Chapter 7: Second Lab

1. This problem uses the M and M candies data set. Be sure to read the background information on the website regarding this data set. You may assume that the Plain M&Ms in this data set represent a simple random sample from all the Plain M&Ms in the world, and that the Peanut Butter M&Ms in this data set represent a simple random sample from all the Peanut Butter M&Ms in the world.

(a) Use the data set to test the hypothesis that the mean of the masses of all Peanut Butter M&Ms in the world equals the mean of the masses of all Plain M&Ms in the world. (The standard deviation of the masses of all the Plain and Peanut Butter M&Ms in the world is not known.)

(b) Find a 95% confidence interval for the mean of the masses of all Plain M&Ms in the world minus the mean of the masses of all Peanut Butter M&Ms in the world.

2. This problem uses the quizzes and finals data set. Be sure to read the background information on the website regarding this data set. You may assume that the exam scores in this data set represent a simple random sample from all possible exams scores students would earn when taking the quizzes or finals in questions.

(a) Use the data set to test the hypothesis that there is, on average, no difference between a student’s Spring 2007 quiz score and that student’s Spring 2007 final exam score (over all possible introductory statistics students who might take these quizzes as part of the course). Again the random variable standard deviations are not known.

(b) Find a 95% confidence interval for the mean of the differences between the Spring 2007 quiz scores and the Spring 2007 final exam scores. (Here, as before, “mean” refers to the theoretical random variable means over all possible introductory statistics students who might take these quizzes as part of the course.)
3. This problem uses the test-taking times 2008 data set (ID 13). Be sure to read the background information on the website regarding this data set. You may assume that the test-taking times data in this data set represent a simple random sample from all of the tests in the course, in each of the two Introductory Statistics sections. This is manifestly not true, but it is a simplifying assumption that we make for the purposes of modeling at present.

(a) Use the data set to test the hypothesis that the mean test-taking time in Section A equals the mean test-taking time in Section B. (Remember that these test-taking times are assumed to represent a simple random sample out of all the test-taking times for the semester. You are testing for the unknown mean test-taking time over all of the tests in the course.)

(b) Find a 95% confidence interval for the mean of the Section B test-taking times minus the mean of the Section A test-taking times.