TPSTAT_CHEN_Zeyu_GROUP_2.1

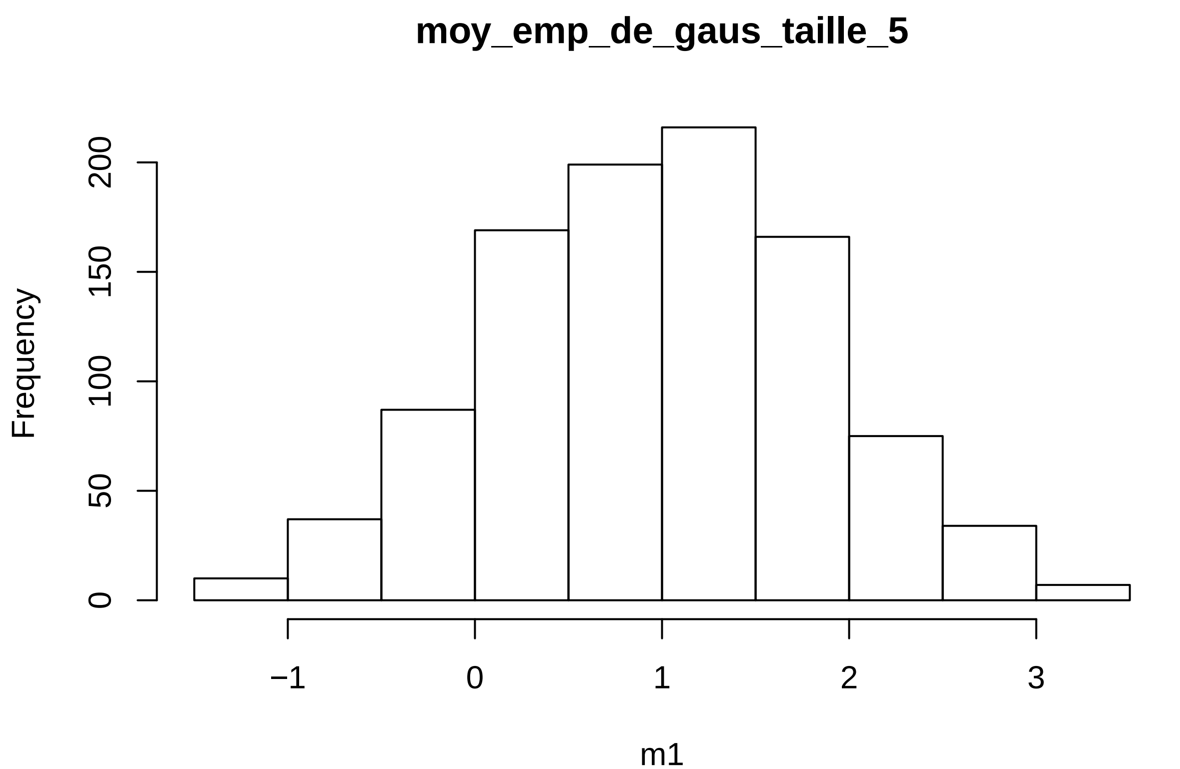
Zeyu

2018/3/5

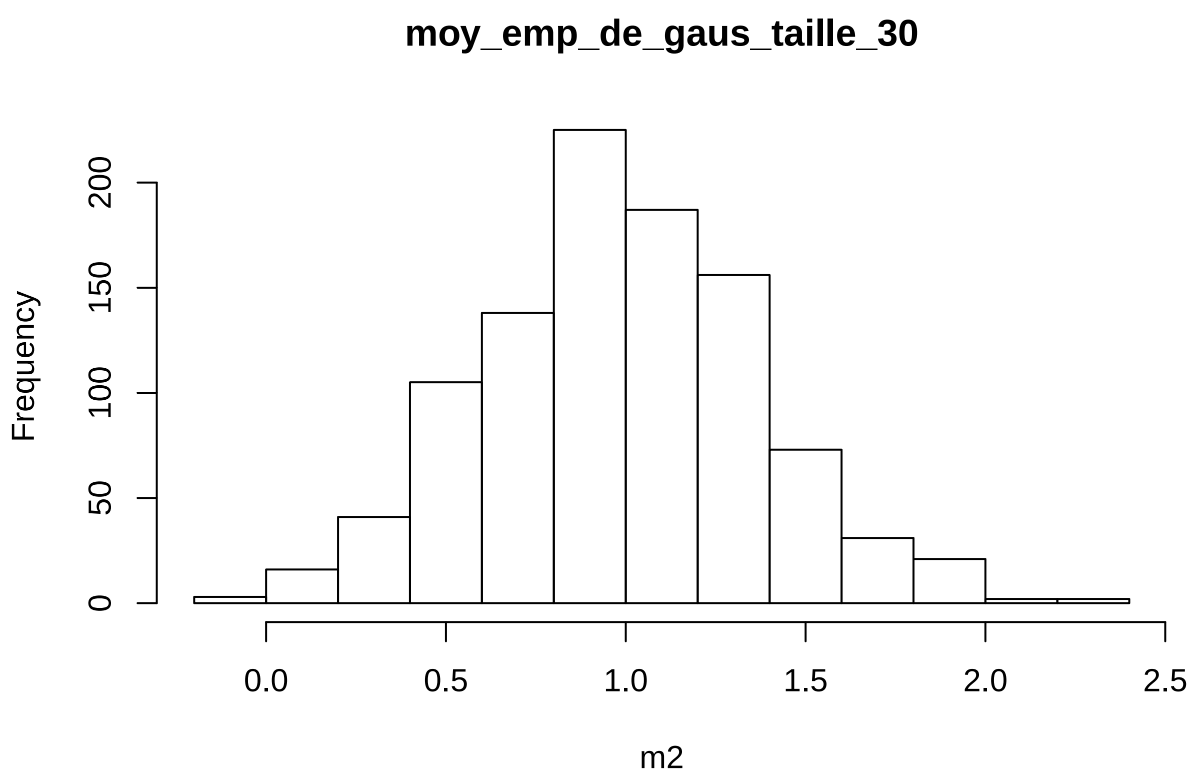
1.Echantillon, Théorème Central Limite, Estimation Monte Carlo

<–1–>

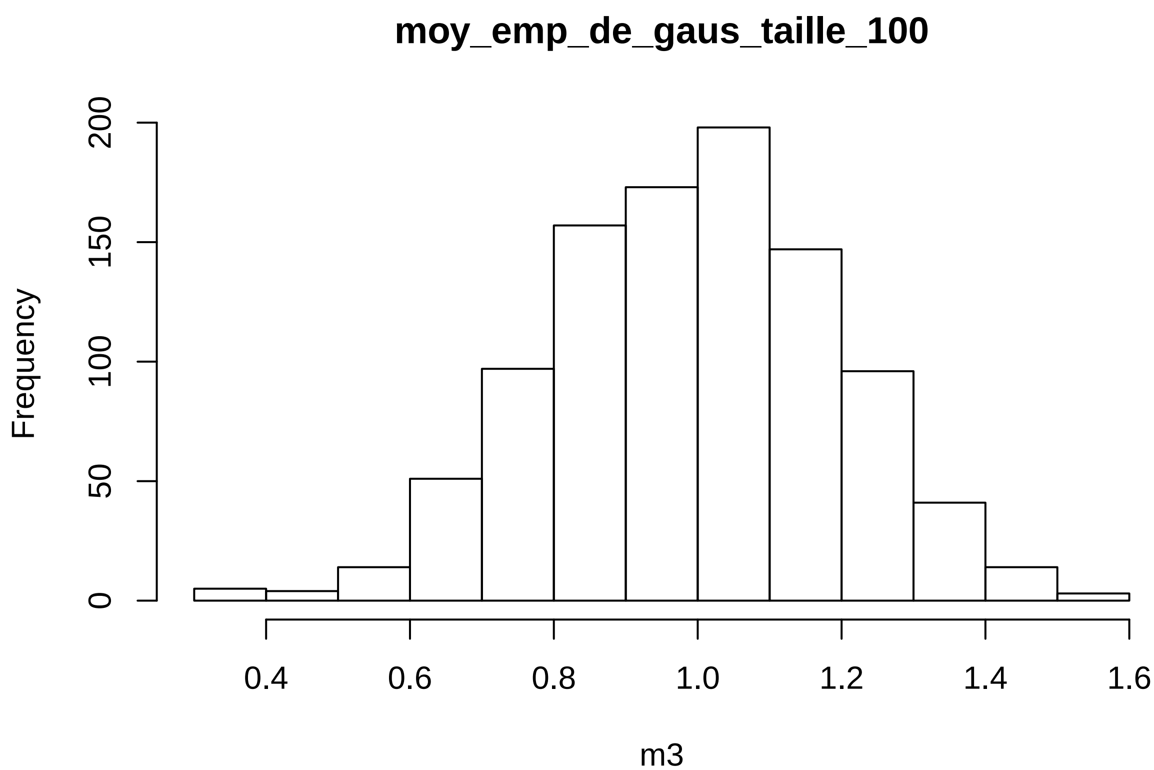
```
# N est la fois d'échantillons
# f est la fonction qu'on va simuler
# n taille d'échantillon, su est mu, o est sigma
sim.fun <-function (N,f,n,u,o)
  {
    sample<-1:N
    for (i in 1:N) {
        sample[i] <-f(n,u,o)
     }
    return(sample)
}

moyenne=function(n,mu,sigma)
  {
  r=rnorm(n,mu,sigma);
  return(mean(r));
}

varience=function(n,mu,sigma)
  {
  r=rnorm(n,mu,sigma);
  return(var(r));
}
m1 = sim.fun(1000,moyenne,5,1,2)
m2 = sim.fun(1000,moyenne,30,1,2)
m3 = sim.fun(1000,moyenne,100,1,2)
v1 = sim.fun(1000,varience,5,1,2)
v2 = sim.fun(1000,varience,30,1,2)
v3 = sim.fun(1000,varience,100,1,2)
hist(m1,freq =T ,main ="moy_emp_de_gaus_taille_5")
```



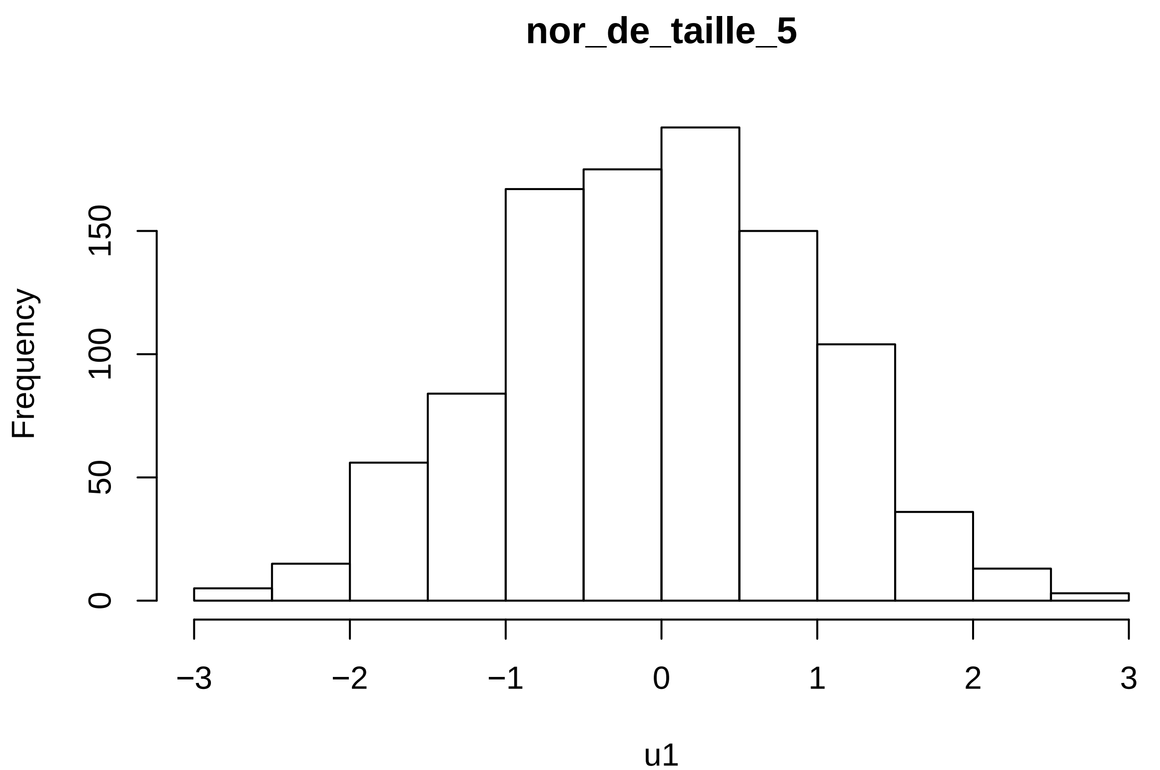

```
hist(m2,freq =T ,main ="moy_emp_de_gaus_taille_30")
```



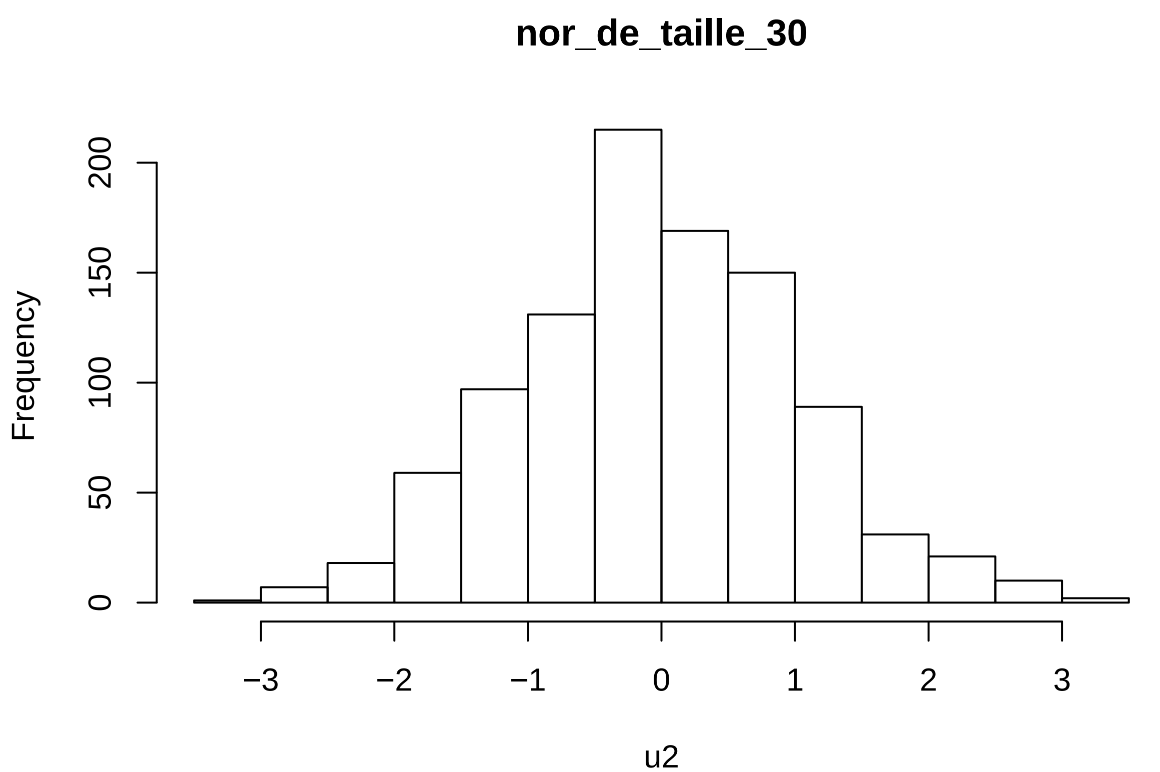

```
hist(m3,freq =T ,main ="moy_emp_de_gaus_taille_100")
```



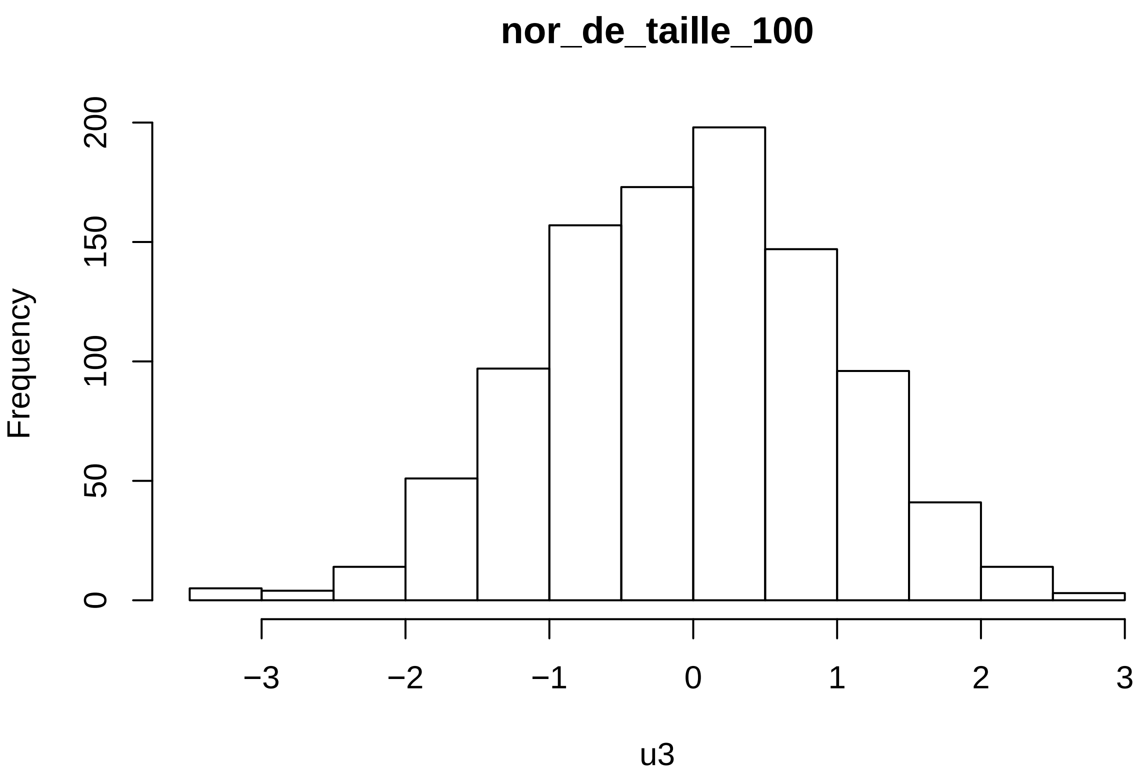

```
renormalisation<-function(N,moyenne_epq,n,mu,sigma)
{
  for ( i in 1:N)
  {
    moyenne_epq[i]=(moyenne_epq[i]-mu)/(sigma/sqrt(n));
  }
  return(moyenne_epq)
}

u1 = renormalisation(1000,m1,5,1,2);
u2 = renormalisation(1000,m2,30,1,2);
u3 = renormalisation(1000,m3,100,1,2);
hist(u1,freq =T ,main ="nor_de_taille_5")
```




```r
hist(u2,freq =T ,main ="nor_de_taille_30")
```




```r
hist(u3,freq =T ,main ="nor_de_taille_100")
```



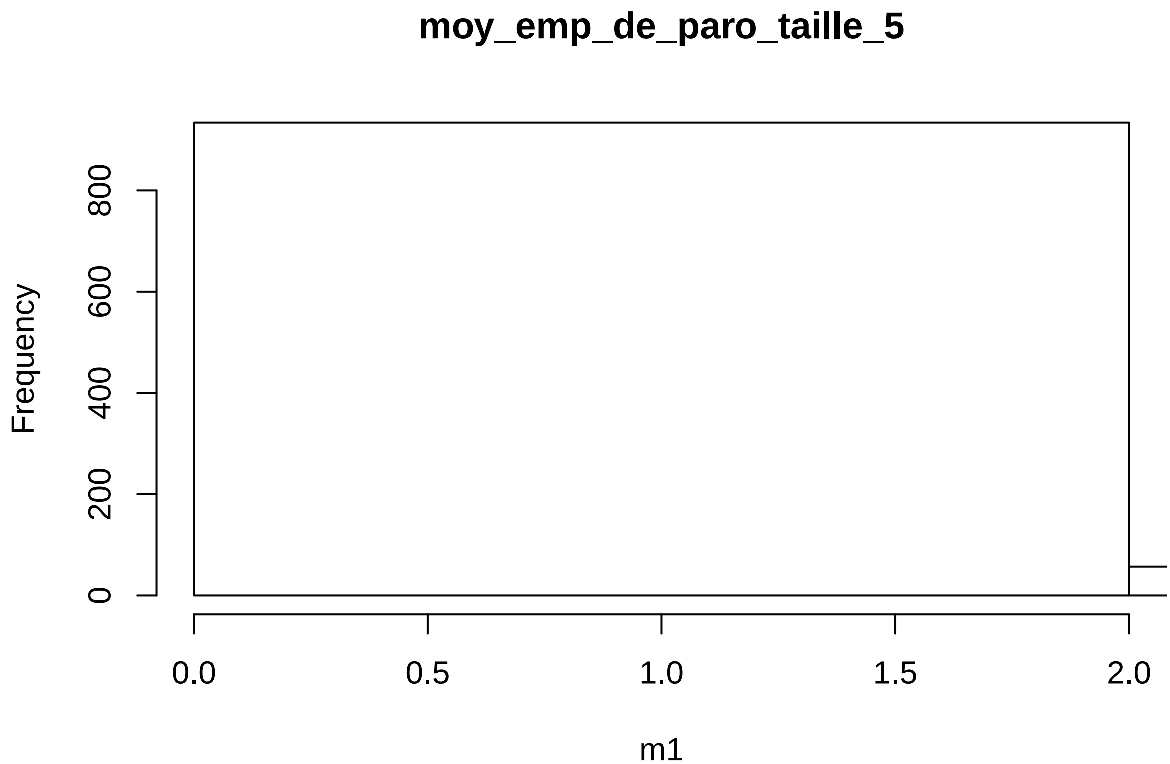
<–**2**–>

```r
library(rmutil)
```

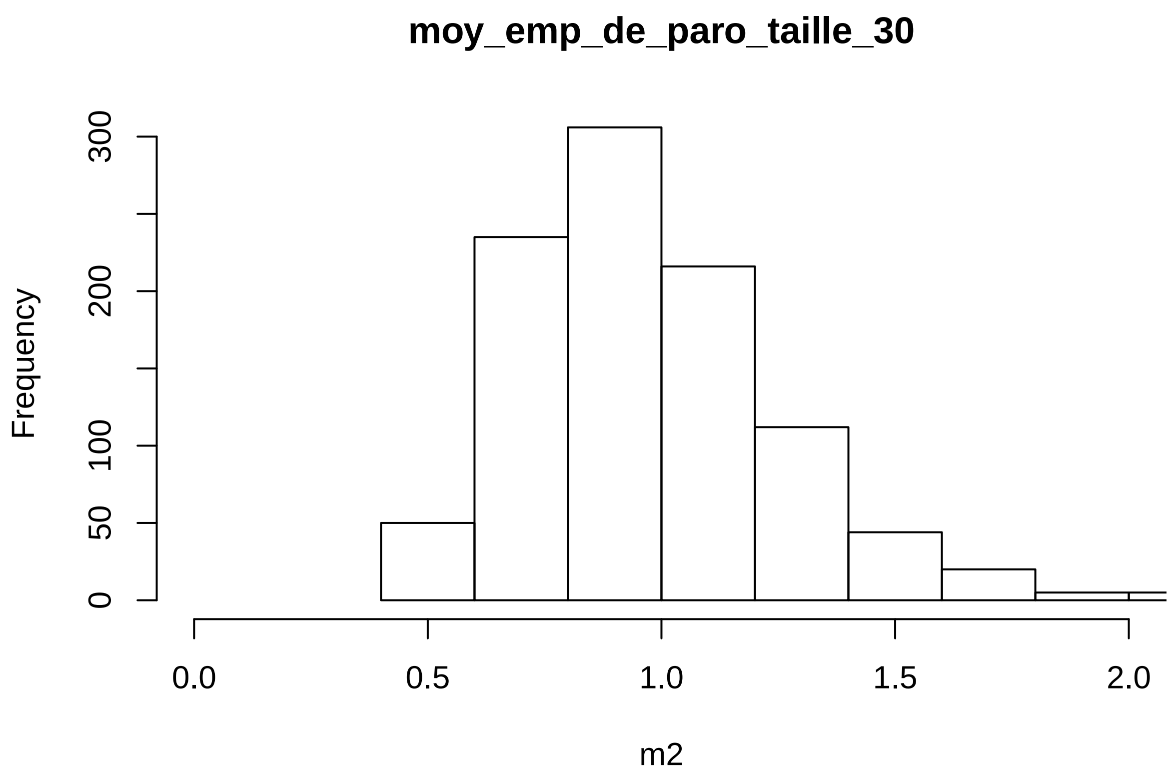
```
##
## Attaching package: 'rmutil'

## The following object is masked from 'package:stats':
##
##     nobs
```

```r
# N est la fois d'échantillons
# f est la fonction qu'on va simuler
# n taille d'échantillon
sim.fun <-function (N,f,n,m,s)
  {
    sample<-1:N
    for (i in 1:N) {
        sample[i] <-f(n,m,s)
     }
    return(sample)
}
moyenne=function(n,m,s)
  {
  r=rpareto(n,m,s);
  return(mean(r));
}

varience=function(n,m,s)
```

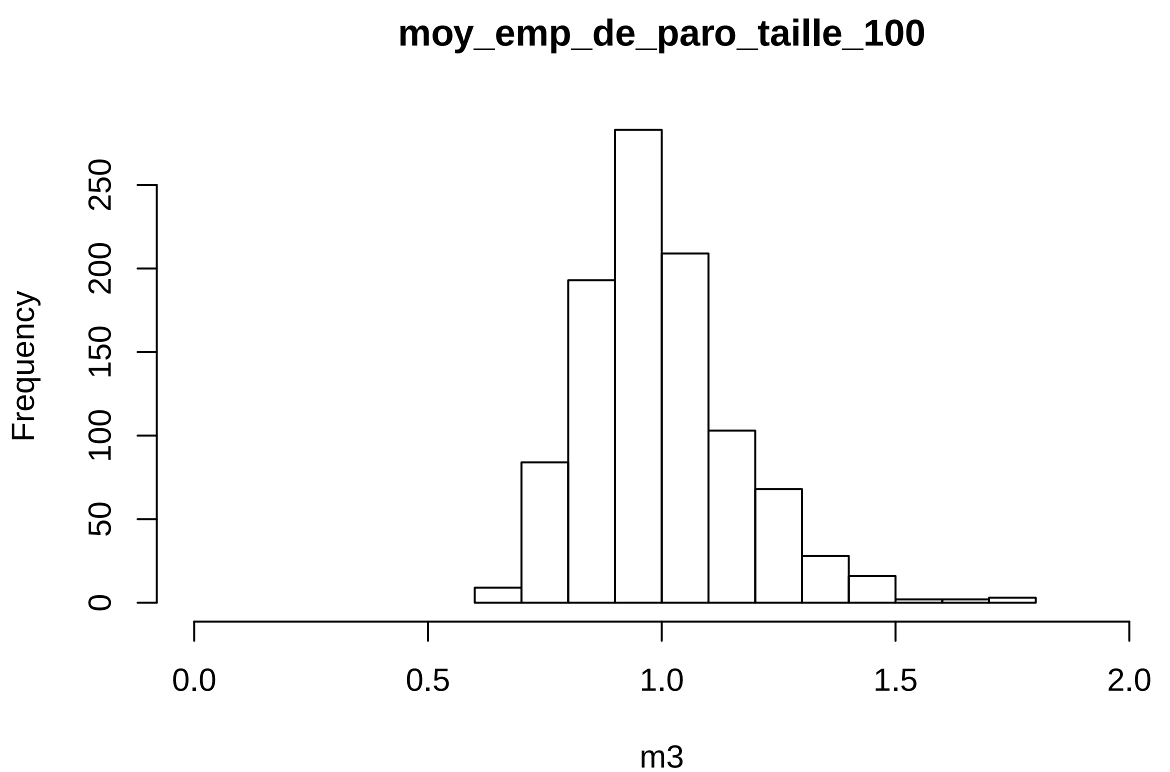
```
  {
  r=rpareto(n,m,s);
  return(var(r));
}
m1 = sim.fun(1000,moyenne,5,1,3)
m2 = sim.fun(1000,moyenne,30,1,3)
m3 = sim.fun(1000,moyenne,100,1,3)
v1 = sim.fun(1000,varience,5,1,3)
v2 = sim.fun(1000,varience,30,1,3)
v3 = sim.fun(1000,varience,100,1,3)
hist(m1,freq =T ,xlim=c(0,2),main ="moy_emp_de_paro_taille_5")
```




```
hist(m2,freq =T ,xlim=c(0,2),main ="moy_emp_de_paro_taille_30")
```




```
hist(m3,freq =T ,xlim=c(0,2),main ="moy_emp_de_paro_taille_100")
```

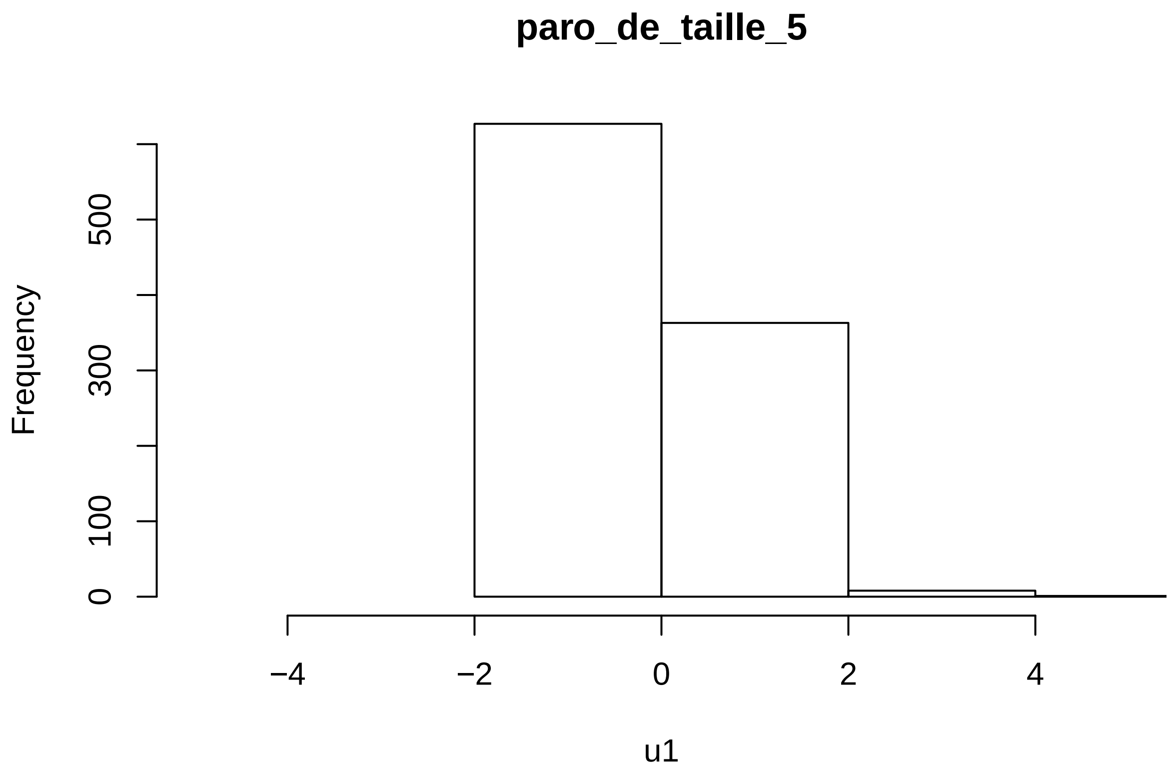



```
renormalisation<-function(N,moy_emp,n,mu,sigma)
{
  for ( i in 1:N)
  {
```

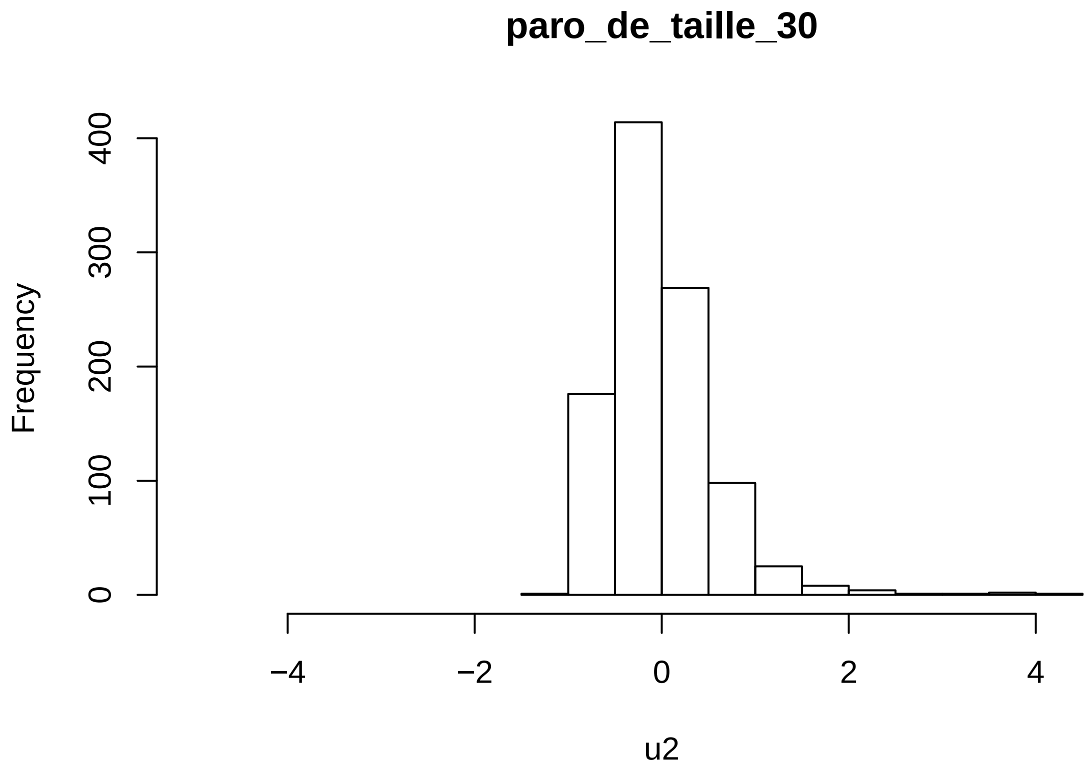
```r
    moy_emp[i]=(moy_emp[i]-mu)/(sigma/sqrt(n));
  }
  return(moy_emp)
}
```

$$a_n = m, b_n = s$$

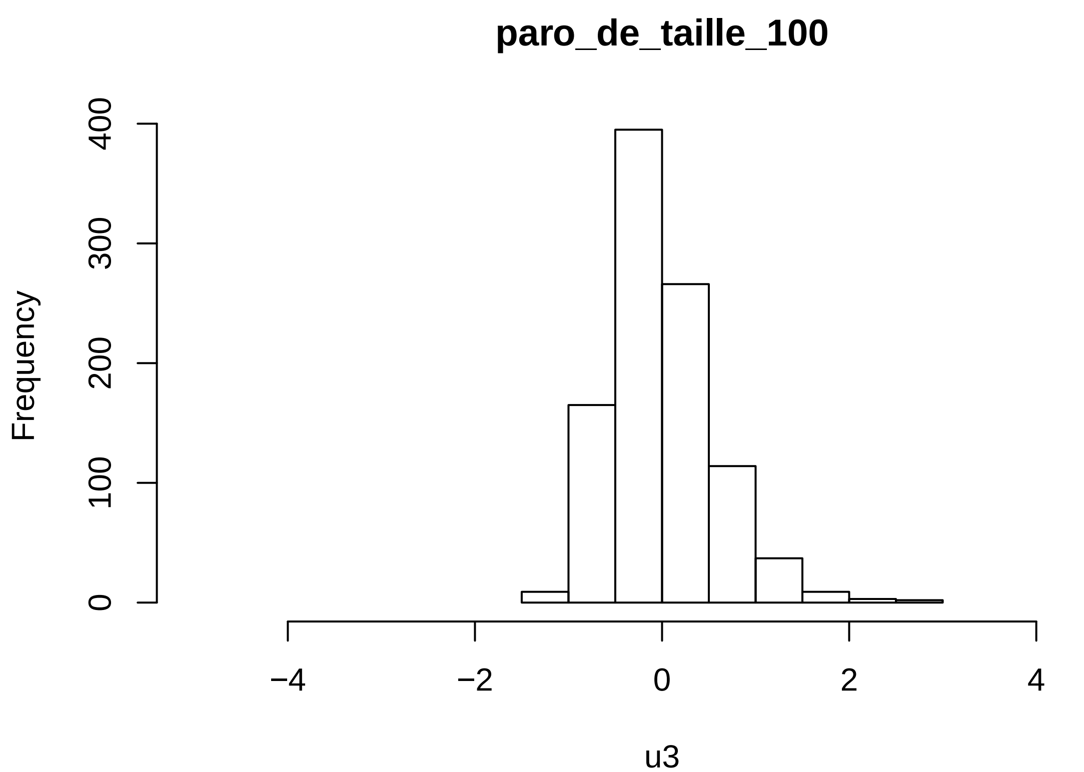
```r
u1 = renormalisation(1000,m1,5,1,3);
u2 = renormalisation(1000,m2,30,1,3);
u3 = renormalisation(1000,m3,100,1,3);
hist(u1,freq =T ,xlim=c(-5,5),main ="paro_de_taille_5")
```




```r
hist(u2,freq =T ,xlim=c(-5,5),main ="paro_de_taille_30")
```

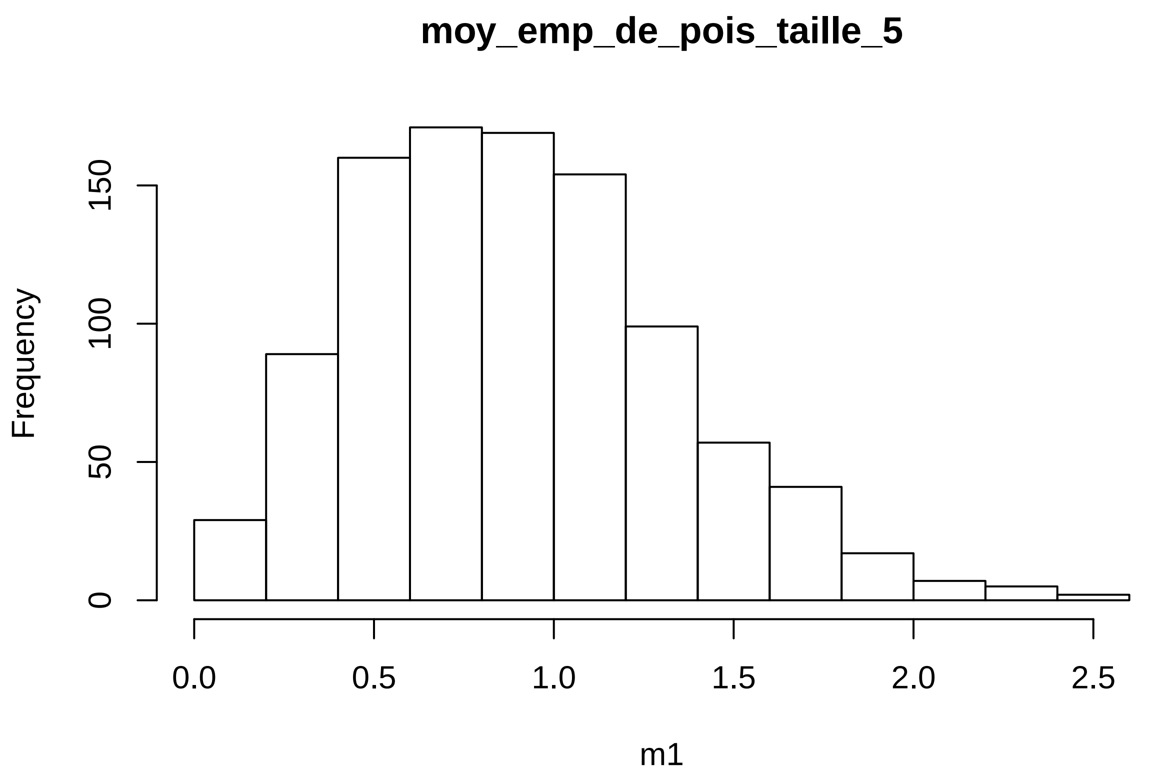



```r
hist(u3,freq =T ,xlim=c(-5,5),main ="paro_de_taille_100")
```

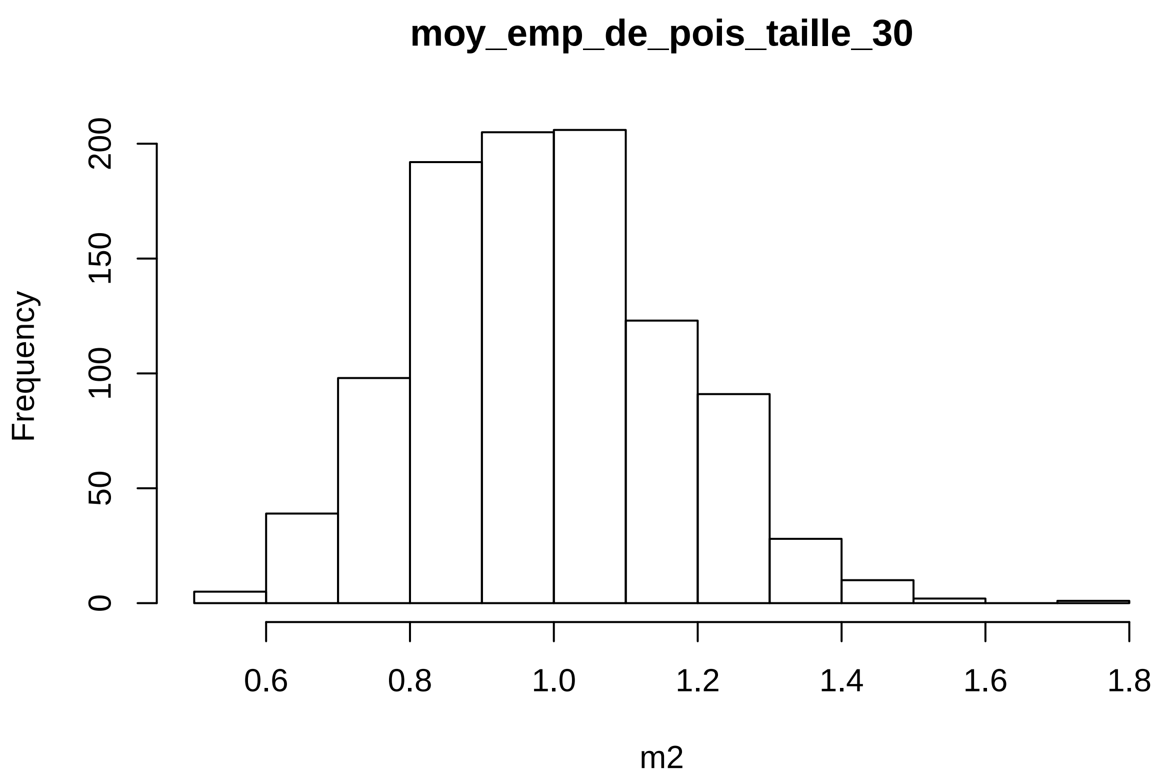


<–3–>

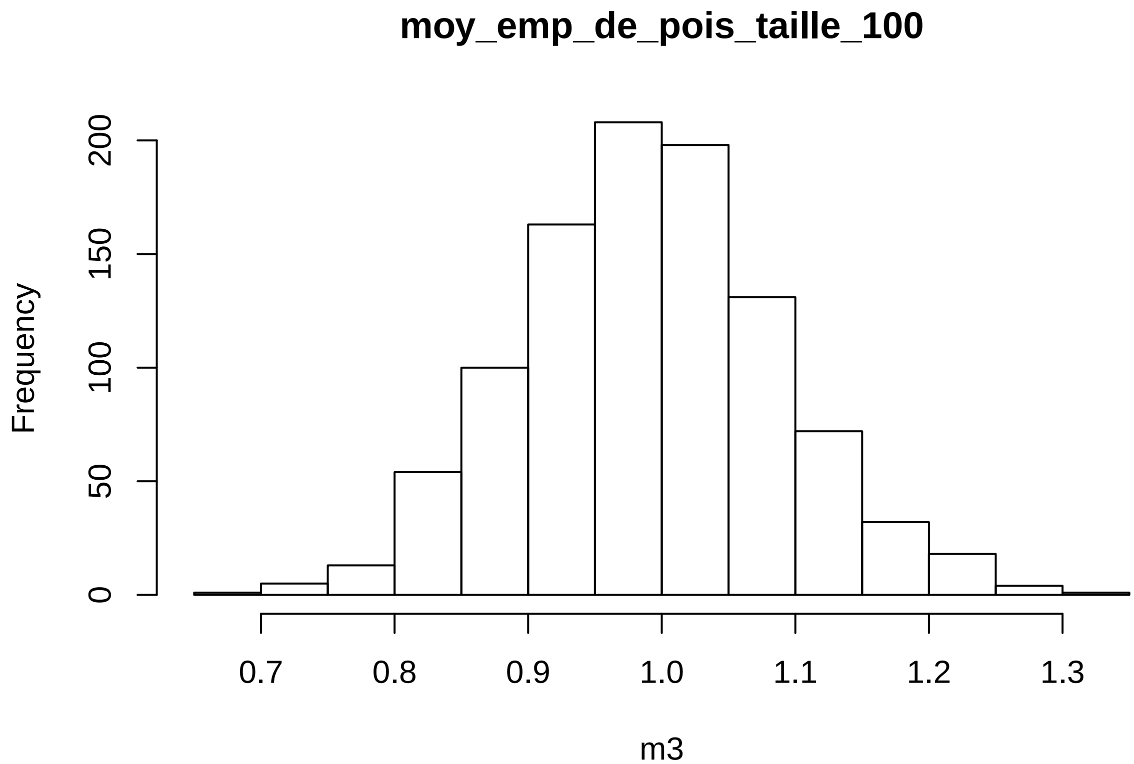
```r
# N est la fois d'échantillons
# f est la fonction qu'on va simuler
# n taille d'échantillon
sim.fun <-function (N,f,n,lamda)
  {
    sample<-1:N
    for (i in 1:N) {
        sample[i] <-f(n,lamda)
     }
    return(sample)
}
moyenne=function(n,lamda)
  {
  r=rpois(n,lamda);
  return(mean(r));
}

varience=function(n,lamda)
  {
  r=rpois(n,lamda);
  return(var(r));
}
m1 = sim.fun(1000,moyenne,5,1)
m2 = sim.fun(1000,moyenne,30,1)
m3 = sim.fun(1000,moyenne,100,1)
v1 = sim.fun(1000,varience,5,1)
v2 = sim.fun(1000,varience,30,1)
v3 = sim.fun(1000,varience,100,1)
hist(m1,freq =T ,main ="moy_emp_de_pois_taille_5")
```



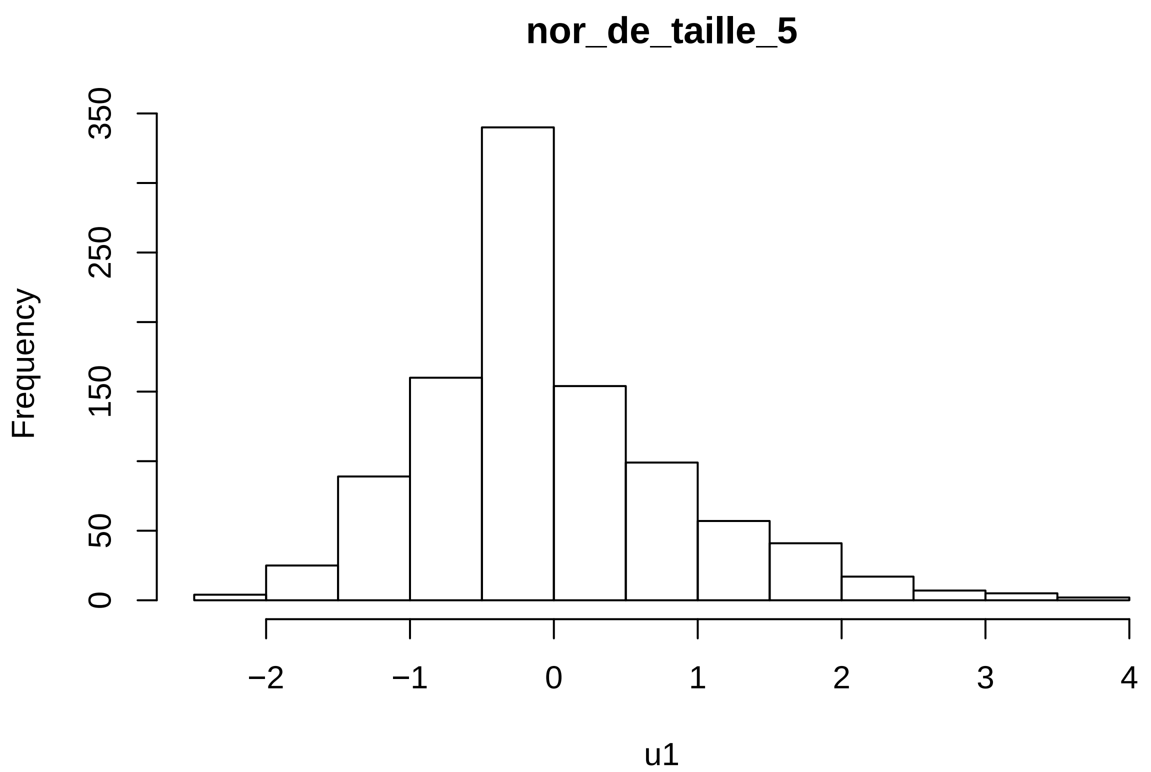

```r
hist(m2,freq =T ,main ="moy_emp_de_pois_taille_30")
```



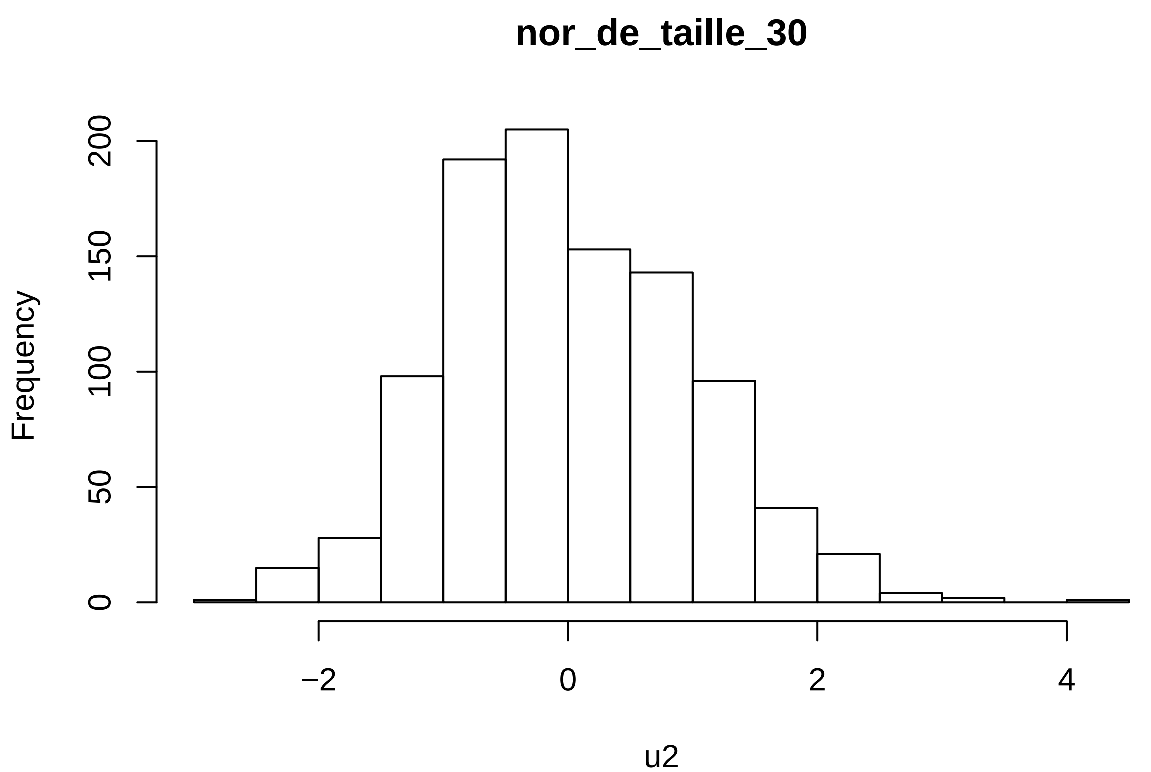

```r
hist(m3,freq =T ,main ="moy_emp_de_pois_taille_100")
```



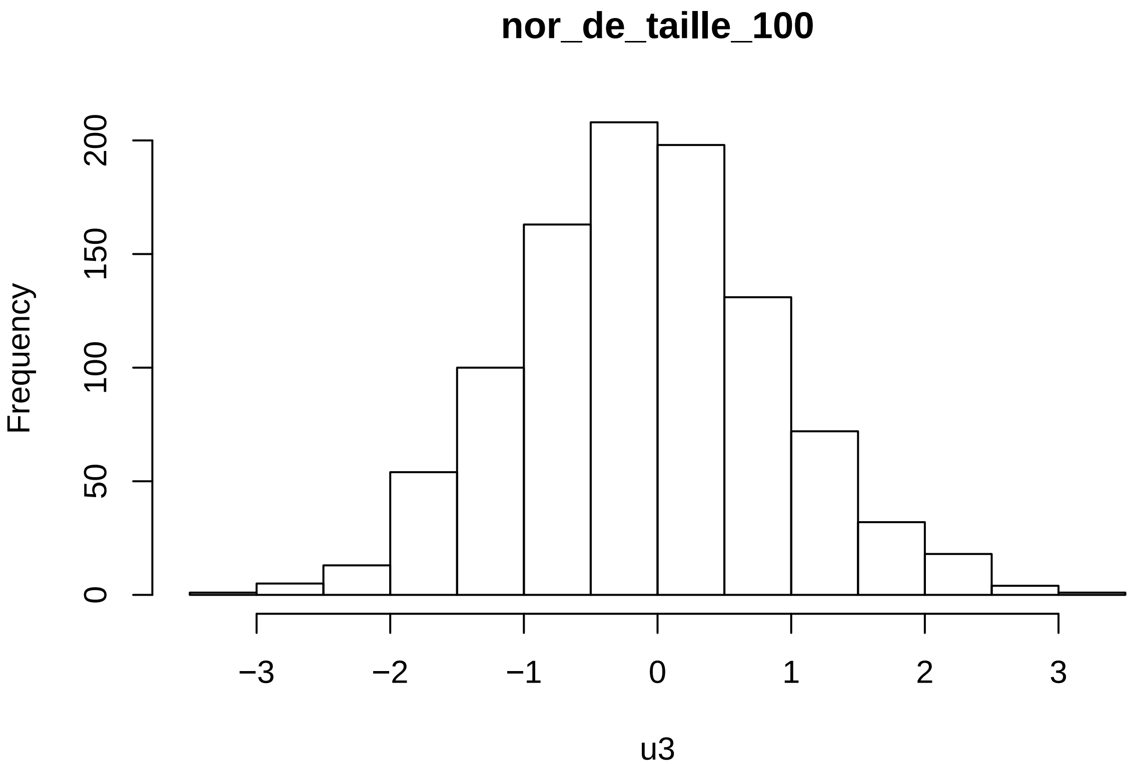

```
renormalisation<-function(N,moy_emp,n,mu,sigma)
{
  for ( i in 1:N)
  {
    moy_emp[i]=(moy_emp[i]-mu)/(sigma/sqrt(n));
  }
  return(moy_emp)
}

u1 = renormalisation(1000,m1,5,1,1);
u2 = renormalisation(1000,m2,30,1,1);
u3 = renormalisation(1000,m3,100,1,1);
hist(u1,freq =T ,main ="nor_de_taille_5")
```




```
hist(u2,freq =T ,main ="nor_de_taille_30")
```




```
hist(u3,freq =T ,main ="nor_de_taille_100")
```



<–4–>

quand on fais N fois échantillon (x1,x2..xn) iid de n'importe quel loi, on trouve que toute moyenne $\bar{X}_n$ des échantillons tends vers une variable aléatoire gausienne. En plus, plus N est grand, plus la distribution d'échantillonage est similaire à gaussienne avec μ = E(x) var = var(x). Par ailleurs,

$$Z_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$$

$$\lim_{n\to\infty} P(Z_n \leq z) = \Phi(z)$$

$\Phi(z)$ suis la **loi normale centrée réduite** $\mathcal{N}(0,1)$

2.Moyenne et dispersion

<–1–>

l'inégalité de Chebychef dans les cas Gaussien est

$$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

l'inégalité de Chebychef dans les cas Poisson est

$$P(|X-\lambda| \geq k\lambda) \leq \frac{1}{k^2}$$

parce que l'espérence et la varience de Poison sont tous λ

<–2–>

a

$$P(|X-\mu| \geq \sigma) \ = E(I_{|X-\mu| \geq k\sigma})$$

avec I_A est une fonction indicatrice

b

```r
#n la taille
#r une échantillon qu'on a obeteun
#mu espérence

  P_deviation <- function(n,r,mu,delta)
  {
    res<- 0
    for (i in 1:n) {
        if (r[i]-mu >= delta)
          res<-res+1
    }
    return(res/n)
  }
  r1<-rnorm(10000,mean = 1,sd = 1)
  r2<-rpareto(10000,m=1,s=2)
  r3<-rpois(10000,lambda=1)
```

On calcule par la fonction P_deviation les probabilités de déviation d'une v.a de sa moyenne

Dans le cas Gaussien avec mean=1 sd=1 $P(|X-1| \geq 2) =$

```r
P_deviation(10000,r1,1,2)
```

```
## [1] 0.0217
```

Dans le cas Pareto avec m=1 s=1 $P(|X-1| \geq 2) =$

```r
P_deviation(10000,r1,1,2)
```

```
## [1] 0.0217
```

Dans le cas Pareto avec lambda=1$P(|X-1| \geq 2) =$

```r
P_deviation(10000,r1,1,2)
```

```
## [1] 0.0217
```

La précison de cette estimation est cinq

c

*quand $\delta = \sigma$, on a borne obtenue par Bienaymé Chebychev qui égale à 1

*quand $\delta = 2\sigma$,on a borne obtenue par Bienaymé Chebychev qui égale à 0.25

*quand $\delta = 3\sigma$,on a borne obtenue par Bienaymé Chebychev qui égale à 0.11

particulierment, si on change σ, les borne changeont pas