An introduction to multivariate and dynamic risk measures

Arthur Charpentier

charpentier.arthur@uqam.ca

http://freakonometrics.hypotheses.org/

Université Catholique de Louvain-la-Neuve,

April 2014

Preambule

This document is freely inspired by

Ekeland, I., Galichon, A. & Henry, M. (2010). Comonotonic Measures of Multivariate Risks. *Mathematical Finance*.

Föllmer, H. & Schied, A. (2004) Stochastic Finance: An Introduction in Discrete Time. 2nd Edition, Walter de Gruyter.

Galichon, A. (2010) The VaR at risk. *International Journal of Theoretical & Applied Finance*, **13**, 503-506.

Galichon, A. & Henry, M. (2012). Dual theory of choice with multivariate risks. *Journal of Economic Theory* 147 (4), 1501-1516

Gilboa, I. (2009). Theory of Decision under Uncertainty. Cambridge University Press

Henry, M. (2010). Mesures de Risques, notes de cours.

1 Introduction and Notations

All (univariate) risk measures - or to be more specific of downside (or upside) risk - are, somehow, related to quantiles. So, in order to derive some general multivariate risk measures, or dynamic ones, we need to understand more deeply what quantile functions are, and why we need them (in this risk measure context).

1.1 Probablistic and Measurable Spaces

Consider some topological space S, metrizable, in the sense that there is a metric d on that space. Assume that S is separable, so that the σ-algebra $\mathcal{S}$ of S is generated by open d-balls, centered on a contable dense subset of S.

Let $\mathcal{M}(S)$ denote the set of all non-negative finite measures on S. Observe that every $\mu \in \mathcal{M}(S)$ can be writen $\mu = \alpha\mathbb{P}$ for some $\alpha \in [0, \infty)$. The set of all probability measures on S is $\mathcal{M}_1(S)$.

Finite-dimensional Probability Spaces

Consider a simple coin tossing model, or a single lottery. Then Ω is isomorphic to the set $\{0,1\}$, that we will call canonical. This setting will be related to lotteries in decision theory, with two possible outcomes.

Jacob Bernoulli and Pierre Simon Laplace stated an indifference principle: if there are n states of world, and if we have no reason to view one as more likely than another, then the canonical measure should be a uniform distribution, and each event will be assigned a $1/n$ probability. Thus, on the set $\Omega=\{0,1\}$, the canonical measure will be $\mathbb{P}=(1/2,1/2)\propto\mathbf{1}$. Actually, the measure is on Borelian sets of Ω, namely

$$\begin{cases}\mathbb{P}(\emptyset)=0\\ \mathbb{P}(\{0\})=1/2\\ \mathbb{P}(\{1\})=1/2\\ \mathbb{P}(\{0\}\cup\{1\})=\mathbb{P}(\Omega)=1\end{cases}$$

On $(\Omega, \mathbb{P})$, on can define measures $\mathbb{Q}$ or sets of measures $\mathcal{Q}$.

This was what we have have one lottery, but one can consider compound lotteries, where the canonical space can now be $\{0,1\}^n$, if we consider sequential simple lotteries.

Infinite-dimensional Probability Spaces

For a continuous state space Ω, the canonical space will be $[0,1]$. A first step before working on that continuous space can be to consider $\{0,1\}^{\mathbb{N}}$. This space is obtained using a binary representation of points on the unit interval, in the sense that

$$x = \sum_{i=1}^{\infty} \frac{x_i}{2^i} \in [0,1] \text{ with } x_i \in \{0,1\}, \text{ for all } i \in \mathbb{N}_\star.$$

The canonical measure is the uniform distribution on the unit interval $[0,1)$, denoted λ. $\lambda([0,1/2))$ corresponds to the probability that $X_1 = 0$, and thus, it should be $1/2$; $\lambda([0,1/4))$ corresponds to the probability that $X_1 = 0$ and $X_2 = 0$, and thus, it should be $1/4$; etc. Thus $\lambda([x, x+h)) = h$, which is the caracterization of the uniform distribution on the unit interval.

In the context of real-valued sequences,

$$L^p = \{\boldsymbol{u} = (u_n) | \ ||\boldsymbol{u}||_p < \infty\}, \text{ where } ||\boldsymbol{u}||_p = \left(\sum_{n\in\mathbb{N}} |u_n|^p\right)^{\frac{1}{p}}$$

where $p \in [1, \infty]$.

Proposition1

Let $(p, q) \in (1, +\infty)^2$ such that $1/p + 1/q = 1$, then L^q is the dual of L^p.

If $b \in L^q$ and $a \in L^p$, the mapping

$$T : L^q \to L^{p\star} : b \mapsto \ell_b \text{ where } \ell_b(a) = \sum_{i\in\mathbb{N}} a_i b_i$$

is an isometric isomorphism. So $L^{p\star} = L^q$.

Consider a linear mapping ℓ from L^p to $\mathbb{R}$, linear in the sense that

$$\ell(af + bg) = a\ell(f) + b\ell(g) \text{ for all } a, b \in \mathbb{R} \text{ and } f, g \in L^p.$$

Assume that this functional is bounded, in the sense that there is M such that $|\ell(f)| \leq M||f||_p$. One can define a norm $||\cdot||$ on the space of such linear mapping. Define

$$||\ell|| = \sup_{||f||=1} \{|\ell(f)|\} \sup_{||f||\leq 1} \{|\ell(f)|\}$$

The space of all of linear mappings (with that norm) is the dual of L^p.

One can prove that the dual of L^p is L^q, in the sense that for all linear mapping ℓ, there is $g \in L^q$ such that

$$\ell(f) = \int f(\omega)g(\omega)d\mathbb{P}(\omega) \text{ for all } f \in L^p.$$

This should not be suprising to see that L^q is the dual of L^p since for $g \in L^q$

$$||g||_q = \sup_{||f||=1} \{|\int fg|\} \sup_{||f||\leq 1} \{|\int fg|\}$$

The optimum is obtain

$$f(x) = |g(x)|^{q-1}\text{sign}(g(x))\frac{1}{||g||_q^{q-1}},$$

which satisfies

$$||f||_p^p = \int |g(x)|^{p(q-1)}\text{sign}(g(x))\frac{d\mu}{||g||_q^{p(q-1)}} = 1.$$

Remark1

L^∞ is the dual of L^1, but the converse is generally not true.

The space L^∞ is the class of functions that are essentially bounded.

$X \in L^\infty$ if there exits $M \geq 0$ such that $|X| \leq M$ a.s. Then define

$$||X||_{L^\infty} = \inf\{M \in \mathbb{R}_+ | \mathbb{P}(|X| \leq M) = 1\}.$$

Given X, define

$$\text{essup}X = \inf\{M \in \mathbb{R} | \mathbb{P}(X \leq M) = 1\}$$

and

$$\text{essinf}X = \inf\{m \in \mathbb{R} | \mathbb{P}(X \geq m) = 1\}$$

Observe that $X \in L^\infty$ if and only if essup $< \infty$, esinf$X < \infty$, and $||X||_{L^\infty} = \text{essup}|X|$

It is also possible to define the essential supremum on a set of random variables (on $(\Omega, \mathcal{F}, \mathbb{P})$). Let Φ denote such a set. Then there exists $\varphi^\star$ such that

$$\varphi^\star \geq \varphi, \; \mathbb{P} - \text{a.s. for all } \varphi \in \Phi.$$

Such as function is a.s. unique, and $\varphi^\star$ is denoted esssupΦ.

Remark2

Given a random variable X, and

$$\Phi = \{c \in \mathbb{R} | \mathbb{P}(X > c) > 0\}$$

then $\text{esssup}\Phi = \text{esssup}X$, which is the smallest constant such that $X \leq c^\star$, $\mathbb{P}$−a.s.

1.2 Univariate Functional Analysis and Convexity

$f:\mathbb{R}\rightarrow\mathbb{R}\cup\{+\infty\}$ is a convex function if for all $x,y\in\mathbb{R}$, with $x\in\text{dom}f$, and $\alpha\in[0,1]$,

$$f(\alpha x+(1-\alpha)y)\leq af(x)+(1-\alpha)f(y).$$

where $\text{dom}f=\{x\in\mathbb{R}|f(x)<+\infty\}$.

Recall that if f is convex, then it is (upper) semi-continuous (and locally Lipschitz) on the interior of $\text{dom}f$. Further, f admits left- and right-hand derivatives, and one can write, for all $x\in\text{dom}f$,

$$f(x+h)=f(x)+\int_x^{x+h}f'_+(y)dy \text{ and } f(x-h)=f(x)+\int_x^{x-h}f'_-(y)dy$$

An other possible definition is the following: f is a convex function is there exists $a:\mathbb{R}\rightarrow\overline{\mathbb{R}}$ such that, for all $x\in\mathbb{R}$,

$$f(x)=\sup_{y\in\mathbb{R}}\{x\cdot y-a(y)\}=a^\star(x)$$

The interpretation is that f should be above the tangent at each point. Thus, they should be above the supremum of all tangeants. This function $a^\star$ will be related to the Legendre-Fenchel transformation of a.

Legendre-Fenchel transformation

The conjugate of function $f:\mathbb{R}^d\rightarrow\overline{\mathbb{R}}$ is function $f^\star$ defined as

$$f^\star(\boldsymbol{s})=\sup_{\boldsymbol{x}\in\mathbb{R}^d}\{\boldsymbol{s}\boldsymbol{x}-f(\boldsymbol{x})\}$$

Note that it is possible to extend this notion to more general spaces E, then $\boldsymbol{s}\in E^\star$ (dual of space E) and $\boldsymbol{s}\boldsymbol{x}$ becomes $<\boldsymbol{s},\boldsymbol{x}>$.

Observe that $f^\star$ is a convex function lower semi-continuous.

Example1

Let $\mathcal{E}$ denote sur nonempty subset of $\mathbb{R}^d$, and define the indicator function of $\mathcal{E}$,

$$\mathbf{1}_{\mathcal{E}}(\boldsymbol{x}) = \begin{cases} 0 \text{ if } \boldsymbol{x} \notin E \\ +\infty \text{ if } \boldsymbol{x} \in E \end{cases}$$

Then

$$\mathbf{1}^{\star}_{\mathcal{E}}(\boldsymbol{s}) = \sup_{\boldsymbol{x} \in \mathcal{E}} \{\boldsymbol{s}\boldsymbol{x}\}$$

which is the support function of $\mathcal{E}$.

Example2

Let $f(x) = \alpha \exp[x]$, with $\alpha \in (0,1)$, then

$$f^{\star}(s) = \begin{cases} +\infty \text{ if } s < 0 \\ 0 \text{ if } s = 0 \\ s[\log s - \log \alpha] - s \text{ if } s > 0 \end{cases}$$

Those functions can be visualized Figure 1.

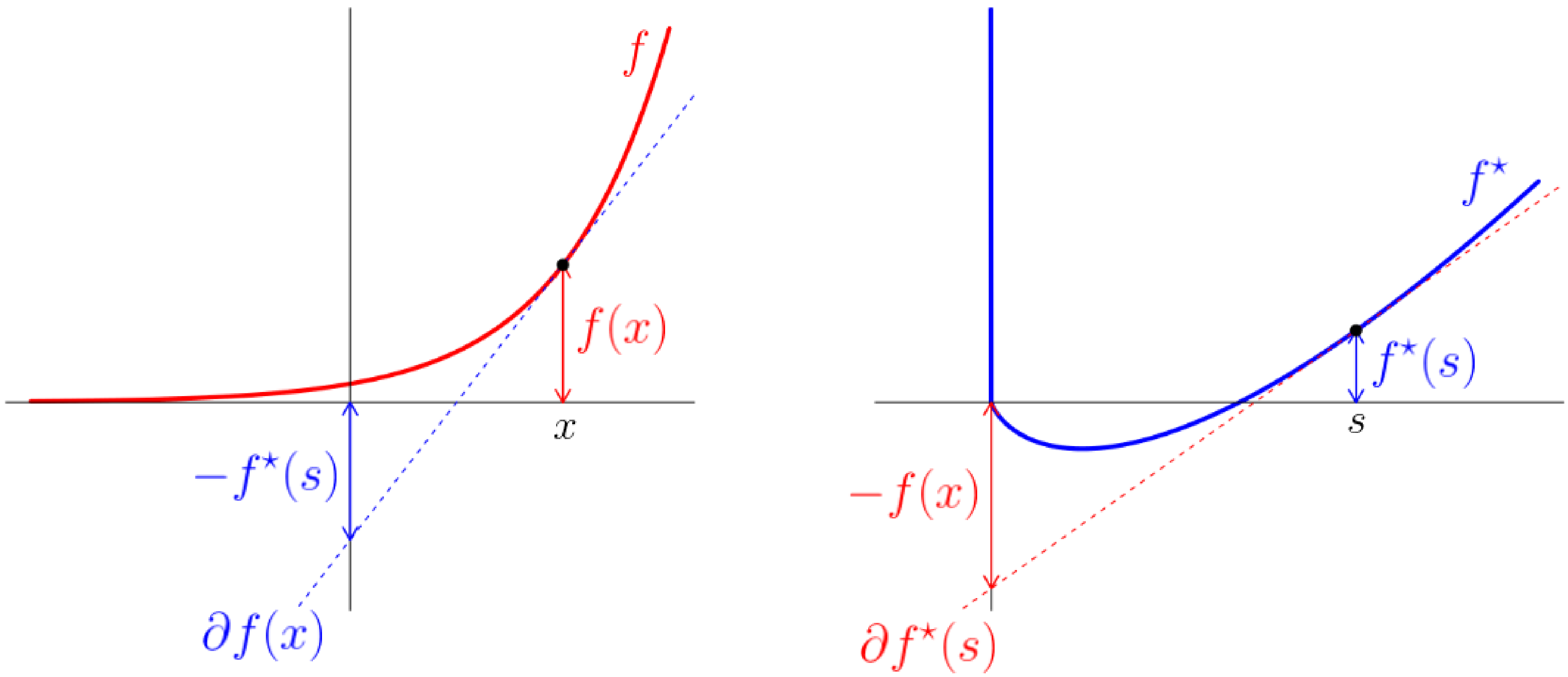


Figure 1: A convex function f and the Fenchel conjugate $f^\star$

If f is 1-coercive, in the sense that $\dfrac{f(\boldsymbol{x})}{||\boldsymbol{x}||} \to \infty$ as $||\boldsymbol{x}|| \to \infty$, then $f^\star$ is finite on $\mathbb{R}^d$.

Proposition2

If $f : \mathbb{R}^d \to \mathbb{R}$ is strictly convex, differentiable, and 1-coercive, then

- $f^\star$ is also finite, strictly convex, differentiable and 1-coercive
- $\nabla f : \mathbb{R}^d \to \mathbb{R}^d$ is also differentiable and

$$f^\star \boldsymbol{s} = \boldsymbol{s}[(\nabla f)^{-1}(\boldsymbol{s})] - f((\nabla f)^{-1}(\boldsymbol{s})).$$

Proposition3

If $f : \mathbb{R}^d \to \mathbb{R}$ is convex, lower semi-continuous then so is $f^\star$, and $f^{\star\star} = f$.

More generally, we have that $f^{\star\star}$ is the largest convex function satisfying $f^{\star\star}(x) \leq f(x)$, which is actually the convex hull of function f.

Definition1

An element $\boldsymbol{s}$ of $\mathbb{R}^d$ such that for any $\boldsymbol{y}$

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \boldsymbol{s}[\boldsymbol{y} - \boldsymbol{x}]$$

is called sub-gradient of f at point $\boldsymbol{x}$. The set of sub-gradients is denoted $\partial f(\boldsymbol{x})$.

Proposition4

As a consequence,

$$\boldsymbol{s} \in \partial f(\boldsymbol{x}) \iff f^\star(\boldsymbol{s}) + f(\boldsymbol{x}) = \boldsymbol{s}\boldsymbol{x}.$$

Proposition5

If $f : \mathbb{R}^d \to \mathbb{R}$ is convex, lower semi-continuous then

$$\boldsymbol{s} \in \partial f(\boldsymbol{x}) \iff \boldsymbol{x} \in \partial f^\star(\boldsymbol{s})$$

that might be denoted - symbolically - $\partial f^\star = [\partial f]^{-1}$.

Corollary1

If $f : \mathbb{R}^d \to \mathbb{R}$ is convex, twice differentiable, and 1-coercice, then $\nabla f^\star(\boldsymbol{s}) = [\nabla f]^{-1}(\boldsymbol{s})$.

Example3

If f is a power function, $f(x) = \dfrac{1}{p}|x|^p$ where $1 < p < \infty$ then

$$f^\star(x^\star) = \frac{1}{q}|x^\star|^q$$

where $\dfrac{1}{p} + \dfrac{1}{q} = 1$.

Example4

If f is the exponential function, $f(x) = \exp(x)$ then

$$f^\star(x^\star) = x^\star \log(x^\star) - x^\star \text{ if } x^\star > 0.$$

Example5

Let X be a random variable with c.d.f. F_X and quantile function Q_X. The Fenchel-Legendre tranform of

$$\Psi(x) = \mathbb{E}[(x-X)_+] = \int_{-\infty} x F_X(z)dz$$

is

$$\Psi^\star(y) = \sup_{x\in\mathbb{R}}\{xy - \Psi(x)\} = \int_0^y Q_X(t)dt$$

on $[0,1]$.

Indeed, from Fubini,

$$\Psi(x) = \int_{-\infty} x\mathbb{P}(X\leq z)dz = \int_{-\infty} x\mathbb{E}(\mathbf{1}_{X\leq z})dz = \mathbb{E}\left(\int_{-\infty} x\mathbf{1}_{X\leq z}dz\right)$$

i.e.

$$\Psi(x) = \mathbb{E}\left([x-X]_+\right) = \int_0^1 [x - Q_X(t)]_+ dt$$

Observe that

$$\Psi^\star(1)=\sup_{x\in\mathbb{R}}\{x-\Psi(x)\}=\lim_{x\uparrow\infty}\int_0^1[x-(x-Q_X(t))_+]dt=\int_0^1 Q_X(t)dt$$

and $\Psi^\star(0)=0$. Now, the proof of the result when $y\in(0,1)$ can be obtained since

$$\frac{\partial xy-\Psi(x)}{\partial x}=y-F_X(x)$$

The optimum is then obtained when $y=F_X(x)$, or $x=Q_Y(y)$.

One can also prove that

$$\left(\inf_\alpha f_\alpha\right)^*(x)=\sup_\alpha f_\alpha^*(x)\text{ and }\left(\sup_\alpha f_\alpha\right)^*(x)\leq\inf_\alpha f_\alpha^*(x).$$

Further, $f=f^{\star\star}$ if and only if f is convex and lower semi-continuous.

And from Fenchel-Young inequality, for any f,

$$<x^\star,x>\leq f(x)+f^\star(x^\star).$$

and the equality holds if and only if $x^\star\in\partial f(x)$.

Example6

The standard expression of Young's inequality is that if $h : \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous strictly increasing function on $[0, m]$ with $h(0) = 0$, then for all $a \in [0, m]$ and $b \in [0, h(m)]$, then

$$ab \leq \int_0^a h(x)dx + \int_0^b h^{-1}(y)dy$$

with the equality if and only if $b = h(a)$ (see Figure 2). A well know corollary is that

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} \text{ when } p \text{ and } q \text{ are conjugates.}$$

The extension is quite natural. Let $f(a) = \int_0^a h(x)dx$, then f is a convex function, and its convex conjugate is $f^\star(b) = \int_0^b h^{-1}(y)dy$, then

$$ab \leq f(a) + f^\star(b).$$

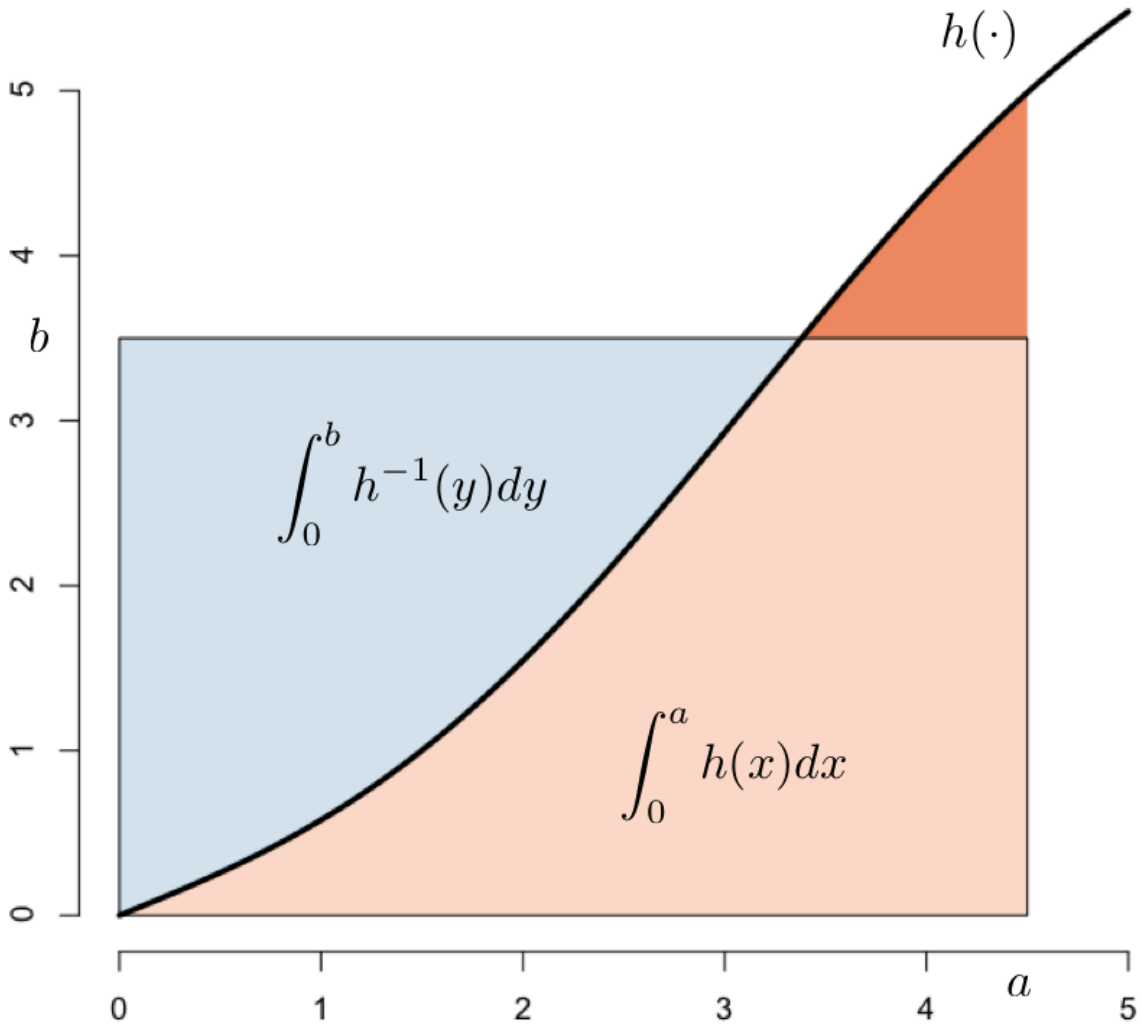


Figure 2: Fenchel-Young inequality

1.3 Changes of Measures

Consider two probability measures $\mathbb{P}$ and $\mathbb{Q}$ on the same measurable space $(\Omega, \mathcal{F})$. $\mathbb{Q}$ is said to be absolutely continuous with repect to $\mathbb{P}$, denoted $\mathbb{Q} \ll \mathbb{P}$ if for all $A \in \mathcal{F}$,

$$\mathbb{P}(A) = 0 \Longrightarrow \mathbb{Q}(A) = 0$$

If $\mathbb{Q} \ll \mathbb{P}$ and $\mathbb{Q} \gg \mathbb{P}$, then $\mathbb{Q} \approx \mathbb{P}$.

$\mathbb{Q} \ll \mathbb{P}$ if and only if there exists a (positive) measurable function φ such that

$$\int h d\mathbb{Q} = \int \int h \varphi d\mathbb{P}$$

for all positive measurable functions h. That function *varphi* is call Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$, and we write

$$\varphi = \frac{d\mathbb{Q}}{d\mathbb{P}}$$

Observe that, generally, $\mathbb{Q} \approx \mathbb{P}$ if and only if φ is stricly positive, and in that case,

$$\frac{d\mathbb{P}}{d\mathbb{Q}} = \left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)^{-1}$$

Let $\mathbb{E}_{\mathbb{P}}(\cdot|\mathcal{F}_0)$ denote the conditional expectation with respect to a probability measure $\mathbb{P}$ and a σ-algebra $\mathcal{F}_0 \subset F$.

If $\mathbb{Q} \ll \mathbb{P}$,

$$\mathbb{E}_{\mathbb{Q}}(\cdot|\mathcal{F}_0) = \frac{1}{\mathbb{E}_{\mathbb{P}}(\varphi|\mathcal{F}_0)}\mathbb{E}_{\mathbb{P}}(\cdot\varphi|\mathcal{F}_0), \text{ where } \varphi = \frac{d\mathbb{Q}}{d\mathbb{P}}.$$

If there is no absolute continuity property between two measures $\mathbb{P}$ and $\mathbb{Q}$ (neither $\mathbb{Q} \ll \mathbb{P}$ nor $\mathbb{P} \ll \mathbb{Q}$), one can still find a function φ, and a $\mathbb{P}$-null set N (in the sense $\mathbb{P}(N) = 0$) such that

$$\mathbb{Q}(A) = \mathbb{Q}(A \cap N) + \int_A \varphi d\mathbb{P}$$

Thus,

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \varphi \text{ on } N^{\mathsf{C}}.$$

1.4 Multivariate Functional Analysis

Given a vector $\boldsymbol{x} \in \mathbb{R}^d$ and $I = \{i_1, \cdots, i_k\} \subset \{1, 2, \cdots, d\}$, then denote

$$\boldsymbol{x}_I = (x_{i_1}, x_{i_2}, \cdots, x_{i_k}).$$

Consider two random vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$. We denote $\boldsymbol{x} \leq \boldsymbol{y}$ if $x_i \leq y_i$ for all $i = 1, 2, \ldots, d$. Then function $h : \mathbb{R}^d \to \mathbb{R}$, is said to be increasing if

$$h(\boldsymbol{x}) \leq h(\boldsymbol{y}) \text{ whenever } \boldsymbol{x} \preceq \boldsymbol{y}.$$

If $f : \mathbb{R}^d \to \mathbb{R}$ is such that $\nabla f : \mathbb{R}^d \to \mathbb{R}^d$ is bijective, then

$$f^\star(\boldsymbol{y}) =< \boldsymbol{y}, (\nabla f)^{-1}(y) > -f((\nabla f)^{-1}(y)) \text{ for all } \boldsymbol{y} \in \mathbb{R}^d.$$

We will say that $\boldsymbol{y} \in \partial f(\boldsymbol{x})$ if and only if

$$< \boldsymbol{y}, \boldsymbol{x} >= f(\boldsymbol{x}) + f^\star(\boldsymbol{y})$$

1.5 Valuation and Neyman-Pearson

Valuation of contingent claims can be formalized as follows. Let X denote the claim, which is a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$, and its price is given be $\mathbb{E}(\varphi X)$, where we assume that the price density φ is a strictly positive random variable, absolutely continuous, with $\mathbb{E}(\varphi) = 1$. The risk of liability $-X$ is measures by $\mathcal{R}$, and we would like to solve

$$\min\{\mathcal{R}(-X) | 0 \in [0, k] \text{ and } \mathbb{E}(\varphi X) \geq a\}$$

In the case where $\mathcal{R}(-X) = \mathbb{E}(X)$, we have a problem that can be related to Neyman-Pearson lemma (see [24], section 8.3 and [33])

2 Decision Theory and Risk Measures

In this section, we will follow [14], trying to get a better understanding of connections between decision theory, and orderings of risks and risk measures. From Cantor, we know that any ordering can be represented by a functional. More specifically,

Proposition6

Let $\preceq$ denote a preference order that is

- complete for every X and y, either $x \preceq y$ or $y \preceq x$
- transitive for every x, y, z such that $x \preceq y$ and $y \preceq z$, then $x \preceq z$
- separable for every x, y such that $x \prec y$, then there is z such that $x \preceq z \preceq y$.

Then $\preceq$ can be represented by a real valued function, in the sense that

$$x \preceq y \iff u(x) \leq u(y).$$

Keep in mind that u is unique up to an increasing transformation. And since there is no topology mentioned here, it is meaningless to claim that u should be continuous. This will require additional assumption, see [6].

Proof. In the case of a finite set $\mathcal{X}$, define

$$u(x) = \text{card}\{y \in \mathcal{X} | y \preceq x\}.$$

In the case of an infinite set, but countable,

$$u(x) = \sum_{\{y_i \in \mathcal{X} | y_i \preceq x\}} \frac{1}{2^i} - \sum_{\{y_i \in \mathcal{X} | x \preceq y_i\}} \frac{1}{2^i}$$

$\square$

2.1 von Neuman & Morgenstern: comparing lotteries

In the previous setting, space $\mathcal{X}$ was some set of alternatives. Assume now that we have lotteries on those alternative. Formally, a lottery is function $P:\mathcal{X}\rightarrow[0,1]$. Consider the case where $\mathcal{X}$ is finite or more precisely, the cardinal of x's such that $P(x)>0$ is finite. Let $\mathcal{L}$ denote the set of all those lotteries on $\mathcal{X}$. Note that mixtures can be considered on that space, in the sense that for all $\alpha\in[0,1]$, and for all $P,\{Q\in\mathcal{L}$, $\alpha P\oplus(1-\alpha)Q\in\mathcal{L}$, where for any $x\in\mathcal{X}$,

$$[\alpha P\oplus(1-\alpha)Q](x)=\alpha P(x)+(1-\alpha)Q(x)$$

It is a standard mixture, in the sense that we have lottery P with probability α and Q with probability $1-\alpha$.

Proposition7

Let $\preceq$ denote a preference order on $\mathcal{L}$ that is

eak order (complete and transitive)

ontinuous for every P, Q, R such that $P \prec Q \prec R$, then there are α, β such that

$$\alpha P \oplus (1-\alpha)R \preceq Q \preceq \beta P \oplus (1-\beta)R.$$

dependent for every P, Q, R and every $\alpha \in (0,1)$

$$P \preceq Q \iff \alpha P \oplus (1-\alpha)R \preceq \alpha Q \oplus (1-\alpha)R,$$

Then $\preceq$ can be represented by a real valued function, in the sense that

$$P \preceq Q \iff \sum_{x\in\mathcal{X}} P(x)u(x) \leq \sum_{x\in\mathcal{X}} Q(x)u(x).$$

Proof. See [18]. □

2.2 de Finetti: comparing outomes

[7] considered the case of bets on canonical space $\{1, 2, \cdots, n\}$. The set of bet outcomes is $\mathcal{X} = \{\boldsymbol{x} = (x_1, \cdots, x_n)\} \in \mathbb{R}^n$.

Proposition8

Let $\preceq$ denote a preference order on $\mathcal{X}$ that is

vial order (complete, transitive and there are $\boldsymbol{x}, \boldsymbol{y}$ such that $\boldsymbol{x} \prec \boldsymbol{y}$,

ontinuous for every $\boldsymbol{x}$, sets $\{\boldsymbol{y}|\boldsymbol{x} \prec \boldsymbol{y}\}$ and $\{\boldsymbol{y}|\boldsymbol{y} \prec \boldsymbol{x}\}$ are open

additive for every $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$,

$$\boldsymbol{x} \preceq \boldsymbol{y} \iff \boldsymbol{x} + \boldsymbol{z} \preceq \boldsymbol{y} + \boldsymbol{z}$$

nonotonic consider $\boldsymbol{x}, \boldsymbol{y}$ such that $x_i \leq y_i$ for all i, then $\boldsymbol{x} \preceq \boldsymbol{y}$

Then $\preceq$ can be represented by a probability vector, in the sense that

$$\boldsymbol{x} \preceq \boldsymbol{y} \iff \boldsymbol{p}\boldsymbol{x} \leq \boldsymbol{p}\boldsymbol{y} \iff \sum_{i=1}^n p_i x_i \leq \sum_{i=1}^n p_i y_i$$

Proof. Since $\boldsymbol{x} \preceq \boldsymbol{y}$ means that $\boldsymbol{x} - \boldsymbol{y} \preceq \boldsymbol{0}$, the argument here is nothing more than a separating hyperplane argument, between two spaces,

$$A = \{\boldsymbol{x} \in \mathcal{X} | \boldsymbol{x} \prec \boldsymbol{0}\} \text{ and } B = \{\boldsymbol{x} \in \mathcal{X} | \boldsymbol{0} \prec \boldsymbol{x}\}$$

□

2.3 Savage Subjective Utility

With von Neuman & Morgenstern, we did focus on probabilities of states of the world. With de Finetti, we did focus on outcomes in each states of the world. Savage decided to focus on acts, which are functions from states to outcomes

$$\mathcal{A} = \mathcal{X}^{\Omega} = \{X : \Omega \to \mathcal{X}\}$$

In Savage model, we do not need a probability measure on $(\Omega, \mathcal{F})$, what we need is a finite additive measure. Function μ, defined on $\mathcal{F}$ - taking values in $\mathbb{R}_+$ - is said to be finitely additive if

$$\mu(A \cup B) = \mu(A) + \mu(B) \text{ whenever } A \cap B = \emptyset.$$

Somehow, σ-additivity of probability measure can be seen as an additional constraint, related to continuity, since in that case, if A_i's are disjoint sets and if

$$B_n = \bigcup_{i=1}^{n} A_i$$

then with σ-additivity,

$$\mu\left(\lim_{n\uparrow\infty} B_n\right) = \mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu\left(A_i\right) = \lim_{n\uparrow\infty} \sum_{i=1}^{n} \mu\left(A_i\right) = \lim_{n\uparrow\infty} \mu(B_n)$$

Actually, a technical assumption is usually added: measure μ should be non-atomic. An atom is a set that cannot be split (with respect to μ). More precisely, if A is an atom, then $\mu(A) > 0$, and if $B \subset A$, then either $\mu(B) = 0$, or $\mu(B) = \mu(A)$.

Now, given $X, Y \in \mathcal{A}$, and $S \subset \Omega$, define

$$S_X^Y(\omega) = \begin{cases} Y(\omega) \text{ if } \omega \in S \\ X(\omega) \text{ if } \omega \notin S \end{cases}$$

Proposition9

Let $\preceq$ denote a preference order on $\mathcal{A} == \mathcal{X}^\Omega$ that is

vial order (complete, transitive and there are X, Y such that $X \prec Y$,

P2 For every $X, Y, Z, Z' \in \mathcal{A}$ and $S \subset \Omega$,

$$S_X^Z \preceq S_Y^Z \iff S_X^{Z'} \preceq S_Y^{Z'}.$$

P3 For every $Z \in \mathcal{A}$, $x, y \in \mathcal{X}$ and $S \subset \Omega$,

$$S_Z^{\{x\}} \preceq S_Z^{\{y\}} \iff x \preceq y.$$

P4 For every $S, T \subset \Omega$, and every $x, y, z, w \in \Omega$ with $x \prec y$ and $z \prec w$,

$$S_y^x \preceq T_y^x \iff S_z^w \preceq T_z^w$$

P6 For every $X, Y, Z \in \mathcal{A}$, with $X \preceq Y$, there exists a partition of Ω, $\{S_1, S_2, \cdots, S_n\}$ such that, for all $i \in \{1, 2, \cdots, n\}$,

$$(S_i)_X^Z \preceq Y \text{ and } X \preceq (S_i)_Y^Z$$

P7 For every $X, Y \in \mathcal{A}$ and $S \subset \Omega$, if for every $\omega \in S$, $X \preceq_S Y(\omega)$, then $X \preceq_S Y$, and if for every $\omega \in S$, $Y(\omega) \preceq_S X$, then $Y \preceq_S X$.

Then $\preceq$ can be represented by a non-atomic finitely additive measure μ on ω and a non-constant function $\mathcal{X} \to \mathbb{R}$, in the sense that

$$X \preceq Y \iff \sum_{\omega \in \Omega} u(X(\omega))\mu(\{\omega\}) \leq \sum_{\omega \in \Omega} u(Y(\omega))\mu(\{\omega\})$$

Notations P2,..., P7 are based on [14]'s notation.

With a more contemporary style,

$$X \preceq Y \iff \mathbb{E}_\mu[u(X)] \leq \mathbb{E}_\mu[u(Y)].$$

2.4 Schmeidler and Choquet

Instead of considering finitely additional measures, one might consider a weaker notion, called non-additive probability (or capacity, in [5]), which is a function ν on $\mathcal{F}$ such that

$$\begin{cases} \nu(\emptyset) = 0 \\ \nu(A) \leq \nu(B) \text{ whenever } A \subset B \\ \nu(\Omega) = 1 \end{cases}$$

It is possible to define the integral with respect to ν. In the case where $\mathcal{X}$ is finite with a positive support, i.e. X takes (positive) value x_i in state ω_i, let σ denote the permutation so that $x_{\sigma(i)}$'s are decreasingly. Let $\tilde{x}_i = x_{\sigma(i)}$ and $\tilde{\omega}_i = \omega_{\sigma(i)}$

$$\mathbb{E}_\nu(X) = \int X d\nu = \sum_{i=1}^{n} [\tilde{x}_i - \tilde{x}_{i+1}] \nu \left(\bigcup_{j \leq i} \{\tilde{\omega}_j\} \right)$$

In the case where $\mathcal{X}$ is continuous, and positive,

$$\mathbb{E}_\nu(X) = \int X d\nu = \int_{\mathcal{X}} \nu(X \geq t) dt$$

(where the integral is the standard Riemann integral).

This integral is nonadditive in the sense that (in general)

$$\mathbb{E}_\nu(X + Y) \neq \mathbb{E}_\nu(X) + \mathbb{E}_\nu(Y).$$

Now, Observe that we can also write (in the finite case)

$$\mathbb{E}_\nu(X) = \int X d = \sum_{i=1}^{n} \tilde{x}_i \left[\nu \left(\bigcup_{j \leq i} \{\tilde{\omega}_j\} \right) - \nu \left(\bigcup_{j < i} \{\tilde{\omega}_j\} \right) \right]$$

There is a probability $\mathbb{P}$ such that

$$\mathbb{P}\left(\bigcup_{j\leq i}\{\tilde{\omega}_j\}\right)=\nu\left(\bigcup_{j\leq i}\{\tilde{\omega}_j\}\right)$$

and thus,

$$\mathbb{E}_{\mathbb{P}}(X)=\int Xd\mathbb{P}$$

Probability $\mathbb{P}$ is related to permutation σ, and if we assume that both variables X and Y are related to the same permutation σ, then

$$\mathbb{E}_{\nu}(X)=\int Xd\mathbb{P}\text{ and }\mathbb{E}_{\nu}(Y)=\int Yd\mathbb{P}$$

so in that very specific case,

$$\mathbb{E}_{\nu}(X+Y)=\int(X+Y)d\mathbb{P}=\int Xd\mathbb{P}+\int Yd\mathbb{P}=\mathbb{E}_{\nu}(X)+\mathbb{E}_{\nu}(Y).$$

The idea that variables X and Y are related to the same permutation means that variables X and Y are comonotonic, since

$$[X(\omega_i) - X(\omega_j)] \cdot [Y(\omega_i) - Y(\omega_j)] \geq 0 \text{ for all } i \neq j.$$

Proposition10

Let $\preceq$ denote a preference order on $\mathcal{X}^\Omega$ that is

vial order (complete, transitive and there are X, Y such that $X \prec Y$,

pendence for every X, Y, Z comonotonic, and every $\alpha \in (0,1)$,

$$X \preceq Y \iff \alpha X \oplus (1-\alpha) Z \preceq \alpha Y \oplus (1-\alpha) Z$$

Then $\preceq$ can be represented by a nonatomic non-additive measure ν on Ω and a non-constant function $u : \mathcal{X} \to \mathbb{R}$, in the sense that

$$X \preceq Y \iff \sum_\omega [E_{X(\omega)} u] d\nu \leq \sum_\omega [E_{Y(\omega)} u] d\nu$$

where $E_{X(\omega)}u = \sum_{x\in\mathcal{X}} X(\omega)(x)u(x)$.

Here ν is unique, and u is unique up to a (positive) linear transformation.

Actually, an alternative expression is the following

$$\int_0^1 u(F_X^{-1}(t))d(t) \leq \int_0^1 u(F_Y^{-1}(t))d(t)$$

2.5 Gilboa and Schmeidler: Maxmin Expected Utility

Consider some non-additive (probability) measure on Ω. And define

$$\text{core}(\nu) = \{\mathbb{P} \text{ probability measure on } \Omega | \mathbb{P}(A) \geq \nu(A) \text{ for all } A \subset \Omega\}$$

The non-additive measure ν is said to me convex if (see [31] and [34]) $\text{core}(\nu) \neq \emptyset$ and for every $h : \Omega \rightarrow \mathbb{R}$,

$$\int_\Omega h d\nu = \min_{\mathbb{P}\in\text{core}(\nu)} \left\{ \int_\Omega h d\mathbb{P} \right\}$$

Conversely, we can consider some (convex) set of probabilities $\mathcal{C}$, and see if using some axiomatic on the ordering, we might obtain a measure that will be the minimum of some integral, with respect to probability measures. [15] obtained the following result

Proposition11

Let $\preceq$ denote a preference order on $\mathcal{X}^\Omega$ that is

- vial order (complete, transitive and there are X, Y such that $X \prec Y$,
- y aversion for every X, Y, if $X \sim Y$, then for every $\alpha \in (0,1)$, $X \preceq \alpha X \oplus (1-\alpha)Y$
- pendence for every X, Y, every constant c, and for every $\alpha \in (0,1)$,

$$X \preceq Y \iff \alpha X \oplus (1-\alpha)c \preceq \alpha Y \oplus (1-\alpha)c$$

Then $\preceq$ can be represented a closed and convex of probability measure $\mathcal{C}$ on Ω and a non-constant function $\mathcal{X} \to \mathbb{R}$, in the sense that

$$X \preceq Y \iff \min_{\mathbb{P}\in\mathcal{C}} \left\{ \int_\Omega [E_{X(\omega)} u] d\mathbb{P} \right\} \leq \min_{\mathbb{P}\in\mathcal{C}} \left\{ \int_\Omega [E_{Y(\omega)} u] d\mathbb{P} \right\}$$

2.6 Choquet for Real Valued Random Variables

In the section where we introduced Choquet's integral, we did assume that X was a positive random variable. In the case where $\mathcal{X} = \mathbb{R}$, two definitions might be considered,

The symmetric integral, in the sense introduced by Šipoš of X with respect to ν is

$$\mathbb{E}_{\nu,\mathsf{s}}(X) = \mathbb{E}_\nu(X_+)E - \nu(\mathbb{X}_-)$$

where $X_- = \max\{-X, 0\}$ and $X_+ = \max\{0, X\}$. This coincides with Lebesgue integral in the case where ν is a probability measure.

Another extention is the one introduced by Choquet,

$$\mathbb{E}_\nu(X) = \mathbb{E}_\nu(X_+) - \mathbb{E}_{\overline{\nu}}(X_-)$$

where $\overline{\nu}(A) = 1 - \nu(A^{\mathsf{C}})$. Here again, this integral coincides with Lebesgue integral in the case where ν is a probability measure. One can write, for the later expression

$$\mathbb{E}_{\nu}(X)=\int_{-\infty}^{0}[\nu(X>x)-1]dx+\int_{0}^{\infty}\nu(X>x)dx$$

2.7 Distortion and Maximum

Definition2

Let $\mathbb{P}$ denote a probability measure on $(\Omega,\mathcal{F})$. Let $\psi:[0,1]\rightarrow[0,1]$ increasing, such that $\psi(0)=0$ and $\psi(1)=1$. Then $(\cdot)=\psi\circ\mathbb{P}(\cdot)$ is a capacity. If ψ is concave, then $\nu=\psi\circ\mathbb{P}$ is a subadditive capacity.

Definition3

Let $\mathcal{P}$ denote a family of probability measures on $(\Omega,\mathcal{F})$.Then $\nu(\cdot)=\sup_{\mathbb{P}\in\mathcal{P}}\{\mathbb{P}(\cdot)\}$ is a capacity. Further, ν is a subadditive capacity and $\mathbb{E}_{\nu}(X)\geq\sup_{\mathbb{P}\in\mathcal{P}}\{\mathbb{E}_{\mathbb{P}}(X)\}$ for all random variable X.

3 Quantile(s)

Definition4

The quantile function of a real-valued random variable X is a $[0,1] \to \mathbb{R}$ function, defined as

$$Q_X(u) = \inf\{x \in \mathbb{R} | F_X(x) > u\}$$

where $F_X(x) = \mathbb{P}(X \leq x)$.

This is also called the upper quantile function, which is right-continuous.

Consider n states of the world, $\Omega = \{\omega_1, \cdots, \omega_n\}$, and assume that $X(\omega_i) = x_i$, $i = 1, 2, \cdots, n$. Then

$$Q_X(u) = x_{(i:n)} \text{ where } \frac{i-1}{n} \leq u < \frac{i}{n}$$

Thus, Q_X is an increasing rearrangement of values taken by X.

Proposition12

For all real-valued random variable X, there exists $U \sim \mathcal{U}([0,1])$ such that $X = Q_X(U)$ a.s.

Proof. If F_X is strictly increasing

$$\mathcal{E}_X = \{x|\mathbb{P}(X = x) > 0\} = \emptyset$$

and F_X as well as Q_X are bijective, with $Q_X = F_X^{-1}$ and $F_X = Q_X^{-1}$. Define U as $U(\omega) = F_X(X(\omega))$, then $Q_X(U(\omega)) = X(\omega)$. And U is uniformely distributed since

$$\mathbb{P}(U \leq u) = \mathbb{P}(F_X(X) \leq u) = \mathbb{P}(X \leq Q_X(u)) = F_X(Q_X(u)) = u.$$

More generally, if F_X is not strictly increasing, for all $x \in \mathcal{E}_X$, define some uniform random variable U_x, on $\{u|Q_X(u) = x\}$. Then define

$$U(\omega) = F_X(X(\omega))\mathbf{1}_{\{X(\omega)\notin\mathcal{E}_X\}} + U_{X(\omega)}\mathbf{1}_{\{X(\omega)\in\mathcal{E}_X\}}$$

$\square$

Proposition13

If $X = h(Y)$ where h is some increasing function, and if Q_Y is the quantile function for Y, then $h \circ Q_X$ is the quantile function for X,

$$Q_X(u) = Q_{h \circ Y}(u) = h \circ Q_Y(u)$$

The quantile function is obtained by means of regression, in the sense that

Proposition14

$Q_X(\alpha)$ can be written as a solution of the following regression problem

$$Q_X(\alpha) \in\subset \text{argmin}q \{\mathbb{E}(s_\alpha(X - q))\} \text{ where } s_\alpha(u) = [\alpha - \mathbf{1}(u \leq 0)] \cdot u.$$

Proposition15

A quantile function, as a function of X, is

- PO positive, $X \geq 0$ implies $Q_X(u) \geq 0$, $\forall u \in [0,1]$.
- MO monotone, $X \geq Y$ implies $Q_X(u) \geq Q_Y(u)$, $\forall u \in [0,1]$.
- PH (positively) homogenous, $\lambda \geq 0$ implies $Q_{\lambda X}(u) = \lambda Q_X(u)$, $\forall u \in [0,1]$.
- TI invariant by translation, $k \in \mathbb{R}$ implies $Q_{X-k}(u) = Q_X(X) - k$, $\forall u \in [0,1]$, i.e. $Q_{X-Q_X(u)}(u) = 0$.
- IL invariant in law, $X \sim Y$ implies $Q_X(u) = Q_Y(u)$, $\forall u \in [0,1]$.

Observe that the quantile function is not convex

Proposition16

A quantile function is neither

CO convex, $\forall\lambda\in[0,1]$, $Q_{\lambda X+(1-\lambda)Y}(u)\not\leq\lambda Q_X(u)+(1-\lambda)Q_Y(u)\ \forall u\in[0,1]$.

SA subadditive, $Q_{X+Y}(u)\not\leq Q_X(u)+Q_Y(u)\ \forall u\in[0,1]$.

Example7

Thus, the quantile function as a risk measure might penalize diversification. Consider a corporate bond, with default probabilty p, and with return $\tilde{r}>r$. Assume that the loss is

$$-\frac{\tilde{r}-r}{1+r}w \text{ if there is no default,}$$

$$w \text{ if there is a default.}$$

Assume that $p\leq u$, then

$$p=\mathbb{P}\left(X>-\frac{\tilde{r}-r}{1+r}w\right)\leq u$$

thus

$$Q_X(u) \leq -\frac{\tilde{r}-r}{1+r}w < 0$$

and X can be seen as acceptable for risk level u.

Consider now two independent, identical bonds, X_1 and X_2. Let $Y = \frac{1}{2}(X_1 + X_2)$. If we assume that the return for Y satifies $\tilde{r} \in [r, 1+2r]$, then

$$\frac{\tilde{r}-r}{1+r} < 1 \text{ i.e. } \frac{\tilde{r}-r}{1+r}w < w.$$

$$Q_{\frac{1}{2}[X_1+X_2]}(u) \geq \frac{w}{2}\left(1 - \frac{\tilde{r}-r}{1+r}\right) > Q_X(u).$$

Thus, if the quantile is used as a risk measure, it might penalize diversification.

Example8

From [12]. Since the quantile function as a risk measure is not subadditive, it is possible to subdivide the risk into n desks to minimize the overall capital, i.e.

$$\inf\left\{\sum_{i=1}^{n} Q_{X_i}(u) \middle| \sum_{i=1}^{n} X_i = X\right\}.$$

If we subdivide the support of X on $\mathcal{X} = \bigcup_{j=1}^{m} [x_{j-1}, x_j)$ such that $\mathbb{P}(X \in [x_{j-1}, x_j)) < \alpha$. Let $X_i = X \cdot \mathbf{1}_{X \in [x_{j-1}, x_j)}$. Then $P(X_i > 0) < \alpha$ and $Q_{X_i}(\alpha) = 0$.

4 Univariate and Static Risk Measures

The quantile was a natural risk measure when X was a loss. In this section, we will define risk measures that will be large when $-X$ is large. And we will try to understand the unlying axiomatic, for some random variable X.

The dual of L^p, with the $||\cdot||_p$-norm is L^q, if $p\in[1,\infty)$, and then, $<\boldsymbol{s},\boldsymbol{x}>=\mathbb{E}(\boldsymbol{s}\boldsymbol{x})$. As we will see here, the standard framework is to construct convex risk measures on L^∞. But to derive (properly) a dual representation, we need to work with a weak topology on the dual of L^∞, and some lower semi-continuity assumption is necessary.

Definition5

The Value-at-Risk of level α is

$$\text{VaR}_\alpha(X)=-Q_X(\alpha)=Q_{1-\alpha}(-X).$$

Risk X is said to be VaR_α-acceptable if $\text{VaR}_\alpha(X)\leq 0$.

More generally, let $\mathcal{R}$ denote a monetary risk measure.

Definition6

A monetary risk measure is a mapping $L^p(\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \overline{\mathbb{R}}$

Definition7

A monetary risk measure $\mathcal{R}$ can be

- PO positive, $X \geq 0$ implies $\mathcal{R}(X) \leq 0$
- MO monotone, $X \geq Y$ implies $\mathcal{R}(X) \leq \mathcal{R}(Y)$.
- PH (positively) homogenous, $\lambda \geq 0$ implies $\mathcal{R}(\lambda X) = \lambda \mathcal{R}(X)$.
- TI invariant by translation, $k \in \mathbb{R}$ implies $\mathcal{R}(X + k) = \mathcal{R}(X) - k$,
- IL invariant in law, $X \sim Y$ implies $\mathcal{R}(X) = \mathcal{R}(Y)$.
- CO convex, $\forall \lambda \in [0,1]$, $\mathcal{R}\lambda X + (1 - \lambda Y)) \leq \lambda \mathcal{R}(X) + (1 - \lambda)\mathcal{R}(Y)$.
- SA subadditive, $\mathcal{R}(X + Y) \leq \mathcal{R}(X) + \mathcal{R}(Y)$.

The interpretation of [TI] is now that $\mathcal{R}(X + \mathcal{R}(X)) = 0$.

And property [PH] implies $\mathcal{R}(0) = 0$ (which is also called the grounded property).

Observe that if $\mathcal{R}$ satisfies [TI] and [CO],

$$\mathcal{R}(\mu + \sigma Z) = \sigma\mathcal{R}(Z) - \mu.$$

Definition8

A risk measure is convex if it satisfies [MO], [TI] and [CO].

Proposition17

If $\mathcal{R}$ is a convex risk measure, normalized (in the sense that $\mathcal{R}(0) = 0$), then, for all $\lambda \geq 0$

$$\begin{cases} 0 \leq \lambda \leq 1, & \mathcal{R}(\lambda X) \leq \lambda\mathcal{R}(X) \\ 1 \leq \lambda, & \mathcal{R}(\lambda X) \geq \lambda\mathcal{R}(X). \end{cases}$$

Definition9

A risk measure is coherent if it satisfies [MO], [TI], [CO] and [PH].

If $\mathcal{R}$ is coherent, then it is normalized, and then, convexity and sub-additivity are equivalent properties,

Proposition18

If $\mathcal{R}$ is a coherent risk measure, [CO] is equivalent to [SA]

Proof. If $\mathcal{R}$ satistfies [SA] then

$$\mathcal{R}(\lambda X + (1-\lambda)Y) \leq \mathcal{R}(\lambda X) + \mathcal{R}((1-\lambda)Y)$$

and [CO] is obtained by [PH].

If $\mathcal{R}$ satistfies [CO] then

$$\mathcal{R}(X+Y) = 2\mathcal{R}\left(\frac{1}{2}X + \frac{1}{2}Y\right) \leq \frac{2}{2}\left(\mathcal{R}(X) + \mathcal{R}(Y)\right)$$

and [SA] is obtained by [PH]. □

Proposition19

If $\mathcal{R}$ is a coherent risk measure, then if $X \in [a, b]$ a.s., then $\mathcal{R}(X) \in [-b, -a]$.

Proof. Since $X - a \geq 0$, then $\mathcal{R}(X - a) \leq 0$ (since $\mathcal{R}$ satisfies [MO]), and $\mathcal{R}(X - a) = \mathcal{R}(X) + a$ by [TI]. So $\mathcal{R}(X) \leq -a$. Similarly, $b - X \geq 0$, so $\mathcal{R}(b - X) \leq 0$ (since $\mathcal{R}$ satisfies [MO]), and $\mathcal{R}(b - X) = \mathcal{R}(-X) - b$ by [TI]. Since R is coherent, $\mathcal{R}(0) = 0$ and $\mathcal{R}(-X) = -\mathcal{R}(X)$. So $\mathcal{R}(b - X) = -\mathcal{R}(X) - b \leq 0$ i.e. $\mathcal{R}(X) \geq -b$. $\square$

Other properties can be mentioned ([E] from [32] and [16])

Definition10

A risk measure is

- E elicitability if there is a (positive) score function s such that

$$\mathbb{E}[s(X - \mathcal{R}(X))] \leq \mathbb{E}[s(X - x)] \text{ for any } x \in \mathbb{R}$$

- QC quasi-convexity, $\mathcal{R}(\lambda X + (1-\lambda)Y) \leq \max\{\mathcal{R}(X), \mathcal{R}(Y)\}$ for any $\lambda \in [0,1]$.

- FP L^p-Fatou property if given $(X_n) \in L^p$ bounded with, $p \in [1, \infty)$, and $X \in L^p$ such that $X_n \stackrel{L^p}{\to} X$, then

$$\mathcal{R}(X) \leq \text{liminf}\{\mathcal{R}(X_n)\}$$

Recall that the limit inferior of a sequence (u_n) is defined by $\liminf_{n\to\infty} x_n := \lim_{n\to\infty} \left(\inf_{m\geq n} x_m \right)$. One should keep in mind that the limit inferior satisfies a superadditivity property, since

$$\liminf_{n\to\infty}(u_n + v_n) \geq \liminf_{n\to\infty}(u_n) + \liminf_{n\to\infty}(v_n).$$

4.1 From risk measures to acceptance sets

Definition11

Let $\mathcal{R}$ denote some risk measure. The associated acceptance set is

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

Proposition20

If $\mathcal{R}$ is a risk measure satisfying [MO] and [TI]

1. $\mathcal{A}_{\mathcal{R}}$ is a closed set
2. $\mathcal{R}$ can be recovered from $\mathcal{A}_{\mathcal{R}}$,

$$\mathcal{R}(X) = \inf\{m|X - m \in \mathcal{A}_{\mathcal{R}}\}$$

3. $\mathcal{R}$ is convex if and only if $\mathcal{A}_{\mathcal{R}}$ is a convex set
4. $\mathcal{R}$ is coherent if and only if $\mathcal{A}_{\mathcal{R}}$ is a convex cone

Proof. (1) Since $X - Y \leq ||X - Y||_\infty$, we get that $X \leq Y + ||X - Y||_\infty$, so if we use [MO] and [TI],

$$\mathcal{R}(X) \leq \mathcal{R}(Y) + ||X - Y||_\infty$$

and similarly, we can write

$$\mathcal{R}(Y) \leq \mathcal{R}(X) + ||X - Y||_\infty,$$

so we get

$$|\mathcal{R}(Y) - \mathcal{R}(X)| \leq ||X - Y||_\infty$$

So risk measure $\mathcal{R}$ is Lipschitz (with respect to the $||\cdot||_\infty$-norm, so $\mathcal{R}$ is continous, and thus, $\mathcal{A}_\mathcal{R}$ is necessarily a closed set.

(2) Since $\mathcal{R}$ satisfies [TI],

$$\inf\{m|X - m \in \mathcal{A}_\mathcal{R}\} = \inf\{m|\mathcal{R}(X - m) \leq 0\} = \inf\{m|\mathcal{R}(X) \leq m\} = \mathcal{R}.$$

(3) If $\mathcal{R}$ is convex then clearly $\mathcal{A}_\mathcal{R}$ is a convex set. Now, consider that $\mathcal{A}_\mathcal{R}$ is a convex set. Let X_1, X_2 and m_1, m_2 such that $X_i - m_i \in \mathcal{A}_\mathcal{R}$. Since $\mathcal{A}_\mathcal{R}$ is convex,

for all $\lambda \in [0,1]$,

$$\lambda(X_1 - m_1) + (1-\lambda)(X_2 - m_2) \in \mathcal{A}_{\mathcal{R}}$$

so

$$\mathcal{R}(\lambda(X_1 - m_1) + (1-\lambda)(X_2 - m_2)) \leq 0.$$

Now, since $\mathcal{R}$ satisfies [TI],

$$\begin{aligned}
\mathcal{R}(\lambda X_1 + (1-\lambda)X_2) &\leq \lambda m_1 + (1-\lambda)m_2 \\
&\leq \lambda \inf\{m|X_1 - m \in \mathcal{A}_{\mathcal{R}}\} + (1-\lambda)\inf\{m|X_2 - m \in \mathcal{A}_{\mathcal{R}}\} \\
&= \lambda\mathcal{R}(X_1) + (1-\lambda)\mathcal{R}(X_2).
\end{aligned}$$

(4) If $\mathcal{R}$ satisfies [PH] then clearly $\mathcal{A}_{\mathcal{R}}$ is a cone. Conversely, consider that $\mathcal{A}_{\mathcal{R}}$ is a cone. Let X and m. If $X - m \in \mathcal{A}_{\mathcal{R}}$, then $\mathcal{R}(\lambda(X-m)) \leq 0$, and $\lambda(X-m) \in \mathcal{A}_{\mathcal{R}}$ so

$$\mathcal{R}(\lambda X) \leq \lambda m \leq \lambda \inf\{m|\mathcal{R}(X) \leq m\} = \lambda\mathcal{R}(X)$$

And if $X - m \notin \mathcal{A}_{\mathcal{R}}$, then $\mathcal{R}(\lambda(X-m)) > 0$, and

$$\mathcal{R}(\lambda X) > \lambda m \geq \lambda \sup\{m|\mathcal{R}(X) \geq m\} = \lambda\mathcal{R}(X)$$

□

Example9

Let $u(\cdot)$ denote a concave utility function, strictly increasing, and

$$\mathcal{R}(X) = u^{-1}\left(\mathbb{E}[u(X)]\right) \text{ is the certain equivalent.}$$

The acceptance set is

$$\mathcal{A} = \{X \in L^\infty | \mathbb{E}[u(X)] \leq u(0)\}$$

which is a convex set.

4.2 Representation of L^∞ risk measures

Let $X \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$.

Let $\mathcal{M}_1(\mathbb{P})$ denote the set of probability measures, $\mathcal{M}_1(\mathbb{P}) = \{\mathbb{Q}|\mathbb{Q} \ll \mathbb{P}\}$, and $\mathcal{M}_{1,f}(\mathbb{P})$ denote the set of additive measures, $\mathcal{M}_{1,f}(\mathbb{P}) = \{\nu|\nu \ll \mathbb{P}\}$.

Definition12

Let $\nu \in \mathcal{M}_{1,f}(\mathbb{P})$, then Choquet's integral is defined as

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

In this section, $\mathbb{Q}$ will denote another measure, which could be a probability measure, or simply a finitely-additive one.

Consider a functional $\alpha : \mathcal{M}_{1,f}(\mathbb{P}) \to \mathbb{R}$ such that $\inf_{\mathbb{Q}\in\mathcal{M}_{1,f}(\mathbb{P})}\{\alpha(\mathbb{Q})\} \in \mathbb{R}$, then for all $\mathbb{Q} \in \mathcal{M}_{1,f}(\mathbb{P})$

$$\mathcal{R} : X \mapsto \mathbb{E}_\mathbb{Q}(X) - \alpha(\mathbb{Q})$$

is a (linear) convex risk measure, and this property still hold by taking the supremum on all measures $\mathbb{Q} \in \mathcal{M}_{1,f}(\mathbb{P})$,

$$\mathcal{R} : X \mapsto \sup_{\mathbb{Q} \in \mathcal{M}_{1,f}(\mathbb{P})} \{\mathbb{E}_{\mathbb{Q}}(X) - \alpha(\mathbb{Q})\}.$$

Such a measure is convex, and $\mathcal{R}(0) = -\inf_{\mathbb{Q} \in \mathcal{M}_{1,f}(\mathbb{P})} \{\alpha(\mathbb{Q})\}$.

Proposition21

A risk measure $\mathcal{R}$ is convex if and only if

$$\mathcal{R}(X) = \max_{\mathbb{Q} \in \mathcal{M}_{1,f}(\mathbb{P})} \{\mathbb{E}_{\mathbb{Q}}(X) - \alpha_{\min}(\mathbb{Q})\},$$

where $\alpha_{\min}(\mathbb{Q}) = \sup_{X \in \mathcal{A}_{\mathcal{R}}} \{\mathbb{E}_{\mathbb{Q}}(X)\}$.

What we have here is that any convex risk measure can be written as a worst expected loss, corrected with some random penalty function, with respect to some given set of probability measures.

In this representation, the risk measure is characterized in terms of finitely additive measures. As mentioned in [?], is we want a representation in terms of probability measures (set $\mathcal{M}_1$ instead of $\mathcal{M}_{1,f}$) additional continuity properties are necessary.

Proof. From the definitions of $\alpha_{\min}$ and $\mathcal{A}_{\mathcal{R}}$, $X - \mathcal{R}(X) \in \mathcal{A}_{\mathcal{R}}$ for all $X \in L^\infty$. Thus,

$$\alpha_{\min}(\mathbb{Q}) \geq \sup_{X \in L^\infty} \{\mathbb{E}_{\mathbb{Q}}[X - \mathcal{R}(X)]\} = \sup_{X \in L^\infty} \{\mathbb{E}_{\mathbb{Q}}[X] - \mathcal{R}(X)\}$$

which is Fenchel's transform of $\mathcal{R}$ in L^∞ Since $\mathcal{R}$ is Lipschitz, it iscontinuous with respect to the L^∞ norm, and therefore $\mathcal{R}^{\star\star} = \mathcal{R}$. Thus

$$\begin{aligned}
\mathcal{R}(X) &= \sup_{\mathbb{Q} \in L^{\infty\star}} \{\mathbb{E}_{\mathbb{Q}}(X) - \mathcal{R}^\star(X)\} \\
&= \sup_{\mathbb{Q} \in L^{\infty\star}} \{\mathbb{E}_{\mathbb{Q}}(X) - \alpha_{\min}(\mathbb{Q})\}
\end{aligned}$$

Hence, we get that

$$\alpha_{\min}(\mathbb{Q}) = \sup_{X \in L^\infty} \{\mathbb{E}_{\mathbb{Q}}(X) - \mathcal{R}(X)\} = \sup_{X \in \mathcal{A}_{\mathcal{R}}} \{\mathbb{Q}(X)\}$$

To conclude, we have to prove that the supremum is attained in the subspace of $L^{\infty\star}$, denoted $\mathcal{M}_{1,f}(\mathbb{P})$. Let μ denote some positive measure,

$$\mathcal{R}^\star(\mu) = \sup_{X\in L^\infty}\{\mathbb{E}_\mu(X) - \mathcal{R}(X)\}$$

but since $\mathcal{R}$ satisfies [TI],

$$\mathcal{R}^\star(\mu) = \sup_{X\in L^\infty}\{\mathbb{E}_\mu(X-1) - \mathcal{R}(X)+1\}$$

Hence, $\mathcal{R}^\star(\mu) = \mathcal{R}^\star(\mu)+1-\mu(1)$, so $\mu(1)=1$. Further

$$\begin{aligned}\mathcal{R}^\star(\mu) &\geq \mathbb{E}_\mu(\lambda X) - \mathcal{R}(\lambda X) \text{ for } \lambda\leq 0\\ &\geq \lambda\mathbb{E}_\mu(X) - \mathcal{R}(0) \text{ for } X\leq 0\end{aligned}$$

so, for all $\lambda\leq 0$, $\lambda\mathbb{E}_\mu(X)\leq\mathcal{R}(0)+\mathcal{R}^\star(\mu)$, and

$$\begin{aligned}\mathbb{E}_\mu(X) &\geq \lambda^{-1}(\mathcal{R}(0)+\mathcal{R}^\star(\mu)),\ \text{for any } \lambda\leq 0,\\ &\geq 0.\end{aligned}$$

So, finally,

$$\mathcal{R}(X) = \sup_{\mathbb{Q}\in\mathcal{M}_{1,f}(\mathbb{P})} \left\{\mathbb{E}_{\mathbb{Q}}(X) - \alpha_{\min}(\mathbb{Q})\right\},$$

where $\alpha_{\min}(\mathbb{Q}) = \sup_{X\in\mathcal{A}_{\mathcal{R}}} \{\mathbb{E}_{\mathbb{Q}}(X)\}$. To conclude, (i) we have to prove that the supremum can be attained. And this is the case, since $\mathcal{M}_{1,f}$ is a closed unit ball in the dual of L^∞ (with the total variation topoplogy). And (ii) that $\alpha_{\min}$ is, indeed the minimal penalty.

Let α denote a penalty associated with $\mathcal{R}$, then, for any $\mathbb{Q}\in\mathcal{M}_{1,f}(\mathbb{P})$ and $X\in L^\infty$,

$$\mathcal{R}(X) \geq \mathbb{E}_{\mathbb{Q}}(X) - \alpha(\mathbb{Q}),$$

and

$$\begin{aligned}\alpha(\mathbb{Q}) &\geq \sup_{X\in L^\infty}\{\mathbb{E}_{\mathbb{Q}}(X) - \mathcal{R}(X)\}\\ &\geq \sup_{X\in\mathcal{A}_{\mathcal{R}}}\{\mathbb{E}_{\mathbb{Q}}(X) - \mathcal{R}(X)\} \geq \sup_{X\in\mathcal{A}_{\mathcal{R}}}\{\mathbb{E}_{\mathbb{Q}}(X)\} = \alpha_{\min}(\mathbb{Q})\end{aligned}$$

$\square$

The minimal penalty function of a coherent risk measure will take only two values, 0 and $+\infty$. Observe that if $\mathcal{R}$ is coherent, then, from [PH], for all $\lambda \geq 0$,

$$\alpha_{\min}(\mathbb{Q}) = \sup_{X\in L^\infty} \{\mathbb{E}_{\mathbb{Q}}(\lambda X) - \mathcal{R}(\lambda X)\} = \lambda\alpha_{\min}(\mathbb{Q}).$$

Hence, $\alpha_{\min}(\mathbb{Q}) \in \{0, \infty\}$, and

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

where

$$\mathcal{Q} = \{\mathbb{Q} \in \mathcal{M}_{1,f}(\mathbb{P}) | \alpha_{\min}(\mathbb{Q}) = 0\}.$$

Proposition22

Consider a convex risk measure $\mathcal{R}$, then $\mathcal{R}$ can be represented by a penalty function on $\mathcal{M}_1(\mathbb{P})$ if and only if $\mathcal{R}$ satisfies [FP].

Proof. $\Longrightarrow$ Suppose that $\mathcal{R}$ can be represented using the restriction of $\alpha_{\min}$ on $\mathcal{M}_1(\mathbb{P})$. Consider a sequence (X_n) of L^∞, bounded, such that $X_n \to X$ a.s. From the dominated convergence theorem, for any $\mathbb{Q} \in \mathcal{M}_1(\mathbb{P})$,

$$\mathbb{E}_{\mathbb{Q}}(X_n) \to \mathbb{E}_{\mathbb{Q}}(X) \text{ as } n \to \infty,$$

so

$$\begin{aligned}
\mathcal{R}(X) &= \sup_{\mathbb{Q}\in\mathcal{M}_1(\mathbb{P})} \{\mathbb{E}_{\mathbb{Q}}(X) - \alpha_{\min}(\mathbb{Q})\} \\
&= \sup_{\mathbb{Q}\in\mathcal{M}_1(\mathbb{P})} \left\{\lim_{n\to\infty} \mathbb{E}_{\mathbb{Q}}(X_n) - \alpha_{\min}(\mathbb{Q})\right\} \\
&\leq \liminf_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{M}_1(\mathbb{P})} \{\mathbb{E}_{\mathbb{Q}}(X_n) - \alpha_{\min}(\mathbb{Q})\} \\
&= \liminf_{n\to\infty} \mathcal{R}(X_n)
\end{aligned}$$

so [FP] is satisfied.

$\Longleftarrow$ Conversely, let us prove that [FP] implies lower semi-continuity with respect to some topology on L^∞ (seen as the dual of L^1). The strategy is to prove that

$$\mathcal{C}_r = \mathcal{C} \cap \{X \in L^\infty | \; ||X||_\infty \leq r\}$$

is a closed set, for all $r > 0$, where $\mathcal{C} = \{X|\mathcal{R}(X) < c \text{ for some } c\}$. Once we have that $\mathcal{R}$ is l.s.c., then Fenchel-Moreau theorem can be invoked, $\mathcal{R}^{\star\star} = \mathcal{R}$, and $\alpha_{\min} = \mathcal{R}^\star$. $\square$

Several operations can be considered on risk measures,

Proposition23

If $\mathcal{R}_1$ and $\mathcal{R}_2$ are coherent risk measures, then $\mathcal{R} = \max\{\mathcal{R}_1, \mathcal{R}_2\}$ is coherent. If $\mathcal{R}_i$'s are convex risk measures, then $\mathcal{R} = \sup\{\mathcal{R}_i\}$ is convex, and further, $\alpha = \inf\{\alpha_i\}$.

Proof. Hence

$$\mathcal{R}(X) = \sup_i \left\{ \sup_{\mathbb{Q}\in\mathcal{M}_{1,f}(\mathbb{P})} \{\mathbb{E}_{\mathbb{Q}}(X) - \alpha_i(\mathbb{Q})\} \right\} = \sup_{\mathbb{Q}\in\mathcal{M}_1(\mathbb{P})} \left\{ \mathbb{E}_{\mathbb{Q}}(X) - \inf_i \{\alpha_i(\mathbb{Q})\} \right\}.$$

□

4.3 Expected Shortfall

Definition13

The expected shortfall of level $\alpha \in (0,1)$ is

$$ES_X(\alpha) = \frac{1}{1-\alpha}\int_\alpha^1 Q_X(u)du$$

If $\mathbb{P}(X = Q_X(\alpha)) = 0$ (e.g. X is absolutely continuous),

$$ES_X(\alpha) = \mathbb{E}(X|X \geq Q_X(\alpha))$$

and if not,

$$ES_X(\alpha) = \mathbb{E}(X|X \geq Q_X(\alpha)) + [\mathbb{E}(X|X \geq Q_X(\alpha)) - Q_X(\alpha)]\left(\frac{\mathbb{P}(X \geq Q_X(\alpha))}{1-\alpha} - 1\right)$$

Proposition24

The expected shortfall of level $\alpha \in (0,1)$ can be written

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

where

$$\mathcal{Q}_\alpha = \left\{\mathbb{Q} \in \mathcal{M}_1(\mathbb{P}) \,\middle|\, \frac{d\mathbb{Q}}{d\mathbb{P}} \leq \frac{1}{\alpha},\ \text{a.s.}\right\}$$

Hence, we can write

$$ES_X(\alpha) = \sup\{\mathbb{E}(X|A)|\mathbb{P}(A) > \alpha\} \geq Q_X(\alpha).$$

Proof. Set $\mathcal{R}(X) = \sup_{\mathbb{Q}\in\mathcal{Q}_\alpha}\{\mathbb{E}_{\mathbb{Q}}(X)\}$. Let us prove that this supremum can be attained, and then, that $\mathcal{R}(X) = ES_X(\alpha)$. Let us restrict ourself here to the case where $\mathbb{E}(X) = 1$ and $X \geq 0$ (the general case can then be derived, since

Let $\tilde{\mathbb{P}}$ denote the distribution of X, so that

$$\begin{aligned}\sup_{\mathbb{Q}\in\mathcal{Q}_\alpha}\{\mathbb{E}_{\mathbb{Q}}(X)\} &= \sup_{\mathbb{Q}\in\mathcal{Q}_\alpha}\left\{\mathbb{E}_{\mathbb{P}}\left(X\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right\} = \sup_{Y\in[0,1/\alpha]}\left\{\mathbb{E}_{\mathbb{P}}\left(Y\frac{d\mathbb{Q}}{d\tilde{\mathbb{P}}}\right)\right\}\\ &= \frac{1}{\alpha}\sup_{Y\in[0,1],\mathbb{E}(Y)=\alpha}\{\mathbb{E}_{\tilde{\mathbb{P}}}(Y)\}\end{aligned}$$

The supremum is then attained for

$$Y_\star = \mathbf{1}_{X>Q_X(1-\alpha)} + \kappa\mathbf{1}_{X=Q_X(1-\alpha)}$$

where κ is chosen to have $\mathbb{E}(Y)=\alpha$, since

$$\mathcal{R}(X) = \mathbb{E}_{\tilde{\mathbb{P}}}\left(\frac{Y}{\alpha}\right) = \mathbb{E}_{\mathbb{P}}\left(\frac{XY}{\alpha}\right) = \mathbb{E}_{\mathbb{Q}_\star}(X).$$

(see previous discussion on Neyman-Pearson's lemma). Thus,

$$\frac{d\mathbb{Q}_\star}{d\mathbb{P}} = \frac{1}{\alpha}\left[\mathbf{1}_{X>Q_X(1-\alpha)} + \kappa\mathbf{1}_{X=Q_X(1-\alpha)}\right]$$

If $\mathbb{P}(X = Q_X(1-\alpha))$, then $\kappa = 0$; if not,

$$\kappa = \frac{\alpha - \mathbb{P}(X > Q_X(1-\alpha))}{\mathbb{P}(X = Q_X(1-\alpha)}.$$

So, if we substitute,

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}_\star}(X) &= \frac{1}{\alpha}\left(\mathbb{E}[X\mathbf{1}_{\{X>Q_X(1-\alpha)\}} + [\alpha - \mathbb{P}(X > Q_X(1-\alpha))]\, Q_X(1-\alpha)]\right) \\
&= \frac{1}{\alpha}\left(\mathbb{E}(X - Q_X(1-\alpha))_+ + \alpha Q_X(1-\alpha)\right) \\
&= \frac{1}{\alpha}\left(\int_{1-\alpha}^{1} (Q_X(t) - Q_X(1-\alpha))_+ dt + \alpha Q_X(1-\alpha)\right) \\
&= \frac{1}{\alpha}\int_{1-\alpha}^{1} Q_X(t)dt = ES_X(\alpha).
\end{aligned}$$

$\square$

Remark3

Observe that if $\mathbb{P}(X = Q_X(1-\alpha)) = 0$, i.e. $\mathbb{P}(X > Q_X(1-\alpha)) = \alpha$, then

$$ES_X(\alpha) = \mathbb{E}(X|X > Q_X(1-\alpha)).$$

Proposition25

If $\mathcal{R}$ is a convex risk measure satisfying [IL], exceeding the quantile of level $1-\alpha$, then $\mathcal{R}(X) \geq ES_X(1-\alpha)$.

Proof. Let $\mathcal{R}$ denote a risk measure satisfying [CO] and [IL], such that $\mathcal{R}(X) \geq Q_X(1-\alpha)$. Given $\varepsilon > 0$, set $A = \{X \geq Q_X(1-\alpha) - \varepsilon\}$ and

$$Y = X\mathbf{1}_{A^\mathsf{C}} + \mathbb{E}(X|A)\mathbf{1}_A.$$

Then $Y \leq Q_X(1-\alpha) - \varepsilon \leq \mathbb{E}(X|A)$ on A^C, so $\mathbb{P}(Y > \mathbb{E}(X|A)) = 0$. On the other hand,

$$\mathbb{P}(Y \geq \mathbb{E}(X|A)) \geq \mathbb{P}(A) > \alpha,$$

so with those two results, we get that $Q_Y(1-\alpha)=\mathbb{E}(X|A)$. And because $\mathbb{R}$ dominates the quantile, $\mathcal{R}(Y)\geq Q_Y(1-\alpha)=\mathbb{E}(X|A)$. By Jensen inequality (since $\mathcal{R}$ is convex),

$$\mathcal{R}(X)\geq\mathcal{R}(Y)\geq\underbrace{\mathbb{E}(X|A)}_{\mathbb{E}(X|Q_X(1-\alpha)+\varepsilon)}$$

for any $\varepsilon>0$. If $\varepsilon\downarrow 0$, we get that

$$\mathcal{R}(X)\geq ES_X(1-\alpha).$$

$\square$

4.4 Expectiles

For quantiles, an asymmetric linear loss function is considered,

$$h_\alpha(t) = |\alpha - \mathbf{1}_{t\leq 0}| \cdot |t| = \begin{cases} \alpha|t| \text{ if } t > 0 \\ (1\text{-}\alpha)|t| \text{ if } t \leq 0 \end{cases}$$

For expectiles - see [27] - an asymmetric quadratic loss function is considered,

$$h_\alpha(t) = |\alpha - \mathbf{1}_{t\leq 0}| \cdot t^2 = \begin{cases} \alpha t^2 \text{ if } t > 0 \\ (1\text{-}\alpha)t^2 \text{ if } t \leq 0 \end{cases}$$

Definition14

The expectile of X with probability level $\alpha \in (0,1)$ is

$$e_X(\alpha) = \underset{e\in\mathbb{R}}{\text{argmin}} \left\{\mathbb{E}\left[\alpha(X-e)_+^2 + (1-\alpha)(e-X)_+^2 +\right]\right\}$$

The associated expectile-based risk measure is $\mathcal{R}_\alpha(X) = e_X(\alpha) - \mathbb{E}(X)$.

Observe that $e_X(\alpha)$ is the unique solution of

$$\alpha\mathbb{E}[(X-e)_+] = (1-\alpha)\mathbb{E}[(e-x)_+]$$

Further, $e_X(\alpha)$ is subadditive for $\alpha \in [1/2, 1]$.

As proved in [20], expectiles are quantiles, but not associated with F_X,

$$G(x) = \frac{\mathbb{P}(X=x) - xF_X(x)}{2[\mathbb{P}(X=x) - xF_X(x)] + (x - \mathbb{E}(x))}$$

Let

$$\mathcal{A} = \{Z|\mathbb{E}_{\mathbb{P}}[(\alpha-1)Z_- + \alpha Z_+] \geq 0\}$$

then

$$e_\alpha(X) = \max\{Z|Z - X \in \mathcal{A}\}$$

Further

$$e_\alpha(X) = \min_{\mathbb{Q}\in\mathcal{S}}\{\mathbb{E}_{\mathbb{Q}[X]}\}$$

where

$$\mathcal{S} = \left\{ \mathbb{Q} \middle| \text{there is } \beta > 0 \text{ such that } \beta \leq \frac{d\mathbb{Q}}{d\mathbb{P}} \leq \frac{1-\alpha}{\alpha}\beta \right\}$$

Remark4

When $\alpha \to 0$, $E_\alpha(X) \to \text{essinf} X$.

Let $\gamma = (1-\alpha)/\alpha$, then $e_\alpha(X)$ is the minimum of

$$e \mapsto \int_0^1 Q_Z(u)du \text{ with } Z = \frac{1_{[e,1]} + \beta 1_{[0,x]}}{1 + (\gamma - 1)e}$$

Let $f(x) = \dfrac{x}{\gamma - (\gamma - 1)x}$. f is a convex distortion function, and $f \circ \mathbb{P}$ is a subadditive capacity. And the expectile can be represented as

$$E_\alpha(X) = \inf_{\mathbb{Q} \in \mathcal{S}} \left\{ \int_0^1 ES_u(X)\nu(du) \right\}$$

where

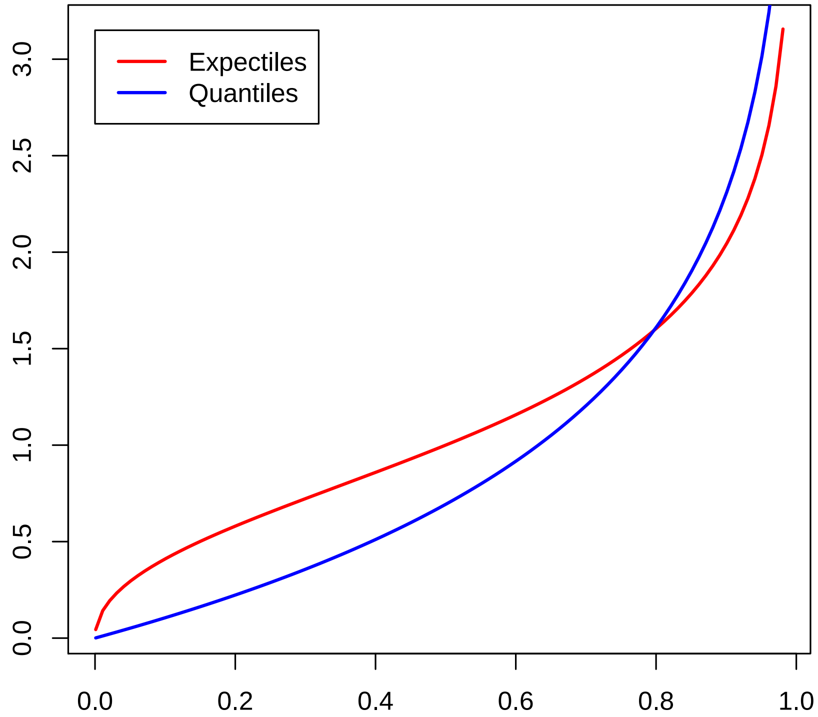
$$\mathcal{S} = \left\{ \mathbb{Q} | \int_0^1 \frac{\mathbb{Q}(du)}{u} \leq \gamma \mathbb{Q}(\{1\}) \right\}.$$

Observe that $X \mapsto E_\alpha(X)$ is continuous. Actually, it is Lipschitz, in the sense that

$$|E_\alpha(X) - E_\alpha(Y)| \leq \sup_{\mathbb{Q}\in\mathcal{S}}\{\mathbb{E}_{\mathbb{Q}}(|X-Y|)\} \leq \gamma||X-Y|vert_1.$$

Example10

The case where $X \sim \mathcal{E}(1)$ can be visualized on the left of Figure 3, while the case $X \sim \mathcal{N}(0,1)$ can be visualized on the right of Figure 3.



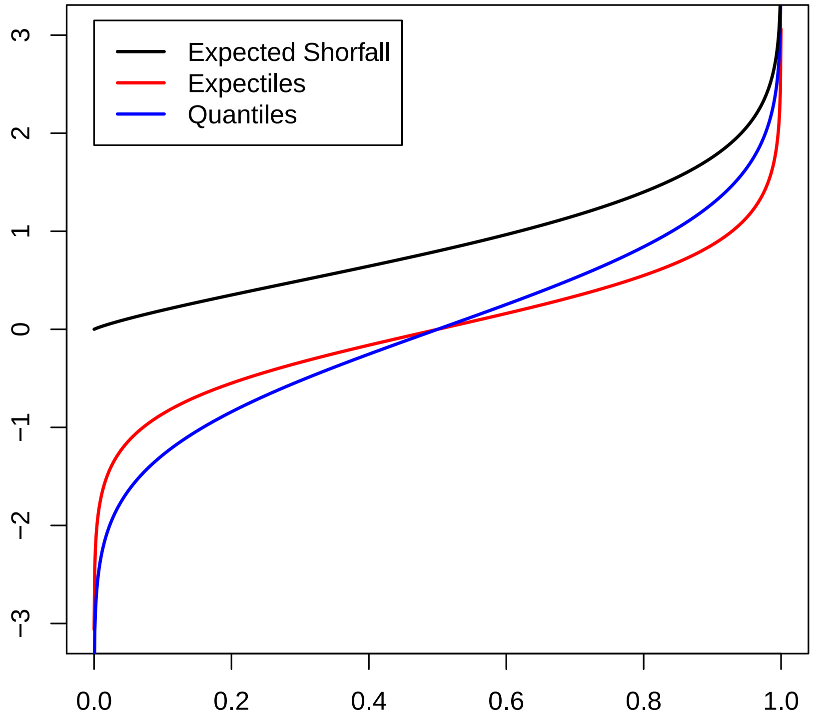


Figure 3: Quantiles, Expected Shortfall and Expectiles, $\mathcal{E}(1)$ and $\mathcal{N}(0,1)$ risks.

4.5 Entropic Risk Measure

The entropic risk measure with parameter α (the risk aversion parameter) is defined as

$$\mathcal{R}_\alpha(X) = \frac{1}{\alpha}\log\left(\mathbb{E}_{\mathbb{P}}[e^{-\alpha X}]\right) = \sup_{\mathbb{Q}\in\mathcal{M}_1}\left\{\mathbb{E}_{\mathbb{Q}}[-X] - \frac{1}{\alpha}H(\mathbb{Q}|\mathbb{P})\right\}$$

where $H(\mathbb{Q}|\mathbb{P}) = \mathbb{E}_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\log\frac{d\mathbb{Q}}{d\mathbb{P}}\right]$ is the relative entropy of $\mathbb{Q} \ll \mathbb{P}$.

One can easily prove that for any $\mathbb{Q} \ll \mathbb{P}$,

$$H(\mathbb{Q}|\mathbb{P}) = \sup_{X\in L^\infty}\left\{\mathbb{E}_{\mathbb{Q}}(-X) - \frac{1}{\alpha}\log\mathbb{E}(e^{-\alpha X})\right\}$$

and the supremum is attained when $X = -\frac{1}{\gamma}\log\frac{d\mathbb{Q}}{d\mathbb{P}}$.

Observe that

$$\frac{d\mathbb{Q}_X}{d\mathbb{P}} = \frac{e^{-\gamma X}}{\mathbb{E}(e^{-\gamma X})}$$

which is the popular Esscher transform

Observe that the acceptance set for the entropic risk measure is the set of payoffs with positive expected utility, where the utility is the standard exponential one, $u(x) = 1 - e^{-\alpha x}$, which has constant absolute risk aversion, in the sense that

$$\frac{-u''(x)}{u'(x)} = \alpha \text{ for any } x.$$

The acceptance set is here

$$\mathcal{A} = \{X \in L^p | \mathbb{E}[u(X)] \geq 0\} = \{X \in L^p | \mathbb{E}_{\mathbb{P}}\left[e^{-\alpha X}\right] \leq 1\}$$

5 Comonotonicity, Maximal Correlation and Optimal Transport

Heuristically, risks X and Y are comonotonic if both suffer negative shocks in the same states $\omega \in \Omega$, so it is not possible to use one to hedge the other. So in that case, there might be no reason to expect that the risk of the sum will ne smaller than the sum of the risks (as obtained with convex or subadditive risk measures).

5.1 Comonotonicity

Definition15

Let X and Y denote two random variables on Ω. Then X and Y are comonotonic random variables if

$$[X(\omega) - X(\omega')] \cdot [Y(\omega) - Y(\omega')] \geq 0$$

for all $\omega, \omega' \in \Omega$.

Proposition26

X and Y are comonotonic if and only if there exists Z, and f, g two increasing functions such that $X = f(Z)$ and $Y = g(Z)$.

Proof. Assume that X and Y are comonotonic. Let $\omega \in \Omega$ and set $x = X(\omega)$, $y = Y(\omega)$ and $z = Z(\omega)$. Let us prove that if there is ω' such that $z = X(\omega') + Y(\omega')$, then necessarily $x = X(\omega')$ and $y = Y(\omega')$.

Since variables are comonotonic, $X(\omega') - X(\omega)$ and $Y(\omega') - Y(\omega)$ have the same signe. But $X(\omega') + Y(\omega') = X(\omega) + Y(\omega)$ implies that $X(\omega') - X(\omega) = -[Y(\omega') - Y(\omega)]$. So $X(\omega') - X(\omega) = 0$, i.e. $x = X(\omega')$ and $y = Y(\omega')$.

So z has a unique decomposition $x + y$, so let us write $z = x_z + y_z$. What we need to prove is that $z \mapsto x_z$ and $z \mapsto y_z$ are increasing functions.

Consider ω_1 and ω_2 such that

$$X(\omega_1) + Y(\omega_1) = z_1 \ \leq \ z_2 = X(\omega_2) + Y(\omega_2)$$

Then

$$X(\omega_1) - X(\omega_2) \ \leq \ -[Y(\omega_1) - Y(\omega_2)].$$

If $Y(\omega_1) > Y(\omega_2)$, then

$$[X(\omega_1) - X(\omega_2)] \cdot [Y(\omega_1) - Y(\omega_2)] \leq -[Y(\omega_1) - Y(\omega_2)]^2 < 0,$$

which contracdicts the comonotonic assumption. So $Y(\omega_1) \leq Y(\omega_2)$. So $z_1 \leq z_2$ necessarily implies that $y_{z_1} \leq y_{z_2}$, i.e. $z \mapsto y_z$ is an increasing function (denoted g here). $\square$

Definition16

A risk measure $\mathcal{R}$ is

CA comonotonic addive if $\mathcal{R}(X+Y) = \mathcal{R}(X) + \mathcal{R}(Y)$ when X and Y are comonotonic.

Proposition27

VaR and ES are comontonic risk measures.

Proof. Let X and Y denote two comonotone random variables. Let us prove that $Q_{X+Y}(\alpha) = Q_X(\alpha) + Q_Y(\alpha)$. From the proposition before, there is Z such that $X = f(Z)$ and $Y = g(Z)$, where f and g are increasing functions. We need to prove that $h \circ Q_Z$ is a quantile of $X + Y$, with $h = f + g$. Observe that $X + Y = h(Z)$, and that h is increasing, so

$$\begin{aligned} F_{X+Y}(h \circ Q_Z(t)) &= \mathbb{P}(h(Z) \leq h \circ Q_Z(t)) \geq \mathbb{P}(Z \leq Q_Z(t)) \\ &= F_Z(Q_Z(t)) \geq t \geq \mathbb{P}(Z < Q_Z(t)) \geq F_{X+Y}(h \circ Q_Z(t)^-). \end{aligned}$$

From those two inequalities,

$$F_{X+Y}(h \circ Q_Z(t)) \geq t \geq F_{X+Y}(h \circ Q_Z(t)^-)$$

we get that, indeed, $h \circ Q_Z$ is a quantile of $X + Y$.

Further, we know that $X = Q_X(U)$ a.s. for some U uniformly distributed on the unit interval. So, if X and Y are comonotonic,

$$\begin{cases} X = f(Z) = Q_X(U) \\ Y = g(Z) = Q_Y(U) \end{cases} \text{with } U \sim \mathcal{U}([0,1]),$$

So if we substitute U to Z and $Q_X + Q_Y$ to h, we just proved that $(Q_X + Q_Y) \circ Id = Q_X + Q_Y$ was a quantile function of $X + Y$. □

5.2 Hardy-Littlewood-Polyá and maximal correlation

In the proof about, we mentioned that if X and Y are comonotonic,

$$\begin{cases} X = f(Z) = Q_X(U) \\ Y = g(Z) = Q_Y(U) \end{cases} \text{with } U \sim \mathcal{U}([0,1]),$$

i.e. X and Y can be rearranged simultaneously.

Consider the case of discrete random variables,

$$\begin{cases} X \in \{x_1, x_2, \cdots, x_n\} \text{ with } 0 \leq x_1 \leq x_2 \leq \cdots \leq x_n \\ Y \in \{y_1, y_2, \cdots, y_n\} \text{ with } 0 \leq y_1 \leq y_2 \leq \cdots \leq y_n \end{cases}$$

Then, from Hardy-Littlewood-Polyá inequality

$$\sum_{i=1}^{n} x_i y_i = \max_{\sigma \in \mathcal{S}(1,\cdots,n)} \left\{ \sum_{i=1}^{n} x_i y_{\sigma(i)} \right\},$$

which can be interpreted as : correlation is maximal when vectors are simultaneously rearranged (i.e. comonotonic). And similarly,

$$\sum_{i=1}^{n} x_i y_{n+1-i} = \min_{\sigma \in \mathcal{S}(1,\cdots,n)} \left\{ \sum_{i=1}^{n} x_i y_{\sigma(i)} \right\},$$

The continuous version of that result is

Proposition28

Consider two positive random variables X and Y, then

$$\int_0^1 Q_X(1-u)Q_Y(u)du \leq \mathbb{E}[XY] \leq \int_0^1 Q_X(u)Q_Y(u)du$$

Corollary1

Let $Y \in L^\infty$ and $X \in L^1$ on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then

$$\max_{\tilde{Y}\sim Y}\{\mathbb{E}[X\tilde{Y}]\} = \mathbb{E}[Q_X(U)Q_Y(U)] = \int_0^1 Q_X(u)Q_Y(u)du$$

Proof. Observe that

$$\max_{\tilde{Y}\sim Y}\{\mathbb{E}[X\tilde{Y}]\} = \max_{\tilde{Y}\sim Y}\left\{\frac{1}{2}\left(-\mathbb{E}[X-\tilde{Y}]^2 + \mathbb{E}[X^2] + \mathbb{E}[\tilde{Y}^2]\right)\right\}$$

thus,

$$\max_{\tilde{Y}\sim Y}\{\mathbb{E}[X\tilde{Y}]\} = \underbrace{\frac{\mathbb{E}[X^2]+\mathbb{E}[\tilde{Y}^2]}{2}}_{=\text{constant}} - \frac{1}{2}\underbrace{\inf_{\tilde{Y}\sim Y}\{\mathbb{E}[X-\tilde{Y}]^2\}}_{\inf_{\tilde{Y}\sim Y}\{||X-\tilde{Y}||_{L^2}\}}$$

$\square$

More generally ([26]), for all convex risk measure, invariant in law,

$$\mathcal{R}(X+Y) \leq \mathcal{R}(Q_X(U)+Q_Y(U)) = \sup_{\tilde{X}\sim X,\tilde{Y}\sim Y}\{\mathcal{R}(\tilde{X}+\tilde{Y})\}$$

Definition17

A risk measure $\mathcal{R}$ is

SC strongly coherent if $\mathcal{R}(X+Y) = \sup_{\tilde{X}\sim X,\tilde{Y}\sim Y}\{\mathcal{R}(\tilde{X}+\tilde{Y})\}$

Proposition29

If a risk measure $\mathcal{R}$ satisfies [CO] and [SC] then$\mathcal{R}$ satisfies [PH].

Proposition30

Consider a risk measure $\mathcal{R}$ on L^p, with $p \in [1, \infty]$. Then the following statements are equivalent

- $\mathcal{R}$ is lower semi-continous and satisfies [CO] and [SC]
- $\mathcal{R}$ is lower semi-continous and satisfies [CO], [CI] and [LI]
- $\mathcal{R}$ is a measure of maximal correlation: let

$$\mathbb{Q} \in \mathcal{M}_1^q(\mathbb{P}) = \left\{\mathbb{Q} \in \mathcal{M}_1(\mathbb{P}) : \frac{d\mathbb{Q}}{d\mathbb{P}} \in L^q\right\}$$

then, for all X,

$$\mathcal{R}(X) = \mathcal{R}_{\mathbb{Q}}(X) = \sup_{Y \sim \frac{d\mathbb{Q}}{d\mathbb{P}}} \{\mathbb{E}[XY]\} = \int_0^1 Q_X(t) q_{\frac{d\mathbb{Q}}{d\mathbb{P}}}(t) dt.$$

Example11

ES_α is a $\mathcal{R}_\mathbb{Q}$-risk measure, with $\dfrac{d\mathbb{Q}}{d\mathbb{P}} \sim \mathcal{U}(1-\alpha, 1)$.

5.3 Distortion of probability measures

There is another interpretation of those maximal correlation risk measures, as expectation (in the Choquet sense) relative to distortion of probability measures.

Definition18

A function $\psi : [0,1] \to [0,1]$, nondecreasing and convex, such that $\psi(0) = 0$ and $\psi(1) = 1$ is called a distortion function.

Remark5

Previously, distortion were not necessarily convex, but in this section, we will only consider convex distortions.

Proposition31

If $\mathbb{P}$ is a probability measure, and ψ a distortion function, then $C : \mathcal{F} \rightarrow [0,1]$ defined as

$$\nu(A) = \psi \circ \mathbb{P}(A)$$

is a capacity, and the integral with respect to ν is

$$\mathbb{E}_\nu(X) = \int X d\nu = \int_{-\infty}^{0} [\psi \circ \mathbb{P}(X > x) - 1]dx + \int_{0}^{+\infty} \psi \circ \mathbb{P}(X > x)dx$$

The fundamental theorem is the following : maximal correlation risk measures can be written as Choquet integral with respect to some distortion of a probability measures.

Assume that X is non-negative, and let

$$\mathcal{R}_{\mathbb{Q}}(X) = \max\left\{\mathbb{E}(XY) \mid Y \sim \frac{d\mathbb{Q}}{d\mathbb{P}}\right\} = \int_0^1 Q_X(t) Q_{\frac{d\mathbb{Q}}{d\mathbb{P}}}(t)dt$$

but since

$$\psi'(1-t) = Q_{\frac{d\mathbb{Q}}{d\mathbb{P}}}(t)$$

we can write

$$\mathcal{R}_{\mathbb{Q}}(X) = \int_0^1 Q_X(t)\psi'(1-t)dt = \int_0^1 \psi(1-t)\ Q_{\frac{d\mathbb{Q}}{d\mathbb{P}}}(t)dt$$

by integration by parts, and then, with $t = F_X(u) = Q_X^{-1}(u)$,

$$\mathcal{R}_{\mathbb{Q}}(X) = \int_0^\infty \psi\left[1 - F_X(u)\right] du = \int_0^\infty \psi\left[\mathbb{P}(X > u)\right] du$$

which is Choquet's expectation with respect to capacity $\psi \circ \mathbb{P}$.

Thus,

$$\mathcal{R}_{\mathbb{Q}}(X) = \max\left\{\mathbb{E}(XY) \mid Y \sim \frac{d\mathbb{Q}}{d\mathbb{P}}\right\}$$

is a coherent risk measure, as a mixture of quantiles, it can be written using a set of scenarios $\mathcal{Q}$,

$$\mathcal{R}_{\mathbb{Q}}(X) = \max\left\{\mathbb{E}_{\tilde{\mathbb{Q}}}(X) \mid \tilde{\mathbb{Q}} \in \mathcal{Q} = \{\tilde{\mathbb{Q}} \in \mathcal{M}_1^q(\mathbb{P} : \mathcal{R}_{\mathbb{Q}}^\star(\tilde{\mathbb{Q}}) = 0)\}\right\}$$

where $\mathcal{R}_{\mathbb{Q}}^\star(\tilde{\mathbb{Q}}) = \sup_{X \in L^p}\left\{\mathbb{E}_{\tilde{\mathbb{Q}}}(X) - \mathcal{R}_{\mathbb{Q}}(X)\right\}$.

Observe that $\mathcal{R}_{\mathbb{Q}}^\star(\tilde{\mathbb{Q}}) = 0$ means that, for all $X \in L^p$, $\mathbb{E}_{\tilde{\mathbb{Q}}}(X) \leq \mathcal{R}_{\mathbb{Q}}(X)$, i.e., for all A, $\psi \circ \mathbb{P}(A) \geq \tilde{\mathbb{Q}}(A)$. Thus,

$$\mathcal{R}_{\mathbb{Q}}(X) = \max\left\{\mathbb{E}_{\tilde{\mathbb{Q}}}(X) \mid \tilde{\mathbb{Q}} \leq \psi \circ \mathbb{P}\right\}$$

where ψ is the distortion associated with $\mathbb{Q}$, in the sense that

$$\psi'(1-t) = Q_{\frac{d\mathbb{Q}}{d\mathbb{P}}}(t)$$

Example12

Let $X \in L^p$, then we defined

$$\mathcal{R}_{\mathbb{Q}}(X) = \sup\left\{\mathbb{E}(X\cdot Y) \mid Y \sim \frac{d\mathbb{Q}}{d\mathbb{P}}\right\}$$

In the case where

$$X \sim \mathcal{N}(0,\sigma_x^2) \text{ and } \frac{d\mathbb{Q}}{d\mathbb{P}} \sim \mathcal{N}(0,\sigma_u^2)$$

then

$$\mathcal{R}_{\mathbb{Q}}(X) = \sigma_x \cdot \sigma_u.$$

From Optimal Transport results, one can prove that the optimal coupling

$$\sup_{\tilde{Y}\sim Y}\{\mathbb{E}(X\tilde{Y})\}$$

is given by $\mathbb{E}(\nabla f(Y)Y)$, where f is some convex function. In dimension 1, the quantile function Q_X (which yields the optimal coupling) is increasing, but in higher dimension, what should appear is the gradient of some comvex function.

5.4 Optimal Transport and Risk Measures

Definition19

A map $T : \mathcal{G} \to \mathcal{H}$ is said to be a transport map between measures μ and ν if

$$\nu(B) = \mu(T^{-1}(B)) = T_{\#}\mu(B) \text{ for every } B \subset \mathcal{H}.$$

Thus

$$\int_{\mathcal{E}} \varphi[T(x)]d\mu(x) = \int_{\mathcal{E}} \varphi[y]d\nu(y) \text{ for all } \phi \in \mathcal{C}(\mathcal{H}).$$

Definition20

A map $T : \mathcal{G} \to \mathcal{H}$ is said to be an optimal transport map between measures μ and ν, for some cost function $c(\cdot,\cdot)$ if

$$T \in \underset{T, T_{\#}\mu=\nu}{\text{argmin}} \left\{ \int_{\mathcal{G}} c(x, T(x))d\mu(x) \right\}$$

The reformulation of is the following. Consider the Fréchet space $\mathcal{F}(\mu,\nu)$.

Definition21

A transport plan between measures μ and ν if a probability measure in $\mathcal{F}(\mu,\nu)$.

Definition22

A transport plan between measures μ and ν if said to be optimal if

$$\gamma \in \underset{\gamma\in\mathcal{F}(\mu,\nu)}{\text{argmin}} \left\{\int_{\mathcal{G}\times\mathcal{H}} c(x,y)d\gamma(x,y)\right\}$$

Consider two measures on $\mathbb{R}$, and define for all $x\in\mathbb{R}$

$$T(x)=\inf_{t\in\mathbb{R}}\{\nu((-\infty,t])>\mu((-\infty,x])\}$$

T is the only monotone map such that $T_{\#}\mu=\nu$.

6 Multivariate Risk Measures

6.1 Which Dimension?

In this section, we consider some $\mathbb{R}^d$ random vector $\boldsymbol{X}$. What could be the risk of that random vector? Should it be a single amount, i.e. $\mathcal{R}(\boldsymbol{X}) \in \mathbb{R}$ or a d-dimensional one $\mathcal{R}(\boldsymbol{X}) \in \mathbb{R}^d$?

6.2 Multivariate Comonotonicity

In dimension 1, two risks X_1 and X_2 are comonotonic if there is Z and two increasing functions g_1 and g_2 such that

$$X_1 = g_1(Z) \text{ and } X_2 = g_2(Z)$$

Observe that

$$\mathbb{E}(X_1 Z) = \max_{\tilde{X}_1 \sim X_1} \{\mathbb{E}(\tilde{X}_1 Z)\} \text{ and } \mathbb{E}(X_2 Z) = \max_{\tilde{X}_2 \sim X_2} \{\mathbb{E}(\tilde{X}_2 Z)\}.$$

For the higher dimension extension, recall that $\mathbb{E}(\boldsymbol{X}\cdot\boldsymbol{Y})=\mathbb{E}(\boldsymbol{X}\boldsymbol{Y}^{\mathsf{T}})$

Definition23

$\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are said to be comonotonic, with respect to some distribution μ if there is $\boldsymbol{Z}\sim\mu$ such that both $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are in optimal coupling with Z, i.e.

$$\mathbb{E}(\boldsymbol{X}_1\cdot\boldsymbol{Z})=\max_{\tilde{\boldsymbol{X}}_1\sim\boldsymbol{X}_1}\{\mathbb{E}(\tilde{\boldsymbol{X}}_1\cdot Z)\}\text{ and }\mathbb{E}(\boldsymbol{X}_2\cdot\boldsymbol{Z})=\max_{\tilde{\boldsymbol{X}}_2\sim\boldsymbol{X}_2}\{\mathbb{E}(\tilde{\boldsymbol{X}}_2\cdot\boldsymbol{Z})\}.$$

Observe that, in that case

$$\mathbb{E}(\boldsymbol{X}_1\cdot\boldsymbol{Z})=\mathbb{E}(\nabla f_1(\boldsymbol{Z})\cdot\boldsymbol{Z})\text{ and }\mathbb{E}(\boldsymbol{X}_2\cdot\boldsymbol{Z})=\mathbb{E}(\nabla f_2(\boldsymbol{Z})\cdot\boldsymbol{Z})$$

for some convex functions f_1 and f_2. Those functions are called Kantorovitch potentials of $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, with respect to μ.

Definition24

The μ-quantile function of random vector $\boldsymbol{X}$ on $\mathcal{X}=\mathbb{R}^d$, with respect to distribution μ is $Q_{\boldsymbol{X}}=\nabla f$, where f is Kantorovitch potential of $\boldsymbol{X}$ with respect to μ, in the sense that

$$\mathbb{E}(\boldsymbol{X}\cdot\boldsymbol{Z})=\max_{\tilde{\boldsymbol{X}}_1\sim\boldsymbol{X}_1}\{\mathbb{E}(\tilde{\boldsymbol{X}}_1\cdot\boldsymbol{Z})\}=\mathbb{E}(\nabla f(\boldsymbol{Z})\cdot\boldsymbol{Z})$$

Example13

Consider two random vectors, $\boldsymbol{X} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_X)$ and $\boldsymbol{Y} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_Y)$, as in [9]. Assume that our baseline risk is Gaussian. More specifically, μ has a $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_U)$ distribution. Then $\boldsymbol{X}$ and $\boldsymbol{Y}$ are μ-comonotonic if and only if

$$\mathbb{E}(\boldsymbol{X} \cdot \boldsymbol{Y}) = \boldsymbol{\Sigma}_U^{-1/2}[\boldsymbol{\Sigma}_U^{1/2}\boldsymbol{\Sigma}_X\boldsymbol{\Sigma}_U^{1/2}]^{1/2}[\boldsymbol{\Sigma}_U^{1/2}\boldsymbol{\Sigma}_Y\boldsymbol{\Sigma}_U^{1/2}]^{1/2}\boldsymbol{\Sigma}_U^{-1/2}.$$

To prove this result, because variables are multivariate Gaussian vectors, $\boldsymbol{X}$ and $\boldsymbol{Y}$ are μ-comonotonic if and only if there is $\boldsymbol{U} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_U)$, and two matrices A_X and A_Y such that $\boldsymbol{X} = A_X\boldsymbol{U}$ and $\boldsymbol{Y} = A_Y\boldsymbol{U}$. [30] proves that mapping $\boldsymbol{u} \mapsto A\boldsymbol{u}$ with

$$A = \boldsymbol{\Sigma}_U^{-1/2}[\boldsymbol{\Sigma}_U^{1/2}\boldsymbol{\Sigma}_X\boldsymbol{\Sigma}_U^{1/2}]^{1/2}\boldsymbol{\Sigma}_U^{-1/2}$$

will tranform probability measure $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_U)$ to probability measure $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$.

Conversely, define

$$\boldsymbol{U} = A_X^{-1}\boldsymbol{X} \text{ and } \boldsymbol{U} = A_Y^{-1}\boldsymbol{Y}$$

Clearly, $\boldsymbol{U} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_U)$, as well as $\boldsymbol{V}$. Observe further that

$$\mathbb{E}(\boldsymbol{U}\cdot\boldsymbol{V}) = A_X^{-1}\mathbb{E}(\boldsymbol{X}\cdot\boldsymbol{Y})A_Y^{-1} = \boldsymbol{\Sigma}_U = \boldsymbol{\Sigma}_U^{1/2}\boldsymbol{\Sigma}_U^{1/2}$$

so by Cauchy-Scharz, $\boldsymbol{U} = \boldsymbol{V}$, a.s. So $\boldsymbol{X}$ and $\boldsymbol{Y}$ are μ-comonotonic.

In the case where μ has a $\mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$ distribution, $\boldsymbol{X}$ and $\boldsymbol{Y}$ are μ-comonotonic if and only if

$$\mathbb{E}(\boldsymbol{X}\cdot\boldsymbol{Y}) = \boldsymbol{\Sigma}_X^{1/2}\boldsymbol{\Sigma}_Y^{1/2}$$

But this is not the only was to define multivariate comontonicity.

6.3 π-Comonotonicity

Following [29], inspired by multivariate rearrangement introduced in [36] or [25], one can befine π-comonotonicity,

Definition25

$\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are said to be π-comonotonic, if there is Z and some increasing functions $g_{1,1}, \cdots, g_{1,d}, g_{2,1}, \cdots, g_{2,d}$, such that

$$(\boldsymbol{X}_1, \boldsymbol{X}_2) = ([g_{1,1}(X_{1,1}), \cdots, g_{1,d}(X_{1,d})], [g_{2,1}(X_{2,1}), \cdots, g_{2,d}(X_{2,d})])$$

6.4 Properties of Multivariate Risk Measures

More generally, let $\mathcal{R}$ denote a multivariate risk measure.

Definition26

A multivariate risk measure is a mapping $L^{p,d}(\Omega, \mathcal{F}, \mathbb{P}) \to \overline{\mathbb{R}}$.

Definition27

A multivariate risk measure $\mathcal{R}$ can be

PO positive, $\boldsymbol{X} \geq \mathbf{0}$ implies $\mathcal{R}(\boldsymbol{X}) \leq 0$

MO monotone, $\boldsymbol{X} \geq \boldsymbol{Y}$ implies $\mathcal{R}(\boldsymbol{X}) \leq \mathcal{R}(\boldsymbol{Y})$.

PH (positively) homogenous, $\lambda \geq 0$ implies $\mathcal{R}(\lambda \boldsymbol{X}) = \lambda \mathcal{R}(\boldsymbol{X})$.

TI invariant by translation, $k \in \mathbb{R}$ implies $\mathcal{R}(\boldsymbol{X} + k\mathbf{1}) = \mathcal{R}(\boldsymbol{X}) - k$,

IL invariant in law, $\boldsymbol{X} \sim \boldsymbol{Y}$ implies $\mathcal{R}(\boldsymbol{X}) = \mathcal{R}(Y)$.

CO convex, $\forall \lambda \in [0,1]$, $\mathcal{R}\lambda \boldsymbol{X} + (1-\lambda)\boldsymbol{Y}) \leq \lambda \mathcal{R}(\boldsymbol{X}) + (1-\lambda)\mathcal{R}(\boldsymbol{Y})$.

SA subadditive, $\mathcal{R}(\boldsymbol{X} + \boldsymbol{Y}) \leq \mathcal{R}(\boldsymbol{X}) + \mathcal{R}(\boldsymbol{Y})$.

One should keep in mind that [LI] means that

$$\mathcal{R}(\boldsymbol{X}) = \sup_{\tilde{\boldsymbol{X}} \sim \boldsymbol{X}} \{\mathcal{R}(\tilde{\boldsymbol{X}})\}.$$

If $\mathcal{R}$ is a convex lower semi-continuous risk measure, then

$$\mathcal{R}(\boldsymbol{X}) = \sup_{\boldsymbol{Y}\in\mathbb{Q}} \{\mathbb{E}(\boldsymbol{X}\cdot\boldsymbol{Y}) - \mathcal{R}^\star(\boldsymbol{Y})\}$$

where $\mathcal{R}^\star$ is the Fenchel transform of $\mathcal{R}$, for some set $\mathbb{Q}$.

Definition28

A multivariate risk measure $\mathcal{R}$ on $L^{p,d}$ is

SC strongly coherent if $\mathcal{R}(\boldsymbol{X}+\boldsymbol{Y}) = \sup_{\tilde{\boldsymbol{X}}\sim\boldsymbol{X},\tilde{\boldsymbol{Y}}\sim\boldsymbol{Y}} \{\mathcal{R}(\tilde{\boldsymbol{X}}+\tilde{\boldsymbol{Y}})\}$

MC a maximal correlation measure if

$$\mathcal{R}(\boldsymbol{X}) = \sup_{\boldsymbol{Y}\in\mathcal{Y}\subset L^{q,d}} \{\mathbb{E}(\boldsymbol{X}\cdot\boldsymbol{Y})\}$$

for some $\mathcal{Y}$.

6.5 μ-Comonotonicity and Strong Coherence

Even if the extention is not unique, the concept of μ-comonotonicity seems to be the natural extension of what we obtained in the univariate case,

Proposition32

Let $\mathcal{R}$ denote a multivariate convex risk measure on $L^{p,d}$, the following statements are equivalent

- $\mathcal{R}$ is strongly coherent
- $\mathcal{R}$ is μ-comonotone additive (for some μ) and invariant in law
- $\mathcal{R}$ is a maximal correlation measure

Proof. As mentioned in the previous section, since $\mathcal{R}$ is a convex lower semi-continuous risk measure, then

$$\mathcal{R}(\boldsymbol{X}) = \sup_{\boldsymbol{Y}\in\mathcal{Y}} \{\mathbb{E}(\boldsymbol{X}\cdot\boldsymbol{Y}) - \mathcal{R}^\star(\boldsymbol{Y})\}$$

where $\mathcal{R}^\star$ is the Fenchel transform of $\mathcal{R}$, for some set $\mathcal{Y}$.

Let us prove that [SC] implies [MC]. If $\mathcal{R}$ satisfies [SC], then it satisfies [LI], and

$$\mathcal{R}(\boldsymbol{X}) = \sup_{\tilde{\boldsymbol{X}}\sim\boldsymbol{X}} \{\mathcal{R}(\boldsymbol{X})\} = \sup_{\boldsymbol{Y}\in\mathcal{Y}} \left\{ \sup_{\tilde{\boldsymbol{X}}\sim\boldsymbol{X}} \{\mathbb{E}(\tilde{\boldsymbol{X}}\cdot\boldsymbol{Y}) - \mathcal{R}^\star(\boldsymbol{Y})\} \right\}$$

Observe that the penalty function $\mathcal{R}^\star$ satisfies [LI] since

$$\begin{aligned}
\mathcal{R}^\star(\boldsymbol{Y}) &= \sup_{\boldsymbol{X}\in L^{p,d}} \{\mathbb{E}(\boldsymbol{X}\cdot\boldsymbol{Y}) - \mathcal{R}(\boldsymbol{X})\} \\
&= \sup_{\boldsymbol{X}\in L^{p,d}} \left\{ \sup_{\tilde{\boldsymbol{X}}\sim\boldsymbol{X}} \{\mathbb{E}(\tilde{\boldsymbol{X}}\cdot\boldsymbol{Y}) - \mathcal{R}(\tilde{\boldsymbol{X}}) \right\} \\
&= \sup_{\boldsymbol{X}\in L^{p,d}} \left\{ \underbrace{\sup_{\tilde{\boldsymbol{X}}\sim\boldsymbol{X}} \{\mathbb{E}(\tilde{\boldsymbol{X}}\cdot\boldsymbol{Y})}_{\text{maximal correlation}} - \mathcal{R}(\boldsymbol{X}) \right\}
\end{aligned}$$

(the maximal correlation satisfies [LI]).

Observe that

$$\mathcal{R}(\boldsymbol{X}) = \sup_{\boldsymbol{Y}\in\mathcal{Y}}\{\mathbb{E}(\boldsymbol{X}\cdot\boldsymbol{Y}) - \mathcal{R}^\star(\boldsymbol{Y})\}$$

can be writen

$$\mathcal{R}(\boldsymbol{X}) = \sup_{\mathbb{Q}\in\mathcal{Q}}\{\mathcal{R}_{\mathbb{Q}}(\boldsymbol{X}) - \mathcal{R}^\star(\boldsymbol{Y})\} \text{ where } \mathcal{Q} = \left\{\mathbb{Q}\left|\frac{d\mathbb{Q}}{d\mathbb{P}} \in \mathcal{Y}\right.\right\}$$

Recall that in the univariate case,

$$\mathcal{R}_{\mathbb{Q}}(X) = \int_0^1 Q_X(t)Q_{\frac{d\mathbb{Q}}{d\mathbb{P}}}(t)dt$$

Conversely, let us prove that [MC] implies [SC]. Consider here X and $\boldsymbol{Y}$ that are μ-comononotone, i.e. there is $\boldsymbol{Z}\sim\mu$ such that

$$\mathbb{E}(\boldsymbol{X}\cdot\boldsymbol{Z}) = \sup_{\tilde{\boldsymbol{Z}}\sim\boldsymbol{Z}}\{\mathbb{E}(\boldsymbol{X}\cdot\tilde{\boldsymbol{Z}})\}$$

and

$$\mathbb{E}(\boldsymbol{Y}\cdot\boldsymbol{Z}) = \sup_{\tilde{\boldsymbol{Z}}\sim\boldsymbol{Z}}\{\mathbb{E}(\boldsymbol{Y}\cdot\tilde{\boldsymbol{Z}})\}$$

As discussed previously, it means that there are convex functions f_X and f_Y such that $\boldsymbol{X} = \nabla f_X(\boldsymbol{Z})$ and $\boldsymbol{Y} = \nabla f_Y(\boldsymbol{Z})$ (a.s.). So $\boldsymbol{X}+\boldsymbol{Y} = \nabla(f_X + f_Y)(\boldsymbol{Z})$, $f_X + f_Y$ being a convexe function lower semi-continuous. So $\boldsymbol{X}+\boldsymbol{Y}$ is comonotonic with both $\boldsymbol{X}$ and $\boldsymbol{Y}$. Thus, we can write

$$\mathbb{E}[(\boldsymbol{X}+\boldsymbol{Y})\cdot\boldsymbol{Z}] = \sup_{\tilde{\boldsymbol{Z}}\sim\boldsymbol{Z}}\{\mathbb{E}[(\boldsymbol{X}+\boldsymbol{Y})\cdot\tilde{\boldsymbol{Z}}]\} = \mathcal{R}(\boldsymbol{X}+\boldsymbol{Y})$$

and

$$\mathbb{E}[(\boldsymbol{X}+\boldsymbol{Y})\cdot\boldsymbol{Z}] = \sup_{\tilde{\boldsymbol{Z}}\sim\boldsymbol{Z}}\{\mathbb{E}[\boldsymbol{X}\cdot\tilde{\boldsymbol{Z}}]\} + \sup_{\tilde{\boldsymbol{Z}}\sim\boldsymbol{Z}}\{\mathbb{E}[\boldsymbol{Y}\cdot\tilde{\boldsymbol{Z}}]\} = \mathcal{R}(\boldsymbol{X})+\mathcal{R}(\boldsymbol{Y})$$

which means that $\mathcal{R}$ satisfies [SC]. $\square$

Example14

Example 12 can be extended in higher dimension.

$$\mathcal{R}_{\mathbb{Q}}(\boldsymbol{X}) = \sup\left\{\mathbb{E}(\boldsymbol{X}\cdot\boldsymbol{Y})\mid \boldsymbol{Y}\sim\frac{d\mathbb{Q}}{d\mathbb{P}}\right\}$$

with

$$\boldsymbol{X}\sim\mathcal{N}(\mathbf{0},\boldsymbol{\Sigma}_x)\text{ and }\frac{d\mathbb{Q}}{d\mathbb{P}}\sim\mathcal{N}(\mathbf{0},\boldsymbol{\Sigma}_u).$$

In that case

$$\mathcal{R}_{\mathbb{Q}}(\boldsymbol{X}) = \mathit{trace}\left([\boldsymbol{\Sigma}_u^{1/2}\boldsymbol{\Sigma}_x\boldsymbol{\Sigma}_u^{1/2}]^{1/2}\right)$$

For instance, if

$$\boldsymbol{X}\sim\mathcal{N}\left(\mathbf{0},\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}\right)\text{ and }\frac{d\mathbb{Q}}{d\mathbb{P}}\sim\mathcal{N}(\mathbf{0},\mathbb{I}).$$

then

$$\mathcal{R}_{\mathbb{Q}}(\boldsymbol{X}) = \sqrt{\sigma_1^2+\sigma_2^2+2\sigma_1\sigma_2\sqrt{1-\rho^2}}.$$

Example15

In example 14 we were solving

$$\mathcal{R}_{\mathbb{Q}}(\boldsymbol{X}) = \sup\left\{\mathbb{E}(\boldsymbol{X}\cdot\tilde{\boldsymbol{Y}}) \mid \tilde{\boldsymbol{Y}} \sim \boldsymbol{Y}\right\}$$

with $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_x)$ and $\boldsymbol{Y} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_u)$, which mean minimizing transportation cost, with a quadratic cost function. The general solution is

$$\mathbb{E}(\nabla f_{\boldsymbol{X}}(\boldsymbol{Y})\cdot\boldsymbol{Y})$$

Thus, here

$$\nabla f_{\boldsymbol{X}}(\boldsymbol{Y}) = \boldsymbol{\Sigma}_u^{-1/2}\left[\boldsymbol{\Sigma}_u^{1/2}\boldsymbol{\Sigma}_x\boldsymbol{\Sigma}_u^{1/2}]^{1/2}\right]\boldsymbol{\Sigma}_u^{-1/2}$$

6.6 Examples of Multivariate Risk Measures

In the univariate case, the expected shortfall $ES_X(\alpha)$ is the maximal correlation measure associated with a baseline risk $U \sim \mathcal{B}(1-\alpha, 1)$.

In the multivariate case, one can define $ES_{\boldsymbol{X}}(\alpha)$ as the maximal correlation measure associated with a baseline risk $\boldsymbol{U} \sim \mathcal{B}(1-\alpha, 1)$. More specifically, $\mathbb{P}(\boldsymbol{U} = \boldsymbol{0}) = \alpha$ while $\mathbb{P}(\boldsymbol{U} = \boldsymbol{1}) = 1-\alpha$. Define

$$f(\boldsymbol{x}) = \max_{c, \mathbb{P}(\boldsymbol{X}^\top \boldsymbol{1} \geq c) = \alpha} \{\boldsymbol{x}^\top \boldsymbol{1} - c\},$$

then f is a convex function, ∇f exists and pushes from the distribution of $\boldsymbol{X}$ to the distribution of $\boldsymbol{U}$. Thus, the maximal correlation is here

$$\mathbb{E}\left(\boldsymbol{X}^\top \boldsymbol{1} \cdot \boldsymbol{1}_{\{\boldsymbol{X}^\top \boldsymbol{1} \geq c\}}\right).$$

Actually, the maximal correlation risk measure is the univariate expected shortfall of the sum,

$$ES_{\boldsymbol{X}}(\alpha) = ES_{\boldsymbol{X}^\top \boldsymbol{1}}(\alpha)$$

7 Dynamic Risk Measures

As we will see in this section, dynamic risk measures should - somehow - be consistent over time: what is preferred at time t should be consistent with what is preferred at another time $s \neq t$). A strong time consistency concept will be related to the dynamic programming principle. In continuous time, such risk measure will be obtained as solutions of some backward stochastic differential equation.

Dynamic risk measures, in discrete or continuous time, will simply denote sequences of *conditional* risk measures, adapted to the underlying filtration. So the first step will be to characterize those conditional risk measures.

7.1 Conditional Risk Measures

Let $\mathcal{G} \subset \mathcal{F}$ denote a sub-σ-algebra.

Definition29

A conditional risk measure can satisfy

$\mathcal{G}$-TI For any $X \in L^\infty$ and $K \in Ł^\infty$ -G-measurable, $\mathcal{R}(X+K) = \mathcal{R} - K$.

$\mathcal{G}$-CV For any $X, Y \in L^\infty$ and $\Lambda \in Ł^\infty$ -G-measurable, with $\Lambda \in [0,1]$,

$$\mathcal{R}(\Lambda X + (1-\Lambda)Y) = \Lambda\mathcal{R}(X) + (1-\Lambda)\mathcal{R}(Y).$$

$\mathcal{G}$-PH For any $X \in L^\infty$ and $\Lambda \in Ł^\infty$ -G-measurable, with $\Lambda \geq 0$, $\mathcal{R}(\Lambda X) = \Lambda\mathcal{R}(X)$.

Definition30

$\mathcal{R}$ is a $\mathcal{G}$-conditional convex risk measure if it satisfies [MO], $\mathcal{G}$-conditional [TI] and [CV], and $\mathcal{R}(0) = 0$. $\mathcal{R}$ is a $\mathcal{G}$-conditional coherent risk measure if it is a $\mathcal{G}$-conditional convex risk measure that satisfies $\mathcal{G}$-conditional [PH].

A risk measure is said to be representable if

$$\mathcal{R}(X) = \operatorname*{esssup}_{\mathbb{Q}\in\mathcal{P}_{\mathcal{G}}} \{-\mathbb{E}_{\mathbb{Q}}(X|\mathcal{G}) - \alpha(\mathbb{Q})\}$$

where α is a random penalty function, associated to $\mathcal{R}$.

If $\mathcal{R}$ is $\mathcal{G}$-conditional convex risk measure, it can be represented using

$$\alpha(\mathbb{Q}) = \operatorname*{esssup}_{X\in L^\infty} \{-\mathbb{E}_{\mathbb{Q}}(X \in |\mathcal{G}) - \mathcal{R}(X)\}$$

If $\mathcal{R} is a \text{G}-conditional coherent risk measure, it can be represented as \mathcal{R}(X) = \operatorname*{esssup}_{\mathbb{Q}\in\mathcal{Q}_{\mathcal{G}}} \{-\mathbb{E}_{\mathbb{Q}}(X|\mathcal{G})\} where \mathcal{Q}_{\mathcal{G}} = \{\mathbb{Q} \in \mathcal{P}_{\mathcal{G}} | \mathbb{E}_{\mathbb{Q}}(X|\mathcal{G}) \geq -\mathcal{R}(X)$ for all $X \in L^\infty\}$.

7.2 On which set(s) of measures will we work with?

In the static setting, we considered random variables defined on probability space $(\Omega, \mathcal{F}, \mathbb{P})$. From now on, we will consider adapted stochastic processes $\boldsymbol{X} = (X_t)$ on the filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$

The $||\cdot||_\infty$ norm on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ is defined as

$$||\boldsymbol{X}||_\infty = \inf\{m \in \mathbb{R} | \sup_t\{|X_t|\} < m\}.$$

Let $\boldsymbol{L}^\infty$ denote the set of all bounded adapted stochastic processes, in the sense that

$$\boldsymbol{L}^\infty = \{\boldsymbol{X} | ||\boldsymbol{X}||_\infty < \infty\}.$$

We now need to extend the form $< X, s >= \mathbb{E}(Xs)$ defined on $L^\infty \times L^1$ on the set of stochastic processes. Set

$$< \boldsymbol{X}, \boldsymbol{s} >= \mathbb{E}\left(\sum_{t\in\mathbb{N}} X_t \Delta a_t\right) = X_0 a_0 + X_1(a_1 - a_0) + X_2(a_2 - a_1) + \cdots$$

This will be used when considering risk evaluation at time 0, but it might be interesting to evaluate risk at some time τ (which can be deterministic, or some stoping time). In that case, define

$$< \boldsymbol{X}, \boldsymbol{s} >_\tau = \mathbb{E}\left(\sum_{t=\tau}^{\infty} X_t \Delta a_t \middle| \mathcal{F}_\tau\right)$$

It is then possible to define

$$\boldsymbol{L}_\tau^\infty = \{\boldsymbol{X} = (0, 0, \cdots, 0, X_\tau, X_{\tau+1}, \cdots) | || \boldsymbol{X} ||_\infty < \infty\}.$$

7.3 Dynamic Risk Measure

Definition31

A dynamic monetary risk measure is a sequence of mappings $\boldsymbol{\mathcal{R}}_\tau = (\mathcal{R}_t)_{t\geq\tau}$, on $\boldsymbol{L}^\infty_\tau$ is a conditional monetary risk measure if

MO If $\boldsymbol{X} \leq \boldsymbol{Y}$, then $\boldsymbol{\mathcal{R}}_\tau(\boldsymbol{X}) \geq \boldsymbol{\mathcal{R}}_\tau(\boldsymbol{Y})$

$\mathcal{F}_\tau$-RG If $A \in \mathcal{F}_\tau$ then $\boldsymbol{\mathcal{R}}_\tau(\mathbf{1}_A\boldsymbol{X}) = \mathbf{1}_A \cdot \boldsymbol{\mathcal{R}}_\tau(\boldsymbol{X})$ (regularity condition)

$\mathcal{F}_\tau$-TI If $K \in L^\infty$ is $\mathcal{F}_\tau$ measurable, then $\boldsymbol{\mathcal{R}}_\tau(\boldsymbol{X}+K) = \boldsymbol{\mathcal{R}}_\tau(\boldsymbol{X}) - K$

Observe that [$\mathcal{F}_\tau$-RG] is actually equivalent to $\boldsymbol{\mathcal{R}}_\tau(\mathbf{0}) = 0$. This condition is weaker than the [$\mathcal{F}_\tau$-PH] property.

Definition32

A dynamic monetary risk measure is a sequence of mappings $\boldsymbol{\mathcal{R}}_\tau = (\mathcal{R}_t)_{t\geq\tau}$, on $\boldsymbol{L}_\tau^\infty$ is a dynmaic convex risk measure if it satisfies [MO], [$\mathcal{F}_\tau$-RG], [$\mathcal{F}_\tau$-TI] and

$\mathcal{F}_\tau$-CV If $\Lambda \in [0,1]$ is $\mathcal{F}_\tau$ measurable, then

$$\boldsymbol{\mathcal{R}}_\tau(\Lambda\boldsymbol{X} + (1-\Lambda)\boldsymbol{X}) \leq \Lambda\boldsymbol{\mathcal{R}}_\tau(\boldsymbol{X}) + (1-\Lambda)\boldsymbol{\mathcal{R}}_\tau(\boldsymbol{Y})$$

Definition33

A dynamic monetary risk measure is a sequence of mappings $\boldsymbol{\mathcal{R}}_\tau = (\mathcal{R}_t)_{t\geq\tau}$, on $\boldsymbol{L}_\tau^\infty$ is a dynamic coherent risk measure if it satisfies [MO], [$\mathcal{F}_\tau$-RG], [$\mathcal{F}_\tau$-TI], [$\mathcal{F}_\tau$-CV] and

$\mathcal{F}_\tau$-PH If $\Lambda \in L\infty$ is positive, and $\mathcal{F}_\tau$-measurable, then $\boldsymbol{\mathcal{R}}_\tau(\Lambda\cdot\boldsymbol{X}) = \Lambda\cdot\boldsymbol{\mathcal{R}}_\tau(\boldsymbol{X})$

Finally, from a dynamic risk measure, it is possible to extend the concept of acceptence set.

Definition34

Given a dynamic monetary risk measure $\boldsymbol{\mathcal{R}}_\tau = (\mathcal{R}_t)_{t\geq\tau}$, on $\boldsymbol{L}_\tau^\infty$. An $(\mathcal{F}_t)$-adapted stochastic process $\boldsymbol{X}$ is considered acceptable if

$$X \in \mathcal{A}_{\boldsymbol{\mathcal{R}}_\tau} \text{ with } \mathcal{A}_{\boldsymbol{\mathcal{R}}_\tau} = \{\boldsymbol{X}|\mathcal{R}_\tau(\boldsymbol{X}) \leq 0\}$$

Based on those definition, it is possible to get a representation theorem for dynamic convex risk measures, following [4] and [11]. Define

$$\mathcal{Q}_\tau = \{\boldsymbol{Z} - (\mathcal{F}_t) - \text{adapted} \mid < \boldsymbol{1}, \boldsymbol{Z} >_\tau = 1\}$$

called set of $(\mathcal{F}_t)$-adapted density processes.

Proposition33

A dynamic convex risk measure $\boldsymbol{\mathcal{R}}_\tau$ that is continuous from above (its acceptence set $\mathcal{A}\boldsymbol{\mathcal{R}}_\tau$ is closed) can be representes as follows,

$$\mathcal{R}_\tau(\boldsymbol{X}) = \sup_{\boldsymbol{Z}\in\mathcal{Q}_\tau} \{< \boldsymbol{X}, \boldsymbol{Z} >_\tau - \alpha_{\min,\tau}(\boldsymbol{Z})\}$$

where the minimal penalty is defined as

$$\alpha_{\min,\tau}(\boldsymbol{Z}) = \sup_{\boldsymbol{Y}\in\mathcal{A}_{\boldsymbol{\mathcal{R}}_\tau}} \{< \boldsymbol{Y}, \boldsymbol{Z} >_\tau\}$$

7.4 On time consistency

In order to get a better understanding of what time consistency could mean, consider the following example.

Example16

Assume that, at time t,

$$\mathcal{R}_t(X) = \underset{\mathbb{Q}\in\mathcal{Q}}{\text{esssup}}\{\mathbb{E}_{\mathbb{Q}}(-X|\mathcal{F}_{\sqcup})\}$$

where $\mathcal{Q}$ is a class of probability measures, i.e. $\mathcal{Q} \subset \mathcal{M}_1(\mathbb{P})$. Here, a worst case scenario is considered, in the sense that if $Z_{\{}\mathbb{Q}$ is the random variable $\mathbb{E}_{\mathbb{Q}}(-X|\mathcal{F}_{\sqcup})$, then the essential supremum is the smallest random variable Z such that $\mathbb{P}(Z \geq Z_{\{}\mathbb{Q}) = 1$ for all $\mathbb{Q} \in \mathcal{Q}$.

We have a two period binomial tree - see Figure 4. It is a simple Heads & Tail game. After 2 Heads or 2 Tails, the payoff is $+4$, while it is -5 with 1 Head and 1 Tail. There are two probabilities, considered by the agent, $\mathcal{Q} = \{\mathbb{Q}_1, \mathbb{Q}_2\}$. Observe that

$$\mathbb{E}_{\mathbb{Q}_i}(\star) = 0 \text{ for } i = 1, 2.$$

There is no *worst case*, for all probabilities in $\mathcal{Q}$, the expected payoff is the same. So, the agent should accept the risk at time 0. Assume that at time 1 the agent wants to re-evaluate the riskiness of the game. The strategy will be to consider conditional probabilities.

If we went up from time 0 to time 1, then

- under $\mathbb{Q}_1$: $\mathbb{E}_{\mathbb{Q}_1}(-\star) = -1$
- under $\mathbb{Q}_2$: $\mathbb{E}_{\mathbb{Q}_1}(-\star) = +2$

and if we went down from time 0 to time 1, then

- under $\mathbb{Q}_1$: $\mathbb{E}_{\mathbb{Q}_1}(-\star) = +2$
- under $\mathbb{Q}_2$: $\mathbb{E}_{\mathbb{Q}_1}(-\star) = -1$

So, the worst case scenario is that the risk is $+2$. Hence, for both knots - i.e. whatever happened at time 1 - the agent should reject the risk.

This is somehow inconsistent.

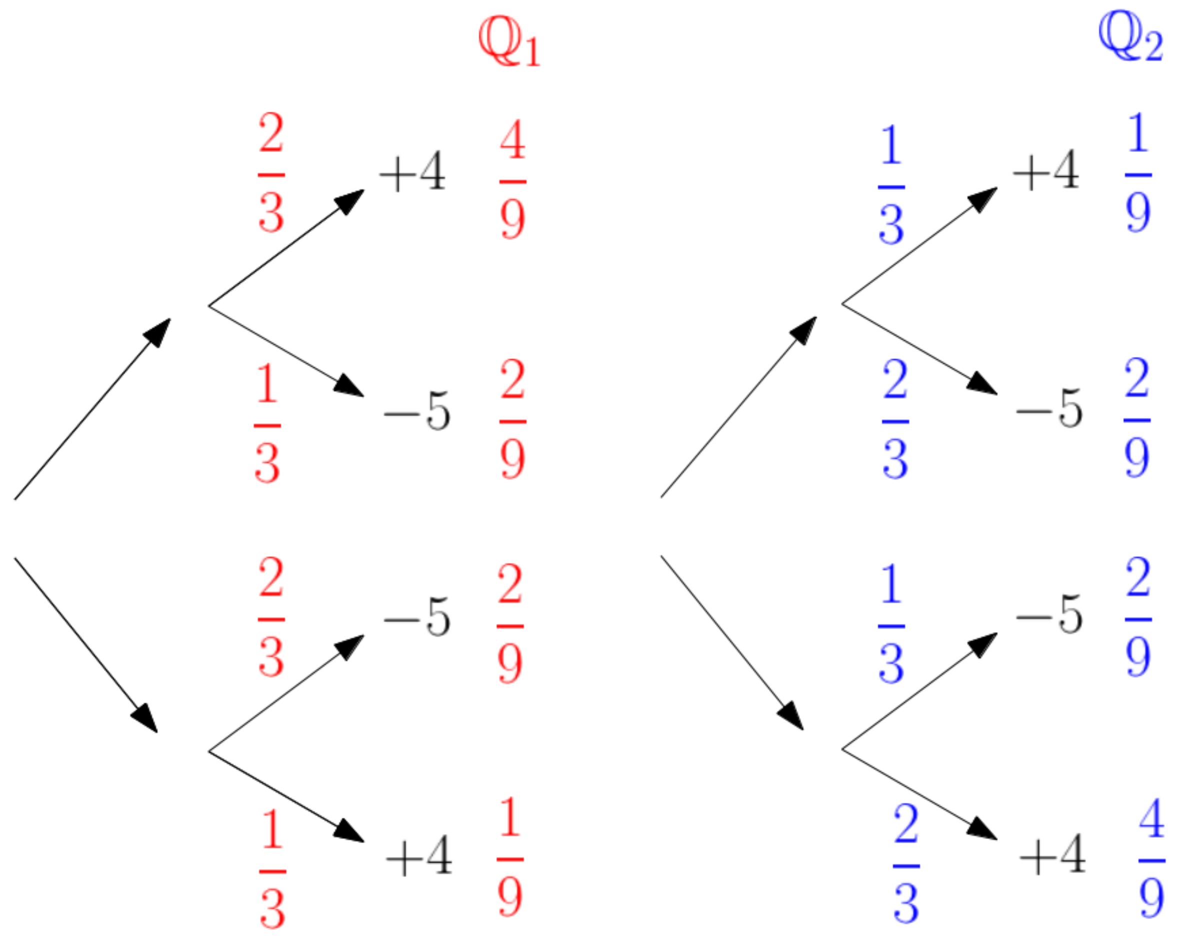


Figure 4: Time insconsistency, with a two period binomial model, an some worst case scenarios over $\mathcal{Q} = \{\mathbb{Q}_1, \mathbb{Q}_2\}$.

Definition35

A dynamic monetary risk measure $\boldsymbol{\mathcal{R}} = (\mathcal{R}_t)$ is said to be (strongly) time consistent if for all stochastic process $\boldsymbol{X}$ and all time t,

$$\boldsymbol{\mathcal{R}}_t(\boldsymbol{X}) = \boldsymbol{\mathcal{R}}_t(\boldsymbol{X} \cdot \mathbf{1}_{[t,\tau]} - \boldsymbol{\mathcal{R}}_\tau(\boldsymbol{X}) \cdot \mathbf{1}_{[\tau,\infty)})$$

where τ is some $\mathcal{F}_t$-stopping time.

We have here the interpretation of the previous exemple: the risk should ne the same

- with a direct computation at time t
- with a two step computation, at times t and τ

Further, as proved in [4]

Proposition34

The dynamic risk measure $\boldsymbol{\mathcal{R}}_t$ is time consistent if and only if $\mathcal{A}_{\boldsymbol{\mathcal{R}}_{[t,T]}} = \mathcal{A}_{\boldsymbol{\mathcal{R}}_{[t,\tau]}} + \mathcal{A}_{\boldsymbol{\mathcal{R}}_{[\tau,T]}}$

A weaker condition can be obtained, to characterize time consistency

Proposition35

A dynamic monetary risk measure $\boldsymbol{\mathcal{R}} = (\mathcal{R}_t)$ is said to be (strongly) time consistent if for all stochastic process $\boldsymbol{X}$ and all time t,

$$\boldsymbol{\mathcal{R}}_t(\boldsymbol{X}) = \boldsymbol{\mathcal{R}}_t(\boldsymbol{X} \cdot \mathbf{1}_{\{t\}} - \boldsymbol{\mathcal{R}}_{t+1}(\boldsymbol{X}) \cdot \mathbf{1}_{[t+1,\infty)})$$

Proof. Let us prove it assuming that $t \in \{0, 1, \dots, T\}$. Consider some stochastic process $\boldsymbol{X}$ and define

$$\boldsymbol{Y} = \boldsymbol{X} \cdot \mathbf{1}_{[t,\tau]} - \boldsymbol{\mathcal{R}}_\tau(\boldsymbol{X}) \cdot \mathbf{1}_{[\tau,\infty)}$$

When $t = T$, then $\boldsymbol{\mathcal{R}}_t(\boldsymbol{X}) = \boldsymbol{\mathcal{R}}_t(\boldsymbol{Y})$. Let us now consider some backward induction. One can write - using the recursive relationship and the [TI] assumption,

$$\begin{aligned}
\boldsymbol{\mathcal{R}}_t(\boldsymbol{Y}) &= \boldsymbol{\mathcal{R}}_t(-\mathbf{1}_{\{\tau=t\}}\boldsymbol{\mathcal{R}_t}(\boldsymbol{X})\mathbf{1}_{[t,\infty)} + \mathbf{1}_{\tau\geq t+1}\boldsymbol{Y}) \\
&= \mathbf{1}_{\{\tau=t\}}\boldsymbol{\mathcal{R}_t}(\boldsymbol{X}) + \mathbf{1}_{\{\tau\geq t+1\}}\boldsymbol{\mathcal{R}_t}(\boldsymbol{Y}) \\
&= \mathbf{1}_{\{\tau=t\}}\boldsymbol{\mathcal{R}_t}(\boldsymbol{X}) + \mathbf{1}_{\{\tau\geq t+1\}}\boldsymbol{\mathcal{R}_t}(\boldsymbol{Y}\mathbf{1}_{\{t\}} - \boldsymbol{\mathcal{R}_{t+1}}(\boldsymbol{Y})\mathbf{1}_{[t+1,\infty)}) \\
&= \mathbf{1}_{\{\tau=t\}}\boldsymbol{\mathcal{R}_t}(\boldsymbol{X}) + \mathbf{1}_{\{\tau\geq t+1\}}\boldsymbol{\mathcal{R}_t}(\boldsymbol{X}\mathbf{1}_{\{t\}} - \boldsymbol{\mathcal{R}_{t+1}}(\boldsymbol{X})\mathbf{1}_{[t+1,\infty)}) \\
&= \boldsymbol{\mathcal{R}}_t(\boldsymbol{X})
\end{aligned}$$

$\square$

In the case of time consistent convex risk measure, it is possible to express the penalty function using some concatenation operator, see [4]

7.5 Entropic Dynamic Risk Measure

As discussed previously, the entropic risk measure is a convex risk measure, related to the exponential utility, $u(x) = 1 - e^{-\gamma x}$. Define the realteive entropy - corresponding to the popular Kullback-Leibler divergence - of $\mathbb{Q}$ with respect to $\mathbb{P}$, with $\mathbb{Q} \ll \mathbb{P}$, defined as

$$H(\mathbb{Q}|\mathbb{P}) = \mathbb{E}\left(\frac{d\mathbb{Q}}{d\mathbb{P}} \log \frac{d\mathbb{Q}}{d\mathbb{P}}\right) = \mathbb{E}_{\mathbb{Q}}\left(\log \frac{d\mathbb{Q}}{d\mathbb{P}}\right)$$

Such a function can be a natural penalty function. More specifically, consider

$$\alpha(\mathbb{Q}) = \frac{1}{\gamma} H(\mathbb{Q}|\mathbb{P})$$

that will penalize for risk aversion. Thus, in the static case, the entropic risk measure was

$$\mathcal{R}(X) = \sup_{\mathbb{Q} \in \mathcal{M}_1(\mathbb{P})} \left\{ \mathbb{E}_{\mathbb{Q}}(-X) - \frac{1}{\gamma} H(\mathbb{Q}|\mathbb{P}) \right\}.$$

Following [8], define

$$\boldsymbol{\mathcal{R}}_t(\boldsymbol{X}) = \frac{1}{\gamma}\log\mathbb{E}(e^{-\gamma\boldsymbol{X}}|\mathcal{F}_t)$$

Proposition36

The dynamic entropic risk measure is a dynamic convex measure that is (strongly) time consistent.

Proof. Observe that

$$\begin{aligned}
\boldsymbol{\mathcal{R}}_t(-\boldsymbol{\mathcal{R}}_{t+1}(\boldsymbol{X})) &= \frac{1}{\gamma}\log\mathbb{E}\left(e^{\frac{\gamma}{\gamma}\log\mathbb{E}[e^{-\gamma\boldsymbol{X}}|\mathcal{F}_{t+1}]}|\mathcal{F}_t\right)\\
&= \frac{1}{\gamma}\underbrace{\log\mathbb{E}\left(\mathbb{E}\left(e^{-\gamma\boldsymbol{X}}|\mathcal{F}_{t+1}\right)vert\mathcal{F}_t\right)}_{=\log\mathbb{E}(e^{-\gamma\boldsymbol{X}}vert\mathcal{F}_t)}
\end{aligned}$$

so we recognize $\boldsymbol{\mathcal{R}}_t(\boldsymbol{X})$.

□

References

[1] Aliprantis, C.D. & Border, K.C. (2006). Infinite Dimensional Analysis: A Hitchhiker's Guide. Springer Verlag.

[2] Bion-Nadal, J. (2008) Dynamic risk measures: time consistency and risk measures from BMO mar-tingales. *Finance & Stochastics* 12(2), 219?244 (2008)

[4] Cheridito P., Delbaen, F. & Kupper, M. (2006). Dynamic monetary risk measures for bounded discrete-time processes *Electronic Journal of Probability*, **11**, 57-106.

[5] Choquet, G. (1953) Theory of capacities. *Annales de l?Institut Fourier* **5** 131?295.

[6] Debreu, G. (1959). Topological Methods in Cardinal Utility Theory. *Cowles Foundation Discussion Papers*, 76

[7] De Finetti, B. (1931). Sul significato soggettivo della probabilita. *Fundamenta Mathematicae* **17**, 298-329

[8] Detlefsen, K. & Scandolo, G. (2005). Conditional dynamic convex risk measures. *Finance & Stochastics* **9**, 539-561.

[9] Ekeland, I., Galichon, A. & Henry, M. (2010). Comonotonic Measures of Multivariate Risks. *Mathematical Finance.*

[10] Föllmer, H. & Schied, A. (2004) *Stochastic Finance: An Introduction in Discrete Time.* 2nd Edition, Walter de Gruyter.

[11] Föllmer H. & Penner, I (2006). Convex risk measures and the dynamics of their penalty functions *Statistics and Decisions*, **24**, 61-96.

[12] Galichon, A. (2010) The VaR at risk. *International Journal of Theoretical & Applied Finance*, **13**, 503-506.

[13] Galichon, A. & Henry, M. (2012). Dual theory of choice with multivariate risks. *Journal of Economic Theory* 147 (4), 1501-1516

[14] Gilboa, I. (2009). Theory of Decision under Uncertainty. Cambridge University Press

[15] Gilboa, I. & Schmeidler, D. (1989). Maxmin expected utility with non-unique prior, *Journal of Mathematical Economics*, **18**, 141-153.

[16] Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association*, **106**, 746-762.

[17] Grigorova, M. (2013). Quelques liens entre la théorie de l'intégrationnon-additive et les domaines de la finance et de l'assurance. PhD Thesis, Paris 7, http://tel.archives-ouvertes.fr/tel-00878599

[18] Herstein, I. N. & Milnor, J. (1953). An Axiomatic Approach to Measurable Utility. *Econometrica*, **21**, 291-297

[19] Huber, P. (1981). Robust Statistics. Wiley.

[20] Jones, M.C. (1993). Expectiles and M-quantiles are quantiles. *Statistics Probability Letters*, **20**, 149-153.

[21] Jouini, E., Schachermayer, W. & Touzi, N. (2006). Law invariant risk measures have the Fatou property, *Advances in Mathematical Economics*, **9** 49-72.

[22] Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, **47** 263–291.

[23] Kusuoka, S. (2001) On law invariant coherent risk measures. *Advances in Mathematical Economics*, **3**, 83-95.

[24] Lehmann, E.L. & Romano, J.P. (2005) *Testing Statistical Hypotheses*. Springer Verlag.

[25] Marshall, A.W. & Olkin, I. (1974). Majorization in multivariate distributions *Annals of Statistics*, **2**, 1189-1200.

[26] Melijson, I. & Nadas, A. (1979). Convex majorization with an application to the length of critical paths. *Journal of Applied Probabiliby*, **16**, 671-677.

[27] Newey, W. K., & Powell, J. L. (1987). Asymmetric least squares esti-mation and testing. *Econometrica* **55**, 819-847.

[28] Penner, I. (2007). Dynamic convex risk measures: time consistency, prudence, and sustainability. PhD Thesis, Humboldt-Universität zu Berlin.

[29] Puccetti, G. & Scarsini, M. (2010) Multivariate comonotonicity. *Journal of Multivariate Analysis* **101**, 291?304.

[30] Rachev, S.T. & Rüschendorf, L. (1998). Mass Transportation Problems. Vol. I. Springer Verlag.

[31] Rosenmüller, J. (1971) On core and value *Methods Operations Research* **9** 84?104

[32] Savage, L. J. (1971). Elicitation of personal probabilities andexpectations. {em Journal of the American Statistical Association, **66**, 783-810.

[33] Schied, A. (2004). On the Neyman-Pearson problem for law-invariant risk

measures and robust utility functionals. *Annals of Applied Probability*, **14**, 1398-1423, 2004.

[34] Schmeidler, D. (1986). Integral Representation Without Additivity. *Proceeding of the American Mathematical Society*, 97, 255-261

[35] Schmeidler. D. (1989) Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571-587.

[36] Sobel, M. (1954) On a generalization of an inequality of Hardy, Littlewood and Pólya. *Proceedings American Mathematical Society*, **5**, 596-602.

[37] Weber, S. (2006) Distribution-invariat risk measures, information and dynamic consistency. *Mathematical Finance*, **16**, 419-441.