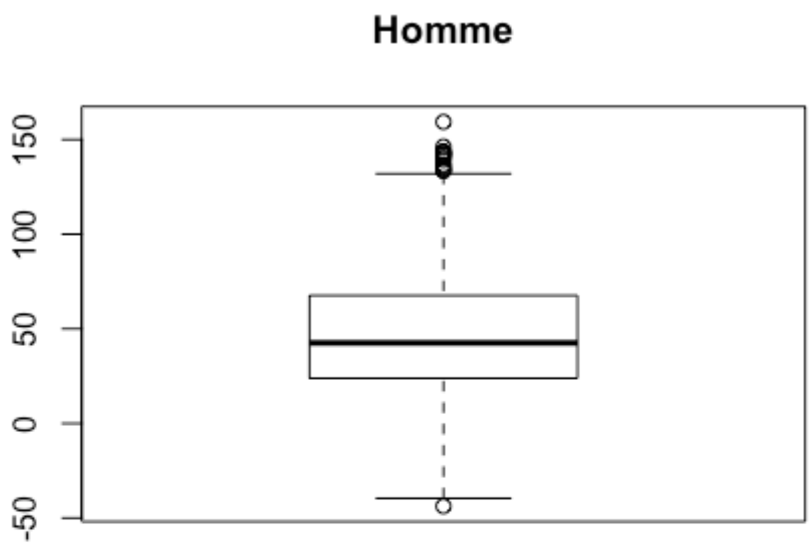
projet

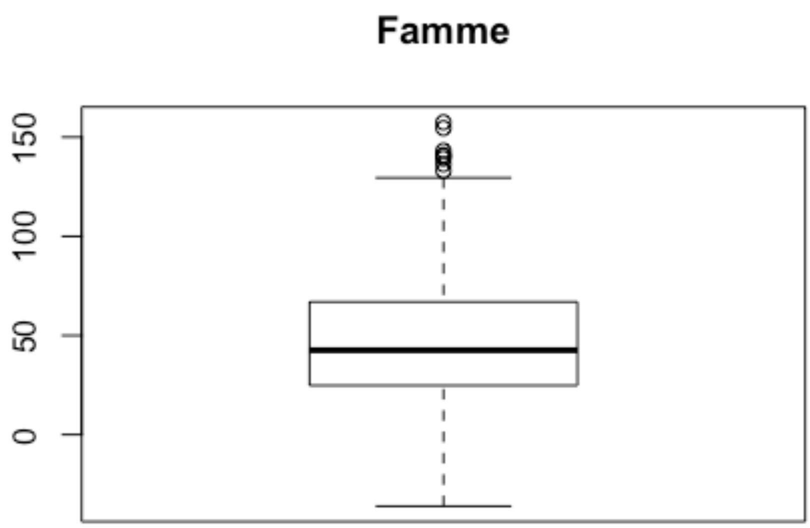
1 Statistique Descriptives

(a) Tracer les boîte des femmes et les hommes

```
rm(list = ls())
 data=read.csv("DB_binome_56.csv",header = TRUE)
 hommes <-subset(data,data['Sexe']==0)
 femmes<-subset(data,data['Sexe']==1)
 boxplot(hommes['Peche'],TRUE,main = "Homme")
```



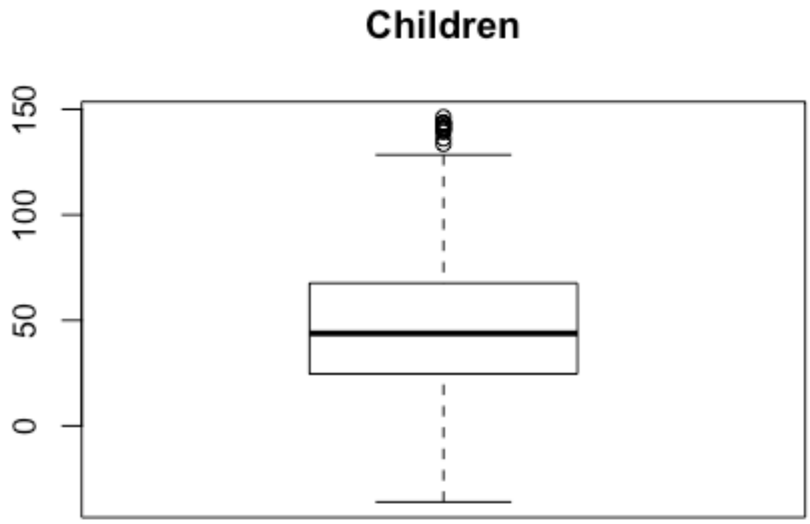

```
 boxplot(femmes['Peche'],TRUE,main = "Famme")
```



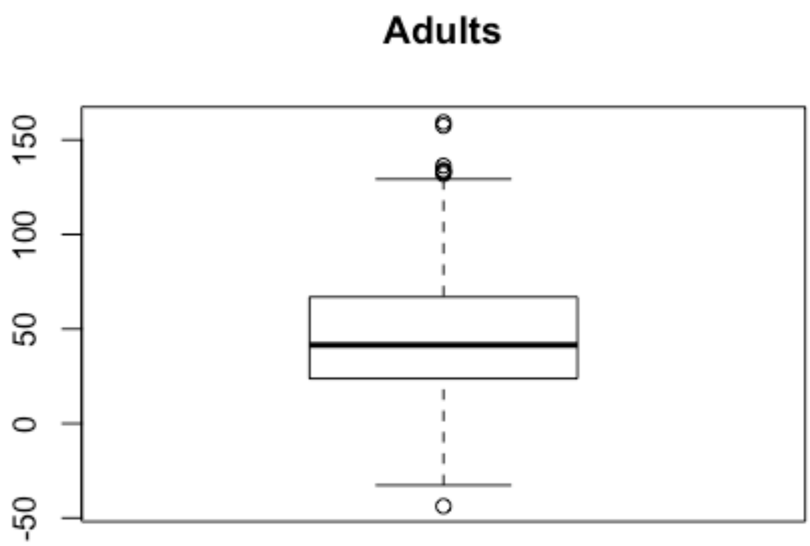
On remarque que il n’y a pasune grande defference entre les hommes et les femmes.

(b)Tracer la quantité de péche en fonction de la tranche d'age

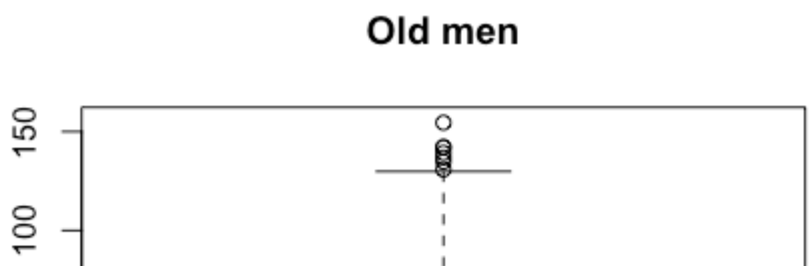
```
rm(list = ls())
data=read.csv("DB_binome_56.csv",header = TRUE)
children<-subset(data,data['Age']==0)
adults<-subset(data,data['Age']==1)
oldmen<-subset(data,data['Age']==2)
boxplot(children['Peche'],TRUE,main = "Children")
```

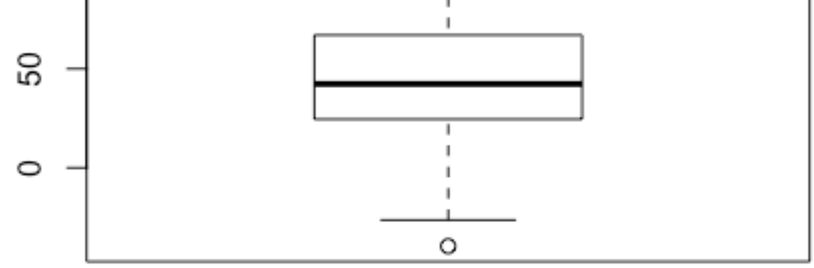



```
boxplot(adults['Peche'],TRUE,main = "Adults")
```




```
boxplot(oldmen['Peche'],TRUE,main = "Old men")
```

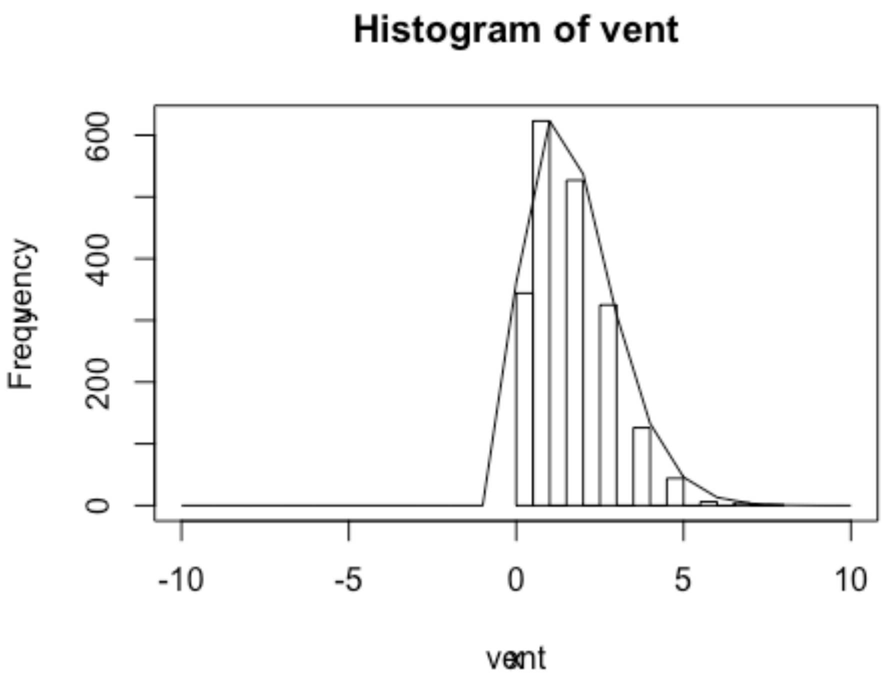




On remarque que il n’y a pasune grande defference entre les gens qui ont les age defferents.

(c)Tracer l’histogramme l’intensité de vent

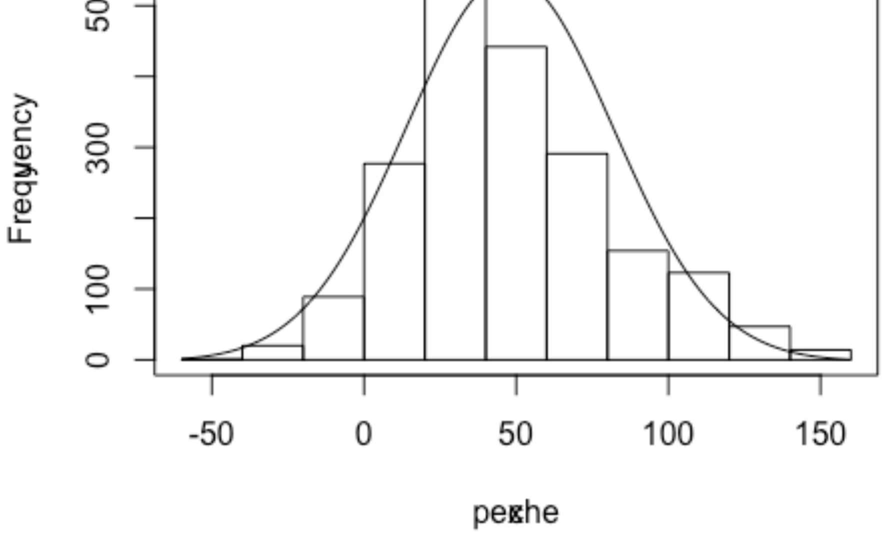
```
 vent<-data[,'Noeuds']
hist(vent,freq=TRUE,xlim=c(-10,10))
lambda=mean(vent)

x=seq(-10,10)
y<-dpois(x,lambda)
par(new=TRUE)
plot(x,y,type="l",xaxt="n",yaxt="n")
```



Elle suive la loi de Poisson . On le justifie par la courbe de “dpoi”

(d)Tracer l’histogramme la quantité de péche

```
 peche<-data[,'Peche']
hist(peche,freq=TRUE,xlim=c(-60,160))

mu=mean(peche)
delta=sqrt(var(peche))
x=seq(-60,160)
y<-dnorm(x,mu,delta)
par(new=TRUE)
plot(x,y,type="l",xaxt="n",yaxt="n")
```



Elle suive la loi de Normale . On le justifie par la courbe de “dnorm”

2

(a)Verifier les observation

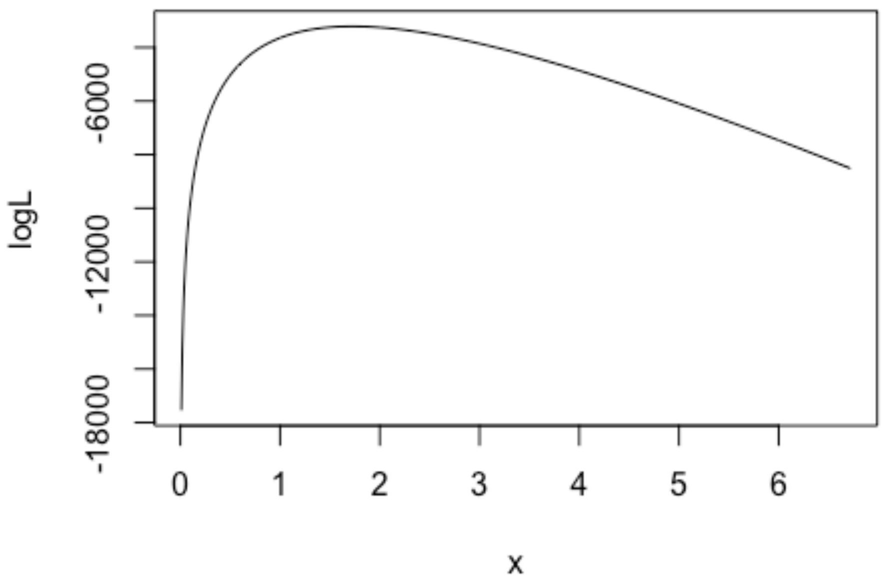
```
nb = length(vent)
lambda = mean(vent)
echant =  rpois(nb , lambda)
ks.test(vent, echant)
## Warning in ks.test(vent, echant): p-value will be approximate in the
## presence of ties
##
##  Two-sample Kolmogorov-Smirnov test
##
## data:  vent and echant
## D = 0.0055, p-value = 1
## alternative hypothesis: two-sided
```

Donc on remarque la p-value est presque 1 > 0.05 , donc il suivre la loi poisson

(b)La vraisemblance

(c) Tracer La log-vraisemblance et déterminer le maximun de vraisemblance

```
lambda=mean(vent)

x<-seq(0.01,lambda+5,0.01)
logL=vector(length=length(x))
n<-length(vent)

max=-100000
maxX=0

for( i in seq(1,length(x))){
 sum=0
 for(j in seq(1,length(vent))){
 sum = sum + vent[j]*log(x[i]) - log(factorial(vent[j]))
 }

 logL[i]=-n*x[i] +sum

 if(logL[i]>max){
   max=logL[i]
   maxX=x[i]
 }
}

plot(x,logL,type="l",xaxt="lambda",yaxt="logL")
```

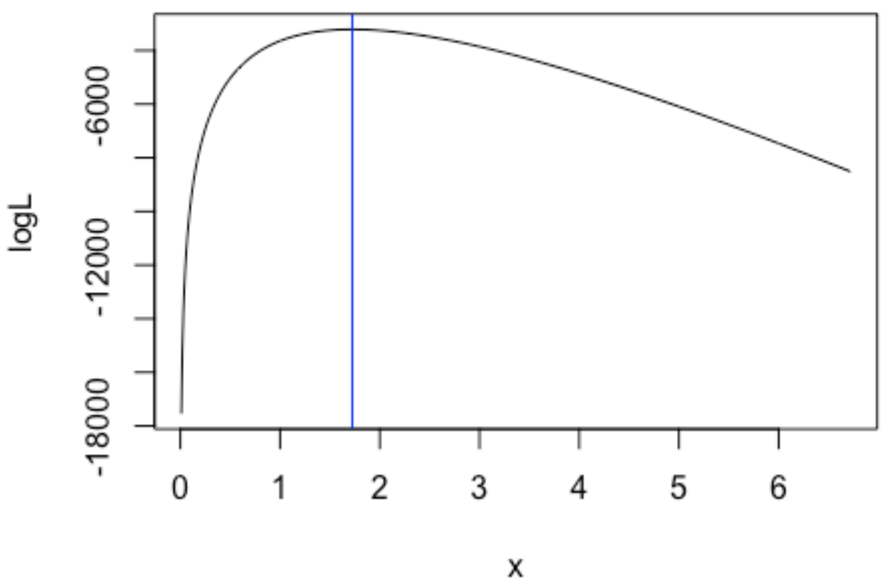


#

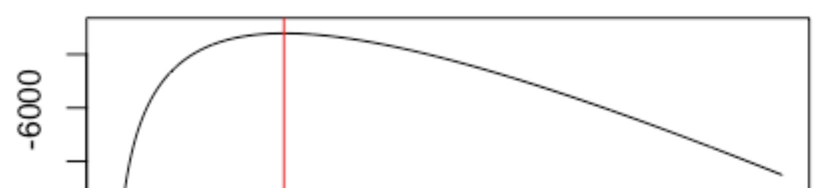
(d) donner une estimation du parametre ,et faire apparaitre sur le graphique

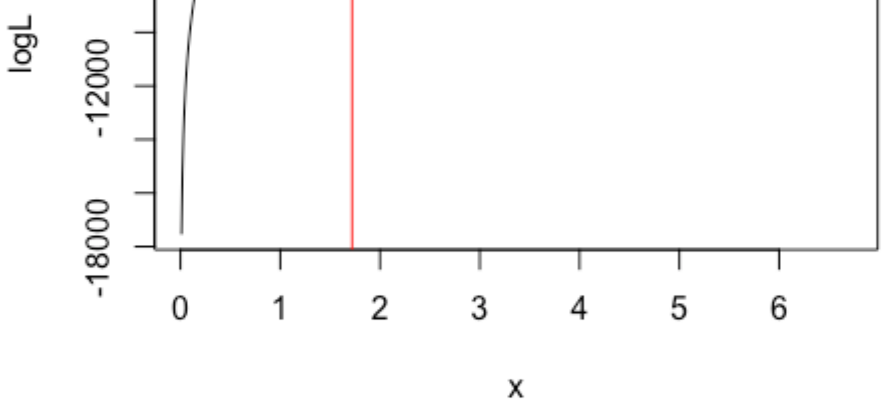
je utilise deux facos pour estimer le parametre

```
plot(x,logL,type="l",xaxt="lambda",yaxt="logL")
par(new=TRUE)

abline(v=lambda,col = "blue")
```




```
plot(x,logL,type="l",xaxt="lambda",yaxt="logL")
par(new=TRUE)

abline(v=maxX,col = "red")
```






```
cat("1er",lambda,"2eme",maxX)
## 1er 1.724 2eme 1.72
```

(e)écrire la vraisemblance théorique du modèle de la quantité de péche

pour lambda = 1.724 , ici les

sont omises du calcul .

```
repartition<-function(xi,sigma){
 m<-qpois(p=0.999999,lambda=lambda)
 proba<-dpois(seq(0,m,1),lambda)
  f=vector(length = m)
   sum<-0
 for(j in seq(1,m)){

     norm<- 1/sqrt(2*pi)/sigma*exp(-((xi-100/(j))/sigma)^2/2)

   f[j]=proba[j]*norm
   sum=sum+f[j]
 }
   return(sum)
}
```

(f)TCL

On pose X=

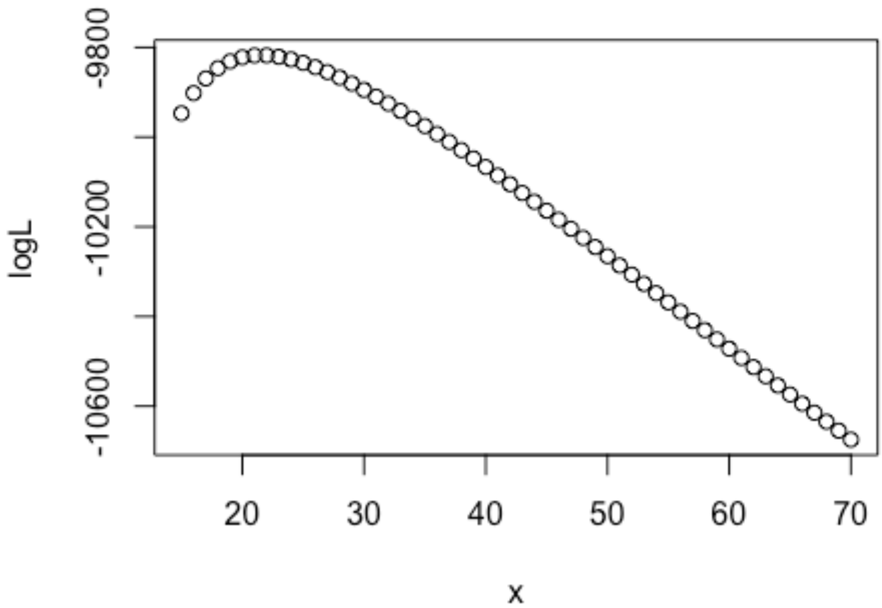
~

(g)Déterminer le maximum de vraissemblance du parametre

, tracer la log-vraissemblance de l’échantillon en fonction de la valeur du avec chaqu’une des méthodes

```
sum<-0
x<-seq(15,70,1)
logL=vector(length=length(x))
 max=-100000
```

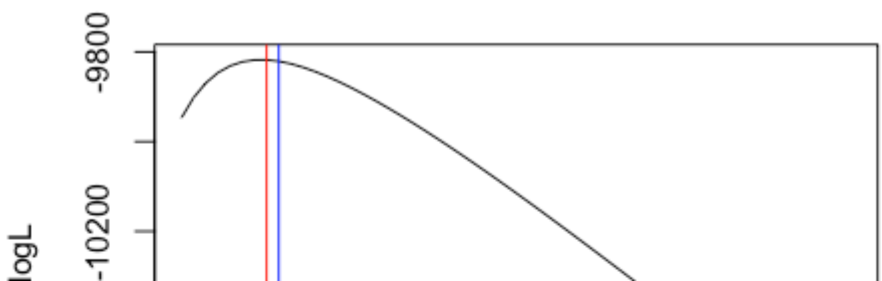
```
 max=-100000
maxX=0
second=-100000
secondX=0

for(j in seq(1,length(x))){
 sum<-0
 for(i in seq(1,length(peche))){
    sum=sum+log(repartition(peche[i],x[j]))
    }

 logL[j]=sum

 if(logL[j]>max){
   max=logL[j]
   maxX=x[j]
 }else if (logL[j]>second){
   second=logL[j]
   secondX=x[j]
 }
 }
plot(x,logL)
```

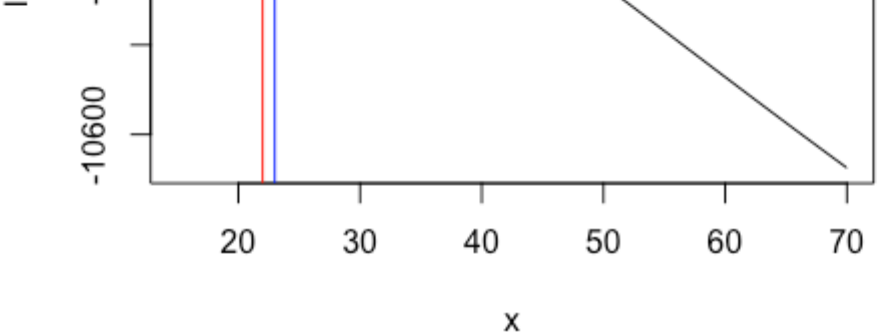


#

(h)Donner une estimation du parametre avec les deux approches et la faire apparaitre sur les graphiques

```
cat("les deux approches",maxX,secondX)
## les deux approches 22 23
cat("estimateur=",maxX/2+secondX/2)
## estimateur= 22.5
plot(x,logL,type="l",xaxt="lambda",yaxt="logL")
par(new=TRUE)

abline(v=maxX,col = "red")
par(new=TRUE)
abline(v=secondX,col = "blue")
```





3

(1)

On a deja donc on a ~ donc ~ avec on pose que , , donc on a

(2)

```
Intervalle=function(lambda, alpha, n) {
 theta1 = lambda + sqrt(lambda/n) * qnorm(alpha/2, mean = 0, sd = 1, log = FALSE)
 theta2 = lambda + sqrt(lambda/n) * qnorm(1-alpha/2, mean = 0, sd = 1, log = FALSE)
 paste("Intervalle de confiance : [ ", theta1, ", ", theta2, "]", sep = "")
}
Intervalle(lambda, 0.05, 2000)
## [1] "Intervalle de confiance : [ 1.66645577784488, 1.78154422215512]"
```

4 Tests

(a)

On pose que H0:la quantité de pêche ne dépend pas du sexe H1:la quantité de péche dépend du sexe

```
quantiteH <-subset(data,data['Sexe']==1)[,'Peche']

quantiteF <-subset(data,data['Sexe']==0)[,'Peche']

t.test(quantiteH,quantiteF,paired = FALSE)
##
##  Welch Two Sample t-test
##
## data:  quantiteH and quantiteF
## t = -0.13155, df = 1817.6, p-value = 0.8954
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.219464  2.814720
## sample estimates:
## mean of x mean of y
##  47.73338  47.93575
```

On a p-value = 0.8953 > 0.05, donc on ne rejette pas l'hypothèse 0. #(b) Test :

et

si

, on rejette le Hypothese

```
trois_Noeuds <-subset(data,data['Noeuds']==3)[,'Peche']
sum=0
for(i in seq(1,length(trois_Noeuds))){

 sum=sum+(trois_Noeuds[i]-100/(3+1))^2
```

```
}

 cat(sum,"inferieur a",20^2*qchisq(p=0.999, df=length(trois_Noeuds)),"superieur
a",20^2*qchisq(p=0.001, df=length(trois_Noeuds)))
## 161348.9 inferieur a 163806 superieur a 100749.8
if(sum>20^2*qchisq(p=1, df=length(trois_Noeuds))){
 print("on rejette H0")
}else if(sum<20^2*qchisq(p=0.005, df=length(trois_Noeuds))){
 print("on rejette H0")
}else{
 print("on ne rejette pas H0")
}
## [1] "on ne rejette pas H0"
```

On remarque que pour ces datas , H0 n’est pas rejet quand alpha = 0.001