Modélisation des Sinistres IARD

Ahmed EL YOUSEFI

IMPORTATION DES DONNÉES :

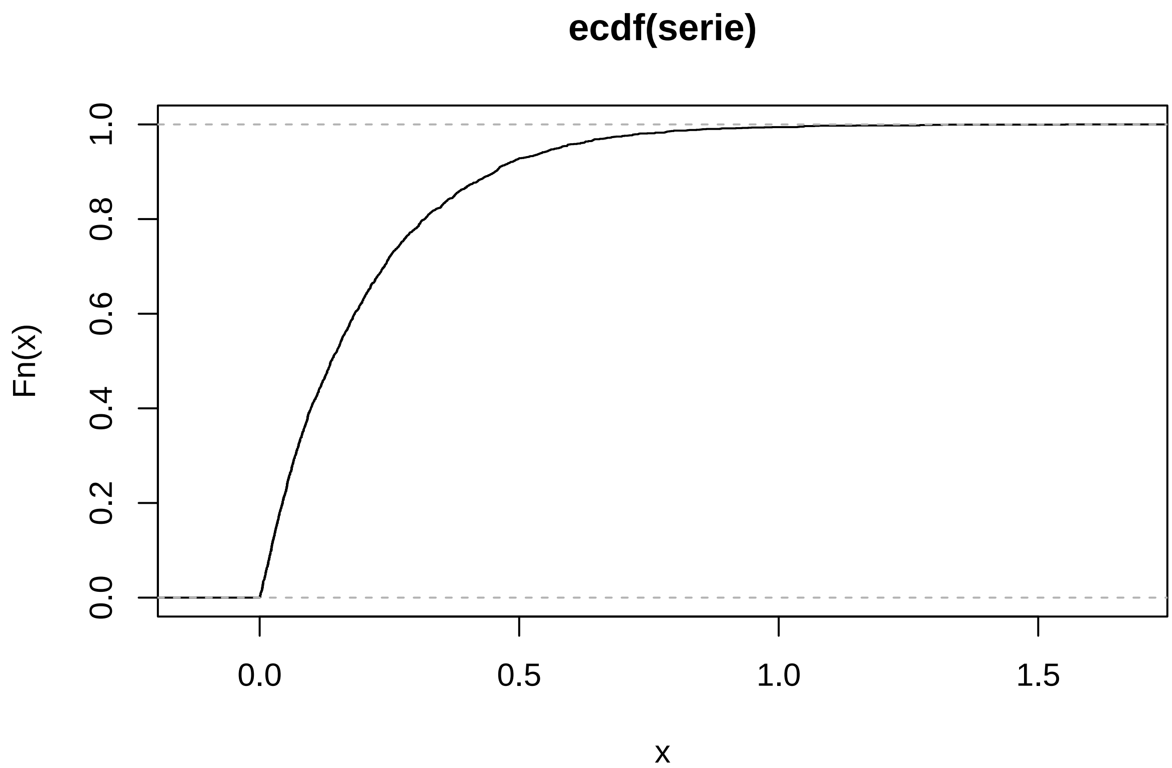
```
setwd("C:/Users/DELL/Desktop/IARD")
table <- read.csv("montantssinistre16.csv",sep=";",dec=",")
serie <- table[,2]
head(table)
```

```
##   X          x
## 1 1 0.07675949
## 2 2 0.04085351
## 3 3 0.17902609
## 4 4 0.06165967
## 5 5 0.11395614
## 6 6 0.22183424
```

ANALYSE EXPLORATOIRE :

Faisons le plotage de la fonction de repartition empirique de l'echantillon :

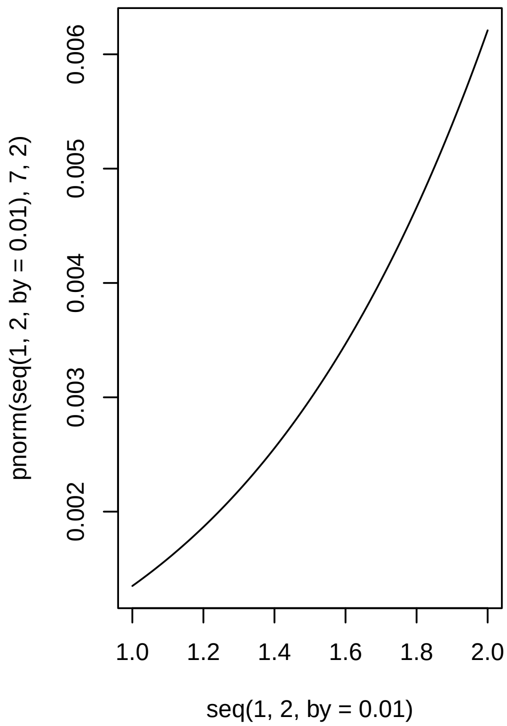
```
plot(ecdf(serie))
```

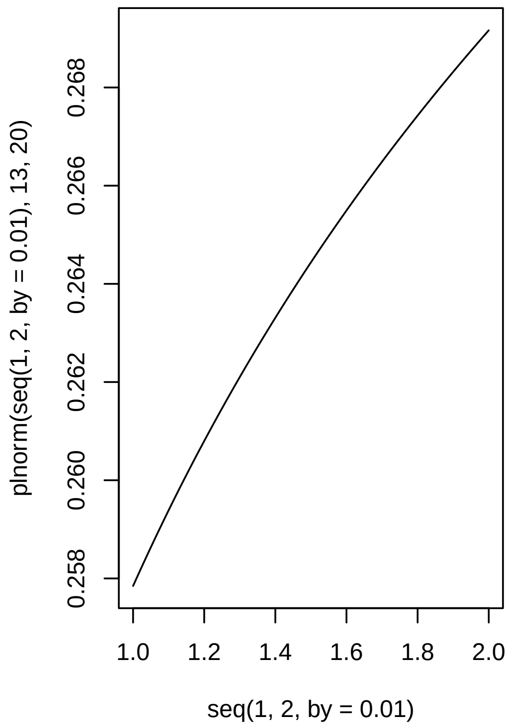


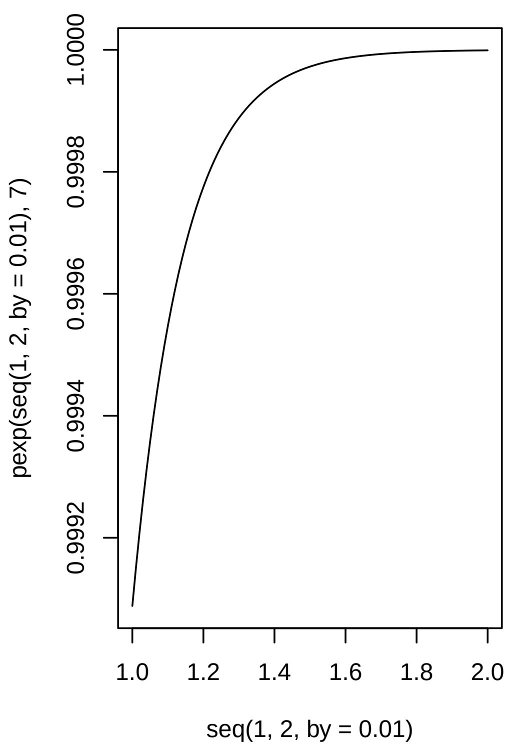
Pour decider une distribution adequate on pourra comparer visualser les fonctions de repartition des trois modeles respectivement de gauche à droite :

- Normal
- Log Normal
- Exponentielle

```
par(mfrow=c(2,2))
plot(seq(1,2,by=0.01),pnorm(seq(1,2,by=0.01),7,2),type="l")
plot(seq(1,2,by=0.01),plnorm(seq(1,2,by=0.01),13,20),type="l")
plot(seq(1,2,by=0.01),pexp(seq(1,2,by=0.01),7),type="l")
```


pnorm(seq(1, 2, by = 0.01), 7, 2)
0.002
0.003
0.004
0.005
0.006
1.0
1.2
1.4
1.6
1.8
2.0
seq(1, 2, by = 0.01)


plnorm(seq(1, 2, by = 0.01), 13, 20)
0.258
0.260
0.262
0.264
0.266
0.268
1.0
1.2
1.4
1.6
1.8
2.0
seq(1, 2, by = 0.01)


pexp(seq(1, 2, by = 0.01), 7)
0.9992
0.9994
0.9996
0.9998
1.0000
1.0
1.2
1.4
1.6
1.8
2.0
seq(1, 2, by = 0.01)

Il est claire que le modele exponentiel est le plus adequate (parmis les trois modeles) en fait : - Le modele normale a une fonction d repartition convexe, ce qui n'est pas le cas pour la fonction de repartition empirique. - Le modele log-normale a une fonction d repartition faiblement concave, ce qui n'est pas le cas pour la fonction de repartition empirique.

MODELE PARAMETRIQUE :

En suite on supposera que la variable : M = montant des sinistre suite loi exponentielle, c-à-d :

$$M \sim (E(\lambda))_{\lambda\ >\ 0}$$

NB : il sera aussi possible d'utiliser les modeles : gamma et weibull (le modele exponentiel n'est qu'une cas particulier de ces deux modeles), mais on préfere d'utilisr le modele exponentiele pour sa simplicité

Estimation de parametre :

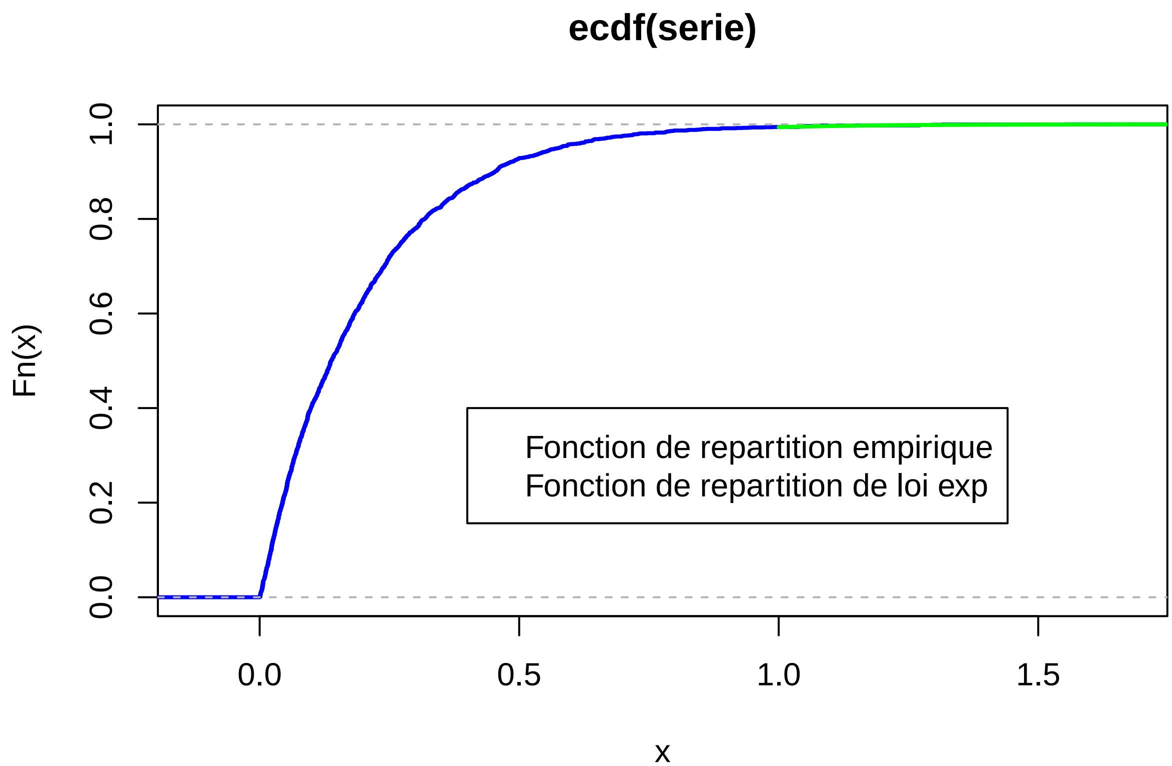
```
library(MASS)
model <- fitdistr(serie , "exponential")
rate <- model$estimate
rate
```

```
##     rate
## 5.116635
```

EVALUATION DE MODELE :

Evaluation visuelle :

```
plot(ecdf(serie),col="blue",lwd=2)
lines(seq(1,2,by=0.01),pexp(seq(1,2,by=0.01),rate=rate),type="l",col="green",lwd=2)
legend(x=0.4,y=0.4,
       legend = c("Fonction de repartition empirique","Fonction de repartition de loi exp") )
```



Cette graphique montre la quasi-egalité entre les deux foncitons de repartitions, ce qui montre que ce choix etait adequate.

Evaluation statistique :

Pour s'assurer on ferra le teste de *Kolmogrov* :

```
ks.test(serie , pexp,rate)
```

```
##
##  Asymptotic one-sample Kolmogorov-Smirnov test
##
## data:  serie
## D = 0.012832, p-value = 0.8433
## alternative hypothesis: two-sided
```

Interpretation :

On obtient un p-value \$ $>$ 5%\$, donc on décide d'accepter l'hypothese H_0 avec un risque de 16%. cela signifie que Le modèle proposée est validé statistiquement.