Homework 9

11/19/2021

Question 1 (8.1)

Compute a jackknife estimate of the bias and the standard error of the correlation statistic in Example 8.2.

```
n <- nrow(law82)
theta.hat <- cor(law82$LSAT,law82$GPA)

theta.jack <- numeric(n)
for(i in 1:n) {
  theta.jack[i] <- cor(law82$LSAT[-i],law82$GPA[-i])
}

bias <- (n-1)*(mean(theta.jack)-theta.hat)
se <- sqrt((n-1)*mean((theta.jack-mean(theta.jack))^2))
```

-Ans: Jackknife estimate for bias is: -0.0029386. Jackknife estimate for standard error is: 0.0533478.

Question 2 (8.3)

Obtain a bootstrap t confidence interval estimate for the correlation statistic in example 8.2 (law data in bootstrap).

```
r <- function(x, i){
  cor(x[i,1],x[i,2])
}

boot.t.ci <- function(x, B=500, R=100, level=0.95, statistic){
  x <- as.matrix(x)
  n <- nrow(x)
  stat <- numeric(B)
  se <- numeric(B)
  boot.se <- function(x, R, f){
    x <- as.matrix(x)
    m <- nrow(x)
    th <- replicate(R, expr={
      i <- sample(1:m, size=m, replace=T)
      f(x[i,])
    })
    return(sd(th))
  }
```

```
  for(b in 1:B){
    j <- sample(1:n, size=n, replace=T)
    y <- x[j,]
    stat[b] <- statistic(y)
    se[b] <- boot.se(y,R = R, f=statistic)
  }
  stat0 <- statistic(x)
  t.stats <- (stat-stat0)/se
  se0 <- sd(stat)
  alpha <- 1-level
  Qt <- quantile(t.stats, c(alpha/2,1-alpha/2),type=1)
  names(Qt) <- rev(names(Qt))
  CI <- rev(stat0-Qt*se0)
}

ci <- boot.t.ci(x = law82,statistic=r, B=2000, R=200)
```

-Ans: The 95% Bootstrap confidence interval for the correlation statistic is: (-0.2531696, 0.1624903).

Question 3

Efron and Tibshirani discuss the $scor(bootstrap)$ test score data on 88 students who took examinations in five subjects [91, Table 7.1], [194, Table 1.2.1]. The first two tests (mechanics, vectors) were closed book and the last three tests (algebra, analysis, statistics) were open book. Each row of the data frame is a set of scores $(x_{i1}, ..., x_{i5})$ for the ith student. Obtain bootstrap estimates of the standard errors for each of the following correlation estimates: $\hat{p}_{12} = \hat{p}(mec, vec)$, $\hat{p}_{34} = \hat{p}(alg, ana)$, $\hat{p}_{35} = \hat{p}(alg, sta)$, $\hat{p}_{45} = \hat{p}(ana, sta)$.

```
r <- function(x, i){
  cor(x[i,1],x[i,2])
}

#1&2
df12 <- cbind(scor$mec, scor$vec)
obj12 <- boot(df12, r, 2000)
se12 <- sd(obj12$t)

#3&4
df34 <- cbind(scor$alg, scor$ana)
obj34 <- boot(df34, r, 2000)
se34 <- sd(obj34$t)

#3&5
df35 <- cbind(scor$alg, scor$sta)
obj35 <- boot(df35, r, 2000)
se35 <- sd(obj35$t)

#4&5
df45 <- cbind(scor$ana, scor$sta)
obj45 <- boot(df45, r, 2000)
se45 <- sd(obj45$t)
```

-Ans:

$\hat{p}$	se estimate
$\hat{p}_{12}$	0.0738261
$\hat{p}_{34}$	0.048408
$\hat{p}_{35}$	0.0609803
$\hat{p}_{45}$	0.0673989

Question 4

Obtain a 95% standard normal bootstrap confidence interval, a 95% basic bootstrap confidence interval, and a percentile confidence interval for the $\hat{p}_{12}$ in Question 3.

```
boot.ci(obj12, type=c("basic","norm","perc"))
```

```
## BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
## Based on 2000 bootstrap replicates
##
## CALL :
## boot.ci(boot.out = obj12, type = c("basic", "norm", "perc"))
##
## Intervals :
## Level      Normal              Basic              Percentile
## 95%   ( 0.4108,  0.7002 )   ( 0.4269,  0.7210 )   ( 0.3858,  0.6799 )
## Calculations and Intervals on Original Scale
```