

STATISTICS 220H1F - The Practice of Statistics – Fall 2007

Course website: <https://ccnet3.utoronto.ca/20079/sta220h1f/>

Important announcements, problem sets, past tests and exams, additional examples, and other course info will be posted on this course website. **Check it regularly.**

This course (in combination with Sta221S) presents the basic statistical methodology used in many fields of application. The emphasis of this course is on concepts and techniques and will be useful to students who seek to gain an understanding of the use of statistics in their own fields.

Required Texts

1. *Introduction to the Practice of Statistics (5th ed.)* - Moore and McCabe (*Freeman*). Do not purchase the 4th edition of the text. It is **highly recommended** that you buy the text bundled together with the Minitab (student version) statistical software CD - do not confuse this with the resource CD included with each text. This Minitab software is for Windows, but can also run on a Power Mac with appropriate Windows emulator software.

2. *Minitab Workbook* – A. Vukov (*Canadian Scholar's Press*)

Course Contents and Lecture Notes

Content, emphasis, etc. of the course is defined by means of the lecture material - *not the text*. The lecture notes are posted on the course website, it is **highly recommended** that you print a copy of them prior to lecture and bring them with you. The lecture notes contains examples and past exams questions. The only way to obtain solutions to those is to attend the lectures. It is important to attend all lectures, as there is normally no simple way to make up for missed lectures (perhaps obtain another student's notes). There will also be many lecture examples using Minitab statistical software, which students will be using.

Tutorials - starting September 17

Tutorials meet one hour per week. In preparation for tutorial, you should do the weekly assignment, posted on the web page. There will be some brief assessment at each tutorial - either there will be a short quiz based on the assignment, or you may be asked to submit a portion of your solutions. The major purpose of the tutorial is to go over the assignment and associated material; no new material will be covered. All records are kept by tutorial so please attend the correct one, as **missing marks** often arise when students attend tutorials they are not enrolled in. If you miss a tutorial/quiz due to illness, late enrolment, etc., bring an appropriate note to your next tutorial - your TA has full discretionary power to deal with such matters. At the end of term, your TA will submit a grade 0 -10 for you. Most students receive at least 9 out of 10, so the tutorial grade is designed to be an encouragement and reward for keeping up and doing the required work. All TA's will have to submit the same distribution of grades (same average, etc.) at the end, so that students are not advantaged or disadvantaged by their TA's approach. **The location of your tutorial will be posted on the board opposite the Statistics Aid Centre (SS2133) and on the course website, prior to 1pm September 17.**

Evaluation

The final course grade will be composed of:

Term test - time and location TBA	35%
Tutorial grade – best ($n-1$) quizzes	10%
Final Exam – multiple choice	55%

The term test will be written in various rooms across campus – check the course website later. There will be an alternative writing time, for those with proof of conflict – but if this applies to you, be sure to submit your course schedule to your lecturer at least 10 days before the test date.

If your final exam grade exceeds your test score, its weight will be increased to 70% (test decreased to 20%). You are allowed to have a one-sided 8.5" x 11" aid sheet on test (two-sided on final exam) and a nonprogrammable calculator. **You must bring your student identification to term tests as well as the final exam.** The final exam will consist of 30 multiple choice questions (4-5 selections each). A correct answer receives 3.33 marks. An unanswered question receives 0.67 marks.

Missed Tests

There are **no make-up tests**. Should you miss the term test due to illness, you are required to submit to your lecturer or to SS6018 (Stats office), within one week, completed by your doctor, the 'U of T Student Medical Certificate', obtainable from your college registrar, the Office of the Faculty Registrar (SS1006), the Stats Dept. office, or the Koffler health service. The test's weight will be shifted to the exam. **If this documentation is not received, your test mark will be zero.**

Computing

Students will be using, on a weekly basis, the Minitab Statistical Computing Software, either at home (if you purchase the student version of Minitab bundled inexpensively with the text), or at the CQUEST computing facilities at the University of Toronto. **No previous computing experience is assumed.**

With this software, you will analyze the data sets used in the text exercises. The data sets can be found on the CD-ROM accompanying the textbook, and at the publisher's web site

www.whfreeman.com/ips5e.

It is also possible to lease Minitab (professional version) for short-term use, or to download a 30-day free trial version, via the www.minitab.com website.

CQUEST computing accounts may be obtained by going to www.cquest.utoronto.ca and clicking on 'Request an account'. If you do not purchase and install Minitab on your home computer, you will need this CQUEST account to use Minitab at CQUEST sites on campus, see above website for rooms and hours of operation (it is not possible to log on to CQUEST's Minitab program from home).

In nearly every problem set, analyses using Minitab will be assigned - **bring the full computer output to tutorial**. Computing problems frequently arise, so do your assignment early, and be prepared to come back another day, if the printer gets jammed, etc. Expect no special consideration by your TA unless you have made repeated verifiable attempts to do the assignment, sought help, etc.

Calculators

Hand calculators are cheap and useful. Any cheap one with square root and one memory button will do. Mean, standard deviation, sum, and sum of squares keys may save you a bit of time on occasion, but we do not recommend the purchase of expensive calculators. Tests and exams will be designed so that those calculators give no advantage. We emphasize the use of Minitab software for doing any tedious or complex calculations. Programmable calculators are **not permitted** on tests and exam.

Statistics Aid Centre (starting September 17)

Your primary source of help with difficulties is your tutor in the scheduled tutorial, but additional assistance can be obtained at the Statistics Aid Centre, Room 2133, in Sidney Smith Hall. A schedule will be posted on the website and outside the Stats Dept Office (SS6018). Your own TA will be on duty there two hours per week, but you can drop in on any Sta220 TA present.

Additional Help

- Student electronic bulletin board (see website) – communicate with other students in the course.
- Department of Statistics – for general administrative queries only: stats@utstat.utoronto.ca
phone: 416-978-3452.
- **Additional aids:** A study guide for the Moore/McCabe text is available (order at the bookstore).
Minitab Manual for Moore/McCabe by Michael J. Evans.
- At the publisher's website, www.whfreeman.com/ips5e, there are additional learning aids such as online quizzes, supplementary exercises, and interactive applets.

Academic Offences

Academic offences are totally unacceptable and harm everyone. Offenders are caught, and **sanctions can be severe** - zero in the course with annotation on the transcript for several years; suspension for a year; even expulsion. Various measures, announced and unannounced, will be taken throughout the year to reduce their incidence and to ensure successful prosecution (photocopying of graded tests, multiple versions of multiple choice exams, etc.).

Important Final Comments

- For most students, statistics is a new subject. Like most new subjects, it seems difficult at first but simple in retrospect. Many new concepts will be introduced and built upon. The course will be far easier and more useful if you master concepts as they are introduced - similar to learning a new language. Doing the exercises is not like calisthenics or jogging, which are good for you no matter how your mind wanders as long as you go through the motions. Think about why a particular method was used, and what has been learned about the original questions motivating the study!
- The lecturers and TA's are there to help. Ask questions! If you have any comments or suggestions about the course, please tell us. We want to know what is good; we want to know when problems are developing. Remember that it takes time for information to travel and for action to be taken. Therefore, if you become aware of any problems, mention this to your instructor right away.

STATISTICS 220H1F Fall 2007 - TENTATIVE LECTURE GUIDE

Textbook (IPS by Moore/McCabe) chapter references are in parentheses, and indicate when a section has been finished. Do many of the odd problems in the text for practice (answers are in the back)

Week 1: Introduction to course, overview. Variables, units, measurement. Frequency, relative frequency, cumulative frequency distributions. Histograms. Shapes of distributions, outliers. Stem and leaf plot. Intro to Minitab (1.1)

Week 2: Summation notation. Sample mean, median, trimmed mean. Standard deviation. The empirical (68-95-99.7) rule. Calculating mean, standard deviation for grouped data. z-scores. Percentiles. Boxplots.

Week 3: Linear and non-linear transformations of data. Density curves and the normal distribution. Normal probability (quantile) plot. (1.2, 1.3) Bi-variate data: Scatterplot, correlation.

Week 4: The least-squares line, coef. of determination, residual plots, outliers & influential observations; lurking variables, association & causality. Comparing boxplots. (2)

Week 5: Collecting data: Randomized experiments & observational studies. Sample vs population characteristics. Simple random samples and variations. Sampling and non-sampling errors(bias). Designing experiments: comparison, randomization, blocking; factorial design (3)

Week 6: Introduction to probability. Sample space, outcomes, and events. Rules of calculation for probabilities. Conditional probability (with a Bayes example via tree diagram). Statistical independence. Two useful counting rules (combinations, sequences). Computer simulations.

Week 7: Discrete random variables and probability functions. Expectation (mean) and variance of random variables and linear combinations (e.g. sums, differences, averages). (4) The binomial distribution, applications, mean and variance (using sums of Bernoulli variables), the sample proportion.

Week 8: Continuous r.v.'s and density functions. The normal distribution and normal approximation of binomial. (5.1) Sampling distributions (worked out directly in a simple case). Distribution of the sample mean (and total). The Central Limit Theorem. (5.2)

Week 9: Margin of error, Confidence Intervals for μ (σ known or large n) and p , and sample size (6.1,8.1). Exact C.I. tables for p . Statistical tests, for μ (σ known) and p . Observed and fixed levels of significance (p -value, α -level). (6.2, 8.1)

Week 10: Decision errors - types I, II, and power (with a z-test calculation). Testing hypotheses via C. I.'s. Statistical vs practical significance. Multiple tests and data snooping. (6.3, 6.4) Test & C.I. for μ with unknown variance: the Student t distribution. Robustness of t procedures (examined via simulations). Power of the t-test (using tables, Minitab).

Week 11: Paired comparison t-test, and a nonparametric (distribution-free) approach - the sign test. (7.1). Two independent samples: large sample z-test & C.I. for comparing the means. Small sample t-tests comparing two population means: pooled/unpooled variance procedures.

Week 12: Comparing designs: paired comparison vs independent groups. Robustness, transformations, power (using Minitab). (7.2) Comparing proportions with z-test. (8.2). Testing equality of two population variances - the F distribution. (7.3)

Week 13: Overview of inferential procedures. Final case study (AZT trials).