Breonna Williams

Kaitlin Smith

MATH 171-50

Statistics

Dr. Wears

5/2/2017

**Final Project**

When we collected a simple random sample of 50 pennies, we first used the pennies that we had of our own and recorded the dates of them. After not having the sufficient amount, we asked a few of our friends or family for the dates of any pennies that they had and recorded the dates of them. When we collected a simple random sample of 50 dimes, we first used the dimes that we had of our own and recorded the dates of them. After not having a sufficient amount, we asked a few of our friends or family for the dates of any dimes that they had and recorded the dates of them. We do not believe that our data collection method could introduce bias because of the sample being random and is not an opinion-based sample. Our data will be exact facts from our sample without any biased results.



This is the boxplot of the simple random sample of 50 pennies that we collected.



This is the boxplot of the simple random sample of 50 dimes that we collected.

For our simple random sample of 50 pennies, the data is skewed to the right with a center of 6.5 years and a spread of 47 years. The distribution may have a couple of outliers being 40 years and 47 years. The distribution has the shape that we have observed because the mean is greater than the median, causing it to skew more to the right.

For our simple random sample of 50 dimes, the data is approximately normal with a center of 11.5 years and a spread of 29 years. The distribution may have a couple of outliers being 24 years and 29 years. This distribution has the shape we observed because the mean is slightly smaller than the median, but not small enough to skew the data.

The five-number summary for our simple random sample of the ages of pennies is (0, 2, 6.5, 21, 47) with a standard deviation of 13.098. Given that for our sample, we have a median of 6.5 and a sample larger mean of 12.5, we can automatically assume a skewed right distribution.

The five-number summary for our simple random sample of the ages of dimes is (0, 4, 11.5, 16, 29) with a standard deviation of 7.391. Given that for our sample, we have a median of 11.5 and a close sample mean of 10.840, we can automatically assume a normal distribution.

We are assuming that this will be a t-confidence interval because we do not know the population standard deviation, but we know our sample standard deviation. Given a sample mean of 12.5, a sample standard deviation of 13.098 and a sample size of 50, we are 95% confident that the mean age (in years) of pennies that we collected is between 8.778 and 16.222 years.

 We are assuming that this will be a t-confidence interval because we do not know the population standard deviation, but we know our sample standard deviation. Given a sample mean of 10.840, a sample standard deviation of 7.391 and a sample size of 50, we are 95% confident that the mean age (in years) of dimes that we collected is between 8.74 and 12.94 years.

 The margin of error (½ \* Length of the Confidence Interval) for each of our estimates is 3.722 for the pennies and 2.1 for the dimes.

To estimate the average age of pennies to within one year with 99% confidence we would take our 99% z\* value of 2.58 and multiply that by our sample standard deviation of 13.098 to get 33.793. We then take 33.793 and divide it by our margin of error of 3.722 to get 9.079. We take 9.079 and square to get a final value of 82.432. To estimate the average age of pennies to within one year with 99% confidence we would have to have a sample size of 83 pennies.

For our sample of pennies, we used the 1.5\*IQR rule to test for outliers and determined that this sample does not have any outliers because the lower outlier benchmark was unrealistic at -28.5 years and our upper outlier benchmark came out to be 49.5 years. Given that our maximum sample item is 47 years, we do not have any outliers. Because each item in our sample represents 2% of our population, a penny that would be considered “rare” would be one that is 0 years old and one that is 47 years old because they are within the lower 2% of our sample and the upper 2% of our sample.

For our sample of dimes, we used the 1.5\*IQR rule to test for outliers and determined that this sample does not have any outliers because the lower outlier benchmark was unrealistic at -14 years and our upper outlier benchmark came out to be 34 years. Given that our maximum sample item is 29 years, we do not have any outliers. Because each item in our sample represents 2% of our population, a dime that would be considered “rare” would be one that is 0 years old and one that is 29 years old because they are within the lower 2% of our sample and upper 2% of our sample.

**Hypothesis Testing**

*Is the mean age of all dimes in circulation the same as the mean age of all pennies in circulation?*

 **P:**

* Population 1: All American Pennies
* 𝜇-1: The mean age (in years) of pennies
	+ n-1 = 50
	+ s-1= 13.098
* Population 2: All American Dimes
* 𝜇-2: The mean age (in years) of dimes
	+ n-2 = 50
	+ s-2 = 7.391

**H:** Ho: 𝜇-1=𝜇-2

 Ha: 𝜇-1 ≠ 𝜇-2

**A:** i) Both samples were collected randomly.

 ii) Both sample sizes are 50 so are sufficiently large, but there are a couple outliers to aware of.

 iii) Skewed right distribution and skew right data in the pennies sample data. Approximately normal distribution and data for the dimes sample data.

**N:** Two Sample T-test for difference in population means

**T:** Test Statistic

 T = 0.780

**O:** Obtain P-value

 P-value = 0.4375

**M:** Because our P-value is not statistically significant (does not mean the 𝛂 = 0.05 significance level), then we have decided to fail to reject the null hypothesis.

**S:** We do not find significant evidence that the true mean age of pennies in circulation is different than the true mean age of dimes in circulation.