STA221H1S - The Practice of Statistics II - Summer 2014

Instructor: Augustin Vukov SS6024A, 416-9784722, Office hrs: TR3-4pm (often available later), and also available at the lecture room after each lecture

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 July 8. Tutorials meet on Tuesdays and Thursdays at 6-7pm. Problem assignments (requiring Minitab) will be posted at the course web site. They are due at tutorial, for discussion and quizzing. *There is an assignment due at the first tutorial.*

Texts/Software:

Required: The Practice of Statistics: Course Notes for STA221, by A. Vukov (Canadian Scholar's Press)
Exercises will also be assigned from your STA220 textbook: Stats: Data and Models, First Canadian Edition, by DeVeaux, Velleman, et al (Pearson)
  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 *Stats: Data and Models* mainly for its exercises and data sets (which may also be downloaded from http://wps.pearsoned.ca/ca_ph_deveaux/204/52457/13429053.cw/index.html). Both texts are available at the campus bookstore - Avoid frustration by ordering online at www.uoftbookstore.com/online. Minitab software will be an *extremely important tool* for this course. The student version 14 of Minitab is available bundled with the SDM text, at a great 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.** It is also possible to lease Minitab (professional version) for short-term use at www.onthehub.com/minitab and to download a 30-day free trial version from the www.minitab.com website.

Statistics Aid Centre (from July 8)

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 1091 in Sidney Smith Hall. Your TA will be on duty there two hours per week. *For further assistance, try using the course online discussion forum.*

Evaluation – tentative for now

Your final course grade will contain a tutorial/assignment component worth 15%. There will be a term test & final exam worth either 35% & 50% respectively, or 15% & 70% respectively, depending on which calculation gives you a higher course grade.

The term test is *tentatively scheduled* for July 24 at 6pm, Room: 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 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 tutorial class when test papers are returned.

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

Lecture 1: Overview. The chi-square goodness of fit test, for categorical data. The Poisson distribution. Testing a hypothesized distribution. Testing for association between variables. Simpson's Paradox. (ch26) [pp1-25]

Lecture 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. (ch27) [pp.26-45]

Lecture 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. (ch27/30) [pp34-60]

Lecture 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. (ch30/31) [pp57-80]

Lecture 5: Multicollinearity. Use of dummy variables. Stepwise regression & Model building. (ch31) [pp80-96]

Lecture 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. (ch28) [pp.97-106]

Lecture 7: **// TERM TEST on Lectures 1- 6 material //**
Planned comparisons (contrasts) among the means: t-test & F-test. Orthogonality and decomposing the Between Group SS. [pp106-116]

Lecture 8: Analyzing a quantitative factor with orthogonal polynomials. Post-hoc comparisons (Bonferroni, Fisher, Tukey, Scheffe methods). (ch28) [pp116-129]

Lecture 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. (ch29) [pp130-153]

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

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

Lecture 12: Introduction to some nonparametric and computer intensive (re-sampling) approaches: the sign test, Wilcoxon tests and multi-sample extensions, Spearman's rank correlation; Bootstrapping standard errors and confidence limits. (ch32/33) [pp188-205]