MANA130083.01 Nonparametric *Spring 2023*

Solution to Homework 5

Date: March 31, 2023 *Scribe: Jianliang He*

***Warning**: This note is only used as a reference solution for the homework, and the solution to each question is not unique. The solution may contain factual and/or typographic errors and comments and criticism are kindly welcomed.*

Problem 1 *Generate 50 pairs of (X, Y) according to the following model: $X \sim U[-1, 1]$ and $Y|X \sim N(\sqrt{1-X^2}, 1)$. 1) Test the null hypothesis H_0 : X and Y are independent using Kendall's τ. Is the null hypothesis rejected at level $\alpha = 0.05$? 2) Test the same null hypothesis using Spearman's correlation. Is the null hypothesis rejected at level $\alpha = 0.05$?*

Solution:

```
corr.Kendall<-function(x,y){
  n = length(x)
  x.sign = apply(matrix(x,nrow = 1), MARGIN = 2,
                 function(data){sign(x-data)})
  y.sign = apply(matrix(y,nrow = 1), MARGIN = 2,
                 function(data){sign(y-data)})
  tau = (sum(x.sign*y.sign)/2)/choose(n,2)
  return(tau)
}

corr.Spearman<-function(x,y){
  x.rank = rank(x); y.rank = rank(y)
  return(cor(x.rank,y.rank))
```

```
}

corr.pvalue <- function(x,y,nrep){
  n <- length(x)
  # Calculate the test statistics
  stats.Kendall <- corr.Kendall(x, y)
  stats.Spearman <- corr.Spearman(x, y)
  # Simulate under the null
  sim.Kendall <- c(); sim.Spearman <- c()
  for (i in 1:nrep) {
    sim.1 <- rnorm(n,0,1); sim.2 <- rnorm(n,0,1);
    sim.Kendall[i] <- corr.Kendall(sim.1,sim.2)
    sim.Spearman[i] <- corr.Spearman(sim.1,sim.2)
  }
  return(list(Kendall=2*min(mean(sim.Kendall<stats.Kendall),
                            mean(sim.Kendall>stats.Kendall)),
              Spearman=2*min(mean(sim.Spearman<stats.Spearman),
                             mean(sim.Spearman>stats.Spearman))))
}

# Calculate the p-value of the statistics under the problem setting
set.seed(217)
X = runif(50,-1,1); Y = rnorm(50,sqrt(1-X^2),1); nrep = 500
corr.pvalue(X,Y,nrep)
```

Both of the test failed to reject the null under the problem setting with respectively p-value of 0.111 and 0.139. Kendall's and Spearman's correlation test both fail to reject the null with $\alpha = 0.05$. □

Problem 2 *Generate 10 pairs of* (X, Y) *according to the following model:* $X \sim U[1, 2]$ *and* $Y|X \sim N(X/2, 1)$. *Test the null hypothesis* H_0 : *X and Y are independent using Kendall's* τ *and*

Spearman's correlation. Which method is more powerful in this setting?

Solution:

```
corr.power <- function(np,nrep){
  power.Kendall <- 0; power.Spearman <- 0;
  for (i in 1:np){
    X = runif(10,1,2); Y = rnorm(10,X/2,0.1)
    pvalue <- corr.pvalue(X,Y,nrep)
    power.Kendall <- power.Kendall+ ifelse(pvalue$Kendall<0.05,1,0)
    power.Spearman <- power.Spearman+ ifelse(pvalue$Spearman<0.05,1,0)
    print(pvalue)
  }
  return(list(Kendall=power.Kendall/np,Spearman=power.Spearman/np))
}
corr.power(200,1000)
```

The estimated power of Kendall's and Spearman's correlation test is approximately 0.04 and 0.07 under the problem setting and the latter one is more powerful.

□