STP598-Assignment 2

Hao Wang
hwang306

September 28, 2017

1 Question 1

2 Question 3

Permuation test

```
#generate a random dataset including x and y variables

set.seed(1)
x <- rnorm(100) + 100
y <- rchisq(100, 7)

#spearman correlation
cor.0 <- cor(x, y, method = "spearman")
cor.test(x, y, method = "spearman")
```

```
##
##  Spearman's rank correlation rho
##
## data:  x and y
## S = 153240, p-value = 0.4255
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##        rho
## 0.08046805
```

```
#run a permutation test for 10,000 times
#number of permutations
R <- 10000
reps <- numeric(R)
#create a long vector first
z <- c(x,y) #pooled sample
K <- length(z)
for (i in 1:R) {
#generate indicies k for the first sample
```

```
k <- sample(K, size = 100, replace = FALSE)
x1 <- z[k]
y1 <- z[-k] #the rest
reps[i] <- cor(x1, y1, method = "spearman") #spearman test statistics
}

#get empirical p vale
p <- mean(c(cor.0, reps) >= cor.0)
p
```

```
## [1] 0.2122788
```

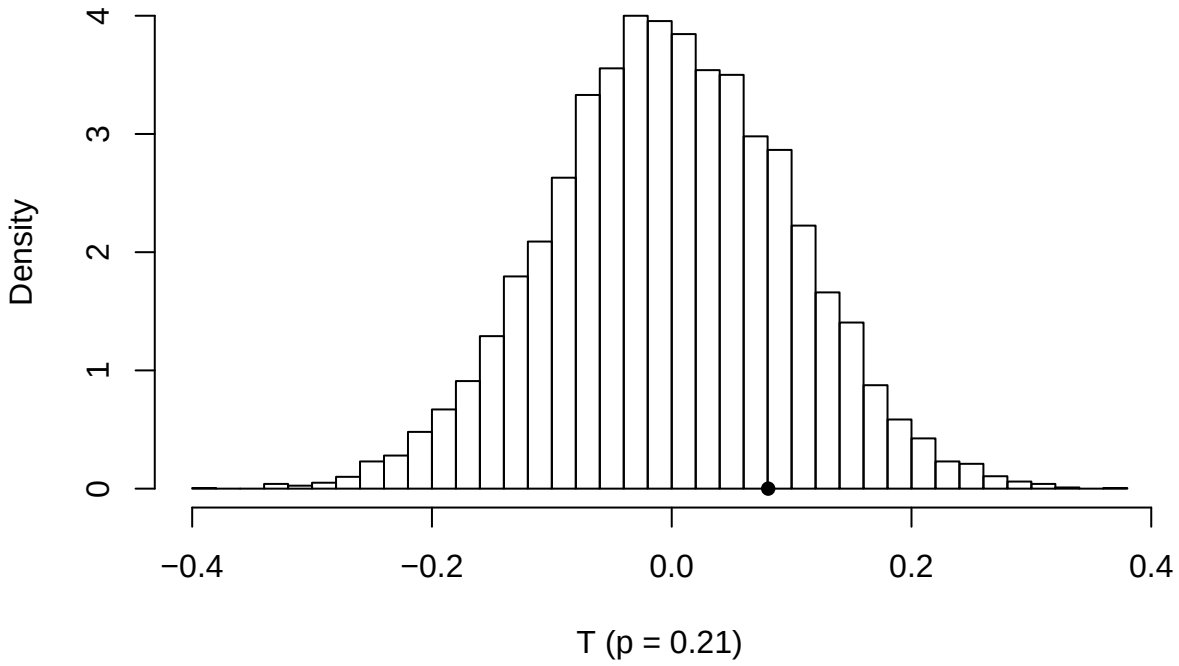
The empirical test value $\hat{p}$ can be obtained by

$$\hat{p} = \frac{\{1 + \sum_{b=1}^{B} I(\hat{\theta}^{(b)}) \geq \hat{\theta}\}}{B+1}$$

In this equation, B is the number of permutations, $\hat{theta}$ is the test statistics. In my case the test statistics is the spearman correlation test value. The p value I got is 0.2122788.

And we can get a histogram of the spearman statistics

```
hist(reps, main = "", freq = FALSE, xlab = "T (p = 0.21)", breaks = "scott")
points(cor.0, 0, cex = 1, pch = 16) #observed T
```



Thus we cannot reject the null: true spearman correlation is 0. In the orginal spearman test, the p value is 0.422.