

Théorie des valeurs extrêmes et estimation de mesures de risque

par

Jonathan EL METHNI
(Université Paris Descartes)

travail en collaboration avec

Gilles STUPFLER
(The University of Nottingham)

Séminaire de Statistique Appliquée du CNAM
1 Février 2019



UNIVERSITÉ
**PARIS
DESCARTES**

- 1 Introduction
- 2 Versions extrêmes des mesures de risque de Wang
- 3 Estimateurs et résultats asymptotiques
- 4 Simulations
- 5 Applications à des données actuarielles
- 6 Conclusions et discussion

- 1 Introduction
 - Théorie des valeurs extrêmes
 - Mesures de risque

- **But de la Théorie des Valeurs Extrêmes** (TVE) $\implies$ étudier et caractériser le comportement des valeurs extrêmes d'un échantillon de variables aléatoires.
- L'exemple d'application historique est l'hydrologie $\implies$ Gumbel années 50.

On suppose la quantité de pluie modélisée par une variable aléatoire X continue ayant pour fonction de répartition $F(x) = \mathbb{P}(X \leq x) = 1 - \mathbb{P}(X \geq x)$.

On dispose de $X_{1,n} \leq \dots \leq X_{n,n}$ un échantillon ordonné de quantités de pluies annuelles.

Les hydrologues souhaitent estimer le niveau de pluie H qui est atteint ou dépassé en moyenne une fois sur T années *i.e.* on veut estimer H tel que

$$1/T = \mathbb{P}(X \geq H) = 1 - F(H)$$

autrement dit on veut estimer

$$H = F^{-1}(1 - 1/T)$$

- H sera appelé niveau de retour correspondant à une période de retour T .
- C'est la quantité standard d'intérêt dans les études environnementales.

Le quantile d'ordre $\beta \in (0, 1)$ de la fonction de répartition est défini par

$$\forall \beta \in (0, 1), \quad q(\beta) = F^{\leftarrow}(\beta) = \inf\{t \in \mathbb{R} \mid F(t) \geq \beta\},$$

où q désigne la fonction quantile de X .

$\implies$ Le niveau de retour $H = q(1 - 1/T)$ est donc un quantile d'ordre $\beta = 1 - 1/T$.

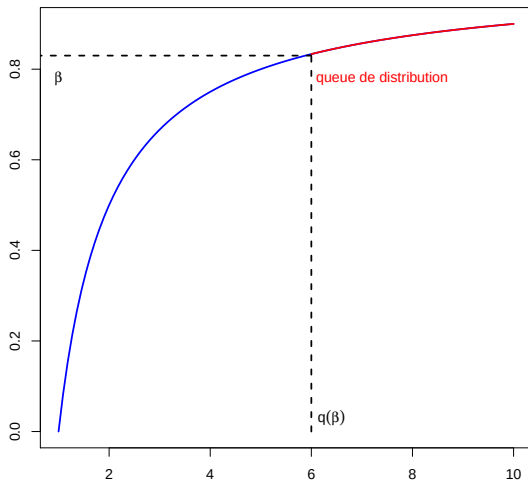
- Que se passe-t-il si la période de retour T est supérieure à la durée d'observation ?
Autrement dit que se passe-t-il si

$$T > n \iff \beta = 1 - 1/T < 1 - 1/n \xrightarrow{n \rightarrow \infty} 1 \quad ?$$

On souhaite estimer des quantiles dit extrêmes $q(\beta_n)$ d'ordre β_n définis par

$$q(\beta_n) = F^{\leftarrow}(\beta_n) \quad \text{avec} \quad \beta_n \rightarrow 1 \quad \text{quand} \quad n \rightarrow \infty$$

- Un niveau de retour dont la période de retour est supérieure à la durée d'observation est un **quantile extrême**.



Que se passe-t-il si $q(\beta_n) > X_{n,n}$?

On peut montrer que l'on a

$$\mathbb{P}(q(\beta_n) > X_{n,n}) = \exp(-n(1 - \beta_n)(1 + o(1)))$$

- 1er cas :

$$\text{Si } n(1 - \beta_n) \rightarrow \infty \text{ alors } \mathbb{P}(q(\beta_n) > X_{n,n}) \rightarrow 0.$$

Un estimateur naturel est la statistique d'ordre $X_{[\lceil n\beta_n \rceil], n}$ où $\lceil \cdot \rceil$ est la fonction **partie entière par excès**.

- 2ème cas :

$$\text{Si } n(1 - \beta_n) \rightarrow 0 \text{ alors } \mathbb{P}(q(\beta_n) > X_{n,n}) \rightarrow 1.$$

Dans ce cas, on ne peut pas estimer $q(\beta_n)$ en inversant simplement la fonction de répartition empirique :

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\{X_i \leq x\} \quad \text{car} \quad \hat{F}_n(x) = 1 \quad \text{pour} \quad x \geq X_{n,n}.$$

Le comportement de $X_{n,n}$ est caractérisé par sa fonction de répartition $F_{X_{n,n}}(x) = F^n(x)$ qui est une loi dégénérée.

Théorème

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées de fonction de répartition F . S'il existe deux suites normalisantes réelles $(a_n)_{n \geq 1} > 0$ et $(b_n)_{n \geq 1} \in \mathbb{R}$ et une loi non-dégénérée $\mathcal{H}_\gamma$ telles que

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{X_{n,n} - b_n}{a_n} \leq x \right) = \lim_{n \rightarrow \infty} F^n(a_n y + b_n) = \mathcal{H}_\gamma(x)$$

alors à une translation et un changement d'échelle près on a

$$\mathcal{H}_\gamma(x) = \exp \left(-(1 + \gamma x)_+^{-1/\gamma} \right),$$

où $\gamma \in \mathbb{R}$ et $z_+ = \max(0, z)$.

- $\mathcal{H}_\gamma$ est appelée fonction de répartition de la **loi des valeurs extrêmes**.
- Si F vérifie le théorème de Fisher-Tippett-Gnedenko on dit que F appartient au domaine d'attraction de $\mathcal{H}_\gamma$.
- Cette loi dépend du seul paramètre de forme γ appelé **indice des valeurs extrêmes** ou **indice de queue**.

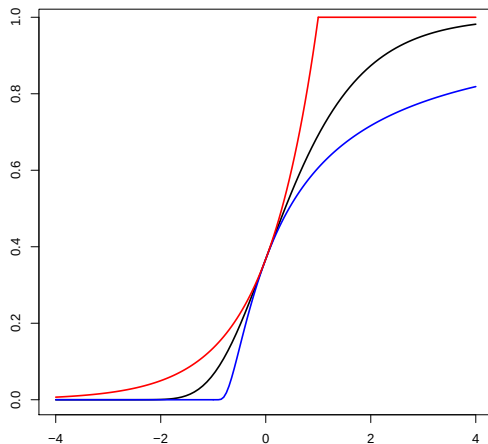
Trois domaines d'attraction

Selon le signe de γ , on distingue trois domaines d'attraction :

- si $\gamma < 0$, on dit que F appartient au domaine d'attraction de **Weibull**. Il contient des lois dont la fonction de survie n'a **pas de queue de distribution**.
- si $\gamma = 0$, on dit que F appartient au domaine d'attraction de Gumbel. Il contient les lois dont la fonction de survie est à décroissance exponentielle, *i.e.* les lois à queues légères.
- si $\gamma > 0$, on dit que F appartient au domaine d'attraction de **Fréchet**. Il contient les lois dont la fonction de survie est à décroissance polynomiale, *i.e.* les lois à **queues lourdes**.

Weibull ($\gamma < 0$)	Gumbel ($\gamma = 0$)	Fréchet ($\gamma > 0$)
Uniforme Beta ReverseBurr	Normale Exponentielle Log-normale Gamma Weibull Logistique Gumbell	Pareto Student Burr Chi-deux Fréchet Log-gamma Log-logistique Cauchy

Trois domaines d'attraction



- Fonction de répartition de la loi des valeurs extrêmes pour $\gamma = -1$, $\gamma = 0$ et $\gamma = 1$

Toutes les lois appartenant au domaine d'attraction de Fréchet peuvent se réécrire

$$F(x) = 1 - x^{-1/\gamma} \ell(x) \quad \text{avec } \gamma > 0,$$

et ℓ est une **fonction à variations lentes** à l'infini i.e. $\forall \lambda \geq 1$,

$$\lim_{x \rightarrow \infty} \frac{\ell(\lambda x)}{\ell(x)} \rightarrow 1.$$

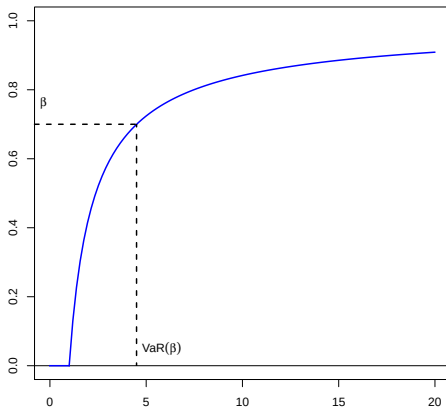
⇒ Le paramètre γ contrôle le comportement de la queue de la fonction de survie et donc celui des valeurs extrêmes.

$\bar{F}(x) = 1 - F(x) = x^{-1/\gamma} \ell(x)$ est une fonction dite à variations régulières d'indice $-1/\gamma$.

Les fonctions à variations lentes sont des fonctions à variations régulières d'indice 0.

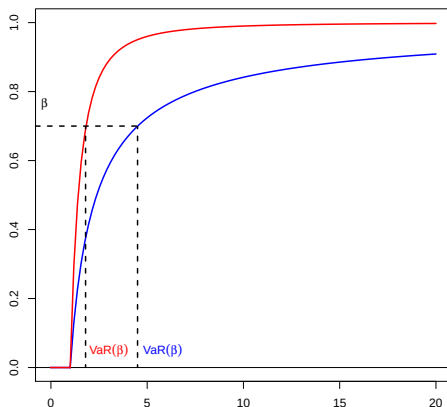
- Soit X une variable aléatoire désignant un montant de pertes. La Value-at-Risk au niveau $\beta \in (0, 1)$ notée $\text{VaR}(\beta)$ introduite en 1993 est définie par

$$\text{VaR}(\beta) = q(\beta)$$



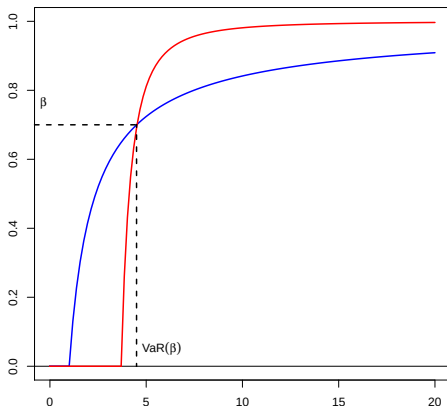
- La $\text{VaR}(\beta)$ est le quantile d'ordre β de la fonction de répartition de la v.a. X .

- Soit X_1 , X_2 deux v.a. de pertes ayant pour fonctions de répartition associées F_1 , F_2 .



- La $\text{VaR}(\beta)$ est différente selon la fonction de répartition F_1 ou F_2 .

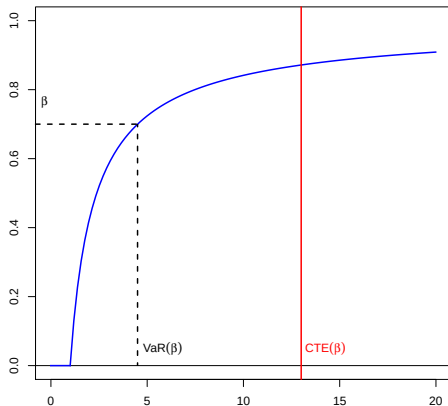
- Soit X_1 , X_2 deux v.a. de pertes ayant pour fonctions de répartition associées F_1 , F_2 .



- Mais les quantiles au niveau β peuvent coïncider...
- Un des principaux reproches fait à la VaR est que des v.a à **queues légères** et à **queues lourdes** (Embrechts *et al.* [1997]) peuvent avoir la même $\text{VaR}(\beta)$ ou $q(\beta)$.

- La Conditional Tail Expectation au niveau $\beta \in (0, 1)$ notée $\text{CTE}(\beta)$ est définie par

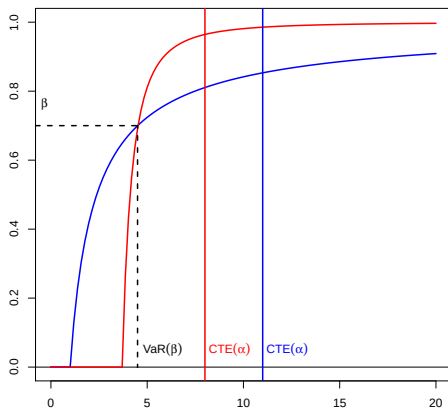
$$\text{CTE}(\beta) = \mathbb{E}(X|X > \text{VaR}(\beta)).$$



- La $\text{CTE}(\alpha)$, donne des informations sur la distribution de X au delà de la $\text{VaR}(\beta)$ et donc contrairement à la $\text{VaR}(\beta)$, sur l'épaisseur de la queue de distribution.

- La Conditional Tail Expectation au niveau $\beta \in (0, 1)$ notée $\text{CTE}(\beta)$ est définie par

$$\text{CTE}(\beta) = \mathbb{E}(X|X > \text{VaR}(\beta)).$$



- Les espérances conditionnelles au-dessus de ce niveau ne coïncident pas.

Pour répondre à la nécessité de principes théoriques et pratiques, Artzner et al. [1999] ont cherché à caractériser ce qui ferait qu'une mesure de risque soit "performante". Dans ce but ils ont introduit la notion de mesure de risque cohérente.

Definition

Une mesure de risque $\mathcal{R}$ est dite cohérente si, pour deux v.a Z et T , elle satisfait les propriétés suivantes :

- **Monotonie** : si $\mathbb{P}(Z \leq T) = 1$, alors $\mathcal{R}(Z) \leq \mathcal{R}(T)$.
- **Homogénéité positive** : $\mathcal{R}(aZ) = a\mathcal{R}(Z)$, pour tout $a > 0$.
- **Invariance par translation** : $\mathcal{R}(Z + b) = \mathcal{R}(Z) + b$, pour tout $b > 0$.
- **Sous-additivité** : $\mathcal{R}(Z + T) \leq \mathcal{R}(Z) + \mathcal{R}(T)$.

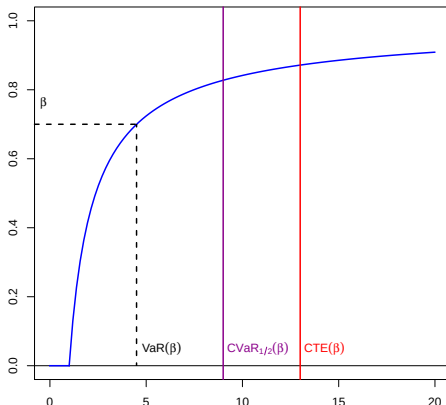
- La VaR n'est pas toujours une mesure de risque cohérente.

⇒ sous-additivité non nécessairement vérifiée.

- La CTE est une mesure de risque cohérente.

- La Conditional-Value-at-Risk au niveau $\beta \in (0, 1)$ notée $\text{CVaR}_\lambda(\beta)$ introduite par Rockafellar et Uryasev [2000] est définie par

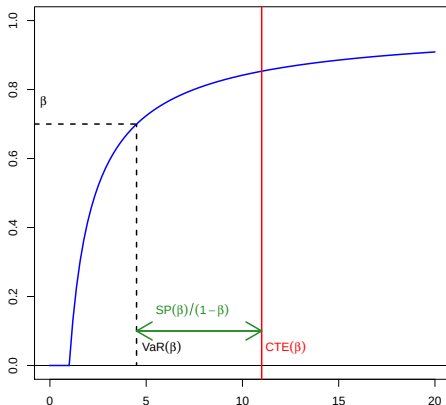
$$\text{CVaR}_\lambda(\beta) = \lambda \text{VaR}(\beta) + (1 - \lambda) \text{CTE}(\beta) \quad \text{avec} \quad 0 \leq \lambda \leq 1.$$



- On peut remarquer que $\text{CVaR}_1(\beta) = \text{VaR}(\beta)$ et que $\text{CVaR}_0(\beta) = \text{CTE}(\beta)$.

- La mesure de risque Stop-loss Premium reinsurance avec un niveau de rétention égal à $\text{VaR}(\beta)$ (voir Cai et Tan [2007]) est définie par :

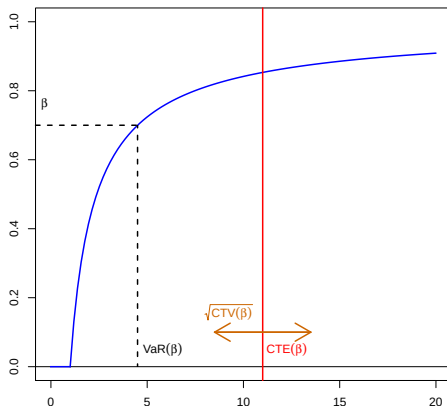
$$\text{SP}(\beta) = (1 - \beta) (\text{CTE}(\beta) - \text{VaR}(\beta)) .$$



- Cette mesure de risque permet ainsi de mettre en évidence les cas dangereux.

- La Conditional Tail Variance au niveau $\beta \in (0, 1)$ notée $CTV(\beta)$ introduite par Valdez [2005] est définie par

$$CTV(\beta) = \mathbb{E}((X - CTE(\beta))^2 | X > VaR(\beta)).$$



- Cette mesure de risque mesure la variabilité conditionnelle de X sachant que $X > VaR(\beta)$ et nous dit à quel point les événements s'éloignent de la $CTE(\beta)$.

Elmethni et al. (2014) ont unifié les définitions des mesures de risques précédentes à l'aide d'une nouvelle mesure de risque.

Le Conditional Tail Moment au niveau $\beta \in (0, 1)$ est défini par

$$\text{CTM}_a(\beta) = \mathbb{E}(X^a | X > \text{VaR}(\beta))$$

où $a \geq 0$ est tel que le moment d'ordre a de Y existe.

Toutes les mesures de risque précédentes au niveau β peuvent être réécrites

Mesures de risque	Réécriture à l'aide de combinaison de $\text{CTM}_a(\beta)$ et de la $\text{VaR}(\beta)$
$\text{CTE}(\beta)$	$\text{CTM}_1(\beta)$
$\text{CVaR}_\lambda(\beta)$	$\lambda \text{VaR}(\beta) + (1 - \lambda) \text{CTM}_1(\beta)$ où $\lambda \in [0, 1]$
$\text{SP}(\beta)$	$(1 - \beta)(\text{CTM}_1(\beta) - \text{VaR}(\beta))$
$\text{CTV}(\beta)$	$\text{CTM}_2(\beta) - \text{CTM}_1^2(\beta)$
$\text{TSD}_\lambda(\beta)$	$\text{CTM}_1(\beta) + \lambda \sqrt{\text{CTM}_2(\beta) - \text{CTM}_1^2(\beta)}$ où $\lambda \geq 0$
$\text{CTS}(\beta)$	$\text{CTM}_3(\beta) / (\text{CTM}_2(\beta) - \text{CTM}_1^2(\beta))^{3/2}$
$\text{GlueVaR}_{\beta, \alpha}^{h_1, h_2}$	$\omega_1 \text{CTM}_1(\beta) + \omega_2 \text{CTM}_1(\alpha) + \omega_3 \text{VaR}(\alpha)$ où $\omega_1 = h_1 - \frac{(h_2 - h_1)(1 - \beta)}{\beta - \alpha}$, $\omega_2 = \frac{(h_2 - h_1)(1 - \alpha)}{\beta - \alpha}$ et $\omega_3 = 1 - \omega_1 - \omega_2 = 1 - h_2$, avec $h_1 \in [0, 1]$, $h_2 \in [h_1, 1]$ et $\alpha \leq \beta$

Soit X une variable aléatoire positive de fonction de répartition F .

- Wang (1996) a introduit une famille de mesure de risque appelée mesures de risque de distorsion.

Pour ce faire il a introduit le concept de fonction de distorsion.

Definition :

Une fonction $g : [0, 1] \rightarrow [0, 1]$ est appelée fonction de distorsion si g est une fonction **croissante** telle que $g(0) = 0$ and $g(1) = 1$.

- Pour plus de simplicité de présentation on supposera que toutes les fonctions de distorsions sont continues à droite.
- C'est une condition très faible qui est valide sur tous les exemples usuels.

Definition :

La mesure de risque de Wang de X , de fonction de distorsion g est définie par :

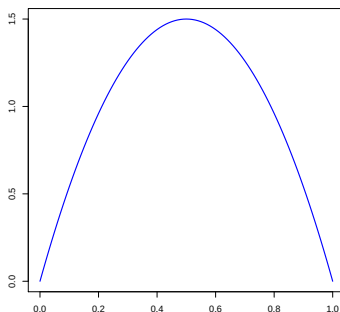
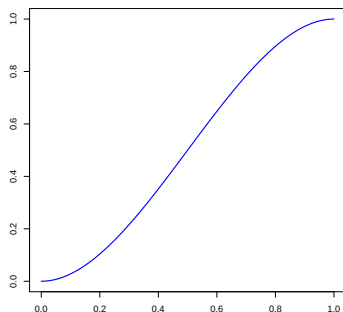
$$R_g(X) = \int_0^\infty g(1 - F(x))dx.$$

On peut réécrire les mesures de risque de Wang sous la forme

$$R_g(X) = \int_0^1 q(1 - \alpha)dg(\alpha).$$

- Si l'on prend $g(x) = x$, on retrouve l'espérance de X .
- Une mesure de risque de Wang peut être aussi vue comme **une version pondérée de l'espérance d'une variable aléatoire X** .
- De plus si g est dérivable, le poids est donné par g' .

Fonction de distorsion (à gauche) et sa dérivée (à droite).



Remarquons également que :

- La mesure de risque R_g est **cohérente** si et seulement si g est **concave** (Wirch et Hardy, 2002).

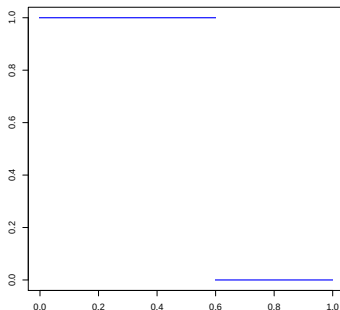
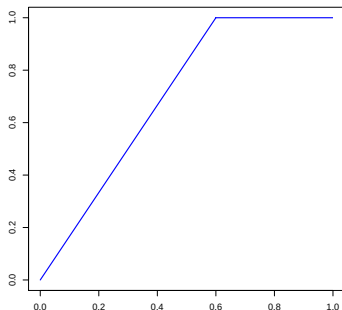
Exemple de mesures de risque de Wang

On rappelle que

$$R_g(X) = \int_0^1 q(1 - \alpha) dg(\alpha).$$

- Tail VaR(β) : moyenne des quantiles de niveau supérieur ou égal à β ,

$$g(x) = \min\{1, x/(1 - \beta)\}$$



Fonction de distorsion (à gauche) et sa dérivée (à droite).

Exemples de mesures de risque de Wang

Mesures de risque $R_g(X)$	Fonctions de distorsion g
VaR au niveau β	$g(x) = \mathbf{1}\{x \geq 1 - \beta\}$ où $0 \leq \beta < 1$
TVaR au dessus du niveau β	$g(x) = \min \left\{ \frac{x}{1 - \beta}, 1 \right\}$ où $0 \leq \beta < 1$
Proportional Hazard transform	$g(x) = x^\alpha$ où $0 < \alpha < 1$
Dual Power	$g(x) = 1 - (1 - x)^{1/\alpha}$ où $0 < \alpha < 1$
MAXMINVAR	$g(x) = (1 - (1 - x)^\alpha)^{1/\alpha}$ où $0 < \alpha < 1$
MINMAXVAR	$g(x) = 1 - (1 - x^{1/\alpha})^\alpha$ où $0 < \alpha < 1$
Gini's principle	$g(x) = (1 + \alpha)x - \alpha x^2$ où $0 < \alpha \leq 1$
Denneberg's absolute deviation	$g(x) = \begin{cases} (1 + \alpha)x & \text{si } 0 \leq x \leq 1/2 \\ \alpha + (1 - \alpha)x & \text{si } 1/2 \leq x \leq 1 \end{cases}$ où $0 < \alpha \leq 1$
Exponential transform	$g(x) = \begin{cases} (1 - \exp(-rx))/(1 - \exp(-r)) & \text{si } r > 0 \\ x & \text{si } r = 0 \end{cases}$
Logarithmic transform	$g(x) = \begin{cases} (\log(1 + rx))/(\log(1 + r)) & \text{si } r > 0 \\ x & \text{si } r = 0 \end{cases}$
Square-root transform	$g(x) = \begin{cases} (\sqrt{1 + rx} - 1)/(\sqrt{1 + r} - 1) & \text{si } r > 0 \\ x & \text{si } r = 0 \end{cases}$
S-inverse shaped transform	$g(x) = a \left(\frac{x^3}{6} - \frac{\delta}{2}x^2 + \left(\frac{\delta^2}{2} + \beta \right)x \right)$ où $a = \left(\frac{1}{6} - \frac{\delta}{2} + \frac{\delta^2}{2} + \beta \right)^{-1}$ avec $0 \leq \delta \leq 1$ et $\beta \in \mathbb{R}$
Wang's transform	$g(x) = \Phi(\Phi^{-1}(x) + \Phi^{-1}(\alpha))$ où Φ est la fonction de répartition d'une $\mathcal{N}(0, 1)$ et $0 \leq \alpha \leq 1$
Beta's transform	$g(x) = \int_0^x \frac{1}{\beta(a, b)} t^{a-1} (1 - t)^{b-1} dt$ où $\beta(a, b)$ est la fonction Beta avec pour paramètres $a, b > 0$

- $\text{VaR}(\beta)$ obtenue pour $g(x) = \mathbb{I}\{x \geq 1 - \beta\}$ où $\mathbb{I}$ est la fonction indicatrice.
- On peut également, dans le cas où F est continue, retrouver les **moments conditionnels** de X (El Methni *et al.*, 2014) :

$$\forall a > 0, \text{CTM}_a(\beta) := \mathbb{E}(X^a | X > q(\beta)) = R_g(X^a)$$

pour $g(x) = \min\{1, x/(1 - \beta)\}$.

⇒ Ainsi que toutes leurs combinaisons du premier tableau.

- J'inclus dorénavant le **CTM** comme une mesure de risque de Wang.

Les mesures de Wang peuvent donc :

- servir à donner une idée du comportement de X sachant que X dépasse un certain niveau.
- et l'information apportée ne se limite pas à celle donnée par un seul quantile de X .

Question

- Comment utiliser (de façon simple) les mesures de Wang pour comprendre le comportement de la queue de distribution de X autrement dit les extrêmes de X ?

2 Versions extrêmes des mesures de risque de Wang

On rappelle que

$$R_g(X) = \int_0^\infty g(1 - F(x))dx.$$

Le premier travail sur le sujet est dû à (Vandewalle et Beirlant, 2006) qui ont considéré les mesures de risque de distorsion de Wang de

$$(X - R)\mathbb{I}\{X > R\} = \max(X - R, 0) \text{ avec } R \uparrow \infty : R_g(X) = \int_R^\infty g(1 - F(x))dx.$$

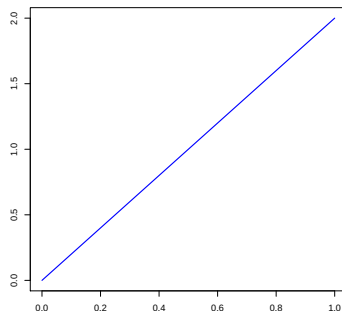
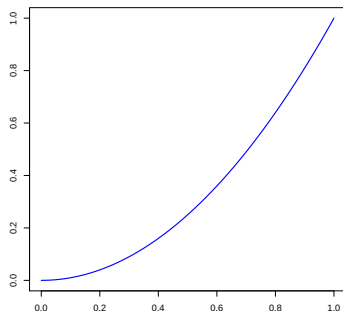
- Leurs travaux sont restreints au cas de fonctions g concaves satisfaisant une condition de variations régulières au voisinage de 0
- Cela exclut de nombreuses mesures de risque (CVaR, GlueVaR) et en particulier la plus simple d'entre toutes : la VaR.

Notre idée est plutôt de considérer :

- Une construction "conditionnelle" c'est-à-dire que l'on va s'intéresser aux extrêmes de X conditionnellement au fait que X soit plus grand qu'un quantile élevé.
- On souhaite également créer un cadre de travail unifiant le plus de mesures de risque possible.

On rappelle que

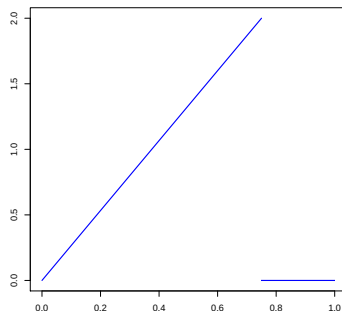
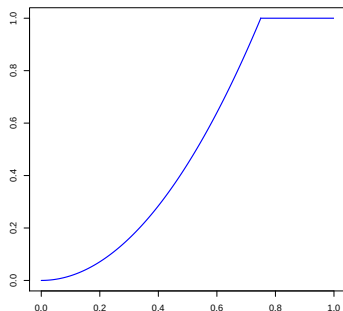
$$R_g(X) = \int_0^1 q(1-\alpha) dg(\alpha).$$



La fonction de distorsion $g(x) = x^2$ (à gauche) et sa dérivée (à droite).

On rappelle que

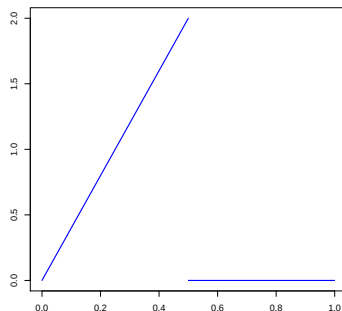
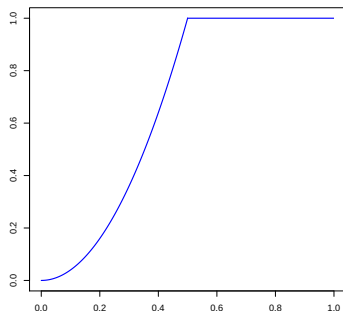
$$R_g(X) = \int_0^1 q(1-\alpha) dg(\alpha).$$



La fonction de distorsion $g_{0.25}$ (à gauche) et sa dérivée (à droite).

On rappelle que

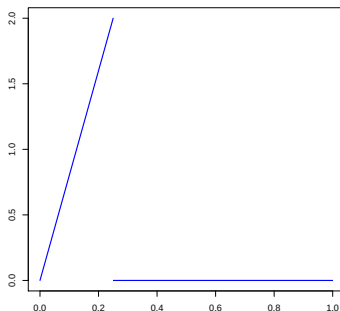
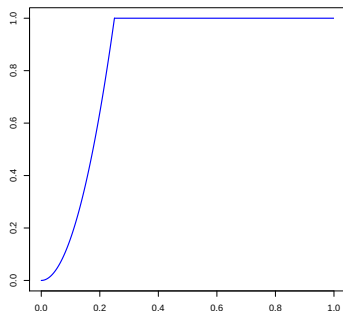
$$R_g(X) = \int_0^1 q(1-\alpha) dg(\alpha).$$



La fonction de distorsion $g_{0.5}$ (à gauche) et sa dérivée (à droite).

On rappelle que

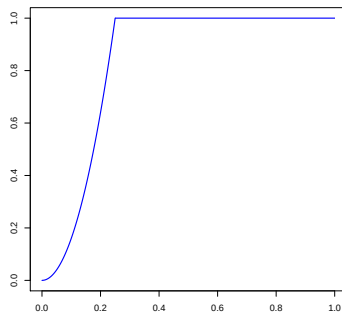
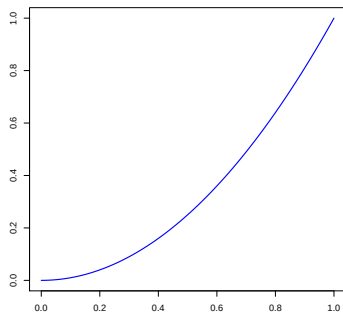
$$R_g(X) = \int_0^1 q(1-\alpha) dg(\alpha).$$



La fonction de distorsion $g_{0.75}$ (à gauche) et sa dérivée (à droite).

On rappelle que

$$R_g(X) = \int_0^1 q(1-\alpha) dg(\alpha).$$



La fonction de distorsion $g(x) = x^2$ (à gauche) et $g_{0.75}$ (à droite).

Soit g une fonction de distorsion. Pour $\beta \in [0, 1]$, on pose

$$g_\beta(y) := \begin{cases} g\left(\frac{y}{1-\beta}\right) & \text{si } y \leq 1-\beta \\ 1 & \text{sinon.} \end{cases}$$

- La fonction g_β est une fonction de distorsion, constante égale à 1 sur $[1-\beta, 1]$.
- La mesure de risque de Wang associée est notée $R_{g,\beta}(X)$.
- De plus si g est concave alors g_β l'est également.

Intérêt : la distribution des poids accordée par g_β aux différents quantiles est celle accordée par g , mais **déplacée** sur $[0, \beta]$.

$\implies$ On ne prend en compte que les grands quantiles dont l'ordre se situe dans $[\beta, 1]$.

En utilisant cette construction on retrouve les mesures de risque classiques :

- $\text{VaR}(\beta)$ ou $q(\beta)$ qui correspond au cas $g(x) = \mathbb{I}\{x = 1\}$
- Lorsque F est continue, le moment d'ordre conditionnel d'ordre a de X au-delà du niveau $q(\beta)$ est

$$\forall a > 0, \text{CTM}_a(\alpha) = R_{g,\beta}(X^a)$$

pour $g(x) = x$.

⇒ Comment estimer les mesures de risque $R_{g,\beta}$ lorsque $\beta \uparrow 1$?

- 3 Estimateurs et résultats asymptotiques
 - Cas intermédiaire
 - Cas extrême

On a

$$\begin{aligned}R_{g,\beta}(X^a) &= \int_0^1 q(1-\alpha)^a dg_\beta(\alpha) \\ &= \int_0^1 q(1-(1-\beta)s)^a dg(s)\end{aligned}$$

Si on a X qui suit une loi de Pareto stricte (sans fonction à variations lentes) soit :

$$\forall x > 1, \quad F(x) = 1 - x^{-1/\gamma} \quad \text{ou} \quad q(\alpha) = (1-\alpha)^{-\gamma}$$

On a alors

$$R_{g,\beta}(X^a) = \int_0^1 (q(\beta)s^{-\gamma})^a dg(s) = q(\beta)^a \int_0^1 s^{-a\gamma} dg(s).$$

Dans notre cadre de travail, l'égalité suivante est toujours vraie **asymptotiquement** pour $\beta_n \uparrow 1$ lorsque $n \rightarrow \infty$ soit

$$R_{g,\beta_n}(X^a) = q(\beta_n)^a \int_0^1 s^{-a\gamma} dg(s) (1 + o(1)).$$

On suppose disposer d'un n -échantillon $(X_1, \dots, X_n)$ de copies indépendantes de X .

$$R_{g, \beta_n}(X^a) = q(\beta_n)^a \int_0^1 s^{-a\gamma} dg(s)(1 + o(1)).$$

- Soit $\beta = \beta_n \uparrow 1$ de sorte que $R_{g, \beta_n}(X^a)$ soit une mesure de risque extrême.
- On remplace $q(\beta_n)$ par son analogue empirique $\hat{q}_n(\beta_n) = X_{[n\beta_n], n}$.
- On remplace γ par un estimateur $\hat{\gamma}_n$.

Notre premier estimateur est alors

$$\hat{R}_{g, \beta_n}^{AE}(X^a) = X_{[n\beta_n], n}^a \int_0^1 s^{-a\hat{\gamma}_n} dg(s).$$

Second estimateur : Fonctionnal Plug-in

On rappelle que l'on a pour $a > 0$:

$$R_{g,\beta_n}(X^a) = \int_0^1 [q(1 - (1 - \beta_n)s)]^a dg(s).$$

Dans l'estimateur précédent nous avons utilisé les deux approximations successives

$$q(1 - (1 - \beta_n)s) \approx q(\beta_n)s^{-\gamma} \approx X_{\lceil n\beta_n \rceil, n} s^{-\hat{\gamma}_n}$$

L'idée est cette fois de ne faire qu'une seule approximation en remplaçant directement la fonction q par son analogue empirique :

$$s \mapsto \hat{q}_n(1 - (1 - \beta_n)s) = X_{\lceil n(1 - (1 - \beta_n)s) \rceil, n}$$

ce qui nous donne notre second estimateur :

$$\hat{R}_{g,\beta_n}^{PL}(X^a) = \int_0^1 X_{\lceil n(1 - (1 - \beta_n)s) \rceil, n}^a dg(s)$$

Puisque les quantiles empiriques utilisés sont des statistiques d'ordre élevé (dont l'ordre tend vers 1 lorsque $n \rightarrow \infty$), il nous faut un cadre adapté à leur étude.

On introduit la fonction “tail quantile” U définie par

$$\forall t > 1, U(t) = q(1 - t^{-1}).$$

Hypothèse (Condition de second ordre $\mathcal{C}_2(\gamma, \rho, A)$)

Il existe $\gamma > 0$ et $\rho \leq 0$ tels que

$$\forall x > 0, \lim_{t \rightarrow \infty} \frac{1}{A(t)} \left(\frac{U(tx)}{U(t)} - x^\gamma \right) = x^\gamma \frac{x^\rho - 1}{\rho}$$

où A est une fonction mesurable, de signe constant, qui converge vers 0.

Lorsque $\rho = 0$, le membre de droite se lit $x^\gamma \log x$.

- Cette condition classique en TVE signifie que U est à variations régulières d'indice γ ou autrement dit qu'elle se comporte **essentiellement comme x^γ** lorsque $x \rightarrow \infty$
- Cette hypothèse contrôle la vitesse de convergence du rapport $U(tx)/U(t)$ vers x^γ lorsque $t \rightarrow \infty$.

Théorème 1 : estimateur AE

Supposons que U vérifie $\mathcal{C}_2(\gamma, \rho, A)$. Supposons également que

$$\beta_n \rightarrow 1, \quad n(1 - \beta_n) \rightarrow \infty \quad \text{et} \quad \sqrt{n(1 - \beta_n)} A((1 - \beta_n)^{-1}) \rightarrow \lambda \in \mathbb{R}$$

et que

$$\sqrt{n(1 - \beta_n)} \left(\hat{\gamma}_n - \gamma, \frac{X_{\lceil n\beta_n \rceil, n}}{q(\beta_n)} - 1 \right) \xrightarrow{d} (\Gamma, \Theta).$$

Soient $g_1, \dots, g_d$ des fonctions de distorsion et $a_1, \dots, a_d > 0$ si

$$\exists \eta > 0, \quad \forall j \in \{1, \dots, d\}, \quad \int_0^1 s^{-a_j \gamma - 1/2 - \eta} dg_j(s) < \infty.$$

alors le vecteur aléatoire $\sqrt{n(1 - \beta_n)} \left(\frac{\hat{R}_{g_j, \beta_n}^{AE}(X^{a_j})}{R_{g_j, \beta_n}(X^{a_j})} - 1 \right)_{1 \leq j \leq d}$ a pour loi jointe

asymptotique

$$\left(a_j \left[\Theta + \frac{\int_0^1 s^{-a_j \gamma} \left(\log(1/s) \Gamma - \lambda \frac{s^{-\rho} - 1}{\rho} \right) dg_j(s)}{\int_0^1 s^{-a_j \gamma} dg_j(s)} \right] \right)_{1 \leq j \leq d}.$$

Théorème 2 : estimateur PL

Supposons que U vérifie $\mathcal{C}_2(\gamma, \rho, A)$. Supposons également que

$$\beta_n \rightarrow 1, \quad n(1 - \beta_n) \rightarrow \infty \quad \text{et} \quad \sqrt{n(1 - \beta_n)} A((1 - \beta_n)^{-1}) \rightarrow \lambda \in \mathbb{R}.$$

Soient $g_1, \dots, g_d$ des fonctions de distorsion et $a_1, \dots, a_d > 0$. Si

$$\exists \eta > 0, \quad \forall j \in \{1, \dots, d\}, \quad \int_0^1 s^{-a_j \gamma - 1/2 - \eta} dg_j(s) < \infty,$$

alors

$$\sqrt{n(1 - \beta_n)} \left(\frac{\widehat{R}_{g_j, \beta_n}^{PL}(X^{a_j})}{R_{g_j, \beta_n}(X^{a_j})} - 1 \right)_{1 \leq j \leq d} \xrightarrow{d} \mathcal{N}(0, V)$$

avec

$$V_{i,j} = a_i a_j \gamma^2 \frac{\int_{[0,1]^2} \min(s, t) s^{-a_i \gamma - 1} t^{-a_j \gamma - 1} dg_i(s) dg_j(t)}{\int_0^1 s^{-a_i \gamma} dg_i(s) \int_0^1 t^{-a_j \gamma} dg_j(t)}.$$

La condition d'intégrabilité

$$\int_0^1 s^{-a\gamma-1/2-\eta} dg(s) < \infty$$

- entraîne en particulier que

$$\int_0^1 s^{-a\gamma-\eta} dg(s) < \infty$$

et donc que la mesure de risque de Wang considérée **existe et est finie**.

- le facteur supplémentaire $s^{-1/2}$ permet d'utiliser un **théorème d'approximation** du processus de quantile empirique $s \mapsto X_{\lceil n(1-(1-\beta_n)s) \rceil, n}$ par un **mouvement brownien**, ce qui constitue le point de départ de la preuve des Théorèmes 1 et 2.

En particulier, si g est de classe $\mathcal{C}^1$ avec $g'(s) \sim s^b$ quand $s \downarrow 0$, la condition d'intégrabilité devient

$$\exists \eta > 0, \int_0^1 s^{b-a\gamma-1/2-\eta} ds < \infty \quad \Leftrightarrow \quad \gamma < \frac{2b+1}{2a}.$$

Dans le cas de l'estimation d'un CTM d'ordre a , on a $g'(x) = 1$ et la condition trouvée est

$$\gamma < \frac{1}{2a}$$

qui est celle obtenue également par El Methni *et al.* (2014).

La condition $\sqrt{n(1-\beta_n)}A((1-\beta_n)^{-1}) \rightarrow \lambda$ permet le contrôle du **biais asymptotique** de nos estimateurs. Elle est classique en TVE.

- Les conditions $\beta_n \rightarrow 1$ et $n(1 - \beta_n) \rightarrow \infty$: signifient qu'on ne prend en compte que des quantiles qui sont d'ordre "élevé, mais pas trop".
- Cela nous assure seulement que nos estimateurs estiment des mesures de risque intermédiaires ("extrêmes mais pas trop") dans le sens où l'ordre du plus petit quantile que l'on prend convergera suffisamment lentement vers 1.
- En d'autres termes on ne peut estimer des mesures de risque qui se trouvent au dessus d'un seuil $q(\beta_n)$ qui dans ce cas appartient à notre échantillon.
- Cette restriction nous empêche d'estimer des mesures de risque aussi extrêmes que l'on voudrait. Nous allons alors introduire un nouvel estimateur nous permettant de d'annuler cette restriction.

On se concentre sur l'élimination de la condition $n(1 - \beta_n) \rightarrow \infty$.

Cas extrême : estimateur de Weissman

Afin de s'affranchir de la condition sur l'ordre β_n de la mesure de Wang, on remarque que $C_2(\gamma, \rho, A)$ entraîne

$$\frac{U(tx)}{U(t)} = \frac{q(1 - 1/tx)}{q(1 - 1/t)} = \frac{(tx)^\gamma \ell(tx)}{t^\gamma \ell(t)} = x^\gamma \frac{\ell(tx)}{\ell(t)} \approx x^\gamma \text{ quand } t \rightarrow \infty.$$

Par conséquent, si on prend β_n et $\delta_n \uparrow 1$, on a sous certaines conditions

$$\forall s \in (0, 1), \quad q(1 - (1 - \delta_n)s) \approx \left(\frac{1 - \beta_n}{1 - \delta_n} \right)^\gamma q(1 - (1 - \beta_n)s)$$

pour n assez grand. Cette équivalence relie les quantiles extrêmes aux quantiles intermédiaires à l'aide de γ .

En d'autres termes, on peut extrapoler l'estimation à des quantiles extrêmes, pourvu que l'on puisse estimer l'indice des valeurs extrêmes γ .

On estime le quantile extrême $q(1 - (1 - \delta_n)s)$ par

$$\forall s \in (0, 1), \quad \hat{q}^W(1 - (1 - \delta_n)s | \beta_n) = \left(\frac{1 - \beta_n}{1 - \delta_n} \right)^{\hat{\gamma}_n} X_{\lceil n(1 - (1 - \beta_n)s) \rceil, n}$$

appelé estimateur de Weissman (1978).

On a

$$\forall s \in (0, 1), \quad q(1 - (1 - \delta_n)s) \approx \left(\frac{1 - \beta_n}{1 - \delta_n} \right)^\gamma q(1 - (1 - \beta_n)s)$$

d'où

$$\forall s \in (0, 1), \quad q(1 - (1 - \delta_n)s)^a \approx \left(\frac{1 - \beta_n}{1 - \delta_n} \right)^{a\gamma} q(1 - (1 - \beta_n)s)^a$$

en intégrant selon dg on obtient

$$\int_0^1 q(1 - (1 - \delta_n)s)^a dg(s) \approx \int_0^1 \left(\frac{1 - \beta_n}{1 - \delta_n} \right)^{a\gamma} q(1 - (1 - \beta_n)s)^a dg(s)$$

soit

$$R_{g, \delta_n}(X^a) = \left(\frac{1 - \beta_n}{1 - \delta_n} \right)^{a\gamma} R_{g, \beta_n}(X^a)(1 + o(1))$$

On a

$$R_{g,\delta_n}(X^a) = \left(\frac{1 - \beta_n}{1 - \delta_n} \right)^{a\gamma} R_{g,\beta_n}(X^a)(1 + o(1))$$

On prend alors :

- un niveau **intermédiaire** $\beta_n \uparrow 1$ tel que $n(1 - \beta_n) \rightarrow \infty$,
- un niveau **extrême** $\delta_n \uparrow 1$ tel que $(1 - \delta_n)/(1 - \beta_n) \rightarrow 0$,
- $\hat{\gamma}_n$ un estimateur de γ ,
- $\hat{R}_{g,\beta_n}(X^a)$ un estimateur de $R_{g,\beta_n}(X^a)$, qui peut être aussi bien l'estimateur AE que l'estimateur PL. On peut alors estimer la mesure de risque $R_{g,\delta_n}(X^a)$ par

$$\hat{R}_{g,\delta_n}^W(X^a; \beta_n) = \left(\frac{1 - \beta_n}{1 - \delta_n} \right)^{a\hat{\gamma}_n} \hat{R}_{g,\beta_n}(X^a)$$

C'est un estimateur de type Weissman, obtenu par **estimation** de la mesure de risque au-delà d'un niveau **intermédiaire** puis par multiplication par un **facteur d'extrapolation**.

Cette technique nous permet d'estimer des mesures de risque de Wang aussi extrêmes que l'on veut et pour cela il nous suffit de savoir estimer des **quantiles extrêmes**.

Théorème 3 : estimateur de type Weissman

Supposons que U vérifie $\mathcal{C}_2(\gamma, \rho, A)$, avec $\rho < 0$. Supposons également que

$\beta_n, \delta_n \rightarrow 1$, $n(1 - \beta_n) \rightarrow \infty$, $(1 - \delta_n)/(1 - \beta_n) \rightarrow 0$ et $\sqrt{n(1 - \beta_n)}A((1 - \beta_n)^{-1}) \rightarrow \lambda \in \mathbb{R}$

Soient $g_1, \dots, g_d$ des fonctions de distorsion et $a_1, \dots, a_d > 0$ si

$$\exists \eta > 0, \forall j \in \{1, \dots, d\}, \int_0^1 s^{-a_j \gamma - 1/2 - \eta} dg_j(s) < \infty \text{ et } \sqrt{n(1 - \beta_n)}(\hat{\gamma}_n - \gamma) \xrightarrow{d} \xi$$

alors sachant que

$$\forall j \in \{1, \dots, d\}, \sqrt{n(1 - \beta_n)} \left(\frac{\hat{R}_{g_j, \beta_n}(X^{a_j})}{R_{g_j, \beta_n}(X^{a_j})} - 1 \right) = O_{\mathbb{P}}(1)$$

on a

$$\frac{\sqrt{n(1 - \beta_n)}}{\log([1 - \beta_n]/[1 - \delta_n])} \left(\frac{\hat{R}_{g_j, \delta_n}^W(X^{a_j}; \beta_n)}{R_{g_j, \delta_n}(X^{a_j})} - 1 \right)_{1 \leq j \leq d} \xrightarrow{d} \begin{pmatrix} a_1 \xi \\ \vdots \\ a_d \xi \end{pmatrix}.$$

- On retrouve le théorème correspondant sur l'estimateur classique de Weissman des quantiles extrêmes en prenant $d = 1$ et $g(x) = \mathbb{1}\{x = 1\}$.
- $\hat{\gamma}_n$ dicte la limite de notre estimateur de la mesure de risque de Wang extrême.

- 4 Simulations
 - Procédure de choix
 - Comparaison de nos deux estimateurs

On a nos deux estimateurs (cas intermédiaire)

$$\widehat{R}_{g,\beta_n}^{AE}(X^a) = X_{[n\beta_n],n}^a \int_0^1 s^{-a\widehat{\gamma}_n} dg(s) \quad \text{et} \quad \widehat{R}_{g,\beta_n}^{PL}(X^a) = \int_0^1 X_{[n(1-(1-\beta_n)s)],n}^a dg(s)$$

Que l'on souhaite extrapoler à l'aide de notre estimateur de type Weissman

$$\widehat{R}_{g,\delta_n}^W(X^a|\beta_n) = \left(\frac{1-\beta_n}{1-\delta_n} \right)^{a\widehat{\gamma}_n} \widehat{R}_{g,\beta_n}(X^a)$$

Nos estimateurs sont basés sur le choix du paramètre β_n , qui intervient :

- dans l'estimation de γ
- dans l'estimation **préliminaire** de la mesure de risque $R_{g,\beta_n}(X)$, qu'on utilise pour l'**extrapolation**.

⇒ La question étant : à partir de quel moment est-on dans la queue de distribution ?

- Le choix du paramètre β_n est **crucial** :
 - choisir β_n trop proche de 1 réduit le biais car on reste dans la queue de distribution mais si le seuil β_n est trop grand on aura alors peu de statistiques d'ordre ce qui augmentera considérablement la variance.
 - choisir β_n trop loin de 1 réduit la variance car utilisera beaucoup de statistiques d'ordre et notre estimateur sera stable mais si le seuil β_n est trop petit on sortira de la queue de distribution ce qui introduira du biais.
- Il convient donc de choisir un paramètre β_n qui soit proche de 1 “mais pas trop”, afin de réaliser un **compromis biais-variance**.

⇒ On souhaite proposer un choix automatique de β_n .

- Pour cela il nous faut un estimateur de γ afin dans un premier temps de l'utiliser comme **outil exploratoire**.

Lorsque l'on veut analyser en pratique un jeu de données d'un point de vue des extrêmes, on commence souvent par tracer un ou plusieurs graphiques d'estimateurs de l'indice de queue γ .

On souhaite y détecter une **plage de stabilité** (en terme d'estimation) dans les données, pour laquelle (en fonction de β)

- la **variance** de l'estimateur est **modérée** ;
- le **biais**, apporté essentiellement par les petites estimations, est **encore négligeable**.

Cette plage de stabilité s'interprète en général comme celle pour laquelle :

- on est toujours **dans les extrêmes** de l'échantillon,
- tout en ayant **assez de variables** pour pouvoir calculer des estimations raisonnables.

Notre idée est de détecter la première plage de stabilité se trouvant dans la partie extrême de l'échantillon (queue de distribution) de $\beta \mapsto \hat{\gamma}_\beta$.

Plusieurs estimateurs :

- Le plus courant est l'estimateur de Hill (1975), qui est aussi l'estimateur du maximum de vraisemblance pour des lois de Pareto,

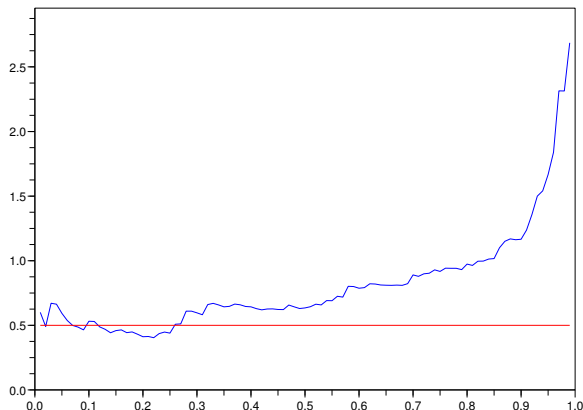
$$\hat{\gamma}_{\beta_n} = \frac{1}{\lceil n(1 - \beta_n) \rceil} \sum_{i=1}^{\lceil n(1 - \beta_n) \rceil} \log(X_{n-i+1,n}) - \log(X_{n - \lceil n(1 - \beta_n) \rceil, n})$$

Si $\beta_n \rightarrow 1$, $n(1 - \beta_n) \rightarrow \infty$ et $\sqrt{n(1 - \beta_n)}A((1 - \beta_n)^{-1}) \rightarrow \lambda$ on a alors

$$\sqrt{n(1 - \beta_n)}(\hat{\gamma}_{\beta_n} - \gamma) \xrightarrow{d} \mathcal{N}(0, \gamma^2)$$

- Avec un nombre fini de quantiles empiriques : Pickands (1975)
- Estimateur du maximum de vraisemblance : Smith (1987)
- Estimateur des moments pondérés : Hosking and Wallis (1987)
- Version débiaisée : Peng (1998), Fraga Alves *et al.* (2003), Caeiro and Gomes (2005)...

et bien d'autres !

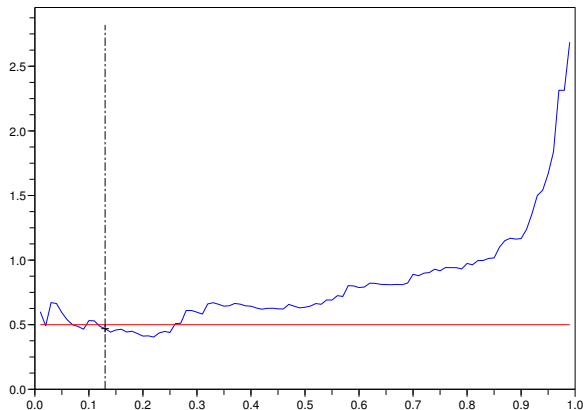


- Trajectoire de l'estimateur de Hill, pour $n = 100$ variables de Pareto où $\gamma = 1/2$.
- Attention en abscisse : $1-\beta$.

Une possibilité est la suivante :

- on choisit $\beta_0 > 0$ et un paramètre de fenêtre $h > 0$;
- pour $\beta_0 < \beta < 1 - h$, on pose $I(\beta, h) = [\beta, \beta + h]$ et on calcule l'écart-type $\sigma(\beta, h)$ du bloc des estimations $\{\hat{\gamma}_\beta, \beta \in I\}$.
- on trouve le dernier β tel que $\sigma(\beta, h)$ atteigne :
 - un minimum local,
 - qui soit inférieur à un critère donné, comme la moyenne ou la médiane de $\{\sigma(\beta, h), \beta \in]\beta_0, 1 - h[\}$ lorsque β varie.
- en ce sens, c'est dans cet intervalle que l'estimation est suffisamment stable pour la première fois dans les extrêmes de l'échantillon .
- on choisit une valeur β^* dans cet intervalle, par exemple :
 - le centre de l'intervalle,
 - celui qui rend l'estimation médiane.

Exemple de procédure de choix



- La trajectoire précédente en prenant $\beta_0 = 0.5$ et $h = 0.1$ et en retenant l'estimation médiane dans la plage de stabilité, on choisit ici $\beta^* = 0.86$.
- Attention en abscisse : $1-\beta$.

On considère deux cas pour la distribution de X :

- la loi de **Fréchet** :

$$F(x) = e^{-x^{-1/\gamma}}, \quad x > 0$$

- la loi de **Burr** :

$$F(x) = 1 - (1 + x^{-\rho/\gamma})^{-1/\rho}, \quad x > 0$$

Ces distributions sont des cas classiques de lois à queue lourde :

- elles ont toutes les deux pour indice des valeurs extrêmes γ ,
- pour la loi de **Fréchet**, on a $\rho = -1$,
- pour la loi de **Burr**, le paramètre de second-ordre est ρ .

Dans chaque cas, on applique notre procédure à :

- 5000 échantillons de $n \in \{100, 300\}$ copies indépendantes de X ;
- Notre procédure de choix est mise en oeuvre avec $\beta_0 = 0.5$ et $h = 0.1$ comme paramètre de fenêtre.
- Nous avons fait tourner nos simulations pour $\gamma \in \{1/6, 1/5, 1/4\}$ et $\rho \in \{-2, -1\}$.

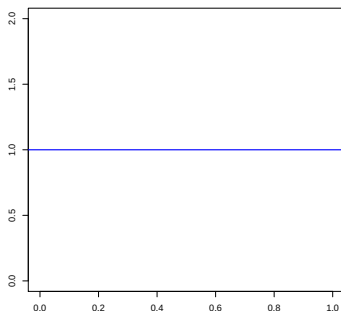
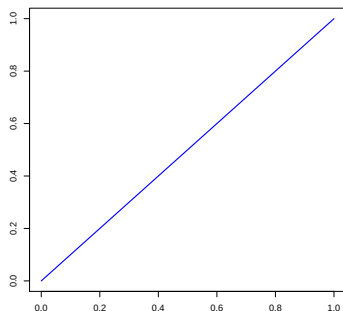
Nous retenons les erreurs quadratiques moyennes relatives pour chaque estimateur :

$$\text{MSE}(\hat{R}_{g,\delta}^W) = \frac{1}{N} \sum_{i=1}^N \left(\frac{\hat{R}_{g,\delta}^W(X|\beta_i^*)}{R_{g,\delta}(X)} - 1 \right)^2$$

avec trois niveaux d'extrapolation $\delta = 0.99, 0.995$ et 0.999 .

Ici β_i^* est le niveau intermédiaire obtenu pour le i^{eme} échantillon.

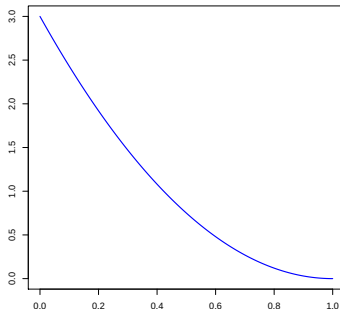
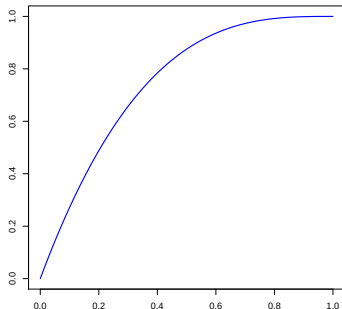
- La CTE pour laquelle $g(x) = x$ affecte le même poids à tous les quantiles.



- Les conditions d'intégrabilité pour l'existence de la mesure de risque (respectivement pour la validité du Théorème 3) sont : $\gamma < 1$ (respectivement $\gamma < 1/2$).

Valeur de γ	δ	Estimateur	Fréchet		Burr $\rho = -1$		Burr $\rho = -2$	
			$n = 100$	$n = 300$	$n = 100$	$n = 300$	$n = 100$	$n = 300$
$\gamma = 1/6$	0.99	AE	0.0325	0.0098	0.0374	0.0133	0.0291	0.0095
		PL	0.0317	0.0097	0.0357	0.0127	0.0286	0.0094
	0.995	AE	0.0457	0.0137	0.0540	0.0191	0.0401	0.0130
		PL	0.0446	0.0135	0.0518	0.0184	0.0395	0.0129
	0.999	AE	0.0891	0.0258	0.1115	0.0386	0.0752	0.0236
		PL	0.0871	0.0255	0.1073	0.0375	0.0741	0.0235
$\gamma = 1/5$	0.99	AE	0.0519	0.0164	0.0627	0.0199	0.0472	0.0140
		PL	0.0502	0.0161	0.0588	0.0191	0.0461	0.0138
	0.995	AE	0.0739	0.0229	0.0915	0.0289	0.0657	0.0191
		PL	0.0717	0.0225	0.0862	0.0277	0.0643	0.0189
	0.999	AE	0.1500	0.0437	0.1952	0.0589	0.1266	0.0349
		PL	0.1461	0.0430	0.1850	0.0569	0.1239	0.0344
$\gamma = 1/4$	0.99	AE	0.0973	0.0285	0.1028	0.0349	0.0834	0.0248
		PL	0.0900	0.0278	0.0944	0.0332	0.0835	0.0246
	0.995	AE	0.1411	0.0402	0.1515	0.0509	0.1190	0.0341
		PL	0.1305	0.0392	0.1395	0.0484	0.1202	0.0337
	0.999	AE	0.3039	0.0787	0.3350	0.1063	0.2492	0.0631
		PL	0.2807	0.0768	0.3102	0.1017	0.2604	0.0622

- La Dual Power (DP) pour laquelle $g(x) = 1 - (1 - x)^{1/\alpha}$ avec $\alpha \in (0, 1)$, qui affecte un poids **plus fort aux grands quantiles**.



Ici $\alpha = 1/3$.

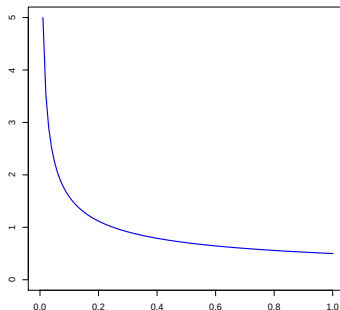
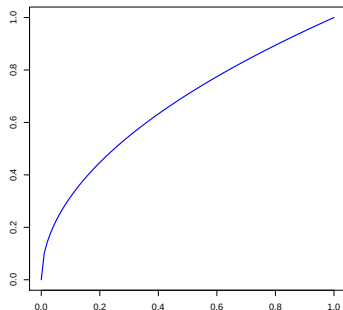
⇒ Lorsque $c = 1/\alpha$ est entier, la mesure Dual Power s'interprète comme l'espérance de $\max(X_1, \dots, X_c)$ si $X_1, \dots, X_c$ sont des copies indépendantes de X .

- Les conditions d'intégrabilité pour l'existence de la mesure de risque (respectivement pour la validité du Théorème 3) sont : $\gamma < 1$ (respectivement $\gamma < 1/2$).

Valeur de γ	δ	Estimateur	Fréchet		Burr $\rho = -1$		Burr $\rho = -2$	
			$n = 100$	$n = 300$	$n = 100$	$n = 300$	$n = 100$	$n = 300$
$\gamma = 1/6$	0.99	AE	0.0487	0.0169	0.0629	0.0215	0.0458	0.0140
		PL	0.0448	0.0160	0.0549	0.0194	0.0443	0.0142
	0.995	AE	0.0653	0.0225	0.0866	0.0295	0.0609	0.0182
		PL	0.0597	0.0212	0.0757	0.0267	0.0586	0.0184
	0.999	AE	0.1177	0.0394	0.1658	0.0549	0.1084	0.0307
		PL	0.1073	0.0371	0.1456	0.0499	0.1033	0.0306
$\gamma = 1/5$	0.99	AE	0.0808	0.0261	0.0988	0.0336	0.0680	0.0211
		PL	0.0743	0.0256	0.0852	0.0304	0.0652	0.0217
	0.995	AE	0.1100	0.0349	0.1376	0.0463	0.0907	0.0276
		PL	0.1004	0.0339	0.1187	0.0417	0.0862	0.0281
	0.999	AE	0.2078	0.0620	0.2723	0.0870	0.1630	0.0468
		PL	0.1879	0.0598	0.2362	0.0785	0.1535	0.0468
$\gamma = 1/4$	0.99	AE	0.1558	0.0449	0.2175	0.0570	0.1327	0.0376
		PL	0.1397	0.0439	0.1707	0.0501	0.1252	0.0388
	0.995	AE	0.2182	0.0602	0.3161	0.0787	0.1818	0.0494
		PL	0.1932	0.0582	0.2471	0.0690	0.1698	0.0503
	0.999	AE	0.4485	0.1086	0.7089	0.1508	0.3561	0.0854
		PL	0.3899	0.1038	0.5482	0.1323	0.3279	0.0852

La Proportional Hazard transform

- La Proportional Hazard (PH) transform pour laquelle $g(x) = x^\alpha$ avec $\alpha \in (0, 1)$, qui affecte un poids **plus fort aux grands quantiles** et qui est telle que $g'(x) \uparrow \infty$ lorsque $x \downarrow 0$.



Ici $\alpha = 2/3$.

- Les conditions d'intégrabilité pour l'existence de la mesure de risque (respectivement pour la validité du Théorème 3) sont : $\gamma < \alpha$ (respectivement $\gamma < \alpha - 1/2$).

Valeur de γ	δ	Estimateur	Fréchet		Burr $\rho = -1$		Burr $\rho = -2$	
			$n = 100$	$n = 300$	$n = 100$	$n = 300$	$n = 100$	$n = 300$
$\gamma = 1/6$	0.99	AE	0.0517	0.0162	0.0618	0.0207	0.0487	0.0141
		PL	0.0395	0.0145	0.0421	0.0157	0.0382	0.0133
	0.995	AE	0.0699	0.0216	0.0848	0.0282	0.0654	0.0184
		PL	0.0534	0.0191	0.0584	0.0215	0.0511	0.0172
	0.999	AE	0.1290	0.0383	0.1612	0.0523	0.1196	0.0311
		PL	0.0993	0.0334	0.1143	0.0406	0.0932	0.0286
$\gamma = 1/5$	0.99	AE	0.0800	0.0272	0.1116	0.0335	0.0756	0.0204
		PL	0.0579	0.0221	0.0670	0.0240	0.0583	0.0186
	0.995	AE	0.1083	0.0363	0.1549	0.0455	0.1010	0.0267
		PL	0.0780	0.0291	0.0941	0.0327	0.0776	0.0239
	0.999	AE	0.2020	0.0644	0.3067	0.0843	0.1829	0.0454
		PL	0.1457	0.0515	0.1916	0.0619	0.1401	0.0397
$\gamma = 1/4$	0.99	AE	0.1920	0.0461	0.2432	0.0678	0.1516	0.0405
		PL	0.1008	0.0347	0.1122	0.0438	0.0927	0.0355
	0.995	AE	0.2669	0.0613	0.3421	0.0921	0.2055	0.0529
		PL	0.1384	0.0453	0.1595	0.0594	0.1242	0.0452
	0.999	AE	0.5454	0.1088	0.7137	0.1727	0.3928	0.0906
		PL	0.2760	0.0796	0.3409	0.1136	0.2330	0.0748

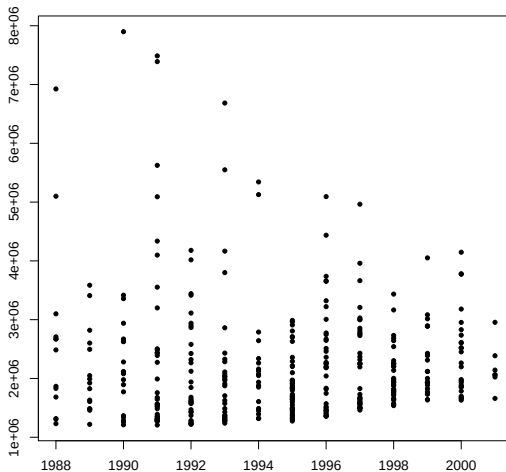
- Nos estimateurs réalisent de bonnes performances pour des valeurs de γ pas trop grandes.
- Comme on pouvait s'y attendre l'estimateur PL semble **plus performant** :
 - cela dans presque tous les cas
 - et plus spécialement lorsque n est **petit**
 - et/ou que la contrainte d'intégrabilité sur γ est **forte** cas : **PH(2/3)**.
- Les MSE augmentent lorsque γ augmente car on a

$$\sqrt{n(1 - \beta_n)}(\hat{\gamma}_{\beta_n} - \gamma) \xrightarrow{d} \mathcal{N}(0, \gamma^2)$$

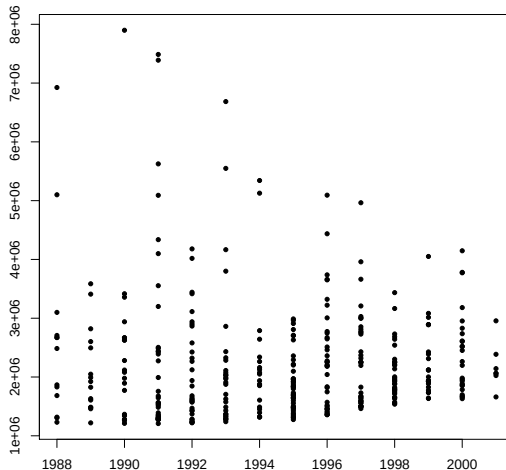
- Les résultats s'améliorent lorsque $|\rho|$ est grand, ce qui est normal car il y a moins de biais.

5 Applications à des données actuarielles

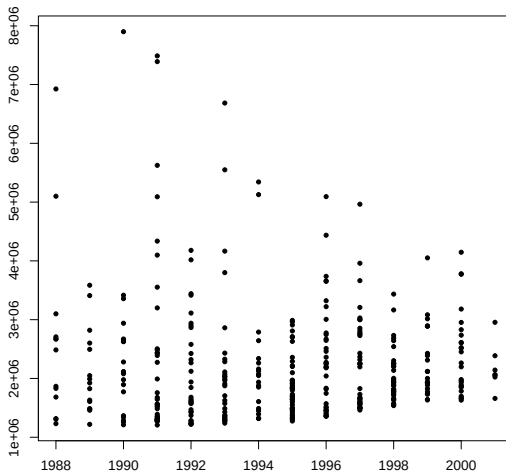
- On s'intéresse ici au jeu de données de réassurance [Secura Belgian](#) sur les sinistres automobile de 1998 à 2001, analysé dans [Vandewalle et Beirlant \(2006\)](#) d'un point de vue extrême.



- Les données consistent en $n = 371$ sinistres qui supérieures ou égales à 1.2 million d'Euros et qui ont été corrigées par l'inflation.



- Notre but ici est de revisiter ce jeu de données et de montrer que l'on trouve des résultats équivalents à ceux de [Vandewalle et Beirlant \(2006\)](#) bien qu'ils aient travaillé dans un contexte différent.



On cherche d'abord à estimer l'indice des valeurs extrêmes γ de ce jeu de données.

La taille de l'échantillon étant faible, on va utiliser l'estimateur de Hill et une **version débiaisée** proposée par Peng (1998) :

$$\hat{\gamma}_{\beta_n}^{BR}(\tau) = \frac{1}{\hat{\rho}_{\beta_1}(\tau)} \hat{\gamma}_{\beta_n} + \left(1 - \frac{1}{\hat{\rho}_{\beta_1}(\tau)}\right) \frac{\hat{\gamma}_{\beta_n}^S}{2\hat{\gamma}_{\beta_n}}$$

où

$$\hat{\gamma}_{\beta_n}^S = \frac{1}{\lceil n(1 - \beta_n) \rceil} \sum_{i=1}^{\lceil n(1 - \beta_n) \rceil} (\log X_{n-i+1,n} - \log X_{n-\lceil n(1 - \beta_n) \rceil,n})^2$$

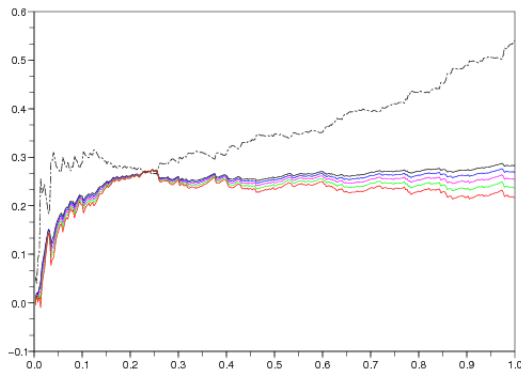
et où $\hat{\rho}_{\beta_1}(\tau)$ est un estimateur de ρ qui dépend d'une fraction $1 - \beta_1$ de l'échantillon différente et d'un paramètre $\tau \geq 0$.

On prend $\beta_1 = 1 - \lceil n^{0.975} \rceil / n \approx 0.0882$, comme recommandé par Caeiro *et al.* (2009).

Cet estimateur est essentiellement une **combinaison** de l'estimateur de Hill et d'un estimateur similaire, et les coefficients de la combinaison sont des estimateurs d'une **fonction simple** de ρ , qui dépendent d'un paramètre additionnel $\tau \geq 0$.

Estimations de γ obtenues par notre procédure de sélection

- Nous utilisons l'estimateur de Hill et des versions débiaisées.
- L'estimateur de Hill semble ici très influencé par le biais.
- Par contre les versions débiaisées donnent des estimations très proches.



- En noir : Hill. En couleurs : $\tau = 1, 3/4, 1/2, 1/4, 0$.

- Nous utilisons l'estimateur de Hill et des versions débiaisées.
- L'estimateur de Hill semble ici très influencé par le biais.
- Par contre les versions débiaisées donnent des estimations très proches.

Estimateurs $\hat{\gamma}_n$	β^*	Estimations de γ
Hill	0.854	0.292
Biais réduit, $\tau = 1$	0.782	0.263
Biais réduit, $\tau = 3/4$	0.792	0.262
Biais-réduit, $\tau = 1/2$	0.792	0.261
Biais réduit, $\tau = 1/4$	0.792	0.260
Biais réduit, $\tau = 0$	0.792	0.258

- On prend pour estimation de γ la médiane des estimations rendues par les versions débiaisées, soit $\hat{\gamma} = 0.261$ pour $\beta^* = 0.792$.
- Dans ce cas on obtient également $\hat{\rho} = -1.064$.

On rappelle que

$$R_g(X) = \int_0^{\infty} g(1 - F(x))dx.$$

- Le cas principal considéré par Vandewalle et Beirlant (2006) est celui du Net Premium au dessus d'un niveau de retention R noté $NP(R)$.
- Il correspond au cas d'une mesure de risque où $g(x) = x$

$$NP(R) = \int_R^{\infty} g(1 - F(x))dx = \int_R^{\infty} [1 - F(x)]dx$$

Que l'on peut donc réécrire

$$NP(q(\beta)) = (1 - \beta)(R_{g,\beta}(X) - \text{VaR}(\beta)) = (1 - \beta)(\text{CTE}(\beta) - \text{VaR}(\beta)) = \text{SP}(\beta)$$

- On peut alors estimer le $NP(R)$ si l'on arrive à faire coïncider le niveau R avec un quantile estimé autrement que l'on arrive à estimer la probabilité $\mathbb{P}(X > R)$.
- Vandewalle et Beirlant (2006) ont pris $R=5$ million d'Euros, qui peut être vu comme une valeur avec une probabilité d'être dépassée de $1 - \beta$ qui vaut environ 0.02.
- Autrement dit R correspond à la **VaR estimée au niveau 98%**.
- Vandewalle et Beirlant (2006) ont estimé le Net Premium à **41.798 milliers d'Euros**.

- On peut alors estimer certaines mesures de risque (l'unité est le millier d'Euros).

δ	Estimateurs	$\widehat{\text{VaR}}$	$\widehat{\text{CTE}}$	$\widehat{\text{SP}}$
0.98	AE	4989 [3505, 6473]	6750 [4742, 8758]	35.220 [24.744, 45.696]
	PL	4989 [3505, 6473]	6864 [4822, 8906]	37.500 [26.346, 48.654]
0.99	AE	5978 [3673, 8283]	8087 [4969, 11205]	21.092 [12.960, 29.224]
	PL	5978 [3673, 8283]	8224 [5053, 11395]	22.459 [13.800, 31.118]
0.995	AE	7163 [3770, 10556]	9690 [5100, 14280]	12.636 [6.6506, 18.621]
	PL	7163 [3770, 10556]	9854 [5186, 14522]	13.455 [7.0817, 19.828]
0.999	AE	10899 [3506, 18291]	14744 [4743, 24745]	3.8452 [1.2371, 6.4533]
	PL	10899 [3506, 18292]	14993 [4823, 25163]	4.0944 [1.3172, 6.8716]

- Autrement dit R correspond à la **VaR estimée au niveau 98%**.
- Vandewalle et Beirlant (2006) ont estimé le Net Premium à **41.798 millions d'Euros**.
- La différence vient du fait que nous avons utilisé un estimateur à biais réduit pour estimer γ alors que eux ont utilisé un estimateur de Hill classique.

6 Conclusions et discussion

Commentaires

- On a construit des versions **extrêmes** des **mesures de risque** de Wang.
- Ces mesures donnent comme **cas particuliers** les **CTM** considérés par El Methni et al. (2014).
- On a constitué un cadre de travail **unifiant** le plus de mesures de risque possibles **au dessus d'un quantile élevé**.

Points clés

- On a ensuite construit deux **estimateurs** de ces mesures de risque :
 - dans le cas **intermédiaire** ;
 - par extrapolation, dans le cas **extrême**.
- Les propriétés asymptotiques de nos estimateurs sont **cohérentes** avec la littérature.
- Nos estimateurs sont **globalement satisfaisants** en pratique pour des $\gamma < 1/2$.
- L'estimateur PL semble **meilleur** dans presque tous les cas et plus spécifiquement lorsque n est **petit** et/ou que la condition d'intégrabilité sur γ est **forte**.

Les travaux de cette présentation ont fait l'objet d'un article de recherche :



El Methni, J., Stupfler, G. (2017) Extreme versions of Wang risk measures and their estimation for heavy-tailed distributions, *Statistica Sinica*, vol. 27, pp. 907–930.

Par la suite

- ⇒ Nous avons étudié deux nouvelles versions de l'estimateur **PL** afin de relaxer les hypothèses restrictives sur γ .
- Ces nouvelles versions mettent moins de poids sur les quantiles les plus extrêmes.
 - Pour cela nous avons considéré une version **Trimmée** et une version **Winsorisée**.

Les travaux de cette présentation ont fait l'objet d'un article de recherche :



El Methni, J., Stupfler, G. (2017) Extreme versions of Wang risk measures and their estimation for heavy-tailed distributions, *Statistica Sinica*, vol. 27, pp. 907–930.

Par la suite

⇒ Nous avons étudié deux nouvelles versions de l'estimateur **PL** afin de relaxer les hypothèses restrictives sur γ .

- Ces nouvelles versions mettent moins de poids sur les quantiles les plus extrêmes.
- Pour cela nous avons considéré une version **Trimmée** et une version **Winsorisée**.



El Methni, J., Stupfler, G. (2017) Improved estimators of extreme Wang distortion risk measures for very heavy-tailed distributions, *Econometrics and Statistics*, to appear.

Vous pouvez trouver les Preprints sur Hal ou sur ma page web.

- Définir et étudier des **estimateurs analogues** pour des lois :
 - à **queue légère**, type loi normale, de Weibull...
 - à **support borné**, comme la loi uniforme, Beta...
- On peut également généraliser ces travaux en présence d'une **covariable**.
- Dans le cas de lois ayant un point terminal fini $\implies$ **estimation de frontière**.
- Proposer des mesures de risque de Wang à biais réduit.
- Essayer de proposer de **nouveaux estimateur de γ** comme combinaison de mesures de risque de Wang.
- Se servir des mesures de risque de Wang afin de détecter la plage de stabilité ou autrement dit le **choix du seuil**.

- Définir et étudier des **estimateurs analogues** pour des lois :
 - à **queue légère**, type loi normale, de Weibull...
 - à **support borné**, comme la loi uniforme, Beta...
- On peut également généraliser ces travaux en présence d'une **covariable**.
- Dans le cas de lois ayant un point terminal fini $\implies$ **estimation de frontière**.
- Proposer des mesures de risque de Wang à biais réduit.
- Essayer de proposer de **nouveaux estimateur de γ** comme combinaison de mesures de risque de Wang.
- Se servir des mesures de risque de Wang afin de détecter la plage de stabilité ou autrement dit le **choix du seuil**.

Merci pour votre attention.



Artzner, P., Delbaen, F., Eber, J.M. & Heath, D. (1999) Coherent measures of risk. *Math. Finance* **9**(3), 203–228.



Caeiro, F., Gomes, M.I. (2006) A new class of estimators of a “scale” second order parameter, *Extremes* **9** : 193–211.



El Methni, J., Gardes, L., Girard, S. (2014) Nonparametric estimation of extreme risks from conditional heavy-tailed distributions, *Scandinavian Journal of Statistics* **41**(4) : 988–1012.



Hill, B.M. (1975) A simple general approach to inference about the tail of a distribution, *Annals of Statistics* **3** : 1163–1174.



Peng, L. (1998). Asymptotically unbiased estimators for the extreme-value index. *Statist. Probab. Lett.* **38**, 107–115.



Vandewalle, B., Beirlant, J. (2006) On univariate extreme value statistics and the estimation of reinsurance premiums, *Insurance : Mathematics and Economics* **38** : 441–459.



Wang, S.S. (1996) Premium calculation by transforming the layer premium density, *ASTIN Bulletin* **26** : 71–92.



Weissman, I. (1978) Estimation of parameters and large quantiles based on the k largest observations, *Journal of the American Statistical Association* **73** : 812–815.