Day13

Olivia Wu

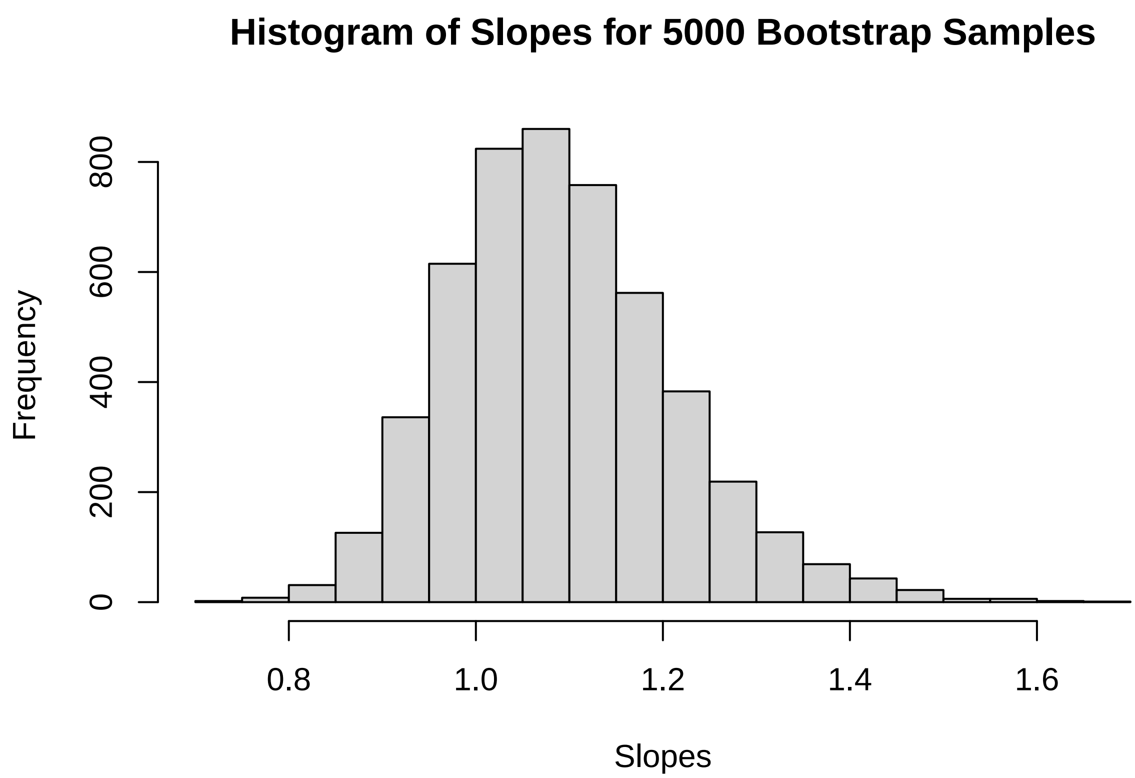
2024-04-15

Problem 4.20

a) We are 90% confident that the true slope of the population regression lies in the interval (0.914, 1.240)

```
##
## Call:
## lm(formula = Length ~ Time, data = HP)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4112 -1.1636 -0.0413  1.0514  3.7743
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.10039    1.06739   1.031    0.308
## Time         1.07711    0.09699  11.105 2.39e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.818 on 44 degrees of freedom
## Multiple R-squared:  0.737,  Adjusted R-squared:  0.7311
## F-statistic: 123.3 on 1 and 44 DF,  p-value: 2.39e-14
```

b) The histogram is slightly skewed right. The slopes tend to be around 1.1.



c) The mean is about 1.09, and the standard deviation is around 0.12. They are both just slightly above the numbers in our computer output of the original model.

```
pt <- mean(boot$Time)
se <- sd(boot$Time)
c(pt,se)
```

```
## [1] 1.0930313 0.1205294
```

d) pt estimate is the initial value

```
sum <- summary(model1)
paste("(",pt-qnorm(0.95)*se,",",pt+qnorm(0.95)*se,")")
```

```
## [1] "( 0.894778066329618 , 1.29128453751388 )"
```

```
paste("(",sum$coefficients[2,1]-qnorm(0.95)*se,",",sum$coefficients[2,1]+qnorm(0.95)*se,")")
```

```
## [1] "( 0.878853036831474 , 1.27535950801574 )"
```

e)

```
lower <- quantile(boot$Time, 0.05)
upper <- quantile(boot$Time, 0.95)
paste("(",lower,",",upper,")")
```

```
## [1] "( 0.918936258904322 , 1.30595285020395 )"
```

f)

```
newlower <- 2*1.07711 - upper
newupper <- 2*1.07711 - lower
paste("(",newlower,",",newupper,")")
```

```
## [1] "( 0.848267149796048 , 1.23528374109568 )"
```

g) The intervals are all very similar.

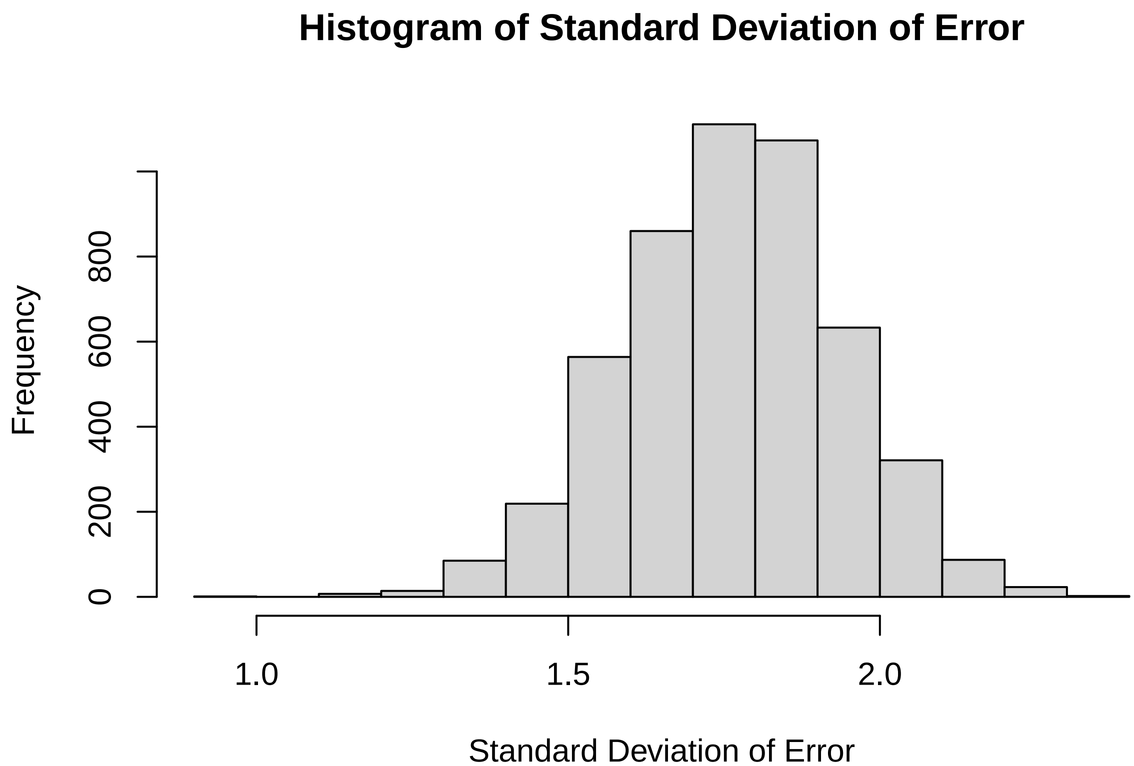
Problem 4.21

Save initial model:

```
model1 <- lm(Length ~ Time, data=HP)
```

Save initial standard deviation of the error term:

```
initial <- summary(model1)$sigma
```



Simulate:

Interval 1:

```
pt <- mean(boot$sigma)
se <- sd(boot$sigma)
paste("(",initial+qnorm(0.05)*se,",",initial+qnorm(0.95)*se,")")
```

```
## [1] "( 1.53302123748635 , 2.10352118266568 )"
```

Interval 2:

```
lower <- quantile(boot$sigma, 0.05)
upper <- quantile(boot$sigma,0.95)
paste("(",lower,",",upper,")")
```

```
## [1] "( 1.47840102433439 , 2.04300885208939 )"
```

Interval 3:

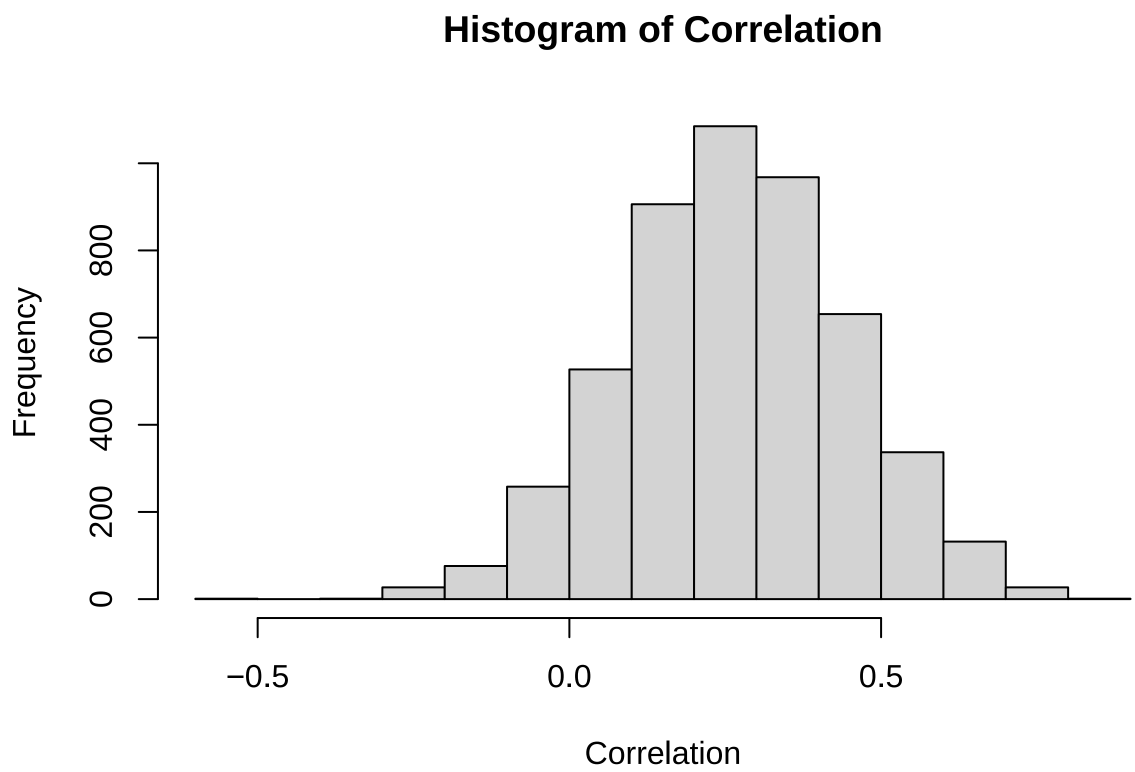
```
paste("(",2*initial - upper,",",2*initial-lower,")")
```

```
## [1] "( 1.59353356806264 , 2.15814139581763 )"
```

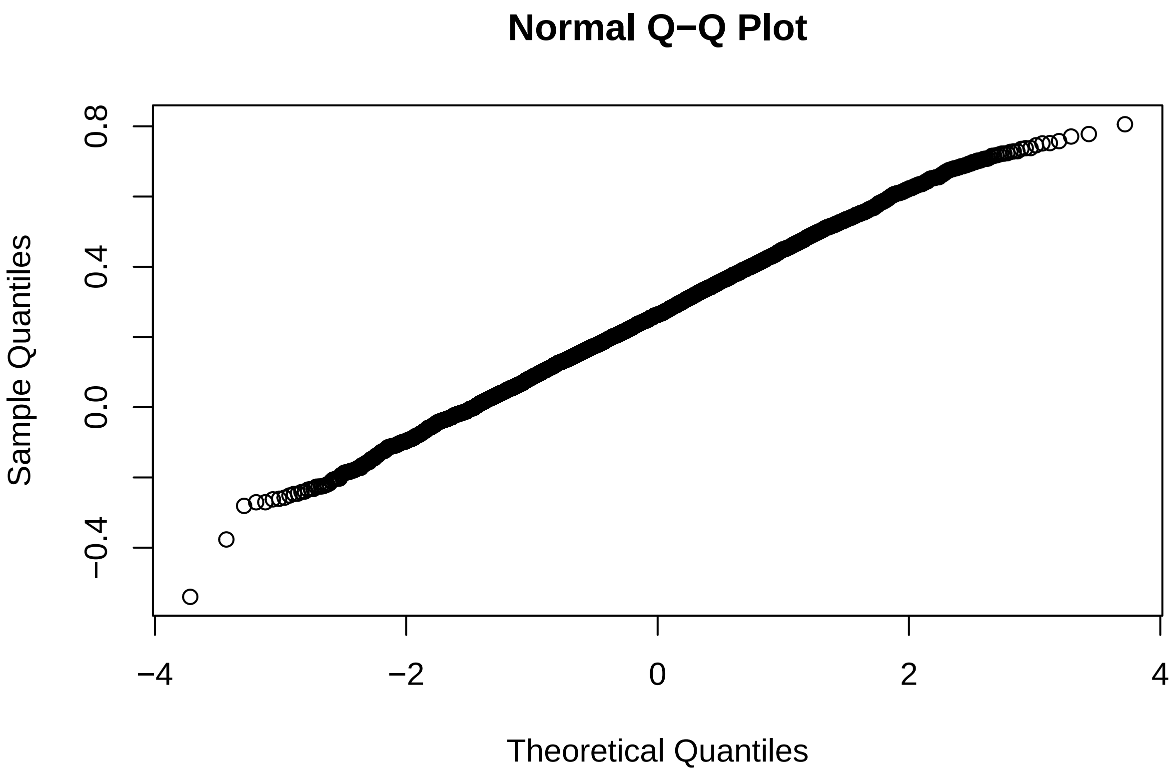
Problem 4.23

a) The initial correlation is $r=$0.244. The histogram is roughly symmetric, and the normal probability plot is roughly linear, suggesting that the boostrap distribution is normal.

```r
initial <- cor(VerbalSAT~GPA, data=SAT)
boot <- do(5000)*cor(VerbalSAT~GPA, data=resample(SAT))
hist(boot$cor, xlab="Correlation", main="Histogram of Correlation")
```




```r
qqnorm(boot$cor)
```



b) The two intervals are (-0.107, 0.596) and (-0.090, 0.615). They are similar.

```
## [1] "( -0.107180497117106 , 0.596089013149533 )"
```

```
## [1] "( -0.0899406794982115 , 0.615010610873399 )"
```

c) 0 is in both confidence intervals. We are not confident that there is a relationship between Verbal SAT scores and GPA.