

# DaigleHomework7.R

*mbair*

*Tue Oct 30 16:12:03 2018*

---

Chris Daigle

Homework 7 Bootstrapping

---

```
# Exercise
#
# Another bootstrap method in a linear regression model is to
# resample residuals instead of (X, Y). The former is called Residual
# Bootstrapping and the latter is called Pairwise Bootstrapping. I beleve that
# boot() is based on the Pairwise Bootstrap. Estimate the standard error and
# bootstrap the 95% confidence intervals for the coeficient of horsepower with
# Residual Bootstrapping.
library(boot)
library(ISLR)

set.seed(1)

df <- Auto
dfBoot <- df[, c(1, 4, 5)]

# Sample Model
X1 <- dfBoot[, 2]
X2 <- dfBoot[, 3]
y <- dfBoot[, 1]

lmFit <- lm(y ~ X1 + X2)

beta0Hat <- lmFit$coefficients[1]
beta1Hat <- lmFit$coefficients[2]
beta2Hat <- lmFit$coefficients[3]

n <- nrow(dfBoot)
# Using robust standard errors
num <- sum(((X1 - mean(X1)) ^ 2) * (lmFit$residuals^2))/n
den <- sum((X1-mean(X1))^2)/(n-1)
se <- sqrt(num)/den

tCoef1 <- sqrt(n) * beta1Hat / se

# Residual Bootstrap Model

# Initialize relevant values
R <- 1000

beta1Star <- rep(NA, R)
```

```

beta2Star <- rep(NA, R)
t1Star <- rep(NA, R)

for (r in 1:R) {
  index <- sample(1:n, size = n, replace = TRUE)

  uHat <- lmFit$residuals[index]
  yHat <- beta0Hat + beta1Hat * X1 + beta2Hat * X2 + uHat

  lmRes <- lm(yHat ~ X1 + X2)
  beta1Star[r] <- lmRes$coefficients[2]

  numStar <- sum(((X1 - mean(X1)) ^ 2) * (lmRes$residuals^2))/n
  denStar <- sum((X1 - mean(X1))^2)/(n-1)
  seStar <- sqrt(numStar)/denStar

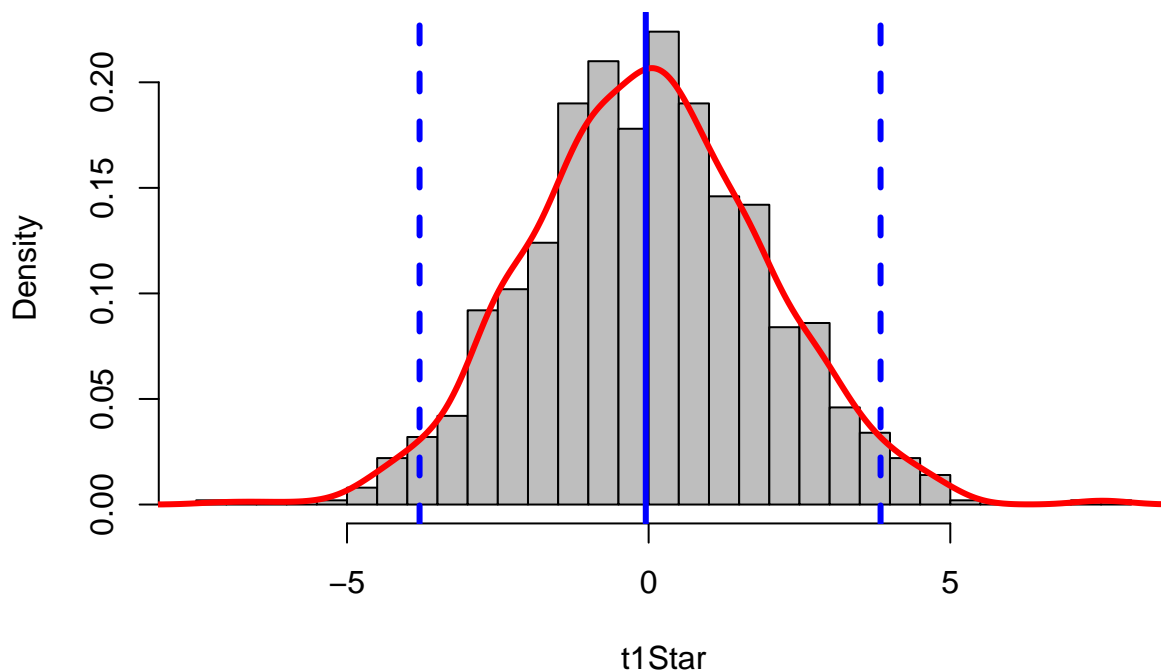
  t1Star[r] <- sqrt(n) * (beta1Star[r] - beta1Hat) / seStar
}

t1Star <- sort(t1Star)
# Get the 2.5% and 97.5% levels
critT1Star <- c(t1Star[25], t1Star[975])

# Plot and cut the values
hist(t1Star, breaks = 30, probability = TRUE, col= "grey",
     main = "Distribution of t* Under Residual Bootstrap")
lines(density(t1Star), col = "red", lwd = 3)
abline(v = c(critT1Star, beta1Hat), col = c("blue", "blue"), lty = c(2, 2, 1), lwd = 3)

```

## Distribution of $t^*$ Under Residual Bootstrap



```

cat(paste("The coefficient of interest is", round(beta1Hat, 3),
          "and the t-statistic is", round(tCoef1, 3), "."))

## The coefficient of interest is -0.047 and the t-statistic is -8.49 .

cat(paste("The 95% critical values are", round(critT1Star[1], 3),
          "and", round(critT1Star[2], 3), "."))

## The 95% critical values are -3.797 and 3.844 .

```