BACS-hw5.R

2022-04-16

#Question 1  
自己上網看上課前十分鐘(討論)  
  
#Question 2  
setwd("C:/Users/eason/Desktop/清大 BACS/資料/")  
verizon <- read.csv("verizon.csv")  
time <- verizon$Time  
  
hyp\_mean <- 7.6  
sample\_n <- length(time)  
sample\_mean <- mean(time)  
sample\_sd <- sd(time)  
sample\_se <- sd(time)/sqrt(length(time))  
  
#a  
#1  
t.test(time, mu=7.6, alternative="greater", conf.level=0.99)

##   
## One Sample t-test  
##   
## data: time  
## t = 2.5608, df = 1686, p-value = 0.005265  
## alternative hypothesis: true mean is greater than 7.6  
## 99 percent confidence interval:  
## 7.683604 Inf  
## sample estimates:  
## mean of x   
## 8.522009

#2 不知道為啥要加sig.level, type=one.sample  
power.t.test(n=length(time), delta=sample\_mean-hyp\_mean, sig.level=0.01, sd=sample\_sd, alternative="one.sided", type= "one.sample")

##   
## One-sample t test power calculation   
##   
## n = 1687  
## delta = 0.9220095  
## sd = 14.78848  
## sig.level = 0.01  
## power = 0.5918705  
## alternative = one.sided

#b  
#1  
t\_value <- (mean(time)-hyp\_mean) / sample\_se  
  
#2  
bootstrap\_null\_alt <- function(sample0, hyp\_mean){  
 resample <- sample(sample0, length(sample0), replace=TRUE)  
 resample\_se <- sd(resample) / sqrt(length(resample))  
   
 t\_stat\_alt <- (mean(resample) - hyp\_mean) / resample\_se  
 t\_stat\_null <- (mean(resample) - mean(sample0)) / resample\_se  
   
 c(t\_stat\_alt, t\_stat\_null)  
}  
  
set.seed(42)  
boot\_t\_stats <- replicate(10000, bootstrap\_null\_alt(time, hyp\_mean))  
t\_alt <- boot\_t\_stats[1,]  
t\_null <- boot\_t\_stats[2,]  
  
#3  
cutoff\_99 <- quantile(t\_null, probs =0.99)  
  
#4  
#bootstrapped\_pvalue 跟直接用pt算不知道為啥不一樣(?)因為bootstrap?//不太懂ecdf怎麼運作的  
null\_probs <- ecdf(t\_null)  
one\_tailed\_pvalue <- 1 - null\_probs(t\_value)  
#bootstrapped\_power   
alt\_probs <- ecdf(t\_alt)  
t\_power <- 1 - alt\_probs(cutoff\_99)