VU University Amsterdam
Faculty of Sciences

Jagiellonian University
Faculty of Mathematics and Computer Science

Joint Master of Science Programme

Dorota Kopycka

Student no. 1824821 (VU), WMiI/69/04/05 (UJ)

Dynamic risk measures. Robust representation and examples

Master's thesis
in STOCHASTICS AND FINANCIAL MATHEMATICS

Supervisors:

dr. Sandjai Bhulai
Department of Mathematics, VU

dr hab. Piotr Kobak
Institute of Mathematics, UJ

Amsterdam – Kraków, 2009

Contents

Introduction

Recent years have witnessed a truly quantitative revolution in the world of finance. Financial markets have to deal with a huge variety of traded derivatives and call for advanced mathematical models. Moreover, the same theory can also be applied to institutions, such as banks or insurance companies, that are constantly forced to cope with different kinds of risk. Consequently, risk management has become a very important research area. In view of that this thesis is dedicated to *risk measure theory*. It consists of two parts. The first two chapters are theoretical. Their goal is to familiarize us with the notion of a measure of risk and to present some representation results concerning them. On the other hand, in Chapter 3 and Chapter 4 we construct two mappings wanted to be measures of risk and investigate their properties.

We start with the classic approach. In Chapter 1 we define a *static risk measure* and present some axioms that are welcome to be met. We also introduce the first examples of risk measures. More precisely, the *Value at Risk* and the *Average Value at Risk* are investigated here. Next we relax assumptions and allow certain random variables to be values of risk measures. In other words, we introduce *conditional measures of risk* and extend definitions of the Value at Risk and the Average Value at Risk to that case.

Chapter 2 is a further generalization. In order to deal with multi-period models, we introduce the notion of a *dynamic risk measure*. Again convexity and coherence are defined in a current framework. First we treat time in a continuous manner and consider measures of risk for final payments. Then we are interested in ones in discrete time, but for general stochastic processes. The main result of that part is a *characterization theorem for coherent risk measures*. It turns out that each of them, under some technical assumptions, can be represented by the essential supremum of conditional expectations over some stable set of probability measures. From a theoretical point of view, it is really a meaningful result. However, it does not answer the question what stable set to take. Due to that, we try to define a dynamic risk measure directly. More precisely, we introduce a mapping that seems to us reasonable and then verify whether it is indeed a dynamic risk measure.

In Chapter 3 we introduce the *Recalculated Conditional Average Value at Risk* and the *Iterated Conditional Average Value at Risk*. The idea is based on [HW 04]. In that paper Mary Wirch and Julia Hardy develop dynamic risk measures only for final payments.

We generalize these results, since our mappings assign risk to stochastic processes. We additionally show that they are coherent and satisfy the time consistency and relevance properties.

Similarly, Chapter 4 is devoted to the other mapping. On a basis of [PR 05] and [M 07] we define the *Pflug–Ruszczyński risk measure*. Although it does not fulfill all desirable axioms, it is interesting because of an easy implementation in a Markovian model. The significant fact is that, applying Bayesian decision theory, the risk measure can be extended for an incomplete information case. That plays an important role in practice.

The results presented in the thesis show that theoretical knowledge about dynamic measures of risk is already very wide. Regardless, there is still lack of risk measures that are rational and can be easily used by financial and insurance institutions. I hope that two examples, the Iterated Conditional Average Value at Risk and the Pflug–Ruszczyński risk measure, are only the beginning of the intensive work on that research field.

Chapter 1

Preliminaries

Consider an investor who wants to decide what financial position he should take. Because of the uncertainty of the future, it is important not only to maximize the income but also to compare risks associated with every possible choice. The aim of this chapter is to make this feasible by introducing the definition of risk and by presenting the most popular methods of its measurement.

1.1. Static risk measures

As a starting point we choose a *static* setting. In other words, we suppose that the investor is interested in rating his position only once, at the beginning (at time $t = 0$).

1.1.1. Definitions and characterization

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a given probability space. At once we note that throughout the whole thesis all equalities, inequalities etc. between random variables are understood in a $\mathbb{P}$-almost sure sense.

A random variable $X\colon \Omega \to \mathbb{R}$, which represents a discounted future value of a financial position, is called a *risk*. By $\mathcal{X}$ we denote the set of all risks to be investigated, i.e., $\mathcal{X} \subset L^0(\Omega, \mathcal{F}, \mathbb{P})$. It will be seen later that we often restrict ourselves to the space $L^p(\Omega, \mathcal{F}, \mathbb{P})$ for some $p \in [1, +\infty]$.

Definition 1.1. *A* (static) measure of risk (risk measure) *is a mapping* $\rho\colon \mathcal{X} \to \bar{\mathbb{R}}$ *satisfying the following conditions:*

- monotonicity: $\rho(X_1) \leq \rho(X_2)$ *for all* $X_1, X_2 \in \mathcal{X}$ *with* $X_1 \geq X_2$,
- translation invariance: $\rho(X + c) = \rho(X) - c$ *for* $X \in \mathcal{X}$ *and* $c \in \mathbb{R}$.

Since the above definition is very broad, it is reasonable to narrow down the class of all risk measures. It can certainly be done in many ways, but analists mostly consider *coherent* or *convex* measures of risk. Historically the notion of coherence was introduced by Philippe Artzner et al. in [ADEH 99]. Next Hans Föllmer and Alexander Schied generalized it by defining convex risk measures (see [FS 04]). In this thesis we mainly concentrate on them. The motivation for that may be found in the later part of this section.

Definition 1.2. *A risk measure $\rho\colon \mathcal{X} \to \bar{\mathbb{R}}$ is* convex *if*

$$\rho(\lambda X_1 + (1-\lambda)X_2) \leq \lambda\rho(X_1) + (1-\lambda)\rho(X_2) \text{ for } X_1, X_2 \in \mathcal{X} \text{ and } 0 \leq \lambda \leq 1.$$

Definition 1.3. *A mapping $\rho\colon \mathcal{X} \to \bar{\mathbb{R}}$ is called a* coherent risk measure *if it is a measure of risk satisfying the following statements:*

- subadditivity: $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$ *for* $X_1, X_2 \in \mathcal{X}$,
- positive homogeneity: $\rho(\lambda X) = \lambda\rho(X)$ *if* $\lambda \geq 0$ *and* $X \in \mathcal{X}$.

It is immediately clear that every coherent risk measure is convex. Conversely, a convex measure of risk for which the positive homogeneity condition is fulfilled is coherent as well.

Some simple properties of risk measures are given by the following proposition:

Proposition 1.4. *For a risk measure $\rho\colon \mathcal{X} \to \bar{\mathbb{R}}$ we have*

$$\rho(X + \rho(X)) = 0 \text{ for } X \in \mathcal{X} \text{ when } |\rho(X)| < +\infty.$$

If additionally ρ is positive homogeneous, then it holds that

(1) $\rho(0) = 0$, *i.e., ρ is* normalized,

(2) $\rho(c) = -c$ *for all* $c \in \mathbb{R}$.

Proof. Fix $X \in \mathcal{X}$ such that $|\rho(X)| < +\infty$. By translation invariance one has

$$\rho(X + \rho(X)) = \rho(X) - \rho(X) = 0.$$

If ρ is positive homogeneous, then $\rho(0) = \rho(2 \cdot 0) = 2\rho(0)$, so

$$\rho(0) = 0.$$

Therefore

$$\rho(c) = \rho(0) - c = -c,\ c \in \mathbb{R}.$$

□

Since

$$\rho(X + \rho(X)) = 0,$$

$\rho(X)$ can be seen as an amount of money that has to be added to the risk X to make it acceptable to the investor. In that connection it is obvious that each measure of risk has to be *translation invariant.* Furthermore, a position giving higher income is less risky, so *monotonicity* seems to be rational as well. *Subadditivity* is less intuitive. But consider a firm with two departments. Under the subadditivity condition it suffices to compute $\rho(X)$ and $\rho(Y)$ for risks X and Y associated with positions of every department separately, because $\rho(X+Y)$ is upper bounded by $\rho(X)+\rho(Y)$. Sometimes it is regarded that the value of the risk measure should be proportional to the risk. It leads us to *positive homogeneity.* However, some people find this axiom too strict. All that proves that the research on convex measures of risk is really worthwhile.

We have already introduced the notion of a coherent risk measure. However, it is still not specified how it can be constructed. Due to that, we present a theorem that characterizes the class of coherent risk measures.

Definition 1.5. *A risk measure $\rho\colon L^p(\Omega,\mathcal{F},\mathbb{P}) \to \bar{\mathbb{R}}$, $p \in [1,+\infty)$, satisfies the L^p-Fatou property if for each bounded sequence $(X_n)_{n\in\mathbb{N}} \subset L^p(\Omega,\mathcal{F},\mathbb{P})$ and $X \in L^p(\Omega,\mathcal{F},\mathbb{P})$ such that $X_n \xrightarrow[n\to\infty]{L^p} X$ the following inequality holds:*

$$\rho(X) \leq \liminf_{n\to\infty} \rho\left(X_n\right).$$

Theorem 1.6. *Let p and q be such that $p \in [1,+\infty)$ and $1/p+1/q=1$. Then a mapping $\rho\colon L^p(\Omega,\mathcal{F},\mathbb{P}) \to \bar{\mathbb{R}}$ is a coherent risk measure satisfying the L^p-Fatou property if and only if there exists a convex $L^q(\mathbb{P})$-closed and $L^q(\mathbb{P})$-bounded set $\mathcal{Q}$ of probability measures that are absolutely continuous with respect to $\mathbb{P}$ such that*

$$\rho(X) = \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}(-X),\ X \in L^p(\Omega,\mathcal{F},\mathbb{P}).$$

For the proof the reader is referred to [I 03] (Theorem 1.1).

The above theorem is a useful tool if we want to decide whether a risk measure is coherent or not. A plain application of it can be found below.

(1) Consider a *negative expectation* ρ_{NE} defined by

$$\rho_{\mathrm{NE}}(X) = \mathbb{E}(-X).$$

A set of probability measures associated with ρ_{NE} is just the singleton $\{\mathbb{P}\}$, thus ρ_{NE} is coherent.

(2) Let ρ_{WC} be given by

$$\rho_{\mathrm{WC}}(X) = \operatorname{ess\,sup}(-X).$$

The mapping ρ_{WC} is called a *worst-case risk measure.* It holds that

$$\rho_{\mathrm{WC}}(X) = \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}(-X)$$

for $\mathcal{Q} = \{\mathbb{Q} \ll \mathbb{P}\}$, so Theorem 1.6 guarantees coherence of ρ_{WC}. Note that if ρ is a coherent risk measure with the Fatou property, one has $\rho \leq \rho_{\mathrm{WC}}$.

1.1.2. Value at Risk

We have already defined two coherent risk measures, a negative expectation and a worst-case measure. But, because of their simplicity, they are not very popular. Here and in the next section we introduce risk measures that are more applied in risk management.

A well-known risk measure that is widely used by financial institutions is the *Value at Risk*. Later on we will see that it is not convex and, as a consequence, increasingly criticized. Now we present its definition.

Definition 1.7. *The* Value at Risk *at a level $\alpha \in (0,1]$ of a risk X is given by*

$$\operatorname{VaR}_\alpha X = -q_X^+(1-\alpha),$$

where $q_X^+(1-\alpha)$ is the upper $(1-\alpha)$-quantile of X (see Proposition A.2).

Proposition 1.8. *For $X \in \mathcal{X}$ and $\alpha \in (0,1]$ the following equalities are satisfied:*

$$\operatorname{VaR}_\alpha X = q_{-X}^-(\alpha) = \inf\left\{x \in \mathbb{R} \mid \mathbb{P}(X + x < 0) \leq 1-\alpha\right\}. \tag{1.1}$$

Proof. By (A.1) we obtain

$$\begin{aligned} q_{-X}^-(\alpha) &= \sup\left\{x \mid \mathbb{P}(-X < x) < \alpha\right\} = \sup\left\{x \mid \mathbb{P}(X \leq -x) > 1-\alpha\right\} \\ &= -\inf\left\{x \mid \mathbb{P}(X \leq x) > 1-\alpha\right\} = -q_X^+(1-\alpha) = \operatorname{VaR}_\alpha X. \end{aligned}$$

In a similar way we get that

$$q_{-X}^-(\alpha) = \inf\left\{x \mid \mathbb{P}(X + x < 0) \leq 1-\alpha\right\}.$$

□

Since

$$\operatorname{VaR}_\alpha X = \inf\left\{x \in \mathbb{R} \mid \mathbb{P}(X + x < 0) \leq 1-\alpha\right\},$$

$\operatorname{VaR}_\alpha X$ can be interpreted as an amount of money that needs to be added to make sure that the probability of a loss is less than or equal to $1-\alpha$. Due to that, Value at Risk is usually computed for a safety level α large enough (close to 1).

We have already mentioned that the Value at Risk is not coherent. Now it is an appropriate moment to verify this assertion.

Remark 1.9. *For a fixed $\alpha \in (0,1]$ $\operatorname{VaR}_\alpha$ is monotone, translation invariant and positive homogeneous, but it is <u>not</u> subadditive.*

Proof.

- *Monotonicity*: Take risks $X_1, X_2 \in \mathcal{X}$ such that $X_1 \geq X_2$. Then

  $$A_{X_2} := \left\{x \mid \mathbb{P}\left(X_2 + x < 0\right) \leq 1-\alpha\right\} \subset \left\{x \mid \mathbb{P}\left(X_1 + x < 0\right) \leq 1-\alpha\right\} =: A_{X_1}$$

  and, as a result of Proposition 1.8,

  $$\operatorname{VaR}_\alpha X_2 = \inf A_{X_2} \geq \inf A_{X_1} = \operatorname{VaR}_\alpha X_1.$$

- *Translation invariance*: A risk $X \in \mathcal{X}$ and $c \in \mathbb{R}$ satisfy

$$\begin{aligned}\mathrm{VaR}_\alpha(X + c) &= \inf\{x \mid \mathbb{P}(X + c + x < 0) \le 1 - \alpha\}\\ &= \inf\{x \mid \mathbb{P}(X + x < 0) \le 1 - \alpha\} - c = \mathrm{VaR}_\alpha X - c.\end{aligned}$$

- *Positive homogeneity*: It can be shown in the same manner as translation invariance.

- <u>*No*</u> *subadditivity*: Let X_1, X_2 be independent random variables Bernoulli distributed. More precisely,

$$X_1 = \begin{cases} 200 & \text{with probability } 0.9 \\ -100 & \text{with probability } 0.1, \end{cases} \qquad X_2 = \begin{cases} 200 & \text{with probability } 0.9 \\ -100 & \text{with probability } 0.1. \end{cases}$$

  Then

$$X_1 + X_2 = \begin{cases} 400 & \text{with probability } 0.81 \\ 100 & \text{with probability } 0.18 \\ -200 & \text{with probability } 0.01. \end{cases}$$

  We know that $\mathrm{VaR}_{0.9} X_1 = \mathrm{VaR}_{0.9} X_2 = -200$. On the other hand, $\mathrm{VaR}_{0.9}(X_1 + X_2) = -100$. Hence

$$\mathrm{VaR}_{0.9}(X_1 + X_2) = -100 > -400 = \mathrm{VaR}_{0.9} X_1 + \mathrm{VaR}_{0.9} X_2.$$

□

1.1.3. Average Value at Risk

We have already shown that the Value at Risk is not subadditive. Because of that the *Average Value at Risk* (often called the *Conditional Value at Risk* or the *Expected Shortfall*) was constructed. Due to its coherence, the Average Value at Risk has become very popular and displaced the Value at Risk.

Definition 1.10. *For a risk $X \in \mathcal{X}$ and $\alpha \in (0, 1)$ we define the* Average Value at Risk *by*

$$\mathrm{AVaR}_\alpha X = \frac{1}{1 - \alpha} \int_\alpha^1 \mathrm{VaR}_\gamma X \, d\gamma.$$

As the name suggests, the Average Value at Risk at a level α is simply the conditional expectation of the Value at Risk at γ given that $\gamma \ge \alpha$.

It is possible to extend the above definition for safety levels $\alpha = 0$ and $\alpha = 1$. Namely, applying Lemma A.6, we get that

$$\begin{aligned}
\operatorname{AVaR}_0 X &:= \int_0^1 \operatorname{VaR}_\gamma \,\mathrm{d}\gamma = -\int_0^1 q_X^+(1-\gamma)\,\mathrm{d}\gamma = -\int_0^1 q_X^+(\gamma)\,\mathrm{d}\gamma = \mathbb{E}(-X),\\
\operatorname{AVaR}_1 X &:= \lim_{\alpha\to 1} \operatorname{AVaR}_\alpha X = \lim_{\alpha\to 1} \frac{1}{1-\alpha}\int_\alpha^1 \operatorname{VaR}_\gamma X\,\mathrm{d}\gamma = \lim_{\alpha\to 1}\operatorname{VaR}_\alpha X = \operatorname{VaR}_1 X\\
&= \inf\{x \mid \mathbb{P}(-X > x) = 0\} = \operatorname{ess\,sup}(-X).
\end{aligned}$$

In practice, computing the Average Value at Risk directly from its definition is quite complicated and takes a lot of time. However, it turns out that it can also be obtained as a solution for a certain optimization problem.

Lemma 1.11. *Let $\alpha \in (0,1)$ and q be a $(1-\alpha)$-quantile of X. Then the following equalities hold*

$$\operatorname{AVaR}_\alpha X = \frac{1}{1-\alpha}\,\mathbb{E}(X-q)^- - q = \inf_{x\in\mathbb{R}}\left(\frac{1}{1-\alpha}\,\mathbb{E}(X+x)^- + x\right),\ X \in \mathcal{X}. \tag{1.2}$$

Moreover, the above infimum is attained at $\operatorname{VaR}_\alpha X$*.*

Proof. By Lemma A.6 we have

$$\begin{aligned}
\mathbb{E}(X-q)^- &= \int_0^1 \left(q_X^+(\gamma)-q\right)^-\,\mathrm{d}\gamma = \int_0^{1-\alpha}\left(q - q_X^+(\gamma)\right)\,\mathrm{d}\gamma = (1-\alpha)q - \int_\alpha^1 q_X^+(1-\gamma)\,\mathrm{d}\gamma\\
&= (1-\alpha)q + \int_\alpha^1 \operatorname{VaR}_\gamma X\,\mathrm{d}\gamma = (1-\alpha)\left(\operatorname{AVaR}_\alpha X + q\right),
\end{aligned}$$

so the first equality is satisfied.

To prove the second one we define

$$\Psi(x) = \mathbb{E}(X-x)^-.$$

Because of Lemma A.6, it holds true that

$$\Psi(x) = \int_0^1 \left(q_X^+(\gamma)-x\right)^-\,\mathrm{d}\gamma.$$

Let F be the cumulative distribution function of X. Then, by Fubini's theorem,

$$\begin{aligned}
\int_{-\infty}^x F(t)\,\mathrm{d}t &= \int_{-\infty}^x \mathbb{E}\,\mathbb{1}_{\{X\le t\}}\,\mathrm{d}t = \mathbb{E}\left(\int_{-\infty}^x \mathbb{1}_{\{X\le t\}}\,\mathrm{d}t\right) = \mathbb{E}\left(\int_{\min\{X,x\}}^x \mathrm{d}t\right) = \mathbb{E}(X-x)^-\\
&= \Psi(x).
\end{aligned}$$

It follows that Ψ is increasing and convex.

Note that if Ψ^* denotes the Fenchel–Legendre transform of Ψ restricted to $[0,1]$, Ψ^* is given by

$$\Psi^* \colon [0,1] \ni y \mapsto \sup_{x\in\mathbb{R}} (xy - \Psi(x)) = \int_0^y q_X^+(\gamma)\,\mathrm{d}\gamma.$$

Indeed,

- For $\underline{y=0}$: $\Psi^*(0) = \sup_x (-\Psi(x)) = -\inf_x \Psi(x) = -\lim_{x\searrow-\infty}\Psi(x) = 0$.
- For $\underline{y=1}$: $x - \Psi(x) = x - \int_0^1 \left(q_X^+(\gamma) - x\right)^- \mathrm{d}\gamma = \int_0^1 \min\left\{x, q_X^+(\gamma)\right\} \mathrm{d}\gamma$, thus the function $x \mapsto x - \Psi(x)$ is increasing. As a consequence, $\Psi^*(1) = \sup_x (x - \Psi(x)) = \lim_{x\nearrow+\infty} \int_0^1 \min\left\{x, q_X^+(\gamma)\right\} \mathrm{d}\gamma = \int_0^1 q_X^+(\gamma)\,\mathrm{d}\gamma$.
- For $\underline{0<y<1}$: we know that the function $f\colon x \mapsto xy - \Psi(x)$ is concave. Moreover, $f'_+(x) = y - F(x)$ and $f'_-(x) = y - F(x-)$. Hence x_0 maximizes f if and only if $F(x_0-) \le y \le F(x_0)$. In other words, x_0 is a maximizer of f if it is a y-quantile of X. Then we write $x_0 = q_X(y)$ and $\Psi^*(y) = \sup_x (xy - \Psi(x)) = x_0 y - \int_0^1 \left(q_X^+(\gamma) - x_0\right)^- \mathrm{d}\gamma = x_0 y - \int_0^y \left(x_0 - q_X^+(\gamma)\right) \mathrm{d}\gamma = \int_0^y q_X^+(\gamma)\,\mathrm{d}\gamma$.

Therefore

$$\begin{aligned}
\inf_{x\in\mathbb{R}} \left(\mathbb{E}(X+x)^- + (1-\alpha)x\right) &= -\sup_{x\in\mathbb{R}} ((1-\alpha)x - \Psi(x)) = -\Psi^*(1-\alpha) \\
&= -\int_0^{1-\alpha} q_X^+(\gamma)\,\mathrm{d}\gamma = -\int_\alpha^1 q_X^+(1-\gamma)\,\mathrm{d}\gamma = \int_\alpha^1 \mathrm{VaR}_\gamma X\,\mathrm{d}\gamma \\
&= (1-\alpha)\,\mathrm{AVaR}_\alpha X.
\end{aligned}$$

□

It has been mentioned that the Average Value at Risk is coherent. Now we develop a theorem verifying this statement.

Theorem 1.12. *For $\alpha \in [0,1)$ and $X \in L^\infty(\Omega,\mathcal{F},\mathbb{P})$ it holds that*

$$\mathrm{AVaR}_\alpha X = \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_\mathbb{Q}(-X), \ \textit{where } \mathcal{Q} = \left\{\mathbb{Q} \ll \mathbb{P} \,\middle|\, \frac{d\mathbb{Q}}{d\mathbb{P}} \le \frac{1}{1-\alpha}\right\}.$$

Proof. First notice that for $\alpha = 0$ it holds that $\mathrm{AVaR}_0 X = \mathbb{E}(-X) = \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_\mathbb{Q}(-X)$ for $\mathcal{Q} = \{\mathbb{P}\}$, so the statement is obvious.

Now assume that $\alpha \in (0,1)$. Let ρ be given by

$$\rho(X) = \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_\mathbb{Q}(-X)$$

for $\mathcal{Q}$ as in the formulation of the theorem.

Take $X \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ with $X < 0$. Define a probability measure $\tilde{\mathbb{P}}$ by the Radon–Nikodym derivative

$$\frac{\mathrm{d}\tilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}} = \frac{X}{\mathbb{E}X}.$$

Then $\tilde{\mathbb{P}}$ is equivalent to $\mathbb{P}$ and

$$\rho(X) = \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\mathbb{Q}}(-X) = \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_{\tilde{\mathbb{P}}} \left(-X \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\tilde{\mathbb{P}}}\right) = \frac{\mathbb{E}(-X)}{1-\alpha} \sup_{\varphi\in A} \mathbb{E}_{\tilde{\mathbb{P}}}\, \varphi$$

for $A = \{0 \le \varphi \le 1 \mid \mathbb{E}\,\varphi = 1-\alpha\}$.

Let q denote a $(1-\alpha)$-quantile of X. One has $\mathbb{P}(X < q) \le 1-\alpha < 1 = \mathbb{P}(X < 0)$ and thus $q < 0$. We define

$$\varphi_0 = \mathbb{1}_{\{X<q\}} + \kappa\, \mathbb{1}_{\{X=q\}}, \text{ where } \kappa = \begin{cases} \frac{1-\alpha-\mathbb{P}(X<q)}{\mathbb{P}(X=q)}, & \mathbb{P}(X=q) > 0 \\ 0, & \mathbb{P}(X=q) = 0. \end{cases}$$

Since $\mathbb{P}(X<q) \le 1-\alpha$ and $1-\alpha-\mathbb{P}(X<q) = 1-\alpha-\mathbb{P}(X\le q)+\mathbb{P}(X=q) \le \mathbb{P}(X=q)$, we get that $0 \le \kappa \le 1$ and also $0 \le \varphi_0 \le 1$. It is easy to see that if $\mathbb{P}(X=q)=0$, then $\mathbb{P}(X<q) = \mathbb{P}(X\le q) = 1-\alpha$. Hence

$$\mathbb{E}\,\varphi_0 = \mathbb{P}(X<q) + \kappa\,\mathbb{P}(X=q) = 1-\alpha$$

and $\varphi_0 \in A$.

We know that for every $\varphi \in A$ it holds that

$$(\varphi_0 - \varphi)\left(\frac{\mathrm{d}\tilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}} - \frac{q}{\mathbb{E}X}\right) = \frac{(\varphi_0-\varphi)(X-q)}{\mathbb{E}X} \ge 0.$$

Therefore

$$\mathbb{E}_{\tilde{\mathbb{P}}}\,\varphi_0 - \mathbb{E}_{\tilde{\mathbb{P}}}\,\varphi = \mathbb{E}\left((\varphi_0-\varphi)\frac{\mathrm{d}\tilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}}\right) \ge \mathbb{E}\left((\varphi_0-\varphi)\frac{q}{\mathbb{E}X}\right) = \frac{q}{\mathbb{E}X}\,\mathbb{E}(\varphi_0-\varphi) = 0.$$

We get that

$$\mathbb{E}_{\tilde{\mathbb{P}}}\,\varphi_0 = \sup_{\varphi\in A} \mathbb{E}_{\tilde{\mathbb{P}}}\,\varphi$$

and then

$$\begin{aligned} \rho(X) &= \frac{\mathbb{E}(-X)}{1-\alpha}\mathbb{E}_{\tilde{\mathbb{P}}}\,\varphi_0 = \frac{\mathbb{E}(-X)}{1-\alpha}\mathbb{E}\left(\varphi_0\frac{\mathrm{d}\tilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}}\right) = -\frac{\mathbb{E}(\varphi_0 X)}{1-\alpha} \\ &= -\frac{1}{1-\alpha}\big(-\mathbb{E}(X-q)^- + q\,\mathbb{P}(X<q) + \kappa q\,\mathbb{P}(X=q)\big) = \frac{1}{1-\alpha}\mathbb{E}(X-q)^- - q. \end{aligned}$$

Finally, by (1.2),

$$\rho(X) = \mathrm{AVaR}_\alpha X.$$

Now take an arbitrary $X \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. Then $c := \sup_{\omega\in\Omega} X + 1 < \infty$ and, by the first part of the proof,

$$\begin{aligned}\mathrm{AVaR}_\alpha X &= \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_\gamma X \,\mathrm{d}\gamma = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_\gamma(X-c)\,\mathrm{d}\gamma - c = \mathrm{AVaR}_\alpha(X-c) - c\\ &= \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_\mathbb{Q}(-X+c) - c = \sup_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}_\mathbb{Q}(-X),\end{aligned}$$

so we are done. □

Corollary 1.13. *For $\alpha \in [0,1)$ the Average Value at Risk $\mathrm{AVaR}_\alpha \colon L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \to \bar{\mathbb{R}}$ is coherent and satisfies the Fatou property.*

Proof. The assertion follows from Theorem 1.6 and Theorem 1.12. □

1.2. Conditional risk measures

In this section we aim to extend the definition of a static risk measure. More precisely, we want to allow the investor to rate his positions only once, but at time $t \in [0, T)$, where $T > 0$ is the terminal date.

1.2.1. Definitions

Suppose that we are given a sub-σ-algebra $\mathcal{G}$ in addition to the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. $\mathcal{G}$ can be interpreted as knowledge about the underlying risk at time t. Then, intuitively, a *conditional measure of risk* should be a mapping from $\mathcal{X}$ to the set of $\mathcal{G}$-measurable random variables. Here we limit ourselves to all L^p-integrable risks, i.e., $\mathcal{X} = L^p(\Omega, \mathcal{F}, \mathbb{P})$.

Let $\overline{L^0}(\Omega, \mathcal{G}, \mathbb{P})$ denote the set of all random variables measurable with respect to $(\mathcal{G}, \bar{\mathcal{B}})$, where $\bar{\mathcal{B}}$ is the Borel σ-algebra of $\bar{\mathbb{R}}$.

Definition 1.14. *A mapping $\rho(\,\cdot\,|\,\mathcal{G}) \colon L^p(\Omega, \mathcal{F}, \mathbb{P}) \to \overline{L^0}(\Omega, \mathcal{G}, \mathbb{P})$, $p \in [1, +\infty)$, is a* conditional risk measure *if it is*

- monotone: *$\rho(X_1|\,\mathcal{G}) \le \rho(X_2|\,\mathcal{G})$ for all $X_1, X_2 \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ with $X_1 \ge X_2$,*
- conditional translation invariant: *$\rho(X + Y|\,\mathcal{G}) = \rho(X|\,\mathcal{G}) - Y$ for $X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ and $Y \in L^p(\Omega, \mathcal{G}, \mathbb{P})$.*

Analogously to the previous section we also define a *convex* (respectively *coherent*) *conditional measure of risk*.

Definition 1.15. *A conditional risk measure* $\rho(\,\cdot\,|\,\mathcal{G})\colon L^p(\Omega,\mathcal{F},\mathbb{P}) \to \overline{L^0}(\Omega,\mathcal{G},\mathbb{P})$ *is* convex *if for all* $X_1, X_2 \in L^p(\Omega,\mathcal{F},\mathbb{P})$ *and* $\Lambda \in L^p(\Omega,\mathcal{G},\mathbb{P})$ *such that* $0 \le \Lambda \le 1$ *it holds that*

$$\rho\left(\Lambda X_1 + (1-\Lambda)X_2 |\,\mathcal{G}\right) \le \Lambda\rho\left(X_1|\,\mathcal{G}\right) + (1-\Lambda)\rho\left(X_2|\,\mathcal{G}\right).$$

Definition 1.16. *A conditional risk measure* $\rho(\,\cdot\,|\,\mathcal{G})\colon L^p(\Omega,\mathcal{F},\mathbb{P}) \to \overline{L^0}(\Omega,\mathcal{G},\mathbb{P})$ *is called* coherent *if it satisfies the following properties:*

- subadditivity: $\rho\left(X_1 + X_2|\,\mathcal{G}\right) \le \rho\left(X_1|\,\mathcal{G}\right) + \rho\left(X_2|\,\mathcal{G}\right)$ *for* $X_1, X_2 \in L^p(\Omega,\mathcal{F},\mathbb{P})$,
- conditional positive homogeneity: $\rho(\Lambda X|\,\mathcal{G}) = \Lambda\rho(X|\,\mathcal{G})$ *if* $\Lambda \in L^p_+(\Omega,\mathcal{G},\mathbb{P})$ *and* $X \in L^p(\Omega,\mathcal{F},\mathbb{P})$.

Recall that we investigated static risk measures under the assumption that the risk measurement takes place only at $t = 0$. Due to the fact that there is no additional information about risk available at this time, we set $\mathcal{G} = \{\emptyset, \Omega\}$. Then, for every $X \in \mathcal{X}$, $\rho(X)$ is $\mathcal{G}$-measurable as a constant and thus $\rho = \rho(\,\cdot\,|\,\mathcal{G})$. It proves that the static risk measure is just a trivial case of the conditional one.

The simplest example of a conditional coherent risk measure is given by

$$\rho(X|\,\mathcal{G}) = \mathbb{E}\left(-X|\,\mathcal{G}\right) \text{ for } X \in L^p(\Omega,\mathcal{F},\mathbb{P}).$$

$\rho(\,\cdot\,|\,\mathcal{G})$ is called a *negative conditional expectation.*

1.2.2. Conditional Value at Risk

Now we present a more complicated example of a conditional risk measure, the *Conditional Value at Risk.* As we will see, it is a natural extension of the Value at Risk.

Let $p \in [1,+\infty)$ and $X \in L^p(\Omega,\mathcal{F},\mathbb{P})$. Theorem A.8 guarantees the existence of a probability kernel $P^{(X,\mathcal{G})}\colon \Omega \times \mathcal{B} \to [0,1]$ from $(\Omega,\mathcal{G})$ to $(\mathbb{R},\mathcal{B})$ such that

$$P^{(X,\mathcal{G})}(\,\cdot\,,B) = \mathbb{P}(X \in B|\,\mathcal{G})(\,\cdot\,) \quad \mathbb{P}\text{-a.s.}$$

for $B \in \mathcal{B}$. Knowing that we can define the *Conditional Value at Risk.*

Definition 1.17. *The* Conditional Value at Risk *of* $X \in L^p(\Omega,\mathcal{F},\mathbb{P})$ *at a level* $\alpha \in (0,1)$ *is given by*

$$\mathrm{VaR}_\alpha(X|\,\mathcal{G})(\omega) = \mathrm{VaR}_\alpha\left(P^{(X,\mathcal{G})}(\omega,\cdot\,)\right).$$

By Proposition 1.8 and law invariance we get

$$\begin{aligned}\mathrm{VaR}_\alpha(X|\,\mathcal{G})(\omega) &= \mathrm{VaR}_\alpha\left(P^{(X,\mathcal{G})}(\omega,\cdot\,)\right) = \mathrm{VaR}_\alpha\left(\mathbb{P}(X \in \cdot\,|\,\mathcal{G})(\omega)\right) \\ &= \inf\left\{x \in \mathbb{R}\,|\,\mathbb{P}(X + x < 0\,|\,\mathcal{G})(\omega) \le 1-\alpha\right\}.\end{aligned} \tag{1.3}$$

It is not immediately clear that $\mathrm{VaR}_\alpha(X|\,\mathcal{G})$ is $\mathcal{G}$-measurable. But we have the following theorem:

Theorem 1.18. *For every $X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ and $\alpha \in (0,1)$ the function* $\mathrm{VaR}_\alpha(X|\,\mathcal{G})\colon \Omega \to \mathbb{R}$ *is $\mathcal{G}$-measurable.*

Moreover, $\mathrm{VaR}_\alpha(\,\cdot\,|\,\mathcal{G})$ *is monotone, conditional translation invariant and conditional positive homogeneous.*

Proof. To show that $\mathrm{VaR}_\alpha(X|\,\mathcal{G})$ is a $\mathcal{G}$-measurable function it suffices to prove that for each $x \in \mathbb{R}$ the following holds

$$(\mathrm{VaR}_\alpha(X|\,\mathcal{G}))^{-1}\left((-\infty, x]\right) \in \mathcal{G}.$$

We have

$$\begin{aligned}
(\mathrm{VaR}_\alpha(X|\,\mathcal{G}))^{-1}\left((-\infty, x]\right) &= \{\omega \mid \ \mathrm{VaR}_\alpha(X|\,\mathcal{G})(\omega) \le x\} \\
&= \{\omega \mid \ \inf\{y \mid \mathbb{P}(X + y < 0\,|\,\mathcal{G})(\omega) \le 1-\alpha\} \le x\} \\
&= \{\omega \mid \mathbb{P}(X + x < 0\,|\,\mathcal{G})(\omega) \le 1-\alpha\} \\
&= \left\{\omega \mid P^{(X,\mathcal{G})}\left(\omega, (-\infty, -x)\right) \le 1-\alpha\right\} \\
&= \left(P^{(X,\mathcal{G})}\left(\,\cdot\,, (-\infty, -x)\right)\right)^{-1}\left([0, 1-\alpha]\right) \in \mathcal{G},
\end{aligned}$$

since $(-\infty, -x) \in \mathcal{B}$ and, as a consequence, $P^{(X,\mathcal{G})}\left(\,\cdot\,, (-\infty, -x)\right)$ is $\mathcal{G}$-measurable.

Now we move on to properties of the Conditional Value at Risk.

- *Monotonicity*: Take risks $X_1, X_2 \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ such that $X_1 \ge X_2$. Then
$$\{x \mid \mathbb{P}\left(X_2 + x < 0\,|\,\mathcal{G}\right)(\omega) \le 1-\alpha\} \subset \{x \mid \mathbb{P}\left(X_1 + x < 0\,|\,\mathcal{G}\right)(\omega) \le 1-\alpha\}$$
and consequently
$$\mathrm{VaR}_\alpha\left(X_2|\,\mathcal{G}\right)(\omega) \ge \mathrm{VaR}_\alpha\left(X_1|\,\mathcal{G}\right)(\omega) \text{ for every } \omega \in \Omega.$$

Conditional translation invariance and conditional positive homogeneity can be proved in a similar way. Here we concentrate on homogeneity.

- *Conditional positive homogeneity*: Let $X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ and $\Lambda \in L^p_+(\Omega, \mathcal{G}, \mathbb{P})$. Applying Proposition 2.13 from [YZ 99] we get that
$$\begin{aligned}
\mathrm{VaR}_\alpha(\Lambda X|\,\mathcal{G})(\omega) &= \inf\{x \mid \mathbb{P}(\Lambda X + x < 0\,|\,\mathcal{G})(\omega) \le 1-\alpha\} \\
&= \inf\{x \mid \mathbb{P}\left(\Lambda(\omega) X + x < 0\,|\,\mathcal{G}\right)(\omega) \le 1-\alpha\}.
\end{aligned}$$
First suppose that $\omega \in \Omega$ is such that $\Lambda(\omega) = 0$. Then
$$\mathrm{VaR}_\alpha(\Lambda X|\,\mathcal{G})(\omega) = 0 = \Lambda(\omega)\,\mathrm{VaR}_\alpha(X|\,\mathcal{G})(\omega).$$
Now assume that $\Lambda(\omega) > 0$. We have
$$\begin{aligned}
\mathrm{VaR}_\alpha(\Lambda X|\,\mathcal{G})(\omega) &= \Lambda(\omega)\inf\{x \mid \mathbb{P}\left(X + x < 0\,|\,\mathcal{G}\right)(\omega) \le 1-\alpha\} \\
&= \Lambda(\omega)\,\mathrm{VaR}_\alpha(X|\,\mathcal{G})(\omega).
\end{aligned}$$

□

Because for the trivial sub-σ-algebra $\mathcal{G} = \{\emptyset, \Omega\}$ $\mathrm{VaR}_\alpha(\,\cdot\,|\,\mathcal{G})$ coincides with VaR_α, the Conditional Value at Risk is not subadditive.

1.2.3. Conditional Average Value at Risk

Like in the previous section, we define the *Conditional Average Value at Risk* by analogy to the Average Value at Risk (see Lemma 1.11).

We use the following notation: $L^0_{\mathcal{G}} = L^0(\Omega, \mathcal{G}, \mathbb{P})$.

Definition 1.19. *For $X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$, $p \in [1, +\infty)$, and $\alpha \in (0,1)$ we define a mapping*

$$\mathrm{AVaR}_\alpha(X|\,\mathcal{G})\colon \Omega \ni \omega \mapsto \inf_{Y \in L^0_{\mathcal{G}}} \mathbb{E}\left(\frac{1}{1-\alpha}(X+Y)^- + Y \,\Big|\, \mathcal{G}\right)(\omega),$$

which is called the Conditional Average Value at Risk.

Again we have a theorem concerning measurability of the risk measure defined above.

Theorem 1.20. *For $X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ and $\alpha \in (0,1)$ we know that $\mathrm{AVaR}_\alpha(X|\,\mathcal{G})\colon \Omega \to \bar{\mathbb{R}}$ is $\mathcal{G}$-measurable.*

The infimum in Definition 1.19 is attained at $\mathrm{VaR}_\alpha(X|\,\mathcal{G})$, i.e.,

$$\mathrm{AVaR}_\alpha(X|\,\mathcal{G}) = \mathrm{VaR}_\alpha(X|\,\mathcal{G}) + \frac{1}{1-\alpha}\mathbb{E}\left((X + \mathrm{VaR}_\alpha(X|\,\mathcal{G}))^- \,|\, \mathcal{G}\right). \tag{1.4}$$

In addition, $\mathrm{AVaR}_\alpha(\,\cdot\,|\,\mathcal{G})$ is coherent.

Proof. Due to Proposition 2.13 in [YZ 99], we know that

$$\begin{aligned}\mathrm{AVaR}_\alpha(X|\,\mathcal{G})(\omega) &= \inf_{Y \in L^0_{\mathcal{G}}} \mathbb{E}\left(\frac{(X+Y)^-}{1-\alpha} + Y \,\Big|\, \mathcal{G}\right)(\omega)\\ &= \inf_{Y \in L^0_{\mathcal{G}}} \mathbb{E}\left(\frac{(X+Y(\omega))^-}{1-\alpha} + Y(\omega) \,\Big|\, \mathcal{G}\right)(\omega).\end{aligned}$$

For $Y \in L^0_{\mathcal{G}}$ the value of $\mathbb{E}\left((X+Y(\omega))^- /(1-\alpha) + Y(\omega)\,|\,\mathcal{G}\right)(\omega)$ depends on Y only by $Y(\omega)$. Hence, by Lemma 1.11,

$$\mathrm{AVaR}_\alpha(X|\,\mathcal{G})(\omega) = \mathrm{VaR}_\alpha(X|\,\mathcal{G})(\omega) + \frac{1}{1-\alpha}\mathbb{E}\left((X + \mathrm{VaR}_\alpha(X|\,\mathcal{G}))^- \,|\, \mathcal{G}\right)(\omega).$$

$\mathrm{VaR}_\alpha(X|\,\mathcal{G})$ and $\mathbb{E}\left((X + \mathrm{VaR}_\alpha(X|\,\mathcal{G}))^- \,|\, \mathcal{G}\right)$ are $\mathcal{G}$-measurable, hence $\mathrm{AVaR}_\alpha(X|\,\mathcal{G})$ is $\mathcal{G}$-measurable as well.

Coherence of $\mathrm{AVaR}_\alpha(\,\cdot\,|\,\mathcal{G})$ can be easily proved by applying Proposition 2.13 from [YZ 99] again. We skip it. ☐

Chapter 2

Dynamic risk measures

Risk measurement is aimed at dealing with the uncertainty of the future and should prevent us from potential losses. However, as time goes by, the world is changing. In that connection there is a need to extend the notion of the conditional risk measure to a *dynamic* setting.

Recall our investor. Suppose that the last of his financial positions will have expired by $T > 0$. Let $\mathcal{T}$ stand for the set of all time instants up to T. In other words, $\mathcal{T} = \{0, 1, \ldots, T\}$ or $\mathcal{T} = [0, T]$, depending on whether we treat time in a discrete manner or not. By $\mathcal{T}_{-}$ we denote the set $\mathcal{T} \setminus T$, i.e., $\mathcal{T}_{-} = \{0, 1, \ldots, T-1\}$ or $\mathcal{T}_{-} = [0, T)$, respectively. The investor would like to assign his positions repeatedly, for every $t \in \mathcal{T}_{-}$, to make his measurements update constantly. Hence we have to construct a process of risk measures that is adapted to accessible information.

First we discuss dynamic risk measures for *final payments*. Next, having already some knowledge about them, we will move on to ones for *general processes*.

2.1. Continuous time dynamic risk measures for final payments

In this section we deal with the case of continuous time, so $\mathcal{T} = [0, T]$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space endowed with a filtration $(\mathcal{F}_t)_{t \in \mathcal{T}}$ such that $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{F}$. Certainly, $\mathcal{F}_t$, $t \in \mathcal{T}$, stands for a situation in the market at time t. As before, we want to rate positions we are interested in, but now they are no longer random variables, but stochastic processes. Here we restrict ourselves to ones with only final payments, i.e., we consider processes $(X_t)_{t \in \mathcal{T}}$ with $X_t = 0$ for $t \neq T$. Then $(X_t)_{t \in \mathcal{T}}$ can simply be identified with the random variable X_T. As a consequence, from a practical point of view, nothing changes and the set of positions $\mathcal{X}$ is a subset of $L^0(\Omega, \mathcal{F}, \mathbb{P})$. Now we are ready to introduce a *dynamic measure of risk for final payments*.

2.1.1. Definitions

Definition 2.1. *A mapping* $\rho\colon \Omega \times \mathcal{T}_- \times \mathcal{X} \ni (\omega, t, X) \mapsto \rho(\omega, t, X) = \rho_t(X)(\omega) \in \bar{\mathbb{R}}$ *is called a* dynamic risk measure for final payments *if the following conditions are met:*

- *the process* $(\rho_t(X))_{t\in\mathcal{T}_-}$ *is* $(\mathcal{F}_t)_{t\in\mathcal{T}_-}$*-adapted,*
- ρ *is* monotone: $\rho_t(X_1) \leq \rho_t(X_2)$, $t \in \mathcal{T}_-$, *if* $X_1, X_2 \in \mathcal{X}$ *and* $X_1 \geq X_2$,
- ρ *is* dynamic translation invariant: *it holds that* $\rho_t(X+Y) = \rho_t(X) - Y$ *for* $t \in \mathcal{T}_-$, $X, Y \in \mathcal{X}$ *such that* Y *is* $\mathcal{F}_t$*-measurable.*

The above is strikingly similar to Definition 1.14. The same holds true for the interpretation, so we do not repeat it. Instead we introduce convexity and coherence in the current framework. As we will see, they are already familiar too.

Definition 2.2. *A dynamic risk measure for final payments* $\rho\colon \Omega \times \mathcal{T}_- \times \mathcal{X} \to \bar{\mathbb{R}}$ *is called* convex *if for* $t \in \mathcal{T}_-$ *and an* $\mathcal{F}_t$*-measurable random variable* $\Lambda \in \mathcal{X}$ *such that* $0 \leq \Lambda \leq 1$ *we have*

$$\rho_t(\Lambda X_1 + (1-\Lambda)X_2) \leq \Lambda\rho_t(X_1) + (1-\Lambda)\rho_t(X_2),\ X_1, X_2 \in \mathcal{X}.$$

Definition 2.3. *A dynamic risk measure for final payments* $\rho\colon \Omega\times\mathcal{T}_-\times\mathcal{X} \to \bar{\mathbb{R}}$ *is* coherent *if it satisfies the following properties:*

- subadditivity*:* $\rho_t(X_1 + X_2) \leq \rho_t(X_1) + \rho_t(X_2)$ *if* $t \in \mathcal{T}_-$ *and* $X_1, X_2 \in \mathcal{X}$,
- dynamic positive homogeneity*:* $\rho_t(\Lambda X) = \Lambda\rho_t(X)$ *for* $t \in \mathcal{T}_-$, $X \in \mathcal{X}$ *and* $\mathcal{F}_t$*-measurable* $\Lambda \in \mathcal{X}$ *with* $\Lambda \geq 0$.

2.1.2. Groundwork

We aim to characterize the class of coherent dynamic risk measures in a similar way as in Theorem 1.6. But we start with developing a representation theorem for convex dynamic measures of risk. We introduce the technical property that is necessary for the formulation.

Definition 2.4. *Let* $p \in [1, +\infty)$. *We say that a dynamic risk measure for final payments* $\rho\colon \Omega \times \mathcal{T}_- \times L^p(\Omega, \mathcal{F}, \mathbb{P}) \to \bar{\mathbb{R}}$ *satisfies the* L^p-Fatou property *if for a bounded sequence* $(X_n)_{n\in\mathbb{N}} \subset L^p(\Omega, \mathcal{F}, \mathbb{P})$ *and* $X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ *such that* $X_n \xrightarrow[n\to\infty]{L^p} X$ *it holds true that*

$$\rho_t(X) \leq \liminf_{n\to\infty} \rho_t(X_n),\ t \in \mathcal{T}_-.$$

Suppose that $\rho\colon \Omega \times \mathcal{T}_- \times \mathcal{X} \to \bar{\mathbb{R}}$ is a dynamic risk measure. Fix $t \in \mathcal{T}_-$. Then

$$\mathcal{A}_t^\rho = \{X \in \mathcal{X} \mid \rho_t(X) \leq 0\}$$

is called an *acceptance set*. The name is reasonable, since if $\rho_t(X) \leq 0$ for a certain position X, then we want to take it.

Theorem 2.5. *Let p and q be such that $p \in [1, +\infty)$ and $1/p + 1/q = 1$. If $\rho\colon \Omega \times \mathcal{T}_- \times L^p(\Omega, \mathcal{F}, \mathbb{R}) \to \bar{\mathbb{R}}$ is a convex dynamic risk measure for final payments, then equivalent are:*

(1) *There exists a mapping $\alpha\colon \Omega \times \mathcal{T}_- \times \mathcal{Q}_q \to (-\infty, +\infty]$ such that*

$$\rho_t(X) = \operatorname*{ess\,sup}_{\mathbb{Q} \in \mathcal{Q}_q} \left(\mathbb{E}_{\mathbb{Q}}(-X \mid \mathcal{F}_t) - \alpha_t(\mathbb{Q}) \right), \; X \in L^p(\Omega, \mathcal{F}, \mathbb{P}),$$

where $\mathcal{Q}_q = \{\mathbb{Q} \ll \mathbb{P} \mid d\mathbb{Q}/d\mathbb{P} \in L^q(\Omega, \mathcal{F}, \mathbb{P})\}$.

More precisely, α is given by

$$\alpha_t(\mathbb{Q}) = \operatorname*{ess\,sup}_{X \in \mathcal{A}_t^\rho} \mathbb{E}_{\mathbb{Q}}(-X \mid \mathcal{F}_t).$$

(2) *For $t \in \mathcal{T}_-$ the acceptance set $\mathcal{A}_t^\rho = \{X \in L^p(\Omega, \mathcal{F}, \mathbb{P}) \mid \rho_t(X) \leq 0\}$ is $\|\cdot\|_p$-closed.*

(3) *The risk measure ρ satisfies the L^p-Fatou property.*

Proof.
(1) $\Rightarrow$ (3): It is obvious.

(3) $\Rightarrow$ (2): For any sequence $(X_n)_{n\in\mathbb{N}} \subset \mathcal{A}_t^\rho$ with $X_n \xrightarrow{L^p} X$ there exists $N \in \mathbb{N}$ such that $(X_n)_{n \geq N}$ is bounded. Since the Fatou property holds, we get that

$$\rho_t(X) \leq \liminf_{n\to\infty} \rho_t(X_n) \leq 0,$$

so $X \in \mathcal{A}_t^\rho$. In that case $\mathcal{A}_t^\rho$ is $\|\cdot\|_p$-closed.

(2) $\Rightarrow$ (1): Define α as in the formulation of the theorem and set

$$\varphi_t(X) = \operatorname*{ess\,sup}_{\mathbb{Q} \in \mathcal{Q}_q} \left(\mathbb{E}_{\mathbb{Q}}(-X \mid \mathcal{F}_t) - \alpha_t(\mathbb{Q}) \right), \; \mathcal{Q}_q = \left\{ \mathbb{Q} \ll \mathbb{P} \,\middle|\, \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \in L^q(\Omega, \mathcal{F}, \mathbb{P}) \right\}. \tag{2.1}$$

Since

$$\alpha_t(\mathbb{Q}) = \operatorname*{ess\,sup}_{X \in \mathcal{A}_t^\rho} \mathbb{E}_{\mathbb{Q}}(-X \mid \mathcal{F}_t) = \operatorname*{ess\,sup}_{X \in L^p_{\mathcal{F}}} \left(\mathbb{E}_{\mathbb{Q}}(-X \mid \mathcal{F}_t) - \rho_t(X) \right),$$

we immediately get that

$$\varphi_t(X) \leq \rho_t(X), \; X \in L^p(\Omega, \mathcal{F}, \mathbb{P}).$$

To prove the opposite inequality it suffices to show that if $Z \in L^p(\Omega, \mathcal{F}_t, \mathbb{P})$ with $Z \geq \varphi_t(X)$, then it also holds that $Z \geq \rho_t(X)$. For the indirect proof suppose that there exist $X \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ and $Z \in L^p(\Omega, \mathcal{F}_t, \mathbb{P})$ such that $Z \geq \varphi_t(X)$, but $Z < \rho_t(X)$ with positive probability. The latter is equivalent to $X + Z \notin \mathcal{A}_t^\rho$, because $\rho_t(X + Z) = \rho_t(X) - Z$. We know that $\mathcal{A}_t^\rho$ is $\|\cdot\|_p$-closed, $\{X+Z\}$ is compact and both are convex. Hence, by the separation theorem for convex sets (see Theorem A.10), there exists a linear continuous functional $l\colon L^p(\Omega,\mathcal{F},\mathbb{P}) \to \mathbb{R}$ such that

$$l(X+Z) < \inf_{Y\in\mathcal{A}_t^\rho} l(Y). \tag{2.2}$$

Observe that

$$l(Y) \geq 0 \text{ if } Y \geq 0. \tag{2.3}$$

Indeed, for $\lambda \geq 0$ and $Y^* \geq 0$, $\lambda Y^* \in \mathcal{A}_t^\rho$. Therefore $l(X+Z) < \inf_{Y\in\mathcal{A}_t^\rho} l(Y) \leq l\left(\lambda Y^*\right) = \lambda\, l\left(Y^*\right)$ and thus $-\infty < l(X+Z) \leq \lim_{\lambda\nearrow+\infty} \lambda\, l\left(Y^*\right)$. Finally $l\left(Y^*\right) \geq 0$.

The Riesz representation theorem (Theorem A.11) guarantees the existence of a function $g \in L^q(\Omega,\mathcal{F},\mathbb{P})$ with

$$l(X) = \mathbb{E}(Xg). \tag{2.4}$$

Because of (2.3), $g \geq 0$. Furthermore, due to (2.2), $l \neq 0$. The latter is equivalent to $\mathbb{P}(g=0) < 1$, so $\mathbb{P}(g>0) > 0$. Therefore we can define a probability measure $\mathbb{Q}$ by its Radon–Nikodym derivative with respect to $\mathbb{P}$ as follows:

$$\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \frac{g}{\mathbb{E}g}.$$

Obviously, $\mathbb{Q} \in \mathcal{Q}_q$. Using (2.4) and (2.2) we get

$$\mathbb{E}_\mathbb{Q}\left(-(X+Z)\right) = -\frac{l(X+Z)}{\mathbb{E}g} > -\inf_{Y\in\mathcal{A}_t^\rho} \frac{l(Y)}{\mathbb{E}g} = \sup_{Y\in\mathcal{A}_t^\rho} \frac{l(-Y)}{\mathbb{E}g} = \sup_{Y\in\mathcal{A}_t^\rho} \mathbb{E}_\mathbb{Q}(-Y). \tag{2.5}$$

Now we want to show that

$$\sup_{Y\in\mathcal{A}_t^\rho} \mathbb{E}_\mathbb{Q}(-Y) = \mathbb{E}_\mathbb{Q}\left(\alpha_t(\mathbb{Q})\right). \tag{2.6}$$

Since $\mathbb{E}_\mathbb{Q}(-Y) = \mathbb{E}_\mathbb{Q}\left(\mathbb{E}_\mathbb{Q}(-Y|\mathcal{F}_t)\right) \leq \mathbb{E}_\mathbb{Q}\left(\alpha_t(\mathbb{Q})\right)$, $Y \in \mathcal{A}_t^\rho$, it holds that

$$\sup_{Y\in\mathcal{A}_t^\rho} \mathbb{E}_\mathbb{Q}(-Y) \leq \mathbb{E}_\mathbb{Q}\left(\alpha_t(\mathbb{Q})\right).$$

Fix $Y_1, Y_2 \in \mathcal{A}_t^\rho$ and set $Y_3 = \mathbb{1}_B\, Y_1 + \mathbb{1}_{B^c}\, Y_2$, where $B = \{\mathbb{E}_\mathbb{Q}(-Y_1|\,\mathcal{F}_t) \geq \mathbb{E}_\mathbb{Q}(-Y_2|\,\mathcal{F}_t)\}$. Then $Y_3 \in \mathcal{A}_t^\rho$ and, because $B \in \mathcal{F}_t$,

$$\mathbb{E}_\mathbb{Q}(-Y_3|\,\mathcal{F}_t) = \mathbb{1}_B\; \mathbb{E}_\mathbb{Q}(-Y_1|\,\mathcal{F}_t) + \mathbb{1}_{B^c}\; \mathbb{E}_\mathbb{Q}(-Y_2|\,\mathcal{F}_t) \geq \mathbb{E}_\mathbb{Q}(-Y_1|\,\mathcal{F}_t) \vee \mathbb{E}_\mathbb{Q}(-Y_2|\,\mathcal{F}_t).$$

Therefore the family $\mathcal{F}_\mathbb{Q} = \{\mathbb{E}_\mathbb{Q}(-Y|\,\mathcal{F}_t) \mid Y \in \mathcal{A}_t^\rho\}$ is directed upwards and, due to Theorem A.12, there exists a sequence $(Y_n)_{n\in\mathbb{N}}$ such that

$$\mathbb{E}_\mathbb{Q}(-Y_n|\,\mathcal{F}_t) \nearrow \operatorname*{ess\,sup}_{Y\in\mathcal{A}_t^\rho} \mathbb{E}_\mathbb{Q}(-Y|\,\mathcal{F}_t) = \alpha_t(\mathbb{Q}).$$

Hence, by the monotone convergence theorem,

$$\begin{aligned}\sup_{Y\in\mathcal{A}_t^\rho}\mathbb{E}_{\mathbb{Q}}(-Y) &\geq \lim_{n\to\infty}\mathbb{E}_{\mathbb{Q}}(-Y_n) = \lim_{n\to\infty}\mathbb{E}_{\mathbb{Q}}\left(\mathbb{E}_{\mathbb{Q}}(-Y_n|\,\mathcal{F}_t)\right) = \mathbb{E}_{\mathbb{Q}}\left(\lim_{n\to\infty}\mathbb{E}_{\mathbb{Q}}(-Y_n|\,\mathcal{F}_t)\right)\\ &= \mathbb{E}_{\mathbb{Q}}\left(\alpha_t(\mathbb{Q})\right),\end{aligned}$$

so (2.6) is satisfied.

Finally, applying consecutively (2.5), (2.6) and (2.1) we get that

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}\left(\varphi_t(X)\right) \leq \mathbb{E}_{\mathbb{Q}}Z < \mathbb{E}_{\mathbb{Q}}(-X) - \sup_{Y\in\mathcal{A}_t^\rho}\mathbb{E}_{\mathbb{Q}}(-Y) &= \mathbb{E}_{\mathbb{Q}}\left(-X-\alpha_t(\mathbb{Q})\right)\\ = \mathbb{E}_{\mathbb{Q}}\left(\mathbb{E}_{\mathbb{Q}}(-X|\,\mathcal{F}_t)-\alpha_t(\mathbb{Q})\right) &\leq \mathbb{E}_{\mathbb{Q}}\left(\varphi_t(X)\right),\end{aligned}$$

which yields a contradiction, thus

$$\varphi_t(X) = \rho_t(X),\ X\in L^p(\Omega,\mathcal{F},\mathbb{P}).$$

□

Now, by Theorem 2.5, we quickly obtain the next result.

Theorem 2.6. *Let p and q be such that $p\in[1,+\infty)$ and $1/p+1/q=1$. Suppose that $\rho\colon\Omega\times\mathcal{T}_-\times L^p(\Omega,\mathcal{F},\mathbb{R})\to\bar{\mathbb{R}}$ is a coherent dynamic risk measure for final payments. Then the following are equivalent:*

(1) *For every $t\in\mathcal{T}_-$ there exists a convex $L^q(\mathbb{P})$-closed set $\mathcal{Q}_q^t$ such that*

$$\rho_t(X) = \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_q^t}\mathbb{E}_{\mathbb{Q}}(-X|\,\mathcal{F}_t),\ X\in L^p(\Omega,\mathcal{F},\mathbb{P}).$$

More precisely, the set $\mathcal{Q}_q^t$ is of the form

$$\mathcal{Q}_q^t = \left\{\mathbb{Q}\ll\mathbb{P}\ \middle|\ \frac{d\mathbb{Q}}{d\mathbb{P}}\in L^q(\Omega,\mathcal{F},\mathbb{P}),\ \alpha_t(\mathbb{Q})=0\right\},$$

where α_t is the Legendre–Fenchel transform of ρ_t given by

$$\alpha_t(\mathbb{Q}) = \operatorname*{ess\,sup}_{X\in\mathcal{A}_t^\rho}\mathbb{E}_{\mathbb{Q}}(-X|\,\mathcal{F}_t).$$

(2) *For $t\in\mathcal{T}_-$ the acceptance set $\mathcal{A}_t^\rho = \{X\in L^p(\Omega,\mathcal{F},\mathbb{P})\mid \rho_t(X)\leq 0\}$ is $\|\cdot\|_p$-closed.*

(3) *ρ satisfies the L^p-Fatou property.*

Proof. The fact that assertion (1) implies assertion (3) is obvious. Moreover, from Theorem 2.5 it follows that the implication (3) ⇒ (2) is true as well.

(2) ⇒ (1): Fix $t \in \mathcal{T}_-$ and define $\mathcal{Q}_q^t$ by

$$\mathcal{Q}_q^t = \left\{ \mathbb{Q} \ll \mathbb{P} \,\Big|\, \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \in L^q(\Omega, \mathcal{F}, \mathbb{P}),\ \alpha_t(\mathbb{Q}) = 0 \right\}.$$

First we verify the convexity of $\mathcal{Q}_q^t$. Take $\mathbb{Q}_1, \mathbb{Q}_2 \in \mathcal{Q}_q^t$ and $\alpha \in [0,1]$. For a probability measure $\mathbb{Q}$, given by $\mathbb{Q} = \lambda\mathbb{Q}_1 + (1-\alpha)\mathbb{Q}_2$, and $X \in \mathcal{A}_t^\rho$ we define a set $A_\mathbb{Q}^X = \{\mathbb{E}_\mathbb{Q}(-X|\,\mathcal{F}_t) > 0\}$. Then $A_\mathbb{Q}^X \in \mathcal{F}_t$ and $\int_{A_\mathbb{Q}^X} \mathbb{E}_\mathbb{Q}(-X|\,\mathcal{F}_t)\,\mathrm{d}\mathbb{Q} \geq 0$. On the other hand,

$$\begin{aligned}\int_{A_\mathbb{Q}^X} \mathbb{E}_\mathbb{Q}(-X|\,\mathcal{F}_t)\,\mathrm{d}\mathbb{Q} &= \int_{A_\mathbb{Q}^X} -X\,\mathrm{d}\mathbb{Q} = \lambda \int_{A_\mathbb{Q}^X} -X\,\mathrm{d}\mathbb{Q}_1 + (1-\lambda)\int_{A_\mathbb{Q}^X} -X\,\mathrm{d}\mathbb{Q}_2 \\ &= \lambda \int_{A_\mathbb{Q}^X} \mathbb{E}_{\mathbb{Q}_1}(-X|\,\mathcal{F}_t)\,\mathrm{d}\mathbb{Q}_1 + (1-\lambda)\int_{A_\mathbb{Q}^X} \mathbb{E}_{\mathbb{Q}_2}(-X|\,\mathcal{F}_t)\,\mathrm{d}\mathbb{Q}_2 \\ &\leq \lambda\alpha_t(\mathbb{Q}_1) + (1-\lambda)\alpha_t(\mathbb{Q}_2) = 0.\end{aligned}$$

Hence $\mathbb{Q}\left(A_\mathbb{Q}^X\right) = 0$. It follows that $0 \leq \alpha_t(\mathbb{Q}) = \operatorname{ess\,sup}_{X\in\mathcal{A}_t^\rho} \mathbb{E}_\mathbb{Q}(-X|\,\mathcal{F}_t) \leq 0$, so $\mathbb{Q} \in \mathcal{Q}_q^t$. Therefore $\mathcal{Q}_q^t$ is indeed convex.

In a similar way we check whether $\mathcal{Q}_q^t$ is $L^q(\mathbb{P})$-closed. Fix a sequence $(\mathbb{Q}_n)_{n\in\mathbb{N}} \subset \mathcal{Q}_q^t$ and $\mathcal{Q}$ such that $\mathrm{d}\mathbb{Q}_n/\mathrm{d}\mathbb{P} \to \mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}$ in L^q as $n \to \infty$. Then we also have the weak convergence in L^q and, in particular,

$$\begin{aligned}0 &\leq \int_{A_\mathbb{Q}^X} \mathbb{E}_\mathbb{Q}(-X|\,\mathcal{F}_t)\,\mathrm{d}\mathbb{Q} = \int_{A_\mathbb{Q}^X} -X\,\mathrm{d}\mathbb{Q} = \lim_{n\to\infty} \int_{A_\mathbb{Q}^X} -X\,\mathrm{d}\mathbb{Q}_n = \lim_{n\to\infty}\int_{A_\mathbb{Q}^X} \mathbb{E}_{\mathbb{Q}_n}(-X|\,\mathcal{F}_t)\,\mathrm{d}\mathbb{Q}_n \\ &\leq \lim_{n\to\infty} \alpha_t(\mathbb{Q}_n) = 0,\ X \in \mathcal{A}_t^\rho,\ A_\mathbb{Q}^X = \{\mathbb{E}_\mathbb{Q}(-X|\,\mathcal{F}_t) > 0\}.\end{aligned}$$

Again $\alpha_t(\mathbb{Q}) = 0$ and $\mathbb{Q} \in \mathcal{Q}_q^t$.

Furthermore, note that $\alpha_t(\mathbb{Q}^*) \in \{0, +\infty\}$, $\mathbb{Q}^* \in \mathcal{Q}_q = \{\mathbb{Q} \ll \mathbb{P}\,|\, \mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P} \in L^q(\Omega,\mathcal{F},\mathbb{P})\}$. It is obvious that $\alpha_t(\mathbb{Q}^*) \geq 0$. Define $A = \{\alpha_t(\mathbb{Q}^*) > 0\}$. Then $A \in \mathcal{F}_t$. Thus, for $\lambda > 0$ and $X \in \mathcal{A}_t^\rho$, $\lambda \mathbb{1}_A X \in \mathcal{A}_t^\rho$ as well and

$$\begin{aligned}\alpha_t(\mathbb{Q}^*) &= \operatorname*{ess\,sup}_{X\in\mathcal{A}_t^\rho} \mathbb{E}_{\mathbb{Q}^*}(-X|\,\mathcal{F}_t) \geq \operatorname*{ess\,sup}_{X\in\mathcal{A}_t^\rho} \mathbb{E}_{\mathbb{Q}^*}(-\lambda\mathbb{1}_A X|\,\mathcal{F}_t) = \lambda\,\mathbb{1}_A \operatorname*{ess\,sup}_{X\in\mathcal{A}_t^\rho} \mathbb{E}_{\mathbb{Q}^*}(-X|\,\mathcal{F}_t) \\ &= \lambda\,\mathbb{1}_A\,\alpha_t(\mathbb{Q}^*).\end{aligned}$$

Letting $\lambda \nearrow +\infty$ we obtain

$$\alpha_t(\mathbb{Q}^*) = \begin{cases} +\infty & \text{on } A \\ 0 & \text{on } A^c. \end{cases}$$

Therefore

$$\operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_q^t} \mathbb{E}_{\mathbb{Q}}(-X|\,\mathcal{F}_t) = \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_q^t} \left(\mathbb{E}_{\mathbb{Q}}(-X|\,\mathcal{F}_t) - \alpha_t(\mathbb{Q})\right) = \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_q} \left(\mathbb{E}_{\mathbb{Q}}(-X|\,\mathcal{F}_t) - \alpha_t(\mathbb{Q})\right).$$

We apply Theorem 2.5, what finishes the proof. □

2.1.3. Technical properties

Theorem 2.6 gives us a characterization of the class of coherent dynamic risk measures. Unfortunately, the set $\mathcal{Q}_q^t$ depends on time. We want to strengthen assumptions of the theorem in order to obtain the representation result with a time-independent set of probability measures. It is the aim of the next section. Here we introduce some technical properties that will be necessary for the formulation.

Time consistency

It is quite evident that each reasonable dynamic risk measure should satisfy a kind of *time consistency* condition. In the literature one can find different definitions. We use the one from [B 05].

Definition 2.7. *Let $\rho\colon \Omega \times \mathcal{T}_- \times \mathcal{X} \to \bar{\mathbb{R}}$, be a dynamic risk measure for final payments. Then ρ is called* time-consistent *if for all stopping times σ, τ with $\sigma \leq \tau$ and $X_1, X_2 \in \mathcal{X}$ the following implication holds:*

$$\rho_\tau(X_1) \leq \rho_\tau(X_2) \;\Rightarrow\; \rho_\sigma(X_1) \leq \rho_\sigma(X_2).$$

The above condition gives us information about riskiness of positions. More precisely, if one position is less risky than others in the future, then it is also less risky at every earlier moment, e.g., today.

Relevance

A risk measure should be sensitive about every loss possibility. In other words, the value of the risk measure of the position that generates loss with positive probability has to be positive with positive probability as well. Formally, we introduce the *relevance* property:

Definition 2.8. *We call the mapping $\rho\colon \Omega \times \mathcal{T}_- \times \mathcal{X} \to \bar{\mathbb{R}}$* relevant *if for $A \in \mathcal{F}$ with $\mathbb{P}(A) > 0$ and $t \in \mathcal{T}_-$ we have*

$$\mathbb{P}\left(\rho_t(-\mathbb{1}_A) > 0\right) > 0.$$

Stability of the set of probability measures

Let $\mathcal{Q}$ be a set of probability measures absolutely continuous with respect to $\mathbb{P}$. By $\mathcal{Q}^e$ we denote the set of all probability measures in $\mathcal{Q}$ that are equivalent to $\mathbb{P}$. Then we define a *density process* of $\mathbb{Q} \in \mathcal{Q}$ with respect to $\mathbb{P}$ by

$$Z_t^{\mathbb{Q}} = \mathbb{E}\left(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} \,\middle|\, \mathcal{F}_t\right),\ t \in \mathcal{T}.$$

It is clear that $\left(Z_t^{\mathbb{Q}}\right)_{t\in\mathcal{T}}$ is a $\mathbb{P}$-martingale. Moreover, $Z_T^{\mathbb{Q}} = \mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}$. Now we are ready to set the definition of *stability* forth.

Definition 2.9. *The set $\mathcal{Q}$ of probability measures absolutely continuous with respect to $\mathbb{P}$ is called* stable (under pasting) *if for $\mathbb{Q}_1, \mathbb{Q}_2 \in \mathcal{Q}^e$ with density processes $\left(Z_t^{\mathbb{Q}_1}\right)_{t\in\mathcal{T}}, \left(Z_t^{\mathbb{Q}_2}\right)_{t\in\mathcal{T}}$ and a stopping time $\tau \in \mathcal{T}$ the process $(Z_t^\tau)_{t\in\mathcal{T}}$, given by*

$$Z_t^\tau = \begin{cases} Z_t^{\mathbb{Q}_1}, & t \le \tau \\ Z_\tau^{\mathbb{Q}_1} \frac{Z_t^{\mathbb{Q}_2}}{Z_\tau^{\mathbb{Q}_2}}, & t > \tau, \end{cases}$$

defines a probability measure $\mathbb{Q}^\tau$ that is an element of $\mathcal{Q}$.

2.1.4. Characterization theorem

Here we present the already announced theorem:

Theorem 2.10. *Let p and q be such that $p \in [1, +\infty)$ and $1/p + 1/q = 1$. Let $\rho\colon \Omega \times \mathcal{T}_- \times L^p(\Omega, \mathcal{F}, \mathbb{P}) \to \bar{\mathbb{R}}$ be a time-consistent relevant coherent dynamic risk measure for final payments. Then equivalent are:*

(1) *There exists a stable convex $L^q(\mathbb{P})$-closed set $\mathcal{Q}_q^* \subset \{\mathbb{Q} \sim \mathbb{P} \mid d\mathbb{Q}/d\mathbb{P} \in L^q(\Omega, \mathcal{F}, \mathbb{P})\}$ such that*

$$\rho_t(X) = \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_q^*} \mathbb{E}_{\mathbb{Q}}(-X \mid \mathcal{F}_t),\ X \in L^p(\Omega, \mathcal{F}, \mathbb{P}),\ t \in \mathcal{T}_-.$$

(2) *For $t \in \mathcal{T}_-$ the acceptance set $\mathcal{A}_t^\rho = \{X \in L^p(\Omega, \mathcal{F}, \mathbb{P}) \mid \rho_t(X) \le 0\}$ is $\|\cdot\|_p$-closed.*

(3) *ρ has the L^p-Fatou property.*

Proof. The implications (1) $\Rightarrow$ (3) and (3) $\Rightarrow$ (2) can be shown similarly as in the proof of Theorem 2.6.

(2) $\Rightarrow$ (1): Due to Theorem 2.6, we know that

$$\rho_t(X) = \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_q^t} \mathbb{E}_{\mathbb{Q}}(-X \mid \mathcal{F}_t),\ X \in L^p(\Omega, \mathcal{F}, \mathbb{P}),\ t \in \mathcal{T}_-, \tag{2.7}$$

where $\mathcal{Q}_q^t = \{\mathbb{Q} \ll \mathbb{P} \mid \mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P} \in L^q(\Omega, \mathcal{F}, \mathbb{P}),\ \alpha_t(\mathbb{Q}) = 0\}$.

From time consistency of ρ it follows that $\mathcal{A}_t^\rho \subset \mathcal{A}_s^\rho$ for $s \leq t$. Therefore we have

$$\mathcal{Q}_q^s \subset \mathcal{Q}_q^t. \tag{2.8}$$

Indeed, fix any $\mathbb{Q} \in \mathcal{Q}_q^s$ and for $X \in \mathcal{A}_t^\rho$ define $A_X = \{\mathbb{E}_\mathbb{Q}(-X \mid \mathcal{F}_t) > 0\}$. Since $A_X \in \mathcal{F}_t$, $\mathbb{1}_{A_X} X \in \mathcal{A}_t^\rho \subset \mathcal{A}_s^\rho$ as well. Hence $0 \leq \mathbb{E}_\mathbb{Q}\left(\mathbb{1}_{A_X} \mathbb{E}_\mathbb{Q}(-X \mid \mathcal{F}_t) \mid \mathcal{F}_s\right) = \mathbb{E}_\mathbb{Q}\left(-\mathbb{1}_{A_X} X \mid \mathcal{F}_s\right) \leq \operatorname{ess\,sup}_{Y \in \mathcal{A}_s^\rho} \mathbb{E}_\mathbb{Q}(-Y \mid \mathcal{F}_s) = \alpha_s(\mathbb{Q}) = 0$. It follows that $\mathbb{E}_\mathbb{Q}(-X \mid \mathcal{F}_t) \leq 0$. Finally we have $0 \leq \alpha_t(\mathbb{Q}) = \operatorname{ess\,sup}_{X \in \mathcal{A}_t^\rho} \mathbb{E}_\mathbb{Q}(-X \mid \mathcal{F}_t) \leq 0$, so $\mathbb{Q} \in \mathcal{Q}_q^t$.

We define $\mathcal{Q}_q^* = \mathcal{Q}_q^0$. Theorem 2.6 guarantees that $\mathcal{Q}_q^*$ is convex and $L^q(\mathbb{P})$-closed. It remains to show that it is stable.

Take $\mathbb{Q}_1, \mathbb{Q}_2 \in \left(\mathcal{Q}_q^*\right)^e$. Let $\left(Z_t^{\mathbb{Q}_1}\right)_{t \in \mathcal{T}}, \left(Z_t^{\mathbb{Q}_2}\right)_{t \in \mathcal{T}}$ denote density processes associated with $\mathbb{Q}_1$ and $\mathbb{Q}_2$, respectively. Assume that

$$Z_T^\tau = Z_\tau^{\mathbb{Q}_1} \frac{Z_T^{\mathbb{Q}_2}}{Z_\tau^{\mathbb{Q}^2}}$$

for a stopping time τ is not an element of $\mathcal{Z}^{\mathcal{Q}_q^*} = \left\{Z_T^\mathbb{Q} \mid \mathbb{Q} \in \mathcal{Q}_q^*\right\}$. Now we proceed similarly as in the proof of Theorem 2.5. By the separation theorem for convex sets (Theorem A.10), there exists a linear continuous functional $l \colon L^q(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ with

$$l\left(Z_T^\tau\right) < \inf_{\mathbb{Q} \in \mathcal{Q}_q^*} l\left(Z_T^\mathbb{Q}\right).$$

Furthermore, due to Theorem A.11, there is a function $f \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ such that

$$l(X) = \mathbb{E}(Xf),\ X \in L^q(\Omega, \mathcal{F}, \mathbb{P}).$$

Then, using (2.7),

$$\mathbb{E}\left(-Z_T^\tau f\right) = -l\left(Z_T^\tau\right) > \sup_{\mathbb{Q} \in \mathcal{Q}_q^*} l\left(-Z_T^\mathbb{Q}\right) = \sup_{\mathbb{Q} \in \mathcal{Q}_q^*} \mathbb{E}\left(-Z_T^\mathbb{Q} f\right) = \sup_{\mathbb{Q} \in \mathcal{Q}_q^*} \mathbb{E}_\mathbb{Q}(-f) = \rho_0(f).$$

However, we also have

$$\begin{aligned}
\mathbb{E}\left(-Z_T^\tau f\right) &= -\mathbb{E}\left(Z_\tau^{\mathbb{Q}_1} \frac{Z_T^{\mathbb{Q}_2}}{Z_\tau^{\mathbb{Q}_2}} f\right) = -\mathbb{E}\left(\mathbb{E}\left(\frac{\mathrm{d}\mathbb{Q}_1}{\mathrm{d}\mathbb{P}} \,\middle|\, \mathcal{F}_\tau\right) \frac{\frac{\mathrm{d}\mathbb{Q}_2}{\mathrm{d}\mathbb{P}}}{\mathbb{E}\left(\frac{\mathrm{d}\mathbb{Q}_2}{\mathrm{d}\mathbb{P}} \mid \mathcal{F}_\tau\right)} f\right) \\
&= -\mathbb{E}_{\mathbb{Q}_1}\left(\mathbb{E}\left(\frac{\mathrm{d}\mathbb{Q}_1}{\mathrm{d}\mathbb{P}} \,\middle|\, \mathcal{F}_\tau\right) \frac{\frac{\mathrm{d}\mathbb{Q}_2}{\mathrm{d}\mathbb{Q}_1}}{\mathbb{E}\left(\frac{\mathrm{d}\mathbb{Q}_2}{\mathrm{d}\mathbb{Q}_1} \frac{\mathrm{d}\mathbb{Q}_1}{\mathrm{d}\mathbb{P}} \mid \mathcal{F}_\tau\right)} f\right) = -\mathbb{E}_{\mathbb{Q}_1}\left(\frac{\frac{\mathrm{d}\mathbb{Q}_2}{\mathrm{d}\mathbb{Q}_1}}{\mathbb{E}_{\mathbb{Q}_1}\left(\frac{\mathrm{d}\mathbb{Q}_2}{\mathrm{d}\mathbb{Q}_1} \mid \mathcal{F}_\tau\right)} f\right) \\
&= -\mathbb{E}_{\mathbb{Q}_1}\left(\frac{\mathbb{E}_{\mathbb{Q}_1}\left(\frac{\mathrm{d}\mathbb{Q}_2}{\mathrm{d}\mathbb{Q}_1} f \mid \mathcal{F}_\tau\right)}{\mathbb{E}_{\mathbb{Q}_1}\left(\frac{\mathrm{d}\mathbb{Q}_2}{\mathrm{d}\mathbb{Q}_1} \mid \mathcal{F}_\tau\right)}\right) = \mathbb{E}_{\mathbb{Q}_1}\left(\mathbb{E}_{\mathbb{Q}_2}(-f \mid \mathcal{F}_\tau)\right).
\end{aligned}$$

Furthermore, by (2.8) and (2.7),

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}_1}\left(\mathbb{E}_{\mathbb{Q}_2}(-f\,|\,\mathcal{F}_\tau)\right) &\leq \mathbb{E}_{\mathbb{Q}_1}\left(\operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_q^\tau}\mathbb{E}_{\mathbb{Q}}(-f\,|\,\mathcal{F}_\tau)\right) = \mathbb{E}_{\mathbb{Q}_1}\left(\rho_\tau(f)\right) \leq \sup_{\mathbb{Q}\in\mathcal{Q}_q^*}\mathbb{E}_{\mathbb{Q}}\left(\rho_\tau(f)\right)\\ &= \rho_0\left(-\rho_\tau(f)\right).\end{aligned}$$

We know that $\rho_\tau\left(-\rho_\tau(f)\right) = \rho_\tau(f)$. Thus, by time consistency, we get that $\rho_0\left(-\rho_\tau(f)\right) = \rho_0(f)$ and finally

$$\rho_0(f) < \mathbb{E}\left(-Z_T^\tau f\right) \leq \rho_0(f),$$

which is a contradiction. Therefore $\mathcal{Q}_q^*$ is stable.

Relevance of ρ implies that $\mathbb{Q}$ is equivalent to $\mathbb{P}$ if $\mathbb{Q}\in\mathcal{Q}_q^*$ (for details see Theorem 3.5 in [D 02]).

Let φ be given by

$$\varphi_t(X) = \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_q^*}\mathbb{E}_{\mathbb{Q}}(-X|\,\mathcal{F}_t).$$

We want to show that for $X\in L^p(\Omega,\mathcal{F},\mathbb{P})$ and $t\in\mathcal{T}_-$ it holds true that

$$\rho_t(X) = \varphi_t(X). \tag{2.9}$$

By (2.8) one immediately gets that $\rho_t(X)\geq\varphi_t(X)$, $X\in L^p(\Omega,\mathcal{F},\mathbb{P})$, $t\in\mathcal{T}_-$. Now we want to show that the opposite inequality holds too. For the indirect proof assume the contrary, i.e., suppose that there exist $X\in L^p(\Omega,\mathcal{F},\mathbb{P})$ and $\varepsilon>0$ such that $\mathbb{P}(A_\varepsilon)>0$, where

$$A_\varepsilon = \left\{\rho_t(X)\geq\varepsilon+\varphi_t(X)\right\}.$$

Then

$$\rho_t\left(X+\varphi_t(X)\right) = \rho_t(X)-\varphi_t(X) \geq \varepsilon\,\mathbb{1}_{A_\varepsilon} = \rho_t(-\varepsilon\,\mathbb{1}_{A_\varepsilon})$$

and, because of time consistency and relevance of ρ,

$$\rho_0\left(X+\varphi_t(X)\right) \geq \rho_0(-\varepsilon\,\mathbb{1}_{A_\varepsilon}) = \varepsilon\rho_0(-\mathbb{1}_{A_\varepsilon}) > 0.$$

On the other hand,

$$\begin{aligned}\rho_0\left(X+\varphi_t(X)\right) &= \sup_{\mathbb{Q}\in\mathcal{Q}_q^*}\mathbb{E}_{\mathbb{Q}}\left(-X-\varphi_t(X)\right)\\ &= \sup_{\mathbb{Q}\in\mathcal{Q}_q^*}\mathbb{E}_{\mathbb{Q}}\left(\mathbb{E}_{\mathbb{Q}}(-X|\,\mathcal{F}_t) - \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_q^*}\mathbb{E}_{\mathbb{Q}}(-X|\,\mathcal{F}_t)\right) \leq 0.\end{aligned}$$

Finally we get that $0<\rho_0\left(X+\varphi_t(X)\right)\leq 0$, which yields a contradiction. So we have shown that equality (2.9) holds. □

We finish this section with a theorem that gives a characterization of coherent dynamic risk measures. It is an analogue of Theorem 1.6 in a dynamic setting.

Theorem 2.11. *Let p and q be such that $p \in [1, +\infty)$ and $1/p + 1/q = 1$. Then the following statements are equivalent:*

(1) *A dynamic risk measure $\rho \colon \Omega \times \mathcal{T}_{-} \times L^p(\Omega, \mathcal{F}, \mathbb{P}) \to \bar{\mathbb{R}}$ is coherent, time-consistent, relevant and has the L^p-Fatou property.*

(2) *There exists a stable convex $L^q(\mathbb{P})$-closed set $\mathcal{Q}_q^*$ of probability measures that are equivalent to $\mathbb{P}$ and have L^q-integrable Radon–Nikodym derivatives with respect to $\mathbb{P}$ such that*

$$\rho_t(X) = \operatorname*{ess\,sup}_{\mathbb{Q} \in \mathcal{Q}_q^*} \mathbb{E}_{\mathbb{Q}} \left(-X \middle| \mathcal{F}_t \right), \; X \in L^p(\Omega, \mathcal{F}, \mathbb{P}),\; t \in \mathcal{T}_{-}. \tag{2.10}$$

Proof.
(2) $\Rightarrow$ (1): Let ρ be defined by (2.10). It is easy to see that ρ is a relevant coherent dynamic risk measure that satisfies the L^p-Fatou property. Time consistency follows from Theorem 5.1 in [ADEHK 07].

(1) $\Rightarrow$ (2): Assume that ρ is a dynamic risk measure for final payments such that all axioms from the formulation of the theorem are satisfied. Then Theorem 2.10 yields that there exists a stable convex $L^q(\mathbb{P})$-closed set $\mathcal{Q}_q^*$ of probability measures equivalent to $\mathbb{P}$ with L^q-integrable Radon–Nikodym derivatives such that for $t \in \mathcal{T}_{-}$ it holds true that

$$\rho_t(X) = \operatorname*{ess\,sup}_{\mathbb{Q} \in \mathcal{Q}_q^*} \mathbb{E}_{\mathbb{Q}}(-X \mid \mathcal{F}_t),\; X \in L^p(\Omega, \mathcal{F}, \mathbb{P}).$$

That finishes the proof. □

2.2. Discrete-time dynamic risk measures for processes

From now on, time is treated in a discrete manner. We assume that there is a time horizon $T \in \mathbb{N}$, so $\mathcal{T} = \{0, 1, \ldots, T\}$ and $\mathcal{T}_{-} = \{0, 1, \ldots, T-1\}$. We no longer restrict ourselves to assign positions with only final payments. Conversely, we intend to consider general stochastic processes that are adapted. The set of them we denote by $\mathcal{X}$. Moreover, $\mathcal{X}^p \subset \mathcal{X}$ is the set of all processes $X = (X_0, X_1, \ldots, X_T)$ such that $X_t \in L^p(\Omega, \mathcal{F}_t, \mathbb{P})$ for $t \in \mathcal{T}$. Furthermore, we suppose that the interest rate $r > -1$ is known and constant. The last assumption can be easily relaxed. Everything else remains the same as in the previous section.

2.2.1. Definitions and technical properties

It is easy to guess what a dynamic risk measure for processes is. The definition is just a generalization of that from Section 2.1. It is also the case for the notion of convexity (coherence). Nevertheless, we state these definitions here, simply for the sake of completeness.

Definition 2.12. *A mapping $\rho\colon \Omega \times \mathcal{T}_- \times \mathcal{X} \ni (\omega, t, X) \mapsto \rho(\omega, t, X) = \rho_t(X)(\omega) \in \bar{\mathbb{R}}$ is called a* dynamic risk measure *if the following conditions are met:*

- *the process $(\rho_t(X))_{t\in\mathcal{T}_-}$ is $(\mathcal{F}_t)_{t\in\mathcal{T}_-}$-adapted,*
- *ρ is* independent of the past: *for $t \in \mathcal{T}_-$ and $X \in \mathcal{X}$ $\rho_t(X)$ does not depend on $X_0, X_1, \dots, X_{t-1}$,*
- *ρ is* monotone: *$\rho_t\left(X^{(1)}\right) \leq \rho_t\left(X^{(2)}\right)$, $t \in \mathcal{T}_-$, if $X^{(1)}, X^{(2)} \in \mathcal{X}$ and $X^{(1)} \geq X^{(2)}$,*
- *ρ is* dynamic translation invariant: *it holds that*

$$\rho_t(X+Y) = \rho_t(X) - \sum_{n=t}^{T} \frac{Y_n}{(1+r)^{n-t}}$$

  for $t \in \mathcal{T}_-$, $X, Y \in \mathcal{X}$ such that $Y = (0, \dots, 0, Y_t, \dots, Y_T)$ and $\sum_{n=t}^{T} Y_n/(1+r)^{n-t}$ is $\mathcal{F}_t$-measurable.

Definition 2.13. *A dynamic risk measure $\rho\colon \Omega \times \mathcal{T}_- \times \mathcal{X} \to \bar{\mathbb{R}}$ is* convex *if*

$$\rho_t\left(\Lambda X^{(1)} + (1-\Lambda)X^{(2)}\right) \leq \Lambda\rho_t\left(X^{(1)}\right) + (1-\Lambda)\rho_t\left(X^{(2)}\right),$$

for $X^{(1)}, X^{(2)} \in \mathcal{X}$, $t \in \mathcal{T}_-$ and $\Lambda \in L^p(\Omega, \mathcal{F}_t, \mathbb{P})$ with $0 \leq \Lambda \leq 1$.

Definition 2.14. *A dynamic risk measure $\rho\colon \Omega \times \mathcal{T}_- \times \mathcal{X} \to \bar{\mathbb{R}}$ is called* coherent *if it satisfies the following properties:*

- subadditivity: *$\rho_t\left(X^{(1)} + X^{(2)}\right) \leq \rho_t\left(X^{(1)}\right) + \rho_t\left(X^{(2)}\right)$, $X^{(1)}, X^{(2)} \in \mathcal{X}$, $t \in \mathcal{T}_-$,*
- dynamic positive homogeneity: *$\rho_t(\Lambda X) = \Lambda\rho_t(X)$ for $X \in \mathcal{X}$ and $\mathcal{F}_t$-measurable $\Lambda \in \mathcal{X}$ with $\Lambda \geq 0$.*

The only property that needs a comment is the *independence of the past.* For any fixed $t \in \mathcal{T}_-$ and a position X $\rho_t(X)$ is evaluated at time t, so all payments $X_0, X_1, \dots, X_{t-1}$, which have already passed, cannot influence the value of $\rho_t(X)$.

In the current framework we also introduce definitions of the *L^p-Fatou property*, *time consistency* and *relevance*.

Definition 2.15. *Let $\rho\colon \Omega \times \mathcal{T}_- \times \mathcal{X}^p \to \bar{\mathbb{R}}$, $p \in [1, +\infty)$, be a dynamic risk measure. We say that ρ satisfies the* L^p-Fatou property *if for $t \in \mathcal{T}_-$, a sequence $(X^{(n)})_{n\in\mathbb{N}} \subset \mathcal{X}^p$ and $X \in \mathcal{X}^p$ with*

$$\sup_{k\geq t} \mathbb{E}\left|X_k^{(n)}\right| \leq 1,\ n \in \mathbb{N},\ \textit{and}\ \sup_{k\geq t}\left|X_k^{(n)} - X_k\right| \xrightarrow[n\to\infty]{L^p} 0$$

it holds that

$$\rho_m(X) \leq \liminf_{n\to\infty} \rho_m\left(X^{(n)}\right),\ m \in \{t, \dots, T-1\}.$$

Definition 2.16. *A dynamic risk measure $\rho\colon \Omega \times \mathcal{T}_- \times \mathcal{X} \to \bar{\mathbb{R}}$ is* time-consistent *if the following condition is met:*

$$\rho_\sigma\left(X + Y \cdot \mathrm{e}_\tau\right) = \rho_\sigma\left(X + (1+r)^{T-\tau}\, Y \cdot \mathrm{e}_T\right)$$

for stopping times σ, τ such that $\sigma \leq \tau$, $X \in \mathcal{X}^p$ and $Y \in L^p(\Omega, \mathcal{F}_\tau, \mathbb{P})$.

Definition 2.17. *A dynamic risk measure $\rho\colon \Omega \times \mathcal{T}_- \times \mathcal{X} \to \bar{\mathbb{R}}$ is* relevant *if for every set $A \in \mathcal{F}$ with $\mathbb{P}(A) > 0$ we have*

$$\mathbb{P}\left(\rho_t\left(-\mathbb{1}_A \cdot \mathrm{e}_T\right) > 0\right) > 0,\ t \in \mathcal{T}_-.$$

2.2.2. Characterization theorem

This section is dedicated to the main result of the current chapter. We formulate and prove the equivalents of Theorem 2.5 and Theorem 2.11 for risk measures for general stochastic processes.

Theorem 2.18. *Let p and q be such that $p \in [1, +\infty)$ and $1/p + 1/q = 1$. Then $\rho\colon \Omega \times \mathcal{T}_- \times \mathcal{X}^p \to \bar{\mathbb{R}}$ is a time-consistent convex dynamic risk measure that has the L^p-Fatou property if and only if it is of the form*

$$\rho_t(X) = \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_q}\left(\mathbb{E}_\mathbb{Q}\left(-\sum_{n=t}^{T}\frac{X_n}{(1+r)^{n-t}}\,\middle|\,\mathcal{F}_t\right) - \alpha_t(\mathbb{Q})\right), \tag{2.11}$$

where

$$\alpha_t(\mathbb{Q}) = \operatorname*{ess\,sup}_{X\in\mathcal{X}^p}\mathbb{E}_\mathbb{Q}\left(-\sum_{n=t}^{T}\frac{X_n}{(1+r)^{n-t}} - \rho_t(X)\,\middle|\,\mathcal{F}_t\right),$$
$$\mathcal{Q}_q = \left\{\mathbb{Q} \ll \mathbb{P}\,\middle|\,\frac{d\mathbb{Q}}{d\mathbb{P}} \in L^p(\Omega, \mathcal{F}, \mathbb{P})\right\}.$$

Proof. Assume that ρ is given by (2.11). Then, obviously, ρ is a dynamic risk measure. It can be easily verified that the axioms are satisfied as well.

Now suppose that ρ is a dynamic risk measure satisfying the axioms. Define φ by

$$\varphi_t(Y) = (1+r)^{T-t}\rho_t(0,\ldots,0,Y),\ Y \in L^p(\Omega,\mathcal{F},\mathbb{P}).$$

Then φ is a convex dynamic risk measure for final payments that has the L^p-Fatou property. From Theorem 2.5 it follows that

$$\varphi_t(Y) = \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_q} \left(\mathbb{E}_{\mathbb{Q}}(-Y \mid \mathcal{F}_t) - \alpha_t^{\varphi}(\mathbb{Q})\right)$$

for $\alpha_t^{\varphi}(\mathbb{Q}) = \operatorname{ess\,sup}_{Y\in\mathcal{A}_t^{\varphi}} \mathbb{E}_{\mathbb{Q}}(-Y \mid \mathcal{F}_t)$ and $\mathcal{Q}_q = \{\mathbb{Q} \ll \mathbb{P} \mid \mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P} \in L^q(\Omega,\mathcal{F},\mathbb{P})\}$. By independence of the past and time consistency we get

$$\begin{aligned}
\rho_t(X) &= \rho_t\left(0,\ldots,0,X_t,\ldots,X_T\right) = \rho_t\left(0,\ldots,0,\sum_{n=t}^{T}(1+r)^{T-n}X_n\right)\\
&= (1+r)^{t-T}\varphi_t\left(\sum_{n=t}^{T}(1+r)^{T-n}X_n\right)\\
&= \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_q}\left(\mathbb{E}_{\mathbb{Q}}\left(-\sum_{n=t}^{T}\frac{X_n}{(1+r)^{n-t}}\,\middle|\,\mathcal{F}_t\right) - \frac{\alpha_t^{\varphi}(\mathbb{Q})}{(1+r)^{T-t}}\right).
\end{aligned}$$

Moreover,

$$\begin{aligned}
\alpha_t(\mathbb{Q}) &= \operatorname*{ess\,sup}_{X\in\mathcal{X}^p}\left(\mathbb{E}_{\mathbb{Q}}\left(-\sum_{n=t}^{T}\frac{X_n}{(1+r)^{n-t}}\,\middle|\,\mathcal{F}_t\right) - \rho_t(X)\right)\\
&= (1+r)^{t-T}\operatorname*{ess\,sup}_{X\in\mathcal{X}^p}\left(\mathbb{E}_{\mathbb{Q}}\left(-\sum_{n=t}^{T}(1+r)^{T-n}X_n\,\middle|\,\mathcal{F}_t\right) - \varphi_t\left(\sum_{n=t}^{T}(1+r)^{T-n}X_n\right)\right)\\
&= (1+r)^{t-T}\alpha_t^{\varphi}(\mathbb{Q}),
\end{aligned}$$

so we are done. □

Theorem 2.19. *Let p and q be such that $p \in [1,+\infty)$ and $1/p + 1/q = 1$. The following are equivalent:*

(1) *A dynamic risk measure $\rho\colon \Omega\times\mathcal{T}_-\times\mathcal{X}^p \to \bar{\mathbb{R}}$ is coherent, time-consistent, relevant and satisfies the L^p-Fatou property.*

(2) *There exists a stable convex $L^q(\mathbb{P})$-closed set $\mathcal{Q}_q^*$ of probability measures that are equivalent to $\mathbb{P}$ and have L^q-integrable Radon–Nikodym derivatives with respect to $\mathbb{P}$ such that*

$$\rho_t(X) = \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_q^*}\mathbb{E}_{\mathbb{Q}}\left(-\sum_{n=t}^{T}\frac{X_n}{(1+r)^{n-t}}\,\middle|\,\mathcal{F}_t\right),\ X\in\mathcal{X}^p,\ t\in\mathcal{T}_-.$$

Proof. It is an analogue of the previous proof. The only difference is we need to use Theorem 2.10 instead of Theorem 2.5. ☐

At the end of this section we develop a simple remark:

Remark 2.20. *Let $p \in [1, +\infty)$ and $\rho\colon \Omega \times \mathcal{T}_- \times \mathcal{X}^p \to \bar{\mathbb{R}}$ be a time-consistent relevant coherent dynamic risk measure with the L^p-Fatou property. For $X \in \mathcal{X}^p$ we define a process $M = (M_t)_{t\in\mathcal{T}_-}$ by*

$$M_t = \frac{\rho_t(X)}{(1+r)^t} - \sum_{n=1}^{t-1} \frac{X_n}{(1+r)^n}.$$

Then M is a $\mathbb{Q}$-supermartingale for $\mathbb{Q} \in \mathcal{Q}_q^*$.

Proof. We define

$$\varphi_t(X) = (1+r)^{T-t}\rho_t(0, \dots, 0, X),\ X \in L^p(\Omega, \mathcal{F}, \mathbb{P}).$$

Then φ is a time-consistent relevant coherent dynamic risk measure for final payments and has the L^p-Fatou property. Moreover,

$$\rho_t(X) = \rho_t\left(0, \dots, 0, \sum_{n=t}^{T}(1+r)^{T-n}X_n\right) = (1+r)^t\varphi_t\left(\sum_{n=t}^{T}\frac{X_n}{(1+r)^n}\right).$$

Hence

$$M_t = \frac{\rho_t(X)}{(1+r)^t} - \sum_{n=1}^{t-1}\frac{X_n}{(1+r)^n} = \varphi_t\left(\sum_{n=t}^{T}\frac{X_n}{(1+r)^n}\right) - \sum_{n=1}^{t-1}\frac{X_n}{(1+r)^n} = \varphi_t\left(\sum_{n=1}^{T}\frac{X_n}{(1+r)^n}\right).$$

Let Z be given as follows:

$$Z = \sum_{n=1}^{T}\frac{X_n}{(1+r)^n}.$$

Since $\varphi_{t+1}\left(-\varphi_{t+1}(Z)\right) = \varphi_{t+1}(Z)$, we also have $\varphi_t\left(-\varphi_{t+1}(Z)\right) = \varphi_t(Z)$. Then, by Theorem 2.11,

$$0 = \varphi_t\left(-\varphi_{t+1}(Z) + \varphi_t(Z)\right) = \operatorname*{ess\,sup}_{\mathbb{Q}\in\mathcal{Q}_q^*}\mathbb{E}_{\mathbb{Q}}\left(\varphi_{t+1}(Z) - \varphi_t(Z)\middle|\,\mathcal{F}_t\right).$$

For $\mathbb{Q} \in \mathcal{Q}_q^*$ and $t \in \mathcal{T}_-$ it holds that

$$\mathbb{E}_{\mathbb{Q}}\left(M_{t+1} - M_t \,|\, \mathcal{F}_t\right) = \mathbb{E}_{\mathbb{Q}}\left(\varphi_{t+1}(Z) - \varphi_t(Z)\,|\,\mathcal{F}_t\right) \le 0,$$

so $(M_t)_{t\in\mathcal{T}_-}$ is indeed a $\mathbb{Q}$-supermartingale. ☐

Chapter 3

Recalculated and Iterated Conditional Average Value at Risk

In the previous chapter we were concentrated on theoretical aspects of dynamic risk measures. We found that if a measure of risk satisfies some technical properties, then it can be represented as the essential supremum of expectations with respect to probability measures from a certain stable set. However, in practice it is really hard to decide what set we should consider. As a consequence, there is a problem with developing a reasonable risk measure. Here we want to deal with this drawback. We do not want to create a set of probability measures, but rather define a function that, as we will see, is a time-consistent coherent dynamic risk measure.

We still suppose that the time is discrete, i.e., $\mathcal{T} = \{0, 1, \dots, T\}$ for a time horizon $T \in \mathbb{N}$ and the interest rate $r > -1$ is known and constant.

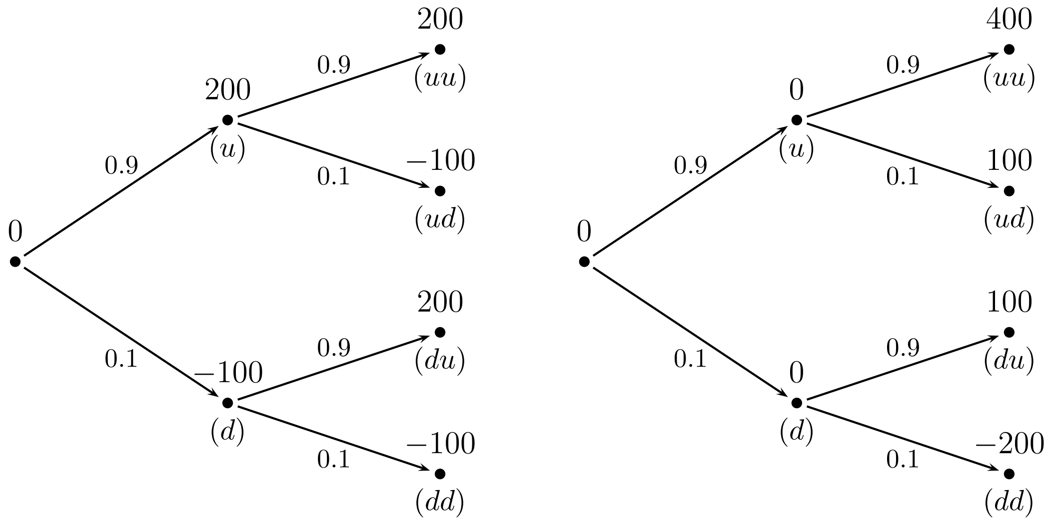
3.1. Basic idea

In this section we define two dynamic risk measures, the *Recalculated Conditional Average Value at Risk* and the *Iterated Conditional Average Value at Risk*. As their names suggest, both of them are based on the Conditional Average Value at Risk. The following idea was introduced in [HW 04]. However, in this paper the Iterated Conditional Tail Expectation was considered. We will see later on that the procedure can be applied to any reasonable conditional measure of risk.

We start with an easy, but illustrative example of a binomial tree with two time steps. For simplicity we assume that there is no interest rate, i.e., r is equal to 0. By X we denote a process of cash flows represented by Figure 3.1a. In that case the natural filtration $(\mathcal{F}_t)_{t\in\{0,1,2\}}$ is given by

$$\mathcal{F}_0 = \{\emptyset, \Omega\}, \quad \mathcal{F}_1 = \sigma\left(\{x\} \times \{u, d\}, \, x \in \{u, d\}\right), \quad \mathcal{F}_2 = \sigma\left(\{(x, y)\}, \, x, y \in \{u, d\}\right).$$

We can simplify the figure by considering the process of only final payments. More precisely, we compute the future net value of cash flows at time 2 and get the tree as in Figure 3.1b. Set the stochastic process presented there by Y.



(a) Process X (b) Process Y

Figure 3.1: Binomial trees

We define the *Recalculated Conditional Average Value at Risk at level* $\alpha \in (0,1)$ by

$$\mathrm{RAVaR}_t^{\alpha} X = \mathrm{AVaR}_{\alpha}\left(Y_2 \middle| \mathcal{F}_t\right),\ t \in \{0,1\}.$$

For $\alpha = 0.9$ we have

$$\mathrm{RAVaR}_0^{0.9} X(\omega) = -70, \quad \mathrm{RAVaR}_1^{0.9} X(\omega) = \begin{cases} -100, & \omega \in \{(uu),(ud)\} \\ 200, & \omega \in \{(du),(dd)\}. \end{cases}$$

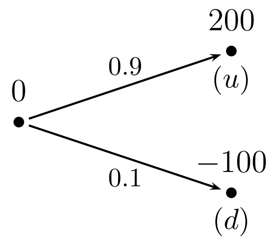
On the other hand, we can proceed in a different way and, also on a basis of the Conditional Average Value at Risk, create another dynamic risk measure. This new measure we will call the *Iterated Conditional Average Value at Risk* and denote by IAVaR. To construct it we use backward induction. First we take

$$\mathrm{IAVaR}_1^{0.9} X = \mathrm{AVaR}_{0.9}\left(X_2 \middle| \mathcal{F}_1\right) - X_1.$$

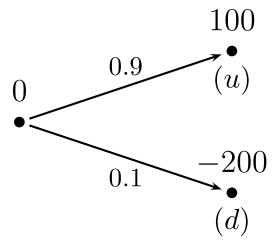
Hence the value of the Iterated Conditional Average Value at Risk at time 1 is equal to the Conditional Average Value at Risk given the sub-σ-algebra $\mathcal{F}_1$ reduced by the income at that time. Due to translation invariance, it makes sense. We have

$$\mathrm{IAVaR}_1^{0.9} X(\omega) = \left\{\begin{matrix} -100, & \omega \in \{(uu),(ud)\} \\ 200, & \omega \in \{(du),(dd)\} \end{matrix}\right\} = \mathrm{RAVaR}_1^{0.9} X(\omega).$$

Furthermore, we want to define $\operatorname{IAVaR}_0^{0.9} X$. Instead of considering the one-period model given by Figure 3.2a as in the case of the Recalculated Conditional Average Value at Risk, we examine the tree from Figure 3.2b.



(a) Initial model



(b) New model

Figure 3.2: One-period model from 0 to 1

The values at nodes (u) and (d) are just equal to $-\operatorname{IAVaR}_1^{0.9} X(\omega)$ for $\omega \in \{(uu),(ud)\}$ and $\omega \in \{(du),(dd)\}$, respectively. Then we have

$$\operatorname{IAVaR}_0^{0.9} X(\omega) = \operatorname{AVaR}_{0.9}\left(-\operatorname{IAVaR}_1^{0.9} X \,|\, \mathcal{F}_0\right)(\omega) = 200 > \operatorname{RAVaR}_0^{0.9} X(\omega).$$

We have introduced two functions that are, as we will see later, dynamic risk measures. Obviously, they do not coincide and, what is more important, they can lead us to different decisions. In our example $\operatorname{RAVaR}_0^{0.9} X = -70 \leq 0$ and we should accept the position X. On the other hand, $\operatorname{IAVaR}_0^{0.9} X = 200 > 0$, so the decision whether to accept X or not depends on our attitude to risk. If we are risk averse, it ought to be rejected.

3.2. Formal definitions

In the previous section we gave an idea of how to define the Recalculated Conditional Average Value at Risk and the Iterated Conditional Average Value at Risk. Now it is an appropriate moment to provide formal definitions. We have already mentioned that the use of the Conditional Average Value at Risk is our choice and each coherent conditional measure of risk can be exploited. Due to that, we present general definitions of a *recalculated risk measure* and an *iterated risk measure*.

As usual, there is a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_{t\in\mathcal{T}}, \mathbb{P})$ given. We additionally assume that $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{F}$. Let $X = (X_t)_{t\in\mathcal{T}} \in \mathcal{X}$ be a stochastic process representing our cash flows. Fix a sequence $(\rho(\,\cdot\,|\,\mathcal{F}_t))_{t\in\mathcal{T}_-}$ of coherent conditional risk measures.

Definition 3.1. *A process $(R_t(X))_{t\in\mathcal{T}_-}$ such that*

$$R_t(X) = \rho\left(\sum_{n=t}^{T} \frac{X_n}{(1+r)^{n-t}} \,\middle|\, \mathcal{F}_t\right), \; t \in \mathcal{T}_-,$$

is called a recalculated risk measure *of X.*

Remark 3.2. *A mapping* $R\colon \Omega \times \mathcal{T}_- \times \mathcal{X} \ni (\omega, t, X) \mapsto R(\omega, t, X) = R_t(X)(\omega) \in \bar{\mathbb{R}}$ *given by*

$$R_t(X) = \rho\left(\sum_{n=t}^{T} \frac{X_n}{(1+r)^{n-t}} \,\middle|\, \mathcal{F}_t\right),\ t \in \mathcal{T}_-,$$

is a coherent dynamic risk measure.

Proof. First of all, note that $(R_t(X))_{t\in\mathcal{T}_-}$ is adapted for any $X \in \mathcal{X}$ and $R_t(X)$ depends only on $X_t, \ldots, X_T$.

If $X^{(1)}, X^{(2)} \in \mathcal{X}$ with $X^{(1)} \geq X^{(2)}$ and $t \in \mathcal{T}_-$, then we have $\sum_{n=t}^{T} X_n^{(1)}/(1+r)^{n-t} \geq \sum_{n=t}^{T} X_n^{(2)}/(1+r)^{n-t}$. By monotonicity of $\rho\,(\,\cdot\,|\,\mathcal{F}_t)$, we get that $R_t\left(X^{(1)}\right) \leq R_t\left(X^{(2)}\right)$.

Similarly, $R_t(X+Y) = R_t(X) - \sum_{n=t}^{T} Y_n/(1+r)^{n-t}$ for $t \in \mathcal{T}_-$ and $X, Y \in \mathcal{X}$ such that $Y = (0, \ldots, 0, Y_t, \ldots, Y_T)$ and $\sum_{n=t}^{T} Y_n/(1+r)^{n-t}$ is $\mathcal{F}_t$-measurable.

Subadditivity and dynamic positive homogeneity of R immediately follow from the fact that $\rho\,(\,\cdot\,|\,\mathcal{F}_t)$ for each $t \in \mathcal{T}_-$ is a coherent conditional measure of risk. $\square$

Definition 3.3. *Let* $(I_t(X))_{t\in\mathcal{T}_-}$ *be given by backward induction as follows*

$$\begin{aligned} I_{T-1}(X) &= \frac{1}{1+r}\,\rho\,(X_T|\,\mathcal{F}_{T-1}) - X_{T-1}, \\ I_t(X) &= \frac{1}{1+r}\,\rho\,(-I_{t+1}(X)|\,\mathcal{F}_t) - X_t,\ t \in \{0, 1, \ldots, T-2\}. \end{aligned} \tag{3.1}$$

Then the above process is called an iterated risk measure *of* X.

Remark 3.4. *A mapping* $I\colon \Omega \times \mathcal{T}_- \times \mathcal{X} \ni (\omega, t, X) \mapsto I(\omega, t, X) = I_t(X)(\omega) \in \bar{\mathbb{R}}$ *that satisfies conditions* (3.1) *for every* $X \in \mathcal{X}$ *is a coherent dynamic risk measure.*

Proof. Independence of the past and adaptedness are obvious. Monotonicity, translation invariance, subadditivity and dynamic positive homogeneity can be shown by backwards induction. To do so, take $X^{(1)}, X^{(2)}, X, Y \in \mathcal{X}$ and a non-negative random variable Λ.

- *Base case:* Assume additionally that $Y = (0, \ldots, 0, Y_{T-1}, Y_T)$, $Y_{T-1} + Y_T/(1+r)$ and Λ are $\mathcal{F}_{T-1}$-measurable. Since Y is adapted, we get that Y_T is $\mathcal{F}_{T-1}$-measurable. By direct calculations it can be verified that the following holds:

$$\begin{aligned} &I_{T-1}\left(X^{(1)}\right) \leq I_{T-1}\left(X^{(2)}\right) \text{ if } X^{(1)} \geq X^{(2)}, \\ &I_{T-1}(X+Y) = I_{T-1}(X) - Y_{T-1} - \frac{Y_T}{1+r}, \\ &I_{T-1}\left(X^{(1)} + X^{(2)}\right) \leq I_{T-1}\left(X^{(1)}\right) + I_{T-1}\left(X^{(2)}\right), \\ &I_{T-1}(\Lambda X) = \Lambda I_{T-1}(X). \end{aligned}$$

- *Inductive step:* Fix $t \in \{0, 1, \ldots, T-2\}$ and suppose that I_{t+1} is a coherent conditional risk measure. Assume that $Y = (0, \ldots, 0, Y_t, \ldots, Y_T)$, $\sum_{n=t}^{T} Y_n/(1+r)^{n-t}$ and Λ are $\mathcal{F}_t$-measurable. Hence $\sum_{n=t+1}^{T} Y_n/(1+r)^{n-t-1}$ is $\mathcal{F}_t$-measurable as well. Therefore

$$
\begin{aligned}
I_t\left(X^{(1)}\right) &= \frac{1}{1+r}\,\rho\left(-I_{t+1}\left(X^{(1)}\right) \mid \mathcal{F}_t\right) - X_t^{(1)} \leq \frac{1}{1+r}\,\rho\left(-I_{t+1}\left(X^{(2)}\right) \mid \mathcal{F}_t\right) - X_t^{(2)} \\
&= I_t\left(X^{(2)}\right) \text{ if } X^{(1)} \geq X^{(2)}, \\
I_t(X+Y) &= \frac{1}{1+r}\,\rho\left(-I_{t+1}(X+Y) \mid \mathcal{F}_t\right) - X_t - Y_t \\
&= \frac{1}{1+r}\,\rho\left(-I_{t+1}\left(X + (0, \ldots, 0, Y_{t+1}, \ldots, Y_T)\right) \mid \mathcal{F}_t\right) - X_t - Y_t \\
&= \frac{1}{1+r}\,\rho\left(-I_{t+1}(X) \mid \mathcal{F}_t\right) - X_t - \sum_{n=t}^{T} \frac{Y_n}{(1+r)^{n-t}} = I_t(X) - \sum_{n=t}^{T} \frac{Y_n}{(1+r)^{n-t}}
\end{aligned}
$$

and

$$
\begin{aligned}
I_t\left(X^{(1)} + X^{(2)}\right) &= \frac{1}{1+r}\,\rho\left(-I_{t+1}\left(X^{(1)} + X^{(2)}\right) \mid \mathcal{F}_t\right) - X_t^{(1)} - X_t^{(2)} \\
&\leq \frac{1}{1+r}\,\rho\left(-I_{t+1}\left(X^{(1)}\right) \mid \mathcal{F}_t\right) - X_t^{(1)} + \rho\left(-I_{t+1}\left(X^{(2)}\right) \mid \mathcal{F}_t\right) - X_t^{(2)} \\
&= I_t\left(X^{(1)}\right) + I_t\left(X^{(2)}\right), \\
I_t(\Lambda X) &= \frac{1}{1+r}\,\rho\left(-I_{t+1}(\Lambda X) \mid \mathcal{F}_t\right) - \Lambda X_t = \frac{\Lambda}{1+r}\,\rho\left(-I_{t+1}(X) \mid \mathcal{F}_t\right) - \Lambda X_t \\
&= \Lambda\, I_t(X).
\end{aligned}
$$

So I_t is also coherent.

□

Definition 3.5. *The* Recalculated Conditional Average Value at Risk at level $\alpha \in (0,1)$ *of X is a recalculated risk measure for a sequence $(\rho(\,\cdot \mid \mathcal{F}_t))_{t \in \mathcal{T}_-}$ such that $\rho(\,\cdot \mid \mathcal{F}_t) = \mathrm{AVaR}_\alpha(\,\cdot \mid \mathcal{F}_t)$, $t \in \mathcal{T}_-$.*

Analogously we define the Iterated Conditional Average Value at Risk.

Definition 3.6. *A dynamic risk measure $(I_t)_{t \in \mathcal{T}_-}$ satisfying* (3.1) *for every $X \in \mathcal{X}$ and $\rho(\,\cdot \mid \mathcal{F}_t) = \mathrm{AVaR}_\alpha(\,\cdot \mid \mathcal{F}_t)$, $t \in \mathcal{T}_-$, is called the* Iterated Conditional Average Value at Risk *at level α.*

3.3. Time consistency and relevance

We have already shown that both the Recalculated Conditional Average Value at Risk and the Iterated Conditional Average Value at Risk are coherent. However, it will be very welcome if they are additionally time-consistent and relevant. In that case, without applying Theorem 2.19 and going through the notion of stability, we have constructed two reasonable dynamic risk measures, which can be used in practice.

We start with considering the Recalculated Conditional Average Value at Risk. Obviously, we have

$$\begin{aligned}\mathrm{RAVaR}_\sigma^\alpha\left(X + (1+r)^{T-\tau}\, Y \cdot \mathrm{e}_T\right) &= \mathrm{AVaR}_\alpha\left(\sum_{n=\sigma}^{T} \frac{X_n}{(1+r)^{n-\sigma}} + \frac{Y}{(1+r)^{\tau-\sigma}} \,\Big|\, \mathcal{F}_\sigma\right)\\ &= \mathrm{RAVaR}_\sigma^\alpha\left(X + Y \cdot \mathrm{e}_\tau\right)\end{aligned}$$

for $\alpha \in (0,1)$, $\sigma \le \tau$ that are stopping times, $X \in \mathcal{X}$ and $Y \in L^0(\Omega, \mathcal{F}_\tau, \mathbb{P})$. Hence the Recalculated Conditional Average Value at Risk is indeed time-consistent.

The case of the Iterated Conditional Average Value at Risk is more complicated. Note that it suffices to show that for each $t \in \mathcal{T}_-$ and a stopping time $\tau \ge t$ it holds that

$$\mathrm{IAVaR}_t^\alpha\left(X + Y \cdot \mathrm{e}_\tau\right) = \mathrm{IAVaR}_t^\alpha\left(X + (1+r)^{T-\tau}\, Y \cdot \mathrm{e}_T\right), \; X \in \mathcal{X},\; Y \in L^0(\Omega, \mathcal{F}_\tau, \mathbb{P}).$$

Again we do backward induction. First assume that $\tau \ge T-1$. Since

$$(1+r)^{T-\tau}\, Y = (1+r) Y\, \mathbb{1}_{\{\tau = T-1\}} + Y\, \mathbb{1}_{\{\tau = T\}}$$

and $Y\,\mathbb{1}_{\{\tau=T-1\}}$ is $\mathcal{F}_{T-1}$-measurable, we have

$$\begin{aligned}&\mathrm{IAVaR}_{T-1}^\alpha\left(X + (1+r)^{T-\tau}\, Y \cdot \mathrm{e}_T\right)\\ &\quad= \frac{1}{1+r}\,\mathrm{AVaR}_\alpha\left(X_T + (1+r)^{T-\tau}\, Y \,|\, \mathcal{F}_{T-1}\right) - X_{T-1}\\ &\quad= \frac{1}{1+r}\,\mathrm{AVaR}_\alpha\left(X_T + Y\,\mathbb{1}_{\{\tau=T\}} \,|\, \mathcal{F}_{T-1}\right) - X_{T-1} - Y\,\mathbb{1}_{\{\tau=T-1\}}\\ &\quad= \mathrm{IAVaR}_{T-1}^\alpha\left(X + Y \cdot \mathrm{e}_\tau\right).\end{aligned}$$

Now suppose that $t \in \{0, 1, \ldots, T-2\}$, $\tau \ge t$ and the following is true

$$\mathrm{IAVaR}_{t+1}^\alpha\left(X + Y \cdot \mathrm{e}_{\tau'}\right) = \mathrm{IAVaR}_{t+1}^\alpha\left(X + (1+r)^{T-\tau'}\, Y \cdot \mathrm{e}_T\right)$$

for every stopping time $\tau' \ge t+1$, $X \in \mathcal{X}$ and $Y \in L^0(\Omega, \mathcal{F}_\tau, \mathbb{P})$. Similarly as before,

$$(1+r)^{T-\tau}\, Y = (1+r)^{T-t}\, Y\,\mathbb{1}_{\{\tau=t\}} + (1+r)^{T-\tau'}\, Y\,\mathbb{1}_{\{\tau \ge t+1\}},$$

where $\tau' = \max\{t+1, \tau\}$. Then τ' is a stopping time such that $\tau' \ge \tau$, so Y is $\mathcal{F}_{\tau'}$-measurable. Moreover, $Y\,\mathbb{1}_{\{\tau=t\}}$ and $Y\,\mathbb{1}_{\{\tau \ge t+1\}}$ are $\mathcal{F}_t$- and $\mathcal{F}_{\tau'}$-measurable, respectively. Therefore, by inductive assumption, one has

$$\begin{aligned}&\mathrm{IAVaR}_{t+1}^\alpha\left(X + (1+r)^{T-\tau}\, Y \cdot \mathrm{e}_T\right)\\ &\quad= \mathrm{IAVaR}_{t+1}^\alpha\left(X + (1+r)^{T-\tau'}\, Y\,\mathbb{1}_{\{\tau \ge t+1\}} \cdot \mathrm{e}_T\right) - (1+r) Y\,\mathbb{1}_{\{\tau=t\}}\\ &\quad= \mathrm{IAVaR}_{t+1}^\alpha\left(X + Y\,\mathbb{1}_{\{\tau \ge t+1\}} \cdot \mathrm{e}_{\tau'}\right) - (1+r) Y\,\mathbb{1}_{\{\tau=t\}}.\end{aligned}$$

Furthermore,

$$\begin{aligned}&\mathrm{IAVaR}_t^\alpha\left(X+(1+r)^{T-\tau}\,Y\cdot \mathrm{e}_T\right)\\&\quad=\frac{1}{1+r}\,\mathrm{AVaR}_\alpha\left(-\,\mathrm{IAVaR}_{t+1}^\alpha\left(X+(1+r)^{T-\tau}\,Y\cdot \mathrm{e}_T\right)\,|\,\mathcal{F}_t\right)-X_t\\&\quad=\frac{1}{1+r}\,\mathrm{AVaR}_\alpha\left(-\,\mathrm{IAVaR}_{t+1}^\alpha\left(X+Y\,\mathbb{1}_{\{\tau\geq t+1\}}\cdot \mathrm{e}_{\tau'}\right)\,|\,\mathcal{F}_t\right)-X_t-Y\,\mathbb{1}_{\{\tau=t\}}\,.\end{aligned}$$

By independence of the past we finally get that

$$\begin{aligned}&\mathrm{IAVaR}_t^\alpha\left(X+(1+r)^{T-\tau}\,Y\cdot \mathrm{e}_T\right)\\&\quad=\frac{1}{1+r}\,\mathrm{AVaR}_\alpha\left(-\,\mathrm{IAVaR}_{t+1}^\alpha\left(X+Y\cdot \mathrm{e}_\tau\right)\,|\,\mathcal{F}_t\right)-X_t-Y\,\mathbb{1}_{\{\tau=t\}}\\&\quad=\mathrm{IAVaR}_t^\alpha\left(X+Y\cdot \mathrm{e}_\tau\right).\end{aligned}$$

Hence also the Iterated Conditional Average Value at Risk is a time-consistent risk measure.

Note that for the Conditional Average Value at Risk given any sub-σ-algebra $\mathcal{G}$ the following statement is satisfied:

$$\begin{matrix}\mathrm{AVaR}_\alpha(X\,|\,\mathcal{G})\geq 0\\ \mathbb{P}\left(\mathrm{AVaR}_\alpha(X\,|\,\mathcal{G})>0\right)>0\end{matrix}\quad\text{for } X\leq 0 \text{ such that } \mathbb{P}(X<0)>0. \tag{3.2}$$

Because $X \leq 0$, it is clear that $\mathrm{VaR}_\alpha(X\,|\,\mathcal{G}) \geq 0$ and $\mathrm{AVaR}_\alpha(X\,|\,\mathcal{G}) \geq 0$. Let A be given by $A = \{\mathrm{VaR}_\alpha(X\,|\,\mathcal{G}) > 0\}$. If $\mathbb{P}(A) > 0$, we are done, because, by (1.4), we have $\mathrm{AVaR}_\alpha(X\,|\,\mathcal{G}) > 0$ on A. Therefore assume that $\mathbb{P}(A) = 0$. Then, by (1.4) as well, we get that $\mathrm{AVaR}_\alpha(X\,|\,\mathcal{G}) = -\mathbb{E}\left(X\,|\,\mathcal{G}\right)/(1-\alpha)$, so $\mathbb{P}\left(\mathrm{AVaR}_\alpha(X\,|\,\mathcal{G}) > 0\right) > 0$.

From (3.2) it immediately follows that both the Recalculated Conditional Average Value at Risk and the Iterated Conditional Average Value at Risk are relevant.

3.4. Artzner game

Consider an example that is due to Philippe Artzner. There are three coins, not necessarily fair. More precisely, the probability of showing heads for each of them is equal to $p \in (0,1)$. There are two possible games. In the first one we get 1, if the third coin shows heads. In the second, we win 1, if no less than two coins show heads. We want to compare the riskiness of these games by using the Iterated Conditional Average Value at Risk. Let $X^{(i)}$, $i = 1, 2$, denote the income process for the i-th game. The streams are presented in Figure 3.3, where an up-move means that a coin shows heads. Again we consider the natural filtration $(\mathcal{F}_t)_{t\in\{0,1,2,3\}}$, i.e.,

$$\begin{aligned}&\mathcal{F}_0=\{\emptyset,\Omega\},\quad \mathcal{F}_1=\sigma\left(\{x\}\times\{u,d\}^2,\,x\in\{u,d\}\right),\\&\mathcal{F}_2=\sigma\left(\{(x,y)\}\times\{u,d\},\,x,y\in\{u,d\}\right),\quad \mathcal{F}_3=\sigma\left(\{(x,y,z)\},\,x,y,z\in\{u,d\}\right).\end{aligned}$$

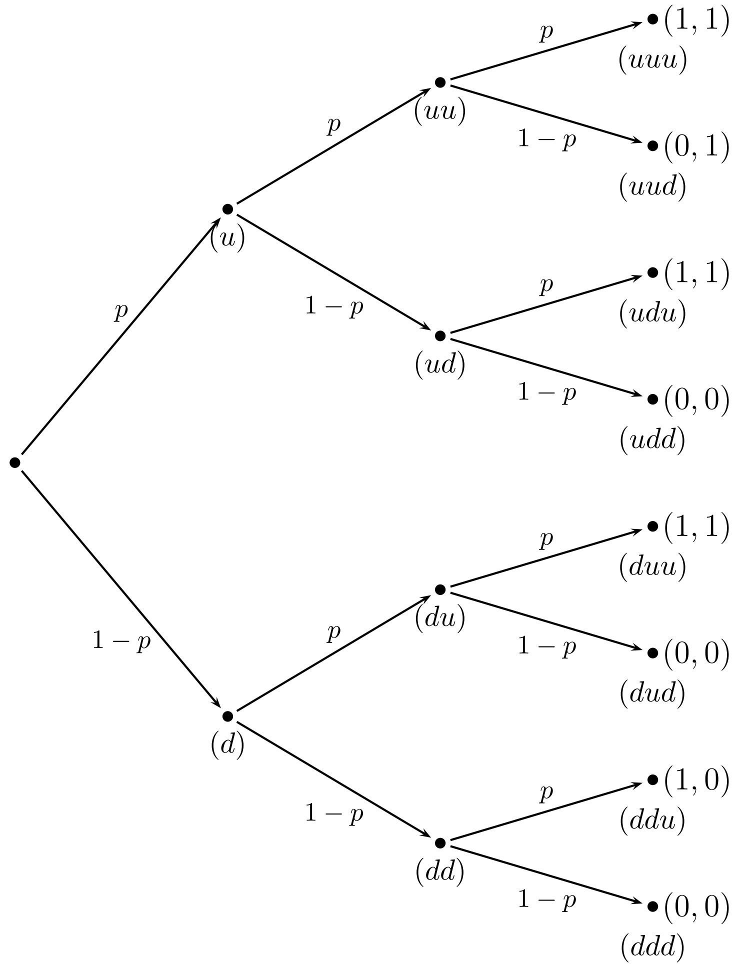


Figure 3.3: Income stream tree for the Artzner game

Note that, due to (1.1) and (1.2), for a random variable X such that

$$X = \begin{cases} a, & \text{with probability } \theta \in (0,1) \\ b, & \text{with probability } 1-\theta, \end{cases} \qquad a > b,$$

it holds that

$$\begin{aligned} \operatorname{VaR}_\alpha X &= \inf\{x \mid \mathbb{P}(X + x < 0) \le 1-\alpha\} = \begin{cases} -b, & \alpha > \theta \\ -a, & \alpha \le \theta, \end{cases} \\ \operatorname{AVaR}_\alpha X &= \operatorname{VaR}_\alpha X + \frac{1}{1-\alpha}\, \mathbb{E}\left(X + \operatorname{VaR}_\alpha X\right)^- \\ &= \begin{cases} -b + \frac{1}{1-\alpha}\, \mathbb{E}(X-b)^- = -b, & \alpha > \theta \\ -a + \frac{1}{1-\alpha}\, \mathbb{E}(X-a)^- = \frac{a(\alpha-\theta)-b(1-\theta)}{1-\alpha}, & \alpha \le \theta. \end{cases} \end{aligned}$$

First we concentrate on the process $X^{(1)}$. It is immediate that

$$\mathrm{IAVaR}_2^{\alpha}\left(X^{(1)}\right) = \frac{1}{1+r}\,\mathrm{AVaR}_{\alpha}\left(X_3^{(1)}\middle|\,\mathcal{F}_2\right) = \begin{cases} 0, & \alpha > p \\ \frac{\alpha-p}{(1+r)(1-\alpha)}, & \alpha \le p. \end{cases}$$

Since $\mathrm{IAVaR}_2^{\alpha}$ is constant, we get that

$$\mathrm{IAVaR}_1^{\alpha}\left(X^{(1)}\right) = \frac{1}{1+r}\,\mathrm{AVaR}_{\alpha}\left(-\,\mathrm{IAVaR}_2^{\alpha}\left(X^{(1)}\right)\mid\mathcal{F}_1\right) = \begin{cases} 0, & \alpha > p \\ \frac{\alpha-p}{(1+r)^2(1-\alpha)}, & \alpha \le p, \end{cases}$$

$$\mathrm{IAVaR}_0^{\alpha}\left(X^{(1)}\right) = \frac{1}{1+r}\,\mathrm{AVaR}_{\alpha}\left(-\,\mathrm{IAVaR}_1^{\alpha}\left(X^{(1)}\right)\right) = \begin{cases} 0, & \alpha > p \\ \frac{\alpha-p}{(1+r)^3(1-\alpha)}, & \alpha \le p. \end{cases}$$

Now we move on to the investigation of $X^{(2)}$. We consider two cases.

- Suppose that $\alpha > p$. Then we have

$$\mathrm{IAVaR}_2^{\alpha}\left(X^{(2)}\right)(\omega) = \frac{1}{1+r}\,\mathrm{AVaR}_{\alpha}\left(X_3^{(2)}\middle|\,\mathcal{F}_2\right)(\omega) = \begin{cases} -\frac{1}{1+r}, & \omega \in \{(u,u)\}\times\{u,d\} \\ 0, & \omega \notin \{(u,u)\}\times\{u,d\}, \end{cases}$$

$$\mathrm{IAVaR}_1^{\alpha}\left(X^{(2)}\right)(\omega) = \frac{1}{1+r}\,\mathrm{AVaR}_{\alpha}\left(-\,\mathrm{IAVaR}_2^{\alpha}\left(X^{(2)}\right)\mid\mathcal{F}_1\right)(\omega) = 0,$$

$$\mathrm{IAVaR}_0^{\alpha}\left(X^{(2)}\right)(\omega) = \frac{1}{1+r}\,\mathrm{AVaR}_{\alpha}\left(-\,\mathrm{IAVaR}_1^{\alpha}\left(X^{(2)}\right)\right)(\omega) = 0.$$

- Now assume that $\alpha \le p$. We compute

$$\begin{aligned}
\mathrm{IAVaR}_2^{\alpha}\left(X^{(2)}\right)(\omega) &= \frac{1}{1+r}\,\mathrm{AVaR}_{\alpha}\left(X_3^{(2)}\middle|\,\mathcal{F}_2\right)(\omega) \\
&= \begin{cases} -\frac{1}{1+r}, & \omega \in \{(u,u)\}\times\{u,d\} \\ \frac{\alpha-p}{(1+r)(1-\alpha)}, & \omega \in \{(u,d),(d,u)\}\times\{u,d\} \\ 0, & \omega \in \{(d,d)\}\times\{u,d\}, \end{cases} \\
\mathrm{IAVaR}_1^{\alpha}\left(X^{(2)}\right)(\omega) &= \frac{1}{1+r}\,\mathrm{AVaR}_{\alpha}\left(-\,\mathrm{IAVaR}_2^{\alpha}\left(X^{(2)}\right)\mid\mathcal{F}_1\right)(\omega) \\
&= \begin{cases} -\frac{(\alpha-p)(\alpha+p-2)}{(1+r)^2(1-\alpha)^2}, & \omega \in \{u\}\times\{u,d\}^2 \\ -\frac{(\alpha-p)^2}{(1+r)^2(1-\alpha)^2}, & \omega \in \{d\}\times\{u,d\}^2, \end{cases} \\
\mathrm{IAVaR}_0^{\alpha}\left(X^{(2)}\right)(\omega) &= \frac{1}{1+r}\,\mathrm{AVaR}_{\alpha}\left(-\,\mathrm{IAVaR}_1^{\alpha}\left(X^{(2)}\right)\right)(\omega) = \frac{(\alpha-p)^2(\alpha+2p-3)}{(1+r)^3(1-\alpha)^3}.
\end{aligned}$$

We summarize all results in Table 3.1. The star in the last row is as follows

$$\star = \begin{cases} >, & \alpha - 2p + 1 < 0 \\ =, & \alpha - 2p + 1 = 0 \\ <, & \alpha - 2p + 1 > 0. \end{cases}$$

IAVaR_t^α	$\omega \in$	$\alpha \geq p$			$\alpha < p$		
		$X^{(1)}$		$X^{(2)}$	$X^{(1)}$		$X^{(2)}$
2	$\{(u,u)\} \times \{u,d\}$	0	$>$	$-\frac{1}{1+r}$	$\frac{\alpha-p}{(1+r)(1-\alpha)}$	$>$	$-\frac{1}{1+r}$
	$\{(u,d),(d,u)\} \times \{u,d\}$	0	$=$	0	$\frac{\alpha-p}{(1+r)(1-\alpha)}$	$=$	$\frac{\alpha-p}{(1+r)(1-\alpha)}$
	$\{(d,d)\} \times \{u,d\}$	0	$=$	0	$\frac{\alpha-p}{(1+r)(1-\alpha)}$	$<$	0
1	$\{u\} \times \{u,d\}^2$	0	$=$	0	$\frac{\alpha-p}{(1+r)^2(1-\alpha)}$	$>$	$-\frac{(\alpha-p)(\alpha+p-2)}{(1+r)^2(1-\alpha)^2}$
	$\{d\} \times \{u,d\}^2$	0	$=$	0	$\frac{\alpha-p}{(1+r)^2(1-\alpha)}$	$<$	$-\frac{(\alpha-p)^2}{(1+r)^2(1-\alpha)^2}$
0	$\{u,d\}^3$	0	$=$	0	$\frac{\alpha-p}{(1+r)^3(1-\alpha)}$	$\star$	$\frac{(\alpha-p)^2(\alpha+2p-3)}{(1+r)^3(1-\alpha)^3}$

Table 3.1: The Iterated Conditional Average Value at Risk for the Artzner game

Since for $\alpha \geq p$ the results are trivial, we comment only those for $\alpha < p$.

At time 2 in the case $\omega \in \{(u,d),(d,u)\} \times \{u,d\}$ both processes have the same risk, because incomes depend only on the third throw. If $\omega \in \{(u,u)\} \times \{u,d\}$, there is a sure income of 1 associated with $X^{(2)}$. Therefore $X^{(2)}$ is less risky. Similarly, for $\omega \in \{(d,d)\} \times \{u,d\}$ we have nothing from $X^{(2)}$, but still a chance for a positive income from $X^{(1)}$, so the second game is more risky.

At time 1 if $\omega \in \{u\} \times \{u,d\}^2$, then the income from $X^{(2)}$ is more probable, so this process has lower risk. For $\omega \in \{d\} \times \{u,d\}^2$ we have a reversed situation.

At time 0 we have no additional information and every possible ordering can occur, depending on parameters α and p.

3.5. Geometric Brownian motion

In general, computing the Iterated Conditional Average Value at Risk can cause problems, because there are no closed formulas for it. However, for some specific processes it is possible. This is the case for a geometric Brownian motion.

Suppose that a process $X = (X_t)_{t\in[0,T]}$ is a geometric Brownian motion with parameters μ and $\sigma > 0$. In other words, X is of the form

$$X_t = \exp\left(\mu t + \sigma W_t\right),$$

where W is a standard Brownian motion. Since we consider discrete time, define $\tilde{X} = (X_0, X_1, \ldots, X_T)$ and a filtration generated by $\tilde{X}$, i.e., $\mathcal{F}_t = \sigma\left(X_0, X_1, \ldots, X_t\right)$, $t \in \mathcal{T}$. Let Y be given by $Y_t = \ln X_t = \mu t + \sigma W_t$, $t \in \mathcal{T}$.

Fix $t \in \mathcal{T}_+$. We want to find a distribution of a random variable $Y_t|\, Y_0, Y_1, \dots, Y_{t-1}$. Since Y_0 is a constant, the above random variable equals $Y_t|\, Y_1, \dots, Y_{t-1}$. We name its density by $f_{Y_t|\, Y_1,\dots,Y_{t-1}}$. We know that

$$f_{Y_t|\, Y_1,\dots,Y_{t-1}}\left(y_t|\, y_1, \dots, y_{t-1}\right) = \frac{f_{Y_1,\dots,Y_t}\left(y_1, \dots, y_t\right)}{f_{Y_1,\dots,Y_{t-1}}\left(y_1, \dots, y_{t-1}\right)},$$

where $f_{Z_1,\dots,Z_n}$ is a joint density of random variables $Z_1, \dots, Z_n$.

Take $n \in \mathbb{N} \cap \mathcal{T}_+$ and note that the two following systems of equations are equivalent

$$\begin{cases} Y_1 = y_1 \\ Y_2 = y_2 \\ \quad\vdots \\ Y_n = y_n, \end{cases} \quad \begin{cases} Y_1 = y_1 \\ Y_2 - Y_1 = y_2 - y_1 \\ \quad\vdots \\ Y_n - Y_{n-1} = y_n - y_{n-1}. \end{cases}$$

Therefore we have

$$f_{Y_1,\dots,Y_n}\left(y_1, \dots, y_n\right) = f\left(y_1\right) \prod_{k=1}^{n-1} f\left(y_{k+1} - y_k\right),$$

where f is a density of a random variable normally distributed with parameters μ and σ. Hence

$$f_{Y_t|\, Y_1,\dots,Y_{t-1}}\left(y_t|\, y_1, \dots, y_{t-1}\right) = \frac{f\left(y_1\right) \prod_{k=1}^{t-1} f\left(y_{k+1} - y_k\right)}{f\left(y_1\right) \prod_{k=1}^{t-2} f\left(y_{k+1} - y_k\right)} = f\left(y_t - y_{t-1}\right).$$

Due to that, $Y_t|\, Y_1, \dots, Y_{t-1}$ has a normal distribution with parameters $\mu + \ln X_{t-1}$ and σ. As a consequence, $X_t|\, X_1, \dots, X_{t-1}$ is log-normal distributed with the same parameters. Knowing that we can come back to our example.

Recall that, by (1.3) and (1.4), we have

$$\begin{aligned} \mathrm{VaR}_\alpha\left(Z|\, \mathcal{G}\right)(\omega) &= \inf\left\{x \mid \mathbb{P}\left(Z + x < 0 \,|\, \mathcal{G}\right)(\omega) \le 1 - \alpha\right\}, \\ \mathrm{AVaR}_\alpha\left(Z|\, \mathcal{G}\right) &= \mathrm{VaR}_\alpha\left(Z|\, \mathcal{G}\right) + \frac{1}{1-\alpha}\, \mathbb{E}\left(\left(Z + \mathrm{VaR}_\alpha\left(Z|\, \mathcal{G}\right)\right)^- \,|\, \mathcal{G}\right). \end{aligned}$$

For a fixed $x \in \mathbb{R}$ we compute

$$\begin{aligned} \mathbb{P}\left(X_t + x < 0 \,|\, \mathcal{F}_{t-1}\right) &= \mathbb{E}\left(\mathbb{1}_{\{X_t + x < 0\}} \,|\, X_1, \dots, X_{t-1}\right) = \int_{\infty}^{-x} f^{\text{log-normal}}_{\mu + \ln X_{t-1}, \sigma}(t)\, \mathrm{d}t \\ &= \begin{cases} 0, & x \ge 0 \\ \Phi\left(\frac{\ln(-x) - \mu - \ln X_{t-1}}{\sigma}\right), & x < 0. \end{cases} \end{aligned}$$

Therefore, for $x < 0$,

$$\mathbb{P}\left(X_t + x < 0 \,|\, \mathcal{F}_{t-1}\right) \le 1 - \alpha \ \text{ if and only if } \ x \ge -X_{t-1} \exp\left(\mu + \sigma \Phi^{-1}(1-\alpha)\right),$$

where Φ stands for a standard normal cumulative distribution function. Then

$$\mathrm{VaR}_\alpha\left(X_t \middle| \mathcal{F}_{t-1}\right) = -X_{t-1}\exp\left(\mu + \sigma\Phi^{-1}(1-\alpha)\right) = -c_\alpha X_{t-1},$$

where $c_\alpha = \exp\left(\mu + \sigma\Phi^{-1}(1-\alpha)\right)$. Moreover,

$$\begin{aligned}
\mathbb{E}&\left(\left(X_t + \mathrm{VaR}_\alpha\left(X_t \middle| \mathcal{F}_{t-1}\right)\right)^- \middle| \mathcal{F}_{t-1}\right) = \mathbb{E}\left(\left(X_t - c_\alpha X_{t-1}\right)^- \middle| \mathcal{F}_{t-1}\right) \\
&= \int_{-\infty}^{c_\alpha X_{t-1}} \left(-t + c_\alpha X_{t-1}\right) f_{\mu+\ln X_{t-1},\sigma}^{\text{log-normal}}(t)\,\mathrm{d}t \\
&= -\exp\left(\mu + \ln X_{t-1} + \frac{\sigma^2}{2}\right)\left(1 - \Phi\left(\frac{-\ln\left(c_\alpha X_{t-1}\right) + \mu + \ln X_{t-1} + \sigma^2}{\sigma}\right)\right) \\
&\quad + c_\alpha X_{t-1}\Phi\left(\frac{\ln\left(c_\alpha X_{t-1}\right) - \mu - \ln X_{t-1}}{\sigma}\right) \\
&= -X_{t-1}\exp\left(\mu + \frac{\sigma^2}{2}\right)\Phi\left(\Phi^{-1}(1-\alpha) - \sigma\right) + (1-\alpha)c_\alpha X_{t-1}.
\end{aligned}$$

It follows that

$$\mathrm{AVaR}_\alpha\left(X_t \middle| \mathcal{F}_{t-1}\right) = -\frac{\Phi\left(\Phi^{-1}(1-\alpha) - \sigma\right)\exp\left(\mu + \frac{\sigma^2}{2}\right)}{1-\alpha} X_{t-1} = -d_\alpha X_{t-1}$$

for $d_\alpha = \Phi\left(\Phi^{-1}(1-\alpha) - \sigma\right)\exp\left(\mu + \sigma^2/2\right)/(1-\alpha)$.

Our aim is to show that

$$\mathrm{IAVaR}_t^\alpha\left(\tilde{X}\right) = -\sum_{n=0}^{T-t}\left(\frac{d_\alpha}{1+r}\right)^n X_t. \tag{3.3}$$

Again we do mathematical induction.

We have

$$\mathrm{IAVaR}_{T-1}^\alpha\left(\tilde{X}\right) = \frac{1}{1+r}\mathrm{AVaR}_\alpha\left(X_T \middle| \mathcal{F}_{T-1}\right) - X_{T-1} = -\left(1 + \frac{d_\alpha}{1+r}\right)X_{T-1}.$$

Now fix $t \in \{0, 1, \ldots, T-2\}$ and suppose that the following holds

$$\mathrm{IAVaR}_{t+1}^\alpha\left(\tilde{X}\right) = -\sum_{n=0}^{T-t-1}\left(\frac{d_\alpha}{1+r}\right)^n X_{t+1}.$$

Then

$$\begin{aligned}
\mathrm{IAVaR}_t^\alpha\left(\tilde{X}\right) &= \frac{1}{1+r}\mathrm{AVaR}_\alpha\left(-\mathrm{IAVaR}_{t+1}^\alpha\left(\tilde{X}\right) \middle| \mathcal{F}_t\right) - X_t \\
&= \frac{1}{1+r}\mathrm{AVaR}_\alpha\left(\sum_{n=0}^{T-t-1}\left(\frac{d_\alpha}{1+r}\right)^n X_{t+1} \middle| \mathcal{F}_t\right) - X_t \\
&= \sum_{n=1}^{T-t}\frac{d_\alpha^{n-1}}{(1+r)^n}\mathrm{AVaR}_\alpha\left(X_{t+1} \middle| \mathcal{F}_t\right) - X_t = -\sum_{n=0}^{T-t}\left(\frac{d_\alpha}{1+r}\right)^n X_t.
\end{aligned}$$

As a consequence, (3.3) holds true.

Chapter 4

Pflug–Ruszczyński risk measure

This chapter is dedicated to the *Pflug–Ruszczyński measure*. Its static version is indeed due to Georg Pflug and Andrzej Ruszczyńki (see [PR 05]). However, the dynamic version was created by André Mundt in his PhD thesis ([M 07]). The measure was constructed intuitively there. Then, under assumption that the model has a Markovian structure and by applying Markov decision theory, a closed formula was obtained. Here we proceed in a different way. First we define a dynamic risk measure explicitly. Later on we motivate our choice by developing an optimization problem for which the measure is a solution.

Again we consider discrete time with a time horizon $T \in \mathbb{N}_+$, i.e., $\mathcal{T} = \{0, 1, \ldots, T\}$, $\mathcal{T}_- = \{0, 1, \ldots, T-1\}$. Additionally by $\mathcal{T}_+$ we denote the set $\{1, \ldots, T\}$.

4.1. Definition and properties

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space endowed with a filtration $(\mathcal{F}_t)_{t\in\mathcal{T}}$ such that $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{F}$. We restrict ourselves to consider only income processes that are adapted and integrable, so $\mathcal{X} = \{(I_0, I_1, \ldots, I_T) \mid I_t \in L^1(\Omega, \mathcal{F}_t, \mathbb{P}), \ t \in \mathcal{T}\}$. Moreover, by $c_t > 0$, $t \in \mathcal{T}$, we denote a discounting factor from time t to 0. It means that an income of 1 at t has a value c_t at time 0. In particular, $c_0 = 1$. The final wealth is discounted with a factor $c_{T+1} > 0$. We also assume that $c_{t+1} < c_t$, $t \in \mathcal{T}$. Due to the time value of money principle, it seems reasonable.

Fix a sequence $(\gamma_t)_{t\in\mathcal{T}}$ such that $\gamma_t \in (0, 1)$, $t \in \mathcal{T}$. For each $t \in \mathcal{T}_-$ define a static risk measure by

$$\rho^{(t)}(X) = \lambda_t \mathbb{E}(-X) + (1 - \lambda_t) \operatorname{AVaR}_{\gamma_t}(X), \ X \in L^1(\Omega, \mathcal{F}, \mathbb{P}), \tag{4.1}$$

where

$$\lambda_t = \frac{c_{t+1}}{c_t}.$$

Note that every $\rho^{(t)}$ is just a convex combination of two coherent risk measures, the negative expectation and the Average Value at Risk. As a consequence, the new measure is coherent as well.

Now we are ready to state the already mentioned definition:

Definition 4.1. *A mapping* $\rho^{\mathrm{PR}}\colon \Omega \times \mathcal{T}_- \times \mathcal{X} \ni (\omega, t, I) \mapsto \rho_t^{\mathrm{PR}}(I)(\omega) \in \bar{\mathbb{R}}$ *such that*

$$\rho_t^{\mathrm{PR}}(I) = -\frac{c_{t+1}}{c_t}\left(I_t + \mathrm{VaR}_{\gamma_t}\left(I_t \middle| \mathcal{F}_{t-1}\right)\mathbb{1}_{\mathcal{T}_+}(t)\right)^+ + \mathbb{E}\left(\sum_{n=t+1}^{T}\frac{c_n}{c_t}\rho^{(n)}\left(I_n \middle| \mathcal{F}_{n-1}\right) \middle| \mathcal{F}_t\right),$$

where each $\rho^{(n)}$ *is given by* (4.1), *is called the* dynamic Pflug–Ruszczyński risk measure.

From the above definition it immediately follows that the process $\left(\rho_t^{\mathrm{PR}}(I)\right)_{t\in\mathcal{T}_-}$ is adapted with respect to the filtration $(\mathcal{F}_t)_{t\in\mathcal{T}_-}$ for every $I \in \mathcal{X}$. It is also clear that ρ^{PR} is independent of the past. Furthermore, we have the following theorem:

Theorem 4.2. *Let* $t \in \mathcal{T}_-$. *Then the following statements are true:*

(1) *Let* $Y = (0, \ldots, 0, Y_{t+1}, \ldots, Y_T) \in \mathcal{X}$ *be a predictable process such that the sum* $\sum_{n=t+1}^{T} c_n Y_n$ *is* $\mathcal{F}_t$*-measurable. Then*

$$\rho_t^{\mathrm{PR}}(I + Y) = \rho_t^{\mathrm{PR}}(I) - \sum_{n=t+1}^{T}\frac{c_n}{c_t}Y_n,\ I \in \mathcal{X}.$$

(2) *For* $I^{(1)}, I^{(2)} \in \mathcal{X}$ *such that* $I_t^{(i)} + \mathrm{VaR}_{\gamma_t}\left(I_t^{(i)} \middle| \mathcal{F}_{t-1}\right)\mathbb{1}_{\mathcal{T}_+}(t) \le 0$, $i = 1, 2$, *it holds that*

$$\begin{aligned}
&\rho_t^{\mathrm{PR}}\left(I^{(1)}\right) \le \rho_t^{\mathrm{PR}}\left(I^{(2)}\right) \ \text{ if } I^{(1)} \ge I^{(2)},\\
&\rho_t^{\mathrm{PR}}\left(I^{(1)} + I^{(2)}\right) \le \rho_t^{\mathrm{PR}}\left(I^{(1)}\right) + \rho_t^{\mathrm{PR}}\left(I^{(2)}\right).
\end{aligned}$$

(3) *If* $\Lambda \in L^\infty(\Omega, \mathcal{F}_{t-1}, \mathbb{P})$ *and* $\Lambda > 0$, *then*

$$\rho_t^{\mathrm{PR}}(\Lambda I) = \Lambda\rho_t^{\mathrm{PR}}(I).$$

Proof. Fix $I \in \mathcal{X}$ and a predictable process $Y = (0, \ldots, 0, Y_{t+1}, \ldots, Y_T) \in \mathcal{X}$ such that $\sum_{n=t+1}^{T} c_n Y_n$ is measurable with respect to $\mathcal{F}_t$. Since

$$\begin{aligned}
\rho_t^{\mathrm{PR}}(I + Y) &= -\frac{c_{t+1}}{c_t}\left(I_t + \mathrm{VaR}_{\gamma_t}\left(I_t \middle| \mathcal{F}_{t-1}\right)\mathbb{1}_{\mathcal{T}_+}(t)\right)^+ + \mathbb{E}\left(\sum_{n=t+1}^{T}\frac{c_n}{c_t}\rho^{(n)}\left(I_n + Y_n \middle| \mathcal{F}_{n-1}\right) \middle| \mathcal{F}_t\right)\\
&= \rho_t^{\mathrm{PR}}(I) - \sum_{n=t+1}^{T}\frac{c_n}{c_t}Y_n,
\end{aligned}$$

assertion (1) holds.

Assertions (2) and (3) are obvious. ∎

In general, dynamic translation invariance, monotonicity, dynamic positive homogeneity and subadditivity (see Definition 2.12 and Definition 2.14) do not hold. To see that, consider the following example. Let $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, $\mathbb{P}(\omega_i) = 1/4$ for each $i = 1, \ldots, 4$ and

$$\mathcal{F}_0 = \{\emptyset, \Omega\}, \quad \mathcal{F}_1 = \sigma(\{\omega_i, \omega_{i+1}\}, i = 1, 3), \quad \mathcal{F}_2 = \sigma(\{\omega_i\}, i = 1, \ldots, 4).$$

Suppose that we are given a constant interest rate $r > 0$. Hence $c_t = (1+r)^{-t}$, $t = 0, 1, 2$. Take $\gamma_1 = 0.9$ and $\xi = \mathbb{1}_{\{\omega_1, \omega_2\}}$. Define a process $Y = (Y_0, Y_1, Y_2)$ via

$$Y_0 = 0, \quad Y_1 = \xi, \quad Y_2 = -\frac{c_1}{c_2}\,\xi.$$

Then Y is adapted and $Y_0 + c_1 Y_1 + c_2 Y_3 = 0$ is $\mathcal{F}_0$-measurable. On the other hand,

$$\rho_0^{\mathrm{PR}}(Y) = \mathbb{E}\left(c_1\rho^{(1)}(Y_1) + c_2\rho^{(2)}(Y_2|\,\mathcal{F}_1)\right) = c_1\rho^{(1)}(\xi) + c_1\mathbb{E}\xi = (c_1 - c_2)\,\mathbb{E}\xi > 0.$$

We also have

$$\begin{aligned}
&\rho_1^{\mathrm{PR}}(0,0,0) = \rho_1^{\mathrm{PR}}(0,1,0) = 0,\\
&\rho_1^{\mathrm{PR}}(0,\xi,0) = -\frac{c_2}{c_1}\left(\xi + \mathrm{VaR}_{\gamma_1}\xi\right)^+ = -\frac{c_2}{c_1}\,\xi,\\
&\rho_1^{\mathrm{PR}}(0,-\xi,0) = -\frac{c_2}{c_1}\left(-\xi + \mathrm{VaR}_{\gamma_1}(-\xi)\right)^+ = -\frac{c_2}{c_1}(1-\xi).
\end{aligned}$$

Hence

$$\begin{aligned}
&\rho_1^{\mathrm{PR}}(0,1,0) > \rho_1^{\mathrm{PR}}(0,\xi,0) \text{ with positive probability, but } 1 \geq \xi,\\
&\rho_1^{\mathrm{PR}}(0,\xi,0) \neq \xi\rho_1^{\mathrm{PR}}(0,1,0),\\
&\rho_1^{\mathrm{PR}}(0,0,0) > \rho_1^{\mathrm{PR}}(0,\xi,0) + \rho_1^{\mathrm{PR}}(0,-\xi,0).
\end{aligned}$$

However, Theorem 4.2 guarantees that the dynamic Pflug–Ruszczyński measure is "almost" a dynamic risk measure in the sense of Definition 2.12. It is due to the fact that the dynamic translation invariance property is satisfied for all predictable processes and ρ^{PR} is monotone for $I \in \mathcal{X}$ with $I_t + \mathrm{VaR}_{\gamma_t}(I_t|\,\mathcal{F}_{t-1})\mathbb{1}_{\mathcal{T}_+}(t) \leq 0$, $t \in \mathcal{T}_-$.

As in the case of the Iterated Conditional Average Value at Risk, from (3.2) it follows that the dynamic Pfug–Ruszczyński risk measure is relevant. Unfortunately, it does not satisfy the time consistency condition. Extending the last example with stopping times $\sigma = 0$ and $\tau = 1$ we get that

$$\begin{aligned}
&\rho_\sigma^{\mathrm{PR}}(\xi\cdot \mathrm{e}_\tau) = \rho_0^{\mathrm{PR}}(0,\xi,0) = \mathbb{E}\left(c_1\rho^{(1)}(\xi)\right) = c_1\rho^{(1)}(\xi) = -c_2\mathbb{E}\xi,\\
&\rho_\sigma^{\mathrm{PR}}\left(\frac{c_\tau}{c_2}\,\xi\cdot \mathrm{e}_2\right) = \rho_0^{\mathrm{PR}}\left(0,0,\frac{c_1}{c_2}\,\xi\right) = \mathbb{E}\left(c_1\rho^{(2)}(\xi|\,\mathcal{F}_1)\right) = -c_1\mathbb{E}\xi.
\end{aligned}$$

Therefore

$$\rho_\sigma^{\mathrm{PR}}(\xi\cdot \mathrm{e}_\tau) \neq \rho_\sigma^{\mathrm{PR}}\left(\frac{c_\tau}{c_2}\,\xi\cdot \mathrm{e}_2\right).$$

We finish the section with a simple remark:

Remark 4.3. *Let $\xi \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and $I \in \mathcal{X}$. We define two processes as follows:*

$$S_t^{(1)} = c_t \rho_t^{\mathrm{PR}}(0, \ldots, 0, \xi),$$
$$S_t^{(2)} = c_t \rho_t^{\mathrm{PR}}(I) + c_{t+1} \left(I_t + \mathrm{VaR}_{\gamma_t} \left(I_t \middle| \mathcal{F}_{t-1} \right) \mathbb{1}_{\mathcal{T}_+}(t) \right)^+ + \sum_{n=1}^{t} c_n \rho^{(n)} \left(I_n \middle| \mathcal{F}_{n-1} \right), \qquad t \in \mathcal{T}_-.$$

Then $S^{(i)}$, $i = 1, 2$, is a martingale with respect to the filtration $(\mathcal{F}_t)_{t \in \mathcal{T}_-}$.

Proof. Since for $t \in \mathcal{T}_-$ it holds that

$$\begin{aligned} S_t^{(1)} &= \mathbb{E}\left(c_T \rho^{(T)} \left(\xi \middle| \mathcal{F}_{T-1} \right) \middle| \mathcal{F}_t \right), \\ S_t^{(2)} &= \mathbb{E}\left(\sum_{n=t+1}^{T} c_n \rho^{(n)} \left(I_n \middle| \mathcal{F}_{n-1} \right) \middle| \mathcal{F}_t \right) + \sum_{n=1}^{t} c_n \rho^{(n)} \left(I_n \middle| \mathcal{F}_{n-1} \right) \\ &= \mathbb{E}\left(\sum_{n=1}^{T} c_n \rho^{(n)} \left(I_n \middle| \mathcal{F}_{n-1} \right) \middle| \mathcal{F}_t \right), \end{aligned}$$

we are done. □

Compare the last remark to Remark 2.20. The latter guarantees that the process is only a supermartingale, if the dynamic risk measure is, in particular, time-consistent. The dynamic Pflug–Ruszczyński risk measure is not, but for it we have even stronger result.

4.2. Motivation

In this section we concern ourselves with motivating the definition of the dynamic Pflug–Ruszczyński measure. Although it could be surprising at first glance, it makes sense to consider such a measure. Our main tool will be Markov decision theory (see Appendix B). That involves the necessity of an assumption that the model is Markovian.

4.2.1. Basic idea

Consider a company for which $I = (I_0, I_1, \ldots, I_T) \in \mathcal{X}$ is an income process. At each time $t \in \mathcal{T}_-$ a manager determines an amount a_t that is going to be consumed at time $t+1$. It follows that a_t is an $\mathcal{F}_t$-measurable random variable. We define an *accumulated wealth process* of the company by

$$\begin{cases} W_0 = I_0, \\ W_{t+1} = W_t^+ + I_{t+1} - a_t, \; t \in \mathcal{T}_-. \end{cases}$$

If for any $t \in \mathcal{T}$ it holds that $W_t < 0$, the company faces a loss of $-W_t$. Then it is obliged to pay immediately $q_t / c_t W_t^-$, where

$$q_t = \frac{c_t - c_{t+1}\gamma_t}{1 - \gamma_t}, \; t \in \mathcal{T}.$$

The number q_t can be seen as an insurance premium at time t. Because $c_{t+1} < c_t$, $t \in \mathcal{T}$, we also get that $c_t < q_t$, $t \in \mathcal{T}$.

Taking $t \in \mathcal{T}_-$ as a starting point, we aim at maximizing the discounted expected utility associated with all future cash flows. It is clear that at time $n > t$ the company has an amount of the discounted value of

$$c_n a_{n-1} - q_n W_n^- + c_{T+1} W_T^+ \mathbb{1}_{\{T\}}(n)$$

at its disposal. As a consequence, we want to maximize

$$\frac{1}{c_t} \left(\sum_{n=t+1}^{T} \left(c_n a_{n-1} - q_n W_n^- \right) + c_{T+1} W_T^+ \right)$$

over all decision vectors $(a_t, \dots, a_{T-1}) \in \mathcal{X}^{(T-t)}$, where

$$\mathcal{X}^{(T-t)} = \{ (X_t, \dots, X_{T-1}) \mid X_n \text{ is } \mathcal{F}_n\text{-measurable}, \, t \leq n \leq T-1 \} .$$

Hence we define a mapping $\rho \colon \Omega \times \mathcal{T}_- \times \mathcal{X} \to \bar{\mathbb{R}}$ by

$$\rho_t(I) = - \operatorname*{ess\,sup}_{(a_t, \dots, a_{T-1}) \in \mathcal{X}^{(T-t)}} \frac{1}{c_t} \mathbb{E} \left(\sum_{n=t+1}^{T} \left(c_n a_{n-1} - q_n W_n^- \right) + c_{T+1} W_T^+ \,\middle|\, \mathcal{F}_t \right) \tag{4.2}$$

for $I \in \mathcal{X}$ and $t \in \mathcal{T}_-$. In the next section we will prove that the above essential supremum is attained and ρ coincides with ρ^{PR} on some subset of $\mathcal{X}$.

4.2.2. Optimization problem solving via Markov decision theory

As we have already mentioned, we want to solve the optimization problem from (4.2). To do so, we need to introduce the Markov environment for the income process I.

Suppose that $Y_0, Y_1, \dots, Y_T$ are stochastically independent random variables such that $\mathcal{F}_t = \sigma(Y_s, \, s \leq t)$, $t \in \mathcal{T}$. Moreover, there is also given a Markov chain $(Z_t)_{t \in \mathcal{T}}$ such that $Z_0 = c \in \mathbb{R}$ and Z_{t+1} depends only on Z_t and Y_{t+1}, $t \in \mathcal{T}_-$. Therefore there is a measurable function $f_t \colon \mathbb{R}^2 \to \mathbb{R}$ with

$$Z_t = f_t (Z_{t-1}, Y_t) .$$

We consider only income processes I such that I_t depends only on Z_t for each $t \in \mathcal{T}$. Then there exists a measurable function $g_t \colon \mathbb{R}^2 \to \mathbb{R}$, $t \in \mathcal{T}_+$, such that the following holds

$$I_t = g_t (Z_{t-1}, Y_t) .$$

The set of all these processes I we denote by $\mathcal{X}^{\mathrm{M}}$.

Now we concentrate on defining a Markov decision process. Let $S \subset \mathbb{R}^2$ be a *state space* equipped with a sub-σ-algebra $\mathcal{S} = \mathcal{B}_S^2$. If $s = (w, z) \in \mathcal{S}$, it means that w and z are realizations of the wealth process and the Markov chain, respectively. An *action space* we denote by A. Then $A \subset \mathbb{R}$ and its sub-σ-algebra equals $\mathcal{A} = \mathcal{B}_A$. An element $a \in A$ is simply an amount of money dedicated to be consumed. There are no restriction on actions, so a *restriction set* $D = S \times A \subset \mathbb{R}^3$. For $t \in \mathcal{T}_-$ we set

$$\Pi^{(T-t)} = \left\{\pi^{(T-t)} = (\pi_t, \ldots, \pi_{T-1}) \mid \pi_n \colon S \to A \text{ is measurable}, \, n = t, \ldots, T-1\right\}.$$

The above is a *set of* $(T-t)$*-step admissible policies.* $E \subset \mathbb{R}$ (equipped with a sub-σ-algebra $\mathcal{E} = \mathbb{B}_E$) is a set of all *disturbances.* A *transition function* $T_t \colon D \times E \to S$ is defined as follows

$$T_t\left(w, z, a, y\right) = \left(w^+ + g_t(z, y) - a, f_t(z, y)\right).$$

By $Q_t \colon D \times \mathcal{E} \to [0, 1]$ with

$$Q_t\left(w, z, a, B\right) = \mathbb{P}\left(Y_t \in B\right)$$

we denote a *transition law.* Finally we define *rewards.* The *one-step reward function* is given by

$$r_t \colon D \ni (w, z, a) \mapsto r_t\left(w, z, a\right) = c_{t+1}a - q_t w^- \in \mathbb{R}.$$

A *terminal reward function* is the following mapping

$$V_T \colon S \ni (w, z) \mapsto V_T\left(w, z\right) = c_{T+1}w^+ - q_T w^- \in \mathbb{R}.$$

For $\pi^{(T)} = (\pi_0, \pi_1, \ldots, \pi_{T-1}) \in \Pi^{(T)}$ we define a Markov decision process by

$$X_0 = s_0 = (w_0, z_0) \in S, \quad X_t = T_t\left(X_{t-1}, \pi_{t-1}\left(X_{t-1}\right), Y_t\right), \, t \in \mathcal{T}_+.$$

We introduce *value functions*:

$$\begin{aligned} &V_{t,\pi^{(T-t)}}(s) = \mathbb{E}\left(\sum_{n=t}^{T-1} r_n\left(X_n, \pi_n\left(X_n\right)\right) + V_T\left(X_T\right) \middle| X_t = s\right), \, \pi^{(T-t)} \in \Pi^{(T-t)}, \\ &V_t(s) = \sup_{\pi^{(T-t)} \in \Pi^{(T-t)}} V_{t,\pi^{(T-t)}}(s). \end{aligned}$$

Fix $t \in \mathcal{T}_+$ and suppose that $y = (y_1, \ldots, y_t) \in E^t$ is a realization of a random vector $(Y_1, \ldots, Y_t)$. Take $\omega \in \{(Y_1, \ldots, Y_t) = y\}$. Since both W and Z are adapted, there exists a measurable function $h_t^{W,Z} \colon E^t \to S$ such that

$$X_t = (W_t, Z_t) = h_t^{W,Z}\left(Y_1, \ldots, Y_t\right)$$

or, equivalently,

$$X_t(\omega) = \left(W_t, Z_t\right)(\omega) = h_t^{W,Z}(y).$$

It follows that for $I \in \mathcal{X}^{\mathrm{M}}$ we have

$$\begin{aligned}
\rho_t(I)(\omega) &= - \operatorname*{ess\,sup}_{a=(a_t,\dots,a_{T-1})} \frac{1}{c_t} \mathbb{E}\left(\sum_{n=t+1}^{T} \left(c_n a_{n-1} - q_n W_n^-\right) + c_{T+1} W_T^+ \,\middle|\, \mathcal{F}_t \right)(\omega) \\
&= - \sup_a \frac{1}{c_t} \mathbb{E}\left(\sum_{n=t}^{T-1} \left(c_{n+1} a_n - q_n W_n^-\right) + q_t W_t^- + c_{T+1} W_T^+ - q_T W_T^- \,\middle|\, X_t = h_t^{W,Z}(y) \right) \\
&= -\frac{q_t}{c_t} \left(h_t^{W,Z}(y)\right)_1^- - \sup_{\pi^{(T-t)}} \frac{1}{c_t} \mathbb{E}\left(\sum_{n=t}^{T-1} r_n\left(X_n, \pi_n\left(X_n\right)\right) + V_T\left(X_T\right) \,\middle|\, X_t = h_t^{W,Z}(y) \right) \\
&= -\frac{q_t}{c_t} \left(h_t^{W,Z}(y)\right)_1^- - \frac{1}{c_t} V_t\left(h_t^{W,Z}(y)\right),
\end{aligned}$$

where $(\,\cdot\,)_1$ stands for the projection onto the first coordinate. In general, it holds that

$$\rho_t(I) = -\frac{q_t}{c_t} W_t^- - \frac{1}{c_t} V_t\left(X_t\right). \tag{4.3}$$

Now it is an appropriate moment for the most important theorem of this section. We apply Markov decision theory and, by using the dynamic programming theorem (see Theorem B.4), we get explicit formulas for the value functions.

Theorem 4.4. *Let $t \in \mathcal{T}$ and $s^* = (w^*, z^*) \in S$. Then the value function is given by*

$$V_t\left(s^*\right) = c_{t+1}\left(w^*\right)^+ - q_t\left(w^*\right)^- - \mathbb{1}_{\mathcal{T}_-}(t) \sum_{n=t+1}^{T} c_n \mathbb{E}\left(\rho^{(n)}\left(I_n \,|\, Z_{n-1}\right) \,|\, Z_t = z^*\right). \tag{4.4}$$

Moreover, the optimal policy $\pi_*^{(T-t)} = \left(\pi_t^*, \dots, \pi_{T-1}^*\right) \in \Pi^{(T-t)}$ and the optimal Markov process $(X_t^*, \dots X_T^*)$ are as follows

$$\begin{aligned}
&\pi_n^*(s) = w^+ - \mathrm{VaR}_{\gamma_n}\left(I_{n+1} \,|\, Z_n = z\right), \; t \le n \le T-1, \\
&X_t^* = s^*, \\
&X_n^* = \left(I_n + \mathrm{VaR}_{\gamma_n}\left(I_n \,|\, Z_{n-1} = Z_{n-1}^*\right), f_n\left(Z_{n-1}^*, Y_n\right)\right), \; t+1 \le n \le T.
\end{aligned}$$

Proof. First note that for $a \in \mathbb{R}$, $s = (w, z) \in S$ and $t \in \mathcal{T}_+$ we have

$$\begin{aligned}
&c_{t+1}\left(w^+ + g_t\left(z, Y_t\right) - a\right)^+ - q_t\left(w^+ + g_t\left(z, Y_t\right) - a\right)^- \\
&= c_{t+1}\left(w^+ + g_t\left(z, Y_t\right) - a\right) - \left(q_t - c_{t+1}\right)\left(w^+ + g_t\left(z, Y_t\right) - a\right)^- \\
&= c_{t+1}\left(w^+ + g_t\left(z, Y_t\right)\right) - c_t a + \left(c_t - c_{t+1}\right)\left(a - \frac{q_t - c_{t+1}}{c_t - c_{t+1}}\left(w^+ + g_t\left(z, Y_t\right) - a\right)^-\right) \\
&= c_{t+1}\left(w^+ + g_t\left(z, Y_t\right)\right) - c_t a + \left(c_t - c_{t+1}\right)\left(a - \frac{1}{1-\gamma_t}\left(w^+ + g_t\left(z, Y_t\right) - a\right)^-\right).
\end{aligned}$$

Therefore

$$
\begin{aligned}
&\sup_a \Big(r_{t-1}(s,a) + \mathbb{E}\Big(c_{t+1}\big(w^+ + g_t(z,Y_t) - a\big)^+ - q_t\big(w^+ + g_t(z,Y_t) - a\big)^-\Big)\Big) \\
&= -q_{t-1}w^- + c_{t+1}\big(w^+ + \mathbb{E}g_t(z,Y_t)\big) - (c_t - c_{t+1})\inf_a \left(-a + \frac{\mathbb{E}\left(w^+ + g_t(z,Y_t) - a\right)^-}{1-\gamma_t}\right) \\
&= -q_{t-1}w^- + c_{t+1}\big(w^+ + \mathbb{E}g_t(z,Y_t)\big) - (c_t - c_{t+1})\big(\mathrm{AVaR}_{\gamma_t} g_t(z,Y_t) - w^+\big) \\
&= -q_{t-1}w^- + c_t w^+ + c_t\left(\lambda_t \mathbb{E}g_t(z,Y_t) - (1-\lambda_t)\,\mathrm{AVaR}_{\gamma_t} g_t(z,Y_t)\right) \\
&= -q_{t-1}w^- + c_t w^+ - c_t\rho^{(t)}\left(g_t(z,Y_t)\right) \\
&= -q_{t-1}w^- + c_t w^+ - c_t\mathbb{E}\left(\rho^{(t)}\left(I_t \mid Z_{t-1}\right) \mid Z_{t-1} = z\right).
\end{aligned}
$$

The proof is by backward induction. Obviously,

$$V_T(s) = c_{T+1}w^+ - q_T w^-.$$

Furthermore, due to Theorem B.4, we obtain

$$
\begin{aligned}
V_{T-1}(s) &= J_{T-1}(s) = \sup_a \left(r_{T-1}(s,a) + \int_E J_T\left(T_T(s,a,y)\right) Q_T(s,a;\mathrm{d}y)\right) \\
&= \sup_a \left(r_{T-1}(s,a) + \int_E V_T\big(w^+ + g_T(z,y) - a, f_T(z,y)\big) Q_T(s,a;\mathrm{d}y)\right) \\
&= -q_{T-1}w^- + c_T w^+ - c_T\mathbb{E}\left(\rho^{(T)}\left(I_T \mid Z_{T-1}\right) \mid Z_{T-1} = z\right).
\end{aligned}
$$

Moreover, the supremum is attained at

$$a^* = \pi^*_{T-1}(w,z) = -\mathrm{VaR}_{\gamma_T}\big(g_T(z,Y_T) + w^+\big) = w^+ - \mathrm{VaR}_{\gamma_T}\left(I_T \mid Z_{T-1} = z\right)$$

and

$$W^*_T = w^+ + g_T(z,Y_T) - a^* = g_T(z,Y_T) + \mathrm{VaR}_{\gamma_T}\left(I_T \mid Z_{T-1} = z\right).$$

Now suppose that the assertion holds for $t+1$, $t \in \{0,\dots,T-2\}$. Then, again by Theorem B.4,

$$
\begin{aligned}
V_t(s) &= J_t(s) = \sup_a \left(r_t(s,a) + \int_E J_{t+1}\left(T_{t+1}(s,a,y)\right) Q_{t+1}(s,a;\mathrm{d}y)\right) \\
&= \sup_a \left(r_t(s,a) + \int_E V_{t+1}\big(w^+ + g_{t+1}(z,y) - a, f_{t+1}(z,y)\big) Q_{t+1}(s,a;\mathrm{d}y)\right) \\
&= \sup_a \Big(r_t(s,a) + \mathbb{E}\Big(c_{t+2}\big(w^+ + g_{t+1}(z,Y_{t+1}) - a\big)^+ - q_{t+1}\big(w^+ + g_{t+1}(z,Y_{t+1}) - a\big)^-\Big)\Big) \\
&\quad - \mathbb{E}\left(\sum_{n=t+2}^{T} c_n \mathbb{E}\left(\rho^{(n)}\left(I_n \mid Z_{n-1}\right) \mid Z_{t+1} = f_{t+1}(z,Y_{t+1})\right)\right) \\
&= -q_t w^- + c_{t+1}w^+ - c_{t+1}\mathbb{E}\left(\rho^{(t+1)}\left(I_{t+1} \mid Z_t\right) \mid Z_t = z\right) \\
&\quad - \sum_{n=t+2}^{T} c_n \mathbb{E}\left(\mathbb{E}\left(\rho^{(n)}\left(I_n \mid Z_{n-1}\right) \mid Z_{t+1}\right) \mid Z_t = z\right) \\
&= -q_t w^- + c_{t+1}w^+ - \sum_{n=t+1}^{T} c_n \mathbb{E}\left(\rho^{(n)}\left(I_n \mid Z_{n-1}\right) \mid Z_t = z\right).
\end{aligned}
$$

Furthermore,

$$\begin{aligned}a^* &= \pi_t^*(w,z) = w^+ - \mathrm{VaR}_{\gamma_{t+1}}\left(g_{t+1}\left(z, Y_{t+1}\right)\right) = w^+ - \mathrm{VaR}_{\gamma_{t+1}}\left(I_{t+1} \middle| Z_t = z\right),\\ W_{t+1}^* &= w^+ + g_{t+1}\left(z, Y_{t+1}\right) - a^+ = g_{t+1}\left(z, Y_{t+1}\right) + \mathrm{VaR}_{\gamma_t}\left(I_{t+1} \middle| Z_t = z\right).\end{aligned}$$

This completes the proof. ☐

Theorem 4.5. *The following holds true:*

$$\rho_t(I) = -\frac{c_{t+1}}{c_t} W_t^+ + \mathbb{E}\left(\sum_{n=t+1}^{T} \frac{c_n}{c_t} \rho^{(n)}\left(I_n \middle| Z_{n-1}\right) \middle| Z_t\right), \; I \in \mathcal{X}^{\mathrm{M}},\, t \in \mathcal{T}_-. \tag{4.5}$$

Proof. The assertion immediately follows from (4.3) and (4.4). ☐

Once again we remind the reader that the assumption of a Markovian structure of the model was essential to obtain the above formula. Because of that, (4.5) holds only on the set $\mathcal{X}^{\mathrm{M}}$. It is obviously a drawback. Nevertheless, most of the discrete-time models (including the most important ones) meet the assumption.

To conclude the section note that for the optimal wealth process W^* we obtain

$$\rho_t(I) = -\frac{c_{t+1}}{c_t}\left(I_t + \mathrm{VaR}_{\gamma_t}\left(I_t \middle| Z_{t-1}\right) \mathbb{1}_{\mathcal{T}_+}(t)\right)^+ + \mathbb{E}\left(\sum_{n=t+1}^{T} \frac{c_n}{c_t} \rho^{(n)}\left(I_n \middle| Z_{n-1}\right) \middle| Z_t\right),$$

which coincides with the definition of the dynamic Pflug–Ruszczyński measure (see Definition 4.1).

4.3. Artzner game

Recall the example from Section 3.4. Name the income processes by $I^{(1)}$ and $I^{(2)}$. They are presented in Figure 3.3. Let Y_i, $i = 1,2,3$, denote a result of the i-th throw. Then

$$Y_i = \begin{cases} 1, & \text{with probability } p \\ 0, & \text{with probability } 1-p, \end{cases}$$

where 1 and 0 mean that the i-th coin shows heads and tails, respectively. We also define a Markov chain by

$$Z_0 = 0, \quad Z_{t+1} = Z_t + Y_{t+1},\, t \in \{0,1,2\}.$$

Then we write

$$I^{(1)} = \left(0,0,Y_3\right), \quad I^{(2)} = \left(0,0,\mathbb{1}_{\{Z_3 \geq 2\}}\right).$$

Since $I_j^{(i)} = 0$, $i = 1, 2$, $j = 0, 1, 2$, we get that

$$\begin{aligned}
\rho_0^{\mathrm{PR}}\left(I^{(1)}\right) &= c_3 \mathbb{E}\left(\rho^{(3)}\left(Y_3 \middle| Z_2\right)\right) = c_3 \rho^{(3)}\left(Y_3\right), \\
\rho_t^{\mathrm{PR}}\left(I^{(1)}\right) &= \frac{c_3}{c_t}\, \mathbb{E}\left(\rho^{(3)}\left(Y_3 \middle| Z_2\right) \middle| Z_t\right) = \frac{c_3}{c_t}\, \rho^{(3)}\left(Y_3\right), \; t = 1, 2.
\end{aligned}$$

Now we move on to the process $I^{(2)}$. We have

$$\begin{aligned}
\rho_0^{\mathrm{PR}}\left(I^{(2)}\right) &= c_3 \mathbb{E}\left(\rho^{(3)}\left(\mathbb{1}_{\{Z_3 \geq 2\}} \middle| Z_2\right)\right) = c_3 \left(2(1-p)\rho^{(3)}\left(Y_3\right) - p\right) p, \\
\rho_1^{\mathrm{PR}}\left(I^{(2)}\right) &= \frac{c_3}{c_1}\, \mathbb{E}\left(\rho^{(3)}\left(\mathbb{1}_{\{Z_3 \geq 2\}} \middle| Z_2\right) \middle| Z_1\right) = \begin{cases} \frac{c_3}{c_1}\left((1-p)\rho^{(3)}\left(Y_3\right) - p\right), & Z_1 = 1 \\ \frac{c_3}{c_1}\, p\, \rho^{(3)}\left(Y_3\right), & Z_1 = 0, \end{cases} \\
\rho_2^{\mathrm{PR}}\left(I^{(2)}\right) &= \frac{c_3}{c_2}\, \mathbb{E}\left(\rho^{(3)}\left(\mathbb{1}_{\{Z_3 \geq 2\}} \middle| Z_2\right) \middle| Z_2\right) = \frac{c_3}{c_2}\, \rho^{(3)}\left(\mathbb{1}_{\{Z_3 \geq 2\}} \middle| Z_2\right) = \begin{cases} -\frac{c_3}{c_2}, & Z_2 = 2 \\ \frac{c_3}{c_2}\rho^{(3)}\left(Y_3\right), & Z_2 = 1 \\ 0, & Z_2 = 0. \end{cases}
\end{aligned}$$

Now we compute $\rho^{(3)}\left(Y_3\right)$. Since

$$\mathrm{AVaR}_{\gamma_3}\left(Y_3\right) = \frac{\gamma_3 - p}{1 - \gamma_3}\, \mathbb{1}_{(0,p]}\left(\gamma_3\right),$$

we obtain

$$\rho^{(3)}\left(Y_3\right) = -\lambda_3 p + \frac{(1-\lambda_3)(\gamma_3 - p)}{1 - \gamma_3}\, \mathbb{1}_{(0,p]}\left(\gamma_3\right).$$

As before, we summarize the results, see Table 4.1. The star in the last row can be ">", "<" or "=", depending on parameters. For the most important case when all three coins are fair ($p = 1/2$), it holds that the process $I^{(1)}$ has a higher risk and then $\star$ denotes ">". Note that all inequalities (except for $t = 0$) are the same as in Section 3.4 for the Iterated Conditional Average Value at Risk. Because of that, we do not make a comment.

ρ_t^{PR}	$\omega \in$	$I^{(1)}$		$I^{(2)}$
2	$\{Z_2 = 2\}$	$\frac{c_3}{c_2}\rho^{(3)}\left(Y_3\right)$	$>$	$-\frac{c_3}{c_2}$
	$\{Z_2 = 1\}$	$\frac{c_3}{c_2}\rho^{(3)}\left(Y_3\right)$	$=$	$\frac{c_3}{c_2}\rho^{(3)}\left(Y_3\right)$
	$\{Z_2 = 0\}$	$\frac{c_3}{c_2}\rho^{(3)}\left(Y_3\right)$	$<$	0
1	$\{Z_1 = 1\}$	$\frac{c_3}{c_1}\rho^{(3)}\left(Y_3\right)$	$>$	$\frac{c_3}{c_1}\left((1-p)\rho^{(3)}\left(Y_3\right) - p\right)$
	$\{Z_1 = 0\}$	$\frac{c_3}{c_1}\rho^{(3)}\left(Y_3\right)$	$<$	$\frac{c_3}{c_1}\, p\, \rho^{(3)}\left(Y_3\right)$
0	Ω	$c_3\rho^{(3)}\left(Y_3\right)$	$\star$	$c_3\left(2(1-p)\rho^{(3)}\left(Y_3\right) - p\right) p$

Table 4.1: The Pflug–Ruszczyński measure for the Artzner game

4.4. Incomplete information

Sometimes it happens that we do not have complete information about the market. Nevertheless, we still want to assign the riskiness of different financial positions. As we will see in this section, we can deal with the problem by extending the definition of the dynamic Pflug–Ruszczyński risk measure. As previously, it is can be done under the assumption that the model is Markovian. Again we assume that there exist a generating Markov chain $(Z_t)_{t\in\mathcal{T}}$ and random variables $Y_0, Y_1, \ldots, Y_T$ such that $\mathcal{F}_t = \sigma(Y_s,\, s \le t)$. However, each Y_t, $t \in \mathcal{T}$, depends now on a parameter $\vartheta \in \Theta \subset \mathbb{R}$. Since ϑ can be unknown, we also treat it as a random variable with some distribution $\mathcal{L}(\Theta)$ on Θ. We additionally suppose that $Y_0, Y_1, \ldots, Y_T$ are independent under ϑ.

Similarly as in Section 4.2.2, we aim at defining a Markov decision process. Due to the incompleteness of information, we want to apply *Bayesian decision theory*. In that connection we have to extend the state space S to $\tilde{S} = S \times \mathcal{P}(\Theta)$, where $\mathcal{P}(\Theta)$ denotes the set of all probability measures on Θ. Moreover, we fix $\mu_0 \in \mathcal{P}(\Theta)$ as the so-called *prior distribution*. It is our initial prediction of the law of ϑ. Then, using the *Bayes operator*, we get the sequence $(\mu_t)_{t\in\mathcal{T}}$ such that μ_{t+1} is an update of μ_t, $t \in \mathcal{T}_-$. Every μ_t, $t \in \mathcal{T}_+$, is called the *posterior distribution*. We do not go into details, but again we are able to define a Markov decision process $X_t = (W_t, Z_t, \mu_t)$. Therefore we can state the following definition:

Definition 4.6. *A mapping*

$$\rho^{\mathrm{B},\mu_0} \colon \Omega \times \mathcal{T}_- \times \mathcal{X}^{\mathrm{M}} \ni (\omega, t, I) \mapsto \rho(\omega, t, I) = -\frac{q_t}{c_t} W_t^-(\omega) - \frac{1}{c_t} V_t\left(X_t(\omega)\right),$$

where V_t is the value function for the process X, is called the Bayesian Pflug–Ruszczyński risk measure.

In general, it is hard to obtain explicit formulas for the value functions. However, in some special cases it is possible. To see that, assume that $Y_0 = 0$ and Y_t given $\vartheta = \theta$, $t \in \mathcal{T}_+$, is Bernoulli distributed with

$$\mathbb{P}\left(Y_t = u \mid \vartheta = \theta\right) = \theta, \quad \mathbb{P}\left(Y_t = d \mid \vartheta = \theta\right) = 1 - \theta.$$

Moreover, as an initial estimate μ_0 of $\mathcal{L}(\Theta)$ take the uniform distribution on the interval $[0,1]$. Note that $\mathcal{U}([0,1]) = \mathrm{Beta}(1,1)$ and if $\mu_t = \mathrm{Beta}(\alpha, \beta)$, then

$$\mu_{t+1} = \mathrm{Beta}\left(\alpha + \mathbb{1}_{\{u\}}\left(Y_{t+1}\right), \beta + \mathbb{1}_{\{d\}}\left(Y_{t+1}\right)\right).$$

We have the following theorem, which is due to André Mundt. Here we only state it. If the reader is interested in the proof, see Proposition 5.1 in [M 07].

Theorem 4.7. *A Bayesian Pflug–Ruszczyński risk measure for* $\mu_0 = \mathcal{U}([0,1])$ *is given by the following formula:*

$$\rho_t^{\mathrm{B},\mathcal{U}([0,1])}(I) = -\frac{c_{t+1}}{c_t} W_t^+ + \sum_{n=t+1}^{T} \frac{c_n}{c_t} \mathbb{E}_{t,\alpha,\beta}\left(\rho_{t,\alpha,\beta}^{(n)}\left(I_n \middle| Y_{t+1},\dots,Y_{n-1},Z_t\right) \middle| Z_t\right), \; I \in \mathcal{X}^{\mathrm{M}},$$

where $\mathbb{E}_{t,\alpha,\beta}$ *and* $\rho_{t,\alpha,\beta}^{(n)}$ *denote the expectation and the risk measure* $\rho^{(n)}$ *with respect to the probability measure* $\mathbb{P}_{t,\alpha,\beta}$ *on* $\sigma(Z_t, Y_{t+1},\dots,Y_T)$ *defined by*

$$\begin{aligned}&\mathbb{P}_{t,\alpha,\beta}\left(Z_t = z,\ Y_{t+1} = y_{t+1},\dots,Y_T = y_T\right)\\ &\qquad = \mathbb{P}(Z_t = z)\prod_{n=t+1}^{T}\mathbb{P}\left(Y_n = y_n \middle| \vartheta = \frac{\alpha + \sum_{k=t+1}^{n-1}\mathbb{1}_{\{u\}}(y_k)}{\alpha+\beta-t-1}\right).\end{aligned}$$

Hence we have the result that is analogous to (4.5), but the probability measure that we use now is much more complicated.

We come back to the Artzner game (see Section 4.3). Again we aim at measuring the riskiness of $I^{(1)} = (0,0,0,Y_3)$ and $I^{(2)} = \left(0,0,0,\mathbb{1}_{\{Z_3\ge 2\}}\right)$. By direct computations we obtain the results presented in Table 4.2. For simplicity we assume that $\gamma_3 > 3/4$. Note that the orderings are still the same as in the cases of the Iterated Conditional Average Value at Risk and the dynamic Pflug–Ruszczyński risk measure.

t	θ_t	$\omega \in$	$I^{(1)}$		$I^{(2)}$
2	$\frac{1+Z_2}{4}$	$\{Z_2=2\}$	$-\frac{3c_4}{4c_2}$	$>$	$-\frac{c_3}{c_2}$
		$\{Z_2=1\}$	$-\frac{c_4}{2c_2}$	$=$	$-\frac{c_4}{2c_2}$
		$\{Z_2=0\}$	$-\frac{c_4}{4c_2}$	$<$	0
1	$\frac{1+Z_1}{3}$	$\{Z_1=1\}$	$-\frac{2c_4}{3c_1}$	$>$	$-\frac{2(3c_3+c_4)}{9c_1}$
		$\{Z_1=0\}$	$-\frac{c_4}{3c_1}$	$<$	$-\frac{c_4}{9c_1}$
0	$\frac{1}{2}$	Ω	$-\frac{c_4}{2}$	$>$	$-\frac{c_3+c_4}{4}$

(a) $\rho_t^{\mathrm{B},\delta_{\theta_t}}$

t	$\omega \in$	$I^{(1)}$		$I^{(2)}$
2	$\{Z_2=2\}$	$-\frac{3c_4}{4c_2}$	$>$	$-\frac{c_3}{c_2}$
	$\{Z_2=1\}$	$-\frac{c_4}{2c_2}$	$=$	$-\frac{c_4}{2c_2}$
	$\{Z_2=0\}$	$-\frac{c_4}{4c_2}$	$<$	0
1	$\{Z_1=1\}$	$-\frac{2c_4}{3c_1}$	$>$	$-\frac{4c_3+c_4}{6c_1}$
	$\{Z_1=0\}$	$-\frac{c_4}{3c_1}$	$<$	$-\frac{c_4}{6c_1}$
0	Ω	$-\frac{c_4}{2}$	$>$	$-\frac{2c_3+c_4}{6}$

(b) $\rho_t^{\mathrm{B},\mathcal{U}([0,1])}$

Table 4.2: The Bayesian Pflug–Ruszczyński measures for the Artzner game

From Table 4.2 one can easily see that

$$\begin{aligned}
&\rho_t^{\mathrm{B},\delta_{\theta_t}}\left(I^{(1)}\right) = \rho_t^{\mathrm{B},\mathcal{U}([0,1])}\left(I^{(1)}\right), \; t\in\{0,1,2\},\\
&\rho_2^{\mathrm{B},\delta_{\theta_2}}\left(I^{(2)}\right) = \rho_2^{\mathrm{B},\mathcal{U}([0,1])}\left(I^{(2)}\right),\\
&\rho_1^{\mathrm{B},\delta_{\theta_1}}\left(I^{(2)}\right) < \rho_1^{\mathrm{B},\mathcal{U}([0,1])}\left(I^{(2)}\right) \text{ on } \{Z_1=1\},\\
&\rho_1^{\mathrm{B},\delta_{\theta_1}}\left(I^{(2)}\right) > \rho_1^{\mathrm{B},\mathcal{U}([0,1])}\left(I^{(2)}\right) \text{ on } \{Z_1=0\},\\
&\rho_0^{\mathrm{B},\delta_{\theta_0}}\left(I^{(2)}\right) > \rho_0^{\mathrm{B},\mathcal{U}([0,1])}\left(I^{(2)}\right),
\end{aligned}$$

where θ_t is given by

$$\theta_t = \begin{cases} \frac{1+Z_2}{4}, & t = 2 \\ \frac{1+Z_1}{3}, & t = 1 \\ \frac{1}{2}, & t = 0. \end{cases}$$

As a consequence, if we start with no knowledge about the parameter ϑ, we attach less risk to $I^{(2)}$ using the Bayesian approach with $\mu_0 = \mathcal{U}([0,1])$ than applying the ordinary Pflug–Ruszczyński risk measure under the assumption that ϑ is known and equals $1/2$.

Appendix A

Some useful facts

In this chapter we present some useful results. For more details and proofs the reader is referred to the literature.

A.1. Probability theory

A.1.1. Quantiles and quantile functions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

First we want to define a *quantile* of a random variable.

Definition A.1. *A real number q_X is called a* α-quantile *(for $\alpha \in [0,1]$) of a random variable X if it holds that $\mathbb{P}(X < q_X) \leq \alpha \leq \mathbb{P}(X \leq q_X)$.*

Proposition A.2. *The set of all α-quantiles of X is an interval $[q_X^-(\alpha), q_X^+(\alpha)]$, where*

$$\begin{aligned} q_X^-(\alpha) &:= \sup \{x \in \mathbb{R} \mid \mathbb{P}(X < x) < \alpha\} = \inf \{x \in \mathbb{R} \mid \mathbb{P}(X \leq x) \geq \alpha\}, \\ q_X^+(\alpha) &:= \inf \{x \in \mathbb{R} \mid \mathbb{P}(X \leq x) > \alpha\} = \sup \{x \in \mathbb{R} \mid \mathbb{P}(X < x) \leq \alpha\}. \end{aligned} \tag{A.1}$$

Real numbers $q_X^-(\alpha)$ and $q_X^+(\alpha)$ are called a lower α-quantile *of X and an* upper α-quantile *of X, respectively.*

Now we go over to *quantile functions.*

Definition A.3. *Let $F\colon (a,b) \to \mathbb{R}$ be an increasing function with*

$$c = \lim_{x \searrow a} F(x), \quad d = \lim_{x \nearrow b} F(x).$$

Then $q\colon (c,d) \to (a,b)$ is called an inverse function *for F if*

$$F\left(q(s)-\right) \leq s \leq F\left(q(s)+\right) \text{ for every } s \in (c,d).$$

Definition A.4. *An inverse function $q\colon (0,1) \to \mathbb{R}$ of a distribution function F is called a* quantile function.

Lemma A.5. *Let q be an inverse function for an increasing function F. Then F is also an inverse function for q.*

See Lemma A.17 in [FS 04] for the proof.

Lemma A.6. *Suppose that U is a random variable uniformly distributed on the interval $[0,1]$ and $F\colon \mathbb{R} \to [0,1]$ is increasing and right-continuous. Let q be an inverse function for F. Then F is a cumulative distribution function for $q \circ U$.*

The proof can be found in [FS 04], see Lemma A.19.

A.1.2. Conditional probability

Suppose that $(S, \mathcal{S})$ and $(T, \mathcal{T})$ are two measurable spaces. Then we define a *probability kernel* as follows:

Definition A.7. *A mapping $P\colon S \times \mathcal{T} \to [0,1]$ is called a* probability kernel from $(S, \mathcal{S})$ to $(T, \mathcal{T})$ *if the following conditions are fulfilled:*

- *a function $P(\,\cdot\,, B)$ is $\mathcal{S}$-measurable for fixed $B \in \mathcal{T}$,*
- *a function $P(s, \,\cdot\,)$ is a probability measure for fixed $s \in S$.*

We have two very useful theorems, which we want to cite here:

Theorem A.8 (Conditional distribution). *Let $(S, \mathcal{S})$ be a Borel space and $(T, \mathcal{T})$ a measurable space. Suppose that ξ and η are random elements in S and T, respectively. Then there exists a unique $\mathbb{P} \circ \eta^{-1}$-a.s. probability kernel P from $(T, \mathcal{T})$ to $(S, \mathcal{S})$ such that*

$$\mathbb{P}(\xi \in \cdot \mid \eta) = P(\eta, \,\cdot\,).$$

For the proof see [K 97], Theorem 5.3.

Theorem A.9. *Let $\big((S_n, \mathcal{B}_{S_n})_{n \in \mathbb{N}}\big)$ be a sequence of Borel spaces. For each $n \in \mathbb{N}$ we define $T_n = \prod_{k=0}^{n} S_k$. Moreover, T stands for $\prod_{k=0}^{\infty} S_k$. Assume that μ is a probability measure on S_0 and $P_n\colon T_n \times \mathcal{B}_{S_{n+1}} \to [0,1]$ is a probability kernel from $(T_n, \mathcal{B}_{T_n})$ to $\big(S_{n+1}, \mathcal{B}_{S_{n+1}}\big)$, $n \in \mathbb{N}$. Then there exists a unique probability measure $\mathbb{P}$ on T such that*

$$\mathbb{P}\left(B_0 \times B_1 \times \cdots \times B_n\right) = \int_{B_0} \mu\left(\mathrm{d}x_0\right) \int_{B_1} P_0\left(x_0; \mathrm{d}x_1\right) \int_{B_2} P_1\left(x_0, x_1; \mathrm{d}x_2\right) \cdots \int_{B_n} P_{n-1}\left(x_0, x_1, \ldots, x_{n-1}; dx_n\right)$$

for $B_0 \times B_1 \times \cdots \times B_n \in T_n$.

The proof can be found [BS 96], see Proposition 7.28.

A.2. Functional analysis

Theorem A.10 (Separation theorem for convex sets). *Let X be a real normed linear space and $Y, Z \subset X$ be non-empty, disjoint and convex. If moreover Y is compact and Z is closed, then there exists a linear continuous functional $l\colon X \to \mathbb{R}$ such that*

$$\sup_{x \in Y} l(x) < \inf_{x \in Z} l(x).$$

Theorem A.11 (Riesz). *Let $(X, \mathcal{F}, \mu)$ be a σ-finite measurable space and p, q be such that $p \in [1, +\infty)$ and $1/p + 1/q = 1$. Suppose that $l\colon L^p(X, \mathcal{F}, \mu) \to \mathbb{R}$ is a bounded linear functional. Then there exists a unique $g \in L^q(X, \mathcal{F}, \mathbb{P})$ such that $\|g\|_q = \|L\|$ and*

$$l(f) = \int_X fg \, d\mu, \ f \in L^p(X, \mathcal{F}, \mathbb{P}).$$

For the proof see Theorem 18.6 in [Y 06].

A.3. Analysis

Theorem A.12 (Essential supremum). *Let $(X, \mathcal{F}, \mathbb{P})$ be a probability space and F be a family of random variables. Then there exists a unique random variable $g\colon \Omega \to \bar{\mathbb{R}}$ such that*

- *$g \geq f$ $\mathbb{P}$-a.s. for all $f \in F$,*
- *if $h \geq f$ $\mathbb{P}$-a.s. for a random variable h and each $f \in F$, then $h \geq g$ $\mathbb{P}$-a.s.*

We call the function g the essential supremum *of F and denote by* $\operatorname{ess\,sup} F$.

Moreover, there exists a sequence $(f_n)_{n \in \mathbb{N}} \subset F$ such that

$$\sup_{n \in \mathbb{N}} f_n = \operatorname{ess\,sup} F.$$

If additionally F is directed upwards, *i.e., for all $f_1, f_2 \in F$ there is $f_3 \in F$ such that $f_3 \geq f_1 \vee f_2$, the sequence $(f_n)_{n \in \mathbb{N}}$ can be chosen in such a way that $f_n \leq f_{n+1}$, $n \in \mathbb{N}$.*

The proof can be found in [N 75], see Proposition VI-1-1.

Appendix B

Markov decision processes

This section is devoted to basic Markov decision theory. We give an overview of some definitions and theorems that are useful in Chapter 4. For more details and proofs the reader is referred to [HL 96].

We concentrate on the case of the finite time horizon. Therefore the set of time instants, denoted by $\mathcal{T}$, equals $\{0, 1, \ldots, T\}$ for some $T \in \mathbb{N}_+$. Let $\mathcal{T}_-$ stands for $\mathcal{T} \setminus \{T\}$. We state a model as follows:

Definition B.1. *By* discrete-time Markov decision (control) model *we call a five-tuple* $\{(S, \mathcal{B}_S), (A, \mathcal{B}_A), \{A(s) \mid s \in S\}, Q_t, r_t\}$ *such that*

- state space: $(S, \mathcal{B}_S)$ *is a Borel space,*
- action (control) set: $(A, \mathcal{B}_A)$ *is a Borel space,*
- $\{A(s) \mid s \in S\}$ *is a family of nonempty measurable sets* $A(s)$ *such that* $A(s) \subset A$ *denotes the* set of feasible actions being in the state $s \in S$, *we define* the set of all feasible state-action pairs *by* $D = \{(s, a) \mid s \in S,\, a \in A(s)\}$,
- transition law: $Q_t \colon D \times \mathcal{S} \to [0, 1]$ *is a stochastic kernel from* $(D, \mathcal{B}_D)$ *to* $(S, \mathcal{B}_S)$, $t \in \mathcal{T}$,
- one-step reward function: $r_t \colon D \to \mathbb{R}$ *is measurable.*

Often instead of rewards we consider *costs*. Then almost everything remains the same, we only have to change maximizing problems to minimizing ones, etc.

We define

$$H_0 = S, \quad H_{t+1} = D \times H_t,\, t \in \mathcal{T}_-.$$

The set H_t, $t \in \mathcal{T}$, is called an *admissible history up to time* t and $h_t \in H_t$ is a *t-history*. Moreover,

$$\bar{H}_0 = H_0, \quad \bar{H}_{t+1} = S \times A \times \bar{H}_t,\, t \in \mathcal{T}_-.$$

Definition B.2. *Let $\pi = (\pi_0, \pi_1, \ldots \pi_{T-1})$ be such that $\pi_t \colon H_t \times \mathcal{B}_A \to [0,1]$ is a probability kernel from $(H_t, \mathcal{B}_{H_t})$ to $(A, \mathcal{B}_A)$ with*

$$\pi_t\left(h_t, A\left(s_t\right)\right) = 1$$

for every $t \in \mathcal{T}_-$. Then π is a policy.

We construct a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let Ω, given by $\Omega = \bar{H}_T$, be equipped with a product σ-algebra $\mathcal{F}$. Fix any probability measure μ on S and a policy $\pi = (\pi_0, \pi_1, \ldots, \pi_{T-1})$. Then, due to Theorem A.9, there exists a unique probability measure $\mathbb{P}$ such that

$$\begin{aligned}
&\mathbb{P}\left(H_T\right) = 1,\\
&\mathbb{P}\left(s_0 \in B\right) = \mu(B),\ B \in \mathcal{B}_S,\\
&\mathbb{P}\left(a_t \in C \middle| h_t\right) = \pi_t\left(h_t, C\right),\ C \in \mathcal{B}_A,\\
&\mathbb{P}\left(x_{t+1} \in B \middle| h_t,\, a_t\right) = Q_t\left(x_t, a_t, B\right),\ h_t = (x_0, a_0, x_1, a_1, \ldots, x_{t-1}, a_{t-1}, x_t) \in H_t.
\end{aligned}$$

Now we introduce a Markov decision process:

Definition B.3. $\left(\Omega, \mathcal{F}, \mathbb{P}, (X_t)_{t\in\mathcal{T}}\right)$ *is called a* Markov decision process (Markov control process).

Note that transition law is sometimes defined by means of the following equations

$$\begin{aligned}
x_0 &= s_0,\ s_0 \in S,\\
x_{t+1} &= F\left(x_t, a_t, \xi_t\right),\ t \in \mathcal{T}_-
\end{aligned}$$

where $(\xi_t)_{t\in\mathcal{T}}$ is a sequence of independent random variables that are independent of the initial state s_0 as well. Then each ξ_t, $t \in \mathcal{T}$, is called a *disturbance*.

Suppose that $(S, A, \{A(s) \mid s \in S\}, Q_t, r_t)$ is a given Markov decision model. The aim is to maximize the function J given by

$$J\left(\pi, s\right) = \mathbb{E}\left(\sum_{t=0}^{T-1} r_t\left(x_t, a_t\right) + c_T\left(x_T\right)\right),$$

where $c_T \colon S \to \mathbb{R}$ is a measurable function called a *terminal reward function*. We also define a *value function* via

$$J^*(s) = \inf_{\pi} J\left(\pi, s\right).$$

We want to find a policy π^* with

$$J\left(\pi^*, s\right) = J^*(s),\ s \in S.$$

Such a policy is the so-called *optimal policy* and the following theorem enables us to find it.

Theorem B.4 (Dynamic programming theorem). *Let $(J_t)_{t\in\mathcal{T}}$ be a sequence of functions defined backwards by*

$$
\begin{aligned}
&J_T(s) = c_T(s), \\
&J_t(s) = \max_{a\in A(s)} \left(r(s,a) + \int_S J_{t+1}(y)\, Q(s,a;dy) \right), \; t \in \{0,1,\ldots,T-1\}\,.
\end{aligned}
\tag{B.1}
$$

If for each $t \in \mathcal{T}$ J_t is measurable and there exists a function $\pi_t^*\colon S \to A$ such that π_t^* is a maximizer of (B.1) *and $\pi_t^*(s) \in A(s)$ for all $s \in S$, then the policy $\pi^* = \left(\pi_0^*, \ldots, \pi_{T-1}^*\right)$ is optimal and*

$$J^*(s) = J\left(\pi^*, s\right),\; s \in S.$$

For the proof see Theorem 3.2.1 in [HL 96].

Bibliography

[ADEH 99] Artzner, P., F. Delbaen, J. M. Eber, D. Heath, *Coherent measures of risk*, Mathematical Finance, 9(3), 203–228 (1999)

[ADEHK 07] Artzner, P., F. Delbaen, J. M. Eber, D. Heath, H. Ku, *Coherent multiperiod risk adjusted values and Bellman's principle*, Annals of Operations Research, 152 (1), 5–22 (2007)

[BS 96] Bertsekas, D. P., S. E. Shreve, *Stochastic Optimal Control: The Discrete-Time Case*, Athena Scientific (1996)

[B 05] Burgert C., *Darstellungssätze für statische und dynamische Risikomaße mit Anwendungen*, PhD thesis (2005)

[D 02] Delbaen, F., *Coherent Risk Measures on General Probability Spaces*, Advances in Finance and Stochastics: Essays in Honour of Dieter Sondermann, Springer-Verlag, 1–37 (2002)

[FS 04] Föllmer, H., A. Schied, *Stochastic Finance: An Introduction in Discrete Time*, Second Edition, Walter de Gruyter (2004)

[HW 04] Hardy, M. R., J. L. Wirch, *The Iterated CTE: A Dynamic Risk Measure*, North American Actuarial Journal, 8 (4), 62–75 (2004)

[HL 96] Hernández-Lerma O., J. B. Lasserre, *Discrete-Time Markov Control Processes: Basic Optimality Criteria*, Springer-Verlag (1996)

[I 03] Inoue, A., *On the worst conditional expectation*, Journal of Mathematical Analysis and Applications, 286 (1), 237–247 (2003)

[K 97] Kallenberg O., *Foundations of Modern Probability*, Springer-Verlag (1997)

[M 07] Mundt A. P., *Dynamic risk management with Markov decision processes*, PhD thesis (2007)

[N 75] Neveu J., *Discrete-parameter martingales*, North-Holland (1975)

[PR 05] Pflug, G. Ch., A. Ruszczyński, *Measuring Risk for Income Streams*, Computational Optimization and Applications, 32 (1–2), 161–178 (2005)

[Y 06] Yeh J., *Real Analysis: Theory of Measure and Integration*, Second Edition, World Scientific (2006)

[YZ 99] Yong J., X. Y. Zhou, *Stochastic controls: Hamiltonian Systems and HJB Equations*, Springer-Verlag (1999)