Homework 5

1. Consider again the file `railtrailproperty.mpj`, which gives data for 104 houses in Northampton, MA that were sold in 2007. Assume that these data can be viewed as sample data from an ongoing process related to home prices in the northeastern United States.

(a) According to Walk Score, a bicycling score of 50 or higher corresponds to a location that is at least “bikeable” (it might also be “very bikeable” or a “biker’s paradise”). Do the data suggest that this Northampton neighborhood is at least bikeable? That is, is the average of the bike scores for the houses in the data significantly greater than this cutoff? Carefully state the hypotheses you are testing and the test that you are using. Use $\alpha = .01$ for your test.

(b) What assumptions are you making in constructing this test? Do they seem to hold here? Do you think that the implications of the result in part (a) are likely to be true, based on what you are seeing here? Can you see any patterns in these data that might help you characterize the actual bikeability pattern in this neighborhood?

(c) According to the U.S. Census American Housing Survey, the “typical” home in the U.S. has three bedrooms (in the sense of the most common number), with 20.4% having more than that number. You can see from the data that 36 of the 104 houses in the sample have more than three bedrooms. Is the observed proportion of homes with more than three bedrooms significantly different from the number provided by the American Housing Survey? Carefully state the hypotheses you are testing and the test that you are using. Use $\alpha = .05$ for your test.

(d) Is the average 2014 price significantly different from the average 2011 price for these houses? (Remember, these values have been corrected for inflation, so this is a very meaningful question.) Carefully state the hypotheses you are testing and the test that you are using. Use $\alpha = .01$ for your test. What assumptions are you making in constructing this test? Do they seem to hold here?

2. Consider again the file `hdtv.mpj`, which is based on data for 50 high-definition televisions. Consider these data to be a random sample of reasonable quality new televisions.

(a) Is the average price of the televisions significantly different from $750? Carefully state the hypotheses you are testing and the test that you are using. Use $\alpha = .05$ for your test.

(b) What assumptions are you making in constructing this test? Do they seem to hold here?

(c) Is the average HD picture quality score for televisions significantly different from the average sound score? Carefully state the hypotheses you are testing and the test that you are using. Use $\alpha = .01$ for your test. What assumptions are you making in constructing this test? Do they seem to hold here?

Homework due: November 27