

ENGRD 2700 - BASIC ENGINEERING PROBABILITY AND STATISTICS

Enrollment Information

10293	★	LEC 001	TR	1:25pm - 2:40pm	Pender, J	●
				Olin Hall 255		
You may opt to enroll in a 1-credit (S/U graded based on attendance) Academic Excellence Workshop (AEW) to be taken in conjunction with this course. AEWs are weekly collaborative problem-solving workshops designed to enhance student understanding of course material. In order to attend the workshops, you must enroll in an AEW section, listed under course number ENGRG 2700.						
10294	★	DIS 201	M	2:55pm - 4:10pm	Pender, J	■
				Frank H T Rhodes Hall 571		
10295	★	DIS 202	R	2:55pm - 4:10pm	Pender, J	■
				Frank H T Rhodes Hall 571		
10296	★	DIS 203	T	8:40am - 9:55am	Pender, J	●
				Frank H T Rhodes Hall 571		
10297	★	DIS 204	T	11:40am - 12:55pm	Pender, J	■
				Frank H T Rhodes Hall 571		
10298	★	DIS 205	M	8:40am - 9:55am	Pender, J	■
				Frank H T Rhodes Hall 571		

Prerequisites: MATH 1910 and 1920.

Staff:

- Professor:
 - **Professor Jamol Pender, 228 Rhodes Hall**, email: jjp274@cornell.edu. Office hours Wednesdays 5-6:30 (Rhodes 261) or by appointment.
- Assistants (PhD)
 - Samuel Bright-Thonney, skb93.
 - Misha Padidar, map454
 - Jian Wang, jw2333
 - Rohan Sarkar, rs2466

- Assistants (MEng)
 - Yishu Zhang ap2295
- UG graders
 - Jenny Li, jl2872
 - Arjun Suri, as2542
 - Rhythm Raghuwanshi, rsr264
 - Hannah Lee, hl742

Office hours and additional contact information will be announced on the class webpage. Click on “Course Tools” then “Contacts.”

Attendance: Attendance at both weekly lectures and one of the recitations is expected. You are responsible for being aware of the announcements and content of both. Extensive data analysis reveals that good attendance is positively correlated with good grades.

Course Website: Many course materials (including lecture notes, section notes, handouts, homework assignments/solutions, announcements, exam solutions, etc) will be posted at our course website via Canvas.

VideoNote: The course is **NOT** being videoed and archived.

Approximate syllabus:

- Introduction; survey; culture; descriptive statistics; plotting.
- Probability models; Bayes theorem.
- Random variables
- Discrete distributions (binomial, Poisson, geometric, hypergeometric)
- Continuous distributions (exponential, normal)
- Central limit theorem and linear combination of random variables
- Normal approximation to binomial and Poisson
- Inference: Point and interval estimation
- Hypothesis testing
- Power and sample size
- Comparisons: Two-sample t-test, paired t-test; new better than old?
- Two variances, two proportions; new less variable than old?
- Introduction to linear regression, inference in regression analysis, regression diagnostics

- Multiple regression (maybe)
- Goodness-of-fit (maybe)

Texts and Class-notes.

- (1) A pdf version of each lecture (perhaps with gaps to be filled in during class) should be posted on the course webpage the night before the lecture. This can either be printed for annotation (so last century) or annotated electronically. I hold the copyright on the course notes and other course materials (homework etc), and buying, selling or reposting course materials is prohibited and illegal. We check course-hero.
- (2) Recommended text: J. L. Devore. Probability and Statistics for Engineering and the Sciences. Any edition, including an online edition is fine. (The cheaper, the better.) The publisher churns editions to kill the used book market. This is an useful book with helpful examples that matches the course well.
- (3) Other texts: Available free as a download from <https://www.openintro.org> is something called *OpenIntro Statistics*; it's not up to the mathematical level of this course but it is good for context.

Software: We will use the statistical software tool R coupled with the RStudio interface. If you would like to use Python, I have no objection to this, however, examples will be done in R. This is **free**, available for all platforms, and used frequently in subsequent courses such as

- ORIE 3120 - Industrial Data and Systems Analysis,
- ORIE 4740 - Data Mining,
- ORIE 4741 - Learning with Big Messy Data
- ORIE 5550 - Applied Time Series Analysis,
- ORIE 5640 - Statistics for Financial Engineering
- ORIE 5650 - Quantitative Methods of Financial Risk Management
- ORIE 6780 - Bayesian Statistics and Data Analysis.

R is dazzlingly professional and powerful but by itself it is a little geeky and so we recommend the shell R-studio (<http://www.rstudio.com>) which is jazzier than just R, more functional and definitely easier to use. If there is a downside to using R-studio rather than R, I haven't found it. The R repository is <http://cran.r-project.org> where specialized libraries exist as add-ons. (It is unlikely you will need extra libraries;

certainly not initially.) The first cycle of recitations will be devoted to getting started on loading and using R and R-studio.

For the stubborn: you may use any computational tool you like to do the homework but only R and R-studio will be supported in this course. Wean yourself from your high school hand calculator or be subject to withering disdain.

Although less necessary than in the past, it comes in handy to know how to write a path for your machine. PC's write paths with `xxx\xxx` and macs (unix based) and linux write them as `xxx/xxx`.

R-resources. Books featuring R have proliferated. For instance the publisher Springer-Verlag has a series called *Use R!*. This is more than you need and I don't really recommend stressing by going into these books but if you need more specialized knowledge, its nice to know these are available. Some examples:

- Alain F. Zuur, Elena N. Ieno, Erik H.W.G. Meesters. *A Beginner's Guide to R*, Springer. Available to download with a Cornell net-id from the Cornell library. It's more than you need but nice to know it's there if you need it.
- Nicholas J. Horton, Ken Kleinman. *Using R and R-Studio for Data Management, Statistical Analysis, and Graphics*. CRC Press. Browse the first chapter. Not in the library.
- Other resources more atuned to our needs: Introductory PDF's collected from hither and yon are posted on the course web site.

R-history. I am usually not big on history but this is one case where good guys seem to have won over rapacious capitalism. The S (not R) system was started in 1976 as a programming language at AT&T Bell Labs by John Chambers for "organizing, visualizing and analyzing data." In the mid '80's, Bell Labs licensed S to a company *Statistical Sciences* which distributed an implementation called S-Plus. This company was later bought by TIBCO Software. The original implementation of S-Plus had severe memory management problems which were apparent when looping or analyzing large data sets. The commercial versions of Splus put all its resources into a gui and did not fix the memory management problem. In the early '90's **R**oss Ihaka and **R**obert Gentleman (University of Auckland) were pissed off enough (my interpretation for their motivation) with the ignored weaknesses

of S-plus that they decided start from scratch and write a new, open-source, modern implementation which they called R. The reason for calling it R is either it was the initial of their first names or it is a letter in the alphabet close to S. Since 1997, R is governed by the R Development Core Team. They vet packages and standards are strict (and complaints are many about rudeness and brevity). See <http://cran.r-project.org>. R has pretty much killed Splus. See <https://en.wikipedia.org/wiki/S-PLUS> if you like history.

Exams policies: To avoid continuing discussions about behavior during the exams, here are some rules. You may bring a single sheet (US-Letter size) with you to the exams. How you stuff this sheet is up to you but only one piece of paper is allowed. You may not use a computer, calculator or books. Cellphones must be turned off. A live cellphone means you want to be removed from the room and accused of an honor code violation. “Off” means off and not in silent or vibrate mode.

Calendar. There are 28 lectures.

- March 12 ENGRD 2700 in class **midterm** in our usual room,
- According to <https://registrar.cornell.edu/exams/fall-final-exam-schedule>, here is our final exam schedule for the 2.5 hour final:

ENGRD 2700 001 Sunday, May 10 9:00 AM
Place to be announced. Do not plan on leaving town early before the final exam.

Grading: Your final grade score will be based on

- class project (10%),
- homework (20%) written up individually,
- Tests (70%)

where

$$\text{test} = \max\{[(.40) * (\text{midterm}) + (.60) * (\text{final})], \text{final}\}.$$

So a strong final exam score makes us forget a mediocre midterm score.

So the algorithm for the final grade score is

$$\begin{aligned} \text{final grade score} &= (.10) * [\text{class project}] \\ &+ (.20) * (\text{homework percentage}) \end{aligned}$$

$$+ (.70) * [\max\{[(.40) * (\text{midterm}) + (.60) * (\text{final})], \text{final}\}].$$

After a final numerical score is computed, the scores are ranked and letter grades are assigned. Some discretion is involved to account for variability in exam's level of difficulty. In a prior iteration, about 30% were given something in the A-range, 44% in the B-range, 25% in the C-range. I get very few complaints.

Re homework: Homework assignments are equally weighted **BUT** your lowest two homework scores will be dropped when we compute the homework composite score at the end of the semester. If you do not hand in an assignment, you get a grade of 0 and in many but not all cases, lowest means the homework was not handed in. Thus you get two free passes. Consequently use them when you have a family emergency, scheduling difficulty, job interview, personal crisis etc. The discerning among you will note the point of this policy is to avoid requests for extensions or waivers. We will not be sympathetic to such requests which create havoc with the TA's and graders and a lot of extra work.

All homework assignments are to be completed by students working alone. You may discuss the homework problems with others if this has educational value but write-ups should be completed alone.

Weekly cycle. Weekly homework assignments will be posted on Cavnas and are typically due one week after it is assigned. Please submit all assignments to Gradescope. **Late homeworks are not accepted.**

Homework will typically be graded within a week and handed back in section. If for some reason you cannot pick up your homework during section, you may pick up your homework at your TA's office hours or during section the following week. Homework solutions will typically be posted shortly after the assignment is due.

Regrades: It is theoretically possible that a mistake could be made in grading homework and exams. If you think the grading on a homework or exam is in error, please write a note explicitly stating what the error is claimed to be and turn in the paper (with note attached—be specific about what you think the grading error is!) to the section TA for regrading. Requests for regrades must be made within one week after the homework or exam is returned to the class. In the past, regrade

requests have sometimes been tainted by cheating so please behave responsibly.

Etiquette even though I'm generally a nice guy:

Email. You may have a small problem and think it is reasonable to email me or one of the TA staff. Multiply a small problem by 100 and it may no longer be small. Before emailing, make a good faith attempt to find the information. Check the information you need is not on the course web site ("where is your office?"), or on a Cornell site ("Is there class on labor day?") or not available by googling ("where do I download R-studio?"). Remember there are office hours where face to face contact and instant feedback is possible. If you choose to contact me or the TA by email put ENGRD 2700 in the subject line but if it is 12:30AM, do not expect 5-minute turn around even if the subject line reads "I'm freaking out about ...". **Finally, I will also not answer emails on Sunday.**

Of course, posting a question to Piazza is also possible.

Cell phones. Either turn off your phone when you enter the class room or if this is unimaginable, at least put it in vibrate or silent mode.

Section meetings, office hours. Section meetings will be used for discussion of some of the statistical issues covered in class and help you master the software. It will give you an opportunity to discuss homework and exams with the TA's and your classmates. It will also be devoted to mechanics of problem solution and software use. Learn what section you are in, who your recitation TA is and then label all work with the section and TA written prominently in the upper right corner. This means you have to remember one number out of the set $\{\text{DIS 201}, \dots, \text{DIS 205}\}$.

Office hours. Every TA and grader will have an office hour posted on the course web site as will I. Try to restrict office hour questions to clarifying concepts or approaches to problem solution. We will experiment with the effectiveness of Zoom (Cornell licensed conferencing software) to supplement face-to-face interaction.

Piazza. We will have a Piazza site which will be monitored 1 hour every other day. The site to sign up is: piazza.com/cornell/spring2020/engrd2700 and the access code is: **probabilityisawesome**. Finally, the link for the class is: piazza.com/cornell/spring2020/engrd2700/home.

AEW & Tutoring: Material from Engineering Learning Initiatives.

0.0.1. *AEW.* An Academic Excellence Workshop (AEW) section is available to be taken in conjunction with this course. AEWs are optional 1-credit supplemental courses which meet for one two-hour collaborative problem-solving session each week throughout the semester. Designed to enhance student understanding, the workshops feature group work on problems at or above the level of course instruction. In the workshops, small-group problem-solving is directed by undergraduate peer educators called facilitators. The AEWs are graded S/U, based on attendance.

Tutoring: One-to-one peer tutoring is available free of charge for engineering students and BEE majors enrolled in this course through Engineering Learning Initiatives' Tutors-on-Call program. Review program policies and request a tutor at: <https://www.engineering.cornell.edu/students/undergraduate-students/academic-opportunities-and-support/tutors-call>

There is also a learning strategies center at <http://lsc.cornell.edu>.

Academic Conduct: Each student in this course is expected to abide by the Cornell University Code of Academic Integrity. Any work submitted by a student in this course for academic credit should be the student's own work, with exceptions/particulars described below.

Each semester brings repeated incidents of depressingly dishonest behavior. Copying during exams or copying homework is an example of cheating. These are not that hard to detect. Accusations of cheating invariably bring responses along the lines of "I'm not that sort of person; that's not who I am" when the evidence points to the contrary.

Do not bring other people's i-clickers to class.

Violations will be handled in accordance with the Code of Academic Integrity available at <http://www.theuniversityfaculty.cornell>.

[edu/academic-integrity/](http://cornell.edu/academic-integrity/) where you will find, among other things, *The Essential Guide to Academic Integrity at Cornell* and even a 14-minute video entitled *Cheating*. If you have any questions about this policy, let's discuss.

Personal or Academic Stress: If you are experiencing undue personal or academic stress at any time during the semester or need to talk with someone about a personal problem or situation, we encourage you to seek support as soon as possible. I am available to talk with you about stresses related to your work in my class. Additionally, I can assist you in reaching out to any one of a wide range of campus resources, including:

- Your college's Academic Advising or Student Services Office.
- Cornell Learning Strategies Center at 255-6310, <http://lsc.cornell.edu>.
- Gannett Health Services at 255-5155, www.gannett.cornell.edu.
- Peer Support provided by Empathy Assistance and Referral Service at 255-EARS, <http://orgsync.rso.cornell.edu/org/ears>.

Disability Related Concerns: Students with either an ongoing or short-term disability are encouraged to contact Student Disability Services (SDS) <https://sds.cornell.edu> for a confidential discussion of their need for academic accommodations. SDS details:

Student Disability Services

Cornell Health, Level 5

110 Ho Plaza, Ithaca, NY 14853-3102;

Phone: 607-254-4545

Email: sds_cu@cornell.edu

Students qualified for extra time on exams or other accommodation should bring me authorization from SDS.