Chapter 6

6.1

Explain the meaning of a .95 confidence interval.

A confidence interval is a statistical measure of the probability coverage a given lies within some interval based upon the observed data.

6.2

If the goal is to compute a .80, or .92, or a .98 confidence interval for μ when σ is known, and sampling is from a normal distribution, what values for c should be used in Equation (6.4)?

```
pretty_kable(
   data.table(
      Confidence = c(.8, .92, .98) )[,
         Val := (1 - (1 - Confidence)/2)][, C := qnorm(Val)] ,
   "Confidence Values")
```

Table 1: Confidence Values

Confidence	Val	C
0.80	0.90	1.28
0.92	0.96	1.75
0.98	0.99	2.33

6.3

```
conf <- function(alpha) {
   c( Lower = (1 - alpha)/2, Upper = 1 - (1 - alpha)/2)
}
```

Assuming random sampling is from a normal distribution with standard deviation $\sigma = 5$, if the sample mean is $\bar{X} = 45$ based on $n = 25$ participants, what is the 0.95 confidence interval for μ?

```
qnorm(conf(.95), mean = 45, sd = 5/sqrt(25))
```

```
   Lower    Upper
43.04004 46.95996
```

6.4

Repeat the previous example, only compute a .99 confidence interval instead.

```r
qnorm(conf(.99), mean = 45, sd = 5/sqrt(25))
```

```
   Lower    Upper
42.42417 47.57583
```

6.5

A manufacturer claims that their light bulbs have an average life span that follows a normal distribution with $\mu = 1,200$ hours and a standard deviation of $\sigma = 25$. If you randomly test 36 light bulbs and find that their average life span is $\bar{X} = 1,150$, does a .95 confidence interval for mu suggest that the claim $mu = 1,200$ is reasonable?

```r
qnorm(conf(.95), mean = 1150, sd = 25/sqrt(36))
```

```
   Lower    Upper
1141.833 1158.167
```

No, 1,200 is outside the bounds of a 95% confidence interval.

6.6

Compute a .95 confidence interval for the mean, assuming normality, for the following situations:

a.) $n = 12, \sigma = 22, \bar{X} = 65$

```r
qnorm(conf(.95), mean = 65, sd = 22/sqrt(12))
```

```
   Lower    Upper
52.55256 77.44744
```

b.) $n = 22, \sigma = 10, \bar{X} = 185$

```r
qnorm(conf(.95), mean = 185, sd = 10/sqrt(22))
```

```
   Lower    Upper
180.8213 189.1787
```

c.) $n = 50, \sigma = 30, \bar{X} = 19$

```r
qnorm(conf(.95), mean = 19, sd = 30/sqrt(50))
```

```
   Lower    Upper
10.68458 27.31542
```

6.7

What happens to the length of a confidence interval for the mean of a normal distribution when the sample size is doubled? In particular, what is the ratio of the lengths?

What is the ratio of the lengths if the sample size is quadrupled?

Answer:

In general, For some *x*,

$CI = \mu \pm C\frac{\sigma}{x\sqrt{n}}$

As *x* increases, the size of the confidence intervals decrease (closer approximations).

6.8

The length of a bolt made by a machine parts company has a normal distribution with standard deviation σ equal to 0.01 mm. The lengths of four randomly selected bolts are as follows:

20.01, 19.88, 20.00, 19.99

a.) Compute a 0.95 confidence interval for the mean.

```
x <- c(20.01, 19.88, 20.00, 19.99)
n <- length(x)
xbar <- mean(x)

qnorm(conf(.95), mean = xbar, sd = 0.01/sqrt(n))
```

```
  Lower   Upper
19.9602 19.9798
```

b.) Specifications requre a mean lengths μ of 20.00 mm for the population of bolts. Do the data indicate that this specification is being met?

No, the 95% confidence interval for the mean does not include 20.

c.) Given that the 0.95 confidence interval contains the value 20, why might it be inappropriate to conclude that the specification is being met?

Just because the confidence interval includes some value, that does not mean that the population paramter is that value.

6.9

The weight of trout sold at a trout farm has a standard deviation of 0.25. Based on a sample of 10 trout, the average weight is 2.10 lbs.

Assume normality and Compute a .99 confidence interval for the mean.

```
qnorm(conf(.99), mean = 2.10, sd = .25/sqrt(10))
```

```
   Lower    Upper 
1.896363 2.303637 
```

6.10

The average bounce of 45 randomly selected tennis balls is found to be $\bar{X} = 1.70$.

Assuming that the standard deviation of the bounce is .30, compute a 0.90 confidence interval for the average bounce assuming normality.

```
qnorm(conf(.90), mean = 1.70, sd = .30/sqrt(45))
```

```
  Lower   Upper 
1.62644 1.77356 
```

6.11

Assuming that the degrees of freedom are 20, find the value t for which:

a.) $P(T \leq t) = 0.995$

```
qt(.995, df = 20)
```

```
[1] 2.84534
```

b.) $P(T \geq t) = 0.025$

```
1 - qt(.025, df = 20)
```

```
[1] 3.085963
```

c.) $P(-t \leq T \leq t) = 0.90$

```
qt( (1 + .90) / 2, df = 20)
```

```
[1] 1.724718
```

6.12

Compute a 0.95 confidence interval if

a.) $n = 10, \bar{X} = 26, s = 9$

```
26 + qt(conf(.95), df = 9) * 9/sqrt(10)
```

```
   Lower    Upper
19.56179 32.43821
```

b.) $n = 18, \bar{X} = 132, s = 20$

```
132 + qt(conf(.95), df = 17) * 20/sqrt(18)
```

```
   Lower    Upper
122.0542 141.9458
```

c.) $n = 25, \bar{X} = 52, s = 12$

```
52 + qt(conf(.95), df = 24) * 12/sqrt(25)
```

```
   Lower    Upper
47.04664 56.95336
```

6.13

Compute a 0.99 confidence interval if

a.) $n = 10, \bar{X} = 26, s = 9$

```
26 + qt(conf(.99), df = 9) * 9/sqrt(10)
```

```
   Lower    Upper
16.75081 35.24919
```

b.) $n = 18, \bar{X} = 132, s = 20$

```
132 + qt(conf(.99), df = 17) * 20/sqrt(18)
```

```
   Lower    Upper
118.3376 145.6624
```

c.) $n = 25, \bar{X} = 52, s = 12$

```
52 + qt(conf(.99), df = 24) * 12/sqrt(25)
```

```
   Lower    Upper
45.28735 58.71265
```

6.14

For a study on self-awareness, the observed values for one of the groups were:

77, 87, 88, 114, 151, 210, 219, 246, 253, 262, 296, 299, 306, 376, 428, 515, 666, 1310, 2611.

Compute a .95 confidence interval for the mean assuming normality.

```
x <- c(77,87,88,114,151,210,219,246,253,262,296,299,306,376,428,515, 666,1310,2611)
xbar <- mean(x)
n <- length(x)

xbar + qt(conf(.95), df = n - 1) * sd(x)/sqrt(n)
```

```
   Lower    Upper
161.5030 734.7075
```

Compare the result to the bootstrap-t confidence interval.

```
winsorize <- function( x, trim = .2) {
   n <- length(x)

   o <- sort(x)
   g <- floor(trim*n)

   o[1:(g+1)] <- o[(g+1)]
   o[(n-g):n] <- o[n-g]

   o
}

trimse<-function(x,tr=.2,na.rm=FALSE){
   if(na.rm)x<-x[!is.na(x)]
   trimse <- sqrt(var(winsorize(x,tr)))/((1-2*tr)*sqrt(length(x)))
   trimse
}

nboot <- 2000
alpha <- .95

sims <- matrix(sample(x, size = x*nboot, replace =T), nrow = nboot)
data <- sims - xbar

top <- apply(data, 1, mean)
bot <- apply(data, 1, trimse )

tval <- top/bot
tval <- sort(tval)

icrit <- round((1-alpha)*nboot)

c(Lower = mean(x)+tval[icrit]*trimse(x),
   Upper = mean(x)-tval[icrit]*trimse(x))
```

```
   Lower    Upper
151.0749 745.1356
```

Why do they differ?

6.15

Rats are subjected to a drug that might affect aggression. Suppose that for a random sample of rats, measures of aggression are found to be:

5, 12, 23, 24, 18, 9, 18, 11, 36, 15

Compute a 0.95 confidence interval for the mean assuming the scores are from a normal distribution.

6.16

Suppose $M = 34$ and the McKean-Schrader estimate of the standard error of M is $S_M = 3$.

Compute a .95 confidence interval for the population median.

6.17

6.18

6.19

6.20

6.21

6.22

6.23

6.24

6.25

6.26

6.27

6.28

6.29

6.30

6.31

6.32

6.33

6.34

6.35

6.36

6.37

6.38