

# Analyse et modélisation des pertes extrêmes

## approches et applications

Jérémie Barde

École d'actuariat  
Université Laval, Québec, Canada

8 août 2024



UNIVERSITÉ  
LAVAL

Faculté des  
sciences et de génie  
École d'actuariat

CIMMUL

Quantact



CIRCEB

# Agenda

---

- 1 Introduction
- 2 Mise en contexte
- 3 Modélisation de la sévérité
- 4 Mesure de risque
- 5 Conclusion

# Introduction

1 Introduction

2 Mise en contexte

3 Modélisation de la sévérité

4 Mesure de risque

5 Conclusion

# Introduction

## Objectifs

---

Le présent travail est fait dans le cadre du cours de Lecture dirigées, ACT-7014, sous la supervision du professeur Etienne Marceau.

Objectifs du travail :

- Reproduire l'article [Brazauskas and Kleefeld \(2016\)](#), « *Modeling Severity and Measuring Tail Risk of Norwegian Fire Claims* ».
- Reproduire l'article [Upretee and Brazauskas \(2022\)](#), « *Computing and Estimating Distortion Risk Measures : How to Handle Analytically Intractable Cases ?* ».
- Appliquer les techniques présentées à la base de données canadienne de périls incendie et à la National Fire Information Database (NFID).

# Mise en contexte

1 Introduction

**2** Mise en contexte

3 Modélisation de la sévérité

4 Mesure de risque

5 Conclusion

# Loi Pareto

## Loi Pareto Lomax

La loi Pareto Lomax est essentiellement une loi Pareto de type I translatée pour que le support commence à 0. Si on a  $X \sim \text{PaI}(x_0 = \lambda, \alpha)$ , alors  $X - x_0 \sim \text{PaLomax}(\alpha, \lambda)$ .

La fonction de répartition est donnée par

$$F_X(x) = 1 - \left( \frac{\lambda}{\lambda + x} \right)^\alpha, \quad x > 0.$$

Caractéristiques de la loi Pareto Lomax :

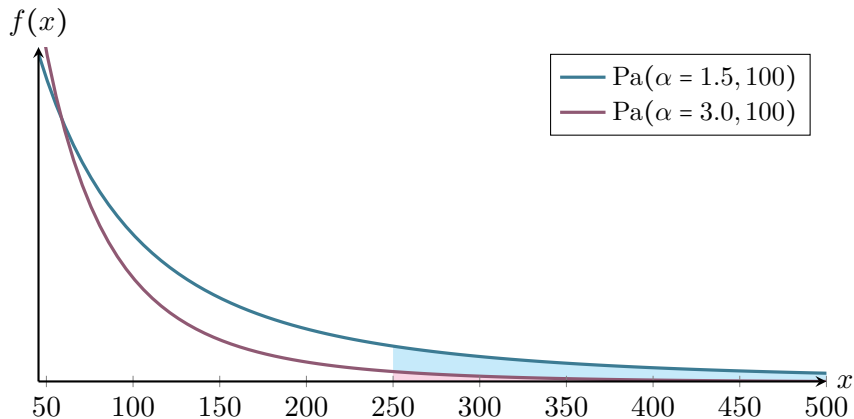
$$E[X] = \frac{\lambda}{\alpha - 1}, \quad \alpha > 1, \quad \text{Var}(X) = \frac{\alpha \lambda^2}{(\alpha - 1)^2 (\alpha - 2)}, \quad \alpha > 2.$$

**Remarque :** Si  $\alpha < 2$  la variance n'existe pas et si  $\alpha < 1$  l'espérance aussi n'existe pas.

# Loi Pareto

## Paramètres alpha

Le paramètre de forme  $\alpha$  donne l'indication du niveau de dangerosité de la distribution.



# Loi Pareto

## Paramètres alpha

Le tableau ci-dessous montre la  $Pr(X > 400)$  selon le paramètre  $\alpha$  avec  $\lambda = 100$

$\alpha$	$Pr(X > 400)$
3.0	0.016
2.0	0.063
1.5	0.125
0.7	0.379

**Remarque** : lorsque  $\alpha$  passe de 3 à 1.5 il y a 8 fois plus de chance d'avoir un sinistre supérieur à 400 et lorsqu'il passe de 3 à 0.7 il y a 24 fois plus de chance.



# Loi Pareto généralisée

Pour obtenir la loi Pareto généralisée, on applique les transformations suivantes à la loi Pareto Lomax :

- Le paramètre  $\alpha = 1/\xi$ .
- Le paramètre  $\lambda = \beta/\xi$ .
- Ajout du paramètre de localisation  $\theta$ .

La fonction de répartition est donnée par

$$G_{\xi,\beta,\theta}(x) = 1 - \left(1 + \frac{\xi}{\beta}(x - \theta)\right)^{-\frac{1}{\xi}}, \quad \xi > 0, x > \theta,$$

où  $\xi$  est le paramètre de forme et  $\beta$  le paramètre d'échelle.

Caractéristiques de la loi Pareto généralisée :

$$E[X] = \frac{\beta}{1-\xi} + \theta, \quad \xi < 1 \quad \text{Var}(X) = \frac{\beta^2}{(1-\xi)^2(1-2\xi)}, \quad \xi < \frac{1}{2}$$

# Notions sur la théorie des valeurs extrêmes

## Motivations

---

La modélisation des valeurs extrêmes est une composante essentielle du domaine de l'actuariat.

Motivations de la modélisation des valeurs extrêmes :

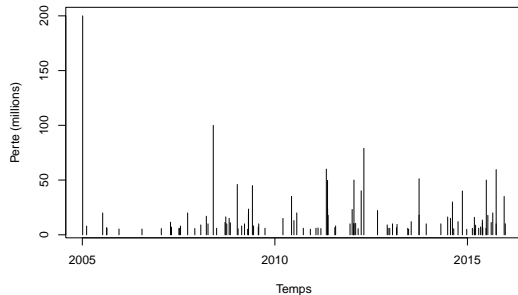
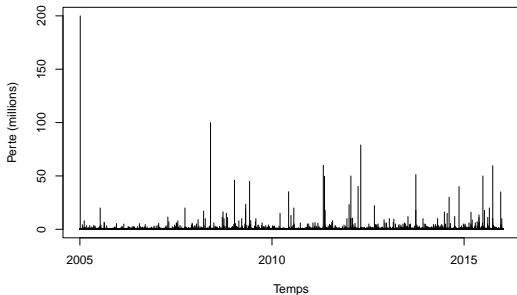
- Impacts significatifs sur les résultats financiers
- Impact significatif sur les réserves des compagnies d'assurance.
- Impacts significatifs sur la prise de décision en matière de gestion des risques.

**Exemple** : le risque d'incendie est réputé pour la possibilité de pertes extrêmes.

# Notions sur la théorie des valeurs extrêmes

## Données extrêmes : incendie

Pertes estimées selon la date du sinistre :



# Notions sur la théorie des valeurs extrêmes

## Distribution généralisée des valeurs extrêmes

La théorie des valeurs extrêmes, [Pickands III \(1975\)](#), est un outil mathématique essentiel pour la modélisation des événements extrêmes.

On s'intéressera à la distribution généralisée de la valeur extrême (GEV), qui est justement utilisée pour décrire le comportement du maximum d'un échantillon de v.a. i.i.d.

La fonction de répartition de cette distribution est donnée par

$$H_{\xi, \sigma, \mu} = \begin{cases} \exp\left\{-\left(\frac{1}{1+\frac{\xi}{\beta}(x-\mu)}\right)^{\frac{1}{\xi}}\right\} & , \xi < 0, 0 < x < -\frac{\beta}{\xi} \quad (\text{Weibull}) \\ \exp\left\{-\exp\left\{-\frac{x-\mu}{\beta}\right\}\right\} & , \xi = 0, x > 0 \quad (\text{Gumbel}) \\ \exp\left\{-\left(\frac{1}{1+\frac{\xi}{\beta}(x-\mu)}\right)^{\frac{1}{\xi}}\right\} & , \xi > 0, x > 0 \quad (\text{Fréchet}). \end{cases}$$

# Théorie des valeurs extrêmes

## Mise en contexte

---

Essentiellement, on considère une suite de v.a. i.i.d  $X_1, X_2, \dots$ . On définit le maximum des  $n$  premières v.a. par

$$M_n = \max\{X_1, \dots, X_n\}.$$

Soit  $x_F = \sup\{x; F(x) < 1\}$ , le point extrême droite de  $F$ . Il est possible de montrer que

$$\lim_{n \rightarrow \infty} M_n = x_F.$$

Pour décrire le comportement asymptotique de la v.a.  $M_n$  on suppose qu'il existe des entiers normalisant  $a_n > 0$  et  $b_n$  tels que

$$\Pr\left(\frac{M_n - b_n}{a_n} \leq x\right) = F(a_n x + b_n)^n \rightarrow H(x), \text{ quand } n \rightarrow \infty$$

où  $H(x)$  est une fonction de répartition non dégénérée.

# Théorie des valeurs extrêmes

## Théorème de Fisher-Tippett

---

Lorsque cette condition est respectée, alors on dit que  $F_X$  est un élément du domaine d'attraction du maximum de la distribution  $H$ , qu'on note

$$F_X \in DAM(H).$$

Le théorème de Fisher-Tippett, [Fisher and Tippett \(1928\)](#), indique que la seule fonction de répartition non dégénérée de  $H$  est la fonction de répartition de  $H_\xi$  de la distribution GEV.

# Théorie des valeurs extrêmes

## Lien entre la loi Pareto généralisée et la distribution généralisée des valeurs extrêmes

Lien entre la loi Pareto généralisée et la distribution GEV :

- Théorème de Pickands-Balkema-de Haan
- Distribution de l'excédent

### Définition (Distribution de l'excédant)

Soit la v.a.  $X$  et la v.a.  $W_\theta$ , qui est défini comme

$$W_\theta = (X - \theta | X > \theta),$$

où  $W_\theta$  représente l'excédent d'un sinistre par rapport à un seuil  $\theta$ .

**Objectif** : Modéliser les pertes en excès d'un seuil  $\theta$ .

# Théorie des valeurs extrêmes

## Lien entre la loi Pareto généralisée et la distribution généralisée des valeurs extrêmes

### Théorème (Pickands–Belkema–de Haan)

Pour tout  $\xi \in \mathbb{R}$ ,  $F \in MDA(H_\xi)$  si et seulement si

$$\lim_{\theta \rightarrow \infty} \sup_{x \geq 0} |F_{W_\theta}(x) - GP_{\xi, \beta(\theta)}(x)| = 0$$

pour une certaine fonction  $\beta$  positive et mesurable.

En d'autre mot,  $H_\xi$  gouverne le comportement de  $M_n$  si et seulement si  $G_\xi$  gouverne le comportement de  $W_\theta$ .

Bases théoriques suffisantes pour dire que pour un seuil  $\theta$  assez élevé on peut approximer la distribution de l'excédent par une loi Pareto généralisée.



# Modélisation de la sévérité

1 Introduction

2 Mise en contexte

**3 Modélisation de la sévérité**

4 Mesure de risque

5 Conclusion

# Reproduction de Brazauskas & Kleefeld (2016)

## Objectifs des auteurs

---

Les auteurs de [Brazauskas and Kleefeld \(2016\)](#) modélisent les pertes de la base de données *Norwegian fire claims* avec les lois suivantes :

- GPD : loi Pareto généralisée
- LNPa2 : loi lognormale-Pareto type I
- WePa2 : loi Weibull-Pareto type I
- LNPa3 : loi lognormale-Pareto généralisée
- WePa3 : loi Weibull-Pareto généralisée
- FT : loi *Folded-t*

Ensuite, les auteurs regardent pour chacun de ces modèles les éléments suivants :

- Test graphique : graphiques quantiles-quantiles.
- Tests quantitatifs : Kolmogorov-Smirnov, Anderson-Darling et les critères AIC et BIC.
- Les mesures *VaR* et *Tail-conditional median* (TCM).
- Le pouvoir prédictif des modèles.

# Reproduction de Brazauskas & Kleefeld (2016)

## Raccordement de lois

Motivations de l'utilisation d'un raccordement de loi :

- Ne pas utiliser la portion empirique pour les données sous le seuil.
- Avoir une meilleure adéquation à celle que l'on obtient avec une seule loi.

### Définition (Raccordement de $n$ lois)

Soit  $X$  une variable qui suit une loi composite et soient  $\{X_j, j = 1, \dots, n\}$  des variables non identiquement distribuées. Alors,

$$f_X(x) = \begin{cases} w_1 \frac{f_{X_1}(x)}{F_{X_1}(\theta_1) - F_{X_1}(\theta_0)} & , \theta_0 < x \leq \theta_1 \\ \dots & \\ w_n \frac{f_{X_n}(x)}{F_{X_n}(\theta_n) - F_{X_n}(\theta_{n-1})} & , \theta_{n-1} < x \leq \theta_n \end{cases}$$

où  $w_i > 0$  et  $\sum_{i=1}^n w_i = 1$ .

# Reproduction de Brazauskas & Kleefeld (2016)

## Raccordement de lois : continuité et dérivabilité

Comme montré par [Brazauskas and Kleefeld \(2016\)](#) il est possible d'assurer la continuité et la dérivabilité aux points  $\theta_i$ .

Pour une combinaison de deux lois il faut résoudre le système d'équations suivant

$$\begin{cases} w \frac{f_{X_1}(\theta)}{F_{X_1}(\theta)} = (1-w) \frac{f_{X_2}(\theta)}{F_{X_2}(\theta)} & \text{(continuité)} \\ w \frac{f'_{X_1}(\theta)}{F_{X_1}(\theta)} = (1-w) \frac{f'_{X_2}(\theta)}{F_{X_2}(\theta)} & \text{(dérivabilité)} \end{cases} .$$

Cela a aussi comme effet de réduire le nombre de paramètres à estimer.

# Reproduction de Brazauskas & Kleefeld (2016)

## Raccordement de lois : exemple

Voici un exemple avec le modèle LNPa2, on a  $X_1$  qui est une loi lognormale et  $X_2$  qui est une loi Pareto I.

La fonction densité de  $X$  est

$$f_X(x) = \begin{cases} w \frac{f_{X_1}(x)}{F_{X_1}(\theta)}, & 0 < x \leq \theta \\ (1-w)f_{X_2}(x), & x > \theta. \end{cases}$$

Pour tenir compte de la continuité et de la dérivabilité on définit les paramètres  $\mu$  et  $w$  comme

$$\mu = \ln(\theta) - \alpha\sigma^2, \quad w = \left( \frac{\theta}{\alpha} \frac{f_{X_1}(\theta)}{F_{X_1}(\theta)} + 1 \right)^{-1}.$$

# Données de la *National Fire Information Database*

## Présentation des données

---

Description des données canadiennes de périls incendie :

- 467 929 cas d'incendies.
- 136 variables explicatives.
- 2009-2015.

Ajustements faits aux données :

- retraits de 360 181 données indisponibles de la variable *pertes estimées*.
- retraits de 5112 données, dont les pertes estimées, sont de 0.

On présentera les résultats obtenus pour les années 2011-2015

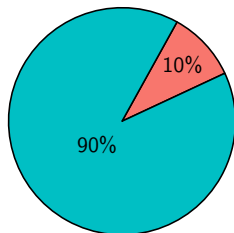
# Données de la *National Fire Information Database*

## Statistiques descriptives

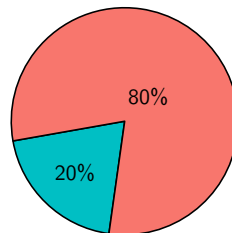
Statistiques descriptives et contribution pour les  $n = 102\,635$  pertes estimées.

Min.	1 <sup>er</sup> Qu.	Médiane	Moyenne	Variance	3 <sup>er</sup> Qu.	Max.
1	1000	5000	75 800	$9.73 \times 10^{11}$	30 000	$2 \times 10^8$

Portion du nombre de pertes  
élevées sur le nombre total de pertes



Portion du montant total des pertes  
élevées sur le montant total de pertes



# Données de la *National Fire Information Database*

## Fréquence relatives des montants de pertes

Fréquences des montants de perte selon l'année.

Intervalles	Nombre de données				
	2011	2012	2013	2014	2015
$[1, 50\,000[$	7581	8083	7531	7234	7428
$[5 \times 10^4, 5 \times 10^5[$	2022	1902	1778	1811	1752
$[5 \times 10^5, 1 \times 10^6[$	388	159	174	145	166
$[1 \times 10^6, 1 \times 10^7[$	107	98	89	83	89
$[1 \times 10^7, 5 \times 10^7[$	4	6	4	6	9
$[5 \times 10^7, \infty[$	1	2	1	0	2



# Données de la *National Fire Information Database*

## Modèles testés

---

Pour la modélisation des montants de perte, on teste les modèles suivants :

- LN-PaG : loi lognomrale-Pareto généralisée.
- GB2-PaG : loi bêta généralisée de type 2-Pareto généralisée.

Pour présenter les résultats on applique les transformations suivantes :  $\alpha = \frac{1}{\xi}$  et  $\lambda = \frac{\beta}{\xi}$

# Données de la *National Fire Information Database*

## Loi bêta généralisée de type 2 : définition et caractéristiques

---

Loi bêta généralisée de type 2 (GB2) de paramètres  $(a, b, p, q)$ . La fonction de densité est

$$f_X(x) = \frac{ax^{ap-1}}{b^{ap}I(p,q)(1 + (\frac{x}{b})^a)^{p+q}}, \quad x > 0.$$

où les paramètres de position  $b$ , d'échelle  $a$  et de forme  $p$  et  $q$  ont comme support  $\mathbb{R}^+$ .

Il est possible d'exprimer la fonction de répartition comme

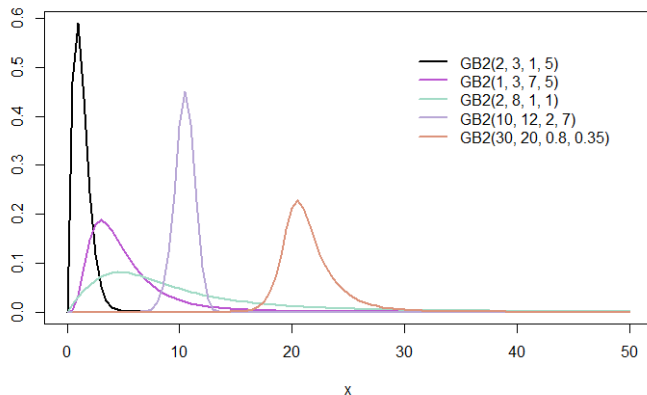
$$F_X(x) = B\left(\frac{(x/d)^a}{1 + (x/d)^a}, p, q\right).$$

où  $B$  est la fonction de répartition d'une loi bêta.

# Données de la *National Fire Information Database*

## Loi bêta généralisée de type 2 : définition et caractéristiques

**Motivation** : Avec ses 4 paramètres, la loi GB2 a l'avantage d'être très flexible.



# Données de la *National Fire Information Database*

## Modélisation des montants de perte

On estime les paramètres en maximisant la fonction de log-vraisemblance.

	2011	2012	2013	2014	2015
LN-PaG					
$\hat{\alpha}$	0.50	0.89	0.80	0.87	0.92
$\hat{\theta}$	399 957	399 841	399 975	399 974	377 058
GB2-PaG					
$\hat{\alpha}$	0.50	0.89	0.80	0.87	0.91
$\hat{\theta}$	400 000	399 848	396 500	399 996	370 000

**Remarques :** Dans [Barde \(2022\)](#) l'auteur détermine un seuil à 350 000 pour ces données.

# Données de la *National Fire Information Database*

## Validation des modèles : graphiques quantiles-quantiles

Pour tester la qualité de l'ajustement du modèle, on commence par un test d'adéquation graphique.

Test choisi : graphique quantiles-quantiles.

Les graphiques quantiles-quantiles des modèles LN-PaG et GB2-PaG sont obtenus avec la définition suivante :

### Définition (Graphique quantiles-quantiles)

Si on a une variable aléatoire  $X$ , alors le graphique quantiles-quantiles est défini comme

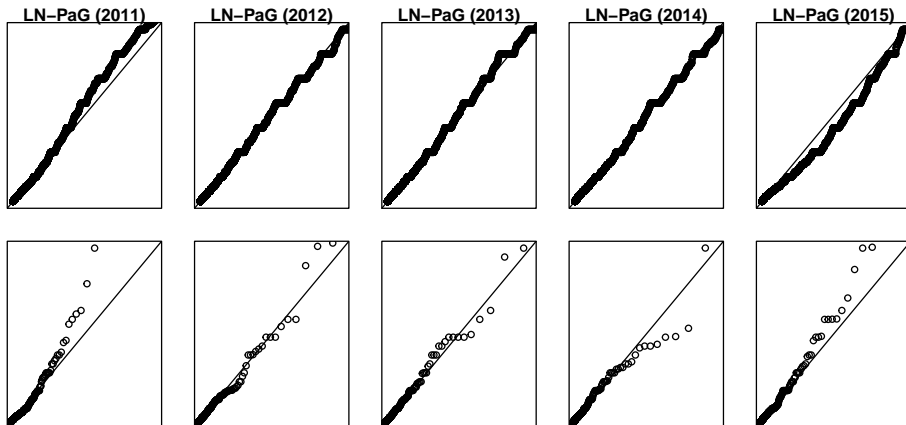
$$\left( F_X^{-1} \left( \frac{i}{n+1} \right), x_{[i]} \right), \quad i = 1, 2, \dots, n,$$

où les  $x_{[i]}$  sont les statistiques d'ordres et  $n$  le nombre d'observations.

# Données de la *National Fire Information Database*

## Validation des modèles : graphiques quantiles-quantiles

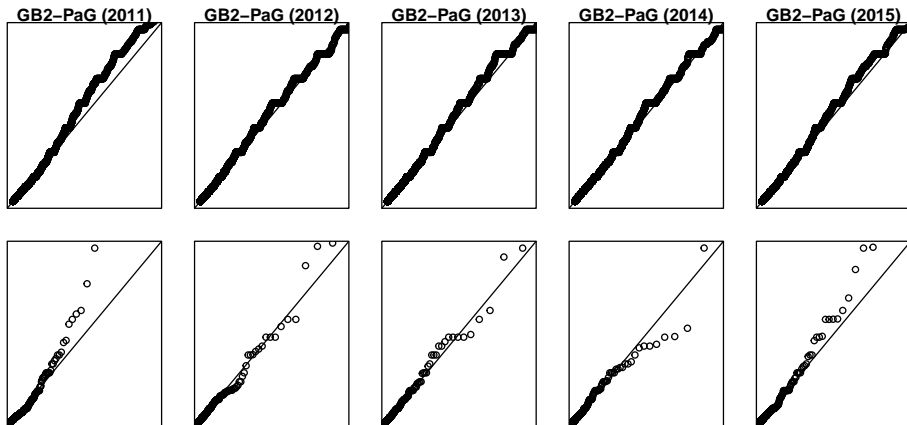
Pour le modèle LN-PaG on obtient les graphiques quantiles-quantiles suivants :



# Données *National Fire Information Database*

## Validation des modèles : graphiques quantiles-quantiles

Pour le modèle GB2-PaG on obtient les graphiques quantiles-quantiles suivants :



# Données de la *National Fire Information Database*

## Validation des modèles : Statistique de Kolmogorov-Smirnov et Anderson-Darling

On fait aussi les tests de Kolmogorov-Smirnov ( $D$ ) et Anderson-Darling ( $A^2$ ) :

	2011	2012	2013	2014	2015
Statistique $D$					
LN-PaG	0.03	0.04	0.04	0.04	0.07
GB2-PaG	0.03	0.04	0.03	0.04	0.04
Statistique $A^2$					
LN-PaG	7.88	8.40	7.24	7.94	57.78
GB2-PaG	7.80	7.69	6.04	6.09	12.07

**Remarque** : le modèle GB2-PaG selon ces tests est meilleur.



# Données de la *National Fire Information Database*

## Validation des modèles : Critère d'information AIC et BIC

On termine les tests d'adéquations avec les critères AIC et BIC :

	2011	2012	2013	2014	2015
Akaike Information Criterion, AIC					
LN-PaG	229719	227048	212612	207036	207296
GB2-PaG	229724	227028	212595	207006	206896
Bayesian Information Criterion, BIC					
LN-PaG	229720	227049	212613	207037	207297
GB2-PaG	229725	227029	212596	207007	206897

**Remarque :** Le modèle GB2-PaG est meilleur selon ces critères sauf pour l'année 2011.

# Données de la *National Fire Information Database*

## Mesure de risques

---

La mesure  $VaR$  d'une v.a  $B$  est définie comme,

$$VaR_{\kappa}(B) = \inf\{x \in \mathbb{R} : F_B(x) \geq \kappa\}, \kappa \in (0,1).$$

Cependant, elle ne donne pas d'information sur des valeurs supérieures à  $VaR_{\kappa}(B)$ .

La mesure *Tail-conditional median* d'une variable aléatoire  $B$  est définie comme,

$$TCM_{\kappa}(B) = \text{median}[B|B > VaR_{\kappa}(B)] = VaR_{\frac{1+\kappa}{2}}(B).$$

Voir [So and Wong \(2012\)](#). Il s'agit simplement de la médiane des montants de pertes au-dessus de la mesure  $VaR$ .

# Données de la *National Fire Information Database*

## Mesure de risques

On calcule les valeurs pour ces deux mesures :

Modèles	Mesure	2011	2012	2013	2014	2015
Empirique	$Var_{0.9}(X)$	300 000	150 000	150 000	150 000	150 000
	$TCM_{0.9}(X)$	494 631	300 000	308 616	300 000	300 450
LN-PaG	$Var_{0.9}(X)$	255 829	143 172	148 252	154 937	182 285
	$TCM_{0.9}(X)$	508 444	300 120	313 850	306 233	327 310
GB2-PaG	$Var_{0.9}(X)$	252 683	135 606	141 059	145 117	137 745
	$TCM_{0.9}(X)$	508 524	291 984	307 738	296 373	301 814

## Données de la *National Fire Information Database*

### Pouvoir prédictif

Afin de comparer le pouvoir prédictif des deux modèles, on utilise les données de l'année précédente pour estimer  $\Pr(X > 100000)$  de l'année actuelle.

	2011	2012	2013	2014	2015
Actuel	0.17	0.12	0.12	0.13	0.12
LN-PaG	0.12	0.18	0.13	0.13	0.14
GB2-PaG	0.12	0.17	0.13	0.13	0.13

**Remarques :** Les prédictions pour 2011 et 2012 sont moins bonnes. Cependant, il faut noter que l'année 2011 avait la perte moyenne la plus élevée.

# Mesure de risque

1 Introduction

2 Mise en contexte

3 Modélisation de la sévérité

**4 Mesure de risque**

5 Conclusion

# Notions préliminaires

## Mesures de risque de distorsion : définition

On a la v.a.  $B$ , non-négative, qui représente les pertes pour un assureur. On définit l'espérance de  $B$  comme

$$E[B] = \int_0^{\infty} (1 - F_B(x)) dx.$$

On peut également définir une fonction de distorsion  $g : [0, 1] \rightarrow [0, 1]$  et calculer « l'espérance déformée »,  $R[B]$ , définit comme

$$R[B] = \int_0^{\infty} g(1 - F_B(x)) dx.$$

Il existe donc une fonction de distorsion  $g$  qui implique que l'assureur préfère les pertes  $K$  aux pertes  $B$  si et seulement si  $R[K] \leq R[B]$ , voir ([Denuit et al., 2006](#)).

# Notions préliminaires

## Mesures de risque de distorsion : Définition alternative

Si la fonction de distorsions  $g$  est dérivable, on peut réécrire « l'espérance déformée » comme

$$R[B] = \int_0^1 F^{-1}(u)g'(1-u)du = \int_0^1 F^{-1}(u)\psi(u)du = E[Bg'(1 - F_B(B))]$$

On accorde plus de poids aux valeurs élevées du domaine de la v.a.  $B$ .

Par exemple pour la mesure  $TVaR$  on a

$$\psi(u) = \begin{cases} 0, & 0 \leq u \leq \kappa \\ \frac{1}{1-\kappa}, & \kappa < u \leq 1 \end{cases}.$$

# Mesure de risques étudié

## Définition

Dans [Upretree and Brazauskas \(2022\)](#) les auteurs s'intéressent aux mesures de risque suivantes :

- *Value-at Risk.*
- *Tail-Value at Risk.*
- *Proportional Hazards Transform :*

$$PHT_r(B) = \int_0^\infty (1 - F(x))^r dx = \int_0^\infty F^{-1}(u)r(1 - u)^{r-1} du.$$

- *Wang Transformation :*

$$WT_\lambda(B) = \int_0^\infty \Phi(\Phi^{-1}[1 - F(x)] + \lambda) dx = \int_0^1 F^{-1}(u)e^{\lambda\Phi^{-1}(u) - \lambda^2/2} du.$$

- *Gini Shortfall*

$$GS_{\kappa,\delta}(B) = TVaR_\kappa(B) + \frac{2\delta}{(1 - \kappa)^2} \int_c^\infty [(1 - \kappa)(1 - F(x)) - (1 - F(x))^2] dx.$$



# Mesures de risque de distorsion

## Forme analytique

Pour les preuves des formes analytiques suivantes, voir ([Barde, 2023](#)).

Mesure Risque	Exponentielle $\text{Exp}(x_0, \beta)$	Pareto I $\text{PaI}(x_0, \alpha)$	Lognormale $\text{LN}(x_0, \mu, \sigma)$
$VaR$	$x_0 - \beta \ln(1 - \kappa)$	$x_0(1 - \kappa)^{-1/\alpha}$	$x_0 + e^{\mu + \sigma \Phi^{-1}(\kappa)}$
$TVaR$	$x_0 - \beta \ln(1 - \kappa) + \beta$	$x_0(1 - \kappa)^{-1/\alpha} \frac{\alpha}{\alpha - 1}, \alpha > 1$	$x_0 + \frac{1}{1 - \kappa} e^{\mu + \sigma^2/2} \Phi(\sigma - \Phi^{-1}(\kappa))$
$PHT$	$x_0 + \frac{\beta}{r}$	$x_0 \frac{r\alpha}{r\alpha - 1}, a > 1/r$	N.D.
$WT$	N.D.	N.D.	$x_0 + e^{\mu + \lambda\sigma + \sigma^2/2}$
$GS$	$x_0 + \beta \ln(1 - \kappa) + \beta(1 + \delta)$	$x_0(1 - \kappa)^{-1/\alpha} \frac{\alpha}{\alpha - 1} \frac{2(\alpha + \delta) - 1}{2\alpha - 1}, \alpha > 1$	N.D.

# Reproduction de Upretee & Brazauskas (2022)

## Objectif et motivation

---

**Objectif et motivation** : obtenir des bornes pour les mesures de risques pour les lois où il n'y a pas de forme analytique.

**Motivation** : l'intégration numérique peut être plus exigeante que d'utiliser la moyenne entre la borne inférieure et supérieure dans une étude de simulation.

# Reproduction de Upretree & Brazauskas (2022)

## Exemple pour la mesure PHT

Il est possible d'écrire la mesure  $PHT$  pour la loi lognormale de la façon suivante

$$PHT_r(B) = x_0 + e^\mu \sigma \int_{-\infty}^{\infty} (1 - \Phi(z))^r e^{\sigma z} dz = x_0 + e^\mu C_{PHT}(r, \sigma).$$

Voire preuve [Barde \(2023\)](#). Le théorème suivant donne les bornes.

### Théorème (Bornes pour la mesure PHT pour la loi lognormale)

On définit  $C_{PHT}(r, \sigma) = \sigma \int_{-\infty}^{\infty} (1 - \Phi(z))^r e^{\sigma z} dz$ , pour  $0 < r < 1$  and  $\sigma > 0$ . Alors on a la double inégalité

$$e^{\sigma x} [(1 - \Phi(x))^r - (1 - \Phi(x))] + e^{\frac{\sigma^2}{2}} \Phi(\sigma - x) \leq C_{PHT}(r, \sigma) < e^{\sigma x} + K_x(r, \sigma),$$

où  $K_x(r, \sigma) = \sigma x^{-r} r^{-\frac{1}{2}} (2\pi)^{\frac{1-r}{2}} e^{\frac{\sigma^2}{2r}} \Phi\left(\frac{\sigma - rx}{\sqrt{r}}\right)$ ,  $\forall x > 0$ .

# Reproduction de Upretee & Brazauskas (2022)

## Problématique des bornes

---

**Problème** : Comment choisir la valeur de  $x$  qui minimisera l'intervalle entre les deux bornes ?

Il s'agit d'un problème d'optimisation pas très simple.

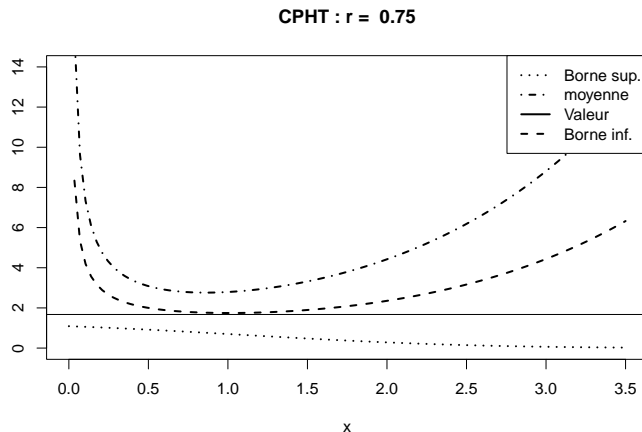
Les auteurs conseillent de regarder graphiquement quelle valeur de  $x$  minimise la moyenne entre les deux bornes.

Il s'agirait d'une bonne approximation selon les auteurs.

# Reproduction de Upretree & Brazauskas (2022)

## Exemple de graphique pour évaluer le bon $x$

Avec les borne on obtient 1.75 alors qu'avec l'intégration numérique on a 1.67.



# Application à la loi Pareto généralisée

## Mesure de risques PHT et GS

On obtient les formes analytiques des mesures de risque  $PHT$  et  $GS$  qui sont

- $PHT_r(Y) = \theta + \frac{\beta}{r-\xi}, \quad r > \xi.$
- $GS_{\kappa,\delta}(Y) = \frac{\beta}{\xi} \left( \frac{(1-\kappa)^{-\xi}}{1-\xi} - 1 \right) + \frac{1}{(1-\kappa)^\xi} \frac{2\delta\beta}{(1-\xi)(2-\xi)} + \theta, \quad \xi < 1.$

On obtient les valeurs suivantes pour la NFID.

$r$	$\widehat{PHT_r(Y)}$
0.75	15 626 068
0.90	1 824 517
1.00	1 270 240

(a) Mesure  $PHT$

$\kappa$	$\widehat{GS_{\kappa,0.25}(Y)}$
0.9	8 781 260
0.95	14 594 203
0.99	47 518 946

(b) Mesure  $GS$

# Application au modèles LN-PaG

## Mesure de risque PHT : modèles LN-PaG

On rappelle la formule pour calculer la mesure  $PHT$ , qui est

$$PHT_r(X) = \int_0^1 F_X^{-1}(u) r(1-u)^r du.$$

Pour le modèle LN-PaG il faut porter une attention particulière à la fonction inverse

$$VaR_\kappa(X) = \begin{cases} \exp\left\{\mu + \sigma\Phi^{-1}\left(\frac{\kappa F_{X_1}(u)}{w}\right)\right\}, & \kappa \leq w \\ \frac{\beta}{\xi} \left(\left(\frac{1-\kappa}{1-w}\right)^{-\xi} - 1\right) + u, & \kappa > w \end{cases}.$$

On aura donc,

$$PHT_r(X) = \int_0^w F_1^{-1}(u) r(1-u)^r du + \int_w^1 F_2^{-1}(u) r(1-u)^r du.$$

# Application au modèles LN-PaG

## Mesure de risque PHT : forme analytique

On obtient la forme analytique suivante :

$$PHT_r(X) = e^{\mu} \sigma \int_{-\infty}^{\Phi^{-1}(F_{X_1}(\theta))} \left( 1 - \frac{w}{F_{X_1}(\theta)} \Phi(z) \right)^r e^{\sigma z} dz - e^{\mu + \sigma \Phi^{-1}(F_{X_1}(\theta))} (1-w)^r \\ + (1-w)^r \left( \frac{\beta}{r-\xi} + \theta \right), \quad r > \xi.$$

Afin de vérifier l'adéquation du modèle LN-PaG on comparera la mesure  $PHT$  avec la valeur empirique de cette mesure, qui est

$$PHT_r(\text{EMP}) = \sum_{i=1}^n x_{(i)} \left( (1 - (i-1)/n)^r - (1 - i/n)^r \right).$$



# Application au modèles LN-PaG

## Mesure de risque PHT : résultats

Les paramètres estimés pour cette base de données sont :

$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{w}$	$\theta$
8.74	2.66	1.36	332 281	0.96	350 000

Ensuite, on obtient les résultats suivants :

$r$	$\widehat{PHT}_r(X)$	$PHT_r(\text{EMP})$
0.90	128 396	116 339
0.95	96 296	92 963
1.00	76 275	75 800

# Conclusion

1 Introduction

2 Mise en contexte

3 Modélisation de la sévérité

4 Mesure de risque

**5 Conclusion**

# Conclusion

---

- L'article de ([Brazauskas and Kleefeld, 2016](#)) a permis :
  - ▶ De découvrir de nouvelles familles pour modéliser les pertes, comme la famille *folded-t*.
  - ▶ De découvrir la mesure de risque  $TCM$ , utile lorsque la mesure  $TVaR$  n'existe pas.
  - ▶ Approfondir la connaissance et la compréhension des deux bases de données.
- L'article de ([Upretree and Brazauskas, 2022](#)) à permis :
  - ▶ Approfondir les connaissances sur les mesures de risque de distorsion.
  - ▶ Découvrir des mesures de risques, par exemple le mesure *Gini Shortfall*.
  - ▶ De développer des aptitudes mathématiques pour obtenir des formes analytiques pour les mesures de risques.

# Remerciements

---

Je remercie sincèrement la Chaire industrielle de recherche sur la construction écoresponsable en bois (CIRCERB) et la chaire d'actuariat pour le financement, qui me permet de me consacrer entièrement à mes études.

Je tiens aussi à remercier les personnes suivantes :

- Prof. Etienne Marceau
- Prof. Hélène Cossette
- Benjamin côté
- André Orelieu Chuisseu Tchuisse

# Merci pour votre attention !

## Des questions ?

# Références I

---

- Barde, J. (2022). Analyse actuarielle du péril incendie. Technical report, Université Laval.
- Barde, J. (2023). Analyse et modélisation des pertes extrêmes : approches et applications. Technical report, Université Laval.
- Brazauskas, V. and Kleefeld, A. (2016). Modeling severity and measuring tail risk of norwegian fire claims. *North American Actuarial Journal*, 20(1) :1–16.
- Denuit, M., Dhaene, J., Goovaerts, M., and Kaas, R. (2006). *Actuarial theory for dependent risks : measures, orders and models*. John Wiley & Sons.
- Fisher, R. A. and Tippet, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical proceedings of the Cambridge philosophical society*, volume 24, pages 180–190. Cambridge University Press.

## Références II

---

- Pickands III, J. (1975). Statistical inference using extreme order statistics. *the Annals of Statistics*, pages 119–131.
- So, M. K. and Wong, C.-M. (2012). Estimation of multiple period expected shortfall and median shortfall for risk management. *Quantitative Finance*, 12(5) :739–754.
- Upretee, S. and Brazauskas, V. (2022). Computing and estimating distortion risk measures : How to handle analytically intractable cases? *North American Actuarial Journal*, pages 1–21.