STA442/2101: Methods of Applied Statistics

University of Toronto, Fall 2016

http://www.utstat.toronto.edu/~brunner/appliedf16

Lecture: Friday 2:10-5:00 in SS2117.

- Instructor: Jerry Brunner
- Office: SS6025
- Phone: 416-978-0673
- email: brunner[at]utstat.toronto.edu

- Office Hours: Wednesday 12:10-2:00 in 341 Gerald Larking and Friday 12:10-1:00 in SS2123.

Note: I do not read my email every day, and the problem tends to get worse as the term progresses. It is much more efficient to talk with me before or after class, or during office hours.

Textbooks: These are available in PDF format free of charge. The texts are mostly for background reading in case you need some review, or you want to go beyond what is covered in lecture.

- *Linear models with R* (2009) by J. Faraway.
- *Linear models in statistics* (2008) by A. C. Renscher and B. G. Schaalje. A strong masters level regression text.
- *Statistical models* (2003) by A. C. Davison. This is the place to look if you want the real truth about almost any applied statistical topic.
- *Introduction to R* by Venables, Smith and others. This free 100 page document is very helpful if you plan to do serious work with R.

Topics: Random matrices and vectors, Normal linear models, Factorial ANOVA, Multiple comparisons, Power and sample size, Random effects, Within-cases designs, Permutation tests, Bootstrapping, Likelihood methods, Logistic regression and other generalized linear models. We will use R.

Prerequisite: For STA442, the prerequisites are STA302 and CSC108 or higher. For STA2101 the prerequisite is a course in linear regression. It is assumed that graduate students also have some computing experience.

Grading: There will be a quiz each week in lecture starting Friday September 30th. There will also be a comprehensive final exam. The quizzes count for 70% of your mark, and the final exam counts 30%.

There will be ten quizzes. The lowest quiz mark will be dropped. There will be an assignment for each quiz. The knowledge you need to do each quiz is a subset of the knowledge you need to do the corresponding assignment. Some (most) of the assignments include a computer part. You will bring printouts to the quiz and answer questions based on the printouts. Possibly, one of the quiz questions will be to hand in a printout. The non-computer parts of the assignments are just to prepare you for the quizzes; they will never be handed in.

Undergraduate and graduate students will take the same weekly quizzes, and the same criteria for marking will be employed. There will be a separate and more challenging final examination for the graduate students.

In spite of the 70-30 weighting of quizzes and the final exam, a good performance on the final can save a student from failing the course. Suppose your average including the final exam is a failing mark (less than 50% for undergraduates). If your mark on the final exam is at least 70%, or your mark on the final is at the class median or above, then you get the minimum passing mark for the course (50% for undergraduates). This rule is intended to give hope to those who have messed up on the quizzes, and encourage them to study for the final exam. It will be applied separately to graduate and undergraduate students, because the definitions of "passing" are different, and the medians will be different too.

Policy for missed work: If you miss a quiz, the mark is zero. However, your lowest quiz mark will be dropped. If you miss a quiz with a valid excuse, your mark on the final exam will be substituted for the missing quiz mark.

Academic Honesty: It is an academic offense to present someone else's work as your own, or to allow your work to be copied for this purpose. To repeat: the person who allows her/his work to be copied is equally guilty, and subject to disciplinary action by the university.

It is fine to discuss the assignments and to learn from each other, but there are clear limits on what is acceptable. It is okay to discuss the meaning of the question. It is okay to discuss general principles related to the question. It is okay, and encouraged, to discuss examples from lecture or textbook that are similar to the question. It is okay to reveal your approach to solving a problem, but only to somebody who has tried the problem and is really stuck. Even then, it is better to ask questions (like "Well, what's the model?" Or "All we've got is the Law of Large Numbers. Does anything in this expresion look like a sample mean?") instead of just giving your answer. A good rule is to never help someone who hasn't started yet.

But above all, **don't copy, and don't let anyone else copy from you**. You are expected to do the work yourself, and then *perhaps* compare answers after you have done so.

- For the non-computer parts of the homework: Never photocopy or photograph anyone's homework, or allow your homework to be photocopied or photographed. Copying by hand is also strongly discouraged. Your "friends" may ask you. You are expected to refuse.
- But you don't have to be rude. Instead of showing them your answer, walk them through an in-class example that is similar to the problem (there probably is such an example). Or make up another example or homework problem yourself. Then you can show all the details without any worries.
- For the computer parts of the homework: Never look at anyone else's printouts or show anyone your printout before the quiz when they might be handed in. Above all, do not allow anyone in the class to see your program file before a computer assignment is due, and do not look at anyone else's.
- You are allowed to compare numerical answers. Suppose one person says "What did you get for the MLE? My answer was 37.2." Three other people say Yeah, that's what I got, and one person groans and opens his laptop. This is all good.
- For some quizzes, you will be asked to bring your printouts to class; maybe you will hand them in, and maybe you will use them to answer some questions. Never, ever, bring a copy of somebody else's printout, or allow anyone to have a copy of yours. Again, your "friends" may ask you. You are expected to refuse.

You might be surprised to know how easy it is to detect copying on computer assignments. Here are some guidelines:

- It is permitted to copy from me. If your work is very similar to what is presented in lecture, office hours or suggested readings, that is okay.
- If two students have computer work that is excessively similar to each other, but *not* similar to what was presented in lecture or office hours, that is evidence of cheating. Of course it's easier to detect if the work is also wrong.
- If you allow anyone to have an *electronic* copy of your computer work, for any reason, you are not only guilty of an academic offence, you have lost your mind.
- *Direct copying of computer code from the internet is prohibited*. You are expected to do the work yourself. An exception is that you may download a publically available R package that does what you need to do.

It is acceptable to get help with your computer assignments from someone outside the class, but the help must be limited to general discussion and examples that are not the same as the assignment. As soon as you get an outside person to actually start working on one of your computer assignments, you have committed an academic offense.

For more detail, the latest version of the student handout "How not to Plagiarize" is available at http://www.writing.utoronto.ca/advice/using-sources/how-not-to-plagiarize The Academic Regulations of the University are outlined in the Code of Behaviour on Academic matters, which can be found in the Arts and Science Calendar or on the web at http://www.governingcouncil.utoronto.ca/policies/behaveac.htm.