Math 181B: Homework 5

Merrick Qiu

Exercise 1

Since $\frac{(n-2)S^2}{\sigma^2} \sim \chi^2_{n-2}$,

$$\begin{aligned}
P\left(\chi^2_{n-2} \leq \chi^2_{1-\alpha,n-2}\right) = 1-\alpha &\implies P\left(\frac{(n-2)S^2}{\sigma^2} \leq \chi^2_{1-\alpha,n-2}\right) = 1-\alpha \\
&\implies P\left(\frac{(n-2)S^2}{\chi^2_{1-\alpha,n-2}} \leq \sigma^2\right) = 1-\alpha \\
&\implies P\left(\sigma^2 \geq \frac{(n-2)S^2}{\chi^2_{1-\alpha,n-2}}\right) = 1-\alpha
\end{aligned}$$

Since $P\left(\sigma^2 \geq \frac{(n-2)S^2}{\chi^2_{1-\alpha,n-2}}\right) = 1-\alpha$, This implies that $\left[\frac{(n-2)S^2}{\chi^2_{1-\alpha,n-2}}, \infty\right]$ is a $100(1-\alpha)\%$ one-sided CI for σ^2

Exercise 2

1. The t-value of the intercept is the estimate divided by the standard error. $18166.1/1003.7 = 18.099$. The probability is $P(t_{50} > |18.099|) = 2 * pt(18.099, 50, lower = F) = 1.346e - 23$.

   We can find the t-value using the inverse cdf of the t-distribution. $P(t_{50} > |7.34e - 09|) \implies t = qt(7.34e - 09/2, 50) = 6.94$. Dividing the estimate by the t-value yields a standard error of $752.8/6.94 = 108.47$.

   Overall the summary looks like:

a	Estimate	Std. Error	t-value	$Pr(> \|t\|)$
Intercept	18166.1	1003.7	18.099	1.346e-23
years	752.8	108.47	6.94	7.34e-09

2. We have $H0 : \beta_1 = 0$ and $H1 : \beta_1 > 0$. Since the p-value is $(7.34e - 09)/2 = 3.67e - 09 < 0.03$ for the slope, we reject the null and there is a positive relationship between salaries and years. The test statistic is 6.94 and the degrees of freedom are 50.

Exercise 3

```
#———PART A———
setwd("C:/Users/merri/Documents/MATH-31H/MATH 181B/Homework 5")
cars = read.csv("cars.csv")

model = lm(mpg ~ hp, data=cars)
summary(model)

#   Estimate Std. Error t value Pr(>|t|)
# (Intercept) 30.09886    1.63392   18.421   < 2e-16 ***
#   hp            -0.06823    0.01012   -6.742 1.79e-07 ***

#PLUGGING IN THE ESTIMATES YIELDS THIS REGRESSION LINE:
# mpg = -0.06823*hp + 30.09886

#———PART B———
# Multiple R-squared:  0.6024,   Adjusted R-squared:  0.5892

# THE R^2 IS ABOUT 0.60, MEANING 60% OF THE VARIANCE
# IN THE MPG IS EXPLAINED BY THE HP.

#———PART C———
# R^2 CANNOT BE USED TO DETERMINE THE APPROPRIATENESS OF LINEAR REGRESSION

#———PART D———
# R = SQRT(0.6024) = 0.7761
# THEREFORE THE PREDICTED MPG IS 3*0.7761=2.3283 STD BELOW THE MEAN

#———PART E———
predict(model,newdata = data.frame(hp=150),interval='confidence', level=0.9)
# fit         lwr       upr
# 1 19.86462 18.70419 21.02504

# THE 90% CI IS (18.70419, 21.02504)
# IF WE WERE TO REPEAT THIS PROCEDURE MANY TIMES,
# 90% OF THE CI WILL CONTAIN THE TRUE AVERAGE MPG FOR 150HP.

#———PART F———
predict(model,newdata = data.frame(hp=150),interval='prediction', level=0.9)
#   fit        lwr       upr
#1 19.86462 13.20626 26.52297

# THE 90% PI IS (13.20626, 26.52297)
# IF WE WERE TO REPEAT THIS PROCEDURE MANY TIMES,
# 90% OF THE PI WILL CONTAIN THE TRUE MPG OF A CAR FOR 150HP.
```