STA410 *Statistical Computation* (L5101 + L2501)

STA2102 *Computational Techniques in Statistics* (L0201)

Winter 2023 | Scott Schwartz

Typical Weekly Schedule

	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Complete Prereading (~3 hours)	✓	✓	✓				
Complete Programming (~4 hours)			✓	✓	✓	✓	✓
Discussion MS 3154 **6:10—9 PM ET**			✓				
Office Hour MS 3154 **9-10 PM ET**			✓				
Office Hour zoom **5-7 PM ET**					✓		
Programming **Due Sunday 11:59 PM ET**							✓

Assignments

	Due Weekly	On Feb 15 and Apr 5	On Mar 1	TBA
	Programming Assignments	**Coding Challenges**	**Midterm Exam**	**Final Exam**
STA130/STA2102	28% = 4\% × 7 (best of 9)	32% = 16%+16%	20%	20%
STA2102 Optional	42% = 6\% × 7 (best of 9)	32% = 16%+16%	13%	13%

Course Policies

Programming Assignments	
Late Enrolments	Week 1-4 **Programming Assignments** may be submitted LATE by Feb 12, 11:59 PM ET with permission from sta410@utoronto.ca
Sickness	The highest 7 of 9 **Programming Assignments** scores are used and there are no due date extensions for sickness
Illness Absence Declaration	
Feb 15 Coding Challenge	will be reweighted to the ***Midterm*** if an illness declaration is submitted through ACORN and forwarded to sta410@utoronto.ca by Feb 15
Apr 5 Coding Challenge	will be reweighted to the ***Final*** if an illness declaration is submitted through ACORN and forwarded to sta410@utoronto.ca by Apr 5
Mar 1 Midterm Exam	will be reweighted to the ***Final*** if an illness declaration is submitted through ACORN and forwarded to sta410@utoronto.ca by Mar 1
Final Exam	Petitions for Deferred ***Final Exams*** may be submitted to FAS
Accomodations/Exceptions	
Accomodations	For course accomodations please contact Accessibility Services or your College Registrar and alert sta410@utoronto.ca
Exceptional Circumstances	For exceptional circumstances please contact your College Registrar

---

Grading

	0 - 49%	**50 - 52%**	**53 - 56%**	**57 - 59%**	**60 - 62%**	**63 - 66%**	**67 - 69%**	**70 - 72%**	**73 - 76%**	**77 - 79%**	**80 - 84%**	**85 - 100%**
GPA	0.0	0.7	1.0	1.3	1.7	2.0	2.3	2.7	3.0	3.3	3.7	4.0
	STA410	STA410					STA2102	STA2102				
	Failing	Passing					Failing	Passing				

Course Topics and Dates

		Discussion Topics / Programming Topics (28%)	**Programming Due**	
Week 1	Jan 9-11	Syllabus, Floating-Point, Pseudorandomness		
	Jan 11-15	Integer Bit Representation, Modulus Recursion	Jan 15, 11:59 PM ET	
Week 2	Jan 16-18	Floating-Point Density, Numeric Errors, Tricks		
	Jan 18-22	Sum Roundoff Error, Catastrophic Cancellation	Jan 22, 11:59 PM ET	Jan 22 Last Day to Enrol
Week 3	Jan 23-25	Singular Values, PCA/PCR, Condition, $Ax=b$		
	Jan 25-29	Gram-Schmidt, Cholesky Factor Decomposition	Jan 29, 11:59 PM ET	
Week 4	Jan 30 - Feb 1	Algorithmic Speed, $X\hat{\beta} \approx y$, Iterative Methods		
	Feb 1-5	Convolution Counting, Automatic Differentiation	Feb 5, 11:59 PM ET	**Week 1-4 Programming**
Week 5	Feb 6-8	Vector/Function Spaces, More on $X\hat{\beta} \approx y$, FFT		**May be submitted LATE**
	Feb 8-12	Interpolation, Radix-2 FFT Recursive Algorithm	Feb 12-19, 11:59 PM ET	**by Feb 12, 11:59 PM ET**
Week 6	Feb 15	6-9 PM ET **In-Person Coding Challenge (16%)**		
	In-Person	You may use your laptop in **MS 3154** (n=250) or	an IIT Lab Workstation	in **SS561** (n=49)
Week 7	Feb 20-25	READING WEEK		
Week 8	Mar 1	6-9 PM ET **In-Person Midterm Exam (20%)**		
	In-Person	**MS 3154** (n=250)		
Week 9	Mar 6-8	Optimization, Information, Variational Inference		
	Mar 8-12	Newton's Method, Fixed-Point Iteration Methods	Mar 12, 11:59 PM ET	
Week 10	Mar 13-15	Fisher Scoring, Non-L_2 Loss, More Optimization		
	Mar 15-19	Iteratively Reweighted LS, nonlinear Gauss-Seidel	Mar 19, 11:59 PM ET	Mar 19 Last Day to Drop
Week 11	Mar 20-22	Integral Approximation and Estimation, Sampling		
	Mar 22-26	TBA	Mar 26, 5:00 PM ET	
Week 12	Mar 27-29	MCMC, Gibbs, Metropolis-Hastings, Hamiltonian		
	Mar 29 - Apr 2	TBA	Apr 2, 11:59 PM ET	
Week 13	Apr 5	6-9 PM ET **In-Person Coding Challenge (16%)**		Apr 6 Last Day to CR/NCR
	In-Person	You may use your laptop in **MS 3154** (n=250) or	an IIT Lab Workstation	in **SS561** (n=49)
	Apr 11-28 TBA	**In-Person Final Exam (20%)**		

Course Components

Quercus Course Page	Announcements are made in Quercus	Quercus will link additional course resources
Markus	Programming Assignments Submissions	Marking and regrades will be done in Markus
UofT Jupyterhub	Course Cloud Coding Environment	Requires a UofT UTORid
Google Colab	Course Cloud Coding Environment	Requires a private google account
IIT Computer Labs	Drop in workstations **SS561** (n=49)	CR325(n=30), RW109(n=24), RW107(n=20)
In-Person Lecture	**MS 3154** Wed 6:10—9:00 PM ET	**WILL NOT BE RECORDED**
In-Person Office Hour	**MS 3154** Wed 9:00—10:00 PM ET	**WILL NOT BE RECORDED**
Online Office Hours	**zoom** Friday 5-7 PM ET	**WILL NOT BE RECORDED**
Discussion Board	~1-2 weekday (only) response time	All non-private course communication
sta410@utoronto.ca	~1-2 day weekday (only) response time	**Private course communication ONLY**
Coding Challenges	Wed 6:10—9:00 PM ET Feb 15 and Apr 5	You may use your laptop in **MS 3154** (n=250) or an IIT Lab Workstation in **SS561** (n=49)
Midterm Exam	**MS 3154** Wed 6:10—9:00 PM ET Mar 1	
Final Exam	TBA: will take place between Apr 11-28	

Course Design

1. **Statistics is a desirable skill set, particularly in conjunction with coding ability**: this course gives you an opportunity to practice and build valuable `Python` (and coding) skills.

   > `Python` is chosen instead of `R` for this course because
   >
   > - `Python` is more general purpose and ubiquitous
   > - `Python` is more like low level programming languages (such as `C`, `C++`, and `Fortran`)
   > - `Python` naturally supports computational algorithm development in addition to guided data analysis
   >
   > Students are expected to code via UofT JupyterHub or google colab. Local `Python` installations (conda, mamba, etc.) and other code editors (VSCode, PyCharm, Atom, pico/nano vi, emacs, etc.) will not be supported.

2. **This course will be based on Q&A discussions and conversations, not lectures**: come to class ready to explain the material to other students, or ask clear questions for clarification. *Comfort with self-learning will lead to professional success and helps create new opportunities. Communication ability is extremely important for career progression.*

   > Students are expected to read and understand the course material to the best of their ability prior to coming to class, and then practice their communication skills during class. The course notes for each week will include a sufficiently self-contained and complete synthesis of selected material from
   >
   > - **Computational Statistics** by James E. Gentle (Springer)
   > - **Computational Statistics** (Second Edition) by G. H. Givens and J. A. Hoeting (Wiley)
   > - Previous STA410/2102 course notes from Keith Knight

3. **Less is more, as this course is interested in building foundations:** we'll cover numeric pitfalls, fundamental linear algebra topics, least squares and beyond, the foundations of modern differentiation-based optimization, and estimation-based integration with MCMC.

   > The empahsis firstly on foundations and then on modern differentiation and integration topics is made **at the expense** of several topics typially found in a ***Statistical Computation*** course
   >
   > - Select linear algebra topics, e.g., sparse matrices, fast hadamard transforms, randomized numerical linear algebra
   > - Sampling from specific statistical distributions and nonparametric inference with bootstrapping
   > - Combinatorial (discrete) optimzation, e.g., with simulated annealing
   > - Constrained optimization, e.g., Lagrange multipliers, and interior and exterior point algorithms (e.g., for quantile regression)
   > - Expectation/minorization-maximization, e.g., for censoring and mixture models
   >
   > but the successful student of this course would certainly be capable (and encouraged) to pursue these topics upon completion of the course.