STAT 432/512 / BSTA 621 - Spring 2009

Dr W J Ewens
wewens@sas.upenn.edu
215 898 7109
324 Leidy Labs (corner 38th. St and Hamilton Walk)

The purpose of this course is to discuss theoretical aspects of estimation theory and hypothesis testing procedures, together with some of their more important applications. The broad nature of the material to be covered is indicated below. It is assumed that each student in the class has taken probability theory to the level discussed in STAT 430, as well as a good introductory course in Statistics, a year of calculus and a good introduction to matrix theory. Any student not having this background should contact the course lecturer Dr Ewens - see contact information above - as soon as possible.

Background knowledge. The following list gives some standard results of Statistics that it is assumed that a student in the class already knows:-

If X has a normal distribution with mean 0, variance 1, then X^2 has a chi-square distribution with one degree of freedom.

If $X_1, X_2, \ldots, X_n$ are independent random variables, each having a chi-square distribution with $\upsilon_1, \upsilon_2, \ldots., \upsilon_n$ degrees of freedom respectively, then $\sum_{i=1}^{n} X_i$ has a chi-square distribution with $\sum_{i=1}^{n} \upsilon_i$ degrees of freedom.

If X has a normal distribution with mean μ, variance σ^2, then $Z = (X-\mu)/\sigma$ has a normal distribution with mean 0, variance 1.

If $X_1, X_2, \ldots, X_n$ are independent random variables, each having a normal distribution with mean μ, variance σ^2, then $\sum_{i=1}^{n} (X_i - \bar{X})^2 / \sigma^2$ has a chi-square distribution with n-1 degrees of freedom.

If $X_1, X_2, \ldots, X_n$ are independent random variables, each having a normal distribution with mean μ, variance σ^2, and if t is defined by $T = (\bar{X} - \mu)\sqrt{n}/S$, where S is defined by $S^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2 / (n-1)$, then T has the "t" distribution with n-1 degrees of freedom. (*)

If $X_{11}, X_{12}, \ldots, X_{1n}$ are independent random variables, each having a normal distribution with mean μ_1, variance σ^2, and if $X_{21}, X_{22}, \ldots, X_{2m}$ are independent random variables, each having a normal distribution with mean μ_2, variance σ^2, and each X_{1i} is independent of each X_{2j}, then F, defined by $F = S_1^2/S_2^2$, where $S_1^2 = \sum_{i=1}^{n} (X_{1i} - \overline{X}_1)^2 /(n-1)$ and $S_2^2 = \sum_{i=1}^{m} (X_{2i} - \overline{X}_2)^2 /(m-1)$, has a F distribution with (n-1,m-1) degrees of freedom. (*)

(*) These results will be proved in class.

Textbook

The course is based on D D Wackerley, W Mendenhall and R L Scheaffer "Mathematical Statistics with Applications" , seventh edition, (Thomson Books/Cole 2008), ISBN 978-0-495-11081-1. However the lectures are self-contained, and although it will be very useful for you have this book, it is not required. A slightly higher level book is "Introduction to Mathematical Statistics", by R.V. Hogg, J. W. McKean and A. T. Craig, (HMC) sixth edition, (Prentice Hall, 2005). Some class material will be at the higher level as given in the HMC book, but this book is *certainly* not required.

References to the book by Wackerley et al. (as WMS) are given below against each topic covered in the course. While it will be assumed that the material covered in WMS chapters 1-6 is known, some of it will nevertheless be reviewed in the first few lectures of the course.

The main topics covered will include estimation theory, including in particular the desirable properties of estimators and how the properties can be achieved, as well as the concepts of sufficient statistics and maximum likelihood estimation, confidence intervals, hypothesis testing theory and the various methods of hypothesis testing, sequential hypothesis testing and distribution-free methods of hypothesis testing, and tests involving linear models.

TOPIC	WMS
Review of basic material	Chapters 1 – 5
Functions of random variables	6.2
Transformation theory	6.4, 6.6
Order statistics	6.7
Results connected with the central limit theorem	7
Various elementary of estimators	8.1-8.4
More advanced properties of estimators. The Cramer-Rao inequality	9.1-9.3
The concepts of sufficiency. The Rao-Blackwell theorem	9.4
The method of moments	9.6
The likelihood function and maximum likelihood estimation	9.7 – 9.8

Concepts of hypothesis testing. Elementary examples	10.1 - 10.9
Neyman-Pearson theory and likelihood ratio tests. Applications.	10.10 - 10.11
Sequential tests.	-------
Distribution –free (non-parametric) tests.	15.1 -15.6
ANOVA	13.1-13.9
Hypothesis testing for linear models.	11.1-11.11

Examinations. There will be a mid-term during the 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 a*t 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.