

## STA221H1S - The Practice of Statistics II - Winter 2012

Instructor: Olga Chilina. Office: SS6025

Office hrs: M 1-2 pm

Email: [olgac@utstat.utoronto.ca](mailto:olgac@utstat.utoronto.ca) or [ochilina@utsc.utoronto.ca](mailto:ochilina@utsc.utoronto.ca)

Webpage: <http://www.utstat.toronto.edu/~olgac/>

### Course Objective

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 16. Tutorials meet on Monday at either 2-3pm or 4-5pm. Be sure to register for a tutorial section at ROSI. Tutorial rooms will be posted on the course web site prior to Jan 16. Problem assignments will be posted on the course web site. They are due at tutorials, for discussion.

### Texts/Software:

**Required:** Stats: Data and Models, First Canadian Edition., by DeVeaux, Velleman, et al (Pearson). Abbreviated SDM. It is strongly recommend to buy the text bundled together with the Minitab software (student release 14), which gives you the Minitab software at a big discount.

Another textbook to be used in this course is The Practice of Statistics (Course Notes for STA221), by A. Vukov (**strongly recommended**). Abbreviated PS. This 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. It refers to the old textbook. So should any exercises be assigned in the class, I will post the corresponding data on the web.

### Statistics Aid Centres

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 one hour per week, but you may drop in on any of the TAs for the course. Schedules will be posted on the course web page. Also, check out the New College Stat Aid Centre, where experienced Statistics TAs will hold regular hours.

### Evaluation

Tutorial Weekly Quizzes: 10%

Midterm Test: 40%

Final Exam: 50%

### Quizzes

Quizzes will be given in tutorial. A typical quiz will be a multiple choice question, you get either 1 mark (for the correct answer) or 0.3 (for attendance). Your TA will record your mark for each quiz. So be sure to attend the correct tutorial, and to know your TA's name.

### Midterm Test/Final Exam

The midterm test is on Feb 29. **Programmable calculators are not permitted on tests and exam.** A one-sided 8-1/2"x 11" aid sheet, hand-written, is allowed on the test (two-sided on final exam). **You must bring your student identification to term tests as well as the final exam.** The day and time for the final exam will be announced later.

### 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 tutorial class when test papers are returned
- (ii) **Regrading requests** will only be considered for **term tests** which are written in **ink**

### **Tentative Lecture Outline**

**Week 1:** Review of the last topics covered in STA220. Comparing counts. Two-way tables. The chi-square statistics, P-value. (SDM: ch. 26)

**Week 2:** Goodness-of-fit. Independence. The Poisson distribution. (SDM: ch. 26, PS: pp 15-18)

**Week 3:** Inferences for regression. Simple linear regression. Regression parameters estimates. Confidence intervals for mean value of  $y$  at specified  $x$ . Prediction intervals (SDM: review of ch. 8-9, ch. 27; PS: pp 26-45)

**Week 4:** Linearizing Transformations (SDM: ch. 10; PS: pp 45-46). Analysis of Variance. The one-way ANOVA. The ANOVA model. The ANOVA table. The ANOVA F-test. (SDM: ch. 28)

**Week 5:** Planned comparisons (contrasts) among the means: t-test & F-test. Orthogonality and decomposing the Between Group SS. Pooled Contrasts. Multiple Comparisons. (PS: pp 106-129)

**Week 6:** Multifactor Analysis of variance. The two-way ANOVA model. The ANOVA table for two-way ANOVA. Interactions. Inference when effects of variables are related. (SDM: ch. 29, PS: pp 134-147)

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

**Week 7: Monday:** Review for the midterm test.

*Wednesday:* // **TERM TEST tentatively set for FEB 29 on weeks 1- 6 material** //

**Week 8:** Inference for multiple regression. Multiple linear regression model. CIs and significance tests for regression coefficients. ANOVA table for multiple regression. (SDM: ch.30, PS: pp 57-67). The General Linear Model (GLM) and transforming to GLM (e.g. exponential models). (PS: pp 78-79)

**Week 9:** Use of dummy variables (indicators). Adjusting for different slopes. Diagnosing regression models: leverage, residuals, influential cases. Building multiple regression models. (SDM: ch. 31)

**Week 10:** Logistic regression model. Inference for logistic regression.

**Week 11:** Rank-based nonparametric tests. Wilcoxon rank sum test, Kruskal-Wallis test, Wilcoxon signed rank test for paired data. Friedman Test for a randomized block design. Rank correlation. (SDM: ch. 32; PS: pp 188-198)

**Week 12:** Bootstrapping. (SDM: ch. 33; PS: pp 199-204) Review for the final exam.