Exam 3

Mitchell Meier

December 4, 2020

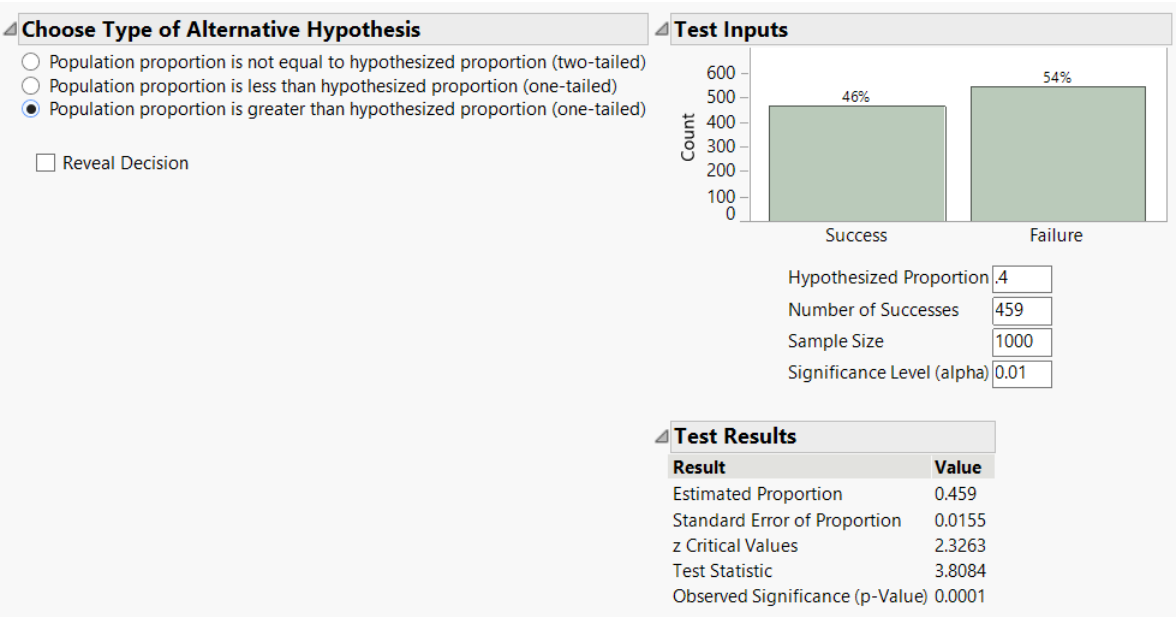
Problem 1

1. (a) $H_0 : \mu = 0.4$
   $H_1 : \mu > 0.4$

   (b) Significance level (p-value) $= 0.0001$

   $0.0001 < 0.01$, $p < \alpha$

   Therefore, we reject the null hypothesis H_0, and can say with 99% confidence that $\mu > 0.4$, or that the proportion of adults who live in Tempe, Arizona, who are college graduates is greater than 40%



2. (a) μ_0 = Mean of weight percent calcium in standard cement
μ_1 = Mean of weight percent calcium in cement doped with lead

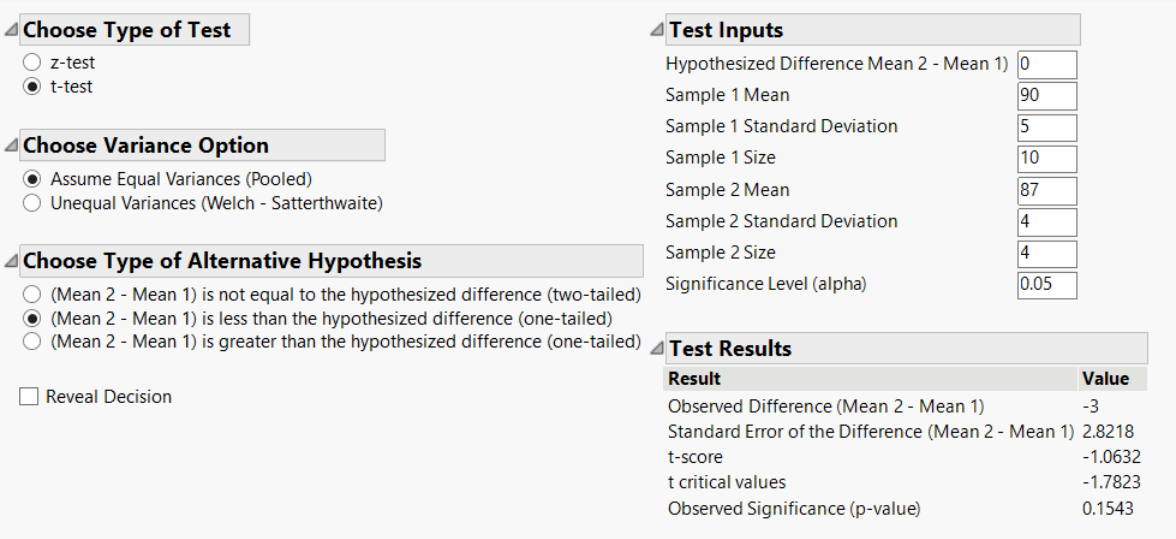
$H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 > \mu_2$

(b) When the α value is not given, we assume $\alpha = 0.05$

Significance level (p-value) = 0.1543

$0.1543 > 0.05$, $p > \alpha$

Therefore, we fail to reject the null hypothesis H_0 and can not say with 95% confidecne that the mean of weight percent calcium in standard cement is greater than the mean of weight percent calcium in cement doped with lead



3. (a) μ_0 = the mean of the diameter of the rods produced by machine 1
μ_1 = the mean of the diameter of the rods produced by machine 2

$H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 \neq \mu_2$

(b) Significance level (p-value) = 0.8184

$0.8184 > 0.05$, $p > \alpha$

Therefore, we fail to reject the null hypothesis H_0 and can not say with 95% confidecne that the mean of the diameter of the rods produced by machine 1 is different than the mean of the diameter of the rods produced by machine 2

Choose Type of Test
- ○ z-test
- ◉ t-test

Choose Variance Option
- ○ Assume Equal Variances (Pooled)
- ◉ Unequal Variances (Welch - Satterthwaite)

Choose Type of Alternative Hypothesis
- ◉ (Mean 2 - Mean 1) is not equal to the hypothesized difference (two-tailed)
- ○ (Mean 2 - Mean 1) is less than the hypothesized difference (one-tailed)
- ○ (Mean 2 - Mean 1) is greater than the hypothesized difference (one-tailed)

☐ Reveal Decision

Test Inputs

Hypothesized Difference Mean 2 - Mean 1)	0
Sample 1 Mean	8.73
Sample 1 Standard Deviation	0.59
Sample 1 Size	15
Sample 2 Mean	8.68
Sample 2 Standard Deviation	0.63
Sample 2 Size	17
Significance Level (alpha)	0.05

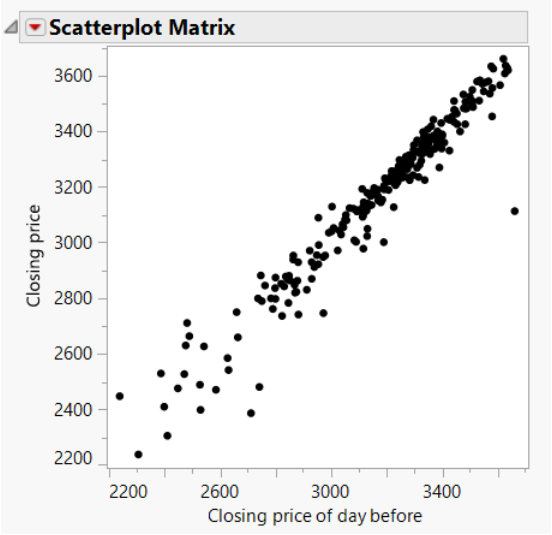
Test Results

Result	Value
Observed Difference (Mean 2 - Mean 1)	-0.05
Standard Error of the Difference (Mean 2 - Mean 1)	0.2158
t-score	-0.2317
t critical values	+/- 2.0426
Observed Significance (p-value)	0.8184

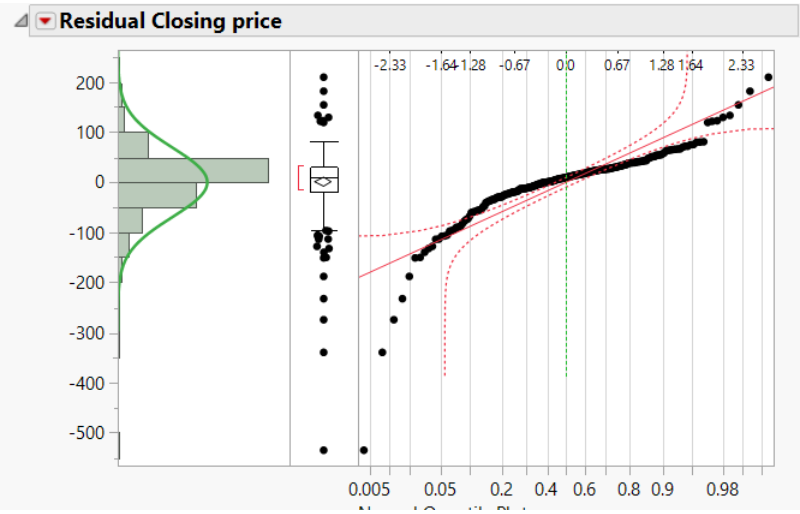
Problem 2

1. The S&P 500, or the "Standard & Poor" 500 is a stock market index that includes all the stocks of the (roughly) top 500 United States corporations. Being a list of the top 500 companies, the S&P 500 is a good representation of the overall United States stock market performance
   The S&P 500 was concieved by two financial companies, Standard and Poor, and was officially introduced on March 4th, 1957. In present day, the S&P Dow Jones Indicies own it, and they are a joint venture between the companies McGraw Hill Financial, CME Group, and News Corp
   To qualify to be in the S&P index, a company must be located in the United States, have a market cap of at least 8.2 billion US dollars, have at least 50% of its stock open to the public, have stock priced at at least 1 dollar per share, have at least 4 consecituve quarters of positive earnings, and have 50% of its assets located in the United States
   The S&P tracks the market capitalization of companies in its index, which is calculated by taking the number of shares distributed by the company and multiplying it by its stock price. The companies proportion of market cap compared to the other companies in the index will determine how much representation they recievce in the index

2. (a) Scatterplot Matrix of the sample data

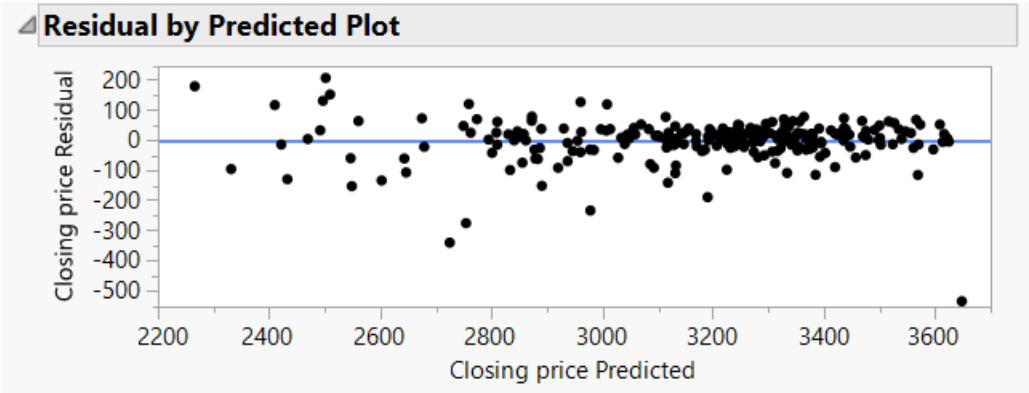


(b) Check the assumption that population distributions are normal:



Shape of the distrubution resembles the shape of a normal distribution in the histogram, therefore making it safe to assume that the populations follow a normal distribution

Check the assumption that the population distributions have the same variance:



Residual plot values average out to around 0 and the residual plot shows no increasing or decreasing relations, therefore making it safe to assume that the populations have the same variance

(c) $\hat{y} = \hat{B}_0 + \hat{B}_1 x$
B_0 = Y intercept , B_1 = scalar value, x = closing price of day before

(d) $B_0 = 96.127$, $B_1 = 0.9697$

$\hat{y} = 96.127 + 0.9697x$

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	19235813	19235813	3957.476
Error	251	1220017	4860.6263	Prob > F
C. Total	252	20455830		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	96.126826	49.13297	1.96	0.0515
Closing price of day before	0.9697208	0.015415	62.91	<.0001*

3. Closing price $= x = 3662.449951$

   Simple Linear Regression Model:
   $\hat{y} = 96.127 + 0.9697 \times 3662.449951$
   $\hat{y} = 3647.604717485$

   If the closing price on December 1st was \$3662.45, then the closing price the next day, December 2nd, would be \$3647.60

4. Yes, this does seem like an accurate model to trust when predicting the closing price of the S&P 500 index if we know the closing price of the day before. The sample data we tested was a large sample size, seemed to follow a pattern when looking at its scatter plot, followed the assumptions of ANOVA, and produced a very small p value, all which suggests the model is accurate to not only the sample data but the population data as well