邹易+202111106+lab1

zouyi

2020/9/30

#设置全局路径

#work1

cor.data <- read.csv('correlation\_data.csv', header = T)  
cor.data <- cor.data[,-1] #去掉第一列  
  
# 计算pearson相关系数  
# corr.a <- cor(a.x,a.y)  
# corr.b <- cor(b.x,b.y)  
# corr.c <- cor(c.x,c.y)  
# corr.d <- cor(d.x,d.y)  
# corr.e <- cor(e.x,e.y)  
# corr.f <- cor(f.x,f.y)  
# corr.g <- cor(g.x,g.y)  
# corr.h <- cor(h.x,h.y)  
corr.pearson <- vector(mode = "numeric", length = 8)  
for (i in seq(2,16,2)) {  
 corr.pearson[i/2] <- cor(cor.data[,i], cor.data[,i-1])  
}  
cat("pearson correlation:", '\n', letters[1:8], '\n', corr.pearson, '\n')

## pearson correlation:   
## a b c d e f g h   
## 0.8350736 0.3130239 0.01134618 -0.3201501 -0.004464113 -0.05662948 0.0298095 -0.09294615

#计算Spearman 秩相关系数  
corr.Spearson <- vector(mode = "numeric", length = 8)  
for (i in seq(2,16,2)) {  
 corr.Spearson[i/2] <- cor(cor.data[,i], cor.data[,i-1], method = "spearman")  
}  
cat("Spearson correlation:", '\n', letters[1:8], '\n', corr.Spearson, '\n')

## Spearson correlation:   
## a b c d e f g h   
## 0.8277667 0.3203891 0.02894975 -0.348998 0.02100691 -0.06771639 -0.008803729 -0.08317638

spearson.corr <- function(x,y){  
 if(length(x) != length(y)){  
 return("The lengths of x and y are not equal")  
 }  
 x.rank <- rank(x)  
 y.rank <- rank(y)  
 n <- length(x)  
 temp <- function(r1,r2){  
 n <- length(r1)  
 re <- sum(r1 \* r2)-n \* ((n + 1) / 2) ^ 2  
 return(re)  
 }  
 # 这是带结的计算公式  
 corr1 <- temp(x.rank,y.rank) / (sqrt(temp(x.rank,x.rank)) \* sqrt(temp(y.rank,y.rank)))  
 #这是不带结的计算公式  
 corr2 <- 1 - 6 \* sum((x.rank - y.rank) ^ 2) / (n^3-n)   
 return(c(corr1,corr2))  
}  
spearson.corr(cor.data$a.x, cor.data$a.y) #用所编写的函数计算a的Spearson

## [1] 0.8277667 0.8277743

#计算Kendall 秩相关系数  
corr.kendall <- vector(mode = "numeric", length = 8)  
for (i in seq(2,16,2)) {  
 corr.kendall[i/2] <- cor(cor.data[,i-1], cor.data[,i], method = "kendall")  
}  
cat("kendall correlation:", '\n', letters[1:8], '\n', corr.kendall, '\n')

## kendall correlation:   
## a b c d e f g h   
## 0.6332125 0.2171152 0.01729049 -0.2375529 0.02719716 -0.04058556 -0.01307256 -0.03327443

kendall.corr <- function(x,y){  
 if(length(x) != length(y)){  
 return("The lengths of x and y are not equal")  
 }  
 n <- length(x)  
 nc = nd = 0  
 for (i in 1:(n-1)) {  
 for (j in (i+1):n) {  
 temp = (x[j] - x[i]) \* (y[j] - y[i])  
 if(temp > 0){  
 nc = nc + 1  
 }  
 if(temp < 0){  
 nd = nd + 1  
 }  
 }  
 }  
 cor1 <- (nc - nd) / choose(n, 2) # 无结  
 cor2 <- (nc - nd) / (nc + nd) # 有结   
 return(c(cor1,cor2))  
}  
  
kendall.corr(cor.data$a.x, cor.data$a.y)

## [1] 0.6315578 0.6348757