**Stat 587 Section 2: Lab 5 self assessment**

**Fish price:**

These data were collected to assess whether average fish price has changed more than expected from inflation alone. If so, that is considered evidence of overfishing. The data (on the data sets page) are different types of fish. You are given the price of fish in 1980 and the inflation-adjusted 1970 price (i.e., the 1970 price adjusted for the overall inflation rate between 1970 and 1980). They are provided in two forms: fishprice.csv is in wide format (one column with adjusted 1970 price, one with 1980 price, 14 rows) and fishlong.csv is in long format (one column with period and one column with price, 28 rows). Use the appropriate file for the question.

1) Calculate the log transformation of the 1970 adjusted price and the 1980 price. No answer needed.

2) Consider the data to be two independent samples. Test the null hypothesis of no difference (between the two periods) in average log price. What is the p-value for this test?

3) Use a paired t-test to test the null hypothesis of no change in average log transformed price. What is the p-value for that test.

4) Which p-value (from part 3 or from part 4) is more appropriate to report? Why?

5) Estimate the mean difference between the log adjusted 1970 and log 1980 prices.

6) Convert the mean difference from question 5 into the values that “fills in” the ?? in this statement: The median fish price in 1980 is ?? times the inflation adjusted 1970 price.

7) Calculate a 95% confidence interval for the ?? in question 6.

**Physical activity in middle school kids:**

There is substantial interest in increasing middle school student’s physical activity (PA). A research group at ISU has developed a school-based program that attempts to do this. A recent study evaluated whether that new program actually works. 22 schools throughout Iowa were recruited for this study. 11 were randomly chosen to receive the new program; the other 11 continued with their traditional activities. Physical activity was measured on each kid in the 7’th grade: once in the early Fall semester before the new program started and again in the late Spring, after a full school year of either the new program or traditional activities. The response variable is the difference in PA, for each kid. There are a total of 679 kids in the study.

8) What is the experimental unit in this study? Briefly explain your choice.

9) What is the observation unit in this study? Briefly explain your choice.

10) If you do a t-test with 679 observations, the error variance has 677 degrees of freedom. Is this an appropriate test? Briefly explain why or why not.

Answers:

Note that you need in the fish price questions, to log transform price, then subtract. If you try to subtract prices, then log transform, you aren’t looking at the ratio of price. More importantly, some price differences are negative, for which the log is undefined. Negative differences of log values are just fine – they represent ratios less than 1.0.

2) p = 0.61

3) p = 0.0068.

4) from the paired t-test. The data are paired by type of fish.

5) mean difference (log 1980 – log 1970): 0.234.

6) 1.26, calculated as exp(0.234)

7) (1.08, 1.48), calculated by exponentiating the endpoints of the log scale confidence interval for the average difference in log price, i.e. (0.0765, 0.391)

8) school – that is what was randomly assigned to the treatment (new or traditional)

9) kid – they provide one row of data (the difference in PA for that kid)

10) No. The proposed analysis violates the assumption of independence. The ou is not the same as the eu.