**HW1 R code and plots sheet**

1 (d)

**R code:**

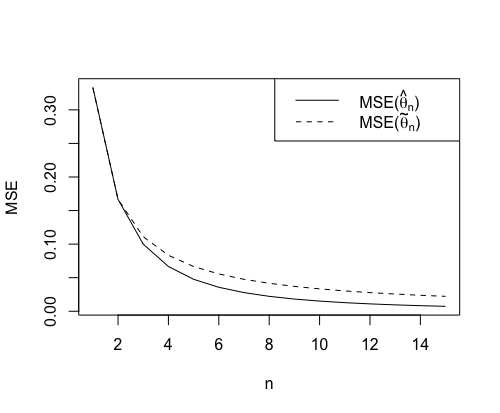
n <- 1:15

plot(x = n, y = 2/((n+1)\*(n+2)), xlab = "n", ylab = "MSE", type = "l")

lines(x = n, y = 1/(3\*n), lty = 2)

legend("topright", legend = c(expression(paste("MSE(",hat(theta)[n],")", sep = "")),

expression(paste("MSE(",tilde(theta)[n],”)”, sep = ""))), lty = 1:2)

**Plot:**

From the plot above, it is clear that I would prefer the estimator with smaller overall MSE, which is .

5.

**R code:**

quakes <- read.table(file = "~/Documents/STAT\_201\_B/Homework/HW1/fijiquakes.dat.txt", header = TRUE)

x <- quakes$mag

xsample <- seq(min(x), max(x), length = 100) # Take 100 samples for x.

Fhat <- apply(outer(x, xsample, "<="), 2, mean) # Calculate the ECDF.

n <- length(x)

# Calculate L(x) and U(x).

epsilon\_n <- sqrt(log(2/0.05)/(2\*n))

L <- sapply(Fhat, FUN=function(x) max(x - epsilon\_n, 0))

U <- sapply(Fhat, FUN=function(x) min(x + epsilon\_n, 1))

# Plot the ECDF and a 95% confidence interval for F.

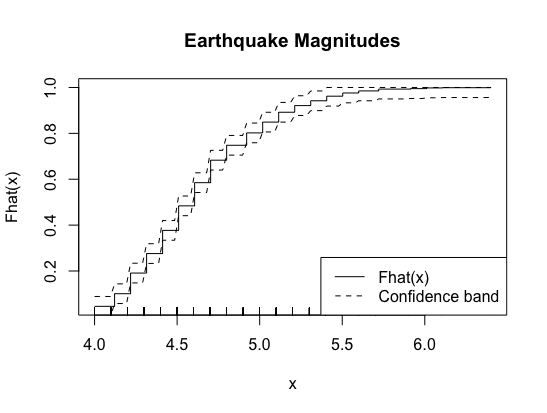
plot(xsample, Fhat, xlab = "x", ylab = "Fhat(x)", type = "s", main = "Earthquake Magnitudes")

rug(x)

lines(xsample, L, lty = 2)

lines(xsample, U, lty = 2)

legend("bottomright", legend = c("Fhat(x)", "Confidence band"), lty = 1:2)

**Plot:**

6.

At first, I searched on Google and found an R package named “boot”. So, I tried to follow the syntax introduced online and get the R code as follow. But to be honest, I cannot understand all of the following codes because some of them can only be used in “boot” package.

**R code:**

library(boot)

clouds <- read.table("~/Documents/STAT\_201\_B/Homework/HW1/clouds.dat", header = TRUE)

seeded <- clouds$Seeded

unseeded <- clouds$Unseeded

theta\_hat <- median(seeded) - median(unseeded)

sample\_median <- function(x, d){return (median(x[d]))}

N <- 100

seeded\_boot <- boot(seeded, sample\_median, N)

unseeded\_boot <- boot(unseeded, sample\_median, N)

se\_hat = sqrt(var(seeded\_boot$t) + var(unseeded\_boot$t))

CI <- c(theta\_hat + qnorm(0.025) \* se\_hat, theta\_hat - qnorm(0.025) \* se\_hat)

print(CI)

**Result:**

After the first run of this code, I got .

Then I tried to make the bootstrap work without using the “boot” package.

**R code:**

clouds <- read.table("~/Documents/STAT\_201\_B/Homework/HW1/clouds.dat", header = TRUE)

seeded <- clouds$Seeded

unseeded <- clouds$Unseeded

theta\_hat <- median(seeded) - median(unseeded)

bootstrap <- function(data, B) {

resample <- lapply(1 : B, function(i) sample(data, replace = TRUE))

sample\_median <- sapply(resample, FUN = function(x) median(x))

variance <- var(sample\_median)

list(resample = resample, median = sample\_median, variance = variance)

}

seeded\_boot <- bootstrap(seeded, 10000)

unseeded\_boot <- bootstrap(unseeded, 10000)

se\_hat = sqrt(seeded\_boot$variance + unseeded\_boot$variance)

CI <- c(theta\_hat + qnorm(0.025)\*se\_hat, theta\_hat - qnorm(0.025)\*se\_hat)

print(CI)

**Result:**

After the first run of this code, I got .