

Advanced Data Analysis 1, Stat 427/527, Fall 2017

Assignment 2: Due Sep 7 Thursday in class

Problem 1. You can use R to perform the t-test for this problem, but don't show the R output in your answer. You can report either the  $t$  statistic or  $p$ -value (or both) in your answer which you get from R output, and interpret them to justify your conclusion, but don't copy the output from `t.test()` function in R. The following data are the total cholesterol levels (TCL) for a sample of 14 young adult males (aged 25 years or less) on the Kaiser Health plan in California:

TCL  
227  
239  
221  
213  
218  
246  
218  
224  
210  
204  
197  
229  
220  
197

Suppose it is believed that the mean TCL of all adult males in the United States is 210. Is it plausible the (population) mean TCL of all young adult males on the Kaiser plan is the same as the U.S. male population mean TCL? Test at the 5% level.

- (A) define the population parameter in context,
- (B) clearly state the hypotheses in notation and in words, is this a two sided test or one sided?
- (C) state assumptions and assess assumptions,
- (D) Discuss the CI and interpret the result in context.

- (E) Derive the test statistic, perform the  $t$  test, what did you conclude from the test? Is it consistent with CI observation?

Problem 2. A bearing used in an automotive application is supposed to have a nominal inside diameter of 1.5 inches. A random sample of 25 bearings is selected and the average inside diameter of these bearings is 1.4975 inches and the sample standard deviation is 0.01 inches.

- Is there evidence, at the significance level of .01, that the nominal inside diameter is less than 1.5 inches?
- What is the p-value of the test?

Problem 3. **Do this problem by R** In the acid data, the first column is the concentration values called “conc”, these are the results of a titration to determine the acidity of a solution in a chemistry class. The second column experiment ”exper” are with two different kinds of acid, called “Acid1” (124 rows) and “Acid2”. In this problem, we only consider “Acid1” experiment.

Download data acid.txt from my website, read in the data into R by using the following code

```
d1<-read.table(file="C:/jenn/teaching/ADA1/data/acid.txt",header=T)
#use your own path
#header=T : data has headline
dnew<-d1[which(d1[,2]=='Acid1'),]
#define a new data set ‘dnew”, in which "exper" are only with Acid1
#new data should have 124 rows
```

The instructor knew in both cases that the correct value for this solution was 0.110. Use a test of hypothesis and corresponding CIs to see if the class is consistent with the instructor’s solution.