UNIVERSITY OF TORONTO
DEPARTMENT OF STATISTICAL SCIENCES

STA221 H1 S – The Practice of Statistics II
Course Outline – summer 2019

lectures (L5101): Tue & Thu 7:10-10 pm SS1087
tutorials (T5101): Tue & Thu 6:10-7 pm RW142 and SS2111

Instructor:	**J.G. Pitt**
Office Location:	**Stewart 103**
Office Telephone:	**none**
Email:	**greg.pitt@utoronto.ca**
Office Hours:	**TR 5:10-5:55 pm, SS lobby and by appointment**
TA(s):	**Spark Tseung**
	Ruoyong Xu

Official Course Description

Continuation of STA220H1 (or similar course), emphasizing major methods of data analysis such as analysis of variance for one factor and multiple factor designs, regression models, categorical and non-parametric methods (Note: STA221H1 does not count as a distribution requirement course).

Prerequisite: STA220H1/ STA288H1/ PSY201H1/ GGR270H1/ EEB225H1/STAB22H3/STA220H5
Exclusion: ECO220Y1/ ECO227Y1/ GGR270Y1/ PSY202H1/ SOC300H1/ SOC202H1/ SOC252H1/ STA261H1/ STA248H1/ STAB27H3/ STA221H5/ ECO220Y5/ ECO227Y5/ STAB57H3/ STA258H5/ STA260H5
Distribution Requirement Status: None
Breadth Requirement: The Physical and Mathematical Universes (5)

Course Objectives

The successful student will master additional concepts of statistics that are understood and applied by the majority of statisticians practicing today. Students should understand these concepts well enough to continue learning on their own after the course has concluded.

Main Text:

deVeaux, Velleman, Bock, Vukov, & Wong: *Stats Data and Models*, 3rd Canadian ed., Pearson

Supplementary Books:

Moore, McCabe, & Craig: *Introduction to the Practice of Statistics*, 7th ed., 2012
Roberts & Russo: A *Student's Guide to Analysis of Variance*, Routledge, 1999
Montgomery: *Design and Analysis of Experiments*, 5th edition, Wiley, 2001

Determination of Grades

quiz 1	15%	July 16, tutorial rooms
mid-term	30%	July 23, Bahen 2165
quiz 2	15%	Aug 06, tutorial rooms
final exam	40%	Aug 13 or later (t.b.a. by Faculty)
TOTAL	100%	

2019 July 29: Last day to drop courses with S section codes from academic record and GPA. After this deadline a mark is recorded for each course, whether course work is completed or not (a 0/zero is assigned for incomplete work), and calculated into the GPA.

Requirements and Criteria
Problem sets, consisting of questions assigned from the book as well as supplementary questions, will be posted from time to time. These questions are for practice and discussion and not for handing in and grading.

The quizzes will be administered during the tutorial time slots. The mid-term will be administered during a longer time period which will begin at tutorial start time. Tests may include any material covered up to the end of the previous class, unless stated otherwise in class.

The final exam will emphasize material in the second half of the course, but may still include questions from earlier material.

For the various tests, students will be permitted to use a non-programmable, non-plotting calculator. If there is any doubt about the permissibility of your particular calculator, please consult in advance.

Students must bring their U of T student ID to all quizzes, tests, and exams.

Conduct of Classes
Ordinary classroom etiquette is expected of all of the students. This includes arriving on time, turning off cell phones and similar devices, and respect for fellow students.

Leaving the classroom while a lecture is in progress is disruptive and should be avoided. If you feel that you will need to leave class before it ends, please sit close to the doors and alert the instructor to your situation.

The class time will be used for lectures and discussion, based mostly on the material in the textbook. However, the instructor may assign additional reading and/or exercises to supplement the book.

Class participation is strongly encouraged: asking questions, comments that relate this course to others that you are taking, pointing out mistakes on the chalkboard, etc.

The course website will be the centre for communication from the instructor to the students. The students are urged to complete the recommended problem sets. Solutions to some of these problems may be posted on the course web-site.

The Statistics Aid Centre can be an important source of help with difficulties. The current location

is Sidney Smith 623B. Your designated TA will be on duty at least one hour per week, but you may drop in at any time if you are willing to wait for other students to be served. Schedules will be posted as soon as they become available. Some additional information may be available at http://www.utstat.utoronto.ca - click on Statistics Aid Centres for the schedule.

Procedures and Rules

communication
The course website will be the centre for communication from the instructor to the students. The students are urged to complete the recommended problem sets. Solutions to some of these problems may be posted on the course web-site.

If we experience a Quercus outage, the instructor will attempt to post essential material at http://utstat.utoronto.ca/~pitt/sta221

The instructor and the TAs may be contacted by email at any time. In general, the TAs should be contacted regarding subject matter, and the instructor regarding administrative matters. Please include the course code (sta221) in the subject line of all email messages. Note that if message traffic becomes excessive, the course email policy may be revised at any time during the term.

missed tests: In the event a student misses the mid-term exam due to illness or domestic situation, the student must contact the Statistics departmental office immediately, and submit a medical certificate indicating type of illness and date of illness (or other applicable documentation for domestic situations) to the instructor. This should be done within 48 hours of the test date, if possible.

If a student misses a quiz or the mid-term for legitimate reasons, the missing points will be earned by weight shifting to the next major test or exam.

programming languages: The official primary programming language for this course is R. R is free to install on a student's personal computer. No other language will be supported by the instructor and TAs.

how to present your work: On assignments and tests, make sure that your final answers are not difficult to find. When asked to supply one answer, do not give two. Show your work (i.e., document your thought processes). Some wrong answers may be awarded partial credit, but not unless you show your work. **Careless rounding and similar sloppiness will result in deductions.**

marking issues: The TAs and the instructor are well aware of the importance of grades to most students, and great care will be taken in the marking of assignments, quizzes, and exams. In the unlikely event that you feel a question has been mis-marked, or the marks have been added up incorrectly, you can submit your test back to the instructor with a note explaining what you believe requires further examination. This must be done within one week of the test being returned to the class.

cancellation policy: In the event of inclement weather, instructor illness, or similar circumstances resulting in class cancellation, any quiz or assignment due date will be postponed until the next class. Minor adjustments to the overall course schedule might be necessary, and these will be posted. Formal rules are in place regarding the rescheduling of final exams, and these will be followed if

necessary.

accessibility: Students with diverse learning styles and needs are welcome in this course. In particular, if you have a disability or health consideration that may require accommodations, please feel free to approach me and/or the Accessibility Services Office as soon as possible. The Accessibility Services staff are available by appointment to assess specific needs, provide referrals and arrange appropriate accommodations. The sooner you let them and me know your needs, the quicker we can assist you in achieving your learning goals in this course.

For more information, please refer to http://www.accessibility.utoronto.ca/

academic integrity is fundamental to learning and scholarship at the University of Toronto. 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 Matters* (http://www.governingcouncil.utoronto.ca/policies/behaveac.htm). It is the rule book for academic behaviour at the U of T, and you are expected to know the rules. Potential offences include, but are not limited to:

In papers and assignments:

- Using someone else's ideas or words without appropriate acknowledgement.
- Copying material word-for-word from a source (including lecture and study group notes) and not placing the words within quotation marks.
- Submitting your own work in more than one course without the permission of the instructor.
- Making up sources or facts.
- Including references to sources that you did not use.
- Obtaining or providing unauthorized assistance on any assignment including
    - working in groups on assignments that are supposed to be individual work,
    - having someone rewrite or add material to your work while "editing".
- Lending your work to a classmate who submits it as his/her own without your permission.

On tests and exams:

- Using or possessing any unauthorized aid, including a cell phone.
- Looking at someone else's answers
- Letting someone else look at your answers.
- Misrepresenting your identity.
- Submitting an altered test for re-grading.

Misrepresentation:

- Falsifying or altering any documentation required by the University, including doctor's notes.
- Falsifying institutional documents or grades.

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 have questions about appropriate research and citation methods, seek out additional information from the instructor, or from other available campus resources like the U of T Writing Website. 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.

Note that because of crowding in the classroom, multiple versions of the quizzes and mid-term exam may be administered. The differences between versions will be slight and should not affect the difficulty of particular problems.

summer courses
Students should note that summer courses present additional challenges that are not found in regular-term courses. The pace will be brisk, and those of you who are accustomed to short lectures during the day may find it difficult to adjust to long lectures in the evenings. Additionally, because of time constraints, students will be asked to learn a few topics via independent study.

topic outline - The following topics will be covered, and in approximately this order:

I. means (SDM2: ch. 20-22; SDM3: ch. 18-20)
Review of the last topics covered in STA220

II. proportions (SDM2: ch. 19; SDM3: ch. 21)

III. count data (SDM2: ch. 23; SDM3 ch. 22)
Comparing counts. Two-way tables. The chi-square statistics, p-value.
Goodness-of-fit. Independence.

IV. regression (SDM2: ch. 24; SDM3: ch. 23)
Review of regression as covered in STA220
Inferences for regression. Simple linear regression. Regression parameters estimates. Confidence intervals relevant to regression. Prediction intervals
Linearizing transformations

V. analysis of variance (SDM2: ch. 25; SDM3: ch. 24)
The one-way ANOVA. The ANOVA model. The ANOVA table. The ANOVA F-test.
Planned comparisons (contrasts) among the means: t-test & F-test. Orthogonality and decomposing the between-group SS. Pooled Contrasts. Multiple Comparisons
Multi-factor Analysis of variance. The two-way ANOVA model. The ANOVA table for two-way ANOVA.
Interactions. Inference when effects of variables are related.

VI. multiple regression (SDM2: ch. 27; SDM3: ch. 26)
Inference for multiple regression. Multiple linear regression model. CIs and significance tests for regression coefficients. ANOVA table for multiple regression.
Generalized Linear Models (GLM) and transforming to GLM (e.g. exponential models)
Binary (dummy) variables. Adjusting for different slopes.
Diagnosing regression models: leverage, residuals, influential cases.
Building multiple regression models.

VII. logistic regression (if time permits)
Logistic regression models. Inference for logistic regression.

VIII. non-parametric tests (SDM2: ch. 29; SDM3: ch. 28)
Rank-based nonparametric tests. Wilcoxon rank sum test, Kruskal-Wallis test, Wilcoxon signed rank test

for paired data. Friedman Test for a randomized block design. Rank correlation.

IX. advanced topics (SDM2: ch. 30; SDM3: ch. 29)
missing data
bootstrapping