ForteLab10.R

andre

2022-06-28

set.seed(2) #this is to set seed. By doing so, the initiation point is always the same, not random.  
  
sampleSize <- 30  
  
# create normal distribution of 20000 observations with mean value 20 and standard deviation of 3 and set this as a student population  
studentPop <- rnorm(20000, mean=20, sd=3)  
  
# The following line of code samples 30 random values from the studentPop variable.  
undergrads <- sample(studentPop, size = sampleSize, replace=TRUE)  
  
# create a sample of graduate students. Sample size is 30, mean is 25, standard deviation is 3. See the mean is 5 years older than the undergraduate sample apparently.   
# This line of code creates a normal distribution of 30 data points with a mean of 25 and a std deviation of 3.  
grads <- rnorm(sampleSize, mean = 25,sd = 3)  
  
# The following code essentially "flips a coin" which will cause the testSample variable to be assigned either the values from the grads variable or the undergrads variable.  
if (runif(1) > 0.5)  
 {testSample <- grads}else{  
 testSample <- undergrads}  
  
mean(testSample)

## [1] 24.89729

mySample <- replicate(100, mean(sample(studentPop, size = sampleSize, replace=TRUE)))  
  
quantile(mySample, probs = c(0.025, 0.975))

## 2.5% 97.5%   
## 19.01696 20.92209

if(mean(testSample) < quantile(mySample, probs = 0.025) | mean(testSample) < quantile(mySample, probs = 0.975)){  
 print("Sample mean is extreme")}else{  
 print("Sample mean is not extreme")}

## [1] "Sample mean is not extreme"