STA_104_HW2.R

pumad

2021-10-21

```
##APPENDIX
#Sungwon Lee
#912978026

# Exercise 8
siblings<-data.frame(hometown=c(rep("rural",24),rep("urban",17)),
                     siblings=c(3,2,1,1,2,1,3,2,2,2,2,5,1,4,1,1,1,1,6,2,2,
                                2,1,1,1,0,1,1,0,0,1,1,1,8,1,1,1,0,1,1,2))

set.seed(1)
nsim=10000
permDiffs=rep(NA,nsim)
for (i in 1:nsim){
  new.dat=data.frame(sib=sample(siblings$siblings, replace = FALSE), home=siblings$hometown)
  permDiffs[i]=diff(tapply(new.dat$sib, new.dat$home, mean))
}

only.rural = subset(siblings, hometown=="rural")
mean.rural = mean(only.rural$siblings)

only.urban = subset(siblings, hometown=="urban")
mean.urban = mean(only.urban$siblings)

actual.diff1 = mean.rural-mean.urban
actual.diff2 = mean.urban-mean.rural

lower= sum(permDiffs<=actual.diff2)
upper=sum(permDiffs>=actual.diff1)
pv.b=(lower+upper)/nsim


# Exercise 10
Apple<-data.frame(Group=c(rep("experimental",5),rep("control",6)),
                  numdata=c(-1.383, -0.674, 0.431, -0.967, -0.431, 0.674, 0, 0.210, 1.383, 0.96
7, -0.210))

n = length(Apple$Group)
p = 462
variable = Apple$numdata

PermSamples = matrix(0, nrow=n, ncol = p)

for(i in 1:p){
  PermSamples[,i] = sample(variable, size = n, replace=FALSE)
}
PermSamples[, 1:5]
```

```
##         [,1]   [,2]   [,3]   [,4]   [,5]
##  [1,] -0.674 -1.383  0.000  0.674 -0.967
##  [2,]  0.431  0.967 -0.431  0.967 -0.674
##  [3,]  0.674  1.383  0.431 -1.383  0.210
##  [4,] -0.967 -0.210 -0.967 -0.210  0.000
##  [5,]  0.967  0.210  1.383 -0.674  0.674
##  [6,] -0.210 -0.431 -1.383  0.210 -0.210
##  [7,]  0.210  0.431  0.210  0.431 -1.383
##  [8,] -0.431  0.000  0.674 -0.967  0.431
##  [9,] -1.383 -0.674 -0.210  1.383 -0.431
## [10,]  1.383  0.674 -0.674 -0.431  1.383
## [11,]  0.000 -0.967  0.967  0.000  0.967
```

```
Perm.test.stat1 = Perm.test.stat2 = rep(0, p)

for (i in 1:p){
  Perm.test.stat1[i] = abs(mean(PermSamples[Apple$Group=='experimental', i] -
                                mean(PermSamples[Apple$Group=='control', i])))

  Perm.test.stat2[i] = abs(median(PermSamples[Apple$Group=='experimental', i] -
                                  median(PermSamples[Apple$Group=='control', i])))
}

test.stat1 = abs(mean(Apple$numdata[Apple$Group=='experimental']) - mean(Apple$numdata[Apple$Gr
oup=='control']))
test.stat2 = abs(median(Apple$numdata[Apple$Group=='experimental']) - median(Apple$numdata[Appl
e$Group=='control']))

test.stat1; test.stat2
```

```
## [1] 1.1088
```

```
## [1] 1.116
```

```
mean(Perm.test.stat1 >= test.stat1)
```

```
## [1] 0.03679654
```

```
mean(Perm.test.stat2 >= test.stat2)
```

```
## [1] 0.05194805
```

```
# Exercise 18
# Wilcoxon
Exp = c(11, 33, 48, 34, 112, 369, 64, 44)
Cont = c(177, 80, 141, 332)
W=wilcox.test(Exp, Cont, conf.int=TRUE,conf.level = 0.95);W
```

```
##
##  Wilcoxon rank sum exact test
##
## data:  Exp and Cont
## W = 5, p-value = 0.07273
## alternative hypothesis: true location shift is not equal to 0
## 95 percent confidence interval:
##  -284   32
## sample estimates:
## difference in location
##                   -102
```

```
# Permutation
Apple<-data.frame(Group=c(rep("experimental",8),rep("control",4)),
                  numdata=c(11, 33, 48, 34, 112, 369, 64, 44, 177, 80, 141, 332))

n = length(Apple$Group)
p = 495
variable = Apple$numdata

PermSamples = matrix(0, nrow=n, ncol = p)

for(i in 1:p){
  PermSamples[,i] = sample(variable, size = n, replace=FALSE)
}
PermSamples[, 1:5]
```

```
##       [,1] [,2] [,3] [,4] [,5]
##  [1,]   64  112  112  141   48
##  [2,]   80   33   44   33   34
##  [3,]  177  332   80   64  141
##  [4,]  141  177   64  369   44
##  [5,]   11   80  332  177   80
##  [6,]   33   34   11  112   33
##  [7,]  112   48   34   80  112
##  [8,]   44   44  141   48   11
##  [9,]  369  141  177   34  332
## [10,]  332   64  369   44   64
## [11,]   48  369   48  332  369
## [12,]   34   11   33   11  177
```

```
Perm.test.stat1 = Perm.test.stat2 = rep(0, p)

for (i in 1:p){
  Perm.test.stat1[i] = abs(mean(PermSamples[Apple$Group=='experimental', i] -
                                mean(PermSamples[Apple$Group=='control', i])))

  Perm.test.stat2[i] = abs(median(PermSamples[Apple$Group=='experimental', i] -
                                  median(PermSamples[Apple$Group=='control', i])))
}

test.stat1 = abs(mean(Apple$numdata[Apple$Group=='experimental']) - mean(Apple$numdata[Apple$Gr
oup=='control']))
test.stat2 = abs(median(Apple$numdata[Apple$Group=='experimental']) - median(Apple$numdata[Appl
e$Group=='control']))

test.stat1; test.stat2
```

```
## [1] 93.125
```

```
## [1] 113
```

```
mean(Perm.test.stat1 >= test.stat1)
```

```
## [1] 0.2222222
```

```
mean(Perm.test.stat2 >= test.stat2)
```

```
## [1] 0.1131313
```