

UNIVERSITÀ DEGLI STUDI DI PADOVA

FACOLTÀ DI SCIENZE STATISTICHE

**CORSO DI LAUREA
IN STATISTICA, ECONOMIA, FINANZA E AZIENDA**

RELAZIONE FINALE

**Convex measures of risk: comparison
of different approaches and applications
to utility maximization and pricing rules.**

RELATORE: CH.MO PROF. MARCO FERRANTE

CORRELATORE: CH.MO PROF. MANUEL MORENO

LAUREANDO: MATTEO BONATO

ANNO ACCADEMICO 2004-05

Contents

1	Introduction	7
2	The Axioms	11
2.1	Axioms	11
2.1.1	Convexity(a) and Sublinearity (d) (d1)	12
2.1.2	Translation invariance (e) and Constancy (e1)	13
2.1.3	*Positivity (c) and *Monotonicity (c1)	15
2.1.4	Lower semi-continuity (b)	15
2.1.5	Law invariance(f)	15
2.1.6	Normalization ($\rho(0) = 0$)	15
2.1.7	Comonotonic additivity	16
3	Coherent Measures of Risk	17
3.1	Coherent measure of risk: Ω finite case	17
3.1.1	Representation Theorems for Coherent Risk Measures	19
3.2	Coherent Measures of risk on General Probability Spaces	19
3.2.1	Notation	19
3.2.2	The General Case	20
3.2.3	The σ -additive Case	22
4	Convex Risk Measures	23
4.1	The Föllmer and Schied convex risk measure	23
4.1.1	Acceptance Sets	24
4.1.2	The representation theorem for convex measures of risk	25
4.1.3	Robust representation of convex measures of risk	26
4.1.4	Risk Measures defined in terms of shortfall risk	29
4.2	The Frittelli and Rosazza Gianin (2002) convex risk measure	30
4.2.1	Notation	31
4.2.2	Convex Risk Measures	32

4.2.3	Representation of convex risk measures	32
4.3	Differences and similarities between the two definitions of convex measure of risk	34
4.3.1	About the axioms	34
4.3.2	Comparison between the two representations	35
5	Exponential utility, loss function and relative entropy	37
5.1	Shortfall Risk and Certain Equivalent	37
5.2	Certainty Equivalent and Exponential Utility Function	39
5.2.1	Main Definitions	41
5.2.2	Duality	42
5.2.3	Generalized distances $\delta(Q, P)$ and examples	42
5.2.4	Computation of the value	43
5.3	Entropic Convex Measure of Risk	44
6	Risk measure and claims pricing	47
6.1	Coherent risk measures and pricing rules	48
6.2	Convex measure of risk and pricing rules	49
6.2.1	Abstract contingent claims	49
6.2.2	Superreplication Price	50
6.2.3	Utility Maximization Price	51
6.3	Interpretation of the result	52
7	Applications	55
7.1	Value at Risk	55
7.2	Average Value at Risk	57
7.3	Examples	58
7.3.1	The continuous case	58
7.3.2	The general case	60
7.3.3	Portfolio risk measure	63
8	Conclusions	67
A	Real Analysis and Measure Theory	69
A.1	Ordered Sets	69
A.2	Lebesgue Measure	70
A.3	The Lebesgue Integral	72
A.4	The $L^p[0, 1]$ Spaces	74
A.5	Metric Spaces	77
A.6	Topological Spaces	78

A.7	Banach Spaces	79
A.8	Measure and Integration: the Radon-Nikodym Theorem	82
A.8.1	The general L^p Spaces	84
B	Principles of Convex Analysis	85
B.1	Convex Functions on \mathbb{R}	85
B.1.1	The Conjugate Function	86
B.1.2	Convex Functions With Values in $\bar{\mathbb{R}}$	88
B.2	Convex Functions On a Linear Space	89
B.3	Duality Theory	90
B.3.1	The Conjugate Function	90
C	Implementation of the programs	93

Chapter 1

Introduction

Research about risk measurement had a great grown in the recent years. Such research touched different but interconnected aspects:

1. axiomatic characterization of risk measures;
2. construction on risk measures;
3. premium principles in insurance context;
4. dynamic risk measures;
5. the relation between risk measures and other economics and financial theories;
6. application of risk measures to financial activities.

The matter of analysis of this work will be the static measures of risk.

In many papers risk was defined in terms of *changes* in value between two dates, probably because risk is related to the variability of the *future* value of a position, due to market changes or more generally to uncertain events. That is why, in their work, Artzener et al. (1999) prefer consider future values only. So the basic object of the study are the random variables in the set of states of nature at a future date, interpreted as possible future values of positions or portfolios currently held. A first measurement of the risk of a position will be whether its future value belongs or does not belong to the subset of *acceptable risks*, as decided by a supervisor such as:

- (a) a *regulator* who takes into account the unfavorable states when allowing a risky position that may draw on the resources of the government;
- (c) an *exchange's clearing firm*, which has to make good on the promises to all parties of transactions being securely completed;

- (c) an investment *manager* who knows that his firm has basically given to its traders an exit option in which the strike “price” consist in being fired in the event of big trading losses on one’s position. For an *unacceptable* position one has to alter the position or look for some commonly accepted instruments that, when added to the current position, make it acceptable. The current cost of getting enough of these instruments is a good candidate for a measure of risk of the initially unacceptable position.

Let call X the space we will assume to be the “habitat” of all the financial position whose riskiness we want to quantify.

Definition 1.1 (Measure of risk) *A (static) risk measure is a functional*

$$\rho : X \rightarrow \mathbb{R}$$

satisfying some properties which seem to be “desirable” from a financial point of view.

We will follow now a deductive approach to risk measures. Historically, in fact, the first axiomatization of the concept of risk measure is due to Artzner et al. (1999) with their path-breaking work *Coherent measure of risk*. The main merit of the work of Artzner et al. (1999) was to try to fix four properties that a risk measure has to satisfy to be “correct”. In their interpretation *coherent*. One of the aim of this work is to generalize the concept of coherent risk measure by defining new kinds of risk measures. Our presentation is deductive because we will present in Chapter 2 a list of (non independent) axioms for ρ and we will show that different risk measures arise with the imposition of different axioms. A particular choice of four of them, for example, lead in [3] to define a coherent risk measure. The other were used in other contexts by other authors and they all have a financial explanation. We present in Chapter 3 the first axiomatic approach to risk measurement and the definition of the axioms of coherence. We will see the case when Ω , the set of states of nature, is finite and when it can be general. In some cases coherence for a measure of risk can be a too strong requirement. Thus some authors weakened one of the axioms a coherent risk measure has to satisfy (the subadditivity axiom) and imposed the axiom of convexity instead. This led to a new kind of risk measures, called, for this reason, *convex*. In Chapter 4 we introduce and analyze two independent approaches to convex measures of risk. We will see the differences and the similarities and we will show that in some ways these two measures coincide. In Chapter 5 we present a classical example of convex measure of risk: the entropic risk measure. We show that in the case of totally incomplete markets, both the approaches lead to the same measure. In Chapter 6

we go beyond and find a relation between the entropic measure of risk and claim pricing. This way of pricing a claim keeps in consideration the preferences of the investors by modelling them using the well-know exponential utility function. To conclude we present in Chapter 7 a simple application of the convex measure of risk known as AVaR. We made a comparison between Value at Risk and AVaR in both theoretical and empirical context. For the empirical analysis we considered the risk measure know as Expected Shortfall, measure that was our matter of investigation in [5].

Chapter 2

The Axioms

We present now a list of axioms that represent possible properties that a risk measure has to satisfy.

2.1 Axioms

- (a) convexity: $Epi(\rho) = \{(x, a) \in X \times \mathbb{R} : \rho(x) \leq a\}$ is convex in $X \times \mathbb{R}$;
- (b) lower semi-continuity: the set $\{x \in X : \rho(x) \leq c\}$ is closed in X for all $c \in \mathbb{R}$;
- (c) *positivity: $x \geq 0 \Rightarrow \rho(x) \leq \rho(0)$, $\forall x \in X$;
- (c1) *monotonicity: $x \geq y \Rightarrow \rho(x) \leq \rho(y)$, $x, y \in X$;
- (c2) *relevance: $x \leq 0$, and $x \neq 0$ implies $\rho(x) > 0$, $x \in X$;
- (d) subadditivity: $\rho(x + y) \leq \rho(x) + \rho(y)$, $\forall x, y \in X$;
- (d1) positive homogeneity: $\rho(ax) = a\rho(x)$, $\forall a \geq 0, \forall x \in X$;
- (e) translation invariance: $\rho(x + a) = \rho(x) - a$, $\forall a \geq 0, \forall x \in X$;
- (e1) constancy: $\rho(a) = -a$, $\forall a \geq 0$;
- (f) law invariance: if $x, y \in X$ have the same distribution w.r.t. P , then $\rho(x) = \rho(y)$ (this is the only axiom that effectively depends on the reference probability P);
- (g) normalization: $\rho(0) = 0$;
- (h) Comonotonic additivity: for comonotonic x and y , which means that $x = f \circ z$ and $y = g \circ z$ for non decreasing f and g and for z, y and $x \in X$ implies that $\rho(x + y) = \rho(x) + \rho(y)$.

Due to financial interpretation of ρ (see discussion of the axiom (e) below), the axioms (c) and (c1) have the inequality sign opposite to what the name of the axiom would suggest (this is the reason why we added the symbol (*) to the denomination of the axiom). An other way to solve this possible arising confusion is to work with $\pi(x) = \rho(-x)$, as done in [18], but we don't really think this is going to be a problem.

We discuss now the financial interpretation of the above axioms.

2.1.1 Convexity(a) and Sublinearity (d) (d1)

Recall that: ρ is convex if and only if

$$\rho(\alpha x + (1 - \alpha)y) \leq \alpha\rho(x) + (1 - \alpha)\rho(y), \quad \forall x, y \in X, \quad \forall \alpha \in [0, 1]; \quad (2.1)$$

ρ is sublinear if it ρ satisfies both axioms (d) subadditivity and (d1) positive homogeneity. In the definition of a *convex risk measure* (see Definitions 4.1 and 4.4) we require the convexity axiom but not necessary the sublinearity axiom. In fact, as we will see, sublinearity is stronger than convexity, and all we do is to weaken this axiom.

Subadditivity has an easy interpretation. Let us suppose that we own two positions which jointly have a positive measure of risk. Hence, we have to add extra cash to obtain a "neutral position". If the subadditivity did not hold, then, in order to deposit less extra cash, it would be sufficient for us to separate in two account two positions. Roughly speaking, it seems reasonable to have a *discount* when we "buy" several positions.

We notice that subadditivity implies that $\rho(nx) \leq n\rho(x)$, $\forall x \in X$, $\forall n \in \mathbb{N}$. The opposite inequality is imposed by the positive homogeneity axiom. However, this last axiom may not be necessary. We will see later why.

In many situations the risk of a position might increase in a non linear way with the size of the position. For example, an additional liquidity risk may arise if a position is multiplied by a large factor. This suggest to relax the condition of positive homogeneity and subadditivity. In the following four items we show why it could be reasonable to impose convexity instead (see [18] and [14]). Convexity means that diversification does not increase the risk, i.e., the risk of a diversified position $\rho(\alpha x + (1 - \alpha)y)$ is less or equal than the weighted average of the individual risks.

- (1) The convexity axiom clearly express the requirement that, as already seen, the *risk is not increased by the diversification* on the position held on the portfolio.

(2) Convexity alone implies the following inequalities (if $\rho(0) = 0$):

$$(a1) \quad \rho(\alpha x) \leq \alpha \rho(x), \quad \forall \alpha \in [0, 1], \quad \forall x \in X;$$

$$(a2) \quad \rho(\alpha x) \geq \alpha \rho(x), \quad \alpha \geq 1, \quad \forall x \in X.$$

The first is an immediate application of (2.1) with $y = 0$, the latter by applying (2.1) with $\frac{1}{\alpha}$ and $y = 0$. Both conditions (a1) and (a2) are justified by *liquid arguments*: Indeed, when α becomes large, the whole position (αx) is less liquid than α times the same position x , hence inequality (a2) seems reasonable. When α is small, the opposite inequality must hold for specular reasons.

While Artzner and al. [3] motivated the axiom (d1) of positive homogeneity because of liquid arguments, the belief that only property (a1) and (a2) were to be required, held to the so called *convex measure of risk*.

- (3) Some authors have argued that positive homogeneity is necessary to preserve the property that a *risk measure should be invariant with respect to the change of the currency*. In the discussion of the translation invariance axiom (e) below, we will see that this is not really the case (see also Remark 3.9 in [20]).
- (4) If $\rho(0) = 0$ it can be easily checked (see Lemma 2) that which ever two axioms, among convexity (a), subadditivity (d) and positive homogeneity (d1), hold true, then the other one holds true as well.

2.1.2 Translation invariance (e) and Constancy (e1)

- (1) The axiom of translation invariance (e) allows for the representation of $\rho(x)$ as a *capital requirement*. It guarantees that $\rho(x)$ is the minimal amount of money to add to the initial position x to make it acceptable:

Lemma 1 $\rho : X \rightarrow \mathbb{R}$ satisfies the axiom of translation invariance (e) if and only if there exists a set $\mathcal{A} \subseteq X$ such that:

$$\rho(x) = \inf\{\alpha \in \mathbb{R} \mid x + \alpha \in \mathcal{A}\}.$$

where $\mathcal{A} \doteq \{x \in X \mid \rho(x) \leq 0\}$

See Lemma 1 in [19] for the proof.

As we will see later (Definition 3.2), the set \mathcal{A} is the so called *acceptance set associated with ρ* . Thus $\rho(x)$ is positive for unacceptable position x , while is negative for acceptable position.

Note that axiom (e) ensure that, for each $x \in X$ we have $\rho(x - \rho(x)) = \rho(x) - \rho(x) = 0$. And this confirm the natural interpretation in terms of acceptance set associated with ρ .

- (2) Note that in the statement of axiom (e) (and the same argument could be used also - and only - for the constancy axiom (e1)) it is required that both sums $x + a$ and $\rho(x) - a$ are well defined. This implies that x and $\rho(x)$ must be expressed in the same unit: the unit of the constant a . If the random variable x (or $x_{\$}$) represents a random amount expressed in \$, then also $\rho(x)$ (or $\rho_{\$}(x_{\$})$) will be a sure amount expressed in \$. Hence a risk measure satisfying the translation invariance axiom *does depend on the particular choice of the currency* Therefore also the acceptable set associated to ρ (as well as the penalty function that will be introduced later) will depend on it.
- (3) Consider two currencies (to be concrete: dollar and pound) and let $\lambda > 0$ be the exchange rate: $1\$ = \lambda\mathcal{L}$. Let $A_{\mathcal{L}}$ be a subset of random variables which are expressed in \mathcal{L} . Then obviously, $x_{\mathcal{L}} \in A_{\mathcal{L}}$ if and only if $\lambda x_{\mathcal{L}} \in (\lambda A_{\mathcal{L}}) \doteq \{y | \exists x_{\mathcal{L}} \in A_{\mathcal{L}} : y = \lambda x_{\mathcal{L}}\}$. Then the elements of the sets $\lambda A_{\mathcal{L}}$ and $A_{\mathcal{L}}$ are the "same" random variable, denominated either in \$ or in \mathcal{L} . Hence if $\mathcal{A}_{\mathcal{L}}$ is the acceptable set associated with $\rho_{\mathcal{L}}$, then $\mathcal{A}_{\$} \doteq \lambda \mathcal{A}_{\mathcal{L}}$ is the acceptance set associated with $\rho_{\$}$.

Remark 1 Let $\lambda > 0$ be the exchange rate: $1\$ = \lambda\mathcal{L}$, let $\rho : X \rightarrow \mathbb{R}$ satisfy the axiom of translation invariance (e), let $\rho_{\mathcal{L}}$ (resp. $\rho_{\$}$) be the risk measure ρ expressed in pound (resp. dollar) and let $\mathcal{A}_{\mathcal{L}}$ (resp. $\mathcal{A}_{\$}$) be the acceptance set associated with $\rho_{\mathcal{L}}$ (resp. $\rho_{\$}$). If $x_{\$} = \lambda x_{\mathcal{L}}$, then:

$$\mathcal{A}_{\$} = \mathcal{A}_{\mathcal{L}} \quad \text{iff} \quad \rho_{\$}(x_{\$}) = \lambda \rho_{\mathcal{L}}(x_{\mathcal{L}}),$$

which is the proper substitute of the positive homogeneity property.

See Remark 2 in [19] for the proof.

- (4) We will see in Lemma 2 that for a convex risk measure (as proposed in [18]) the translation invariance axiom (e) is equivalent to the, self evident, constancy axiom (e1).

Remark 2 It is an easy exercise to prove that axiom (e) and (e1) together implies $\rho(0) = 0$.

Proof: $\rho(a) = \rho(0 + a) = \rho(0) - a = -a \Leftrightarrow \rho(0) = 0 \forall a \in \mathbb{R}$. \square

2.1.3 *Positivity (c) and *Monotonicity (c1)

Consider the axiom:

$$(c^\circ) \quad x \leq 0 \Rightarrow \rho(x) \geq \rho(0).$$

Pay attention to the fact that there is no symmetry between the axiom (c°) and (c) . It may be easily checked (see Remark 8 (iv)) that if ρ is convex (in the sense of [18]) then (c) implies (c°) but the converse implication is false, as shown by the simple counterexample $\rho(x) \doteq |x|, x \in \mathbb{R}$.

The interpretation of the axiom (c) (as well as $(c1)$) follows immediately from the financial meaning of a risk measure: suppose that $x \geq 0$, then the position is clearly acceptable and so $\rho(x) \leq 0$. Note that $-\rho(x)$ is the maximum amount of money which we can withdraw from the position. We will discuss this deeply in the next chapter.

2.1.4 Lower semi-continuity (b)

This axiom is technical and it is required essentially to achieve the adequate functional representation in Theorem 4.6. As we will see, this axiom is imposed in [19] and is a consequence in [14].

2.1.5 Law invariance(f)

In addition to the more "classical" axiom (a)-(e), law invariance is also recurrent in literature (see for example, Kusuoka (2001) and Wang and al. (1997)).

On the one hand, the financial motivation of law invariance is intuitive. Indeed, it is desirable to have risk measures which "allow the same riskiness" to financial positions that are identically distributed with respect to the probability P .

On the other hand, note that the definition of law invariance depends on the probability measure P given a priori, hence it is reasonable to expect that in the representation of law invariant coherent or convex risk measures the set \mathcal{P} of generalized scenarios will be dependent on P .

2.1.6 Normalization ($\rho(0) = 0$)

As we have already seen, this condition arises in case the axioms of translation invariance (e) and (e1) hold true. What we want to emphasize is that, if the risk measure ρ is normalized in the sense that $\rho(0) = 0$, then the quantity $\rho(x)$ can be interpreted as a "marginal requirement", i.e., as the minimal amount of capital which, if added to the position at the beginning of the given period and invested into a risk-free asset, makes the discounted position x acceptable. We will see

that the under this condition, both the approaches to the definition of convex risk measures in [14] and [18] lead to the same conclusion in term of the representation form that a convex measure has to satisfy.

2.1.7 Comonotonic additivity

This property can be interpreted as the fact that the risk of two random variables depending on the same underlying source of risk (z) is additive.

Chapter 3

Coherent Measures of Risk

We present now the first line of research that was started by a group of scholar: Artzner, Delbaen, Eber and Heath. The axiomatic definition of coherent risk measures was introduced in their path-breaking paper [3]. Delbaen, furthermore, extended the definition on coherent risk measures to general probability spaces.

3.1 Coherent measure of risk: Ω finite case

Here we briefly introduce the concept of coherent risk measure developed by Artzner et al. in [3].

NOTATION.

- (a) We shall call Ω the sets of states of nature, and assume it finite. Considering Ω as the set of outcomes of an experiment, we compute the final net worth of a position for each element of Ω . It is a random variable denoted by X . Its negative part, $\max(-X, 0)$, is denote by X^- and the supremum of X^- is denoted by $\|X^-\|$, if no possible confusion arises. The random variable identically equal to 1 is denoted by $\mathbf{1}$. The indicator function of the state ω is denoted by $\mathbf{1}_{\{\omega\}}$.
- (b) Let \mathcal{G} the set of all risks, that is the set of all real-valued functions on Ω . Since Ω is supposed to be finite, \mathcal{G} is isomorphic to \mathbb{R}^n , where $n = \text{card}(\Omega)$. The cone of nonnegative elements in \mathcal{G} will be denoted by L_+ , its negative part by L_- .
- (c) We call $\mathcal{A}_{i,j}$, $j \in J_i$, as a set of final net worth, expressed in currency i , which in country i , are accepted by a regulator/supervisor j .

- (d) We shall denote $\mathcal{A}_j = \bigcap_{i \in J_j} \mathcal{A}_{i,j}$ and use the generic notation \mathcal{A} in the listening of axioms below.
- (e) Differently from Artzner and al. paper, to simplify the notation and without loss of generality, in the furthering we will always assume that the risk free interest rate is zero.

Axiom 3.1 *The acceptance set \mathcal{A} contains L_+*

Axiom 3.2 *The acceptance set \mathcal{A} does not intersect the set L_{--} where*

$$L_{--} = \{X \mid \text{for each } \omega \in \Omega, X(\omega) < 0\}.$$

It will be interesting to consider the stronger axiom.

Axiom 3.3 *The acceptance set \mathcal{A} satisfies $\mathcal{A} \cap L_- = \emptyset$*

This axiom reflects the risk aversion of the regulator.

Axiom 3.4 *The acceptance set \mathcal{A} is convex.*

Axiom 3.5 *The acceptance set \mathcal{A} is a positively homogeneous cone.*

Definition 3.1 (Risk Measure) *A measure of risk is a mapping from \mathcal{G} into \mathbb{R} .*

When positive, the number $\rho(X)$ assigned by the measure ρ to the risk X will be interpreted as the minimum extra cash the agent has to *add* to the risky position X to make it acceptable. If it's negative, the cash amount $-\rho(X)$ can be withdrawn from the position or it can be received as restitution, as in the case of organized markets for financial futures.

We define a correspondence between acceptance sets and measures of risk.

Definition 3.2 (Risk measures associated with an acceptance set) *The risk measure associated with the acceptance set \mathcal{A} is the mapping from \mathcal{G} into \mathbb{R} denoted by $\rho_{\mathcal{A}}$ and defined by*

$$\rho_{\mathcal{A}}(X) = \inf\{m \mid m + X \in \mathcal{A}\} \quad (3.1)$$

Definition 3.3 (Acceptance set associated with risk measure) *The acceptance set associated with a risk measure ρ is denoted by \mathcal{A}_{ρ} and defined by*

$$\mathcal{A}_{\rho} = \{X \in \mathcal{G} \mid \rho(X) \leq 0\}. \quad (3.2)$$

We now define the set of four axioms a coherent risk measure has to satisfy:

Definition 3.4 (Coherence) *A risk measure satisfying the following axioms: (c1) *monotonicity, (d) subadditivity, (d1) positive homogeneity and (e) translation invariance is called coherent.*

We notice that for $\lambda > 0$ axioms (d),(d1),(c1) and (c2) remain satisfied by the measure $\lambda \cdot \rho$ if satisfied by the measure ρ . It is not the case for Axiom (e).

In this approach the acceptance set is the fundamental object and we have discussed the axioms mostly in terms of the associated risk measure. The following prepositions show that this was reasonable.

Proposition 3.1 *If the set \mathcal{B} satisfies Axioms 3.1,3.2,3.3,3.4 and 3.5, the risk measure $\rho_{\mathcal{B}}$ is coherent. Moreover $\mathcal{A}_{\rho_{\mathcal{B}}} = \bar{\mathcal{B}}$ is the closure of \mathcal{B}*

Proposition 3.2 *If a risk measures ρ is coherent, then the acceptance set \mathcal{A}_{ρ} is closed and satisfies Axioms 3.1,3.2,3.3,3.4 and 3.5. Moreover $\rho = \rho_{\mathcal{A}_{\rho}}$*

3.1.1 Representation Theorems for Coherent Risk Measures

In this section we show a general representation for coherent risk measures: *any* coherent risk measures arises as the supremum of the expected negative of final net worth for some collection of "generalized scenarios" or probability measures on states of the world. We continue to suppose that Ω is a finite set, otherwise we would also get finitely additive measure as scenarios.

The σ -algebra, 2^{Ω} , is the class of all subsets of Ω .

Proposition 3.3 *A risk measure ρ is coherent if and only if there exist a family \mathcal{P} of probability measures on the set of finite states of the nature, such that:*

$$\rho(X) = \sup\{E_P[-X] | P \in \mathcal{P}\}. \tag{3.3}$$

3.2 Coherent Measures of risk on General Probability Spaces

The aim of this section is to show that a coherent risk measures can be extended to arbitrary probability spaces. We will follow, in the exposition, the results presented in [11].

3.2.1 Notation

Throughout the section we will work with a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. With $L^{\infty}(\Omega, \mathcal{F}, \mathbb{P})$ (or $L^{\infty}(\mathbb{P})$ or L^{∞} if no confusion is possible), we mean the space of bounded real valued random variables. The space $L^0(\Omega, \mathcal{F}, \mathbb{P})$ (or $L^0(\mathbb{P})$ or simply L^0) denotes the space of all equivalence class of real valued random variables. The space L^0 is equipped with topology of convergence in probability. The space

$L^\infty(\mathbb{P})$, equipped with the usual L^∞ norm, is the dual of the space of integrable (equivalence classes of) random variables $L^1(\Omega, \mathcal{F}, \mathbb{P})$ (also denoted by $L^1(\mathbb{P})$ or L^1 if no confusion is possible). Let us recall that the dual of $L^\infty(\mathbb{P})$ is the Banach space $\mathbf{ba}(\Omega, \mathcal{F}, \mathbb{P})$ of all bounded, finitely additive measures μ on (Ω, \mathcal{F}) with the property that $\mathbb{P}(A) = 0$ implies $\mu(A) = 0$.

3.2.2 The General Case

In this subsection we will show that the main theorems in the paper [3] can be generalized to the case of general probability spaces. The main difficulty consists in replacing the finite dimension space \mathbb{R}^Ω by the space of bounded measurable functions, $L^\infty(\mathbb{P})$. In this setting the definition of a coherent risk measure as given in [3] can be written as:

Definition 3.5 *A mapping $\rho : L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$ is called a coherent risk measure if the properties of Positivity, Subadditivity, Positive Homogeneity and Translation Invariance hold.*

Although the properties listed in the definition of a coherent measure have a direct interpretation in mathematical finance, it is mathematically more convenient to work with the related submodular function, ψ , or with the associated supermodular function, ϕ .

Definition 3.6 (Submodular) *A mapping $\psi : L^\infty \rightarrow \mathbb{R}$ is called submodular if:*

1. For $X \leq 0$ we have that $\psi(X) \leq 0$.
2. If X and Y are bounded random variables then $\psi(X + Y) \leq \psi(X) + \psi(Y)$.
3. For $\lambda \geq 0$ and $X \in L^\infty$ we have $\psi(\lambda X) = \lambda\psi(X)$.

The submodular function is called translation invariant if moreover

4. For $X \in L^\infty$ and $a \in \mathbb{R}$ we have that $\psi(X + a) = \psi(X) + a$.

Definition 3.7 (Supermodular) *A mapping $\phi : L^\infty \rightarrow \mathbb{R}$ is called supermodular if*

1. For $X \leq 0$ we have that $\phi(X) \geq 0$.
2. If X and Y are bounded random variables then $\phi(X + Y) \geq \phi(X) + \phi(Y)$.
3. For $\lambda \geq 0$ and $X \in L^\infty$ we have $\phi(\lambda X) = \lambda\phi(X)$.

The supermodular function is called translation invariant if moreover

4. For $X \in L^\infty$ and $a \in \mathbb{R}$ we have that $\phi(X + a) = \phi(X) + a$.

Remark 3 If ρ is a coherent risk measure and if we put $\psi(X) = \rho(-X)$ we get a translation invariant submodular functional.

The following theorem is an immediate application of the bipolar theorem from functional analysis.

Theorem 3.1 Suppose that $\rho : L^\infty \rightarrow \mathbb{R}$ is a coherent risk measure with associated sub(super)modular function $\psi(\phi)$. There is a convex $\sigma(\mathbf{ba}(\mathbb{P}), L^\infty(\mathbb{P}))$ -closed set $\mathcal{P}_{\mathbf{ba}}$ of finitely additive probabilities, such that:

$$\psi(X) = \sup_{\mu \in \mathcal{P}_{\mathbf{ba}}} E_\mu[X] \quad \text{and} \quad \phi(X) = \inf_{\mu \in \mathcal{P}_{\mathbf{ba}}} E_\mu[X]$$

Remark on notation. There is a one-to-one correspondence between

1. coherent risk measures ρ ,
2. the associated supermodular function $\phi(X) = -\rho(X)$,
3. the associated submodular function $\psi(X) = \rho(-X)$,
4. the weak* closed convex set of finitely additive probability measures $\mathcal{P}_{\mathbf{ba}} \subset \mathbf{ba}(\mathbb{P})$,
5. $\|\cdot\|_\infty$ closed convex cones $C \subset L^\infty$ such that $L_+^\infty \subset C$.

The relation between C and ρ is given by

$$\rho(X) = \inf\{\alpha \mid X + \alpha \in C\}.$$

The set C is called set of acceptable positions, see [3].

Remark on the interpretation of the probability space. The σ -algebra \mathcal{F} describe all the events that becomes known at the end of an observed period. The interpretation of the probability \mathbb{P} seems to be more difficult. The measure \mathbb{P} describes the probability that events may occur. However, in economics and finance, such probabilities are subjective and depend on the preference of the regulators, and we may argue that the class of negligible sets and consequently the class of probability measures that are equivalent to \mathbb{P} remain the same. This can be expressed saying that only the knowledge of event with probability zero is important. So we only need agreement on the “ possibility ” that events might occur, not the actual value of the probability.

In view of this, there are two natural spaces of random variables on which we can define a probability measure. Only these two space remain the same when we change the underlying probability to an equivalent one. These two spaces are $L^\infty(\Omega, \mathcal{F}, \mathbb{B})$ and $L^0(\Omega, \mathcal{F}, \mathbb{B})$. The space L^0 cannot be given a norm and cannot be turned into a locally convex space.

3.2.3 The σ -additive Case

The previous subsection gave a characterization of translation invariant submodular functionals (or equivalently coherent risk measures) in terms of finitely additive probabilities. The characterization in terms of σ -additive measure requires additional hypothesis. E.g if μ is a purely finitely additive measure, the expression $\phi(X) = E_\mu[X]$ gives a translation invariant submodular functional. This functional cannot be described by a σ -additive probability measure. So we need extra condition.

Definition 3.8 *The translation invariant supermodular mapping $\phi : L^\infty \rightarrow \mathbb{R}$ is said to satisfy the Fatou property if $\phi(X) \geq \limsup \phi(X_n)$, for any sequence, $(X_n)_{n \geq 1}$, of functions, uniformly bounded by 1 and converging to X in probability.*

So we obtain:

Theorem 3.2 *For a translation invariant supermodular mapping ϕ , the following 4 properties are equivalent:*

1. *There is an $L^1(\mathbb{P})$ -closed, convex set of probability measures \mathcal{P}_σ , all of them being absolutely continuous with respect to \mathbb{P} and such that for $X \in L^\infty$:*

$$\phi(X) = \inf_{\mathbb{Q} \in \mathcal{P}_\sigma} E_{\mathbb{Q}}[X].$$

2. *The convex cone $C = \{X | \phi(X) \geq 0\}$ is weak*, i.e. $\sigma(L^\infty(\mathbb{P}), L^1(\mathbb{P}))$ closed.*
3. *ϕ satisfies the Fatou property.*
4. *If X_n is a uniformly bounded sequence that decreases to X a.s., then $\phi(X_n)$ tends to $\phi(X)$.*

Chapter 4

Convex Risk Measures

As anticipated and motivated previously, convex measures or risk were introduced as a generalization of coherent ones. They were firstly proposed by Heath (there called “shareholder risk measure” or “weak coherent measures of risk”) in finite sample spaces and later in general probability spaces by Föllmer and Schied (2002a) and, independently, by Frittelli and Rosazza Gianin (2002). All above notation of “convex risk measures” are based in the convexity axiom. However, they differ from each other because of the different selection of the other axiom. Even if the choice if the others axioms could be different, we will show that the representation reached in [14] and [18] is the same, and the two measures coincide.

To simplify the notation and without loss of generality, in the sequel we will always assume that the risk free rate is zero, i.e. we do not need to discount the future value of a position to get to the present value.

4.1 The Föllmer and Schied convex risk measure

We present now the definition of convex risk measures given by Föllmer and Schied. We will follow, in the exposition their work [14], and we always refer to this for the proofs. Their approach is, in some sense, the most obvious. In fact, starting from the definition of coherent risk measure in [3], they simple weaken the axiom of sublinearity (subadditivity (d) and translation invariance (d1)) and imposed convexity (a) instead. The whole way of proceeding is very similar to that used in [3] and we can say that [14] is the natural development of what started in [3].

Let X be a convex set of functions on the set Ω of possible scenarios. We assume that $0 \in X$ and that X is closed under the addition of constants.

Definition 4.1 (Convex Risk Measure) *A map $\rho : X \rightarrow \mathbb{R}$ is called convex*

risk measure if it satisfies the condition of convexity (a), *monotonicity (c1), and translation invariance (e).

4.1.1 Acceptance Sets

Definition 4.2 Let X be a linear space of functions on a given set Ω of possible scenarios. We assume that X contains all constants functions. Any risk measure $\rho : X \rightarrow \mathbb{R}$ induces an acceptance set \mathcal{A}_ρ defined as

$$\mathcal{A}_\rho \doteq \{x \in X | \rho(x) \leq 0\}.$$

Conversely, for a given class \mathcal{A} of acceptable position, we can introduce an associated risk measure $\rho_{\mathcal{A}}$ by defining:

$$\rho_{\mathcal{A}}(x) \doteq \inf\{m \in \mathbb{R} | m + x \in \mathcal{A}\} \quad (4.1)$$

The following two propositions summarize the relation existing between a convex risk measure and its acceptance set \mathcal{A}_ρ . They are similar to those funded for coherent measures of risk (see [3] and [11] for more details).

Proposition 4.1 Suppose $\rho : X \rightarrow \mathbb{R}$ is a convex measure of risk with associated acceptance set \mathcal{A}_ρ . Then

$$\rho_{\mathcal{A}_\rho} = \rho.$$

Moreover, $\mathcal{A} \doteq \mathcal{A}_\rho$ satisfies the following properties.

1. \mathcal{A} is convex and non-empty.
2. If $x \in \mathcal{A}$ and $y \in X$ satisfies $y \geq x$, then $y \in \mathcal{A}$.
3. If $x \in \mathcal{A}$ and $y \in X$, then

$$\{\lambda \in [0, 1] | \lambda x + (1 - \lambda)y \in \mathcal{A}\}$$

is closed in $[0, 1]$.

4. If the risk measure ρ is coherent, then \mathcal{A} is a convex cone.

Example 1 (Value at Risk) Value at Risk at level $\gamma > 0$,

$$\text{VaR}_\gamma(x) \doteq \inf\{m | P[x + m < 0] \leq \gamma\},$$

is not a convex measure of risk. This can be seen the example in [3], p. 218, since the acceptance set is not convex.

Proposition 4.2 *Assume that \mathcal{A} is a non-empty convex subset of X which satisfies property 2 of Proposition 4.1, and denote by $\rho_{\mathcal{A}}$ the functional associated to \mathcal{A} via 4.1. If $\rho_{\mathcal{A}}(0) > -\infty$, then*

- $\rho_{\mathcal{A}}$ is a convex measure of risk.
- \mathcal{A} is a subset of $\mathcal{A}_{\rho_{\mathcal{A}}}$. Moreover, if \mathcal{A} satisfies property 3 of Proposition 4.1, then $\mathcal{A} = \mathcal{A}_{\rho_{\mathcal{A}}}$
- If \mathcal{A} is a cone, then $\rho_{\mathcal{A}}$ is a coherent measures of risk.

4.1.2 The representation theorem for convex measures of risk

As in [3] for the case of coherent risk measure, here we give a representation theorem for convex measure of risk. We first consider the special case in which X is the space of all real-valued functions on some finite set Ω , while \mathcal{P} is the set of all probability measure on Ω .

Theorem 4.1 *Suppose X is the space of all real-valued functions on a finite set Ω . Then $\rho : X \rightarrow \mathbb{R}$ is a convex measure of risk if and only if there exist a "penalty function" $\alpha : \mathcal{P} \rightarrow (-\infty, +\infty]$ such that:*

$$\rho(Z) = \sup_{Q \in \mathcal{P}} (E_Q[-Z] - \alpha(Q)). \quad (4.2)$$

The function α satisfies $\alpha(Q) \geq -\rho(0)$ for any $Q \in \mathcal{P}$, and it can be taken to be convex and lower semi-continuous on \mathcal{P} .

Note that this theorem includes the structure theorem for coherent measure of risk as a special case. Indeed, it is easy to see that ρ satisfies the property of positive homogeneity, i.e. ρ will be a coherent measure of risk, if and only if the above penalty function $\alpha(\cdot)$ in (4.2) takes only value 0 and $+\infty$. In this case, our theorem implies the representation (3.3) in terms of the set

$$\mathcal{Q} = \{Q \in \mathcal{P} | \alpha(Q) = 0\}.$$

In the proof of this theorem in [14], the assumption that Ω is finite was only used to obtain the closedness of the acceptance set \mathcal{A}_{ρ} . In the case where X is given as the space $L^{\infty}(\Omega, \mathcal{F}, P)$ of bounded functions on a general probability space (Ω, \mathcal{F}, P) , we will have to *assume* the closedness of \mathcal{A}_{ρ} in a suitable topology, but then the previous argument goes through. Thus we obtain the following extension of Delbaen's representation theorem for coherent measure of risk on general probability spaces; see Theorem 3.2. Note that by defining $X = L^{\infty}(\Omega, \mathcal{F}, P)$, we fix a priori, on the probability space (Ω, \mathcal{F}) , the measure P .

Theorem 4.2 Suppose $X = L^\infty(\Omega, \mathcal{F}, P)$, \mathcal{P} is the set of probability measure $Q \ll P$, and $\rho : X \rightarrow \mathbb{R}$ is a convex measure of risk. Then the following properties are equivalent.

1. There is a "penalty function" $\alpha : \mathcal{P} \rightarrow (-\infty, +\infty]$ such that:

$$\rho(x) = \sup_{Q \ll P} (E_Q[-x] - \alpha(Q)) \text{ for all } x \in X. \quad (4.3)$$

2. The acceptance set \mathcal{A}_ρ associated with ρ is weak*-closed, i.e. $\sigma(L^\infty(P), L^1(P))$ -closed.
3. ρ possesses the Fatou property: If a sequence $(x_n)_{n \in \mathbb{N}} \subset X$ is uniformly bounded and converges to some $x \in X$, then $\rho(x) \leq \liminf_n \rho(x_n)$.
4. If the sequence $(x_n)_{n \in \mathbb{N}} \subset X$ decrease to x , then $\rho(x_n) \rightarrow \rho(x)$.

Proposition 4.3 Suppose $\rho : L^\infty(\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$ is a convex measure of risk possessing a representation of the form (4.3) and take \mathcal{P} as in Theorem 4.2. Then the representation (4.3) holds as well in term of the penalty function

$$\alpha_0(Q) = \sup_{x \in L^\infty} (E_Q[-x] - \rho(x)) = \sup_{x \in \mathcal{A}_\rho} E_Q[-x] \quad (4.4)$$

Moreover, it is minimal in the sense that $\alpha_0(Q) \leq \alpha(Q)$ for all $Q \in \mathcal{P}$ if the representation (4.3) holds for $\alpha(\cdot)$. In addition,

$$\alpha_0(Q) = \sup_{x \in \mathcal{A}_\rho} E_Q[-x] = \sup_{x \in \mathcal{A}} E_Q[-x] \quad (4.5)$$

if ρ is defined as in (4.1) via a given acceptance set \mathcal{A} .

4.1.3 Robust representation of convex measures of risk

We have seen the definition of convex risk measures as generalization of coherent ones. We have seen an extension from the case of possible finite scenarios to the case of a general probability space, too. In [11] and [14] financial positions are modelled as function of the space L^∞ with respect to a fixed probability measures P on a measurable space (Ω, \mathcal{F}) . In their other work [15], Föllmer and Schied, characterize measures of risk in a situation of uncertainty, without referring to a given probability measure. We will present now this extension ¹.

From now on we assume that X is the linear space of all bounded measurable function on a measurable space (Ω, \mathcal{F}) . We denote by $\mathcal{M}_1 \doteq \mathcal{M}_1(\Omega, \mathcal{F})$ the class of all probability measures on (Ω, \mathcal{F}) . Moreover, we introduce the larger class

¹For more details and proofs see [15]

$\mathcal{M}_{1,f} \doteq \mathcal{M}_{1,f}(\Omega, \mathcal{F})$ of all finitely additive and non-negative set functions Q on \mathcal{F} which are normalized to $Q[\Omega] = 1$. We emphasize that no probability measure on (Ω, \mathcal{F}) is fixed in advance.

In this general context, the following characterization of coherent risk measure is essentially well known; see e.g. Theorem 3.2 or [11].

Proposition 4.4 *A functional $\rho : X \rightarrow \mathbb{R}$ is a coherent measure of risk if and only if there exists a subset \mathcal{Q} of $\mathcal{M}_{1,f}$ such that:*

$$\rho(x) = \sup_{Q \in \mathcal{Q}} E_Q[-x], \quad x \in X. \quad (4.6)$$

Moreover, \mathcal{Q} can be chosen as a convex set for which the supremum in (4.6) is attained.

The first goal is to obtain an analogous result for *convex* risk measures.

Let $\alpha : \mathcal{M}_{1,f} \rightarrow \mathbb{R} \cup \{+\infty\}$ be any functional which is bounded from below and which is not identically equal to $+\infty$. For each $Q \in \mathcal{M}_{1,f}$ the functional $X \rightarrow E_Q[-x] - \alpha(Q)$ is convex, monotone, and translation invariant on X , and these three properties are preserved when taking the supremum over $Q \in \mathcal{Q}$. Hence

$$\rho(x) \doteq \sup_{Q \in \mathcal{M}_{1,f}(\Omega)} (E_Q[-x] - \alpha(Q)) \quad (4.7)$$

defines a convex risk measures of risk on X . The function α will be called a *penalty function* for ρ on $\mathcal{M}_{1,f}$, and we will say that ρ is *represented* by α on $\mathcal{M}_{1,f}$.

Theorem 4.3 *Any convex measure of risk ρ on X is of the form*

$$\rho(x) = \max_{Q \in \mathcal{M}_{1,f}} (E_Q[-x] - \alpha_{\min}(Q)), \quad x \in X, \quad (4.8)$$

where the penalty functional α_{\min} is given by

$$\alpha_{\min} \doteq \sup_{x \in \mathcal{A}_\rho} E_Q[-x], \quad \text{for } Q \in \mathcal{M}_{1,f}.$$

Moreover, $\alpha_{\min}(Q)$ is the minimal penalty function which represents ρ , i.e. any penalty function α for which 4.7 satisfies $\alpha(Q) \geq \alpha_{\min}(Q)$ for all $Q \in \mathcal{M}_{1,f}$.

Remark 4 1. As done in Section 4.1.2, we can obtain an equivalent representation for α_{\min} as in (4.4) and (4.5), for all $Q \in \mathcal{M}_{1,f}$.

2. The representation (4.6) is a particular case of the representation theorem for convex measures of risk, since it corresponds to the penalty function

$$\alpha(Q) = \begin{cases} 0 & \text{if } Q \in \mathcal{Q} \\ +\infty & \text{otherwise} \end{cases}$$

The following corollary describes the minimal penalty function when dealing with coherent risk measures.

Corollary 1 *The minimal penalty function α_{min} of a coherent risk measure ρ takes only the values 0 and $+\infty$. In particular,*

$$\rho(x) = \max_{Q \in \mathcal{Q}_{max}} E_Q[-x], \quad x \in X,$$

for the weakly closed convex set

$$\mathcal{Q}_{max} \doteq \{Q \in \mathcal{M}_{1,f} | \alpha_{min}(Q) = 0\},$$

and \mathcal{Q}_{max} is the largest set for which the representation of the form (4.6), seen as a particular case of convex risk measure, holds.

In the sequel, we are particularly interested in the situation where a convex measure of risk ρ admits a representation in term of σ -additive probability measure, i.e., it can be represented by a penalty function α which is infinite outside the set $\mathcal{M}_1 = \mathcal{M}_1(\Omega, \mathcal{F})$:

$$\rho(x) = \sup_{Q \in \mathcal{M}_1} (E_Q[-x] - \alpha(Q)). \quad (4.9)$$

A representation (4.9) in terms of probability measures is closely related to certain continuity properties of ρ .

Remark 5 *A convex measure of risk ρ which admits a representation (4.9) on \mathcal{M}_1 is continuous from above in the sense that*

$$x_n \downarrow x \quad \Longrightarrow \quad \rho(x_n) \uparrow \rho(x).$$

Moreover, continuity from above is equivalent to the lower semi-continuity with respect to bounded pointwise convergence: If (x_n) is a bounded sequence in X which converges pointwise to $x \in X$, then

$$\rho(x) \leq \liminf_{n \rightarrow \infty} \rho(x_n).$$

The following proposition gives a sufficient condition that shows that every penalty function for ρ is concentrated on the set \mathcal{M}_1 of probability measures. This condition in "continuity from below" rather than from above.

Proposition 4.5 *Let ρ be a convex measure of risk which is continuous from below in the sense that*

$$\rho(x_n) \downarrow \rho(x) \quad \text{whenever} \quad x_n \uparrow x,$$

and suppose that α is any penalty function on $\mathcal{M}_{1,f}$ representing ρ . Then α is concentrated on probability measures in the usual sense, i.e.,

$$\alpha(Q) < \infty \quad \Longrightarrow \quad Q \text{ is } \sigma\text{-additive.}$$

Remark 6 Any convex measure of risk ρ that is continuous from below is also continuous from above, as can be seen by combining Proposition 4.5 and Remark 5. Thus a straightforward argument yields that $\rho(x_n) \rightarrow \rho(x)$ whenever (x_n) is a bounded sequence in X which converges pointwise to x .

4.1.4 Risk Measures defined in terms of shortfall risk

In this section, we will establish a relation between convex measure of risk and utility function.

Suppose that a risk-averse investor assesses the downside risk of a financial position $x \in X$ by taking the expected utility $E[u(-x^-)]$ derived from the *shortfall* x^- , or by considering the expected utility $E[u(x)]$ of the position itself. Recall that if an agent is risk-averse then u , the utility function, is strictly concave. If the focus is on the downside risk, then it's natural to change the sign and replace u by the loss function $l(x) \doteq -u(-x)$. Then l is a strictly convex and increasing function, and the maximization of the expected utility is equivalent to minimizing the expected loss $E[l(-x)]$ or the *shortfall risk* $E[l(x^-)]$. In order to unify the discussion of both cases, we do not insist on *strict* convexity. In particular, l may vanish $(-\infty, 0]$, and in this case the shortfall risk takes the form

$$E[l(x^-)] = E[l(-x)].$$

Definition 4.3 A function $l : \mathbb{R} \rightarrow \mathbb{R}$ is called a loss function if it is increasing and not identically constant.

In this section we will only consider *convex* loss function. Let u_0 an interior point in the range of l . A position $x \in L^\infty(\Omega, \mathcal{F}, P)$ will be called *acceptable* if the expected loss is bounded by u_0 . Thus, we consider the class

$$\mathcal{A} \doteq \{x \in L^\infty(\Omega, \mathcal{F}, P) | E_P[l(-x)] \leq u_0\}. \quad (4.10)$$

of acceptable position. The set \mathcal{A} satisfies the first two properties of Proposition 4.1 and thus define a convex measure of risk $\rho \doteq \rho_{\mathcal{A}}$. Since l is continuous as a finitely value convex function on \mathbb{R} , ρ possesses the Fatou property and, hence, representation of the form (4.3).

In a robust representation, we define the *loss functional* L :

$$L(x) = \sup_{Q \in \mathcal{Q}} E_Q[l(-x)]$$

where \mathcal{Q} is a set of probability measures on (Ω, \mathcal{F}) . As done before, a position x is acceptable if $L(x)$ does not exceed a given bound x_0 . So let us consider the convex class:

$$\mathcal{A}_L \doteq \{x \in X | L(x) = \sup_{Q \in \mathcal{Q}} E_Q[l(-x)] \leq u_0\}$$

of acceptable positions, where, as above, x_0 is in the range of l . Applying a well known result we can conclude that ρ admits a representation of the form

$$\rho(x) = \sup_{Q \in \mathcal{M}_1} (E_Q[-x] - \alpha_L(Q)).$$

Thus, also in this case, the problem is reduced to compute of a suitable penalty function.

In both cases, the corresponding penalty function $\alpha_0(\cdot)$ can be expressed in terms of the Fenchel-Legendre transform

$$l^*(z) \doteq \sup_{x \in \mathbb{R}} (zx - l(x))$$

of l ;

In this context, the penalty function can be expressed in the following form:

Theorem 4.4 *Suppose that \mathcal{A} is the acceptance set given by (4.10). Then, for $Q \ll P$, the minimal penalty function of $\rho = \rho_{\mathcal{A}}$ is given by:*

$$\alpha_0(Q) = \sup_{x \in \mathcal{A}} E_Q[-x] = \inf_{\lambda > 0} \frac{1}{\lambda} \left(x_0 + E_P[l^*(\lambda \frac{dQ}{dP})] \right). \quad (4.11)$$

and in a robust context:

Theorem 4.5 *The convex risk measure corresponding to the acceptance set \mathcal{A} can be represented in terms of penalty function*

$$\alpha_L(P) = \inf_{\lambda > 0} \frac{1}{\lambda} \left(x_0 + \inf_{Q \in \mathcal{Q}} E_Q[l^*(\lambda \frac{dP}{dQ})] \right). \quad (4.12)$$

where dP/dQ is a generalized density in the sense of the Lebesgue decomposition. Thus, $\alpha_L(P) < \infty$ only if $P \ll Q$ for at least some $Q \in \mathcal{Q}$.

Remark 7 *A bit of confusion could arise because in the two theorems above the roles of Q and P are one the opposite of the other. In fact in Theorem 4.4 we have $Q \ll P$, $P \in \mathcal{P}$. In Theorem 4.5 we take $Q \in \mathcal{Q}$, $\mathcal{Q} \subset \mathcal{M}_1$ as the measure that dominates the measure P . This to be coherent with the notation used when dealing with robust representation of convex measure of risk.*

4.2 The Frittelli and Rosazza Gianin (2002) convex risk measure

Independently from [14], Frittelli and Rosazza Gianin gave in their work [18] a characterization of a convex measure of risk. Their approach is different for the

4.2 The Frittelli and Rosazza Gianin (2002) convex risk measure 31

choice of the others axioms (clearly the axiom of convexity is in both [14] and [18]). Frittelli, moreover, bases almost all of his results on the duality between ρ and the penalty function $F(\alpha$ in [14]).

We will follow, in the exposition, [18] and [19]. For the proofs see [19].

4.2.1 Notation

Let T be a, fixed in advance, future date and let X be an ordered locally convex topological vector space that represents the “habitat” of all the financial positions whose riskiness we want to quantify.

Assume that X is endowed with a topology τ for which X and its topological dual space X' , of all continuous linear functionals on X , form a dual system. Although most of the results hold in an ordered locally convex topological vector space, we will assume for simplicity that:

$$\begin{aligned} X &= L^p(\Omega, \mathcal{F}, P), & 1 \leq p \leq \infty \\ X' &\subseteq L^1(\Omega, \mathcal{F}, P), \end{aligned}$$

where (Ω, \mathcal{F}, P) is a probability space.

If the sample space Ω is finite (say, its cardinality is n), then $X = X' = \mathbb{R}^n$. Other examples of possible settings are: $X = L^p(\Omega, \mathcal{F}, P)$ and $X' = L^q(\Omega, \mathcal{F}, P)$ where $p \in (1, +\infty)$, p and q are conjugate, and τ is the norm topology in $L^p(\Omega, \mathcal{F}, P)$; or $X = L^\infty(\Omega, \mathcal{F}, P)$ and $X' = L^1(\Omega, \mathcal{F}, P)$ and $\tau = \sigma(L^\infty, L^1)$.

We denote by $\mathbf{1}$ the random variable P -a.s. equal to 1, with “ \leq ” the natural preorder on the vector space X given by inequalities that hold P -a.s. We notice that, in this case, a probability measure is fixed a priori, too.

Let X'_+ the set formed with all the positive continuous linear functionals on X , that is

$$X'_+ \doteq \{x' \in X' \mid x'(x) \leq 0 \ \forall x \in X : x \leq 0\},$$

and

$$\mathcal{Z} \doteq \{x' \in X'_+ : x(\mathbf{1}) = 1\}$$

be the set of all probability densities in X' . By the Radon-Nikodym theorem (see Theorem A.10), we may identify any probability density $x' \in \mathcal{Z}$ with its associated probability measure P' by setting $\frac{dP'}{dP} = x'$. Hence, $x'(\cdot)$ is simply the expected value $E_{P'}[\cdot]$, namely

$$x'(x) = E_P[x'x] = E_{P'}[x], \text{ if } x' \in \mathcal{Z}.$$

4.2.2 Convex Risk Measures

As done above, we now give the definition of convex risk measure arising in this context.

Definition 4.4 *A functional $\rho : X \rightarrow \mathbb{R}$ is a convex risk measure if ρ satisfies axioms: (a) convexity, (b) lower semi-continuity and (g) normalization.*

Remark 8 *A convex risk measure ρ satisfies :*

- (i) $\rho(\alpha x) \leq \alpha \rho(x)$, $\forall \alpha \in [0, 1]$, $\forall x \in X$;
- (ii) $\rho(\alpha x) \geq \alpha \rho(x)$, $\forall \alpha \in (-\infty, 0] \cup [1, +\infty]$, $\forall x \in X$;
- (iii) $\rho(x - y) \geq -\rho(-x) - \rho(y)$, $\forall x, y \in X$;
- (iv) if $(x \geq 0 \Rightarrow \rho(x) \leq 0)$ then $(x \leq 0 \Rightarrow \rho(x) \geq 0)$ (i.e. $(c) \Rightarrow (c^o)$).

Lemma 2 *For a convex risk measure, the following couples of axioms are equivalent:*

- (c) *positivity and (c1) *monotonicity;
- (d) subadditivity and (d1) positivity;
- (e) translation invariance and (e1) constancy.

Remark 9 *Let $\rho : X \rightarrow \mathbb{R}$ satisfies the translation invariance axiom (e). Then:*

- (i) *The l.s.c. axiom (b) is equivalent to $\{x \in X : \rho(x) \leq 0\}$ is closed in X .*
- (ii) *The following statements are equivalent:*

- α) ρ is convex (axiom (a));
- β) ρ is quasi convex (i.e.: $\{x \in X : \rho(x) \leq c\}$ is convex for all $c \in \mathbb{R}$);
- γ) $\{x \in X : \rho(x) \leq 0\}$ is convex;

4.2.3 Representation of convex risk measures

With the following result (see Frittelli and Rosazza Gianin [18] Theorem 6 and Corollary 7), we provide the characterization of convex and sublinear measures of risk.

Theorem 4.6

1. $\rho : X \rightarrow \mathbb{R}$ is a convex measure of risk if and only if there exists a convex functional $F : X' \rightarrow \mathbb{R} \cup \{+\infty\}$, satisfying $\inf_{x' \in X} F(x') = 0$, such that

$$\rho(x) = \sup_{x' \in \mathcal{P}} \{x'(x) - F(x')\} < +\infty, \quad \forall x \in X, \quad (4.13)$$

where $\mathcal{P} = \{x' \in X' : F(x') < +\infty\}$ is the effective domain of F .

2. $\rho : X \rightarrow \mathbb{R}$ is a sublinear and lower semi-continuous risk measure (i.e. ρ satisfies axioms (b),(d) and (d1)) if and only if ρ is representable as in 4.13 with $F \equiv 0$ on \mathcal{P} , i.e.

$$\rho(x) = \sup_{x' \in \mathcal{P}} \{x'(x)\} < +\infty, \quad \forall x \in X, \quad (4.14)$$

This shows that the representation in (4.13)(similar but more general than (3.2)) holds true with axioms that are much weaker than the coherence ones. Although only the convexity and the lower semi-continuity axioms are necessary to represent ρ as in (4.13), one might be interested in other properties. The following result shows how further axioms come into play in the representation (4.13) and (4.14).

Corollary 2 *If $\rho : X \rightarrow \mathbb{R}$ is a convex risk measure, then:*

- (i) ρ satisfies (c) *positivity iff we have $\mathcal{P} \subseteq X'_+$ in (4.13);
- (ii) ρ satisfies (e) translation invariance iff we have $\mathcal{P} \subseteq \{x' \in X' : x'(\mathbf{1}) = 1\}$ in (4.13);
- (iii) ρ satisfies (c) and (e) iff we have $\mathcal{P} \subseteq \mathcal{Z}$ in (4.13).

If ρ is sublinear and l.s.c.(i.e. satisfies axioms (b),(d) and (d1)), then:

- (iv) ρ satisfies (c) *positivity iff we have: $\mathcal{P} \subseteq X'_+$ in (4.14).
- (v) ρ satisfies (c) *positivity and (e) translation invariance iff we have: $\mathcal{P} \subseteq \mathcal{Z}$ (this is exactly the case of coherent risk measures) in (4.14).

From Corollary 2 we deduce the following:

Corollary 3 *$\rho : X \rightarrow \mathbb{R}$ is a convex risk measure satisfying the axiom (c) *positivity and (e) translation invariance iff there exists a convex set of probability measures \mathcal{P} and a convex functional $F : \mathcal{P} \rightarrow \mathbb{R} \cup \{+\infty\}$ satisfying $\inf\{F(Q)|Q \in \mathcal{P}\} = 0$ and*

$$\rho(x) = \sup_{Q \in \mathcal{P}} \{E_Q[-x] - F(Q)\} < +\infty, \quad \forall x \in X. \quad (4.15)$$

The representation in (4.15) has an easily financial interpretation (as well as representation in (4.2)). Indeed ρ is the supremum over a set \mathcal{P} of scenarios of the expected loss “correct” with a “penalty” term $F(\alpha$ in [14]) which depends on the scenarios. Moreover, while the set \mathcal{P} of possible scenarios could be exogenously determined, for example by some regulatory institutions or by the market itself, the functional F (or α) could be determined by the investors (by mean of their preferences and utility or loss functions).

4.3 Differences and similarities between the two definitions of convex measure of risk

We have seen these two different approaches to convex risk measures. These two ways of proceeding clearly start from different assumptions but in the end they converge to the same representation. We are now going to analyze what are the differences and the similarities between both approaches. Moreover, we will be able to state that the two convex risk measures coincides². The obvious starting point is the different axiomatization, but we will later focus our attention on the so called “penalty function” $\alpha(\cdot)$ (or equivalently $F(\cdot)$).

4.3.1 About the axioms

We have already repeatedly said that these two definitions of convex risk measure differ because the choice of the other axioms to put together with the one of convexity. We will show that the necessary and sufficient axioms to define a convex measure of risk are: (a) convexity and (b) lower semi-continuity. Thus different definitions (and consequently representations) arise if we impose or not other axioms. And clearly, this further imposition depends on what one is interested in.

The assumption of these further axioms only modifies the functional F and the set \mathcal{P} over which the supremum is taken.

In [14], to define a convex risk measure, were used the axioms: (a) convexity, (e) translation invariance and (c1) *monotonicity. And this led to a representation for ρ as the supremum over a set \mathcal{P} of all probability on Ω (generalize scenarios), finite, of the expected value of the worst cases ”corrected” with a penalty function. This penalty function can be taken to be convex and lower semi-continuous. In the case of general probability spaces the supremum has to be taken over the set \mathcal{P} , the set of all the probability measure $Q \ll P$, with $X = L^\infty(\Omega, \mathcal{F}, P)$.

²Starting form now, when referring to a convex measure of risk, we will consider ρ as defined in representations (4.2) or (4.15)

In [18] to define a convex measure of risk, were used the axioms: (a) convexity, (b) lower semi-continuity and (g) normalization. The result in Theorem 4.6 leads to a representation that is a very well-known result from convex analysis (see Theorem B.6). This representation is more general than (4.3), and an interpretation in term of supremum over a set \mathcal{P} of all probability on Ω (general case), of an expected value corrected by a penalty function is not immediate. Thus, to get such a representation, as done in Corollary 3, we have to impose other conditions. In particular to get (4.15), ρ , convex risk measure as defined in Definition 4.4, has satisfy the axioms (c) *positivity and (e) translation invariance. Moreover the set \mathcal{P} of probability measures such that $Q \ll P$, $Q \in \mathcal{P}$, has to be convex.

We recall that, by Lemma 2(c), for a convex risk measure as defined in Definition 4.4 the axioms of (c) *positivity and (c1) * monotonicity are equivalent. Then these axioms hold in both representations.

4.3.2 Comparison between the two representations

So, starting from the consideration above, we will see how the two representation coincides. We will deal with formulas (4.3) and (4.15). In fact, in this case, we have a general probability space, while in (4.2) Ω was supposed to be finite. Then $x \in L^\infty(\Omega, \mathcal{F}, P)$ in (4.3) and $x \in L^p(\Omega, \mathcal{F}, P)$, $1 \leq p \leq +\infty$ in (4.15). Let assume in this last setting: $x \in L^\infty(\Omega, \mathcal{F}, P)$, $x' \in L^1(\Omega, \mathcal{F}, P)$ and $\tau = \sigma(L^\infty, L^1)$.

Both the representations are in the form of the supremum over a set \mathcal{P} of the expected value of the worst cases minus a penalty function:

- (a) \mathcal{P} is the set of probability measures $Q : Q \ll P$. In (4.15) we have $\mathcal{P} \subset \mathcal{Z}$, with \mathcal{Z} the set of all probability densities $x' \in X'$. Recall, see subsection 4.2.1, that $x' = \frac{dQ}{dP}$, $x'(x) = E_Q[x]$, if $x' \in \mathcal{Z}$. Then $Q \ll P$.
- (b) the supremum is calculated over \mathcal{P} ;
- (c) the expected value is computed with respect to Q , $Q \ll P$;
- (d) the penalty function is $\alpha : \mathcal{P} \rightarrow (-\infty, +\infty]$ (resp. F).

We have seen that, starting from Definition 4.4, by adding additional conditions (axioms), we get to a representation for ρ equivalent to (4.2). It would be natural to state that, admitting the same representation, these two ways of defining a convex measure of risk coincide. Actually, they differ for the axioms (g) normalization and (b) lower semi-continuity.

We will show that axiom (b) is common in both definition and that axiom (g) can be imposed in Definition 4.1 without loss of generality. In fact, it only modifies the lower bound of the penalty function.

About lower semi-continuity

In Definition 4.4 axiom (b) lower semi-continuity is nothing but a technical requirement to apply duality theorem and get a representation as in (4.13). We can see representation (4.3) as a particular case of (4.13). In Theorem 4.2 we have that ρ possesses the Fatou property, and then is lower semi-continuous.

About the normality

In representation (4.3) we have $\alpha(Q) \geq -\rho(0)$. If we take, as in Definition 4.4, $\rho(0) = 0$, then

$$\alpha(Q) \geq -\rho(0) = 0$$

or equivalently

$$0 = \rho(0) = \sup_{Q \in \mathcal{P}} \{E_Q[0] - \alpha(Q)\} = \sup_{Q \in \mathcal{P}} \{-\alpha(Q)\} = -\inf_{Q \in \mathcal{P}} \{\alpha(Q)\}$$

and thus

$$\inf_{Q \in \mathcal{P}} \alpha(Q) = 0$$

We have now seen the fact that the imposition of axiom (g) of normalization only change the inferior bound of the penalty function.

Both the definitions imply the lower semi continuity of ρ . The only difference is the axiom (g) of normalization. However to impose $\rho(0) = 0$ is not a loss of generality and moreover has a sensible financial meaning.

Remark 10 *Recall that, by Remark 2, to get axiom (g) normality it is sufficient to impose axiom (e1) constancy in Definition 4.1.*

Chapter 5

Exponential utility, loss function and relative entropy

In this chapter we will as application of the convex measures of risk: the entropic risk measures. We will show, starting from the two different definitions of convex risk measures, two ways to obtain the entropic convex risk measure. We will also see that, in a particular context, both coincide. We first recall the notion of *shortfall risk* from [14]. We will then see a relationship between the convex risk measure defined by the acceptance set in (4.10) and the Certainty Equivalent. Starting from [17] we will present the notion of “Dynamic Certainty Equivalent” and it will lead us to a convex risk measure.

5.1 Shortfall Risk and Certain Equivalent

Recall that, for a risk measure defined in terms of shortfall risk, the set of acceptable position is

$$\mathcal{A} \doteq \{x \in L^\infty(\Omega, \mathcal{F}, P) \mid E_P[l(-x)] \leq u_0\}.$$

where u_0 is an interior point in the range of l and $x \in L^\infty(\Omega, \mathcal{F}, P)$; $l : \mathbb{R} \rightarrow \mathbb{R}$ is an increasing convex loss function; the expected loss of a position $x \in X$ is $E_P[l(-x)]$. Let us take $l(u) = e^u$ and $u_0 = 1$ so that

$$\begin{aligned} \rho(x) &= \inf\{m \in \mathbb{R} \mid E_P[e^{-m-x}] \leq 1\} = \\ &= \inf\{m \in \mathbb{R} \mid e^{-m} E_P[e^{-x}] \leq 1\} = \\ &= \inf\{m \in \mathbb{R} \mid E_P[e^{-x}] \leq e^m\} = \\ &= \inf\{m \in \mathbb{R} \mid m \geq \ln E_P[e^{-x}]\} = \end{aligned}$$

$$= \ln E_P[e^{-x}].$$

We know, by (4.4) that

$$\alpha(Q) = \sup_{x \in L^\infty(P)} \{E_Q[-x] - \ln E_P[e^{-x}]\} = H(Q, P) \quad (5.1)$$

where the relative entropy of Q with respect to P is defined as

$$H(Q, P) \doteq \begin{cases} E_Q \left[\ln \frac{dQ}{dP} \right] = E_P \left[\frac{dQ}{dP} \ln \frac{dQ}{dP} \right] & \text{if } Q \ll P, \\ +\infty & \text{otherwise.} \end{cases}$$

Remark 11 *Jensen's inequality applied to the strictly convex function $h(u) = u \ln u$ yields*

$$H(Q, P) = E \left[h \left(\frac{dQ}{dP} \right) \right] \geq h(1) = 0. \quad (5.2)$$

with the equality if and only if $Q = P$.

Let us prove that the supremum in (5.1) is less or equal to the relative entropy of Q with respect to P (for the opposite inequality see the proof of Lemma 3.31, p.127 in [16]).

Assume that $H(Q, P) < +\infty$, i.e. $Q \ll P$. Let us take $x \in X$ such that $E_P[e^{-x}] < +\infty$ and define a probability measure P^x in the following way:

$$\frac{dP^x}{dP} = \frac{e^{-x}}{E_P[e^{-x}]}. \quad (5.3)$$

By the assumptions on x , P^x is equivalent to the probability measure P and it holds

$$\ln \frac{dQ}{dP} = \ln \frac{dQ}{dP^x} + \ln \frac{dP^x}{dP}. \quad (5.4)$$

Integrating with respect to Q , we obtain

$$H(Q, P) = H(Q, P^x) + E_Q[-x] - \ln E_P[e^{-x}]. \quad (5.5)$$

Since $H(Q, P) \geq 0$ by (5.2), we have proved that $H(Q, P)$ is larger than or equal to both suprema on the right of (5.1).

On the other hand, when dealing with a robust representation of convex measures, we have

$$\alpha_L(Q) = \inf_{Q \in \mathcal{Q}} H(P, Q). \quad (5.6)$$

So we have defined a classical example of convex risk measures: the *entropic risk measure*:

$$\rho(x) = \sup_{Q \in \mathcal{P}} \{E_Q[-x] - H(Q, P)\}. \quad (5.7)$$

We can also see that, starting from the negative exponential utility function $u(x) = -e^{-x}$, we get

$$\rho(x) = \ln E_P[e^{-x}] = \ln E_P[-u(x)] = -\mathbb{E}(x) \quad (5.8)$$

where $\mathbb{E}(x)$ is the so called *Certain Equivalent* of the claim x . $\mathbb{E}(x)$ is the value such that

$$u(\mathbb{E}(x)) = E_P[u(x)]$$

Solving we get

$$\begin{aligned} -e^{-\mathbb{E}(x)} &= E_P[-e^{-x}] \\ \mathbb{E}(x) &= -\ln E_P[e^{-x}] = -\ln E_P[-u(x)] = -\rho(x) \end{aligned}$$

5.2 Certainty Equivalent and Exponential Utility Function

We will follow, in the exposition, Frittelli's work [17]. A (non empty) family Σ of adapted stochastic processes on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_{t \in [0, T]}, P)$ represents prices in the market. Set $\mathcal{I} = [0, T]$ and $\mathcal{F} = \mathcal{F}_T$. We will assume for simplicity (differently from [17]) that the risk-free interest rate is zero. We will consider the family χ of price processes and note that, by construction, χ contains at least the constant process equal to 1.

We denote by \mathcal{M} the set of probability measures Q absolutely continuous with respect to P such that all processes in χ are (\mathcal{F}_t, Q) -martingales. We assume the existence of a martingale measure equivalent to P .

With $u : \mathcal{D} \rightarrow \mathbb{R}$ we always denote a *non decreasing* real function defined on an interval $\mathcal{D} \subseteq \mathbb{R}$ with nonempty interior and taking value $-\infty$ on the external points of \mathcal{D} . We denote by $L^- = L^-(\Omega, \mathcal{F}_T, P)$ the set of lower bounded random variables. A T -claim ω is an element of L^- .

Let ω be a time T -claim that we want to price. The function u is the time- T utility. We can distinguish several alternatives:

- Totally incomplete market

In this case the possibility of trading in the available market assets does not provide any help for hedging (not even partially) the risk carried by ω . In this case the subjective value of ω is traditionally assigned by the certain amount $\pi(\omega) \in \mathbb{R}$ whose utility is equal to the expected one of the claim ω :

$$u(\pi(\omega)) = E[u(\omega)]. \quad (5.9)$$

The agent can't take advantage of the presence of market securities.

- Complete market

If a bounded claim ω is attainable by a self-financing strategy in the traded assets or if the market is complete, the value of the claim is independent of agents preferences and it is univocally assigned by the formula:

$$\pi(\omega) = E_Q[\omega]$$

where Q is any martingale measure (eventually unique if the market is complete).

- Incomplete market

In incomplete market, to determine the value of the claim, the agent has to take into consideration his subjective preferences. However, he may partially hedge the risk carried by ω by trading on the available securities. The presence of these securities will affect the pricing of the claim. Indeed, the no arbitrage principle imposes restrictions on the admissible prices: in order to prevent arbitrage opportunities, the value of ω must lie in the interval

$$\left[\inf_{Q \in \mathcal{M}} E_Q[\omega], \sup_{\mathcal{M}} E_Q[\omega] \right]$$

The aim is to construct a Theory of Value based on agents preferences and coherent with the no arbitrage principle. The idea is to embedding incomplete market asset pricing via utility maximization

Let x_0 be the initial capital of an agent. For a given random variable (or a real number) y we define the budget constraints set as

$$\Theta(y) \doteq \{z : z \preceq y\}.$$

The maximum attainable utility from x_0 with the T claim ω are respectively given by

$$V_0(x_0) \doteq \sup_{z \in \Theta(x_0)} E[u(z)] \quad (5.10)$$

$$V(\omega) \doteq \sup_{z \in \Theta(\omega)} E[u(z)] \quad (5.11)$$

Clearly $V_0(\cdot) : \mathcal{D} \rightarrow \mathbb{R}$ as a real function of a real variable is not decreasing. One possible interpretation of the above maximization is problem is that the agent holding the contract corresponding to the T -claim ω may sell in the market this contract and buy another contract corresponding to any T -claim $z \in \Theta(\omega)$, since claims in $\Theta(\omega)$ have prices less than or equal to ω . In the spirit of Equation (5.9), $x_0 = \pi(\omega)$ is the value of ω if it satisfies the equation:

$$V_0(\pi(\omega)) = V(\omega).$$

Definition 5.1 Given a partial preorder \preceq on L^- , consider $\Theta(x_0), V_0(x_0), V(\omega)$ as defined in Equations (5.9)-(5.11). Define the value $\pi(\omega) \in \mathbb{R}$ of ω as the solution of the equation

$$V_0(\pi(\omega)) = V(\omega). \quad (5.12)$$

A natural candidate for the partial preorder, such that the value $\pi(\omega)$ assigned in Equation (5.12) is compatible with no arbitrage principle, is given by:

Definition 5.2 (Market Preorder) The market preorder is the partial preorder on the set of claims defined by:

$$z \preceq \omega \Leftrightarrow E_Q[z] \leq E_Q[\omega] \quad Q \in \mathcal{M}.$$

The value $\pi(\omega)$ of the claim ω is given in Definition 5.1 with the partial preorder assigned by the market preorder.

We will only analyze the case of totally incomplete market because the rest of the cases is beyond our interests.

5.2.1 Main Definitions

Frittelli defines the *Dynamic Certainty Equivalent* as the value $\pi(\omega)$ of the T -claim ω at time 0. Proposition 4 in [17] guarantees the existence and the uniqueness of $\pi(\omega)$, as the solution of the equation (5.12).

We present a property of the value $\pi(\omega)$ we will find useful later (see part (e) of Proposition 7 in [17]).

Proposition 5.1 If $\mathcal{M} = \{Q : Q \ll P\}$ (totally incomplete market) then $\pi(\omega)$ is the solution of

$$E[u(\pi(\omega))] = E[u(\omega)]$$

and if u is strictly increasing then $\pi(\omega) = \mathbb{E}(\omega)$.

There are three cases: $\mathbb{E}(\omega)$ is additive with respect to a constant, $\mathbb{E}(\omega)$ is positively homogeneous and the case of a linear utility function. The first case is of our interest because, by the Nagumo-Kolmogoroff-De Finetti theorem (see [8]), we know that if u is exponential, $\mathbb{E}(\omega)$ is additive with respect to a constant. Then, from Proposition 8 (b) in [17], we get:

Proposition 5.2 If the Certainty Equivalent is additive with respect to a constant ($\mathbb{E}(\omega + k) = \mathbb{E}(\omega) + \mathbb{E}(k)$, $k \in \mathbb{R}$), or equivalently if the utility is exponential, then

$$\pi(\omega + k) = \pi(\omega) + \pi(k).$$

5.2.2 Duality

The computation of the value $\pi(\omega)$ is based in the Fenchel Duality theorem.

Recall that the *concave conjugate* $u^* : \mathbb{R} \rightarrow \mathbb{R} \cup \{-\infty\}$ of u is given by:

$$u^*(x^*) \doteq \inf_{x \in \mathbb{R}} \{xx^* - u(x)\}, \quad x^* \in \mathbb{R}$$

Definition 5.3 *Let $Q \ll P$ and set*

$$\Delta(Q, P; x) = \min_{\lambda \in (0, +\infty)} \left\{ \lambda x - E \left[u^* \left(\lambda \frac{dQ}{dP} \right) \right] \right\}. \quad (5.13)$$

Then by Theorem 10 in [17], we have the following characterization of $\pi(\omega)$:

Corollary 4 *The value $\pi(\omega)$ is the solution of*

$$\inf_{Q \in \mathcal{M}} \{\Delta(Q, P; \pi(\omega))\} = \inf_{Q \in \mathcal{M}} \{\Delta(Q, P, E_Q[\omega])\}$$

5.2.3 Generalized distances $\delta(Q, P)$ and examples

In the following example each utility function determines by duality a “generalized distance” between probability measures. We need to calculate $\Delta(Q, P; x)$ to determine $\pi(\omega)$.

When u is strictly increasing we define the following quantity:

$$\delta(Q, P; x) = u^{-1}(\Delta(Q, P; x)) - x. \quad (5.14)$$

The advantages of this simple transformation are shown the following proposition (see Proposition 13 in [17] for more details):

Proposition 5.3 (a) *If the Certainty Equivalent is additive with respect to constants, then*

$$\delta(Q, P; x_1) = \delta(Q, P; x_2), \quad \forall x_1, x_2 \in \mathcal{D}.$$

(b) *In case (a) the functional $\delta(\cdot, P; x)$ is convex.*

(c) $\delta(Q, P; x) \geq 0 \forall Q \ll P \forall x \in \mathcal{D}$.

(d) *If $Q = P$ then $\delta(Q, P; x) = 0 \forall x \in \mathcal{D}$.*

(e) *In case (a) we have*

$$\delta(Q, P; 1) = 0 \Leftrightarrow Q = P.$$

If the utility function is sufficiently regular, formula (5.13) can be rewritten more explicitly and the computation of Δ and δ simplified. Thus we have:

$$\begin{aligned} \Delta(Q, P; x) &= E[u(I(\lambda^* \varphi))], \\ \delta(Q, P; x) &= u^{-1}(E[u(I(\lambda^* \varphi))]) - x. \end{aligned}$$

where: let $Q \ll P$ and $\varphi = \frac{dQ}{dP}$; $x \in \text{int}(\mathcal{D})$, suppose that $u : \mathcal{D} \rightarrow \mathbb{R}$ is a strictly increasing, strictly concave, differentiable function. We denote by $I = (u')^{-1}$ the inverse function of u' and by $\lambda^* = \lambda(x, \varphi)$ the unique solution of the equation

$$E[\varphi I(\lambda\varphi)] = x,$$

We present now an example where the utility function is exponential.

Example 2 (Exponential Utility) Let $Q \ll P$, $\varphi = \frac{dQ}{dP}$ and set $\delta(Q, P) = \delta(Q, P; 1)$. Let $u(x) = -e^{-x}$, $\mathcal{D} = (-\infty, +\infty)$. Then

$$\delta(Q, P) = H(Q, P)$$

where $H(Q, P)$ is the relative entropy.

5.2.4 Computation of the value

From Corollary 4 and the definition of δ given in Equation (5.14) we get:

Corollary 5 For a strictly increasing u , the value $\pi(\omega)$ is the solution of:

$$\begin{aligned} & \inf_{Q \in \mathcal{M}} \{\delta(Q, P, \pi(\omega)) + \pi(\omega)\} \\ &= \inf_{Q \in \mathcal{M}} \{\delta(Q, P, E_Q[\omega]) + E_Q[\omega]\} \end{aligned}$$

Applying Corollary 5 and Proposition 5.3 we have (see Corollary 15 in [17] for more details):

Corollary 6 When the utility is exponential, then the value $\pi(\omega)$ is the solution of:

$$\pi(\omega) = \inf_{Q \in \mathcal{M}} \{E_Q[\omega] + H(Q, P)\} - \inf_{Q \in \mathcal{M}} \{H(Q, P)\} \quad (5.15)$$

where $H(Q, P)$ is the relative entropy.

Remark 12 Assume that $\mathcal{M} = \{Q : Q \ll P\}$. Then we can derive, from Proposition 5.1 and Equation (5.15) the inverse duality relationship between the certainty equivalent and the generalized distance if \mathbb{E} is additive:

$$\mathbb{E}(\omega) = \inf_{Q \ll P} \{E_Q[\omega] + H(Q, P)\}.$$

We recall that, if the utility is exponential, then \mathbb{E} is additive with respect to a constant.

If, in (5.15) we take the opposite, then

$$\begin{aligned}
-\pi(\omega) &= -\inf_{Q \in \mathcal{M}} \{E_Q[\omega] + H(Q, P) - \inf_{Q' \in \mathcal{M}} \{H(Q', P)\}\} \\
&= \sup_{Q \in \mathcal{M}} \{E_Q[-\omega] - H(Q, P) + \inf_{Q' \in \mathcal{M}} \{H(Q', P)\}\} \\
&= \sup_{Q \in \mathcal{M}} \{E_Q[-\omega] - \left(H(Q, P) - \inf_{Q' \in \mathcal{M}} \{H(Q', P)\} \right)\} \\
&= \rho(\omega)
\end{aligned}$$

where

$$F(Q) \doteq H(Q, P) - \inf_{Q' \in \mathcal{M}} \{H(Q', P)\}. \quad (5.16)$$

Hence, the Dynamic Certainty Equivalent $\pi(\omega)$ is the opposite of $\rho(\omega)$, the entropic convex risk measure, where in the representation (4.15) we have: $\mathcal{P} = \mathcal{M}$ and $\inf_{Q \in \mathcal{M}} F(Q) = 0$. In fact:

$$\inf_{Q \in \mathcal{M}} F(Q) = \inf_{Q \in \mathcal{M}} \{H(Q, P)\} - \inf_{Q' \in \mathcal{M}} \{H(Q', P)\} = 0.$$

5.3 Entropic Convex Measure of Risk

Up to now, we have seen two examples of convex entropic risk measures. Both arise from two different, but very similar, definition of convex risk measure. So, using the negative exponential utility function, we can define a relationship between:

Entropic Convex Risk Measure in [14] \longrightarrow Certainty Equivalent
 Entropic Convex Risk Measures in [19] \longrightarrow Dynamic Certainty Equivalent

And summarizing:

in [14] we have

$$\rho(x) = -\mathbb{E}(x) = \sup_{Q \in \mathcal{P}} \{E_Q[-x] - H(Q, P)\} \quad (5.17)$$

where

$$x \in L^\infty(\Omega, \mathcal{F}, P), \quad \mathcal{P} = \{Q : Q \ll P\}$$

and in [19] we have

$$\rho(x) = -\pi(x) = \sup_{Q \in \mathcal{M}} \{E_Q[-x] - H(Q, P) + \inf_{Q' \in \mathcal{M}} \{H(Q', P)\}\} \quad (5.18)$$

where

$$x \in L^\infty(\Omega, \mathcal{F}, P),$$

and \mathcal{M} is the set of probability measure Q absolutely continuous with respect to P such that Q is a martingale measure.

We recall that if in Definition 4.1 we impose $\rho(0) = 0$ ((g) normazation) then we have, by (4.3), $\inf_{Q \ll P} \alpha(Q) = 0$. Moreover, if we consider in (5.18) $\mathcal{M} =$

$\{Q : Q \ll P\} = \mathcal{P}$ (totally incomplete market), with \mathcal{P} as in (4.3), then the two representations coincide. We can show it in two alternative ways.

1. We see that, with $\mathcal{P} = \mathcal{M}$, (5.17) and (5.18) only differ for the penalty function. So, we have that

$$H(Q, P) = H(Q, P) - \inf_{Q' \in \mathcal{P}} H(Q', P) \Leftrightarrow \inf_{Q' \in \mathcal{P}} H(Q', P) = 0 \quad (5.19)$$

But, if $\mathcal{M} = \mathcal{P} = \{Q : Q \ll P\}$, then $P \ll P$, too. This implies $P \in \mathcal{M}$. Then we have

$$\inf_{Q \in \mathcal{P}} H(Q, P) = 0$$

and the equation holds.

2. Taking a look at Remark 12, we have:

$$\begin{aligned} \mathbb{E}(\omega) &= \inf_{Q \ll P} \{E_Q[\omega] + H(Q, P)\} \\ &= \inf_{Q \ll P} \{-(E_Q[-\omega] - H(Q, P))\} \\ &= - \sup_{Q \ll P} \{E_Q[-\omega] - H(Q, P)\} \\ &= -\rho(\omega) \end{aligned}$$

where $\rho(\omega)$ defined as in (5.7).

This last equation expresses exactly the duality relationship between the relative entropy and the so called “free energy”,

$$\ln E_P[e^\omega] = \sup_{Q \ll P} \{E_Q[\omega] - H(Q, P)\} = -\mathbb{E}(-\omega)$$

This is exactly, for $-\omega = x$, what we have found in (5.8). And the two risk measures, in terms of opposite of the Certainty Equivalent, coincide.

Note that in this particular case axiom (g) normalization holds in both representations and its imposition is not a necessary requirement to get the equivalence between the two approaches.

Chapter 6

Risk measure and claims pricing

We have seen that there exists a close relation between the entropic convex risk measures and the pricing rules of a T -claim ω in the case of incomplete or totally incomplete market. This way of given a value keep in consideration the preferences of the agents. We used, in particular, the exponential utility function. We can now go beyond.

Our aim is, in fact, to derive a concrete interpretation of ρ given by its relation with the value(price) of a claim. We will get to define ρ as a “price”. We will consider coherent and convex measures of risk. Our setting will be a constrained (incomplete) financial market and totally incomplete financial market. As said before, in an incomplete market, perfect replication of a claim is usually not possible. The superreplicating price is the minimal quantity that an agent has to invest to find a strategy that dominates the claim payoff with certainty. And it has been characterized as the essential supremum on the set of equivalent martingale measures of the expectation of the actualized discounted payoff. It was also shown (see El Karoui and Quenez(1991)) that the price for a claim may vary between the superreplicating price for buyers h_{low} and the superreplicating price for the sellers h^{up} , and that any price chosen in the open interval (h_{low}, h^{up}) does not lead to an arbitrage opportunity.

On the other hand, in the case of totally incomplete market, we have already seen that the price of a claim coincides with the Certainty Equivalent.

6.1 Coherent risk measures and pricing rules

We study now a relation between ρ considered as coherent measure of risk and the superreplicating price of a claim. We follow the notation of Delbaen and Schachermayer, 1994, [9] and [10] in the exposition.

Let $(\Omega, (\mathcal{F}_{t,0 \leq t}), P)$ be a filtered probability space and let $S : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}^d$ be a càdlàg locally bounded, adapted process. We suppose that the set

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

is non-empty. Since S is locally bounded, M is a closed convex subset of L^1 . We also suppose that $\exists Q \in M, Q \approx P$, 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. We know that, by Theorem 3.2, ρ , coherent measure of risk, admit the representation

$$\rho(x) = \sup_{Q \in \mathcal{P}} E_Q[-x].$$

We recall the following result in [9]. If $x \in L^\infty$ then the quantity

$$p(x) = \sup_{Q \in M} E_Q[x] \tag{6.1}$$

is called the superhedging (or superreplicating) 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 minimum price that can be charged for x . The minimum price is:

$$m(x) = \inf_{Q \in M} E_Q[x]. \tag{6.2}$$

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. The requirement $(\forall x \in L^\infty; \rho(x) \leq p(-x))$ is, by the Hahn-Banach theorem (see Theorem A.8), equivalent to $\mathcal{P} \in M$.
- (b) If $\mathcal{P} \cap M = \emptyset$ then, by weak compactness of \mathcal{P} , the Hahn-Banach theorem gives us an element $-x \in L^\infty$ so that:

$$\sup_{Q \in \mathcal{P}} E_Q[-x] < \inf_{Q \in M} E_Q[-x].$$

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.

6.2 Convex measure of risk and pricing rules

We will follow the paper of Rouge and El Karoui (2000) that, as Frittelli in [17], mix pricing and utility maximization. In the model, a small investor (who does not influence market prices) is confronted with the problem of selling contingent claim while performing maximization of utility. The price of the contingent claim is defined as the smallest amount of money p to add to his initial wealth x that allows him to achieve the same expected utility he would have had with initial wealth x without selling the claim at time T . Recall that, when positive, the number $\rho(x)$ assigned by the measure ρ can be interpreted (see [3]) as the minimal extra cash the agent has to add to the risky position x to make it acceptable.

6.2.1 Abstract contingent claims

Let (Ω, \mathcal{F}, P) be a probability space whose role is to give the null set- i.e., those $A \in \mathcal{F}$ such that $P[A] = 0$. In a financial market, let $C \geq 0$ P-a.s. be a random payoff (or *claim*) of date T . Denote by L_0^+ the set of claims C , L_0^- the set $\{-C, C \in L_0^+\}$, and $L_0 = L_0^+ \cup L_0^-$. Suppose that an agent is given a preference relation \succeq on the set of the pairs $(x, C) \in \mathbb{R} \times L_0$ (initial endowment, possibly nonpositive, and terminal agreement to buy or sell), compatible with the usual order on \mathbb{R} and preorder on L_0 , namely, a transitive relation on $\mathbb{R} \times L_0$ such that $x' \geq x, C' \geq C \Rightarrow (x', C') \succeq (x, C)$.

An agent with an initial endowment x wishes to sell at time 0 a claim $C \in L_0^+$. He may choose either of the following:

1. delivering the claim C at time T in exchange for an additional endowment of y at time 0; that is, he chooses $(x + y, -C)$;
2. not delivering anything: $(x, 0)$.

For him to prefer the first alternative, the quantity y has to be such that $(x + y, -C) \succeq (x, 0)$. We thus define a *price to sell* C as

$$pr(x, C) = \inf\{y \geq 0, (x + y, -C) \succeq (x, 0)\}. \quad (6.3)$$

The *price to buy* C is the quantity that someone is willing to pay at time 0 to get C at time T . We take it nonpositive for convenience, and define it as

$$-pr(x, -C) = \sup\{y \geq 0, (x - y, C) \succeq (x, 0)\}. \quad (6.4)$$

Both definition (6.3) and (6.4) may be summarized in the following, now with $C \in L_0$

$$pr(x, C) = \inf\{y \in \mathbb{R}, (x + y, -C) \succeq (x, 0)\} \quad (6.5)$$

Conversely, we call $p : \mathbb{R} \times L_0 \rightarrow \mathbb{R}$ a compatible pricing function if it defines a compatible preference relation through

$$(x', C') \succeq (x, C) \Leftrightarrow x' - p(x', -C') \geq x - p(x, -C).$$

6.2.2 Superreplication Price

Given its initial endowment x , the financial agent may choose between time 0 and T an investment strategy denoted π (if no confusion is possible with the Value π in [17] and in the previous chapter) is a set of admissible strategy \mathcal{A} . $X_t^{x, \pi}$ represent the agent's wealth at time T .

In this setting, a first example of compatible pricing function is given by hedging consideration.

The *seller's cost* of a claim $C \in L_0^+$, denoted h^{up} , is the smallest initial amount of wealth for which there exists a superreplicating portfolio strategy (with the convention $\inf \emptyset = +\infty$):

$$h^{up}(C) = \inf\{x \geq 0, \exists \pi \in \mathcal{A}, X_T^{x, \pi} \geq C\}. \quad (6.6)$$

Symmetrically, define the *buyer's cost* (or *lower hedging price*, *buyer's price*) as

$$h_{low}(C) = \sup\{x \geq 0, \exists \pi \in \mathcal{A}, X_T^{-x, \pi} \geq -C\}. \quad (6.7)$$

We may once again give a unified definition: Call h the hedging price of $C \in L_0$ if

$$h(C) = \inf\{x \in \mathbb{R}, \exists \pi \in \mathcal{A}, X_T^{x, \pi} \geq C\}. \quad (6.8)$$

so that for $C \in L_0^+$, $h^{up}(C) = h(C)$ and $h_{low}(C) = -h(-C)$. The hedging price h is a pricing function compatible with the usual order on \mathbb{R} and the preorder on L_0 .

Let us introduce some vocabulary. We say that the contingent claim $C \in L_0^+$ is *sellable* (resp. *buyable*) if there exists x and a portfolio $\pi \in \mathcal{A}$ such that $X_T^{x, \pi} \geq C$ P-a.s. (resp. $X_T^{-x, \pi} \geq -C$ P-a.s). In such a case, $h^{up} < \infty$, and any claim is always buyable since $h_{low}(C) \geq 0 \forall C \in L_0^+$. If the preceding inequalities can be written as equalities, then the claim C is said to be *replicable* for a seller (resp. for a buyer). We finally say that a claim is *tradeable* if both the seller and the buyer may replicate it, and then $h_{low} = h^{up}$.

The *arbitrage-free interval* for a claim $C \in L_0^+$ is the interval $[h_{low}, h^{up}]$. Any claim C may be sold or bought for a price in this interval without giving rise to an arbitrage opportunity.

6.2.3 Utility Maximization Price

As done in [17], the more natural way to define the agent's preferences is to model his attitude toward risk by a utility function u (concave and strictly increasing). The maximal expected utility of $(x, C) \in \mathbb{R} \times L_0^+$ is

$$\hat{U}(x, C) = \max_{\pi \in \mathcal{A}} E_P[u(X_T^{x, \pi} + C)] \quad (6.9)$$

and we define the reference relation \succeq by $(x', C') \succeq (x, C) \Leftrightarrow \hat{U}(x', C') \geq \hat{U}(x, C)$. We define the price according to Equation (6.5), and denote it by p .

In [26] it is shown, under some hypothesis on the set of admissible strategy \mathcal{A} that, in the vocabulary of Frittelli(2000) [17], the price $p(x, C)$ derived from utility maximization is a value coherent with the no-arbitrage principle. That is

$$h_{low}(c) \leq p(x, C) \leq h^{up}(C).$$

Moreover, if a claim C is tradeable, its price from utility $p(x, C)$ is equal to its arbitrage price.

Since not only superreplicating strategies ($X_T^{x, \pi} \geq 0$ P-a.s.) are considered, there is the necessity to compute expected utility of portfolios taking nonpositive terminal values. This is not possible for the usual utility function, such as power or logarithmic functions. Because of its simplicity and link with the relative entropy, we choose u of the negative exponential type. For simplicity, even in this case, the risk aversion coefficient will be considered equal to 1. The choice of an exponential utility function obliges to impose the following condition:

Assumption 1 *All the claims $C \in L_0^+$ we shall consider will now be bounded.*

Then we have the following theorem (see Theorem 2.1 in [26]):

Theorem 6.1 *The price of the claim C is given by:*

$$p(x, C) = \sup_{Q_T} \{E_{Q_T}[C] - H(Q_T, P)\} - \sup_{Q_T} \{-H(Q_T, P)\} \quad (6.10)$$

where Q_T runs through the set of probabilities $Q_T \approx P$ such that

$$E_{Q_T}[X_T^{x, \pi}] \leq x$$

an $H(Q_T, P)$ is the relative entropy of the probability measure Q_T with respect to P .

Notice that the price is independent from the initial endowment x .

Differently from Equation (2.13) in [26], we have considered in (6.10) the free-risk interest rate equal to be equal to zero.

6.3 Interpretation of the result

We got in the subsection above a formulation (Eq. (6.10)) for the price of claim C compatible with the no-arbitrage principle. This formulation takes on consideration the preferences of the agent by modelling them with his utility function. In this case the exponential utility function has been adopted. Note that Equation (6.10) is very close to the *value* π of Equation (5.15).

To unify the notation and not to create confusion, we will denote by $x \in \mathbb{R}$ the initial endowment and by $\omega \in L^\infty$ the T -claim.

Let us consider the buyer price of the claim $\omega \in L^\infty$ defined as the solution of $p_b = p_b(\omega)$ of equation

$$V(x - p_b + \omega) = V(x)$$

where $x \in \mathcal{R}$ is the initial wealth and V is defined as in (5.11). Due to the particular properties of the exponential utility function, it can be shown that $p_b(\omega)$ is independent from the initial endowment x and that the buyer price coincide with the Dynamic Certainty Equivalent $p_b(\omega) = \pi(\omega)$.

If $\omega \geq 0$ we have, by Equations (6.3)-(6.5) that $p_s(x, \omega) = pr(x, \omega)$ and $p_b(x, \omega) = -pr(x, -\omega)$, where p_s is the seller price as defined in (6.3). We know that $\pi(\omega)$ is the buyer's price $p_b(\omega)$ and $p(\omega)$ as in (6.10) is the seller's price of the positive claim ω . Then we have the following relation:

$$\pi(\omega) = -p(-\omega) \tag{6.11}$$

We could have got this result starting from the formulations for $\pi(\omega)$ and $p(\omega)$. In fact:

$$\begin{aligned} \pi(\omega) &= \inf_Q \{E_Q[\omega] + H(Q, P) - \inf_{Q'} H(Q', P)\} \\ &= \inf_Q \{-(E_Q[-\omega] - H(Q, P) + \inf_{Q'} H(Q', P))\} \\ &= -\sup_Q \{E_Q[-\omega] - H(Q, P) + \inf_{Q'} H(Q', P)\} \\ &= -p(-\omega) \end{aligned}$$

We will try now to understand the economic meaning of the relation between the value $\pi(\omega)$ and the entropic convex measure of risk $\rho(\omega)$ of the T -claim ω .

We know that for ρ , convex measure of risk admitting representation (4.15), axioms (c) * positivity holds and implies $\omega \geq 0 \Rightarrow \rho(\omega) \leq 0$.

Let us assume the time- T claim ω to be nonnegative. We will use the exponential utility function with unitary risk aversion coefficient to model the agent's preferences toward risk. We suppose that the free-risk interest rate is zero.

We know, by previous considerations, that $\rho(\omega)$ and $\pi(\omega)$ are in the following relation:

$$\rho(\omega) = -\pi(\omega)$$

This last equation shows that the buyer's price $\pi(\omega)$ of the nonnegative T -claim ω coincide with the quantity $-\rho(\omega)$, that is, the cash amount that can be withdrawn from the position or that can be received as a restitution, as in the case of organized market for financial futures.

Thus the price the agent pays at time 0 to obtain the T -claim ω is nothing but the opposite of the number $\rho(\omega)$, that is, the measure of the risk of the claim, keeping in consideration the agent's preferences.

In the same way we can define a relation between ρ and $pr(x, \omega)$. Recall that, for $\omega \in L_0$

$$pr(x, \omega) = \inf\{y \in \mathbb{R} : (x + y, -\omega) \succeq (x, 0)\}$$

is the minimal amount one has to add to his initial wealth at time 0 to deliver at time T the claim ω . In the case when $\omega \in L_0^+$ we have seen that $\pi(\omega)$ is the buyer's price of the T -claim ω . i.e.

$$\pi(\omega) = -\rho(\omega) = -pr(x, -\omega) = \sup\{y \geq 0 : (x - y, \omega) \succeq (x, 0)\}$$

We now study the case when $\omega \in L_0^-$. We know that, by the axiom (c^0) , this implies $\rho(\omega) \geq 0$. Then we have

$$\rho(\omega) = p(-\omega) = p_s(-\omega) \tag{6.12}$$

This last equation says that the riskier (negative) the position is, the higher is the price that the agent wants to be given to sell the position $-\omega$.

In the general case, when $\omega \in L_0$, we have

$$\rho(\omega) = -\pi(\omega) = pr(x, -\omega) = \inf\{y \in \mathbb{R} : (x + y, \omega) \succeq (x, 0)\} \tag{6.13}$$

It means that $\rho(\omega)$ is the minimal amount that needs to be added (or withdrawn if negative) to the initial wealth x of the agent to make the possession of the claim ω at time T *acceptable*. As done with coherent and convex measures of risk, we say that the possession of a claim is acceptable if $\omega \in \mathcal{A}$, with

$$\mathcal{A} = \{\omega \in X : \exists y \in \mathbb{R}, (x + y, \omega) \succeq (x, 0)\}$$

We present a simple example. Let us assume that an agent wants to borrow some money from a bank. The interest rate the bank asks to give him the money is nothing but the cost for the agent to get the money. Clearly, the riskier is the situation of the agent (he is in a short position) the higher is the interest rate the bank asks. This shows the positive relation between the riskiness of the (short) position and the price the bank asks to accept it.

Chapter 7

Applications

The aim of this chapter is to give a simple application of the convex measures of risk. The setting will be restricted to financial measurement of risk. We will first recall the well-known measure of risk called Value at Risk. Then we will show why it is not an adequate instrument to measure the risk of a financial position and we will introduce the convex measure of risk called AVar (or Expected Shortfall). Then we will make a theoretical and empirical comparison between VaR and AVar.

7.1 Value at Risk

Value at Risk (VaR) was introduced in 1994 and became one of the most important tools for risk management in the financial industry and part of the regulator mechanism. Even if it is the most widely used risk measure nowadays it can present several and significant problems. To define it we need to recall the definition of quantile.

Let $\alpha \in (0, 1)$ be some fixed small probability or confidence level, in practice usually but not necessarily below 5 percent. Often used values are 0.01 and 0.02.

Definition 7.1 (Quantile) *Given $\alpha \in]0, 1[$ the number q is called quantile- α of the random variable X on (Ω, \mathcal{F}, P) if one of the three equivalent properties holds:*

1. $P[X \leq q] \geq \alpha \geq P[X < q]$
2. $P[X \leq q] \geq \alpha$ e $P[X \geq q] \geq 1 - \alpha$
3. $F_X \geq \alpha$ e $F_X(q^-) \leq \alpha$ con $F_X(q^-) = \lim_{x \rightarrow q, x < q} F(x)$, where F_X is the cumulative distribution function of X .

More precisely

$x_{(\alpha)} = q_{(\alpha)}(X) = \inf\{x \in \mathbb{R} : P[X \leq x] \geq \alpha\}$ is the lower α -quantile of X

$x^{(\alpha)} = q^{(\alpha)}(X) = \inf\{x \in \mathbb{R} : P[X \leq x] > \alpha\}$ is the upper α -quantile of X

We use the notation in x if the dependence from X is clear, otherwise we use the notation q .

Note that $x^\alpha = \sup\{x \in \mathbb{R} : P[X \leq x] \leq \alpha\}$.

From

$$\{x \in \mathbb{R} : P[X \leq x] > \alpha\} \subset \{x \in \mathbb{R} : P[X \leq x] \geq \alpha\}$$

is clear that $x_{(\alpha)} \leq x^{(\alpha)}$. Moreover, it is easy to see that

$$x_{(\alpha)} = x^{(\alpha)} \text{ if and only if } P[X \leq x] = \alpha \text{ for at least } x, \quad (7.1)$$

and in the case $x_{(\alpha)} < x^{(\alpha)}$

$$\{x \in \mathbb{R} : \alpha = P[X \leq x]\} = \begin{cases} [x_{(\alpha)}, x^{(\alpha)}), & P[X = x^{(\alpha)}] > 0 \\ [x_{(\alpha)}, x^{(\alpha)}], & P[X = x^{(\alpha)}] = 0 \end{cases} \quad (7.2)$$

As function of α , $q^{(\alpha)}(X)$ is the right-continuous inverse of the distribution function $F(X)$. In this section we will see some properties of $q^{(\alpha)}(\cdot)$, viewed as a functional on the space of financial position.

We give now a formal definition of the Value at Risk:

Definition 7.2 Fix some level $\alpha \in (0, 1)$. For a financial position x , we define its Value at Risk at level α as

$$VaR_\alpha \doteq -q^\alpha(x) = q_{1-\alpha}(-x) = \inf\{m | P[x + m < 0] \leq \alpha\}.$$

In financial term, $VaR_\alpha(x)$ is the smallest amount of capital which, if added to x and invested in the free-risk asset, keeps the probability of a negative outcome below the level α . We will not insist on the drawbacks of VaR. We just present some properties and explain why VaR is not an adequate measure of risk.

Theorem 7.1 VaR satisfies (c1) *monotonicity, (d1) positive homogeneity, (e) translation invariance, (f) law invariance and (h) comonotonic additivity.

Proof See the proof of Theorem 3.1.1 in [6].

From this theorem we can see that VaR is not coherent (see for counterexample [5]) and its acceptances set is typically not convex. So VaR is *not* a convex measure of risk.

Consider VaR as a measure of risk on the linear space $X = L^2(\Omega, \mathcal{F}, P)$. Let then consider a Gaussian subspace X_0 , i.e. a linear space $X_0 \subset X$ consisting of normally distributed random variables. It can be shown (see for example Remark 4.34 in [16]) that VaR_α does satisfy the axiom of convexity if restricted to the Gaussian subspace X_0 and if α belongs to $(0, \frac{1}{2}]$.

7.2 Average Value at Risk

The aim of this section is to present a risk measure, defined on the space $X = L^\infty$ which, in contrast with VaR is convex or even coherent. The solution is a measure of risk defined in terms of the Value at Risk, but does satisfy the axiom of a coherent risk measure.

Definition 7.3 *The Average Value at Risk at level $\alpha \in (0, 1)$ of a position $x \in X$ is given by*

$$\text{AVaR}_\alpha(x) = \frac{1}{\alpha} \int_0^\alpha \text{VaR}_\alpha(x) dx \quad (7.3)$$

Sometimes, the Average Value at Risk is also called “Conditional Value at Risk” or the “Expected Shortfall”, and one writes $\text{CVaR}_\alpha(x)$ or $\text{ES}_\alpha(x)$. These terms are motivated by formulas (7.7) and (7.4). They can be in some way misleading. In fact “Conditional Value at Risk” might be used also to denote the Value at Risk with respect to a conditional distribution, and “Expected Shortfall” might be understood as the expectation of the Shortfall x^- .

Proposition 7.1 *Suppose that $x \in X$ and that q is the α -quantile for x , i.e., $q \in [q_{(\alpha)}(x), q^{(\alpha)}(x)]$. Then*

$$\text{AVaR}_\alpha(x) = \frac{1}{\alpha} E[(q - x)^+] - q \quad (7.4)$$

$$= \frac{1}{\alpha} \inf_{s \in \mathbb{R}} (E[(s - x)^+] - \alpha s) \quad (7.5)$$

Proof. See the proof of Proposition 4.37 in [16]. The following theorem states the coherence of AVaR.

Theorem 7.2 *For $\alpha \in (0, 1)$, AVaR_α is a coherent measure of risk which is continuous from below. It has the representation*

$$\text{AVaR}_\alpha(x) = \max_{Q \in \mathcal{Q}_\alpha} E_Q[-x], \quad x \in X \quad (7.6)$$

where \mathcal{Q}_α is the set of all probability measure $Q \ll P$ whose density dQ/dP is P -a.s. bounded by $1/\alpha$.

Proof. See Theorem 4.39 in [16].

Corollary 7 For all $x \in X$,

$$\begin{aligned} AVaR_\alpha(x) &\geq E[-x | -x \geq VaR_\alpha(x)] \\ &\geq VaR_\alpha(x). \end{aligned} \quad (7.7)$$

Moreover the two inequalities are in fact identities if

$$P[x \leq q^{(\alpha)}(x)] = \alpha, \quad (7.8)$$

which is the case if x has a continuous distribution.

A last result shows that, under suitable hypothesis, AVaR is the best conservative approximation to VaR_α in the class of all *distribution invariant* convex measures of risk which are continuous from above.

Theorem 7.3 On an atomless probability space, $AVaR_\alpha$ is the smallest distribution invariant convex measure of risk which is continuous from above and dominates VaR_α .

Proof. See the proof of Theorem 4.46 in [16].

7.3 Examples

In this section we will present some concrete applications of the risk measures discussed before. We will distinguish the case when the random variable representing the value of a position is continuous and the general case.

7.3.1 The continuous case

We will denote by $x \in X$ the random variable representing the profit (or loss) on an investment, at a fixed time horizon (one month or one year for example). It can be for instance the random return on a stock, an index or any other portfolio, measured in absolute or relative terms. We will focus on situation in Finance. Positive values of x will be interpreted as profits and negative value as losses. We will assume, in this subsection, that x is a continuous random variable with distribution function $F = F_x$.

We know that, if x a continuous random variable, then for a fixed level α , lower and upper quantile, coincide. This fact allows us to consider AVaR, as defined in (7.3), equivalent to the risk measure called ES in [1]. Recall that ES admits a representation of the the form

$$ES_\alpha(x) = -\frac{1}{\alpha} \int_0^\alpha q_{(u)}(x) du. \quad (7.9)$$

. Starting from now, while dealing with continuous random variable, we will use ES as convex measure of risk. The following theorem shows a quite easy analytical way to calculate Es.

Theorem 7.4 (Th. 3.2.2 [6]) *Let x be a continuous random variable with cdf F and pdf $f(x) = \frac{\partial F}{\partial x}(x)$. Then*

$$ES_{\alpha}(x) = -\frac{1}{\alpha} \int_{-\infty}^{q_{(\alpha)}(x)} x \cdot f(x) dx. \quad (7.10)$$

Proof. See [6]

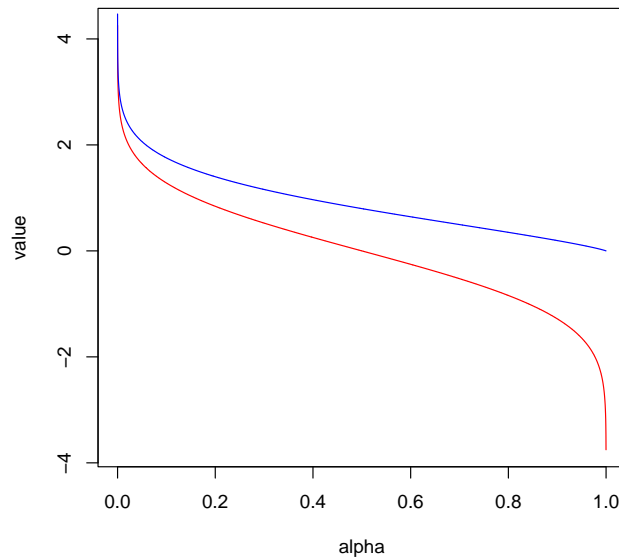


Figure 7.1: $\text{Var}_{\alpha}(x)$ and $\text{ES}_{\alpha}(x)$ as a function of α if $x \sim N(0, 1)$. The upper (blue) line is ES, the lower (red) line is VaR.

Let us consider the simple case of a normal distributed random variable.

Example 3 *Consider a random variable x such that $x \sim N(0, 1)$. It can be considered as a simple investment with mean profit 0, unit variance and a profit is equally likely as a loss.*

So we have $F(x) = \Phi(x)$, the standard normal distribution function and $f(x) = \phi(x)$, the standard normal density function. Fix $\alpha \in (0, 1)$:

$$\text{VaR}_{\alpha}(x) = -F^{-1}(\alpha) = -z_{(\alpha)}$$

$$\begin{aligned}
ES_\alpha(x) &= -\frac{1}{\alpha} \int_{-\infty}^{z(\alpha)} x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx \\
&= -\frac{1}{\alpha\sqrt{2\pi}} \left[-\exp\left(-\frac{1}{2}x^2\right) \right]_{x=-\infty}^{x=z(\alpha)} \\
&= \frac{\exp\left(-\frac{1}{2}z(\alpha)^2\right)}{\alpha\sqrt{2\pi}}
\end{aligned}$$

Note that in the above notation $z(\alpha) = \Phi^{-1}(\alpha)$ are the usual quantiles of the standard normal distribution that are tabulated.

Looking at Figure 7.1 we can see that $\text{Var}_\alpha(x)$ and $\text{ES}_\alpha(x)$ are both decreasing in α , and that $\text{Var}_\alpha(x) < \text{ES}_\alpha(x)$ for all $\alpha \in (0, 1)$. $\text{ES}_\alpha(x)$ is positive for all $\alpha \in (0, 1)$, because it is always $-E[x] = 0$, to which it converges as α goes to 1. We also see that $\text{Var}_\alpha(x)$ and $\text{ES}_\alpha(x)$ can get arbitrarily large for arbitrarily small values of α , reflecting the fact that the possible loss is not bounded. Note that in the region of interest, that usually is α between 0 and 3 %, $\text{Var}_\alpha(x)$ and $\text{ES}_\alpha(x)$ are above 2, and that both risk measure increase exponentially if α decrease.

7.3.2 The general case

Up to now, we every made the assumption that x is a continuous random variable with a continuous cdf F . Off course, this need not to be the case in reality and in practice there are many example of return-distribution that are discrete. Examples are portfolios of not-traded loans and portfolio of derivatives as options.

Therefore, we will assume in this section that x can be any random variable, possibly discrete.

We have seen that, in a general case we have $x_{(\alpha)} \leq x^{(\alpha)}$. Then we cannot state that, for a fixed $\alpha \in (0, 1)$, $\text{AVar}_\alpha(x)$ and $\text{ES}_\alpha(x)$ coincide. Moreover dealing with discrete distribution creates problems in the estimation of Var .

Suppose that we want to estimate the lower α -quantile $x_{(\alpha)}$ for some random variable x . Let some sample (x_1, \dots, x_n) , drawn from independent copies of x , be given. Denote by $x_{1:n} \leq \dots \leq x_{n:n}$ the components of the ordered n -tuple (x_1, \dots, x_n) . Denote by $\lfloor x \rfloor$ the integer part of the number $x \in \mathbb{R}$, hence

$$\lfloor x \rfloor = \max\{k \in \mathbb{Z} : k \leq x\}.$$

Then the order statistic $x_{\lfloor n\alpha \rfloor : n}$ appears as a natural estimator for $x_{(\alpha)}$. Nevertheless, it is well known that in the case of a non-unique quantile (i.e. $x_{(\alpha)} < x^{(\alpha)}$) the quantity $x_{\lfloor n\alpha \rfloor : n}$ does not converge to $x_{(\alpha)}$. See for example Theorem 1 in [13] which says that

$$1 = P[x_{\lfloor n\alpha \rfloor : n} \leq x_{(\alpha)} \text{ infinitely often}] = P[x_{\lfloor n\alpha \rfloor : n} \leq x^{(\alpha)} \text{ infinitely often}].$$

Surprisingly, we get a well-determined limit when we replace the single order statistic by an average over the left tail of the sample.

Proposition 7.2 *Let $\alpha \in (0, 1)$ be fixed, x a real random variable with $E[x^-] < \infty$ and x_1, x_2, \dots an independent sequence of random variables with the same distribution as x . Then with probability 1*

$$-\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^{\lfloor n\alpha \rfloor} x_{i:n}}{\lfloor n\alpha \rfloor} = ES_{(\alpha)}(x) \quad (7.11)$$

If x is integrable, then the convergence in (7.11) holds in L^1 , too.

Proof. See proof of Theorem 4.1 in [1].

As seen above, this result

$$\lim_{n \rightarrow \infty} x_{\lfloor n\alpha \rfloor:n} = x_{(\alpha)}$$

does not hold in general, but only

$$\liminf_{n \rightarrow \infty} x_{\lfloor n\alpha \rfloor:n} = x_{(\alpha)} \quad \text{and} \quad \limsup_{n \rightarrow \infty} x_{\lfloor n\alpha \rfloor:n} = x^{(\alpha)}$$

To get to an estimator for VaR we present the following result:

Theorem 7.5 *Given a random sample x_1, \dots, x_n from a certain distribution F and for a fixed $\alpha \in (0, 1)$, then if F is continuous, the estimator*

$$\widehat{VaR}_\alpha(x) = -x_{\lfloor n\alpha \rfloor:n} \quad (7.12)$$

converges to $VaR_\alpha(x)$ as $n \rightarrow \infty$.

Proof. See [1].

So, starting from now, we will use, because of its easy computation, Expected Shortfall (and not AVaR) as example of convex risk measure. We first introduce the following proposition. See Corollary 3.3 in [1] for the proof.

From this last Theorem we can notice that this way of estimating VaR works well only in the case of continuous distributions.

Proposition 7.3 *If x is a real-valued random variable with $E[x^-] < \infty$, then the mapping $\alpha \rightarrow ES_\alpha$ is continuous in $(0, 1)$.*

One problem with VaR is that when applied to discontinuous distributions, may be its sensitivity to small changes in the confidence level α . In other word, it is not continuous with respect to the confidence interval α . In contrast, from Proposition 7.3 we know that, the risk measured by ES_α will not change dramatically when there is a switch in the confidence interval. In practice for many investments it is

not really a constraints to assume that the underlying distribution is continuous. This means that we do not have to worry about the convergence of the estimator for VaR.

For the random sample we simply take in practice n , (usually daily) observation for the price at closure of the investment under consideration. Then we compute the log-returns and denote them by r_1, \dots, r_n . Here n is the number of observations or the length of the observation period (in days). The larger n , the better estimation for VaR and ES are obtained in general. We will not discuss the estimator error for VaR and Es, it goes beyond our aim. One can see for example [30].

We analyzed the time series of the Ftse index of the London Stock Exchange from 2/1/1996 to 29/12/2000. In total we have 1304 observation. After the computation of the log returns, we computed the values of VaR_α and ES_α for $\alpha \in (0, 1)$. We used the formulation given in Proposition 7.2 in order to obtain an estimation of ES and Theorem 7.5 to obtain an estimation for VaR.

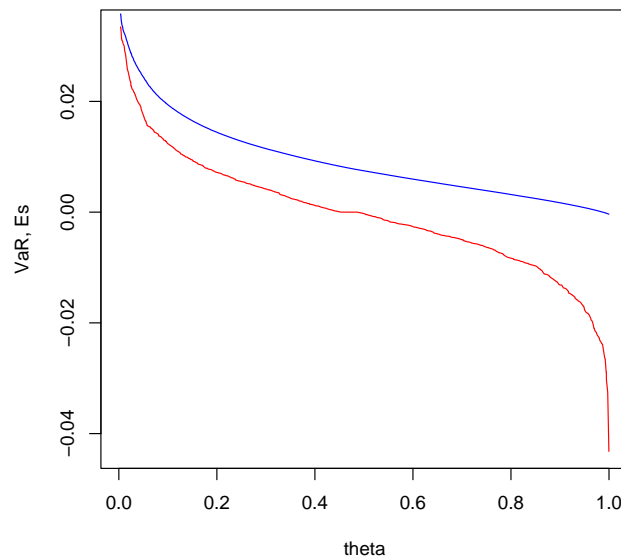


Figure 7.2: $\hat{\text{VaR}}_\alpha$ and $\hat{\text{ES}}_\alpha$ as a function of α . The upper (blue) line is ES, the lower (red) line is VaR.

We can see in Figure 7.2 that $\hat{\text{VaR}}_\alpha$ and $\hat{\text{ES}}_\alpha$ are both decreasing in α , $\hat{\text{VaR}}_\alpha < \hat{\text{ES}}_\alpha$ for all $\alpha \in (0, 1)$. $\hat{\text{VaR}}_\alpha$ is positive for values of α below circa 0.45. On the other hand $\hat{\text{ES}}_\alpha$ converges to the opposite of the sample mean. The sample mean

is 0.0004011565.

We will not spend more time on the interpretation of the figure above from a financial point of view. What we prefer to emphasize is the shape of the two curves. What we see is that \hat{ES}_α produces a beautiful convex curve, and this is due to its continuity and convexity property, as discussed above.

On the contrary \hat{VaR}_α shows a less smooth line with more distortion. We can see in Figure 7.3 that shows the curves for $\alpha \in (0, 0.08)$. This gives us a clear incentive to doubt about the continuity in α of \hat{VaR}_α , as already discussed.

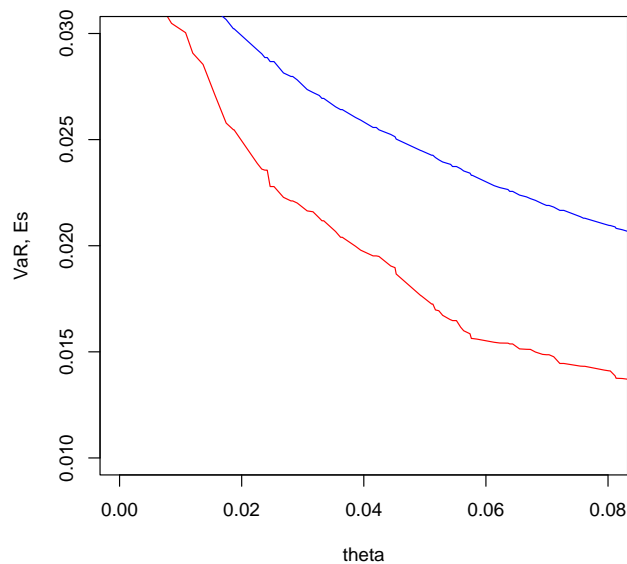


Figure 7.3: \hat{VaR}_α and \hat{ES}_α as a function of α on $(0, 0.8)$ for the returns of the FTSE. The upper (blue) line is \hat{ES}_α , the lower (red) line is \hat{VaR}_α .

7.3.3 Portfolio risk measure

In this subsection we will show the importance, from a financial point of view, of the convexity of a risk measure. We created in fact a portfolio weighted with two assets.

Let's call r the total return of a portfolio. We know that, taking n assets, say a_1, \dots, a_n , computed the total returns for every asset r_1, \dots, r_n and the relative weights $\omega_1, \dots, \omega_n$ such that $\sum_{i=1}^n \omega_i = 1$, the total return of a weighted portfolio

composed with that assets is

$$r = \sum_{i=1}^n r_i \omega_i \quad (7.13)$$

We imposed the condition $\omega_i \in [0, 1], \forall i \in 1, \dots, n$. This means that only long positions are admissible. Remember now the property of ES and VaR. ES is convex or even subadditive, VaR is in general not convex. This simple example will show why we insist that much on convexity. We think that if ρ , measure of risk is convex, then

$$\rho(r) \leq \sum_{i=1}^n \omega_i \rho(r_i) \quad (7.14)$$

where r, ω_i and $r_i, i \in 1, \dots, n$ are as defined above. This means obtaining less risk by diversification.

The portfolio we created is composed with two assets: Intel and Coca-Cola. Both the titles are part of the 30 titles that compose the Dow Jones Index. The period we considered goes from 03/02/1995 to 31/01/2005. We considered the returns of the portfolio of the form $r = \omega r_1 + (1 - \omega)r_2$. Then we computed, fixed $\alpha = 0.01$, $\hat{E}S_\alpha$ and $\hat{V}aR_\alpha$ by letting the value of the weight ω varying in $[0, 1]$.

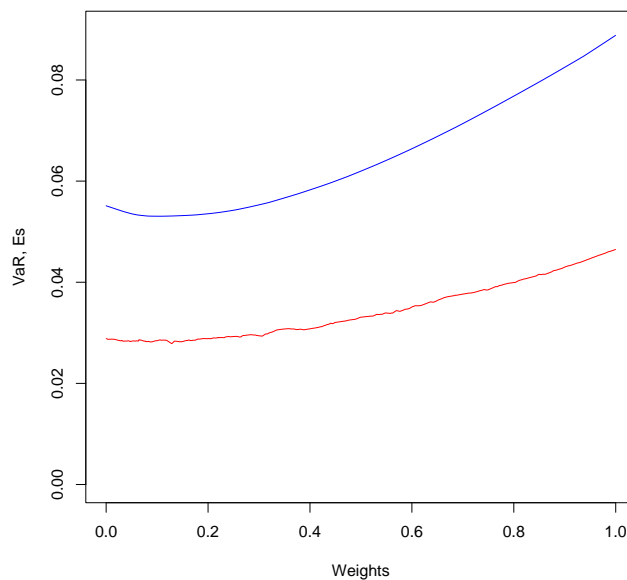


Figure 7.4: $\hat{V}aR_{0.01}$ (blue line) and $\hat{E}S_{0.01}$ (red line) as a function of $\omega \in [0, 1]$, for the returns of the portfolio composed with the returns of the titles Intel and Coca-Cola.

From Figure 7.4 it is immediately clear that ES is convex, because it provides a convex curve, with an unique global minimum. On the other hand the problem with VaR is clearly illustrated: the curve is not convex, not smooth and has several local minima. This is why VaR is not a suitable instrument in optimization problems.

Chapter 8

Conclusions

In this work we start analyzing the meaning of a measure of risk and we see some nice properties that could satisfy. We focus on the class of the convex measures of risk. Starting from this new definition a measure for the risk of a position, we model the preferences of the agent involved with the market (incomplete or totally incomplete) using the well-known concept of utility function (resp. loss function). We obtain, via utility maximization, a type of convex measure which we call entropic convex measure of risk. Moreover we try to give an economic explanation to this measure of risk and found a link between measure of risk and the price of a claim.

Due to its relation with the relative entropy, we adopt here, the exponential utility function. We impose, for simplicity, the risk aversion coefficient to be equal to one. Moreover consider the free risk interest rate to be equal to zero.

One of the possibly development of this matter could be to try to impose different conditions, i.e., considering a different risk aversion coefficient or even a different utility function; consider the discount factor not to be deterministic but try to model it with, for example, a stochastic model. On the other hand it could be interesting to mix other axioms with the one of convexity and defining new classes of measures of risk.

The gap between research and real economic world is great, and VaR is still the most common risk measure used. Things are slowly changing (some regulators started to use ES instead of VaR), but it will take some years.

Appendix A

Real Analysis and Measure Theory

We will deal in this first appendix with concepts and instruments belonging to pure mathematics. In particular we need some definitions from real and functional analysis. We follow in the exposition mainly [27] and [24].

A.1 Ordered Sets

Definition A.1 (Relation) *A relation is any subset of a Cartesian product. For instance, a subset of $A \times B$, called a binary relation from A to B , is a collection of ordered pairs (a, b) with first components from A and second components from B , and, in particular, a subset of $A \times A$ is called a relation on A . For a binary relation R , one often writes aRb to mean that (a, b) is in R .*

Definition A.2 (Totally Ordered Sets) *A relation \leq is a total order on a set S (\leq totally orders S) if the following properties hold:*

1. *Reflexivity: $a \leq a$ for all $a \in S$;*
2. *Weak antisymmetry: $a \leq b$ and $b \leq a$ implies $a = b$;*
3. *Transitivity: $a \leq b$ and $b \leq c$ implies $a \leq c$;*
4. *Comparability (Trichotomy law): For any $a, b \in S$, either $a \leq b$ or $b \leq a$.*

The first three are the axioms of a partial order, while addition of the trichotomy law defines a total order.

Recall that:

Definition A.3 (Trichotomy Law) Every real number is negative, 0, or positive. The law is sometimes stated as “For arbitrary real numbers a and b , exactly one of the relations $a < b, a = b, a > b$ holds” (Apostol 1967, p. 20).

Definition A.4 (Preorder) A relation “ \leq ” is called preorder on a set S if it satisfies the property of: Reflexivity and Transitivity.

A preorder that also has antisymmetry is clearly a partial order.

A.2 Lebesgue Measure

Let Ω be a nonempty point set, and \mathcal{A} be a class of subsets of Ω . Let \emptyset be the empty set. Consider the following properties:

1. $\emptyset \in \mathcal{A}$ and $\Omega \in \mathcal{A}$
2. If $A \in \mathcal{A}$ then $A^c \in \mathcal{A}$
3. \mathcal{A} is closed under finite unions and finite intersections: i.e., if A_1, \dots, A_n are all in \mathcal{A} , then $\bigcup_{i=1}^n A_i$ and $\bigcap_{i=1}^n A_i$ are in \mathcal{A} as well;
4. \mathcal{A} is closed under countable unions and countable intersections: i.e., if A_1, A_2, A_3, \dots is a countable sequences of events in \mathcal{A} , then $\bigcup_{i=1}^{\infty} A_i$ and $\bigcap_{i=1}^{\infty} A_i$ are also both in \mathcal{A} .

Definition A.5 \mathcal{A} is an algebra if it satisfies (1),(2) and (3) above. It is a σ -algebra if it satisfies (1),(2) and (4) above.

Definition A.6 (Countably additive measure) We say that m is a countably additive measure if it's a nonnegative extended real-valued function whose domain of definition is a σ -algebra \mathcal{M} of sets (of real numbers) and we have $m(\bigcup E_n) = \sum m(E_n)$ for each sequence $\{E_n\}$ of disjoint sets in \mathcal{M} .

Let m be a countably additive measure defined for all sets in a σ -algebra \mathcal{M} .

1. If A and B are two sets in \mathcal{M} with $A \subset B$ then $m(A) \leq m(B)$. This property is called *monotonicity*.
2. Let $\{E_n\}$ be any sequence of sets in \mathcal{M} . Then $m(\bigcup E_n) \leq \sum mE_n$. This property of a measure is called *countable subadditivity*.
3. If there is a set $A \in \mathcal{M}$ s.t. $m(A) < \infty$, then $m(\emptyset) = 0$.

A *finitely additive* measure has the same definition except that is defined on an algebra and the property in the definition above is only required to hold for finite unions. Note the slight abuse of terminology: a finitely additive measure is not necessarily a measure.

Definition A.7 (Counting measure) Let $n(E)$ be ∞ for an infinite set E and be equal to the number of elements in E for finite sets. $n(\cdot)$ is a countably additive set function which is translation invariant and defined for all sets of real numbers. Let us call this measure the counting measure.

For each set A of real numbers consider the countable collections $\{I_n\}$ of open intervals which cover A , and for each such collection consider the sum of the lengths of the intervals in the collection. Since the lengths are positive numbers, this sum is uniquely defined independently of the orders of the terms. Then

Definition A.8 (Outer measure) We define the outer measure $m^*(A)$ of A as:

$$m^*(A) = \inf_{ACI_n} \sum l(I_n).$$

While the outer measure has the advantage that it's defined for all sets, it is not countable additive. It becomes countable additive, however, if we suitably reduce the family of sets on which it is defined. Perhaps the best way of doing this is to use the following definition due to Carathéodory:

Definition A.9 A set E is said to be measurable if for each set A we have $m^*(A) = m^*(A \cap E) + m^*(A \cap E^c)$.

If E is a measurable set, we define the *Lebesgue measure* $m(E)$ to be the outer measure of E . Thus m is the set function obtained by restricting the set function m^* to the family \mathcal{M} of measurable set. Two important properties of Lebesgue measure are summarized by the following proposition:

Proposition A.1 Let $\{E_i\}$ be a sequence of measurable sets. Then

$$m\left(\bigcup E_i\right) \leq \sum m(E_i).$$

If the sets E_n are piecewise disjoint, then

$$m\left(\bigcup E_i\right) = \sum m(E_i).$$

Proposition A.2 Let $\{E_i\}$ be an infinite decreasing sequence of measurable sets, with $m(E_1)$ finite. Then

$$m\left(\bigcap_{i=1}^{\infty} E_i\right) = \lim_{n \rightarrow \infty} m(E_n)$$

Since not all sets are measurable, it is of great importance to know that sets which arise naturally in certain constructions are measurable. If we start with a function f the most important sets which arise from it are those listed in the following properties:

Proposition A.3 *Let f be an extended real-valued function whose domain is measurable. Then the following statements are equivalent:*

- i. For each real number α , the set $\{x : f(x) > \alpha\}$ is measurable.*
- ii. For each real number α , the set $\{x : f(x) \geq \alpha\}$ is measurable.*
- iii. For each real number α , the set $\{x : f(x) < \alpha\}$ is measurable.*
- iv. For each real number α , the set $\{x : f(x) \leq \alpha\}$ is measurable.*

This statements imply

- v. For each extended real number α , the set $\{x : f(x) = \alpha\}$ is numerable*

Definition A.10 *An extended real-valued function f is said to be (Lebesgue) measurable if its domain is measurable and if it satisfies one of the first four statements of Proposition A.3*

A property is said to hold *almost everywhere* (abbreviated a.e.) if the set of points where it fails to hold is a set of measure zero. Thus in particular we say that $f = g$ a.e. if f and g have the same domain and $m(\{x : f(x) \neq g(x)\}) = 0$. Similarly we say that f_n converges to g almost everywhere if there is a set E of measure zero s.t. $f_n(x)$ converges to $g(x)$ for each $x \notin E$.

A.3 The Lebesgue Integral

The function χ_E defined by

$$\chi_E(x) = \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases}$$

is called *indicator (or characteristic) function* of E . A linear combination

$$\varphi(x) = \sum_{i=1}^n a_i \chi_{E_i}(x)$$

is called *simple function* if the sets E_i are measurable.

If φ vanishes outside a set of finite measure, we define the integral of φ by

$$\int \varphi(x) dx = \sum_{i=1}^n a_i m(A_i)$$

where $\{a_1, \dots, a_n\}$ is the set of non-zero values of φ and $A_i = \{x : \varphi(x) = a_i\}$. In a more compact way, and with E measurable set, we define

$$\int_E \varphi = \int \varphi \cdot \chi_E$$

Definition A.11 (The Lebesgue integral) *If f is a bounded measurable function defined on a measurable set E with $m(E)$ finite, we define the (Lebesgue) integral of f over E by*

$$\int_E f(x) dx = \inf \int_E \psi(x) dx$$

for all simple functions $\psi \geq f$.

Theorem A.1 (Bounded Convergence) *Let $\{f_n\}$ be a sequence of measurable functions defined on a set E of finite measures, and suppose that there is a real number M s.t. $|f_n(x)| \leq M$ for all n , for all x . If $f(x) = \lim f_n(x)$ for each $x \in E$, then*

$$\int_E f = \lim \int_E f_n.$$

If $\{f_n\}$ is a sequence of measurable functions which converges a.e. to f , then, as we will see, the Fatou's Lemma, the Monotone Convergence Theorem and the Lebesgue Convergence Theorem all state that under suitable hypothesis we can assert something about $\int f$ in terms of $\int f_n$.

Theorem A.2 (Fatou's Lemma) *If $\{f_n\}$ is a sequence of non-negative measurable functions and $f_n(x) \rightarrow f(x)$ a.e. on a set E , then*

$$\int_E f \leq \liminf \int_E f_n$$

Theorem A.3 (Monotone Convergence) *Let $\{f_n\}$ be an increasing sequence of non-negative measurable functions, and let $f(x) = \lim f_n(x)$. Then*

$$\int_E f = \lim \int_E f_n$$

Definition A.12 *A non-negative measurable function f is called integrable over the measurable set E if*

$$\int_E f < \infty$$

Theorem A.4 (Lebesgue Convergence) *Let g be integrable over E and let $\{f_n\}$ be a sequence of measurable functions s.t. $|f_n| \leq g$ on E and for almost all $x \in E$ we have $f(x) = \lim f_n(x)$. Then*

$$\int_E f = \lim \int_E f_n.$$

Theorem A.5 Let $\{g_n\}$ be a sequence of integrable functions which converges a.e. to an integrable function g . Let $\{f_n\}$ a sequence of measurable functions s.t. $|f_n| \leq g_n$ and $\{f_n\}$ converge to f a.e. If

$$\int g = \lim \int g_n,$$

then

$$\int f = \lim \int f_n$$

Suppose that $\{f_n\}$ is a sequence of measurable functions s.t. $\int f_n \rightarrow 0$. What can we say about the sequence $\{f_n\}$? Perhaps the most important property of such a sequence is that for each positive η the measure of the sets $\{x : |f_n| > \eta\}$ must tend to zero. This leads us to the following definition:

Definition A.13 (Convergence in Measure) A sequence $\{f_n\}$ of measurable functions is said to converge to f in measure if, given $\varepsilon > 0$, there is an N s.t. $\forall n \geq N$ we have

$$m(\{x : |f(x) - f_n(x)| \geq \varepsilon\}) < \varepsilon.$$

A.4 The $L^p[0, 1]$ Spaces

In this section we will see some spaces of functions of a real variable.

Definition A.14 (L^p Spaces) Let p be a positive real number. A measurable function defined on $[0, 1]$ is said to belong to the space $L^p = L^p[0, 1]$ if

$$\int_0^1 |f|^p < \infty$$

Thus L^1 consists precisely of the Lebesgue integrable functions on $[0, 1]$. Since $|f + g|^p \leq 2^p(|f|^p + |g|^p)$, we see that the sum of two functions on L^p is again in L^p . Since αf is in L^p whenever f is, we have $\alpha f + \beta g$ in L^p whenever f and g are.

Definition A.15 (Linear Space) A space X of real-valued function is called linear space (or vector space) if it has the property that

$$\alpha f + \beta g \in X \quad \forall f, g \in X, \quad \forall \alpha, \beta \in \mathbb{R}$$

Thus the L^p spaces are linear spaces.

For a function $f \in L^p$ we define

$$\|f\| = \|f\|_p = \left(\int_0^1 |f|^p \right)^{1/p}.$$

with the property that $\|f\| = 0$ iff $f = 0$ and $\forall \alpha \in \mathbb{R}$, $\|\alpha f\| = |\alpha| \|f\|$.

Definition A.16 (Normed Linear Space) A linear space is said to be a normed linear space if we have assigned a nonnegative real number $\|f\|$ to each f s.t.

$$\|\alpha f\| = |\alpha| \|f\|$$

$$\|f + g\| \leq \|f\| + \|g\|$$

$$\|f\| = 0 \Leftrightarrow f \equiv 0$$

Unfortunately, the norm of the L^p spaces does not satisfy the last requirement, from $\|f\| = 0$ we can only conclude that $f = 0$ a.e. We will, however, consider two measurable functions to be equivalent if they are equivalent a.e.; and, if we do not distinguish between equivalent functions, then the L^p spaces are normed linear spaces.

It is convenient to denote L^∞ the space of all bounded measurable functions on $[0, 1]$ (or rather all measurable functions which are bounded except possibly on a subset of measure zero). Again we identify functions which are equivalent. Then L^p is a linear space, and it becomes a normed linear space if we define:

$$\|f\| = \|f\|_\infty = \text{ess sup } |f(t)|,$$

where $\text{ess sup } f(t)$ is the infimum of $\sup g(t)$ as g ranges over all functions which are equal to f a.e. Thus

$$\text{ess sup } f(t) = \inf\{M : m(t : f(t) > M) = 0\}$$

Proposition A.4 (Hölder Inequality) If p and q are nonnegative extended real numbers s.t. $\frac{1}{p} + \frac{1}{q} = 1$, and if $f \in L^p$ and $g \in L^q$, then $f \cdot g \in L^1$ and

$$\int |fg| \leq \|f\|_p \cdot \|g\|_q.$$

Equality holds iff, for some nonzero constants α and β , we have $\alpha|f|^p = \beta|g|^q$ a.e.

Proposition A.5 (Minkowski Inequality) If f and g are in L^p , then so is $f+g$ and

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

The notion of convergence for a sequence of real numbers generalizes to give us a notion of convergence for sequences in a linear normed space.

Definition A.17 (Convergence) A sequence $\{f_n\}$ in a normed linear space is said to be convergent to an element f if, given $\epsilon > 0$, there is an N s.t. for all $n > N$ we have $\|f - f_n\| < \epsilon$. If f_n converges to f we write $f = \lim_{n \rightarrow \infty} f_n$ or $f_n \rightarrow f$.

Another way of formulating the convergence of f_n to f is by noting that $f_n \rightarrow f$ if $\|f - f_n\| \rightarrow 0$. Convergence in the space L^p $1 \leq p < \infty$, is often referred to as *convergence in the mean of order p* .

Definition A.18 (Cauchy sequence) A sequence $\{f_n\}$ in a normed linear space is said to be a Cauchy sequence if, given $\epsilon > 0$, there is an N s.t. for all $n > N$ and $m > N$ we have

$$\|f_n - f_m\| < \epsilon$$

Definition A.19 (Completeness) A normed linear space is called complete if every Cauchy sequence in the space converges, that is, if for each Cauchy sequence $\{f_n\}$ in the space there is an element f in the space s.t. $f_n \rightarrow f$. A complete normed linear space is called Banach space.

Definition A.20 A series $\{f_n\}$ in a normed linear space is said to be summable to a sum s if s is in the space and the sequence of partial sums of the series converges to s ; that is

$$\left\| s - \sum_{i=1}^n f_i \right\| \rightarrow 0.$$

Proposition A.6 A normed linear space X is complete if and only if every absolutely summable series is summable.

Theorem A.6 (Riesz-Fischer) The L^p spaces are complete.

Definition A.21 (Linear Functional) We define a linear functional on a normed linear space X to be a mapping F of the space X into the set of real numbers s.t.

$$F(\alpha f + \beta g) = \alpha F(f) + \beta F(g).$$

We say that the linear functional is bounded if there is a constant M s.t. $|F(f)| \leq M \cdot \|f\|$ for all f in X . The smallest constant M for which the inequality is true is called the norm of F . Thus

$$\|F\| = \sup \frac{|F(f)|}{\|f\|},$$

as f ranges over all nonzero elements of X .

Proposition A.7 Each function g in L^q defines a bounded linear functional F on L^p by

$$F(f) = \int fg.$$

We have $\|F\| = \|g\|_q$.

To conclude the present section, let us show that for $1 \leq p \leq \infty$ the converse of this proposition holds, i.e., we obtain every bounded linear functional on L^p in this manner. The following lemma holds:

Lemma 3 *Let g be an integrable function on $[0, 1]$, and suppose that there is a constant M s.t.*

$$\left| \int fg \right| \leq M \|f\|_p$$

for all bounded measurable functions f . Then g is in L^q , and $\|g\|_q \leq M$.

We are now in position to give the following characterization of the bounded linear functionals on L^p for $1 \leq p \leq \infty$:

Theorem A.7 (Riesz Representation) *Let F be a bounded linear functional on L^p , $1 \leq p \leq \infty$. Then there is a function g in L^q s.t.*

$$F(f) = \int fg.$$

We have also $\|F\| = \|g\|_q$.

A.5 Metric Spaces

The system of real numbers has two types of property. The first type consists of the algebraic, dealing with addition, multiplication, etc. The other type consists of properties having to do with the notion of distance between two numbers and with the concept of limit. These latter properties are called topological or metric, and here we want to introduce these properties in general spaces, where the notion of distance is defined.

Definition A.22 (Metric) *A metric space (X, ρ) is a nonempty set X of elements (which we call points) together with a real-valued function ρ defined on $X \times X$ s.t. $\forall x, y$ and $z \in X$:*

- i. $\rho(x, y) \geq 0$;
- ii. $\rho(x, y) = 0 \Leftrightarrow x = y$;
- iii. $\rho(x, y) = \rho(y, x)$;
- iv. $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$.

The function ρ is called a metric

A function f on a metric space (X, ρ) into a metric space (Y, σ) is a rule which associates to each $x \in X$ a unique $y \in Y$. We also call f a mapping of X into Y , or a function.

Definition A.23 (Continuity) The function f is said to be continuous at x if, for every $\epsilon > 0$, there is a $\delta > 0$ so that if $\rho(x, y) < \delta$, then $\sigma(f(x), f(y)) < \epsilon$. The function is called continuous if it is continuous at each $x \in X$.

Definition A.24 (Homeomorphism) A one-to-one mapping f of X onto Y is called a homeomorphism between X and Y if f is continuous and the mapping f^{-1} inverse to f is also continuous.

The spaces X and Y are said to be *homeomorphic* if there is an homeomorphism between them.

Not all properties in a metric spaces are preserved under a homeomorphism.

Definition A.25 (Isometry) A homeomorphism which leaves distances unchanged, that is, one for which

$$\sigma(h(x_1), h(x_2)) = \rho(x_1, x_2)$$

for all x_1 and x_2 in X , is called an isometry between X and Y .

Example 4 Let $(X, \rho) = (Y, \sigma) = (\mathbb{R}, d(\cdot, \cdot))$, with $d(x, y) = |x - y|$. Then the application $h(x) = x + t$, $t \in \mathbb{R}$, is trivially an isometry.

A.6 Topological Spaces

Definition A.26 (Topology) A topological space (X, κ) is a nonempty set X of points together with a family κ of subsets (which we will call open) possessing the following properties:

- i. $X \in \kappa$, $\emptyset \in \kappa$;
- ii. O_1 and $O_2 \in \kappa$ imply $O_1 \cap O_2 \in \kappa$;
- iii. $O_\alpha \in \kappa$ implies $\bigcup_\alpha O_\alpha \in \kappa$.

The family κ is called a topology for the set X .

The properties in this definition are all satisfied by open sets in a metric space (X, ρ) , and hence to each metric space (X, ρ) we can associate a topological space (X, κ) , where κ is the family of open sets in (X, ρ) ¹. A topological space which is associated in this manner to some metric space is called *metrizable*, and the metric ρ is said to be a metric for the topological space.

Definition A.27 (Base) A collection \mathcal{B} of open sets of a topological space X is called a base for the topology κ of X if for each open sets $O \in \kappa$ and each $x \in O$ there is a set $B \in \mathcal{B}$ s.t. $x \in B \subset O$.

¹Recall that $U \in \kappa$ is an open set if and only if, $\forall x \in U, \exists \epsilon > 0$ such that $\{y : \rho(x, y) < \epsilon\} \subseteq U$.

A.7 Banach Spaces

We are now going to see a class of spaces which are endowed with both a topological and an algebraic structure.

Definition A.28 (Vector Space) *A set X of elements is called a vector space (or linear space or linear vector space) over the reals if we have a function $+$: $X \times X \rightarrow X$ and a function \cdot : $\mathbb{R} \times X \rightarrow X$ which satisfy the following conditions:*

- i. $x + y = y + x$.*
- ii. $(x + y) + z = x + (y + z)$.*
- iii. There is a vector $\underline{0} \in X$ s.t. $x + \underline{0} = x$, $\forall x \in X$.*
- iv. $\lambda(x + y) = \lambda x + \lambda y; \forall \lambda \in \mathbb{R}, \forall x, y \in X$.*
- v. $(\lambda + \mu)x = \lambda x + \mu x; \forall \lambda, \mu \in \mathbb{R}, \forall x \in X$.*
- vi. $\lambda(\mu x) = (\lambda\mu)x; \forall \lambda, \mu \in \mathbb{R}, \forall x \in X$.*
- vii. $0 \cdot x = \underline{0}, 1 \cdot x = x$.*

We call $+$ *addition* and \cdot *multiplication* by scalars. It should be noted that the element $\underline{0}$ defined in (iii) is unique. The element $(-1)x$ is called *negative* of x and written $-x$.

Definition A.29 (Norm) *A nonnegative real-valued function $\|\cdot\|$ defined on a vector space is called norm if*

- i. $\|x\| = 0 \Leftrightarrow x = \underline{0}$.*
- ii. $\|x + y\| \leq \|x\| + \|y\|$.*
- iii. $\|\alpha x\| = |\alpha| \|x\|$.*

A normed vector space becomes a metric space if we define a metric ρ by $\rho(x, y) = \|x - y\|$. When we speak about metric properties in a normed space we are referring to this metric.

Definition A.30 (Banach Space) *If a normed vector space is complete in this metric, it is called Banach space.*

Definition A.31 (Linear Operator) *A mapping A of a vector space X into a vector space Y is called linear operator, or a linear transformation if*

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A x_1 + \alpha_2 A x_2$$

for all $x_1, x_2 \in X$ and all real α_1, α_2 . If X, Y are normed vector spaces, we call a linear operator bounded if there is a constant M s.t. for all x we have $\|Ax\| \leq M\|x\|$. We call the least such M the norm of A and denote it by $\|A\|$. Thus

$$\|A\| = \sup_{x \in X, x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

Proposition A.8 *The space \mathbb{B} of all bounded linear operators from a normed vector space X to a Banach space Y is itself a Banach space.*

A linear functional on a vector space X is a linear operator from X to the space \mathbb{R} of real numbers. Thus a linear functional is a real-valued function on X s.t. $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$. The first question with which we will be concerned is that of extending a linear functional from a subspace to the whole space X in such a manner that various properties of the functional are preserved. The principal result in this direction is the following:

Theorem A.8 (Hans-Banach) *Let p be a real-valued function defined on the vector space X satisfying $p(x+y) \leq p(x) + p(y)$ and $p(\alpha x) = \alpha p(x)$ for each $\alpha \geq 0$. Suppose that f is a linear functional defined on a subspace S and that $f(s) \leq p(s) \forall s \in S$. Then there is a linear functional F defined on X s.t. $F(x) \leq p(x) \forall x$, and $F(s) = f(s) \forall s \in S$.*

Now we introduce the definition of duality, a concept which we will often deal with in this thesis.

Definition A.32 (Dual) *The space of bounded linear functional on a normed space X is called the dual (or conjugate) of X and it is denoted by X^* .*

Since \mathbb{R} is complete, the dual X^* of any normed space X is a Banach space by Proposition A.8. Two normed vector spaces are said to be isometrically isomorphic if there is a one-to-one linear mapping of one of them onto the other which preserves norms. From an abstract point of view, isometrically isomorphic spaces are identical, the isomorphism merely amounting to a renaming of the elements. We saw in section A.4 that the dual of L^p was (isometrically isomorphic to) L^q for $1 \leq p < \infty$ where q is such that $\frac{1}{p} + \frac{1}{q} = 1$ and that there was a natural representation of the bounded linear functionals on L^p by elements of L^q . A similar representation does not hold for bounded linear functionals on $L^\infty[0, 1]$.

If we consider the dual X^{**} of X^* , then to each $x \in X$ there corresponds an element $\varphi_x \in X^{**}$ defined by $\varphi_x(f) = f(x)$. We have $\|\varphi_x\| = \sup_{\|f\|=1} f(x)$. Since $f(x) \leq \|f\| \|x\|$, we have $\|\varphi_x\| \leq \|x\|$. We can also prove that $\|\varphi_x\| = \|x\|$ (see Proposition 6 chap.10 in [27]). Since φ is clearly a mapping, φ is an isometric

isomorphism of X onto some linear subspaces $\varphi[X]$ of X^{**} . The mapping φ is called the natural isomorphism of X into X^{**} , and if $\varphi[X] = X^{**}$ we say that X is *reflexive*.

Thus L^p is reflexive if $1 < p < \infty$. Since there are functionals on L^∞ which are not given by integration with respect to a function in L^1 , it follows that L^1 is not reflexive. It should be observed that X may be isometric with X^{**} without being reflexive.

Just as the notion of metric space generalizes to that of a topological space, so the notion of a normed linear space generalizes to that of a topological vector space:

Definition A.33 (Topological Vector Space) *A linear vector space X with a topology κ on it is called a topological vector space if addition is a continuous function from $X \times X$ into X and multiplication by scalars is a continuous function from $\mathbb{R} \times X$ into X .*

Definition A.34 (Weak Topology) *If X is any vector space and \mathcal{F} a collection of linear functionals on X , we define the weak topology generated by \mathcal{F} to be the weakest topology s.t. each $f \in \mathcal{F}$ is continuous.*

If X is a normed vector space and the functionals in \mathcal{F} are all continuous (that is, if $\mathcal{F} \in X^*$), then the weak topology generated by \mathcal{F} is weaker (has fewer open sets) than the norm topology of X . We usually call the metric topology generated by the norm the strong topology of X and the weak topology on X generated by X^* the *weak topology* of X . Thus we speak of strongly closed and strongly open sets when referring to the strong topology and weakly open and weakly closed sets when for the weak topology. Every weakly closed set is strongly closed but not conversely. Every strongly convergent sequence is weakly convergent.

If we apply the notion of weak topology to the dual X^* of a normed space X , we see that the weak topology of X^* is the weakest topology for X^* s.t. all of the functionals in X^{**} are continuous. The weak topology turns out to be less useful than the weak topology for X^* generated by X (or more precisely, by $\varphi[X]$ where φ is the natural embedding of X into X^{**}). This topology is called the *weak* topology* for X^* and is even weaker than the weak topology. Thus a weak* closed subset of X^* is weakly closed, and weak convergence implies weak* convergence.

We have already seen what is a convex function. We will see now what is a convex set.

Definition A.35 (Convexity) *A subset K of a vector space X is said to be convex if*

$$\forall x, y \in K \Rightarrow \lambda x + (1 - \lambda)y \in K, \forall \lambda \in [0, 1].$$

The set $\{z : z = \lambda x + (1 - \lambda)y, \forall \lambda \in [0, 1]\}$ is called the *line segment* joining x and y . The points x and y are its endpoints, and a point z for which $\lambda \in (0, 1)$ is called an interior point of the segment. Thus a set K is convex if and only if whenever it contains x and y it contains the segment joining x and y .

We give now some properties of convex sets.

Lemma 4 *If K_1 and K_2 are convex sets, so also are the sets $K_1 \cap K_2, \lambda K_1, K_1 + K_2$.*

A point x_0 is said to be an *internal point* of a set K if the intersection with K of each line through x_0 contains an open interval about x_0 .

Definition A.36 *The support function $p(x|C) = p(x)$ of a convex set C in \mathbb{R}^k is defined by*

$$p(x) = \sup\{y(x) | y \in C\}.$$

Lemma 5 *If K is a convex set containing 0 as an interior point, then the support function p has the following properties:*

- i. $p(\lambda x) = \lambda p(x) \quad \lambda \geq 0$.*
- ii. $p(x + y) \leq p(x) + p(y)$.*
- iii. $\{x : p(x) < 1\} \subset K \subset \{x : p(x) \leq 1\}$.*

Definition A.37 *A topological vector space is called locally convex if we can find a base for the topology consisting of convex sets.*

Definition A.38 (Cone) *Given a vector space V , a set $C \subset V$ is called cone if and only if for all $x \in C$ and $\lambda \in \mathbb{R}$ we have $\lambda x \in C$.*

A cone C is called *positively homogeneous* if and only if for every pair $x, y \in \text{int}(C)$ there exists a linear definitely positive mapping A such that does not affect C and such that $A(x) = y$.

A.8 Measure and Integration: the Radon-Nikodym Theorem

We will deal now with some results in measure theory, results that we will use when defining a convex risk measure.

Definition A.39 *A measurable space is a set Ω together with a collection \mathcal{F} of subset of Ω which is a σ -algebra. The elements of \mathcal{F} are called measurable sets.*

Definition A.40 Let $\pi \in \mathbf{ba}(\Omega, \mathcal{F}, P)$, $\pi \geq 0$. Then π is said to be purely finitely additive if the only countably additive nonnegative set function $\xi \in \mathbf{ba}(\Omega, \mathcal{F}, P)$ such that $\xi \leq \pi$ is $\xi = 0$.

Definition A.41 Let (X, \mathcal{B}) a fixed measurable space, and let μ and ν two measures defined on it. A measure ν is said to be absolutely continuous with respect to a measure μ if

$$\mu(A) = 0 \Rightarrow \nu(A) = 0, \text{ for each } A \in \mathcal{B}$$

We use the symbolism $\nu \ll \mu$ for ν absolutely continuous w.r.t. μ .

Definition A.42 If $\nu \ll \mu$ and $\mu \ll \nu$, then ν and μ have the same class of null sets, and ν, μ are said to be mutually equivalent, denoted by $\mu \equiv \nu$.

On the other hand

Definition A.43 Let μ, ν be two measures on (Ω, \mathcal{F}) . If there is a set $B \in \mathcal{F}$ such that $\mu(B) = 0$ and $\nu(B^c) = 0$ (or equivalently $\nu(A) = \nu(A \cap B)$, $A \in \mathcal{F}$), then ν and μ are called mutually singular or orthogonal, denoted by $\mu \perp \nu$.

Clearly this relation is symmetric (i.e. $\mu \perp \nu \Leftrightarrow \nu \perp \mu$), in contrast with absolute continuity.

Whenever we are dealing with more than a measure on a measurable space (X, \mathcal{B}) , the term 'almost everywhere' becomes ambiguous, and we must specify almost everywhere with respect to μ or a.e. with respect to ν etc... These are abbreviated μ -a.e., ν -a.e.. If $\nu \ll \mu$ and a property holds μ -a.e., then it holds ν -a.e..

Theorem A.9 (Lebesgue decomposition) Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, and ν be a given σ -finite measure on \mathcal{F} . Then ν can be uniquely expressed as $\nu = \nu_1 + \nu_2$ where $\nu_1 \ll \mu$ and $\nu_2 \perp \mu$.

See [23] for the proof.

Theorem A.10 (Radon-Nikodym) Let (X, \mathcal{B}, μ) a σ -finite measure space, and let ν be a measure defined on \mathcal{B} which is absolutely continuous w.r.t. μ . Then there is a nonnegative measurable functions f s.t. for each set $E \in \mathcal{B}$ we have

$$\nu(E) = \int_E f d\mu.$$

The function f is unique in the sense that if g is any measurable function with this property then $g = f$ μ -a.e..

The function f given by Theorem A.10 is called the *Radon-Nikodym derivative* of ν with respect to μ . It is denoted by $\left[\frac{d\nu}{d\mu} \right]$

Definition A.44 (Probability Space) A triple (Ω, \mathcal{F}, P) on the domain Ω , where (Ω, \mathcal{F}) is a measurable space, \mathcal{F} are the measurable subsets of Ω , and P is a measure on Ω with $P[\Omega] = 1$ is called probability space.

Definition A.45 (Atom) The set $A \in \mathcal{F}$ is called an atom of (Ω, \mathcal{F}, P) , if $P[A] > 0$ and if each $B \in \mathcal{F}$ with $B \subseteq A$ satisfies either $P[B] = 0$ or $P[B] = P[A]$.

Definition A.46 (Atomless probability space) A probability space (Ω, \mathcal{F}, P) is called atomless if it does not contain any atom.

A.8.1 The general L^p Spaces

If (X, \mathcal{B}, μ) is a measure space, we denote by $L^p(\mu)$ the space of all measurable functions on X for which $\int |f|^p d\mu < \infty$, considering two functions in L^p to be equivalent if they are equal a.e.. As in Section A.4 we define $L^\infty(\mu)$ to be the space of bounded measurable functions. For $1 \leq p < \infty$ we set

$$\|f\|_p = \left\{ \int |f|^p d\mu \right\}^{\frac{1}{p}},$$

and for $p = \infty$ we set

$$\|f\|_\infty = \text{ess sup } |f|.$$

Note that the space $L^\infty(\mu)$ depends on the choice of μ to determine the norm and the classes of equivalent functions, but that this only requires knowing what the set of zero measure are.

The Hölder and Minkowski inequalities and the Riesz-Fisher theorem follow just as in Section A.4, and we summarize them in the following theorem:

Theorem A.11 For $1 \leq p \leq \infty$ the space $L^p(\mu)$ are Banach spaces, and if $f \in L^p(\mu)$, $g \in L^q(\mu)$, with $\frac{1}{p} + \frac{1}{q} = 1$, then $fg \in L^1(\mu)$ and

$$\int |fg| d\mu \leq \|f\|_p \|g\|_q.$$

Appendix B

Principles of Convex Analysis

In this chapter we will give some basic definitions and results of convex analysis. For more details see [28], [16] and [24].

B.1 Convex Functions on \mathbb{R}

In this section we will designate by I a (closed, open or half-open, finite or infinite) interval on \mathbb{R} .

Definition B.1 (Convexity) *Let f be a function $I \rightarrow \mathbb{R}$.*

(a) *f is said to be convex if*

$$f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b)$$

for all $a, b \in I$ and all $\lambda \in \mathbb{R}$ s.t. $0 \leq \lambda \leq 1$.

(b) *f is said to be strictly convex if it is convex and the strict inequality holds in (a) whenever $a \neq b$.*

If we look at the graph of f , this condition can be formulated geometrically by saying that each point on the chord between $(x, f(x))$ and $(y, f(y))$ is above the graph of f . An important property of the chords of a convex functions is given by the following lemma:

Lemma 6 *If f is convex on (a, b) and if x, y, x', y' are points of (a, b) with $x \leq x' < y'$ and $x < y \leq y'$, then the chord over (x', y') has larger slope than the chord*

over (x, y) ; that is,

$$\frac{f(y) - f(x)}{y - x} \geq \frac{f(y') - f(x')}{y' - x'}$$

Theorem B.1 *Let $f : I \rightarrow \mathbb{R}$ be convex. Then f has a right derivative and a left derivative at every point of $\text{int}(I)$, and f'_- and f'_+ are non-decreasing on $\text{int}(I)$. If $c \in \text{int}(I)$, we have*

$$f'_-(c) \leq f'_+(c)$$

and

$$f(x) \geq f(c) + f'_-(c)(x - c), \quad f(x) \geq f(c) + f'_+(c)(x - c)$$

for all $x \in I$.

It is not really difficult to prove the following inequality.

Proposition B.1 (Jensen Inequality) *Let f be a convex function on $(-\infty, \infty)$ and h an integrable function on $[0, 1]$. Then*

$$\int f(h(t))dt \geq f\left[\int h(t)dt\right].$$

This inequality has a geometric interpretation worth mentioning. Since the point $\lambda x_1 + (1 - \lambda)x_2$ is the centroid of masses λ and $(1 - \lambda)$ at x_1 and x_2 , we can say that a function is convex if its value at the centroid of a two-point mass is less than the weighted average of its value at the two points. The Jensen inequality is a generalization of this fact: If we define a mass distribution μ in the line by setting $\mu(a, b] = m(\{t : a < f(t) \leq b\})$, then the $\int f(t)dt$ is the centroid of this mass and $\int \varphi(f(t))dt = \int \varphi(x)d\mu$ is the weighted average of φ .

B.1.1 The Conjugate Function

A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is convex if and only if there exists a function $g : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ such that

$$f(x) = \sup_{y \in \mathbb{R}} [xy - g(y)] \tag{B.1}$$

for all $x \in \mathbb{R}$. The above function g is called the *conjugate* of f .

An alternative way to define the conjugate of a function is the following:

Definition B.2 (Fenchel-Legendre transform) *The Fenchel-Legendre transform of a function f on \mathbb{R} is defined as*

$$f^*(y) \doteq \sup_{x \in \mathbb{R}} \{yx - f(x)\}, \quad y \in \mathbb{R}$$

If $f \neq +\infty$, then f^* is a convex and lower semi-continuous as the supremum of the affine functions $y \rightarrow yx - f(x)$. In particular, f^* is a convex function which is continuous on its effective domain. If f is itself a convex function, then f^* is also called the *conjugate function* of f .

Proposition B.2 *Let f be a convex function.*

(a) *For all $x, y \in \mathbb{R}$,*

$$xy \leq f(x) + f^*(y) \quad (\text{B.2})$$

with equality if x belongs to the interior of $\text{dom} f$ and if $y \in [f'_-(x), f'_+(x)]$.

(b) *If f is lower semi-continuous, then $f^{**} = f$, i.e.,*

$$f(x) = \sup_{y \in \mathcal{R}} \{xy - f^*(y)\}$$

See [16] for the proof.

We now summarize some basic properties of the functions f and f^* . See [14] for the proof.

Lemma 7 *Let us assume f and f^* as defined above. Then*

1. $f^*(0) = -\inf_{x \in \mathbb{R}} f(x)$ and $f^*(y) \geq -f(0)$ for all z .
2. The set $N \doteq \{y \in \mathbb{R} | f^*(y) = -f(0)\}$ is nonempty, $y_1 \doteq \inf N \geq 0$, and $f^*(y) = \sup_{x \geq 0} (xy - f(x))$ for $y \geq y_1$. In particular, f^* is non-decreasing in $[y_1, \infty)$.
3. $y_0 \doteq \inf\{y \in \mathbb{R} | f^*(y) < \infty\} \in [0, \infty)$.
4. $\frac{f^*(y)}{y} \rightarrow \infty$ as $y \uparrow \infty$

When the function is concave, as could be the utility function, we have the notion of *concave conjugate*.

Definition B.3 (Concave Conjugate) *Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be a concave function. Then we define the concave conjugate $g^* : \mathbb{R} \rightarrow \mathbb{R} \cup \{-\infty\}$ of g by:*

$$g^*(y) \doteq \inf_{x \in \mathbb{R}} \{xy - g(x)\}, \quad y \in \mathbb{R}. \quad (\text{B.3})$$

Let $x \in \text{int}(\mathcal{D})$ and suppose that $g : \mathcal{D} \rightarrow \mathbb{R}$ is a strictly concave, differentiable function and denote with $I = (u')^{-1}$ the inverse function of u' . Then

$$g^*(y) = yI(y) - u(I(y)).$$

Remark 13 We know that if f is a convex function, $g = -f$ is a concave function. Then the relation between conjugate and concave conjugate is the following:

$$f^*(y) = -g^*(-y)$$

where f, g are as defined above and $y \in \mathbb{R}$

Proof.

$$\begin{aligned} f^*(y) &= \\ \sup_{x \in \mathbb{R}} \{yx - f(x)\} &= \sup_{x \in \mathbb{R}} \{yx + g(x)\} \\ \sup_{x \in \mathbb{R}} \{-[(-y)x - g(x)]\} &= - \inf_{x \in \mathbb{R}} \{(-y)x - g(x)\} \\ &= -g^*(-y). \quad \square \end{aligned}$$

And in the case as $l(x) = -u(-x)$, where l is the loss function and u is the utility function, we have

$$l^*(y) = -u^*(y)$$

The proof follows the proof given above.

B.1.2 Convex Functions With Values in $\bar{\mathbb{R}}$

We now consider more general functions, with values in $\bar{\mathbb{R}} := \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$. So we can now provide a generalization of the concept of convex functions.

Definition B.4 A function $f : \mathbb{R} \rightarrow \bar{\mathbb{R}}$ is said to be convex if for all $x, y, \lambda, \mu, \nu \in \mathbb{R}$ such that $f(x) < \mu, f(y) < \nu, 0 < \lambda < 1$

$$f(\lambda x + (1 - \lambda)y) < \lambda\mu + (1 - \lambda)\nu$$

Definition B.5 (a) The effective domain of a convex function $f : \mathbb{R} \rightarrow \bar{\mathbb{R}}$, denoted by $\text{dom}(f)$, is the set $\{x \in \mathbb{R} \mid f(x) < +\infty\}$

(b) A proper convex function on \mathbb{R} is a convex function $\mathbb{R} \rightarrow \bar{\mathbb{R}}$ which is not identically $+\infty$.

(c) An improper convex function on \mathbb{R} is a convex function on \mathbb{R} which is not proper.

Now we give the definition of some concepts closely related with the one of convexity.

Definition B.6 Let f be a function $I \rightarrow \bar{\mathbb{R}}$.

(a) f is said to be quasi-convex if

$$f(\lambda a + (1 - \lambda)b) \leq f(b)$$

for all $a, b \in \mathbb{R}$ with $f(a) \leq f(b)$ and all $\lambda \in (0, 1)$.

(b) f is said to be strictly quasi-convex if

$$f(\lambda a + (1 - \lambda)b) < f(b)$$

for all $a, b \in \mathbb{R}$ with $f(a) < f(b)$ and all $\lambda \in (0, 1)$.

A strictly quasi-convex function is not necessarily quasi convex.

B.2 Convex Functions On a Linear Space

In this section we designate by V a linear space over \mathbb{R} and by E a linear topological space over \mathbb{R} , both containing more than one point.

Definition B.7 (The Epigraph) Let X be a set and f a function $X \rightarrow \bar{\mathbb{R}}$. The epigraph $\text{epi}(f)$ of f is the set

$$\{(x, \lambda) \in X \times \mathbb{R} \mid f(x) \leq \lambda\}.$$

In the sequel, properties of f will sometimes be described in terms of property of $\text{epi}(f)$.

If X is a topological space, we endow $X \times \mathbb{R}$ with the product topology. *Closedness* of $\text{epi}(f)$ turns out to correspond with *lower semi-continuity* of f

Let X be a topological space.

Definition B.8 (Lower Semi-Continuity) Let f be a function $X \rightarrow \bar{\mathbb{R}}$. f is said to be lower semi-continuous at a if for each $K \in \mathbb{R}$, $K < f(a)$ there exists a neighborhood U such that $f(U) > K$. f is said to be lower semi-continuous if f is lower semi-continuous at each point of X .

Remark 14 (a) A continuous function is lower semi-continuous.

(b) If $a \in X$ is an accumulation point of X and $f(a) = +\infty$, and if f is lower semi-continuous at a , then

$$\lim_{x \rightarrow a} f(x) = +\infty$$

(c) If $f(a) = -\infty$, then f is lower semi-continuous at a .

Theorem B.2 *Let f be a function $X \rightarrow \bar{\mathbb{R}}$. The following conditions are equivalent:*

- (a) f is lower semi-continuous.
- (b) $\{x \in X \mid f(x) > \lambda\}$ is open for each $\lambda \in \mathbb{R}$.
- (c) $\{x \in X \mid f(x) \leq \lambda\}$ is closed for each $\lambda \in \mathbb{R}$.
- (d) $\text{epi}(f)$ is closed (as subset of $X \times \mathbb{R}$).

B.3 Duality Theory

In this section we denote with E as a normed linear space (containing more than one point) over \mathbb{R} , with norm $x \rightarrow \|x\|$, and with E' the dual of E . The separation theorem implies (see [28]) that for each $x \in E$, $x \neq 0$ there exists $x' \in E'$ such that $x'(x) \neq 0$.

B.3.1 The Conjugate Function

Definition B.9 (Conjugate) (a) *The conjugate (or dual or polar) of a function $f : E \rightarrow \bar{\mathbb{R}}$ is the function $f^* : E' \rightarrow \bar{\mathbb{R}}$ defined by*

$$f^*(x') = \sup_{x \in E} \{x'(x) - f(x)\} \quad (x' \in E').$$

(b) *The conjugate of a function $g : E' \rightarrow \bar{\mathbb{R}}$ is the function $g^* : E \rightarrow \bar{\mathbb{R}}$ defined by*

$$g^*(x) = \sup_{x' \in E'} \{x'(x) - g(x')\} \quad (x \in E).$$

(c) *The bipolar (or biconjugate) f^{**} of a function f from E to $\bar{\mathbb{R}}$ or from E' to $\bar{\mathbb{R}}$ is the conjugate $(f^*)^*$ of the conjugate of f .*

Remark 15 *If $f^*(x')$ is finite, then it equals to the smallest number α satisfying*

$$f(x) \geq x'(x) - \alpha$$

whenever $x \in E$.

We give now some simple properties of the conjugate function.

- (a) If f, h are functions from E to $\bar{\mathbb{R}}$ such that $f \leq h$, then $f^* \geq h^*$.
- (b) $(+\infty)^* = -\infty$.

(c) If there is a point where $f : E \rightarrow \bar{\mathbb{R}}$ has the value $-\infty$, then $f^* = +\infty$. In particular $(-\infty)^* = +\infty$.

Note that (b) and (c) imply that the formula $f^{**} = f$ is generally *not* true. One can show that for all $f : E \rightarrow \bar{\mathbb{R}}$ we have $f^{**} \leq f$.

(d) If $\{f_\alpha | \alpha \in A\}$ is an arbitrary collection of functions $E \rightarrow \bar{\mathbb{R}}$, then

$$\left(\inf_{\alpha} f_{\alpha}\right)^* = \sup_{\alpha} f_{\alpha}^*$$

$$\left(\sup_{\alpha} f_{\alpha}\right)^* \leq \inf_{\alpha} f_{\alpha}^*$$

In the last inequality, equality does not hold in general.

(e) If f is a function $E \rightarrow \bar{\mathbb{R}}$ and $\lambda > 0$, then

$$(\lambda f)^*(x') = \lambda f^*\left(\frac{x'}{\lambda}\right) \quad (x' \in E').$$

(f) If f is a function $E \rightarrow \bar{\mathbb{R}}$ and $\alpha \in \bar{\mathbb{R}}$, then

$$(f + \alpha)^* = f^* - \alpha$$

(g) If f is a function $E \rightarrow \bar{\mathbb{R}}$ and $x \in E$, $x' \in E'$, then

$$f_x^*(x') = f^*(x') + x'(x)$$

where the function f_x is defined as $f_x(y) = f(y - x)(y \in E)$.

(h) If f is a function $E \rightarrow \bar{\mathbb{R}}$, then

$$\inf\{f(x) | x \in E\} = -f^*(0).$$

Theorem B.3 *Let f be a function $E \rightarrow \bar{\mathbb{R}}$. Then f^* is a lower semi-continuous convex function on E' (with the norm topology).*

Theorem B.4 *Let f be a function $E \rightarrow \bar{\mathbb{R}}$. For each $x \in E$, $x' \in E'$,*

$$f^*(x') \geq x'(x) - f(x)$$

hence

$$f(x) + f^*(x') \geq x'(x) \tag{B.4}$$

whenever the left-hand side is defined. The (B.4) is called Fenchel's inequality. (Cf. B.1.1)

Theorem B.5 *Let f be a function $E \rightarrow \bar{\mathbb{R}}$, and let x be a point of E where f is finite. Then*

$$x' \in \partial f(x) \Leftrightarrow f^*(x') = x'(x) - f(x).$$

The following theorem is the well-known 1-1 correspondence between closed convex functions f on X and closed convex functions f^* on X' .

Theorem B.6 (Rockafellar [25], Theorem 5) *If the function $f : X \rightarrow \mathbb{R} \cup \{+\infty\}$ is convex and lower semi-continuous, then $f = f^{**}$, i.e.*

$$f(x) = \sup_{x' \in X'} \{x'(x) - f^*(x')\}, \quad \forall x \in X. \quad (\text{B.5})$$

Theorem B.7 (Fenchel duality theorem - Luenberger [22], Th.1, Ch. 7.12)

Let $f : L^\infty \rightarrow \mathbb{R} \cup \{+\infty\}$ be convex, $g : L^\infty \rightarrow \mathbb{R} \cup \{-\infty\}$ be concave and set $C = \{z \in L^\infty : f(z) < +\infty\}$, $D = \{z \in L^\infty : g(z) > -\infty\}$. Suppose that $C \cap D$ contains points in the relative interior of C and D and either the epigraph $[f, C]$ or $[g, D]$ has no empty interior in the product topology of $L^\infty \times \mathbb{R}$. If $\sup_{z \in L^\infty} g(z) - f(z)$ is finite then

$$\sup_{z \in L^\infty} g(z) - f(z) = \min_{\mu \in (L^\infty)^*} f^*(\mu) - g^*(\mu)$$

where f^* (resp. g^*) is the convex (resp. concave) conjugate functional:

$$f^* : (L^\infty)^* \rightarrow \mathbb{R}, \quad f^*(\mu) = \sup_{z \in L^\infty} \{\mu(z) - f(z)\},$$

$$g^* : (L^\infty)^* \rightarrow \mathbb{R}, \quad g^*(\mu) = \inf_{z \in L^\infty} \{\mu(z) - g(z)\},$$

Appendix C

Implementation of the programs

We present in this chapter the simple R programming we used to implement the empirical analysis and obtain the estimations and the graphics.

To get the estimation for \hat{VaR} and \hat{ES} in Figure 7.1 we first generated a variable a from a uniform in $(0,1)$ with 10000 observation. Then we sorted it and called it $alpha$. This to get different estimation of \hat{VaR} and \hat{ES} for $alpha \in (0,1)$. The we computed formulas in Example 3 for these different values of $alpha$.

```
##### VaR and Es in the case that the returns are distributed as a N(0,1) #####
```

```
rm(list=ls())

a<-runif(10000, 0,1)
alpha<-sort(a)
z<-rnorm(10000)
value<--qnorm(alpha, mean=0, sd=1)

es<-array(0, 10000)
for (i in 1:10000)
  es[i]<-(exp(-(value[i])^2/2))/(alpha[i]*sqrt(2*3.14))

plot(alpha, value, type="l", col=4)
points(alpha, es, type="l", col=3)
```

To get estimations for $\hat{\text{VaR}}$ and $\hat{\text{ES}}$ for the returns of the FTSE index, we first obtained, starting from the basis points, the log-returns and we sorted them in increasing order. Then, as done before, we generated and sorted the vector of variables θ , uniformly distributed in $(0,1)$. To get the estimations for different θ we just applied the formulas in Equation (7.11) and Theorem 7.5.

```

%% VaR and ES for the serie of the returns of FTSE %%
rm(list=ls())
p<-scan("c:/ftse.txt")
n_p<-length(p)
rend<-array(0,c(n_p-1,1))
for (i in 2:n_p)
  rend[i-1]<-log(p[i])-log(p[i-1])
theta_rand<-runif(1000, 0,1)
theta<-sort(theta_rand)
m<-length(theta)
n<-length(rend)

r_o<-sort(rend)
v_a_r_1<-array(0, c(m,1))
es_1<-array(0,c(m,1))

for(i in 1:m)
  v_a_r_1[i]<--r_o[floor(n*theta[i])]

for(j in 1:m)
  es_1[j]<--sum(r_o[1:floor(n*theta[j])])/(floor(n*theta[j]))

plot(theta, v_a_r_1, type="l", col=7,xlab="theta", ylab="VaR, Es")
points (theta, es_1, type="l", col=8)

```

And to conclude, we created the weighted portfolio. We used the data we found in the site finance.yahoo.it. Unfortunately, the available data are in decreasing order. So we had to sort them. To get different portfolio for different weights, we created a vector of variables θ -rand distributed as a uniform in $(0,1)$. We

sorted it and called it *theta*. Then we obtained the matrix of the possible returns for different values of *theta* in (0,1). We fixed α , the confidence level, to be 0.01. We sorted all that value in increasing order and computed $\text{VaR}_{0.01}^{\hat{}}$ and $\text{ES}_{0.01}^{\hat{}}$ as done above.

```
rm(list=ls()) p1_d<-scan("c:/intel.txt", dec=",")
n1<-length(p1_d)
p1<-array(0, c(n1,1))
  for( i in 1:n1 )
    p1[n1+1-i]<-p1_d[i]

r1<-array(0, c(n1-1,1))
for (i in 2:n1)
r1[i-1]<-log(p1[i])-log(p1[i-1])
n<-n1-1

p2_d<-scan("c:/coca.txt",dec=",")

p2<-array(0, c(n1,1))
  for( i in 1:n1 )
    p2[n1+1-i]<-p2_d[i]

r2<-array(0, c(n1-1,1))
for (i in 2:n1)
  r2[i-1]<-log(p2[i])-log(p2[i-1])

r1_o<-sort(r1)
r2_o<-sort(r2)

theta_rand<-runif(1000, 0,1)
  theta<-sort(theta_rand)
m<-length(theta)
alpha<-0.01
level<-floor(alpha*n)

c<-array(0, c(n,m))
```

```
for( i in 1:m )
  for( j in 1:n)
    c[j,i]<- r1[j]*theta[i]+ r2[j]*(1-theta[i])

c_ordered<-array(0, c(n,m))
for (i in 1:m)
  c_ordered[,i]<-sort(c[,i])

v_a_r<-array(0, c(m,1))
es<-array(0, c(m,1))

for (j in 1:m)
v_a_r[j]<--c_ordered[level,j]

for(j in 1:m)
  es[j]<--sum(c_ordered[1:level,j])/(level)

plot(theta, es, type="l", col=3, xlab="Weights", ylab="VaR, Es",
ylim=c(0,0.09))
points(theta, v_a_r, type="l", col=4)
```


Bibliography

- [1] Acerbi c., D. Tasche (2002) *On the coherence of Expected Shortfall*. Journal of Banking and Finance **26** n.7
- [2] Artzner P., F. Delbaen, J.M. Eber and D. Heat (1999) *Coherent measure of risk*. Math. Finance **4** (1999) n.3, 203-228
- [3] Artzner P., F. Delbaen, J.M. Eber, D. Heat and H.Ku (2002) *Coherent Multi-period Risk Measurement*. Preprint, E.T.H., Zurich
- [4] Bellini F. and M. Frittelli (2002) *On the existence of the minimax martingale measure*. Mathematical Finance **12** n.1, 1-21
- [5] Bonato M. (2003) *Coerenti misure di rischio: Expected Shortfall come alternativa al VaR*. Tesi di Laurea Triennale, Facoltà di Scienze Statistiche, Univesità degli Studi di Padova.
- [6] von Boxtel E. (2004) *Risk Measures*. Bachelor's Thesis, Universiteit Van Tilburg
- [7] Cheng S., Y. Liu and S. Wang () *Progress in Risk Measurements*. Advanced Modelling and Optimization **6** n.1, 1-20
- [8] De Finetti, B. (1932) *Sul concetto di media*. Giornale degli Economisti **2** 369-396
- [9] Delbaen F. and Schachermayer, W. (1994) *A general version of the fundamental theorem of asset pricing*. Math. Ann. **300** 463-520.
- [10] Delbaen F. (2000) *Coherent Risk Measures*. Lecture notes, Scuola Normale Superiore, Pisa, Italy
- [11] Delbaen F. (2002) *Coherent risk measures on general probability spaces*. Advanced in finance and stochastics, 1-37, Springer, Berlin

-
- [12] El KARoui N. and M.C. Quenez (1991) *Dynamic Programming and Pricing of Contingent Claims In Incomplete Markets*. SIAM J.Contr.Opt. 33(1), 29-66
- [13] Feldman D. and H.G. Tucker (1966) *Estimation of non-unique quantile*. Ann. Math. Stat. **37** 451-457.
- [14] Follmer H. and A.Schied (2002) *Convex measures of risk and trading constraints*. Finance & Stochastics **6** n.4, 429-447
- [15] Follmer H. and A.Schied (2002b) *Robust preferences and convex measures of risk* . Advance in Finance and Stochastics, K.Sandmann nad P.J. Schönbucher eds., Springer-Verlag, 39-56
- [16] Follmer H. and A.Schied (2002c) *Stochastic Finance. An Introduction in Discrete Time*. Walter de Gruyter, Berlin-New York
- [17] Frittelli M., (2000) *Introduction to a theory of value coherent to the no arbitrage principle*. Finance & Stochastics **4** n.3, 275-297.
- [18] Frittelli M., M.R. Gianin (2002) *Puttin order in risk measures*. Journal of Banking and Finance **26** (2002) n.7, 1437-1486
- [19] Frittelli M., M.R. Gianin (2004) *Dynamic Convex Risk Measures*.Risk Mesure in the 21-th Century G.Szego ed., John Wiley & Sons (2004)
- [20] Goovaerts, M.J., Kaas, R., Dhaene, J., (2003). *A unified approach to generate risk measures*. Austin bulletin 33 (2) 173-191
- [21] Kusuoka S. (2001) *On law invariant coherent risk measures*. in: Advanced in Mathematical Economics, Vol.3, Spriner, Tokio, 83-95.
- [22] Luenberger, D.G. (1969) *Optimization by Vector Space Method*. New York: Wiley Sons Inc.
- [23] Rao M.M (2004) *Measure Theory and Integration* Marcel Dekker, Inc. New York
- [24] Rockafellar,R.T (1970) *Convex Analysis*. Princeton, N. J.: Princeton University Press
- [25] Rockafellar,R.T (1974) *Conjugate Duality and Optimization* Regional Conference Series in Applied Mathematics, Vol. 16 SIAM, Philadelphia.
- [26] Rouge R. and N. El Karoui (2000) *Pricing via utility maximization and entropy*. Math. Finance **10** n.2, 259-276

-
- [27] Royden H.L. (1968) *Real Analysis*. Second Edition. MacMillan Publishing Co., Inc. New York
- [28] (1984) Tiel J. van *Convex Analysis. An Introductory Text*. John Wiley & Sons.
- [29] Wang S.S., V.R. Young, H.H. Panjer(1997) *Axiomatic Characterization of Insurance Prices*. *Insurance: Mathematics and Economics*, **21**, 173-183.
- [30] Ymai, Y. and T. Yoshida (2002) *On the validity of Value-at-Risk: Comparative analysis with Expected Shortfall*. *Monetary and economic studies* 20, Bank of Japan.

This thesis has been typed in T_EX.

I would like to thank all the people that helped me in these five years of university. My parents because without their efforts I wouldn't ever have been studying till now. I would like to remember all my friends at the faculty of Statistics: Alberto, Nicola, Stefano, Andrea, Filippo, Betta, Daniela, Thomas, Angela, Gessica, Claudia and all the others. I would like to thank professor Marco Ferrante for his help in these last three years.

Thanks to professor Manuel Moreno for his help to write this thesis.

Y para terminar, me gustaría recordar todas las estupendas personas que he conocido en mi cinco meses en Barcelona: Mario, Bruno, André, Julian, Anina, Catarina y todos los otros. Nunca me olvidaré de vosotros.