

tpStat2

Partie 1 : Analyse de la vitesse du vent

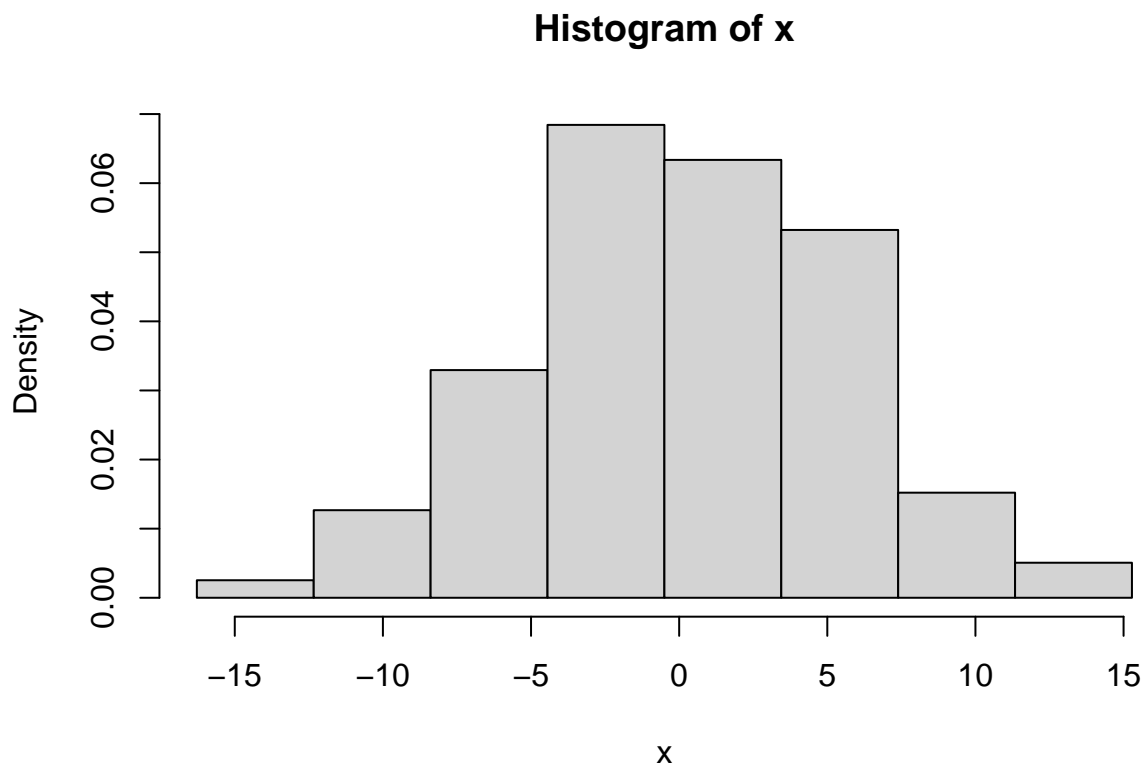
Question 1 :

Commençons par tracer l'histogramme de classes à même largeur de A1 :

```
source("histolarg.R")
source("histoeff.R")
vent<-read.table("vent.csv",sep=";",header=T)
attach(vent)

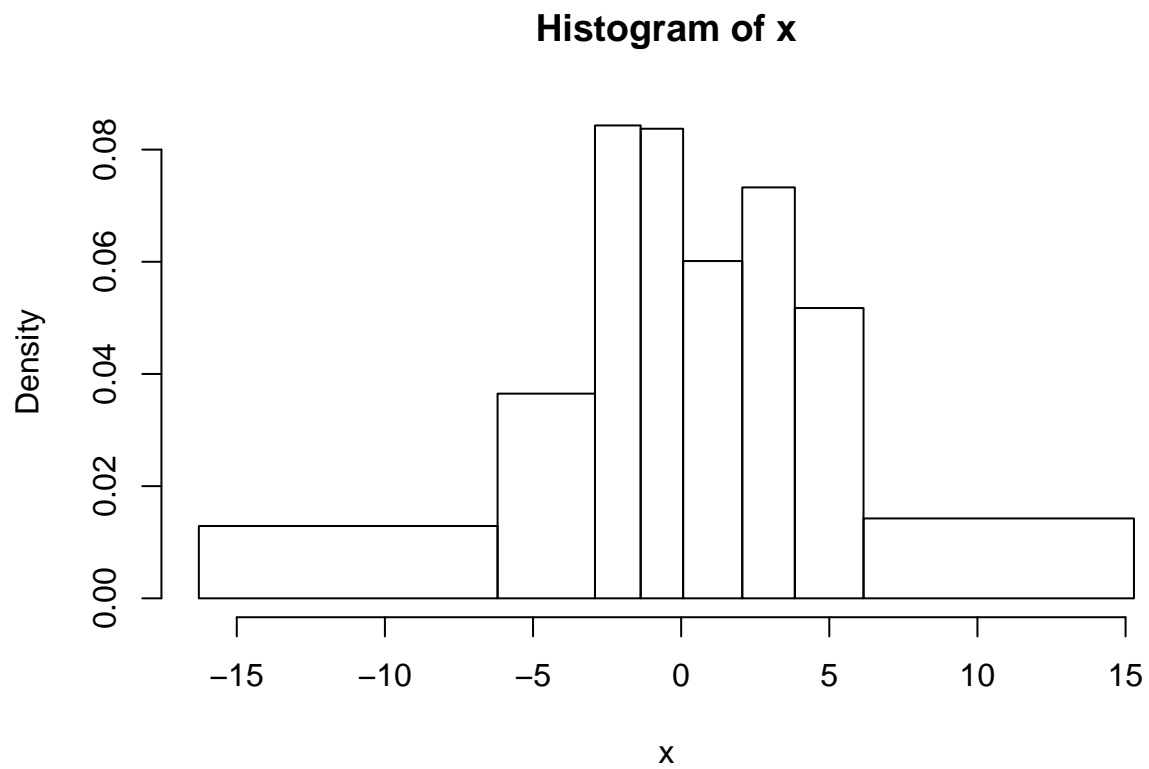
n<-length(A1)

histolarg(A1)
```



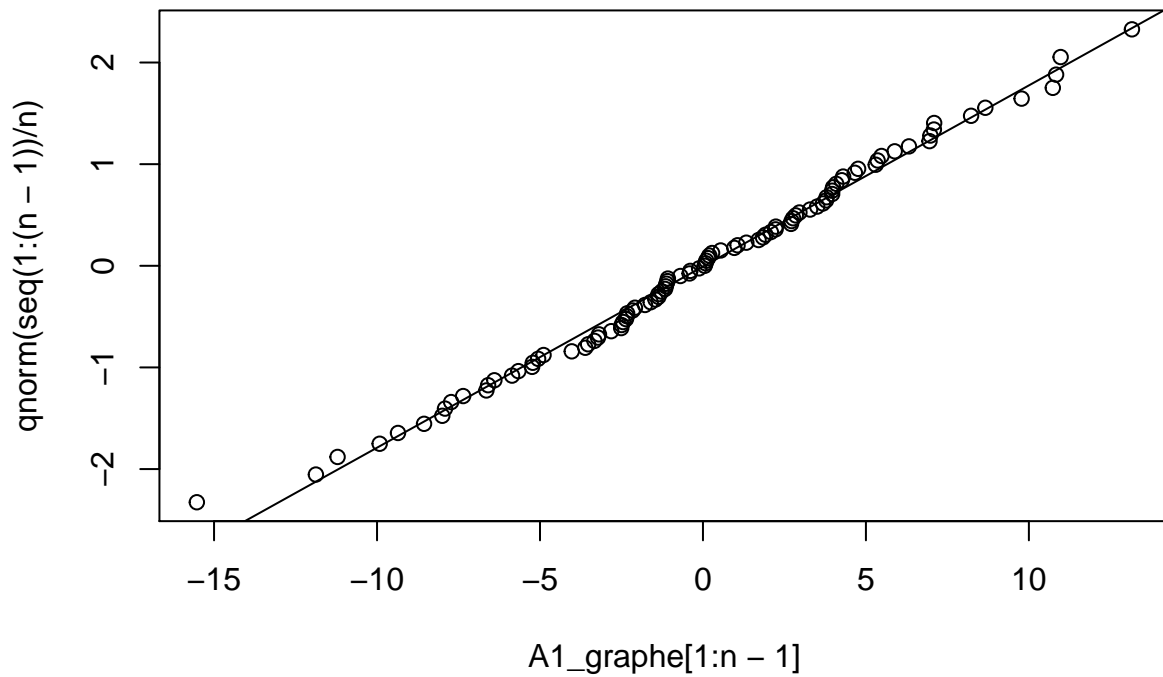
Traçons maintenant son histogramme à classes de même effectif :

```
histoeff(A1)
```



L'allure de ces deux courbes nous mettent bien sur la voie d'une loi normale centrée, mais pour vérifier cette hypothèse traçons le graphe de probabilité :

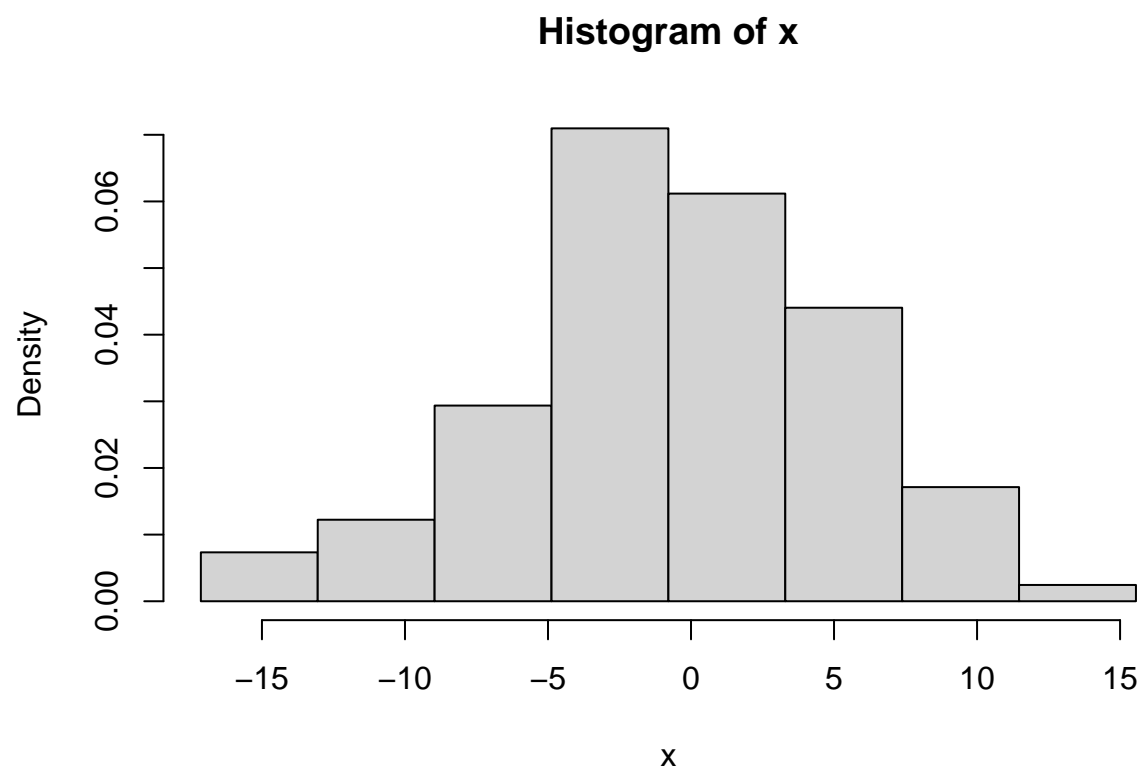
```
A1_graphe <- sort(A1)
plot(A1_graphe[1:n-1], qnorm(seq(1:(n-1))/n))
reg1<-lm(qnorm(seq(1:(n-1))/n)~A1_graphe[1:(n-1)])
abline(reg1)
```



Ici en calculant la régression linéaire on s'aperçoit que le graphe de probabilité est très proche d'une droite donc A1 suit une loi normale centrée.

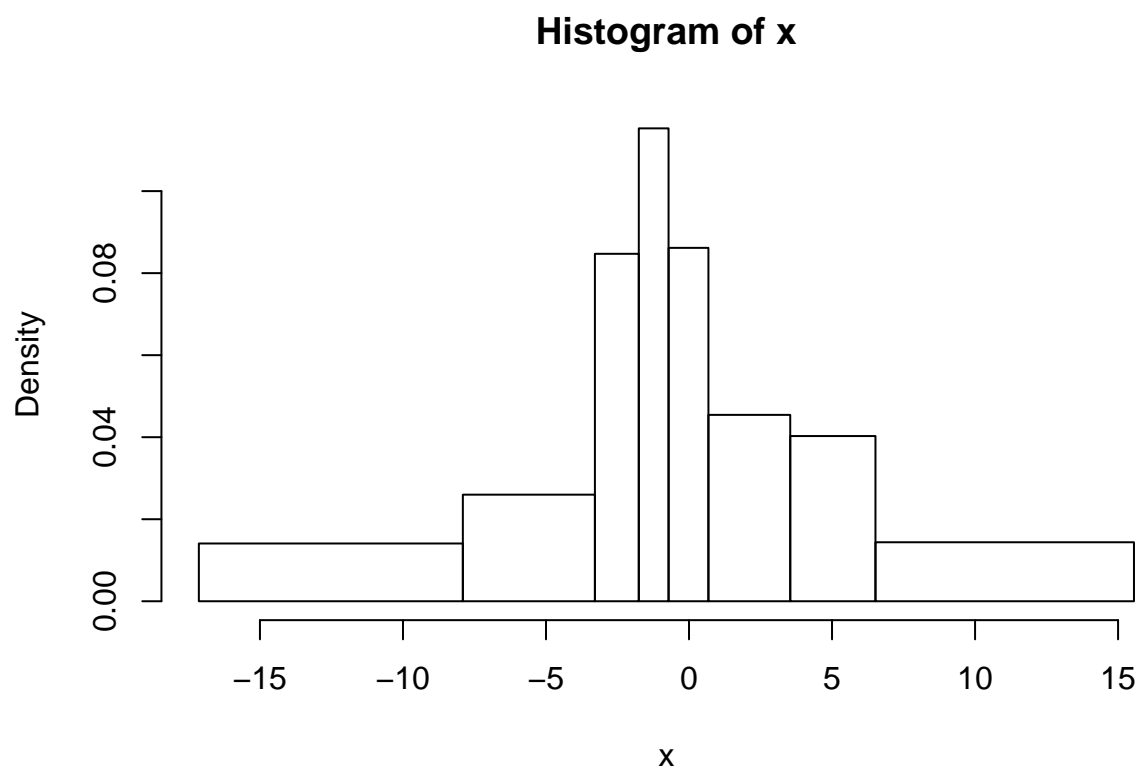
On fait de même pour A2 et on trace l'histogramme à classes de même largeur ce qui donne :

```
histolarg(A2)
```



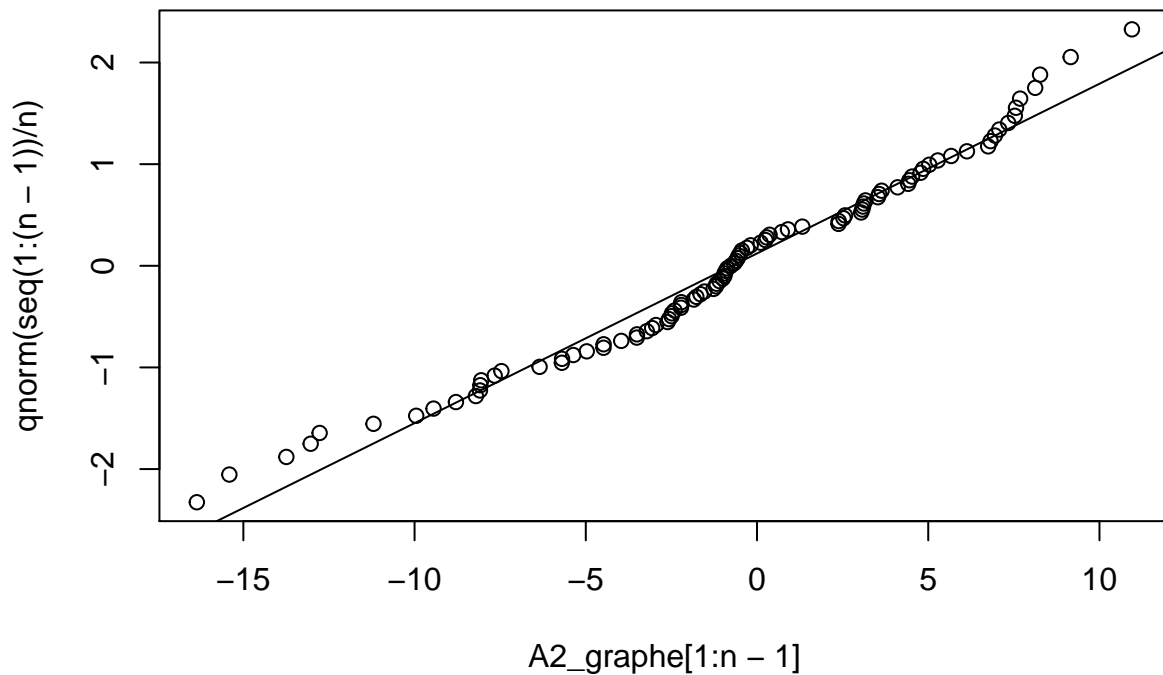
Puis l'histogramme à classes de même effectif ce qui donne :

```
histoeff(A2)
```



Enfin on trace le graphe de probabilité :

```
A2_graphe <- sort(A2)
plot(A2_graphe[1:n-1], qnorm(seq(1:(n-1))/n))
reg2<-lm(qnorm(seq(1:(n-1))/n)~A2_graphe[1:(n-1)])
abline(reg2)
```



Là encore en calculant la régression linéaire on s'aperçoit que le graphe de probabilité est très proche d'une droite donc A2 suit bien une loi normale centrée.

D'autres part comme les mesures des coordonnées A1 et A2 sont indépendantes et du même ordre de grandeur on peut estimer que A1 et A2 suivent la même loi.

Comme A1 et A2 suivent la même loi normal, calculons la regression linéaire sur l'ensembles des mesures A1 et A2 puis à partir de celle-ci calculons une estimation de sigma.

```
A<-c(A1, A2)
A<-sort(A)
n<-length(A)

reg<-lm(qnorm(seq(1:(n-1))/n)~A[1:(n-1)])

sigma<- 1/(reg$coefficients[2])

cat("estimation de sigma au carré : ", sigma^2,"\n")
```

```
## estimation de sigma au carré : 33.49484
```

Question 2 :

$A1 \sim \mathcal{N}(0, \sigma^2)$ donc $\frac{A1}{\sigma^2} \sim \mathcal{N}(0, 1)$ et ainsi $\frac{A1^2}{\sigma^2} \sim \mathcal{X}_1^2$

De même on obtient $\frac{A2^2}{\sigma^2} \sim \mathcal{X}_1^2$

Par conséquent $\frac{A_1^2}{\sigma^2} + \frac{A_2^2}{\sigma^2} \sim \chi_2^2$

Comme $\frac{X^2}{\sigma^2} = \frac{A_1^2}{\sigma^2} + \frac{A_2^2}{\sigma^2}$ on a $f_{\frac{X^2}{\sigma^2}}(x) = f_{\chi_2^2}(x) = \frac{2^{-1}}{1} \exp(\frac{-x}{2}) 1_{\mathbb{R}_+}(x)$

Donc $\frac{X^2}{\sigma^2} \sim \mathcal{E}(\frac{1}{2})$

Question 3 :

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= P\left(\frac{X^2}{\sigma^2} \leq \frac{x^2}{\sigma^2}\right) \\ F_X(x) &= F_{\frac{X^2}{\sigma^2}}\left(\frac{x^2}{\sigma^2}\right) \end{aligned}$$

Par dérivation on obtient :

$$\begin{aligned} f_X(x) &= \frac{2x}{\sigma^2} \times f_{\frac{X^2}{\sigma^2}}\left(\frac{x^2}{\sigma^2}\right) 1_{\mathbb{R}_+}(x) \\ &= \frac{2x}{\sigma^2} \times \frac{1}{2} \exp\left(\frac{-x^2}{2\sigma^2}\right) 1_{\mathbb{R}_+}(x) \\ f_X(x) &= \frac{x}{\sigma^2} \exp\left(\frac{-x^2}{2\sigma^2}\right) 1_{\mathbb{R}_+}(x) \end{aligned}$$

Question 4 :

On a $f_X(x) = \frac{x}{\sigma^2} \exp\left(\frac{-x^2}{2\sigma^2}\right) 1_{\mathbb{R}_+}(x)$

d'où $F_X(x) = 1 - \exp\left(\frac{-x^2}{2\sigma^2}\right)$

$\ln[1 - F_X(x)] = \frac{-1}{2\sigma^2} \times x^2 + 0$

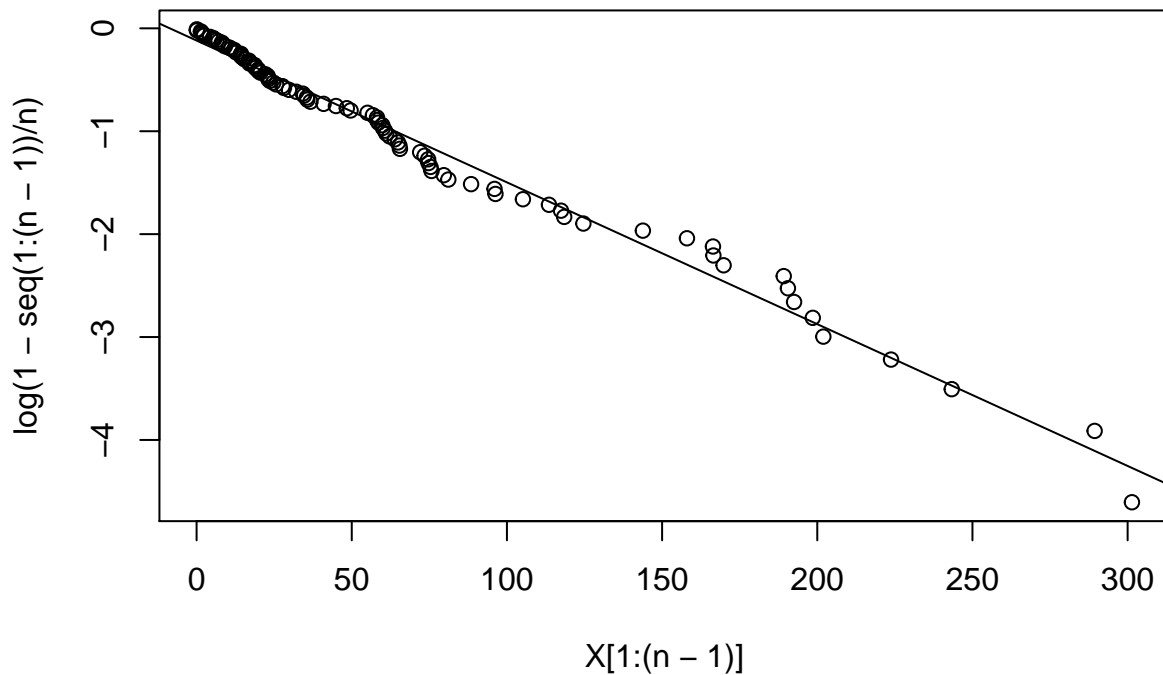
On obtient donc le graphe de probabilité suivant : $G((x_i^*)^2, \ln(1 - \frac{i}{n}))$

Ce qui donne :

```
X <- sort(A1^2 + A2^2)
n <- length(X)

plot(X[1:(n-1)], log(1-seq(1:(n-1))/n))

reg_ray<-lm(log(1-seq(1:(n-1))/n)~X[1:(n-1)])
abline(reg_ray)
```



En calculant la régression linéaire on s'aperçoit que le graphe de probabilité est très proche d'une droite donc X suit bien la loi de rayleigh.

On estime σ^2 graphiquement pour obtenir σ_g^2 :

```
sigma_carre <- -1/(2*reg_ray$coefficients[2])
sigma_carre
```

```
## X[1:(n - 1)]
##      36.28843
```

Question 5 :

Le calcul de $\mathbb{E}[X]$ s'obtient par intégration par parties :

$$\begin{aligned}
 \mathbb{E}[X] &= \int_0^{+\infty} x \frac{x}{\sigma^2} \exp^{-\frac{x^2}{2\sigma^2}} dx \\
 &= [-x \exp^{-\frac{x^2}{2\sigma^2}}]_0^{+\infty} - \int_0^{+\infty} -\exp^{-\frac{x^2}{2\sigma^2}} dx \\
 &= 0 + \int_0^{+\infty} \exp^{-\frac{x^2}{2\sigma^2}} dx \\
 &= \sigma \sqrt{\frac{\pi}{2}}
 \end{aligned}$$

On obtient $\mathbb{E}[X^2]$ de manière analogue :

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_0^{+\infty} x^2 \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \\
&= [-x^2 \exp\left(-\frac{x^2}{2\sigma^2}\right)]_0^{+\infty} - \int_0^{-\infty} -2x \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \\
&= 0 + 2\sigma^2 \int_0^{+\infty} \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \\
&= 2\sigma^2 [-\exp\left(-\frac{x^2}{2\sigma^2}\right)]_0^{+\infty} \\
&= 2\sigma^2
\end{aligned}$$

On peut alors calculer la variance $Var[X]$:

$$\begin{aligned}
Var[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= 2\sigma^2 - \left(\sigma\sqrt{\frac{\pi}{2}}\right)^2 \\
&= \left(\frac{4-\pi}{2}\right)\sigma^2
\end{aligned}$$

Un estimateur par la methode des moments est donc $\tilde{\sigma}^2 = \frac{2}{4-\pi}S_n^2$ Cependant démontrons que cet estimateur est biaisé :

$$\begin{aligned}
\mathbb{E}[\tilde{\sigma}^2] &= \mathbb{E}\left[\frac{2}{4-\pi}S_n^2\right] \\
&= \frac{2}{4-\pi}\mathbb{E}[S_n^2] \\
&= \frac{2}{4-\pi}(\mathbb{E}[X^2] - \mathbb{E}[\bar{X}_n^2]) \\
&= \frac{2}{4-\pi}(Var[X] + \mathbb{E}[X]^2 - Var[\bar{X}_n] - \mathbb{E}[\bar{X}_n]^2) \\
&= \frac{2}{4-\pi}(Var[X] + \mathbb{E}[X]^2 - \frac{Var[\bar{X}_n]}{n} - \mathbb{E}[X]^2) \\
&= \frac{2}{4-\pi}\left(1 - \frac{1}{n}\right)Var[X] \\
&= \frac{2}{4-\pi}\left(\frac{n-1}{n}\right)\left(\frac{4-\pi}{2}\right)\sigma^2 \\
&= \frac{n-1}{n}\sigma^2 \neq \sigma^2
\end{aligned}$$

Il faut donc multiplier $\tilde{\sigma}^2$ par $\frac{n}{n-1}$ pour obtenir un estimateur sans biais :

$$\tilde{\sigma}'^2 = \frac{2}{4-\pi}S_n'^2 = \frac{2n}{(4-\pi)(n-1)}S_n^2$$

Question 6 :

On calcule la fonction de vraisemblance, sachant que les X_i sont indépendants :

$$\begin{aligned}
\mathcal{L}(\sigma^2; x_1; \dots; x_n) &= \prod_{i=1}^n f_X(x_i) \\
&= \prod_{i=1}^n \frac{x_i}{\sigma^2} \exp\left(-\frac{x_i^2}{2\sigma^2}\right)
\end{aligned}$$

On passe au logarithme

$$\begin{aligned}\ln(\mathcal{L}(\sigma^2; x_1; \dots; x_n)) &= \sum_{i=1}^n \ln\left(\frac{x_i}{\sigma^2} \exp\left(-\frac{x_i^2}{2\sigma^2}\right)\right) \\ &= \sum_{i=1}^n \left(\ln(x_i) - \ln(\sigma^2) - \frac{x_i^2}{2\sigma^2}\right)\end{aligned}$$

L'estimateur du maximum de vraisemblance $\hat{\sigma}_n^2$ peut s'obtenir en dérivant cette expression par rapport à σ^2

$$\begin{aligned}\frac{\partial}{\partial \sigma^2} \ln(\mathcal{L}(\sigma^2; x_1; \dots; x_n)) &= 0 \\ \sum_{i=1}^n \left(\frac{-1}{\sigma^2} + \frac{x_i^2}{2\sigma^4}\right) &= 0 \\ \frac{1}{\sigma^2} \sum_{i=1}^n \left(\frac{x_i^2}{2\sigma^2} - 1\right) &= 0 \\ \sum_{i=1}^n \frac{x_i^2}{2\sigma^2} &= n\end{aligned}$$

Donc $\hat{\sigma}_n^2 = \frac{1}{2n} \sum_{i=1}^n x_i^2$ On peut alors montrer que cet estimateur n'est pas biaisé

$$\begin{aligned}\mathbb{E}[\hat{\sigma}_n^2] &= \mathbb{E}\left[\frac{1}{2n} \sum_{i=1}^n x_i^2\right] \\ &= \frac{1}{2n} \sum_{i=1}^n \mathbb{E}[x_i^2] \\ &= \frac{1}{2} \mathbb{E}[X^2] \\ &= \frac{1}{2} (\text{Var}[X] + \mathbb{E}[X]^2) \\ &= \sigma^2 \frac{4 - \pi}{4} + \frac{1}{2} (\sigma^2 \frac{\pi}{2}) \\ &= \sigma^2\end{aligned}$$

Pour obtenir l'efficacité de cet estimateur, on a besoin de sa variance :

$$\begin{aligned}\text{Var}[\hat{\sigma}_n^2] &= \text{Var}\left[\frac{1}{2n} \sum_{i=1}^n x_i^2\right] \\ &= \frac{1}{4n^2} \sum_{i=1}^n \text{Var}[x_i^2] \\ &= \frac{1}{4n} \text{Var}[X^2] \\ &= \frac{1}{4n} (\mathbb{E}[X^4] - \mathbb{E}[X^2]^2)\end{aligned}$$

Or

$$\begin{aligned}\mathbb{E}[X^4] &= \int_0^{+\infty} x^4 f_X(x) dx \\ &= \int_0^{+\infty} \frac{x^5}{\sigma^2} \exp(-\frac{x^2}{2\sigma^2}) dx\end{aligned}$$

$$\text{on pose } u = \frac{x}{\sigma}, \quad du = \frac{dx}{\sigma}, \quad dx = \sigma du$$

$$\begin{aligned}&= \int_0^{+\infty} \frac{\sigma^5 u^5}{\sigma^2} \exp(-\frac{u^2}{2}) \sigma du \\ &= \sigma^4 \int_0^{+\infty} u^5 \exp(-\frac{u^2}{2}) du \\ &= -\sigma^4 ([u^4 \exp(-\frac{u^2}{2})]_0^{+\infty} - \int_0^{+\infty} 4u^3 \exp(-\frac{u^2}{2}) du) \\ &= -\sigma^4 ([4u^2 \exp(-\frac{u^2}{2})]_0^{+\infty} - \int_0^{+\infty} 8u \exp(-\frac{u^2}{2}) du) \\ &= \sigma^4 [-8 \exp(-\frac{u^2}{2})]_0^{+\infty} \\ &= 8\sigma^4\end{aligned}$$

Enfin

$$\begin{aligned}\text{Var}[\hat{\sigma}_n^2] &= \frac{1}{4n} (\mathbb{E}[X^4] - \mathbb{E}[X^2]^2) \\ &= \frac{1}{4n} (8\sigma^4 - 2^2 \sigma^4) \\ &= \frac{\sigma^4}{n}\end{aligned}$$

On a aussi besoin de la quantité d'information de Fisher.

$$\begin{aligned}\mathcal{I}(\sigma^2) &= -\mathbb{E}[\frac{\partial^2}{\partial \sigma^2} \ln \mathcal{L}(\sigma^2; X_1; \dots; X_2)] \\ &= -\mathbb{E}[\sum_{i=1}^n (\frac{1}{\sigma^4} - \frac{x_i^2}{\sigma^6})] \\ &= -(\frac{n}{\sigma^4} - \frac{\mathbb{E}[X^2]}{\sigma^6}) \\ &= -(\frac{n}{\sigma^4} - \frac{2n}{\sigma^4}) \\ &= \frac{n}{\sigma^4}\end{aligned}$$

On peut alors calculer l'efficacité.

$$\begin{aligned}\text{Eff}(\hat{\sigma}_n^2) &= \frac{1}{\frac{n}{\sigma^4} \frac{\sigma^4}{n}} \\ &= 1\end{aligned}$$

Ainsi cet estimateur est efficace.

Question 7 :

On a pour l'estimateur par la méthode des moments $\tilde{\sigma}_n^2$:

```

X_carre <- A1^2 + A2^2
X <- sqrt(X_carre)
X<-sort(X)
X_carre<-sort(X_carre)
n <- length(X)

2/(4-pi)*((n-1)/n)*var(X)

```

```
## [1] 38.85148
```

Puis l'estimateur débiaisé $\hat{\sigma}_n'^2$ est :

```
2/(4-pi)*var(X)
```

```
## [1] 39.24392
```

Enfin l'estimateur du maximum de vraisemblance $\hat{\sigma}_n^2$ est :

```

sigma_carre_chapeau<-sum(X_carre)/(2*n)

sigma_carre_chapeau

```

```
## [1] 32.66009
```

L'estimateur du maximum de vraisemblance étant efficace, c'est celui-là que l'on choisira.

Question 8 :

Soit $\alpha \in]0, 1[$, soit $(a, b) \in \mathbb{R}^+$,

On sait que $\frac{X^2}{\sigma^2}$ suit une loi exponentielle de paramètre 1/2 (d'après les questions précédentes) donc $\sum_{k=0}^n X_i^2/\sigma^2$ suit une loi Gamma $G(n, 1/2)$.

$$\mathbb{P}(a \leq \frac{1}{\sigma^2} \sum_{k=1}^n X_i^2 \leq b) = \mathbb{P}\left(\frac{\sum_{k=1}^n X_i^2}{b} \leq \sigma^2 \leq \frac{\sum_{k=1}^n X_i^2}{a}\right) = F_{G(n, 1/2)}(b) - F_{G(n, 1/2)}(a) = 1 - \alpha$$

On pose $F_{G(n, 1/2)}(b) = 1 - \alpha/2$ et $F_{G(n, 1/2)}(a) = \alpha/2$.

Or, la loi $G(n, 1/2)$ est aussi la loi du chi-2 à $2n$ degrés de libertés χ_{2n}^2 . Donc avec la table de loi du χ^2 on a:

- $a = z_{2n, 1-\alpha/2}$
- $b = z_{2n, \alpha/2}$

Et donc l'intervalle de confiance est,

$$\left[\frac{\sum_{k=1}^n X_i^2}{z_{2n, \frac{\alpha}{2}}}, \frac{\sum_{k=1}^n X_i^2}{z_{2n, 1-\frac{\alpha}{2}}} \right]$$

Pour $\alpha = 0,05$

```
somme = 0
for (i in 1:100){
  somme <- somme + X[i]**2
}

borneinf <- somme / qchisq(1 - 0.05/2, 200)
bornesup <- somme / qchisq(0.05/2, 200)
print(borneinf)
```

```
## [1] 27.0973
```

```
print(bornesup)
```

```
## [1] 40.14071
```

Question 9 :

On pose les hypothèses suivantes:

- H_0 : Terrain constructible ce qui équivaut à $\sigma^2 \leq \sigma_0^2 = \frac{162}{\pi} m/s$
- H_1 : Terrain n'est pas constructible ce qui équivaut à $\sigma^2 > \sigma_0^2 = \frac{162}{\pi} m/s$

On a :

$$\alpha = \sup_{\sigma^2 \leq \sigma_0^2} P(\hat{\sigma}_n^2 > I_\alpha) = P\left(\frac{2n}{\sigma^2} \hat{\sigma}_n^2 > \frac{2n}{\sigma^2} I_\alpha\right)$$

On sait que :

$$\frac{2n}{\sigma^2} \hat{\sigma}_n^2 = \frac{\sum_{i=1}^n X_i^2}{\sigma^2} = Y \rightarrow \chi_{2n}^2$$

Donc :

$$\alpha = \sup_{\sigma^2 \leq \sigma_0^2} (1 - F_Y\left(\frac{2n}{\sigma^2} I_\alpha\right))$$

On a :

$$I_\alpha = \frac{\sigma_0^2}{2n} z_{2n, \alpha}$$

$$\text{Ainsi la zone critique est : } W = \left[\frac{2n}{\sigma_0^2} \hat{\sigma}_n^2 > z_{2n, \alpha}\right]$$

Faisons la modélisation R:

```
sigma_0_carre = 2 / pi * 81
stat = 2*100 / sigma_0_carre * sigma_carre_chapeau

quant = qchisq(0.95, 2*100)
cat("stat = ", stat, "quantile = ", quant)
```

```
## stat = 126.6725 quantile = 233.9943
```

On a p-valeur :

$$p\text{valeur} = P_{H_0}(Y < stat)$$

```
pval = pchisq(stat, 2*100)
cat("pval = ", pval)
```

```
## pval = 1.277368e-05
```

Partie 2 : Vérifications expérimentales à base de simulations

Question 1 :

Afin de simuler un échantillon de taille n de la loi $\mathcal{R}(\sigma^2)$ pour σ^2 quelconque on a besoin de plusieurs étapes. La première est de simuler deux échantillons de taille n et de loi $\mathcal{N}(0, 1)$ afin d'obtenir $A1$ et $A2$. Ensuite il faut élever au carré chaque valeurs de ces échantillons afin d'obtenir $A1^2$ et $A2^2$. Finalement il suffit de prendre la racine carré de la somme de ces deux échantillons pour générer X .

En R cela donne :

```
A1_simu <- rnorm(n, 0, sigma)
A1_simu_carre <- A1^2

A2_simu <- rnorm(n, 0, sigma)
A2_simu_carre <- A2^2

X_simu_carre <- A1_simu_carre + A2_simu_carre
X_simu <- sqrt(X_simu_carre)

X_simu

##      [1]  2.9546451  1.5390173  4.3457824  4.1309676  3.7939818  3.8675577
##      [7]  3.0239101 19.0570705 12.9044691  5.6678407  8.6458692  4.8593085
##     [13]  0.3120709 12.5700757  8.7006991  4.7848101 17.0109951 14.9587775
##     [19]  8.0947268 12.8992279  2.3113839  7.0446687  3.9303616  4.3200086
##     [25]  7.6557876  5.9747666  4.1118012  7.4217626  7.8178613  8.9254961
##     [31]  0.2909006  8.6816678 10.6562197  7.7363781  7.5376237  1.3292542
##     [37]  4.3549670  1.1778999  6.0573142  3.4856941  2.5282743  8.4884157
##     [43]  6.9570705  4.8090136  5.3063709  7.9922456  5.4568606 15.5990400
##     [49]  2.9113041  3.5158008 11.1639901  7.6255610  9.8135692  8.6379697
##     [55]  1.2840252  5.8832586  5.0613973 14.2109302  8.0557514  4.7946621
##     [61] 14.0906281  4.1243466  2.0779234  4.5351978 13.0313943  7.6247431
##     [67] 11.9918488 13.8752337  5.8519515  8.0843795  9.4049977  6.3979697
##     [73]  3.3428225  3.5850968  7.7487507  1.4845765  9.0065579 10.2546240
##     [79] 13.8036095  5.9591402  3.7850251  8.5711831  6.7014727  3.7852067
##     [85]  5.2606519 13.7547935  2.8493409 17.3607615  3.2398028 10.8363725
##     [91]  9.7981152  4.9729382 10.8833731  2.3874861  7.8845537  4.4646963
##     [97]  4.4529584  2.6592398  4.7074503  7.7891638
```

Question 2 :

```
#simulation de X au carre
simu_rayleigh <- function(n, sigma)
{
```

```

A1_simu <- rnorm(n, 0, sigma)
A1_simu_carre <- A1_simu^2

A2_simu <- rnorm(n, 0, sigma)
A2_simu_carre <- A2_simu^2

X_simu_carre <- A1_simu_carre + A2_simu_carre

return(X_simu_carre)
}

#Fonction qui affiche le pourcentage de fois ou sigma est dans l'IC
pourcentage_rayleigh<- function (alpha, m, n, sigma)
{
  nb_fois<-0

  for (i in 1:m)
  {
    X <- sqrt(simu_rayleigh(n, sigma))
    ech <- sort(X)
    sigma_chapeau = sum(ech^2)/(2*n)

    a = 2*n*sigma_chapeau/qchisq(1- alpha/2, 2*n)
    b = 2*n*sigma_chapeau/qchisq(alpha/2, 2*n)

    nb_fois=nb_fois + ( (sigma^2 >= a) & (sigma^2 <= b) )
  }

  cat("pourcentage de fois où sigma dans l'IC:", nb_fois/m*100, "\n")
}

pourcentage_rayleigh(0.01,10,10,10)

## pourcentage de fois où sigma dans l'IC: 100

pourcentage_rayleigh(0.05,10,10,10)

## pourcentage de fois où sigma dans l'IC: 60

pourcentage_rayleigh(0.1,10,10,10)

## pourcentage de fois où sigma dans l'IC: 90

pourcentage_rayleigh(0.5,10,10,10)

## pourcentage de fois où sigma dans l'IC: 60

```

```
cat("\n")
```

```
pourcentage_rayleigh(0.01,100,10, 10)
```

```
## pourcentage de fois où sigma dans l'IC: 100
```

```
pourcentage_rayleigh(0.05,100,10,10)
```

```
## pourcentage de fois où sigma dans l'IC: 93
```

```
pourcentage_rayleigh(0.1,100,10,10)
```

```
## pourcentage de fois où sigma dans l'IC: 90
```

```
pourcentage_rayleigh(0.5,100,10,10)
```

```
## pourcentage de fois où sigma dans l'IC: 52
```

```
cat("\n")
```

```
pourcentage_rayleigh(0.01,1000,10, 10)
```

```
## pourcentage de fois où sigma dans l'IC: 99.2
```

```
pourcentage_rayleigh(0.05,1000,10,10)
```

```
## pourcentage de fois où sigma dans l'IC: 95.9
```

```
pourcentage_rayleigh(0.1,1000,10,10)
```

```
## pourcentage de fois où sigma dans l'IC: 89.3
```

```
pourcentage_rayleigh(0.5,1000,10,10)
```

```
## pourcentage de fois où sigma dans l'IC: 50.7
```

```
cat("\n")
```

```
pourcentage_rayleigh(0.01,1000,100,10)
```

```
## pourcentage de fois où sigma dans l'IC: 99
```

```
pourcentage_rayleigh(0.05,1000,100,10)
```

```
## pourcentage de fois où sigma dans l'IC: 94.9
```



```
pourcentage_rayleigh(0.1,1000,100,10)
```

```
## pourcentage de fois où sigma dans l'IC: 90.6
```

```
pourcentage_rayleigh(0.5,1000,100,10)
```

```
## pourcentage de fois où sigma dans l'IC: 51.7
```

```
cat("\n")
```

```
pourcentage_rayleigh(0.01,1000,1000,10)
```

```
## pourcentage de fois où sigma dans l'IC: 99
```

```
pourcentage_rayleigh(0.05,1000,1000,10)
```

```
## pourcentage de fois où sigma dans l'IC: 94.8
```

```
pourcentage_rayleigh(0.1,1000,1000,10)
```

```
## pourcentage de fois où sigma dans l'IC: 90.2
```

```
pourcentage_rayleigh(0.5,1000,1000,10)
```

```
## pourcentage de fois où sigma dans l'IC: 47.6
```

```
cat("\n")
```

```
pourcentage_rayleigh(0.01,1000,1000,20)
```

```
## pourcentage de fois où sigma dans l'IC: 98.7
```

```
pourcentage_rayleigh(0.05,1000,1000,20)
```

```
## pourcentage de fois où sigma dans l'IC: 94
```

```
pourcentage_rayleigh(0.1,1000,1000,20)
```

```
## pourcentage de fois où sigma dans l'IC: 91.5
```

```
pourcentage_rayleigh(0.5,1000,1000,20)
```

```
## pourcentage de fois où sigma dans l'IC: 49.5
```

```
cat("\n")
```

```
pourcentage_rayleigh(0.01,1000,1000,30)
```

```
## pourcentage de fois où sigma dans l'IC: 99
```

```
pourcentage_rayleigh(0.05,1000,1000,30)
```

```
## pourcentage de fois où sigma dans l'IC: 94.6
```

```
pourcentage_rayleigh(0.1,1000,1000,30)
```

```
## pourcentage de fois où sigma dans l'IC: 89.4
```

```
pourcentage_rayleigh(0.5,1000,1000,30)
```

```
## pourcentage de fois où sigma dans l'IC: 52.1
```

Tout d'abord on remarque que le choix de σ^2 n'a pas un gros impacte sur les pourcentages que l'on observe.

Concernant le choix n de la taille de l'échantillon on se rend compte qu'à partir du moment où celui-ci est suffisamment grand, augmenter la taille n'a plus de réel impacte sur les pourcentages.

Enfin on remarque que m est le facteur qui a le plus d'impacte sur les pourcentages pour un α fixé, plus celui-ci est grand plus les pourcentages se rapprochent de la valeur $1 - \alpha$.

Question 3 :

```
#Fonction qui calcule et affiche le biais et l'erreur quadratique pour chaque estimateur
compare_estimateurs<- function ( m, n, sigma)
{
  estimateur1 <-NULL
  estimateur2 <-NULL
  estimateur3 <-NULL
  estimateur4 <-NULL

  for (i in 1:m)
  {
    X_simu_carre <- simu_rayleigh(n, sigma)
    X_simu<- sqrt(X_simu_carre)
    X_simu<-sort(X_simu)
    X_simu_carre<-sort(X_simu_carre)

    reg_ray<-lm(log(1-seq(1:(n-1))/n)~X_simu_carre[1:(n-1)])

    sigma_g <- -1/(2*reg_ray$coefficients[2])

    estimateur1<-c(estimateur1, sigma_g)
```

```

sigma_chapeau = sum(X_simu_carre)/(2*n)

estimateur2<-c(estimateur2, sigma_chapeau)

sigma_tilde = (2/(4-pi))*(n-1)/n*var(X_simu)

estimateur3<-c(estimateur3, sigma_tilde)

sigma_tilde_corr = (2/(4-pi))*var(X_simu)

estimateur4<-c(estimateur4, sigma_tilde_corr)

}

biais1<- mean(estimateur1-sigma^2)
biais2<- mean(estimateur2-sigma^2)
biais3<- mean(estimateur3-sigma^2)
biais4<- mean(estimateur4-sigma^2)

errquad1<-mean((estimateur1-sigma^2)^2)
errquad2<-mean((estimateur2-sigma^2)^2)
errquad3<-mean((estimateur3-sigma^2)^2)
errquad4<-mean((estimateur4-sigma^2)^2)

cat("biais estimateur 1 estimé : ", biais1, "\n")
cat("biais estimateur 2 estimé :",biais2, "\n")

cat("biais estimateur 3 estimé :",biais3, "\n")
cat("biais estimateur 4 estimé :",biais4, "\n")

cat("\n")

cat("Erreur quadratique estimé estimateur 1 : ", errquad1, "\n")
cat("Erreur quadratique estimé estimateur 2 : ",errquad2, "\n")

cat("Erreur quadratique estimé estimateur 3 : ",errquad3, "\n")
cat("Erreur quadratique estimé estimateur 4 : ",errquad4, "\n")

}

compare_estimateurs(1000,10,10)

```

```

## biais estimateur 1 estimé : -8.960172
## biais estimateur 2 estimé : -0.08500382
## biais estimateur 3 estimé : -9.533587
## biais estimateur 4 estimé : 0.5182371
##
## Erreur quadratique estimé estimateur 1 : 1427.813
## Erreur quadratique estimé estimateur 2 : 1065.791
## Erreur quadratique estimé estimateur 3 : 2080.473
## Erreur quadratique estimé estimateur 4 : 2456.545

```

```
compare_estimateurs(1000,50,10)
```

```
## biais estimateur 1 estimé : -4.112
## biais estimateur 2 estimé : 0.007175375
## biais estimateur 3 estimé : -2.128039
## biais estimateur 4 estimé : -0.1306516
##
## Erreur quadratique estimé estimateur 1 : 308.4347
## Erreur quadratique estimé estimateur 2 : 203.1404
## Erreur quadratique estimé estimateur 3 : 447.9533
## Erreur quadratique estimé estimateur 4 : 461.7255
```

```
compare_estimateurs(1000,500,10)
```

```
## biais estimateur 1 estimé : -0.9025897
## biais estimateur 2 estimé : -0.01088388
## biais estimateur 3 estimé : -0.003197031
## biais estimateur 4 estimé : 0.1971974
##
## Erreur quadratique estimé estimateur 1 : 34.27476
## Erreur quadratique estimé estimateur 2 : 18.78903
## Erreur quadratique estimé estimateur 3 : 44.51202
## Erreur quadratique estimé estimateur 4 : 44.72948
```

```
compare_estimateurs(1000,5000,10)
```

```
## biais estimateur 1 estimé : -0.2230957
## biais estimateur 2 estimé : -0.03570422
## biais estimateur 3 estimé : -0.05834148
## biais estimateur 4 estimé : -0.03834915
##
## Erreur quadratique estimé estimateur 1 : 3.866121
## Erreur quadratique estimé estimateur 2 : 2.037706
## Erreur quadratique estimé estimateur 3 : 4.773714
## Erreur quadratique estimé estimateur 4 : 4.77369
```

On observe bien que L'estimateur 2 est sans biais tout comme l'estimateur 4, cependant à la différence de celui-ci l'estimateur 2 converge plus rapidement vers 0 en terme d'erreur quadratique.

On en déduit donc que le meilleur estimateur est l'estimateur 2. L'estimateur 2 étant $\hat{\sigma}_n^2$ cela n'est pas surprenant car c'est également le seul estimateur efficace.

Question 4 :

```
convergence_faible<- function ( m, n, sigma, epsilon)
{
  nb_fois <- 0
  for (i in 1:m)
  {
    X_simu_carre <- simu_rayleigh(n, sigma)
```

```

X_simu_carre<-sort(X_simu_carre)

sigma_chapeau = sum(X_simu_carre)/(2*n)

ecart = abs(sigma_chapeau - sigma^2)

nb_fois=nb_fois + (ecart > epsilon)
}

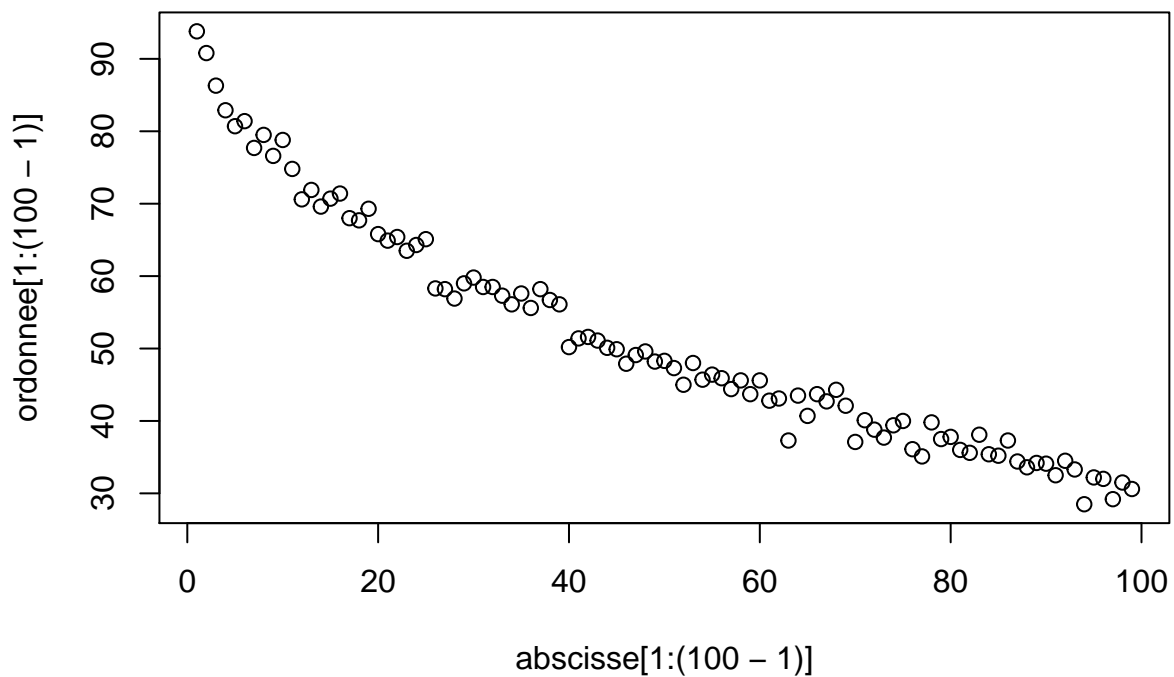
return(nb_fois/m*100)
}

abscisse <- NULL
ordonnee <- NULL

for (i in 1:100)
{
  abscisse<-c(abscisse, i)
  ordonnee<-c(ordonnee, convergence_faible(1000, i, 10, 10))
}

plot(abscisse[1:(100-1)], ordonnee[1:(100-1)])

```



On observe que plus n est grand plus le pourcentage de fois où l'écart en valeur absolue entre $\hat{\sigma}_n^2$ et σ^2 est supérieur à ϵ se rapproche de 0.

On en conclut donc que $\hat{\sigma}_n^2$ est bien faiblement convergent.

Question 5 :

```
convergence_asymptotique<- function ( m, n, sigma)
{
  estimateur <- NULL

  for (i in 1:m)
  {
    X_simu_carre <- simu_rayleigh(n, sigma)
    X_simu_carre<-sort(X_simu_carre)

    sigma_chapeau = sum(X_simu_carre)/(2*n)

    estimateur<-c(estimateur, sigma_chapeau)
  }

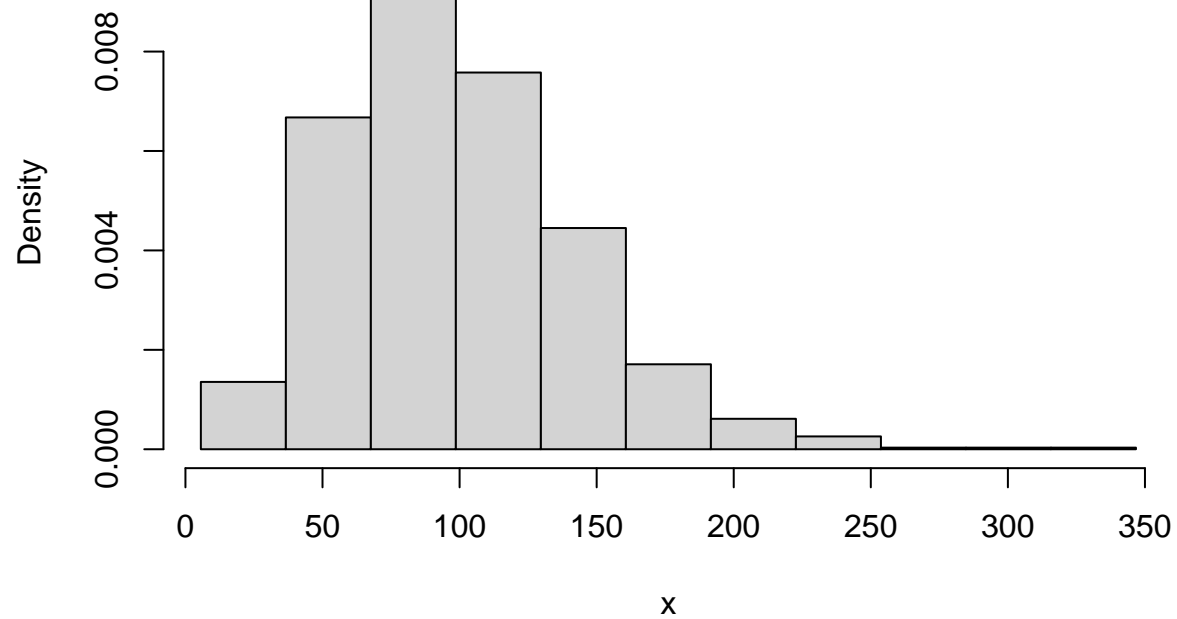
  estimateur<-sort(estimateur)

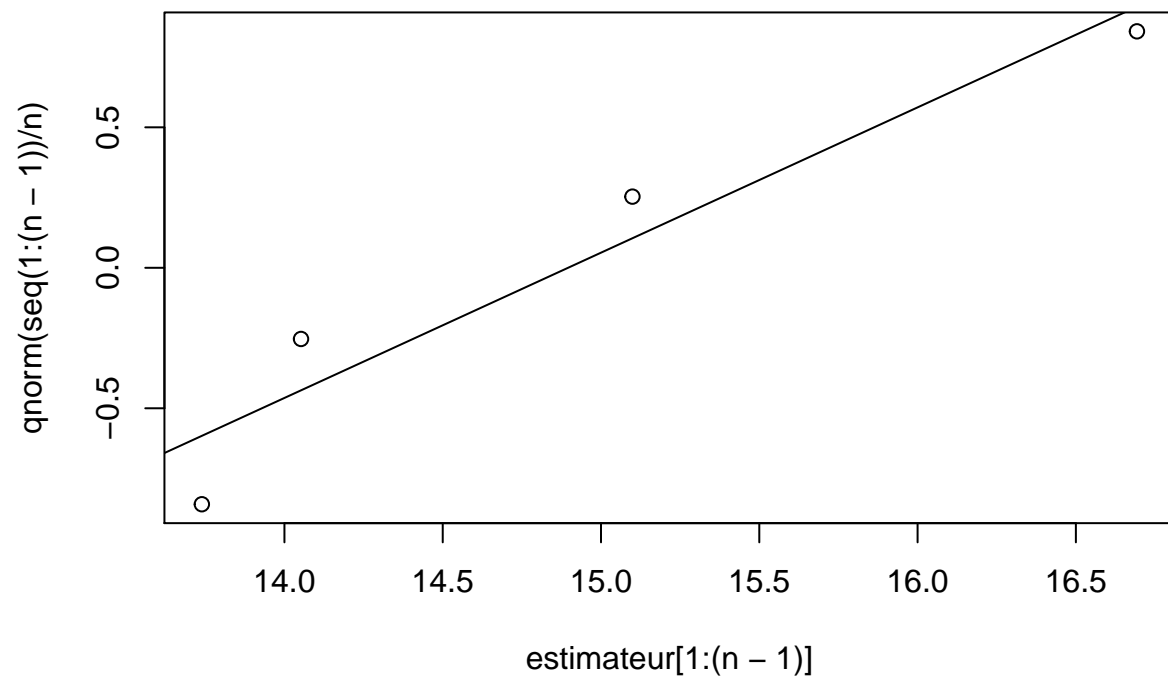
  histolarg(estimateur)
  plot(estimateur[1:(n-1)], qnorm(seq(1:(n-1))/n))
  abs<-estimateur[1:n-1]
  ord<-qnorm(seq(1:(n-1))/n)
  reg<-lm(ord~abs)

  abline(reg)
}

convergence_asymptotique(1000, 5, 10)
```

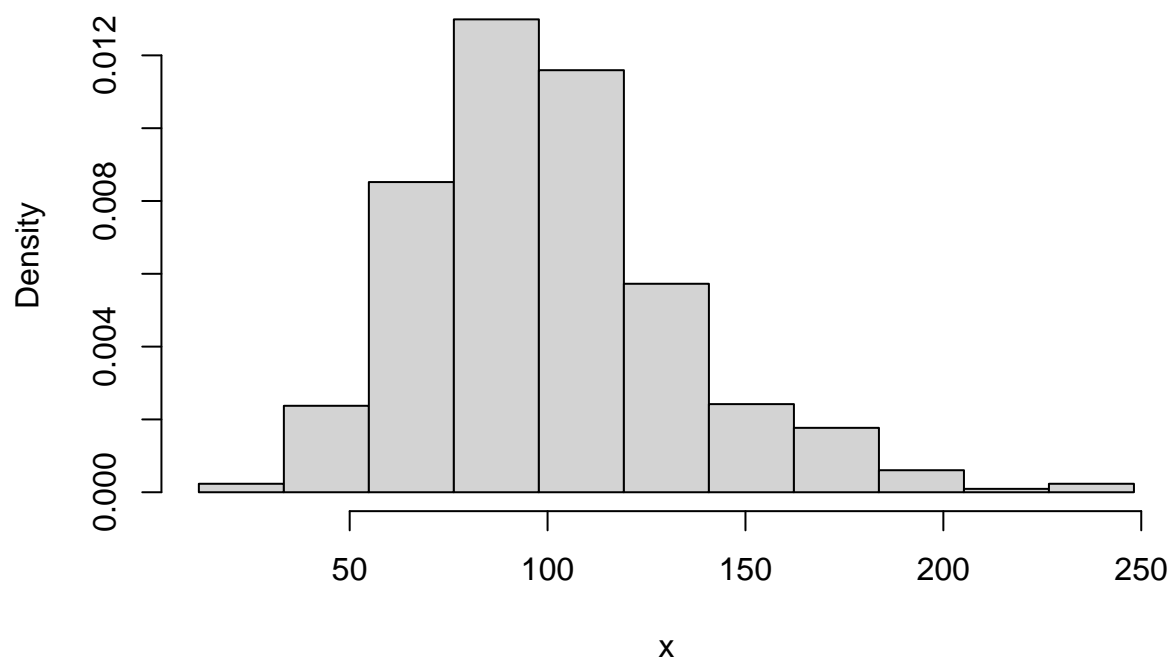
Histogram of x

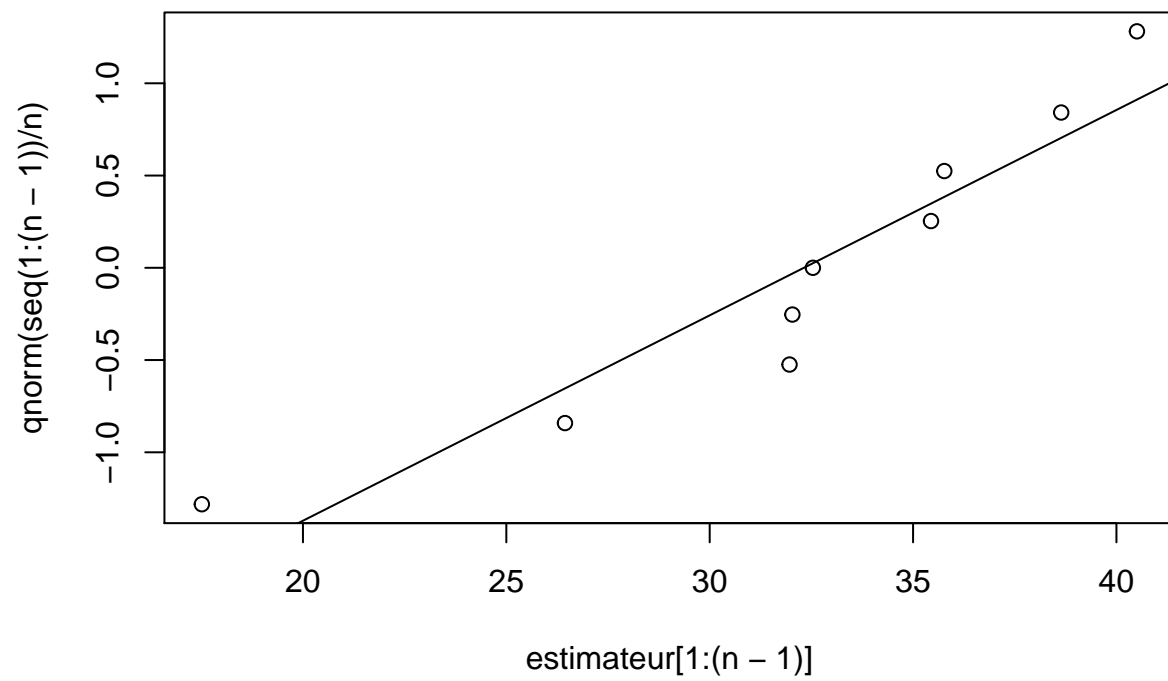




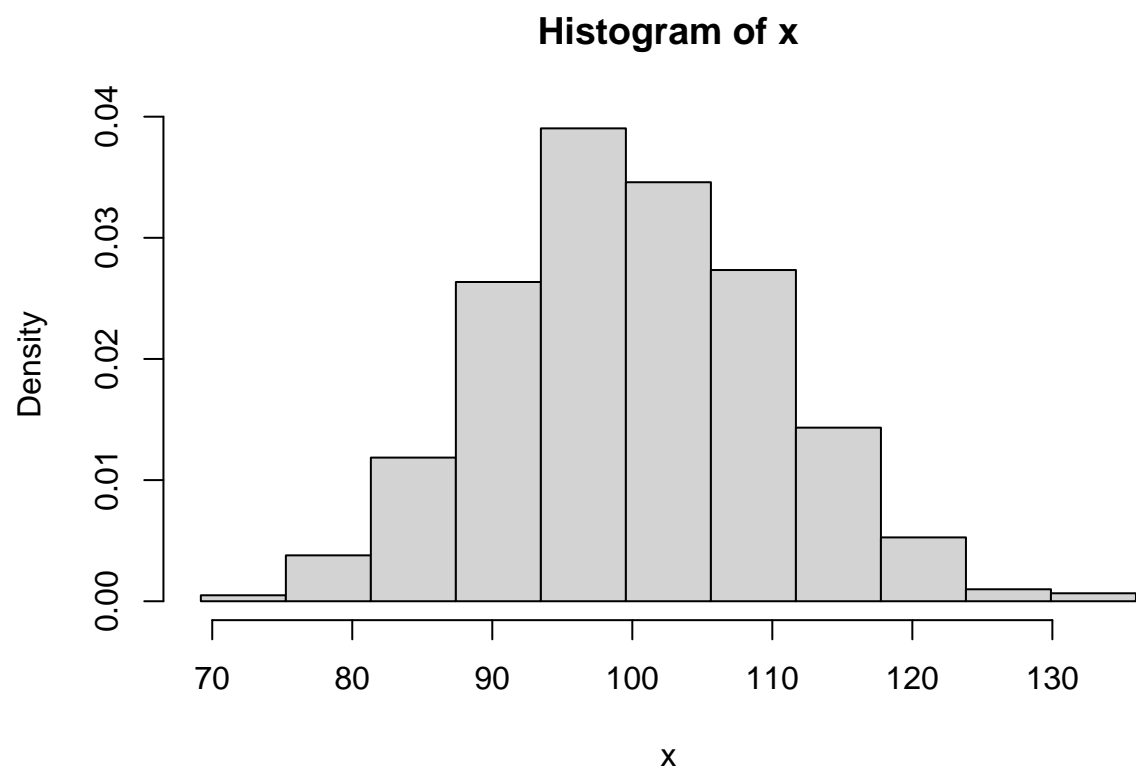
```
convergence_asymptotique(1000, 10, 10)
```

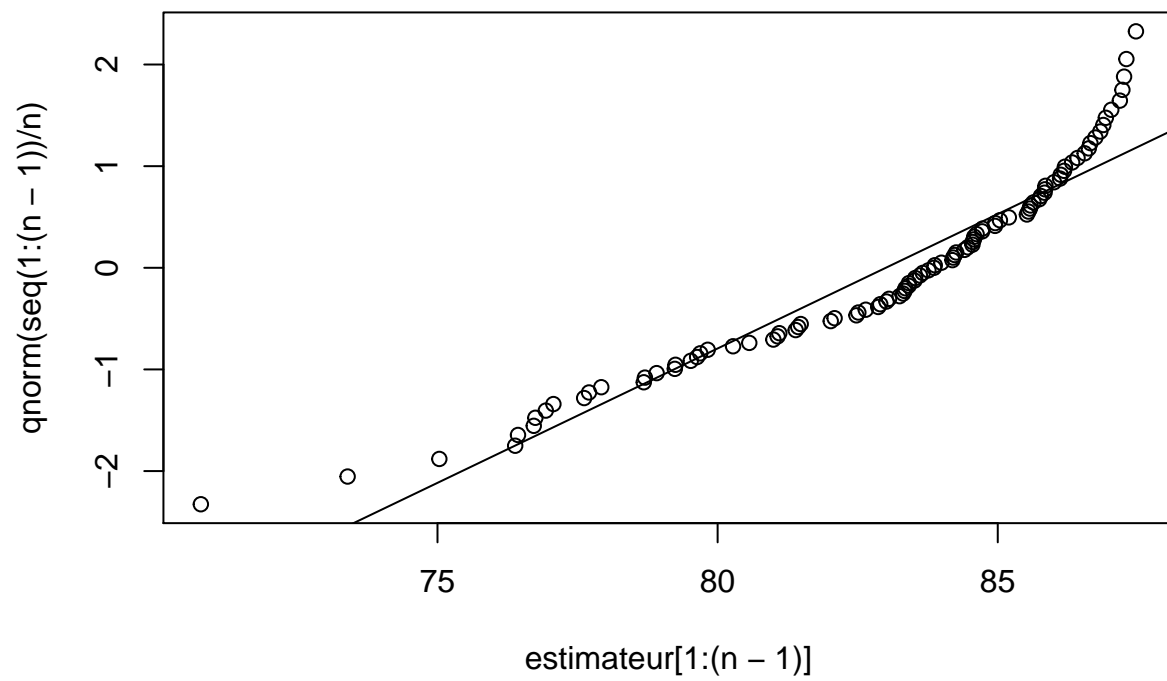

Histogram of x



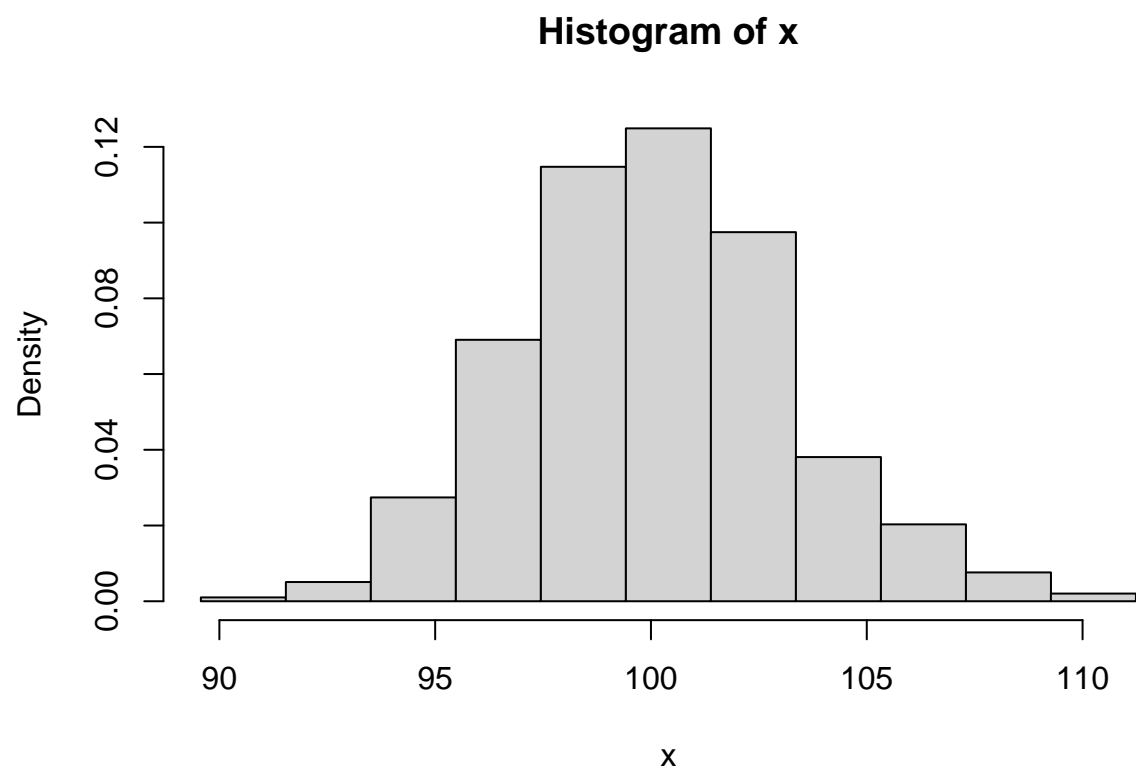


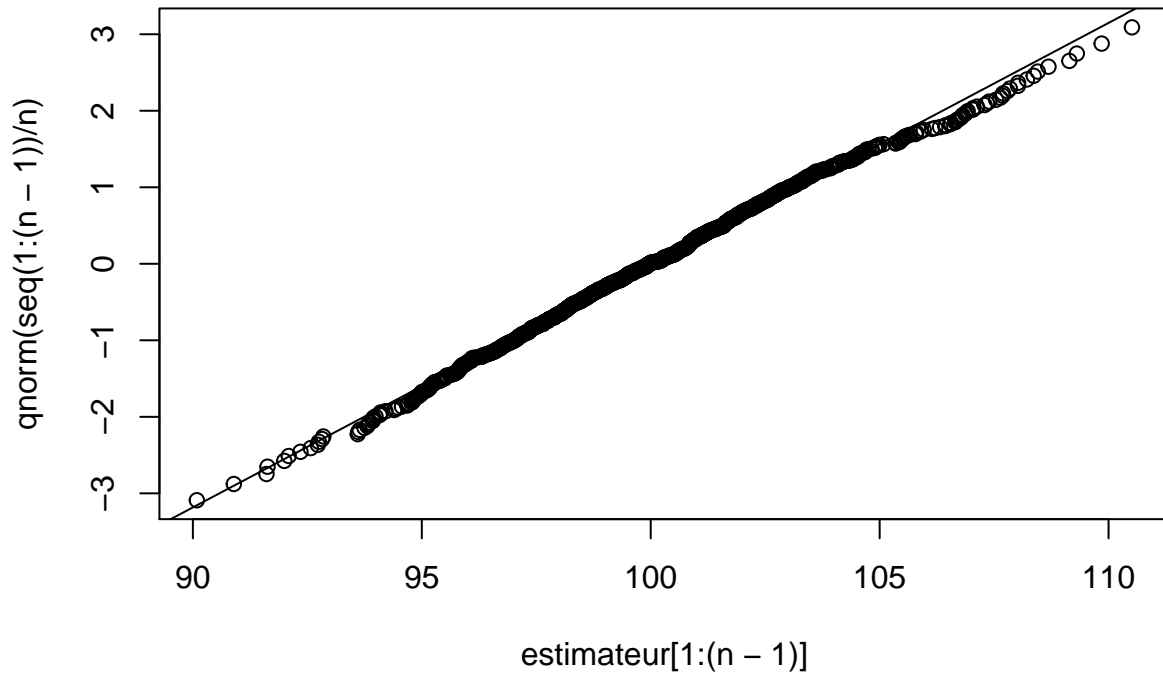
```
convergence_asymptotique(1000, 100, 10)
```





```
convergence_asymptotique(1000, 1000, 10)
```





On observe que plus n est grand plus les graphes de probabilités se rapproche d'une droite.

Cela signifie que plus n est grand plus la répartition des valeurs de $\hat{\sigma}_n^2$ suit une loi normale et confirme bien le théorème centrale limite.

En effet on sait que $\hat{\sigma}_n^2$ se construit à partir de $\sum_{i=1}^n X_i^2$, or les X_i^2 sont des variables aléatoires et indépendantes. Donc d'après le théorème centrale limite $\sum_{i=1}^n X_i^2$ suit approximativement une loi normale et donc $\hat{\sigma}_n^2$ suit également cette loi.