Kush Parmar, Jo-Anne Rivera, Simerjeet Mudhar

I pledge my honor that I have abided by the stevens honor system- Kush, Jo-Anne, Simmerjeet

A. Based on the MPG data (MPG1sample at Canvas) of Exercise 6.30, work out the following confidence intervals at the confidence level 0.95.

1. Assume the observations are from a normal population. Identify the population random variable X.

The random variable X in this scenario is the a random mpg from the given sample population.

2. The confidence interval of the population mean in the context that the population standard deviation is 3.5.

```
> 43.17-((1.96 * 3.5)/(sqrt(20)))
[1] 41.63606
> 43.17+((1.96 * 3.5)/(sqrt(20)))
[1] 44.70394
```

[41.63606, 44.70394]

3. The confidence interval of the population mean in the context of unknown population standard deviation.

```
t_lower=mean(x)-qt(1-0.05/2,length(x)-1)*(sd(x))/sqrt(length(x))
t_higher=mean(x)-qt(0.05/2,length(x)-1)*(sd(x))/sqrt(length(x))
t_interval=c(t_lower,t_higher)
t_interval
```

```
[1] 41.10375 45.23625
```

[41.1038,45.237]

4. Compare margin of errors of the above two intervals, identify the one which is more accurate, and interpret your finding.

```
> z_error= (qnorm(1-0.05/2)*3.5)/sqrt(length(x))
> t_error= (qt(1-0.05/2,length(x)-1)*(sd(x)))/sqrt(length(x))
> z_error
[1] 1.533914
> t_error
[1] 2.066255
```

We can see that the margin of error on the t test is greater than the margin of error on the z test. In the z test we had our standard deviation, in the t test we did not. The results are to be expected because the z test is typically more reliable than the t test. The numbers that we got indicate that we might be off 1.534 or 2.066 mpg depending on what test at a confidence level of 95%.

5. The confidence interval of the population variance.

$\left(\frac{\sqrt{(n-1)}}{\sqrt{b}}s \leq \sigma \leq \frac{\sqrt{(n-1)}}{\sqrt{a}}s\right)$, where $a = X^{2}_{1-\alpha/2, n-1}$ $and\ b = x^{2}_{\alpha/2, n-1}$

$a = X^{2}_{0.975,\,19}\ and\ b = X^{2}_{0.025,\,19},\ a = 8.907,\ b = 32.852,\ s = 19.49168$

$$\left(\frac{\sqrt{19}(19.49)}{\sqrt{32.852}} \leq \sigma \leq \frac{\sqrt{19}(19.49)}{\sqrt{8.907}}\right)$$

```
> (sqrt(19)*var(x))/(sqrt(32.852))
[1] 14.82331
> (sqrt(19)*var(x))/(sqrt(8.907))
[1] 28.46823
```

[14.82331, 28.46823]

B. Transform the MPG data into a binary data through comparing whether the observation goes above 41. Set the confidence level 0.90.

1. Identify the Bernoulli population random variable Y corresponding to the binary data.

The Bernoulli population random variable Y is either 1 if the observation is above 41, or 0 if the observation is under 41.

2. Construct the confidence interval of the population proportion based on the first 15 observations.

```
> phat + qnorm(1-alpha/2)*sqrt(phat*(1-phat))/sqrt(n)
[1] 0.9698798
> phat + qnorm(alpha/2)*sqrt(phat*(1-phat))/sqrt(n)
[1] 0.6301202
```

[0.6301202 , 0.9698798]

3. Construct the confidence interval of the population proportion based on all 20 observations.

```
> na = length(a)
> la = phata + qnorm(alpha/2)* sqrt(phata*(1-phata))/sqrt(na)
> la
[1] 0.5907377
> nb = length(b)
> ra = phata + qnorm(1-alpha/2)* sqrt(phata*(1-phata))/sqrt(na)
> ra
[1] 0.9092623
```

4. Report margin of errors of the above two intervals, identify the one which is more accurate, and interpret your finding.
5.

```
y=ifelse(new_x>41.0,1,0)
phat= mean(y)
z=ifelse(x>41.0,1,0)
phat_1=mean(z)

error_2=+qnorm(1-0.1/2)*sqrt(phat*(1-phat))/sqrt(15)
error_2


error_3=qnorm(1-0.1/2)*sqrt(phat_1*(1-phat_1))/sqrt(20)
error_3
```

```
> error_2
[1] 0.1698798
> error_3
[1] 0.1592623
```

The margin of error from our findings come from the fact that the sample sizes are different. The interval with the higher amount of sample size is more trust worthy as there are more degrees of freedom. Which in this case, error_3 is more accurate as the error margin is lower.