STA 237 - Probability, Statistics and Data Analysis I, L0101
M @ 3-5 PM Toronto Time on Bb Collaborate

Instructor: Karen H. Wong - karen.huynhwong@utoronto.ca

Office Hours: Will be posted on Quercus, held on Bb Collaborate/MS Teams

Course Description: An introduction to probability using simulation and mathematical frameworks, with emphasis on the probability needed for more advanced study in statistical practice. Topics covered include probability spaces, random variables, discrete and continuous probability distributions, probability mass, density, and distribution functions, expectation and variance, independence, conditional probability, the law of large numbers, the central limit theorem, sampling distributions. Computer simulation will be taught and used extensively for calculations and to guide the theoretical development.

Prerequisites: (MAT135H1, MAT136H1)/MAT137Y1/MAT157Y1/(MATA30H3, MATA36H3)/(MATA31H3, MATA37H3)/MAT135Y5/MAT137Y5/MAT157Y5
Prerequisites are strictly enforced by the department and prerequisite waivers are not accepted.

Textbooks:

1. *A Modern Introduction to Probability* by *Dekking et. al.*, available through the library here
2. *Modern Mathematical Statistics with Applications*, 2nd ed. by *Devore and Berk* available here.

Course Structure: During these unprecedented and challenging times, this course will operate differently from the regular sitting. Our course will be a combination of synchronous and asynchronous lectures with irregular tutorial activities. Synchronous lectures *will be recorded* and made available on our Quercus page, but tutorials are self-paced individual and group work, and are **mandatory**. Note that all course times listed are in Toronto time!

Suggested Weekly Routine

Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Monday
Video lectures posted. Watch at your own pace, take your own notes. Post your questions and answers on discussion board. Work on tutorial, if applicable.		*Attempt 5-10 suggested problems each day. Post your questions to the discussion board. Finish up tutorial work.*		REST DAY	*Review notes from the week's videos to prepare for sync session.*	*Work through R material from sync session. Start on tutorial, if applicable.*

Everyone learns at a different pace. In addition to our course times ($\sim$ 3-4 hours per week), students should expect to commit *at least an additional 4 hours* as self-study, or 7-8 hours per week. Self-study includes: creating course notes for yourself, working on suggested problems, reviewing class material, familiarizing yourself with R, posting on the discussion board, and completing tutorial work.

Quercus Page: All lecture slides, video recordings, quizzes/tests/assignments, and other course material will be posted on Quercus under Modules. Course materials provided on Quercus are for the use of students *currently enrolled in this course only.* **Distributing course materials to anyone outside of the course is considered unauthorized use.**

Grading: There will be multiple evaluations in form of assignments and midterms. The grade breakdown is

Weekly Discussion	12%	Throughout Winter Term
Syllabus Quiz	2%	Jan. 28
Individual Assignments (2)	8% each	Feb. 12, Apr. 1 @ 10:00 PM
Tutorial Activities (Best 5 of 6)	5% each	Weeks of: Jan. 18, Feb. 1, Feb. 22, Mar. 8, Mar. 15, Apr. 5
Midterm	20%	March 5: 3-5 AM/PM EST
Final Assessment	25%	TBD

Note: For all assessments with R, you are expected to submit the accompanying ***knitted R Markdown*** *file with your code, comments, and relevant outputs.*

Discussion Board: There will be weekly topic threads on the Quercus Discussion Board as a space for students to discuss textbook problems, course topics, and exchange peer support. The instructor and TAs will be monitoring and contributing regularly to keep the discussions going, guide students back on track, or answer any lingering questions. The expectation is that students contribute by working collaboratively with each other and build on your understanding of concepts by answering each others' questions. The grading scheme will be posted and discussed during our first class.

Assignments: Only select problems on the assignments will be evaluated. **Assignments will be submitted through Crowdmark.**. Late assignments will receive a penalty of 5% for every hour interval that the assignment is late. For example, if an assignment is submitted 10 minutes after the due date, there would be a penalty of 5% on the assignment grade (i.e. 90% $\rightarrow$ 85%).

Tutorials: Will usually alternate between individual or group tutorial activities. Group tutorials will usually two components: an *independent* pre-tutorial to be completed and submitted individually in preparation for the group activity where you collaborate with other classmates. The tutorials should take no longer than 2-3 hours, depending on individual progress and fluency with course material.

Tutorial Weeks: Jan. 18, Feb. 1, Feb. 22, Mar. 8, Mar. 15, Apr. 5

Midterm & Final Assessment: Assessment information will be available on Quercus approximately 1 week prior to the date of the assessment. The midterm will be a timed assessment on Friday March 5, 3-5 PM EST. A *mirrored time slot* on the same day at **3-5 AM EST** will also be offered. A survey will be available a week before for you to register for your time slot.

Any time conflicts with the midterm must be communicated as soon as you are aware, and **no later than 2 weeks prior to the midterm**. No accommodations will be offered for late notice of time conflicts.

Minimum Passing Requirement: Students must complete and submit at minimum:

- 3 of the 6 mandatory tutorials
- 1 assignment
- Both the midterm and the final assessment.

If you miss more than these minimum passing requirements, even with accommodation, you will not be able to pass this course. There are no make-up tutorials. If there are extenuating circumstances that will affect your performance in the course in the long term, it is your responsibility to contact your college registrar and the instructor as early as possible.

Accommodations: Accommodation requests for assignments must be made by email to the course instructor (karen.huynhwong@utoronto.ca) **at least 24 hours before** the assignment due date. Extensions will be arranged where possible. There are no accommodations for tutorials due to the self-paced time frame and the group component.

Missed Work: Missed term work (tutorials up to a maximum of 3, and midterm) due to legitimate reasons (e.g., injury, illness, or other exceptional circumstances) should be communicated to the instructor within 1 week following the assessment **with** the absence declaration form on ACORN completed and attached. For documented missed work:

- Missed tutorial work will be redistributed among the tutorials.
- Missed assignment will be distributed among the other assignment, midterm, and final assessment, proportional to weight.
- Missed midterm: A make-up midterm will be scheduled after March 8 for students who missed the midterm with documentation.

Grading: Marking schemes/solutions will be provided (except the final assessment). These should be reviewed before requesting a reread. If you still have concerns about your grading, send me an email with a brief explanation **no later than 1 week** after the assessments have been released back to you. Late requests will not be accepted.

Course Conduct:

- **Email:** Any administrative or personal concerns regarding the course should be addressed to the course instructor (karen.huynhwong@utoronto.ca). Questions regarding course material and concepts should be left for office hours or the discussion board. I will usually respond within 48 hours but may take up to 5 business days.

- **Programming Languages:** We will be using R Statistical Sofware regularly, R statistical software which is available for free download. An installation guide will be provided in Week 1.
  - If you have problems with setting up R, you may find it easier to use through R Studio Cloud which offers free 15 project hours per month.
  - By accessing through JupyterHub (Select 'R Studio', or if that's not readily available, click 'New', and 'R Studio' on drop down menu).
  - Instructions on using R as well as demos will be done regularly during our synchronous sessions, and starter code will be provided where appropriate.

  By the end of the course, you are expected to be able to run simple simulations on your own, write your own functions, and more generally, understand how various code works, apply and adapt them to solve problems, and predict the output that will be generated. Check the weekly agenda to see if you will need to have R ready to use during the synchronous lecture times.

Accessibility Services: The University of Toronto provides accommodations through accessibility services to students with diverse learning styles and needs. If you have a disability or health consideration that may require accommodations, please feel free to reach out to Accessibility Services at 416-978-8060 or through accessibility.utoronto.ca.

Academic Integrity: Participating honestly, respectfully, responsibly, and fairly in this academic community ensures that the U of T degree that you earn will be valued as a true indication of your individual academic achievement, and will continue to receive the respect and recognition it deserves. Familiarize yourself with the University of Toronto's Code of Behaviour on Academic Matter. Potential offences include, but are not limited to:

- Obtaining or providing unauthorized assistance on any test/assignment/tutorial including:
  - working in groups on individual assessments, including giving hints to the answer!
  - having someone rewrite, edit, or add material to your independent work
  - researching for inspiration, hints, or answers to any graded problem
  - posting active assessment questions on discussion boards/private tutoring companies for hints/solutions
- Lending your work to a classmate who submits it as their own with or without your permission

The University of Toronto treats cases of academic misconduct very seriously. All suspected cases of academic dishonesty will be investigated following the procedures outlined in the Code. The consequences for academic misconduct can be severe, including a failure in the course and a notation on your transcript. **If you have any questions about what is or is not permitted in this course, please do not hesitate to contact the instructor.** If you are experiencing personal challenges that are having an impact on your academic work, please speak to the instructor or seek the advice of your college registrar.

Tentative Schedule:

Week Of	Topics
1: Jan. 11	Introduction to the Course, Getting to Know R & R Studio. Recurring definitions, Events ($\cup, \cap, ^c$), Probability Axioms
2: Jan. 18	Counting, Conditional Probability, Law of Total Probability, Bayes' Rule
3: Jan. 25	Counting, Intro do Discrete Random Variables and Properties
4: Feb. 1	Properties of Transformed RVs, Estimation via Simulation
5: Feb. 8	Inequalities, Common Discrete Dist^{n}s: Bernoulli, Binomial, Poisson
6: Feb. 15	WINTER READING BREAK
7: Feb. 22	Intro to Continuous Distributions, Common Dist^{n}s
8: Mar. 1	Transformations of RVs, MIDTERM
9: Mar. 8	Univariate Transformations, Moment Generating Functions
10: Mar. 15	Bivariate Distributions (Discrete & Continuous): Joint vs Marginal
11: Mar. 22	Conditional Bivariate Dist^n, Dist^n Features
12: Mar. 29	Sampling Distributions, Simulated Sampling Dist^{n}s
13: Apr. 5	Central Limit Theorem and Applications