

STATISTICS 501 - SYLLABUS AND COURSE REQUIREMENTS - Spring 2009

Dr. W. J. Ewens

wewens@sas.upenn.edu

215 898 7109

324 Leidy Labs, (corner 38th St. and Hamilton Walk)

The aim of this course is to discuss non-parametric (distribution-free) statistical procedures. Comparisons with the corresponding distribution-dependent statistical procedures will often be made. It is assumed that each student in the class has taken an introductory course in Statistics and year of calculus. Any student not having this background should contact Dr Ewens – see contact information above - as soon as possible.

Textbooks

“Non-Parametric Statistical Methods”, by M. Hollander and D. A. Wolfe, second edition, (Wiley, 1999) is required, since the course will closely follow the material in this book. “The Analysis of Cross-classified Categorical Data”, by S. E. Fienberg second edition, (Springer, 2007) is recommended but not required, since only about one quarter of the course will cover material from this book, and material in the course from this book not discussed by Hollander and Wolfe will be taught in a self-contained way. References to these books (as HW and F) are given below against each topic covered in the course.

The course will begin with a review of various probability distributions and standard statistical inference procedures and principles. The topics covered in the course are listed below. The course will *not* be taught in conjunction with a computer package, but instead will focus on ideas and general concepts. However, students may use statistical packages to answer homework questions, if they wish.

TOPIC

HW or F

Reviews

Review – what is statistics?

The binomial and the normal distributions.

Review of hypothesis testing theory.

Introductory hypothesis testing examples.

Testing for the value of a binomial parameter p .

HW, 20-29

Normal approximation to the binomial. The central limit theorem.

One-sample location problems: non-parametric approaches.

HW, 35-87

More hypothesis testing theory – power and robustness.

Asymptotic relative efficiency calculations.

HW, 104-105

Review of estimation theory.

Unbiasedness, consistency.

Estimation of a binomial parameter.

HW, 29-34

Other classical estimation procedures.

Walsh averages. Hodges-Lehmann estimation.

HW, 52-55, 79-83, 125-132, 139-140

Review of order statistics theory.

Two-sample problems.

Testing for the equality of two binomial parameters.

F, 1 – 23, HW 458-484

Two-sample location problems.

HW, 106-125

Permutation tests.

More estimation procedures.

HW, 125-135

Asymptotic relative efficiency calculations.

HW, 139-140

More complicated examples.

Tests related to dispersion parameters.

HW, 141-158

Kolmogorov-Smirnov procedures.

HW, 179-186

Non-parametric alternatives to ANOVA.

HW, 189-351

Testing for independence of two random variables.

HW, 363-383, 394-408

Hypothesis testing in regression.

HW 415-457

Log linear models.

F, 13-15

Three- and many-way contingency tables.

F, 27-51

Choice of models.

F, 56-68

Logit models.

F, 97-102

Examinations. There will be a mid-term during normal class hours in the week after the mid-semester break, and a final exam in the normal end-of-semester exam period. The final exam will have more weight than the mid-term exam, and although it will tend to focus on the material covered in the second half of the semester, will cover the material in the entire semester. Further details about the exams will be announced in class.

Homework. Homework problems will be handed out each week. Homework performance will count towards the final grade. Homework material as well as the topics covered in the course will be discussed in class, and participation in these discussions will also count towards the final grade.

Office hours. I am available at *all times* to discuss any aspects of this course. Contact information – email is best – is given above. *Never* hesitate to contact me or come to see me at any time.

