

HW 3: One Sample Location Problems (Part 2)

Instructions: Work must be shown to receive full credit. You may work with others on the homework, but you must write and turn in your own copy. **This does not mean that you can simply copy someone else's work!!** Also, make sure your homework is neat, stapled, and all answers are written in complete sentences!! Come and see me if you have any questions.

On problems that require the use of R, PLEASE give me the RELEVANT R code and output to for each problem so I can assess partial credit. I may take off for including unnecessary R output. If one problem refers back to output from another problem, make sure to cite that output in your answer. Incorrect one-sentence answers will get little or no credit.

NOTE: If a problem asks you to perform a hypothesis test, make sure to give the hypotheses, test statistic, p-value, and a conclusion in the terms of the problem. Also, if the problem asks you to perform a confidence interval, make sure to interpret the confidence interval.

“By Hand” Problems: For hypothesis tests, you may use R to find the p-value. For confidence intervals, you may use R to find the multiplier.

1. Suppose that you are interested in purchasing a house in a particular area of a city and are interested in the average size of the homes in that area. In a random sample of 200 homes, you find a sample mean of 2127.94 square feet and a standard deviation of 387.276 square feet. Furthermore, you calculated a 99% confidence interval for the true mean size to be (2056.72, 2199.16). Why is it unnecessary to check for normality in this setting?
2. In an attempt to increase business on Monday nights, a restaurant offers a free dessert with every dinner order. Before the offer, the mean number of dinner customers on Monday was 150. The following data represents the number of diners on a random sample of 12 days while the free dessert offer was in effect.

206 169 191 152 212 139 142 151 174 220 192 153

Using the Wilcoxon signed rank test, it is possible to conclude that the typical number of diners on a Monday night increased while the free dessert offer was in effect? Be sure to state the hypotheses, calculate the test statistic, find the p-value in R, interpret the p-value, and state your conclusion in terms of the problem.

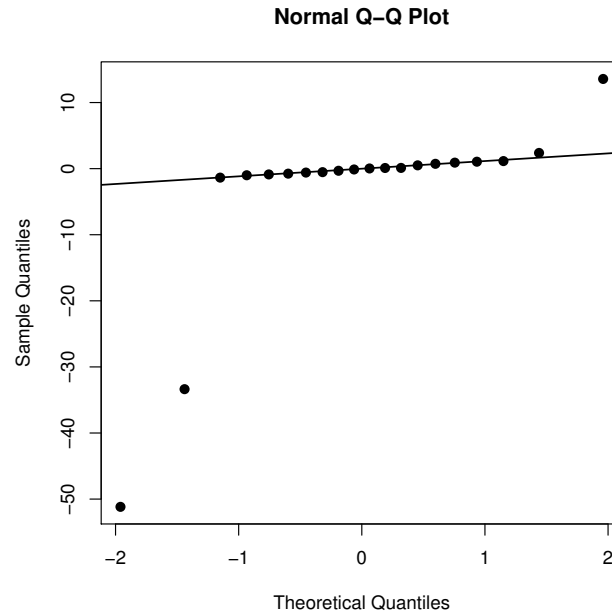
3. The following are the prices of 8 randomly selected 43 inch televisions.

\$279.99 \$299.99 \$649.99 \$379.99 \$449.99 \$379.99 \$399.99 \$289.99

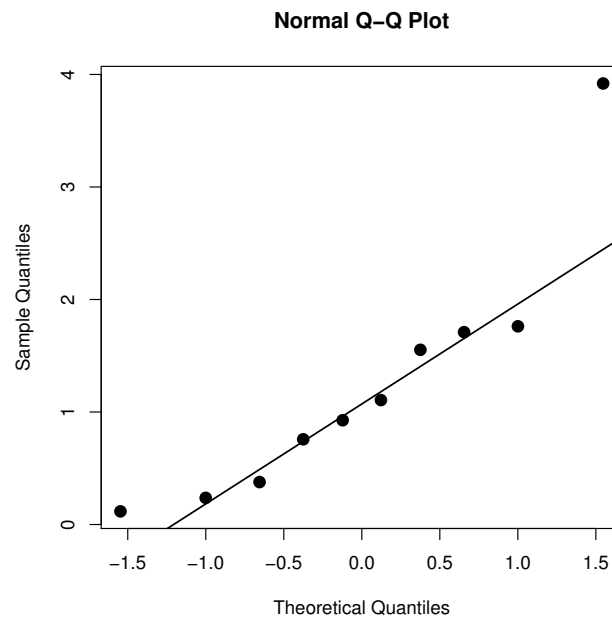
A friend of mine claims that a typical 43 inch TV is less than \$400. Using the Wilcoxon signed rank test, test my friend's claim. Be sure to state the hypotheses, calculate the test statistic, find the p-value in **R**, interpret the p-value, and state your conclusion in terms of the problem.

4. The following graphs are Normal probability plots for three different sets of data. Suppose we would like to construct a confidence interval for the mean in each case. For each setting, comment on whether a t tool or a non-parametric procedure should be used and explain why you made your decision.

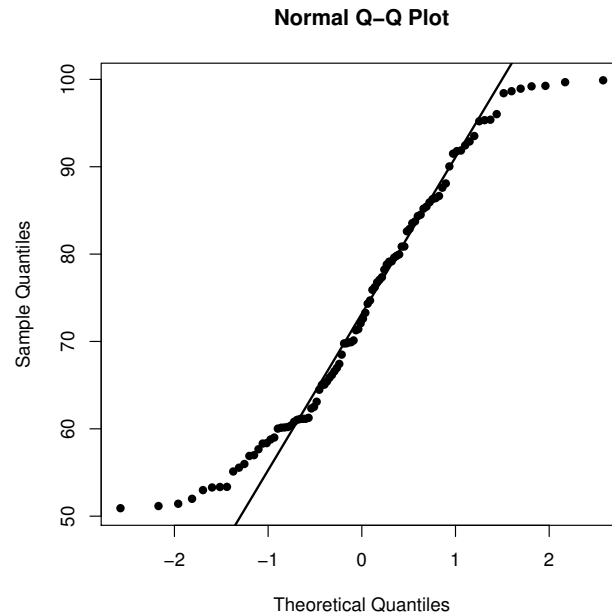
(a) In this setting, $n = 20$



(b) In this setting, $n = 10$



(c) In this setting, $n = 100$



“R” Problems:

5. A roller coaster enthusiast is interested in examining the speed and height of roller coasters across the country. He took a random sample of 20 roller coasters and measured their top speed (mph) and height of initial drop (feet). The data can be found on *Moodle* (coastersamp.xlsx) where **Speed** represents the maximum speed of the coaster (mph) and **Drop** represents the height (in feet) of the initial drop.
 - (a) The roller coaster enthusiast believes that the typical maximum speed of a roller coaster exceeds 55 mph. Use R and the Wilcoxon signed rank test to examine the validity of this claim. Be sure to include the R input, output, and full hypothesis test write-up from the results.
 - (b) In (a), a non-parametric method was used in place of a t tool. Construct Normal probability plots for both maximum speed and height of initial drop and comment on whether a t tool or a non-parametric method is most appropriate.
6. The following data gives the salaries (in millions of dollars) of 16 randomly selected Major League Baseball starting pitchers during the 2015 season.

23.5	19.5	12.0	9.8	9.0	6.7	4.27	3.2
2.8	2.5	1.5	0.8	0.52	0.51	0.51	0.51

- (a) A friend claims the typical salary is less than 2 million dollars. Use R and the Wilcoxon signed rank test to examine the validity of this claim. Be sure to include the R input, output, and full hypothesis test write-up from the results.
- (b) Construct a Normal Probability plot of the data and comment on which procedure (t or non-parametric) is most appropriate.