M440B HW 4

Erick Castillo

2/13/2021

Problem 39

Using the information provided on page 284, all instanced with $\hat{\theta} - \theta_0$ can be replaced with $\frac{\hat{\theta}}{\theta_0}$ to get the desired answer.
Then the following equalities are $P(\frac{\hat{\theta}}{\theta_0} \leq \underline{\delta}) = \frac{\alpha}{2}$ and $P(\frac{\hat{\theta}}{\theta} \leq \bar{\delta}) = 1 - \frac{\alpha}{2}$.
Putting both of these together results in $P(\underline{\delta} \leq \frac{\hat{\theta}}{\theta_0} \leq \bar{\delta}) = 1 - \alpha$.
Solving for θ_0 results in $P(\frac{\hat{\theta}}{\bar{\delta}} \leq \theta_0 \leq \frac{\hat{\theta}}{\underline{\delta}}) = 1 - \alpha$.
This means that the bootstrap interval for $\frac{\hat{\theta}}{\theta_0}$ has an interval:

$$(\frac{\hat{\theta}}{\bar{\delta}}, \frac{\hat{\theta}}{\underline{\delta}})$$

.

Problem 40

- The probability $P(|\hat{\theta} - \theta_0| > 0.01)$ can be rewritten to consider the bootstrapping process. In this case, it would be the proportion of bootstrap samples for which the bootstrap estimator has a distance from the actual estimate that is greater than 0.01. As an inequality it would be the proportion of bootstrap samples where $|\theta_j^* - \hat{\theta}| > 0.01$ for all $j = 1, ..., B$.

- From office hours I am aware that $E(|\hat{\theta} - \theta_0|)$ is the average error. The error inside the absolute value can be rewritten to express the distance between the bootstrap estimator from the actual estimate. It can be visualized as follows
$$\frac{1}{B}\sum_{j=1}^{B}|\theta_j^* - \hat{\theta}|$$

- The given probability $P(|\hat{\theta} - \theta_0| > \Delta) = 0.5$ can once again be rewritten to express the bootstrapping process. In the bootstrapping process it would read: the proportion where the distance between the bootstrap estimator and the actual estimate is greater than Δ, where the unknown value Δ gives a gives a proportion of 0.5. A more succinct way to write the above would be, the proportion of bootstrap samples is 0.5 whenever $|\theta_j^* - \hat{\theta}| > \Delta$ for all $j = 1, ..., B$ and some Δ.

Bootstrapping HW

Part 1:

```
#create a vector of the data:
heartrates = c(160,184,173,176,168,156,160,160,174,166)
```

Part 2:

```
#determine the sample mean and the variance:
ybar = mean(heartrates)
s.2 = var(heartrates)
ybar
```

```
## [1] 167.7
```

```
s.2
```

```
## [1] 80.01111
```

Part 3:

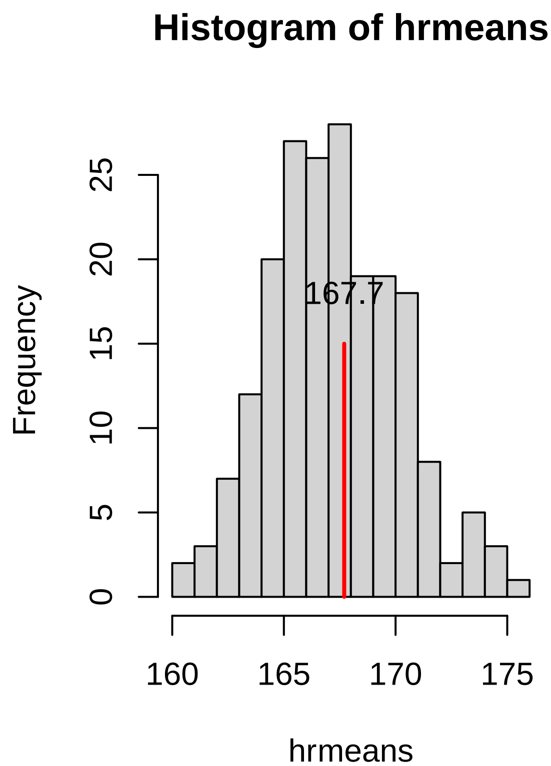
```
#create 200 bootstrap samples and compute the mean* and var* of each one.
hrmeans = numeric(200)
hrvars = numeric(200)
for (i in 1:200){
  this.samp = heartrates[sample(10,10,replace=TRUE)]
  hrmeans[i] = mean(this.samp)
  hrvars[i] = var(this.samp)
}
```

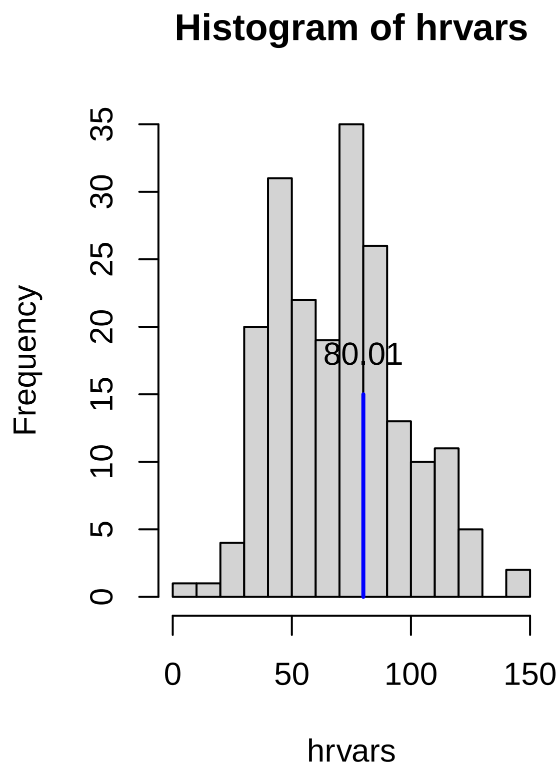
Part 4:

```
#histogram of the 200 bootstrap sample means and variances:

par(mfrow=c(1,2))
hist(hrmeans,nclass=15)
lines(c(ybar,ybar), c(0,15), col = "red", lwd = 2)
text(ybar, 18 , round(ybar, 2))


hist(hrvars,nclass=15)
lines(c(s.2,s.2), c(0,15), col = "blue", lwd = 2)
text(s.2, 18 , round(s.2, 2))
```





Part 5:

```
#Obtain and sort 200 values of ybar*-ybar
bsmeanpivq <- hrmeans-167.7
sorted.means = sort(bsmeanpivq)

#finding a and b given after sorting
a1 = (sorted.means[5]+sorted.means[6])/2
b1 = (sorted.means[195]+sorted.means[194])/2
a1 #value of a
```

```
## [1] -5.45
```

```
b1 #value of b
```

```
## [1] 5.8
```

```
#the confidence interval would be (ybar-b1, ybar-a1):
lower1 = ybar-b1
upper1 = ybar-a1

c(lower1, upper1) #the confidence interval
```

```
## [1] 161.90 173.15
```

Part 6:

```
#find the 95% bootstrap confidence
bsvarpivq<-hrvars/80.0111
sorted.vars = sort(bsvarpivq)

#The above suggests that second smallest value is the 0.025 quantile and 19th largest is the 0.975 quan
a2 = (sorted.vars[5]+sorted.vars[6])/2
b2 = (sorted.vars[195]+sorted.vars[194])/2
a2 #value of a
```

```
## [1] 0.3319678
```

```
b2 #value of b
```

```
## [1] 1.53812
```

```
lower2 = s.2/b2
upper2 = s.2/a2

c(lower2,upper2) #the 95% confidence interval
```

```
## [1]  52.01877 241.02068
```

Part 7:

```
#classical approach to find the CI for sigma^2
orig.upper = sum((heartrates-mean(heartrates))^2)/qchisq(0.025, df=9)
orig.lower = sum((heartrates-mean(heartrates))^2)/qchisq(0.975, df=9)
c(orig.lower,orig.upper) # the confidence interval for s^2
```

```
## [1]  37.85464 266.66523
```

```
#classical approach to find the CI for mu
t.test(heartrates,conf.level=.95)
```

```
##
##  One Sample t-test
##
## data:  heartrates
## t = 59.287, df = 9, p-value = 5.568e-13
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  161.3012 174.0988
## sample estimates:
## mean of x
##     167.7
```