Risk measures for income streams

Georg Ch. Pflug [1] and Andrzej Ruszczynski [2]

Abstract.

A new measure of risk is introduced for a sequence of random incomes adapted to some filtration. This measure is formulated as the optimal net present value of a stream of adaptively planned commitments for consumption.

The calculation of the new measure is done by solving a stochastic dynamic linear optimization problem, which, in case of a finite filtration, reduces to a simple deterministic linear program.

We show properties of the new measure by exploiting the convexity and duality structure of the stochastic dynamic linear problem. The measure depends on the full distribution of the income process (not only on its marginal distributions) as well as on the filtration, which is interpreted as the available information about the future.

1 Introduction: The one-period case

Let I be a random income variable defined on some probability space $(\Omega, \mathcal{F}_I, P)$. The risk contained in I is caused by the lack of information about its exact value. A variable, but predictable value of I is riskless. If a natural catastrophe, e.g. a flood, were completely predictable, there would be no risk and no company would insure against it.

If a decision maker were clairvoyant, he/she would face no risk since he/she would see the future in a deterministic way and would be able to adapt to it. For us, normal humans, some but not all information about the future may be available. The amount of information available may be expressed in terms of some σ-algebra $\mathcal{F} \subseteq \mathcal{F}_I$. The extreme cases are the clairvoyant ($\mathcal{F} = \mathcal{F}_I$) and the totally uninformed ($\mathcal{F} = \mathcal{F}_0 = \{\Omega, \emptyset\}$).

[1] Department of Statistics and Decision Support Systems, University of Vienna, 1090 Wien-Vienna, Universitaetsstrasse 5; e-mail: georg.pflug@univie.ac.at

[2] Department of Management Science and Information Systems and RUTCOR, Rutgers University, Piscataway, NJ 08854; e-mail rusz@rutcor.rutgers.edu

The ultimate goal of engaging in risky entrprises with uncertain income opportunities is consumption. Consumption, however, can only be realized after deciding about the amount one wants to commit for this purpose (to buy a house, a car etc.).

Suppose that the decision maker decides to commit an amount a. In this case, he/she risks not achieving this decided target, since I may be less than a. However, he/she may insure against the shortfall event, i.e. the event that $I < a$. Insurance comes at a price of $\mathbb{E}(q[I-a]^-)$, for $q > 1$.[3] The costs for insurance decrease the possible consumption.

If, on the other hand, some surplus is left after consumption, this surplus is discounted by a factor $d < 1$, since saving does not provide the same satisfaction as the consumption committed for.

The expected net present value (ENPV) of the consumption and savings is therefore

$$\mathbb{E}(a + d[I-a]^+ - q[I-a]^-).$$

A rational decision maker maximizes the ENPV w.r.t. the available information $\mathcal{F}$; i.e. his/her utility functional is

$$\mathcal{U}_{\mathcal{F}}(I) = \max\{\mathbb{E}(a + d[I-a]^+ - q[I-a]^-) : a \text{ is } \mathcal{F} \text{ measurable } \}. \qquad (1)$$

It is evident that $\mathcal{F}_1 \subseteq \mathcal{F}_2$ implies that $\mathcal{U}_{\mathcal{F}_1}(I) \leq \mathcal{U}_{\mathcal{F}_2}(I)$, i.e., more information gives more utility.

Since I was supposed to be $\mathcal{F}_I$-measurable and

$$a + d[v-a]^+ - q[v-a]^- = v - (1-d)[v-a]^+ - (q-1)[v-a]^-,$$

one sees that

$$\mathcal{U}_{\mathcal{F}_I}(I) = \mathbb{E}(I)$$

and

$$\mathcal{U}_{\mathcal{F}}(I) \leq \mathbb{E}(I)$$

for any other sub σ-algebra $\mathcal{F}$ of $\mathcal{F}_I$.

The risk $\mathcal{R}$ contained in the random variable I and the information $\mathcal{F}$ is defined as the difference between the maximal utility (the utility of the clairvoyant) and the actual utility.

$$\mathcal{R}_{\mathcal{F}}(I) = \mathbb{E}(I) - \mathcal{U}_{\mathcal{F}}(I). \qquad (2)$$

Necessarily, $\mathcal{R}_{\mathcal{F}}(I) \geq 0$. Evidently, the clairvoyant has no risk and the totally uninformed has the maximal risk in this setup. The risk of the latter

[3] We use the notation $[x]^+ = \max(x, 0)$ and $[x]^- = \max(-x, 0)$.

is connected to the notion of the conditional-value-at risk ($\mathbb{CV@R}$). Recall that the $\mathbb{CV@R}$ is defined as

$$\mathbb{CV@R}_\alpha(I) = \max\{a - \frac{1}{\alpha}\mathbb{E}([I-a]^-) : a \in \mathbb{R}\}$$

(see Rockefellar and Uryasev [6]). It is known that

$$\begin{aligned}\mathbb{CV@R}_\alpha(I) &= \frac{1}{\alpha}\int_0^\alpha G^{-1}(p)\,dp \\ &= \mathbb{E}(I|I \le G^{-1}(\alpha)) - \left(\frac{G(G^{-1}(\alpha)) - \alpha}{\alpha}\right) G^{-1}(\alpha),\end{aligned}$$

where $G(u) = \mathbb{P}\{I \le u\}$ and $G^{-1}(p) = \inf\{u : G(u) \ge p\}$ (see [4]).

Lemma 1 *For the totally uninformed, i.e.* $\mathcal{F}_0 = \{\Omega, \emptyset\}$*, we have*

$$\mathcal{U}_{\mathcal{F}_0}(I) = d\mathbb{E}(I) + (1-d)\mathbb{CV}@R_\alpha(I)$$

and

$$\mathcal{R}_{\mathcal{F}_0}(I) = (1-d)[\mathbb{E}(I) - \mathbb{CV}@R_\alpha(I)],$$

where $\alpha = (1-d)/(q-d)$*.*

Proof. Take a closer look at the function

$$\begin{aligned}U_a(v) &= a + dv - a]^+ - q[v-a]^- \\ &= a + d(v-a) + d[v-a]^- - q[v-a]^- \\ &= dv + (1-d)a - (q-d)[v-a]^-. \end{aligned} \quad (3)$$

Using (3), we find that

$$\begin{aligned}\mathcal{U}_{\mathcal{F}_0}(I) &= d\mathbb{E}(I) + (1-d)\max\left\{\mathbb{E}\Big(a + \frac{q-d}{1-d}[I-a]^-\Big) : a \in \mathbb{R}\right\} \\ &= d\mathbb{E}(I) + (1-d)\mathbb{CV@R}_\alpha(I)],\end{aligned}$$

with $\alpha = (1-d)/(q-d)$.

□

Notice that $v \mapsto U_a(v)$ is a concave, monotonic utility function for every fixed a. Recall the following orderings for random variables.

Definition 1 Let $I^{(1)}$ and $I^{(2)}$ two random income variables.

- We say that first order stochastic dominance ($I^{(1)} \prec_{FSD} I^{(2)}$) holds, if $\mathbb{E}[U(I^{(1)}] \le \mathbb{E}[U(I^{(2)})]$ for all monotonic, integrable functions U.

- We say that second order stochastic dominance ($I^{(1)} \prec_{SSD} I^{(2)}$) holds, if $\mathbb{E}[U(I^{(1)}] \leq \mathbb{E}[U(I^{(2)})]$ for all monotonic and concave, integrable functions U.
- We say that concave dominance ($I^{(1)} \prec_{CC} I^{(2)}$) holds, if $\mathbb{E}[U(I^{(1)}] \leq \mathbb{E}[U(I^{(2)})]$ for all concave integrable functions U.

Obviously, since all U_a are monotonic and concave by (3), it follows that

$$\mathcal{U}_{\mathcal{F}_0}(I) = \max\{\mathbb{E}(U_a(I) : a \in \mathbb{R}\}$$

is monotonic w.r.t. second order stochastic dominance $\prec_{SSD}$ and a fortiori with first order stochastic dominance $\prec_{FSD}$ and concave dominance $\prec_{CC}$. By a similar argument, $\mathcal{R}_{\mathcal{F}_0}$ is antitonic w.r.t. $\prec_{CC}$.

More generally, if $I^{(1)}$ and $I^{(2)}$ are defined on the same probability space, and all the conditional distributions satisfy $(I^{(1)}|\mathcal{F}) \prec_{SSD} (I^{(2)}|\mathcal{F})$, then $\mathcal{U}_{\mathcal{F}}(I^{(1)}) \leq \mathcal{U}_{\mathcal{F}}(I^{(2)})$. Similarly, if $(I^{(1)}|\mathcal{F}) \prec_{CC} (I^{(2)}|\mathcal{F})$, then $\mathcal{R}_{\mathcal{F}}(I^{(1)}) \geq \mathcal{R}_{\mathcal{F}}(I^{(2)})$.

It is necessary to require the ordering of all conditional distributions.

Example 1 Let the probability space have three points, ω_1, ω_2, ω_3, each having probability 1/3. Let $I^{(1)}(\omega_1) = 1.01$, $I^{(1)}(\omega_2) = 1.015$, $I^{(1)}(\omega_3) = 1.03$; $I^{(2)}(\omega_1) = 1.01501$, $I^{(2)}(\omega_2) = 1.0301$, $I^{(2)}(\omega_3) = 1.0101$. Choose $q = 1.2$, $d = 0.93$ and $\mathcal{F} = \{\{\omega_1, \omega_2\}, \omega_3\}$. Then

$$\mathcal{U}(I^{(1)}) = 1.0175 > \mathcal{U}(I^{(2)}) = 1.0174,$$

but $I^{(1)} \prec_{SSD} I^{(2)}$.

Notice that $\mathcal{U}_{\mathcal{F}}$ is translation-equivariant, i.e. for all constant b

$$\mathcal{U}_{\mathcal{F}}(I + b) = \mathcal{U}_{\mathcal{F}}(I) + b. \tag{4}$$

This follows directly from the definition.

In contrast, $\mathcal{R}_{\mathcal{F}}(I)$ is translation-invariant, i.e. for all constant b

$$\mathcal{R}_{\mathcal{F}}(I + b) = \mathcal{R}_{\mathcal{F}}(I). \tag{5}$$

Since $U_{\lambda a}(\lambda v) = \lambda U_a(v)$, $\mathcal{U}_{\mathcal{F}}$ and $\mathcal{R}_{\mathcal{F}}$ are (positively) homogeneous, i.e.

$$\mathcal{U}_{\mathcal{F}}(\lambda I) = \lambda \mathcal{U}_{\mathcal{F}}(I)$$

$$\mathcal{R}_{\mathcal{F}}(\lambda I) = \lambda \mathcal{R}_{\mathcal{F}}(I).$$

$\mathcal{U}$ is concave and $\mathcal{R}$ is convex in the following sense: If I_1 and I_2 are two income variables (they may be dependent), then

$$\mathcal{U}_{\mathcal{F}}(p + (1-p)I_2) \geq p\mathcal{U}_{\mathcal{F}}(I_1) + (1-p)\mathcal{U}_{\mathcal{F}}(I_2) \tag{6}$$

and

$$\mathcal{R}_{\mathcal{F}}(pI_1 + (1-p)I_2) \leq p\mathcal{R}_{\mathcal{F}}c(I_1) + (1-p)\mathcal{R}_{\mathcal{F}}(I_2). \tag{7}$$

The proof of (6) goes as follows: Suppose that $\mathcal{U}_{\mathcal{F}}(I_1) = \mathbb{E}(U_{a_1}(I_1))$, $\mathcal{U}_{\mathcal{F}}(I_2) = \mathbb{E}(U_{a_2}(I_2))$, then, using (3),

$$\begin{aligned} \mathcal{U}_{\mathcal{F}}(pI_1 + (1-p)I_2) &\geq \mathbb{E}(U_{pa_1+(1-p)a_2}(pI_1 + (1-p)I_2) \\ &\geq \mathbb{E}pU_{a_1}(I_1) + (1-p)U_{a_2}(I_2)) \\ &= p\mathcal{U}_{\mathcal{F}}(I_1) + (1-p)\mathcal{U}_{\mathcal{F}}(I_2). \end{aligned}$$

(7) is easily deduced from that.

If we compound I_1 and I_2 with probability p, i.e.,

$$I = \begin{cases} I_1 & \text{with probability } p, \\ I_2 & \text{with probability } 1-p, \end{cases}$$

$\mathbb{E}(U_a(I)) = p\mathbb{E}(U_a(I_1)) + (1-p)\mathbb{E}(U_a(I_2))$ and therefore

$$\mathcal{U}_{\mathcal{F}}(I) \leq p\mathcal{U}_{\mathcal{F}}(I_1) + (1-p)\mathcal{U}_{\mathcal{F}}(I_2).$$

Artzner, Delbaen, Eber and Heath [1] have introduced the notion of a coherent risk measure as a measure being translation-equivariant (they call it translation-invariant), positive homogeneos, convex in the sense of (7) and monotonic w.r.t. pointwise ordering. Thus $-\mathcal{U}_{\mathcal{F}}$ is a coherent risk measure in the sense of [1], but $\mathcal{R}_{\mathcal{F}}$ is not since it is translation invariant in the sense of (5).

2 Risk of multiperiod income streams

Suppose now that $I_1, I_2, \ldots, I_T$ is a stream of random incomes which arrive at times $1, 2, \ldots, T$. We denote by $(\Omega, \mathcal{F}.\mathbb{P})$ the probability space on which these random variables are defined. Together with that, a filtration $\{\mathcal{F}_t\}$, $t = 1, \ldots, T$, is defined, so that I_t is $\mathcal{F}_t$-measurable, for each $t = 1, \ldots, T$. The σ-subfield $\mathcal{F}_t$ represents the information available at time t. We take the convention that $\mathcal{F}_0 = \{\emptyset, \Omega\}$.

Analogously to the static case, let a_t be the amount to be consumed at time t. The decision about a_t must be made at time $t-1$, so a_t must be $\mathcal{F}_{t-1}$-measurable. The consumption of one unit at time t gives a NPV of $c_t \geq 0$. The shortfall costs are $q_t \geq 0$. The expected shortfall costs are immediately subtracted from the consumption before period t (this can be interpreted as an insurance cost). Any surplus occuring in period t increases the income of the next period. The final surplus is discounted by a factor $d \geq 0$. We make the following assumptions about the sequences $\{c_t\}$, $\{q_t\}$ and the constant d:

$$\begin{aligned} c_t &\leq q_t, \quad t = 1, \ldots, T, \\ c_{t+1} &\leq c_t, \quad t = 1, \ldots, T-1, \\ d &\leq c_T. \end{aligned} \tag{8}$$

Let K_t be the (random) surplus carried from period t to period $t+1$. We have $K_0 = 0$ and

$$K_t = [K_{t-1} + I_t - a_t]^+, \quad t = 1, \ldots, T. \tag{9}$$

The shortfall M_t at period t is given by

$$M_t = [K_{t-1} + I_t - a_t]^-. \tag{10}$$

Our objective is to maximize the expected consumption minus the expected shortfall costs. This can be written as the following optimization problem:

$$\mathcal{U}(I_1, I_2, \ldots, I_T) = \max \mathbb{E}\Big[\sum_{t=1}^{T} (c_t a_t - q_t M_t) + d K_T \Big] \tag{11}$$

$$\text{s.t. } a_t \text{ is } \mathcal{F}_{t-1}\text{-measurable for } t = 1, \ldots, T. \tag{12}$$

We introduce the dynamic risk measure of the sequence $\{I_t\}$ as

$$\mathcal{R}(I_1, \ldots, I_T) = \mathcal{U}(\mathbb{E}I_1, \ldots, \mathbb{E}I_T) - \mathcal{U}(I_1, \ldots, I_T). \tag{13}$$

We shall prove in the next section that it is always non-negative, and that it posesses most of the properties of the risk measure in the static case.

In order to analyze problem (11)–(12) we shall formalize it as a stochastic control problem. We denote by $\mathcal{X}_t$ the space of $\mathcal{F}_t$-measurable random variables having a finite expected value: $\mathcal{X}_t = \mathcal{L}^1(\Omega, \mathcal{F}_t, \mathbb{P})$. We also use the notation $\mathbb{E}_t\{\cdot\}$ for $\mathbb{E}\{\cdot|\mathcal{F}_t\}$.

Problem (11)–(12) can be now written as follows: find random variables $a_t \in \mathcal{X}_{t-1}$, $M_t \in \mathcal{X}_t$, and $K_t \in \mathcal{X}_t$, $t = 1, \ldots, T$, so as to

$$\max\ \mathbb{E}\Big[\sum_{t=1}^{T}(c_t a_t - q_t M_t) + dK_T\Big] \tag{14}$$

$$\text{s.t. } K_t = K_{t-1} + I_t - a_t + M_t, \quad t = 1, \ldots, T, \tag{15}$$

$$K_t \geq 0, \quad M_t \geq 0, \quad t = 1, \ldots, T, \tag{16}$$

where $K_0 = 0$ and the constraints (15)–(16) are understood in the 'almost sure' sense.

We can view (14)–(16) as a linear programming problem in abstract spaces. Let us introduce Lagrange multipliers $\lambda_t \in \mathcal{L}^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$ associated with the constraints (15), $t = 1, \ldots, T$. The lagrangian takes on the form

$$L(a, M, K, \lambda) = \mathbb{E}\sum_{t=1}^{T}(c_t a_t - q_t M_t) + d\mathbb{E}K_T \tag{17}$$

$$- \mathbb{E}\sum_{t=1}^{T}\lambda_t(K_t - K_{t-1} - I_t + a_t - M_t). \tag{18}$$

The dual functional is defined as

$$D(\lambda) = \sup_{(a,M,K)\in X_0} L(K, a, M, \lambda),$$

where

$$X_0 = \{(a, M, K) : a_t \in \mathcal{X}_{t-1},\ M_t \in \mathcal{X}_t,\ M_t \geq 0,\ K_t \in \mathcal{X}_t,\ K_t \geq 0,\ t = 1, \ldots, T\}. \tag{19}$$

We have

$$\begin{aligned} L(a, M, K, \lambda) &= \mathbb{E}\sum_{t=1}^{T}(c_t - \lambda_t)a_t + \mathbb{E}\sum_{t=1}^{T}(\lambda_t - q_t)M_t + \mathbb{E}\,(d - \lambda_T)K_T \\ &+ \mathbb{E}\sum_{t=1}^{T-1}(\lambda_{t+1} + \lambda_t)K_t + \mathbb{E}\sum_{t=1}^{T}\lambda_t I_t \\ &= \mathbb{E}\sum_{t=1}^{T}(c_t - \mathbb{E}_{t-1}\lambda_t)a_t + \mathbb{E}\sum_{t=1}^{T}(\lambda_t - q_t)M_t + \mathbb{E}\,(d - \lambda_T)K_T \\ &+ \mathbb{E}\sum_{t=1}^{T-1}(\mathbb{E}_t\lambda_{t+1} - \lambda_t)K_t + \mathbb{E}\sum_{t=1}^{T}\lambda_t I_t, \end{aligned}$$

where we have manipulated (by conditioning) the coefficients in front of a_t, M_t and K_t to obtain elements of the corresponding dual spaces $\mathcal{L}^\infty(\Omega, \mathcal{F}_{t-1}, \mathbb{P})$, $\mathcal{L}^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$, and $\mathcal{L}^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$. It follows that $D(\lambda) < +\infty$ if and only if the following conditions are satisfied:

$$\mathbb{E}_{t-1}\lambda_t = c_t, \quad t = 1, \ldots, T, \tag{20}$$
$$\lambda_t \leq q_t, \quad t = 1, \ldots, T, \tag{21}$$
$$\lambda_T \geq d, \tag{22}$$
$$\lambda_t \geq \mathbb{E}_t \lambda_{t+1}, \quad t = 1, \ldots, T-1, \tag{23}$$

and the dual problem is to find

$$\min\ \mathbb{E} \sum_{t=1}^{T} \lambda_t I_t \tag{24}$$

subject to (20)–(23).

Kuhn–Tucker optimality conditions and duality relations hold for our model (14)–(16), similarly to the finite-dimensional case.

Theorem 1 *The processes $\hat{a}_t$, $\hat{M}_t$, and $\hat{K}_t$, $t = 1, \ldots, T$, constitute an optimal solution of (14)–(16) if and only if there exists multipliers $\hat{\lambda}_t \in \mathcal{L}^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$, $t = 1, \ldots, T$, such that conditions (20)–(23) are satisfied together with the complementary slackness conditions (understood in the 'almost sure' sense):*

$$\hat{M}_t(q_t - \hat{\lambda}_t) = 0, \quad t = 1, \ldots, T, \tag{25}$$
$$\hat{K}_T(\hat{\lambda}_T - d) = 0, \tag{26}$$
$$\hat{K}_t(\hat{\lambda}_t - \mathbb{E}_t \hat{\lambda}_{t+1}) = 0, \quad t = 1, \ldots, T-1. \tag{27}$$

Proof. Consider the affine operator $G = (G_1, \ldots, G_T)$ involved in (15):

$$G_t(a, M, K) = K_t - K_{t-1} - I_t + a_t - M_t. \quad t = 1, \ldots, T.$$

We treat it as an operator from the space on which (a, M, K) are defined (the product of the corresponding $\mathcal{L}^1$ spaces) to $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_T$. Since the image of the set (19) under G contains a neighborhood of 0 in $\mathcal{X}$, our result follows from [2, Thm. 4, §1.1]. □

Theorem 2 *Suppose that conditions (8) hold. Then for every sequence $I_1, \ldots, I_T$ such that $\mathbb{E}|I_t| < +\infty$, $t = 1, \ldots, T$, the optimal values of problems (14)–(16) and (20)–(24) are finite and equal.*

Proof. A feasible solution to the primal problem (14)–(16) is given by $a_t = \mathbb{E}I_t$, with the other variables determined by (9)–(10). The objective value at this point provides a lower bound for the optimal value of the dual problem. The feasible set of the dual problem, given by (20)–(23), is convex, closed and bounded in $\mathcal{L}^\infty(\Omega, \mathcal{F}_1, \mathbb{P}) \times \cdots \times \mathcal{L}^\infty(\Omega, \mathcal{F}_T, \mathbb{P})$. Hence, it is weakly* compact (Alaoglu theorem, see[3, Thm. 6, p. 179]). Therefore the dual problem has an optimal solution, $\hat{\lambda}$. Then every solution (a, M, K) of the conditions (25)–(27) which satisfies equation (15) is, by Theorem 1, an optimal solution of the primal problem. Such a solution exists, because we can determine K and M from (25)–(27), and then choose a (which is not constrained) to ensure (15). □

It is clear that the optimal Lagrange multipliers $\hat{\lambda}_t(\omega)$ can be interpreted as the (random) costs of a unit of a credit at time t and scenario ω. With such costs it is not profitable to borrow and to lend at each time t.

3 Properties of the dynamic risk measure

Trivially, the functionals $\mathcal{U}$ and $\mathcal{R}$ are homogeneous. $\mathcal{U}$ is monotonic in the following sense: If two income processes $(I_t^{(1)})$ and $(I_t^{(2)})$ are defined on the same probability space $(\Omega, (\mathcal{F}_t), \mathbb{P})$ with the same filtration $\mathcal{F}_t$ and if $I_t^{(1)} \le I_t^{(2)}$ a.s. for all t, then $\mathcal{U}(I_1, \dots, I_T) \le \mathcal{U}(I_1, \dots, I_T)$. More generally, if all conditional distributions $(I_t | \mathcal{F}_{t-1})$ satisfy $I_t^{(1)} | \mathcal{F}_{t-1} \prec_{SSD} I_t^{(2)} | \mathcal{F}_{t-1}$, then $\mathcal{U}(I_1, \dots, I_T) \le \mathcal{U}(I_1, \dots, I_T)$.

Finally, $\mathcal{U}$ is translation equivariant in the following sense:

$$\mathcal{U}(I_1 + b_1, \dots, I_T + b_T) = \mathcal{U}(I_1, \dots, I_T) + c_1 b_1 + c_2 b_2 + \dots c_T b_T,$$

where $b_1, \dots, b_T$ are constants. We shall also show in this section that $\mathcal{U}$ is concave, so it makes sense to call $-\mathcal{U}$ coherent in the sense of [1].

Let us start from the following observation.

Lemma 2 *Suppose that conditions (8) hold and that that each I_t it $\mathcal{F}_{t-1}$-measurable and integrable, $t = 1, \dots, T$. Then*

$$\mathcal{U}(I_1, \dots, I_t) = \sum_{t=1}^{T} c_t \mathbb{E}\{I_t\}.$$

Proof. The solution

$$a_t = I_t, \quad M_t = 0, \quad K_t = 0, \quad t = 1, \dots, T,$$

is feasible for the primal problem (14)–(16), while the solution

$$\lambda_t = c_t, \quad t = 1, \ldots, T, \tag{28}$$

is feasible for the dual problem (20)–(24). Both have the same objective values, $\sum_{t=1}^{T} c_t \mathbb{E}\{I_t\}$, and, by virtue of Theorem 2, they are optimal for their problems. □

As a conclusion from this result we obtain a basic property of our risk measure.

Theorem 3 *Suppose that conditions (8) hold. Then for every sequence $I_1, \ldots, I_T$ such that $\mathbb{E}|I_t| < +\infty$, $t = 1, \ldots, T$, the risk measure (13) is finite and non-negative.*

Proof. Under conditions (8) the deterministic solution (28) is feasible for (20)–(24). Since a feasible solution for a dual problem always provides an upper bound for the primal problem, for every sequence $I_1, \ldots, I_T$ such that $\mathbb{E}|I_t| < +\infty$, $t = 1, \ldots, T$, we have

$$\mathcal{U}(I) \le D(c) = \sum_{t=1}^{T} c_t \mathbb{E}\{I_t\} = \mathcal{U}(\mathbb{E}\{I\}),$$

where the last equality follows from Lemma 2. □

Theorem 4 *Let $\mathcal{B}_t$, $t = 1, \ldots, T$, be σ-subalgebras such that $\mathcal{F}_{t-1} \subseteq \mathcal{B}_t \subseteq \mathcal{F}_t$, $t = 1, \ldots, T$. Then for every sequence $I, \ldots, I_T$, with $\mathbb{E}|I_t| < \infty$ we have*

$$\mathcal{R}(\mathbb{E}\{I_1|\mathcal{B}_1\}, \ldots, \mathbb{E}\{I_T|\mathcal{B}_T\}) \le \mathcal{R}(I_1, \ldots, I_t). \tag{29}$$

Proof. By theorem 2 both $\mathcal{U}(I_1, \ldots, I_t)$ and $\mathcal{U}(\mathbb{E}\{I_1|\mathcal{B}_1\}, \ldots, \mathbb{E}\{I_T|\mathcal{B}_T\})$ are finite. Let $\hat{\mu}_t$, $t = 1, \ldots, T$, be the optimal solution of the dual problem (20)–(24) with the income stream $\mathbb{E}\{I_t|\mathcal{B}_t\}$, $t = 1, \ldots, T$. Then the multipliers $\bar{\mu}_t = \mathbb{E}\{\mu_t|\mathcal{B}_t\}$, $t = 1, \ldots, T$, are also optimal solutions of this problem. Indeed, the feasibility follows from

$$\mathbb{E}_t \bar{\mu}_t = \mathbb{E}_t \mathbb{E}\{\mu_t|\mathcal{B}_t\} = \mathbb{E}_t \mu_t = c_t, \quad t = 1, \ldots, T,$$

and the optimality is guaranteed by

$$\mathbb{E} \sum_{t=1}^{T} \mu_t \mathbb{E}\{I_t|\mathcal{B}_t\} = \mathbb{E} \sum_{t=1}^{T} \bar{\mu}_t \mathbb{E}\{I_t|\mathcal{B}_t\}.$$

The multipliers $\bar{\mu}_t$ are also feasible for (20)–(24) with the income stream I_t, $t = 1, \dots, T$. Therefore,

$$\mathcal{U}(I_1, \dots, I_t) \le \mathbb{E} \sum_{t=1}^{T} \bar{\mu}_t I_t = \mathbb{E} \sum_{t=1}^{T} \bar{\mu}_t \mathbb{E}\{I_t | \mathcal{B}_t\}.$$

Combining the last two relations and using (13) we obtain the required result. □

A simple interpretation of Theorem 4 is that the additional information, represented by $\mathcal{B}_t$, reduces risk. In particular, if each I_t becomes known at the preceding period, there is no risk at all, as we have shown it in Lemma 2.

Also, combining two income streams cannot increase risk.

Theorem 5 *Let $I = (I_1, \dots, I_T)$ and $J = (J_1, \dots, J_T)$ be two streams of integrable incomes. Then for every $\gamma \in (0,1)$*

$$\mathcal{R}(\gamma I + (1-\gamma)J) \le \gamma \mathcal{R}(I) + (1-\gamma)\mathcal{R}(J),$$

that is, the functional $\mathcal{R}(\cdot)$ is convex.

Proof. The result follows from Theorem 2. Let us denote by Λ the set of multipliers defined by (20)–(23). We have

$$\begin{aligned}
\mathcal{U}(\gamma I + (1-\gamma)J) &= \min_{\lambda \in \Lambda} \Big[\gamma \mathbb{E} \sum_{t=1}^{T} \lambda_t I_t + (1-\gamma)\mathbb{E} \sum_{t=1}^{T} \lambda_t J_t \Big] \\
&\ge \gamma \min_{\lambda \in \Lambda} \; \mathbb{E} \sum_{t=1}^{T} \lambda_t I_t + (1-\gamma) \min_{\lambda \in \Lambda} \; \mathbb{E} \sum_{t=1}^{T} \lambda_t J_t \\
&= \gamma \mathcal{U}(I) + (1-\gamma)\mathcal{U}(J).
\end{aligned}$$

Since Lemma 2 implies that

$$\mathcal{U}(\gamma \mathbb{E} I + (1-\gamma)\mathbb{E} J) = \gamma \mathcal{U}(\mathbb{E} I) + (1-\gamma)\mathcal{U}(\mathbb{E} J),$$

our result follows. □

4 Finite filtrations

Let us consider in more detail the case when the filtration $\mathcal{F} = (\mathcal{F}_1, \dots, \mathcal{F}_T)$ is finite. This filtration generates partitions of the probability space Ω, which may be represented by a rooted tree of height T. Each node of the tree at layer t stands for an atom of the σ-algebra $\mathcal{F}_t$. Subtrees represent subpartitions.

Suppose that the nodes of this tree are numbered $\{0, 1, 2, \ldots N\}$, with 0 being the root. Let

$$\mathcal{N} = \{1, 2, \ldots, N\}$$

be the node set (not including the root). We assume that there are $N_0 - 1$ nonterminal nodes in $\mathcal{N}$ and that

$$\mathcal{T} = \{N_0, \ldots, N\}$$

is the set of terminal nodes. If n is a node in $\mathcal{N}$, then $n-$ denotes its predecessor and $t(n)$ denotes its time stage (its distance from the root).

The nodes of the tree are marked by the probabilities of the corresponding elements of the partitions. Evidently, such a tree represents the filtered probability space $(\Omega, (\mathcal{F}_t)_{t=1,\ldots,T}, P)$.

An income stream $I = (I_t)$, which is adapted to the filtration $\mathcal{F}_t$ assigns values I_n to each node $n \in \mathcal{N}$.

We call such a valuated tree an *income stream tree*.

The commitment decisions are made at the nonterminal nodes (including the root), i.e. a is a vector of length N_0 with components $a_0, \ldots, a_{N_0-1}$.

The calculation of the dynamic utility functional $\mathcal{U}_{\mathcal{F}}$ turns out to be a standard linear program defined on income stream trees. It reads

$$\begin{aligned}
\max_{a,M,K} \ & \sum_{n\in\mathcal{N}} p_n c_{t(n)} a_{n-} - \sum_{n\in\mathcal{N}} p_n q_{t(n)} M_n + \sum_{n\in\mathcal{T}} p_n d K_n \\
\text{s.t. } & K_n + a_{n-} - M_n = I_n, \quad t(n) = 1, \quad n \in \mathcal{N}, \\
& K_n - K_{n-} + a_{n-} - M_n = I_n, \quad t(n) > 1, \quad n \in \mathcal{N}, \\
& M_n \geq 0, \quad K_n \geq 0 \quad n \in \mathcal{N}, \\
& a_n \geq 0, \quad n \in (\mathcal{N} \backslash \mathcal{T}) \cup \{0\}.
\end{aligned} \tag{30}$$

This linear program has $N_0 + 2N$ nonnegative variables and N equality constraints. Its optimal value is $\mathcal{U}(I)$. The risk is defined as

$$\mathcal{R}(I) = \sum_{n\in\mathcal{N}} c_{t(n)} p_n I_n - \mathcal{U}(I).$$

Let (z_n) be the vector of dual variables of (30). We introduce the notation $n+$ for the set of all successors of the node $n \in \mathcal{N} \backslash \mathcal{T}$. Setting $z_n = p_n y_n$, the dual has the following form:

$$
\begin{aligned}
\min_y \quad & \sum_{n\in\mathcal{N}} p_n y_n I_n \\
\text{s.t. } & y_n \geq \frac{1}{p_n} \sum_{m\in n+} p_m y_m, \quad n \in \mathcal{N}\backslash\mathcal{T}, \\
& c_n \leq \frac{1}{p_n} \sum_{m\in n+} p_m y_m, \quad n \in \mathcal{N}\backslash\mathcal{T}, \\
& y_n \leq q_n, \quad n \in \mathcal{N}, \\
& y_n \geq d, \quad n \in \mathcal{T}.
\end{aligned}
\tag{31}
$$

The dual process y_n is a submartingale.

5 Examples

Example 2 This example is due to Philippe Artzner. Suppose a fair coin is thrown three times. Consider two situations:

Situation 1: The income is 1 at the final stage, if more heads than tails were counted.

Situation 2: The income is 1 at the final stage, if the last throw shows heads.

The corresponding income stream tree is shown in Figure 2, where an upmove means heads and a downmove means tails:

Evidently, the two cases leads to exactly the same marginal income distributions at each stage. On the other hand, Situation 1 is more predictable and should lead to a smaller risk.

We calculated the linear program (30 with the specification

$$c = [1, 0.95, (0.95)^2], \quad q = [1.2, 1.2 \cdot 0.95, 1.2 \cdot (0.95)^2], \quad d = (0.93)^2,$$

and we have obtained the following results:

$$\mathcal{U}(I^{(1)}) = 0.4419, \mathcal{U}(I^{(2)}) = 0.4325,$$

$$\mathcal{U}(\mathbb{E}(I^{(1)})) = \mathcal{U}(\mathbb{E}(I^{(2)})) = 0.5 \cdot (0.95)^2 = 0.4512$$

and therefore

$$\mathcal{R}(I^{(1)}) = 0.0093 < \mathcal{R}(I^{(2)}) = 0.0188.$$

This analysis shows that process 2 is riskier than process 1, indeed.

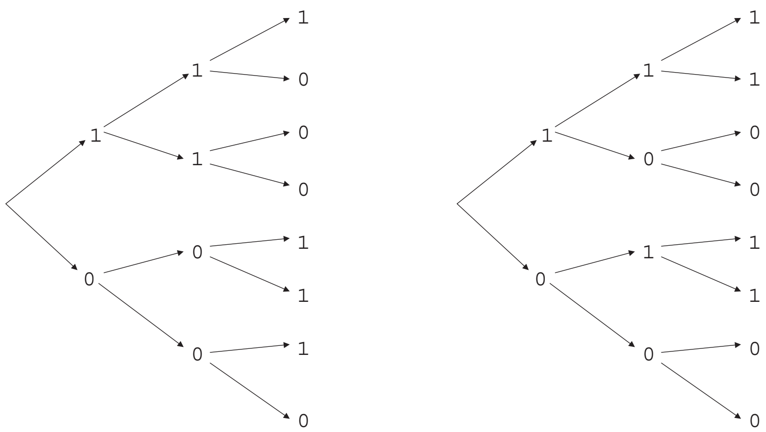


Figure 1: Left: Tree 1 Right: Tree 2.

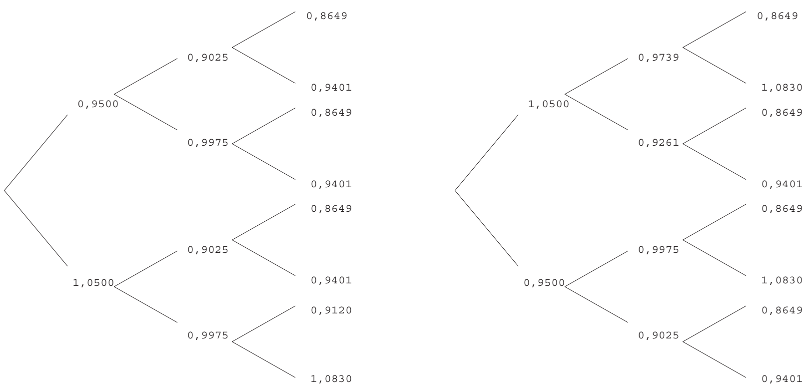


Figure 2: The dual submartingale processes: Left: Tree 1; Right: Tree 2

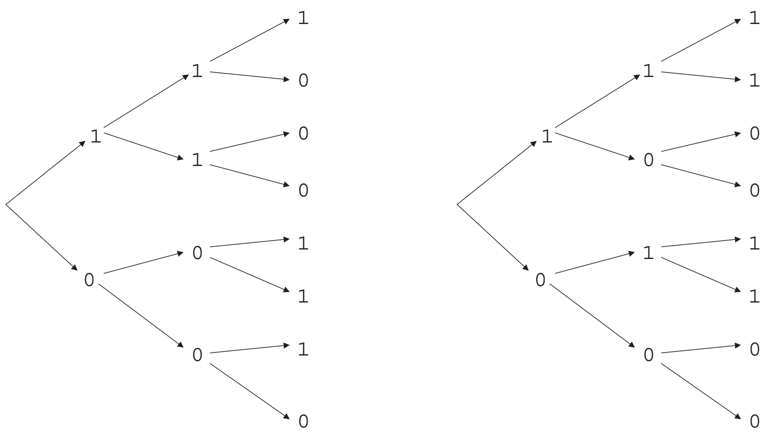


Figure 3: Left: Tree 3; Right: Tree 4.

It is also interesting to look at the dual variables $y^{(1)}$ and $y^{(2)}$ given by (31). They generate a dual process, which lives on the same tree as the income process. It is illustrated in Figure 5.

Example 3 We modify Example 2 in such a way that a positive income may also occur at stages 1 and 2. Consider the following income trees:

Assuming that all arc probabilities are 0.5 one gets the result

$$\mathcal{U}(I^{(3)}) = 1.3919, \mathcal{R}(I^{(3)}) = 0.0344$$

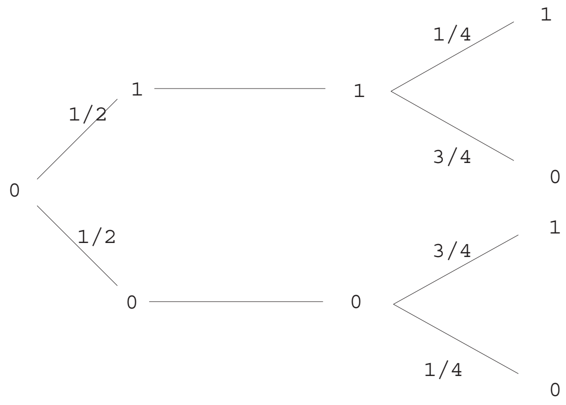
$$\mathcal{U}(I^{(4)}) = 1.3775, \mathcal{R}(I^{(4)}) = 0.0487$$

Since the predicability occurs earlier in tree 3, its risk is smaller. It is important to notice that hiding some information leads to larger risk: Suppose that the outcome of throw 2 is not revealed. In this case, the tree changes to Tree 3a.

The utility and risk for tree 3a are

$$\mathcal{U}(I^{(3a)}) = 1.3825, \mathcal{R}(I^{(3a)}) = 0.0438.$$

As expected, the risk of tree 3a is larger than the risk of tree 3.



References

[1] P. Artzner, F. Delbaen, J.-M. Eber and D. Heath, Coherent measures of risk, *Mathematical Finance* 9 (1999), pp. 203–228.

[2] Ioffe A.D., Tihomirov V.D. , *Theory of Extremal Problems*, North–Holland, Amsterdam, 1979.

[3] Kantorovich L.V., Akilov G.P., *Functional Analysis*, Pergamon Press, Oxford, 1982.

[4] W. Ogryczak and A. Ruszczyński, Dual Stochastic Dominance and Related Mean–Risk Models, Rutcor Research Report 10-2001, Rutgers University, Piscataway 2001 (to appear in *SIAM Journal on Optimization*).

[5] G. Ch. Pflug. Some remarks on the Value-at-Risk and the Conditional Value-at-Risk. In: *Probabilistic Constrained Optimization - Methodology and Applications* (S. Uryasev editor) Kluwer Academic Publishers, ISBN 0-7923-6644-1 pp. 272 - 281, 2000

[6] R. T. Rockafellar and S. Uryasev, Optimization of Conditional Value-at-Risk, *Journal of Risk*, to appear. (Report 99-4, Center for Applied Optimization, University of Florida, 1999).