

Stat 235

Lab 4

WOOSAREE, Arun

Lab EL12

TA: Jessa Marley

November 20, 2018

# 1

## 1 a)

Keeping other parameters constant, changing the confidence level yields the following:

Confidence Level	Margin of Error
0.90	0.300308
0.95	0.357839
0.99	0.470280

Table 1: Confidence levels and the corresponding Margins of Error

As seen in Table 1 above, the Margin of Error increases as the Confidence Level is increased. This makes sense because the margin of error depends on the z value, which also increases as the Confidence Level ( $1 - \alpha$ ) increases.

## 1 b)

Confidence Level	Observed Fraction of Intervals That Failed to Cover the Hypothesized Population Mean
0.90	0.11
0.95	0.06
0.99	0.02

Table 2: Observed fraction of intervals that failed to cover the hypothesized population mean at various confidence levels.

Theoretically, the confidence level and the fraction of intervals that failed to cover the hypothesized mean should add up to 1. Here, we see that the values are reasonable, adding up to 1.01 in all cases. Of course, a small difference is expected since we are using experimental data.

# 2

$$H_0 : \mu = 64 \quad vs. \quad H_A : \mu \neq 64$$

**2 a)**

Level of Significance	Number of Samples That Led to the Rejection of $H_0$	Observed Fraction of Samples
0.10	89	0.89
0.05	94	0.94
0.01	98	0.98

Table 3: Samples that led to the rejection of  $H_0$  at different levels of significance.

As the level of significance increases, the number of samples that led to the rejection of  $H_0$  also increases. This makes sense, because the the probability of making a Type I error is  $\alpha$ , (or the level of significance) which happens when the null hypothesis is incorrectly rejected. With a higher tolerance for error, we would expect the amount of incorrect rejections to increase, which would increase the total fraction of samples rejected.

**2 b)****90% confidence interval null hypothesis**

90 % of the time, the population mean is 64 ksi.

**95% confidence interval null hypothesis**

95 % of the time, the population mean is 64 ksi.

**99% confidence interval null hypothesis**

99 % of the time, the population mean is 64 ksi.

The null hypothesis  $H_0$  is rejected when the test statistic does not fall within the confidence interval with significance  $\alpha$ . The confidence interval is calculated as the mean  $\pm$  the confidence level  $1 - \alpha$

**3****3 a)**

Alloy 1	
Mean	65.09
Standard Error	0.360980466
Median	64.6
Mode	63.8
Standard Deviation	1.977171438
Sample Variance	3.909206897
Kurtosis	0.042639157
Skewness	0.718164135
Range	8.2
Minimum	61.7
Maximum	69.9
Sum	1952.7
Count	30
Confidence Level(95.0%)	0.738287948

Table 4: Summary statistics for Alloy 1

The confidence interval is calculated as follows:

$$65.09 \pm 0.738287948 \approx (64.352, 65.828)$$

Alloy 2	
Mean	65.27333333
Standard Error	0.167601973
Median	65
Mode	64.9
Standard Deviation	0.917993815
Sample Variance	0.842712644
Kurtosis	9.565960304
Skewness	2.914366915
Range	4.5
Minimum	64.5
Maximum	69
Sum	1958.2
Count	30
Confidence Level(95.0%)	0.342784524

Table 5: Summary statistics for Alloy 2

The confidence interval is calculated as follows:

$$65.27333333 \pm 0.342784524 \approx (64.931, 65.616)$$

Alloy 2 appears to be stronger, since it has a higher mean, median and mode compared to Alloy 1. **do we mention the confidence interval here**

### 3 b)

For both of the alloys, there isn't any strong evidence that the mean strength is below the required threshold value of 64. For both of the 95% confidence intervals, the lower bound is above 64, so the chance for a mean strength below 64 is low.

## 4

### 4 a)

#### Alloy 1

$$H_0 : \mu \leq 64 \quad vs. \quad H_A : \mu > 64$$

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{65.09 - 64}{0.917993815/\sqrt{30}} \approx t_{29}$$

$$p - value : pt(3.019553976, df = 29, lower.tail = FALSE) = 0.0026185$$

Because the p-value obtained is extremely low, we reject  $H_0$ . i.e. the data strongly suggests that Alloy 1 exceeds the threshold value of 64.

#### Alloy 2

$$H_0 : \mu \leq 64 \quad vs. \quad H_A : \mu > 64$$

$$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{65.27333333 - 64}{0.917993815/\sqrt{30}} = 7.597364802 \sim t_{29}$$

$$p - value : pt(7.597364802, df = 29, lower.tail = FALSE) = 1.121033 \times 10^{-8}$$

Again, this p-value is extremely low, so we reject  $H_0$ . i.e. the data strongly suggests that Alloy 2 exceeds the threshold value of 64.

**4 b)**

The following assumptions must be true for a t-test:

1. The samples are independent and random
2. The samples come from a normal population, OR from a population with sample size 30 or greater (The Central Limit Theorem guarantees samples are normally distributed when the population size is  $\geq 30$ )

The assumptions outlined above hold for the test above, as we were told that the rods were randomly selected, and the sample size of the population is 30.

**5****5 a)**

We'll do the two-tailed t-test as follows:

Since, the population variances are unknown, we'll assume unequal variances.

$$H_0 : \mu_1 = \mu_2 \quad vs. \quad H_A : \mu_1 \neq \mu_2$$

Using the “t-Test: Two-Sample Assuming Unequal Variances” tool, we obtain the following:

	Alloy 1	Alloy 2
Mean	65.09	65.27333
Variance	3.909206897	0.842713
Observations	30	30
Hypothesized Mean Difference	0	
df	41	
t Stat	-0.460646232	
$P(T \leq t)$ one-tail	0.32374327	
t Critical one-tail	1.682878002	
$P(T \leq t)$ two-tail	0.647486541	
t Critical two-tail	2.01954097	

Table 6: My caption. The tool uses  $\alpha = 0.05$

The test statistic  $t_0 = -0.460646232$  follows a t-distribution, with corresponding p-value: 0.647486541 obtained from the table above.

$$t_{\alpha/2, \min\{n-1, m-1\}} = t_{0.05/2, 29} = pt(0.025, df = 29, lower.tail = TRUE) = 0.5098869$$

$H_0$  should be rejected if  $t_0$  is greater than the value above, which isn't the case. Similarly, with the “judgement approach” we find that the p-value is above 0.1, which is weak to no evidence against  $H_0$ . Therefore, we fail to reject  $H_0$ . i.e. there is not sufficient evidence to support that there is a difference in the mean strengths of Alloy 1 and Alloy 2 rods.

## 5 b)

The assumptions to make the tests in part (a) valid are as follows:

1. The samples are independent and random
2. The samples come from a normal population, OR from a population with sample size 30 or greater (The Central Limit Theorem guarantees samples are normally distributed when the population size is  $\geq 30$ )
3. the populations have unequal variances

The first two assumptions outlined hold for the test above, as we were told that the rods were randomly selected, and the sample size of the population is 30. **The population variances are unknown so we don't know if the 3rd assumption is valid?!**

## 6

### 6 a)

Do the data provide evidence that the treatment increased the mean strength of the ALLOY 2 rods? Answer the question by carrying out an appropriate test in Excel. In particular, state the null and alternative hypotheses in terms of the population parameters, obtain the value of the test statistic, specify the distribution of the test statistic under the null hypothesis, and obtain the p-value of the test. What do you conclude?

We'll do the one-tailed t-test for paired data as follows:

$$H_0 : \mu_2 - \mu_{2+treatment} \geq 0 \quad vs. \quad H_A : \mu_2 - \mu_{2+treatment} < 0$$

Using the “t-Test: Paired Two Sample for Means” tool, we obtain the following:

	Variable 1	Variable 2
Mean	65.27333	66.823333
Variance	0.842713	0.3521954
Observations	30	30
Pearson Correlation	0.86516	
Hypothesized Mean Difference	0	
df	29	
t Stat	-16.9038	
$P(T \leq t)$ one-tail	7.41E-17	
t Critical one-tail	1.699127	
$P(T \leq t)$ two-tail	1.48E-16	
t Critical two-tail	2.04523	

Table 7: My caption. The value  $\alpha = 0.05$  was used

The test statistic  $t_0 = -16.9038$  follows a t-distribution, with corresponding p-value:  $7.41 \times 10^{-17}$  obtained from the table above.

We immediately notice that the p-value is so incredibly small that we can reject  $H_0$ . i.e. the data strongly suggests that the treatment increased the mean strength of the Alloy 2 rods.

### 6 b)

Use the Descriptive Statistics feature in the Data Analysis menu to obtain a 95% two - sided confidence interval for the mean change in strength o f ALLOY 2 rods after the treatment. First create a new variable , EFFECT , defined as the difference in strength between ALLOY 2 + TREATMENT rods and ALLOY 2 rods. Is the interval consistent with the test outcome in part (a)? Explain briefly.

Mean	1.55
Standard Error	0.091695
Median	1.6
Mode	2
Standard Deviation	0.502236
Sample Variance	0.252241
Kurtosis	0.942104
Skewness	-0.9204
Range	2.2
Minimum	0.1
Maximum	2.3
Sum	46.5
Count	30
Confidence Level (95.0%)	0.187538

Table 8: My caption

The two-sided confidence interval is found as follows:

$$1.55 \pm 0.187538 \approx (1.362, 1.738)$$

This interval is consistent with the test outcome in part (a), because it indicates with high confidence that the population mean of Alloy 2 + treatment is between 1.362 and 1.738 ksi stronger than the population mean of Alloy 2 (sans treatment). This agrees with the test outcome in part (a), which indicated with high confidence that the population mean of Alloy 2 + treatment is higher than the population mean of Alloy 2 (without treatment).

### 6 c)

What are the assumptions necessary to make the test in part (a) and confidence interval in part (b) valid? Do the assumptions hold? Explain briefly.



The following assumptions are necessary to make the test in part (a) and confidence interval in part (b) valid:

1. The samples are random
2. The samples come from a normal population, OR from a population with sample size 30 or greater (The Central Limit Theorem guarantees samples are normally distributed when the population size is  $\geq 30$ )
3. the samples are paired?

The first two assumptions hold, as we were told that the rods were randomly selected, and the sample size of the population is 30. **IS there more to be said here**

## 6 d)

Is the effect of the treatment independent of the initial strength of the rods? In order to answer the question, obtain the plot of the variable EFFECT versus ALLOY 2 measurements. What do you conclude?

Output: Change in strength vs Scatter plot of Alloy 2 strength. Remember to put the response on the y-axis. All relationships have error, look at the middle data cloud to determine shape.

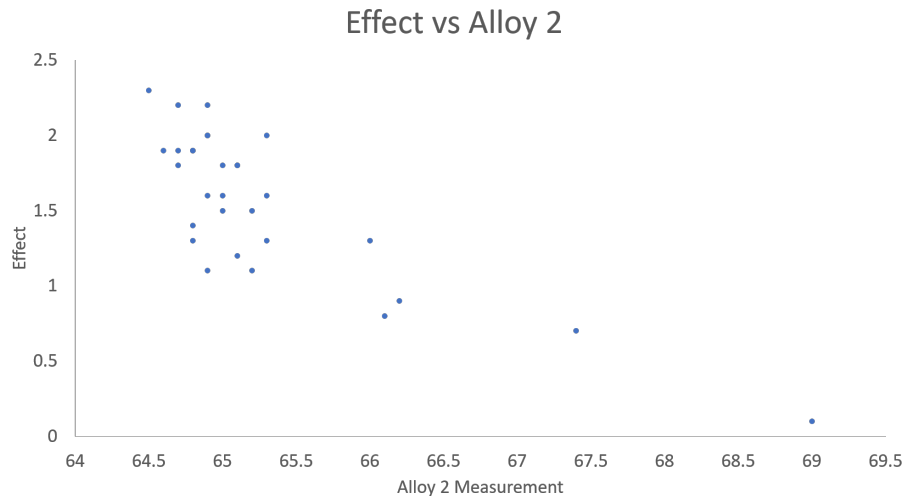


Figure 1: My caption

From the scatterplot above, we can clearly see that with higher strength, the effect is less. This suggests that the effect of the treatment is not independent of the initial strength of the rods. The relationship appears to be linear.