

## STATISTICS 250H1 – Statistical Concepts (Fall 2008)

Course homepage accessible from the U of T Portal at <http://portal.utoronto.ca>

Lecturer: Augustin Vukov, Office hrs: Thursday: 1-2pm and Friday: 12-1pm, in SS6025 (416-9784722)

This course 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. For further studies in Statistics, you need to also take STA255, which will provide more of the theoretical and mathematical development required.

### Tutorials - starting Sept 18

Tutorials meet one hour per week. **In preparation for each tutorial, you should do the weekly assignment, posted on the course web page. Assignment #1 is due at your first tutorial.** One homework assignment is due at each tutorial. 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. **The location of your tutorial will be posted on the web page, prior to 7pm Sept 17.**

There will be some brief assessment (0-5 grade) for each tutorial – perhaps a short quiz based on the assignment; perhaps you'll be asked to submit your solutions for one or two of the assigned questions; perhaps your TA will have other good ideas. At the end of term, your TA will submit a grade 0 -10 for you. This grade is meant to be an encouragement and reward for keeping up with the required work. Make a good, consistent effort and you can expect to receive at least 9 marks out of 10 (regardless of whether your TA gives easy quizzes or hard quizzes). **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.**

### Course Contents

Content, emphasis, etc. of the course is defined by means of the lectures - *not the text*. It is important to attend all lectures, as there is normally no simple way to make up for missed lectures. **Important announcements, problem sets, and other course info will be posted on the course web homepage (see above). Check it regularly.**

### Required Texts

**Introduction to the Practice of Statistics (5<sup>th</sup> ed.)** - Moore and McCabe (*Freeman*). Do not purchase the 4<sup>th</sup> or the new 6<sup>th</sup> edition of the text.

**Highly recommended:** The above textbook bundled together with Minitab statistical software (giving you the Minitab software at a greatly reduced price). Note that Minitab runs only on Windows operating system.

**Additional references:** *Minitab Workbook* - A. Vukov (*Canadian Scholar's Press*)

You can avoid some frustration, by ordering your books online from the bookstore: [www.uoftbookstore.com/online/](http://www.uoftbookstore.com/online/) At the publisher's web site, [www.whfreeman.com/ips5e](http://www.whfreeman.com/ips5e), there are 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 (order at the bookstore).

### Statistics Aid Centre: SS2133 (from Sept 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 at the website. Your own TA will be on duty there one hour per week, but you can drop in on any STA250 TA present. (and maybe even some STA220 TAs)

Also check out the **New College Aid Centre (room 55B)**, <http://www.utstat.utoronto.ca/~cao/newcollege.htm>, which will have one or two highly qualified Stats Dept TA's (e.g. Shelley Cao) present on a fixed weekly schedule, to help with several intro stats courses including STA250.

### Additional Help

**Student online discussion forums (at web site)** – communicate with other students in the course.

**Dept of Statistics** – for general administrative queries only: stats@utstat.utoronto.ca (phone: 416-9783452)

### Evaluation

Your tutorial grade is worth 10% of your final grade. The remaining portion of your grade will be equal to either  $\{0.35 \times \text{Term Test} + 0.55 \times \text{Final Exam}\}$ , or  $\{0.15 \times \text{Term Test} + 0.75 \times \text{Final Exam}\}$ , whichever is higher! The term test is scheduled for Date: TBA, Rooms: TBA, at 10am-12noon or 11am-1pm, depending on your tutorial. **Programmable calculators are not permitted on tests and exam.** Though tests/exams place a very minor emphasis on formulae, you will be allowed a one-sided 8-1/2"x 11" hand-written aid sheet on the term test (two-sided on final exam). **Be sure to bring your student identification to the term test as well as the final exam.** The final exam may contain multiple-choice questions.

### 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 A. Vukov 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 (SS 1006), 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.**

### 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 to get keys with special statistical calculations. Tests and exams will be designed so that those calculators give no significant advantage.

### Computing

**You will need access to a statistical computing package, from day one, in the course.** You will need it to analyze data sets, do simulations, etc. I want you to spend your time learning statistics, not computing languages, so I have chosen Minitab for the course due to its menu-driven, user-friendly nature and considerable capabilities. Minitab (student release 14) is available bundled at a greatly discounted price together with the text – a good choice for students! It has also been installed at the CQUEST computing labs on campus – **sign up for an account at [www.cquest.utoronto.ca](http://www.cquest.utoronto.ca)**. We strongly recommend using the Windows machines in Ramsay Wright CQUEST labs, not the Linux machines in the Earth Sciences CQUEST lab. You may well be able to learn how to use it through trial and error, along with Minitab's Help function and tutorials (on the Minitab menu, choose **Help>Tutorials**); and additional resources are available at <http://www.minitab.com/resources/tutorials/default.aspx>. Another reference book, available at the bookstore, is the **Minitab Workbook** (by A. Vukov). At [www.minitab.com](http://www.minitab.com), you can download a free 30-day trial (professional) version for free. You can also find instructions at [www.minitab.com/education](http://www.minitab.com/education) for short-term lease of Minitab (e-academy semester rental).

While it is best if we all stick to one package, and **I will use Minitab for lecture and test examples**, you may choose to use another one for your assignments if you wish, but note that free internet-based packages simply do not compare with the high-quality professional ones like Minitab. MS Excel statistical add-in packages are more costly than (student) Minitab. Web based [www.statcrunch.com](http://www.statcrunch.com) is nearly free, and not half bad.

The data sets for the text exercises can be found on the CD-ROM accompanying the textbook, both in Minitab portable worksheet format (\*.mtp), and in other formats, e.g. text format, which may be useful for some other computing packages. These data files are also available at the publisher's web site [www.whfreeman.com/ips5e](http://www.whfreeman.com/ips5e).

In nearly every problem set, exercises using Minitab will be assigned - **bring the full computer output to tutorial**

Computing problems (jammed printer, etc) frequently arise, so *do your assignment early, and be prepared if necessary to come back another day*. Sometimes the CQUEST rooms can get extremely busy, so be sure to check out all the CQUEST locations on campus (see [www.cquest.utoronto.ca](http://www.cquest.utoronto.ca)), and not just the nearest one.

#### Academic Offences

**Academic offences are totally unacceptable** and harm everyone. E.g., we have found that some students will alter a test paper after it is graded, and try re-submitting it! 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, etc.). Also, please 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 over your test paper right away.*

#### 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 instructor(s) right away.

#### TENTATIVE LECTURE GUIDE

IPS chapter references are in parentheses. Do many odd numbered exercises for practice - answers are in the back of IPS

**Lecture 1:** Types of data. Descriptive graphs: stem plots, histograms, boxplots. Measures of centre and dispersion. Transformations of data. (chap 1)

**Lecture 2:** Normal distribution. Bivariate data. Correlation. (chap 1, 2)

**Lecture 3:** Fitting by least-squares. Residual plots. Sample surveys and probability samples. (chap 2, 3)

**Lecture 4:** Experimental design concepts. Some basics of probability. (chap 3, 4)

**Lecture 5:** Random variables, expected values, variance. Binomial distribution. (chap 4, 5)

**Lecture 6:** Central Limit Theorem. Estimation and confidence intervals. (chap 5, 6)

**Lecture 7:** Significance tests (for  $\mu$  and  $p$ , one-sample). P-values. Type I, type II errors. (chap 6)

**Lecture 8:** The t-distribution: testing means in single samples, paired data and independent samples. (chap 7)

**Lecture 9:** Categorical data: the z-test for proportions and the chi-square test for goodness of fit and association. (chap 8, 9)

**Lecture 10:** The regression model. Inference for regression. Introduction to multiple regression: an example. (chap 10, 11)

**Lecture 11:** Multiple regression and the General Linear Model. One-way ANOVA (chap 11, 12)

**Lecture 12:** Two-way and multi-way ANOVA. Non-parametric tests. (chap 13, 15)

**Lecture 13:** Catch up week (finish off earlier topics). Introduction to bootstrapping and randomization tests. Final case study (AZT trials). Overview of inferential procedures. (chap 14)