

```

data = read.csv("/Users/shuhualiang/Documents/Davis/STA 144/statepps.csv")

## a
Psi = data$landarea/sum(data$landarea)
New = cbind(data,Psi)

set.seed(32)
index = sample(1:length(data[,1]),5,replace = TRUE,prob = Psi)
samp = New[index,]

samp$Psi

## b
n = 5
t.hat.psi = sum(samp$counties/samp$Psi)/n
t.hat.psi

V.t.hat.psi = (1/n)*(1/(n-1))*sum((samp$counties/samp$Psi - t.hat.psi)^2)
V.t.hat.psi

## c
SE.t.hat.psi = sqrt(V.t.hat.psi)
SE.t.hat.psi

## d

list = c("Washington","Oregon","California","Idaho","Montana","Wyoming")

ind = sapply(1:6, function(i){
  which(data$state == list[i])
})

New.pop = data[ind,]
New.pop

Psi.new = New.pop$landarea/(sum(New.pop$landarea))
Psi.new

Pi = matrix(0,6,6)
for (i in 1:length(New.pop$state)){
  for (j in 1:(length(New.pop$state))){
    Pi[i,j] = Psi.new[i]*Psi.new[j]/(1-Psi.new[i]) +
    Psi.new[j]*Psi.new[i]/(1-Psi.new[j])
    Pi[i,i] = 0
  }
}

rownames(Pi) <-c("WA","OR","CA","ID","MT","WY")
colnames(Pi) <-c("WA","OR","CA","ID","MT","WY")

```

```

Pi_i = colSums(Pi)
sum(Pi_i)

## e
install.packages("gtools")
library(gtools)

WoR = combinations(6,2,1:6, repeats.allowed = FALSE)

t.hat.HT = sapply(1:15, function(i){
  sum(ti[WoR[i,]]/Pi_i[WoR[i,]])
})
t.hat.HT

####
V.hat.HT.tHT = sapply(1:15, function(i){
  first = sum((1-Pi_i[WoR[i,][1]])*(ti[WoR[i,][1]]/Pi_i[WoR[i,][1]])^2, (1-
Pi_i[WoR[i,][2]])*(ti[WoR[i,][2]]/Pi_i[WoR[i,][2]])^2)
  second = (Pi[WoR[i,][1],WoR[i,][2]]-Pi_i[WoR[i,][1]]*Pi_i[WoR[i,][2]])/
Pi[WoR[i,][1],WoR[i,][2]]*ti[WoR[i,][1]]/Pi_i[WoR[i,][1]]*ti[WoR[i,][2]]/
Pi_i[WoR[i,][2]]
  V.hat.HT.tHT = sum(first,second)
  return(V.hat.HT.tHT)
})

V.hat.HT.tHT

SE.tHT = sqrt(V.hat.HT.tHT)
SE.tHT

###
State1 = c(list[WoR[,1]])
State2 = c(list[WoR[,2]])

Output = as.data.frame(cbind(State1,State2,t.hat.HT,V.hat.HT.tHT,SE.tHT))

```