lab4.2 - Confidence Intervals - Nhi Nguyen

Background Information:

```
load('yrbss.rda')
```

Calculating confidence intervals: 1.

```
yrbss.complete = yrbss[complete.cases(yrbss$weight), ]
sample.size = 30
set.seed(5011)

sample.rows = sample(1:nrow(yrbss.complete), sample.size)

yrbss.sample = yrbss.complete[sample.rows, ]

summary(yrbss.sample$weight)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   46.72   56.70   67.36   69.63   72.58  113.40
```

QUES a:

```
mean.sample = mean(yrbss.sample$weight) #calculate the mean of sample
sd.sample = sd(yrbss.sample$weight) #calculate the standard deviation of the sample
```

QUES b:

```
z = qnorm(0.975,)
z
```

```
## [1] 1.959964
```

QUES c: margin error of x:

```
margin.error = z * sd.sample / sqrt(sample.size)
margin.error
```

```
## [1] 6.535656
```

```
margin.left = mean.sample - margin.error
margin.right = mean.sample + margin.error

cat("(", margin.left,",", margin.right,")")
```

```
## ( 63.09234 , 76.16366 )
```

QUES d: when the sd value is getting bigger, the confidence interval is getting larger

2. QUES a:

```
z.90 = qnorm(1-(1-0.9)/2,)

margin.error = z.90 * sd.sample / sqrt(sample.size)
margin.error
```

```
## [1] 5.484895
```

```
margin.left = mean.sample - margin.error
margin.right = mean.sample + margin.error

cat("(", margin.left,",", margin.right,")")
```

```
## ( 64.1431 , 75.1129 )
```

QUES b:

```
z.99 = qnorm(1-(1-0.99)/2,)

margin.error = z.99 * sd.sample / sqrt(sample.size)
margin.error
```

```
## [1] 8.589307
```

```
margin.left = mean.sample - margin.error
margin.right = mean.sample + margin.error

cat("(", margin.left,",", margin.right,")")
```

```
## ( 61.03869 , 78.21731 )
```

QUES c: 90% CI: (64.1431 , 75.1129) 95% CI: (63.09234 , 76.16366) 99% CI: (61.03869 , 78.21731) 90% CI is the biggest interval, followed by 95% CI then 99% CI.

QUES d: the larger the confidence level is, the bigger the confidence interval is respectively. however, when the CL reach the absolute value (100%), the CI varies with every value because it doesnt have meaningful value for prediction. CI with CL 90% is the most informative.

3. QUES a:

```
yrbss.complete = yrbss[complete.cases(yrbss$weight), ]
sample.size = 30
set.seed(5011)

sample.rows = sample(1:nrow(yrbss.complete), sample.size)

yrbss.sample = yrbss.complete[sample.rows, ]

t.test(yrbss.sample$weight, conf.level = 0.95) $conf.int
```

```
## [1] 62.80802 76.44798
## attr(,"conf.level")
## [1] 0.95
```

QUES b:

```
yrbss.complete = yrbss[complete.cases(yrbss$weight), ]
sample.size = 50
set.seed(5011)

sample.rows = sample(1:nrow(yrbss.complete), sample.size)

yrbss.sample = yrbss.complete[sample.rows, ]

t.test(yrbss.sample$weight, conf.level = 0.95) $conf.int
```

```
## [1] 63.59814 74.09826
## attr(,"conf.level")
## [1] 0.95
```

```
yrbss.complete = yrbss[complete.cases(yrbss$weight), ]
sample.size = 100
set.seed(5011)

sample.rows = sample(1:nrow(yrbss.complete), sample.size)

yrbss.sample = yrbss.complete[sample.rows, ]

t.test(yrbss.sample$weight, conf.level = 0.95) $conf.int
```

```
## [1] 63.23826 70.25754
## attr(,"conf.level")
## [1] 0.95
```

```
yrbss.complete = yrbss[complete.cases(yrbss$weight), ]
sample.size = 300
set.seed(5011)

sample.rows = sample(1:nrow(yrbss.complete), sample.size)

yrbss.sample = yrbss.complete[sample.rows, ]

t.test(yrbss.sample$weight, conf.level = 0.95) $conf.int
```

```
## [1] 66.89920 71.18906
## attr(,"conf.level")
## [1] 0.95
```