

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE.  
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE.

UNIVERSITE MOULOUD MAMMERI, TIZI-OUZOU.  
FACULTE : DES SCIENCES  
DEPARTEMENT : MATHÉMATIQUES



## MEMOIRE DE MASTER

en  
Mathématiques appliquées

Option: Processus aléatoires et statistique de la décision

Thème

# Mesures de Risque de Distorsion et Optimisation de Portefeuille

Présenté par

MIHOUBI Warda

Devant le jury d'examen composé de :

<i>M<sup>r</sup></i> BOUDIBA Mohand Arezki	Maître de conférence A	U.M.M.T.O	Président
<i>M<sup>r</sup></i> BERKOUN Youcef	Maître de conférence A	U.M.M.T.O	Rapporteur
<i>M<sup>r</sup></i> DOUKI Hamid	Maître Assistant A	U.M.M.T.O	Examineur

Soutenu le 30/ 09 / 2013

# Remerciements

*En préambule de ce mémoire, je tiens à remercier mes parents, mon frère et ma soeur ainsi que mon fiancé qui m'ont toujours soutenue, encouragé et stimulé tout au long de mes études.*

*Je tiens à exprimer toute ma reconnaissance à **Mr Berkoun Youcef** pour l'honneur qu'il m'a fait en assurant la direction et le suivi scientifique et technique du présent mémoire. Je le remercie pour sa grande contribution à l'aboutissement de ce travail, et pour sa disponibilité.*

*Je remercie également **Mr Boudiba Mohand Arezki** pour l'honneur qu'il me fait en acceptant de présider le jury de ce mémoire.*

*Mes remerciements chaleureux s'adressent également à **Mr Douki Hamid** pour avoir accepté d'examiner ce travail.*

*Je remercie tous mes amis ainsi que tous ceux qui ont contribué de près ou de loin à la réalisation de ce modeste mémoire.*

# Table des matières

<b>Introduction</b>	<b>5</b>
<b>1 Notion de Mesures de Risque</b>	<b>7</b>
1.1 Mesure de risque . . . . .	7
1.2 Caractéristiques d'une mesure de risque . . . . .	7
1.2.1 Mesure de risque cohérente . . . . .	8
1.2.2 Mesure de risque monétaire . . . . .	8
1.2.3 Mesure de risque comonotone additive . . . . .	9
1.3 Mesures de risque les plus usuelles . . . . .	14
1.3.1 La Value at Risk (VaR) . . . . .	14
1.3.2 La Tail-Value-at-Risk ou TVaR . . . . .	15
1.3.3 La Conditional Tail Expectation ou CTE . . . . .	16
1.3.4 L'expected shortfall ou ES . . . . .	16
1.4 Lien entre ces mesures de risque . . . . .	16
<b>2 Mesures de risque de distorsion</b>	<b>19</b>
2.1 Intégrale de Stieltjes . . . . .	19
2.2 Intégrale de Choquet . . . . .	19
2.2.1 Propriétés de l'intégrale de Choquet . . . . .	20
2.3 Fonction de distorsion . . . . .	20
2.4 Mesures de risque de distorsion . . . . .	21
2.4.1 Exemple de mesure de risque de distorsion . . . . .	21
2.5 Propriétés des mesures de risque de distorsion . . . . .	26
<b>3 Estimation d'une mesure de risque de distorsion</b>	<b>31</b>
3.1 Estimation empirique d'une mesure de risque de distorsion . . . . .	31
3.2 Statistiques d'ordre . . . . .	31
3.2.1 Densité conjointe de $n$ statistiques d'ordre . . . . .	31
3.2.2 Densité de la $k^{ième}$ statistique d'ordre . . . . .	32
3.2.3 Densité conjointe de deux statistiques d'ordre . . . . .	34
3.2.4 Quantile empirique . . . . .	34
3.3 Estimateur Linéaire . . . . .	35
3.3.1 Propriétés des L-estimateurs . . . . .	36
3.4 Estimation empirique d'une mesure de risque de distorsion . . . . .	41
3.4.1 Estimateur empirique d'une mesure de risque de distorsion . . . . .	41
3.4.2 Propriétés élémentaires de l'estimateur empirique d'une mesure de risque de distorsion . . . . .	42
3.4.3 Propriétés asymptotiques de l'estimateur empirique d'une mesure de risque de distorsion . . . . .	43

3.5	Estimation de la Variance de l'estimateur empirique d'une mesure de risque distorsion . . . . .	44
3.5.1	La méthode bootstrap . . . . .	44
3.5.2	Méthode non paramétrique delta . . . . .	46
3.6	Approximation de l'estimateur empirique par la méthode de Monte Carlo . . . . .	49
3.6.1	Description de la méthode de Monte Carlo . . . . .	49
3.6.2	Convergence de la méthode de Monte Carlo . . . . .	50
3.7	Simulation . . . . .	51
3.7.1	Distribution de Pareto . . . . .	51
3.7.2	Estimation par la méthode des noyaux . . . . .	52
<b>4</b>	<b>Application: Optimisation de portefeuille</b>	<b>55</b>
4.1	Définition de portefeuille . . . . .	55
4.2	Diversification de portefeuille . . . . .	55
4.3	Optimisation de portefeuille . . . . .	55
4.4	Reformulation de la Tail-Value-at-Risk (TVaR) . . . . .	56
4.5	Discrétisation et linéarisation du problème d'optimisation . . . . .	57
4.5.1	Discrétisation . . . . .	57
4.5.2	Linéarisation . . . . .	58
4.6	Résolution numérique . . . . .	59
	<b>Conclusion</b>	<b>61</b>
	<b>Bibliographie</b>	<b>63</b>
	<b>Annexe</b>	<b>67</b>

# Introduction

Les institutions financières et les assurances sont confrontées à des risques multiples (faillite bancaire, inondations, ...). Il est donc impératif d'identifier ces risques et de les analyser le plus finement possible. Mesurer le risque s'avère donc de nos jours inévitable puisque cette information est nécessaire dans le cadre de la détermination du niveau des réserves et des fonds propres dont doit disposer une banque ou une compagnie d'assurances afin qu'elle soit solvable.

Les entreprises du secteur de la banque et de l'assurance gèrent aujourd'hui des risques importants et de natures diverses. Elles sont soumises à des obligations réglementaires de maîtrise du risque, et cherchent également par elles-mêmes à éviter la faillite. Un des outils intervenant dans cette gestion compliquée est la mesure de risque, qui doit permettre de fournir un indicateur pertinent permettant de générer du profit en limitant le risque de subir des pertes trop importantes. Il est donc important de pouvoir décider du choix de cette mesure, et de savoir comment elle peut être utilisée dans un programme général de maximisation du profit. Il existe de nombreuses méthodes permettant de quantifier le risque. En 1996 le comité de Bâle a proposé la Value-at-Risk (VaR), cette mesure de risque présente quelques inconvénients car elle n'est pas sous additive donc elle n'est pas cohérente. Certains auteurs ont proposé de remplacer la VaR par des mesures de risque alternatives telles que : l'Expected Shortfall (ES) et la Tail-Value-at-Risk (TVaR).

La VaR et ses alternatives sont étroitement liées aux mesures de risque de distorsion introduite par Wang(1995) qui sont une des plus importantes familles des mesures de risque.

La résolution du problème de cohérence va permettre à une institution financière de pouvoir faire des choix mathématiques cohérents en ce qui concerne ses investissements, la première application importante est la gestion de portefeuilles. En investissant d'une façon intelligente, l'institution cherche à minimiser son risque selon, bien sûr, sa propre mesure de risque. Le problème d'optimisation de portefeuille sous contrainte de mesure de risque est sujet bien connu notamment en gestion d'actifs, où il trouve bon nombre d'applications. Ce travail s'articule autour de 4 chapitres.

Le premier chapitre définit les différentes mesures de risque. Une attention particulière est portée aux caractéristiques particulières que l'on peut attendre des mesures de risque pour pouvoir être utilisées à des fins de solvabilité. Nous exposons le cas des mesures de risque cohérentes qui ont été introduites pour remédier à certains défauts de la VaR.

Le deuxième chapitre présente les mesures de risque de distorsion ainsi que leurs propriétés.

Au troisième chapitre, on s'intéresse à l'estimateur empirique d'une mesure de risque de distorsion basé sur les L-estimateur. Nous donnons par la suite les propriétés probabilistes de cet

estimateur, ainsi que l'estimation de sa variance par deux méthodes à savoir : la méthode basée sur le Bootstrap et la méthode non paramétrique delta .

Dans la dernière partie, nous donnerons une application importante dans la gestion de portefeuilles, plus particulièrement l'optimisation de portefeuille qui sera basée sur une méthode numérique.

# Chapitre 1

## Notion de Mesures de Risque

Nous allons aborder dans ce chapitre le concept de mesure de risque. Ensuite nous allons donner les mesures de risque les plus usuelles ainsi que leurs propriétés.

### 1.1 Mesure de risque

Soit  $\mathcal{V}$  l'ensemble des variables aléatoires à valeurs réelles.

**Définition 1.** *Mesure de risque*

*Une mesure de risque  $\rho$  est une fonction définie sur l'espace des variables aléatoires, et prenant ses valeurs dans  $\mathbb{R}$ .*

$$\rho : \mathcal{V} \rightarrow \mathbb{R}$$

### 1.2 Caractéristiques d'une mesure de risque

La définition d'une mesure de risque est très générale puisque toute fonctionnelle réelle positive d'une variable aléatoire peut être considérée comme étant une mesure de risque. Aussi, en pratique, on exige de telles mesures qu'elles disposent de propriétés mathématiques dont la transcription conceptuelle permette de les juger. En pratique, on exige fréquemment qu'une mesure de risque  $\rho$  possède une partie des caractéristiques suivantes :

**1. Invariance en loi :**

$$\forall X, Y \in \mathcal{V}, \quad X \stackrel{\ell}{=} Y \text{ implique } \rho(X) = \rho(Y)$$

**2. Invariance par translation:**

$$\forall X \in \mathcal{V}, \forall \alpha \in \mathbb{R}, \quad \rho(X + \alpha) = \rho(X) + \alpha$$

**3. Sous-additivité:**

$$\forall X, Y \in \mathcal{V}, \quad \rho(X + Y) \leq \rho(X) + \rho(Y)$$

#### 4. Homogénéité positive:

$$\forall \lambda \geq 0, \forall X \in \mathcal{V}, \quad \rho(\lambda X) = \lambda \rho(X)$$

#### 5. Monotonie:

$$\forall X, Y \in \mathcal{V} \quad X \geq Y \text{ implique } \rho(X) \geq \rho(Y)$$

#### 6. Convexité:

$$\forall \beta \in [0,1], \forall X, Y \in \mathcal{V}, \quad \rho(\beta X + (1 - \beta)Y) \leq \beta \rho(X) + (1 - \beta) \rho(Y)$$

#### Remarque 1.

- La monotonie signifie que si une variable aléatoire génère moins de pertes qu'une autre, alors son critère de risque est inférieur.
- L'invariance par translation signifie que les éléments déterministes additifs ne sont pas sujets à la mesure de risque. La propriété implique :  $\rho(X - \rho(X)) = 0$  Ainsi,  $\rho(X)$  est ce qu'il faut retirer à  $X$  pour annuler son risque sous la mesure  $\rho$ .
- L'homogénéité positive signifie que les éléments déterministes multiplicatifs n'influencent pas la mesure de risque, l'homogénéité implique :  $\rho(0) = 0$

### 1.2.1 Mesure de risque cohérente

La notion de mesure de risque cohérente a été introduite par Artzner et al en 1999.

**Définition 2.** Une mesure de risque est dite cohérente si elle est :

- invariante par translation
- sous-additive
- homogène
- monotone

### 1.2.2 Mesure de risque monétaire

**Définition 3.** une mesure de risque est dite monétaire si elle est :

- invariante par translation
- monotone

#### Remarque 2.

Une mesure de risque est dite convexe si elle est :

- monétaire
- convexe

#### Proposition 1.

• Si  $\rho$  est une mesure de risque monétaire et homogène, alors la convexité et la sous-additivité sont des notions équivalentes.

**Preuve 1.**

Si  $\rho$  est homogène c'est à dire que:

$$\forall \lambda \geq 0, \forall X \in \mathcal{V}, \quad \rho(\lambda X) = \lambda \rho(X)$$

On a:

$$\begin{aligned} \rho(\beta X + (1 - \beta)Y) &\leq \beta \rho(X) + (1 - \beta)\rho(Y) \quad (\text{la convexité}) \\ &= \rho(\beta X) + \rho((1 - \beta)Y) \quad (\text{l'homogénéité positive}). \end{aligned}$$

Donc

$$\rho(\beta X + (1 - \beta)Y) \leq \rho(\beta X) + \rho((1 - \beta)Y)$$

D'où la sous-additivité de la mesure  $\rho$ .

**1.2.3 Mesure de risque comonotone additive**

Rappelons qu'un vecteur aléatoire  $(X_1, X_2)$ , de fonctions de répartition marginales  $F_1, F_2$ , est un vecteur comonotone s'il existe une variable aléatoire  $U$  de loi uniforme sur  $[0; 1]$  telle que  $(X_1, X_2)$  a la même loi que  $(F_1^{-1}(U), F_2^{-1}(U))$

**Définition 4.** (*Mesure de risque comonotone*) [47]

On appelle mesure de risque comonotone additive toute mesure de risque  $\rho$  telle que :  $\rho(X_1 + X_2) = \rho(X_1) + \rho(X_2)$  pour tout vecteur comonotone  $(X_1, X_2)$ .

**Définition 5.** (*Inverse généralisé d'une fonction de répartition*)

Soit  $X$  une variable aléatoire de fonction de répartition  $F$ . On appelle inverse généralisé de  $F$  la fonction :

$$\begin{aligned} F_X^{-1}(\alpha) &= \inf\{x \in \mathbb{R} / F(x) \geq \alpha\}; \quad 0 < \alpha < 1 \\ &= \sup\{x \in \mathbb{R} / F(x) \leq \alpha\} = Q(\alpha) \end{aligned}$$

**Définition 6.** (*Fonction de répartition empirique*)

La fonction de répartition empirique de l'échantillon  $X_1, X_2, \dots, X_n$  est définie sur  $\mathbb{R}$  par :

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{X_i \leq x\}}$$

$$\text{Où } \mathbb{I}_A = \begin{cases} 1, & \text{Si } x \in A \\ 0, & \text{Sinon} \end{cases}.$$

On note que  $F_n(x)$  est l'estimateur sans biais de  $F(x)$  ( $\mathbb{E}[F_n(x)] = F(x)$ ).

Si les  $X_i$  sont ordonnées par valeurs croissantes  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ :

$\forall x \in \mathbb{R}$ , on a:

$$F_n(x) = \begin{cases} 0 & \text{Si } x < x_{(1)} \\ \frac{i-1}{n} & \text{Si } x_{(i-1)} \leq x < x_{(i)} \\ 1 & \text{Si } x \geq x_{(n)}. \end{cases}$$

**Exemple 1.**

Soit un échantillon de la variable aléatoire  $X$  donné par:  $X = (4, -8, 11, 28, -5, 3)$  et de fonction de répartition  $F(x)$ . pour avoir sa fonction de répartition empirique on doit d'abord trier l'échantillon par ordre croissant:  $-8 < -5 < 3 < 4 < 11 < 28$ .

$$F_n(x) = \begin{cases} 0, & x < -8; \\ \frac{1}{6}, & \text{si } -8 < x < -5 \\ \frac{2}{6}, & \text{si } -5 < x < 3 \\ \frac{3}{6}, & \text{si } 3 < x < 4 \\ \frac{4}{6}, & \text{si } 4 < x < 11 \\ \frac{5}{6}, & \text{si } 11 < x < 28 \\ 1, & \text{si } x \geq 28. \end{cases}$$

Voici le graphe qui représente cette fonction de répartition empirique:

**Proposition 2.** (Propriétés élémentaires de la fonction de répartition empirique)

- **Biais de l'estimateur  $F_n(x)$ :**

$F_n(x)$  est un estimateur sans biais de  $F(x)$ .

soit la suite  $(F_n)$  définie par

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{x_i \leq x}, \quad x \in \mathbb{R} \text{ la fonction de répartition empirique}$$

$$Y_i = I_{X_i \leq x} = \begin{cases} 1, & x_i \leq x \\ 0, & \text{Sinon.} \end{cases}$$

$$P(Y_i = 1) = P(X_i \leq x) = F(x)$$

$\sum_{i=1}^n (Y_i)$  suit une loi binomiale  $B(n, p)$

$$E(F_n(x)) = E\left(\frac{1}{n} \sum_{i=1}^n Y_i\right) = \frac{np}{n} = p = F(x)$$

Donc, pour tout point  $x$ ,  $F_n(x)$  est un estimateur sans biais de  $F(x)$ .

- **Variance de l'estimateur  $F_n(x)$ :**

Il est facile de montrer que, pour tout  $x$ , la variance de l'estimateur  $F_n(x)$  est donnée par:

$$V(F_n(x)) = \frac{1}{n^2} V\left(\sum_{i=1}^n y_i\right) = \frac{np(1-p)}{n^2} = \frac{F(x)(1-F(x))}{n} \rightarrow 0 \text{ quand } n \rightarrow \infty$$

Il est simple de vérifier que  $F_n(x) \xrightarrow{P.S} F(X)$

Dans ce qui suit, on va citer les théorèmes fondamentaux qui sont liés à cette fonction .

**Proposition 3.** (Propriétés asymptotiques de  $F_n$ )

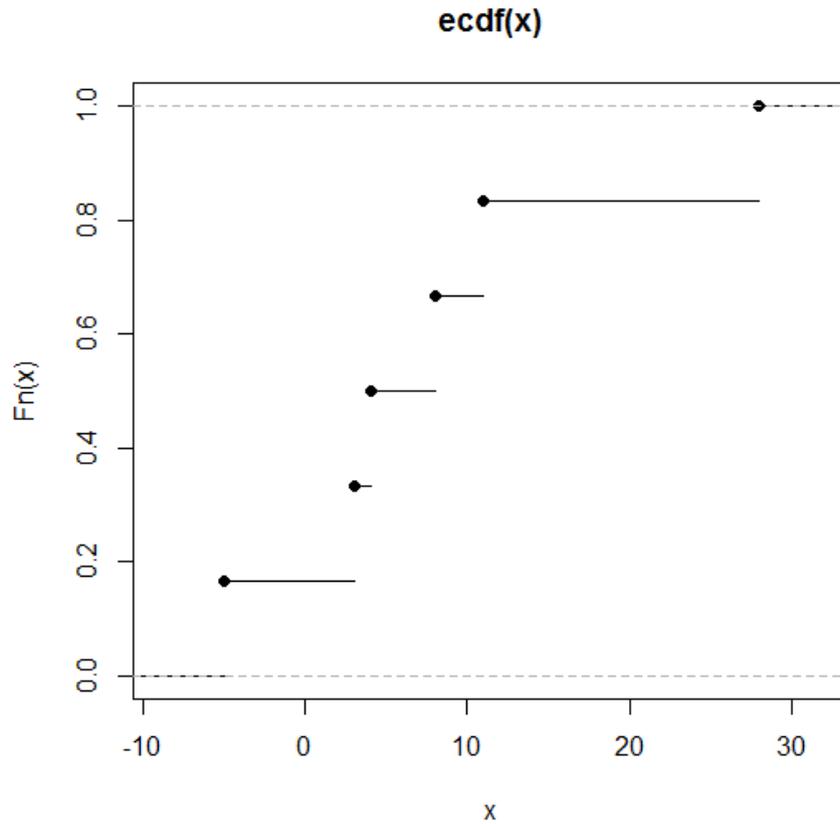


FIG. 1.1 – Fonction de répartition empirique

La fonction de répartition empirique  $F_n(x)$  converge presque sûrement vers  $F(X)$  par:

- **La loi forte des grands nombres:**

Soit  $(X_i)_{i \geq 1}$  une suite de réalisations de la variable aléatoire  $X$ . On suppose que  $\mathbb{E}(|X|) < +\infty$ , alors pour presque tout  $\omega$

$$\mathbb{E}(X) = \lim_{n \rightarrow +\infty} \frac{1}{n} (X_1(\omega), \dots, X_n(\omega))$$

- **Le théorème central limite:**

• Soit la v.a  $S_n$  résultant de la somme de  $n$  variables aléatoire indépendantes et de même loi, telle que  $\mathbb{E}(X^2) < \infty$ :

$$Z_n = \frac{S_n - nm}{\sigma\sqrt{n}} \rightarrow^L N(0,1), \quad \text{quant, } n \rightarrow \infty$$

• En ce qui concerne la fonction de répartition empirique on a: soit  $F_n$  la fonction de répartition empirique définie comme suit:

$$F_n = \frac{1}{n} \sum_{i=1}^n 1_{(X_i \leq x)} = \frac{1}{n} \sum_{i=1}^n Y_i$$

donc  $F_n(x) = \bar{Y}$  p.s

selon le TLC

$$\frac{\bar{Y} - F(x)}{\sqrt{\frac{F(x)(1-F(x))}{n}}} \rightarrow^L N(0,1)$$

donc

$$\bar{Y} \rightarrow^L N\left(F(x), \sqrt{\frac{F(x)(1-F(x))}{n}}\right)$$

• **Théorème de Glivenko Cantelli:**

Soit  $(X_1, X_2 \dots X_n)$  un  $n$  échantillon issu d'une population de fonction de répartition  $F$ .

Soit  $F_n$  la fonction de répartition empirique de cet échantillon alors.

$$\Delta_n = \sup_{-\infty < x < +\infty} |F_n(x) - F(x)| \rightarrow_{n \rightarrow \infty} 0$$

**Exemple:**

nous allons illustrer ce théorème par les graphes suivants:

Le premier graphe pour  $n=100, 1000$  et  $10000$ .

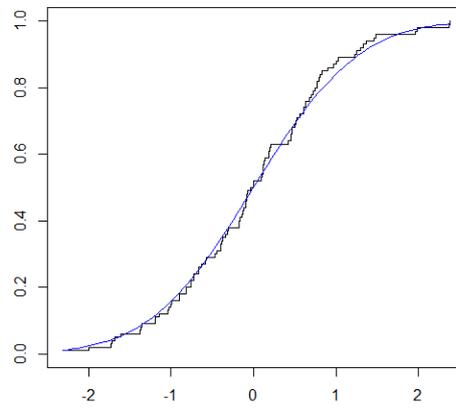


FIG. 1.2 – Fonction de répartition empirique d'un échantillon de loi normale de taille  $n=100$

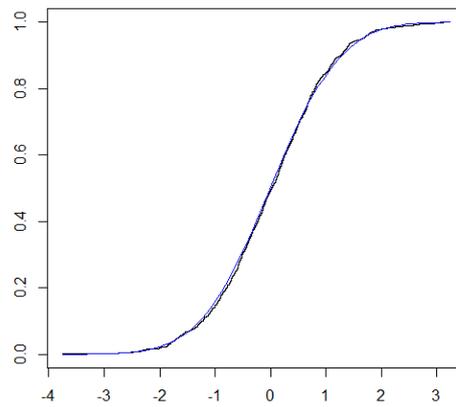


FIG. 1.3 – Fonction de répartition empirique d'un échantillon de loi normale de taille  $n=1000$

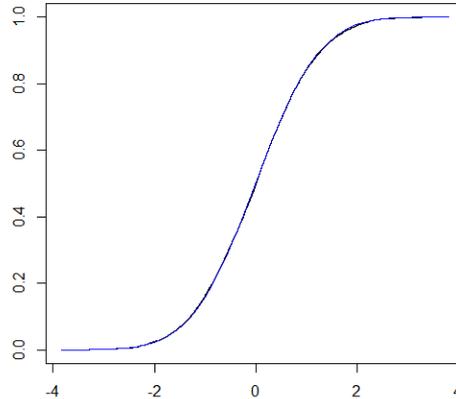


FIG. 1.4 – Fonction de répartition empirique d'un échantillon de loi normale de taille n=10000

## 1.3 Mesures de risque les plus usuelles

### 1.3.1 La Value at Risk (VaR)

La VaR ou la value at risk, est l'un des derniers nés des mesures de risque et des plus largement utilisés de nos jours, est devenue courante dans les banques suite à la réglementation définie par les comités de Bâle successifs (1974).

Elle correspond à la notion de Sinistre Maximum Probable (SMP), bien connue dans l'assurance non-vie. Elle est égale à la perte maximale que peut subir une banque ou une compagnie d'assurance dans des conditions normales de marché, sur une période de temps donnée avec une probabilité  $\alpha$

#### Définition 7.

On appelle Value-at-Risk de niveau  $\alpha \in [0,1]$ , notée  $VaR_\alpha(X)$  le quantile de niveau  $\alpha$  i.e.,

$$\begin{aligned} VaR_\alpha(X) &= \inf\{x, \mathbb{P}(X \leq x) \geq \alpha\} \\ &= F_X^{-1}(\alpha) \end{aligned}$$

#### Proposition 4.

La Value-at-Risk vérifie les propriétés suivantes:

- *monotonie*

$$\begin{aligned} X &\leq Y \\ \mathbb{P}(X > t) &\leq \mathbb{P}(Y > t) \forall t \\ \inf\{t, \mathbb{P}(X > t) \leq 1 - \alpha\} &\leq \inf\{t, \mathbb{P}(Y > t) \leq 1 - \alpha\} \\ \inf\{t, \mathbb{P}(X \leq t) \geq \alpha\} &\leq \inf\{t, \mathbb{P}(Y \leq t) \geq \alpha\} \\ VaR_\alpha(X) &\leq VaR_\alpha(Y) \end{aligned}$$

- *Homogénéité positive:*  $\forall \lambda \geq 0$

$$\begin{aligned} VaR_\alpha(\lambda X) &= \inf\{t, \mathbb{P}(\lambda X \leq t) \geq \alpha\} \\ &= \inf\{\lambda t', \mathbb{P}(X \leq t') \geq \alpha\} \\ &= \lambda \inf\{t', \mathbb{P}(X \leq t') \geq \alpha\} \\ &= \lambda VaR_\alpha(X) \end{aligned}$$

- *Invariance par translation:*  $\forall c \in \mathbb{R}$

$$\begin{aligned} VaR_\alpha(X + c) &= \inf\{t, \mathbb{P}(X + c \leq t) \geq \alpha\} \\ &= \inf\{t' + c, \mathbb{P}(X \leq t') \geq \alpha\} \\ &= \inf\{t', \mathbb{P}(X \leq t') \geq \alpha\} + c \\ &= VaR_\alpha(X) + c \end{aligned}$$

- *Additivité comonotone:* (47)

La Value at Risk a un certain nombre de bonnes propriétés mathématiques parmi les quelles le fait que pour toute fonction  $g$  croissante et continue à gauche, on a :  $VaR_\alpha(g(X)) = g(VaR_\alpha(X))$ . Il découle de cette propriété en prenant  $g = F_1^{-1} + F_2^{-1}$  et  $X = U$ , que la VaR est comonotone additive puisque pour tout  $\alpha \in ]0, 1[$ , on a :

$$VaR_\alpha((F_1^{-1} + F_2^{-1})(U)) = (F_1^{-1} + F_2^{-1})(VaR_\alpha(U)).$$

### Proposition 5.

La VaR n'est pas cohérente. Plus spécifiquement, elle n'est pas convexe car elle n'est pas sous-additive.

#### Preuve:

Nous allons le démontrer par le contre-exemple suivant:

Soit deux variables aléatoires mesurant des pertes  $X_i, i = 1, 2$ .

Elles admettent deux réalisations équiprobables  $X_i = 1$  ou  $X_i = 2$ . On a donc  $VaR_{0.5}(X_i) = 1$ .

Leur réalisation conjointe admet deux scénarios équiprobables, soit  $(X_1, X_2) = (1, 2)$ , soit  $(X_1, X_2) = (2, 1)$ , tels que  $X_1 + X_2 = 3$  pour tous les scénarios.

On a donc  $VaR_{0.5}(0.5(X_1 + X_2)) = 1.5$ .

et donc

$1.5 = VaR_{0.5}(0.5X_1 + 0.5X_2) > 0.5VaR_{0.5}(X_1) + 0.5VaR_{0.5}(X_2) = 1$  ce qui montre que la VaR n'est pas sous-additive.

Pour remédier à ces défauts de la VaR, d'autres mesures ont été proposées. Nous en présentons trois très proches, dont l'idée commune est de quantifier la perte lorsque la VaR est dépassée.

### 1.3.2 La Tail-Value-at-Risk ou TVaR

#### Définition 8.

La Tail-Value-at-Risk au niveau  $\alpha$  est définie par:

$$TVaR_\alpha(X) = \frac{1}{1 - \alpha} \int_\alpha^1 VaR_\mu(X) d\mu$$

La Tail Value at Risk, ou la TVaR est la moyenne des VaR de niveau supérieur à  $\alpha$ .

**Remarque:**

Nous avons:

- $TVaR_0(X) = E[X]$ .

- $TVaR_\alpha(X) = \frac{1}{1-\alpha} \{E[X] - \int_0^\alpha VaR_\zeta(X) d\zeta\}$ . et comme  $VaR_\alpha(X)$  est une fonction croissante, on en déduit que la TVaR est une fonction croissante en  $\alpha$ . D'où  $TVaR_\alpha(X) \geq TVaR_0(X) = E[X]$ .

- La TVaR est invariante par translation et homogène car la VaR l'est.

**1.3.3 La Conditional Tail Expectation ou CTE****Définition 9.**

La Conditional Tail Expectation au niveau  $\alpha$ , notée  $CTE[X; \alpha]$  est:

$$CTE[X; \alpha] = \mathbb{E}[X | X > VaR_\alpha(X)]$$

La Conditional Tail Expectation est la perte attendue sachant que la VaR au niveau  $\alpha$  est dépassée.

**1.3.4 L'expected shortfall ou ES****Définition 10.**

L'Expected shortfall au niveau  $\alpha$ , noté  $ES[X; \alpha]$  est:

$$ES[X; \alpha] = \mathbb{E}[(X - VaR_\alpha(X))_+]$$

**Proposition 6.**

L'Expected shortfall est une mesure de risque cohérente c'est à dire:

- Invariante par translation.
- Homogène positive.
- Monotone.
- Sous-additive.

**1.4 Lien entre ces mesures de risque**

Ces mesures sont très liées, comme le montre la proposition suivante:

**Proposition 7. (12)**

Pour tout  $\alpha \in [0,1]$ , les identités suivantes sont vérifiées:

$$TVaR_\alpha(X) = VaR_\alpha(X) + \frac{1}{1-\alpha} ES[X; \alpha] \tag{1.1}$$

$$CTE[X; \alpha] = VaR_\alpha(X) + \frac{1}{F_X(VaR_\alpha(X))} ES[X; \alpha] \tag{1.2}$$

D'après ces deux identités on a:

$$CTE_\alpha(X) = TVaR_{F_X(VaR_\alpha(X))}(X), \quad \forall \alpha \in [0,1] \quad (1.3)$$

**Proposition 8.** Pour des risques dont la fonction de répartition est continue on a:

$$CTE_\alpha(X) = TVaR_\alpha(X), \quad \alpha \in [0,1]$$

**Preuve:**

Il suffit de noter que si  $F_X$  est continue alors

$$F_X(VaR_\alpha(X)) = \alpha$$

$$\begin{aligned} CTE_\alpha(X) &= E[X|X \geq VaR_\alpha(X)] \\ &= \int_{\Omega} x P_{\{X \geq VaR_\alpha(X)\}}(dw) \\ &= \int_{\Omega} x \frac{P[dw, X \geq VaR_\alpha(X)]}{P[X \geq VaR_\alpha(X)]} \\ &= \frac{1}{P[X \geq VaR_\alpha(X)]} \int_{\{X \geq VaR_\alpha(X)\}} x dP \\ &= \frac{1}{1-\alpha} \int_{VaR_\alpha(X)}^{+\infty} x f(x) dx \\ &= \frac{1}{1-\alpha} \int_{\alpha}^1 F_X^{-1}(u) du \quad (F_X(x) = u) \\ &= \frac{1}{1-\alpha} \int_{\alpha}^1 VaR_u(X) du \end{aligned}$$

**Proposition 9.**

Si la TVaR concide avec la CTE alors:

1. la TVaR est sous-additive.
2. la TVaR est monotone.

**Preuve:**

1. Si les risques sont continus, la TVaR et la CTE concident.

Donc:

$$\begin{aligned} TVaR_\alpha(X+Y) &= CTE[X+Y, \alpha] \\ &= \mathbb{E}[X|X+Y > VaR_\alpha(X+Y)] + \mathbb{E}[Y|X+Y > VaR_\alpha(X+Y)] \\ &\leq \mathbb{E}[X|X > VaR_\alpha(X)] + \mathbb{E}[Y|Y > VaR_\alpha(Y)] \\ &= TVaR_\alpha(X) + TVaR_\alpha(Y) \end{aligned}$$

D'ou la sous-additivité de La Tail-Value-at-Risk

2. La TVaR est monotone, puisque lorsque  $X \leq Y$

$$\begin{aligned} TVaR_\alpha(Y) &= \mathbb{E}[Y|Y > VaR_\alpha(Y)] \\ &\geq \mathbb{E}[Y|X > VaR_\alpha(Y)] \\ &\geq \mathbb{E}[X|X > VaR_\alpha(Y)] \\ &= TVaR_\alpha(X) \end{aligned}$$

**Remarque:**

La Tail-Value-at-Risk ou TVaR est cohérente pour les risques dont la fonction de répartition est continue.

**Proposition 10.** *La TVaR est la plus petite mesure de risque cohérente majorant la VaR, i.e.,*

$$TVaR_\alpha(X) \geq VaR_\alpha(X)$$

**Preuve:**

$$\begin{aligned} TVaR_\alpha(X) &= CTE_\alpha(X) \\ &= E[X/X > VaR_\alpha(X)] \\ &= E[X + VaR_\alpha(X) - VaR_\alpha(X)/X > VaR_\alpha(X)] \\ &= VaR_\alpha(X) + E[X - VaR_\alpha(X)/X > VaR_\alpha(X)] \\ &\geq VaR_\alpha(X) \end{aligned}$$

## Chapitre 2

# Mesures de risque de distorsion

**Définition 11.** (Fonction à variation bornée)

Soit  $f$  une fonction définie de  $[a,b]$  vers  $\mathbb{R}$  et  $S$  l'ensemble de toutes les subdivisions de  $[a,b]$ . Pour chaque subdivision  $\sigma = (a = x_0, x_1, \dots, x_n = b) \in S([a,b])$ , on définit  $V(f, \sigma)$  par:

$$V(f, \sigma) = \sum_{i=1}^n |f(x_i) - f(x_{i-1})|$$

On appelle variation totale de  $f$  la valeur  $v_a^b \in \overline{\mathbb{R}}$  définie par:

$$V_a^b(f) = \sup_{\sigma \in S([a,b])} V(f, \sigma)$$

On dit que  $f$  est à variation bornée si  $v_a^b(f)$  est fini.

### 2.1 Intégrale de Stieltjes

L'intégrale de Stieltjes constitue une généralisation de l'intégrale de Riemann. En effet, considérons deux fonctions réelles bornées et définies sur un intervalle fermé  $[a,b]$ , ainsi qu'une subdivision  $a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$  de cet intervalle. L'intégrale de Stieltjes de la fonction  $f$  par rapport à  $g$  est:

$$\int_a^b f(x) dg(x)$$

Si les fonctions  $f$  et  $g$  possèdent un point de discontinuité en commun, alors l'intégrale n'existe pas. Cependant, si  $f$  est continue et  $g' = dg/dx$ , alors

$$\int_a^b f(x) dg(x) = \int_a^b f(x) g'(x) dx$$

Plus généralement, si  $f$  est continue et  $g$  à variation bornée, cette intégrale est bien définie.

### 2.2 Intégrale de Choquet

**Définition 12.** Soit  $\Psi$  un espace topologique et  $P(\Psi)$  l'ensemble de tous les sous-ensembles de  $\Psi$ . On appelle mesure non-additive dans  $(\Psi, P(\Psi))$  la fonction  $\mu : P(\Psi) \rightarrow [0,1]$  qui satisfait:

$$i) \mu(\emptyset) = 0$$

$$ii) \forall A, B \in P(\Psi), A \subset B, \text{ alors, } \mu(A) \leq \mu(B)$$

**Définition 13.** Soit  $f$  une fonction continue bornée de  $\Psi$  vers  $\mathbb{R}$ . L'intégrale de Choquet de  $f$  par rapport à  $\mu$  est définie par:

$$\int_{x \in \Psi} f(x) d\mu = \int_0^{+\infty} \mu(f^{-1}]t, +\infty[) dt + \int_{-\infty}^0 [\mu(f^{-1}]t, +\infty[) - \mu(\Psi)] dt \quad (2.1)$$

Où les deux dernières intégrales sont des intégrales impropres de Riemann

### 2.2.1 Propriétés de l'intégrale de Choquet

1. L'intégrale de Choquet n'est pas additive, en général

$$\int_{x \in \Psi} f d\mu + \int_{x \in \Psi} g d\mu \neq \int_{x \in \Psi} (f + g) d\mu$$

2. L'intégrale de Choquet est positivement homogène :

$$\int_{x \in \Psi} \alpha f(x) d\mu = \alpha \int_{x \in \Psi} f(x) d\mu$$

pour tout  $\alpha \in \mathbb{R}^+$

3. Soit  $\mu$  une mesure non-additive convexe. L'intégrale de Choquet par rapport à  $\mu$  est sur-additive :

$$\int_{x \in \Psi} f(x) + g(x) d\mu \geq \int_{x \in \Psi} f(x) d\mu + \int_{x \in \Psi} g(x) d\mu$$

En conséquence, l'intégrale de Choquet est concave : pour tout  $\alpha, \beta \geq 0$  avec  $\alpha + \beta = 1$

$$\int_{x \in \Psi} \alpha f(x) + \beta g(x) d\mu = \alpha \int_{x \in \Psi} f(x) d\mu + \beta \int_{x \in \Psi} g(x) d\mu$$

4. Soit  $\mu$  une mesure non-additive concave. L'intégrale de Choquet par rapport à  $\mu$  est sous-additive :

$$\int_{x \in \Psi} f(x) + g(x) d\mu \leq \int_{x \in \Psi} f(x) d\mu + \int_{x \in \Psi} g(x) d\mu$$

5. L'intégrale de Choquet sur les fonctions positives est croissante : si  $\mu \leq \mu'$ , alors pour toute fonction  $f$  continue bornée de  $\Psi$  dans  $\mathbb{R}^+$

$$\int_{x \in \Psi} f(x) d\mu \leq \int_{x \in \Psi} f d\mu'$$

## 2.3 Fonction de distorsion

**Définition 14.** On appelle fonction de distorsion toute fonction croissante  $g : [0,1] \longrightarrow [0,1]$  telle que  $g(0) = 0$  et  $g(1) = 1$ .

## 2.4 Mesures de risque de distorsion

**Définition 15.** On appelle mesure de risque de distorsion la quantité:

$$R(X) = \int_0^{+\infty} g(1 - F_X(x))dx = \int_0^1 F_X^{-1}(1 - \alpha)dg(\alpha) = \int_0^1 VaR_{1-\alpha}(X)dg(\alpha)$$

Où  $g$  est une fonction de répartition sur  $[0; 1]$ , appelée fonction de distorsion.

**Proposition 11.**

Il est possible de réécrire cette expression sous la forme:

$$R(X) = \int_0^{+\infty} g(1 - F(x))dx - \int_{-\infty}^0 [1 - g(1 - F(x))]dx$$

**Preuve:** On a:

$$R(X) = \int_0^1 F_X^{-1}(x)dH(x) = - \int_{-\infty}^0 H(F_X^{-1}(x))dx + \int_0^{+\infty} [1 - H(F_X^{-1}(x))]dx$$

Où  $H(\mu) = 1 - g(1 - \mu)$ . Si on prend  $\bar{F}_X(x) = 1 - F_X(x) = P(X > x)$  on aura:

$$R(X) = \int_0^{+\infty} g(\bar{F}_X(x))dx - \int_{-\infty}^0 [1 - g(\bar{F}_X(x))]dx$$

$$R(X) = \int_0^{+\infty} g(1 - F(x))dx - \int_{-\infty}^0 [1 - g(1 - F(x))]dx$$

**Remarque:**

Dans le cas particulier où  $g$  est la fonction de répartition de la loi uniforme sur  $[0, 1]$ .

$$R(X) = \int_0^{+\infty} [1 - F(x)]dx - \int_{-\infty}^0 F(x)dx = \mathbb{E}(X)$$

### 2.4.1 Exemple de mesure de risque de distorsion

**Définition 16.** (Mesure de risque de Wang)

La mesure de risque de Wang associée à la fonction de distorsion  $g$ , notée  $R_g$ , est définie par:

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

**Remarques:**

1. Si  $g(q) = q$  alors la mesure de risque de Wang correspond à l'espérance mathématique de  $X$  i.e.  $R_g(X) = \mathbb{E}[X]$
2. Si  $\forall q \in [0, 1], g(q) \geq q$ , alors  $R_g(X) \geq \mathbb{E}[X]$

3. Si  $\forall q \in [0,1], g_1(q) \leq g_2(q)$ , alors  $R_{g_1}(X) \leq R_{g_2}(X)$

**Proposition 12.** *Quel que soit le risque  $X$ , la mesure de risque de Wang associée à la fonction de distorsion  $g$  peut s'écrire:*

$$R_g(X) = \int_0^1 VaR_{1-\alpha}(X) dg(\alpha)$$

*Ainsi, la mesure de risque de Wang est une moyenne pondérée de la VaR.*

**Exemple 2.**

*Si nous considérons la fonction de distorsion  $g : [0,1] \rightarrow [0,1]$  définie par:*

$$g_\alpha(x) = \mathbb{I}(x \geq 1 - \alpha)$$

*pour une valeur  $\alpha \in [0,1]$  fixée, on obtient*

$$g_\alpha(\bar{F}_X(x)) = \mathbb{I}(\bar{F}_X(x) \geq 1 - \alpha)$$

*On a :  $\forall x \in \mathbb{R}$  et  $\alpha \in [0,1]$*

$$F_X(x) < \alpha \Leftrightarrow F_X^{-1}(\alpha) > x$$

*Donc*

$$g_\alpha(\bar{F}_X(x)) = \begin{cases} 1, & \text{Si } x < F_X^{-1}(\alpha); \\ 0, & \text{Si } x \geq F_X^{-1}(\alpha); \end{cases}$$

*On obtient*

$$R_{g_\alpha}(X) = \int_0^{F_X^{-1}(\alpha)} dx = F_X^{-1}(\alpha) = VaR_\alpha(X)$$

*ce qui montre que la VaR au niveau de probabilité  $\alpha$  est une mesure de Wang.*

*Voici le graphe qui représente la fonction de distorsion associée à la VaR.*

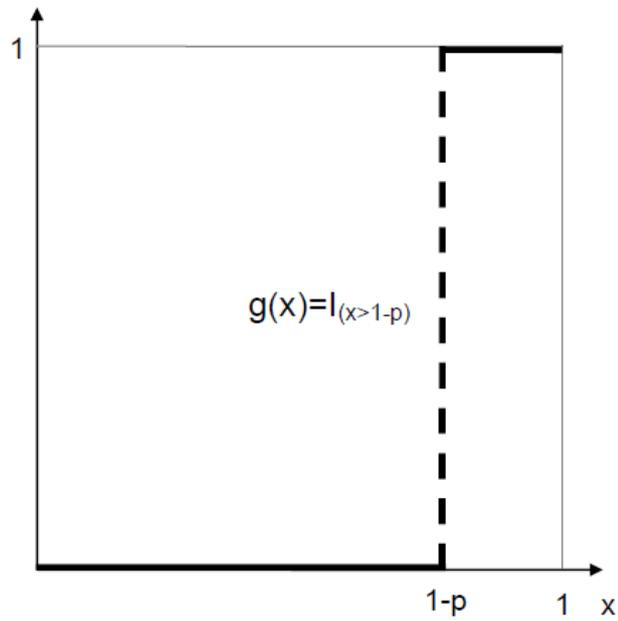


FIG. 2.1 – Fonction de Distorsion associée à la VaR

**Exemple 3.**

De la même manière, si on considère la fonction de distorsion

$$g_{\alpha}(x) = \min\left\{\frac{x}{1-\alpha}, 1\right\}$$

pour  $\alpha \in [0,1]$  fixé, on obtient

$$g_{\alpha}(\bar{F}_X(x)) = \min\left(\frac{1-F_X(x)}{1-\alpha}, 1\right)$$

$$g_{\alpha}(\bar{F}_X(x)) = \begin{cases} 1, & \text{Si } x < F_X^{-1}(\alpha) \\ \frac{1-F_X(x)}{1-\alpha} & \text{Si } x \geq F_X^{-1}(\alpha) \end{cases}$$

Donc

$$\begin{aligned}
R_{g_\alpha}(X) &= \int_0^{F_X^{-1}(\alpha)} dx + \frac{1}{1-\alpha} \int_{F_X^{-1}(\alpha)}^{+\infty} \bar{F}_X(x) dx \\
&= \int_0^{F_X^{-1}(\alpha)} dx + \frac{1}{1-\alpha} \int_{F_X^{-1}(\alpha)}^{+\infty} (1 - F_X(x)) dx \\
&= \int_0^{F_X^{-1}(\alpha)} dx + \frac{1}{1-\alpha} \int_{Q_\alpha(X)}^{+\infty} (1 - F_X(x)) dx \\
&= Q_\alpha(X) + \frac{1}{1-\alpha} \mathbb{E}[X - Q_\alpha(X)]_+ \\
&= Q_\alpha(X) + \frac{1}{1-\alpha} ES_\alpha[X]
\end{aligned}$$

$$R_{g_\alpha}(X) = TVaR_\alpha(X)$$

Dans ce cas,  $g_\alpha$  est une fonction de répartition, correspondant à une loi uniforme sur  $[0,1]$ .

Le graphe suivant représente la fonction de distorsion associée à la TVaR.

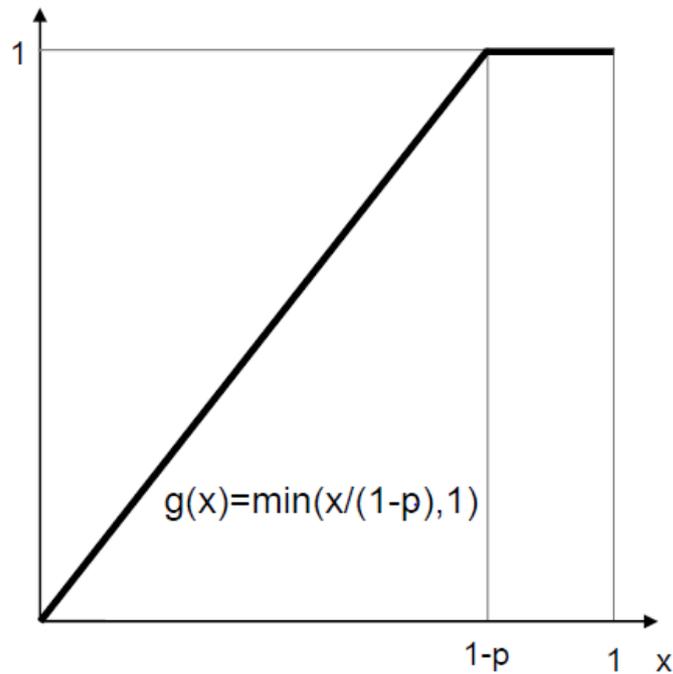


FIG. 2.2 – Fonction de Distorsion associée à la TVaR

**Proposition 13.** *L'Expected Shortfall et la Conditional Tail Expectation ne sont pas des mesures*

de risque de Wang.

**Preuve:**

Supposons que  $ES[X, \alpha]$  puisse effectivement être représentée sous la forme d'une mesure de risque de Wang, associée à une fonction de distorsion  $g_\alpha$ , i.e.  $ES[X, \alpha] = R_{g_\alpha}[X]$  quel que soit le risque  $X$ .

Prenons  $X \sim U(0,1)$ . Dans ce cas,

$$\begin{aligned} ES[X, \alpha] &= \int_\alpha^1 (1-x) dx \\ &= \frac{1}{2}(1-\alpha)^2 \\ &= R_{g_\alpha}(X) \\ &= \int_0^1 g_\alpha(1-x) dx. \end{aligned}$$

Considérons à présent  $Y \sim B(q)$  pour  $0 < q \leq 1 - \alpha$ . On obtient

$$ES[Y, \alpha] = q = R_{g_\alpha}(Y) = g_\alpha(q)$$

Ceci implique que  $g_\alpha(q) = q$ . En insérant ce résultat dans  $\int_0^1 g_\alpha(1-x) dx$  on obtient

$$\begin{aligned} \frac{1}{2}(1-\alpha)^2 &= \int_0^\alpha g_\alpha(1-x) dx + \int_\alpha^1 (1-x) dx \\ &\geq \frac{1}{2}(1-\alpha)^2 + \alpha(1-\alpha) \end{aligned}$$

ce qui nous mène à une contradiction puisque  $0 < \alpha < 1$ .

Pour montrer que la CTE n'est pas une mesure de risque de Wang. Il suffit de procéder de la même manière que pour l'ES.

**Exemple 4.**

La famille beta des mesures de risque de distorsion, proposée par Wirch et Hardy(1999), utilise la fonction beta incomplète:

$$g(F_X(x)) = \beta(a, b; F_X(x)) = \int_0^{F_X(x)} \frac{1}{\beta(a, b)} t^{\alpha-1} (1-t)^{b-1} dt = S_\beta(F_X(x))$$

Avec  $\beta(a,b) = \frac{\tau(a)\tau(b)}{\tau(a+b)}$

**Remarque:**

La mesure de risque beta-distorsion est concave si et seulement si  $a \leq 1$  et  $b \geq 1$ .

**Exemple 5.** Soit

$$g(x) = 1 - (1 - x)^\xi, \xi \geq 1$$

nous obtenons

$$R_g(X) = \int_{x \geq 0} (1 - (F_X(x))^\xi) dx$$

Lorsque  $\xi$  est un entier,  $R_g(X)$  s'interprète comme la valeur attendue du maximum  $M_\xi = \max\{X_1, \dots, X_\xi\}$  d'un ensemble de  $\xi$  variables aléatoires indépendantes et de même loi que  $X$ . En effet, la fonction de queue de  $M_\xi$  est donnée par:

$$Pr[M_\xi > x] = 1 - Pr[X_1 \leq x, \dots, X_\xi \leq x] = 1 - [F_X(x)]^\xi$$

et donc

$$R_g(X) = \mathbb{E}[M_\xi]$$

**Exemple 6.**

Considérons la fonction de distorsion

$$g(x) = x^{\frac{1}{\xi}}, \xi \geq 1$$

On obtient la mesure de risque  $PH_\xi$  définie par:

$$PH_\xi[X] = R_g(X) = \int_{x \geq 0} [\bar{F}_X(x)]^{1/\xi} dx$$

qui est bien une mesure de Wang. Pour  $\xi = 1$ , on a:  $PH_1[X] = \mathbb{E}[X]$

## 2.5 Propriétés des mesures de risque de distorsion

Les principales propriétés des mesures de risque de distorsion sont résumées dans le résultat suivant.

**Proposition 14.**

1. Les mesures de risque de distorsion sont monotones:

$$Si \ X \geq 0 \ \text{alors} \ R_g(X) \geq 0$$

2. Les mesures de risque de distorsion sont homogènes.

$$R_g(\lambda X) = \lambda R_g(X), \quad \forall \lambda \geq 0$$

3. Les mesures de risque de distorsion sont invariantes par translation.

$$R_g(X + c) = R_g(X) + c, \quad \forall c \in \mathbb{R}$$

4. soit  $\tilde{g}(x) = 1 - g(1 - x)$ , on a:  $R_g(-X) = -R_{\tilde{g}}(X)$

5. En général les mesures de risque de distorsion ne sont pas additives:

$$R_g(X + Y) \neq R_g(X) + R_g(Y)$$

6. les mesures de risque de distorsion sont sous-additives si et seulement si la fonction de distorsion  $g$  est concave:

$$R_g(X + Y) \leq R_g(X) + R_g(Y)$$

La preuve est donnée dans Wirch et Hardy (1999). Par conséquent, les mesures de risque de distorsion correspondant à des fonctions de distorsion concaves sont cohérentes.

**Preuve:**

Les propriétés (1) et (2) sont faciles à vérifier.

**Preuve 3:**

$$\begin{aligned} R_g(X + c) &= \int_{-\infty}^0 [g(\bar{F}_{X+c}(x)) - 1] dx + \int_0^c g(\bar{F}_{X+c}(x)) dx + \int_c^{+\infty} g(\bar{F}_{X+c}(x)) dx \\ &= \int_{-\infty}^0 [g(\bar{F}_X(x - c)) - 1] dx + \int_0^c g(\bar{F}_X(x - c)) dx + \int_c^{+\infty} g(\bar{F}_X(x - c)) dx \end{aligned}$$

Si on pose  $x = c + \mu$  on aura

$$\begin{aligned} R_g(X + c) &= \int_{-\infty}^{-c} [g(\bar{F}_X(\mu)) - 1] d\mu + \int_{-c}^0 g(\bar{F}_X(\mu)) d\mu + \int_0^{+\infty} g(\bar{F}_X(\mu)) d\mu \\ &= \int_{-\infty}^0 [g(\bar{F}_X(\mu)) - 1] d\mu + \int_0^{+\infty} g(\bar{F}_X(\mu)) d\mu + \int_{-c}^0 d\mu \\ &= R_g(X) + c \end{aligned}$$

**Preuve 4:**

$$\begin{aligned}
R_g(-X) &= \int_{-\infty}^0 [g(\overline{F}_{-X}(x)) - 1]dx + \int_0^{+\infty} g(\overline{F}_{-X}(x))dx \\
&= \int_{-\infty}^{\mu} [g(1 - \overline{F}_X(-x) + P[X = x]) - 1]dx + \int_{-\infty}^{\mu} g(1 - \overline{F}_X(-x) + P[X = x])dx \\
&= \int_{-\infty}^{\mu} [g(1 - \overline{F}_X(-x)) - 1]dx + \int_{-\infty}^{\mu} g(1 - \overline{F}_X(-x))dx
\end{aligned}$$

On pose  $x = -\mu$

$$R_g(-X) = \int_{-\infty}^{\mu} [g(1 - \overline{F}_X(\mu)) - 1]d\mu + \int_{-\infty}^{\mu} g(1 - \overline{F}_X(\mu))d\mu = -R_{\bar{g}}(X)$$

**Preuve 5: (contre exemple)**

Soit  $g(x) = x^2$ , la distribution des variables aléatoires discrètes et indépendantes  $X$  et  $Y$  est définie comme suit:  $P(1,1) = P(1, -1) = P(-1,1) = P(-1,-1) = 0.25$ .

Les distributions marginales de  $X$  et  $Y$  sont:  $P(1) = P(-1) = 0.5$ .

Donc  $R_g(X) = R_g(Y) = -0.5$  et  $R_g[X + Y] = 0.25$ . Ce qui donne:

$$R_g(X) + R_g(Y) = -1 \neq R_g(X + Y) = 0.25$$

**Proposition 15.** (31)

Soit  $X$  une variable aléatoire à valeurs réelles de fonction de répartition  $F$ , et  $g$  une fonction dérivable avec  $g(0) = 0$  et  $g(1) = 1$  alors:

$$\begin{aligned}
H(X) &= \int_{-\infty}^0 [g(1 - F(x)) - 1]dx + \int_0^{+\infty} g(1 - F(x))dx \\
&= \int_0^1 F^{-1}(x)\psi(x)dx
\end{aligned}$$

Où,  $\psi(x) = g'(1 - x)$

**Preuve:**

$$\begin{aligned}
H(X) &= \int_{-\infty}^0 [g(1 - F(X)) - 1]dx + \int_0^{+\infty} g(1 - F(x))dx \\
&= - \int_{-\infty}^0 \left[ \int_{-\infty}^x g'(1 - F(y))dF(y) \right]dx + \int_0^{+\infty} \left[ \int_x^{+\infty} g'(1 - F(y))dF(y) \right]dx \\
&= - \int_{-\infty}^0 [g'(1 - F(y)) \int_y^0 dx]dF(y) + \int_0^{+\infty} [g'(1 - F(y)) \int_0^y dx]dF(y) \\
&= \int_{-\infty}^0 yg'(1 - F(y))dF(y) + \int_0^{+\infty} yg'(1 - F(y))dF(y) \\
&= \int_{-\infty}^{+\infty} yg'(1 - F(y))dF(y) \\
&= \int_0^1 F^{-1}(s)g'(1 - s)ds
\end{aligned}$$

**Exemple:**

Quand  $g(x) = x^{\frac{1}{\xi}}$ ,  $\xi \geq 1$

$PH_{\xi}(X)$  est donnée par la proposition précédente avec

$$\psi(x) = \xi^{-1}(1 - x)^{\frac{1}{\xi}-1}$$

Donc

$$PH_{\xi}(X) = \xi^{-1} \int_0^1 (1 - x)^{\frac{1}{\xi}-1} F^{-1}(x)dx$$



## Chapitre 3

# Estimation d'une mesure de risque de distorsion

### 3.1 Estimation empirique d'une mesure de risque de distorsion

Soit  $X_1, \dots, X_n$   $n$  variables aléatoires indépendantes et identiquement distribuées (iid) de densité commune  $f$  et de fonction de répartition  $F$ .

### 3.2 Statistiques d'ordre

**Définition 17.**

On appelle statistiques d'ordre associées à un échantillon  $X_1, \dots, X_n$ , notées  $X_{(1)}, \dots, X_{(n)}$ , les variables aléatoires ordonnées:

$$\min(X_1, \dots, X_n) = X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)} = \max(X_1, \dots, X_n)$$

#### 3.2.1 Densité conjointe de $n$ statistiques d'ordre

**Lemme 1.**

Soit  $X_1, \dots, X_n$   $n$  variables aléatoires (iid) de fonction de répartition commune  $F$  continue, alors la densité de  $X_{(1)}, \dots, X_{(n)}$  est donnée par:

$$f_{(X_{(1)}, \dots, X_{(n)})}(x_1, x_2, \dots, x_n) = n! \prod_{i=1}^n f(x_i) \quad \text{avec} \quad x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$$

### 3.2.2 Densité de la $k^{\text{ième}}$ statistique d'ordre

**Lemme 2.**

La loi de la variable aléatoire  $X_{(k)}$ , pour  $1 \leq k \leq n$  est donnée par :

$$\begin{aligned} F_{X_{(k)}}(x) &= P[X_{(k)} \leq x] \\ &= \sum_{j=k}^n C_n^j [F(x)]^j [1 - F(x)]^{n-j} \end{aligned}$$

La densité de  $X_{(k)}$  est

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} [F(x)]^{k-1} [1 - F(x)]^{n-k} f(x)$$

**Démonstration:**

$$\begin{aligned} F_{X_{(k)}}(x) &= P[X_{(k)} \leq x] \\ &= P[\text{au moins } k \text{ variables sont inférieures à } x] \\ &= \bigcup_{j=k}^n P[k \text{ variables } \leq x, (n-k) \text{ variables } > x] \\ &= \sum_{j=k}^n C_n^j [F(x)]^j [1 - F(x)]^{n-j} \end{aligned}$$

On a:

$$P[x \leq X_{(k)} \leq x + \Delta x] = \frac{n!}{(k-1)!(n-k)!} [F(x)]^{k-1} [F(x + \Delta x) - F(x)] [1 - F(x + \Delta x)]^{n-k}$$

Or

$$f_{X_{(k)}}(x) = \lim_{\Delta x \rightarrow 0} \frac{P[x \leq X_{(k)} \leq x + \Delta x]}{\Delta x}$$

Ce qui donne

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} [F(x)]^{k-1} f(x) [1 - F(x)]^{n-k}$$

**Remarque:**

Soient  $U_{(1)}, \dots, U_{(n)}$  les statistiques d'ordre associées à des variables aléatoires  $U_1, \dots, U_n$  de même loi uniforme sur  $[0,1]$ .

Alors:

$$\begin{aligned} f_{U_{(i)}}(u) &= \frac{n!}{(i-1)!(n-i)!} [F(u)]^{i-1} f(u) [1-F(u)]^{n-i} \\ &= \frac{n!}{(i-1)!(n-i)!} u^{i-1} (1-u)^{n-i} \end{aligned}$$

et

$$\mathbb{E}(U_{(i)}) = \frac{i}{n+1}$$

**Démonstration:**

$$\begin{aligned} \mathbb{E}(U_{(i)}) &= \int_0^1 u f_{U_{(i)}}(u) du \\ \mathbb{E}(U_{(i)}) &= \int_0^1 \frac{n!}{(i-1)!(n-i)!} u^i (1-u)^{n-i} du \\ &= \frac{n!}{(i-1)!(n-i)!} \int_0^1 u^i (1-u)^{n-i} du \\ &= \frac{n!}{(i-1)!(n-i)!} \int_0^1 u^{i+1-1} (1-u)^{n-i+1-1} du \\ &= \frac{n!}{(i-1)!(n-i)!} \beta(i+1, n-i+1) \end{aligned}$$

On a:

$$\beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

$$\Gamma(n) = (n-1)!$$

Donc

$$\begin{aligned}
\mathbb{E}(U_{(i)}) &= \frac{n!}{(i-1)!(n-i)!} \beta(i+1, n-i+1) \\
&= \frac{n!}{(i-1)!(n-i)!} \frac{\Gamma(i+1)\Gamma(n-i+1)}{\Gamma(n+2)} \\
&= \frac{n!}{(i-1)!(n-i)!} \frac{i!(n-i)!}{(n+1)!} \\
&= \frac{i}{n+1}
\end{aligned}$$

### 3.2.3 Densité conjointe de deux statistiques d'ordre

La densité conjointe de  $(X_{(r)}, X_{(s)})$  pour  $r < s$  est donnée par :

$$f_{(X_{(r)}, X_{(s)})}(x, y) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!} [F(x)]^{r-1} f(x) [F(y) - F(x)]^{s-r-1} f(y) [1 - F(y)]^{n-s}$$

Avec  $-\infty < x < y < +\infty$

### 3.2.4 Quantile empirique

Soit  $X_1, \dots, X_n$  un n-échantillon de fonction de répartition  $F$ , et  $X_{(1)}, \dots, X_{(n)}$  l'échantillon ordonné.

Soit  $p \in ]0, 1[$ , on appelle quantile empirique d'ordre  $p$ , la variable aléatoire notée  $\widehat{X}_{([np])}$  définie par:

$$\widehat{X}_{([np])} = \begin{cases} X_{(np)}, & \text{Si } np \in \mathbb{N}; \\ X_{([np]+1)}, & \text{Si non;} \end{cases}$$

Où  $[np]$  désigne la partie entière de  $np$

Comme tenu du fait que  $F_n$  est une fonction en escalier continue à gauche, on notera que

$$F_n^{-1}(p) = \widehat{X}_{([np])}, \quad \forall p \in ]\frac{np-1}{n}, \frac{np}{n}]$$

Et

$$F_n(X_{(i)}) = U_{(i)}$$

**Proposition 16.** (12)

Soit  $X_1, \dots, X_n$  un échantillon indépendant et identiquement distribué de loi absolument continue de densité  $f$ . Si  $f(Q(\alpha)) > 0$ , alors

$$\sqrt{n}(Q_n(\alpha) - Q(\alpha)) \xrightarrow{L} N\left(0, \frac{\alpha(1-\alpha)}{f^2(Q(\alpha))}\right), \quad \text{Lorsque } n \rightarrow \infty$$

Pour terminer avec les propriétés du quantile empirique  $Q_n(\alpha)$ , rappelons le résultat suivant.

**Proposition 17.** (12) (Représentation de Bahadur)

Soit  $X_1, \dots, X_n$  un échantillon de densité marginale  $f$  continue. Si  $f(Q(\alpha)) > 0$ , alors

$$Q_n(\alpha) = Q(\alpha) + \frac{\alpha - F_n(Q(\alpha))}{f(Q(\alpha))} + Z_n$$

où,  $Z_n = O(n^{-3/4}[\log(\log n)]^{3/4})$  p.s  
et  $n^2 Z_n \xrightarrow{P} 0$

### 3.3 Estimateur Linéaire

**Définition 18.**

Soit  $X_1, \dots, X_n$  un échantillon de taille  $n \geq 1$  de  $X$  et  $X_{(1)} < X_{(2)} < \dots < X_{(n)}$  la statistique d'ordre.

On appelle  $L$ -estimateur ou (estimateur linéaire) tout estimateur de la forme:

$$L = \sum_{i=1}^n C_{ni} X_{(i)}$$

Où les  $C_{ni}$  sont des coefficients réels.

**Exemple:**

On peut prendre:

$$\begin{aligned} C_{n1} &= -1 \\ C_{nn} &= 1 \\ C_{ni} &= 0, \quad \forall i \neq 1, n \\ L &= X_{(n)} - X_{(1)} \end{aligned}$$

Généralement on écrit  $L$  sous la forme suivante:

$$L = \frac{1}{n} \sum_{i=1}^n J\left(\frac{i}{n+1}\right) X_{(i)}$$

Où  $C_{ni} = \frac{1}{n}J\left(\frac{i}{n+1}\right)$ , et  $J$  est une fonction à variation bornée.

### 3.3.1 Propriétés des L-estimateurs

**Proposition 18.**

#### 1. Espérance d'un L-estimateur

$$\mathbb{E}[L] = \mathbb{E}\left[\sum_{i=1}^n C_{ni}X_{(i)}\right] = \sum_{i=1}^n C_{ni}\mathbb{E}[X_{(i)}]$$

*Il est pratiquement difficile de calculer les moments d'une statistique d'ordre, hormis dans le cas uniforme ou exponentiel.*

*Pour cela, on a essayé d'approximer l'espérance de la statistique d'ordre  $X_{(i)}$  par deux approches :*

(a) **Première approche:** (Représentation de Bahadur (1966))

*Soit  $X_1, \dots, X_n$  un échantillon de taille  $n \geq 1$  de  $X$  de densité marginale  $f$  et de fonction de répartition  $F$  dérivable.*

*Soit  $F(\xi_p) = p$  le quantile d'ordre  $p$  avec  $0 < p < 1$ , si  $f(\xi_p) > 0$  alors :*

$$\widehat{X}_{(np)} = \xi_p - \frac{F_n(\xi_p) - p}{f(\xi_p)} + R_n$$

*Et  $\sqrt{n}R_n \xrightarrow{P} 0$  quant  $n \rightarrow \infty$*

*Ce qui donne en loi*

$$\widehat{X}_{(np)} \simeq \xi_p - \frac{F_n(\xi_p) - p}{f(\xi_p)}$$

*On obtient*

$$\mathbb{E}[\widehat{X}_{(np)}] \simeq \xi_p - \frac{1}{f(\xi_p)}(\mathbb{E}[F_n(\xi_p)] - p)$$

*Avec*

$$\begin{aligned}
\mathbb{E}[F_n(\xi_p)] &= \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{X_i \leq \xi_p\}}\right] \\
&= \frac{1}{n} \sum_{i=1}^n \mathbb{E}[\mathbb{I}_{\{X_i \leq \xi_p\}}] \\
&= P(X_i \leq \xi_p) \\
&= F(\xi_p) \\
\mathbb{E}[F_n(\xi_p)] &= p
\end{aligned}$$

En remplaçant  $\mathbb{E}[F_n(\xi_p)]$  par sa valeur, on aura

$$\begin{aligned}
\mathbb{E}[\widehat{X}_{([np])}] &\simeq \xi_p - \frac{1}{f(\xi_p)}(p - p) \\
&\simeq \xi_p \\
&\simeq F^{-1}(p)
\end{aligned}$$

Donc

$$\mathbb{E}[\widehat{X}_{([np])}] \simeq F^{-1}(p)$$

(b) **Deuxième approche:** (Développement limité)

Soient  $U_{(1)}, \dots, U_{(n)}$  les statistiques d'ordre associées à des variables aléatoires  $U_1, \dots, U_n$  de même loi uniforme sur  $[0, 1]$ , on a:

$$U_{(i)} = F(X_{(i)}) \quad \text{alors} \quad X_{(i)} = F^{-1}(U_{(i)})$$

$$\text{et } \mathbb{E}[U_{(i)}] = \frac{i}{n+1} = p_i$$

Un développement limité de  $F^{-1}$  au voisinage du point  $p_i$  nous donne :

$$X_{(i)} \simeq F^{-1}(p_i) + (F^{-1})'(p_i)(U_{(i)} - p_i)$$

Alors

$$\begin{aligned}
\mathbb{E}[X_{(i)}] &\simeq F^{-1}(p_i) + (F^{-1})'(p_i)(\mathbb{E}[U_{(i)}] - p_i) \\
&\simeq F^{-1}(p_i) + (F^{-1})'(p_i)(p_i - p_i) \\
&\simeq F^{-1}(p_i)
\end{aligned}$$

*Donc*

$$\mathbb{E}[X_{(i)}] \simeq F^{-1}(p_i)$$

*On obtient*

$$\begin{aligned}
\mathbb{E}[L] &= \sum_{i=1}^n C_{ni} \mathbb{E}[X_{(i)}] \\
&= \frac{1}{n} \sum_{i=1}^n J\left(\frac{i}{n+1}\right) \mathbb{E}[X_{(i)}] \\
&\simeq \frac{1}{n} \sum_{i=1}^n J\left(\frac{i}{n+1}\right) F^{-1}(p_i)
\end{aligned}$$

*Donc*

$$\mathbb{E}[L] \simeq \int_0^1 J(t) F^{-1}(t) dt$$

## 2. Variance d'un L-estimateur

$$\begin{aligned}
V(L) &= V\left(\sum_{i=1}^n C_{ni} X_{(i)}\right) \\
&= \sum_{i=1}^n \sum_{j=1}^n C_{ni} C_{nj} \text{Cov}(X_{(i)}, X_{(j)}) \\
&= \frac{1}{n} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n J\left(\frac{i}{n+1}\right) J\left(\frac{j}{n+1}\right) \text{Cov}(X_{(i)}, X_{(j)})
\end{aligned}$$

Pour approximer la covariance entre  $X_{(i)}$  et  $X_{(j)}$ , on va utiliser les approches précédentes.

(a) **Première approche:** (Développement limité )

On a

$$X_{(i)} = F^{-1}(U_{(i)}) \text{ et } \mathbb{E}[U_{(i)}] = p_i$$

Un développement limité de  $F^{-1}$  au voisinage du point  $p_i$  nous donne :

$$X_{(i)} \simeq F^{-1}(p_i) + (F^{-1})'(p_i)(U_{(i)} - p_i)$$

Par conséquent

$$\begin{aligned} \text{Cov}(X_{(i)}, X_{(j)}) &\simeq \text{Cov}(F^{-1}(p_i) + (F^{-1})'(p_i)(U_{(i)} - p_i), F^{-1}(p_j) + (F^{-1})'(p_j)(U_{(j)} - p_j)) \\ &\simeq (F^{-1})'(p_i)(F^{-1})'(p_j)\text{Cov}(U_{(i)} - p_i, U_{(j)} - p_j) \\ &\simeq (F^{-1})'(p_i)(F^{-1})'(p_j)\text{Cov}(U_{(i)}, U_{(j)}) \end{aligned}$$

Or

$$\text{Cov}(U_{(i)}, U_{(j)}) = \mathbb{E}[U_{(i)}U_{(j)}] - \mathbb{E}[U_{(i)}]\mathbb{E}[U_{(j)}]$$

$$\text{Et } U_{(i)} = \mathbb{I}_{\{X \leq x\}}, U_{(j)} = \mathbb{I}_{\{Y \leq y\}}$$

$$\mathbb{E}[U_{(i)}] = P[X \leq x] = F(x) = p_i \quad \text{alors } x = F^{-1}(p_i)$$

$$\mathbb{E}[U_{(j)}] = P[Y \leq y] = F(Y) = p_j \quad \text{alors } y = F^{-1}(p_j)$$

Et

$$\mathbb{E}[U_{(i)}U_{(j)}] = P[X \leq x, Y \leq y] = P[X \leq F^{-1}(p_i), Y \leq F^{-1}(p_j)]$$

- Si  $F^{-1}(p_i) < F^{-1}(p_j)$  alors  $X \leq F^{-1}(p_i)$  implique  $Y \leq F^{-1}(p_j)$

Ce qui donne

$$\mathbb{E}[U_{(i)}U_{(j)}] = F(F^{-1}(p_i)) = F(x) = p_i$$

- Or si  $F^{-1}(p_j) < F^{-1}(p_i)$  alors  $Y \leq F^{-1}(p_j)$  implique  $X \leq F^{-1}(p_i)$

$$\mathbb{E}[U_{(i)}U_{(j)}] = F(F^{-1}(p_j)) = F(y) = p_j$$

On constate que

$$\mathbb{E}[U_{(i)}U_{(j)}] = F(\min(F^{-1}(p_i), F^{-1}(p_j)))$$

Alors

$$\begin{aligned} \text{Cov}(U_{(i)}, U_{(j)}) &= \mathbb{E}[U_{(i)}U_{(j)}] - \mathbb{E}[U_{(i)}]\mathbb{E}[U_{(j)}] \\ &= F(\min(F^{-1}(p_i), F^{-1}(p_j))) - p_i p_j \end{aligned}$$

La covariance entre  $X_{(i)}$  et  $X_{(j)}$  devient

$$\text{Cov}(X_{(i)}, X_{(j)}) \simeq (F^{-1})'(p_{(i)})(F^{-1})'(p_{(j)})F(\min(F^{-1}(p_i), F^{-1}(p_j))) - p_i p_j$$

(b) **Représentation de Bahadur :**

$$\widehat{X}_{([np])} \simeq \xi_p - \frac{F_n(\xi_p) - p}{f(\xi_p)}$$

p.s

Et

$$\begin{aligned} \text{Cov}(\widehat{X}_{([np_i])}, \widehat{X}_{([np_j])}) &\simeq \text{Cov}\left(\frac{F_n(\xi_{p_i}) - p_i}{f(\xi_{p_i})}, \frac{F_n(\xi_{p_j}) - p_j}{f(\xi_{p_j})}\right) \\ &\simeq \frac{1}{f(\xi_{p_i})} \frac{1}{f(\xi_{p_j})} \text{Cov}(F_n(\xi_{p_i}), F_n(\xi_{p_j})) \end{aligned}$$

$$\begin{aligned} \text{Cov}(F_n(\xi_{p_i}), F_n(\xi_{p_j})) &= \frac{1}{n} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(\mathbb{I}_{\{X_{(i)} \leq \xi_{p_i}\}}, \mathbb{I}_{\{X_{(k)} \leq \xi_{p_j}\}}) \\ &= \frac{1}{n} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n F(\min(\xi_{p_i}, \xi_{p_j})) - F(\xi_{p_i})F(\xi_{p_j}) \end{aligned}$$

Donc

$$\begin{aligned}
Cov(\widehat{X}_{([np_i]), \widehat{X}_{([np_j])})} &\simeq \frac{1}{f(\xi_{p_i})} \frac{1}{f(\xi_{p_j})} Cov(F_n(\xi_{p_i}), F_n(\xi_{p_j})) \\
&\simeq \frac{1}{n} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \frac{1}{f(\xi_{p_i})} \frac{1}{f(\xi_{p_j})} F(\min(\xi_{p_i}, \xi_{p_j})) - F(\xi_{p_i})F(\xi_{p_j})
\end{aligned}$$

D'après ces deux approches

$$\begin{aligned}
V(L) &= \frac{1}{n} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n J\left(\frac{i}{n+1}\right) J\left(\frac{j}{n+1}\right) Cov(X_{(i)}, X_{(j)}) \\
&\simeq \frac{1}{n} \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n J\left(\frac{i}{n+1}\right) J\left(\frac{j}{n+1}\right) (F^{-1})'(p_{(i)}) (F^{-1})'(p_{(j)}) F(\min(F^{-1}(p_i), F^{-1}(p_j))) - p_i p_j \\
&\simeq \int_0^1 \int_0^1 \min(F(x), F(y)) - F(x)F(y) J(x)J(y) d(F^{-1}(x)) d(F^{-1}(y))
\end{aligned}$$

### 3.4 Estimation empirique d'une mesure de risque de distorsion

Soit  $X_1, \dots, X_n$   $n$  variables aléatoires indépendantes et identiquement distribuées(iid) de densité commune  $f$ , et de fonction de répartition  $F$ .

#### 3.4.1 Estimateur empirique d'une mesure de risque de distorsion

**Définition 19.**

L'estimateur empirique de la mesure de distorsion  $R(X)$  est défini par:

$$\hat{R}(X) = - \int_{-\infty}^0 [1 - g(1 - F_n(x))] dx + \int_0^{+\infty} g(1 - F_n(x)) dx$$

**Remarque :**

On pose  $\overline{F}_n(x) = 1 - F_n(x)$ , alors

$$\begin{aligned}
\hat{R}(X) &= - \int_{-\infty}^0 [1 - g(\bar{F}_n(x))]dx + \int_0^{+\infty} g(\bar{F}_n(x))dx \\
&= - \int_{-\infty}^{X_{(1)}} [1 - g(\bar{F}_n(x))]dx - \sum_{i=2}^k \int_{X_{(i-1)}}^{X_{(i)}} [1 - g(\bar{F}_n(x))]dx - \int_{X_{(k)}}^0 [1 - g(\bar{F}_n(x))]dx \\
&\quad + \int_0^{X_{(k+1)}} g(\bar{F}_n(x))dx + \sum_{i=k+1}^{n-1} \int_{X_{(i)}}^{X_{(i+1)}} g(\bar{F}_n(x))dx + \int_{X_{(n)}}^{+\infty} g(\bar{F}_n(x))dx \\
&= - \sum_{i=2}^k (X_{(i)} - X_{(i-1)})[1 - g(1 - \frac{i}{n})] - (0 - X_{(k)})[1 - g(1 - \frac{k}{n})] \\
&\quad + (X_{(k+1)} - 0)g(1 - \frac{k}{n}) + \sum_{i=k+1}^{n-1} (X_{(i+1)} - X_{(i)})g(1 - \frac{i}{n}) \\
&= \sum_{i=1}^n X_{(i)} [g(\frac{n-i+1}{n}) - g(\frac{n-i}{n})] \\
&= \sum_{i=1}^n C_{ni} X_{(i)}
\end{aligned}$$

Donc l'estimateur empirique d'une mesure de risque de distorsion est un L-estimateur avec

$$C_{ni} = \int_{\frac{i-1}{n}}^{\frac{i}{n}} \psi(x)dx$$

### 3.4.2 Propriétés élémentaires de l'estimateur empirique d'une mesure de risque distorsion

- Esperance:

$$\begin{aligned}
\mathbb{E}[\hat{R}(X)] &= \mathbb{E}[\sum_{i=1}^n C_{ni} X_{(i)}] \\
&= \sum_{i=1}^n C_{ni} \mathbb{E}(X_{(i)}) \\
&\approx \sum_{i=1}^n (\int_{\frac{i-1}{n}}^{\frac{i}{n}} \psi(x)dx) F^{-1}(p_i)
\end{aligned}$$

- Variance:

$$\begin{aligned}
V[\hat{R}(X)] &= \sum_{i=1}^n \sum_{j=1}^n C_{ni} C_{nj} \text{Cov}(X_{(i)}, X_{(j)}) \\
&\approx \sum_{i=1}^n \sum_{j=1}^n \left( \int_{\frac{i-1}{n}}^{\frac{i}{n}} \psi(x) dx \right) \left( \int_{\frac{j-1}{n}}^{\frac{j}{n}} \psi(x) dx \right) (F^{-1})'(p_{(i)}) (F^{-1})'(p_{(j)}) F(\min(F^{-1}(p_i), F^{-1}(p_j))) - p_i p_j
\end{aligned}$$

### 3.4.3 Propriétés asymptotiques de l'estimateur empirique d'une mesure de risque distorsion

**Théorème 1.** (Shorack and Wellner 1986)

Si  $\psi$  est une fonction continue sur  $[0,1]$  qui satisfait:

$$|\psi(x)| \leq cx^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1$$

Pour des constantes  $\alpha, \beta > 0$  et  $c < \infty$  et si  $\mathbb{E}[|X|^\gamma] < \infty$  pour  $\gamma > \frac{1}{\alpha}$  et  $\gamma > \frac{1}{\beta}$  alors:

$$\hat{R}(X) \xrightarrow[n \rightarrow \infty]{p.s.} R(X)$$

.

**Preuve:** voir Shorack and Wellner (1986) ou Zitikis (2002)

**Théorème 2.** (31)

Soit  $\psi$  une fonction continue sur  $[0,1]$  telle que:

$$|\psi(x)| \leq cx^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1$$

- $\forall \alpha, \beta > \frac{1}{2}$ , et  $c < \infty$
- $\forall \gamma > 1/\alpha - \frac{1}{2}$  et  $\gamma > 1/\beta - \frac{1}{2}$

Si  $\mathbb{E}[|X|^\gamma] < \infty$ , alors:

$$\sqrt{n}(\hat{R}(X) - R(X)) \xrightarrow{L} N(0, \sigma^2), \quad \text{quand, } n \rightarrow \infty$$

avec  $\sigma^2 = \int_0^1 \int_0^1 (\min(s,t) - st) \psi(s) \psi(t) d(F^{-1}(s)) d(F^{-1}(t)) < \infty$

Par l'utilisation des résultats du théorème (2), on peut construire les bornes de confiance asymptotique au niveau  $(1 - \alpha)$  de  $R$  qui sont :

$$\hat{R} \pm \frac{z_{\alpha/2}\sigma}{\sqrt{n}}$$

Où  $z_{\alpha/2}$  est le quantile d'ordre  $(1 - \alpha/2)$  de  $N(0,1)$ .

Ces bornes ne s'appliquent pas dans la pratique puisque  $\sigma^2$  dépend de la fonction  $F$  qui est inconnue.

### 3.5 Estimation de la Variance de l'estimateur empirique d'une mesure de risque distorsion

Pour l'estimation de la variance de  $\hat{R}$ , on peut utiliser des méthodes non paramétriques tel que le bootstrap et la méthode non paramétrique delta.

#### 3.5.1 La méthode bootstrap

En statistique le terme bootstrap fait référence à une technique d'inférence qui s'appuie sur le rééchantillonnage massif des données. Popularisé par *Bradley Efron*, professeur à l'université de Stanford (Californie), cette méthode permet d'effectuer des tests statistiques et de calculer des intervalles de confiance en se servant des données pour approximer une loi de probabilité inconnue.

Pour introduire le concept de bootstrap supposons que nous avons un échantillon de  $n$  variables aléatoires indépendantes et identiquement distribuées de fonction de répartition  $F$  inconnue. Nous nous intéressons à un paramètre  $\theta = \theta(F)$  estimé empiriquement par  $\hat{\theta} = \theta(F_n)$ , où  $F_n$  est la fonction de répartition empirique.

L'idée de base du bootstrap est de rééchantillonnage à plusieurs reprises (autrement dit à effectuer des tirages avec remise ) à partir de l'échantillon original.

Supposons maintenant qu'un échantillon de taille  $n$  est tirée de  $X$ , noté par  $F_n^{*(1)}$  où le symbole \* indique rééchantillonnage. Après avoir répété cette méthode  $R$  fois, on obtient une série d'échantillon bootstrap  $F_n^{*(1)} \dots F_n^{*(R)}$  ensuite les estimations correspondantes  $\theta_1^*, \dots, \theta_R^*$  où  $\theta_i^* = \theta(F_n^{*(i)})$ .

Soient

$$X_n = (X_{(1)}, X_{(2)}, \dots, X_{(n)})'$$

$$C = (C_1, C_2, \dots, C_n)'$$

Alors

$$\hat{R}[X] = C_1 X_{(1)} + C_2 X_{(2)} + \cdots + C_n X_{(n)} = C' \chi_n$$

Pour la classe des L-estimateurs *Hutson et Ernst (2000)* proposent l'estimateur bootstrap exact de la moyenne et de la variance.

**Théorème 3.** (*Hutson et Ernst (2000)*)

Les estimateurs bootstrap exact de  $\mathbb{E}[X_r|F]$  et  $V(X_r|F)$ ,  $1 \leq r \leq n$  sont respectivement

$$\mathbb{E}[X_r|F_n] = \sum_{j=1}^n W_{j(r)} X_{(j)} = \hat{\mu}_r$$

$$V(X_{(r)}|F_n) = \sum_{j=1}^n W_{j(r)} (X_{(j)} - \hat{\mu}_r)^2$$

Où

$$W_{j(r)} = r \binom{n}{r} [B(\frac{j}{n}; r, n-r+1)] - [B(\frac{j-1}{n}; r, n-r+1)]$$

Et

$$B(x; a, b) = \int_0^x t^{a-1} (1-t)^{b-1} dt$$

Hutson et Ernst (2000) donnent également la formule de l'estimateur bootstrap exact de  $Cov(X_{(r)}, X_{(s)}|F)$

**Théorème 4.** (*Hutson et Ernst (2000)*)

L'estimateur bootstrap exact de  $Cov(X_{(r)}, X_{(s)}|F)$  noté par  $\Sigma_n$  est donné par

$\forall 1 \leq r \leq n$  et  $r < s$

$$Cov(X_{(r)}, X_{(s)}|F_n) = \sum_{j=2}^n \sum_{j=1}^{j-1} V_{ij(rs)} (X_{(i)} - \hat{\mu}_r)(X_{(j)} - \hat{\mu}_s) + \sum_{j=1}^n V_{jj(rs)} (X_{(j)} - \hat{\mu}_r)(X_{(j)} - \hat{\mu}_s)$$

Où

$$V_{ij(rs)} = \int_{\frac{i-n}{n}}^{\frac{j}{n}} \int_{\frac{i-n}{n}}^{\frac{j}{n}} f_{rs;n}(\mu_r, \mu_s) d\mu_r d\mu_s$$

et

$$V_{jj(r,s)} \int_{\frac{j-n}{n}}^{\frac{j}{n}} \int_{\frac{j-n}{n}}^{\mu_s} f_{rs;n}(\mu_r, \mu_s) d\mu_r d\mu_s$$

et

$f_{rs;n}(\mu_r, \mu_s) = C_{rs} \mu_r^{r-1} (\mu_s - \mu_r)^{s-r-1} (1 - \mu_s)^{n-s}$  est la distribution conjointe de deux statistiques d'ordre uniforme  $U_{(r)}$  et  $U_{(s)}$

Avec

$$C_{rs} = \frac{n!}{(r-1)!(s-r-1)!(n-s)!}$$

Donc la variance bootstrap exacte de  $\hat{R}[X]$  est

$$\text{Var}(C' \chi_n | F_n) = C' \Sigma_n C$$

### 3.5.2 Méthode non paramétrique delta

Une autre façon d'estimer la variance est la méthode non paramétrique delta (ou tout simplement la méthode delta) qui emploie la fonction d'influence de l'estimateur.

**Définition 20.** (Fonctionnelle)

On appelle fonctionnelle  $T$  sur un ensemble  $\mathfrak{F}$  de fonctions une application linéaire de  $\mathfrak{F}$  dans  $\mathbb{C}$ . On note

$$\begin{aligned} T : \mathfrak{F} &\longrightarrow \mathbb{C} \\ f &\longmapsto \langle T, f \rangle \end{aligned}$$

**Définition 21.** (Fonction d'influence)

La fonction d'influence de la fonctionnelle  $T$  en  $F$  au point  $x_0$  est définie par la limite suivante, si elle existe

$$L_{T,F}(x_0) = \lim_{\epsilon \rightarrow 0} \frac{T((1-\epsilon)F + \epsilon G_{\delta_{x_0}}) - T(F)}{\epsilon}$$

Où  $G_{\delta_{x_0}}$  est la fonction de répartition associée à la masse de Dirac  $\delta_{x_0}$ , plus précisément, on a  $G_{\delta_{x_0}}(t) = \mathbb{I}_{x_0 \leq t}$ .

**Définition 22.** (Fonction de Heaviside)

La fonction heaviside ou fonction échelon, du nom de Oliver Heaviside est une fonction  $H$  discontinue définie par :

$$\forall x \in \mathbb{R}$$

$$H(x) = \begin{cases} 0, & \text{Si } x < 0 \\ 1, & \text{Si } x \geq 0 \end{cases}$$

C'est une primitive de la fonction de Dirac  $\delta$ . Certains auteurs donnent  $H(0) = 0$ , d'autre  $H(0) = 1$ ,  $H(0) = 0,5$  est souvent utilisé parceque la fonction utilisée est ainsi symétrique donc

$$H(x) = \begin{cases} 0, & \text{Si } x < 0 \\ \frac{1}{2}, & \text{Si } x = 0 \\ 1, & \text{Si } x > 0 \end{cases}$$

On retrouve l'estimation de la variance par la méthode delta dans *Staudte et Sheather (1990)* ou *Hampel et al (1986)*.

Soit la fonctionnelle  $t$  définie par:

$$t(G) \approx t(F) + \int L_t(x|F)dG(x)$$

Où  $L_t$  est la première dérivée de  $t$  par rapport à  $F$  qui est appelée la fonction d'influence (IF) définie par :

$$L_t(x|F) = \lim_{\epsilon \rightarrow 0} \frac{t((1-\epsilon)F + \epsilon H_x) - t(F)}{\epsilon}$$

Où  $H_x$  est la fonction de heaviside .

Afin d'estimer la variance de la mesure de risque, on pose  $G = F_n$  donc

$$t(F_n) \approx t(F) + \int L_t(x|F)dF_n = t(F) + \frac{1}{n} \sum_{i=1}^n L_t(x_i|F)$$

En appliquant le théorème central limite on obtient la normalité asymptotique de  $t(F_n)$  qui est

$$t(F_n) - t(F) \xrightarrow[n \rightarrow \infty]{L} N(0, V_L(F))$$

Où

$$V_L(F) = n^{-1} \text{Var}(L_t(x|F)) = n^{-1} \int L_t^2(x|F)dF(x)$$

Donc

$$V_L(F) = \frac{1}{n^2} \sum_{j=1}^n L_t^2(x_j|F_n)$$

La fonction d'influence et la variance asymptotique sont données par

$$L_t(x|F) = \int_{-\infty}^x \Psi(F(y))dy - \int_{-\infty}^{+\infty} (1 - F(y))\Psi(F(y))dy$$

$$V_L(F) = n^{-1} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Psi(F(y))\Psi(F(z))[F(\min(y,z)) - F(y)F(z)]dydz$$

Son estimateur est

$$V_L(F_n) = n^{-1} \sum_{j=1}^{n-1} \sum_{i=1}^{n-1} \Psi\left(\frac{j}{n}\right)\Psi\left(\frac{i}{n}\right)\left[\min\left(\frac{i}{n}, \frac{j}{n}\right) - \frac{i}{n} \frac{j}{n}\right](X_{(i+1)} - X_{(i)})(X_{(j+1)} - X_{(j)})$$

C'est la formule utilisé dans *Jones et Zitikis (2003)* pour l'estimation de la variance, notée par  $\hat{\sigma}^2$ :

$$\hat{\sigma}^2 = \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} c_n(i,j)\Psi(i/n)\Psi(j/n)[(X_{(i+1)} - X_{(i)})(X_{(j+1)} - X_{(j)})]$$

Où

$$c_n(i,j) = \min\left(\frac{i}{n}, \frac{j}{n}\right) - \frac{i}{n} \frac{j}{n}$$

**Exemple:**

**Proportional Hazard Transform**

L'estimateur de la mesure de risque  $PH_\xi(X)$  noté par  $\widehat{PH}_\xi(X)$  est obtenu avec les coefficients:

$$C_{ni} = \xi^{-1} \int_{(i-1)/n}^{i/n} (1-s)^{\frac{1}{\xi}-1} ds, \quad i = 1, \dots, n, \quad \xi \geq 1$$

En appliquant le théorème (2), on obtient La normalité asymptotique de  $PH_\xi(X)$  dans le corollaire suivant:

**Corollaire 1.** (3)

Si  $1 < \xi < 2$  et  $\mathbb{E}[|X|^\gamma] < \infty$  pour  $\gamma > 2\xi/(2-\xi)$ , on a:

$$\sqrt{n}(\widehat{PH}_\xi(X) - PH_\xi(X)) \longrightarrow^L N(0, \sigma_\xi^2), \quad \text{quand, } n \longrightarrow \infty$$

où

$$\sigma_{\xi}^2 = \xi^{-2} \int_0^1 \int_0^1 (\min(s,t) - st)(1-s)^{\frac{1}{\xi}-1}(1-t)^{\frac{1}{\xi}-1} d(F^{-1}(s))d(F^{-1}(t))$$

L'intervalle de confiance de niveau  $(1-\alpha)$  pour  $PH_{\xi}$  est donné par :

$$[\widehat{PH}_{\xi} - \frac{z_{\alpha/2}\sigma_{\xi}}{\sqrt{n}}, \widehat{PH}_{\xi} + \frac{z_{\alpha/2}\sigma_{\xi}}{\sqrt{n}}]$$

Ces bornes sont valables seulement si  $F^{-1}$  est connue.

Si  $F^{-1}$  n'est pas connue on remplace  $\sigma_{\xi}^2$  par son estimateur  $\hat{\sigma}_{\xi}^2$  qui est donné par :

$$\hat{\sigma}_{\xi}^2 = \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} c_n(i,j) [\xi^{-1}(1-i/n)^{\frac{1}{\xi}-1}] [\xi^{-1}(1-j/n)^{\frac{1}{\xi}-1}] [(X_{(i+1)} - X_{(i)})(X_{(j+1)} - X_{(j)})]$$

## 3.6 Approximation de l'estimateur empirique par la méthode de Monte Carlo

Le terme méthode de Monte-Carlo, désigne toute méthode visant à calculer une valeur numérique en utilisant des procédés aléatoires, c'est-à-dire des techniques probabilistes. Le nom de ces méthodes, qui fait allusion aux jeux de hasard pratiqués à Monte-Carlo, a été inventé en 1947 par *Nicholas Metropolis*, et publié pour la première fois en 1949 (Metropolis-Ulam, the Monte-Carlo method. Journal of the American Statistical Association).

La méthode de Monte-Carlo est une méthode d'approximation au sens statistique du terme. Il n'y a pas de définition précise de ce qu'est une technique de type Monte Carlo, mais la description la plus habituelle consiste à dire que les méthodes de ce type se caractérisent par l'utilisation du hasard pour résoudre des problèmes concernant le calcul d'une valeur numérique.

### 3.6.1 Description de la méthode de Monte Carlo

Pour utiliser une méthode de Monte-Carlo, on doit tout d'abord mettre sous la forme d'une espérance la quantité que l'on cherche à calculer. C'est souvent simple (calcul d'intégrale par exemple) mais peut-être plus compliqué (équations aux dérivées partielles par exemple).

Pour approximer une intégrale  $I = \int_a^b g(x)dx$ , l'idée est de générer un échantillon  $X_1, \dots, X_N$  de densité  $f(x)$  de loi uniforme sur  $[a,b]$ .

$$f(x) = \frac{1}{b-a}$$

Donc

$$\begin{aligned} I &= \int_a^b \frac{g(x)}{f(x)} f(x) dx \\ &= (b-a) \int_a^b g(x) f(x) dx \end{aligned}$$

On approxime alors  $E(g(x))$  par

$$E(g(x)) \approx \frac{1}{N} \sum_{i=1}^N g(X_i)$$

Ce qui donne

$$I \approx \frac{b-a}{N} \sum_{i=1}^N g(X_i)$$

### 3.6.2 Convergence de la méthode de Monte Carlo

C'est la loi forte des grands nombres qui permet de justifier la convergence de la méthode de Monte Carlo.

**Théorème 5.** (*Loi forte des grands nombres*)

Soit  $(X_i)_{i \geq 1}$  une suite de réalisations de la variable aléatoire  $X$ . On suppose que  $\mathbb{E}(|X|) < +\infty$ , alors pour presque tout  $\omega$

$$\mathbb{E}(X) = \lim_{n \rightarrow +\infty} \frac{1}{n} (X_1(\omega), \dots, X_n(\omega))$$

En ce qui concerne l'estimateur empirique de la mesure de risque de distorsion on a

$$\hat{R}(X) = \int_0^1 F_n^{-1}(x) dx$$

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{(X_i \leq x)}$$

$$\text{Donc } F_n(x) = \begin{cases} 0 & \text{Si } x < x_{(1)} \\ \frac{i}{n} & \text{Si } x_{(i)} \leq x < x_{(i+1)} \\ 1 & \text{Si } x \geq x_{(n)} \end{cases}$$

Cela implique que  $F_n^{-1}(\frac{i}{n}) = X_{(i)}$

En utilisant la méthode de Monte Carlo, on obtient

$$\begin{aligned}
\hat{R}(X) &= \int_0^1 F_n^{-1}(x) dx \\
&= \int \frac{F_n^{-1}(x)}{f(x)} f(x) dx, \quad f \text{ étant la fonction de densité de la loi uniforme sur } [0,1] \\
&= \int F_n^{-1}(x) f(x) dx \\
&= \mathbb{E}[F_n^{-1}(x)] \\
&\approx \frac{1}{n} \sum_{i=1}^n F_n^{-1}\left(\frac{i}{n}\right) \\
&\approx \frac{1}{n} \sum_{i=1}^n X_{(i)}
\end{aligned}$$

### 3.7 Simulation

Pour simuler cela nous allons utiliser deux échantillons. Un échantillon de loi connue (Pareto) et un échantillon de loi inconnue et de fonction de densité inconnue que nous allons estimer par la méthode des noyaux.

#### 3.7.1 Distribution de Pareto

Une variable aléatoire  $X$  est dite par définition suivre une loi de Pareto si sa fonction de répartition est donnée par:

$$P(X \leq x) = 1 - \left(\frac{x_0}{x}\right)^\alpha = F(x) \quad \text{pour } x \geq x_0 \geq 0 \quad \alpha \in \mathbb{R}^+$$

- La fonction de densité d'une loi de Pareto est alors :

$$\begin{aligned}
f(x) &= dF(x) \\
f(x) &= \frac{d}{dx} \left[ 1 - \left(\frac{x_0}{x}\right)^\alpha \right] \\
&= \alpha \frac{x_0^\alpha}{x^{\alpha+1}}
\end{aligned}$$

Où  $\alpha \in \mathbb{R}_+$  et  $x \geq x_0 \geq 0$

- L'espérance est donnée par :

$$\begin{aligned}
\mathbb{E}(X) &= \int_{x_0}^{+\infty} x f(x) dx \\
&= \int_{x_0}^{+\infty} x \alpha \frac{x_0^\alpha}{x^{\alpha+1}} dx \\
&= \frac{\alpha x_0}{\alpha - 1} \quad \text{si } \alpha > 1
\end{aligned}$$

• La variance est donnée par :

$$V(X) = \frac{\alpha x_0^2}{(\alpha - 1)^2(\alpha - 2)} \quad \text{si } \alpha \leq 2$$

### Exemple 7.

Dans ce premier exemple nous allons générer un échantillon de 100 variables aléatoires de loi de Pareto de paramètre (1,2) en utilisant la fonction `rpareto`, puis nous allons approximer  $\hat{R}(X)$ . ce qui donne la valeur de 2,469

### 3.7.2 Estimation par la méthode des noyaux

Le concept de noyau a d'abord été introduit par ROSENBLATT(1956), mais c'est CACOULOS(1966) qui a été le premier à utiliser le terme "noyau" pour désigner la fonction que l'on utilise dans les méthodes non paramétriques.

#### Noyau

##### Définition 23.

Soit  $K : \mathbf{R} \rightarrow \mathbf{R}$ , on dit que  $K$  est un noyau si et seulement si:

$$\int k(u) du = 1$$

- $K$  est dit positif si  $K(u) \geq 0 \forall u$ .
- $K$  est dit symétrique si  $K(u) = K(-u) \forall u$

#### Estimateur à noyau

##### Définition 24.

Un estimateur à noyau noté  $f_n$  de la fonction  $f$  est défini par:

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) \quad (3.1)$$

où  $(h_n)_{n \geq 1}$  est une suite de réels positifs appelés **paramètres de lissage** ou **largeur de la fenêtre**, qui tend vers 0 quand  $n$  tend vers l'infini.

- si le noyau  $K$  est une fonction de densité alors l'estimateur à noyau  $f_n$  est lui aussi une fonction de densité.

### Exemples de noyaux

Voici quelques exemples de noyaux les plus utilisés:

- le **noyau rectangulaire**:  $K(u) = \frac{1}{2}\mathbb{I}_{[-1,+1]}(u)$ . C'est celui qui donne l'estimateur de type histogramme appelé **noyau de Rosenblatt**.
- le **noyau triangulaire**:  $K(u) = (1 - |u|)\mathbb{I}_{[-1,+1]}(u)$
- le **noyau d'Epanechnikov**:  $\frac{3}{4}(1 - u^2)\mathbb{I}_{[-1,+1]}(u)$
- le **noyau de Tukey ou biweight**:  $K(u) = \frac{15}{16}(1 - u^2)^2\mathbb{I}_{[-1,+1]}(u)$
- le **noyau gaussien**:  $K(u) = \frac{1}{\sqrt{2\pi}} \exp^{-u^2/2}$ ,  $u \in \mathbb{R}$

### Exemple 8.

*Dans ce deuxième exemple nous allons utiliser un échantillon de variables aléatoires de taille  $n = 40$  de loi inconnue. On va d'abord estimer la fonction de densité de cet échantillon par la méthode des noyaux en utilisant le noyau gaussien, puis nous allons approximer  $\hat{R}(X)$   
L'estimation de la densité par la méthode des noyaux donne le graphe suivant:*

L'approximation de  $\hat{R}(X)$  donne la valeur de 36,35.

**density.default(x = x, kernel = "gaussian", n = 40)**

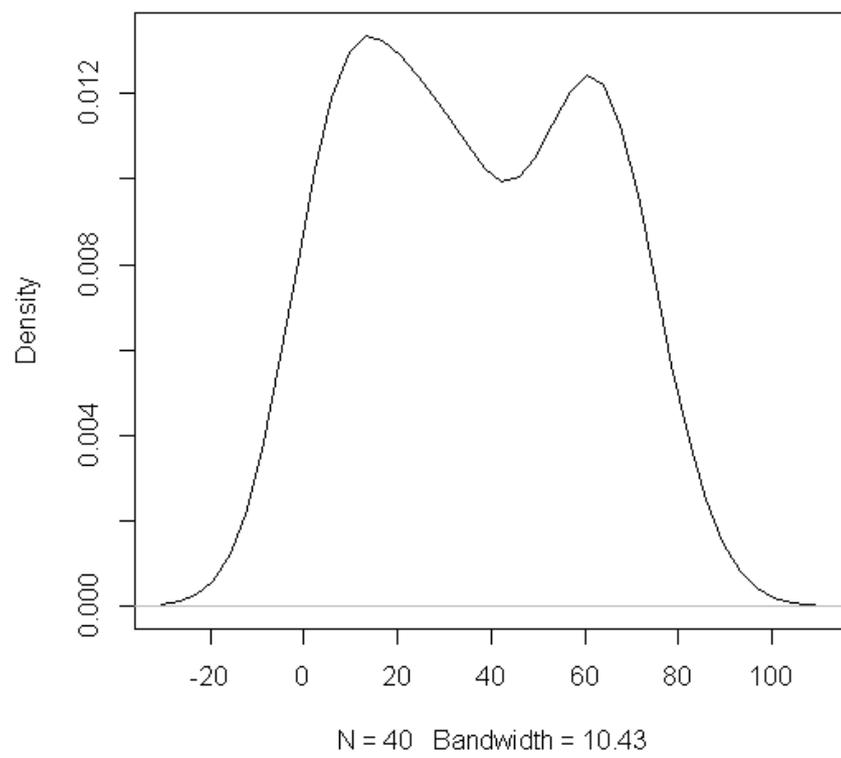


FIG. 3.1 – Estimation de la fonction de densité en utilisant le noyau gaussien

## Chapitre 4

# Application: Optimisation de portefeuille

### 4.1 Définition de portefeuille

**Définition 25.**

*Un portefeuille en finance désigne une collection d'actifs financiers (titre ou action) détenus par un établissement ou un individu. Un portefeuille comprend le plus souvent, des actifs financiers (titre de créance, action,...) mais il peut également se composer d'actifs immobiliers, de matières premières ou d'autres supports d'investissement.*

### 4.2 Diversification de portefeuille

**Définition 26.**

*La notion de diversification fait référence à la diversité des titres qui composent un portefeuille. Un portefeuille ne contenant qu'un seul titre n'est pas diversifié. La diversification de portefeuille doit permettre de se protéger contre les risques associés à la détention d'un nombre limité de titres, d'une seule catégorie d'actifs financiers ou d'un seul marché.*

### 4.3 Optimisation de portefeuille

Le problème d'optimisation de portefeuille, bien connu des gestionnaires d'actifs, consiste à composer à l'aide d'un ensemble d'actifs le portefeuille le plus rentable étant donné un enjeu de contraintes. Celles ci peuvent concerner, comme dans notre cas, un niveau de risque maximal, une contrainte de budget, etc. Le problème d'optimisation de portefeuille pose bien des questions, notamment au sujet de l'estimation de la loi jointe des rendements des actifs et de la méthode numérique de résolution.

Considérons un portefeuille constitué de  $d$  actifs.

Nous souhaitons optimiser ce portefeuille en imposant des contraintes sur le risque encouru par l'investisseur. Bien que la VaR soit une mesure de risque populaire, elle ne présente hélas ni les propriétés souhaitables dans un problème d'optimisation (elle n'est ni convexe, ni différentiable), ni les propriétés désirables d'une mesure de risque cohérente (elle n'est pas sous-additive). Pour une telle mesure de risque, la recherche d'un minimum global devient donc une tâche compliquée. Pour ces raisons, nous avons privilégié comme mesure de risque la Tail-Value-at-Risk (TVaR) du portefeuille, mesure de risque cohérente et convexe.

Dans un premier temps, nous définissons les problèmes d'optimisation. Nous verrons ensuite comment ces problèmes d'optimisation de portefeuille peuvent se ramener à des problèmes linéaires, pouvant être traités avec des méthodes algébriques telles que l'algorithme du Simplex. Notons  $\xi = (\xi^{(1)}, \dots, \xi^{(d)})$  le vecteur ligne des rendements du portefeuille,  $x$  le vecteur colonne des poids, dont le composant  $x_i$  représente le pourcentage du portefeuille investi dans l'actif  $i \in \{1, \dots, d\}$ . Considérons une mesure de risque  $\rho(x)$  générale. Cette mesure de risque peut être la VaR, la TVaR, ou toute autre mesure de risque. Nous souhaitons optimiser le portefeuille, c'est-à-dire trouver le vecteur d'allocation  $x^*$  qui optimise les performances du portefeuille, performances mesurées par son rendement futur espéré et son risque  $\rho(x)$ .

Globalement, deux problèmes d'optimisation d'allocation du portefeuille sont à considérer. Le premier problème est la maximisation du rendement espéré du portefeuille  $r(x)$  pour un certain niveau de risque  $\rho_c$  fixé :

$$L_1 = \begin{cases} \max_x r(x) \\ \rho(x) \leq \rho_c \\ x_i \geq 0, \quad i \in \{1, \dots, d\} \\ \sum_i x_i = 1 \end{cases}$$

Dans ce problème d'optimisation, on impose au gestionnaire du portefeuille un niveau de risque maximal. Le gérant va alors chercher l'allocation des actifs lui permettant de maximiser son espérance de rendement sans dépasser la limite de risque. Le second problème est la minimisation du risque pour un certain niveau de rendement  $r$  fixé :

$$L_2 = \begin{cases} \min_x \rho(x) \\ r(x) \geq r \\ x_i \geq 0, \quad i \in \{1, \dots, d\} \\ \sum_i x_i = 1 \end{cases}$$

Nous explicitons maintenant la mesure de risque  $\rho(x)$  choisie, à savoir la Tail-Value-at-Risk du portefeuille.

#### 4.4 Reformulation de la Tail-Value-at-Risk (TVaR)

Nous allons exprimer la Tail-Value-at-Risk d'une façon différente, mais équivalente, à la définition donnée précédemment. Notons  $f(x, y)$  le rendement journalier en termes de pertes du portefeuille, fonction du vecteur poids  $x$  des  $d$  positions et du vecteur aléatoire  $y$  des rendements quotidiens de perte des  $d$  actifs. Quand le vecteur position  $x$  est fixé, la perte du portefeuille  $f(x, y)$  est une variable aléatoire dont la distribution dépend de la distribution jointe de  $y$ . Supposons qu'il existe une densité de probabilité  $p(y)$  pour le vecteur aléatoire  $y$ . La probabilité que  $f(x, y)$  ne dépasse pas un seuil  $u$  donné est :

$$p[f(x, y) \leq \mu] = \int_{f(x, y) \leq \mu} p(y) dy$$

La TVaR pour une probabilité  $\alpha \in ]0,1[$  à été définie par Rockafellar et Uryasev (2001) de la manière suivante:

$$TVaR_\alpha(x) = \frac{1}{1-\alpha} \int_{f(x,y) \leq VaR_\alpha(x)} f(x,y)p(y)dy$$

Et peut également se réécrire sous la forme :

$$TVaR_\alpha(x) = VaR_\alpha(x) + \frac{1}{1-\alpha} \int_{y \in \mathbb{R}^d} [f(x,y) - VaR_\alpha(x)]^+ p(y)dy$$

L'optimisation d'un portefeuille se basant sur la Tail-Value-at-Risk nécessite le calcul de cette dernière équation. Or, un calcul exact n'est, en pratique, pas envisageable du fait de l'impossibilité d'évaluer avec exactitude la densité de probabilité  $p(y)$  modélisant le risque joint des actifs constituant le portefeuille.

## 4.5 Discrétisation et linéarisation du problème d'optimisation

### 4.5.1 Discrétisation

Dans un premier temps, nous allons montrer qu'il est possible de calculer directement la TVaR du portefeuille, sans connaître la VaR du portefeuille. Soit la fonction  $F_\alpha(x,u)$  définie de la manière suivante:

$$F_\alpha(x,u) = u + \frac{1}{1-\alpha} \int_{y \in \mathbb{R}^d} [f(x,y) - u]^+ p(y)dy$$

Rockafellar et Uryasev (2000) ont prouvé le théorème suivant :

**Théorème 6.** (Rockafellar et Uryasev (2000))

La fonction  $F_\alpha(x,u)$  est convexe et différentiable par rapport à  $x$ . La Tail-Value-at-Risk est obtenue par minimisation de la fonction  $F_\alpha(x,u)$  :

$$TVaR_\alpha(x) = \min_{u \in \mathbb{R}} F_\alpha(x,u)$$

De plus, la valeur  $u^*$  minimisant  $F_\alpha(x,u)$  est la Value-at-Risk du portefeuille pour la probabilité  $\alpha$ .

Le calcul de la TVaR est donc équivalent à un problème de minimisation, dont la solution est directement égale à la VaR du portefeuille. Cependant, nous ne connaissons pas la densité de probabilité sous-jacente  $p(y)$  du rendement du portefeuille. Une manière de simplifier ce problème est de considérer un grand nombre de réalisations  $(y_k)_{1 \leq k \leq q}$  des facteurs de risque. Ces réalisations peuvent être obtenues de différentes manières, par bootstrapping ou par simulations de Monte-Carlo. L'approximation du problème est donc la suivante :

$$F_\alpha(x,u) \approx \widetilde{F}_\alpha(x,u) = u + \frac{1}{q(1-\alpha)} \sum_{k=1}^q [f(x,y_k) - u]^+$$

Donc une expression analytique de  $p(y)$  n'est plus nécessaire. Nous pouvons également linéariser  $\widetilde{F}_\alpha(x,u)$ , car un problème d'optimisation linéaire peut être résolu grâce à des méthodes algébriques, telles que l'algorithme du Simplex ou la méthode du point intérieur. Ces méthodes permettent de nous affranchir des nombreux problèmes pratiques que posent les optimiseurs numériques non linéaires (temps de calcul long, etc...).

#### 4.5.2 Linéarisation

Nous pouvons introduire des variables auxiliaires  $(z_k)_{0 \leq k \leq q}$  de sorte que la fonction  $\widetilde{F}_\alpha(x,u)$  peut être remplacée par la fonction linéaire

$$\widetilde{F}_\alpha^l(x,u) = u + \frac{1}{q(1-\alpha)} \sum_{k=1}^q z_k$$

soumise à l'ensemble des contraintes linéaires :

$$z_k \geq f(x,y_k) - u, \quad z_k \geq 0, \quad k = 1, \dots, q, \quad u \in \mathbb{R}$$

Finalement, la TVaR peut être calculée par minimisation de la fonction linéaire  $\widetilde{F}_\alpha^l(x,u)$  par rapport à  $u$ , dont la solution  $u^*$  est égale à la Value-at-Risque du portefeuille.

En résumé, en utilisant une discrétisation et une linéarisation du calcul de la TVaR, nous avons obtenus les problèmes  $(\widetilde{L}_1)$  et  $(\widetilde{L}_2)$ , approximation des problèmes  $(L_1)$  et  $(L_2)$ , suivants :

$$\widetilde{L}_1 = \begin{cases} \max_x r(x) \\ \widetilde{F}_\alpha^l(x,u) \leq \Pi \in \mathbb{R} \\ z_k \geq f(x,y_k) - u, \quad k \in \{1, \dots, q\} \\ z_k \geq 0 \\ x_i \geq 0 \quad i \in \{1, \dots, d\} \\ \sum_i x_i = 1 \end{cases}$$

$$\tilde{L}_2 = \begin{cases} \min_x \tilde{F}_\alpha^l(x, u) \\ r(x) = r \in \mathbb{R} \\ z_k \geq f(x, y_k) - u, \quad k \in \{1, \dots, q\} \\ z_k \geq 0 \\ x_i \geq 0 \quad i \in \{1, \dots, d\} \\ \sum_i x_i = 1 \end{cases}$$

Les problèmes d'optimisation de portefeuille ( $L_1$ ) et ( $L_2$ ) deviennent alors des problèmes d'optimisation linéaire. De telles optimisations peuvent être résolues algébriquement par la méthode du Simplex par exemple. Cette méthode repose sur une transformation successive de la matrice du problème autour d'un pivot judicieusement choisi.

## 4.6 Résolution numérique

Il existe une procédure numérique pour résoudre  $L_1$  qui est fondée sur la représentation duale  $\mathcal{A}[Y] = \inf\{\mathbb{E}[YZ] : Z \in \mathcal{Z}\}$  donc ce problème peut être résolu par la procédure duale itérative :

1. initialiser  $\tilde{\mathcal{Z}} = \emptyset$
2. **Problème externe:** résoudre :

$$\begin{cases} \max_x \mathbb{E}[\xi] \cdot x \\ \text{sous les contraintes :} \\ \mathcal{A}[Y_x Z] \geq q, \quad \forall Z \in \tilde{\mathcal{Z}} \\ x^T \mathbb{I} = 1 \\ x \geq 0 \end{cases}$$

3. **Problème interne:** avec la solution  $x$ , résoudre  $v = \inf\{\mathbb{E}[Y_x Z] : Z \in \mathcal{Z}\}$ . Si  $v \geq q$ , alors stop. Sinon ajouter la fonction  $Z = \operatorname{argmin}\{\mathbb{E}[Y_x Z] : Z \in \mathcal{Z}\}$  à l'ensemble  $\tilde{\mathcal{Z}}$  et revenir à 2.

**Proposition 19.** (41)

*La procédure duale itérative s'arrête seulement aux points optimaux.*

**Proposition 20.** (41)

*Pour un espace de probabilités fini, La procédure duale itérative s'arrête après un nombre fini d'étape à la solution.*



# Conclusion

La gestion du risque dans la banque a évolué d'une façon spectaculaire ces dernières années. La VaR et la TVaR ont fait leur apparition aux côtés de la traditionnelle variance, d'un point de vue théorique. La conception de mesure de risque a évolué, notamment par la recherche d'une rationalisation de ses propriétés. De nouvelles classes de mesures de risque ont pu être déterminées dans un cadre rigoureux, mesures de risque cohérentes ou mesures de risque de Wang. Ce mémoire est consacré à l'analyse de ces mesures de risque. L'objectif principal était d'estimer les mesures de risque de distorsion par la méthode empirique. L'étude des propriétés asymptotiques de l'estimateur empirique, dont notamment la normalité asymptotique, permet de construire les bornes de confiances pour ces mesures de risque.

Dans ce nouveau cadre nous avons cherché à résoudre le problème de l'allocation optimale de richesse pour l'agent qui espère un certain rendement et veut limiter sa prise de risque.

Il serait intéressant d'étendre cette étude à de nouvelles mesures de risque de distorsion et de voir leur application en finance.



# Bibliographie

- [1 ]Carlo Acerbi. *Spectral Measures of Risk: A coherent representation of subjective risk aversion*. *Journal of Banking and Finance*, 26: 1505-1518, 2002.
- [2 ]Carlo Acerbi and Prospero Simonetti . *Portofolio optimisation with spectral measures of risk*. EconPapers, 2002.
- [3 ]C. Acerbi and D. Tasche. *On the coherence of Expected Shortfall*. *Journal of Banking and Finance*, 2(7) :1487-1503, 2002.
- [4 ]Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. *Coherent Measures of Risk*. *Mathematical finance*, 9(3):203-228, 1999.
- [5 ]Arnold, B.C, Balakrishnan. N, and Nagaraya. H.N. *A first cours in order statistics*. Wiley New york. 1992.
- [6 ]Alejandro Balbas, Beatriz Balbas, and Antonio Heras. *Optimal reinsurance with general risk measures*. *Insurance: Mathematics and Economics*, 44: 374-384, 2009.
- [7 ]Alejandro Balbas, José Garrido, Silvia Mayoral. *Propreties of Distortion Risk Measures*. 2009.
- [8 ]Jean Cristophe Boueté, Jean-François Chassagneaux. *Mesure de risque et allocation optimale de capital*. *Mémoire d'Actuariat*. ENSAE, 2005.
- [9 ]Catherine Bruneau. *Econométrie financière. Note de cour, Université Paris 1 Panthéon-Sorbonne*. 2010.
- [10 ]Vytautas Brazauskas - Bruce L. Jones - Ricardas Zitikis. *Robustification and performance evaluation of empirical risk measures and other vector-valued estimators*. *METRON - International Journal of Statistics*. vol. LXV, n. 2, pp. 175-199, 2007.
- [11 ]Vytautas Brazauskas - Bruce L. Jones - Ricardas Zitikis. *Nested L-statistics and their use in comparing the riskiness of portfolios*. *Scandinavian Actuarial Journal*. 3, 162 -179, 2007.
- [12 ]Arthur Charpentie. *Mesure de risque*. Université Rennes1 France. 2010.
- [13 ]Grzegorz Darkiewicz, Jan Dhaene, and Marc Goovaerts. *Distortion Risk Measures for Sums of Random Variable*.
- [14 ]R.A. Dana *Measures of Risk*. 2009.
- [15 ]F. Dalbaen. *Coherent Risk Measures*. *Cattedra Galileiana Scuola Normale Superiore di Pisa*, 2000.
- [16 ] M. Denuit and Arthur Charpentie. *Mathématiques de l'assurance non-vie: pricipes fondamentaux de théorie du risque*. Tome 1. *Economica*, 2004.
- [17 ]Jan Dhaene , Alexander Kukush, Daniel Linder, Qihe Tang. *Remarks on quantiles and distortion risk measures*. *Eur.Actuar.J*. 2:319-328, 2012.
- [18 ]Amit Dhurandhar and Paul Gader. *Output Distribution of Choquet Integral*.
- [19 ]Kevin Dowd, David Blake. *After VaR: the theory estimation and insurance application of quantile based risk measures*. *The journal of risk and insurance*, 73(2), 193-229. 2006.
- [20 ]N. Forcadel. *Mathématiques du risque*. *Notes de cours, Université Paris Dauphine*. 2012.

- [21] William Gehin. *Modélisation des queues de distribution des rendements des actifs financiers. Application à la mesure et à la gestion du risque de marché. Mémoire d'Actuariat EURIA*. 2011.
- [22] Vladimir Georgescu. *Measuring Risk with Extreme Value Theory. Mémoire Master. Humboldt University. Berlin*. 2009.
- [23] Jean Goubault-Lanecq. *Une introduction aux capacités, aux jeux et aux prévisions. France*. 2007.
- [24] Ani Guerdjikova. *The Choquet Integral. Springer*. 2010. C. Gouriéroux and J.M. Zakoïn. *Mesures de risque . Notes de cour. ENSAE*. 2009.
- [25] Ani Guerdjikova. *The Choquet Integral. Springer*. 2010.
- [26] Nadine Guillotin-Plantard. *Introduction au calcul stochastique. France* 2009.
- [27] Mary.R. Hardy. *An introduction to Risk Measure for actuarial applications. The casualty actuarial society and the society of actuaries. C-25-07*. 2006.
- [28] Antonio Heras, Beatriz Balbas, and Jose Luis Vilar. *Conditional Tail Expectation and premium calculation. ASTIN Bulletin*, 42: 325-342, 2012.
- [29] Mohamed Houkari. *Mesures Spectrales de Risque et Optimisation de Portefeuille. ENSAE*. 2006.
- [30] Mohamed Houkari. *Mesures de risque et couverture des marges nettes de taux d'intérêt des ressources non échéancées. Application aux dépôts à vue en ALM bancaire. Thèse de Doctorat. Université Claude Bernard Lyon 1*. 2008.
- [31] Bruce L.jones and Ricardas Zitikis. *Empirical Estimation of Risk Measures and Related Quantities, north american actuarial journal*. 8, 114-117, 2004.
- [32] Bruce L.jones - Ricardas Zitikis. *Testing for the order of risk measures: an application of L-statistics in actuarial science. METRON-north american actuarial journal . vol. LXIII, n. 2, pp. 193-211, 2003*
- [33] P. Jorion. *Value at Risk. McGraw-Hill, Oxford*. 2007.
- [34] Tomas Kaiser, and Vytautas Brazauskas. *Interval estimation of actuarial risk measures north american actuarial journal, 10(4),249-268 2006*.
- [35] Joseph.H.T. Kim and Mary.R. Hardy .*Estimating the variance of bootstrapped risk measures. ASTIN Bulletin, 39(1), 199-223. 2009*.
- [36] Joseph.H.T.Kim .*Biais correction for estimated Distortion Risk Measure using the bootstrap. Waterloo*. 2009.
- [37] Roger Koenker, Stephen Portnoy .*L-Estimation for Linear Models. Working Paper , University of Illinois at Urbana-Champaign.. 1986*.
- [38] Shigeo Kusuola .*On law invariant coherent risk measure. Advances in Mathematical Economics, 3: 83-95. 2001*.
- [39] Harry Markowitz .*Portofolio selection. The journal of finance, 7(1): pp. 77-91. 1952*.
- [40] G.Nolain, Y.Salhi, S.Werlé .*Optimisation de portefeuille selon le critère de la Value-at-Risk. ISFA. Université Lyon 1 .. 2007*.
- [41] Georg Ch. Pflug. *Mémoire: Mesure de risque multipériodes cohérentes appliquées au compte à terme. Ecole Polytechnique de Louvain* 2010.
- [42] Christophe pochet . *Mémoire: Mesure de risque multipériodes cohérentes appliquées au compte à terme. Ecole Polytechnique de Louvain* 2010.
- [43] Hadda Saidane . *Estimation des Mesures de Risque. Mémoire de magistère. Université Mouloud Mammeri, Tizi-Ouzou*. 2012.
- [44] Ekaterina N. Sereda, Efin M. Broustein, Svetozar T. Rachev, Frank J. Fabozzi, Wei Sun, and Sloyaro,S. *Distortion risk measures in protfolio optimisation. In J.Guerard the handbook*

*of potofolio construction: contemporary application of marko witz techniques, springer, new york, 649-674 2009.*

- [45] Naima Temame. *Estimation du quantile extreme de la VaR. Mémoire de magistere. Université Mouloud Mammeri, Tizi-Ouzou . 2011.*
- [46] Gérald Tenenbaum. *Calcul Integrale et théorie de la mesure. Notes de cour. Université Henri Poincarré-Nancy 1. 1994.*
- [47] Pierre-Emmanuel Thérond. *Mesure et gestion des risques d'assurance: analyse critique des futurs référentiels prudentiel et d'information financière. Thèse de Doctorat. Université Claude Bernard Lyon 1.. 2006.*
- [48] Hideatsu Tsukahara. *Mathématiqueal finance and related ISSES: Application of distortion risk measures . 2010.*
- [49] Hideatsu Tsukahara. *Estimation of Distortion Risk Measures. Int.Statistical Inst: Proc. 58th World Statistical Congress, Dublin, p. 5035. 2011.*
- [50] Hideatsu Tsukahara. *Comparative Analysis of VaR and Some Distortion Risk Measures. 6th BFS World Congress, Toronto. 2010.*
- [51] Li Zhu, Haijun Li. *Tail Distortion Risk and its asymptotic Analysis. 2011.*



# Annexe

- Graphe de la fonction de répartition empirique

```
X = c(4,-8,11,28,-5,3)
Fn= ecdf(X)
plot (Fn)
```

- Illustration du théorème de Glivenko Cantelli

```
X =rnorm(n)
n= length(X)
plot(sort(x), 1:n/n, type="s", ylim=c(0,1), xlab="", ylab="")
curve(pnorm(x,0,1), add=T, col="blue")
```

- Approximation de  $\hat{R}(X)$  par la méthode de Monte Carlo

### Exemple 7

```
library(VGAM)
X=rpseudo(100,2,5)
Fn=ecdf(X)
y=sort(X)
R(X)= mean (y)
R(X)
```

### Exemple 8

```
X=c(10,16,33,55,59,62,63,61,64,60,66,67,69,70,71,30,15,2,4,3,1,50,34,12,23,30,14,6,
8,9,24,45,30,54,32,42,44,11,80,25)
f= density(X, kernel="gaussian", n=40)
plot(f)
y=sort(X)
R(X)= mean (y)
R(X)
```