Test 1 Information

Time, Location, Coverage

Test 1 will be given in class on Tuesday, September 30. The test covers chapters 1, 2, and 4 (except section 4.5). You are responsible for material covered in the text, in the problem sets, and in class.

Format

Test questions will be designed to try to see how well you understand the material, not how well you can perform various procedures mindlessly.

A variety of question formats may be used. You may be required to compute numerical statistics; produce graphs by hand or explain how to get a computer to produce them (you will not need your laptop, however); or to analyze data or numerical or graphical summaries of data. Some items may be be be tested using "short answers" (a couple sentences to a paragraph), multiple choice, or true/false.

Instructions

Read through these prior to coming to the test and follow them when you take your test.

1. Always **show your work and explain your reasoning.** Answers without work or reasoning will not receive full credit.
   - Use mathematical **notation** (especially the equals sign) correctly.
   - Don’t be afraid to **use words** in your explanations.
   - **If you get an unreasonable answer, be sure to say so.** Give a brief explanation about how you know your answer is wrong (for example, “the mean I calculated is less than 10, but I can see from the graph that it should be at least 20”). Then go on to other problems and come back and try to fix the error if you have time at the end of the test period.
   - Even if you cannot do a problem completely, **show me what you do know.**

2. You may use your calculator during the exam, but **for each number you write on the exam, it must be clear where it came from.** For example, if you got .25 by multiplying .5 by .5, I want to see .5 · .5 = .25 on your paper (or words indicating the same).

3. Short answer questions will be graded based on **truth, accuracy, clarity, significance, and brevity.** In short, I’m looking for high quality answers. (Example: If you are asked to give an example of something, pick the best example you can think of, one that makes the issue especially clear.)

4. Test restrictions.
   - The test is **closed book** No notes are allowed.
   - Do not write in purple on the exam. (The exam will be graded in purple.)
Content

Here is a list of things you should be sure you know how to do. It is not intended to be an exhaustive list, but it is an important list.

You should be able to:

- Understand, use and explain the statistical vocabulary/terminology. Be sure to focus on important distinctions being made by terms like statistic vs. parameter, sample vs. population, categorical vs. quantitative, etc.

- Compute numerical summary statistics (mean, median, range, standard deviation, quartiles, 5-number summary, etc.) and know what they tell you – and don’t tell you – about a data set.

- Understand how to make and interpret graphical representations of data (histogram, density plot, boxplot, bar graph, scatter plot) and when each might be appropriate or inappropriate to use.

- Use R to compute numerical summaries, make plots, create randomization distributions, and compute p-values.

- Understand the issues involved in collecting good data and the design of studies, including the distinctions between observational studies and experiments, use of paired designs, blinding, etc.

- Use the 4-step process for conducting a hypothesis test, including
  - expressing null and alternative hypotheses
  - determining a p-value from a randomization distribution
  - expressing the logic of a p-value in words (in the context of a particular example).
  - how to get RStudio to generate a randomization distribution

You do not need to know how to use `fisher.test()` and `binom.test()`. We will talk about these and other ways to get p-values without simulations later.

Note that the test will be a sample from the possible topics; it is not possible to cover everything on the test.
Example Problems

The following problems come from old tests or quizzes I have given. They are provided here to give you some examples of types of questions I have asked in the past. The problems sets (including the extra problems) also provide a good source of problems to use when preparing for the exam.

1. Below are 5 boxplots and 5 histograms from 5 data sets. Match them up. (Boxplots and histograms were given.)

2. Each of the distributions below has the same scale. (Histograms were given.)
   a) Which has the largest standard deviation?
   b) Which has the smallest mean?

3. Compute the mean and standard deviation of the following small data set: 1, 2, 4, 5, 8. Do this “by hand” showing all of your work and using a calculator only to perform arithmetic operations like addition, subtraction, multiplication, division, square roots, etc. That is, do not use any statistical functions on your calculator or RStudio.

4. What is the difference between ____________ and ____________? Give an example of each that clearly illustrates the difference. (The blanks could be filled in with any of our “distinction pairs”.)

5. You are interested in knowing whether moderate exercise for one month reduces cholesterol.
   a) Outline a randomized comparative experiment that attempts to find out. Be brief, but answer the following:
      • What are your explanatory and response variables?
      • Identify which variables are categorical and which are quantitative.
      • What makes your study an experiment and not an observational study?
   b) What advantage does your study have over an observational study designed to answer the same question?
   c) As part of your preliminary analysis of the data, you are asked to present a graphical summary of the data.
      • What type of plot will you make?
      • Write down the RStudio command to make this plot.
      • Sketch what it might look like if exercise has a large effect on reducing cholesterol levels.

6. Many of my tests have problems something like the You Want To Know activity.