Syllabus1112013

STAT 111 – Fall 2013

SYLLABUS AND OTHER INFORMATION

Lecturer: Dr W J Ewens

wewens@sas.upenn.edu

Office: Room 324 Leidy Labs (the "Zoology building"), at the corner of 38th Street and Hamilton Walk. Office hours are 24/7. Make an appointment via the email address above. You will have to swipe your Penn card in the elevator to get to the third floor.

Sections The course is divided into two sections, Section 1 and Section 2. The lectures in the two sections are identical. When you register for this course you also register for the section you plan to attend.

Lectures All lectures are given in Steinberg-Dietrich Hall, room 350. Section 1 lectures are MW 11 am - 12 noon, Section 2 lectures MW 2 pm - 3 pm.

Fall mid-term break and Thanksgiving arrangements The Fall mid-term break is Th-Fri October 10-11. Thus there will be no recitation classes on October 11. Also, for the week Nov 25-29, Tuesday and Wednesday classes will it not be held, and instead lecture and recitation classes normally scheduled for Thursday and Friday of that week will be moved respectively to Tuesday Nov 26 and Wednesday Nov 27. Thus there will be STAT 111 recitation classes but no lecture on Wednesday Nov 27.

Recitation classes Except for October 11 and November 29 (see above), recitation classes are held on Fridays, starting Friday September 6. (The recitation class for the week Nov 25-29 will be held on Wednesday Nov 27 - again, see above.) These classes are compulsory - homeworks will be handed out to you in recitation classes and your answers will be due in the recitation class one week later. When you register for this course you also register for the recitation class you plan to attend. Students attending Section 1 lectures must register for either recitation section 201, 202, 203 or 204. Students attending Section 2 lectures must register for either recitation section 205, 206, 207 or 208.

Recitation classes for Section 1 students are held in Huntsman Hall, room G86. The times for these are as follows: recitation section 201 is at 11 am, recitation section 202 is at 12 noon, recitation section 203 is at 1 pm and recitation section 204 is at 2 pm.

Recitation classes for Section 2 students are held in Huntsman Hall, room G88. The times for these are as follows: recitation recitation section 205 is at 11 am, recitation section 206 is at 12 noon, recitation section 279 is at 1 pm and recitation section 208 is at 2 pm.

Homeworks See above - homeworks will be handed out on Fridays in recitation classes, and your answers will be due in the following recitation class one week later. (A special arrangement will be made for the week of Thanksgiving.) Apart from indicating your name on any homework,with your family name in CAPS, also indicate the recitation section that you will come to in the following week (when your graded homework will be given back to you).

Homework 1 will be handed out in recitation classes on Friday September 6. it will also be posted on canvas (as will all homeworks).

Exams There will be one mid-term exam, to be given **6 - 8 pm Monday October 21**. The location of this exam is still being finalized, and you will be given the relevant information when it becomes available. The final exam is **17 December, 3 - 5 pm**. More details as to the location of this exam will be given later, when it becomes available. The timing of both exams is set by the university and cannot be changed.

Assessment The assessment in this course is by homeworks (10%), the mid-term exam (30%) and the final exam (60%).

Textbook The suggested textbook for the course is Downing and Clark, "E-Z Statistics", Barron, 2009, ISBN 13: 978-0-7641-3978-9. However this book is **not** required, since it is used only as a general guide to the course material and the course is not firmly based on it.

Web resource The web resource in this course is "Canvas". If you do not already have a Wharton canvas account (for example if you are an SAS student), you can access Canvas using a "Wharton class" account. Instructions for obtaining such an account can be found at **http://supportcenteronline.com./link/portal/632/655/Article/5363/Canvas - Accounts -Passwords** under the "Enrolled Non-Wharton Students Who Need Accounts" section. This should be done as soon as possible.

If you have any questions about using web resources please email Peichao Peng at ppeichao@wharton.upenn.edu or else contact the Wharton Computing Student support office at **https://spike.wharton.upenn.edu/support** or else call them at 215 898 8600.

JMP The course will in part be given in association with use of the statistical package JMP 10. You should either buy and install this package on your computer or (better) use the Wharton computers that have it installed. You will not be able to use these computers until you have created a Wharton account (see instructions above under "Web resource). It might help you to buy a copy of the fifth edition of the JMP manual "JMP Start Statistics", SAS. We won't use JMP much, but it is useful to get used to the idea of using statistical packages.

If you have any questions about JMP, please email Yichao Lu at luyichao1123@gmail.com

Course description The content of this course falls into two broad categories, namely probability theory and statistics. The reason why we discuss probability theory will be given in the first lecture. A more detailed list of the topics covered within these two categories is given in the syllabus on the next two pages. **References to corresponding material in the textbook by Downing and Clark for these topics are given in parentheses (....), as for example (DC107-118).** Note that some material in the course is not covered by Downing and Clark, that sometimes the approach taken in class to some topics differs from that in Downing and Clark, and that sometimes material given in class contradicts (incorrect) material in Downing and Clark. Therefore the references to Downing and Clark are only a general guide to the material that will be covered in class.

SYLLABUS

1. Probability theory

1.1 The relation between probability theory and Statistics. Deductions (implications) and inductions (inferences).
1.2 Events and their probabilities. (DC 32–34)
1.3 Unions, intersections and complements of events and their probabilities. (DC 34–40).
1.4 Independence of events. (DC 38).
1.5 Conditional probabilities. (DC 75–86).
1.6 Discrete random variables and their probability distributions. Examples.(DC 87–106).
1.7 The concept of the mean of a discrete random variable. (DC 93–95).
1.8 The variance and standard deviation of a discrete random variable. (DC 95–99).
1.9 Many *iid* random variables. The mean and variance of a sum and an average of many random variables. Examples in the "fair die" case. The mean and the variance of the difference between two random variables. (DC 180).
1.10 *Into the unknown.* The thumbtack example. More on the concept of a parameter and the binomial parameter θ. The binomial distribution. (DC 107–118).
1.11 The mean and variance of a binomial random variable. (DC 116).
1.12 The proportion of successes in the binomial context. The mean, variance and standard deviation of this proportion.
1.13 The " two - standard - deviation rule". (DC 149).
1.14 Continuous random variables and their density functions. (DC 131–140).
1.15 The mean and variance of a continuous random variable. The mean as a "balance point". (DC 138–140).
1.16 The normal distribution. The use of normal distribution charts. (DC 143–155).
1.17 The "Z" transform and the standard normal distribution. The use of normal distribution charts. (DC 147–151).
1.18 The central limit theorem and its uses. (DC 192–198).
1.19 Approximating the binomial distribution by the normal distribution. (DC 193).

2. Statistics

2.1 Estimation

2.1.2 Introduction.
2.1.2 Estimation of the binomial parameter θ and the precision of this estimate. The "margin of error". (DC 265–268).
2.1.3 The average ($\bar{y}$) as an estimate of a mean μ. The "unfair die" case again. (DC 205–207).
2.1.4 s^2 as an estimate of a variance σ^2. (DC 209–215). s^2/n as an estimate of the variance of $\bar{Y}$ $(= \sigma^2/n)$.
2.1.5 An approximate 95% confidence interval for the mean. (DC 213–215).
2.1.6 Regression. (DC 289–300).

2.2 Hypothesis testing (DC 227–245)

2.2.1 General principles of hypothesis testing. (DC 13–17, 227–230).
2.2.2 The concept of the p-value. (DC 230).
2.2.3 The one-sample t test. (DC 164–166, 231–235, 398).
2.2.4 Testing hypotheses about the binomial parameter θ. (DC 235–236).
2.2.5 The 2×2 table used for testing for association.
2.2.6 Further examples of hypothesis testing. (DC 236–240, 243–245, 277–284, 289–304)
2.2.7 Practical considerations. (DC 271–274)