Insurance: Mathematics and Economics 54 (2014) 41–48

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Generalized quantiles as risk measures

Fabio Bellini [a], Bernhard Klar [b], Alfred Müller [c,*], Emanuela Rosazza Gianin [a]

[a] Dipartimento di Statistica e Metodi Quantitativi, Università di Milano Bicocca, Italy
[b] Institut für Stochastik, Karlsruher Institut für Technologie (KIT), Germany
[c] Department of Mathematics, University of Siegen, Germany

HIGHLIGHTS

- We investigate the use of generalized quantiles as risk measures.
- We show that expectiles are interesting coherent risk measures.
- We show that expectiles have interesting robustness properties and discuss their relation to quantiles.

ARTICLE INFO

Article history:
Received February 2013
Received in revised form
October 2013
Accepted 28 October 2013

Keywords:
Expectile
Coherence
Elicitability
Kusuoka representation
Robustness
Wasserstein metric

ABSTRACT

In the statistical and actuarial literature several generalizations of quantiles have been considered, by means of the minimization of a suitable asymmetric loss function. All these generalized quantiles share the important property of *elicitability*, which has received a lot of attention recently since it corresponds to the existence of a natural backtesting methodology. In this paper we investigate the case of M-quantiles as the minimizers of an asymmetric convex loss function, in contrast to Orlicz quantiles that have been considered in Bellini and Rosazza Gianin (2012). We discuss their properties as risk measures and point out the connection with the zero utility premium principle and with shortfall risk measures introduced by Föllmer and Schied (2002). In particular, we show that the only M-quantiles that are coherent risk measures are the *expectiles*, introduced by Newey and Powell (1987) as the minimizers of an asymmetric quadratic loss function. We provide their dual and Kusuoka representations and discuss their relationship with CVaR. We analyze their asymptotic properties for $\alpha \to 1$ and show that for very heavy tailed distributions expectiles are more conservative than the usual quantiles. Finally, we show their robustness in the sense of lipschitzianity with respect to the Wasserstein metric.

1. Introduction

It is well known that the quantiles q_α of a random variable X may be defined as the minimizers of a piecewise-linear loss function. Namely,

$$q_\alpha(X) = \underset{x\in\mathbb{R}}{\operatorname{argmin}} \left\{\alpha E[(X-x)^+] + (1-\alpha)E[(X-x)^-]\right\},$$

where we use the notation $x^+ := \max\{x, 0\}$ and $x^- := \max\{-x, 0\}$. This property is very important in the statistical as well as in the actuarial literature; it lies in the heart of the quantile regression (see, for example Koenker, 2005), and it is the starting point of the approach of Rockafellar and Uryasev (2002) for the computation of the Conditional Value at Risk (CVaR henceforth). In the statistical literature several *generalized quantiles* x_α have been introduced, by considering more general loss functions: Newey and Powell (1987) introduced the *expectiles* as the minimizers of a piecewise quadratic loss function, Chen (1996) considered power loss functions (*L^p-quantiles*), Breckling and Chambers (1988) considered generic loss functions (*M-quantiles*). In the context of the evaluation of point forecasts, Gneiting (2011) introduced the general notion of *elicitability* for a functional that is defined by means of a similar loss minimization process; the relevance of elicitability in connection with backtesting has been recently addressed by Embrechts and Hofert (2013) while the relationship between coherency and elicitability has been considered in Ziegel (2013). In this paper, we will adopt the following general asymmetric formulation:

$$x_\alpha(X) = \underset{x\in\mathbb{R}}{\operatorname{argmin}} \{\alpha E[\Phi_1((X-x)^+)] + (1-\alpha)E[\Phi_2((X-x)^-)]\}, \quad (1)$$

with convex loss functions Φ_1 and Φ_2.

Under suitable conditions (see Proposition 1), the minimizer in (1) is unique and satisfies a first-order condition; in this case, x_α can be equivalently defined as the unique solution of

$$E[\psi(X - x_\alpha)] = 0,$$

* Corresponding author. Tel.: +49 2717403587.
E-mail addresses: fabio.bellini@unimib.it (F. Bellini), bernhard.klar@kit.edu (B. Klar), derfla66@web.de, mueller@mathematik.uni-siegen.de (A. Müller), emanuela.rosazza1@unimib.it (E. Rosazza Gianin).

0167-6687/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.insmatheco.2013.10.015

for an appropriate function $\psi : \mathbb{R} \to \mathbb{R}$, and is thus a special case of the *zero utility premium principle*, well known in the actuarial literature (see for example Deprez and Gerber, 1985, and the references therein). The same object has been studied from different points of view under different names: Föllmer and Schied (2002) called a slightly more general version *shortfall risk measure* and Ben-Tal and Teboulle (2007) called it *ψ-mean certainty equivalent.* Generalized quantiles share several good properties of the usual quantiles: for instance, translation equivariance and monotonicity with respect to the usual stochastic order. It is then quite natural to further investigate the properties of these statistical functionals from the point of view of the axiomatic theory of risk measures. We characterize the generalized quantiles that have positive homogeneity and convexity properties, and show that the only generalized quantiles that are coherent risk measures are the expectiles e_α (with $\alpha \geq \frac{1}{2}$), that have been introduced by Newey and Powell (1987) as the minimizers of an asymmetric quadratic loss

$$e_\alpha(X) = \operatorname*{argmin}_{x\in\mathbb{R}} \left\{\alpha E[((X-x)^+)^2] + (1-\alpha)E[((X-x)^-)^2]\right\}.$$

Related generalizations of the notion of quantiles have also arisen in connection with the Haezendonck–Goovaerts risk measures (see for example Goovaerts et al., 2012; Bellini and Rosazza Gianin, 2012, and the references therein). An interesting alternative formulation has been suggested by Jaworski (2006), and a brief comparison will be discussed in Section 2. Section 3 is devoted to the study of expectiles, which are becoming increasingly popular also in the econometric literature (see for instance Kuan et al., 2009; De Rossi and Harvey, 2009; Embrechts and Hofert, 2013, and the references therein). By applying the results of Föllmer and Schied (2002), we compute their dual representation and their Kusuoka representation. We show that they can be seen as the supremum of 'bilateral' CVaR that have also been considered by Pflug and Ruszczynski (2004). We then discuss the problem of the comparison of expectiles with quantiles. We can prove an asymptotic result that shows that for distributions with Paretian tails, for high values of the tail index α, expectiles are larger than quantiles in the infinite variance case while the opposite inequality holds in the finite variance case. Finally, we discuss robustness properties of expectiles. It has been recently argued by several authors that an appropriate notion of robustness for risk measures is continuity with respect to the Wasserstein distance (see for example Stahl et al., 2012). We prove that actually expectiles are Lipschitz continuous with respect to the Wasserstein metric (see also Pflug and Wozabal (2007) or Pichler (forthcoming) for related results for CVaR or more general coherent risk measures).

2. Generalized quantiles

Let $\Phi_1, \Phi_2 : [0, +\infty) \to [0, +\infty)$ be convex, strictly increasing functions satisfying

$$\Phi_i(0) = 0 \quad \text{and} \quad \Phi_i(1) = 1. \tag{2}$$

Given a function Φ as above, a probability space $(\Omega, \mathcal{F}, P)$ and the space L^0 of all random variables X on $(\Omega, \mathcal{F}, P)$, we recall that the *Orlicz heart* M^Φ

$$M^\Phi := \left\{X \in L^0 : E\left[\Phi\left(\frac{|X|}{a}\right)\right] < +\infty, \text{ for every } a > 0\right\}$$

is a Banach space with respect to the *Luxemburg norm* $\|\cdot\|_\Phi$, defined as

$$\|Y\|_\Phi := \inf\left\{a > 0 : E\left[\Phi\left(\frac{|Y|}{a}\right)\right] \leq 1\right\}.$$

See Rao and Ren (1991) for an exhaustive treatment on Orlicz spaces. To simplify the notation, $X \geq Y$ will often stand for $X \geq Y$, P-a.s., while $X \geq_{st} Y$ will denote that X dominates Y in the first-order stochastic dominance, that is $P(X \leq x) \leq P(Y \leq x)$ holds for any $x \in \mathbb{R}$.

We consider now the minimization problem

$$\pi_\alpha(X) := \inf_{x\in\mathbb{R}} \pi_\alpha(X, x), \tag{3}$$

where

$$\pi_\alpha(X, x) := \alpha E\left[\Phi_1((X-x)^+)\right] + (1-\alpha) E\left[\Phi_2((X-x)^-)\right]. \tag{4}$$

Following Breckling and Chambers (1988), we call any minimizer

$$x_\alpha^* \in \operatorname{argmin} \pi_\alpha(X, x)$$

a *generalized quantile*.

Since the defining minimization problem is convex, generalized quantiles can be characterized by means of a first-order condition (f.o.c. henceforth). In the next proposition we give a general statement that does not require differentiability properties of the Φ_i.

Proposition 1. *Let $\Phi_1, \Phi_2 : [0, +\infty) \to [0, +\infty)$ be convex, strictly increasing and satisfy (2). Let $X \in M^{\Phi_1} \cap M^{\Phi_2}$, $\alpha \in (0, 1)$ and $\pi_\alpha(X, x)$ as in (4).*

(a) *$\pi_\alpha(X, x)$ is finite, non-negative, convex, and satisfies*

$$\lim_{x\to-\infty} \pi_\alpha(X, x) = \lim_{x\to+\infty} \pi_\alpha(X, x) = +\infty;$$

(b) *the set of minimizers is a closed interval:*

$$\operatorname{argmin} \pi_\alpha(X, x) := [x_\alpha^{*-}; x_\alpha^{*+}];$$

(c) *$x_\alpha^* \in \operatorname{argmin} \pi_\alpha(X, x)$ if and only if*

$$\begin{cases} \alpha E\left[\mathbf{1}_{\{X>x_\alpha^*\}}\Phi'_{1-}((X-x_\alpha^*)^+)\right] \\ \quad \leq (1-\alpha) E\left[\mathbf{1}_{\{X\leq x_\alpha^*\}}\Phi'_{2+}((X-x_\alpha^*)^-)\right] \\ \alpha E\left[\mathbf{1}_{\{X\geq x_\alpha^*\}}\Phi'_{1+}((X-x_\alpha^*)^+)\right] \\ \quad \geq (1-\alpha) E\left[\mathbf{1}_{\{X<x_\alpha^*\}}\Phi'_{2-}((X-x_\alpha^*)^-)\right], \end{cases} \tag{5}$$

where Φ'_{i-} and Φ'_{i+} denote the left and right derivatives of Φ_i;

(d) *if Φ_1 and Φ_2 are strictly convex, then $x_\alpha^{*-} = x_\alpha^{*+}$.*

Proof. (a) If $X \in M^{\Phi_1} \cap M^{\Phi_2}$, then $(X-x)^+ \in M^{\Phi_1}$ and $(X-x)^- \in M^{\Phi_2}$ for any $x \in \mathbb{R}$. Indeed, by convexity and monotonicity of Φ_1, it follows that $E\left[\Phi_1((X-x)^+)\right] \leq E[\Phi_1(|X|+|x|)] \leq \frac{1}{2}E[\Phi_1(2|X|)] + \frac{1}{2}\Phi_1(2|x|) < +\infty$. Similarly, $E\left[\Phi_2((X-x)^-)\right] < +\infty$ for any $x \in \mathbb{R}$. So, $\pi_\alpha(X, x)$ is finite for any $x \in \mathbb{R}$. Non-negativity and convexity of $\pi_\alpha(X, x)$ are trivial. By the Monotone Convergence Theorem, it follows that $\lim_{x\to-\infty} \pi_\alpha(X, x) = \lim_{x\to+\infty} \pi_\alpha(X, x) = +\infty$; (b) is a direct consequence of (a).

(c) From the convexity of $\pi_\alpha(X, x)$ it follows that x_α^* is a minimizer if and only if $0 \in \partial\pi_\alpha(X, x_\alpha^*) := [\frac{\partial^-\pi_\alpha}{\partial x}, \frac{\partial^+\pi_\alpha}{\partial x}]$. Define $r(x) = E[\Phi_1((X-x)^+)]$ and $l(x) = E[\Phi_2((X-x)^-)]$. Let $y > x$. We have that

$$\begin{aligned} \frac{r(y)-r(x)}{y-x} &= \frac{E[\Phi_1((X-y)^+)] - E[\Phi_1((X-x)^+)]}{y-x} \\ &=: E[H_X(x, y)], \end{aligned}$$

where the random variable $H_X(x, y)$ fulfils

$$H_X(x, y) = \begin{cases} \dfrac{\Phi_1(X-y)-\Phi_1(X-x)}{y-x} & \text{if } X \geq y \\ -\dfrac{\Phi_1(X-x)}{y-x} & \text{if } x < X < y \\ 0 & \text{if } X \leq x. \end{cases}$$

Since Φ_1 is convex and $X \in M^{\Phi_1}$ we have that $y \mapsto H_X(x, y)$ is non-positive and increasing, and therefore for all y with $|y-x| \leq 1$ it holds that

$$E|H_X(x, y)| \leq E|H_X(x, x-1)| < \infty.$$

From the dominated convergence theorem we get

$$r'_+(x) = \lim_{y\to x^+} \frac{r(y)-r(x)}{y-x}$$
$$= E\left[\lim_{y\to x^+} H_X(x,y)\right] = -E[1_{\{X>x\}}\Phi'_{1-}((X-x)^+)].$$

The same argument gives

$$r'_-(x) = -E[1_{\{X\geq x\}}\Phi'_{1+}((X-x)^+)]$$

and similarly

$$l'_+(x) = E[1_{\{X\leq x\}}\Phi'_{2+}((X-x)^-)]$$
$$l'_-(x) = E[1_{\{X<x\}}\Phi'_{2-}((X-x)^-)]$$

Since $\pi_\alpha(X,x) = \alpha r(x) + (1-\alpha)l(x)$, we get

$$\frac{\partial^-\pi_\alpha}{\partial x}(X,x) = -\alpha E\left[\mathbf{1}_{\{X\geq x\}}\Phi'_{1+}\left((X-x)^+\right)\right] + (1-\alpha)E\left[\mathbf{1}_{\{X<x\}}\Phi'_{2-}((X-x)^-)\right]$$
$$\frac{\partial^+\pi_\alpha}{\partial x}(X,x) = -\alpha E\left[\mathbf{1}_{\{X>x\}}\Phi'_{1-}\left((X-x)^+\right)\right] + (1-\alpha)E\left[\mathbf{1}_{\{X\leq x\}}\Phi'_{2+}((X-x)^-)\right]$$

from which (c) follows immediately.

(d) If Φ_1 and Φ_2 are strictly convex, then for each t the function

$$g(t,x) := \alpha\Phi_1((t-x)^+) + (1-\alpha)\Phi_2((t-x)^-)$$

is strictly convex in x. It follows that $\pi_\alpha(X,x) = E[g(X,x)]$ is strictly convex in x, hence its minimizer is unique. ■

In the general case the f.o.c. that identifies generalized quantiles is thus given by two inequalities, and cannot be reduced to a single equality; the simplest example of this situation is that of the usual quantiles.

Example 2 (*Quantiles*)**.** For $\Phi_1(x) = \Phi_2(x) = x$, generalized quantiles reduce to the usual quantiles. Indeed, the f.o.c. (5) becomes

$$\begin{cases} \alpha E\left[\mathbf{1}_{\{X>x^*_\alpha\}}\right] \leq (1-\alpha)E\left[\mathbf{1}_{\{X\leq x^*_\alpha\}}\right] \\ \alpha E\left[\mathbf{1}_{\{X\geq x^*_\alpha\}}\right] \geq (1-\alpha)E\left[\mathbf{1}_{\{X<x^*_\alpha\}}\right] \end{cases}$$

or, equivalently,

$$P(X < x^*_\alpha) \leq \alpha \leq P(X \leq x^*_\alpha).$$

Under additional smoothness assumptions on the functions Φ_1 and Φ_2 or on the distribution of X, the f.o.c. (5) reduces to a single equation.

Corollary 3. *Under the assumptions of Proposition 1, let Φ_1 and Φ_2 be differentiable. If $\Phi'_{1+}(0) = \Phi'_{2+}(0) = 0$ or if the distribution of X is continuous, then the f.o.c. (5) reduces to*

$$\alpha E\left[\Phi'_1\left(X-x^*\right)^+\right] = (1-\alpha)E\left[\Phi'_2\left(X-x^*\right)^-\right]. \tag{6}$$

Proof. If Φ_1 and Φ_2 are differentiable, then $\Phi'_{i+} = \Phi'_{i-}$. Hence (5) becomes

$$\begin{cases} \alpha E\left[\mathbf{1}_{\{X>x^*_\alpha\}}\Phi'_1((X-x^*_\alpha)^+)\right] \\ \quad \leq (1-\alpha)E\left[\mathbf{1}_{\{X\leq x^*_\alpha\}}\Phi'_2((X-x^*_\alpha)^-)\right] \\ \alpha E\left[\mathbf{1}_{\{X\geq x^*_\alpha\}}\Phi'_1\left((X-x^*_\alpha)^+\right)\right] \\ \quad \geq (1-\alpha)E\left[\mathbf{1}_{\{X<x^*_\alpha\}}\Phi'_2((X-x^*_\alpha)^-)\right]. \end{cases}$$

If moreover $\Phi'_{1+}(0) = \Phi'_{2+}(0) = 0$ or if X is a continuous random variable (hence $P(X = x^*_\alpha) = 0$), then the two inequalities above reduce to

$$\alpha E\left[\Phi'_1((X-x^*_\alpha)^+)\right] = (1-\alpha)E\left[\Phi'_2\left((X-x^*_\alpha)^-\right)\right]. \quad ■$$

A prototypical case is that of expectiles, which will be investigated in detail in the following section. The preceding corollary shows that in contrast to quantiles, expectiles can always be identified as the solution of an equation, even when the distribution of X is not continuous.

Example 4 (*Expectiles*)**.** For $\Phi_1(x) = \Phi_2(x) = x^2$, generalized quantiles reduce to expectiles. Since Φ_1 and Φ_2 are differentiable and $\Phi'_{1+}(0) = \Phi'_{2+}(0) = 0$, the f.o.c. (5) is given by

$$\alpha E\left[\left(X-x^*_\alpha\right)^+\right] = (1-\alpha)E\left[\left(X-x^*_\alpha\right)^-\right].$$

Any solution of the equality above is called α-expectile of X and denoted by $e_\alpha(X)$.

In all cases in which the hypotheses of Corollary 3 are satisfied, generalized quantiles are defined implicitly by means of Eq. (6). Letting

$$\psi(t) := \begin{cases} -(1-\alpha)\Phi'_2(-t) & t<0 \\ \alpha\Phi'_1(t) & t\geq 0, \end{cases} \tag{7}$$

we have that $\psi : \mathbb{R} \to \mathbb{R}$ is strictly increasing, $\psi(0) = 0$ and x^*_α is the unique solution of the equation

$$E[\psi(X-x^*_\alpha)] = 0.$$

Thus generalized quantiles are special cases of the *zero utility premium principle*, very well known in the actuarial literature (see for example Deprez and Gerber, 1985), although we remark that in our setting the function ψ is not necessarily convex. Apart from an unessential sign change, Ben-Tal and Teboulle (2007) called the same quantities *u-mean certainty equivalents*. Föllmer and Schied (2002) introduced *shortfall risk measures*, defined by

$$\rho_{FS}(X) = \inf\{m\in\mathbb{R} \mid E[l(-m-X)] \leq x_0\},$$

where $l : \mathbb{R} \to \mathbb{R}$ is nondecreasing and convex. In particular, for $x_0 = 0$ the shortfall risk measure can also be equivalently defined as the unique solution of

$$E[l(-\rho_{FS}(X) - X)] = 0.$$

It follows that generalized quantiles can also be seen as special cases of shortfall risk measures; we will exploit the results of Föllmer and Schied (2002) in the next section, in order to compute the dual representation of expectiles.

In the following proposition, we collect some elementary properties of generalized quantiles.

Proposition 5. *Let $\Phi_1, \Phi_2 : [0,+\infty) \to [0,+\infty)$ be convex, strictly increasing and satisfy (2). Let $X \in M^{\Phi_1} \cap M^{\Phi_2}$, $\alpha \in (0,1)$ and let $\pi_\alpha(X,x)$ be as in (4). We denote with $y^{*-}_\alpha, y^{*+}_\alpha$ the (lower and upper) generalized quantiles of Y. Then the following holds.*

(a) *translation equivariance: if $Y = X + h$ with $h \in \mathbb{R}$ then*
$$[y^{*-}_\alpha; y^{*+}_\alpha] = [x^{*-}_\alpha + h; x^{*+}_\alpha + h];$$
(b) *positive homogeneity: if $\Phi_1(x) = \Phi_2(x) = x^\beta$, with $\beta \geq 1$, then*
$$Y = \lambda X \quad \text{for } \lambda > 0 \Rightarrow [y^{*-}_\alpha; y^{*+}_\alpha] = [\lambda x^{*-}_\alpha; \lambda x^{*+}_\alpha];$$
(c) *monotonicity: if $X \geq_{st} Y$, then $x^{*-}_\alpha \geq y^{*-}_\alpha$ and $x^{*+}_\alpha \geq y^{*+}_\alpha$;*
(d) *constancy: if $X = c$, P-a.s., then $x^{*-}_\alpha = x^{*+}_\alpha = c$;*
(e) *internality: if $X \in L^\infty$, then $[x^{*-}_\alpha; x^{*+}_\alpha] \in [\operatorname{ess\,inf}(X); \operatorname{ess\,sup}(X)]$;*
(f) *monotonicity in α: if $\alpha_1 \leq \alpha_2$, with $\alpha_1, \alpha_2 \in (0,1)$, then $x^{*-}_{\alpha_1} \leq x^{*-}_{\alpha_2}$ and $x^{*+}_{\alpha_1} \leq x^{*+}_{\alpha_2}$.*

Proof. (a) and (b) follow immediately from

$$\pi_\alpha(X+h,x)=\pi_\alpha(X,x-h),\quad \text{for any } h\in\mathbb{R},$$

and

$$\pi_\alpha(\lambda X,\lambda x)=\lambda^\beta\pi_\alpha(X,x),\quad \text{for any } \lambda>0,$$

respectively.

(c) From (5), it is easy to see that the functions $g^-(X,x,\alpha):=\frac{\partial^-\pi_\alpha}{\partial x}$ and $g^+(X,x,\alpha):=\frac{\partial^+\pi_\alpha}{\partial x}$ are nonincreasing in X and α and nondecreasing in x. Moreover, we have that

$$x_\alpha^{*-}=\inf\{x\in\mathbb{R}:g^+(X,x,\alpha)\geq 0\},$$
$$x_\alpha^{*+}=\sup\{x\in\mathbb{R}:g^-(X,x,\alpha)\leq 0\}.$$

Suppose now that $X\geq_{\text{st}}Y$. Since $g^+(Y,x,\alpha)\geq g^+(X,x,\alpha)$ and $g^-(Y,x,\alpha)\geq g^-(X,x,\alpha)$, it follows that $x_\alpha^{*-}\geq y_\alpha^{*-}$ and $x_\alpha^{*+}\geq y_\alpha^{*+}$.

(d) If $X=c$, P- a.s., then it is straightforward to check that $\pi_\alpha(X,x)=0$ holds if and only if $x=c$.

(e) follows immediately from (c) and (d), while (f) can be proved similarly to (c). ∎

In the next proposition we characterize those generalized quantiles that have reasonable properties in the sense of the axiomatic theory of risk measures. In particular, we characterize generalized quantiles that are positively homogeneous and generalized quantiles that are convex. As a consequence, we will show that the only generalized quantiles that are coherent risk measures are the expectiles with $\alpha\geq\frac{1}{2}$. Similar characterizations have been obtained by Ben-Tal and Teboulle (2007) and Weber (2006) under slightly different hypotheses, so we provide an independent proof.

Proposition 6. *Let $\Phi_1,\Phi_2:[0,+\infty)\to[0,+\infty)$ be strictly convex and differentiable with $\Phi_i(0)=0$, $\Phi_i(1)=1$ and $\Phi_{1+}'(0)=\Phi_{2+}'(0)=0$. Let $\alpha\in(0,1)$ and*

$$x_\alpha^*(X)=\operatorname*{argmin}_{x\in\mathbb{R}}\left\{\alpha E\left[\Phi_1(X-x)^+\right]+(1-\alpha)E\left[\Phi_2(X-x)^-\right]\right\}.$$

(a) *$x_\alpha^*(X)$ is positively homogeneous if and only if $\Phi_1(x)=\Phi_2(x)=x^\beta$, with $\beta>1$.*
(b) *$x_\alpha^*(X)$ is convex if and only if the function $\psi:\mathbb{R}\to\mathbb{R}$ defined in (7) is convex; it is concave if and only if ψ is concave.*
(c) *$x_\alpha^*(X)$ is coherent if and only if $\Phi_1(x)=\Phi_2(x)=x^2$ and $\alpha\geq\frac{1}{2}$.*

Proof. (a) Due to the assumptions made on Φ_1 and Φ_2, the minimizer $x_\alpha^*(X)$ is the unique solution of

$$E[\psi(X-x_\alpha^*)]=0. \tag{8}$$

Assume now that x_α^* is positively homogeneous. Let $\delta<0<\gamma$ and

$$\overline{X}=\begin{cases}\delta, & \text{with prob. } \overline{p}\\ \gamma, & \text{with prob. } 1-\overline{p},\end{cases}$$

with

$$\overline{p}=\frac{\psi(\gamma)}{\psi(\gamma)-\psi(\delta)}.$$

Since

$$E[\psi(\overline{X})]=\overline{p}\psi(\delta)+(1-\overline{p})\psi(\gamma)=0,$$

it follows immediately by (8) and by the uniqueness of $x_\alpha^*(\overline{X})$ that

$$x_\alpha^*(\overline{X})=0.$$

If x_α^* is positively homogeneous, then we must have for every $\lambda>0$

$$x_\alpha^*(\lambda\overline{X})=0,$$

that is

$$\overline{p}\psi(\lambda\delta)+(1-\overline{p})\psi(\lambda\gamma)=0.$$

Denoting by ψ_1 and ψ_2 the restrictions of ψ to the domains $(-\infty,0)$ and $(0,+\infty)$, respectively, the previous equation can be written as

$$\frac{\psi_1(\lambda\delta)}{\psi_2(\lambda\gamma)}=\frac{\psi_1(\delta)}{\psi_2(\gamma)}$$

for every $\delta<0<\gamma$ and for every $\lambda>0$.

For $\gamma=1$ we get

$$\psi_1(\lambda\delta)=\frac{\psi_1(\delta)\psi_2(\lambda)}{\psi_2(1)},$$

which is a Pexider functional equation. From Theorem 4 in Aczel (1966) it follows that $\psi_1(-\lambda)=\psi_1(-1)\lambda^c$ and $\psi_2(\lambda)=\psi_2(1)\lambda^c$, for some $c>0$.

The condition $\Phi_1(x)=\Phi_2(x)=x^\beta$, with $\beta>1$, follows by integrating Eq. (7). The reverse implication is due to Proposition 5(b).

(b) The convexity of the zero utility principles is well known in the actuarial literature (see for example Deprez and Gerber, 1985). For the sake of completeness we provide a direct proof. Due to the f.o.c. $x_\alpha^*(X)$ is the unique solution of

$$E[\psi(X-x_\alpha^*(X))]=0;$$

furthermore, the term on the left side of the equality above is a nonincreasing function of $x_\alpha^*(X)$. If ψ is convex, it follows that

$$\begin{aligned}&E[\psi(\lambda X+(1-\lambda)Y-\lambda x_\alpha^*(X)-(1-\lambda)x_\alpha^*(Y))]\\&\quad=E[\psi(\lambda(X-x_\alpha^*(X))+(1-\lambda)(Y-x_\alpha^*(Y)))]\\&\quad\leq\lambda E[\psi(X-x_\alpha^*(X))]+(1-\lambda)E[\psi(Y-x_\alpha^*(Y))]=0,\end{aligned}$$

from which we get

$$x_\alpha^*(\lambda X+(1-\lambda)Y)\leq x_\alpha^*(\lambda X)+(1-\lambda)x_\alpha^*(Y). \tag{9}$$

If ψ is concave, then the opposite inequality holds, so that

$$x_\alpha^*(\lambda X+(1-\lambda)Y)\geq x_\alpha^*(\lambda X)+(1-\lambda)x_\alpha^*(Y). \tag{10}$$

It remains to prove the reverse inequality in (9) (as well as in (10)). We prove it by contradiction, with an argument similar to Weber (2006). Hence, suppose that ψ is not convex; then there exist $x,y\in\mathbb{R}$ such that $\psi(\frac{x+y}{2})>\frac{\psi(x)}{2}+\frac{\psi(y)}{2}$. Therefore there must exist $z\in\mathbb{R}$ and $\alpha\in(0,1)$ such that

$$\begin{aligned}&\alpha\psi(z)+(1-\alpha)\left(\frac{\psi(x)}{2}+\frac{\psi(y)}{2}\right)<0\\&\quad\leq\alpha\psi(z)+(1-\alpha)\psi\left(\frac{x+y}{2}\right).\end{aligned}$$

Consider now two random variables X,Y satisfying

$$P(X=z,Y=z)=\alpha,$$

$$P(X=x,Y=y)=P(X=y,Y=x)=\frac{1-\alpha}{2}.$$

It is clear that $E[\psi(X)]=E[\psi(Y)]<0$, while $E[\psi(\frac{X+Y}{2})]\geq 0$. Hence, $x_\alpha^*(X),x_\alpha^*(Y)<0$ and $x_\alpha^*(\frac{X+Y}{2})\geq 0$, so that x_α^* cannot be convex.

(c) From (b), ψ has to be convex. This implies that Φ_2' has to be concave and Φ_1' has to be convex. Moreover, from (a) it follows that Φ_1 and Φ_2 should necessarily satisfy $\Phi_1(x)=\Phi_2(x)=x^\beta$ for some $\beta>1$. Putting together the two conditions above, we get $\beta=2$. To ensure convexity of ψ we have to require additionally that $\alpha\geq 1/2$. ∎

Bellini (2012) showed by means of order theoretic comparative static techniques that expectiles with $\alpha\geq\frac{1}{2}$ are the only generalized quantiles (in the symmetric $\Phi_1=\Phi_2$ case) that are isotonic with respect to the increasing convex order, in accordance with the results of Bäuerle and Müller (2006).

Before moving to a deeper study of expectiles, we mention that several alternative generalized quantiles that cannot be cast in the form of the minimization problem (3) have been considered in the literature. Bellini and Rosazza Gianin (2012) defined Orlicz quantiles as the minimizers of the following loss function:

$$\pi_\alpha(X,x) := \alpha \left\| (X-x)^+ \right\|_{\Phi_1} + (1-\alpha) \left\| (X-x)^- \right\|_{\Phi_2},$$

while Jaworski (2006) considered instead loss functions of the type

$$\pi_\alpha(X,x) := E\left[\Phi_1(X-x)^+\right] + \Phi_2(x),$$

that depend only on the right tail of the distribution. For the sake of comparison, we recall that Orlicz quantiles are always positively homogeneous, while generalized quantiles have this property only in the case $\Phi_1(x) = \Phi_2(x) = x^p$. Orlicz quantiles in general lack the property of monotonicity with respect to the $\leq_{st}$ order, as was shown by means of a counterexample in Bellini and Rosazza Gianin (2012). Jaworski's quantiles are monotonic and (under additional assumptions) convex, but they do not satisfy translation invariance nor positive homogeneity.

Another remarkable issue is that the value function

$$\pi_\alpha(X) = \inf_{x\in\mathbb{R}} \pi_\alpha(X,x)$$

is a *convex deviation measure* in the sense of Rockafellar et al. (2006) in the case of generalized quantiles, while it is a *coherent deviation measure* in the case of Orlicz quantiles (see Bellini and Rosazza Gianin, 2012; Rockafellar et al., 2006).

3. Expectiles

We devote this section to the study of expectiles $e_\alpha(X)$. As we have shown in Proposition 6, they are the only generalized quantiles that are coherent risk measures (in the case of $\alpha \geq \frac{1}{2}$). The f.o.c. for expectiles can be written in several equivalent ways:

$$\alpha E\left[(X - e_\alpha(X))^+\right] = (1-\alpha) E\left[(X - e_\alpha(X))^-\right] \tag{11}$$

or also

$$e_\alpha(X) - E[X] = \frac{2\alpha - 1}{1-\alpha} E[(X - e_\alpha(X))^+] \tag{12}$$

(see Newey and Powell, 1987). We remark that Eqs. (11) and (12) have a unique solution for all $X \in L^1$, so expectiles are well defined on L^1, although the loss function (4) can assume the value $+\infty$. From Propositions 5 and 6 we know that expectiles satisfy translation equivariance, positive homogeneity, monotonicity with respect to the $\leq_{st}$ order and subadditivity (when $\alpha \geq \frac{1}{2}$). In the next proposition we collect some further immediate properties.

Proposition 7. *Let $X, Y \in L^1$ and let $e_\alpha(X)$ be the α-expectile of X. Then:*

(a) *$X \leq Y$ and $P(X < Y) > 0$ imply that $e_\alpha(X) < e_\alpha(Y)$ (strong monotonicity);*
(b) *if $\alpha \leq \frac{1}{2}$, then $e_\alpha(X+Y) \geq e_\alpha(X) + e_\alpha(Y)$;*
(c) *$e_\alpha(X) = -e_{1-\alpha}(-X)$.*

Proof. If $X < Y$ on a set A with $P(A) > 0$, then at least one of the sets $A \cap \{X > e_\alpha(X)\}$ and $A \cap \{X < e_\alpha(X)\}$ must have strictly positive measure. This implies that either $E\left[(Y - e_\alpha(X))^+\right] > E\left[(X - e_\alpha(X))^+\right]$ or $E\left[(Y - e_\alpha)^-\right] < E\left[(X - e_\alpha)^-\right]$. Since $e_\alpha(X)$ satisfies (11), we have that

$$\alpha E\left[(Y - e_\alpha(X))^+\right] - (1-\alpha) E\left[(Y - e_\alpha(X))^-\right] > 0,$$

which implies $e_\alpha(X) < e_\alpha(Y)$.

(b) can be seen as a consequence of Proposition 6. We also provide an alternative direct proof. Denote, for simplicity, $e_\alpha(X) := x$ and $e_\alpha(Y) := y$. If $\alpha \leq \frac{1}{2}$, we have

$$\begin{aligned}
&\alpha E[(X+Y-x-y)^+] - (1-\alpha)E[(X+Y-x-y)^-] \\
&\quad = (2\alpha - 1)E[(X+Y-x-y)^+] \\
&\qquad + (1-\alpha)E[X-x] + (1-\alpha)E[Y-y] \\
&\quad = (2\alpha - 1)E[(X+Y-x-y)^+] \\
&\qquad - (2\alpha - 1)E[(X-x)^+] - (2\alpha-1)E[(Y-y)^+] \\
&\quad = (2\alpha - 1)\left\{E[(X+Y-x-y)^+] \right. \\
&\qquad \left. - E[(X-x)^+] - E[(Y-y)^+]\right\} \geq 0.
\end{aligned}$$

Since the function

$$g(X,x,\alpha) := \alpha E[(X-x)^+] - (1-\alpha)E[(X-x)^-]$$

is nonincreasing in x, it follows that $e_\alpha(X+Y) \geq e_\alpha(X) + e_\alpha(Y)$. For $\alpha \geq \frac{1}{2}$, the opposite inequality holds.

(c) The expectile $e_{1-\alpha}(-X)$ satisfies the f.o.c.

$$(1-\alpha)E\left[(-X - e_{1-\alpha}(-X))^+\right] = \alpha E\left[(-X - e_{1-\alpha}(-X))^-\right].$$

This is equivalent to

$$\alpha E\left[(X + e_{1-\alpha}(-X))^+\right] = (1-\alpha) E\left[(X + e_{1-\alpha}(-X))^-\right].$$

Due to the uniqueness of the solution of (11), it follows that $e_{1-\alpha}(-X) = -e_\alpha(X)$. ■

So expectiles satisfy a strict monotonicity property that is stronger than the usual monotonicity property of coherent risk measures. It is easy to see that for example CVaR is not strictly monotone, since it depends only on the right tail of the distribution. Strict monotonicity will be reflected in the dual representation of the next section by the strict positivity of the generalized scenarios.

3.1. Dual representation

Since the expectiles are special cases of shortfall risk measures or u-mean certainty equivalents, their dual representation can be obtained by means of the results in Föllmer and Schied (2002) or in Ben-Tal and Teboulle (2007), as was noticed by Müller (2010). The following result provides the dual representation of expectiles (as maximal — or minimal — expected value over a set of scenarios) as well as the explicit expression of the optimal scenario.

Proposition 8. *Let $X \in L^1$, and let $e_\alpha(X)$ be the α-expectile of X. Then*

$$e_\alpha(X) = \begin{cases} \max\limits_{\varphi \in \mathcal{M}_\alpha} E[\varphi X], & \text{if } \alpha \geq \dfrac{1}{2} \\ \min\limits_{\varphi \in \mathcal{M}_\alpha} E[\varphi X], & \text{if } \alpha \leq \dfrac{1}{2} \end{cases}, \tag{13}$$

where the set of scenarios is

$$\mathcal{M}_\alpha = \left\{ \varphi \in L^\infty : \varphi > 0 \text{ a.s.}, E_P[\varphi] = 1, \frac{\operatorname{ess\,sup} \varphi}{\operatorname{ess\,inf} \varphi} \leq \beta \right\},$$

with $\beta = \max\left\{\frac{\alpha}{1-\alpha}, \frac{1-\alpha}{\alpha}\right\}$. Moreover, the optimal scenario $\overline{\varphi}$ in the dual representation (13) is given by

$$\overline{\varphi} := \frac{\alpha 1_{\{X > e_\alpha\}} + (1-\alpha) 1_{\{X \leq e_\alpha\}}}{E[\alpha 1_{\{X > e_\alpha\}} + (1-\alpha) 1_{\{X \leq e_\alpha\}}]}. \tag{14}$$

Proof. The f.o.c. for an α-expectile can be written as $E[\psi(X - e_\alpha(X))] = 0$, where ψ is given by

$$\psi(t) = \begin{cases} 2\alpha t, & \text{if } t \geq 0 \\ 2(1-\alpha)t, & \text{if } t < 0. \end{cases}$$

Let us assume that $\alpha \geq \frac{1}{2}$. In this case, ψ is convex. From Proposition 4.104 and Theorem 4.106 in Föllmer and Schied (2004), we get that

$$e_\alpha(X) = \max_{\varphi \in \mathcal{M}} \{E[\varphi X] - \alpha_{\min}(\varphi)\},$$

where

$$\mathcal{M} = \left\{\varphi \in L^\infty \text{ s.t. } \varphi \geq 0 \text{ a.s.}, E_P[\varphi] = 1\right\}$$

and

$$\alpha_{\min}(\varphi) = \inf_{\lambda > 0} \left\{\frac{1}{\lambda} E[\psi^*(\lambda\varphi)]\right\}.$$

Since ψ is positively homogeneous, its convex conjugate ψ^* is given by

$$\psi^*(t) = \begin{cases} 0, & \text{if } 2(1-\alpha) \leq \psi \leq 2\alpha \\ +\infty, & \text{otherwise.} \end{cases}$$

Hence, it follows that

$$\alpha_{\min}(\varphi) = \begin{cases} 0, & \text{if } 2(1-\alpha) \leq \lambda\varphi \leq 2\alpha \text{ for some } \lambda > 0 \\ +\infty, & \text{otherwise} \end{cases}$$

which implies the dual representation (13). If $\alpha \leq \frac{1}{2}$, we get from Proposition 7 (c)

$$e_\alpha(X) = -e_{1-\alpha}(-X) = -\max_{\varphi \in \mathcal{M}_{1-\alpha}} E[\varphi(-X)] = \min_{\varphi \in \mathcal{M}_\alpha} E[\varphi X].$$

It remains to check the optimality of $\overline{\varphi}$, that is $e_\alpha(X) = E[\overline{\varphi} X]$. We first note that

$$\frac{\operatorname{ess\,sup} \overline{\varphi}}{\operatorname{ess\,inf} \overline{\varphi}} = \max\left\{\frac{\alpha}{1-\alpha}, \frac{1-\alpha}{\alpha}\right\} = \beta,$$

i.e. $\overline{\varphi} \in \mathcal{M}_\alpha$. The optimality of $\overline{\varphi}$ then follows from

$$\begin{aligned} E[\overline{\varphi} X] &= \frac{E[\alpha X 1_{\{X > e_\alpha\}} + (1-\alpha) X 1_{\{X \leq e_\alpha\}}]}{E[\alpha 1_{\{X > e_\alpha\}} + (1-\alpha) 1_{\{X \leq e_\alpha\}}]} \\ &= e_\alpha(X) \\ &\quad + \frac{E[\alpha(X - e_\alpha(X)) 1_{\{X > e_\alpha\}} + (1-\alpha)(X - e_\alpha(X)) 1_{\{X \leq e_\alpha\}}]}{E[\alpha 1_{\{X > e_\alpha\}} + (1-\alpha) 1_{\{X \leq e_\alpha\}}]} \\ &= e_\alpha(X), \end{aligned}$$

where the last equality is due to the first-order condition. ■

From the dual representation (13) it is possible to derive the so-called Kusuoka representation (see for example Kusuoka (2001) or Pichler and Shapiro (2013) and the references therein), that is basically the representation of a law invariant coherent risk measure as a supremum of convex combinations of CVaR. As a consequence, we have a lower bound for the expectiles. More specifically, we get the following result.

Proposition 9. *Let $X \in L^1$, $\alpha \geq \frac{1}{2}$ and $\beta = \frac{\alpha}{1-\alpha}$. Then*

$$e_\alpha(X) = \max_{\gamma \in \left[\frac{1}{\beta}, 1\right]} \left\{(1-\gamma) CVaR_{\frac{\beta - \frac{1}{\gamma}}{\beta - 1}} + \gamma E[X]\right\}.$$

In particular,

$$e_\alpha(X) \geq \frac{E[X]}{2\alpha} + \left(1 - \frac{1}{2\alpha}\right) CVaR_\alpha(X).$$

Proof. From the dual representation (13)

$$e_\alpha(X) = \max_{\varphi \in \mathcal{M}_\alpha} E[\varphi X] = \max_{\gamma \in \left[\frac{1}{\beta}, 1\right]} \max_{\varphi \in \mathcal{M}_{\gamma, \beta\gamma}} E[\varphi X],$$

where for $0 < \gamma < 1 < \delta$ we define

$$\mathcal{M}_{\gamma,\delta} := \{\varphi \in L^\infty, E[\varphi] = 1, \gamma < \varphi < \delta\}$$

and we note that

$$\begin{aligned} \max_{\varphi \in \mathcal{M}_{\gamma,\delta}} E[\varphi X] &= \max_{\varphi \in \mathcal{M}_{\gamma,\delta}} E[\gamma X + (\varphi - \gamma) X] \\ &= \gamma E[X] + (1-\gamma) \max_{\varphi \in \mathcal{M}_{\gamma,\delta}} E\left[\left(\frac{\varphi - \gamma}{1-\gamma}\right) X\right] \\ &= \gamma E[X] + (1-\gamma) \max_{\widetilde{\varphi} \in \mathcal{M}_{0, \frac{\delta-\gamma}{1-\gamma}}} E[\widetilde{\varphi} X] \\ &= \gamma E[X] + (1-\gamma) CVaR_{\frac{\delta-1}{\delta-\gamma}}(X). \end{aligned}$$

The last equality follows from

$$CVaR_\alpha(X) = \max_{\widetilde{\varphi} \in \mathcal{M}_{0, \frac{1}{1-\alpha}}} E[\widetilde{\varphi} X].$$

Since in our case $\delta = \beta\gamma$, we get

$$e_\alpha(X) = \max_{\gamma \in \left(\frac{1}{\beta}, 1\right]} \left\{(1-\gamma) CVaR_{\frac{\beta - \frac{1}{\gamma}}{\beta - 1}} + \gamma E[X]\right\};$$

putting $\gamma = \frac{1}{2\alpha}$ gives the desired inequality. ■

3.2. Robustness properties of expectiles

The important topic of robustness properties of risk measures has recently found increasing interest; see e.g. Krätschmer et al. (2012a,b) and Stahl et al. (2012). As already observed in these papers, one cannot expect robustness properties with respect to weak convergence for reasonable risk measures like CVaR or, more generally, for coherent risk measures. This also applies to expectiles, as they also contain the expectation as a limiting case, and expectation is not continuous with respect to weak convergence. Therefore we need a stronger notion of convergence, and a natural candidate is the classical Wasserstein distance. We review here some of its properties, as can be found e.g. in Bickel and Freedman (1981) or Dudley (2002).

The Wasserstein distance of two probability measures P and Q is defined as the smallest expected difference between random variables with these distributions, i.e.

$$d_W(P, Q) := \inf\{E[|X - Y|] : X \sim P,\ Y \sim Q\}. \tag{15}$$

We will also write $d_W(X, Y) := d_W(P_X, P_Y)$ for random variables X, Y with distributions P_X, P_Y. Many equivalent characterizations of this distance are available in the literature. The following ones will be useful in our context. First,

$$d_W(X, Y) = \int_{-\infty}^{\infty} |F_X(t) - F_Y(t)| dt = \int_0^1 |F_X^{-1}(\alpha) - F_Y^{-1}(\alpha)| d\alpha,$$

where F_X denotes the cumulative distribution function of X, and F_X^{-1} the quantile function. Second,

$$d_W(X, Y) = \sup\{|E[f(X)] - E[f(Y)]| : \|f\|_L \leq 1\}, \tag{16}$$

where

$$\|f\|_L := \sup_{x,y \in \mathbb{R}: x \neq y} \frac{|f(x) - f(y)|}{|x - y|}$$

denotes the Lipschitz-seminorm.

We also recall that convergence in the Wasserstein distance is stronger than convergence in distribution; more precisely,

$d_W(X_n, X) \to 0$ if and only if

$X_n \to X$ in distribution and $E[X_n] \to E[X]$.

Thus, convergence in the Wasserstein distance is a special case of the ψ-weak convergence as considered by Krätschmer et al. (2012a) with $\psi(x) = |x|$.

We are now ready to state (and prove) the following quantitative result about robustness of expectiles with respect to the Wasserstein distance. Pflug and Wozabal (2007) proved a similar result for CVaR, while a discussion of the same property for general coherent risk measures can be found in Pichler (forthcoming).

Theorem 10. *For all $X, Y \in L^1$ and all $\alpha \in (0, 1)$ it holds that $|e_\alpha(X) - e_\alpha(Y)| \leq \beta d_W(X, Y)$, where $\beta = \max\{\frac{\alpha}{1-\alpha}; \frac{1-\alpha}{\alpha}\}$.*

Proof. Take now $\alpha > \frac{1}{2}$ and define the functions

$$f(x) := x - E[X]; \qquad g(x) := \frac{2\alpha - 1}{1-\alpha} E[(X - x)^+].$$

The Lipschitz-seminorms of f and g are then equal to

$$\|f\|_L = 1 \quad \text{and} \quad \|g\|_L = \frac{2\alpha - 1}{1-\alpha}.$$

According to Eq. (12), the expectile $e_\alpha(X)$ of X solves the equation

$$\begin{aligned} f(e_\alpha(X)) &= e_\alpha(X) - E[X] \\ &= \frac{2\alpha-1}{1-\alpha} E[(X - e_\alpha(X))^+] = g(e_\alpha(X)). \end{aligned} \tag{17}$$

Assume now $d_W(X, Y) \leq \epsilon$ and w.l.o.g. $e_\alpha(Y) \geq e_\alpha(X)$. By (16) it follows that $|E[Y] - E[X]| \leq \epsilon$ and $|E[(Y - e_\alpha(X))^+] - E[(X - e_\alpha(X))^+]| \leq \epsilon$. Hence,

$$E[X] = E[Y] + \epsilon_1 \quad \text{and}$$
$$E[(X - e_\alpha(X))^+] = E[(Y - e_\alpha(X))^+] + \epsilon_2$$

with $|\epsilon_1|, |\epsilon_2| \leq \epsilon$. From (17), we deduce that

$$e_\alpha(X) - E[Y] + \epsilon_1 = \frac{2\alpha-1}{1-\alpha}(E[(Y - e_\alpha(X))^+] + \epsilon_2). \tag{18}$$

Replacing X by Y, and using (17) again, we get

$$e_\alpha(Y) - E[Y] = \frac{2\alpha-1}{1-\alpha} E[(Y - e_\alpha(Y))^+]. \tag{19}$$

Subtracting Eq. (18) from (19) and rearranging terms yields

$$\begin{aligned} e_\alpha(Y) - e_\alpha(X) &= \epsilon_1 + \frac{2\alpha-1}{1-\alpha}(E[(Y - e_\alpha(Y))^+] \\ &\quad - E[(Y - e_\alpha(X))^+] - \epsilon_2) \\ &\leq \epsilon \cdot \left(1 + \frac{2\alpha-1}{1-\alpha}\right) \\ &= \epsilon \cdot \frac{\alpha}{1-\alpha}, \end{aligned}$$

where the inequality follows from $e_\alpha(Y) \geq e_\alpha(X)$.

Let now $\alpha < \frac{1}{2}$. The result follows immediately by applying the inequality just proved to $e_{1-\alpha}(-X)$, since $e_\alpha(X) = -e_{1-\alpha}(-X)$ (by Proposition 7(c)) and $d_W(-X, -Y) = d_W(X, Y)$. The case $\alpha = 1/2$ is trivial as, in that case, $e_\alpha(X) = E[X]$. ∎

3.3. Comparing expectiles with quantiles

In the statistical literature (see for example Jones, 1994) it is often argued that typically the expectiles e_α are closer to the center of a distribution than the corresponding quantiles q_α. For example, this is indeed the case for the uniform or the standard normal distribution. We are not aware of a general comparison result; however Koenker (1993) has derived an example of a distribution where the expectiles $e_\alpha(X)$ and the quantiles $q_\alpha(X)$ *coincide* for all $\alpha \in (0, 1)$. It turns out that this distribution is a Pareto-like distribution with tail index $\beta = 2$, which in particular means that it has infinite variance. Here we say that a distribution is Pareto-like with tail index β if

$$F_X(x) = 1 - L(x) \cdot x^{-\beta}$$

for some function L which is slowly varying at infinity, i.e. with

$$\lim_{x\to\infty} \frac{L(tx)}{L(x)} = 1 \quad \text{for all } t > 0,$$

see e.g. Bingham et al. (1987).

We can prove an asymptotic comparison result (as $\alpha \to 1$) that shows that actually Koenker's case $\beta = 2$ discriminates between situations in which asymptotically expectiles are larger than quantiles ($\beta < 2$) and situations in which quantiles are larger than expectiles ($\beta > 2$). In other words, at least for large α, we can say that expectiles are a more conservative risk measure than the quantiles for (extremely) heavy tailed distributions.

Theorem 11. *Assume that X is a Pareto-like distribution with tail index $\beta > 1$. Then*

$$\frac{\overline{F}(e_\alpha(X))}{\beta - 1} \sim 1 - \alpha \quad \text{as } \alpha \to 1. \tag{20}$$

If $\beta < 2$, then there exists some $\alpha_0 < 1$ such that for all $\alpha > \alpha_0$ it holds $e_\alpha(X) > q_\alpha(X)$. If $\beta > 2$, then there exists some $\alpha_0 < 1$ such that for all $\alpha > \alpha_0$ it holds $e_\alpha(X) < q_\alpha(X)$.

Proof. From Eq. (11) it is easy to see that if X has a Pareto-like distribution, then $e_\alpha(X) \to +\infty$ as $\alpha \to 1$. From the Karamata theorem (see for example Proposition 1.5.10 in Bingham et al., 1987) it holds that when $x \to +\infty$,

$$E[(X - x)^+] \sim \frac{x\overline{F}(x)}{\beta - 1}.$$

Since by Eq. (12), any expectile solves

$$e_\alpha(X) - E[X] = \frac{2\alpha-1}{1-\alpha} E[(X - e_\alpha(X))^+],$$

we can conclude that as $\alpha \to 1$

$$e_\alpha(X) - E[X] \sim \frac{e_\alpha(X)}{1-\alpha} \frac{\overline{F}(e_\alpha(X))}{\beta - 1},$$

that gives

$$e_\alpha(X) \sim \frac{e_\alpha(X)}{1-\alpha} \frac{\overline{F}(e_\alpha(X))}{\beta - 1},$$

from which the first part of the theorem follows. Since clearly $\overline{F}(q_\alpha(X)) \sim 1 - \alpha$, then

$$\frac{\overline{F}(e_\alpha(X))}{\overline{F}(q_\alpha(X))} \sim \beta - 1, \quad \text{as } \alpha \to 1,$$

from which also the remaining part of the statement follows. ∎

Acknowledgments

We thank Georg Pflug for very helpful hints on how to compute the dual representation, and Valeria Bignozzi and Alois Pichler for helpful discussions. We are very grateful to an anonymous referee for many helpful comments and in particular for an improvement of the statement of Theorem 11.

References

Aczel, J., 1966. Lectures on Functional Equations and Their Applications. Academic Press.
Bäuerle, N., Müller, A., 2006. Stochastic orders and risk measures: consistency and bounds. Insurance: Mathematics & Economics 38, 132–148.
Bellini, F., 2012. Isotonicity results for generalized quantiles. Statistics & Probability Letters 82, 2017–2024.
Bellini, F., Rosazza Gianin, E., 2012. Haezendonck–Goovaerts risk measures and Orlicz quantiles. Insurance: Mathematics & Economics 51, 107–114.
Ben-Tal, A., Teboulle, M., 2007. An old–new concept of convex risk measures: the optimized certainty equivalent. Mathematical Finance 17, 449–476.
Bickel, P., Freedman, D., 1981. Some asymptotic theory for the bootstrap. Annals of Statistics 9, 1196–1217.
Bingham, N.H., Goldie, C.M., Teugels, J.L., 1987. Regular Variation. Cambridge University Press.
Breckling, J., Chambers, R., 1988. M-quantiles. Biometrika 75, 761–772.
Chen, Z., 1996. Conditional L^p quantiles and their application to the testing of symmetry in non-parametric regression. Statistics & Probability Letters 29, 107–115.
Deprez, O., Gerber, H.U., 1985. On convex principles of premium calculation. Insurance: Mathematics & Economics 4, 179–189.
De Rossi, G., Harvey, H., 2009. Quantiles, expectiles and splines. Journal of Econometrics 152, 179–185.
Dudley, R.M., 2002. Real Analysis and Probability. In: Cambridge Studies in Advanced Mathematics..
Embrechts, P., Hofert, M., 2013. Statistics and Quantitative Risk Management for Banking and Insurance. ETH, Zurich.
Föllmer, H., Schied, A., 2002. Convex measures of risk and trading constraints. Finance and Stochastics 6, 429–447.
Föllmer, H., Schied, A., 2004. Stochastic Finance. de Gruyter.
Gneiting, T., 2011. Making and evaluating point forecasts. Journal of the American Statistical Association 106, 746–762.
Goovaerts, M.J., Linders, D., Van Weert, K., Tank, F., 2012. On the interplay between distortion-, mean value- and Haezendonck risk measures. Insurance: Mathematics & Economics 51, 10–18.
Jaworski, P., 2006. On a subjective approach to risk measurement. Quantitative Finance 6, 495–511.
Jones, M.C., 1994. Expectiles and M-quantiles are quantiles. Statistics & Probability Letters 20, 149–153.
Koenker, R., 1993. When are expectiles percentiles? Econometric Theory 9, 526–527.
Koenker, R., 2005. Quantile Regression. Cambridge University Press.
Krätschmer, V., Schied, A., Zähle, H., 2012a. Comparative and qualitative robustness for law invariant risk measures. Available at arXiv:1204.2458.
Krätschmer, V., Schied, A., Zähle, H., 2012b. Qualitative and infinitesimal robustness of tail dependent statistical functionals. Journal of Multivariate Analysis 103, 35–47.
Kuan, C., Yeh, J., Hsu, Y., 2009. Assessing value at risk with CARE, the conditional autoregressive expectile models. Journal of Econometrics 150, 261–270.
Kusuoka, S., 2001. On law invariant coherent risk measures. Advances in Mathematical Economics 3, 83–95.
Müller, A., 2010. Expectiles as risk measure. In: Presentation at DagStat conference.
Newey, W., Powell, J., 1987. Asymmetric least squares estimation and testing. Econometrica 55, 819–847.
Pflug, G., Ruszczynski, A., 2004. A risk measure for income processes. In: Szegö, G. (Ed.), Risk Measures for the 21st Century. Wiley and Sons, pp. 249–269.
Pflug, G., Wozabal, D., 2007. Ambiguity in portfolio selection. Quantitative Finance 7, 435–442.
Pichler, A., 2013. Insurance pricing under ambiguity. European Actuarial Journal, forthcoming.
Pichler, A., Shapiro, A., 2013. Uniqueness of Kusuoka representations. ArXiv preprint, arXiv:1210.7257v4.
Rao, M.M., Ren, Z.D., 1991. Theory of Orlicz Spaces. Marcel Dekker Inc., New York.
Rockafellar, R.T., Uryasev, S., 2002. Conditional value-at-risk for general loss distributions. Journal of Banking and Finance 26, 1443–1471.
Rockafellar, R.T., Uryasev, S., Zabarankin, M., 2006. Generalized deviations in risk analysis. Finance and Stochastics 10, 51–74.
Stahl, G., Zheng, J., Kiesel, R., Rülicke, R., 2012. Conceptualizing robustness in risk management, SSRN: http://ssrn.com/abstract=2065723.
Weber, S., 2006. Distribution invariant risk measures, information, and dynamic consistency. Mathematical Finance 16, 419–441.
Ziegel, J.F., 2013. Coherence and elicitability. ArXiv preprint, arXiv:1303.1690v2.