Draft: Coherent Risk Measures

Prof. Freddy DELBAEN

Pisa - February, 28 - March, 8 2000

1 Introduction

These notes grew out of a course I taught at a Cattedra Galileiana at the Scuola Normale di Pisa, March 2000. The aim of these lectures was to show that it is possible to translate problems from Risk Management into mathematics and back.

Part of the course was devoted to an analysis of Value at Risk and its relation to quantiles. A detailed discussion of this can be found in two papers by Artzner, Delbaen, Eber and Heath, **ADEH1** and **ADEH2**. It will not be repeated here. We will rather concentrate on the mathematics behind the concept of coherent risk measures. They were introduced in the two mentioned papers and the mathematical theory was further developped in Delbaen (1999), **D2**. Further use of coherent risk measures can be found in the papers (and their references) by Jaschke, **Ja** and **JaK**, Tasche, **AT** and Föllmer-Schied, **FS**. The reader should consult the cited web-sites to find several papers dealing with this subject. The paper by Föllmer and Schied introduces a generalisation of coherent risk measures.

In section 2 we introduce the notation and recall some basic facts from functional analysis.

Section 3 gives a short description of Value at Risk. The aim of this section is to give a precise definition of what is usually called VaR. It is pointed out that VaR is not subadditive. This property is the mathematical equivalent of the diversification effect. For risk measures that are not sub-additive it may happen that diversified portfolio require more regulatory capital than less diversified portfolios. Especially in the area of Credit Risk the subadditivity property plays a fundamental role.

Section 4 introduces the concept of coherent risk measures. Basically we only deal with coherent risk measures satisfying the Fatou property. Examples are given and relations with weak compact sets of L^1 are pointed out. The example on Credit Risk shows that tail expectation (sometimes also called Worst Conditional Mean or Tailvar) is better behaved than VaR. The reader should carefully read the proof given in that section. Especially for distributions that are not continuous there are slight corrections needed in order to have coherent risk measures. We do not discuss more severe measures of risk, although we could have given examples that show that tail expectation is not yet good enough. However since there is no best risk measure, I did not pursue this discussion. The characterisation theorem permits to give many other examples of coherent risk measures. The interested reader can have a look at Delbaen (1999), **D2**, to see how Orlicz space theory can be used in the construction of coherent risk measures. We also show how convex analysis can be used. The reader familiar with Rockafellar's book, **Ro**, and with Phelps's monograph, **Ph**. can certainly find much more points in common than the ones mentioned here.

In section 5 we mention the connection with convex game theory. The basic references here are Schmeidler, bf Schm1, **Schm2**, and Delbaen, **D1**. The important relation with comonotonicity (Schmeidler's theorem) is mentioned but not proved.

Section 6 shows how coherent risk measures are related to VaR. The main result is that tail expectation is the smallest coherent risk measure, only depending on the distribution of the underlying random variable, that dominates VaR. In a recent paper Kusuoka (2000), **Ku** could characterize this family of law invariant coherent risk measures: he gives a similar proof of our result. The reader should have a look at this paper.

Section 7 gives some application of the theory to financial mathematics. This chapter was not treated during the Pisa lecturers. However I thought that it was useful to mention these results here. These observations were not published before. They show how the existence of risk neutral measures or martingale measures that have bounded densities (or more generally $p-$integrable densities), is related to a property of coherent risk measures. The basic results can be rephrased as follows. There is a martingale measure with density smaller than k if and only if for each (at no cost) attainable claim X, the tail expectation at level $1/k$ of X is positive, i.e. the position X requires extra capital. This interpretation links arbitrage theory with risk management. I am pretty sure that this relation can be developped further.

Section 8 deals with the problem of capital allocation. In our earlier papers, we emphasized that applications to performance measurement and capital allocation were among the driving forces to develop the theory. Denault, **De**, looks for axiomatics regarding this problem and wants to characterize the capital allocation via the Shapley value. The reader can look up this development in his paper. I tried to give other approaches, especially the use of Aubin's result on fuzzy games, **Au**, finds a nice interpretation and leads to the introduction of the subgradient. Here again the duality theory plays a fundamental role.

The last section deals with the definition of coherent risk measures on the space of all random variables. This extension is not obvious and poses some mathematical problems. The approach given here is much simpler than the one presented during the lectures.

I would like to use this occasion to express my thanks to the Scuola Normale for inviting me. Especially I want to thank the director Prof. Settis, Dr. Gulminelli and the Associazione Amici della Scuola Normale. It was a great honour to occupy the Cattedra Galileiana. I also want to thank Prof. Da Prato, Preside della Classe di Scienze, and Prof. Pratelli. The many discussions I had with them made this visit even more rewarding. Prof. Pratelli and his team arranged everything in such a way that nothing could go wrong.

As always, lectures only make sense if there is an audience. I thank the students of the Scuola Normale as well as the many practitioners for their interest in the subject and for the many questions they posed.

Special thanks go to Sara Biagini. She carefully took notes and later turned them into a readable text. Working with her was very pleasant.

Zürich, June 2001

Freddy Delbaen

2 The Model

We consider a very simple model in which only two dates (today and tomorrow) appear. The multiperiod model is the subject of ongoing reserach and will not be presented here. For simplicity we also suppose that all (random) sums of money available tomorrow have already been discounted. This is equivalent to assume that the interest rate is zero. Moreover we fix once and for all a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Expectation of a random variable X with respect to $\mathbb{P}$ will be denoted by $\mathbb{E}[X]$. When more than one probability measure is involved, we will explicitly mention it in the integral and we will write $\mathbb{E}_{\mathbb{P}}[X]$ or $\mathbb{P}(X)$.

In finance, replacing a probability with an equivalent one is quite frequent. From a mathematical point of view, we must pay attention since theorems and properties which depend on variance, higher moments and integrability conditions obviously depend also on the probability measure one is working with. There are two spaces that do not depend on the particular probability measure chosen. The first one is the space of (almost surely) bounded random variables L^∞ endowed with the norm:

$$\|X\|_\infty = \operatorname{ess\,sup}|X(\omega)|\ ,$$

where by ess sup of a random variable Y we denote the number $c = \inf\{\mathrm{r} \,|\, \mathbb{P}[\mathrm{Y} > \mathrm{r}] = 0\}$. We will denote by L^∞ the set of random variables X for which there exists a constant $c \in \mathbb{R}$ such that $\mathbb{P}(|X| > c) = 0$. The second invariant space is L^0, that is the space of all random variables. Such a space is usually endowed with the topology of convergence in probability that is

$$X_n \xrightarrow{\mathbb{P}} X \quad \text{iff} \quad \forall \epsilon > 0 \quad \mathbb{P}[|X_n - X| > \epsilon] \to 0\,,$$

or, equivalently, iff

$$\mathbb{E}[|X_n - X| \wedge 1] \to 0\,,$$

where by $a \wedge b$, we denote the minimum between a and b.

Many theorems in measure theory refer to convergence almost surely, altough they remain valid when convergence a.s. is replaced by convergence in probability. This is the case for the dominated convergence theorem of Lebesgue, Fatou's lemma, etc. We will use these extensions without further notice.

We will denote by $L^1(\Omega, \mathcal{F}, \mathbb{P})$ (or simply with L^1) the space of integrable random variables on $(\Omega, \mathcal{F}, \mathbb{P})$.The dual space of L^1 is L^∞ and the duality (L^1, L^∞) will play a special role. The dual space of L^∞ is $\mathbf{ba}(\Omega, \mathcal{F}, \mathbb{P})$ or just **ba** if no confusion can arise. It is the space of bounded finitely additive measures μ such that $\mathbb{P}[A] = 0$ implies $\mu(A) = 0$.

We will frequently make use of the standard duality theory from functional analysis. The reader can find the relevant theorems in Dunford-Schwartz, **DuS**. The following theorem is very useful when checking whether sets in a dual space are weak* closed. The theorem is sometimes called the Banach-Dieudonné theorem, sometimes it is referred to as the Krein-Sḿulian theorem.

Theorem 1 *Let E be a Banach space with dual space E^*. Then a convex wet $C \subset E^*$ is weak* closed if and only if for each n, the set $W_n = C \cap \{e^* \mid \|e^*\| \leq n\}$ is weak* closed.*

Of course, since convex sets that are closed for the so-called Mackey topology are already weak* closed, we can check whether the sets W_n are Mackey closed. Most of the time, the description of the Mackey topology is not easy, but in the case of L^∞ we can make it more precise. Without giving a proof, we recall that on bounded sets of L^∞, the so-called Mackey topology coincides with the topology of convergence in probability. Checking whether a bounded convex set is weak* closed is then reduced to checking whether is it closed for the convergence in probability. More precisely we have the following lemma, which seems to be due to Grothendieck.

Lemma 1 *Let $\mathcal{A} \subseteq L^\infty$ be a convex set. Then $\mathcal{A}$ closed for the $\sigma(L^\infty, L^1)$ topology iff for each n, the set $W_n = \{X \mid X \in \mathcal{A}, \|X\|_\infty \le n\}$ is closed with respect to convergence in probability.*

We will use two more theorems that play a fundamental role in convex analysis, these are the Bishop-Phelps theorem and James's characterisation of weakly compact sets (see Diestel's book, **Di**, for a proof of these non-trivial results).

Theorem 2 *(Bishop-Phelps) Let $B \subset E$ be a bounded closed convex set of a Banach space E. The set $\{e^* \in E^* \mid e^* \text{ attains its supremum on } B\}$ is norm dense in E^*.*

Theorem 3 *(James) Let $B \subset E$ be a bounded convex set of a Banach space E. The set B is weakly compact if and only if we have that each $e^* \in E^*$ attains its maximum on B. More precisely for each $e^* \in E^*$ there is $b_0 \in B$ such that $e^*(b_0) = \sup_{b\in B} e^*(b)$.*

3 Value at Risk

The philosophy behind the concept of VaR is the following: fix a threshold probability α (say 1%) and define a position as acceptable if and only if the probability to go bankrupt is smaller than α. The main problem with VaR is that it does not distinguish between a bankruptcy of, say, 1 million or 1 hundred million Euro. Anyway, VaR is the most widely used instrument to control risk and in order to study its properties we need a more precise definition.

Definition 1 *Let X be a random variable and $\alpha \in [0, 1]$.*

- *q is called an α–quantile if:*

$$\mathbb{P}[X < q] \le \alpha \le \mathbb{P}[X \le q] \,,$$

- *the largest α–quantile is:*

$$q_\alpha(X) = \inf\{x \mid \mathbb{P}[X \leq x] > \alpha\},$$

- *the smallest α–quantile is:*

$$q_\alpha^-(X) = \inf\{x \mid \mathbb{P}[X \leq x] \geq \alpha\}.$$

As easily seen, $q_\alpha^- \leq q_\alpha$ and q is an α–quantile if and only if $q_\alpha^- \leq q \leq q_\alpha$.

Definition 2 *Given a position X and a number $\alpha \in [0,1]$ we define*

$$\mathrm{VaR}_\alpha(X) := -q_\alpha(X)$$

*and we call X VaR-***acceptable** *if* $\mathrm{VaR}_\alpha(X) \leq 0$ *or, equivalently, if $q_\alpha \geq 0$.*

We can think of the VaR as the amount of extra-capital that a firm needs in order to reduce the probability of going bankrupt to α. A negative VaR means that the firm would be able to give back some money to its shareholders or that it could change its activities, e.g. it could accept more risk.

Remark 1 VaR has the following properties:

1. $X \geq 0 \Longrightarrow \mathrm{VaR}_\alpha(X) \leq 0$,
2. $X \geq Y \Longrightarrow \mathrm{VaR}_\alpha(X) \leq \mathrm{VaR}_\alpha(Y)$,
3. $\mathrm{VaR}_\alpha(\lambda X) = \lambda \mathrm{VaR}_\alpha(X)$, $\forall \lambda \geq 0$,
4. $\mathrm{VaR}_\alpha(X + k) = \mathrm{VaR}_\alpha(X) - k$, $\forall k \in \mathbb{R}$.

In particular, we have $\mathrm{VaR}_\alpha(X + \mathrm{VaR}_\alpha(X)) = 0$.

VaR also has the nice property that it is defined on the whole space L^0. Therefore it can, in principle, be calculated on any random variable. The problem with such a level of generality is that VaR_α loses convexity properties. As an example, consider the case of a bank which has given a \$ 100 loan to a client whose default probability is equal to 0.008. If $\alpha = 0.01$, it is easy to see that $\mathrm{VaR}_\alpha(X) \leq 0$. Consider now another bank which has given two loans of \$ 50 each and for both, the default probability is equal to 0.008. In case the default probabilities of the two loans are independent, $\mathrm{VaR}_\alpha(X)$ is \$ 50. Hence we have that diversification, which is commonly considered as a way to reduce risk, can lead to an increase of VaR. Therefore we argue that VaR is not a good measure of risk. This is the main reason why we are interested in studying other types of risk measures.

4 Coherent Risk Measures

Definition 3 *A coherent risk measure is a function* $\rho : L^\infty \longrightarrow \mathbb{R}$ *such that*

1. $X \geq 0 \Longrightarrow \rho(X) \leq 0$,
2. $\rho(\lambda X) = \lambda \rho(X)$, $\forall \lambda \geq 0$,
3. $\rho(X + k) = \rho(X) - k$, $\forall k \in \mathbb{R}$,
4. $\rho(X + Y) \leq \rho(X) + \rho(Y)$.

Point *4* (sub-additivity) is the one which is not satisfied by VaR, even if it seems to be a reasonable assumption. In fact, subadditivity of a risk measure is a mathematical way to say that diversification leads to less risk. See **ADEH1** and **ADEH2** for a discussion of the axiomatics.
The following two properties are immediate consequences of the definition:

1. $a \leq X \leq b \Longrightarrow -b \leq \rho(X) \leq -a$,
2. $\rho(X + \rho(X)) = 0$.

Example 1 Let us take a family $\mathcal{P}$ of probability measures $\mathbb{Q}$ all absolutely continuous with respect to $\mathbb{P}$. We identify $\mathbb{Q}$ and $\frac{d\mathbb{Q}}{d\mathbb{P}}$, the Radon-Nikodým derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$. Of course, $\frac{d\mathbb{Q}}{d\mathbb{P}}$ is integrable with respect to $\mathbb{P}$. We can therefore identify $\mathcal{P}$ with a subset of $L^1(\Omega, \mathcal{F}, \mathbb{P})$. If we define

$$\rho_{\mathcal{P}}(X) = \sup \{\mathbb{E}_{\mathbb{Q}}[-X] \mid \mathbb{Q} \in \mathcal{P}\}$$

then this $\rho_{\mathcal{P}}$ is indeed a coherent risk measure. Moreover we will show that **any coherent risk measure has such a form**.

Just to make things easier, we add a continuity axiom to the definition of a coherent risk measure.

5. *(The Fatou property) Given a sequence* $(X_n)_{n\geq 1}$, *such that* $\|X_n\|_\infty \leq 1$, *then*

$$X_n \xrightarrow{\mathbb{P}} X \quad \text{implies} \quad \rho(\mathrm{X}) \leq \liminf \rho(\mathrm{X_n})$$

It is possible to show (in a way similar to the proof of Fatou's lemma) that the Fatou property is equivalent to a monotonicity property:

$$0 \leq X_n \leq 1, \quad X_n \downarrow 0 \Longrightarrow \rho(X_n) \uparrow 0.$$

Definition 4 *We say that property (*) is satisfied if*

$$-1 \leq X_n \leq 0, \quad X_n \uparrow 0 \Longrightarrow \rho(X_n) \downarrow 0.$$

Remark 2 Property (*) implies the Fatou property.

Proof. Let $0 \leq Y_n \leq 1$, $Y_n \downarrow 0$, then $0 = Y_n + (-Y_n)$ implies $0 \leq \rho(Y_n) + \rho(-Y_n)$. By *1* and property *(*)* we also have $0 \geq \rho(Y_n) \geq -\rho(-Y_n) \to 0$ so that the Fatou property holds. □

Since the subadditivity inequality we used in the proof of the lemma, does not hold in the other direction, we get that property *(*)* might be strictly stronger than the Fatou property (and this is indeed the case).

Remark 3 As an application of Fatou's lemma, one can show that $\rho_{\mathcal{P}}$ satisfies the continuity axiom.

Proof. If $X_n \xrightarrow{\mathbb{P}} X$ and $\|X_n\|_\infty \leq 1$ then for every $\mathbb{Q} \in \mathcal{P}$ we have:

$$\mathbb{E}_{\mathbb{Q}}[-X] \leq \operatorname{liminf} \mathbb{E}_{\mathbb{Q}}[-\mathrm{X_n}] \leq \operatorname{liminf} \rho_{\mathcal{P}}(\mathrm{X_n})$$

and therefore $\rho_{\mathcal{P}}(X) \leq \operatorname{liminf} \rho_{\mathcal{P}}(\mathrm{X_n})$. □

In working with a family $\mathcal{P}$, we can replace it with its convex L^1-closed hull, so that, from now on, we will take $\mathcal{P}$ to be convex and L^1-closed.

Example 2 We consider $\mathcal{P} = \{\mathbb{P}\}$. In this case, $\rho_{\mathcal{P}} = \mathbb{E}_{\mathbb{P}}[-X]$. A position X is acceptable iff its average $\mathbb{E}_{\mathbb{P}}[X]$ is nonnegative. Clearly, such a risk measure is too tolerant.

Example 3 Let us consider $\mathcal{P} = \{\mathbb{Q} \mid \text{probability on } (\Omega, \mathcal{F}), \mathbb{Q} \ll \mathbb{P}\}$. In this case $\rho_{\mathcal{P}} = \text{ess sup}(-X)$ and $\rho_{\mathcal{P}}(X) \leq 0$ if and only if $X \geq 0$. Hence a position is acceptable if and only if it is nonnegative. The family $\mathcal{P}$ is too large and therefore $\rho_{\mathcal{P}}$ is too severe. Anyway this $\rho_{\mathcal{P}}$ provides an example of a coherent risk measure that satisfies the Fatou property but does not verify property (*). If we consider $X_n = -e^{-nx}$ defined on $[0,1]$ with the Lebesgue σ-algebra and the Lebesgue measure, we have that $X_n \uparrow 0$, almost surely, while $\text{esssup}(-X_n) = 1$.

Example 4 Let us now see what happens for the convex closed set $\mathcal{P}_k = \{\mathbb{Q} \mid \frac{d\mathbb{Q}}{d\mathbb{P}} \leq k\}$. Obviously we only need to investigate the case $k > 1$; indeed, $\frac{d\mathbb{Q}}{d\mathbb{P}} \leq 1$ implies that $\mathbb{Q} = \mathbb{P}$, i.e. $\mathcal{P}_1$ reduces to the singleton $\{\mathbb{P}\}$.

Theorem 4 *If X has a continuous distribution and $\alpha = 1/k$, then*

$$\rho_{\mathcal{P}_k}(X) = \mathbb{E}_{\mathbb{P}}[-X \mid X \leq q_\alpha(X)] \geq -q_\alpha(X) = \mathrm{VaR}_\alpha(X).$$

Proof. Since X has a continuous distribution, we get $\mathbb{P}[X \leq q_\alpha(X)] = \alpha = 1/k$. Define now $\mathbb{Q}_0$ such that $\frac{d\mathbb{Q}_0}{d\mathbb{P}} = k1_A$ with $A = \{X \leq q_\alpha(X)\}$. Since $\mathbb{Q}_0 \in \mathcal{P}_k$ and $\mathbb{E}_{\mathbb{Q}_0}[-X] = \mathbb{E}_{\mathbb{P}}[-X \mid A]$ we have $\rho_{\mathcal{P}_k}(X) \geq \mathbb{E}_{\mathbb{P}}[-X \mid X \leq q_\alpha(X)]$. By considering now an arbitrary $\mathbb{Q} \in \mathcal{P}_k$, we have

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}[-X] &= \int (-X) d\mathbb{Q} = \int_A (-X) \frac{d\mathbb{Q}}{d\mathbb{P}} d\mathbb{P} + \int_{A^c} (-X) \frac{d\mathbb{Q}}{d\mathbb{P}} d\mathbb{P} \\
&= k \int_A (-X) d\mathbb{P} + \int_A (-X) \left(\frac{d\mathbb{Q}}{d\mathbb{P}} - k \right) d\mathbb{P} + \int_{A^c} (-X) \frac{d\mathbb{Q}}{d\mathbb{P}} d\mathbb{P} \\
&\leq k \int_A (-X) d\mathbb{P} + (-q_\alpha(X)) \int_A \left(\frac{d\mathbb{Q}}{d\mathbb{P}} - k \right) d\mathbb{P} + (-q_\alpha(X)) \int_{A^c} \frac{d\mathbb{Q}}{d\mathbb{P}} d\mathbb{P} \\
&\leq k \int_A (-X) d\mathbb{P} + (-q_\alpha(X))[\mathbb{Q}(A) - k\mathbb{P}(A) + \mathbb{Q}(A^c)] \\
&= k \int_A (-X) d\mathbb{P} \\
&= \mathbb{E}_{\mathbb{Q}_0}[-X].
\end{aligned} \tag{1}$$

This ends the proof. □

Later we will prove that $\rho_{\mathcal{P}_k}$ is the smallest distribution invariant coherent risk measure greater than VaR. When the distribution of X has a discontinuity at q_α, the probability measure $\mathbb{Q}_0$ such that $\frac{d\mathbb{Q}_0}{d\mathbb{P}} = k1_{\{X<q_\alpha\}} + \beta 1_{\{X=q_\alpha\}}$ (with a suitably chosen $\beta, 0 \leq \beta \leq 1$) does the job. It implies that

$$\rho_{\mathcal{P}_k}(X) = \frac{1}{\alpha} \left(\int_{X<q_\alpha} (-X) d\mathbb{P} + (\alpha - \mathbb{P}[X < q_\alpha])(-q_\alpha) \right).$$

We now give an example of bad performance of VaR against ρ_k in Credit Risk.

Example 5 Let us imagine there is a bank which lends \$ 1 to 150 clients, who are independent and with the same default probability p of 1.2%. For each client i let us put $Z_i = 0$ if he/she does not default and $Z_i = 1$ if he/she defaults. So we first suppose $(Z_i)_i$ are independent Bernoulli random variables with $\mathbb{P}[Z_i = 1] = 1.2\%$. The number

$Z = \sum_i Z_i$ represents the total number of defaults and therefore the bank's loss. It has the binomial distribution:

$$\mathbb{P}[Z = k] = \left(\begin{array}{c} 150 \\ k \end{array} \right) p^k (1-p)^{150-k} .$$

With $\alpha = 1\%$ we have $VaR_\alpha = 5$ and tail expectation $\rho_{1/\alpha} = 6.287$.
If we now suppose that the clients are **dependent**, things change. A simple way of obtaining a well-behaved dependence structure is by replacing $\mathbb{P}$ with a new probability measure $\mathbb{Q}$ defined as:

$$d\mathbb{Q} = c\, e^{\epsilon Z^2} d\mathbb{P} ,$$

where Z and $\mathbb{P}$ are the same as before, ϵ is positive and c is a normalising constant. Now $\mathbb{Q}[Z_i = 1]$ increases with ϵ: if we take ϵ so that $\mathbb{Q}[Z = 1] = 1.2\%$ (taking $p = 1\%$ and $\epsilon = 0.03029314$) then we obtain $VaR_\alpha = 6$ and tail expectation $\rho_{1/\alpha} = 14.5$.

We notice that VaR is not able to detect the difference between the two cases, which are better differentiated by tail expectation.

This can be explained as follows. VaR only looks at a quantile, it does not tell us how great the losses are. However, tail expectation takes an average over the worst cases and therefore takes into account the tail distribution of the losses. The probability $\mathbb{Q}$ allows to introduce loans whose defaults are dependent on a common economic factor. It reflects the situation that if a substantial percentage defaults, the conditional probability that others default as well, is very high.

Example 6 We could also consider the following family (where $k > 1$ and $p > 1$):

$$\mathcal{P}_{p,k} = \left\{ \mathbb{Q} \mid \left\| \frac{d\mathbb{Q}}{d\mathbb{P}} \right\|_p \leq k \right\}$$

The following theorem holds:

Theorem 5 *There exists a constant $c = 1 \wedge (k-1)$ such that for all $X \in L^\infty$, $X \leq 0$ we have:*

$$c\|X\|_q \leq \rho_{\mathcal{P}_{p,k}} \leq k\|X\|_q$$

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

While the second inequality is just an application of Hölder's inequality, the first one needs a proof. The intererested reader can check Delbaen (1999).

Remark 4 If k tends to 1, c goes to 0 and the family $\mathcal{P}_{p,k}$ shrinks to $\{\mathbb{P}\}$. That c tends to zero has to be expected since the L^p and the L^1 norms are not equivalent.

Remark 5 Actually, if $p = q = 2$ we have:

$$\left\|\frac{d\mathbb{Q}}{d\mathbb{P}} - 1\right\|_2^2 = E\left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)^2\right] - 1 \leq k^2 - 1$$

so that the densities go to 1 in L^2 as k tends to 1. If $p \geq 2$ we can use the same argument (remember that $\|.\|_2 \leq \|.\|_p$) and if $p < 2$, Clarkson's inequality for L^p-norms must be used.

Example 7 This example is related to work of T. Fischer, see **Fi**. He suggested, among other constructions, the following coherent risk measures. For $X \in L^\infty$ we define

$$\rho(X) = -\mathbb{E}[X] + \alpha\|(X - \mathbb{E}[X])^-\|_p.$$

The reader can verify that for $0 \leq \alpha \leq 1$ and $1 \leq p \leq \infty$, this defines a coherent risk measure. This measure can also be found using a set of probability measures. So let

$$\mathcal{P} = \{1 + \alpha(g - \mathbb{E}[g]) \mid g \geq 0; \|g\|_q \leq 1\}.$$

Here of course $q = p/(p-1)$, with the usual interpretation if $p = 1, \infty$. Clearly the set $\mathcal{P}$ is a convex L^1-closed set of functions h that have expectation equal to 1. We still have to check the positivity of such functions. This is easy since, by $g \geq 0$ and $\alpha \in [0, 1]$, we have

$$1 + \alpha(g - \mathbb{E}[g]) \geq 1 - \alpha\,\mathbb{E}[g] \geq 1 - \|g\|_q \geq 0.$$

We will check that

$$\rho(X) = \sup\{\mathbb{E}[h(-X)] \mid h \in \mathcal{P}\}.$$

To see this, take $h = 1 + \alpha(g - \mathbb{E}[g])$ where $g = \frac{((X-\mathbb{E}[X])^-)^{(p-1)}}{\|(X-\mathbb{E}[X])^-\|_p^{(p-1)}}$. This is the standard way to obtain the $p-$norm by integrating against a function with $q-$norm equal to 1. In case $p = 1$ and therefore $q = \infty$, we take for g the indicator function of the set where $X < \mathbb{E}[X]$. For this choice of g and h we get for $X = -Y \in L^\infty$:

$$\begin{aligned}
\mathbb{E}[h\,Y] &= \mathbb{E}[Y] + \mathbb{E}[h\,(Y - \mathbb{E}[Y])] \\
&= \mathbb{E}[Y] + \mathbb{E}[(h - 1 + \alpha\,\mathbb{E}[g])(Y - \mathbb{E}[Y])] = \mathbb{E}[Y] + \alpha\,\|(Y - \mathbb{E}[Y])^+\|_p.
\end{aligned}$$

For an arbitrary $1 + \alpha(g - \mathbb{E}[g]) = h \in \mathcal{P}$ we have, by Hölder's inequality:

$$\mathbb{E}[h\,Y] \leq \mathbb{E}[Y] + \|h - 1 + \alpha\,\mathbb{E}[g]\|_q\,\|(Y - \mathbb{E}[Y])^+\|_p \leq -\mathbb{E}[X] + \alpha\,\|(X - \mathbb{E}[X])^-\|_p.$$

4.1 Characterization of coherent risk measures

Let ρ be a coherent risk measure, $\rho : L^\infty \longrightarrow \mathbb{R}$ and let us assume that the Fatou property holds. Let $\mathcal{A}$ be the set of the acceptable positions, i.e. $\mathcal{A} = \{X \mid \rho(X) \leq 0\}$. We note that $\mathcal{A}$ is a convex cone (by the subadditivity and positive homogeneity properties of coherent risk measures). The next theorem focuses on the relations between ρ and $\mathcal{A}$:

Theorem 6 *If ρ satisfies the Fatou property, then:*

1. *$\mathcal{A}$ is closed for the weak* topology $\sigma(L^\infty, L^1)$ and $\mathcal{A} \supseteq L^\infty_+$;*
2. *$\rho(X) = \inf\{\alpha \mid X + \alpha \in \mathcal{A}\}$;*
3. *$\mathcal{A} = \{X \mid \exists \alpha \geq 0, \exists Y \ s.t. \quad \rho(Y) = 0 \ , \ X = Y + \alpha \ \}$.*

Conversely, if $\mathcal{A}$ is a convex cone, closed in the $\sigma(L^\infty, L^1)$ topology and containing L^∞_+, then $\tilde{\rho}(X) := \inf\{\alpha \mid \ X + \alpha \in \mathcal{A}\}$ defines a coherent risk measure with the Fatou property.

Proof. We start from point *1*. Let us call W the intersection of $\mathcal{A}$ with the unit ball of L^∞. By the Krein-Smulian theorem, if W is closed in the weak* topology, then $\mathcal{A}$ is also closed. We take a sequence $(X_n)_n \in W$ such that $X_n \xrightarrow{\mathbb{P}} X$. But then $\rho(X) \leq \liminf \rho(X_n) \leq 0$: so $X \in W$, that is W is closed under convergence in probability.

Points *2* and *3* are easy exercises.

In order to show the converse, we consider the following:

$$\mathcal{A}^o = \{f \mid f \in L^1 \text{ and } \forall X \in \mathcal{A} \, : \, \mathbb{E}[Xf] \geq 0\}$$

which is, by definition, the polar cone of $\mathcal{A}$. $\mathcal{A}^o$ is L^1 closed and it is contained in L^1_+ because $\mathcal{A} \supseteq L^\infty_+$. We define $\mathcal{P}$ to be the closed convex set $\{f \in \mathcal{A}^o \,|\, \mathbb{E}[f] = 1\}$, which is, by the way, a basis of the cone $\mathcal{A}^o$. This means that $\mathcal{A}^o = \cup_{\lambda \geq 0} \lambda \mathcal{P}$. The bipolar theorem guarantees that:

$$\begin{aligned} \mathcal{A} &= \{X \mid \forall f \in \mathcal{A}^o : \ \mathbb{E}[Xf] \geq 0\} \\ &= \{X \mid \forall f \in \mathcal{P} : \ \mathbb{E}[Xf] \geq 0\} \end{aligned}$$

and therefore:

$$\begin{aligned} \tilde{\rho}(X) &= \inf\{\alpha \mid X + \alpha \in \mathcal{A}\} \\ &= \inf\{\alpha \mid \forall f \in \mathcal{P} : \quad \mathbb{E}[(X + \alpha)f] \geq 0\} \\ &= \inf\{\alpha \mid \forall f \in \mathcal{P} : \quad \mathbb{E}[(-X)f] \leq \alpha\} \\ &= \sup\{\mathbb{E}[(-X)f] \mid f \in \mathcal{P}\} \end{aligned}$$

□

Remark 6 We have in fact established a one-to-one correspondence between:
(a) convex closed sets $\mathcal{P}$ consisting of probabilities which are absolutely continuous with respect to $\mathbb{P}$,
(b) $\sigma(L^\infty, L^1)$-closed convex cones $\mathcal{A}$, containing L^∞_+,
(c) coherent risk measures ρ with the Fatou property.

Theorem 7 *(On weak compactness)*
For closed convex sets $\mathcal{P}$ of probabilities, the following are equivalent:

1. $\mathcal{P}$ is weakly compact;

2. $\mathcal{P}$ is weakly sequentially compact;

3. the set $\{\frac{d\mathbb{Q}}{d\mathbb{P}} \,|\, \mathbb{Q} \in \mathcal{P}\}$ is uniformly integrable;

4. (de la Vallée-Poussin criterion for uniform integrability) there exists $\phi : \mathbb{R}^+ \to \mathbb{R}$, increasing, convex continuous, with $\phi(0) = 0$ such that :
$\lim_{x\to\infty} \frac{\phi(x)}{x} = +\infty$ and $\sup_{\mathbb{Q}\in\mathcal{P}} \mathbb{E}[\phi(\frac{d\mathbb{Q}}{d\mathbb{P}})] < \infty$.

We do not prove this variant of the Dunford-Pettis theorem. It is a basic result in $L^1 - L^\infty$ theory.

Examples.
(a) $\phi(x) = x^p$, $p > 1$; together with point *4* this implies that $\mathcal{P}_{p,k}$ is a weakly compact family; we also have that $\mathcal{P}_k$ is weakly compact;
(b) $\phi(x) = (x + 1)\log(x + 1) - x$; this is another example that can be used in connection with Orlicz space theory. See Delbaen (1999).
According to the above, for coherent risk measures the following are equivalent:

1. if $-1 \le X_n \le 0$ and $X_n \uparrow 0$, then $\rho(X_n) \downarrow 0$;

2. if $(A_n)_{n\ge1}$ is a family of measurable sets such that $A_n \downarrow \emptyset$, then $\sup_{\mathbb{Q}\in\mathcal{P}} \mathbb{Q}[A_n] \to 0$;

3. the set $\{\frac{d\mathbb{Q}}{d\mathbb{P}} \,|\, \mathbb{Q} \in \mathcal{P}\}$ is uniformly integrable;

4. $\mathcal{P}$ is weakly compact.

Remember that point *1* is stronger than the Fatou property!

Theorem 8 *$\mathcal{P}$ is weakly compact iff for every $X \in L^\infty$ there is $\mathbb{Q} \in \mathcal{P}$ such that $\rho(X)$ is exactly $\mathbb{Q}[-X]$ (i.e. $\rho(X)$ is not only a supremum, but also a maximum).*

Proof. A direct application of James's theorem. $\square$

Theorem 9 *If $\mathcal{P}$ is weakly compact then:*

$$\|X_n\|_\infty \leq 1\,, \quad X_n \xrightarrow{\mathbb{P}} X \quad \text{implies} \quad \rho(\mathrm{X}) = \lim_{\mathrm{n}\to\infty} \rho(\mathrm{X_n})$$

Proof. A direct application of the property that $\mathcal{P}$ is uniformly integrable. $\square$

We give now an application to a Credit Risk situation.

Example 8 Suppose that $(X_n)_n$ are i.i.d and that $\|X_n\|_\infty \leq 1$. The random variable X_i stands for the credit risk loss corresponding to the i-th person (the group is supposed to be independent). Let $S_n = X_1 + \ldots + X_n$. The problem is calculating the whole capital needed to face the risk. Therefore we need $\rho(S_n)$ and the capital we will charge to each person will be $\frac{1}{n}\rho(S_n) = \rho(\frac{S_n}{n})$. Suppose now that $\mathcal{P}$ is weakly compact. By the law of large numbers,

$$\frac{S_n}{n} \xrightarrow{a.s.} \mathbb{E}[X_1]$$

so that

$$\rho\left(\frac{S_n}{n}\right) \longrightarrow \rho(\mathbb{E}[X_1]) \equiv -\mathbb{E}[X_1]$$

If we do not have independence, but the correlation coefficients tend to zero when n goes to infinity, the previous result still holds. Indeed if

$$\lim_{k\to\infty} \sup_n \left(\mathbb{E}[X_n X_{k+n}] - \mathbb{E}[X_n]\mathbb{E}[X_{n+k}]\right) \to 0,$$

then by Bernstein's theorem, $\frac{S_n}{n}$ tends to $\mathbb{E}[X_1]$ in probability if $n \to \infty$. We leave the interpretation of this result to the reader.

4.2 Utility theory and risk measures

A function $u : L^\infty \to \mathbb{R}$ is called a quasi-concave monotone utility function if two properties are satisfied:
(a) weak monotonicity: $X \geq Y$ then $u(X) \geq u(Y)$
(b) $\forall \alpha \in \mathbb{R}$ the set $\{X \,|\, u(X) \geq \alpha\}$ is convex. A utility function induces a preference

relation: $X \succeq Y$ (X is at least as good as Y) iff $u(X) \geq u(Y)$.

We say that u is strongly monotone if $X \geq Y$ and $\mathbb{P}[X > Y] > 0$ imply the "strict preference" $u(X) > u(Y)$.

If we start with a coherent risk measure ρ, we obtain a utility function by simply putting $u = -\rho$. Thanks to subadditivity and homogeneity of ρ, such a u is quasi-concave.

Definition 5 *A coherent risk measure ρ is called relevant if $X \in L^\infty$, $X \leq 0$ and $\mathbb{P}[X < 0]$ imply $\rho(X) < 0$.*

Theorem 10 *The following are equivalent:*

1. *ρ is relevant;*
2. *$\forall A \in \mathcal{F} \quad \mathbb{P}[A] > 0 \Rightarrow \rho(-I_A) > 0$;*
3. *$\forall A \in \mathcal{F} \quad \mathbb{P}[A] > 0 \Rightarrow \exists \mathbb{Q} \in \mathcal{P} \quad$ s.t. $\quad \mathbb{Q}[A] > 0$;*
4. *(Halmos-Savage) there exists a probability $\mathbb{Q}_o \in \mathcal{P}$ such that:*

$$\mathbb{P}[A] > 0 \quad \textit{implies} \quad \mathbb{Q}_o[A] > 0.$$

Proof. The first three equivalences are fairly obvious: the only part requiring a proof is the implication *1 ⇒ 4*. We repeat the proof of this result due to Halmos-Savage. The argument is a classical exhaustion.
Put $\mathcal{C} = \{ \{\omega : \frac{d\mathbb{Q}}{d\mathbb{P}}(\omega) > 0\} \,|\, \mathbb{Q} \in \mathcal{P}\}$ that is the collection of the supports of the probabilities in $\mathcal{P}$. We observe that $\mathcal{C}$ is closed under finite unions, because $\mathcal{P}$ is convex. It is even closed under countable unions. If $(A_n)_n$ is a sequence in $\mathcal{C}$ and each A_n is the support of a $\mathbb{Q}_n$, the probability $\mathbb{Q} := \sum_{n=1}^\infty \frac{1}{2^n}\mathbb{Q}_n$ belongs to $\mathcal{P}$ (because the latter is convex closed) and the support of $\mathbb{Q}$ is exactly $\cup_n A_n$. Let $(B_n)_n$ be a sequence in $\mathcal{C}$ so that $\mathbb{P}(B_n) \to \sup\{\mathbb{P}(A) \mid A \in \mathcal{C}\}$. Since $\mathcal{C}$ is stable for countable unions, we have that $B = \cup_n B_n \in \mathcal{C}$ and $\mathbb{P}(B) = \alpha$. Let $\mathbb{Q}^o$ be a probability measure in $\mathcal{P}$ with $\{\frac{d\mathbb{Q}^o}{d\mathbb{P}} > 0\} = B$. We claim that $\mathbb{P}(B) = 1$. Indeed, if $\mathbb{P}(B^c) > 0$, we could have the existence of $\mathbb{Q} \in \mathcal{P}$ with $\mathbb{Q}(B^c) > 0$ and $(\mathbb{Q}^o + \mathbb{Q})/2 \in \mathcal{P}$ would have a support with measure bigger than α. Thus $\mathbb{Q}^o$ is an element of $\mathcal{P}$, equivalent to $\mathbb{P}$. □

As regards utility functions induced by risk measures, we have that relevance does not imply strict monotonicity. For instance, take an atomless space Ω and consider the set $\mathcal{P}_2 = \{\mathbb{Q} \mid \frac{d\mathbb{Q}}{d\mathbb{P}} \leq 2\}$. Then ρ_2 is relevant (because $\mathbb{P}$ itself belongs to $\mathcal{P}_2$)

but the derived u is not strictly monotone. If A is such that $0 < \mathbb{P}[A] < \frac{1}{2}$, then $u(I_A) = -\rho(A) = \inf\{\mathbb{Q}[A] : \mathbb{Q} \in \mathcal{P}\} = 0 \quad (\frac{1}{\mathbb{P}[A^c]} I_{A^c}$ is an element of $\mathcal{P}$!). Of course we have $u(I_A) = u(0) = 0$.

4.3 Operations on risk measures

i) Maximum of two risk measures.

Let ρ_1 and ρ_2 be two distinct risk measures. Just to give an interpretation, they could stand for two different measures of risk calculated for the same company. The first one is the risk measure of the manager and the second the shareholder's one. If both groups must be pleased, it is natural to ask for a risk measure which is more severe than each of the two, that is:

$$\rho \equiv \rho_1 \vee \rho_2$$

We leave it to the reader to check that ρ is indeed a coherent risk measure that also satisfies the Fatou property if ρ_1 and ρ_2 do. If we call $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{A}$ the sets of acceptable positions (the first induced by ρ_1, the second by ρ_2 and the third by ρ respectively) and we define $\mathcal{P}_1$, $\mathcal{P}_2$ and $\mathcal{P}$ to be the related families of probabilities, we have:

$$\begin{aligned} \mathcal{A} &= \mathcal{A}_1 \cap \mathcal{A}_2 \\ \mathcal{P} &= conv\,(\mathcal{P}_1, \mathcal{P}_2) \end{aligned}$$

We will indeed show that $\mathcal{P}$ is closed.

Actually:

$$\mathcal{A} = \{X \,|\, \rho(X) \leq 0\} \equiv \{X \,|\, \rho_1(X) \leq 0 \,\text{and}\, \rho_2(\text{X}) \leq 0\} = \mathcal{A}_1 \cap \mathcal{A}_2 \,.$$

Since the acceptable set characterizes the risk measure, we can find the corresponding set $\mathcal{P}$:

$$\begin{aligned} X \in \mathcal{A} \;&\Leftrightarrow\; X \in \mathcal{A}_1 \quad \text{and} \quad X \in \mathcal{A}_2 \\ &\Leftrightarrow\; \forall \mathbb{Q}_1 \in \mathcal{P}_1 : \; \mathbb{Q}_1[X] \geq 0 \text{ and } \forall \mathbb{Q}_2 \in \mathcal{P}_2 : \; \mathbb{Q}_2[X] \geq 0 \\ &\Leftrightarrow\; \forall \mathbb{Q} \in \mathcal{P} : \quad \mathbb{Q}[X] \geq 0 \end{aligned}$$

We now prove that $\mathcal{P} = conv(\mathcal{P}_1, \mathcal{P}_2)$ is closed.

Let $(Y_n)_n$ be a sequence in $\mathcal{P}$ converging in L^1 norm to a certain $\mathbb{R}$ (remember that we identify probabilities with their densities). By definition there exist $\mathbb{P}_n \in \mathcal{P}_1$ and $\mathbb{Q}_n \in \mathcal{P}_2$ and $t_n \in [0,1]$ such that $Y_n = t_n\mathbb{P}_n + (1-t_n)\mathbb{Q}_n$. We may suppose that $t_n \to t \in [0,1]$ (if not, extract a converging subsequence).

There are now two possible cases:
a) if t_n or $1-t_n$ tend to 0, then we have either $\mathbb{Q}_n \to \mathbb{R}$ or $\mathbb{P}_n \to \mathbb{R}$ and then $\mathbb{R} \in \mathcal{P}_1$ or $\mathbb{R} \in \mathcal{P}_2$.
b) $0 < t < 1$. By dropping a finite number of terms, we may suppose that there is a number $c \in (0,1)$ such that $c \le t_n \le 1-c$. Now:

$$\mathbb{P}_n[A] \le \frac{1}{t_n} Y_n[A] \le \frac{1}{c}(Y_n[A])$$

and therefore the sequence $\left(\frac{d\mathbb{P}_n}{d\mathbb{P}}\right)_n$ is dominated by the strongly convergent sequence $\left(\frac{dY_n}{d\mathbb{P}}\right)_n$. It is therefore uniformly integrable and hence a relatively weakly compact sequence. We may, by selecting a subsequence, suppose that $\mathbb{P}_n \to \mathbb{P}_o$ weakly and since $\mathcal{P}_1$ is convex closed, we have $\mathbb{P}_o \in \mathcal{P}$. Similarly we get $\mathbb{Q}_o \in \mathcal{P}_2$. Finally $\mathbb{R} = t\mathbb{P}_o + (1-t)\mathbb{Q}_o$ belongs to $conv(\mathcal{P}_1, \mathcal{P}_2)$. □

ii) Convex Convolution of risk measures.

With the previous notation, if ρ_1 and ρ_2 are given, with their $\mathcal{A}_1$, $\mathcal{P}_1$ and $\mathcal{A}_2$, $\mathcal{P}_2$, we can construct another risk measure ρ by taking $\mathcal{P} = \mathcal{P}_1 \cap \mathcal{P}_2$ and $\mathcal{A} = \overline{conv\,(\mathcal{A}_1, \mathcal{A}_2)}^{\sigma(L^\infty, L^1)}$. We must show that $\mathcal{A}$ and $\mathcal{P}$ correspond:

Proposition 1 *$\mathcal{A}$ and $\mathcal{P}$ correspond, i.e. $\mathcal{P} = \mathcal{P}_\mathcal{A}$, where*

$$\mathcal{P}_\mathcal{A} := \{\mathbb{Q} \mid \ \mathbb{Q} \text{ a probability such that } \forall X \in \mathcal{A} : \quad \mathbb{Q}[X] \ge 0\}.$$

Proof. We first show that $\mathcal{P} \supseteq \mathcal{P}_\mathcal{A}$. If $X \notin \mathcal{A}$ then by the Hahn Banach theorem (remember that the dual space of L^∞ with the weak* topology is exactly L^1) there exists an $f \in L^1$ such that $\mathbb{E}[fX] < 0$ and $\mathbb{E}[fY] \ge 0$ for every $Y \in \mathcal{A}$. Since $\mathcal{A}$ contains I_A for every $A \in \mathcal{F}$, f will be almost surely nonnegative. Now, f can be assumed to be already normalized, so we have obtained a $\mathbb{Q} \in \mathcal{P}_1 \cap \mathcal{P}_2$ which is strictly negative on X.

Now we show that $\mathcal{P} \subseteq \mathcal{P}_\mathcal{A}$. If $X \in \mathcal{A}$ we have to prove that $\mathbb{Q}[X] \ge 0$ for every $\mathbb{Q} \in \mathcal{P}$. Let us start with $X \in conv\,(\mathcal{A}_1, \mathcal{A}_2) = \mathcal{A}_1 + \mathcal{A}_2$, where the equality holds because the $\mathcal{A}_i$ are convex cones. Then if $\mathbb{Q} \in \mathcal{P}$, $\mathbb{Q}$ belongs to both $\mathcal{P}_i$ and taking into account that X can be written as $Y + Z$ with $Y \in \mathcal{A}_1$ and $Z \in \mathcal{A}_2$, we have that $\mathbb{Q}[X] \ge 0$. Rewritten, this means that $0 \le \int X \frac{d\mathbb{Q}}{d\mathbb{P}}$ for every $X \in conv\,(\mathcal{A}_1, \mathcal{A}_2)$ and for every $\mathbb{Q} \in \mathcal{P}$. By fixing

$\mathbb{Q}$, the set $\{Y \in L^\infty \,|\, Q[Y] \geq 0\}$ is weakly* closed and contains $conv\,(\mathcal{A}_1, \mathcal{A}_2)$: therefore it contains the weak* closure of the latter set, that is, it contains $\mathcal{A}$. □

Remark 7 In case $\mathcal{P}_1 \cap \mathcal{P}_2 = \emptyset$ we get that $\mathcal{A}_1 + \mathcal{A}_2$ is dense in L^∞ for the weak* topology $\sigma(L^\infty, L^1)$. We will see by an example that this does not imply that $L^\infty = \mathcal{A}_1 + \mathcal{A}_2$.

Proposition 2 *Suppose $\mathcal{P}_1 \cap \mathcal{P}_2 \neq \emptyset$. Let $\bar{\rho}$ have the Fatou property and let it be smaller than ρ_1 and ρ_2: then $\bar{\rho} \leq \rho_{\mathcal{P}_1 \cap \mathcal{P}_2}$.*

Proof. Let $\bar{\rho}$ be given by $\overline{\mathcal{P}}$. Then $\overline{\mathcal{P}} \subset \mathcal{P}_1$ and $\overline{\mathcal{P}} \subset \mathcal{P}_2$ because $\bar{\rho} \leq \rho_1$ and $\bar{\rho} \leq \rho_2$. Therefore $\overline{\mathcal{P}} \subset \mathcal{P}_1 \cap \mathcal{P}_2$ and hence $\bar{\rho} \leq \rho_{\mathcal{P}_1 \cap \mathcal{P}_2}$. □

If now we would like to define a coherent risk measure $\tilde{\rho}$, with the property that it is the biggest coherent measure such that $\tilde{\rho} \leq \rho_1 \wedge \rho_2$, we can take the following construction:

$$\begin{aligned} \tilde{\rho}(X) &= \inf\{t\rho_1(X_1) + (1-t)\rho_2(X_2) \,|\, X = tX_1 + (1-t)X_2 \quad 0 \leq t \leq 1\} \\ &= \inf\{\rho_1(tX_1) + \rho_2((1-t)X_2) \,|\, X = (tX_1) + ((1-t)X_2)\} \\ &= \inf\{\rho_1(Y) + \rho_2(X - Y) \,|\, Y \in L^\infty\} \end{aligned}$$

This risk measure is usually denoted as $\rho_1 * \rho_2$ and it is called the convex convolution of ρ_1 and ρ_2. The convex convolution can be characterized using the duality $(L^\infty, \mathbf{ba})$. So let us introduce $\mathcal{P}_1^{\mathbf{ba}} = \{\mu \in \mathbf{ba} \mid \mu \geq 0\,,\, \mu(\Omega) = 1,\, \forall f \in \mathcal{A}_1\ \mathbb{E}_\mu[f] \geq 0\}$ and similarly we have $\mathcal{P}_2^{\mathbf{ba}}$. The reader can check that $\rho_1 * \rho_2$ is defined using the set $\mathcal{P}_1^{\mathbf{ba}} \cap \mathcal{P}_2^{\mathbf{ba}}$. The acceptance set of $\rho_1 * \rho_2$ is $\overline{conv(\mathcal{A}_1, \mathcal{A}_2)}$ where the closure is taken for the **norm** topology of L^∞. Also $\mathcal{P}_1^{\mathbf{ba}} = \overline{\mathcal{P}_1}$ where the closure is taken using the $\sigma(\mathbf{ba}, L^\infty)$ topology. This equality is in fact a restatement of the Fatou property. So we obtain that $\rho_1 * \rho_2$ has the Fatou property if and only if the following holds (where the bar indicates $\sigma(\mathbf{ba}, L^\infty)$ closure):

$$\overline{\mathcal{P}_1^{\mathbf{ba}} \cap \mathcal{P}_2^{\mathbf{ba}} \cap L^1} = \mathcal{P}_1^{\mathbf{ba}} \cap \mathcal{P}_2^{\mathbf{ba}}\,.$$

This is equivalent to: $\overline{\mathcal{P}_1 \cap \mathcal{P}_2} = \overline{\mathcal{P}_1} \cap \overline{\mathcal{P}_2}$, where again the bar indicates $\sigma(\mathbf{ba}, L^\infty)$ closure. So we get:

Proposition 3 *$\rho_{\mathcal{P}_1 \cap \mathcal{P}_2}$ and $\rho_1 * \rho_2$ coincide if and only if $\rho_1 * \rho_2$ has the Fatou property. This is the case when for instance $\mathcal{P}_1$ (or $\mathcal{P}_2$) is weakly compact.*

Example 9 Let $(A_n)_{n \geq 1}$ be a measurable partition of Ω into sets with $\mathbb{P}[A_n] > 0$. For each n, let e_n be the measure with density $\frac{I_{A_n}}{\mathbb{P}[A_n]}$. The sets $\mathcal{P}_1$ and $\mathcal{P}_2$ are defined as follows:

$$\mathcal{P}_1 = \overline{conv}\,(e_1, e_3, e_4, \dots)$$

$$\mathcal{P}_2 = \overline{conv}\,(e_1, \frac{e_2 + ne_n}{1+n}, n \geq 3)\,.$$

Clearly, $\mathcal{P}_1 \cap \mathcal{P}_2 = \{e_1\}$ and $\mathcal{P}_1^{\mathbf{ba}} \cap \mathcal{P}_2^{\mathbf{ba}}$ contains, besides the vector e_1, the adherent points in **ba** of the sequence $(e_n, n \geq 1)$. The measure $\rho_1 * \rho_2$ is therefore not the same as $\rho_{\mathcal{P}_1 \cap \mathcal{P}_2}$ and $\rho_1 * \rho_2$ does not have the Fatou property.

Example 10 We take the same sequences as in the previous example but this time we define:

$$\mathcal{P}_1 = \overline{conv}\,(e_3, e_4, \dots)$$

$$\mathcal{P}_2 = \overline{conv}\,(\frac{e_2 + ne_n}{1+n}, n \geq 3)\,.$$

Clearly, $\mathcal{P}_1 \cap \mathcal{P}_2 = \emptyset$ and $\mathcal{A}_1 + \mathcal{A}_2$ is $\sigma(L^\infty, L^1)$ dense in L^∞. However, $\mathcal{A}_1 + \mathcal{A}_2$ is not norm dense in L^∞, since $\overline{\mathcal{P}_1} \cap \overline{\mathcal{P}_2} \neq \emptyset$. Indeed, $\mathcal{P}_1$ and $\mathcal{P}_2$ are closed convex sets that cannot be strictly separated.

Example 11 We consider a finite Ω (to avoid topological difficulties) and we suppose that the regulator said the positions $Y_1, \dots, Y_n$ are acceptable. In this context positions are just vectors in $\mathbb{R}^{|\Omega|}$. The minimal convex cone $\mathcal{A}_i$ containing $L^\infty_+ = \{X \geq 0\}$ and Y_i is the set $\{X + \lambda Y_i \,|\, \lambda \geq 0\,,\, X \geq 0\}$: the purpose is to construct a risk measure under which each of the originally given positions $(Y_i)_{i=1}^n$, is still acceptable. Therefore we take $\mathcal{A} = conv\,(\mathcal{A}_i\,, i \leq n)$ so that our risk measure ρ will be $\rho_{(\cap_i \mathcal{P}_i)} = \rho_1 * \dots * \rho_n$.

We have:

$$\begin{aligned}
\rho(X) &= \inf \left\{ \rho_1(X_1) + \dots \rho_n(X_n) \,|\, X = \sum_{i=1}^n X_i \right\} \\
&= \inf \left\{ \alpha_1 + \dots + \alpha_n \mid \exists \lambda_i \in \mathbb{R}^+ \, \exists f_i \in \mathbb{R}_+^{|\Omega|} \quad \alpha_i + X_i = f_i + \lambda_i Y_i\,,\, X = \sum_{i=1}^n X_i \right\} \\
&= \inf \left\{ \alpha \mid X + \alpha \geq \sum_{i=1}^n \lambda_i Y_i \quad \text{where} \quad \lambda_i \geq 0 \right\}
\end{aligned}$$

We notice that the specification of the values of $\rho(Y_i)$ is not required and that the risk measure can be equal to $-\infty$ (which is the case if $\cap_{i \leq n} \mathcal{P}_i = \emptyset$).

The problem of calculating $\rho(X)$ can be restated as a linear program:

$$\begin{cases} \max_{\mathbb{Q}} \mathbb{Q}[-X] \\ \sum_\omega Q(\omega) = 1 \text{ , } \mathbb{Q}(\omega) \geq 0 \\ \mathbb{Q}[Y_i] \geq 0 \end{cases}$$

and the preceeding equality is the usual dual-primal linear program relation.

iii) Product of risk measures.

Let $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ and $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ be two probability spaces. We consider the product space $(\Omega, \mathcal{F}, \mathbb{P}) = (\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \mathbb{P}_1 \otimes \mathbb{P}_2)$ and we would like to define a risk measure, the most liberal one, given two measures of risk ρ_1 and ρ_2 defined on Ω_1 and Ω_2 respectively.

For a probability measure $\mathbb{Q}$ on Ω, we define $\mathbb{Q}_1$ and $\mathbb{Q}_2$ to be the marginal probabilities of $\mathbb{Q}$ on Ω_1 and Ω_2 (that is, $\mathbb{Q}_1[A_1] = \mathbb{Q}[A_1 \times \Omega_2]$ and similarly for $\mathbb{Q}_2$). If as usual $\mathcal{P}_i$ and $\mathcal{A}_i$ represent the family of probabilities and the set of acceptable positions for ρ_i, we define:

$$\begin{aligned} \tilde{\mathcal{P}}_1 &= \{\mathbb{Q} \mid \mathbb{Q} \ll \mathbb{P} \,;\, \mathbb{Q}_1 \in \mathcal{P}_1\} \\ \tilde{\mathcal{P}}_2 &= \{\mathbb{Q} \mid \mathbb{Q} \ll \mathbb{P} \,;\, \mathbb{Q}_2 \in \mathcal{P}_2\} \,. \end{aligned}$$

We suppose for simplicity that ρ_1 and ρ_2 are relevant, the general case is left to the reader. If $f \in \mathcal{A}_1$, a "reasonable" request is that $\tilde{f}(\omega_1, \omega_2) = f(\omega_1)$ should be acceptable; the same should hold for $g \in \mathcal{A}_2$. So we put

$$\begin{aligned} \tilde{\mathcal{A}}_1 &= \{f + h \mid f \in \mathcal{A}_1 \,,\ h \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}),\, h \geq 0\} \\ \tilde{\mathcal{A}}_2 &= \{g + h \mid g \in \mathcal{A}_2 \,,\ h \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}),\, h \geq 0\} \\ \mathcal{A} &= \tilde{\mathcal{A}}_1 + \tilde{\mathcal{A}}_2 \,. \end{aligned}$$

The set $\mathcal{A}$ clearly is a convex cone. However, it is also $\sigma(L^\infty, L^1)$ closed. To see this we again use the Krein-Smulian theorem. So let us suppose $(\phi_n)_n \in \mathcal{A}$, $\|\phi_n\|_\infty \leq 1$ and $\phi_n \xrightarrow{\mathbb{P}} \phi$. We have to show that $\phi \in \mathcal{A}$. Each ϕ_n can be written as $\phi_n = f_n + g_n + h_n$, where $f_n \in \mathcal{A}_1$, $g_n \in \mathcal{A}_2$ and $h_n \geq 0$. Take $\mathbb{Q}_1 \in \mathcal{P}_1$, $\mathbb{Q}_2 \in \mathcal{P}_2$, $\mathbb{Q}_1 \sim \mathbb{P}_1$, $Q_2 \sim \mathbb{P}_2$ and let $\mathbb{Q} = \mathbb{Q}_1 \otimes \mathbb{Q}_2$. Of course, $\mathbb{Q} \in \tilde{\mathcal{P}}_1 \cap \tilde{\mathcal{P}}_2$.
Furthermore $\mathbb{Q}_1$ and $\mathbb{Q}_2$ are the marginal probabilities of $\mathbb{Q}$, so that there is no conflict in the notation. We clearly have $1 \geq \mathbb{E}_{\mathbb{Q}}[f_n + g_n + h_n] \geq \mathbb{E}_{\mathbb{Q}_1}[f_n] + \mathbb{E}_{\mathbb{Q}_2}[g_n]$.
Both terms are nonnegative since $f_n \in \mathcal{A}_1$ and $g_n \in \mathcal{A}_2$. Therefore, $\mathbb{E}_{\mathbb{Q}_1}[f_n]$ and $\mathbb{E}_{\mathbb{Q}_2}[g_n]$

are between 0 and 1. We may and do suppose that $\mathbb{E}_{\mathbb{Q}_1}[f_n]$ and $\mathbb{E}_{\mathbb{Q}_2}[g_n]$ converge (if not, we take a subsequence). Since $f_n + g_n + h_n \leq 1$, we also get $f_n + g_n \leq 1$ and hence $f_n + \mathbb{E}_{\mathbb{Q}_2}[g_n] \leq 1$. Indeed for $\mathbb{Q}$, f_n and g_n are independent and the inequality results by taking conditional expectation with respect to $\mathcal{F}_1 \otimes \{\emptyset, \Omega_2\}$. Since $\mathbb{E}_{\mathbb{Q}_2}[g_n] \geq 0$, we get $f_n \leq 1$. Similarly, we get $g_n \leq 1$. We now replace f_n and g_n by respectively $f_n \vee (-2)$ and $g_n \vee (-2)$. Necessarily we have $f_n \vee (-2) \geq f_n$ and therefore $f_n \vee (-2) \in \mathcal{A}_1$, also $g_n \vee (-2) \in \mathcal{A}_2$. But this requires a correction of h_n. So we get:

$$\phi_n = f_n \vee (-2) + g_n \vee (-2) + h_n - (-2 - f_n)^+ - (-2 - g_n)^+ .$$

The function $h_n - (-2 - f_n)^+ - (-2 - g_n)^+$ is still nonnegative. To see this, we essentially have the following two cases.

On the set $\{f_n < -2\} \cap \{g_n < -2\}$ we have:

$$h_n - (-2 - f_n)^+ - (-2 - g_n)^+ = (h_n + f_n + g_n) + 4 \geq -1 + 4 > 0 .$$

On the set $\{f_n \geq -2\} \cap \{g_n < -2\}$ we have:

$$\begin{aligned} h_n - (-2 - f_n)^+ - (-2 - g_n)^+ &= h_n + 2 + g_n \\ &= (h_n + f_n + g_n) + (2 - f_n) \\ &\geq -1 + 1 \geq 0 \end{aligned}$$

The other cases are either trivial or similar.

So we finally may replace the functions as indicated and we may suppose that $\phi_n = f_n + g_n + h_n$, where $-2 \leq f_n \leq 1$, $-2 \leq g_n \leq 1$, $f_n \in \mathcal{A}_1$, $g_n \in \mathcal{A}_2$, $h_n \geq 0$.

Since the sequences $(f_n)_n$ and $(g_n)_n$ are bounded, we can take convex combination of them, (still denoted by the same symbols), that converge in probability. So finally we get $f_n \to f$, $g_n \to g$ in probability. Of course this implies $f \in \mathcal{A}_1$ and $g \in \mathcal{A}_2$. But then we necessarily have that $h_n = \phi_n - f_n - g_n$ converges in probability, say to a function h. Of course, $h \geq 0$. So finally we get $\phi = f + g + h$ with $f \in \mathcal{A}_1$, $g \in \mathcal{A}_2$ and $h \geq 0$.

The polar cone of $\mathcal{A}$ can be described by the sets $\tilde{\mathcal{P}}_1$ and $\tilde{\mathcal{P}}_2$. Indeed

$$\begin{aligned} \mathcal{P} &= \{\mathbb{Q} \mid \mathbb{Q} \text{ a probability and } \forall u \in \mathcal{A} \quad \mathbb{Q}[u] \geq 0\} \\ &= \{\mathbb{Q} \mid \mathbb{Q} \text{ a probability and } \forall f \in \mathcal{A}_1 \quad \mathbb{Q}[f] \geq 0 \text{ and } \forall g \in \mathcal{A}_2 \quad \mathbb{Q}[g] \geq 0\} \\ &= \{\mathbb{Q} \mid \mathbb{Q} \text{ a probability and } \forall f \in \mathcal{A}_1 \quad \mathbb{Q}_1[f] \geq 0 \text{ and } \forall g \in \mathcal{A}_2 \quad \mathbb{Q}_2[g] \geq 0\} \\ &= \{\mathbb{Q} \mid \mathbb{Q}_1 \in \mathcal{P}_1 \,,\ \mathbb{Q}_2 \in \mathcal{P}_2\} \\ &= \tilde{\mathcal{P}}_1 \cap \tilde{\mathcal{P}}_2. \end{aligned}$$

Moreover

$$\begin{aligned}\rho(X) &= \inf\{\alpha \mid \alpha + X \in \mathcal{A}\} \\ &= \inf\{\alpha \mid \exists f \in \mathcal{A}_1\,,\, \exists g \in \mathcal{A}_2 \quad X + \alpha \geq f + g\}.\end{aligned}$$

The previous lines also imply that the sets $\tilde{\mathcal{A}}_1$ and $\tilde{\mathcal{A}}_2$ are $\sigma(L^\infty, L^1)$ closed. Their polars are precisely given by $\tilde{\mathcal{P}}_1$ and $\tilde{\mathcal{P}}_2$ respectively. Indeed:

$$\{\mathbb{Q} \mid \mathbb{Q} \text{ proba. and } \forall u \in \tilde{\mathcal{A}}_1 \quad \mathbb{Q}[u] \geq 0\} = \{\mathbb{Q} \mid \mathbb{Q} \text{ proba. and } \forall u \in \mathcal{A}_1 \quad \mathbb{Q}[u] \geq 0\}$$

and the latter is equal to $\{\mathbb{Q} \mid \forall f \in \mathcal{A}_1 \quad \mathbb{Q}_1[u] \geq 0\}$, which is exactly $\tilde{\mathcal{P}}_1$. Therefore we get that: $\tilde{\mathcal{A}}_1 = \{\phi \mid \forall \mathbb{Q} \in \tilde{\mathcal{P}}_1 \quad \mathbb{Q}[\phi] \geq 0\}$.

Remark 8 Even if $\mathcal{P}_1$ and $\mathcal{P}_2$ consist of a single point, the family $\mathcal{P}$ can be very "big". For instance, let's take $\Omega_1 = \Omega_2 = T$, where T is the one dimensional torus (that is, S^1). On T we consider the Borel σ-algebra and as reference probability we take the normalized Lebesgue measure m, while $\mathcal{P}_1$ and $\mathcal{P}_2$ will coincide with $\{m\}$. If we take the product space $T \times T$ and we consider the set $A_\epsilon = \{(e^{i\theta}, e^{i\phi}) \mid |e^{i\theta} - e^{i\phi}| \leq \epsilon\}$ then $\lim_{\epsilon \to 0} m(A_\epsilon) = 0$; and by taking $\mathbb{Q}_\epsilon$ equal to the uniform distribution on A_ϵ we have that Q_ϵ belongs to $\mathcal{P}$, for each ϵ. But the family $(\mathbb{Q}_\epsilon)_\epsilon$ is not uniformly integrable: therefore $\mathcal{P}$ is not at all small, it isn't even weakly compact! It is still an unsolved problem to characterize the extreme points of the convex set of measures on $T \times T$ so that the marginals are m.

5 Convex games

The aim of this section is to investigate the relations between convex games and coherent risk measures. We start with a couple of definitions.

Definition 6 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A convex game on Ω is a function $w : \mathcal{F} \to [0, 1]$ such that:*

$$w(\Omega) = 1$$

$$w(A) + w(B) \geq w(A \cap B) + w(A \cup B)$$

$$\mathbb{P}(A) = 0 \textit{ implies } w(A) = 0$$

$$\textit{if } A_n \uparrow A \textit{ then } w(A_n) \uparrow w(A)\,.$$

This definition is not quite standard. Usually, $v(A) = w(\Omega) - w(A^c)$ is used as the characteristic function of a game. The reader can translate the above properties into the language of v. The last property is a continuity property which will enable us to use the duality (L^1, L^∞) instead of $(L^\infty, \mathbf{ba})$.

For properties of convex games, we refer to Schmeidler, **Schm1**, and Delbaen, **D1**.

Definition 7 *For a convex game w we define the core of w as*

$$\mathcal{C}(w) = \{\mu \mid \mu \in \mathbf{ba}\,,\, \mu \geq 0, \mu(\Omega) = 1 \text{ and } \forall A \quad \mathbb{Q}(A) \leq w(A)\}\,.$$

The $\sigma-$core is defined as:

$$\mathcal{C}^\sigma(w) = \{\mu \mid \mu \in L^1\,,\, \mu \geq 0\,,\, \mu[\Omega] = 1 \text{ and } \forall A \quad \mathbb{Q}(A) \leq w(A)\}\,.$$

Standard results on convex games allow us to write, for $X \in L^\infty_+$, the following equality:

$$\sup\{\mathbb{E}_\mu[X] \mid \mu \in \mathcal{C}\} = \int_0^\infty w(X > \alpha)\, d\alpha\,.$$

This relation, basically due to Choquet, **Ch**, can be found in Schmeidler **Schm1** and Delbaen, **D1**. Now, if we denote the right hand side by $\rho(-X)$, we can extend ρ to a function on L^∞. Indeed for any $X \in L^\infty$ and any β such that $X + \beta \geq 0$ we can see that:

$$\rho(-(X+\beta)) - \beta = \int_0^\infty w(X + \beta > \alpha)\, d\alpha\ - \beta$$

does not depend on the particular choice of β and the quantity always equals $\sup\{\mathbb{E}_\mu[X] \mid \mu \in \mathcal{C}\}$. Therefore ρ is a coherent risk measure. Because of the continuity property of w we get that $0 \leq X_n \leq 1$ and $X_n \uparrow X$ imply that $\int_0^\infty w(X_n > \alpha)\, d\alpha$ tends to $\int_0^\infty w(X > \alpha)\, d\alpha$ and hence ρ satisfies the Fatou property. It is therefore defined by a closed convex set $\mathcal{P} \subset L^1$ (remember that the set $\mathcal{P}$ is defined as $\{\mathbb{Q} \mid \mathbb{Q}$ probability and $\mathbb{E}_\mathbb{Q}[-f] \leq \rho(f)$ for all $f \in L^\infty\})$. The relation $w(A) = \rho(-I_A) \geq \mathbb{Q}(A)$ for each $A \in \mathcal{F}$ implies that $\mathcal{P} \subset \mathcal{C}^\sigma(w)$.

So we get, for $X \geq 0$:

$$\begin{aligned}\rho(-X) &= \sup\{\mathbb{E}_\mu[X] \mid \mu \in \mathcal{C}(w)\}\\ &= \sup\{\mathbb{E}_\mathbb{Q}[X] \mid \mathbb{Q} \in \mathcal{P}\}\,.\end{aligned}$$

But this means that $\mathcal{P}$ is $\sigma(\mathbf{ba}, L^\infty)$ dense in $\mathcal{C}(w)$ and finally, because $\mathcal{P}$ is convex closed in L^1, that $\mathcal{P} = \mathcal{C}(w) \cap L^1 = \mathcal{C}^\sigma(w)$. It follows that $\mathcal{C}^\sigma(w)$ is non-empty.

Theorem 11 *If $A_1 \subset A_2 \ldots \subset A_n$ is a finite nondecreasing family then there exists $\mathbb{Q} \in \mathcal{C}^\sigma(w)$ with $\mathbb{Q}(A_i) = w(A_i)$ for all $i \le n$.*

Proof. The proof of this theorem is not easy. It relies on the theorem of Bishop-Phelps. We take $X = \sum_{i=1}^n I_{A_i}$: this X belongs to L^∞ and we consider $0 < \epsilon < \frac{1}{8}$. By the Bishop-Phelps theorem there is $Y \in L^\infty$, with $\|X - Y\|_\infty < \epsilon$ and Y attains its supremum on $\mathcal{P}$. Of course we may replace Y by $Y + \epsilon$ and hence we get $Y \ge 0$. This means that there exists $\mathbb{Q}^o \in \mathcal{P}$ such that:

$$\mathbb{E}_{\mathbb{Q}^o}[Y] = \sup\{\mathbb{E}_{\mathbb{Q}}[Y] \mid \mathbb{Q} \in \mathcal{P}\} = \int_0^\infty w(Y > \alpha)\, d\alpha$$

This also implies $\int_0^\infty \mathbb{Q}^o(Y > \alpha)\, d\alpha = \int_0^\infty w(Y > \alpha)\, d\alpha$. But since $\mathbb{Q}^o \in \mathcal{P}$ we have $\mathbb{Q}^o(Y > \alpha) \le w(Y > \alpha)$ and therefore we obtain that for almost every α we necessarily have $\mathbb{Q}^o(Y > \alpha) = w(Y > \alpha)$. Now for each $0 \le k < n$ we take $k + \frac{1}{4} < \alpha < k + \frac{3}{4}$ with the above property and, since for such α we necessarily have $\{Y > \alpha\} = A_{k+1}$, we get $\mathbb{Q}^o[A_{k+1}] = w(A_{k+1})$ for $k = 0 \ldots n-1$. □

Remark 9 The conclusion of the theorem was already known for $\mu \in \mathcal{C}(w)$ (see Delbaen, **D1**). That $\mathbb{Q}$ can be chosen in $\mathcal{C}^\sigma(w)$ is new.

Example 12 If $0 < \beta < 1$ then $w(A) = \mathbb{P}(A)^\beta$ defines a convex game. If $\beta = 0$, the σ-core is the whole family of absolutely continuous probabilities, whereas if $\beta = 1$, $\mathcal{C}^\sigma$ is the singleton $\{\mathbb{P}\}$. We also have for nonnegative X: $\rho(-X) = \int_0^\infty \mathbb{P}(\{X > x\})^\beta\, dx$. These risk measures were introduced by Delbaen, **D2**. They are related to the so-called Lorentz spaces, a generalisation of the more familiar Orlicz spaces.

Example 13 More generally, we may show that if $f : [0,1] \to [0,1]$ is a nondecreasing concave function such that $f(0) = 0$ and $f(1) = 1$, then $w(A) = f(\mathbb{P}(A))$ defines a convex game. The set $\mathcal{C}^\sigma$ is weakly compact iff f is continuous at 0. An example of such a function is:

$$f(x) = \begin{cases} xk \text{ for } x \le \frac{1}{k} \\ 1 \text{ for } \frac{1}{k} \le x \le 1\,, \end{cases}$$

where of course $k \ge 1$. The reader can check that $\mathcal{C}^\sigma$ is $\mathcal{P}_k$ of example 4 in section 4. The sets $\mathcal{P}_{p,k}$ cannot be obtained via convex games, the related risk measures are not comonotone, see **D2** for a proof.

Let us now give the precise definition of comonotonicity.

Definition 8 *Two random variables X, Y, defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ are comonotone if on the product space*

$$(\Omega \times \Omega, \mathcal{F} \otimes \mathcal{F}, \mathbb{P} \otimes \mathbb{P})$$

the random variable $Z(\omega_1, \omega_2) = (X(\omega_1) - X(\omega_2))(Y(\omega_1) - Y(\omega_2))$ is nonnegative.

Example 14 If we take a position X and two increasing functions $f, g : \mathbb{R} \to \mathbb{R}$, then the variables $f(X)$ and $g(X)$ are comonotone.

.

Remark 10 If X and Y are comonotone and square integrable, the covariance $Cov\,(X, Y)$ is nonnegative.

Proof. By integrating Z on the product space we get:

$$0 \leq \int Z\, d(\mathbb{P} \otimes \mathbb{P}) = 2\left(\int XY\, d\mathbb{P} \;-\; \int X\, d\mathbb{P} \int Y\, d\mathbb{P}\right).$$

$\square$

According to its behaviour with respect to comonotone variables, a risk measure ρ is called comonotone if $\rho(X + Y) = \rho(X) + \rho(Y)$ for every comonotone couple X, Y. As regards the relation between convex games and risk measures, we state the following important result. For a proof we refer to Schmeidler's paper, **Schm2**.

Theorem 12 *(Schmeidler 1986) A coherent risk measure ρ originates from a convex game (that is $\mathcal{P}$ is the σ-core of a convex game) iff ρ is comonotone with the Fatou property.*

6 Relation with VaR

In this section we focus on the relation between risk measures and VaR. We recall that if α belongs to the interval $(0, 1)$ the family $\mathcal{P}_{1/\alpha} = \{\mathbb{Q} \mid \frac{d\mathbb{Q}}{d\mathbb{P}} \leq 1/\alpha\}$ is well defined; the corresponding ρ_α is such that if $\mathbb{P}$ is atomless and if the distribution of X is continuous, then $\rho_\alpha(X) = \mathbb{E}[-X \mid X \leq q_\alpha(X)]$ (Recall that $q_\alpha(X)$ is the α-quantile of X).

The measure ρ_α is minimal in the sense that it is the minimum in the class of coherent risk measures satisfying the Fatou property, only depending on the distribution and bigger than VaR_α. That is, the following theorem holds:

Theorem 13 *Suppose that $\mathbb{P}$ is atomless; let ρ be a coherent risk measure with the Fatou property and verifying the additional property that if X and Y are identically distributed, then $\rho(X) = \rho(Y)$. If for every $X \in L^\infty$, $\rho(X)$ is bigger than* $\mathrm{VaR}_\alpha(\mathrm{X})$*, then $\rho \geq \rho_\alpha$.*

Proof. We first prove that for every X, $\rho(X) \geq \mathbb{E}[-X \mid X \leq q_\alpha(X) + \epsilon]$. Let $A = \{\omega \mid X(\omega) \leq q_\alpha(X) + \epsilon\}$, by definition of q_α we have that $\mathbb{P}[A] > \alpha$. Let Y be the random variable coinciding with X on A^c and with the number $\mathbb{E}[X \mid A]$ on A. $\mathrm{VaR}_\alpha(Y)$ is then equal to $\mathbb{E}[-X \mid A]$ and since $\frac{I_A}{\mathbb{P}[A]}$ is a density in $\mathcal{P}_{\frac{1}{\alpha}}$ we deduce from $\mathrm{VaR}_\alpha(Y) \leq \rho(Y)$, that $\mathbb{E}[-X \mid A] \leq \rho(Y)$. Now, let's call ν the distribution of X given A (with the inherited σ algebra). The hypothesis of the absence of atoms on Ω implies in particular the absence of atoms on A. This fact guarantees the existence on A of a sequence of independent, ν-distributed random variables. Let us denote by X_n the random variable coinciding with the n-th element of such an i.i.d. sequence on A and with X (and therefore with Y) outside A. The $(X_n)_n$ have the same distribution, equal to the distribution of X and by the law of large numbers, $\frac{X_1+\cdots+X_n}{n}$ converges almost surely to Y. Remembering that the Fatou property holds, we finally obtain:

$$\rho(Y) \leq \liminf_{n\to\infty} \rho\left(\frac{X_1 + \ldots X_n}{n}\right) \leq \liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \rho(X_i) \equiv \rho(X_1) \equiv \rho(X)$$

Thus we get $\rho(X) \geq \rho(Y) \equiv \mathbb{E}[-X \mid X \leq q_\alpha(X) + \epsilon]$. If X has a distribution function continuous at q_α, we can pass to the limit, obtaining $\rho(X) \geq \rho_\alpha(X)$. What if the distribution of X is not continuous? We can use the following approximation result:

Proposition 4 *If $\mathbb{P}$ is atomless and $X \in L^\infty$, there exists a sequence $(X_n)_n$ so that:*

1. *$X \leq X_n \leq X + \frac{1}{n}$;*
2. *$X_n \downarrow X$;*
3. *each X_n has a continuous distribution.*

Below we will give a sketch of this proposition. We first continue the proof of the theorem. Given X arbitrary, we can find an approximating sequence X_n as in the proposition and then we have that both $\rho(X_n)$ and $\rho_\alpha(X_n)$ tend to $\rho(X)$ and $\rho_\alpha(X)$ respectively (because coherent risk measures are continuous with respect to the uniform L^∞ topology). Passing to the limit in the already established inequality $\rho(X_n) \geq \rho_\alpha(X_n)$, gives $\rho(X) \geq \rho_\alpha(X)$ for all $X \in L^\infty$.

We now give a sketch of the proof of the previous proposition. The (obvious) details are left to the reader. Let $\{a_k \mid k \in \mathbb{N}\}$ be the discontinuity set of the distribution function F_X

of X and let U_k stand for the set $\{X = a_k\}$. Then $\mathbb{P}[U_k] > 0$ and for each k we can construct a variable $Y^k : U_k \to [0,1]$ with the uniform distribution under $\mathbb{P}[\cdot|U_k]$. This follows from the non-atomicity of the probability space. Now, let's define $X_n = X + \frac{1}{n}\sum_{k\geq 1} Y^k I_{U_k}$. It is easily seen that each X_n has a continuous distribution and that the sequence $(X_n)_n$ has the required properties.

Remark 11 Recently Kusuoka could characterize the coherent risk measures that are law invariant. His characterisation gives an alternative proof of the above result, see **Ku**.

As a general result, under the hypotheses of absence of atoms, there is no smallest coherent risk measure that dominates VaR. As usual we say that ρ dominates VaR if for all $Y \in L^\infty$ we have that $\rho(Y) \geq \mathrm{VaR}_\alpha(Y)$.

Theorem 14 *If $\mathbb{P}$ is atomless we have, for each $0 < \alpha < 1$, that*

$$\mathrm{VaR}_\alpha(X) = \inf\{\rho(X) \mid \rho \textit{ coherent with the Fatou property and } \rho \geq \mathrm{VaR}_\alpha\}\}$$

The theorem says that if we take the infimum over **all** risk measures that dominate VaR (and not only the ones depending just on distributions) we obtain VaR, wich as we already saw, is not a coherent risk measure (remember, it's not subadditive). We omit the rather technical proof of this theorem. The reader can find it in Delbaen, **D2**.

7 Application of coherent risk measures to mathematical finance

We follow the notation of Delbaen and Schachermayer, 1994, **DS**. So let $(\Omega, (\mathcal{F}_t)_{0\leq t}, \mathbb{P})$ be a filtered probability space and let $S : \mathbb{R}_+ \times \Omega \to \mathbb{R}^d$ be a càdlàg locally bounded, adapted process. We suppose that the set

$$M^a = \{\mathbb{Q} \mid \mathbb{Q} \text{ probability } \mathbb{Q} \ll \mathbb{P}\,,\ S \text{ is a } \mathbb{Q} \text{ local-martingale}\}$$

is non-empty. Since S is locally bounded, M^a is a closed convex subset of L^1. We also suppose that $\exists \mathbb{Q} \in M^a$, $\mathbb{Q} \sim \mathbb{P}$, which is equivalent to the no arbitrage property "no free lunch with vanishing risk". Now let $\mathcal{P}$ be a closed convex set defining the coherent risk measure ρ. We suppose that $\mathcal{P}$ is weakly compact. Before we compare the relation between $\mathcal{P}$ and M^a let us recall the following result of **DS**. If $X \in L^\infty$ then the quantity

$$p(X) = \sup_{\mathbb{Q}\in M^a} \mathbb{E}_\mathbb{Q}[X]$$

is called the superhedging price of X. If an investor would have $p(X)$ at his disposal, he would be able to find a strategy H so that $H\cdot S$ is bounded and so that $p(X)+(H\cdot S)_\infty \geq X$. This means that after having sold X for the price $p(X)$ he could, by cleverly trading, hedge out the risky position $-X$.

The quantity $p(X)$ is also the maximum price that can be charged for X. The minimum price is:

$$m(X) = \inf_{\mathbb{Q}\in M^a} \mathbb{E}_{\mathbb{Q}}[X]\,.$$

No agent would be willing to sell X for less than $m(X)$ and no agent would be willing to buy X for more than $p(X)$. We now look at two special cases:

(a) We suppose that for all X we have $\rho(X) \leq p(-X)$. This means that for any position X (after having sold $-X$) the necessary capital becomes smaller than the superhedging price of $-X$. This seems reasonable since with $p(-X)$ the selling agent can hedge out the risk. This requirement ($\forall X \in L^\infty\,;\ \rho(X) \leq p(-X)$) is, by the Hahn-Banach theorem, equivalent to $\mathcal{P} \subset M^a$.

(b) If $\mathcal{P} \cap M^a = \emptyset$ then, by weak compactness of $\mathcal{P}$, the Hahn Banach theorem gives us an element $-X \in L^\infty$ so that:

$$\sup_{\mathbb{Q}\in\mathcal{P}} \mathbb{E}_{\mathbb{Q}}[-X] < \inf_{\mathbb{Q}\in M^a} \mathbb{E}_{\mathbb{Q}}[-X]\,.$$

This means that having sold $-X$ the position X requires a capital equal to $\rho(X)$ but this capital is less than the minimum quantity for which $-X$ can be sold. In such a case a regulator, requiring $\rho(X)$, seems to have no understanding of the financial markets.

We can push this analysis a little bit further. We first introduce some notation. Let W be the space $\{(H\cdot S)_\infty \mid H\cdot S \text{ bounded}\}$. It can be shown, see **DS**, that W is a weak* closed subspace of L^∞ and of course $W^\perp = \{f \mid \mathbb{E}[f(H\cdot S)_\infty] = 0 \quad \text{for all } f \in W\}$. Clearly M^a is the intersection of $W^\perp$ with the set of probability measures. This means that $M^a \cap \mathcal{P} = \emptyset$ is equivalent to $W^\perp \cap \mathcal{P} = \emptyset$. By the Hahn-Banach theorem there exists $X \in L^\infty$ so that $\mathbb{E}[Xf] = 0$ for all $f \in W^\perp$ and $\inf_{X\in\mathcal{P}} \mathbb{E}_{\mathbb{Q}}[X] > 0$. But of course this means $X \in W^{\perp\perp} = W$ and hence there is a strategy H so that $(H\cdot S)$ is bounded, $(H\cdot S)_\infty = X$ and $\inf_{\mathbb{Q}\in\mathcal{P}} \mathbb{E}_{\mathbb{Q}}[X] > 0$. This means that there exists $X \in W$ with $-\sup_{\mathbb{Q}\in\mathcal{P}} \mathbb{E}_{\mathbb{Q}}[-X] > 0$, i.e. $\rho(X) < 0$. This would mean that a position X can be completely hedged (by the strategy H) at no cost and the controlling agent or supervisor allows to reduce the capital.

From this we deduce the following theorem.

Theorem 15 *If $\mathcal{P}$ is a weakly compact convex set of probability measures defining the coherent risk measure ρ, then $\mathcal{P} \cap M^a \neq \emptyset$ if and only if for all $X \in W$ we have $\rho(X) \geq 0$.*

We leave it to the reader to rephrase this condition for the examples $\mathcal{P}_k$, $\mathcal{P}_{p,k}$ of section 4. It leads to necessary and sufficient conditions for the existence of local martingale measures with bounded densities (or p-integrable densities).

In a similar way the condition $\mathcal{P} \subset M^a$ is equivalent to $\mathcal{P} \subset W^\perp$. Therefore $\rho(X) = 0$ for all $X \in W$. This means that something that can be replicated does not require extra capital.

8 Capital allocation problem

Let, as usual, $\rho : L^\infty \to \mathbb{R}$ be a coherent risk measure with the Fatou property. Imagine that a firm is divided in N trading units and let their future incoming money flow be denoted by $X_1, \ldots, X_N$, all belonging to L^∞. The total capital required to face the risk is $\rho(\sum_{i=1}^N X_i) = k$ and we have to find a "fair" way to allocate $k_1, \ldots, k_N$ so that $k_1 + \ldots + k_N = k$. Another point of view of the allocation problem is to distribute the gain of diversification over the different business units of a financial institution.

8.1 Game theoretic approach.

In the previous setting, we define $k_1, \ldots, k_N$ to be a fair allocation if:

1. $\sum_{i=1}^N k_i = \rho(\sum_{i=1}^N X_i)$
2. $\forall J \subseteq \{1, \ldots, N\}$ we have $\sum_{j \in J} k_j \leq \rho(\sum_{j \in J} X_j)$.

The existence of a fair allocation is in fact equivalent to the non-emptiness of the core of a "balanced" game. So it is no surprise that the following theorem uses the same technique as the Bondareva-Shapley theorem in game theory.

Theorem 16 *If ρ is coherent then there exists a fair allocation.*

Proof. Let $m = 2^N$ and let $\phi : \mathbb{R}^N \to \mathbb{R}^m$ be the following linear map:

$$\phi((k_i)_i) = \left(\left(\sum_{j \in J} k_j \right)_{\emptyset \neq J \subseteq \{1, \ldots, N\}}, \left(- \sum_{j=1}^N k_j \right) \right)$$

We have to find k so that $\phi(k)_J \leq \rho(\sum_{j\in J} X_j)$ and $\sum_{i\leq N} k_i = \rho(\sum_{i\leq N} X_i)$.

Let $P = \left\{((x_J)_J, x) \mid x_J \leq \rho\left(\sum_{j\in J} X_j\right),\ x \leq -\rho\left(\sum_{i\leq N} X_i\right)\right\}$. The problem is reduced to showing that $\phi(\mathbb{R}^N) \cap P$ is non empty. If it were empty, by the separating hyperplane theorem, there would be $((\alpha_J)_J, \alpha)$ such that:

1. $\sum_J \alpha_J \left(\sum_{j\in J} k_j\right) - \alpha \sum_{i\leq N} k_i = 0$;
2. $\sum_J \alpha_J\, \rho(\sum_{j\in J} X_j) - \alpha\rho(\sum_{i\leq N} X_i) < 0$;
3. $\alpha_J \geq 0,\ \alpha \geq 0$.

Condition *1* can be written as: for each j, we have $\sum_{J\ni j} \alpha_J = \alpha$. If $\alpha = 0$, then all the α_J would be 0 but this is impossible by point *2*. Therefore we can normalize: we may suppose $\alpha = 1$. Hence we have found positive $(\alpha_J)_J$ such that $\sum_{J\ni j} \alpha_J = 1$ and verifying $\sum_J \alpha_J\, \rho(\sum_{j\in J} X_j) < \rho(\sum_{i\leq N} X_i)$. By coherence, it is a contradiction, since we may write:

$$\begin{aligned}
\rho\left(\sum_{i\leq N} X_i\right) &= \rho\left(\sum_j \left(\sum_{J\in j} \alpha_J\right) X_j\right)\\
&= \rho\left(\sum_J \alpha_J \left(\sum_{j\in J} X_j\right)\right)\\
&\leq \sum_J \alpha_J\, \rho\left(\sum_{j\in J} X_j\right)
\end{aligned}$$

So there is a fair allocation. □

8.2 Fuzzy game approach, weak* gradients.

The basic papers regarding this approach are Aubin, **Au**, Artzner-Ostroy, **AO** and Billera-Heath, **BH**. An allocation $k_1, \ldots, k_N$ with $k = k_1 + \ldots k_n = \rho(\sum_{j=1}^N X_j)$ is now called fair (or fair for fuzzy games) if $\forall \alpha_j,\ j = 1, \ldots, N,\ 0 \leq \alpha_j \leq 1$ we have:

$$\sum_{j=1}^N \alpha_j k_j \leq \rho\left(\sum_{j=1}^N \alpha_j X_j\right)$$

Rearranging the relations we can write:

$$\sum_{j=1}^N (1-\alpha_j) k_j \geq \rho\left(\sum_{j=1}^N X_j\right) - \rho\left(\sum_{j=1}^N \alpha_j X_j\right)$$

thus, setting $\beta_j = 1 - \alpha_j$ we get:

$$\rho\left(\sum_{j=1}^{N} X_j - \sum_{j=1}^{N} \beta_j X_j\right) \geq \rho\left(\sum_{i=1}^{N} X_j\right) - \sum_{j=1}^{N} \beta_j k_j$$

If we restrict ρ to the linear combinations of $(X_j)_j$, then we can write $\rho(\sum \alpha_j X_j) = \rho(\alpha)$, where $\alpha \in [0,1]^N$. Therefore, the first term in the previous inequality can be seen as a perturbation of ρ around the point $(1, \dots, 1)$ by $-(\beta_1, \dots, \beta_N)$; if ρ were differentiable everywhere, k_j could be the j-th component of the gradient in $(1, \dots, 1)$. Since convex functions are not necessarily differentiable (or even weakly differentiable), we introduce the subgradient.

Let $f : L^\infty \to \mathbb{R}$ be convex and let X be an element of L^∞. The weak* subgradient of f at X is given by the set:

$$\nabla f(X) = \{g \mid g \in L^1 \, s.t. \forall Y \in L^\infty \quad f(X+Y) \geq f(X) + \mathbb{E}[gY]\}.$$

This set can be empty!

We now focus on the subgradient of ρ. As we will see, the weak* gradient of ρ can be empty at some points of L^∞. This will follow from James's theorem on weak compactness. However, it is a consequence of the Bishop-Phelps theorem that the set $\{X \mid \nabla\rho(X) \neq \emptyset\}$ is norm dense in L^∞.

Theorem 17 *Let ρ be a coherent risk measure with the Fatou property and given by the family $\mathcal{P}$. Then $g \in \nabla\rho(X)$ iff $-g \in \mathcal{P}$ and $\rho(X) = \mathbb{E}[gX] = \mathbb{E}[(-g)(-X)]$.*

Proof. First, let us suppose that there exists $\mathbb{Q} \in \mathcal{P}$ with density h, such that $\rho(X) = \mathbb{Q}[-X] = \mathbb{E}[h(-X)]$. Then we obviously get for all $Y \in L^\infty$:

$$\rho(X+Y) \geq \mathbb{Q}[-(X+Y)] \geq \mathbb{Q}[-X] + \mathbb{Q}[-Y] = \rho(X) + \mathbb{E}[(-h)Y],$$

which implies $-h \in \nabla\rho(X)$.

Conversely, suppose that $g \in \nabla\rho(X)$. Then, $-g$ is positive and $\mathbb{E}[-g] = 1$. In fact, if $Y \geq 0$, we obtain $\rho(X) \geq \rho(X+Y) \geq \rho(X) - \mathbb{E}[(-g)Y]$, which implies $\mathbb{E}[(-g)Y] \geq 0$ for all positive Y. Now, by taking $Y = c$, where $c \in \mathbb{R}$, we have:

$$\rho(X) - c = \rho(X+c) \geq \rho(X) - \mathbb{E}[(-g)c]$$

and consequently we have, for all c, $-c \geq \mathbb{E}[-c(-g)]$, that is $\mathbb{E}[-g] = 1$. So, $-g$ is a probability density and to conclude we must prove that it belongs to $\mathcal{P}$ and that $\rho(X) =$

$\mathbb{E}[(-g)(-X)]$.

Let's take $\lambda \in \mathbb{R}^+$: then $\rho(X + \lambda Y) \geq \rho(X) - \lambda\mathbb{E}[(-g)Y]$ and, dividing by λ, we get:

$$\rho\left(\frac{X}{\lambda} + Y\right) \geq \frac{\rho(Y)}{\lambda} - \mathbb{E}[(-g)Y]\,.$$

Letting λ tend to $+\infty$, since ρ in continuous with respect to the uniform convergence, we have $\rho(Y) \geq \mathbb{E}[(-g)(-Y)]$. Hence $-g \in \mathcal{P}$. To show that $\rho(X) = \mathbb{E}[gX]$, let's take $Y = -X$. We have the relation:

$$0 \geq \rho(X + Y) \geq \rho(X) + \mathbb{E}[(-g)X]\,,$$

which implies $\mathbb{E}[(-g)(-X)] \geq \rho(X)$. Together with the already proved converse inequality, we get $g \in \nabla\rho(X)$. □

Corollary 1 *$\mathcal{P}$ is weakly compact iff for every X the subgradient of ρ at X is non empty.*

Proof. This is a simple consequence of James's theorem. □

Corollary 2 *The set $\{X \mid \nabla\rho(X) \neq \emptyset\}$ is norm dense in L^∞.*

Proof. This corollary follows immediately from the Bishop-Phelps theorem. □

Theorem 18 *If $X = X_1 + \ldots + X_n$ and if $-\mathbb{Q} \in \nabla\rho(X)$ then the allocation $k_i = \mathbb{E}_{\mathbb{Q}}[-X_i]$ is a fair allocation in the sense of fuzzy games.*

Proof. We omit the proof. It is almost a repetition of the introductory calculations above. □

Proposition 5 *Let the subgradient $\nabla\rho(X)$ be the singleton $\{-\mathbb{Q}\}$ with $\mathbb{Q} \in \mathcal{P}$. Then if $\mathcal{P}$ is weakly compact, we have for every Y:*

$$\lim_{\epsilon \to 0} \frac{\rho(X + \epsilon Y) - \rho(X)}{\epsilon} = -\mathbb{Q}[Y]$$

Proof. Since $\mathcal{P}$ is weakly compact, for each ϵ, we can take $\mathbb{Q}_\epsilon \in \mathcal{P}$ so that $\rho(X + \epsilon Y) = \mathbb{Q}_\epsilon[-(X + \epsilon Y)]$. Then we have $\mathbb{Q}_\epsilon \to \mathbb{Q}$ weakly. In fact, let $\tilde{\mathbb{Q}}$ be an adherent point; we can write $\tilde{\mathbb{Q}}[-X] = \lim_{\epsilon \to 0} \mathbb{Q}_\epsilon[-(X + \epsilon Y)] = \lim_{\epsilon \to 0} \rho(X + \epsilon Y) = \rho(X)$, where the first equality holds because $X + \epsilon Y \overset{\|\cdot\|_\infty}{\longrightarrow} X$ and $\tilde{\mathbb{Q}}$ is weakly adherent. So by uniqueness $\tilde{\mathbb{Q}}$ is equal to $\mathbb{Q}$ and $\mathbb{Q}_\epsilon \overset{weakly}{\longrightarrow} \mathbb{Q}$. Using this result, we get:

$$\rho(X + \epsilon Y) - \rho(X) = \mathbb{Q}_\epsilon[-(X + \epsilon Y)] - \rho(X) \leq \mathbb{Q}_\epsilon[-\epsilon Y]$$

and dividing by ϵ:

$$\frac{\rho(X + \epsilon Y) - \rho(X)}{\epsilon} \leq \mathbb{Q}_\epsilon[-Y]\,.$$

Letting ϵ tend to zero, we finally find:

$$\limsup_{\epsilon \to 0} \frac{\rho(X + \epsilon Y) - \rho(X)}{\epsilon} \leq \mathbb{Q}[-Y]\,.$$

Conversely, since $-\mathbb{Q} \in \nabla\rho(X)$ the following is true:

$$\rho(X + \epsilon Y) - \rho(X) \geq \mathbb{Q}[-\epsilon Y]\,;$$

hence:

$$\liminf_{\epsilon \to 0} \frac{\rho(X + \epsilon Y) - \rho(X)}{\epsilon} \geq \mathbb{Q}[-Y]$$

□

Example 15 An example of this has been given by Uwe Schmock in a paper written for Swiss Reinsurance, **Schmo**. He proposed to use $\mathbb{E}[-X_i \mid X < q_\alpha(X)]$ as a capital allocation method. The previous theory shows that this is a very natural way. Indeed the risk measure corresponds to the weakly compact set $\mathcal{P}_{1/\alpha}$ of section 4. If X has a continuous distribution, or more generally when $\mathbb{P}[X \leq q_\alpha(X)] = \alpha$, then $\nabla\rho(X) = \{-1/\alpha\, I_A\}$, where $A = \{X \leq q_\alpha(X)\}$. So this example fits in the above framework of differentiability. The differentiability here is on the space L^∞. If only differentiability is required on the linear span of the random variables $X_1, \ldots, X_n$, things change. For more information on this topic the reader should consult the paper by Tasche, **Ta**.

Remark 12 The hypothesis that $\mathcal{P}$ is weakly compact cannot be omitted. Actually, let us consider the probability space $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mathbb{P})$ where $\mathbb{N}$ is the set of natural numers (including 0) and where $\mathbb{P}\{n\} = \frac{1}{2^{n+1}}$. For $\mathcal{P}$ we take the set of **all** probabilities on $\mathbb{N}$. Then $\mathcal{P}$ is not weakly compact. One way of seeing this simple fact is by taking the family of densities $\left(X_n = 2^{n+1} I_{\{n\}}\right)_n$ which is not uniformly integrable. We now define X (or, better, $-X$) in the following way: $(-X)(0) = 1$ and $(-X)(n) = 1 - \frac{1}{n}$ if $n \geq 1$. Then it is immediately seen that $\rho(X) = 1$ and that $\nabla\rho(X) = \{\delta_0\}$ (i.e. the Dirac measure in 0). If we now define Y by: $(-Y)(0) = 0$ and $(-Y)(n) = (-X)(n) = 1 - \frac{1}{n}$ if $n \geq 1$, then for all $\epsilon > 0$ we get $\rho(X + \epsilon Y) = 1 + \epsilon$. So, $\frac{\rho(X+\epsilon Y) - \rho(X)}{\epsilon} = 1$ whereas $\delta_0[-Y] = 0$.

Going back to the Capital Allocation problem, we again consider the global firm position X as the sum of N trading units positions, $X = X_1 + \ldots + X_N$. If $\nabla\rho(X) = \{-\mathbb{Q}\}$ then $k_i = \mathbb{Q}[-X_i]$ defines a fair allocation. In fact, for every Y we have:

$$\rho(X+Y) \geq \rho(X) - \mathbb{Q}[Y]$$

that is $\mathbb{Q}[-Y] \leq \rho(X+Y) - \rho(X) \leq \rho(Y)$.

If we put $Y = \sum_{i=1}^{N} \alpha_i X_i$, then $\sum_{i=1}^{N} \alpha_i k_i \leq \rho(\sum_{i=1}^{N} \alpha_i X_i)$ for all nonnegative α_i, not only for $\alpha_i \in [0,1]$. If moreover $\mathcal{P}$ is weakly compact then $k_i = \lim_{\epsilon \to 0} \frac{\rho(X+\epsilon X_i) - \rho(X)}{\epsilon}$ which can be interpreted as the marginal contribution of X_i to the risk of the global position X.

9 The extension of risk measures to L^0

As we already said, L^0 is invariant under change of probability measure and it deserves investigation. We can state the following:

Theorem 19 *If $\mathbb{P}$ is atomless, then there exists no functional $\rho : L^0 \to \mathbb{R}$ such that:*

1. $\rho(X+a) = \rho(X) - a \quad \forall a \in \mathbb{R}\,;$

2. $\rho(X+Y) \leq \rho(X) + \rho(Y)\,;$

3. $\rho(\lambda X) = \lambda \rho(X) \quad \forall \lambda \in \mathbb{R}\,, \lambda \geq 0\,;$

4. $X \geq 0 \Rightarrow \rho(X) \leq 0\,.$

This is a consequence of the analytic version of the Hahn-Banach theorem and of the fact that a continuous linear functional on L^0 must be necessarily null if $\mathbb{P}$ is atomless.
In extending coherent risk measures, we may substitute the real line with the real extended line, so that infinite values would be allowed. Anyway, a $(-\infty)$-risky position is meaningless, because it implies that any sum of money can be taken away without modifying the absolutely riskless position. On the contrary, a $+\infty$ position makes sense: it represents an absolutely risky position, which no amount of money can cover. So let us consider $\rho : L^\infty \to \mathbb{R}$ defined by $\mathcal{P} \subset L^1$. We first define for arbitrary random variables $X \in L^0$:

$$\rho(X \wedge n) = \sup_{\mathbb{Q} \in \mathcal{P}} \mathbb{Q}[-(X \wedge n)]\,.$$

We remark that the truncation is necessary to prevent the integral from being $-\infty$ (in practice, we want to avoid the influence of "too" large benefits). We then define:

$$\rho(X) = \lim_{n\to+\infty} \rho(X \wedge n)\,.$$

Unfortunately, this ρ can turn out to be $-\infty$. For instance, one could take $X > 0$ but non integrable, so that every $\rho(X \wedge n)$ is finite, while the limit is not and it has the wrong sign.

We note that the following implications hold:

$$\forall X \in L^0\,,\ \rho(X) > -\infty \Leftrightarrow \forall X \in L^0,\ X \geq 0 \Rightarrow \lim_{n\to+\infty} \inf_{\mathbb{Q}\in\mathcal{P}} \mathbb{Q}[X \wedge n] < +\infty$$

If the first inequality holds, obviously the second one is true; to prove the converse, we note that the newly defined ρ is subadditive and that $X^+ = X + X^-$: then $\rho(X) \geq \rho(X^+) > -\infty$.

So we have already proved the equivalence between the first two points of the following theorem:

Theorem 20 *The following conditions are equivalent:*

1. $\forall X \in L^0$ $\rho(X) > -\infty$;

2. $\forall X \geq 0$, $\phi(X) = \lim_{n\to+\infty} \inf_{\mathbb{Q}\in\mathcal{P}} \mathbb{Q}[X \wedge n] < +\infty$;

3. $\exists \gamma > 0$ *such that* $\forall A \in \mathcal{F}$ $\quad \mathbb{P}[A] \leq \gamma \Rightarrow \inf_{\mathbb{Q}\in\mathcal{P}} \mathbb{Q}[A] = 0$;

4. $\forall f \geq 0$ $\exists \mathbb{Q} \in \mathcal{P}$ *such that* $\mathbb{Q}[f] < +\infty$;

5. $\exists \gamma > 0$ *such that* $\forall A \in \mathcal{F}$, $\quad P[A] \leq \gamma$, $\exists \mathbb{Q} \in \mathcal{P}$ *with* $\mathbb{Q}[A] = 0$;

6. $\exists \gamma > 0$, $\exists k$ *such that* $\forall A \in \mathcal{F}$, *with* $\mathbb{P}[A] \leq \gamma$, $\quad \exists \mathbb{Q} \in \mathcal{P}$ *with the properties:*

$$\begin{cases} \mathbb{Q}[A] = 0 \\ \frac{d\mathbb{Q}}{d\mathbb{P}} \leq k\,. \end{cases}$$

Proof. We need to prove the equivalences from point *2* to point *6* and the scheme is: *3*$\Leftrightarrow$*2* $\Rightarrow$ *6* $\Rightarrow$ *5* $\Rightarrow$ *4* $\Rightarrow$ *2*. Let's start .
(*2* $\Rightarrow$ *3*)
By contradiction, if *3* is false then for every n we can find A_n with $\mathbb{P}[A_n] \leq 2^{-n}$ so that $\inf_{\mathbb{Q}\in\mathcal{P}} \mathbb{Q}[A_n] \geq \epsilon_n > 0$. Then we define $f = \sum_{n\geq 1} I_{A_n} \frac{n}{\epsilon_n}$. By Borel-Cantelli's lemma the sum is finite almost surely. Now we can write:

$$\mathbb{Q}\left[f \wedge \frac{N}{\epsilon_N}\right] \geq \mathbb{Q}\left[\left(I_{A_N}\frac{N}{\epsilon_N}\right) \wedge \frac{N}{\epsilon_N}\right] \geq N$$

and therefore $\inf_{\mathbb{Q}\in\mathcal{P}} \mathbb{Q}[f \wedge \frac{N}{\epsilon_N}] \geq N$; letting N tend to infinity, we contradict *2*.
(*3* $\Rightarrow$ *2*)
Let's fix a positive f: since it is real valued, there exists K such that $\mathbb{P}[\{f > K\}] < \gamma$ and taking $n > K$ we get

$$\inf_{\mathbb{Q}\in\mathcal{P}} \mathbb{Q}[f \wedge n] \equiv \inf_{\mathbb{Q}\in\mathcal{P}} \mathbb{Q}[f \wedge nI_{\{f>K\}} + f \wedge nI_{\{f\leq K\}}] \leq \inf_{\mathbb{Q}\in\mathcal{P}} (\mathbb{Q}[f \wedge nI_{\{f>K\}}] + K) \leq K$$

The implications *6* $\Rightarrow$ *5* $\Rightarrow$ *4* $\Rightarrow$ *2* are easy exercises: the real challenge is proving the arrow *3* $\Rightarrow$ *6.*
(*3* $\Rightarrow$ *6*)
Let $k > \frac{2}{\gamma}$ and let A with $\mathbb{P}[A] < \frac{\gamma}{2}$ be given. We will show *6* by contradiction. So let us take $H_k = \{f \mid |f| \leq k,\ f = 0 \text{ on } A\}$. If H_k and $\mathcal{P}$ were disjoint we could, by the Hahn-Banach theorem, strictly separate the closed convex set $\mathcal{P}$ and the weakly compact convex set H_k. This means that there exists an element $X \in L^\infty$, $\|X\|_\infty \leq 1$ so that

$$\sup\{\mathbb{E}[Xf] \mid f \in H_k\} < \inf\{\mathbb{E}_{\mathbb{Q}}[X] \mid \mathbb{Q} \in \mathcal{P}\}. \tag{2}$$

We will show that this inequality implies that $\|XI_{A^c}\|_1 = 0$. Indeed if not, we would have $\mathbb{P}[\{I_{A^c}|X| > \frac{2}{\gamma}\|XI_{A^c}\|_1\}] \leq \frac{\gamma}{2}$ and hence for each $\epsilon > 0$ there is a $\mathbb{Q} \in \mathcal{P}$ so that $\mathbb{Q}[A \cup \{|X| > \frac{2}{\gamma}\|XI_{A^c}\|_1\}] \leq \epsilon$. This implies that the right side of (2) is bounded by $\frac{2}{\gamma}\|XI_{A^c}\|_1$. However, the left side is precisely $k\|XI_{A^c}\|_1$. This implies $k\|XI_{A^c}\|_1 < \frac{2}{\gamma}\|XI_{A^c}\|_1$, a contradiction to the choice of k. Therefore $X = 0$ on A^c. But then property *3* implies that the right side is 0, whereas the left side is automatically equal to zero. This is a contradiction to the strict separation and the implication *3* $\Rightarrow$ *6* is therefore proved.

$\square$

Remark 13 The proof presented here is much easier than the original proof. Of course, there is some cheating in the sense that the statement *6* is directly verified without giving any reason why it could be true.

10 Bibliography

ADEH1 Artzner, Ph., F. Delbaen, J.-M. Eber, and D. Heath (1997) *Thinking Coherently*, RISK **10**, 1997, November, pp. **68–71**

ADEH2 Artzner, Ph., F. Delbaen, J.-M. Eber, and D. Heath (1998) *Coherent Measures of Risk* Mathematical Finance, 1999, ,pp **203-228**.

AT Acerbi, C., Tasche, D. (2001)*Expected Shortfall: a natural coherent alternative to Value at Risk.* Wilmott (to appear), http://www-m4.mathematik.tu-muenchen.de/m4/pers/tasche/ (for the abstract)

AO Artzner, P. and Ostroy (1984)*Gradients, subgradients and economic equilibria*, Adv. in Appl. Math., **4**, 1984, pp. 245–259

Au Aubin, J.P. (1974)*Cœur et équilibres des jeux flous à paiements latéraux*, C.R.Acad.Sci Paris Sér A, 279, 1974,pp. 891–894

BH Billera, L.J. and Heath, C.C. (1982) *Allocation of costs: a set of axioms yielding a unique procedure*, Mathematics of Operations Research **1**, 1982, pp. 32–39

Ch, Choquet, G (1953), *Theory of capacities*, Ann. Inst. Fourier, **5**, pp. 131–295.

D1, Delbaen, F. (1974).*Convex Games and Extreme Points*, Journ. Math. Anal. Appli., **45**, pp. 210–233.

D2, Delbaen, F. (1999),*Coherent Risk Measures on General Probability Spaces*, preliminary version on http:www.math.ethz.ch/ delbaen.

DS, Delbaen, F. and Schachermayer, W. *A general version of the fundamental theorem of asset pricing*, Math. Ann. **300**, 1994, pp. 463–520.

Den, Denault, M. (2000),*Coherent Allocation of Risk Capital*, Working Paper, H.E.C. Montréal.

Di, Diestel, J. (1975),*Geometry of Banach Spaces-Selected Topics.*, Lecture Notes in Mathematics, **485**, Berlin Heidelberg New York, Springer Verlag.

DuS, Dunford, N. and Schwartz, J. (1958), *Linear Operators, Vol I*, New York, Interscience.

Fi, Fischer, T. (2001),*Coherent Risk Measures depending on Higher Moments*, Ruprecht-Karls-Universität Heidelberg, Institut für International Vergleichende Wirstschafts- und Sozialstatistik (preprint).

FS, Föllmer, H. and Schied, A. (2001),*Convex Measures of Risk and Trading Constraints*, Humboldt University, Department of Mathematics (preprint).

Ja, Jaschke, S (2001),*Quantile-VaR is the Wrong Measure to Quantify Risk for Regulatory Purposes*, http://wws.mathematik.hu-berlin.de/ jaschke/papers/ (preliminary version).

JaK, Jaschke, S and Küchler, U. (1999),*Coherent Risk Measures, Valuation Bounds, and* (μ, σ) *Portfolio Optimisation*, http://wws.mathematik.hu-berlin.de/ jaschke/papers/ (preliminary version).

Ku, Kusuoka, S. (2000),*Law Invariant Coherent Risk Measures*, Graduate School of Mathematical Sciences, Tokyo University (preprint).

Ph, Phelps, R.,R. (1993),*Convex Functions, Monotone Operators and Differentiability, 2nd Edition*, Lecture Notes in Mathematics,**1364**, Berlin Heidelberg New York, Springer Verlag.

Ro, Rockafellar, R. (1970),*Convex Analysis*, Princeton Mathematical Series,**28**, Princeton, New Jersey, Princeton University Press.

Schm1, Schmeidler, D. (1972),*Cores of Convex Games*, J. Math. Anal. Appli.,**40**,pp. 214–225.

Schm2, Schmeidler, D. (1986), *Integral Representation without Additivity*, Proc. Amer. Math. Soc.,**97**, pp. 255–261.

Schmo, Schmock, U. (1999), private communication.

Ta, Tasche, D. (1999), *Risk contributions and performance measurement* (1999), **http://www.math.ethz.ch/ tasche/** or **http://www-m1.mathematik.tu-muenchen.de/m4/Papers/Tasche/**.