STATISTICS 220F - The Practice of Statistics (Summer 2011)
Course homepage accessible from the U of T Portal at http://portal.utoronto.ca

This course (in combination with STA221H) presents the basic statistical methodology used in many fields of application. It emphasizes concepts and techniques and will be useful to students who seek to gain an understanding of the use of statistics in their own fields. Students seeking a more theoretical or mathematical development of the subject should consider STA250+255 or STA 257+261. However, STA220+221 (or STA220+JBS229) will be accepted in lieu of STA250H as a prerequisite for STA255, after which, (limited) further studies in Statistics may be pursued.

Tutorials - starting May 19

Tutorials meet Tuesday and Thursday 6-7pm. The major purpose of the tutorial is to go over the assigned homework questions and associated material; no new material will be covered. In preparation for tutorial, you should do the weekly assignment, posted at the course web page. *Assignment #1 (posted at the course web page) is due at your first tutorial.*

   There will be some brief assessment at each tutorial (0 – 5 scale) based on the assigned homework. This assessment may be based on a brief quiz and/or classroom discussion/ participation and/or portions of your written homework (which we generally do not collect). This assessment is meant to encourage and reward you for keeping up with the work, not to provide very meaningful feedback on your achievement level. Tutorial quizzes will be very brief (5 minutes perhaps), as they detract from the more valuable business of the tutorial: *discussion and review.*

   At the end of term, your TA will submit a 0 -10 grade for you. All TAs will be required to submit similar grades at course end, regardless of whether the quiz average is 2 or 5, so there will be no disadvantage due to your TA's particular approach. Make a good consistent effort throughout the term, as we expect of the majority of students, and you can expect to receive at least 9 marks out of 10.   **The location of your tutorial will be posted on the course web page** prior to May 19. Records are kept by tutorial so please attend the one you are properly enrolled in, *to prevent missing marks.*

   If you miss a tutorial/quiz due to illness, late enrolment, etc., please discuss with your TA, and not your lecturer. Your TA has full discretionary power to deal with, and adjust tutorial grades for, such things. If ill, bring some proof.

Course Content

Content, emphasis, etc. of the course is defined by means of the lecture - *not the texts.* It is important to attend all lectures, as there is normally no simple way to make up for missed lectures. There will also be many lecture examples using Minitab, the statistical software chosen for this course. *Important announcements, problem sets, handouts, practice tests, tutorial and aid centre info, etc., will be posted on the course homepage,* or occasionally emailed out. **So check the web site (via the portal) and your UTORmail account regularly.**

Required Texts

1. Introduction to the Practice of Statistics, 6th ed., by Moore & McCabe (Freeman). Do not purchase the 5th edition of the text. *We strongly recommend* that you buy the text *bundled together* with the Minitab software (student release 14), which gives you the Minitab software at *a big discount.* Don't confuse the Minitab installation CD with the textbook resource CD included with every text. Note that Minitab software runs only on the Windows operation system.   (We will abbreviate this text as **IPS**)

2. Minitab Workbook (Fall 2009 or Summer 2010, but not earlier, since earlier versions do not correspond to the 6th ed of IPS), by A. Vukov (Canadian Scholar's Press)

You can avoid some frustration, by ordering your books online from the bookstore: www.uoftbookstore.com/online/ The publisher's web site, www.whfreeman.com/ips6e, had additional learning aids such as online quizzes, supplementary exercises, and interactive applets. A study guide (by M. A. Fligner) for the Moore & McCabe text is

also available (but you have to order it, at the bookstore).

Additional help from your TAs (from May 19)

Your primary source of help with difficulties is your TA in the scheduled tutorial, but additional assistance will be available at the **Statistics Aid Centre**, Room 2133, in Sidney Smith Hall and/or the **CQUEST** computing site in the Ramsey Wright building. A schedule will be posted at the course website. Your own TA will be available for help at one or both of these sites for a total of two hours per week, but you can drop in on any Sta220 TA.

Additional Help

There are electronic discussion forums at the course web page where you can communicate with other students in the course. For very general administrative queries, the Statistics departmental office can be reached at stats@utstat.utoronto.ca or 416-978-3452.

Evaluation

Your tutorial grade is worth 10% of your final grade. The remaining portion of your grade will be equal to either {0.35 x Term Test + 0.55 x Final Exam}, *or* {0.15 x Term Test + 0.75 x Final Exam}, whichever is higher.

Term Test: Tues, June 7, 6:00-7:40pm*, Locations: TBA

Final Exam: Tues, June 28, at 6 or 7pm (3 hours), Location: TBA - date and time to be confirmed

*followed by lecture. Though tests/exams place a very minor emphasis on formulae, you are allowed a one-sided 8-1/2"x 11" (standard letter size) hand-written aid sheet on the term test (two-sided on final exam). **Programmable calculators are not permitted on tests or exam. 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 (5 choices each). A correct answer receives 3.33 marks. An unanswered question receives 0.67 marks. The term test is, for the most part, not multiple-choice. Both will be similar in format to some old term tests and final exams that will be posted at the course web page.

Missed Tests

There are **no make-up tests**. Should you miss the term test due to illness, you are required to submit, within one week, **a 'U of T Student Medical Certificate'**, completed by your doctor, to your lecturer or to SS6018 (Stats office). You can obtain this form from your college registrar, the Office of the Faculty Registrar (SS1006), 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.**

Calculators

Hand calculators are cheap and useful. Any cheap one with a 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 to get keys with special statistical calculations. 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.

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 - go to www.cquest.utoronto.ca for info about accounts, rooms, hours of operation, etc. You will use this software to analyze data sets used in the text exercises. These data sets can be found on the CD-ROM accompanying the textbook, and at the publisher's web site www.whfreeman.com/ips6e. It is also possible to lease Minitab (professional version) for short-term use at www.minitab.com/academic (click on *Pricing and Licensing*), and to download a 30-day free trial version from 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 (unfortunately it is not possible to log on to CQUEST's Minitab program from your home). Stick to the CQUEST rooms in Ramsey Wright Bldg.

You will use Minitab on every assignment. **Always bring to tutorial the full computer output,** along with your written answers. 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 for 'technical' problems' unless you have made repeated verifiable attempts to do the assignment, sought help, etc.

Academic Offences

Academic offences are unacceptable and harm everyone. E.g., some students have been known to alter a test paper after it is graded, and then re-submit it hoping for more marks - really! But, 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.). And please carefully note the following:

- ❖ Requests for a test remark will be considered only if you **_write your test in ink_**.
- ❖ **_Grading oversights_** such as addition errors and overlooked work must be reported to your TA _immediately_ upon receiving your test paper at class. _So check it over as soon as you get it back._

Feel free to discuss weekly assignments with others, but in the end, be sure to do all the computations, work and final write-up by yourself.

Accessibility Needs

The University of Toronto is committed to accessibility. If you require accommodations for a disability, or have any accessibility concerns, please contact Accessibility Services as soon as possible: disability.services@utoronto.ca or http://studentlife.utoronto.ca/accessibility

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 the objectives of the study, why a particular method was used, and what has been learned about the original questions that motivated 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 instructors right away.

Textbook (IPS 6th edition by Moore/McCabe) chapter/section references are in parentheses. At the same time, read the corresponding chapters in the *Minitab Workbook*.
Do many of the odd problems in IPS for practice (and the answers for these are in the back)

Lecture 1: Introduction to course, overview. Categorical & quantitative variables. Cases (individuals). Worksheet organization of data. Relative frequency distribution. Bar charts & Pie Charts for categorical variables. Dotplots, histograms, stem and leaf plots. Shapes of distributions, outliers. Introduction to Minitab. (1.1)

Lecture 2: Summation notation. Sample mean, median, mode (modal class), trimmed mean. Resistant measures. Standard deviation. The empirical rule. Calculating mean &standard deviation for grouped data. Percentiles. 5-number summary and boxplots. Comparing boxplots. Z-scores. Linear and non-linear transformations of data. (1.2, 1.3)

Lecture 3: Density curves and the normal distribution. Normal probability (quantile) plot. Bivariate data: Scatterplots & correlation. Cautions. (1.3, 2.1, 2.2)

Lecture 4: The least-squares line, coefficient of determination, residual plots, outliers & influential observations; lurking variables, association & causality. (2.3, 2.4, 2.6)

Lecture 5: Collecting data: Randomized experiments & observational studies. Designing experiments: comparison, randomization, blocking; factorial design. Sample vs population characteristics. Random samples (SRS, stratified, systematic, cluster & multistage). Sampling and non-sampling errors. Bias. (3)

Lecture 6: Probability and relative frequency. 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, sequential operations). Estimating probabilities by computer simulation. (4.1, 4.2, 4.5)

Lecture 7: Discrete random variables and probability functions. Expectation (mean) and variance of random variables and linear combinations (e.g. sums, differences, averages). The binomial distribution, applications, mean and variance (using sums of Bernoulli variables), the sample proportion. (4.3, 4.4, 5.1)

Lecture 8: Continuous random variables and density functions. The normal distribution and normal approximation of binomial. Sampling distributions. Distribution of the sample mean (and total). The Central Limit Theorem. (4.3, 5.2)

Lecture 9: Margin of error, Confidence Intervals for μ (σ known / large n) and p (large sample), and sample size. Exact and Plus 4 confidence intervals for p. Statistical tests of hypothesis: observed and fixed levels of significance (p-value, α-level). Tests for μ (σ known) and p. (6.1, 6.2, 8.1)

Lecture 10: Decision errors - types I & II, and power (with a z-test calculation). Testing hypotheses via confidence intervals. Statistical vs. practical significance. Multiple tests and data snooping. Test & C.I. for μ with unknown variance: the Student t distribution. Robustness of t procedures (via simulations). Power of the t-test (using graphs or Minitab). (6.3, 6.4, 7.1)

Lecture 11: Paired comparison t-test, and a nonparametric (distribution-free) approach - the sign test. 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. (7.1, 7.2)

Lecture 12: Comparing designs: paired comparison vs. independent groups. Robustness, transformations, power (via Minitab). Comparing proportions with z-test. Final case study (video: the AZT trials). Overview of inferential

procedures.   (7.2, 7.3, 8.2)