Assignment 4 Question 1 b)

Sheen Thusoo

30/11/2021

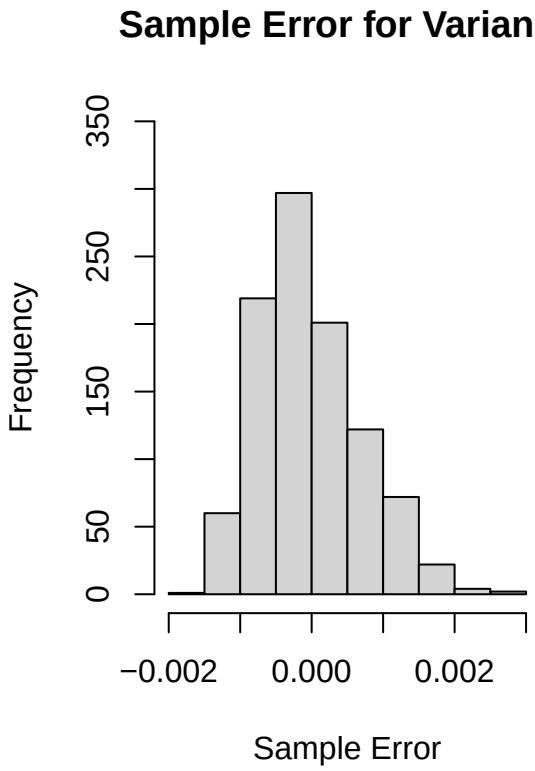
For each sample calculate the variance and the IQR. Then construct two histograms (in a single row) of the sample error for each attribute.

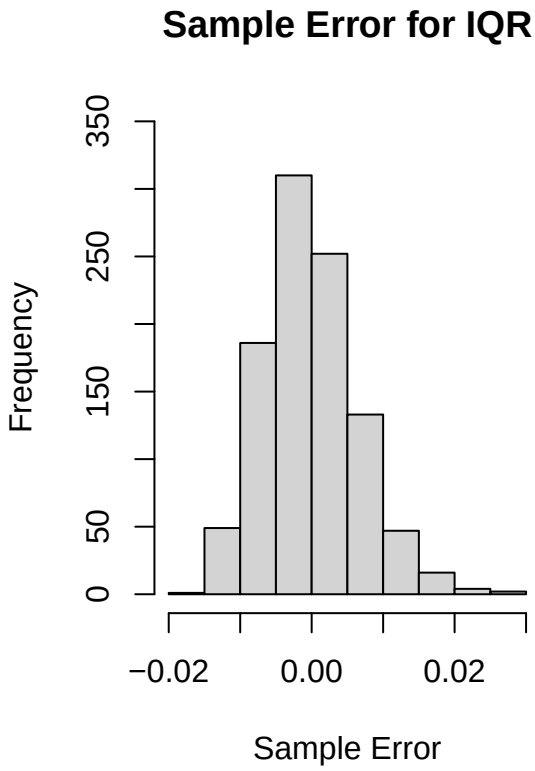
```
data <- read.csv('EconomicMobility.csv')
```

```
VarIQR <- function(pop, N) {
  variance <- var(pop) * (N-1)/N
  iqr <- IQR(pop)
  c(variance, iqr)
}

N <- length(data[,1])
result <- VarIQR(data$Mobility, N)
```

```
SRSWOR <- function(pop, sampSize) {
  sample(pop, sampSize, replace = FALSE)
}

M <- 1000
n <- 100

samples <- sapply(1:M, function(i) SRSWOR(data$Mobility, n))
sample_attributes <- apply(samples, MARGIN=2, FUN=function(s) {
  VarIQR(s, length(s))
})
```

```
samp_error_var <- apply(sample_attributes, MARGIN=2, FUN=function(s) {
  s[1] - result[1]
})
samp_error_iqr <- apply(sample_attributes, MARGIN=2, FUN=function(s) {
  s[2] - result[2]
})

par(mfrow=c(1,2))
hist(samp_error_var, ylim=c(0,350), main="Sample Error for Variance", xlab="Sample Error")
hist(samp_error_iqr, ylim=c(0,350), main="Sample Error for IQR", xlab="Sample Error")
```


Sample Error for Variance
Frequency
0
50
150
250
350
−0.002
0.000
0.002
Sample Error


Sample Error for IQR
Frequency
0
50
150
250
350
−0.02
0.00
0.02
Sample Error