TP_PMS

MOHAMEDAHMED-Mohmedlemine Hamza-Aboutni louis-Sassier

2023-04-11

Partie1 : Assurance maritime

question1 :

Loi expontielle :

- Graphres des probabilites :

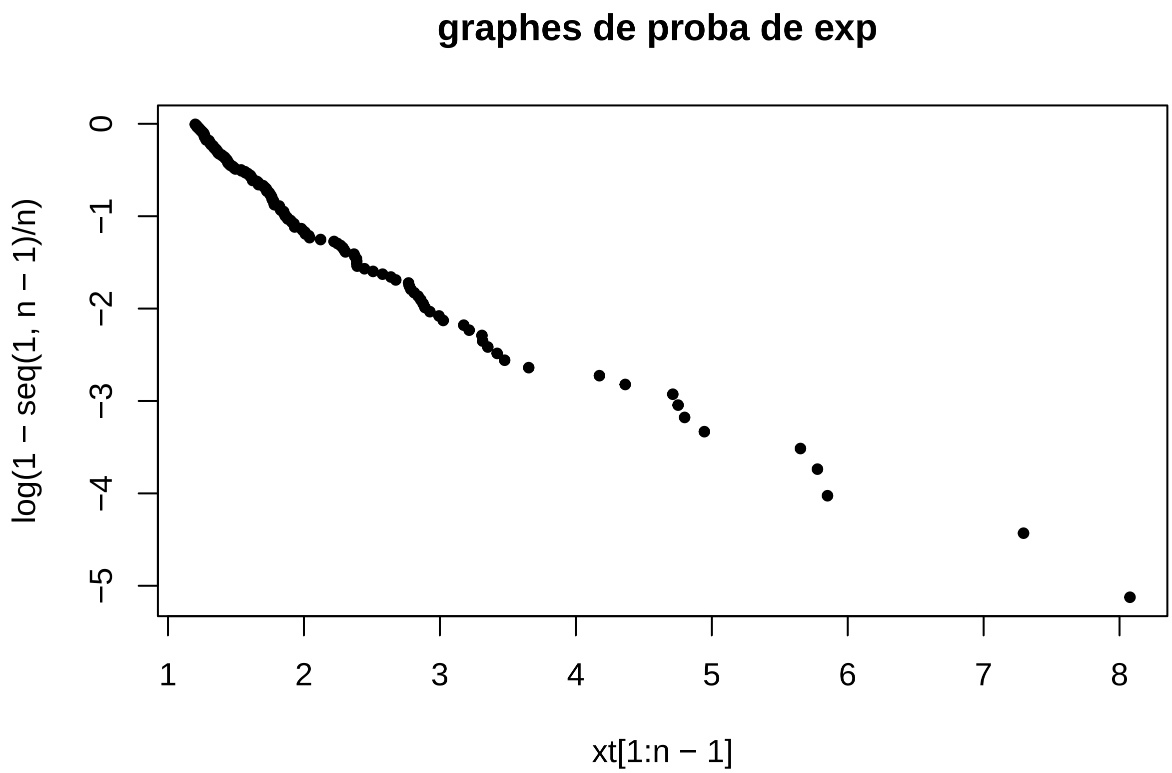
$$\forall x \in \mathbb{R}_+ \ \mathbb{F}(x) = \mathbb{P}(X \leq x) = 1 - e^{-\lambda x}$$

$$\forall x \in \mathbb{R}_+ \ \ln(1 - \mathbb{F}(x)) = -\lambda x$$

le graphe de probabilites de loi expontielle est le nuage des points :

$$\boxed{\{x_i^*, \ln(1 - \frac{i}{n})\} \ pour \ i \in \{1, 2, 3, ..., n-1\}}$$

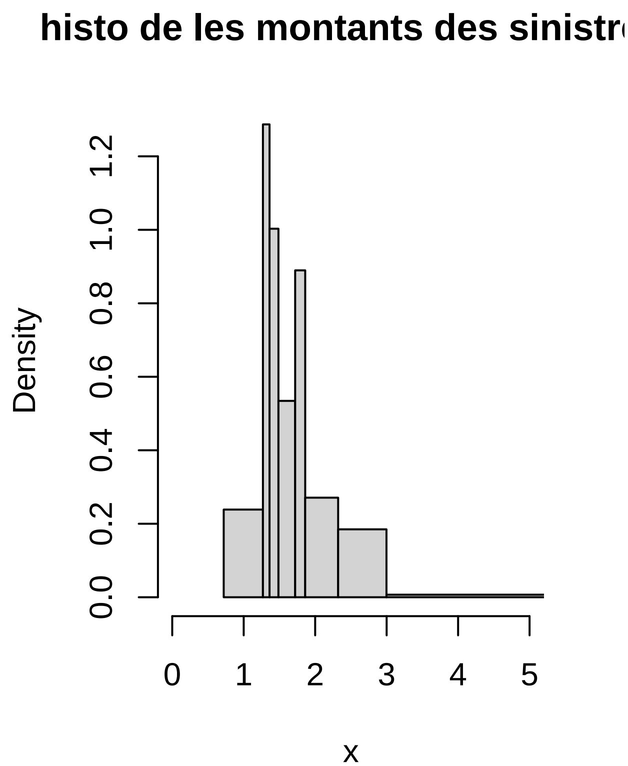
```
x <-scan("sinistres.csv") # pour charger le tableau de données en R
n <- length(x)
xt <- sort(x) # pour triéé x
plot(xt[1:n-1],log(1-seq(1,n-1)/n),main="graphes de proba de exp ",xlim=c(xt[1],xt[n-1]),pch=20)
```

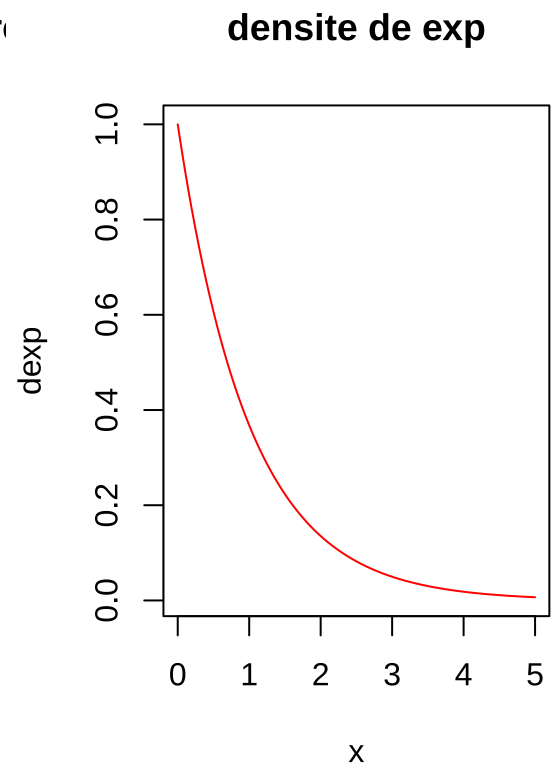


Le non alignement des points dans le graphe ci-dessus montre bien que la loi exponentielle n'est pas un modèle plausible pour décrire les montants des sinistres.

- Histogrammes pour la loi exponentielle:

```
x <-scan("sinistres.csv")
n <- length(x)
k<-round(1+log2(n))
xt<-sort(x)
a0=xt[1]-0.025*(xt[n]-xt[1])
ak=xt[n]+0.025*(xt[n]-xt[1])
h<-(ak-a0)/n
bornes <- c(a0,quantile(x,seq(1,k-1)/k),ak)
par(mfcol=c(1,2))
hist(x,prob=T , breaks=bornes, main="histo de les montants des sinistres",xlim=c(0,5)) # histogramme a
plot(dexp,main="densite de exp" ,xlim=c(0,5), col="red",pch=20) # on affiche la dénsite de la loi exp
```





l'histogramme donne une approximation de la densite de la loi correspondante. le graphe ci-dessus montre bien que la loi exponentielle n'est pas un modèle plausible pour décrire les montants des sinistres.

Loi normale :

- Graphres des probabilites : si La variable aléatoire X suit une loi normale de moyenne μ et d'écart type σ $X \sim \mathcal{N}(\mu, \sigma^2)$ Alors:

$$\frac{X-m}{\sigma} \sim \mathcal{N}(0,1)$$

$$\forall x \in \mathbb{R} \quad \mathbb{F}(x) = \mathbb{P}(X \leq x) = \mathbb{P}(\frac{X-m}{\sigma} \leq \frac{x-m}{\sigma})$$

Donc

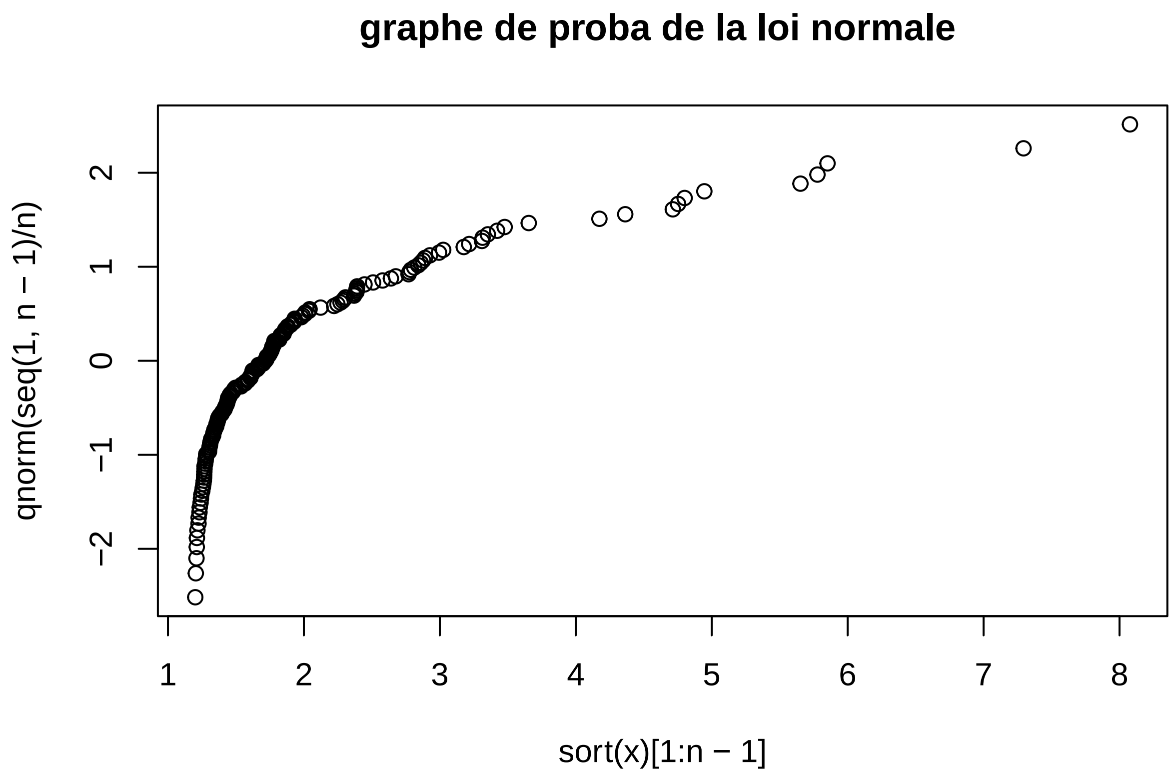
$$\forall x \in \mathbb{R} \quad \Phi^{-1}(\mathbb{F}(x)) = \frac{1}{\sigma}x - \frac{m}{\sigma}$$

avec Φ est la fonction de répartion de la loi $\mathcal{N}(0,1)$

Donc le graphe de probabilites de loi normale est le nuage des points :

$$\{x_i^*, \Phi^{-1}(\frac{i}{n})\} \ pour \ i \in \{1,2,3,...,n-1\}$$

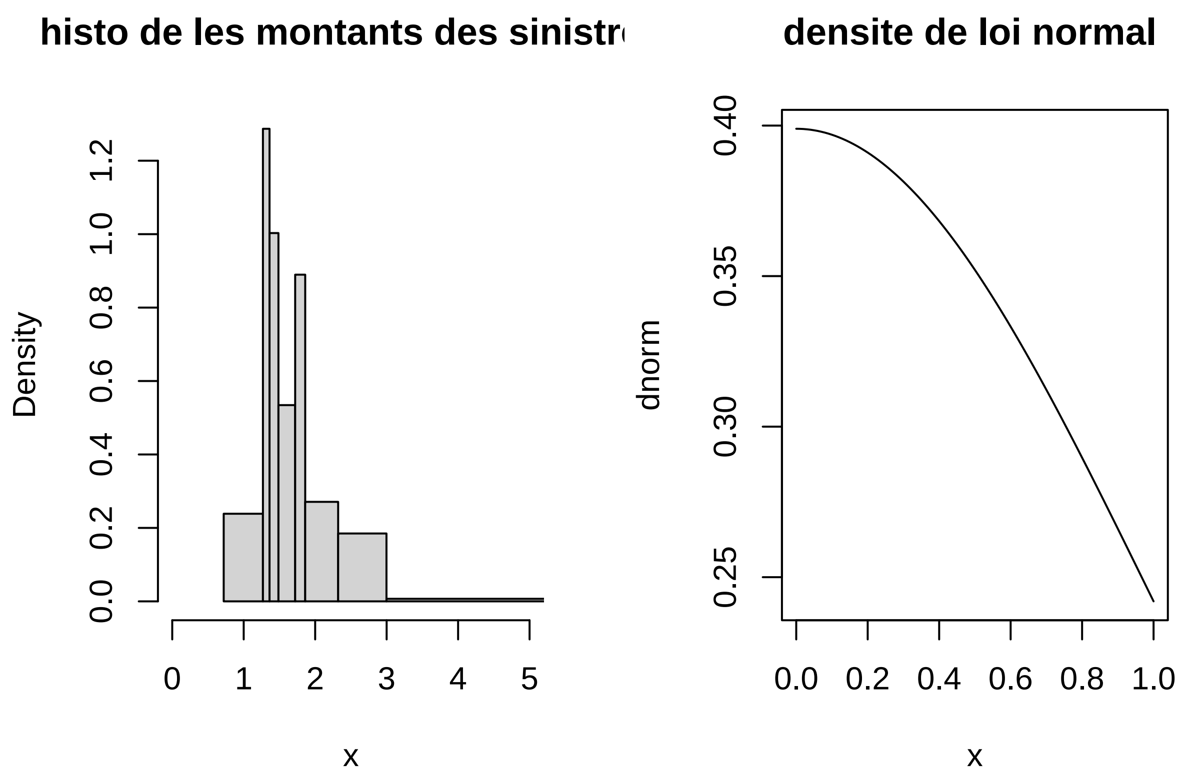
```
plot(sort(x)[1:n-1],qnorm(seq(1,n-1)/n),main="graphe de proba de la loi normale ")
```



Le non alignement des points dans le graphe ci-dessus montre bien que la loi normalle n'est pas un modèle plausible pour décrire les montants des sinistres.

- Histogrammes pour la loi normale :

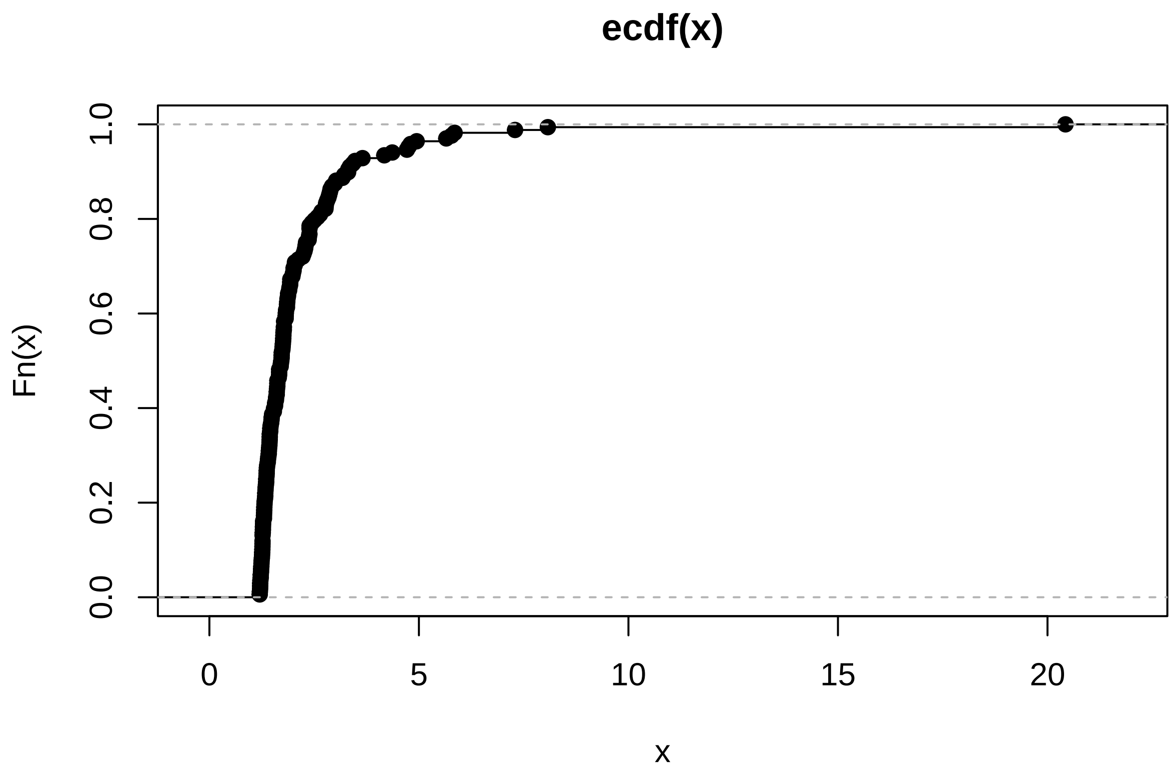
```
par(mfcol=c(1,2))
hist(x,prob=T , breaks=bornes, main="histo de les montants des sinistres",xlim=c(0,5))
plot(dnorm,main="densite de loi normal ",pch=20) # on affiche le graphe de la dénsite de al loi normale
```



l'histogramme donne une approximation de la densite de la loi correspondante . le graphe ci-dessus montre bien que la loi normale n'est pas un modèle plausible pour décrire les montants des sinistres.

Question 2

```
# la fonction de répartition émpirique
plot(ecdf(x))
```




```r
# la moyenne empiriqué
moyenne <- mean(x)
moyenne
```

```
## [1] 2.167946
```

```r
# la mediane empirique
med<-median(x)
med
```

```
## [1] 1.72
```

```r
#le min est le max
min(x)
```

```
## [1] 1.201
```

```r
max(x)
```

```
## [1] 20.43
```

```r
#"écart-type  empirique"
sd<-sqrt((n-1)/n)*sd(x)
sd
```

```
## [1] 1.799712
```

```r
# la variance émpirique
((n-1)/n)*var(x)
```

```
## [1] 3.238964
```

```r
# le coefficient de variation empirique
cv<-sd/moyenne
cv
```

```
## [1] 0.8301461
```

- Les caractéristiques essentielles des cés données :

D'aprés le calcul des indicateurs statistiques précedent on la cv=0.83 donc il y'a une variabilité important entre les montants de sinistres de plus on la médiane=1.72 et la moyenne =2.16 donc il y'a des données aberrantes dans les montatns de sinistres .

- Les principes explication que les modéles usuels de loi normale et exponentielle ne sont pas adaptés : les lois normale et exponentielle ne sont pas adaptés car d'aprés le tracage de graphes de probabilités on ne trouvent pas des points qasiment alignées.

Question3:

La fonction de répartition F de X:

Soit $x \in \mathbb{R}$

- Si $x < b$:

On a :

$$\mathbb{F}(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

comme x<b Donc :

$$\mathbb{F}(x) = 0 \quad \forall\ x < b$$

- Si non:

$$\mathbb{F}(x) = \int_{-\infty}^{x} f(t)dt = \int_{b}^{x} f(t)dt$$

$$\mathbb{F}(x) = \int_{a}^{b} \frac{ab^a}{t^{a+1}}dt = 1 - \left(\frac{b}{x}\right)^a$$

Donc :

$$\mathbb{F}(x) = \begin{cases} 0 & x < b \\ 1 - (\frac{b}{x})^a & si\ non \end{cases}$$

Espérance de X :

- X admet une espérance finie si et seulement si la fonction:

$$x \mapsto xf(x)\ est\ intégrable\ sur\ \mathbb{R}$$

$$\Leftrightarrow \quad \int_{b}^{+\infty} |\frac{ab^a}{x^a}|dx < +\infty$$

$$\Leftrightarrow \ \ a > 1\ (D'aprés\ L'intégrale\ de\ reimanne)$$

- Si $a > 1$:

$$E[X] = \int_{b}^{+\infty} \frac{ab^a}{x^a} = \frac{ab}{a-1}$$

$$E[X] = \frac{1}{1-a}[0 - b^{1-a}] = \frac{ab}{a-1}$$

$$\boxed{E[X] = \frac{ab}{a-1}}$$

Variance de X :

- X admet une variance finie ssi :

$$x \mapsto x^2 f(x) \ est \ intégrable \ sur \ \mathbb{R}$$

$$\Leftrightarrow \int_b^{+\infty} |\frac{ab^a}{x^{a-1}}|dx \prec +\infty$$

$$\Leftrightarrow \ a > 2 \ (D'aprés \ L'intégrale \ de \ reimanne)$$

- Si $a > 2$:

$$V(X) = E(X^2) - (E(X))^2$$

  or ona :

$$E(X^2) = ab^a \int_b^{+\infty} x^{1-a} dx = \frac{ab^2}{a-2}$$

Ainsi

$$V(X) = \frac{ab^2}{a-2} - \frac{ab}{a-1}^2 = \frac{ab^2}{(a-2)(a-1)^2}$$

$$\boxed{V(X) = \frac{ab^2}{(a-2)(a-1)^2}}$$

Question 4:

- Le graphe de probabilité:

On a

$$\forall x \geq b \quad \mathbb{F}(x) = 1 - \left(\frac{b}{x}\right)^a$$

$$\forall x \geq b \ \ln(1 - \mathbb{F}(x)) = a \ln(\left(\frac{b}{x}\right) = -a \ln(x) + a \ln(b)$$

$$\forall x \geq b \ \ln(1 - \mathbb{F}(x)) = -a \ln(x) + a \ln(b)$$

Donc ona :

$$\boxed{\forall x \geq b \ \ln(1 - \mathbb{F}(x)) = -a \ln(x) + a \ln(b)}$$

Donc le graphe de probabilité de la loi $P(a, b)$ est : les nuages des points

$$(\ln(x_i^*), \ln(1 - \frac{i}{n}))\forall i \in \{1, .., n-1\}$$

```
xt <-sort(x) # pour trieé x
abscisse <- log(xt[1:n-1])
y <- log(1-seq(1,n-1)/n)
reg_s <- lm(y~abscisse)
pente <- reg_s$coefficients[2] # le pente  de la droite de regression
ord <- reg_s$coefficients[1] # l'ordonne a l'origine de la droite de regression
ag <- (-1)*pente # ag:l'estimateur graphique de a
bg <- exp(ord/ag) # bg :  l'estimateur graphique de bg
ag
```

```
## abscisse
## 2.339442
```

```r
bg
```

```
## (Intercept)
##    1.226742
```

Question5

les estimateur par la methode EMM on a :

$$\begin{cases} E(X) = \frac{ab}{a-1} & => b = \frac{a-1}{a}E(X)(1) \\ V(X) = \frac{ab^2}{(a-2)(a-1)} & (2) \end{cases}$$

Donc : en ramplacnat l'expression de b (1) dans (2) on trouve que a est la soluiton de l'equation :

$$a^2 - 2a - \frac{(E(X))^2}{V(X)} = 0$$

Donc en estimant E(X) par la moyenne empirique et la V(X) par la variance estiméé ona :

l'EMM de a est la solution de l'equation du seconde dégree :

$$a^2 - 2a - \frac{(\bar{x}_n)^2}{(s'_n)^2} = 0$$

on pose $\alpha = \frac{(\bar{x}_n)^2}{(s'_n)^2}$

- la résolution de l'équation du séconde degrée $a^2 - 2a - \alpha = 0$ donne deux solutions :

$$\begin{cases} a = 1 + \sqrt{1+\alpha} \\ a = 1 - \sqrt{1+\alpha} < 0 \quad \textit{donc regeter car } a > 1 \end{cases}$$

donc : L'EMM de a est :

$$\boxed{\tilde{a}_n = 1 + \sqrt{1 + \frac{(\bar{x}_n)^2}{(s'_n)^2}}}$$

est l'EMM de b est :

$$\tilde{b}_n = \frac{\tilde{a}_n - 1}{\tilde{a}_n}\bar{x}_n$$

$$\boxed{\tilde{b}_n = (1 - \frac{1}{1 + \sqrt{1 + \frac{(\bar{x}_n)^2}{(s'_n)^2}}})\bar{x}_n}$$

Question 6 :

les estimateur par la methode EMV : On calclue tout d'abord la fonction de maximum de vraisemblance $l(a, b; x_1, ..., x_n)$

soient $(x_1, ..., x_n) \in \mathbb{R}^{\ltimes}$:

$$l(a, b; x_1, ..., x_n) = \prod_{i=1}^{n} f_X(x_i) \qquad \textit{car les } (X_i)_i \textit{ sont i.i.d}$$

$$l(a,b;x_1,....,x_n) = \prod_{i=1}^{n} \frac{ab^a}{x_i^{1+a}} \mathbf{1}_{\{x_i \geq b\}}$$

$$l(a,b;x_1,....,x_n) = a^n (b^{na}) \frac{\prod_{i=1}^{n} \mathbf{1}_{\{x_i \geq b\}}}{\prod_{i=1}^{n} x_i^{1+a}}$$

- l'EMV de b $\hat{b}_n$:

pour maximiser $l(a,b;x_1,....,x_n)$ en fonction de b il faut chosir b de telle sorte que on maximise $\prod_{i=1}^{n} \mathbf{1}_{\{x_i \geq b\}}$

Donc l'EMV de b est :

$$\boxed{\hat{b}_n = \min_{1 \leq i \leq n} (x_i)}$$

- calcul de l'expression de $\hat{a}_n$: ona :

$$l(a,b;x_1,....,x_n) = a^n (b^{na}) \frac{\prod_{i=1}^{n} \mathbf{1}_{\{x_i \geq b\}}}{\prod_{i=1}^{n} x_i^{1+a}}$$

pour $b = \min_{1 \leq i \leq n}(x_i) = \hat{b_n}$ ona :

$$l(a,b;x_1,....,x_n) = a^n \hat{b_n}^{na} \frac{n}{\prod_{i=1}^{n} x_i^{1+a}}$$

$$\ln(l(a,b;x_1,....,x_n)) = n\ln(a) + an\ln(\hat{b_n}) + \ln(n) - \sum_{i=1}^{n} \ln(x_i)(1+a)$$

$$\frac{\partial \ln(l(a,b;x_1,....,x_n))}{\partial a} = \frac{n}{a} + n\ln(\hat{b_n}) + - \sum_{i=1}^{n} \ln(x_i)$$

Donc l'EMV de a est la soluion de l'equation :

$$\frac{\partial \ln(l(a,b;x_1,....,x_n))}{\partial a} = 0$$

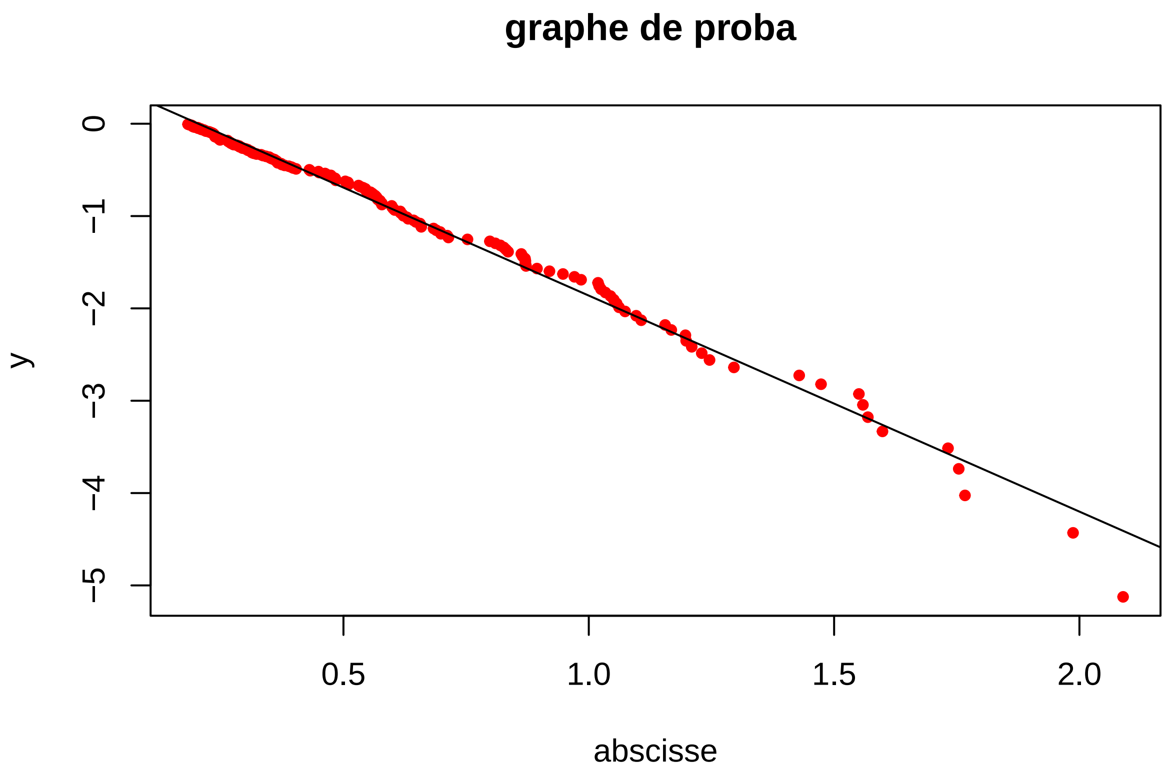
Donc :

$$\hat{a}_n = \frac{n}{\sum_{i=1}^{n} \ln(\frac{x_i}{\hat{b}_n})}$$

Donc on a les estimateures par Méthode de maximum de vraiseemblance de a et b sont :

$$\begin{cases} \hat{a}_n = \frac{n}{\sum_{i=1}^{n} \ln(\frac{x_i}{b})} \\ \hat{b}_n = \min_{1 \leq i \leq n}(x_i) \end{cases}$$

```
# le graphe de proba de la loi Pa(a,b)
reg_s <- lm(y~abscisse)
xt <-sort(x) # pour trieé x
abscisse <- log(xt[1:n-1])
y <- log(1-seq(1,n-1)/n)
plot(abscisse,y,main="graphe de proba ",col="red",pch=20)
abline(reg_s)
```



Question 7 :
D'apres le graphe de proba on a les points sont presque aligneés donc la loi Pa(a,b) est une modéle plausible pour les montants de sinistres.

```
# calcules des estimateures graphiques
reg_s <- lm(y~abscisse)
pente <- reg_s$coefficients[2] # le pente  de la droite de regression
ord <- reg_s$coefficients[1] # l'ordonne a l'origine de la droite de regression
ag <- (-1)*pente # ag:l'estimateur graphique de a
bg <- exp(ord/ag) # bg :  l'estimateur graphique de bg
ag
```

```
## abscisse
## 2.339442
```

```
bg
```

```
## (Intercept)
##    1.226742
```

```
# calcules des estimateures par la methode des momentes
a_EMM <- 1+sqrt(1+(mean(x))^{2}/var(x))
b_EMM <- ((a_EMM -1)/a_EMM)*mean(x)
a_EMM
```

```
## [1] 2.562831
```

```
b_EMM
```

```
## [1] 1.322028
```

```r
# calcule des estimateures par la méthode EMV
b_EMV <- min(x)
a_EMV <-  n/(sum(log(x/b_EMV)))
a_EMV
```

```
## [1] 2.19037
```

```r
b_EMV
```

```
## [1] 1.201
```

Partie2 :Vérifications expérimentales à base de simulations

Question 1 :

Loi de Y=ln(X/b) on calcul la fonction de répartition $\mathbb{F}_Y(x)$ $\forall x \in \mathbb{R}$

- si x<0 :

$$\mathbb{F}_Y(x) = \mathbb{P}(Y \leq x) = \mathbb{P}\left(\ln\left(\frac{X}{b}\right) \leq x\right)$$

$$\mathbb{F}_Y(x) = \mathbb{P}(X \leq be^x)$$

  or x<0 d'ou $e^x < 1$ donc $be^x < b$ or $\forall\ y < b \ \ P(X \leq y) = 0$ donc :

$$\forall \ \ x < 0 \quad \mathbb{F}_Y(x) = 0$$

- si $x \geq 0$:

$$\mathbb{F}_Y(x) = \mathbb{P}(Y \leq x) = \mathbb{P}(X \leq be^x)$$

$$\mathbb{F}_Y(x) = \mathbb{F}_X(be^x)$$

$$\mathbb{F}_Y(x) = 1 - (\frac{b}{be^x})^a$$

$$\mathbb{F}_Y(x) = 1 - e^{-ax}$$

  Donc :

$$\forall\ x \in \mathbb{R} \qquad \mathbb{F}_Y(x) = \begin{cases} 0 & si\ x < 0 \\ 1 - e^{-ax} & si\ non \end{cases}$$

  Donc :

$$\mathbb{Y} \sim \varepsilon(a) : loi \ \ exponentielle \ \ \ de \ \ \ parametre \ \ a$$

```r
# Pour simuler une loi de P(a,b) on simule la loi exp(a)
sim_echantillon_n <-function(a,b,n)
{
  y <- rexp(n,rate=a)  # simulation d'une échantillon de taille n de la loi exp
  x <- b*exp(y)   # une échnatillon de taille n de la loi Pa(a,b)

  return(x)
}

# Test de notre fonction
l<-sim_echantillon_n(3,4,10)
l
```

```
##  [1]  5.502175  4.203733  6.722508  4.442524  4.493755  4.804792  4.229354
##  [8] 13.851832  5.533200  4.995839
```

question 2

a):L'intervalle de cofiance de seuil α pour a : soit $Y_1, ..., Y_n$ des VA tq $\forall \ \ i \in \ \{1, 2, ..., n\} \quad Y_i = \ln\left(\frac{X_i}{b}\right)$

Commme les $(X_i)_i$ sont i.i.d donc $(Y_i)_i$ sont i.i.d de méme loi $\varepsilon(a)$

D'aprés le TCL:Théoréme centrale limite

$$on \ a \quad \sqrt{n}\frac{\bar{Y}_n - E(Y_1)}{\sqrt{var(Y_1)}} \sim \mathcal{N}(0,1) \quad lorsque \ \ n \ \ tend \ vers + \infty$$

Donc :

$$\mathbb{P}\left(\left|\frac{\bar{Y}_n - E(Y_1)}{\sqrt{var(Y_1)}}\right| \leq u_\alpha\right) = 1 - \alpha$$

or :

$$Y_1 \sim \varepsilon(a) \qquad Donc \quad E(Y_1) = \frac{1}{a} \qquad et \quad var(Y_1) = \frac{1}{a^2}$$

Donc :

$$\left|\frac{\bar{Y}_n - E(Y_1)}{\sqrt{var(Y_1)}}\right| \leq u_\alpha \quad ssi \qquad \left|\frac{1}{n}\sum_{i=1}^{n}\ln(\frac{X_i}{b}) - \frac{1}{a}\right| \leq \frac{u_\alpha}{a\sqrt{n}}$$

$$or \quad \hat{a}_n = \frac{n}{\sum_{i=1}^{n}\ln(\frac{x_i}{b})}$$

$$Donc \qquad \left|\frac{\bar{Y}_n - E(Y_1)}{\sqrt{var(Y_1)}}\right| \leq u_\alpha \quad ssi \qquad |a - \hat{a}_n| \leq \frac{u_\alpha}{a\sqrt{n}}$$

Donc un intérvalle de cofiance de seuil α de a est :

$$IC_\alpha(a) = \left[\hat{a}_n - \frac{u_\alpha}{\sqrt{n}}|\hat{a}_n| \ \ , \ \ \hat{a}_n + \frac{u_\alpha}{\sqrt{n}}|\hat{a}_n|\right]$$

```r
# on definit une fonction qui returne
#les deux bornes de L'intervalle de cofaince de seuil alpha de a
#prend en paramétre b et  l'échantillon x et le seuil alpha
intervalle_cofaince <- function(b,x,alpha)
{
  n <- length(x)
  u_aplha <- qnorm(1-alpha/2)
  a_EMV <- n/(sum(log(x/b)))
  borne_min <- a_EMV - (u_aplha/(sqrt(n)))*abs(a_EMV)
  borne_max <-  a_EMV + (u_aplha/(sqrt(n)))*abs(a_EMV)
  ic <- c(borne_min,borne_max)
  return(ic)
}

# on test notre  fonction
ech <- sim_echantillon_n(5,2,100)
ic <- intervalle_cofaince(2,ech,0.05)
ic
```

b):

```
## [1] 3.786283 5.632289
```

```r
# on definit  une fonction simu_n_echantillon_m_fois
# qui simule m echantiloon de la loi Pa(a,b) de taille n
# et on return un vecteur de 0 et 1
# tq 1 : si a est dans IC et 0 si non
simu_n_echantillon_m_fois<- function(a,b,n,m,alpha)
{
  val_p<-c()

  for(i in 1:m)
  {
    ech <- sim_echantillon_n(a,b,n)
    ic <- intervalle_cofaince(b,ech,alpha)
    if(a>=ic[1] && a <=ic[2])
    {
      val_p <-c(val_p,1) # si on trouve a dans IC on ajout 1

    }  else{
      val_p <-c(val_p,0) # si non on ajout 0
    }

  }

  return(val_p)

}

# on dfinit notre fonctino de test
# qui return la proportion des IC de seuil alpaha contant a
test <-function(a,n,m,alpha)
{
  val_p <- simu_n_echantillon_m_fois(a,2,n,m,alpha)
  proption <- mean(val_p)
  return(proption)
}

test(5,1000,100,0.05)
```

```
## [1] 0.99
```

```r
test(4,1000,30,0.01)
```

```
## [1] 0.9
```

```r
test(8,1000,50,0.15)
```

```
## [1] 0.92
```

```r
test(7,1000,30,0.15)
```

```
## [1] 0.7
```

```r
test(3,2000,60,0.10)
```

```
## [1] 0.9666667
```

- Commentaire : D'aprés les différents résulats retourner par la fonction test on trouve bien que la proprtion des intervalles des cofiance de seuil α

obtenus contient la vrai valeur de a est bien : $1-\alpha$

Question 3

```
milleur_estimateur <- function(a,b,n,m)
{
  # les estimateures graphiques
  a_graphe <- c() # on initialise le vecteur qui contient l'estimateur graphique a_g
  b_graphe <-c() # on initialise le vecteur qui contient l'estimateur graphique b_g

  # les estimateures par EMM
  a_EMM_l <-c()  # on initialise le vecteur qui contient l'estimateur par EMM de a a_EMM
  b_EMM_l <-c() # on initialise le vecteur qui contient l'estimateur par EMM de b b_EMM

  # les estimateures par EMV
  a_EMV_l <-c() # on initialise le vecteur qui contient l'estimateur par EMV  de a a_EMV
  b_EMV_l <-c() # on initialise le vecteur qui contient l'estimateur par EMV de b b_EMV

  # On simule mnt m échantillon :
  for(i in 1:m)
  {

    #on récupére l'echantillon :
    ech <-sim_echantillon_n(a,b,n)

    # les estimateures  graphiques
    ordd<-log(1-seq(1,n-1)/n)
    abs<-log(sort(ech))[1:n-1]
    reg <- lm(ordd~abs)
    pente <- reg$coefficients[2]
    ord <- reg$coefficients[1]
    ag <- (-1)*pente
    bg <- exp(ord/ag)

    a_graphe <- c(a_graphe,ag)
    b_graphe <- c(b_graphe,bg)

    # les estimateures des EMM
    a_EMM <- 1+sqrt(1+(mean(ech))^{2}/var(ech))
    b_EMM <-((a_EMM -1)/a_EMM)*mean(ech)

    a_EMM_l <- c( a_EMM_l,a_EMM)
    b_EMM_l <- c(b_EMM_l,b_EMM)

    # les estimateures des EMV
    b_EMV <-min(ech)
    a_EMV <- n/(sum(log(ech/b_EMV)))

    a_EMV_l <- c(a_EMV_l,a_EMV)
    b_EMV_l <- c(b_EMV_l,b_EMV)

    }
```

```r
  # calclues des  biais
  bias_ag <- mean(a_graphe)-a
  bias_bg <- mean(b_graphe)-b
  bias_a_EMM <- mean(a_EMM_l)-a
  bias_b_EMM <- mean(b_EMM_l)-b
  bias_a_EMV <- mean(a_EMV_l)-a
  bias_b_EMV <- mean(b_EMV_l)-b
  # calclues des  EMQ
  EQM_ag <- var(a_graphe)*((n-1)/n) + bias_ag^{2}
  EQM_bg <- var(b_graphe)*((n-1)/n) + bias_bg^{2}
  EQM_a_EMM <- var(a_EMM_l)*((n-1)/n) + bias_a_EMM^{2}
  EQM_b_EMM <- var(b_EMM_l)*((n-1)/n) + bias_b_EMM^{2}
  EQM_a_EMV <- var(a_EMV_l)*((n-1)/n) + bias_a_EMV^{2}
  EQM_b_EMV <- var(b_EMV_l)*((n-1)/n) + bias_b_EMV^{2}



  # le meilleur  estimateur de a est celui de biais min et var min
  cat( "l'affichage de biais  des estimateures de a","\n")
  cat("\n")
  cat("le bias de l'estimateur  a_graphe est  ",bias_ag,"\n")
  cat("les bias de  l'estimateur  a_EMM  est:  ",bias_a_EMM,"\n")
  cat("les bias de  l'estimateur  a_EMV  est ",bias_a_EMV,"\n")
  cat("\n")
  cat("l'affichage de EQM des estimateures de a ","\n")
  cat("\n")
  cat("L'erreur quadratique moyenne de l'estimateur  a_graphe est  ",EQM_ag,"\n")
  cat("L'erreur quadratique moyenne de l'estimateur  a_EMM est :  ",EQM_a_EMM,"\n")
  cat("L'erreur quadratique moyenne de l'estimateur  a_EMV est  ",EQM_a_EMV,"\n")
  cat("\n")
  cat("\n")
  # le meilleur   estimateur de b :
  cat("l'affichage de biais  des estimateures de b","\n")
  cat("les bias de l'estimateur   b_graphe est  ",bias_bg,"\n")
  cat("les bias de  l'estimateur  b_EMM  est ",bias_b_EMM,"\n")
  cat("les bias de  l'estimateur  b_EMV  est ",bias_b_EMV,"\n")
  cat("\n")
  cat("l'affichage de EQM des estimateures de b ","\n")
  cat("\n")
  cat("L'erreur quadratique moyenne de l'estimateur  b_graphe est   ",EQM_bg,"\n")
  cat("L'erreur quadratique moyenne de l'estimateur  b_EMM est  ",EQM_b_EMM,"\n")
  cat("L'erreur quadratique moyenne de l'estimateur  b_EMV est  ",EQM_b_EMV,"\n")
}
# On tset la fonction :
cat("Test 1","\n")
```

```
## Test 1
```

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

```r
milleur_estimateur(3,2,100,50)
```

```
## l'affichage de biais  des estimateures de a
##
```

```
## le bias de l'estimateur  a_graphe est   0.108097
## les bias de  l'estimateur  a_EMM  est:   0.3757345
## les bias de  l'estimateur  a_EMV  est  0.09193957
##
## l'affichage de EQM des estimateures de a
##
## L'erreur quadratique moyenne de l'estimateur  a_graphe est   0.1050365
## L'erreur quadratique moyenne de l'estimateur  a_EMM est :   0.3492525
## L'erreur quadratique moyenne de l'estimateur  a_EMV est   0.07330301
##
##
## l'affichage de biais  des estimateures de b
## les bias de l'estimateur   b_graphe est   -0.006391803
## les bias de  l'estimateur  b_EMM  est  0.07777306
## les bias de  l'estimateur  b_EMV  est  0.00550629
##
## l'affichage de EQM des estimateures de b
##
## L'erreur quadratique moyenne de l'estimateur  b_graphe est    0.002833892
## L'erreur quadratique moyenne de l'estimateur  b_EMM est   0.01550136
## L'erreur quadratique moyenne de l'estimateur  b_EMV est   5.20393e-05
```

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

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

```
cat("Test2","\n")
```

```
## Test2
```

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

```
milleur_estimateur(4,3,1000,100)
```

```
## l'affichage de biais  des estimateures de a
##
## le bias de l'estimateur  a_graphe est   0.0421605
## les bias de  l'estimateur  a_EMM  est:   0.1366087
## les bias de  l'estimateur  a_EMV  est  0.009921271
##
## l'affichage de EQM des estimateures de a
##
## L'erreur quadratique moyenne de l'estimateur  a_graphe est   0.03896384
## L'erreur quadratique moyenne de l'estimateur  a_EMM est :   0.118629
## L'erreur quadratique moyenne de l'estimateur  a_EMV est   0.01910892
##
##
## l'affichage de biais  des estimateures de b
## les bias de l'estimateur   b_graphe est   0.004013541
## les bias de  l'estimateur  b_EMM  est  0.02482596
## les bias de  l'estimateur  b_EMV  est  0.0009558613
##
## l'affichage de EQM des estimateures de b
##
## L'erreur quadratique moyenne de l'estimateur  b_graphe est    0.0004398773
## L'erreur quadratique moyenne de l'estimateur  b_EMM est   0.004411829
## L'erreur quadratique moyenne de l'estimateur  b_EMV est   1.782069e-06
```

- Conclusion : D'aprés les différents resultats retourner par la fonction milleur_estimateur : on trouve que l'estimateur de biais miniamale et de

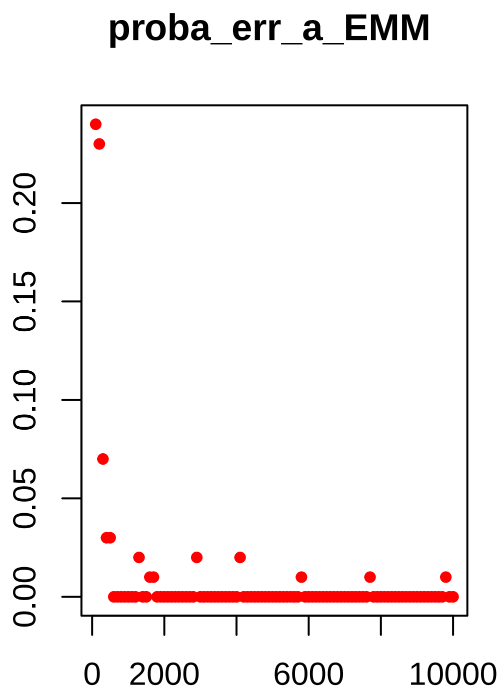
variance minimale est celuide la méthode EMV

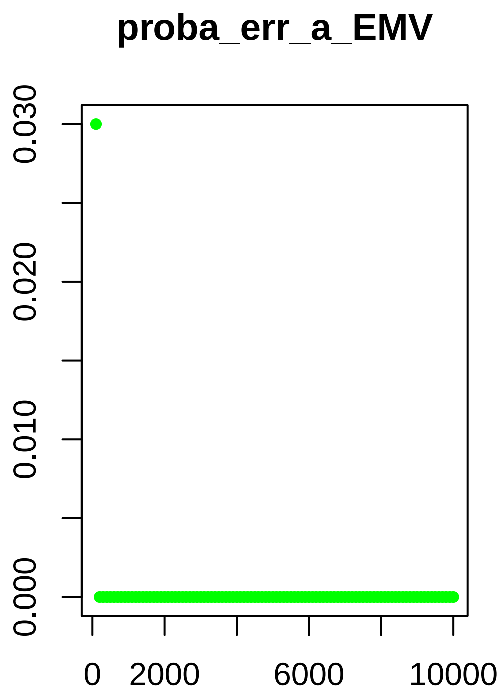
Donc les mielleurs estimateurs de a et b : sont les estimatures $\hat{a}_n$ et $\hat{b}_n$.

Question 4 :

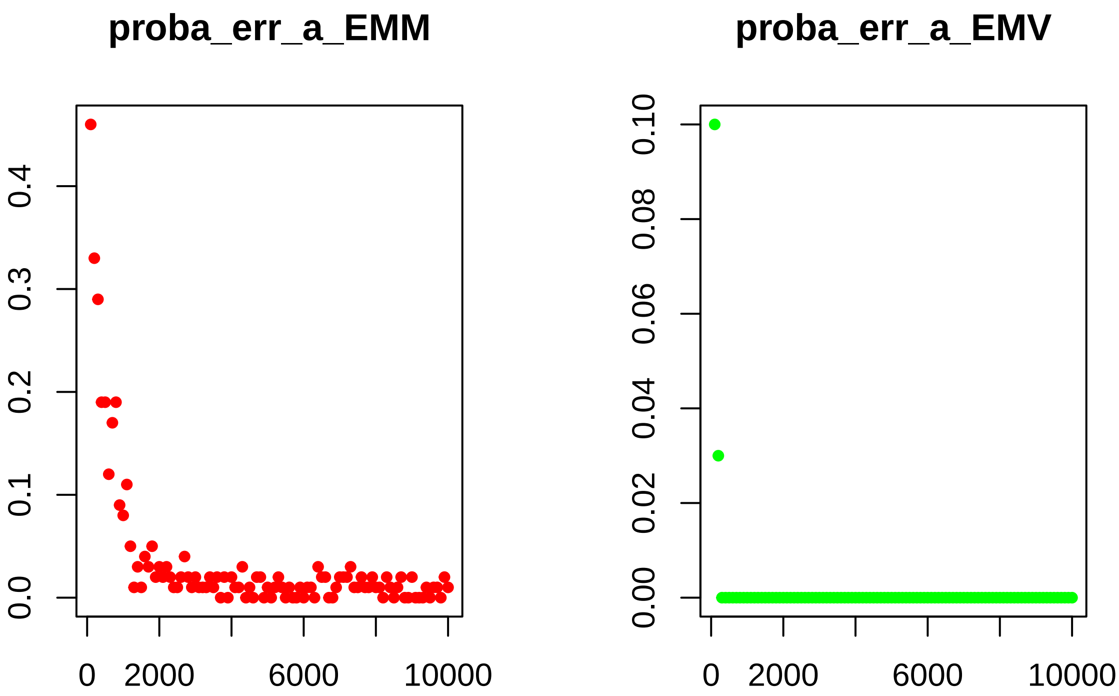
```
# on definit une fonction  qui return la P(|Tn-a|>epsilon)
proba_err <-function(a,b,n,m,epsilon)
{

  val_err_a_EMM <-c()
  val_err_a_EMV <-c()

  for(i in 1:m)
  {
    ech <- sim_echantillon_n(a,b,n) # ech est unne échantillon de taille  n de la loi Pa(a,b)
    a_EMM <- 1+sqrt(1+(mean(ech))^{2}/var(ech)) # l'estimateur par EMM de a
    a_EMV <- n/(sum(log(ech/min(ech))))  # l'éstimateur par EMVV de a
    err_a_EMM <- abs(a_EMM-a)
    err_a_EMV  <- abs(a_EMV -a)

    # calcul de val_err_a_EMM
    if(err_a_EMM > epsilon)
    {
      val_err_a_EMM <-c(val_err_a_EMM,1) # si  abs(a_EMM - a ) > epsilon on ajoute 1 a val_err_a_EMM
    }
    else
    {
        val_err_a_EMM <-c(val_err_a_EMM,0) # si non on ajoute 0
    }

    # calcly de val_err_EMV
    if(err_a_EMV > epsilon)
    {
      val_err_a_EMV <-c(val_err_a_EMV,1) # si  abs(a_EMV - a ) > epsilon on ajoute 1 a val_err_a_EMV
    }
    else
    {
        val_err_a_EMV <-c(val_err_a_EMV,0)  # si non on ajoute 0
    }

  }

  p_a_EMM <-mean(val_err_a_EMM)  # on deduit une estimation de la P(|a_EMM-a|>epsilon)
  p_a_EMV <- mean(val_err_a_EMV)  # on deduit une estimation de la P(|a_EMV-a|>epsilon)

  return(c(p_a_EMM,p_a_EMV))  # on return les deux proba dans un vecteur de taille 2



}
```

```
proba_err_n_EMM<-function(a,b,n,m,epsilon)
{
    err_a_EMM <-c() # on stocke ici l'ensembles de  P(|a_EMM-a|>epsilon) avec n varie

    for(i in seq(100,n,100))
    {
        prob_err <- proba_err(a,b,i,m,epsilon)
        err_a_EMM <- c(err_a_EMM,prob_err[1])

    }
    return(err_a_EMM)

}

proba_err_n_EMV <- function(a,b,n,m,epsilon)
{
    err_a_EMV <-c()  # on stocke ici l'ensembles de  P(|a_EMV-a|>epsilon) avec n varie

    for(i in seq(100,n,100))
    {
        prob_err <- proba_err(a,b,i,m,epsilon)
        err_a_EMV <-c(err_a_EMV,prob_err[2])

    }
    return(err_a_EMV)

}

# deux courbes pour epsilon=0.8
par(mfcol=c(1,2))
plot(seq(100,10000,100),proba_err_n_EMM(3,2,10000,100,0.8),col="red",xlim=c(100,10000),pch=20,main="prol
plot(seq(100,10000,100),proba_err_n_EMV(3,2,10000,100,0.8),col="green",xlim=c(100,10000),pch=20,main="p
```

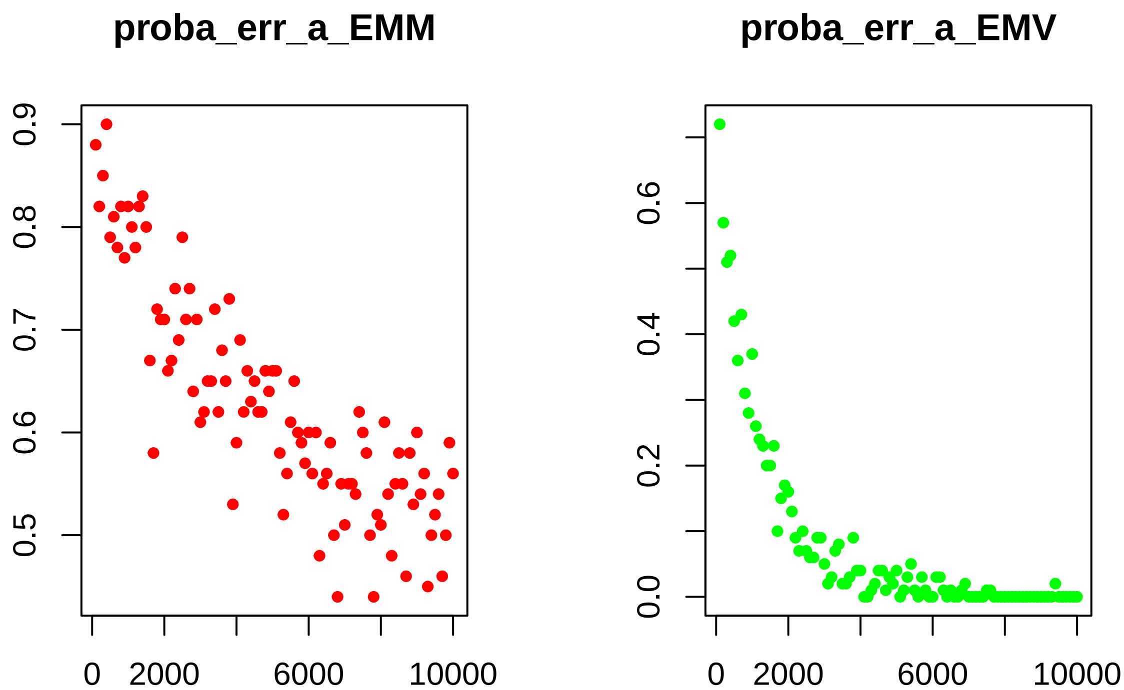





```
# deux courbes pour epsilon=0.5
par(mfcol=c(1,2))
plot(seq(100,10000,100),proba_err_n_EMM(3,2,10000,100,0.5),col="red",xlim=c(100,10000),pch=20,main="prol
plot(seq(100,10000,100),proba_err_n_EMV(3,2,10000,100,0.5),col="green",xlim=c(100,10000),pch=20,main="p
```




```
# deux courbes pour epsilon=0.1
par(mfcol=c(1,2))
plot(seq(100,10000,100),proba_err_n_EMM(3,2,10000,100,0.1),col="red",xlim=c(100,10000),pch=20,main="prol
plot(seq(100,10000,100),proba_err_n_EMV(3,2,10000,100,0.1),col="green",xlim=c(100,10000),pch=20,main="p
```

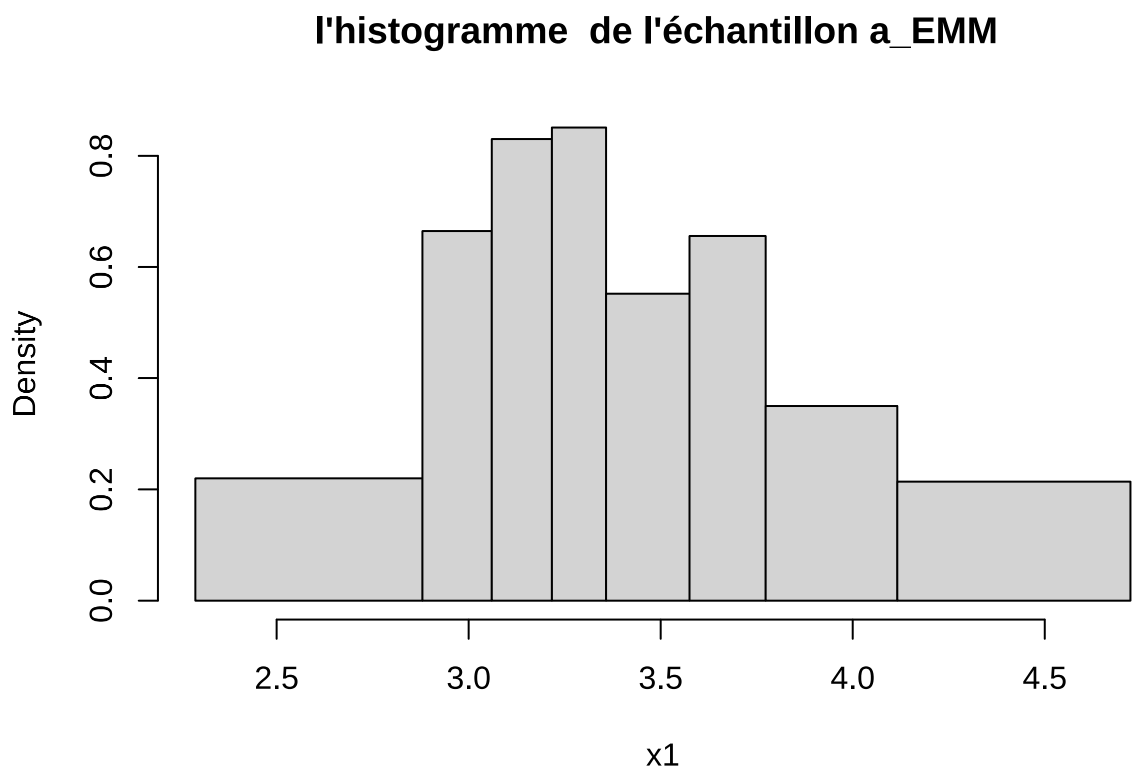


D'aprés le graphe de de $\mathbb{P}(|\tilde{a}_n - a| > \epsilon)$ en fonction de n et le graphe de $P(|\hat{a}_n - a| > \epsilon)$ en fonction de n on vérifie bien la convergence en probabiité de l'estimateur EMM de a et l'estimateur EMV de a vers a et de plus on touve que l'éstimateur qui convergent le plus vite vers a est l'éstimatur par EMV .

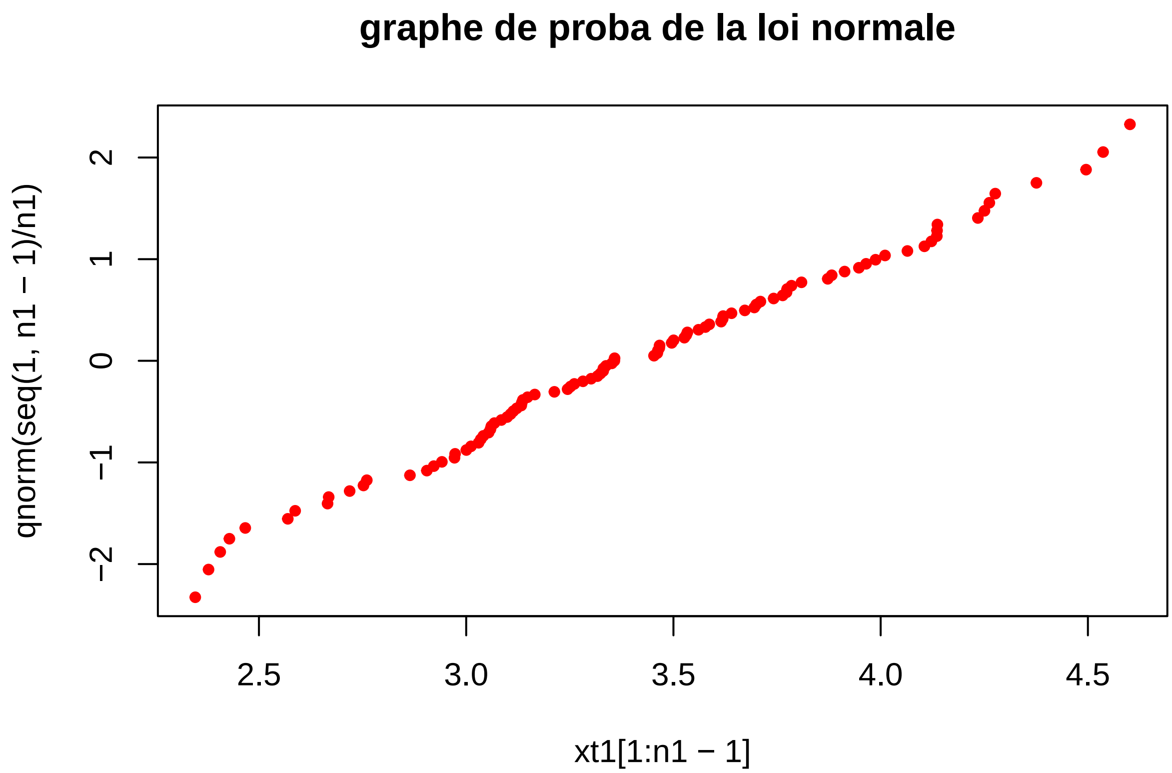
Question 5 :

```
normalite<-function(a,b,n,m)
{
  val_a_EMM <-c()
  val_a_EMV <-c()


  for(i in 1:m)
  {
    ech <- sim_echantillon_n(a,b,n)
    a_EMM <- 1+sqrt(1+(mean(ech))^{2}/var(ech))
    b_EMV <-min(ech)
    a_EMV <- n/(sum(log(ech/b_EMV)))
    val_a_EMM<-c(val_a_EMM,a_EMM)
    val_a_EMV<-c(val_a_EMV,a_EMV)

  }


  return(c(val_a_EMM,val_a_EMV))
```

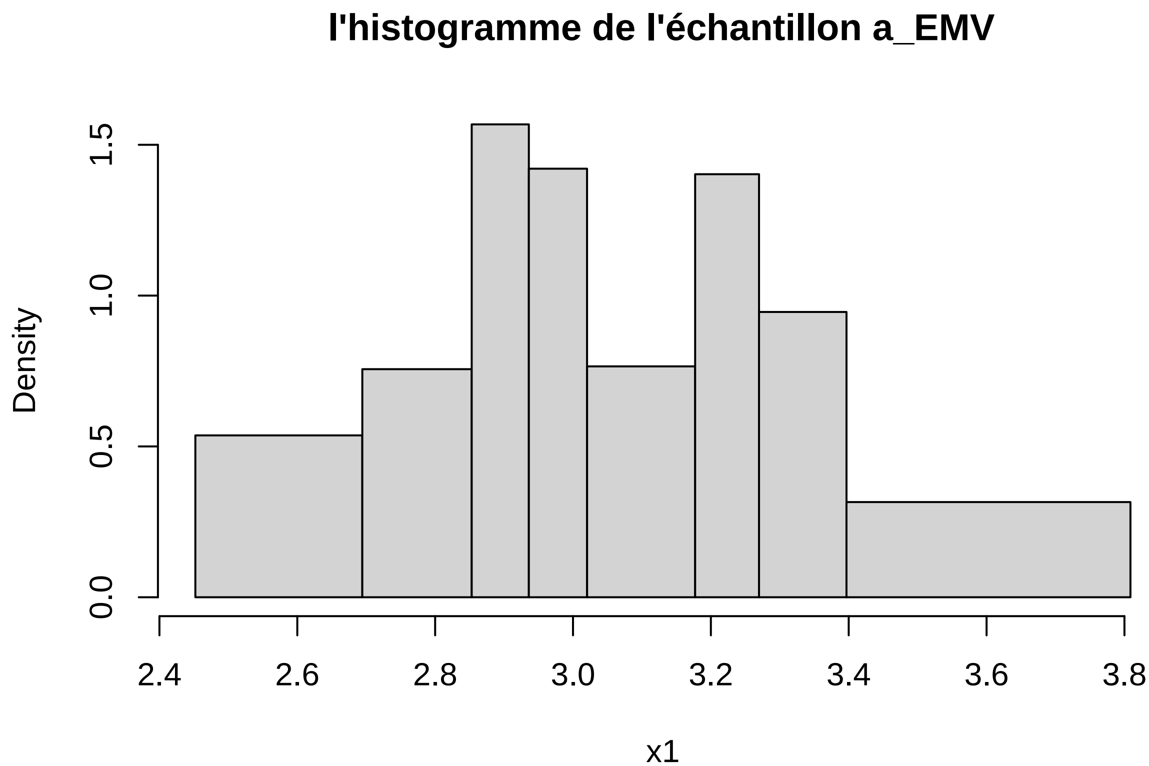
```
}
# on fixe les valeures de a et b
a <- 3
b <- 4
x <-normalite(a,b,100,100)
# l'histogramme de l'échantillon a_EMM
x1<-x[1:100]
n1 <- length(x1)
k1<-round(1+log2(n1))
xt1<-sort(x1)
a0<-xt1[1]-0.025*(xt1[n1]-xt1[1])
ak<-xt1[n1]+0.025*(xt1[n1]-xt1[1])
h<-(ak-a0)/n1
bornes1 <- c(a0,quantile(x1,seq(1,k1-1)/k1),ak)
indice_der<-length(bornes1)
hist(x1,prob=T,breaks=bornes1,xlim=c(bornes1[1],bornes1[indice_der]),main="l'histogramme  de l'échantill
```



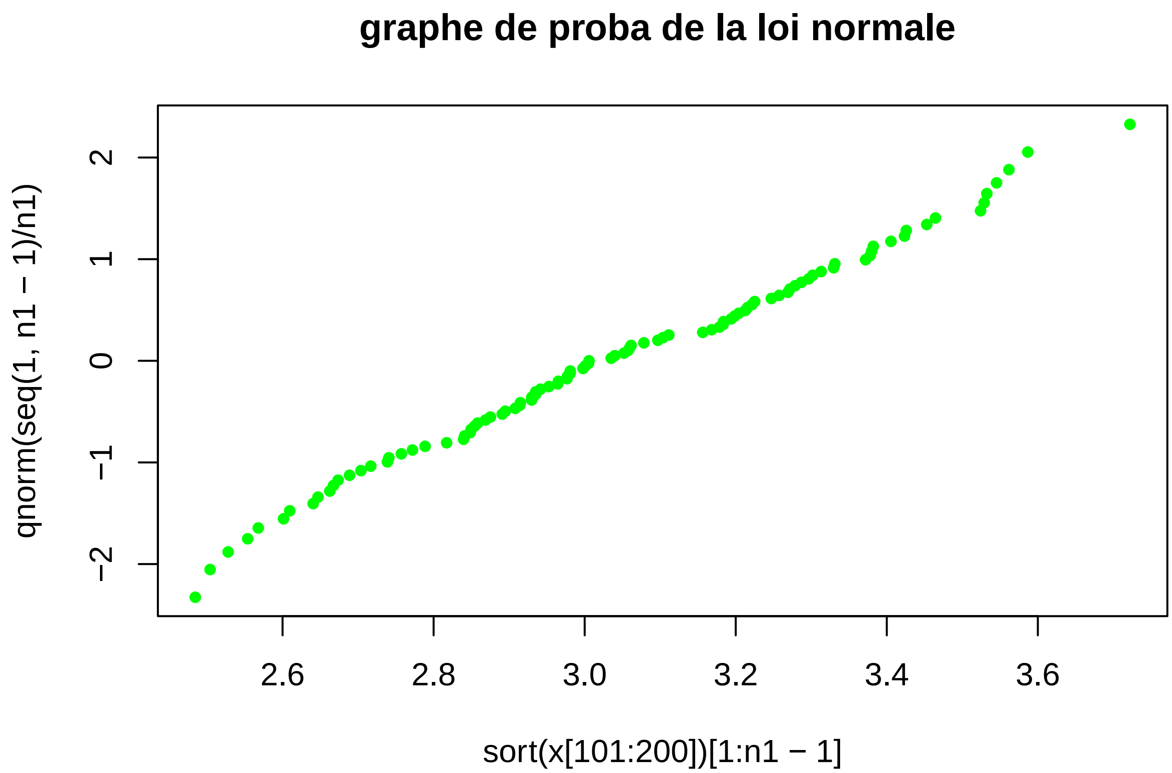

```
# le graphe de proba  de  l'échantillon  a_EMV
plot(xt1[1:n1-1],qnorm(seq(1,n1-1)/n1),main="graphe de proba de la loi normale ",pch=20,col="red")
```




```
# l'histogramme de l'échantillon a_EMV
x1<-x[101:200]
n1 <- length(x1)
k1<-round(1+log2(n1))
xt1<-sort(x1)
a0<-xt1[1]-0.025*(xt1[n1]-xt1[1])
ak<-xt1[n1]+0.025*(xt1[n1]-xt1[1])
h<-(ak-a0)/n1
bornes1 <- c(a0,quantile(x1,seq(1,k1-1)/k1),ak)
indice_der<-length(bornes1)
hist(x1,prob=T,breaks=bornes1,xlim=c(bornes1[1],bornes1[indice_der]),main=" l'histogramme de l'échantil
```




```
#le graphe de proba de l'échantillon a_EMV
plot(sort(x[101:200])[1:n1-1],qnorm(seq(1,n1-1)/n1),main="graphe de proba de la loi normale ",pch=20,col
```



- Conclusion :

D'aprés le tracage des histogramme et des graphe de proba de la loi normale pour les échnatillons des m estimations $\tilde{a}_n$ et $\hat{a}_n$ on deuire le cv des estimateures $\tilde{a}_n$ et $\hat{a}_n$ vers la loi normale $\mathcal{N}(\mu, \sigma^2)$