TP_2

EL_Hadrami_N'DOYE_Florence

22/01/2021

```
source("TP2Etudiant.R")
```

```
## Loading required package: xtable
```

EXERCICE

1. Soit $N = 6$ et $n = 3$. Afin d'étudier les distributions des estimateurs, nous supposons connues les valeurs pour toute la population :

$$y_1 = 98 \quad y_3 = 154 \quad y_5 = 190$$
$$y_2 = 102 \quad y_4 = 133 \quad y_6 = 175$$

(a) : Déterminons les probabilités d'inclusion du premier et du second ordre pour ces deux plans de sondages.

Plan1

- Probabilité d'inclusion du premier ordre

```
n <- 3
N <- 6
x <- plan1[,1:3]
k <- c(1:6)
# nombre d'echantillons
sizeech <- nrow(plan1)
appartenancepik <- matrix(rep(NA,sizeech * N),nrow = N ,ncol=sizeech)
rownames(appartenancepik) <- paste("k=",k)
colnames(appartenancepik) <- paste("ech",1:sizeech)
pik <-  matrix(rep(NA,sizeech * N),nrow = N ,ncol=sizeech)
rownames(pik) <- paste("k=",k)
colnames(pik) <- paste("ech",1:sizeech)
  for (i in 1:sizeech){
    for(j in 1:N){
        appartenancepik[j,i] <- un_dedans(x[i,],k[j])
        appartenancepik[appartenancepik==TRUE] <- 1
        appartenancepik[appartenancepik==FALSE] <- 0
        pik[j,i] <- round(appartenancepik[j,i] / n,2)
    }

  }
  head(appartenancepik,6)
```

```
##      ech 1 ech 2 ech 3 ech 4 ech 5 ech 6 ech 7 ech 8
## k= 1     1     1     1     1     0     0     0     0
```

```
## k= 2     0     0     0     0     1     1     1     1
## k= 3     1     1     0     0     1     1     0     0
## k= 4     0     0     1     1     0     0     1     1
## k= 5     1     0     1     0     1     0     1     0
## k= 6     0     1     0     1     0     1     0     1
```

```
  head(pik,6)
```

```
##      ech 1 ech 2 ech 3 ech 4 ech 5 ech 6 ech 7 ech 8
## k= 1  0,33  0,33  0,33  0,33  0,00  0,00  0,00  0,00
## k= 2  0,00  0,00  0,00  0,00  0,33  0,33  0,33  0,33
## k= 3  0,33  0,33  0,00  0,00  0,33  0,33  0,00  0,00
## k= 4  0,00  0,00  0,33  0,33  0,00  0,00  0,33  0,33
## k= 5  0,33  0,00  0,33  0,00  0,33  0,00  0,33  0,00
## k= 6  0,00  0,33  0,00  0,33  0,00  0,33  0,00  0,33
```

- Probabilité d'inclusion du second ordre

```
sizek <- choose(N,2)
# couple i,jech
coupleij <- t(combn(k,2))
appartenancepikl <- matrix(rep(NA,sizek * sizeech),nrow = sizek ,ncol=sizeech)
rownames(appartenancepikl) <- paste(coupleij[,1],coupleij[,2],sep = ",")
colnames(appartenancepikl) <- paste("ech",1:sizeech)
pikl <-  matrix(rep(NA,sizek * sizeech),nrow = sizek ,ncol=sizeech)
rownames(pikl) <- paste(coupleij[,1],coupleij[,2],sep = ",")
colnames(pikl) <- paste("ech",1:sizeech)
  for (i in 1:sizek){
    for(j in 1:sizeech){
        appartenancepikl[i,j] <- deux_dedans(x[j,],coupleij[i,1],coupleij[i,2])
        appartenancepikl[appartenancepikl==TRUE] <- 1
        appartenancepikl[appartenancepikl==FALSE] <- 0
        pikl[i,j] <- round(appartenancepikl[i,j] / n,2)
    }


  }
 head(appartenancepikl,6)
```

```
##     ech 1 ech 2 ech 3 ech 4 ech 5 ech 6 ech 7 ech 8
## 1,2     0     0     0     0     0     0     0     0
## 1,3     1     1     0     0     0     0     0     0
## 1,4     0     0     1     1     0     0     0     0
## 1,5     1     0     1     0     0     0     0     0
## 1,6     0     1     0     1     0     0     0     0
## 2,3     0     0     0     0     1     1     0     0
```

```
 head(pikl,6)
```

```
##     ech 1 ech 2 ech 3 ech 4 ech 5 ech 6 ech 7 ech 8
## 1,2  0,00  0,00  0,00  0,00  0,00  0,00     0     0
## 1,3  0,33  0,33  0,00  0,00  0,00  0,00     0     0
## 1,4  0,00  0,00  0,33  0,33  0,00  0,00     0     0
## 1,5  0,33  0,00  0,33  0,00  0,00  0,00     0     0
## 1,6  0,00  0,33  0,00  0,33  0,00  0,00     0     0
## 2,3  0,00  0,00  0,00  0,00  0,33  0,33     0     0
```

Plan2

- Probabilité d'inclusion du premier ordre

```
ind <- plan2[,1:3]
sizeind <- nrow(ind)
appartenancepikp2 <- matrix(rep(NA,sizeind * N),nrow = N ,ncol=sizeind)
rownames(appartenancepikp2) <- paste("k=",k)
colnames(appartenancepikp2) <- paste("ech",1:sizeind)
pikp2 <-  matrix(rep(NA,sizeind * N),nrow = N ,ncol=sizeind)
rownames(pikp2) <- paste("k=",k)
colnames(pikp2) <- paste("ech",1:sizeind)
  for (i in 1:sizeind){
    for(j in 1:N){
        appartenancepikp2[j,i] <- un_dedans(ind[i,],k[j])
        appartenancepikp2[appartenancepikp2==TRUE] <- 1
        appartenancepikp2[appartenancepikp2==FALSE] <- 0
        pikp2[j,i] <- round(appartenancepikp2[j,i] / n,2)
    }

  }
  head(appartenancepikp2,6)
```

```
##      ech 1 ech 2 ech 3
## k= 1     1     0     1
## k= 2     0     1     0
## k= 3     0     1     1
## k= 4     1     0     0
## k= 5     0     0     1
## k= 6     1     1     0
```

```
  head(pikp2,6)
```

```
##      ech 1 ech 2 ech 3
## k= 1  0,33  0,00  0,33
## k= 2  0,00  0,33  0,00
## k= 3  0,00  0,33  0,33
## k= 4  0,33  0,00  0,00
## k= 5  0,00  0,00  0,33
## k= 6  0,33  0,33  0,00
```

- Probabilité d'inclusion du second ordre

```
appartenancepiklp2 <- matrix(rep(NA,sizek * sizeind),nrow = sizek ,ncol=sizeind)
rownames(appartenancepiklp2) <- paste(coupleij[,1],coupleij[,2],sep = ",")
colnames(appartenancepiklp2) <- paste("ech",1:sizeind)
piklp2<-  matrix(rep(NA,sizek * sizeind),nrow = sizek ,ncol=sizeind)
rownames(piklp2) <- paste(coupleij[,1],coupleij[,2],sep = ",")
colnames(piklp2) <- paste("ech",1:sizeind)
  for (i in 1:sizek){
    for(j in 1:sizeind){
        appartenancepiklp2[i,j] <- deux_dedans(ind[j,],coupleij[i,1],coupleij[i,2])
        appartenancepiklp2[appartenancepiklp2==TRUE] <- 1
        appartenancepiklp2[appartenancepiklp2==FALSE] <- 0
        piklp2[i,j] <- round(appartenancepiklp2[i,j] / n,2)
    }
```

```
 }
head(appartenancepiklp2,6)
```

```
##     ech 1 ech 2 ech 3
## 1,2     0     0     0
## 1,3     0     0     1
## 1,4     1     0     0
## 1,5     0     0     1
## 1,6     1     0     0
## 2,3     0     1     0
```

```
head(piklp2,6)
```

```
##     ech 1 ech 2 ech 3
## 1,2  0,00  0,00  0,00
## 1,3  0,00  0,00  0,33
## 1,4  0,33  0,00  0,00
## 1,5  0,00  0,00  0,33
## 1,6  0,33  0,00  0,00
## 2,3  0,00  0,33  0,00
```

(b): La valeur de $\bar{y_u}$

```
y <- c(98,102,154,133,190,175)
ybarmu <- mean(y)
ybarmu
```

```
## [1] 142
```

(c) Soit $\bar{y}$ la moyenne des valeurs de l'échantillon. Pour chacun des plans trouvons

Plan1

i. $\mathbf{E}[\bar{y}]$:

```
# Moyenne de chaque echantillon du plan p1
meanechtp1 <- rep(NA,sizeech)
for (i in 1:sizeech){
  meanechtp1[i] <- mean(y[x[i,]])
}
meanechtp1
```

```
## [1] 147,3333 142,3333 140,3333 135,3333 148,6667 143,6667 141,6667 136,6667
```

```
# Moyenne de chaque echantillon du plan p1
meanechtp2 <- rep(NA,sizeind)
for (i in 1:sizeind){
  meanechtp2[i] <- mean(y[ind[i,]])
}
meanechtp2
```

```
## [1] 135,3333 143,6667 147,3333
```

```
probp1 <- plan1[,4]
espybarp1 <- sum(probp1 * meanechtp1)
espybarp1
```

```
## [1] 142
```

ii. $\mathbf{V}[\bar{y}]$:

Formule de la variance

$\mathrm{V}[\bar{y}] = E[\bar{y^2}] - \mathrm{E}[\bar{y}]^2$

```
varybarp1 <- sum(meanechtp1^2*probp1) - espybarp1^2
varybarp1
```

```
## [1] 18,94444
```

iii.Biais[$\bar{y}$] :

$\mathrm{Biais}[\bar{y}] = \mathrm{E}[\bar{y}] - \bar{y}_\mu$

```
biaisybarp1 <- espybarp1 - ybarmu
biaisybarp1
```

```
## [1] 0
```

iiv.EQM[$\bar{y}$] :

Comme le biais est nul donc $\mathrm{EQM}[\bar{y}] = \mathrm{V}[\bar{y}]$

```
eqmybarp1 <- varybarp1
```

Plan2

i. E[$\bar{y}$] :

```
probp2 <- plan2[,4]
espybarp2 <- sum(probp2 * meanechtp2)
espybarp2
```

```
## [1] 142,5
```

ii. V[$\bar{y}$] :

```
varybarp2 <- sum(meanechtp2^2*probp2) - espybarp2^2
varybarp2
```

```
## [1] 19,36111
```

iii.Biais[$\bar{y}$] :

```
biaisybarp2 <- espybarp2 - ybarmu
biaisybarp2
```

```
## [1] 0,5
```

iiv.EQM[$\bar{y}$] :

```
eqmybarp2 <- varybarp2 - biaisybarp2^2
eqmybarp2
```

```
## [1] 19,11111
```

(d)Lequel des plans est le meilleur ? Pourquoi ?

Le plan2 est le meilleur car il minimise le EQM

2. Pour la population utilisée comme exemple en classe, formée de 8 individus.

$$U = \{1, 2, 3, 4, 5, 6, 7, 8\}$$

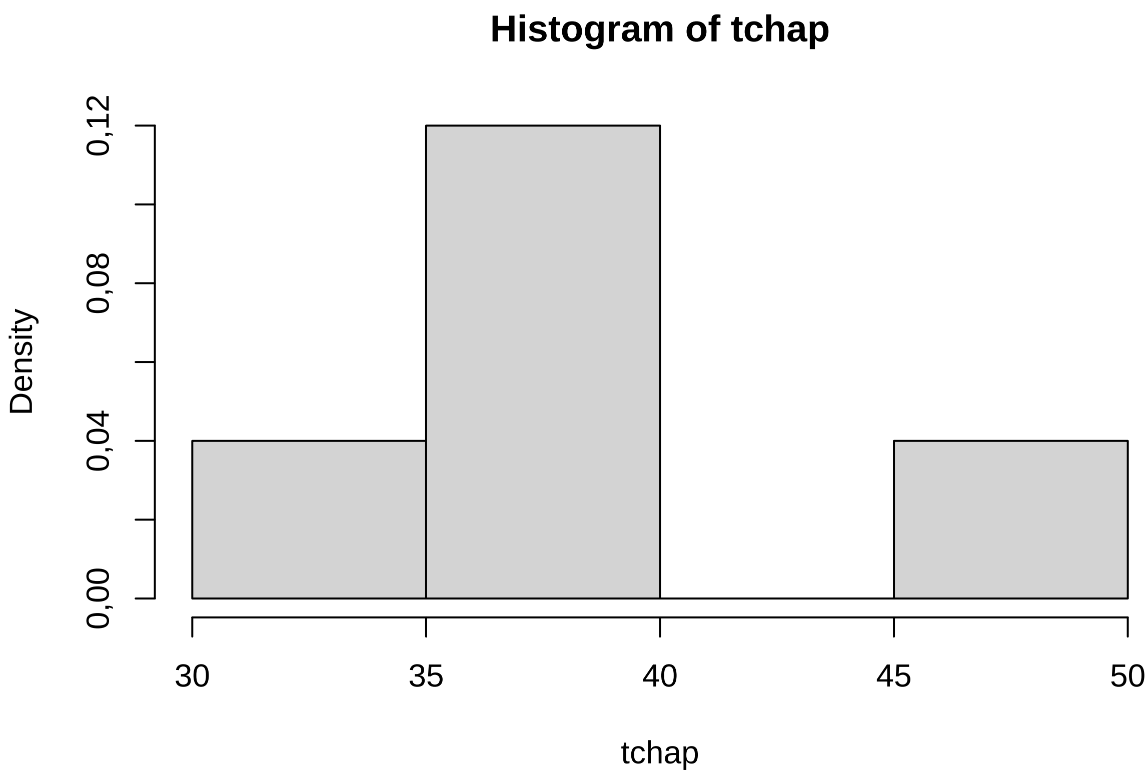
(a) Déterminons la probabilité de sélection π_i , pour chaque unité i.

Le nombre total des individus est de 8. Donc la probabilité de selectionnée chaque individu est de $\frac{1}{8}$

$$\boxed{\pi_i = \frac{1}{8}}$$

(b) Quelle est dans ce contexte la distribution de $\widehat{t} = 8\bar{y}$?

```r
# extraction des echantillons
echtp3 <- plan[,1:4]
meanechtp3 <- apply(echtp3,1,mean)
tchap <- 8 * meanechtp3
hist(tchap,probability = TRUE)
```




```r
shapiro.test(tchap)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  tchap
## W = 0,93998, p-value = 0,6658
```

Le test de normalité donne un p-value > 0.05 donc la distrubition de $\widehat{t}$ est une loi normale.