

**STA221H1S - The Practice of Statistics II - Winter 2008**  
(<http://ccnet.utoronto.ca/20081/sta221h1s>)

This course is a continuation of STA220H, with emphasis on the basic statistical methodologies needed in a broad variety of fields: regression, experimental design and analysis of variance, chi-square and non-parametric procedures. The emphasis is on understanding the concepts and careful application of the basic techniques, using realistic data sets and Minitab software (available at the CQUEST computing labs on campus and also purchasable for home PC).

**Tutorials**

Tutorials begin Jan 14. Tutorials meet on Monday at either 2-3pm or 4-5pm. Register at the course web site indicated above as a student, and you will be able to see your assigned tutorial, and then find the room from the listing under the tutorials side-link. Problem assignments (requiring Minitab) will be posted at the course web site. They are due at tutorial (for hand-in, discussion, or quizzing). Records are kept by tutorial so be sure to attend the correct tutorial, and to know your TA's name. The location of your tutorial will be posted prior to Jan 14 at the course web site.

**Texts/Software:**

***Required:*** *The Practice of Statistics (Course Notes for STA221), by A. Vukov*

***Recommended:*** *Introduction to the Practice of Statistics by Moore & McCabe, 5<sup>th</sup> ed. (IPS)*

The former is a complete set of course notes based on the lectures by A Vukov, including relevant Minitab commands, some statistical tables, and sample exam questions. I will reference IPS (your sta220 text) only for its exercises and data sets (data sets may be downloaded from [www.whfreeman.com/ips5e](http://www.whfreeman.com/ips5e)). Both are available at the campus bookstore - Save 5% and avoid frustration by ordering online at [www.uoftbookstore.com/online](http://www.uoftbookstore.com/online). **Earlier editions of either book are not acceptable.** *The student version 14 of Minitab (software) is available bundled with the Moore & McCabe text, at a bargain price. If you do not purchase Minitab for home use, please request, as soon as possible, a computing account at [www.cquest.utoronto.ca](http://www.cquest.utoronto.ca).* It is also possible to lease Minitab for a term (check at [www.minitab.com](http://www.minitab.com)).

**Statistics Aid Centre (from Jan 14)**

Your primary source of help with difficulties is your TA in the scheduled tutorial, but additional assistance can be obtained at the Statistics Aid Centre, Room 2133, in Sidney Smith Hall. Your own TA will be on duty there one hour per week, but you may drop in on any of the TAs for the course. Some TAs will also be available at the CQUEST computing room in SS, for several hours per week. Schedules will be posted at the course web page. Also, check out the Aid Centre at New College room 55B, where a very experienced Statistics TA will hold regular hours (about 8 hrs per week).

**Evaluation**

Your final course grade will be composed of a term test (35%), multiple choice Final Exam (55%), and tutorial evaluation (10%). However, if your final exam grade exceeds your test grade, its weight will be increased from 55% to 75% (test decreased to 15%). The term test is tentatively scheduled for Feb 25, 2008 at either 2:10-4pm or 3:10-5pm, depending on your tutorial (your sitting is the one that overlaps your tutorial time), in rooms: TBA. **Programmable calculators are not permitted on tests and exam.** A one-sided 8-1/2"x 11" aid sheet, hand-written, is allowed on tests (two-sided on final exam). **You must bring your student identification to term tests as well as the final exam.** The exam will likely consist of 30 multiple choice questions (5 choices each) - correct answers get 3.33 marks, unanswered questions get 0.67 marks.

**Missed Tests**

There are **no make-up tests**. Should you miss the term test due to illness, you must submit to your lecturer or to SS6018 (Stats office), within one week, completed by yourself and 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 then be shifted to the final exam. **If this documentation is not received, your test mark will be zero.**

**Academic Offences**

**Academic offences are 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 insure successful prosecution when they do occur (e.g. photocopying of students' tests, multiple versions of multiple choice exams). In addition, please note the following:

(i) **Oversights in marking on a test paper** (e.g. addition error, overlooked work) must be brought to the attention of the T.A. **immediately** - during the period when test papers are returned.

(ii) **Regrading requests** will only be considered for **term tests** which are written in **ink**

### **Tentative Lecture Outline**

[Course Notes *page references* in brackets] (*IPS chapter references* in parentheses)

**Week 1:** Overview. The chi-square goodness of fit test, for categorical data. The Poisson distribution. Testing a hypothesized distribution. Examining and testing for association. Simpson's Paradox. (ch9) [pp1-25]

**Week 2:** Regression with one explanatory variable: the regression model, estimating the coefficients by least-squares, the ANOVA and F-test. Inference for parameters. Residual plots. Linearizing transformations. (ch2,10) [pp.26-45]

**Week 3:** Prediction and Confidence Intervals at specified  $x$ . Testing lack of fit with replicates at some  $x$ 's. Inference for the correlation. Introduction to regression with several explanatory variables. (ch10) [pp34-60]

**Week 4:** The essentials of multiple regression (using Minitab output): ANOVA table,  $R^2$ , global F-test, residual plots, inference for individual parameters, prediction/confidence intervals. Interaction and polynomial models. The General Linear Model (GLM) and transforming to GLM (e.g. exponential models). Testing portions of a model via the Extra SS principle. (ch11) [pp57-80]

**Week 5:** Multicollinearity. Use of dummy variables. Stepwise Regression & Model Building. [pp80-96]

**Week 6:** The completely randomized design and one-way ANOVA. Variation between and within samples, ANOVA table, F-test, CI's. Minitab output. Residual plots and transformations. Relation of one-way ANOVA to the two-sample t-test. (ch12) [pp.97-106]

### **READING WEEK: NO CLASSES**

**Week 7:** // **TERM TEST tentatively scheduled on FEB 25 on weeks 1- 6 material** // Planned comparisons (contrasts) among the means: t-test & F-test. Orthogonality and decomposing the (Between)Treatment SS. (ch12) [pp106-116]

**Week 8:** Analyzing a quantitative factor with orthogonal polynomials. Post-hoc comparisons (Bonferroni, Fisher, Tukey, Scheffe methods). (ch12) [pp116-129]

**Week 9:** Orthogonal comparisons for a 2x2 factorial design. The general two-factor factorial arrangement. Interaction and additivity. ANOVA summary table, tests, CIs. Planned/post-hoc comparisons for the factors or treatments. (ch13) [pp130-153]

**Week 10:** Three factor experiments. The GLM and missing observations. Randomized block design. Factorial experiments run in complete blocks. Latin square design. [pp153-168]

**Week 11:** Analysis of Covariance. Fixed & random factors, crossing & nesting. Expected MS's and F-ratios. The split-plot and other designs. [pp169-187]

**Week 12:** Introduction to some nonparametric and computer intensive (re-sampling) approaches: the sign test, Wilcoxon tests and multi-sample extensions, Spearman's rank correlation; the Bootstrap. (ch14 & 15) [pp188-205]

**Week 13:** Catch-up time. Additional topics, time permitting (e.g. randomization tests, logistic regression,...). Overview