Kylie Gannon

Dr. Emerson-Stonnell

Math 171-50

3 November 2019

T Intervals to Determine the True Mean Temperature at Lancer Park in Fall

This paper will examine the data taken by the Longwood Environmental Observatory (LEO) over the span of two years at Lancer Park. The purpose of examining this data is to determine the average high temperature at Lancer Park in the fall. The data given is safe to use to find this information because it is taken from a simple random sample and therefore represents the population. I chose to examine fall because it is my favorite season.

The parameter is the true mean high temperature (in Celsius) at Lancer Park in the fall. Based off the sample listed in the appendix and the question posed, the correct test to use in order to find the parameter is a t interval. This is since the problem is a confidence interval and not a hypothesis test, which is revealed by the fact that we are not testing a claim. This is also because there is no population standard deviation given, which rules out the z interval. I will perform this test at the 0.90 and 0.95 confidence levels. A confidence level is how “confident” a person is that the parameter is represented by the confidence interval. The higher the confidence level, the more likely the parameter is within the confidence interval.

First, I will perform the test at the 0.90 confidence level. As mentioned previously, the parameter is the true mean high temperature (in Celsius) at Lancer Park in the fall. The test is the t interval. Certain conditions must be met for the test to be performed. Again, there is no population standard deviation given. The data came from a simple random sample. When the boxplot is graphed, it shows no outliers. When the histogram is graphed, it reveals that the data is close to normal. Due to this, n can be equal to any size, so it is alright that n is equal to 25. Since the conditions were checked, the data can be plugged into the calculator. Since I am given the sample, the data is typed into list 1. Then I went to stat and test and scrolled down to t interval. Since the data was in L1, I selected data and typed in my confidence level of 0.90. If the data wasn’t given in a list, you would select stats instead of data and must list out the sample mean, the sample standard deviation, and n. When I compute the confidence interval, it will give me this data anyways because I entered in all the data from the sample. When I calculated the confidence interval, the interval was 20.498 <μ < 24.898. The sample mean was 22.6976. The sample standard deviation was 6.4295, and n=25. From this information I could conclude that we are 90% confident that the true mean high temperature (in Celsius) at Lancer Park in the fall is between 20.498 and 24.898 degrees Celsius. In other words, if this test was performed many, many times using repeated samples and the same 90% confidence level, 90% of the intervals would contain the true mean high temperature (in Celsius) at Lancer Park in the fall.

Next, I will perform the test at the 0.95 confidence level. The process is the same as the previous test until the data is plugged into the calculator. Instead of typing in 0.90 as the confidence level, I must type in 0.95. I still used the t interval as my test and used the data function. When I selected enter, I found that my interval was 20.044 <μ < 25.352. The sample mean, sample standard deviation, and n were all the same as the previous test. The confidence interval in context was: we are 95% confident that the true mean high temperature (in Celsius) at Lancer Park in fall is between 20.044 and 25.352 degrees Celsius. In other words, if this test was performed many, many times using repeated samples and the same 95% confidence level, 95% of the intervals would contain the true mean high temperature (in Celsius) at Lancer Park in the fall.

There is a problem with these tests, however. Even though the days that the Longwood Environmental Observatory takes the temperature at Lancer Park is random, the temperature could be different from the rest of the season on some of the days that they recorded the temperature. In other words, LEO may have received some data in their sample that were not quite outliers, but still differed from the rest of the data. This could potentially influence the data.

The confidence intervals used, with confidence levels of 90% and 95% are similar because in the context of the problem they both have the same n, sample mean, and sample standard deviation. The confidence intervals are also the same because they both are affected by the confidence level and margin of error. These confidence intervals are different, however, because of their different confidence levels and margin of errors. The first confidence interval, at a 90% confidence level, has a smaller interval than the 95% confidence level from the second confidence interval. When a confidence interval has a smaller confidence level, there is a narrower confidence interval. On the other hand, a larger confidence level causes the width of the interval to increase and for the interval to be more accurate. In other words, a smaller confidence level causes a smaller confidence interval. This means that a person would have a shorter confidence interval, but the drawback is that the interval is less likely to contain the population parameter being examined. Likewise, a larger confidence level causes a wider interval. A person would have a more accurate confidence interval, at the expense of it being wider. In the context of this problem, a person should use the smaller 90% confidence level if they are concerned with a more precise interval. Inversely, a person should use the larger 95% confidence level if they want a more accurate confidence interval.

The margin of error is the largest expected difference between the actual population parameter and the sample estimate of that parameter. The margin of error differs for each confidence interval. For the first confidence interval with a confidence level of 90%, the margin of error is ±2.2004. I calculated this by subtracting the sample mean (22.6976) from the value of the interval that is greater than μ (24.989). For the second confidence interval with a confidence level of 95%, the margin of error is ±2.6544. I calculated this by subtracting the sample mean (22.6976) from the value of the interval that is greater than μ (25.352). These calculations show that a smaller confidence level like that of 90% corresponds with a smaller margin of error, and a larger confidence level such as 95% corresponds with a larger margin of error. This is due to shorter width of intervals that occur with smaller confidence levels, and the longer intervals that occur with higher confidence levels. Therefore, in the context of this problem, if a person wished to have a smaller margin of error and a more precise interval, they should use the 90% confidence level. If a person wished to have a larger margin of error and a more accurate confidence interval, they should use the 95% confidence level.

Two t intervals were calculated to determine the true mean high temperature (in Celsius) at Lancer Park in fall. The result of the first test was: if the test was performed many times using repeated samples and the same 90% confidence level, 90% of the intervals would contain the true mean high temperature (in Celsius) at Lancer Park in the fall. The second test stated: if this test was done many times using repeated samples and the same 95% confidence level, 95% of the intervals would contain the true mean high temperature (in Celsius) at Lancer Park in the fall.

Appendix

