## Assignment \#5 - Due Friday, April 5, by 4pm

*Submit your homework to Liam's or Hongzhi's mailbox anytime (according to your discussion sections) prior to the due date/time. The mailboxes are to the left as you enter the Medical Science Center (1300 University Ave.) from the main University Ave. entrance.
*No late homework will be accepted for credit!
*Please staple multiple pages together.
*On the top of the first page, remember to write your name and the section number of the discussion $(341,342,343)$ that you ATTEND. This will help us in returning your graded homework.
*If a problem asks you to use R, include a copy of the code and output. Please edit your code and output to be only the relevant portions.
*If a problem does not specify how to compute the answer, you may use any appropriate method.

1. The population mean for a Statistics test in a class that is usually taken by juniors and seniors is 75 (out of 85). A random sample of 15 sophomores is selected from the course. The research question is whether or not sophomores are different from the average student in this class. The scores of the 15 sophomores are

$$
76,76,66,49,84,72,72,50,56,62,65,71,60,82,73
$$

A QQ-plot of the data is shown below


Use the information above to answer the following questions
(a) State the null and alternative hypotheses.
(b) Explain why a T-test is suitable.
(c) Using the test described in part (b) and $\alpha=0.05$, compute the observed test statistics.
(d) Draw a conclusion in the context of the problem based on the results from part (c) using the rejection region method.
(e) Compute a $95 \%$ confidence interval for the sophomores and draw a conclusion in the context of the problem based on the interval.
2. A random sample of soil specimens was taken from a large geographic area. The specimens can be assumed to be independent. The amount of organic matter, as a percent, was determined for each specimen. The data are below:

$$
0.14,0.32,1.17,1.45,3.5,5.02,5.09,5.22
$$

A soil scientist wants to know whether the population mean percent organic matter is different than $4 \%$. A significance level of $\alpha=0.05$ is chosen.
(a) State hypotheses appropriate to the research question.
(b) Graph the data as you see fit. Why did you choose the graph(s) that you did and what does it (do they) tell you?
(c) Regardless of your conclusion from (b), use the bootstrap to perform a test of the hypotheses you stated in (a). Use $B=8000$ resamplings. Compute the p -value, and make a reject or not reject conclusion. Then state the conclusion in the context of the problem. In other words, does it seem the mean organic matter level is different than $4 \%$ ?
(d) Regardless of your conclusion from part (b), use a $T$-test to perform a test of the hypotheses you stated in (a). Compute the p-value, and make a reject or not reject conclusion. Then state the conclusion in the context of the problem. In other words, does it seem the mean organic matter level is different than $4 \%$ ?. (I recommend doing this part with a hand calculator and statistical tables as practice for exam conditions, but you may check your answers using $R$ if you wish.)
(e) Compare your answers from parts (c) and (d). Which method do you think is better? Are you surprised at the similarity or dissimilarity? What do you think explains this?
3. A study is conducted regarding shatterproof glass used in automobiles. Twenty-six glass panes are coated with an anti-shattering film. Then a 5 -pound metal ball is fired at 70 mph at each pane. Five of the panes shatter. We wish to determine whether, in the population of all such panes, the probability the glass shatters under these conditions is different from $\pi=0.2$.
(a) State the appropriate null and alternative hypotheses.
(b) Check the conditions for trusting the conclusion of the test, and calculate the observed value of an appropriate test statistic.
(c) Calculate the rejection region and draw a conclusion, given the significance level $\alpha=0.05$.
(d) Calculate the p-value.
(e) Compute the power of the test if the true $\pi$ was in fact 0.3 .
4. An animal's maintenance caloric intake is defined as the number of calories per day required to maintain its weight at a constant value. We wish to discover whether the median maintenance caloric intake, $m$, for a population of rats is less than $10 \mathrm{~g} /$ day. We draw a SRS of 17 rats , feed each rat 10 g of dry food per day for 30 days, and find that 4 of the rats lost weight, while the rest gained weight.
(a) State null and alternative hypotheses in terms of $m$.
(b) Let $B$ be the number of rats in a SRS of size 17 that exhibit daily caloric demands more than $10 \mathrm{~g} /$ day. If $H_{0}$ is true, what is the distribution of $B$ ?
(c) What is the value of $B$ observed in the study?
(d) Use the sign test to calculate the p -value and draw a conclusion using $\alpha=0.05$.

