Assignment 1

Alfonso Rojas-Alvarez

September 9, 2017

Prepare the dataset:

```
library(ggplot2)
library(foreign)
library(Hmisc)
library(grid)
library(gridExtra)
library(easyGgplot2)
rm(list = ls())
```

Modify this to your path

```
setwd("/Users/Alfonso/Google Drive/UT/Fall 2017/RD")
```

Load Dataset:

```
students <- read.dta("students.dta")
```

Load Variable Labels:

```
var.labels <- attr(students, "var.labels")
data.key <- data.frame(var.name=names(students), var.labels)
```

Clear Mising gkschid:

```
STAR_kindergarteners <- students[!(is.na(students$gkschid)),]
```

Population Parameters:

```
mu <- mean(STAR_kindergarteners$gktreadss)
mu
```

```
## [1] 436.7253
```

```
sigma <- sd(STAR_kindergarteners$gktreadss)
sigma
```

```
## [1] 31.70626
```

Set Sample Size:

```
n <- 160
```

Take Sample 1

```
sample1 <- STAR_kindergarteners[sample(nrow(STAR_kindergarteners), n),]

xbar1 <- mean(sample1$gktreadss)
xbar1

## [1] 436.1187

sigma1 <- sd(sample1$gktreadss)
sigma1

## [1] 31.55097
```

Sampling Error:

```
diff1 <- mu - xbar1
diff1

## [1] 0.6065912
```

Confidence Interval:

```
c1 <- xbar1 - qnorm(0.975) * sigma1/sqrt(n)
c2 <- xbar1 + qnorm(0.975) * sigma1/sqrt(n)

sample1_CI <- c(c1, c2)
sample1_CI

## [1] 431.2300 441.0075
```

Hypothesis test:

```
z1 <- (xbar1 - mu)/(sigma1/sqrt(n))
z1

## [1] -0.2431887
```

```r
alpha <- 0.05
critical_values <- c(-qnorm(1-alpha/2), qnorm(1-alpha/2))
reject1 <- isTRUE(!-qnorm(1-alpha/2) <  z1 & z1 < qnorm(1-alpha/2))
reject1

## [1] FALSE
```

Take Sample 2

```r
sample2 <- STAR_kindergarteners[sample(nrow(STAR_kindergarteners), n),]
xbar2 <- mean(sample2$gktreadss)
xbar2

## [1] 438.475

sigma2 <- sd(sample2$gktreadss)
sigma2

## [1] 32.16505
```

Sampling Error:

```r
diff2 <- mu - xbar2
diff2

## [1] -1.749659
```

Sample Variation:

```r
sample_variation <- xbar1 - xbar2
sample_variation

## [1] -2.35625
```

Confidence Interval:

```r
c1 <- xbar2 - qnorm(0.975) * sigma2/sqrt(n)
c2 <- xbar2 + qnorm(0.975) * sigma2/sqrt(n)

sample2_CI <- c(c1, c2)
sample2_CI

## [1] 433.4911 443.4589
```

Hypothesis test:

```
z2 <- (xbar2 - mu)/(sigma2/sqrt(n))
z2

## [1] 0.6880644

alpha <- 0.05
critical_values <- c(-qnorm(1-alpha/2), qnorm(1-alpha/2))
reject2 <- isTRUE(!-qnorm(1-alpha/2) <  z2 & z2 < qnorm(1-alpha/2))
reject2

## [1] FALSE
```

Bootstrap Method

```
ggplot2.multiplot(a,b, c, d, cols=2)
```