsiwoejJQ6trk`

	+		-
50%	Homework		
10%	Attendance & Participation		
40%	Exams (20% per)		

The base percentage-score-to-letter-grade conversion for this course is as follows:

	+		-
A	98%	93%	90%
B	88%	83%	80%
C	78%	73%	70%
D	68%	63%	—

Any lower scores will result in a grade of "R". These minimums may be lowered without notice but will not be raised.

3.1 Homework

Written homework assignments will test your knowledge of the material covered in class. Homework assignment will be made available at least one week before the due date.

All writing assignments are to be submitted on Canvas on the day it is due. Any assignment turned in after the due date will be deducted 5 percentage points for the first day it is late, and 20 percentage points for every day after that. If you have extenuating circumstances, just communicate with us and we'll do what we can to accommodate.

Homework will be submitted as a `pdf` document which will need to be generated using an editor of the students choosing. Due to the reliance of this course on mathematical notation the use of LaTeX is encouraged, If using another editor, students should utilize the editor's equation editor to ensure the correct answers are being communicated. To assist those students using LaTeX, the source code for each assignment will be provided in addition to the `pdf`.

> **Cheating policy**
>
> All work must be your own. Unauthorized collaboration or plagiarism will result in a negative grade (e.g., a homework worth 100 points will be factored in as a -100 points towards your final grade) and will be reported to your academic advisor and dean.
>
> See Section 5.

3.1.1 Regrading of Assignments

In the event that you feel an assignment was graded incorrectly, you must make the request in writing within one week of receiving the graded material. Your request should

include a reference to the section of the assignment you believe was incorrectly graded. By asking for your assignment to be regraded, the whole assignment will be regraded and it is possible that fewer points may be awarded than the previous grade in light of new consideration.

3.2 Attendance & Participation

Attendance will be taken, and we will have occasional in-class exercises that serve to reinforce the concepts we have covered. These exercises will not be graded, but participation will be expected in order to receive a complete grade for that day. You are allowed three "dropped" attendance grades without penalty. These can be used for any purpose.

All classes will be live. All efforts will be made to record lectures, but will only be made available in extreme circumstances.

3.3 Exams

The two exams will test your knowledge of the material from the class. The first midterm will be held in class, and the second midterm will be held during the university's scheduled time. Dates for the exams will be announced in class, and can be found on the course schedule when available.

The exams will not be cumulative: midterm 2 will cover material encountered after midterm 1. That having been said, later material in the class may build upon the techniques covered in the earlier material.

4 Expectations

- Please note, the use of cell phones is not permitted during lecture. These should be left in your backpack or not brought to class.
- Laptop computers may be used for taking notes or for in-class assignments but cannot be used for other activities during lecture.
- No student may record any classroom or laboratory activity without the express written consent of the instructor. If a student believes that he/she is disabled and needs to record or tape classroom activities, he/she should contact the Office of Equal Opportunity Services, Disability Resources to request an appropriate accommodation.

Please keep in mind that these guidelines are necessary to maintain an environment that is safe and conducive for learning.

4.1 Expectations of Students and Instructors

The instructors and teaching assistants have the right to expect the following of students: (Adapted and modified from those developed by Howard Culbertson at Southern Nazarene University).

1. Students will arrive to class on time and will be prepared for the lecture.
2. Students will turn in assignments on time (see policy on assignment due dates).
3. Students will immediately inform the instructor or the teaching assistants if extenuating circumstances prevent the student from attending a lecture.
4. Students will follow the code of conduct regarding academic integrity, cheating, plagiarism, and collaboration as outlined in the syllabus.
5. Students will seek assistance when they need it.
6. If contacted by the instructor or teaching assistant, students will respond within 24 hours during the week and 48 hours on weekends.

The students have the right to expect the following of the instructor and teaching assistants:

1. A syllabus that describes class procedures, policies, and a course description will be provided.
2. Class sessions that will start and end on time.
3. Any changes to the course schedule will be provided to the students within 48 hours of the change.
4. The instructor will be available outside class either during their posted office hours or during other pre-arranged times.
5. If contacted by a student, the primary instructors or teaching assistants will respond within 24 hours during weekdays and 48 hours on weekends.

5 Academic Integrity

(See: https://www.cmu.edu/policies/student-and-student-life/academic-integrity.html)

Students in all CMU programs, because they are members of an academic community dedicated to the achievement of excellence, are expected to meet the highest standards of personal, ethical, and moral conduct possible. These standards require personal integrity, a commitment to honesty without compromise, as well as truth without equivocation, and a willingness to place the good of the community above the good of the self. Obligations once undertaken must be met, commitments kept. Rarely can the life of a student in an academic community be so private t hat it will not affect the community as a whole or that the standards above do not apply. The discovery, advancement, and communication of knowledge are not possible without a commitment to these standards. Creativity cannot exist without acknowledgment of the creativity of others. New knowledge cannot be developed without credit for prior knowledge. Without the ability to trust that these principles will be observed, an academic community cannot exist. The commitment of

its faculty, staff and students to these standards contributes to the high respect in which the CMU degree is held. Students must not destroy that respect by their failure to meet these standards. Students who cannot meet them should voluntarily withdraw from this course.

5.1 Use of Generative AI in this course

To best support your own learning, you should complete all graded assignments in this course yourself, without any use of generative artificial intelligence (AI). Please refrain from using AI tools to generate any content (text, video, audio, images, code, etc.) for an assignment or classroom exercise. Passing off any AI generated content as your own (e.g., cutting and pasting content into written assignments, or paraphrasing AI content) constitutes a violation of CMU's academic integrity policy[2]. If you have any questions about using generative AI in this course please email or talk to us.

5.2 Cheating and Plagiarism

Students in at CMU are engaged in preparation for professional activity of the highest standards. Each profession constrains its members with both ethical responsibilities and disciplinary limits. To assure the validity of the learning experience a university establishes clear standards for student work. In order to deter and detect plagiarism, online tools and other resources are used in this class. In any presentation, creative, artistic, or research, it is the ethical responsibility of each student to identify the conceptual sources of the work submitted. Failure to do so is dishonest and is the basis for a charge of cheating or plagiarism, which is subject to disciplinary action.

Cheating includes but is not necessarily limited to:

1. Plagiarism, explained below.
2. Submission of work that is not the student's own for papers, assignments, or exams.
3. Submission or use of falsified data.
4. Theft of or unauthorized access to an exam.
5. Use of an alternate, stand-in or proxy during an examination.
6. Use of unauthorized material including textbooks, notes or computer programs in the preparation of an assignment or during an examination.
7. Supplying or communicating in any way unauthorized information to another student for the preparation of an assignment or during an examination.
8. Collaboration in the preparation of an assignment. Unless specifically permitted or required by the instructor, collaboration will usually be viewed by the university as cheating. Each student, therefore, is responsible for understanding the policies of the department offering any course as they refer to the amount of help and collaboration permitted in preparation of assignments.

---

[2]https://www.cmu.edu/policies/student-and-student-life/academic-integrity.html

9. Submission of the same work for credit in two courses without obtaining the permission of the instructors beforehand.

Plagiarism includes, but is not limited to, failure to indicate the source with quotation marks or footnotes where appropriate if any of the following are reproduced in the work submitted by a student:

1. A phrase, written or musical.
2. A graphic element.
3. A proof.
4. Specific language.
5. An idea derived from the work, published or unpublished, of another person.

Any disciplinary actions regarding charges of cheating or plagiarism will follow the procedures of the home university of the student involved.

Collaboration vs. Cheating   Collaboration is defined by Merriam-Webster's Collegiate Dictionary (10th edition) as "to work jointly with others or together, especially in an intellectual endeavor." Much of the work that is performed in this laboratory (and in biomedical research as a whole) is collaborative in nature. Therefore, collaboration in this class is encouraged during the execution of the labs. In addition, discussion regarding the content of homework assignments is also encouraged.

You are encouraged to discuss the course material, concepts, and assignments with other students in the class. **However, each student must eventually submit his/her own unique work (e.g., laboratory notebook, final report).** If any collaboration was used to complete an assignment, record the names of the collaborators and the nature of the collaboration. Any attempt to submit work that is not the student's own work will be considered an act of cheating. In addition, any student who knowingly supplies their homework assignment for review to another student is violating the cheating policy and will be subject to disciplinary action.

> ANY VIOLATION OF THIS POLICY WILL NOT BE TOLERATED AND THE PENALTY WILL BE FAILURE IN THE COURSE.

If you have any questions regarding this policy, contact the instructors.

6 Resources

6.1 Accommodations for Students with Disabilities

If you have a disability and have an accommodations letter from the Disability Resources office, we encourage you to discuss your accommodations and needs with me as early in the semester as possible. We will work with you to ensure that accommodations are

provided as appropriate. If you suspect that you may have a disability and would benefit from accommodations but are not yet registered with the Office of Disability Resources, we encourage you to contact them at access@andrew.cmu.edu.

6.2 Statement of Support for Students' Health & Well-being

Take care of yourself. Do your best to maintain a healthy lifestyle this semester by eating well, exercising, avoiding drugs and alcohol, getting enough sleep and taking some time to relax. This will help you achieve your goals and cope with stress. All of us benefit from support during times of struggle. There are many helpful resources available on campus and an important part of the college experience is learning how to ask for help. Asking for support sooner rather than later is almost always helpful. If you or anyone you know experiences any academic stress, difficult life events, or feelings like anxiety or depression, we strongly encourage you to seek support. Counseling and Psychological Services (CaPS) is here to help: call 412-268-2922 and visit their website at http://www.cmu.edu/counseling/. Consider reaching out to a friend, faculty or family member you trust for help getting connected to the support that can help.

> If you or someone you know is feeling suicidal or in danger of self-harm, call someone immediately, day or night:
>
> CaPS: 412-268-2922
> Re:solve Crisis Network: 888-796-8226
> Suicide and Crisis Lifeline: 988
>
> If the situation is life threatening, call the police
> On campus: CMU Police: 412-268-2323
> Off campus: 911

6.3 Diversity Statement

We must treat every individual with respect. We are diverse in many ways, and this diversity is fundamental to building and maintaining an equitable and inclusive campus community. Diversity can refer to multiple ways that we identify ourselves, including but not limited to race, color, national origin, language, sex, disability, age, sexual orientation, gender identity, religion, creed, ancestry, belief, veteran status, or genetic information. Each of these diverse identities, along with many others not mentioned here, shape the perspectives our students, faculty, and staff bring to our campus. We, at CMU, will work to promote diversity, equity and inclusion not only because diversity fuels excellence and innovation, but because we want to pursue justice. We acknowledge our imperfections while we also fully commit to the work, inside and outside of our classrooms, of building and sustaining a campus community that increasingly embraces these core values.

Each of us is responsible for creating a safer, more inclusive environment.

Unfortunately, incidents of bias or discrimination do occur, whether intentional or unintentional. They contribute to creating an unwelcoming environment for individuals and groups at the university. Therefore, the university encourages anyone who experiences or observes unfair or hostile treatment on the basis of identity to speak out for justice and support, within the moment of the incident or after the incident has passed. Anyone can share these experiences using the following resources:

Center for Student Diversity and Inclusion: csdi@andrew.cmu.edu, (412) 268-2150

Report-It online anonymous reporting platform: reportit.net username: tartans password: plaid

All reports will be documented and deliberated to determine if there should be any following actions. Regardless of incident type, the university will use all shared experiences to transform our campus climate to be more equitable and just.