

# **BIEB 100: Biostatistics Lecture & Lab**

**Course Information (updated 2 January 2012)**

**Winter Quarter 2012**

**Lecture: Tues/Thurs 9:30-10:50, Ledden Auditorium**

**Note that the online html version is always the current version. Changes will be there first.  
This pdf version may lag behind.**

**Lab A01: Thursday 12:00-12:50PM, ERCA 117**

**Lab A02: Thursday 1:00-1:50PM, ERCA 117**

**Lab A03: Thursday 2:00-2:50PM, ERCA 117**

**Lab A04: Friday 10:00-10:50AM, ERCA 117**

**Lab A05: Friday 11:00-11:50AM, ERCA 117**

**Lab A06: Friday 1:00-1:50PM, WLH 2114 – Laptops only**

**Lab A07: Friday 2:00-2:50PM, WLH 2114 – Laptops only**

**Course website:**

**<http://www.biology.ucsd.edu/labs/rifkin/courses/bieb100/win2012/bieb100winter2012.html>**

Please note: information in this syllabus is subject to change. Any schedule changes will be announced in class and posted on the website. Make sure to frequently check the website to keep updated about the weekly readings, assignments, and exam schedule. All updates will be noted in the [changeLog](#) section.

## Philosophy

Educational research has shown that student engagement is the most important component of learning. Learning involves wrestling with new concepts, practicing new techniques, and rethinking old assumptions. My role as the professor and the role of your TAs is to facilitate your learning by providing targeted materials for you to use, structuring your exposure to a logical sequence of topics, clarifying misconceptions, and helping you work through difficult material. You should see this course as a collaboration. How much you get out of it and how well you do in it will depend upon how much you engage with the material and how much discipline you apply towards mastering new concepts and deploying them.

Students often come to an introductory statistics course expecting to memorize a series of recipes and formulae to apply in well-defined situations. Statistics is not a cookbook for using formulae but rather a process for getting reliable answers to interesting questions – it embodies the scientific method. Statistical formulae are sometimes (but not always, as we will see) the last step in this process and make no sense without the preceding components. Moreover, the wide variety of statistical tests out there to handle disparate kinds of data are all based on a common logic. Understanding this logic transforms statistics from a hodge-podge of individual, unrelated, arcane formulae into a conceptually coherent and flexible framework for answering all kinds of questions and for thinking about the world.

Every day you are exposed to numbers and data and conclusions drawn from them (or despite them). They are rarely as clean and clear as a textbook problem and yet they form the basis for crucial public and private decisions, political and social debate, and scientific discoveries. Statistics and the scientific method is a way of evaluating these numbers and drawing conclusions from them in a reliable and justifiable way.

This course is designed to help you develop competence in three related areas.

The first is *statistical literacy*. This involves:

- 1) awareness of data, how they are used and misused, and the consequences of each
- 2) understanding basic statistical concepts
- 3) knowing the basics of how to collect data and summarize it
- 4) knowing how to relate statistical results back to the context of the original question
- 5) being able to communicate the results to others

The second is *statistical reasoning*. This involves:

- 1) applying statistical ideas
- 2) making sense of statistical information
- 3) interpreting graphical displays and numerical summaries
- 4) checking whether intuitive conclusions from data really hold up

The third is *statistical thinking*. This involves the following habits:

- 1) consider how to collect meaningful and relevant data to answer a particular question
- 2) keep in mind the statistical process as a whole and how the different pieces fit together

- 3) be skeptical, particularly about how data is obtained
- 4) think about the variables involved – are they the right ones for the question? how will they behave? are there other variables that might be important?
- 5) always keep the context in mind. How does the data relate back to the original question?
- 6) understand the relevance of statistics, especially when you want to make a causal claim or say something is unusual.
- 7) think beyond the textbook.

My hope is that you will come out of this course with a healthy and statistically grounded skepticism about the numbers and data that pervade our society and that you will know what questions need to be answered about them before you will accept any conclusions based on them. We will get to that point by carefully studying how data is produced, what kinds of data there are, how to properly assess probabilities, how to test hypotheses, and how to estimate effects.

### **Important dates**

31 January 2012	Exam 1
1 March 2012	Exam 2
20 March 2012	Final exam. 8:00-11:00AM

### **Grading**

This course will be graded on mastery of the material and not on a curve. The bulk of the grade will come from the three exams. These exams will consist of multiple choice questions, short answer questions, and potentially some longer problems involving a series of steps. These questions will be designed based on the cognitive domain of Bloom's taxonomy ([http://en.wikipedia.org/wiki/Bloom%27s\\_Taxonomy#Cognitive](http://en.wikipedia.org/wiki/Bloom%27s_Taxonomy#Cognitive)). Exams will have a mix of questions specifically testing higher and lower order mastery of the material. A **C** on an exam would reflect competence in knowledge and comprehension. A **B** would reflect an ability to apply knowledge to a new situation. An **A** would reflect the ability to interpret evidence, make good statistical judgments, and demonstrate statistical thinking.

The grading breakdown for components of the course is:

Exam 1	20%
Exam 2	20%
Final	30%
Clicker participation	8%
Online reading homework	8%
Lab participation and homework	14%

A (80-100%) B (60-80%) C (40-60%)

+/- will be assigned depending upon the distributions of scores within each of these letter grades.

### **Textbook**

We will use an online textbook for this course. This online system has interactive exercises that are integrated with the text. These will allow you to continually check your understanding of the material with the goal of making your studying more efficient and productive. The system also gives me feedback in two ways. First, I can look at the results of these exercises to see whether there are particular concepts and topics that are confusing. Second, there are boxes at the end of each section where you can respond directly about the material - which parts were straightforward and which parts were confusing. I can then use this feedback to focus the lecture time on clarifying confusing topics.

To sign up for the online textbook, go to:

<http://oli.web.cmu.edu/openlearning/forstudents/academiccourses>

The course key is **BIEB100Win12**. Please use the password RAFisher.

There is a \$25 fee. When you sign up for the course it will tell you how to pay.

Interactive exercises are sprinkled throughout the textbook. These will be required homework unless otherwise noted. Because I will use the results of these exercises to shape lectures, they will be due at 7AM the day before the lecture. For example, you will need to complete the reading and homework for a Thursday lecture by Wednesday 7AM. Late homework will not be accepted. As with all classes, you will find that the lectures will be far more useful to you and that you will be able to ask better questions if you have already thought about the material.

### **Contact Information**

*Professor*

**Scott Rifkin**

Ecology, Behavior, and Evolution

Division of Biology

Muir 2105

sarifkin@ucsd.edu

*Teaching Assistants*

**Lis Cordner**

Biological Sciences

ecordner@ucsd.edu

**Gunnar Poplawski**

Biological Sciences

ghpoplawski@ucsd.edu

**Phillip Kyriakakis**

Biological Sciences

pkyriaka@ucsd.edu

**Ping Gong**

General Biology

pgong@ucsd.edu

**Office Hours** Please see course website site for details.

### **Software**

Homework assignments, the online textbook, and labs will use both web applets and the statistical computer language R.

You can download R (version 2.14.0) from here: <http://www.r-project.org/>

Please also download and install RStudio from here: <http://rstudio.org>.

RStudio is simply an environment within which you can run R. It has some nice features in that it arranges your windows for plotting, exploring datasets, finding help, running commands, etc., all in one place. To see the difference, you can run R alone (just click on the Application (Mac) or executable (Windows)) and compare it to RStudio. Again, the main point of RStudio is just to collect all the R windows together so they are a little easier to navigate. It also has a few menus that make it easier to do things like import data.

During the first two weeks of the course, you will be expected to learn to do basic operations in R. This will be through a set of online tutorials, lab practice, and homework exercises. Putting in the time upfront to master these skills will make the rest of the course go much more smoothly.

If you do not have a personal computer, you can install and run R from a USB stick on a computer in one of the computer clusters: <http://personal.bgsu.edu/~mrizzo/Rmisc/usbR.htm>. You can run RStudio using the zip/Tarball distribution from the RStudio website. The computers in the ERCA 117 computer lab will have R loaded on them.

We will also use web applets (often written in java) throughout the course. If your web browser cannot run them, you may need to install java, although this should come standard on modern computers. Some of the applets will be written in the Wolfram Computable Document Format. To use these, you will need to install the free CDF Player from: <http://www.wolfram.com/cdf-player/>

In summary, you will need: R, RStudio, java, CDF Player.

### **Homework**

There will be a few different types of homework. The online textbook includes both short comprehension questions and longer exercises. These will be due at 7AM on the day before class. You should expect the labs to have lab reports or exercises which will be due in the following lab session. These may include short experiments to do at home or close readings of scientific papers and the popular press. For some topics in the course there will be additional problems to work through to solidify understanding, particularly of the mechanics of doing tests. Some assignments will lend themselves to online submission through a web form; others will require paper.

## **Labs**

**Labs are not optional.** Lab attendance counts toward your final grade. Missing more than 3 labs results in a FAIL for the whole course. Labs will be a combination of problem solving, computer assignments, and experiments. Labs A01-A05 will take place in a computer room in ERCA; students who sign up for Labs A06 and A07 must have their own laptops. The teaching assistants will be responsible for the labs.

## **Cheating**

Students are expected to do their own work, as outlined in the [UCSD Policy on Academic Integrity](#). Cheating will not be tolerated, and any student caught engaging in academic dishonesty will fail the course. He or she may also be suspended from UCSD. All exams will be closed-book and closed-notes, so all personal materials must be stowed under your seat. Only exams written in non-erasable pen will be considered for regrades.

## **Computers and phones in class and sections**

Please do not use your computers or phones or other distraction-enabled devices in class. Laptops are not a good tool for taking notes in class. They do not have the versatility necessary for capturing text, sketches, arrows, etc. that are important for note-taking in statistics. Bring paper and something to write with to take notes in class. We will use computers for some exercises in sections, but you are expected to be courteous to your fellow students and TA and use them only for the intended purposes. Students using computers for purposes other than those related to the section will be asked to leave and will not receive credit for that day. Phones are only to be used for emergencies and outside the classrooms.

## **i>clicker2**

We will use the i>clicker2 personal response system in class. The i>clicker2 (as opposed to the i>clicker) allows for numeric input which is quite useful for a statistics class. If you do not already have one, you can pick the device up from the bookstore, but please get the right version. Each clicker has a unique serial number on the back of the remote. Place a piece of scotch tape over that bar code and ID to preserve it. In order to receive credit for your votes, you will need to register your clicker online by the end of the second week of class. You must have responded to at least one question in order to complete the registration properly. Registration is at <http://www.iclicker.com/registration>. Complete the fields with your first name, last name, student ID, and remote ID. Your student ID should be the 9 digit PID on your ID card starting with A. The clickers serve a few purposes. First, they are great tools for generating data in class. Second, they give immediate feedback on whether a concept was clear or not. Third, they will be used to measure attendance and participation.

There will be three types of clicker questions. If there is a blue square in the corner of the screen, you will get credit only if you get the right answer. These questions will generally be reviews of the reading material and homework. Each one will be worth a small fraction of the class participation grade. If there is a red triangle, then you will get credit for clicking in regardless of your answer. Please do your best. If you click in for 75% of these over the quarter you will receive full credit for this portion of the class participation grade. Sometimes I will ask

only a portion of the class to click in. These questions will be marked with a green 'P'. If you are not in the portion that is asked to click in, you will not be penalized

### Feedback

Learning is a collaborative endeavor and so your feedback is crucial. There are five main avenues for feedback. Your answers to the online textbook homework (including the dedicated space to ask questions) will help me to focus the lectures on clarifying the concepts that are most confusing. Clicker questions in class provide immediate feedback to me and to you about whether a concept is clear or needs to be explored further. There will be a [form](#) on the website that will ask you for feedback on the lectures and have space for you to ask outstanding questions (anonymously or not as you prefer). The discussion board on TED will be monitored by the instructors. And lastly, all of the instructors will hold office hours.

### Data generating supplies

Class demonstrations will occasionally require tools. Please go to <http://www.random.org/integers/?mode=advanced> and print out a set of 1000 random digits from 0 to 9. (For easy printing choose 15 columns in Part 1, “Bare bones text” in Part 3, and “generate your own” in Part 4). Please also use this website to print out a set of 100 random numbers from 1 to 200. Please bring these sheets and a coin to every class. If we will need a different tool for a demonstration or for lab (e.g. dice, deck of cards, ruler, etc.) it will be posted on the website and announced in class.

### Regrades

If an error has been made in the grading of your homework or exam, you may submit it to the instructor within *one week* of distribution for a regrade. Do not go to a TA for an exam regrade.

**Your exam must first be submitted to the instructor with a written description of the grading error.** Regrade requests will not be processed without a written description of the grading error. **No regrades will be given for homework or exams written in non-permanent ink or pencil.**

### Missed exams

There are no makeup exams. If you must miss an exam, you will be required to provide official documentation of an unavoidable emergency (e.g. serious illness, etc.). Without such documentation, you will receive a failing grade for that exam. If you miss two exams or if you are unable to provide official documentation of an unavoidable emergency, you will receive an incomplete for the course. If you miss an exam, the percentage of your grade from the other exams will be increased accordingly. If you know beforehand that you cannot be present in the classroom for an exam and (1) can present a valid excuse (such as a note from your coach for a sporting event) at least one week in advance of the exam and (2) are receiving a passing grade in the course as of one week before the exam, then a sealed copy of the exam will be given to the coach, musical director, etc. who will administer the exam **at the class exam time**. If no UCSD official can be found to administer the exam, then it will count as a missed exam. If the exam cannot be taken at the class exam time, then it will count as a missed exam. **No exam can be taken before or after the scheduled exam date and time.**

**Web resources**

The course will be using several resources on the web. The class website (url at the beginning of this syllabus) is the hub for all class information. There will be links to the other resources from there such as R tutorials and applets for demonstrating particular topics. Relevant datasets will also be posted online. Grade information will be posted on the course TED website and we will use the discussion board on TED as a forum for you to discuss topics and ask questions. This will be monitored during reasonable hours by one of your TAs. The online textbook is discussed above. If you experience problems with TED, please contact: [iwdc@ucsd.edu](mailto:iwdc@ucsd.edu).