HW2

Problem 1.

In a Survey of 4722 people, 15% are overweight.

Calculate and interpret a CI using 99% confidence level, for the proportion of people that are overweight.

Confidence Level of 99% =>α=0.01

$$Z_{\alpha/2} = Z_{0.01/2} = 2.576$$

$$CI = \left[\bar{X} - Z_{\alpha/2} * \sqrt{\frac{\bar{X} * (1-\bar{X})}{n}}, \bar{X} + Z_{\alpha/2} * \sqrt{\frac{\bar{X} * (1-\bar{X})}{n}}\right]$$

$$CI = \left[0.15 - 2.576 * \sqrt{\frac{0.15 * (1-0.15)}{4722}}, 0.15 + 2.576 * \sqrt{\frac{0.15 * (1-0.15)}{4722}}\right]$$

$$CI = [0.1366, 0.1634]$$

We can say with a 99% confidence that the proportion of American youngsters who are seriously overweight is between 13.66% and 16.34%

Problem 2.

Sample size of 14.

Sample mean of 8.48

Sample Stan Dev 0.79

Calculate a 95% lower confidence bound

We need to use the t-distribution with n-1 degrees of freedom.

Confidence Level of 95% =>α=0.05

Since we are asked for only the lower confidence bound, we don't need to divide alpha by 2

$$t_{\alpha,n-1} = t_{0.05,14-1} = t_{0.05,13} = 1.771$$

$$Lower\ Conf\ Level = \bar{X} - t_{\alpha,n-1} * \frac{S}{\sqrt{n}}$$

$$Lower\ Conf\ Level = 8.48 - 1.771 * \frac{0.79}{\sqrt{14}}$$

$$Lower\ Conf\ Level = 8.106$$

We assume that the population is normal distributed

Alejandro C Parra Garcia

Problem 3.

95% CI for the speed of light given 100 measurements.

The sample Average is 852.4 and the sample Sta. Dev. Is 79.01055

Confidence Level of 95% =>α=0.05

$$Z_{\alpha/2} = Z_{0.05/2} = 1.96$$

$$CI = \left[\bar{X} - Z_{\alpha/2} * \frac{S}{\sqrt{n}}, \bar{X} + Z_{\alpha/2} * \frac{S}{\sqrt{n}}\right]$$

$$CI = \left[852.4 - 1.96 * \frac{79.01055}{\sqrt{100}}, 852.4 + 1.96 * \frac{79.01055}{\sqrt{100}}\right]$$

$$CI = [836.9139, 867.8861]$$

Problem 4.

Treat A. Successful: For Small Stones 81/87=0.9310 | For Large Stones 192/263=0.730

Treat B. Successful: For Small Stones 234/270=0.8667 | For Large Stones 55/80=0.6875

The success rate for each treatment for each of the 3 cases:

	Treatment A		**Treatment B**	
Small Stones	81/87	93.10%	234/270	86.67%
Large Stones	192/263	73.00%	55/80	68.75%
Both	273/350	78.00%	289/350	82.57%

If we look at each data separately, we can conclude that treatment A is better that treatment B in small stones 93.10% over 86.67% and in large stones 73% over 68.75%. But if we look at the combined data, we can see that treatment B is better with an 82.57% over a 78% in treatment A.

So, if we only looked at the combined data, we conclude that B is better than treatment A. But that will be a mistake since A perform better in each test individually. So how does this make sense?

As in the table above we can calculate the combine success rate by just adding the successful outcomes of each test and divide them by the added total test for each type of test. But there is another way to think about it.

The success rate for the combine stones is a weighted average of the results of each individual test. So, we can see it in action:

For treatment A:

$$\frac{0.9310 * 87 + 0.73 * 263}{87 + 263} = \frac{272.987}{350} = 0.77996 \sim 0.78 => 78.00\%$$

For treatment B:

$$\frac{0.8667 * 270 + 0.6875 * 80}{270 + 80} = \frac{289.009}{350} = 0.82574 \sim 0.8257 => 82.57\%$$

This is important since the number of tests done in each case is different for each treatment, this means that the importance of each successful test in the overall score for the combine test is different for each treatment.

In treatment A, the test for large stones was conducted 263 times while the test for small stones only 87 times, so in the success score for the combine results the large stones are going to have more relevance $\left[\frac{263}{263+87} = 75.14\%\right]$ this indicates that 75.14% of the combine score comes from the large stones. If we look at the success rate for each time of stones for treatment A, we can see that for large stones is 73% while for small stones is 93.1%. So, the weighted average get pulls closer to the 73%.

In case of the treatment B, the opposite occurs, there was 270 tests for small stones and 80 for large stones. So, the small stones test is going to have more importance $\left[\frac{270}{270+80} = 77.14\%\right]$ this indicates that 77.14% of the combine score comes from the small stone's tests. As before, if we now check the scores for each time of test for treatment B, we can see that for the small stone test the score is 86.67% while for the large stone test the score is 68.75%. So, the weighted average gets pull closer to the 86.67%.

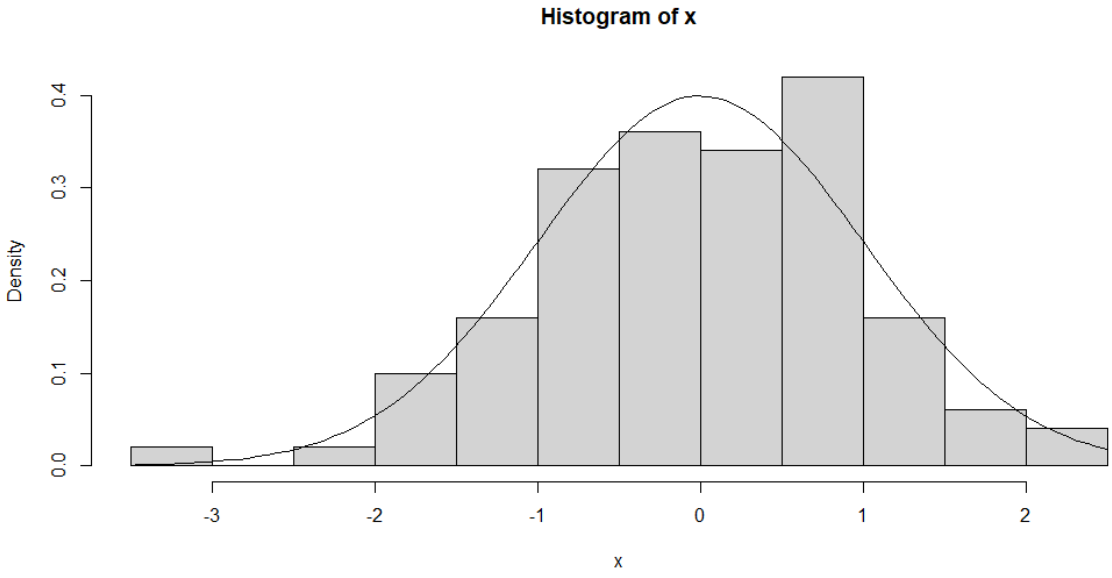
That's why the combine success rate is higher for treatment B. Even though treatment A performs better in each individual test.

Problem 5.

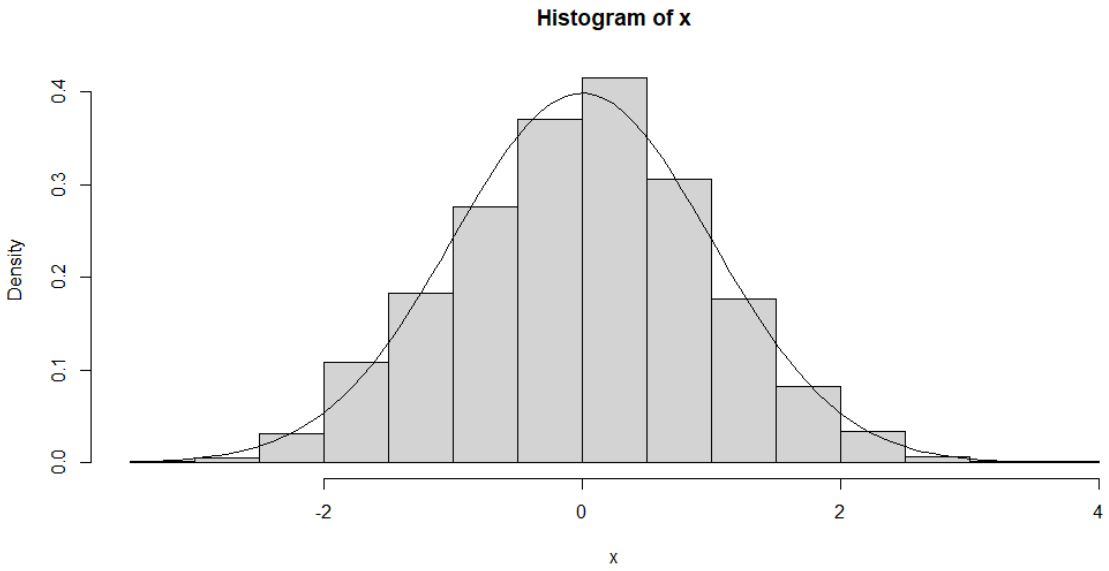
a)

b)

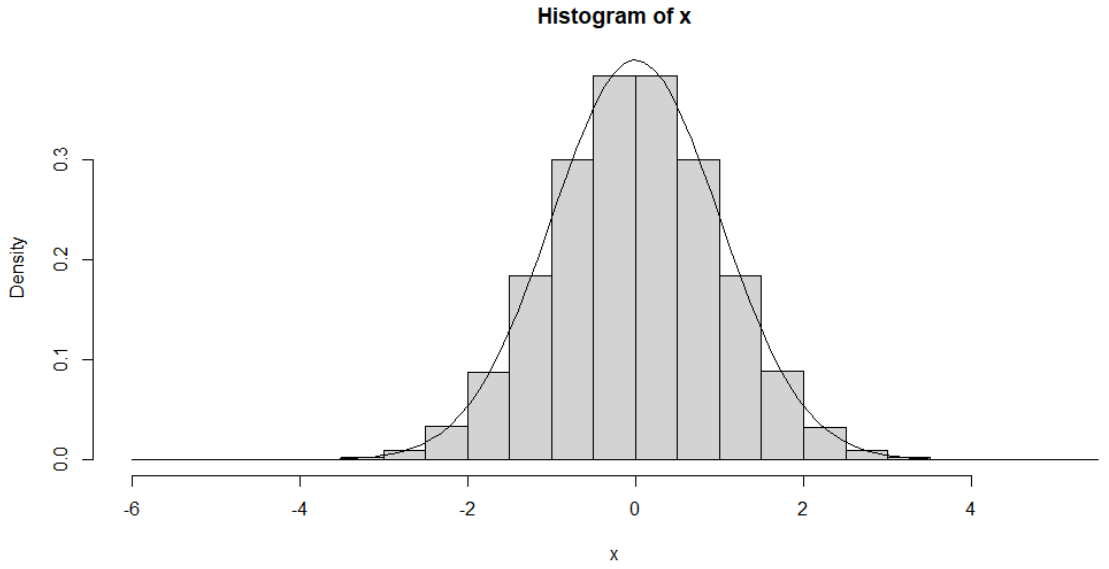
n=100



n=2,000


Histogram of x
Density
0.0
0.1
0.2
0.3
0.4
-2
0
2
4
x

n=5,000,000


Histogram of x
Density
0.0
0.1
0.2
0.3
-6
-4
-2
0
2
4
x