Syllabus111 2015

STAT 111 – Fall 2015

GENERAL INFORMATION AND SYLLABUS

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.

Dr Ewens' office hours: Dr Ewens' office hours are open. Do not hesitate to make an appointment at any time via the email address above. NOTE: do NOT send emails via the "reply" or "reply all" mechanism to a message which was sent to the entire class. Doing so might send your message to the entire class. Instead, initiate a new separate email message when contacting Dr Ewens in reply to a message sent to the entire class.

As noted above, Dr Ewens' office is in Leidy Labs. Leidy Labs has a security system which means that you will have to swipe your Penn card in the elevator/staircase to get to the third floor to reach Dr Ewens' office.

Sections Lectures are given through three sections, Sections 1, 2 and 3. On any given day the lectures in the three sections are identical, so that if on some day you cannot go to your normally scheduled section you can go to another section meeting on that same day. When you register for this course you also register for the section you will attend.

Lectures Section 1 meets Tues-Thurs at 11:00 am - 11:50 am, Section 2 meets Tues-Thurs at 2 pm- 2:50 pm, and Section 3 meets Tues-Thurs 3:00 pm - 3:50 pm. All lectures are given in Steinberg-Dietrich Hall room 351.

Announcements Important announcements will often be given in class. If for some reason you miss a class it is up to you to find out from a friend if any important announcements were made in the class that you missed. So far as is possible these announcements will also be posted on "Canvas". (For more on "Canvas" see **Web Resource: "Canvas"** below.)

Recitation classes Except for October 9, September 25 and November 27 (see below), recitation classes are held on Fridays. The first recitation class will be on **Friday September 4**. The recitation class for the week November 23-27 will be held on **Wednesday November 25** - see below. Homeworks will be handed out to you in recitation classes and your answers will be due in the recitation class one week after each homework is handed out. (Homeworks will also be posted on Canvas - see below). 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. Students attending Section 3 lectures must register for either recitation section 209, 210, 211 or 212.

Recitation classes are given by teaching assistants (TAs) and will be held in ground-floor rooms in Huntsman Hall. The times and places for these are as follows:

Recitation class 201 is at 11 am, room G86
Recitation class 202 is at 11 am, room G90
Recitation class 203 is at 12 noon, room G86
Recitation class 204 is at 1 pm, room G86
Recitation class 205 is at 12 noon, room G90
Recitation class 206 is at 12 noon, room G88
Recitation class 207 is at 1 pm, room G88
Recitation class 208 is at 2 pm, room G88
Recitation class 209 is at 11 am, room G88
Recitation class 210 is at 1 pm, room G90
Recitation class 211 is at 2 pm, room G86
Recitation class 212 is at 2 pm, room G90

The TAs for these classes are as follows:

Recitation classes 201, 203 and 204: Xin Lu Tan (xtan@wharton.upenn.edu)
Recitation classes 202, 210 and 211: Alanna Cruz-Benzedu (alannacr@sas.upenn.edu)
Recitation classes 205, 207 and 208: Ville Satopaa (satopaa@wharton.upenn.edu)
Recitation classes 206, 209 and 212: Linjun Zhang (linjunz@wharton.upenn.edu)

TAs are there to help you. You can contact them at the above email addresses.

Fall mid-term break and Thanksgiving arrangements The Fall mid-term break is Th-Fri October 8-9. Thus there will be no recitation classes on October 9. Also, for the week Nov 23-27, that is the week that includes the Thanksgiving break, normal Tuesday and Wednesday activities will *not* be held, and instead any activity normally scheduled for Thursday Nov 26 and Friday of Nov 27 will be moved respectively to Tuesday Nov 24 and Wednesday Nov 25. In practice this does not affect STAT 111 lectures: lectures will be held on Tuesday November 24. However it does mean that the STAT 111 recitation classes for that week will be held on **Wednesday Nov 25**.

Pope's visit Because of the Pope's visit the university will be closed on Friday September 25. Thus there will be no STAT 111 recitation classes that day. Plans for handing in homeworks which would otherwise be due on that day will be announced in class.

Homeworks See above - homeworks will normally be handed out on Fridays in recitation classes, and answers should be handed in at recitation class one week later. (Special arrangements will be made later for the week of the mid-term break and the week of Thanksgiving.) When handing in any homework, write your first name first and in full (i.e. not just initials), and then your family name written last and in CAPS, for example Mary SMITH. (Asian students: please also follow this convention.) Also, indicate your recitation section clearly on your homework.

Homework 1 will be handed out *in class* on August 27 and should be handed in during recitation classes on Friday September 4.

Homework and exam point scores are posted on "Canvas" - see more on Canvas below. You should check regularly that all your homework scores are entered corectly on Canvas.

Exams There will be one mid-term exam and one final exam. The date and time for the mid-term exam are not yet finalized, but it is almost certain that the exam will be given **6 - 8 pm Monday October 19**. Thus keep this date and time clear of other activities. The location of this exam is still being finalized, and you will be told the location several days before the exam. The final exam will be given **Thursday 17 December, 3 - 5 pm**. More information 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. If one or both of these times is **impossible** for you, contact Dr Ewens immediately.

Assessment The assessment in this course is by homeworks (10%), the mid-term exam (40%) and the final exam (50%). Some of the questions on the mid-term and final exams will be questions previously set in homeworks. Thus the homeworks in effect carry a higher percentage value of the overall score than is suggested by the above.

Textbook There is no required textbook for this course. Printed notes will be available at no cost to you, and these can serve as a textbook. If however you do want to buy a textbook you should get Downing and Clark, "E-Z Statistics", Barron, 2009, ISBN 13: 978-0-7641-3978-9. This book should be available in the Penn bookstore. 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. (It also contains some errors.)

Calculator You will need a hand calculator for this course. All that is needed is a calculator doing the elementary operations of addition, subtraction, multiplication and division as well as taking square roots. You do not need a graphing calculator or one doing operations like taking logarithms.

Web resource: "Canvas" The web resource in this course is "Canvas". This is available to all Penn students at **https://canvas.upenn.edu** You will need your pennkey authentification (username and pasword) to use Canvas.

For a copy of the class notes, and also for the syllabus/information document and for weekly homeworks, use the "Files" link in canvas. For point score information, use the "Grades" link.

For questions about using Canvas you can contact the Wharton Computing Student support office at 215 898 8600 or at **https://spike.wharton.upenn.edu/support** Also you can contact Ville Satopaa at satopaa@wharton.upenn.edu for problems concerning Canvas.

JMP The course will in part be given in association with use of the statistical package JMP. You should either buy and install this package on your computer or (much better, since buying JMP is expensive) use the (free) Wharton computers that have it installed.

Note for non-Wharton students. If you do not have a Wharton computing account you will need to establish one to be able to access Wharton computers. To create an account, go to https://app.wharton.upenn.edu/accounts/ .

Alternatively, Penn students can get a JMP license through e-academy at
http://www.onthehub.com/jmp/ for $30 for a 6 month license or $50 for a year license.

If you have any questions about JMP, contact Ville Satopaa at satopaa@wharton.upenn.edu

Disabilities If you are registered through the Weingarten Center for special arrangements for exams etc., please contact Dr Ewens as soon as possible.

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 below. **References to the pages for the corresponding material in the textbook by Downing and Clark for these topics are given in parentheses, 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

INTRODUCTION

1 Statistics and probability theory

1.1 What is Statistics?
1.2 The relation between probability theory and Statistics
1.3 An example

PROBABILITY THEORY

2. Events (DC 32–34)
2.1 What are events?
2.2 Notation
2.3 Unions, intersections and complements of events (DC 34–40).

3 Probabilities of events (DC 35–40)
3.1 Probabilities of derived events
3.2 Mutually exclusive events
3.3 Independence of events. (DC 79-80).
3.4 Examples of probability calculations involving unions and intersections
3.5 Conditional probabilities of events (DC 75–86).
3.6 An unfair die

4 Probability: one discrete random variable

4.1 Random variables and data
4.2 Definition: one discrete random variable (DC 87–92)
4.3 The probability distribution of a discrete random variable (DC 87–106).
4.4 Parameters
4.5 The binomial distribution (DC 107-118)
4.6 The mean of a discrete random variable (DC 93–95).
4.7 The variance of a discrete random variable (DC 95–99).

5 Many random variables

5.1 Introduction
5.2 Notation
5.3 Independently and identically distributed random variables
5.4 The mean and variance of a sum and of an average

5.5 Two generalizations
5.6 The proportion of successes in n binomial trials
5.7 The standard deviation and the standard error
5.8 Means and averages

6 Continuous random variables (DC 131–140).

6.1 Definition
6.2 The mean and variance of a continuous random variable (DC 138–140).
6.3 The normal distribution (DC 143–155).
6.4 The standardization procedure (DC 147–151).
6.5 Numbers that you will see often (DC 230)
6.6 Sums, averages and differences of independent normal random variables
6.7 The Central Limit Theorem (DC 192-198)
6.8 The normal distribution and the binomial distribution (DC 193)
6.9 The chi-square distribution (DC 161–164).

STATISTICS

7 Introduction

8 Estimation (of a parameter)

8.1 Introduction
8.2 General principles of the estimation of a parameter
8.3 Estimation of the binomial parameter θ (DC 265–268).
8.4 Estimation of a mean (μ) (DC 205–207).
8.5 The 95% confidence interval for a mean μ (DC 216–217)
8.6 Estimation of a variance
8.7 Notes on the above example
8.8 Estimating the difference between two binomial parameters
8.9 Estimating the difference between two means
8.10 Regression. (DC 289–300).

9 Hypothesis testing (DC 227–245)

9.1 Introduction (DC 13–15, 231–236)
9.2 Two approaches to hypothesis testing
. 9.2.1 Both approaches, Step 1
. 9.2.2 Both approaches, Step 2
. 9.2.3 Both approaches, Step 3

. 9.2.4 Approach 1, Step 4, the medicine example
. 9.2.5 Approach 1, Step 5, the medicine example
. 9.2.6 Approach 1, Step 4, the coin example
. 9.2.7 Approach 1, Step 5, the coin example
. 9.2.8 Approach 2, Step 4, the medicine and the coin examples
. 9.2.9 Approach 2, Step 5, the medicine example
. 9.2.10 Approach 2, Step 5, the coin example
9.3 The hypothesis testing procedure and the concepts of deduction and induction
9.4 Tests for the equality of two binomial parameters (DC 240–242)
9.5 Tables bigger than two-by-two (DC 243–245)
9.6 Another use of chi-square: testing for a specified probability distribution (DC 246–247)

10 Tests on means

10.1 The one-sample t test (DC 232–233)
10.2 The two-sample t test (DC 236–239)
10.3 The paired two-sample t test (DC 239–240)
10.4 t tests in regression (DC 299)
10.5 Non-parametric (= distribution-free) tests (DC 277)
. 10.5.1 Introduction
. 10.5.2 The non-parametric alternative to the one-sample t test: the Wilcoxon
. signed-rank test (DC 282–284)
. 10.5.3 The non-parametric alternative to the two-sample t test: the Wilcoxon
. rank-sum test (DC 280–281)