Mesures de risque

Arthur Charpentier

Université Rennes 1

arthur.charpentier@univ-rennes1.fr

http ://freakonometrics.blog.free.fr/

UNIVERSITE DE RENNES 1

Journées d'Études Statistique, Luminy, Novembre 2010.

1 Introduction

1.1 La variance comme mesure de risque ?

1.1.1 Le risque en statistique

En statistique inférentielle, on souhaite prendre une décision. en se basant sur un critère θ inconnu, mais que l'on peut estimer.

Pour juger de la pertinence de la prise de décision, on se donne classiquement une fonction de coût, $L : \Theta \times \mathcal{D} \rightarrow \mathbb{R}$, parfois appelée erreur.

Exemple1

Coût quadratique (ou L^2) en économétrie (Legendre, Gauss) : *somme des carrés des erreurs*.

Coût absolu (ou L^1, Laplace, Boscovich).

On introduit alors également une fonction de coût moyen, ou de risque,

$$R(\theta, d) = \mathbb{E}[L(\theta, d(X))] = \int L(\theta, d(x)) f_\theta(x) dx$$

où la règle de décision est $d(x)$, pour chaque résultat d'une expérience aléatoire.

La fonction de coût la plus classique est le coût quadratique,

$$L(\theta, d) = [\theta - d]^2.$$

On en déduit un critère usuel de mesure d'incertitude associé à un estimateur $\hat{\theta}$, le mean square error,

$$\text{mse}(\widehat{\theta}) = \mathbb{E}\left[(\hat{\theta} - \theta)^2\right] = \text{Var}(\hat{\theta}) + \left(\text{biais}(\hat{\theta}, \theta)\right)^2$$

Aussi, pour un estimateur sans biais, c'est la variance que permet de quantifier l'erreur associée à cette estimation.

1.1.2 Le risque en finance

> PORTFOLIO SELECTION*
>
> Harry Markowitz
>
> *The Rand Corporation*
>
> The process of selecting a portfolio may be divided into two stages. The first stage starts with observation and experience and ends with beliefs about the future performances of available securities. The second stage starts with the relevant beliefs about future performances and ends with the choice of portfolio. This paper is concerned with the second stage. We first consider the rule that the investor does (or should) maximize discounted expected, or anticipated, returns. This rule is rejected both as a hypothesis to explain, and as a maximum to guide investment behavior. We next consider the rule that the investor does (or should) consider expected return a desirable thing *and* variance of return an undesirable thing.

Harry Markowitz ([24] ou [25]) en gestion de portefeuilles.

Les agents cherchent à maximiser l'espérance d'utilité du rendement de leur portefeuille X, $\mathbb{E}(u(X))$.

Si les variations de rendements sont faibles, $X = \mathbb{E}(X) + \epsilon$ et

$$u(X) \approx u(\mathbb{E}(X)) + u'(\mathbb{E}(X))\epsilon + \frac{u''(\mathbb{E}(X))}{2}\epsilon^2$$

i.e.

$$\mathbb{E}(u(X)) \approx u(\mathbb{E}(X)) + \frac{u''(\mathbb{E}(X))}{2}\text{Var}(X)$$

car $\mathbb{E}(\epsilon) = 0$ et $\text{Var}(X) = \text{Var}(\epsilon) = \mathbb{E}(\epsilon^2))$.

Si l'agent est averse au risque, $u'' \leq 0$

Si $X \sim \mathcal{N}(\mu, \sigma^2)$, et que les agents ont utilité exponentielle (CARA), $u(x) = -\exp[-\theta x]$,

$$\mathbb{E}(u(X)) = \mathbb{E}(-\exp(-\theta X)) = -\exp\left(-\theta \mathbb{E}(X) + \frac{\theta^2}{2}\text{Var}(X)\right).$$

The concepts "yield" and "risk" appear frequently in financial writings. Usually if the term "yield" were replaced by "expected yield" or "expected return," and "risk" by "variance of return," little change of apparent meaning would result.

Variance is a well-known measure of dispersion about the expected. If instead of variance the investor was concerned with standard error, $\sigma = \sqrt{V}$, or with the coefficient of dispersion, σ/E, his choice would still lie in the set of efficient portfolios.

L'hypothèse de normalité des rendements n'est pas nouvelle,

I.—*The Mathematical Theory of Banking.*[1]

By F. Y. Edgeworth, Esq., M.A.

[Read before the British Association, September, 1886.]

Probability is the foundation of banking. The solvency and profits of the banker depend upon the probability that he will not be called upon to meet at once more than a certain amount of his liabilities. There is involved not only the calculation of averages which is an affair of arithmetic, but also the doctrine of deviations from an average, the theory of errors, which has exercised the ablest mathematicians.

[2] By *law of error* I mean always the relation between frequency and deviation which is expressed by the curve $y = \frac{1}{\sqrt{\pi}\,c} e^{-\frac{x^2}{c^2}}$

Remarque1

On s'intéresse souvent aux mesures de risque monétaires, i.e. dans la même unité que X.

On préfèrera alors l'écart-type à la variance.

L'idée que le risque peut être quantifier à l'aide d'une variance est associé à l'idée de mutualisation eassurance.

Si $S_n = X_1 + \cdots + X_n$ désigne la charge total payée sur n polices d'assurance sur une année

$$\frac{S_n - n\mathbb{E}(X_i)}{\sqrt{n}\sqrt{\text{Var}(X_i)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

i.e. en multipliant par 4 la taille du portefeuille, on multiplie par 4 le chiffre d'affaire mais le risque (correspondant à l'écart-type) n'est multiplié que par 2.

$\Longrightarrow$ intérêt à la mutualisation des risques sur des portefeuilles aussi grands que possibles.

Mais Paul Samuelson ([31]) propose le jeu suivant, de pile ou face,

gain de 200 si *pile* et perte de 100 si *face*. On décide à l'avance du nombre de lancers n que l'on fera.

$$\text{Var}(X_n) = \frac{300^2}{4n} \to 0 \text{ lorsque } n \to \infty.$$

Remarque2

Avec $n = 100$ lancers, $\mathbb{P}(X_{100} > 0) = \mathbb{P}(34 \textit{ pile} \text{ sur } 100 \text{ lancers}) \approx 99.91\%$.

Mais personne n'est prêt à jour $n = 100$ parties

SAFETY FIRST AND THE HOLDING OF ASSETS[1]

By A. D. Roy

THE PRINCIPLE OF SAFETY FIRST AND LIMITED KNOWLEDGE

In the economic world, disasters may occur if an individual makes a net loss as the result of some activity, if his resources are eroded by the process of inflation to, say, 70 per cent of their former worth, or if his income is less than what he would almost certainly obtain in some other occupation. For large numbers of people some such idea of a *disaster* exists, and the principle of Safety First asserts that it is reasonable, and probable in practice, that an individual will seek to reduce as far as is possible the chance of such a catastrophe occurring.[2] At every moment an individual's whole property is necessarily vulnerable to chance events, and the result of the current exposure to risk will determine his stake at fate's next throw.

[2] Such a principle has been applied to the theory of risk in insurance companies. See Cramér [**1**].

Given the values of m and σ for all feasible choices of action, there will exist a functional relationship between these quantities, which will be denoted by $f(\sigma, m) = 0$.[4] Since it is not possible to determine with this information, the precise probability of the final return being d or less for a given pair of values of m and σ, the only alternative open is a calculation of the upper bound of this probability. This can be done by an appeal to the Bienaymé-Tchebycheff inequality (see, for instance, H. Cramér [2]). Thus, if the final return is a random variable ξ, then[5]

$$P(|\xi - m| \geqslant m - d) \leqslant \frac{\sigma^2}{(m-d)^2}.$$

Then *a fortiori*

$$P(m - \xi \geqslant m - d) = P(\xi \leqslant d) \leqslant \frac{\sigma^2}{(m-d)^2}.$$

If then in default of minimising $P(\xi \leqslant d)$, we operate on $\sigma^2/(m-d)^2$, this is equivalent to maximising $(m-d)/\sigma$. In the subsequent analysis, we shall maximise this quantity and thus approach as near as is possible, under the circumstances, the true principle of Safety First. If ξ was distributed normally with mean m and standard deviation σ, then this line of conduct would minimise the probability of disaster itself; this fact is both interesting and reassuring.

1.2 De la comparaison des risques aux mesures de risques

cf. cours précédant : des propriétés sur un préordre $\preceq$ de comparaison entre risques permettaient de construire une mesure de risque $\mathcal{R}$, au sens où

$$X \preceq Y \Longleftrightarrow \mathcal{R}(X) \leq \mathcal{R}(Y)$$

Comme l'a montré J.M. Tallon,

Théorème1

$\preceq \subset L \times L$ satisfait les 3 axiomes de préordre complet et transitif, de continuité et d'indépendance si et seulement si il existe $u : X \to \mathbb{R}$ telle que, pour tout $\mathbb{P}, \mathbb{Q} \in L$

$$\mathbb{P} \succsim \mathbb{Q} \quad \text{ssi} \quad \sum_{x \in X} P(x)u(x) \geq \sum_{x \in X} Q(x)u(x).$$

De plus, u est unique à une transformation affine positive près.

$\preceq$ peut être caractérisé par $\mathcal{R}(X) = \mathbb{E}_{\mathbb{P}}(u(X))$, où u est une fonction d'utilité (von Neumann et Morgenstern).

Théorème2

$\preceq$ satisfait P1 à P7 si et seulement si il existe une mesure μ sur S et une fonction non constante, bornée $u : X \rightarrow \mathbb{R}$ tels que, pour tout f et g

$$f \succeq g \Leftrightarrow \int_S u(f(s))d\mu(s) \geq \int_S u(g(s))d\mu(s)$$

μ est unique et u est définie à une fonction linéaire positive près.

Sous l'axiomatique de Savage, il existe μ (ou $\mathbb{Q}$, dite subjective) telle que $\mathcal{R}(X) = \mathbb{E}_{\mathbb{Q}}(u(X))$.

Approche classique en théorie de la décision ([32] ou [27]),

Mais [3] a proposé une aximatique sur $\mathcal{R}$, introduisant une notion de cohérence, puis de convexité des mesures de risques.

Remarque1

X désigne un montant de perte.

$\mathcal{R}(X)$ est le capital à déternir pour faire face aux pertes X.

$\Longrightarrow$ de grandes valeurs de $\mathcal{R}(X)$ indiqueront que X est "*dangereux*".

2 Approche axiomatique des mesures de risques et mesures de risques usuelles

Définition1

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

- invariance en loi, $X \stackrel{\mathcal{L}}{=} Y \Longrightarrow \mathcal{R}(X) = \mathcal{R}(Y)$
- croissance $X \geq Y \Longrightarrow \mathcal{R}(X) \geq \mathcal{R}(Y)$,
- invariance par translation $\forall k \in \mathbb{R}, \Longrightarrow \mathcal{R}(X+k) = \mathcal{R}(X) + k$,
- homogénéité $\forall \lambda \in \mathbb{R}_+$, $\mathcal{R}(\lambda X) = \lambda \cdot \mathcal{R}(X)$,
- sous additivité $\mathcal{R}(X+Y) \leq \mathcal{R}(X) + \mathcal{R}(Y)$,
- convexité $\forall \beta \in [0,1]$, $\mathcal{R}(\beta\lambda X + [1-\beta]Y) \leq \beta \cdot \mathcal{R}(X) + [1-\beta] \cdot \mathcal{R}(Y)$.

Par la suite, nous ne nous intéresserons qu'aux mesures de risques invariantes en loi.

Remarque3

En science actuarielle, les mesures de risques ont été introduite sous le nom "*premium principles*", cf [26] ou [29].

Propriété d'invariance par translation ([28]), homogénéité ([33]) ou convexité ([10]).

Proposition1

Si $\mathcal{R}$ est invariante par translation $\mathcal{R}(X - \mathcal{R}(X)) = 0$.

Définition2

– une mesure de risque est dite monétaire si elle est monotone et invariante par translation
– une mesure de risque est dite convexe si elle est monétaire et convexe
– une mesure de risque est dite cohérente si elle est monétaire, homogène et sous-additive

Remarque4

Un mesure cohérente est toujours normalisée à $\mathcal{R}(0) = 0$, par homogénéité.

Si une mesure convexe est normalisée par $\mathcal{R}(0) = 0$, alors $\forall\lambda \in [0, 1]$

$$\mathcal{R}(\lambda X) = \mathcal{R}(\lambda X + [1 - \lambda]0) \leq \lambda\mathcal{R}(X)$$

et $\forall\lambda \in [1, +\infty)$,

$$\mathcal{R}(X) = \mathcal{R}\left(\frac{1}{\lambda}\lambda X\right) = \mathcal{R}\left(\frac{1}{\lambda}\lambda X + \left(1 - \frac{1}{\lambda}\right)0\right) \leq \frac{1}{\lambda}\mathcal{R}(\lambda X)$$

i.e. $\mathcal{R}(\lambda X) \geq \lambda\mathcal{R}(X)$.

Corollaire1

Si $\mathcal{R}$ est une mesure de risque monétaire et homogène, alors la convexité et la sous-additivité sont des notions équivalentes.

Définition3

Si $\mathcal{R}$ est une mesure de risque, on définie la région de risques acceptables pour la mesure $\mathcal{R}$ comme

$$\mathcal{A} = \{X, \mathcal{R}(X) \leq 0\}.$$

Réciproquement, si $\mathcal{A}$ est une région de risques acceptables, la mesure de risque induite $\mathcal{R}$ est

$$\mathcal{R}(X) = \inf\{m, X - m \in \mathcal{A}\}.$$

Proposition2

Si $\mathcal{R}$ est une mesure de risque monétaire alors $\mathcal{R}$ est convexe si et seulement si $\mathcal{A}$ est convexe.

Proposition3

Si $\mathcal{R}$ est une mesure de risque monétaire alors $\mathcal{R}$ est positivement homogène si et seulement si $\mathcal{A}$ est un cône.

On dispose d'un théorème de représentation suivant

Théorème3

$\mathcal{R}$ est une mesure de risque monétaire convexe si et seulement si $\forall X$ bornée $(X \in Ł^\infty)$,

$$\mathcal{R}(X) = \max_{\mathbb{Q}\in\mathcal{M}} \{\mathbb{E}_{\mathbb{Q}} - \alpha(\mathbb{Q})\}$$

où $\mathcal{M}$ est l'ensemble des mesures additives et normalisées, et

$$\alpha(\mathbb{Q}) = \sup_{X\in\mathcal{A}} \{\mathbb{E}_{\mathbb{Q}}(X)\}.$$

On notera que $\mathcal{M}$ contient plus que des mesures de probabilité.

Remarque5

α est la conjugée de Legendre-Fenchel de $\mathcal{R}$: si f est une fonction convexe, on pose

$$f^\star(x) = -\inf_{\lambda}\{f(x) - \lambda x\} = \sup_{\lambda}\{\lambda x - f(x)\}$$

telle que $(f^\star)^\star = f$. Alors $\mathcal{R}(X) = \sup_{\mathbb{Q}\in\mathcal{B}} \{\mathbb{E}_{\mathbb{Q}} - \mathcal{R}^\star(\mathbb{Q})\}$ où $\mathcal{B}$ est l'ensemble des mesures bornéées ($\mathcal{M}$ est la boule unité de $\mathcal{B}$).

Remarque6

Les mesures cohérentes peuvent s'écrire

$$\mathcal{R}(X) = \max_{\mathbb{Q} \in \mathcal{Q}} \{\mathbb{E}_{\mathbb{Q}}(X)\}$$

où $\mathcal{Q} = \{\mathbb{Q} \in \mathcal{M}, \alpha(\mathbb{Q}) = 0\}$.

Notons que des réécritures de la propriété de sous-additivité ont été proposées dans la littérature (en particulier [30] ou [11]).

- additivité pour les risques comonotones si, $\forall X$ et Y comonotones, $\mathcal{R}(X+Y) = \mathcal{R}(X) + \mathcal{R}(Y)$,
- corrélation maximale (par rapport à une mesure μ) si $\forall X$,

$$\mathcal{R}(X) = \sup\{\mathbb{E}(X \cdot U) \text{ où } U \sim \mu\}$$

- cohérence forte si $\forall X$ et Y, $\sup\{\mathcal{R}(\tilde{X}+\tilde{Y})\} = \mathcal{R}(X) + \mathcal{R}(Y)$, pour $\tilde{X} \stackrel{\mathcal{L}}{=} X$ et $\tilde{Y} \stackrel{\mathcal{L}}{=} Y$.

Remarque7

La VaR et la TVaR sont addivites pour les risques comonotones.

Proposition4

Si $\mathcal{R}$ est une mesure de risque convexe, les trois propriétés suivantes sont équivalentes

- $\mathcal{R}$ est fortement cohérente,
- $\mathcal{R}$ est additive pour des risques comontones,
- $\mathcal{R}$ est une mesure de corrélation maximale.

Enfin, [20] a montré le résultat suivant dès lors que $\mathcal{R}$ est invariante en loi,

Proposition5

Une mesure de risque cohérente $\mathcal{R}$ est additive pour des risques comonotones si et seulement s'il exsite une fonction décroissante positive g sur $[0, 1]$ telle que

$$\mathcal{R}(X) = \int_0^1 g(t)F^{-1}(1-t)dt$$

où $F(x) = \mathbb{F}(X \leq x)$.

2.1 L'équivalent certain

Définition4

Soit u une fonction d'utilité concave et croissante, alors l'équivalent certain $\mathcal{R}(X)$ associé à une perte X est

$$u(\mathcal{R}(X)) = \mathbb{E}[u(X)] \text{ soit } \mathcal{R}(X) = u^{-1}\left(\mathbb{E}[u(X)]\right)$$

Exemple2

utilité exponentielle, $u(x) = 1 - \exp[-\theta x]$, CARA, i.e. $-u''(c)/u'(c) = \theta$. Alors l'équivalent certain associé à une perte X est $\mathcal{R}(X) = \dfrac{1}{\theta} \log\left(\mathbb{E}_{\mathbb{P}}[e^{-\theta X}]\right)$

mesure de risque entropique ([12], [23], [8])

$$\mathcal{R}(X) = \sup_{\mathbb{Q}<<\mathbb{P}} \left\{ \mathbb{E}_{\mathbb{Q}}[-X] - \frac{1}{\theta} H(\mathbb{Q}|\mathbb{P}) \right\}$$

$H(\mathbb{Q}|\mathbb{P})$ est l'entropie relative (distance de Kullback-Leibler), i.e. $H(\mathbb{Q}|\mathbb{P}) = \mathbb{E}_{\mathbb{Q}}\left(\log \dfrac{d\mathbb{Q}}{d\mathbb{P}}\right)$.

2.2 La Value-at-Risk

La Value-at-Risk est apparu (sous ce nom) dans les années 90, en réponse à de nombreux désastres qui ont touché les marchés de capitaux à cette période (see [18]).

BOX 1–3

THE ORIGINS OF VAR

Till Guldimann can be viewed as the creator of the term *value at risk* while head of global research at J.P. Morgan in the late 1980s. The risk-management group had to decide whether *fully hedged* meant investing in long bonds, thus generating stable *earnings*, or investing in cash, thus keeping the market *value* constant. The bank decided that "value risks" were more important than "earnings risks," paving the way for VAR.

At that time, there was much concern about managing the risks of derivatives properly. The Group of Thirty, which had a representative from J.P. Morgan, provided a venue for discussing best risk-management practices. The term found its way through the G-30 report published in July 1993. Apparently, this was the first widely publicized appearance of the term *value at risk*.

Définition5

On appelle Value-at-Risk de niveau $\alpha \in (0,1)$ le quantile de niveau α,

$$\mathcal{R}_\alpha(X) = \text{VaR}(X;\alpha) = x_\alpha \text{ où } \mathbb{P}(X \leq x_\alpha) = \alpha,$$

ou encore

$$\text{VaR}(X;\alpha) = \inf\{x, \mathbb{P}(X \leq x) \geq \alpha\} = F_X^{-1}(\alpha) = Q(\alpha).$$

Remarque8

$\mathcal{R}_\alpha(X)$ est une croissante en α.

Lemme1

$\forall \alpha \in (0,1)$, si g est strictement *croissante* et continue à gauche,

$$\text{VaR}(g(X),\alpha) = F_{g(X)}^{-1}(\alpha) = g\left(F_X^{-1}(\alpha)\right) = g(\text{VaR}(X;\alpha)),$$

alors que si g est fonction strictement *décroissante*, continue à droite, et si F_X est bijective,

$$\text{VaR}(g(X),\alpha) = F_{g(X)}^{-1}(\alpha) = g\left(F_X^{-1}(1-\alpha)\right) = g(\text{VaR}(X;1-\alpha)).$$

Toutefois, la VaR n'est pas sous-additive.

Exemple1

risques indépendants suivant des lois de Pareto, $X \sim \mathcal{P}ar(1,1)$ et $Y \sim \mathcal{P}ar(1,1)$, i.e.

$$\mathbb{P}(X > t) = \mathbb{P}(Y > t) = \frac{1}{1+t}, \quad t > 0.$$

Alors

$$\text{VaR}(X;\alpha) = \text{VaR}(Y;\alpha) = \frac{1}{1-\alpha} - 1.$$

De plus, on peut

$$\mathbb{P}[X + Y \leq t] = 1 - \frac{2}{2+t} + 2\frac{\ln(1+t)}{(2+t)^2}, \quad t > 0.$$

Or

$$\mathbb{P}\big[X + Y \leq 2\text{VaR}[X;\alpha]\big] = \alpha - \frac{(1-\alpha)^2}{2}\ln\left(\frac{1+\alpha}{1-\alpha}\right) < \alpha$$

donc $\forall \alpha$

$$\text{VaR}(X;\alpha) + \text{VaR}(Y;\alpha) < \text{VaR}(X+Y;\alpha).$$

2.3 La Tail-Value-at-Risk

Définition1

La Tail Value-at-Risk notée $\text{TVaR}(X;\alpha)$ est définie par

$$\text{TVaR}(X;\alpha) = \frac{1}{1-\alpha}\int_\alpha^1 \text{VaR}(X;t)\,dt.$$

$\Longrightarrow$ la Tail-VaR est la moyenne des VaR de niveau supérieur à α.

Remarque1

il existe une fonction de répartition $\widetilde{F}_X$ (transformée de Hardy-Littlewood de F_X, [15]), telle que $\forall\alpha$,

$$\widetilde{F}_X^{-1}(\alpha) = \text{TVaR}(X;\alpha).$$

Si $\widetilde{X}$ a pour loi $\widetilde{F}_X$,

$$\text{TVaR}(X;\alpha) = \text{VaR}(\widetilde{X};\alpha).$$

La Tail-VaR de X est donc la VaR de la transformée de Hardy-Littlewood de X.

Notons que, $\text{TVaR}[X;0] = \mathbb{E}[X]$. Et comme

$$\text{TVaR}[X;\alpha] = \frac{1}{1-\alpha}\left\{\mathbb{E}[X] - \int_0^\alpha \text{VaR}[X;\xi]\, d\xi\right\}. \qquad (1)$$

on en déduit que la TVaR est une fonction croissante en α.

De plus

$$\text{TVaR}[X;\alpha] \geq \text{TVaR}[X;0] = \mathbb{E}[X].$$

$\Longrightarrow$ la Tail-VaR contient toujours un chargement de sécurité.

La CTE est la la perte attendue sachant que la VaR au niveau α est dépassée, i.e. "*perte moyenne dans les pires* $1-\alpha\%$ *des cas*".

Définition2

La **Conditional Tail Expectation** au niveau α, notée $\text{CTE}[X;\alpha]$, est

$$\text{CTE}[X;\alpha] = \mathbb{E}\Big[X\Big|X > \text{VaR}[X;\alpha]\Big].$$

La CVaR est la valeur moyenne des pertes qui excèdent la VaR, i.e. il s'agit de l'excédent moyen de sinistre au-delà de la VaR,

Définition3

La **Conditional-VaR** au niveau α, notée $\text{CVaR}[X;\alpha]$, est

$$\begin{aligned}\text{CVaR}[X;\alpha] &= \mathbb{E}\Big[X-\text{VaR}[X;\alpha]\Big|X > \text{VaR}[X;\alpha]\Big]\\ &= e_X\big(\text{VaR}[X;\alpha]\big)\\ &= \text{CTE}[X;\alpha]-\text{VaR}[X;\alpha].\end{aligned}$$

L'ES est la prime stop-loss dont la franchise est $\text{VaR}[X;\alpha]$,

Définition4

L'Expected shortfall au niveau α, notée $\text{ES}[X;\alpha]$, est

$$\text{ES}[X;\alpha] = \mathbb{E}\big[(X - \text{VaR}[X;\alpha])_+\big].$$

Proposition1

Quel que soit le niveau de probabilité $\alpha \in (0,1)$, les identités suivantes sont vérifiées :

$$\text{TVaR}[X;\alpha] = \text{VaR}[X;\alpha] + \frac{1}{1-\alpha}\text{ES}[X;\alpha], \quad (2)$$

$$\text{CTE}[X;\alpha] = \text{VaR}[X;\alpha] + \frac{1}{\overline{F}_X(\text{VaR}[X;\alpha])}\text{ES}[X;\alpha]. \quad (3)$$

Proposition1

La CTE et la TVaR coïncident pour des risques dont la fonction de répartition est continue, i.e.

$$\text{CTE}[X;\alpha] = \text{TVaR}[X;\alpha], \qquad \alpha \in (0,1). \quad (4)$$

La TVaR est invariante par translation et homogène.

De la même manière, l'homogénéité de la VaR garantit l'homogénéitéde la TVaR.

Lemme1

Soient le risque X et le niveau de perte x tels que $\overline{F}_X(x) > 0$. Quel que soit l'événement aléatoire E tel que $\mathbb{P}[E] = \overline{F}_X(x)$, on a $\mathbb{E}[X|E] \leq \mathbb{E}[X|X > x]$.

Cette proposition garantie que la TVaR est sous-additive lorsque les risques sont continus (la TVaR et la CTE coïncident)

$$
\begin{aligned}
\text{TVaR}[X+Y;\alpha] &= \mathbb{E}\big[X\big|X+Y > \text{VaR}[X+Y;\alpha]\big] \\
&\quad +\mathbb{E}\big[Y\big|X+Y > \text{VaR}[X+Y;\alpha]\big] \\
&\leq \mathbb{E}\big[X\big|X > \text{VaR}[X;\alpha]\big] \\
&\quad +\mathbb{E}\big[Y\big|Y > \text{VaR}[Y;\alpha]\big] \\
&= \text{TVaR}[X;\alpha] + \text{TVaR}[Y;\alpha].
\end{aligned}
$$

De la même manière, la TVaR est monotone, puisque lorsque $\mathbb{P}[X \leq Y] = 1$

$$\begin{aligned} \text{TVaR}[Y;\alpha] &= \mathbb{E}\big[Y\big|Y > \text{VaR}[Y;\alpha]\big] \\ &\geq \mathbb{E}\big[Y\big|X > \text{VaR}[X;\alpha]\big] \\ &\geq \mathbb{E}\big[X\big|X > \text{VaR}[X;\alpha]\big] \\ &= \text{TVaR}[X;\alpha]. \end{aligned}$$

Proposition6

La TVaR est cohérente pour les risques continus, et coïncide alors avec la CTE.

Proposition7

La TVaR est la plus petite mesure de risque majorant la VaR qui soit cohérente

2.4 La transformée d'Esscher

La mesure de risque d'Esscher consiste à prendre la *prime pure* de la transformée d'Esscher du risque initial.

Définition5

La mesure de risque d'Esscher de paramètre $h > 0$ du risque X, notée $\text{Es}[X; h]$, est

$$\text{Es}[X; h] = \frac{\mathbb{E}[X \exp(hX)]}{M_X(h)} = \frac{d}{dh} \ln M_X(h).$$

Remarque2

$\text{Es}[X; h]$ est la valeur espérée de la transformée d'Esscher X_h de X,

$$\text{Es}[X; h] = \mathbb{E}[X_h] = \int_{\xi \in \mathbb{R}^+} \xi dF_{X,h}(\xi).$$

Proposition2

$\text{Es}[X; h]$ est une fonction croissante de h.

Aussi $\forall h > 0$, $\text{Es}[X;h] \geq \text{Es}[X;0] = \mathbb{E}[X]$ i.e. la mesure de risque d'Esscher contient un chargement de sécurité.

La mesure de risque d'Esscher n'est pas homogène (sauf dans le cas trivial $h = 0$). Elle est invariante par translation mais pas monotone.

Exemple2

Considérons les risques X et Y tels que

$$\Pr[X = 0, Y = 0] = \Pr[X = 0, Y = 3] = \Pr[X = 6, Y = 6] = \frac{1}{3}.$$

Dans ce cas, $\Pr[X \leq Y] = 1$, mais

$$\text{Es}[X;1/2] = 5.4567 > \text{Es}[Y;1/2] = 5.2395.$$

Exemple3

le théorème fondamental de valorisation d'actifs ([2], [17] et [9]) nous dit qu'il n'y a pas d'opportunité d'arbitrage s'il existe une mesure risque neutre équivalente à la mesure originale $\mathbb{P}$ telle que le prix d'un actif est l'espérance (sous cette probabilité) du payoff actualisé. I.e le prix d'une option européenne d'achat est

$$e^{-rT}\mathbb{E}_{\mathbb{Q}}((S_T - K)_+).$$

ce prix pouvait être écrit comme une mesures de risques d'Esscher sous certaines hypothèses sur la diffusion du sous-jacent (S_t), [13].

2.5 Les mesures de risque de Wang

Définition6

Nous appellerons fonction de distorsion toute fonction croissante $g : [0,\ 1] \rightarrow [0,\ 1]$ telle que $g(0) = 0$ et $g(1) = 1$.

Remarque9

Comme [38], on considère ici des variables aléatoires (strictement) positives, i.e. $F(0) = \mathbb{P}(X \leq 0) = 0$. Le cas général sera traité dans la partie suivante.

Définition6

La mesure de risque de Wang associée à la fonction de distorsion g, notée $\mathcal{R}_g$, est

$$\mathcal{R}_g\left(X\right) = \int_0^{\infty} g\left(1 - F_X(x)\right)\ dx = \int_0^{\infty} g\left(\overline{F}_X(x)\right)\ dx \qquad (5)$$

Remarque3

La mesure de risque associée à la fonction de distorsion $g(q) = q$ correspond à l'espérance mathématique $\mathbb{E}[X]$.

Si $g(q) \geq q \; \forall q \in [0,1]$, alors $\mathcal{R}_g[X] \geq \mathbb{E}[X]$.

Si $g_1(q) \leq g_2(q) \; \forall q \in [0,1]$, $\mathcal{R}_{g_1}[X] \leq \mathcal{R}_{g_2}[X]$.

En substituant $\int_0^{\overline{F}_X(x)} dg(\alpha)$ à $g\left(\overline{F}_X(x)\right)$ on obtient

Proposition3

La mesure de risque de Wang associée à la fonction de distorsion g peut s'écrire

$$\mathcal{R}_g[X] = \int_0^1 \text{VaR}[X; 1-\alpha] \, dg(\alpha). \tag{6}$$

Ainsi, les mesures de risque de Wang sont des moyennes pondérées de VaR.

Exemple4

Si

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

alors

$$\mathcal{R}_{g_\alpha}[X] = \text{VaR}[X; \alpha]$$

i.e. la VaR au niveau α est une mesure de Wang particulière correspondant à une fonction de distorsion passant de 0 à 1 en $1 - \alpha$. Ici, g_α est la fonction de répartition associée à une masse de Dirac en $1 - \alpha$.

Exemple5

Si

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

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

$$\mathcal{R}_{g_\alpha}[X] = \frac{1}{1-\alpha}\int_0^{1-\alpha} \text{VaR}[X; 1 - \xi]d\xi = \text{TVaR}[X; \alpha].$$

Ici g_α est une fonction de répartition, correspondant à une loi uniforme sur $[0, 1 - \alpha]$.

En revanche, ce n'est généralement pas le cas de l'expected shortfall ou de la CTE.

Exemple6

En prenant

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

nous obtenons

$$\mathcal{R}_g[X] = \int_{x\geq 0} (1 - \{F_X(x)\}^\xi) dx.$$

Si ξ est entier, $\mathcal{R}_g[X]$ peut être interprétée comme la valeur attendue du maximum $M_\xi = \max\{X_1, \ldots, X_\xi\}$ d'un ensemble de ξ variables aléatoires indépendantes et de même loi que X. i.e. $\mathcal{R}_g[X] = \mathbb{E}[M_\xi]$.

Exemple7

Considérons la fonction de distorsion

$$g(x) = x^{1/\xi}, \quad \xi \geq 1.$$

La mesure de risque PH, introduite dans [37] et [36] est donnée par

$$\text{PH}_\xi[X] = \mathcal{R}_g[X] = \int_{x \geq 0} \{\overline{F}_X(x)\}^{1/\xi} dx.$$

Notons que pour $\xi = 1$, $\text{PH}_1[X] = \mathbb{E}[X]$.

mesure de risque $\mathcal{R}$	fonction de distorsion g
VaR	$g(x) = \mathbb{I}[x \geq p]$
Tail-VaR	$g(x) = \min\{x/p, 1\}$
PH	$g(x) = x^p$
Dual Power	$g(x) = 1 - (1-x)^{1/p}$
Gini	$g(x) = (1+p)x - px^2$
Transformation exponentielle	$g(x) = (1-p^x)/(1-p)$

Table 1 – Principales mesures de risques par distortion, où p est une constante comprise entre 0 et 1.

Proposition4

Les mesures de risque de Wang sont homogènes, invariantes par translation et monotones.

Si la fonction de distorsion g est concave, la fonction $x \mapsto g(\overline{F}_X(x))$ est continue à droite et est donc la fonction de survie d'une certaine variable aléatoire.

Alors $\mathcal{R}_g\left[X\right]$ est effectivement une espérance mathématique, d'une variable Y dont la fonction de survie vaut $\overline{F}_Y(y) = g(\overline{F}_X(y)))$.

Proposition5

Lorsque la fonction de distorsion est concave, la mesure de risque correspondante est sous-additive.

Corollaire1

Les mesures de risque de Wang correspondant à des fonctions de distorsion concaves sont cohérentes.

Proposition6

Quel que soit le niveau de probabilité $0 < \alpha < 1$ et le risque X, nous avons

$$\text{TVaR}[X;\alpha] = \min\left\{\mathcal{R}_g[X]\middle|g \text{ est concave et } \mathcal{R}_g[X] \geq \text{VaR}[X;\alpha]\right\}. \tag{7}$$

2.6 Les mesures de risque par distorsion

Les mesures de Wang sont un cas particulier de mesures de risques par distorsion

Définition7

On appelle mesure de risque par distortion la quantité

$$\mathcal{R}(X) = \int_0^1 F^{-1}(1-u)dg(u)$$

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

Proposition8

$\mathcal{R}(X)$ peut se réécrire

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

Remarque10

Si $g(q) = q \ \forall q \in [0,1]$

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

Remarque11

Soit $\mathbb{Q}$ la *mesure* induite par la transformation g sur $\mathbb{P}$, i.e.

$$\mathbb{Q}([a,b]) = G(\mathbb{P}([a,b])).$$

La croissance sur g sur $[0,1]$ permet de construire une capacité.

Exemple8

Considérons la fonction de répartition $G(x) = x^k$. On appelera mesure de risque à hasard proportionnel la mesure induite par cette transformation,

$$\mathcal{R}(X;k) = \int_0^1 F^{-1}(1-u)ku^{k-1}du =$$

Lorsque $k < 1$, la fonction G est concave.

Remarque12

Les mesures spectrales ([1]) : une fonction spectrale (ou fonction d'aversion pour le risque) est $\phi : [0,1] \to \mathbb{R}_+$, décroissante, telle que $\int_0^1 \phi(t)dt = 1$. La mesure de risque spectrale induite est

$$\mathcal{R}(X) = \int_0^1 F_X^{-1}(t)\phi(t)dt.$$

Ces mesures de risques sont cohérentes.

Les mesures de distortion de fonction de distortion g concave sont des mesures spectrales, avec $\phi = g'$.

3 Comparaison entre risques et mesures de risques

Il existe un lien fondamental entre la comparaison des risques, et les mesures des risques.

3.1 Ordre induit par la Value-at-Risk

Définition7

Y, X sera considéré moins dangereux que Y sur base de la comparaison des VaR, noté $X \preceq_{\text{VaR}} Y$, si

$$\text{VaR}[X;\alpha] \leq \text{VaR}[Y;\alpha], \forall \alpha \in (0,1).$$

La relation $\preceq_{\text{VaR}}$ a été introduite par [21].

Elle est davantage connue sous de dominance stochastique à l'ordre 1, et est parfois notée $\preceq_{st}$, $\preceq_1$ ou $\preceq_{FSD}$ (cf [22] ou [27]).

Proposition7

Etant données deux variables aléatoires X et Y,

$$\begin{aligned} X \preceq_{\text{VaR}} Y &\iff F_X(t) \geq F_Y(t), \forall t \in \mathbb{R}, \\ &\iff \overline{F}_X(t) \leq \overline{F}_Y(t), \forall t \in \mathbb{R}. \end{aligned}$$

Proposition8

Etant données deux variables aléatoires X et Y,

$$X \preceq_{\text{VaR}} Y \Longleftrightarrow \mathbb{E}[g(X)] \leq \mathbb{E}[g(Y)] \text{ pour toute fonction croissante } g,$$

pour autant que les espérances existent.

Démonstration. Il suffit de se rappeler que si $U \sim \mathcal{U}(0,1)$ alors $X \overset{\mathcal{L}}{=} \text{VaR}[X;U]$. Aussi, pour tout g :

$$\mathbb{E}[g(X)] = \mathbb{E}\Big[g\big(\text{VaR}[X;U]\big)\Big] = \int_0^1 g\big(\text{VaR}[X;u]\big)du.$$

Le résultat annoncé s'obtient alors simplement en écrivant

$$\begin{aligned} \mathbb{E}[g(X)] &= \int_0^1 g\big(\text{VaR}[X;u]\big)du \\ &\leq \int_0^1 g\big(\text{VaR}[Y;u]\big)du = \mathbb{E}[g(Y)]. \end{aligned}$$

$\square$

Proposition9

Etant données deux variables aléatoires X et Y,

$$X \preceq_{\text{VaR}} Y \Longleftrightarrow \mathbb{E}[g(X)] \leq \mathbb{E}[g(Y)] \text{ pour toute fonction } g \text{ telle que } g' \geq 0,$$

pour autant que les espérances existent.

$X \preceq_{\text{VaR}} Y$ lorsque les fonctions de survie de X et de Y se dominent mutuellement. Une condition *suffisante* est que les densités de probabilité ne se croisent qu'une seule fois.

Proposition10

Quels que soient les risques X et Y, si $f_X(t) \geq f_Y(t)$ pour $t < c$ et $f_X(t) \leq f_Y(t)$ pour $t > c$ alors $X \preceq_{\text{VaR}} Y$.

Nous pourrions nous intéresser à la comparaison des risques sachant qu'ils excèdent un certain niveau t, i.e.

$$[X|X > t] \preceq_{\text{VaR}} [Y|Y > t]$$

quel que soit le niveau t. Cela n'est pas forcément vrai lorsque $X \preceq_{\text{VaR}} Y$,

$$X \preceq_{\text{VaR}} Y \not\Rightarrow [X|X > t] \preceq_{\text{VaR}} [Y|Y > t] \text{ pour tout } t.$$

Exemple3

Considérons par exemple le cas où $X \sim \mathcal{U}(0,3)$ et Y possède la densité

$$f_Y(x) = \frac{1}{6}\mathbb{I}_{]0,1]}(x) + \frac{1}{2}\mathbb{I}_{]1,2]}(x) + \frac{1}{3}\mathbb{I}_{]2,3[}(x).$$

Alors $X \preceq_{\text{VaR}} Y$, mais $[X|X > 1] \sim \mathcal{U}(1,3)$ et $[Y|Y > 1]$ possède la densité

$$f_Y^*(x) = \frac{3}{5}\mathbb{I}_{]1,2]}(x) + \frac{2}{5}\mathbb{I}_{]2,3[}(x)$$

de telle sorte que

$$[Y|Y > 1] \preceq_{\text{VaR}} [X|X > 1].$$

Proposition11

Etant donnés deux risques X et Y, $[X|X>t] \preceq_{\text{VaR}} [Y|Y>t]$ quel que soit $t \in \mathbb{R}$ si, et seulement si

$$\Longleftrightarrow \quad t \mapsto \frac{\overline{F}_Y(t)}{\overline{F}_X(t)} \text{ est croissante}$$

$$\Longleftrightarrow \quad \overline{F}_X(u)\overline{F}_Y(v) \geq \overline{F}_X(v)\overline{F}_Y(u) \quad \text{quels que soient} \quad u \leq v.$$

On peut rapprocher cette comparaison des taux de hasard, comme le montre le résultat suivant.

Proposition12

Etant donnés deux risques X et Y, $[X|X>t] \preceq_{\text{VaR}} [Y|Y>t]$ quel que soit $t \in \mathbb{R}$ si, et seulement si $r_X(t) \geq r_Y(t)$ quel que soit t.

Remarque4

La méthode de comparaison de lois de probabilité dont il est question dans ces propriétés est souvent appelée l'ordre du taux de hasard (hazard rate order) et notée $\preceq_{hr}$).

Supposons qu'on déflate le taux de hasard d'un facteur ξ, i.e. que l'on passe d'un taux r_X à un taux r_{X^*} donné par

$$r_{X^*}(t) = \frac{r_X(t)}{\xi} \leq r_X(t), \text{ pour } \xi \geq 1,$$

nous obtenons

$$\overline{F}_{X^*}(t) = \exp\left(-\int_0^t \frac{r_X(s)}{\xi} ds\right) = \{\overline{F}_X(t)\}^{1/\xi}.$$

Dès lors, $PH[X;\xi] = \mathbb{E}[X^*]$. Le mesure de risque PH consiste donc à remplacer le risque initial X par un risque transformé X^* dont le taux de hasard a été déflaté, et à calculer ensuite l'espérance associée à X^*. Nous avons

$$[X|X > t] \preceq_{\text{VaR}} [X^*|X^* > t] \text{ quel que soit } t > 0.$$

3.2 Rapport de vraisemblance et principe d'Esscher

Nous pourrions encore songer à imposer

$$[X|a \leq X \leq a+h] \preceq_{\text{VaR}} [Y|a \leq Y \leq a+h]$$

quel que soit le niveau a et l'accroissement $h > 0$. Ceci correspond à la situation d'un réassureur qui aurait à couvrir la tranche $(a, a+h]$ d'un risque X, i.e. qui s'exposerait à une perte de

$$X_{(a,a+h]} = \begin{cases} 0 \text{ si } X < a \\ X - a \text{ si } a \leq X < a+h \\ h \text{ si } a+h \leq X, \end{cases}$$

où a est la rétention et h la portée.

On peut établir le résultat suivant.

Proposition2

Considérons les variables aléatoires X et Y, toutes deux continues ou discrètes, possédant les fonctions de densité f_X et f_Y. Si

$$\frac{f_X(t)}{f_Y(t)} \quad \text{décroît sur l'union des supports de } X \text{ et de } Y \tag{8}$$

(en prenant par convention $a/0$ égal à $+\infty$ lorsque $a > 0$), ou, de manière équivalente, si

$$f_X(u)f_Y(v) \geq f_X(v)f_Y(u) \quad \forall u \leq v. \tag{9}$$

Alors, $[X|a \leq X \leq a+h] \preceq_{\text{VaR}} [Y|a \leq Y \leq a+h]$ quel que soit le niveau a et l'accroissement $h > 0$

Démonstration. Considérons $a < b$. L'inégalité stochastique $[X|a \leq X \leq b] \preceq_{\text{VaR}} [Y|a \leq Y \leq b]$ garantit que

$$\frac{\Pr[u \leq X \leq b]}{\Pr[a \leq X \leq b]} \leq \frac{\Pr[u \leq Y \leq b]}{\Pr[a \leq Y \leq b]} \quad \text{lorsque} \quad u \in [a,b].$$

Il suit alors

$$\frac{\Pr[a \leq X < u]}{\Pr[u \leq X \leq b]} \geq \frac{\Pr[a \leq Y < u]}{\Pr[u \leq Y \leq b]} \quad \text{lorsque} \quad u \in [a, b].$$

C'est-à-dire

$$\frac{\Pr[a \leq X < u]}{\Pr[a \leq Y < u]} \geq \frac{\Pr[u \leq X \leq b]}{\Pr[u \leq Y \leq b]} \quad \text{lorsque} \quad u \in [a, b].$$

En particulier, pour $u < b \leq v$,

$$\frac{\Pr[u \leq X < b]}{\Pr[u \leq Y < b]} \geq \frac{\Pr[b \leq X \leq v]}{\Pr[b \leq Y \leq v]}.$$

Dès lors, lorsque X et Y sont continues,

$$\frac{\Pr[a \leq X < u]}{\Pr[a \leq Y < u]} \geq \frac{\Pr[b \leq X \leq v]}{\Pr[b \leq Y \leq v]} \quad \text{lorsque} \quad a < u \leq b \leq v.$$

Si nous passons à la limite pour $a \to u$ et $b \to v$ nous obtenons (9). La preuve dans le cas discret est similaire. $\square$

Les conditions (8) et (9), apparemment techniques et peu intuitives, sont généralement faciles à établir dans les modèles paramétriques.

Remarque5

La méthode de comparaison de lois de probabilité dont il est question dans la Proposition 2 est encore appelée ordre du rapport de vraisemblance, et notée $\preceq_{lr}$.

Notons X_h la transformée d'Esscher de X. Le rapport des densités de probabilités associées à X et X_h est proportionnel à $\exp(-hx)$, qui est clairement décroissant en x. Ceci indique que

$$[X|a \leq X \leq b] \preceq_{\text{VaR}} [X_h|a \leq X_h \leq b],$$

quels que soient $a < b$.

Le résultat suivant montre sous quelles conditions les mesures de risque d'Esscher relatives à deux risques X et Y sont uniformément ordonnées.

Proposition13

Si $[X|a \leq X \leq b] \preceq_{\text{VaR}} [Y|a \leq Y \leq b]$ quels que soient $a < b$ alors $\text{Es}[X;h] \leq \text{Es}[Y;h]$ $\forall h > 0$.

Démonstration. Nous savons en vertu de la Proposition 2 que l'inégalité

$$f_X(u)f_Y(v) \geq f_X(v)f_Y(u)$$

est satisfaite $\forall u \leq v$. En multipliant les deux membres de cette inégalité par

$$\frac{\exp(hu)}{M_X(h)}\frac{\exp(hv)}{M_Y(h)}$$

on obtient la même inégalité pour les fonctions de densité de X_h et Y_h, d'où on tire que $[X_h|a \leq X_h \leq b] \preceq_{\text{VaR}} [Y_h|a \leq Y_h \leq b]$, ce qui donne le résultat annoncé. $\square$

3.3 Comparaison uniforme des TVaR

Nous introduisons ici une nouvelle méthode de comparaison des risques, basée sur les Tail-VaR.

Définition8

Quelles que soient les variables aléatoires X et Y de moyenne finie, X sera considéré moins dangereux que Y sur base de la comparaison des TVaR, ce qui se notera dorénavant $X \preceq_{\text{TVaR}} Y$, lorsque

$$\text{TVaR}[X;\alpha] \leq \text{TVaR}[Y;\alpha] \text{ pour tout } \alpha \in [0,1].$$

La restriction à des risques de moyenne finie garantit l'existence des TVaR, et donc la correction de la définition de $\preceq_{\text{TVaR}}$. Dorénavant, nous ne comparons donc que des risques dont la prime pure est finie (il est bon de noter que de nombreux résultats que nous établirons dans la suite de cette section dépendent de cette hypothèse). Cette restriction distingue également $\preceq_{\text{VaR}}$ et $\preceq_{\text{TVaR}}$: en effet, $\preceq_{\text{VaR}}$ est définie quelles que soient les variables en présence, ce qui n'est pas le cas pour $\preceq_{\text{TVaR}}$.

La relation $\preceq_{\text{TVaR}}$ est très ancienne (elle est la digne héritière de la relation dite de majorization entre vecteurs numériques, étudiée dans les années 1930). Les actuaires l'appellent encore ordre stop-loss (noté $\preceq_{sl}$), notion intimement liée à la dominance stochastique du deuxième ordre des économistes (souvent notée $\preceq_2$ ou $\preceq_{SSD}$). La relation $\preceq_{\text{TVaR}}$ est mieux connue des probabilistes sous le nom d'ordre convexe croissant (noté $\preceq_{icx}$).

Nous utiliserons encore la relation notée $\preceq_{CX}$, qui restreint $\preceq_{\text{TVaR}}$ aux couples de variables aléatoires de même moyenne.

Définition9

Quelles que soient les variables aléatoires X et Y,

$$X \preceq_{CX} Y \iff \begin{cases} \mathbb{E}[X] = \mathbb{E}[Y], \\ X \preceq_{\text{TVaR}} Y, \end{cases}$$

$$\iff \begin{cases} \text{TVaR}[X;0] = \text{TVaR}[Y;0], \\ \text{TVaR}[X;\alpha] \leq \text{TVaR}[Y;\alpha] \text{ pour tout } \alpha \in (0,1). \end{cases}$$

La relation $\preceq_{CX}$ est connue sous le nom d'ordre convexe parmi les probabilistes (noté $\preceq_{cx}$). Elle est intimement liée à d'autres relations, comme l'ordre de Lorenz.

Remarque6

Notons que $\preceq_{\text{VaR}}$ ne permettait pas de comparer des variables aléatoires de même moyenne. En effet,

$$\left.\begin{array}{l} X \preceq_{\text{VaR}} Y \\ \mathbb{E}[X] = \mathbb{E}[Y] \end{array}\right\} \Rightarrow X \overset{\mathcal{L}}{=} Y.$$

Afin de prouver ce résultat, il suffit d'examiner l'identité

$$\mathbb{E}[Y] - \mathbb{E}[X] = \int_{x=0}^{+\infty} \underbrace{\{\Pr[Y > x] - \Pr[X > x]\}}_{\geq 0 \text{ pour tout } x} dx = 0,$$

qui entraîne $\Pr[Y > x] = \Pr[X > x]$ pour tout x et achève la vérification.

3.3.1 Tail-VaR et primes stop-loss

Les relations $\preceq_{\text{TVaR}}$ et $\preceq_{CX}$ introduites ci-dessus peuvent encore être interprétées en termes de primes stop-loss, comme le montrent les résultats suivants.

Proposition3

Soient deux variables aléatoires X et Y telles que $\mathbb{E}(X)=\mathbb{E}(Y)$ Alors, chacune des deux inégalités suivantes est équivalente à $\mathbb{E}[(X-t)_+]\leq\mathbb{E}[(Y-t)_+]$ $\forall t\in\mathbb{R}$:

$$\int_0^p F_X^{-1}(u)\,du \geq \int_0^p F_Y^{-1}(u)\,du \qquad \forall p\in[0,1]$$

et

$$\int_p^1 F_X^{-1}(u)\,du \leq \int_p^1 F_Y^{-1}(u)\,du \qquad \forall p\in[0,1].$$

Corollaire2

Soient X et Y deux risques de même moyenne μ finie. Alors,

$$X\preceq_{CX}Y \Longleftrightarrow \mathbb{E}[(X-t)_+]\leq\mathbb{E}[(Y-t)_+]\forall t\in\mathbb{R}.$$

On peut montrer qu'un résultat du même type vaut pour $\preceq_{\text{TVaR}}$, à savoir

$$X \preceq_{\text{TVaR}} Y \Longleftrightarrow \mathbb{E}[(X-t)_+] \leq \mathbb{E}[(Y-t)_+] \forall t \in \mathbb{R}.$$

Aussi, $\preceq_{CX}$ et $\preceq_{\text{TVaR}}$ peuvent s'interpréter à l'aide des primes de réassurance relatives à un traité stop-loss.

3.3.2 Tail-VaR et fonctions convexes

$\preceq_{\text{VaR}}$ était liée aux fonctions croissantes, $\preceq_{CX}$ et $\preceq_{\text{TVaR}}$ sont liées aux fonctions convexes et convexes croissantes,

Proposition14

Etant données deux variables aléatoires X et Y de moyennes finies,

$$X \preceq_{CX} Y \Longleftrightarrow \mathbb{E}[g(X)] \leq \mathbb{E}[g(Y)] \text{ pour toute fonction convexe } g,$$

pour autant que les espérances existent, et

$$X \preceq_{CX} Y \Longleftrightarrow \mathbb{E}[g(X)] \leq \mathbb{E}[g(Y)] \text{ pour toute fonction } g \text{ telle que } g'' \geq 0,$$

pour autant que les espérances existent.

Proposition15

Etant données deux variables aléatoires X et Y de moyennes finies, on a

$$X \preceq_{\text{TVaR}} Y \Longleftrightarrow \mathbb{E}[g(X)] \leq \mathbb{E}[g(Y)] \text{ pour toute fonction convexe croissante } g,$$

pour autant que les espérances existent, et

$$X \preceq_{\text{TVaR}} Y \Longleftrightarrow \mathbb{E}[g(X)] \leq \mathbb{E}[g(Y)] \text{ pour toute fonction } g \text{ telle que } g' \geq 0 \text{ et } g" \geq 0$$

pour autant que les espérances existent.

La caractérisation suivante a été obtenue par Strassen en 1965 ([35]),

Caractérisation1

Etant donnés deux risques X et Y, $X \preceq_{CX} Y$ si, et seulement si, on peut trouver deux variables aléatoires $\widetilde{X}$ et $\widetilde{Y}$ telles que $\widetilde{X} \stackrel{\mathcal{L}}{=} X$, $\widetilde{Y} \stackrel{\mathcal{L}}{=} Y$ et $\mathbb{E}[\widetilde{Y}|\widetilde{X}] = \widetilde{X}$.

Corollaire3

En particulier, ajouter un aléa supplémentaire centré ϵ à un risque X rend la situation plus risquée. Formellement, quel que soit le risque X et la variable aléatoire ϵ indépendante de X telle que $\mathbb{E}[\epsilon]=0$,

$$X \preceq_{CX} X+\epsilon.$$

Ici aussi il existe des conditions *suffisantes* simples pour $\preceq_{CX}$ et $\preceq_{\text{TVaR}}$,

Proposition16

Soient X et Y deux variables aléatoires telles que $\mathbb{E}[X]=\mathbb{E}[Y]$. S'il existe une constante c telle que $F_Y(x)\geq F_X(x)\forall x<c$, et $F_Y(x)\leq F_X(x)\forall x>c$, alors $X\preceq_{CX} Y$.

De la même manière, on peut prouver le résultat suivant.

Proposition17

Soient X et Y deux variables aléatoires telles que $\mathbb{E}[X]\leq\mathbb{E}[Y]$. S'il existe une constante c telle que $F_Y(x)\geq F_X(x)\forall x<c$ et $F_Y(x)\leq F_X(x)\forall x>c$, alors $X\preceq_{\text{TVaR}} Y$.

4 Estimation de la Value-at-Risk

Remarque13

l'estimation du quantile (puis des mesures de risques) à partir d'un échantillon indépendant. Le cas des processus est abordé dans [14].

4.1 Estimation paramétrique

Hypothèse classique en finance : rendements gaussiens.

Si $X \sim \mathcal{N}(\mu, \sigma)$, le quantile de niveau α est

$$q(\alpha) = \mu + \Phi^{-1}(\alpha)\sigma,$$

où $\Phi^{-1}(\alpha)$ est tabulé, i.e. 1.64 si $\alpha = 90\%$ 1.96 si $\alpha = 95\%$, ou encore 2.57 si $\alpha = 99.5\%$.

Exemple9

Modèle Gaussien intéressant intéressant pour les risques agrégés dans des modèles à facteurs. Supposons que $X_1, \cdots, X_d$ soient d risques et que

$$X_i = a_i + b_i Z + u_i$$

où les bruits u_i sont supposés indépendants entre eux, et de Z, avec $u_i \sim \mathcal{N}(0, \sigma_i^2)$, avec $Z \sim \mathcal{N}(0,1)$, alors $X = X_1 + \cdots + X_d$ est gaussien.

De plus, les VaR pour chacuns des risques sont

$$\text{VaR}_i(\alpha) = a_i + \Phi^{-1}(\alpha)\sqrt{b_i^2 + \sigma_i^2}$$

mais surtout

$$\text{VaR}(\alpha) = a + \Phi^{-1}(\alpha)\sqrt{b^2 + \sigma^2}$$

où $a = a_1 + \cdots + a_d$, $b = b_1 + \cdots + b_d$, et $\sigma^2 = \sigma_1^2 + \cdots + \sigma_d^2$. On notera que

$$\text{VaR}(\alpha) - \sum_{i=1}^{d} \text{VaR}_i(\alpha) = \Phi^{-1}(\alpha)\left(\underbrace{\sqrt{b^2+\sigma^2} - \sum_{i=1}^{d}\sqrt{b_i^2+\sigma_i^2}}_{\text{négatif}}\right).$$

Définition8

Etant donné $\{X_1, \cdots, X_n\}$, l'estimateur gaussien du quantile de niveau α est

$$\widehat{q}_n(\alpha) = \widehat{\mu} + \Phi^{-1}(\alpha)\widehat{\sigma}, \text{ où } \widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ et } \widehat{\sigma} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \widehat{\mu})^2}.$$

Toutefois, il peut être intéressant d'utiliser des approximations comme [7], i.e.

$$Q(X;\alpha) \sim \mathbb{E}(X) + z_\alpha \sqrt{V(X)}, \tag{10}$$

où

$$\widehat{z}_\alpha = \Phi^{-1}(\alpha) + \frac{\zeta_1}{6}[\Phi^{-1}(\alpha)^2 - 1] + \frac{\zeta_2}{24}[\Phi^{-1}(\alpha)^3 - 3\Phi^{-1}(\alpha)] - \frac{\zeta_1^2}{36}[2\Phi^{-1}(\alpha)^3 - 5\Phi^{-1}(\alpha)],$$

où ζ_1 désigne la skewness de X, et ζ_2 la kurtosis en excès,

$$\zeta_1 = \frac{\mathbb{E}([X - \mathbb{E}(X)]^3)}{\mathbb{E}([X - \mathbb{E}(X)]^2)^{3/2}} \text{ et } \zeta_1 = \frac{\mathbb{E}([X - \mathbb{E}(X)]^4)}{\mathbb{E}([X - \mathbb{E}(X)]^2)^2} - 3. \tag{11}$$

Définition9

Etant donné un échantillon $\{X_1, \cdots, X_n\}$, l'estimation de Cornish-Fisher du quantile de niveau α est

$$\widehat{q}_n(\alpha) = \widehat{\mu} + \widehat{z}_\alpha \widehat{\sigma}, \text{ where } \widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ and } \widehat{\sigma} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n} (X_i - \widehat{\mu})^2},$$

où

$$z_\alpha = \Phi^{-1}(\alpha) + \frac{\widehat{\zeta}_1}{6}[\Phi^{-1}(\alpha)^2 - 1] + \frac{\widehat{\zeta}_2}{24}[\Phi^{-1}(\alpha)^3 - 3\Phi^{-1}(\alpha)] - \frac{\widehat{\zeta}_1^2}{36}[2\Phi^{-1}(\alpha)^3 - 5\Phi^{-1}(\alpha)],$$

avec $\widehat{\zeta}_1$ l'estimateur usuelle de la skewness, et $\widehat{\zeta}_2$ l'estimateur usuelle de la kurtosis

$$\widehat{\zeta}_1 = \frac{\sqrt{n(n-1)}}{n-2} \frac{\sqrt{n}\sum_{i=1}^{n}(X_i - \widehat{\mu})^3}{\left(\sum_{i=1}^{n}(X_i - \widehat{\mu})^2\right)^{3/2}} \text{ et } \widehat{\zeta}_2 = \frac{n-1}{(n-2)(n-3)}\left((n+1)\widehat{\zeta}_2' + 6\right)$$

où

$$\widehat{\zeta}_2' = \frac{n\sum_{i=1}^{n}(X_i - \widehat{\mu})^4}{\left(\sum_{i=1}^{n}(X_i - \widehat{\mu})^2\right)^2} - 3.$$

Remarque14

Ce modèle a est beaucoup utilisé en gestion de portefeuille *moyenne*-VaR, en utilisant les comoments d'ordre 3 et 4.

En fait, de manière plus générale, si $F_X \in \mathcal{F} = \{F_\theta, \theta \in \Theta\}$ (que l'on supposera continue), $q_X(\alpha) = F_\theta^{-1}(\alpha)$, et donc un estimateur naturel du quantile de niveau α est

$$\widehat{q}_X(\alpha) = F_{\widehat{\theta}}^{-1}(\alpha), \tag{12}$$

où $\widehat{\theta}$ est un estimateur de θ (par maximum likelihood, par la méthode des moments, etc).

Exemple10

Pour une loi de Pareto

$$F(x) = \mathbb{P}(X \leq x) = 1 - \left(\frac{1}{x}\right)^{\alpha}, \text{ pour } x \geq 1, \alpha \geq 0.$$

La vraisemblance s'écrit

$$\log \mathcal{L}(\alpha) = \sum_{i=1}^{n} \log\left(\frac{\alpha}{x_i^{\alpha+1}}\right) = n \log \alpha - (\alpha+1) \sum_{i=1}^{n} \log x_i,$$

dont le maximum est obtenu pour

$$\widehat{\alpha}_1 = \left(\frac{1}{n}\sum_{i=1}^{n} \log(x_i)\right)^{-1}$$

En notant que $\mathbb{E}(X) = \dfrac{\alpha}{\alpha-1}$ pour $\alpha > 1$, on en déduit l'estimateur de la méthode des moments, $\widehat{\alpha}_2 = \dfrac{\overline{X}}{\overline{X}-1}$. Notons nfin que $\log[1-F(x)] = \log \overline{F}(x) = 1 - \alpha \log x$, autrement dit, l'estimateur (par moindre carrés) de la pente de la droite de régression peut être utilisé,

$$\widehat{\alpha}_3 = -\frac{-n\sum_{i=1}^n \log X_{i:n}\cdot\log\frac{n-i}{n} + \sum_{i=1}^n \log X_{i:n}\cdot\sum_{i=1}^n \log\frac{n-i}{n}}{n\sum_{i=1}^n [\log X_i]^2 - \left[\sum_{i=1}^n \log X_i\right]^2}.$$

Ces trois estimateurs permettent de construire quatre estimateurs de quantiles.

$$\widehat{Q}_p = (1-p)^{-1/\widehat{\alpha}}.$$

Mais les modèles paramétriques induisent inévitablement des erreurs de modèles.

La Figure de gauche montre l'estimation d'un quantile sous hypothèse de loi Gamma et de loi lognormale, et celle de droite l'estimation d'un quantile sous hypothèse de normalité et de loi de Student.

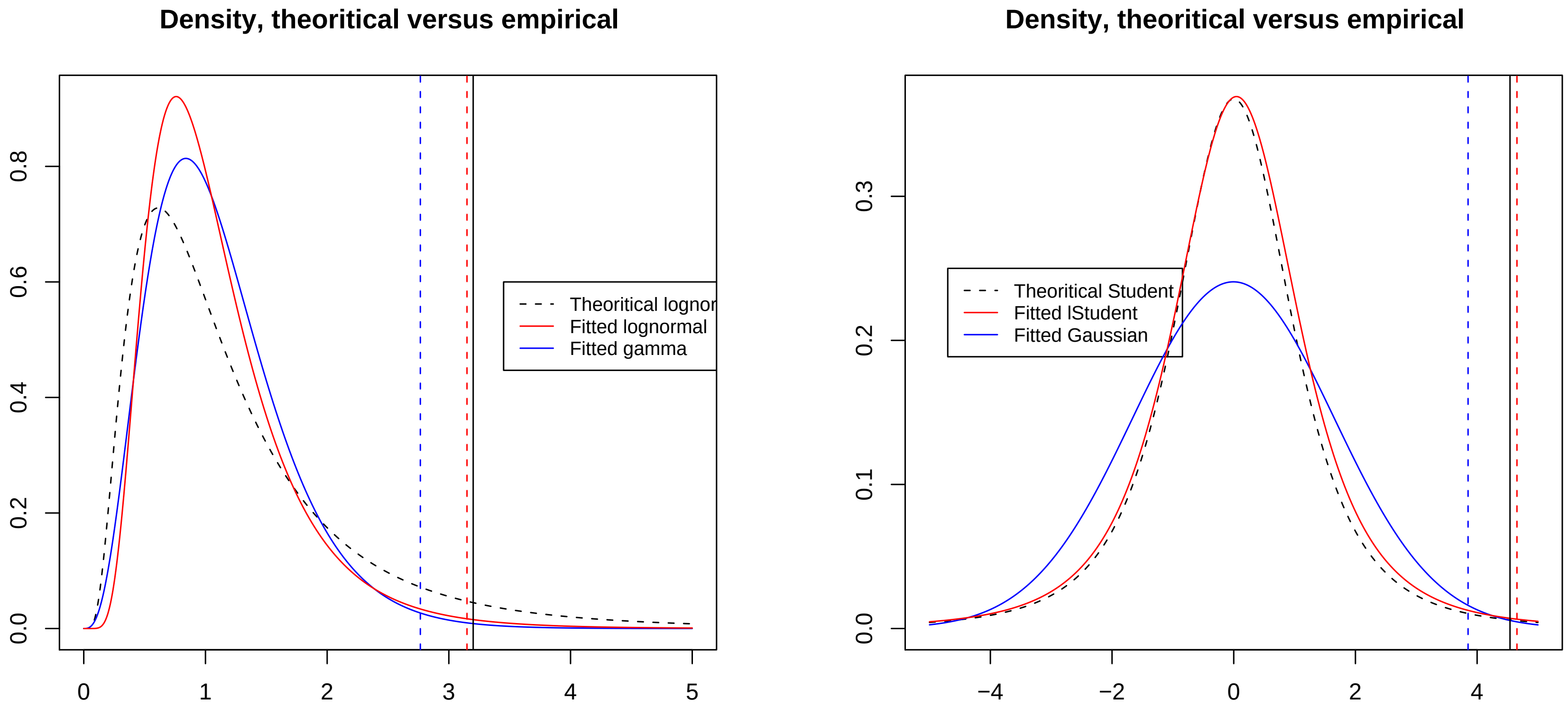

Figure 1 – Estimation de Value-at-Risk et erreur de modèle.

4.2 Estimation nonparamétrique

Afin d'éviter les erreurs dans le choix de la famille paramétrique $\mathcal{F}$, il est naturel d'envisager des estimateurs nonparamétriques.

Définition10

La fonction de répartition empirique F_n, constuite à partir de $\{X_1, \ldots, X_n\}$ est

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

A x fixé, $nF_n(x)$ suit une loi binomiale, centrée sur $nF(x)$ et de variance $nF(x)[1 - F(x)]$.

Par la loi des grands nombres, $F_n(x)$ converge presque sûrement vers $F(x)$ quand $n \to \infty$, et le théorème de Glivenko-Cantelli assure que la convergence est uniforme, i.e. $\sup\{|F_n(x) - F(x)|\}$ tend presque sûrement vers 0 lorsque $n \to \infty$.

De plus, on a une normalité asymptotique à l'aide du théorème central limite,

$$\sqrt{n}[F_n(x) - F(x)] \xrightarrow{\mathcal{L}} \mathcal{N}(0, F(x)[1 - F(x)]),$$

lorsque $n \to \infty$.

Et plus généralement, si l'on considère une convergence au sens des distributions finidimensionnelles, $\sqrt{n}[F_n(x) - F(x)]$ converge vers un processus Gaussien $G(x)$ centré, de fonction de covariance

$$\text{Cov}(G(x_1), G(x_2)) = F(x_1)[1 - F(x_2)] \text{ pour } x_1 < x_2.$$

Ce processus admet la même loi que $(B_{F(x)})$ où $(B_t)_{t\in[0,1]}$ est un pont Brownien.

On en déduit alors la loi de Kolmogorov Smirnov, qui assure que $\sqrt{n}\sup\{|F_n(x) - F(x)|\}$ converge vers $\sup\{B_t\}$.

Définition11

La fonction quantile empirique, construite à partir de $\{X_1, \ldots, X_n\}$ est $Q_n(\alpha) = \inf\{x, F_n(x) \geq \alpha\} = F_n^{-1}(\alpha)$.

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

$$Q_n(\alpha) = X_{k:n} \text{ où } \frac{k-1}{n} < \alpha \leq \frac{k}{n}.$$

Remarque15

De même que la moyenne peut être obtenu par minimisation de la norme L^2, les quantiles peuvent être obtenus par minimisation.

$$Q_n(\alpha) = \underset{x\in\mathbb{R}}{\text{argmin}} \frac{1}{n}\sum_{i=1}^{n} H_\alpha(X_i - x)$$

où $H_\alpha(x) = x[\alpha - \mathbf{1}(x < 0)]$.

Le théorème de Glivenko Cantelli garantie la convergence (forte) de $Q_n(\alpha)$ vers $Q(\alpha)$.

Proposition9

Si $X_1, \cdots, X_n$ est un échantillon i.i.d. de loi absolument continue de densité f, et que $f(Q(\alpha)) > 0$, alors

$$\sqrt{n}[Q_n(\alpha) - Q(\alpha)] \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{\alpha[1-\alpha]}{f^2(Q(\alpha))}\right),$$

lorsque $n \to \infty$.

De manière plus générale, on notera que $\sqrt{n}[Q_n(\alpha) - Q(\alpha)]$ converge vers un processus Gaussien $G(\alpha)$ centré, de fonction de covariance

$$\text{Cov}(G(\alpha_1), G(\alpha_2)) = \frac{\alpha_1[1-\alpha_2]}{f(Q(\alpha_1))f(Q(\alpha_2))} \text{ pour } \alpha_1 < \alpha_2.$$

Pour terminer avec les propriétés du quantile empirique $Q_n(\alpha)$ rappelons le résultat suivant, correspondant à l'expansion de Bahadur-Kiefer

Proposition10

Si $X_1, \cdots, X_n$ est un échantillon i.i.d. de loi absolument continue admettant une densité f telle que $f(Q(\alpha)) > 0$, alors

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

où, presque sûrement, $Z_n = O\left(n^{-3/4}[\log n]^{1/2}[\log(\log n)]^{1/4}\right)$ selon ([5]), amélioré sous la forme $Z_n = O(n^{-3/4}[\log(\log n)]^{3/4})$ par [19].

Définition12

La fonction de répartition empirique lissée, construite à partir de $\{X_1, \ldots, X_n\}$ est

$$\widehat{F}_n(x) = \int_{-\infty}^{+\infty} \mathbb{K}\left(\frac{t-x}{h}\right) dF_n(t) \quad = \quad \frac{1}{nh}\sum_{i=1}^{n}\int_{-\infty}^{x} k\left(\frac{X_i - t}{h}\right) dt$$

$$= \quad \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right)$$

où $K(x) = \displaystyle\int_{-\infty}^{x} k(t)dt$, k étant un noyau et $h > 0$ une fenêtre de lissage.

Classiquement, deux techniques assez différentes ont été considérées dans la littérature.

La première idée est de considérer une combinaison linéaire de statistiques d'ordre.

Le quantile empirique est simplement

$$Q_n(p) = F_n^{-1}(p) = X_{[np]:n} \text{ où } [\cdot] \text{ désigne la partie entière.} \quad (13)$$

il ne dépend que d'une unique observation. Il est alors naturel de considérer un lissage entre deux observations, si np n'est pas entier. Le quantile empirique pondéré est alors défini par

$$Q_n(p) = (1-\gamma)X_{[np]:n} + \gamma X_{[np]+1:n} \text{ où } \gamma = np - [np].$$

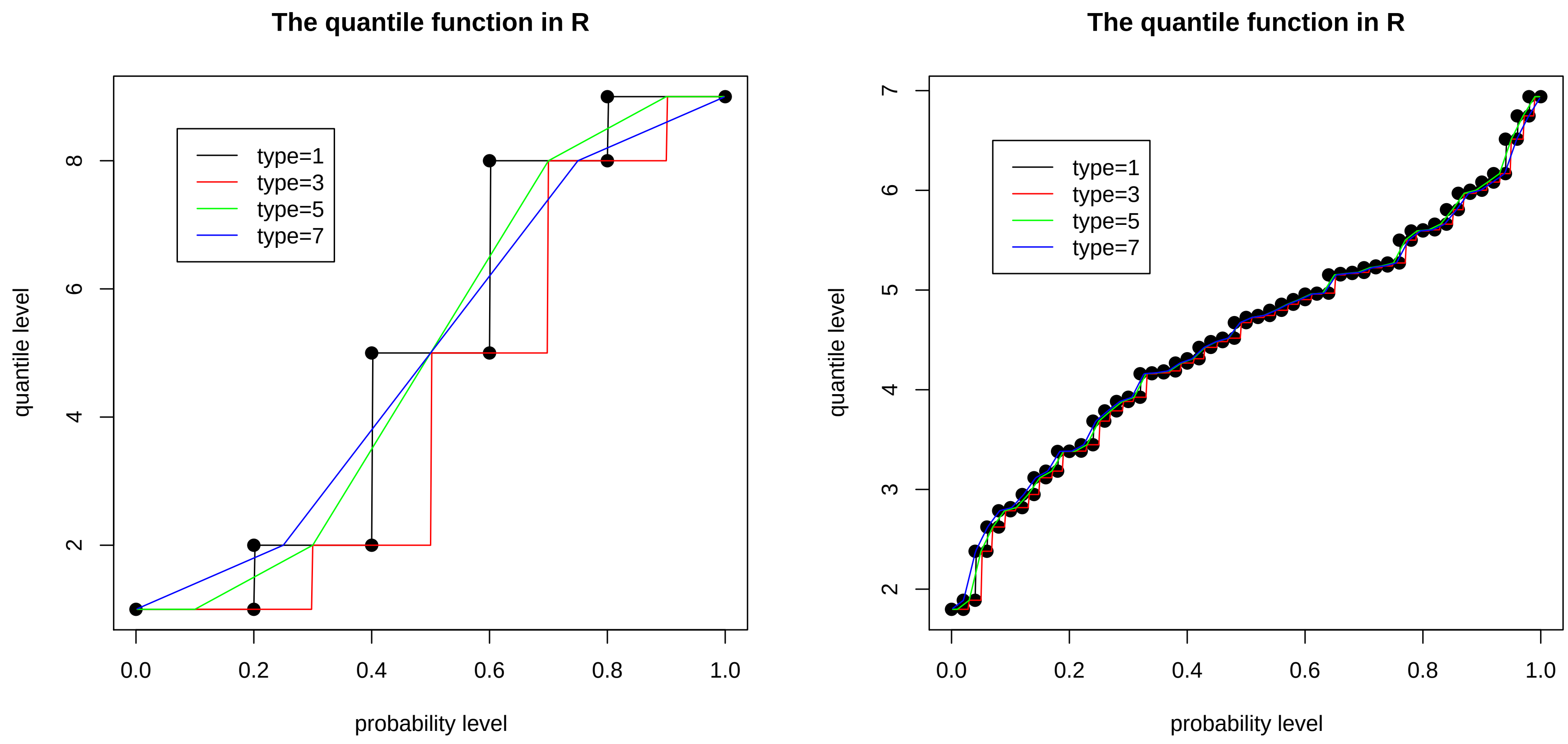

Figure 2 – Les quantiles empiriques usuels, sous R.

Afin d'augementer l'efficience, il est possible, plus généralement de considérer des L statistiques,

$$Q_n(p) = \sum_{i=1}^{n} W_{i,n,p} X_{i:n} = \sum_{i=1}^{n} W_{i,n,p} F_n^{-1}\left(\frac{i}{n}\right) = \int_0^1 F_n^{-1}(t) k(p,h,t) dt \qquad (14)$$

où F_n est la fonction de répartition empirique et où k est un noyau et h est une fenêtre. On peut alors écrire,

$$Q_n(p) = \sum_{i=1}^{n} \left[\int_{\frac{i-1}{n}}^{\frac{i}{n}} k\left(\frac{t-p}{h}\right) dt \right] X_{i:n} = \sum_{i=1}^{n} \left[K\left(\frac{\frac{i}{n}-p}{h}\right) - K\left(\frac{\frac{i-1}{n}-p}{h}\right)\right] X_{i:n} \qquad (15)$$

où $\mathbb{K}(x) = \displaystyle\int_{-\infty}^{x} k(t)dt$. L'idée est alors de donner davantage de poids aux statistiques d'ordre $X_{i:n}$ pour lesquelles i est proche de pn.

Par exemple, l'estimateur de Harrell-Davis (introduit dans [16]) est défini par

$$Q_n(p) = \sum_{i=1}^{n} \left[\int_{\frac{(i-1)}{n}}^{\frac{i}{n}} \frac{\Gamma(n+1)}{\Gamma((n+1)p)\Gamma((n+1)q)} y^{(n+1)p-1}(1-y)^{(n+1)q-1} \right] X_{i:n}.$$

Une approche alternative respose sur l'écriture $F \circ Q(\alpha) = \alpha$.

Si $\widehat{F}_n$ est un estimateur continu de F, alors un estimateur naturel de $Q(\alpha)$ est $\widehat{Q}_n(\alpha)$ tel que $\widehat{F}_n \circ \widehat{Q}_n(\alpha) = \alpha$, obtenu par l'algorithme de Gauss-Newton. [4], [34] ou [6] on suggéré cette approche.

4.3 Estimation semiparamétrique

Une approche alternative est d'utiliser la théorie des valeurs extrêmes (cf. exposé de demain matin).

Une approche peut être d'utiliser le théorème de Pickands-Balkema-de Haan : si u est suffisant grand la loi conditionnelle de $Y-u$ sachant $Y>u$ suit une de Pareto généralisée, de paramètres ξ et β (ces derniers pouvant être estimés par maximum de vraisemblance). Aussi, en posant $u=Y_{n-k:n}$, avec k suffisement grand, notons $\widehat{\beta}_k$ et $\widehat{\xi}_k$ les estimateurs du maximum de vraisemblance de la loi de Pareto généralisé sur l'échantillon $\{Y_{n-k+1:n}-Y_{n-k:n},...,Y_{n:n}-Y_{n-k:n}\}$,

$$\widehat{Q}(Y,\alpha)=Y_{n-k:n}+\frac{\widehat{\beta}_k}{\widehat{\xi}_k}\left(\left(\frac{n}{k}(1-\alpha)\right)^{-\widehat{\xi}_k}-1\right) \tag{16}$$

Une alternative est l'estimateur de Hill, si $\xi>0$,

$$\widehat{Q}(Y,\alpha)=Y_{n-k:n}\left(\frac{n}{k}(1-\alpha)\right)^{-\widehat{\xi}_k}, \tag{17}$$

$$\text{où } \widehat{\xi}_k = \frac{1}{k}\sum_{i=1}^{k} \log Y_{n+1-i:n} - \log Y_{n-k:n}.$$

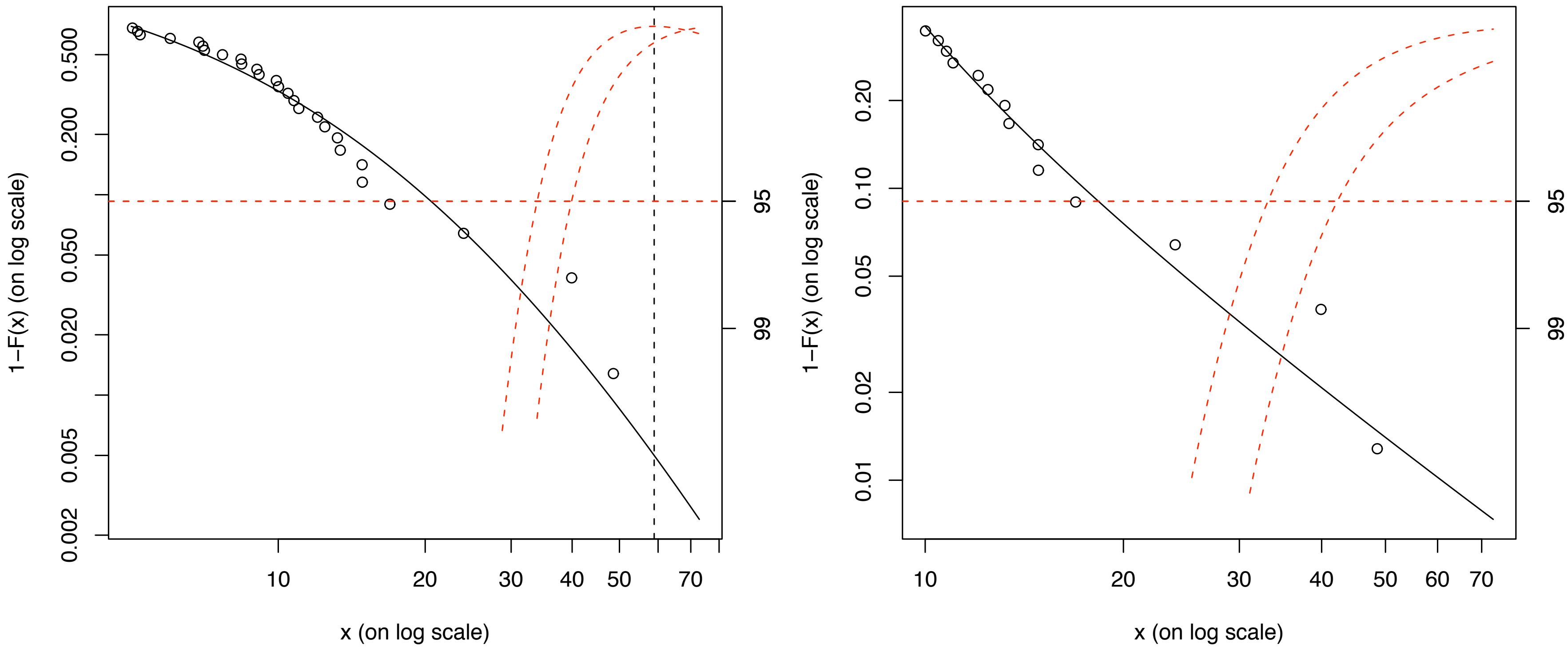

Figure 3 – Estimation de VaR et $TVaR$ à l'aide de la théorie des valeurs extrêmes.

5 Estimation des mesures de risques

Considérons une mesure de risque par distortion, c'est à dire de la forme $\mathcal{R}(X) = \int_0^1 F^{-1}(1-u)dG(u)$, où F est supposée strictement croissante, et continue.

L'estimateur naturel de cette mesure de risque est alors

$$\widehat{\mathcal{R}}(\{X_1, \cdots, X_n\}) = \int_0^1 \widehat{F}^{-1}(1-u)dG(u).$$

Etant données que $\widehat{F}^{-1}$ est une fonction en escalier, alors

$$\mathcal{R}(\{X_1, \cdots, X_n\}) = \sum_{i=1}^{n} \left(G\left(\frac{i}{n}\right) - G\left(\frac{i-1}{n}\right)\right) X_{n-i+1:n},$$

qui est simplement un L estimateur.

On peut alors, tout naturellement, espérer retrouver facilement des propriétés comme celles obtenues pour l'estimation nonparamétrique de la VaR. Par exemple

Proposition11

Si $X_1, \cdots, X_n$ est un échantillon i.i.d.,

$$\sqrt{n}\,(\mathcal{R}(X) - \mathcal{R}(\{X_1, \cdots, X_n\})) \to \int_0^1 \frac{B_{1-t}}{f(F^{-1}(1-t))} dG(t),$$

où $(B_t)_{t\in[0,1]}$ est un pont brownien. La loi limite est centrée, de variance asymptotique

$$\int_0^1 \int_0^1 \frac{\min\{t,u\} - ut}{f(F^{-1}(1-t))f(F^{-1}(1-u))} dG(t)dG(u).$$

Remarque16

Si G est dérivable, la variance asymptotique se réécrit

$$\int_0^1 \int_0^1 [\min\{F(t), F(u)\} - F(t)F(u)]g(1-F(t))g(1-F(u))dtdu.$$

[1] C. Acerbi. Spectral measures of risk : a coherent representation of subjective risk aversion. *Journal of Banking and Finance*, 2(7) :1505–1518, 2002.

[2] Kenneth J. Arrow and Gerard Debreu. Existence of an equilibrium for a competitive economy. *Econometrica*, 22(3) :pp. 265–290, 1954.

[3] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical finance*, 9(3) :203–228, 1999.

[4] A. Azzalini. A note on the estimation of a distribution function and quantiles by a kernel method. *Biometrika*, 68(1) :326–328, 1981.

[5] R. R. Bahadur. A note on quantiles in large samples. *The Annals of Mathematical Statistics*, 37(3) :pp. 577–580, 1966.

[6] Song Xi Chen and Cheng Yong Tang. Nonparametric Inference of Value-at-Risk for Dependent Financial Returns. *Journal of Financial Econometrics*, 3(2) :227–255, Spring 2005.

[7] Edmund A. Cornish and Ronald A. Fisher. Moments and cumulants in the specification of distributions. *Revue de l'Institut International de Statistique*, 5(4) :307–320, 1937.

[8] I. Csiszár. *I-divergence geometry of probability distributions and minimization problems*. *Ann. Probability*, 3 :146–158, 1975.

[9] Freddy Delbaen and Walter Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300 :463–520, 1994. 10.1007/BF01450498.

[10] Olivier Deprez and Hans U. Gerber. On convex principles of premium calculation. *Insurance : Mathematics and Economics*, 4(3) :179 – 189, 1985.

[11] I. Ekeland, A. Galichon, and M. Henry. *Comontonic Measures of Multivariate Risks*. Available at SSRN : http ://ssrn.com/abstract=1115729, 2009.

[12] Hans U. Gerber. *An introduction to mathematical risk theory*, volume 8 of *S.S. Heubner Foundation Monograph Series*. University of Pennsylvania Wharton School S.S. Huebner Foundation for Insurance Education, Philadelphia, Pa., 1979. With a foreword by James C. Hickman.

[13] H.U. Gerber and E.S.W. Shiu. Option pricing by esscher transforms (with discussions). *Transactions of the Society of Actuaries*, 46, 1994.

[14] C. Gouriéroux and J.M. Zakoïan. *Mesures de risques*. Notes de cours, ensae, 2009.

[15] G. H. Hardy and J. E. Littlewood. A maximal theorem with function-theoretic applications. *Acta Math.*, 54(1) :81–116, 1930.

[16] Frank E. Harrell and C. E. Davis. A new distribution-free quantile estimator. *Biometrika*, 69(3) :635–640, 1982.

[17] J. Michael Harrison and Stanley R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11(3) :215 – 260, 1981.

[18] P. Jorion. *Value-at-Risk*. McGraw-Hill, Oxford, 2007.

[19] J. Kiefer. On Bahadur's representation of sample quantiles. *Ann. Math. Statist.*, 38 :1323–1342, 1967.

[20] Shigeo Kusuoka. On law invariant coherent risk measures. In *Advances in mathematical economics, Vol. 3*, volume 3 of *Adv. Math. Econ.*, pages 83–95. Springer, Tokyo, 2001.

[21] E. L. Lehmann. Ordered families of distributions. *Ann. Math. Statist.*, 26 :399–419, 1955.

[22] Shaked M. and Shanthikumar J.G. *Stochastic Orders.* Springer, 2006.

[23] Michael Mania and Martin Schweizer. Dynamic exponential utility indifference valuation. *Ann. Appl. Probab.*, 15(3) :2113–2143, 2005.

[24] Harry Markowitz. Portfolio selection. *The Journal of Finance*, 7(1) :pp. 77–91, 1952.

[25] Harry M. Markowitz. *Mean-variance analysis in portfolio choice and capital markets.* Basil Blackwell, Oxford, 1987.

[26] Goovaerts M.J., De Vylder F., and Haezendonck J. *Insurance Premiums.* North-Holland, 1984.

[27] A. Müller and D. Stoyan. *Comparison methods for stochastic models and risks.* Wiley series in Probability and Statistics, 2002.

[28] Axel Reich. Premium principles and translation invariance. *Insurance : Mathematics and Economics*, 3(1) :57 – 66, 1984.

[29] Axel Reich. Properties of premium calculation principles. *Insurance : Mathematics and Economics*, 5(1) :97 – 101, 1986.

[30] L. Rüschendorf. Law invariant convex risk measures for portfolio vectors. *Statistics & Decisions*, 24 :97–108, 2006.

[31] Paul A. Samuelson. The "fallacy" of maximizing the geometric mean in long sequences of investing or gambling. *Proc. Nat. Acad. Sci. U.S.A.*, 68 :2493–2496, 1971.

[32] David Schmeidler. Integral representation without additivity. *Proc. Amer. Math. Soc.*, 97(2) :255–261, 1986.

[33] Klaus D. Schmidt. Positive homogeneity and multiplicativity of premium principles on positive risks. *Insurance : Mathematics and Economics*, 8(4) :315 – 319, 1989.

[34] Simon J. Sheather and J. S. Marron. Kernel quantile estimators. *J. Amer. Statist. Assoc.*, 85(410) :410–416, 1990.

[35] V. Strassen. The existence of probability measures with given marginals. *Ann. Math. Statist.*, 36 :423–439, 1965.

[36] Shaun Wang. Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance Math. Econom.*, 17(1) :43–54, 1995.

[37] Shaun Wang. Ordering of risks under PH-transforms. *Insurance Math. Econom.*, 18(2) :109–114, 1996.

[38] Shaun S. Wang and Virginia R. Young. Risk-adjusted credibility premiums using distorted probabilities. *Scand. Actuar. J.*, (2) :143–165, 1998.