Kusuoka Representation of Higher Order Dual Risk Measures

Darinka Dentcheva* Spiridon Penev† Andrzej Ruszczyński‡

March 22, 2010

Abstract

We derive representations of higher order dual measures of risk in $\mathcal{L}^p$ spaces as suprema of integrals of Average Values at Risk with respect to probability measures on (0, 1] (Kusuoka representations). The suprema are taken over convex sets of probability measures. The sets are described by constraints on the dual norms of certain transformations of distribution functions. For $p = 2$, we obtain a special description of the set and we relate the measures of risk to the Fano factor in statistics.
Keywords: Lorenz curve, quantile functions, average value at risk, coherent measures of risk, Fano factor, optimization, duality

1 Introduction

Recently mathematical models of risk attract much attention. In many practical applications presence of uncertainty is paramount and influences crucial quantities involved in decision making. Examples of applications, where the stochastic nature of the system cannot be neglected, are the design and operation of a network that should carry uncertain traffic, insurance and investment policy that should perform safely, supplying distribution centers to cover uncertain demand, production yield optimization in semiconductor and chemical industry, etc. One way of dealing with that is to optimize the system on average, if the average performance is representative for our decision problem. This is usually not the case when high uncertainty is involved. Another way suggests constraining probabilities of undesirable events, i.e. formulating probabilistic or chance constraints. In the presence of high uncertainty, it is beneficial to use risk functionals, which assign to a random variable X

*Stevens Institute of Technology, Hoboken, NJ; Email: darinka.dentcheva@stevens.edu
†The University of New South Wales, Sydney; Email: spiro@maths.unsw.edu.au
‡Rutgers University, Piscataway, NJ; Email: rusz@business.rutgers.edu

a nonnegative number $\varrho(X)$ representing a "safe equivalent" that offsets X. These functionals capture the entire distribution of X and account for undesirable events in an aggregate way. Precursors of risk measures were mean–risk models in finance, such as the Markowitz *mean–variance model* [16, 17], or signal to noise measures used in engineering and statistics, such as the *Fano factor* [8] or the *index of dispersion* (see, e.g., [6]).

A systematic theory of measures of risk was initiated by Artzner, Delbaen, Eber and Heath in [3]. The central concept of this theory is that of a coherent measure of risk, which we present here in a more general setting analyzed in [24]. Let $p \in [1, \infty]$ and let $(\Omega, \mathcal{F}, P)$ be a probability space. The notation $\mathcal{L}^p(\Omega, \mathcal{F}, P)$, or simply $\mathcal{L}^p(\Omega)$, is used for the space of measurable functions $X : \Omega \to \mathbb{R}$ such that $|X|^p$ is integrable. We use $\overline{\mathbb{R}}$ to denote the extended real line $\mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$.

A *coherent measure of risk* is a functional $\varrho : \mathcal{L}^p(\Omega) \to \overline{\mathbb{R}}$ satisfying the following axioms:

Convexity: $\varrho(\gamma X + (1-\gamma)Y) \leq \gamma\varrho(X) + (1-\gamma)\varrho(Y)$ for all X, Y and $\gamma \in [0, 1]$.

Monotonicity: If $Y \geq X$ P-a.s., then $\varrho(Y) \leq \varrho(X)$.

Translation Equivariance: If $a \in \mathbb{R}$ then $\varrho(X + a) = \varrho(X) - a$ for all X.

Positive Homogeneity: If $t > 0$ then $\varrho(tX) = t\varrho(X)$ for any X.

This concept was introduced in [3] for functionals ϱ defined on $\mathcal{L}^\infty(\Omega)$. It extended the earlier axiomatic approach of [11], which did not have the monotonicity axiom. For further developments, we refer the reader to [23, 24, 25] and the references therein. The essence of the convexity axiom is to model the benefits of diversification: risk of a combination of outcomes is no larger than combination of risks. The monotonicity axiom reflects our preference to larger outcomes. The translation property allows to remove the effect of certain gains or losses from the risk evaluation. Finally, the homogeneity axiom makes the risk analysis independent on the units in which the outcomes are expressed. Several important classes of coherent measures of risk were proposed and analyzed in the literature (see. e.g, [1, 4, 9, 12, 18, 19, 20, 22, 24, 25] and the references therein).

Our objective is to consider a particular family of coherent measures of risk, higher moment dual risk measures, and to derive their Kusuoka representation.

The symbol $\mathcal{P}(I)$ denotes the set of probability measures on an interval $I \subset \mathbb{R}$. Given a random variable $X \in \mathcal{L}^p(\Omega)$, $p \in (1, +\infty]$, we define $X_+ = \max\{X, 0\}$.

2 Average Value at Risk and Kusuoka Representations

A general assumption throughout the paper is the law-invariance of the measures of risk, i.e., we assume that $\varrho(\cdot)$ has the same value for random variables with the same distributions. For a random variable $X \in \mathcal{L}^1(\Omega)$, with distribution function $F_X(\eta) = P\{X \leq \eta\}$, we consider the left-continuous inverse of the cumulative distribution function defined as follows:

$$F_X^{(-1)}(\alpha) = \inf\{\eta : F_X(\eta) \geq \alpha\} \quad \text{for} \quad 0 < \alpha < 1.$$

It is clear that $F_X^{(-1)}(\alpha)$ is the left α-quantile of X. The *Value at Risk* of X at level α is defined as $\mathrm{VaR}_\alpha(X) = -F_X^{(-1)}(\alpha)$. Next, we define the *absolute Lorenz function* $F_X^{(-2)}(\cdot) : [0, 1] \to \overline{\mathbb{R}}$, introduced in [15], as the cumulative quantile function:

$$F_X^{(-2)}(\alpha) = \int_0^\alpha F_X^{(-1)}(t)\,dt \quad \text{for} \quad 0 < \alpha \leq 1, \tag{1}$$

and $F_X^{(-2)}(0) = 0$. Relative Lorenz functions are widely used in economics for comparison of positive random variables, relative to their (positive) expectations (see [2, 10] and the references therein). This function is defined as $\alpha \mapsto F_X^{(-2)}(\alpha)/\mathbb{E}[X]$. It is just a normalized absolute Lorenz curve due to representation of the expectation as integral of quantiles, i.e., $F_X^{(-2)}(1) = \mathbb{E}[X]$.

In the theory of measures of risk a special role is played by the functional called the Average Value at Risk and denoted $\mathrm{AVaR}(\cdot)$ (see [1, 22]). The *Average Value at Risk* of X at level α is defined as

$$\mathrm{AVaR}_\alpha(X) = -\frac{1}{\alpha} F_X^{(-2)}(\alpha) = \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_t(X)\,dt. \tag{2}$$

We call $\varrho : \mathcal{L}^p(\Omega) \to \overline{\mathbb{R}}$ a *Kusuoka measure of risk*, if there exists a convex set $\mathcal{M} \subset \mathcal{P}(I)$, where $I = (0, 1]$, such that for all X we have

$$\varrho(X) = \sup_{\mu \in \mathcal{M}} \int_0^1 \mathrm{AVaR}_\alpha(X)\,\mu(d\alpha).$$

A fundamental result in the theory of coherent measures of risk is the following theorem : *For a nonatomic space Ω, every law invariant, finite-valued coherent measure of risk on $\mathcal{L}^p(\Omega)$, $p \in [1, \infty]$ is a Kusuoka measure.* In the case of $p = \infty$ this result was proved in [13]. The proof for $p \in [1, \infty)$ can be found in [25, Thm. 6.24].

Kusuoka representations are useful in the analysis of models involving law invariant measures of risk. For example, in [5] analytical results for models involving the Average Value at Risk are extended to general law invariant measures of risk with the use of Kusuoka representations. Furthermore, Kusuoka representations allow to extend statistical estimators of Lorenz curves to law invariant measures of risk [7].

The Lorenz curve is closely related to stochastic orders and coherent measures of risk. For a random variable $X \in \mathcal{L}^p(\Omega)$ we consider the second order distribution function:

$$F_X^{(2)}(\eta) = \int_{-\infty}^{\eta} F_X(\alpha)\, d\alpha = \mathbb{E}\big[(\eta - X)_+\big] \quad \text{for } \eta \in \mathbb{R}. \tag{3}$$

We observe that both $F_X^{(2)}(\cdot)$ and $F_X^{(-2)}(\cdot)$ are well defined for any random variable $X \in \mathcal{L}^p(\Omega)$ and convex. Recall that for a function $f : \mathbb{R}^n \to \mathbb{R}$ the Fenchel conjugate $f^* : \mathbb{R}^n \to \mathbb{R}$ is defined as $f^*(s) = \sup_x\{\langle s, x\rangle - f(x)\}$. The following result [20, Thm. 3.1] establishes relations between (1) and (3): *For every integrable random variable X we have*

$$F_X^{(-2)} = [F_X^{(2)}]^* \quad \text{and} \quad F_X^{(2)} = [F_X^{(-2)}]^*. \tag{4}$$

In particular, from equations (2) and (4) we obtain an extremal representation of $\mathrm{AVaR}_\alpha(X)$ (see also [22]):

$$\begin{aligned}\mathrm{AVaR}_\alpha(X) &= -\frac{1}{\alpha}[F_X^{(2)}]^*(\alpha) = -\frac{1}{\alpha}\sup_{\eta\in\mathbb{R}}\big\{\eta\alpha - F^{(2)}(\eta)\big\}\\ &= \inf_{\eta\in\mathbb{R}}\left\{\frac{1}{\alpha}\mathbb{E}[(\eta - X)_+] - \eta\right\}.\end{aligned} \tag{5}$$

This extremal representation of $\mathrm{AVaR}_\alpha(X)$ was generalized in [12] to suggest the following higher moment measures of risk:

$$\inf_{\eta\in\mathbb{R}}\left\{\frac{1}{\alpha}\|(\eta - X)_+\|_p - \eta\right\}, \quad p > 1. \tag{6}$$

They are special cases of a more general family considered in [4].

Let us note the relation of the Lorenz curve to the second order stochastic dominance relation, which is a fundamental relation used both in statistics and economics. The stochastic dominance relation of order $k = 1, 2$ is defined as follows:

$$X \succeq_{(k)} Y \quad \Leftrightarrow \quad F_X^{(k)}(\eta) \le F_Y^{(k)}(\eta) \text{ for all } \eta \in \mathbb{R}. \tag{7}$$

We may fully characterize the second order dominance relation by using the conjugate function $F_X^{(-2)}(\cdot)$:

$$X \succeq_{(2)} Y \quad \Leftrightarrow \quad F_X^{(-2)}(\alpha) \ge F_Y^{(-2)}(\alpha) \quad \text{for all } 0 \le \alpha \le 1. \tag{8}$$

A risk measure ϱ on $\mathcal{L}^p(\Omega)$ is *consistent with stochastic dominance* if the following implication is true: $X \succeq_{(2)} Y \Rightarrow \varrho(X) \leq \varrho(Y)$. This concept was introduced in [18] and analyzed in [19, 20], where several classes of risk measures consistent with the second order stochastic dominance were identified. Leitner [14] extends the Kusuoka theorem for $p = \infty$ to the case of general probability spaces and coherent risk measures consistent with the second order stochastic dominance.

3 Kusuoka Representation of Higher Order Measures

Let $p \in (1, \infty)$, q be such that $1/p + 1/q = 1$, and choose a constant $c \geq 1$. We consider the risk measure (6) with $\alpha = 1/c$ and denote it by $\varrho_{c,p}(X)$, for $X \in \mathcal{L}^p(\Omega)$. Formally,

$$\varrho_{c,p}(X) = \inf_{\eta \in \mathbb{R}} \big\{ c\|(\eta - X)_+\|_p - \eta \big\}. \tag{9}$$

We shall show that it has a Kusuoka representation with the set

$$\mathcal{M}_q = \left\{ \mu \in \mathcal{P}(I) : \int_0^1 \Big| \int_\alpha^1 \frac{\mu(dt)}{t} \Big|^q d\alpha \leq c^q \right\}. \tag{10}$$

Theorem 1. *The functional $\varrho_{c,p}$ given by* (9) *with $p \in (1, \infty)$ has a Kusuoka representation of the following form:*

$$\varrho_{c,p}(X) = \sup_{\mu \in \mathcal{M}_q} \int_0^1 \mathrm{AVaR}_\alpha(X)\, \mu(d\alpha).$$

Proof. Define the function

$$\varphi_\mu(\alpha) = \int_\alpha^1 \frac{\mu(dt)}{t}. \tag{11}$$

We observe that $\varphi_\mu(\cdot)$ is nonnegative and nonincreasing. Moreover, we have

$$\int_0^1 \varphi_\mu(\alpha)\, d\alpha = \int_0^1 \int_\alpha^1 \frac{\mu(dt)}{t} d\alpha = \int_0^1 \int_0^t d\alpha \frac{\mu(dt)}{t} = \int_0^1 \mu(dt) = 1.$$

The second equation is obtained by changing the order of integration. These observations imply that φ can be viewed as a density function.

Let $\mathcal{B}$ be the σ-algebra of Borel sets on I, and ℓ denote the Lebesgue measure on I. We denote the Banach space $\mathcal{L}^p(I, \mathcal{B}, \ell)$ by $\mathcal{L}^p(I)$. The norms in $\mathcal{L}^p(I)$ and in $\mathcal{L}^p(\Omega)$ are written $\|\cdot\|_p$; it will not lead to any confusion. We observe that the condition $\mu \in \mathcal{M}_q$ is equivalent to $\|\varphi_\mu\|_q \leq c$. Thus, φ_μ can be identified with an element in $\mathcal{L}^q(I)$. Another way to represent the set $\mathcal{M}_q$ is the following:

$$\mathcal{M}_q = \Big\{\mu \in \mathcal{P}(I) : \|\varphi_\mu\|_q \leq c\Big\}.$$

For any probability measure $\mu \in \mathcal{P}(I)$, we define the functional $T_X : \mathcal{P} \to \mathbb{R}$ as follows

$$T_X(\mu) = -\int_0^1 \frac{1}{\alpha} F_X^{(-2)}(\alpha)\, \mu(d\alpha).$$

Notice that $\mu(d\alpha)/\alpha = -d\varphi_\mu(\alpha)$, with $\varphi_\mu(\cdot)$ defined in (11). Integrating by parts and using (1) we obtain

$$T_X(\mu) = \int_0^1 F_X^{(-2)}(\alpha)\, d\varphi_\mu(\alpha) = -\int_0^1 \varphi_\mu(\alpha) F_X^{(-1)}(\alpha)\, d\alpha.$$

The postulated Kusuoka representation of $\varrho_{c,p}(X)$ is equivalent to the following optimization problem:

$$\max_{\mu \in \mathcal{M}_q} T_X(\mu). \tag{12}$$

We denote the cone of nonnegative elements in $\mathcal{L}^q(I)$ by Q.

The following problem is a relaxation of problem (12):

$$\begin{aligned} \min\ & \langle \varphi, F_X^{(-1)} \rangle \\ \text{subject to } & \|\varphi\|_q \leq c, \\ & \langle \varphi, \mathbb{1} \rangle = 1, \\ & \varphi \in Q. \end{aligned} \tag{13}$$

Here $\langle \cdot, \cdot \rangle$ refers to the pairing between the spaces $\mathcal{L}^p(I)$ and $\mathcal{L}^q(I)$, and $\mathbb{1}$ denotes the identity function in $\mathcal{L}^p(I)$. The relaxation is due to the fact that we do not require in (13) that the function $\varphi(\cdot)$ is nonincreasing. With this requirement added, one-to-one correspondence between measures $\mu \in \mathcal{M}_q$ and feasible functions φ in (13) exists. However, we shall show that the optimal solution of (13) is in fact nonincreasing, so that the corresponding $\mu \in \mathcal{M}_q$ is well defined and solves problem (12).

The set defined by the first two constraints of (13),

$$C = \{\varphi \in \mathcal{L}^q(I) : \|\varphi\|_q \leq c, \langle \varphi, \mathbb{1} \rangle = 1\},$$

is non-empty whenever $c \geq 1$. Moreover, it is closed and convex. Problem (13) fits into the problem setting of [21, Example 4', p. 26], with $\varphi \in Q$ as a generalized operator constraint. In this setting, let $u \in \mathcal{L}^q(I)$ be the perturbation of the right hand side of the operator constraint. We consider the function:

$$g(u) = \inf\{\langle \varphi, F_X^{(-1)} \rangle : \varphi \in C,\ \varphi - u \in Q\}. \tag{14}$$

In order to apply [21, Theorem 17], we have to show that g is bounded in a neighborhood of 0. Let $u \in \mathcal{L}^q(I)$ be such that $\|u\|_q \leq \varepsilon < 1$ and let φ_0 be the solution associated with $g(0)$. For $\gamma \in (0, 1)$, we define

$$\varphi_u(\alpha) = \begin{cases} u(\alpha) & \text{if } u(\alpha) \geq \varphi_0(\alpha), \\ \gamma u(\alpha) + (1-\gamma)\varphi_0(\alpha) & \text{if } u(\alpha) < \varphi_0(\alpha). \end{cases}$$

We shall show that we can choose $\gamma \in (0, 1)$, so that φ_u is feasible for the problem on the right hand side of (14). In the first constraint we have

$$\begin{aligned} \langle \varphi_u, \mathbb{1} \rangle \ &= \int\limits_{u(\alpha)\geq\varphi_0(\alpha)} u(\alpha)\, d\alpha + \gamma \int\limits_{u(\alpha)<\varphi_0(\alpha)} u(\alpha)\, d\alpha + (1-\gamma) \int\limits_{u(\alpha)<\varphi_0(\alpha)} \varphi_0(\alpha)\, d\alpha \\ &= \langle u, \mathbb{1} \rangle + (1-\gamma) \int\limits_{u(\alpha)<\varphi_0(\alpha)} \big(\varphi_0(\alpha) - u(\alpha)\big)\, d\alpha. \end{aligned}$$

We observe that for $\gamma = 1$, these relations yield $\langle \varphi_u, \mathbb{1} \rangle = \|u\|_1 \leq \|u\|_q \leq \varepsilon < 1$. On the other hand, for $\gamma = 0$, we obtain $\langle \varphi_u, \mathbb{1} \rangle > \langle \varphi_0, \mathbb{1} \rangle = 1$. Therefore, we can choose $\gamma \in (0, 1)$ such that

$$\langle \varphi_u, \mathbb{1} \rangle = 1. \tag{15}$$

We also observe that $\varphi_u(\alpha) - u(\alpha) \geq 0$ for all $\alpha \in I$, thus implying

$$\varphi_u - u \in Q. \tag{16}$$

Having in mind that $c > \varepsilon$, we get

$$\begin{aligned} \|\varphi_u\|_q^q \ &\leq \|u\|_q^q P\big(u(\alpha) \geq \varphi_0(\alpha)\big) + \|\varphi_0\|_q^q P\big(u(\alpha) < \varphi_0(\alpha)\big) \\ &\leq \varepsilon^q P\big(u(\alpha) \geq \varphi_0(\alpha)\big) + c^q P\big(u(\alpha) < \varphi_0(\alpha)\big) \leq c^q. \end{aligned}$$

This inequality, together with (16) and (15), implies that φ_u is a feasible solution for the problem at the right hand side of (14). Thus, we have the following estimate:

$$g(u) \leq \langle \varphi_u, F_X^{(-1)} \rangle \leq \|\varphi_u\|_q \|F_X^{(-1)}\|_p \leq c\|F_X^{(-1)}\|_p.$$

Therefore, the assumptions of [21, Theorem 17] are satisfied. The subdifferential of the first constraint function $\|\varphi\|_q - c$ in (13) is

$$\partial\|\varphi\|_q = \big\{s \in \mathcal{L}^p(I) : \|s\|_p \leq 1,\ \langle \varphi, s \rangle = \|\varphi\|_q\big\}. \tag{17}$$

Recall that the normal cone to Q at φ is given by

$$\mathcal{N}_Q(\varphi) = \{y \in \mathcal{L}^p(I) : \langle y, \psi - \varphi \rangle \le 0 \ \forall \psi \in Q\}.$$

We have that $y \in \mathcal{N}_Q(\varphi)$ if and only if

$$y(\alpha) \le 0 \text{ and } \varphi(\alpha) y(\alpha) = 0 \text{ for almost all } \alpha \in I. \tag{18}$$

By virtue of [21, Theorem 17], an element $y \in \mathcal{N}_Q(\varphi)$ and a subgradient $s \in \partial \|\varphi\|_q$ exist, along with a nonnegative number ν and a number λ, such that

$$F_X^{(-1)} + \nu s - \lambda \mathbb{1} + y = 0. \tag{19}$$

Applying the functional φ to both sides of this equation, we obtain

$$\langle \varphi, F_X^{(-1)} \rangle = -\nu \langle \varphi, s \rangle + \lambda = -\nu \|\varphi\|_q + \lambda. \tag{20}$$

We consider two cases.

Case 1: $\nu = 0$. From (19) and (18) we conclude that

$$F_X^{(-1)}(\alpha) - \lambda = -y(\alpha) \ge 0 \text{ for almost all } \alpha \in I. \tag{21}$$

This implies that X is bounded from below almost surely and $\lambda \le \operatorname{essinf}(X)$. Denote $\bar{\alpha} = P(X = \lambda)$. From (21) we deduce that $y(\alpha) < 0$ for $\alpha \in (\bar{\alpha}, 1]$. By (18), $\varphi(\alpha) = 0$ for $\alpha \in (\bar{\alpha}, 1]$. As $\langle \varphi, \mathbb{1} \rangle = 1$, $\bar{\alpha}$ must be positive. Therefore $\lambda = \operatorname{essinf}(X)$ and $\bar{\alpha} = P(X = \operatorname{essinf}(X))$. The smallest value of $\|\varphi\|_q$ is attained at

$$\hat{\varphi}(\alpha) = \begin{cases} 1/\bar{\alpha} & \text{if } \alpha \in [0, \bar{\alpha}], \\ 0 & \text{if } \alpha \in (\bar{\alpha}, 1]. \end{cases}$$

This solution is feasible in problem (13) if and only if $\bar{\alpha} \ge c^{-p}$. As all optimality conditions are satisfied, $\hat{\varphi}$ is also optimal. We observe that it is a nonincreasing function of α, and thus corresponds to a measure $\mu \in \mathcal{M}_q$ according to formula (11). Owing to (20), the optimal value of problem (12) in this case takes on the form

$$-\langle \hat{\varphi}, F_X^{(-1)} \rangle = -\lambda = -\operatorname{essinf}(X).$$

We shall show that λ is a minimizer in the problem on the right hand side of (9). Consider the function $h(\eta) = c\|(\eta - X)_+\|_p - \eta$. It is convex. If $\eta < \operatorname{essinf}(X)$, then $h(\cdot)$ is differentiable with $h'(\eta) = -1$. Consider $\eta > \operatorname{essinf}(X)$. Then $h(\cdot)$ is differentiable as well with

$$h'(\eta) = c \frac{\mathbb{E}[Y^{p-1}]}{\big(\mathbb{E}[Y^p]\big)^{(p-1)/p}} - 1, \tag{22}$$

where $Y = (\eta - X)_+$. As $P(\lambda < X < \eta) \to 0$ as $\eta \downarrow \lambda$, we have

$$\mathbb{E}[Y^p] = \bar{\alpha}(\eta - \lambda)^p + o((\eta - \lambda)^p).$$

Substituting this into the formula for the derivative in the numerator (with $p - 1$) and in the denominator, we conclude that

$$\lim_{\eta \downarrow \operatorname{essinf}(X)} h'(\eta) = c\bar{\alpha}^{1/p} - 1. \tag{23}$$

If $\bar{\alpha} \geq c^{-p}$ then this limit is nonnegative. Therefore, the subdifferential of $h(\cdot)$ at λ contains 0, and thus λ is a minimizer of $h(\cdot)$. The value of the risk measure equals in this case $\varrho_{c,p}(X) = -\lambda = -\operatorname{essinf}(X)$ and is the same as the optimal value of problem (13).

Case 2: $\nu > 0$. Equations (19) and (17) imply that

$$\|\lambda\mathbb{1} - y - F_X^{(-1)}\|_p = \nu\|s\|_p \leq \nu. \tag{24}$$

We have the following chain of relations:

$$\begin{aligned}
\nu\|\varphi\|_q &= \langle \varphi, \nu s\rangle && \text{using (17)}\\
&= \langle \varphi, \lambda\mathbb{1} - y - F_X^{(-1)}\rangle && \text{using (19)}\\
&\leq \|\varphi\|_q \|\lambda\mathbb{1} - y - F_X^{(-1)}\|_p && \text{using Hölder's inequality}\\
&\leq \nu\|\varphi\|_q. && \text{using (24)}
\end{aligned}$$

Consequently, we have equality everywhere in the chain and in (24). We conclude that

$$\|s\|_p = 1,$$
$$\nu = \|\lambda\mathbb{1} - y - F_X^{(-1)}\|_p,$$
$$\left(\frac{\varphi}{\|\varphi\|_q}\right)^q = \left(\frac{\lambda\mathbb{1} - y - F_X^{(-1)}}{\|\lambda\mathbb{1} - y - F_X^{(-1)}\|_p}\right)^p.$$

Furthermore, $\lambda\mathbb{1} - y - F_X^{(-1)} \geq 0$, because $\varphi \geq 0$. As $\nu > 0$, the first constraint in (13) is active, i.e., $\|\varphi\|_q = c$. Transforming the last equation, we get

$$\varphi = c\left(\frac{\lambda\mathbb{1} - y - F_X^{(-1)}}{\|\lambda\mathbb{1} - y - F_X^{(-1)}\|_p}\right)^{p-1}. \tag{25}$$

We shall verify that $\lambda\mathbb{1} - y - F_X^{(-1)} = (\lambda\mathbb{1} - F_X^{(-1)})_+$. Indeed, if $y(\alpha) = 0$, then $0 \le \lambda - y(\alpha) - F_X^{(-1)}(\alpha) = \lambda - F_X^{(-1)}(\alpha)$. If $y(\alpha) < 0$, then (18) entails $\varphi(\alpha) = 0$. Using (25), we obtain $\lambda - F_X^{(-1)}(\alpha) = y(\alpha) < 0$. We conclude that

$$\nu = \|(\lambda\mathbb{1} - F_X^{(-1)})_+\|_p > 0, \tag{26}$$

$$\varphi = c\left(\frac{(\lambda\mathbb{1} - F_X^{(-1)})_+}{\|(\lambda\mathbb{1} - F_X^{(-1)})_+\|_p}\right)^{p-1}. \tag{27}$$

We observe that it is a nonincreasing function of α, and thus corresponds to a measure $\mu \in \mathcal{M}_q$ according to formula (11). We also notice that $\lambda > \text{essinf}(X)$ in this case. From the constraint $\langle\varphi, \mathbb{1}\rangle = 1$ we obtain

$$1 = \frac{c}{\|(\lambda\mathbb{1} - F_X^{(-1)})_+\|_p^{p-1}} \int_0^1 (\lambda\mathbb{1} - F_X^{(-1)})_+^{p-1}\, d\alpha. \tag{28}$$

We infer from (22) and (28) that $h'(\lambda) = 0$. Therefore, λ is again the optimal value of η in problem (9). As $\lambda > \text{essinf}(X)$, it follows the limit in (23) is nonpositive. This is equivalent to the inequality $\bar{\alpha} \le c^{-p}$.

Substituting λ and c in (20), we obtain the optimal value of problem (13):

$$\langle\varphi, F_X^{(-1)}\rangle = -c\|(\lambda - F_X^{(-1)})_+\|_p + \lambda.$$

Therefore, the optimal value of problem (12) is equal to the value of the risk measure (9). □ □

Remark 1. *When $p = 2$, the measure in Case 2 takes on the form*

$$\varrho_2(X) = -\eta_X + \frac{\|(\eta_X\mathbb{1} - F_X^{(-1)})_+\|_2^2}{\|(\eta_X\mathbb{1} - F_X^{(-1)})_+\|_1},$$

where η_X is the optimal solution of the problem on the right hand side of (9). If, for a given random variable X, we define $Y = (\eta_X\mathbb{1} - F_X^{(-1)})_+$, then the resulting risk measure is

$$\varrho_2(X) = \mathbb{E}(Y) - \eta_X + \frac{\text{Var}(Y)}{\mathbb{E}(Y)}.$$

Notice that the last term on the right hand side is the Fano factor, which is a specific expression for a noise-to-signal ratio.

Remark 2. *Observe that η_X may not represent a quantile of X. For example, if we consider X to have the uniform distribution on* $(0, 1)$*, we obtain that*

$$\|(\eta_X\mathbb{1} - F_X^{(-1)})_+\|_2 = 1/3\eta_X^3 - 1/3[(\eta_X - 1)_+]^3,$$
$$\|(\eta_X\mathbb{1} - F_X^{(-1)})_+\|_1 = 1/2\eta_X^2 - 1/2[(\eta_X - 1)_+]^2.$$

Assuming that $\eta_X < 1$, *the equation* $h'(\eta_X) = 0$ *becomes:*

$$\sqrt{\frac{1}{3}\eta_X^3} = \frac{c}{2}\eta_X^2,$$

which yields $\eta_X = \frac{4}{3c^2}$. *The requirement* $\eta_X < 1$ *implies that* $c > 2/\sqrt{3}$.

If $p = 2$, we can derive another equivalent description of the set of probability measures in the Kusuoka representation of (9).

Lemma 1. *The set* $\mathcal{M}_2$ *has the following equivalent description:*

$$\mathcal{M}_2 = \left\{\mu \in \mathcal{P}(I) : \int_0^1 \frac{1}{t}\, d\big(F_\mu^2(t)\big) \le c\right\},$$

where F_μ *denotes the cumulative distribution function of the probability measure* μ, *i.e.,* $F_\mu(t) = \mu([0, t])$.

Proof. Integrating by parts, we obtain

$$\begin{aligned}
&\int_0^1 \left[\int_\alpha^1 \frac{1}{t}\, dF_\mu(t)\right]^2 d\alpha \\
&\quad = \alpha\Big[\int_\alpha^1 \frac{1}{t}\, dF_\mu(t)\Big]^2\Big|_0^1 + \int_0^1 2\alpha\Big[\int_\alpha^1 \frac{1}{t}\, dF_\mu(t)\Big]\frac{1}{\alpha}\, dF_\mu(\alpha) \\
&\quad = 2\int_0^1 \int_\alpha^1 \frac{1}{t}\, dF_\mu(t)\, dF_\mu(\alpha) = 2\int_0^1 \int_0^t dF_\mu(\alpha)\, \frac{1}{t}\, dF_\mu(t) \\
&\quad = 2\int_0^1 \frac{F_\mu(t)}{t}\, dF_\mu(t) = \int_0^1 \frac{1}{t}\, d\big(F_\mu^2(t)\big),
\end{aligned}$$

as postulated. □ □

Remark 3. *The measure of risk given by the formula (9) is convex and law-invariant. This implies, that it is consistent with stochastic dominance of second order, i.e., if a random variable* X *dominates a random variable* Y, *then* $\varrho_{c,p}(X) \le \varrho_{c,p}(Y)$ *for any* $p > 1$.

Acknowledgements

The first author was partially by the NSF grant DMS-0604060. The second author was partially supported by a research grant PS15611 of The University of New South Wales. The third author was partially supported by the NSF grant DMS-0603728.

References

[1] Acerbi, C., Tasche, D.: On the coherence of expected shortfall. *Journal of Banking and Finance* 26, 1487–1503 (2002)

[2] Arnold, B. C.: *Majorization and the Lorenz Order: A Brief Introduction*. Lecture Notes in Statistics 43, Springer-Verlag, Berlin (1980)

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

[4] P. Cheridito and. T.H. Li, Risk measures on Orlicz hearts, *Mathematical Finance* 19 (2009) 189–214.

[5] S. Choi, A. Ruszczyński and Y. Zhao, A Multi-Product Risk-Averse Newsvendor with Law-Invariant Coherent Measures of Risk, *Operations Research*, accepted for publication.

[6] D. R. Cox and P. A. W. Lewis, *The Statistical Analysis of Series of Events*, Methuen, London, 1966.

[7] D. Dentcheva and S. Penev, Shape-restricted inference for Lorenz curves using duality theory, *Statistics & Probability Letters* 80 (2010) 403–412.

[8] U. Fano, Ionization field of radiations. II The fluctuations of the number of ions, *Physical Review* 72 (1947) 26–29.

[9] T. Fischer, Examples of Coherent Risk Measures Depending on One-Sided Moments. *Working Paper TU Darmstadt* 2001.

[10] J. L. Gastwirth, The estimation of the Lorenz curve and Gini index, *The Review of Economics and Statistics* 54 (1972) 306-–316.

[11] M. Kijima and M. Ohnishi, Mean–risk analysis of risk aversion and wealth effects on optimal portfolios with multiple investment possibilities, *Annals of Operations Research* 45 (1993) 147–163.

[12] P. Krokhmal, Higher moment coherent risk measures, *Quantitative Finance* 7 (2007) 373-387.

[13] Kusuoka, S.: On law invariant coherent risk measures, Adv. Math. Econ. 3 (2001) 83–95.

[14] J. Leitner, A short note on second order stochastic dominance preserving coherent risk measures, Matheatical Finance 15 (2005) 649–651.

[15] Lorenz, M. (1905) Methods of measuring concentration of wealth, Journal of the American Statistical Association, 9, 209-219.

[16] H. M. Markowitz, Portfolio selection, Journal of Finance, 7 (1952), 77–91.

[17] H. M. Markowitz, Mean–Variance Analysis in Portfolio Choice and Capital Markets, Blackwell, Oxford, 1987.

[18] Ogryczak, W. and Ruszczyński, A., From stochastic dominance to mean-risk models: Semideviations and risk measures, European J. Operations Research, 116 (1999), 33-50.

[19] Ogryczak, W., Ruszczyński, A.: On consistency of stochastic dominance and mean–semideviation models, Mathematical Programming, 89, 217–232 (2001)

[20] Ogryczak, W. and Ruszczyński, A. (2002), Dual stochastic dominance and related mean-risk models. SIAM J. Optim., 13, 1, 60-78.

[21] Rockafellar, R. T.: Conjugate Duality and Optimization, CBMS-NSF Regional Conference Series in Applied Mathematics 16 SIAM, Philadelphia, 1974.

[22] Rockafellar, R. T., Uryasev, S.: Conditional value-at-risk for general loss distributions. Journal of Banking and Finance 26, 1443–1471 (2002)

[23] Rockafellar, R. T., Uryasev, S., Zabarankin, M.: Generalized deviations in risk analysis, *Finance and Stochastics* 10(2006) 51–74.

[24] Ruszczynski, A., Shapiro, A.: Optimization of Convex Risk Functions, Mathematics of Operations Research 31, 433–452 (2006)

[25] Shapiro, A., D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*, SIAM Publications, Philadelphia, 2009.