Kelly Jarratt

11/21/19

H Math 171

Individual Project #2

Two Sample Hypothesis Test for High Temperatures at Lancer Park

The purpose of this paper is to determine if there is a difference between the average high temperatures in fall versus spring at Lancer Park. The data used in this project was obtained from the Longwood Environmental Observatory, which collected the temperatures in Celsius over several years. I chose to use the high temperatures because I had chosen the low temperatures for the last project. I also thought it would be interesting to see how high the temperatures reached at Lancer Park over the past few years for both fall and spring seasons.

The hypotheses for this test are H₀: μ₁ = μ₂ (this is the null hypothesis), and Ha: μ₁ ≠ μ₂ (this is the alternative hypothesis). The alternative hypothesis is the one that is tested. The parameters for this test are μ₁, which equals the average high temperatures (in Celsius) at Lancer Park during the spring and μ₂, which equals the average high temperatures (in Celcius) at Lancer Park during the fall. I used μ₁ and μ₂ because I am testing for two means. I decided to use a 2-sample t-test because I am testing a claim without knowing σ (the sample standard deviation). I am also using a two-sample t-test because there are two independent samples; the temperatures were recorded during different seasons and could not have interfered with each other.

In order to use a 2-sample t-test, I first needed to determine if the sample was a simple random sample. I know that my sample was a simple random sample because it stated that it was a simple random sample in the problem. If my sample was not a simple random sample, I should not have been doing the test because of the various biases that may occur. The next step is to make sure that sigma, σ , was unknown. I know that sigma was unknown in my sample because the problem did not state the σ. If sigma was known, I should have been using a two-sample z-test instead of two-sample t- test. The final step is to determine whether or not my sample had any outliers or skew. Both of my samples did not have any outliers, but both samples did have a light left skew. If my sample were to have any outliers, I should not have been doing the test. Since both of my samples did have a slight left skew, I had to make sure the sample had a population of at least greater than or equal to fifteen. My population included twenty-five temperatures, so I was okay to proceed with the test.

To perform the actual test, I had to plug in the spring data from the Excel sheet into list 1 on my calculator and the fall data into list 2. To plug in a list, I hit stat, edit, and then entered the numbers for both of the data. After plugging in the numbers, I hit stat, tests, and then went down to hit two-sample t-test. I did not have to find out the statistics my first, as I just used the data I already plugged into my calculator, which spit out the statistics for me. However, I did have to determine the significance level that I wanted to use for the test. I chose to use a .05 significance level (𝛼), which is the level of evidence that must be present to determine whether to reject or accept the alternative hypothesis. I then punched in the .05 significance level into the calculator and hit enter. The calculator gave me the statistics for the .05 significance level which included x̄₁ = 20.736 (the average mean of sample 1), S₁ = 6.6311 (the standard deviation of sample 1), n₁ = 25 (sample size of sample 1), x̄₂ = 22.6976 (the average mean of sample 2), S₂ = 6.4295 (the standard deviation of sample 2), n₂ = 25 (the sample size of sample 2). The calculator also calculated t = -1.0619 (the z-score) and the p-value = 0.2936 (the probability of rejecting the null hypothesis). Because the p-value (0.2936) was greater than the significance level (.05), I had to reject Ha. For the .05 significance level, the data showed that there was no significant evidence that the true average mean for high temperatures (in degrees Celsius) for Lancer Park in the fall was different than the true average mean for high temperatures (in degrees Celsius) in the spring.

After completing the tests, I concluded that the average high temperatures in Lancer Park for spring and fall were not different. There was not enough evidence to support the alternative hypothesis which stated that there was a difference in high temperatures for Lancer Park in the spring versus the fall. In simple terms, this means that I cannot conclude that there was a difference between the two, since there was not a significant amount of mathematical evidence that supported that claim.

Appendix:

|  |  |
| --- | --- |
| spring | 27.99 |
| spring | 25.7 |
| spring | 24.86 |
| spring | 26.14 |
| spring | 16.01 |
| spring | 29.36 |
| spring | 20.45 |
| spring | 23.27 |
| spring | 12.51 |
| spring | 29.83 |
| spring | 16.72 |
| spring | 21.63 |
| spring | 11.77 |
| spring | 30.51 |
| spring | 19.68 |
| spring | 23.27 |
| spring | 8.89 |
| spring | 19.17 |
| spring | 20.81 |
| spring | 5.2 |
| spring | 15.04 |
| spring | 16.52 |
| spring | 23.8 |
| spring | 23.27 |
| spring | 26 |

|  |  |
| --- | --- |
| fall | 26.54 |
| fall | 19.07 |
| fall | 16.08 |
| fall | 26.08 |
| fall | 7.88 |
| fall | 17.71 |
| fall | 28.84 |
| fall | 29.72 |
| fall | 22.73 |
| fall | 31.04 |
| fall | 20.6 |
| fall | 19.23 |
| fall | 20.23 |
| fall | 30.76 |
| fall | 29.91 |
| fall | 24.81 |
| fall | 27.08 |
| fall | 31.99 |
| fall | 19.71 |
| fall | 21.13 |
| fall | 18.55 |
| fall | 28.22 |
| fall | 18.43 |
| fall | 9.27 |
| fall | 21.83 |