**Homework 2**

Name: Quan Yuan UNI: qy2205

**2.1**

**(a)** Yes, the conclusion is warranted. The confidence interval for the slope is [0.452886, 1.05721], which not includes 0. So at 5% significance level, the slope is significantly different than 0. (if H0: slope =0, we will reject the null hypothesis)

**(b)** The result of the regression model highly depends on the data we use, the real relationship may like . But when we only use data with x much bigger than 0, the regression line may have negative intercept. Besides, the value of is not important since we don’t need to consider the sales volume in an area with population equals to 0.

**2.4**

**(a)** Since is unknown, we could get,

So, based on the above formula, we could get

According to the result in 1.19

Therefore, the confidence interval for is

Does it include zero? No

Why might the director of admissions be interested in whether the confidence interval includes zero? Because if the confidence interval include 0, we cannot reject the null hypothesis test that , which means there is no relationship between ACT and GPA.

**(b)** a = 0.01

First set hypothesis:

Then we calculate t-statistic

Since

So, we reject H0, there is a linear relationship between ACT and GPA.

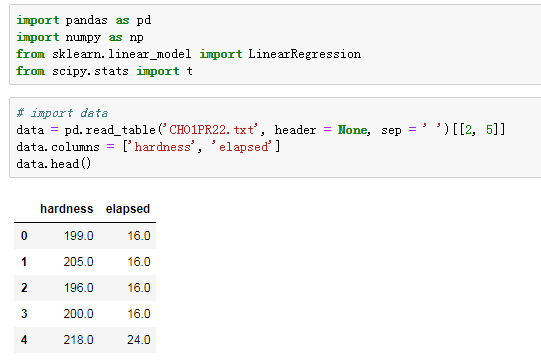
**(c)**

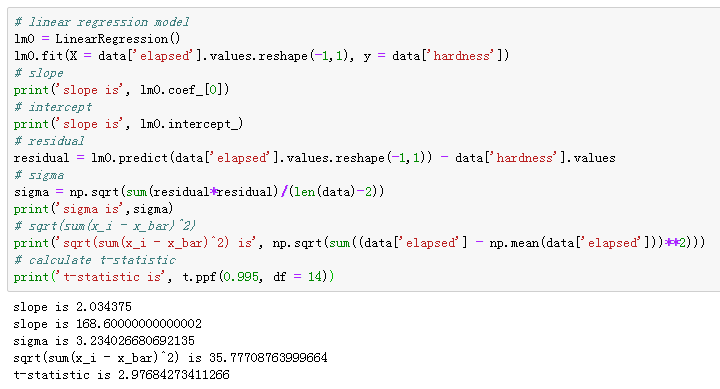
equals to 0.0029 means the probability of is 0.29% which is very small. So we can reject the null hypothesis.

**2.7**

**(a)** For calculating the change, we need to estimateand also get 99% confidence interval. Since we still don’t know the standard deviation, so the confidence interval is

We use Python to calculate the required value.





After that, we can calculate the confidence interval is [1.77, 2.30], which means when the elapsed time increases by one hour, the probability of the change in the mean hardness within 1.77 to 2.30 is 99%.

**(b)** Set hypothesis:

Test Statistics:

Observed t:

So we accept the null hypothesis.

**(c)**

According to the table B.5 in Appendix,

**2.12**

The difference is because the observed values will always have random errors that follow normal distribution.

**2.13**

**(a)** The 95% confidence interval of is

Which means the probability of the mean freshman GPA for students whose ACT score is 28 in the range of 3.063 to 3.34 is 95%.

**(b)** The 95% prediction interval of is

Which means the probability of Mary Jones’s GPA in the range of 1.972 to 4.43 is 95%.

**(c)** Yes, the prediction interval wider than the confidence interval. The confidence interval is an inference on a parameter, therefore, it is intended to cover the value of the parameter. The prediction interval describes the value for a random variable and therefore must have a wider interval to allow for non-parameterized variables to impact the predicted value.

**(d)** The 95% confidence band is

So

It is wider than the interval in part (a) because it is representing the confidence intervals for the entire regression line, not just at a single point .

**2.51**

**2.52**

Based on (2.31)

Since are

is a special case when in

**2.63**

We use Python to solve this problem.

Since we don’t know , the confidence interval is

You can see the code and print result below.

