

HW 2: One Sample Location Problems (Part 1)

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. In the following examples, identify the each variable as either categorical or quantitative.
 - (a) attitude toward abortion
 - (b) gas mileage
 - (c) number of yellow *Skittles* in a bag
 - (d) t-shirt size.
2. A random sample of 100 women from the General Social Survey showed that the mean number of children reported 1.85 with a standard deviation of 1.5.
 - (a) What is the point estimate for the mean number of children?
 - (b) What is the standard error for the mean number of children?
 - (c) What is the multiplier for a 95% confidence interval of the mean number of children? (Show R code and output).
 - (d) Construct a 95% confidence interval for the mean number of children
 - (e) Interpret the interval in context.
3. State whether each of the following changes would make a confidence interval wider or narrower. Use the setting above for some context.
 - (a) Changing from a 95% level of confidence to a 90% level of confidence.
(Hint: Calculate t^* for $n = 100$ at 95% and 90% confidence. Decreasing the level of confidence does what to the multiplier?)

- (b) Changing from a sample size of 100 to a sample size of 200.
(Hint: Compute the standard error using the two different sample sizes. What happens to SE as n increases?)
 - (c) Changing from a standard deviation of 1.5 inches to a standard deviation of 0.5 inches. (Hint: Compute the SE in both cases with the original $n = 100$.)
4. In sports, an athlete's *ability* is a true but unknown value that describes what the player would do if given an infinite number of opportunities in the same context (e.g., same physical condition, location, opponents). An athlete's *performance* is an observed value that describes what the athlete actually did in a specific context. It is possible to use an athlete's *performance* to estimate his or her *ability*. Thus, a confidence interval constructed from an athlete's *performances* provides an interval of plausible values for an athlete's *ability*. In the 1986-1987 regular season, Magic Johnson of the Los Angeles Lakers led the NBA in assists. His average over 80 games was 12.2 assists per game, with a standard deviation of 3.96 assists per game.
- (a) Calculate a 95% confidence interval for Magic's ability to assist. Show all your work.
 - (b) Interpret the confidence interval in context.
 - (c) How could you reduce the margin of error in the interval you calculated in part (a)? Are there any drawbacks to your suggestions? Explain.
5. The mean weight of all 20 year-old-women is 128 pounds (<http://www.kidsgrowth.com>). A random sample of 40 vegetarian women who are 20 years old showed a sample mean of 122 pounds with a standard deviation 15 pounds. Is there enough evidence to conclude that the average 20-year old vegetarian woman weighs less than the average 20-year woman?
- (a) State the hypotheses of the test.
 - (b) Calculate the test-statistic by hand.
 - (c) Find the p-value of the test. Show R input and output.
 - (d) Interpret the p-value.
 - (e) State the conclusion of the test.
6. A random sample of 38 MLB (Major League Baseball) pitchers had an average velocity of 91.618 mph on fastballs with a standard deviation of 2.373 mph (Source: www.fangraphs.com). Does the data suggest that the average velocity of a MLB pitcher's fastball exceeds 90 mph?
- (a) State the hypotheses of the test.
 - (b) Calculate the test-statistic by hand.
 - (c) Find the p-value of the test. Show R input and output.
 - (d) Interpret the p-value.

- (e) State the conclusion of the test.

“R” Problems:

7. A random sample of 10 independent healthy people showed the following body temperatures (in degrees Fahrenheit):

98.5 98.2 99.0 96.3 98.3 98.7 97.2 99.1 98.7 97.2

Test the hypothesis that the average body temperature is not 98.6° F. Provide R code and output for data entry and the appropriate test. Write-up the results of the test (hyp, test stat, p-value + interpretation, conclusion) using the R output and no by hand work.

8. Felix Hernandez of the Seattle Mariners won the 2010 Cy Young Award as the American League’s best pitcher. Here are his strikeout totals for each of his 34 starts:

4 5 9 6 7 5 3 7 6 6 7 9 5 9 9 8 11
6 9 3 8 3 4 3 13 7 11 9 8 9 5 8 5 5

- (a) Use R to calculate the mean and standard deviation of Hernandez’s strikeout totals (performances).
- (b) Use R to construct a histogram of Hernandez’s strikeout totals. Make sure to properly label the axes.
- (c) Use R to calculate a 98% confidence interval for Hernandez’s ability to strike out batters. (No by hand work is required. Provide input and output.)
- (d) Interpret the confidence interval in context.
9. A realtor is interested in examining the price and size of homes for sale Murrells Inlet, South Carolina. Therefore, she randomly selected 40 homes that are currently for sale in that area. The data can be found on *Moodle* (MIhomes.xlsx).
- (a) Use R to calculate the mean and standard deviation for the size of a home for sale in Murrells Inlet.
- (b) Use R to calculate a 95% confidence interval for the average size of a home for sale in Murrells Inlet, SC.
- (c) Use R to calculate the mean and median price of a home for sale in Murrells Inlet, SC.
- (d) Use R to construct a boxplot of the sales price of a home in Murrells Inlet.
- (e) Use R to test the claim that the average price of a home for sale in Murrells Inlet is more than \$300,000. Be sure to provide a write-up of the results (hyp, test stat, p-value + interpretation, conclusion).