STAT300 – Intermediate Statistics for Applications

The following is a guide to the course STAT300, "Intermediate Statistics for Applications", as offered in the first Winter Term of the 2016/2017 academic year.

Aims and objectives: The course aims to be a second course in statistical science, reinforcing and extending ideas encountered in a typical first course in the discipline. The course will expose learners to a wide range of applied statistical methodology, complementing concepts appearing in their first course. Detailed learning objectives for the course will be available on-line on the course web-page.

Pre-requisites: One of STAT200, STAT241, STAT251, COMM291, BIOL300, FRST231, ECON325, or equivalent.

Co-requisites: None.

Lecturers: Lang Wu

Lectures and learning support: M/W/F 12-1 PM in room ESB 1012.

Assessment: Completion of lab activities (10%), a fifty-minute midterm exam (20%), responses to clicker questions (5%), on-line WeBWorK homeworks (5%), a 2.5-hour unseen examination (40%), and two assignments (10% each). Clickers must be registered via the STAT300 page on Connect (connect.ubc.ca). It is necessary to pass the final examination to pass the course. The usual university rules for extenuating circumstances and plagiarism apply. Dates for the setting and completion of the written assignments and on-line WeBWorK homeworks will be announced on the first day of classes. Specific details regarding assessment regulations for the course will be available on-line on the course web-page.

Teaching methods: Classes of approximately fifty minutes of duration will occur three times a week, with sets of notes being available from the course web-page in advance. In all sessions an in–class activity will replace at least part of the lecture component. Guided reading or other activities may be set at the end of one lecture to be completed prior to the next. On-line "pencasts" are available covering some of the course material. There will be required lab sessions most weeks. The web page will include detailed material covering the course content, plus other sundry resources like solutions to exercises when appropriate and an on–line forum. To access the forum, your user name is the first eight characters of the your name (including middle name if applicable, and not case sensitive), the password is S followed by the first seven digits of your student number. A calculator will be necessary for many of the activities.

Programme of work: The study time should total around eight hours per week. So in addition to the contact hours, it is essential that learners spend no less than four hours per week on self–study for the course. It is suggested at least two hours per week are spent on revising and assimilating the material covered in the lectures or on guided reading, and at least two hours should be spent attempting the exercises and assignments that are set.

Feedback: After all assignments have been submitted and marked, individual feedback will be provided in the form of brief notes on marked work. Detailed written comments will also be provided on the course web–page where appropriate.

Recommended texts: There is no core text, but there are numerous books that cover at least some of the material in this course, and it is suggested you try the UBC library stock to find those that suit you. There are few books that aspire to support a second course in Statistics. A good one though is

> Ramsey, F.L. and Schafer, D.W. (2002): The Statistical Sleuth: A Course in Methods of Data Analysis (2nd edition). Brookes/Cole.

It is likely that the textbook used for a pre-requisite course will cover some of the material in this course. In particular, later chapters of

> Moore, D.S. and McCabe, G.P. (2012): An Introduction to the Practice of Statistics. (7th edition). Freeman.

include content relevant to this course. Similarly other introductory texts are useful in containing parts of the content of the course, such as

- Walpole, R.E, Myers, R.M., Myers, S.L. and Ye, K. (2007): Probability and Statistics for Engineers and Scientists. Pearson/Prentice Hall.
- Whitlock, M. and Schluter, D. (2008): The Analysis of Biological Data. Roberts and Company.

There are useful books available electronically via the library (more details in library rep session). These include the following which provide details for implementing methods used using the statistical software package R:

- Ekstrom, C. T. (2012): The R Primer. Chapman and Hall/CRC
- Hay-Jahans, C. (2012): An R Companion to Linear Statistical Models. Chapman and Hall/CRC
- Hothorn, T. and Everitt, B.S. (2010): A Handbook of Statistical Analyses Using R. (2nd edition) Chapman and Hall/CRC

There follows a provisional guide to the lecture slots available. It is possible the material covered in the classes will differ slightly from the description below.

1. Introduction, motivation. Pre-test.
2. Library representative talk. Review of fundamental ideas.
3. Review activities.
4. Nonparametric methods: The sign test.
5. The rank sum test.
6. The Kruskal–Wallis test.
7. Permutation tests.
8. The power of hypothesis tests.
9. Investigating the fit of a model.
10. The Chi-squared test of goodness–of–fit.
11. Goodness–of–fit for contingency tables.
12. Fisher's exact test.
13. Probability plots for model fitting: Normal scores plots
14. Experimental design review: response variables, factors, blocking.
15. ANOVA: Review of concepts.

16. Analysing variance by breakdown of sums of squares.
17. Multiple comparisons
18. Midterm test
19. Interaction in two–way ANOVA
20. Inference in two–way ANOVA
21. Contrasts
22. Selected alternative designs: nested, factorial and fractional designs
23. Introduction to the bootstrap
24. Bootstrap testing and interval estimation
25. Review of regression concepts
26. Sums of squares in regression
27. Properties of estimators in regression
28. Multiple linear regression
29. Curve fitting via regression
30. Residuals in regression
31. ANCOVA
32. Odds ratios for 2x2 tables
33. Introduction to logistic regression
34. Introduction to time series: descriptive methods
35. Smoothing time series

BD