

Enjeux climatique en assurance : Fondamentaux de l'assurance

Thibault MONNET

2024-02-09

Dauphine 2024

1. Histoire

L'histoire fait très rapidement apparaître que l'entraide et la solidarité, se sont traduits par la création de solutions économiques pour encadrer ces notions.

D'un point de vue philosophique, l'assurance dans sa forme première pourrait être aussi vieille que la société humaine et elle existera, en s'adaptant, jusqu'à sa fin.

Si l'histoire permet de savoir d'où nous venons, elle apporte pour les risques naturels (dont les risques climatiques) un éclairage indéniable. En effet, l'assurance moderne permet de remonter assez facilement jusqu'au années 1980, bien que le monde ait évolué et que les risques assurés également... Pour trouver des éléments plus anciens, quand bien même l'on travaillerait dans une société vieille de 200 ans, cela se complique.

L'histoire c'est avant tout une profondeur de données et des descriptifs permettant de reconstituer le passé pour en tirer des leçons. On peut citer les éruptions volcaniques (le Vésuve 79 AD), les crues de l'Arno à Florence de 1333, 1547 ou encore 1844... L'homme prend des notes de ces catastrophes, car elles le marque et certainement aussi afin d'aider les générations futures. Aujourd'hui, la collecte de ces données est un travail de fourmis, toutefois il est indispensable pour mieux comprendre notre monde.

Avez-vous une idée de la probabilité d'observer 1 évènement bicentennale sur une période de 30 ans ? \ D'en observer au moins 1 sur une période de 200 ans ? Et d'en observer 2 sur une période de 50 ans ?

```
P <- 1/200
tirage <- 30
reussite <- 1
round(pnbinom(tirage-reussite,reussite,P)*100,2)
```

```
## [1] 13.96
```

Je vous invite à lire les travaux de Jérémy DESARTHE [[@Desarthe2014](#)] sur les ouragans et submersions dans les Antilles française entre le XVII^{ème} et le XX^{ème} siècle.

2. Les fondamentaux

Dans cette partie du cours, une première partie théorique consistera à poser quelques définitions nécessaires à la compréhension de l'assurance, différencier l'assurance vie de l'assurance dommages (ou IARD) et faire

le lien entre ces couvertures et les risques climatiques. La seconde étape va consister à définir et manipuler les indicateurs classiques qui sont utilisés dans la profession (IARD).

Contrat, exposition, fréquence sinistre, coût moyen, dossiers avec ou sans suite, prime pure, ratio S/P ou S/C, ratio combiné... Des notions techniques permettant d'analyser les risques assurés et nécessaires à l'élaboration d'une stratégie globale de l'entreprise.

Ces critères sont d'autant plus importants qu'il faudra traduire les pertes engendrées par un événement extrême sur ces notions (et parfois d'autres !)

2.A L'assurance : un contrat

Cela implique un engagement légal et donc une connaissance juridique : Code des Assurances et de la mutualité.

Nous y reviendrons, une partie des tarifs sont réglementés, notamment sur la partie assurance des biens face aux Catastrophes Naturelles.

Les différents types de sociétés et de contrats ayant été introduits dans le cours, la suite s'attardera sur l'assurance IARD, ou assurance de biens et responsabilités.

2.B L'assurance IARD : des indicateurs clés

A noter que les indicateurs qui vont être présentés ici peuvent servir en assurance de personnes, la définition exacte du calcul pouvant évoluer...

Les notions de base **Portefeuille** : photo à un instant t des contrats assurés, au niveau du marché on parlera aussi de *Parc*.

Exposition : cette notion fait référence à la part du portefeuille exposée à un risque. En modélisation climatique, où l'on fait le lien entre un événement à un instant t et la sinistralité, 1 contrat exposé compte pour 1. En tarification on va étudier la "vidéo de l'année" et dans ce cas l'exposition se compte au prorata-temporis, sur 1 an si le contrat est assuré 6 mois, il comptera pour 0.5 car il aura été exposé à 50%.

Sinistre : toute déclaration par un assuré d'un dommage

Sinistre sans suite : il s'agit d'un sinistre n'étant pas couvert par le contrat et n'engendrant pas de dépense pécuniaire par l'assurance

Franchise : définie contractuellement il s'agit du montant restant à la charge de l'assuré en cas de sinistre

coût unitaire : c'est le montant de dédommagement à la charge du porteur de risque

coût moyen : moyenne des coûts unitaires

ratio S/P ou **ratio S/C** : rapport entre la charge des sinistres et les primes (ou cotisations)

Prime Pure : c'est le coût individuel du risque, soit ce que coûte à un assureur, par contrat, le fait d'assurer tout son portefeuille

```
library(data.table)
```

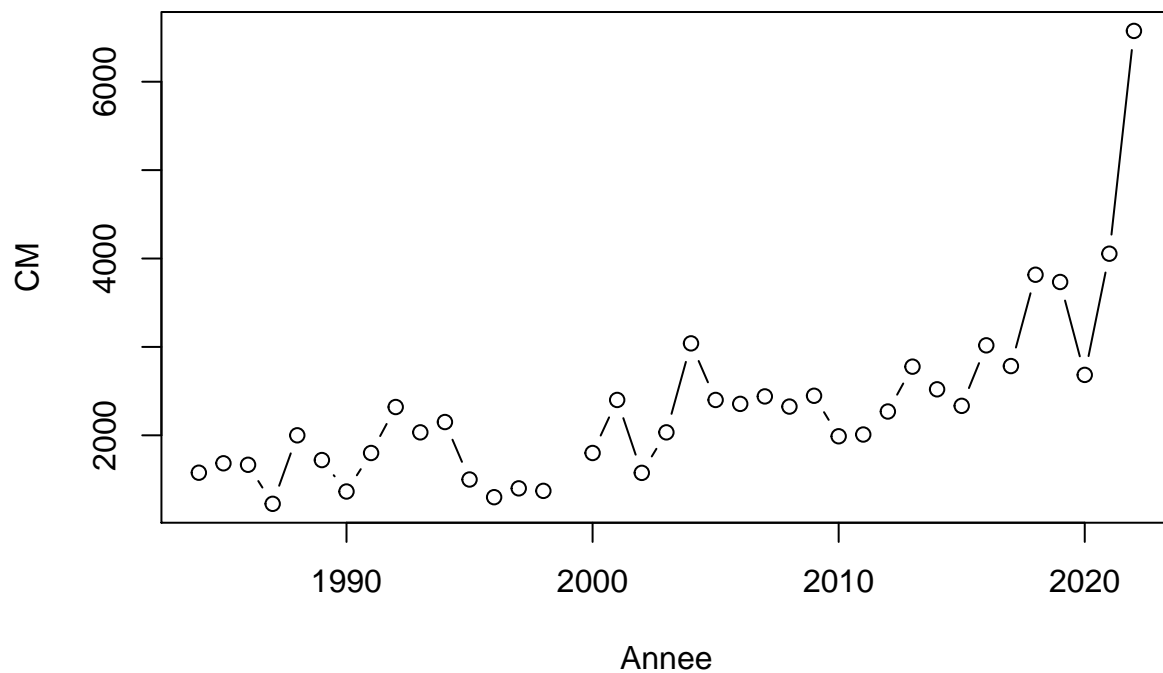
```
## Warning: package 'data.table' was built under R version 4.0.5
```

```
library(readr)

# Chargement des données
DT <- data.table(
  read_delim("TD/TD1-Grele.csv", ";", escape_double = FALSE,
    locale = locale(decimal_mark = ","), trim_ws = TRUE))

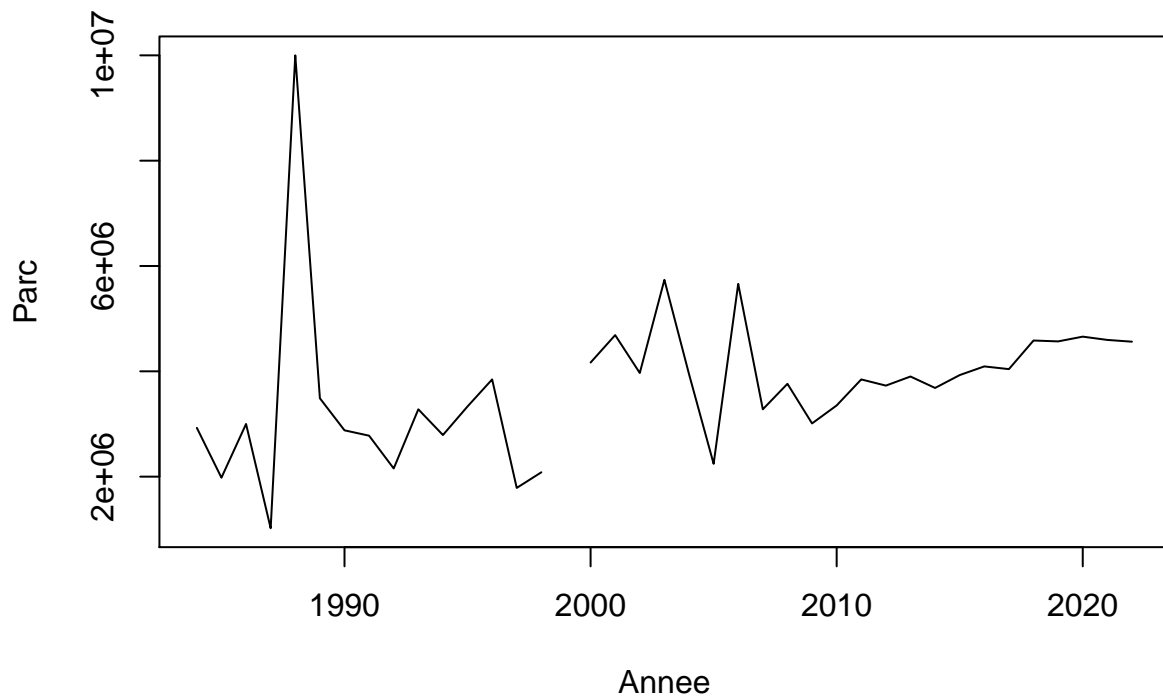
# Calculer les indicateurs manquants
## Cout Moyen : Prime Pure / Fréquence (x1000 car la fréquence donnée pour 1000 contrats)
DT$CM <- DT$PP / DT$Frequence * 1000
plot(data = DT, CM ~ Annee, main = "Evolution du Coût Moyen dans le temps","b")
```

Evolution du Coût Moyen dans le temps



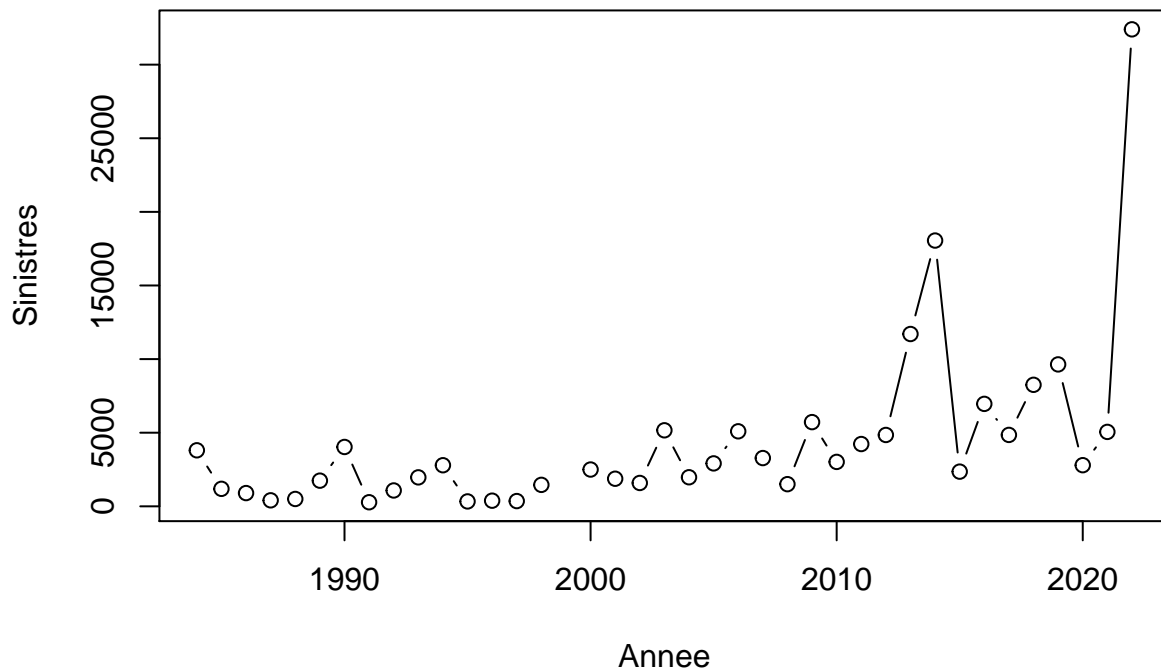
```
## Parc ou garanties souscrites : ici on parle de Dommages aux Biens (DAB) des Particuliers
DT$Parc <- DT$Charge_An_Surv / DT$PP * 100000
plot(data = DT, Parc ~ Annee, main = "Evolution du Parc assuré dans le temps", type = "l")
```

Evolution du Parc assuré dans le temps



```
# NB : ce calcul renvoie des valeurs non exacte du fait des arrondis, notamment en 1988, la fréquence é
## Nombre de sinistre : dépend du parc, donc il y aura un biais
DT$Sinistres <- DT$Parc * DT$Frequence / 1000
plot(data = DT, Sinistres ~ Annee, main = "Estimation du nombre annuel de sinistres grêle", type = "b")
```

Estimation du nombre annuel de sinistres grêle



On notera suite à cet exercice toute la difficulté d'exploiter des données nationales dont on ne dispose pas du meilleur détail.

Cela montre également la limite des données, notamment sur un historique ancien, pourtant nécessaire pour construire un modèle statistique sur le périmètre climatique.

2.C Les risques climatiques

On retiendra qu'il existe un régime spécifique pour les catastrophes naturelles, ce régime fixe légalement les primes, dont une partie est collectée par la CCR (Caisse Centrale de Réassurance) un réassureur de l'état Français. En contrepartie, les assureurs cèdent 50% de leur charge sinistre liée aux évènements de type Cat Nat. Toutefois, si la CCR et son fond (financé par le Marché) venait à faire défaut, l'état interviendrait pour compenser les pertes.

Pour les événements n'entrant pas dans ce régime, notamment les Tempêtes, les assureurs sont libres de rédiger les Conditions Générales de sorte qu'ils peuvent inclure ou exclure certains faits.

3 TD - Modélisation de la Prime Pure annuelle Grêle

```
fit <- lm(data = DT, CM ~ 1+Annee)
summary(fit)
```

```
##
## Call:
## lm(formula = CM ~ 1 + Annee, data = DT)
```

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -799.6 -450.5 -137.3  377.2 3097.2
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.195e+05  2.002e+04  -5.966 7.72e-07 ***
## Annee        6.080e+01  9.996e+00   6.082 5.40e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 701.4 on 36 degrees of freedom
## (1 observation deleted due to missingness)
## Multiple R-squared:  0.5068, Adjusted R-squared:  0.4931
## F-statistic: 36.99 on 1 and 36 DF,  p-value: 5.4e-07
```

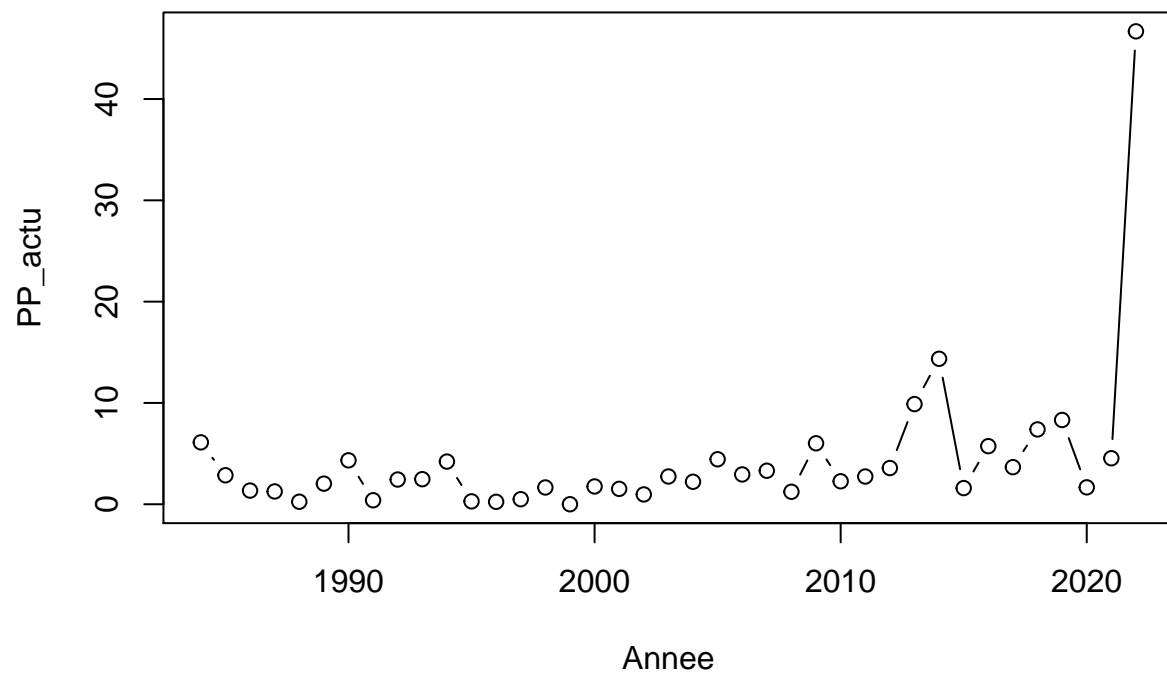
```
# plot(data = DT, CM ~ Annee)
# abline(fit,col = 'blue')
```

```
# Actualisation : idéalement il faudrait comparer l'évolution de la régression à celle des indices des
base100 <- data.table(Annee = 1984:2022, CM_lisse = predict(fit,newdata = data.frame(Annee=1984:2022)))
base100$ref15 <- base100$CM_lisse / base100[Annee==2015]$CM_lisse * 100
```

```
# A défaut on peut s'appuyer sur l'évolution du lissage
base100$euros22 <- base100[Annee==2022]$CM_lisse / base100$CM_lisse
```

```
DT <- merge(DT, base100[,list(Annee,euros22)], by = "Annee", all.x = T)
DT$PP_actu <- DT$PP * DT$euros22
```

```
plot(data=DT,PP_actu ~ Annee, type = "b")
```



A partir de ces données actualisées (malgré un $R\hat{A}^2$ de 0,55 sur l'ajustement) nous pouvons nous arrêter

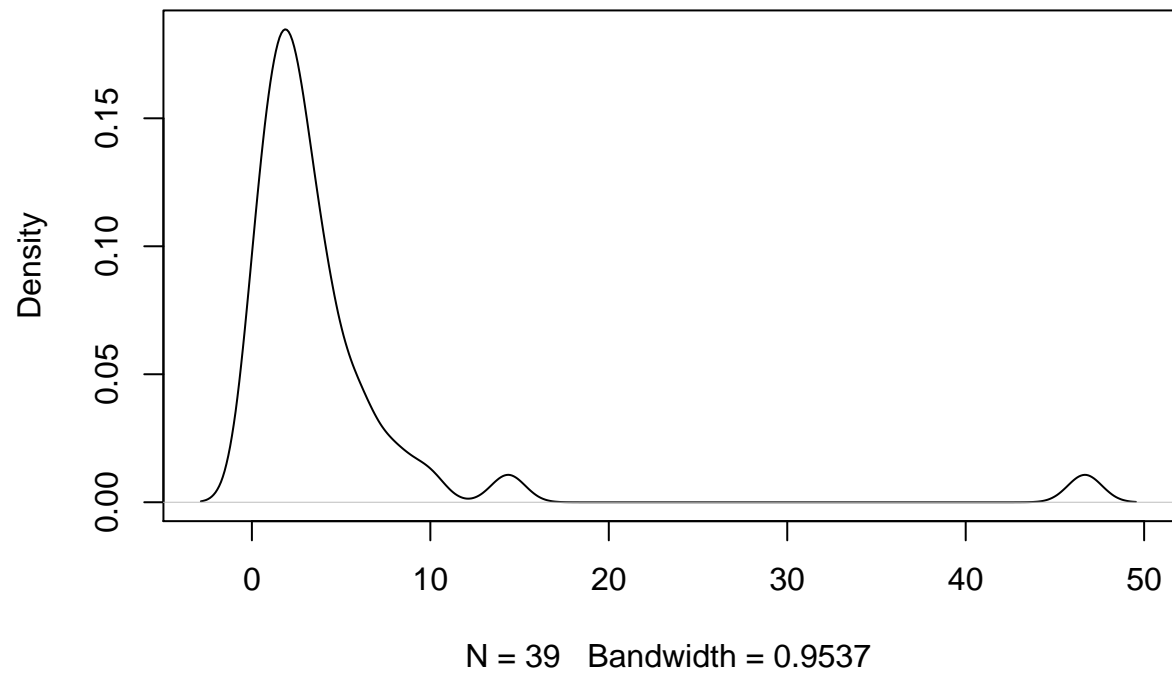
```
plot(density(DT$PP_actu))
```

```
library(fitdistrplus)
```

```
## Warning: package 'fitdistrplus' was built under R version 4.0.5
```

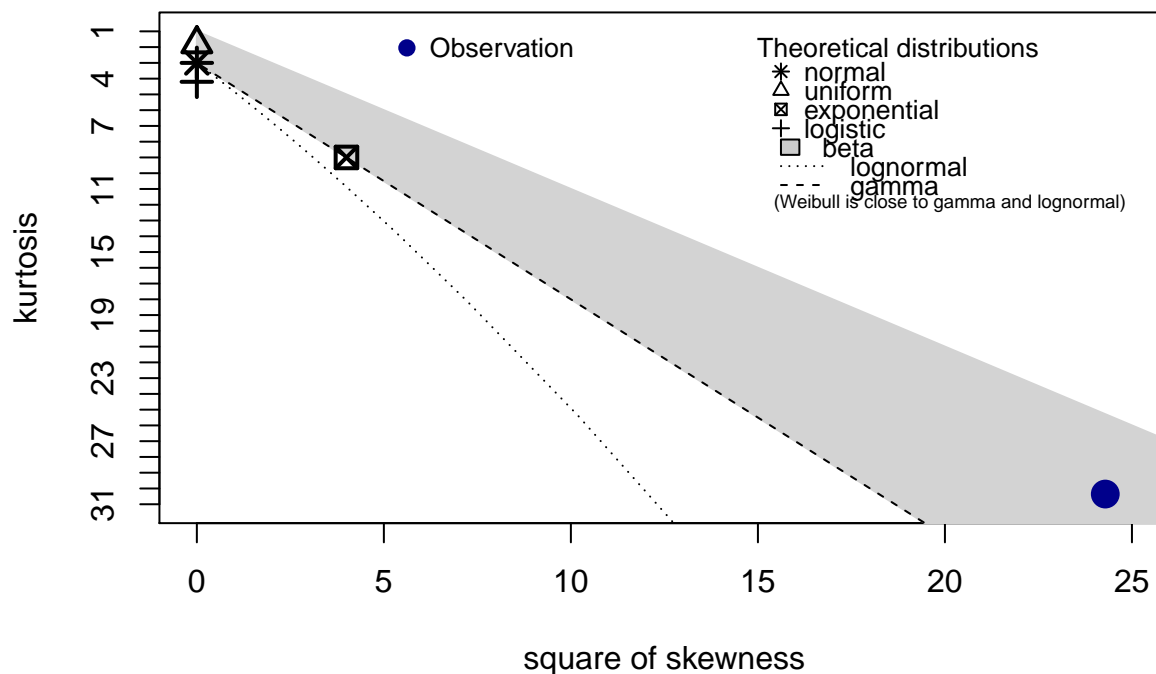
```
## Warning: package 'survival' was built under R version 4.0.5
```

density.default(x = DT\$PP_actu)



```
descdist(DT$PP_actu)
```

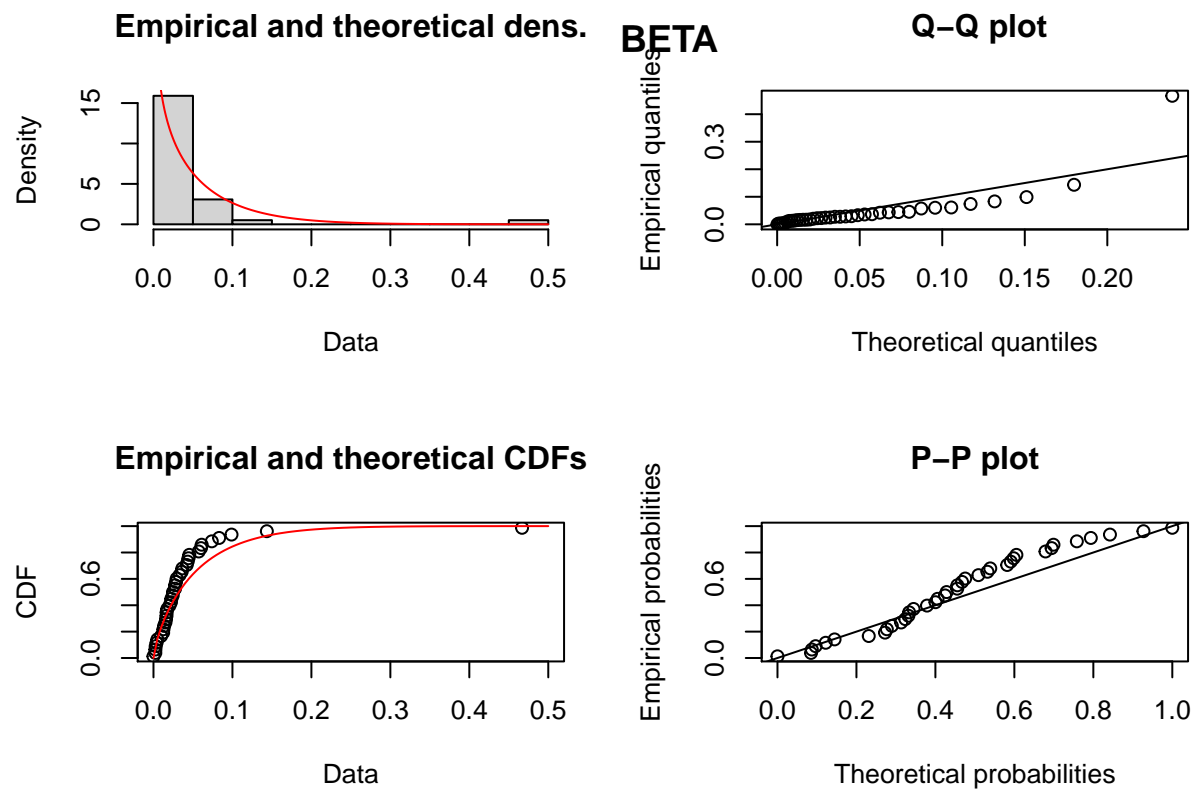

Cullen and Frey graph



```
## summary statistics
## -----
## min: 0    max: 46.68
## median: 2.474746
## mean: 4.35702
## estimated sd: 7.547418
## estimated skewness: 4.928172
## estimated kurtosis: 30.35378
```

Il ressort que la loi à utiliser pour chaque région pourrait-être une loi Beta

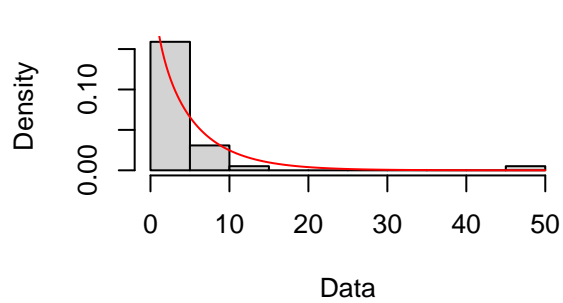
```
# La loi Beta est définie pour x compris entre 0 et 1
# Proposition 1 : définir une Prime Pure maximale puis recentrer ()
# Proposition 2 : modifier l'échelle sur la base des min et max observés
# Fonction de calcul des paramètres d'une loi beta, avec transformation de l'input
# 1 - PP max 100€ et PP min 0,1€
maxi <- max(DT$PP_actu, 100)
mini <- min(DT$PP_actu, 0,1)
xscaled <- (DT$PP_actu-mini)/maxi
fit.beta <- fitdist(xscaled, "beta", method = "mse")
plot(fit.beta); title("BETA")
```



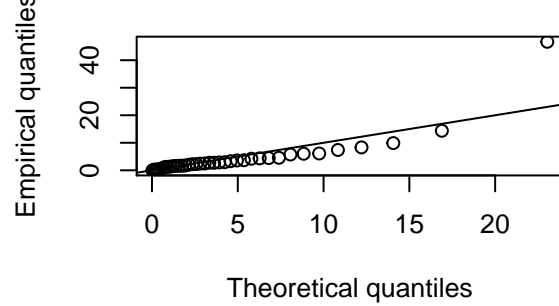
```
# summary(fit.beta)
```

```
fit.gamma <- fitdist(DT$PP_actu, "gamma", method = "mse")
plot(fit.gamma); title("GAMMA")
```

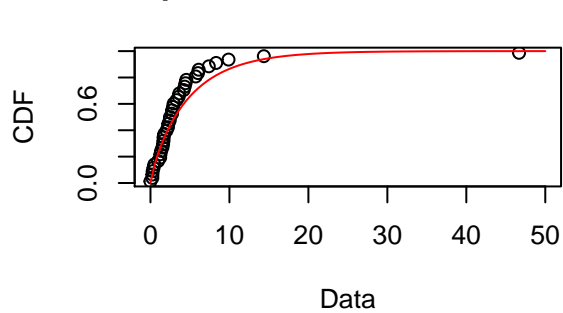
Empirical and theoretical dens. GAMMA



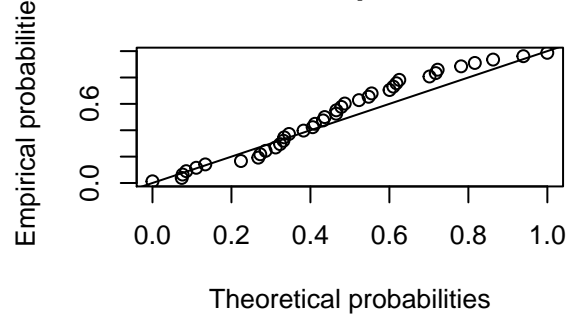
Q-Q plot



Empirical and theoretical CDFs



P-P plot

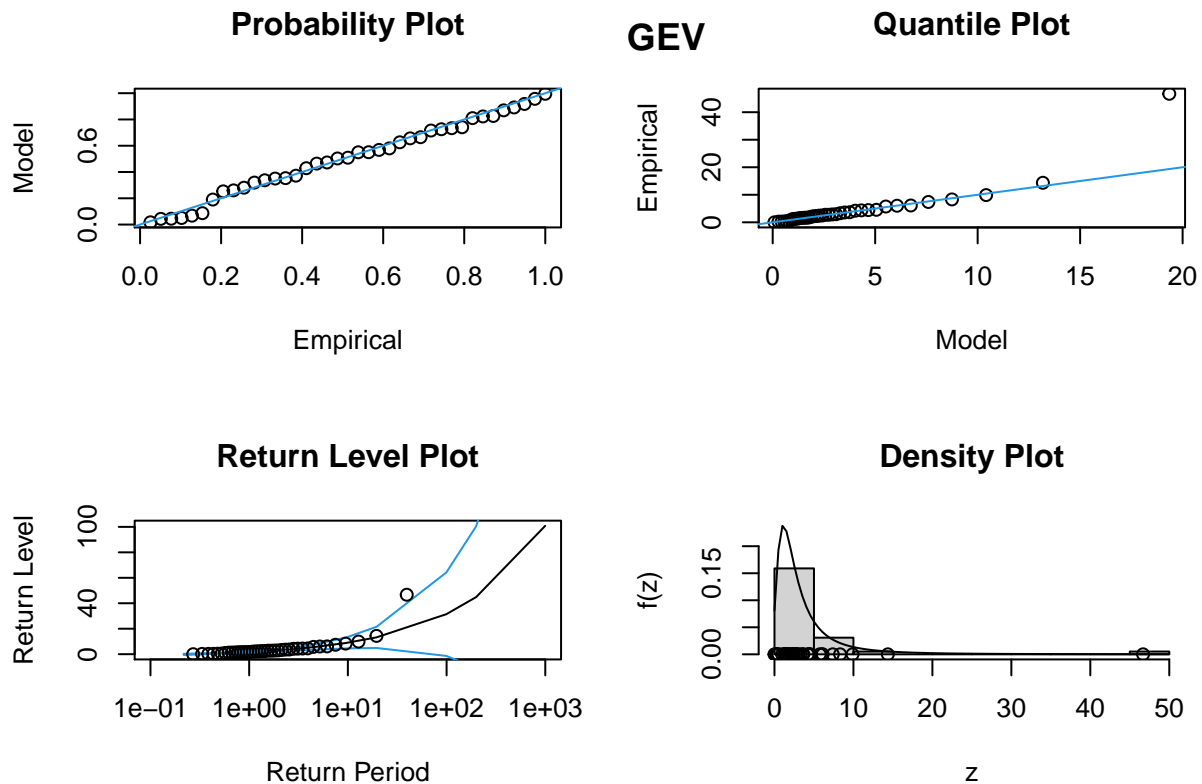


```
# summary(fit.gamma)
```

```
fit.gev <- ismev::gev.fit(DT$PP_actu)
```

```
## $conv
## [1] 0
##
## $nllh
## [1] 93.3314
##
## $mle
## [1] 1.7356760 1.7089241 0.4893811
##
## $se
## [1] 0.3184964 0.3005547 0.1695616
```

```
ismeov::gev.diag(fit.gev); title("GEV")
```



```
# summary(fit.gev)
```

Graphiquement la méthode qui semble être la meilleure est l'ajustement par loi GEV.

Toutefois les bibliothèques utilisées pour ajuster les paramètres de ces lois ne permettent pas d'obtenir d'indicateurs comparable.

Il existe d'autres bibliothèques pour la famille des GEV, l'ajustement des paramètres obtenu est similaire, toutefois la comparaison reste délicate.

Etant donné la matière (les risques climatiques) et la volatilité forte de ce domaine, nous privilégierons une approche GEV.

La suite du TD ci après permet d'obtenir une idée de la période de retour de l'exercice 2022, il permet aussi d'avoir une idée, selon chaque modèle, de la prime pure probable à une période de retour de 200 ans et de 500 ans.

Enfin, pour faire le lien avec l'exercice initial du cours, ce dernier bloc propose d'évaluer la probabilité d'observer sur 39 ans (1984 - 2022) une année aussi extrême que 2022, selon la période de retour estimée dans le TD.

Rappel : le lien entre période de retour et probabilité est : $P(X \geq x) = 1 - 1/PdR$

Ainsi on cherche, à partir des paramètres ajustés de nos lois, le quantile pour une probabilité de 0.

```
param <- data.table(rbind(
  data.frame(law = "beta", PP_200ans = round(maxi * qbeta(p = 0.995, fit.beta$estimate[1], fit.beta$estimate[2]) + mini,
    PP_500ans = round(maxi * qbeta(p = 0.998, fit.beta$estimate[1], fit.beta$estimate[2]) + mini,
    PdR_2022 = round(1 / (1 - pbeta(q = (DT$PP_actu[DT$Annee == "2022"] - mini) / maxi, fit.beta$estimate[1], fit.beta$estimate[2]),
  data.frame(law = "gamma", PP_200ans = round(qgamma(p = 0.995, fit.gamma$estimate[1], fit.gamma$estimate[2]) + mini,
```

```

PP_500ans = round(qgamma(p = 0.998, fit.gamma$estimate[1], fit.gamma$estimate[2]), digits = 0)
PdR_2022 = round(1 / (1 - pgamma(q = DT$PP_actu[DT$Annee == "2022"], fit.gamma$estimate[1],
data.frame(law = "GEV", PP_200ans = fExtremes::qgev(0.995, mu = fit.gev$vals[1,1], beta = fit.gev$vals[1,2], xi = 1)
PP_500ans = fExtremes::qgev(0.998, mu = fit.gev$vals[1,1], beta = fit.gev$vals[1,2], xi = 1)
PdR_2022 = round(1 / (1 - fExtremes::pgev(DT$PP_actu[DT$Annee == "2022"], mu = fit.gev$vals[1,1], beta = fit.gev$vals[1,2], xi = fit.gev$vals[1,3])[1]), digits = 0)))

param

```

```

##      law PP_200ans PP_500ans PdR_2022
## 1:  beta  28.75000  33.17000   12253
## 2:  gamma 28.34000  33.53000    5037
## 3:   GEV 44.86918  71.30502    216

```

```

P <- 1/param$PdR_2022
tirage <- 2022-1984+1
reussite <- 1
round(pnbinom(tirage-reussite,reussite,P)*100,1)

```

```
## [1] 0.3 0.8 16.6
```

On notera l'écart entre les lois "usuelles" et la théorie des valeurs extrêmes, particulièrement visible sur les PP observable tous les 200 ou 500 ans. L'approche par loi Beta ou loi Gamma reste très proche, 40 centimes d'écarts sur un coût du risque à 200 ans est assez faible.

Maintenant, l'approche par les GEV montre un tout autre résultat, avec un ratio de 1.6 environ à 200 ans et supérieur à 2.1 à 500 ans... cela doit être bien maîtrisé pour une intégration dans le tarif.

Si l'on s'attarde sur les Périodes de Retour calculées pour la PP extrême de 2022, l'écart entre Beta et Gamma et bien plus important (du simple au double) et là encore, la théorie des valeurs extrêmes nous donne à réfléchir... 216 ans contre 5000 ans (Gamma) et 12000 (Beta) on parle de facteur x20 / x50...

Enfin, sous hypothèse de stabilité et que nos données dans le temps soit bien iid (ou encore toutes choses égales par ailleurs) nous pouvons estimer que la probabilité de voir une PP aussi forte sur les 39 dernières années serait de 16.6% avec l'approche par la théorie des valeurs extrêmes, avec les autres lois, cette proba devient 0.3% ou 0.8%, soit très peu probable et pourtant observé !

A ce stade, la théorie des valeurs extrêmes nous permet de dire que les orages de grêles de 2022 restent du domaine du probable, contrairement à une approche plus classique. Toutefois, un faisceau d'indices (hausse de la PP depuis 2013, multiplication des événements "atypiques", fortes chaleurs 2022 propices aux orages...) laisse penser que le changement climatique pourrait contribuer à une amplification des phénomènes de grêle extrême...

A ce jour, quelques études émergent et vont dans ce sens. Les réalisateurs de ces études étant parfois partis prenante dans la tarification des risques (réassurance), il faut être prudent dans la lecture de leurs conclusions. Il semble cependant évident que nombre de scientifiques vont s'atteler à ces recherches, ainsi d'ici très peu de temps, nous pourrions associer, ou non, un impact du changement climatique sur ce phénomène.

NB : si d'ici la fin de la décennie nous observons à nouveau des pertes aussi fortes, l'approche de la théorie des valeurs extrêmes croisée avec l'approche par la loi binomiale négative (donc d'observer 2 succès sur la période) renvoie une probabilité légèrement en deça de 2%... Cela entre dans le domaine du rare et remettrait en cause, au moins partiellement, cette étude et ce modèle.