

Superquantiles and Their Applications to Risk, Random Variables, and Regression

R. Tyrrell Rockafellar

Department of Mathematics, University of Washington, Seattle, Washington 98195,
rtr@math.washington.edu

Johannes O. Royset

Operations Research Department, Naval Postgraduate School, Monterey, California 93943,
joroyset@nps.edu

Abstract Superquantiles (also called conditional values-at-risk) are useful tools in risk modeling and optimization, with expanding roles beyond these areas. This tutorial provides a broad overview of superquantiles and their versatile applications. We see that superquantiles are as fundamental to the description of a random variable as the cumulative distribution function (cdf), they can recover the corresponding quantile function through differentiation, they are dual in some sense to superexpectations, which are convex functions uniquely defining the cdf, and they also characterize convergence in distribution. A superdistribution function defined by superquantiles leads to higher-order superquantiles as well as new measures of risk and error, with important applications in risk modeling and generalized regression.

Keywords random variables; quantiles; superquantiles; superexpectations; superdistributions; conjugate duality; stochastic dominance; measures of risk; value-at-risk; conditional value-at-risk; generalized regression

1. Introduction

Superquantiles have for some time been recognized as important tools in risk analysis and stochastic optimization. However, the concept is better known under a variety of names such as “conditional value-at-risk,” “average value-at-risk,” “tail value-at-risk,” and “expected shortfall,” with occasional minor variations in definitions. A purpose of this tutorial is to describe the expanding role of superquantiles beyond the original area of financial engineering (see, e.g., Rockafellar and Uryasev [29, 30], Acerbi and Tasche [1], Föllmer and Schied [13]), and we therefore adopt an application neutral term already promoted in Rockafellar and Royset [27, 28], Rockafellar and Uryasev [31], and Rockafellar et al. [33].

The prominence of superquantiles in risk analysis derives from its “coherency” (Artzner et al. [4]) and “regularity” (Rockafellar and Uryasev [31]) when viewed as a measure of risk of a random variable. These properties make superquantiles well suited as scalar representations of a random variable in risk-averse decision making. Moreover, superquantiles are computationally attractive because of a trade-off formula that is easily incorporated in stochastic optimization models. In fact, a superquantile of a random variable is fundamentally more stable under (parametric) perturbation of that variable than corresponding quantiles, failure probabilities, and chance constraints. Stability becomes especially important in applications, where perturbations almost always need to be considered due to incomplete distributional information about a random variable and other approximations. Superquantiles are therefore widely used in financial engineering (Yamai and Yoshida [41], Alexander et al. [2], Wang and Uryasev [39], Balbas et al. [5], Uryasev et al. [37]), but increasingly also in

other application areas such as structural engineering (Rockafellar and Royset [27], Minguez et al. [20], Haukaas et al. [15]), military operations (Commander et al. [8], Kalinchenko et al. [16], Molyboha and Zabaranin [21]), natural resources (Yamout et al. [42], Webby et al. [40]), supply chains (Tomlin [36], Verderame and Floudas [38]), and energy systems (Carrion et al. [7], Conejo et al. [9]), to mention a few. We refer to Rockafellar [26], Sarykalin et al. [35], and Krokmal et al. [19] for earlier reviews of risk measures and superquantiles, as well as for a more comprehensive list of references.

In this tutorial we briefly summarize the use of superquantiles in risk analysis and optimization, but we go beyond these applications and describe their fundamental role in characterizing random variables, convergence in distribution, and stochastic dominance. Superquantiles also play an important role in statistics, where they define a generalized regression technique that extends traditional quantile regression. Many of the insights come through convex analysis, which we highlight at numerous occasions.

Section 2 defines superquantiles of random variables and the corresponding distribution and quantile functions, which are intimately connected with superquantiles. Section 3 describes briefly the use of superquantiles in risk modeling and optimization. Section 4 shows the expanding role of superquantiles in the description of random variables. Section 5 shows that superquantiles generate a distribution function that gives rise to higher-order superquantiles as well as measures of risk and other quantities. Section 6 utilizes these results in an application to generalized regression. This tutorial is mainly based on results from Rockafellar and Royset [28] and Rockafellar et al. [33], with supporting material drawn from Rockafellar and Uryasev [31].

2. Definitions

A real-valued random variable X is traditionally characterized by its *cumulative distribution function*, which is the right-continuous function $F_X: \mathbb{R} \rightarrow [0, 1]$ given by

$$F_X(x) = \text{prob} \{X \leq x\} \quad \text{for } x \in \mathbb{R}.$$

An equivalent characterization is in terms of the left-continuous *quantile function* $Q_X: (0, 1) \rightarrow \mathbb{R}$ expressed as

$$Q_X(p) = \min \{x \mid F_X(x) \geq p\} \quad \text{for } p \in (0, 1). \quad (1)$$

Consequently, the p -quantile $Q_X(p)$ is the lowest x such that $\text{prob} \{X > x\} \leq 1 - p$. Of course, $Q_X(p) = F_X^{-1}(p)$ if F_X is strictly increasing such that its inverse exists. Although F_X and Q_X are nondecreasing, they may be discontinuous, as illustrated in Figure 1.

The correspondence between distribution functions and quantile functions is one-to-one, with F_X recoverable from Q_X by the formula

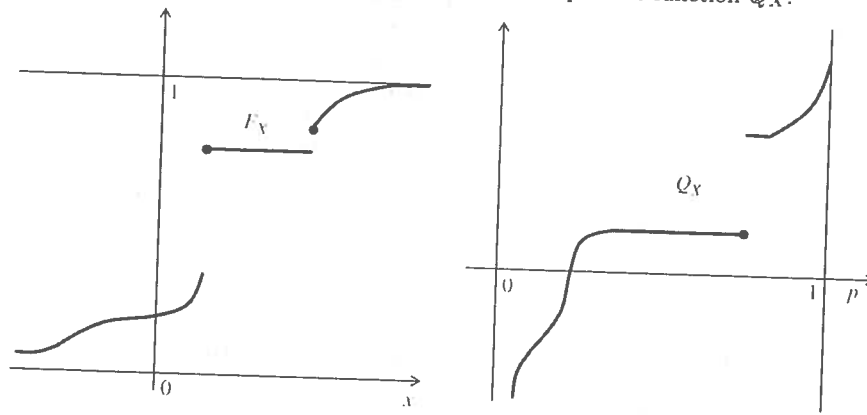
$$F_X(x) = \begin{cases} \max \{p \mid Q_X(p) \leq x\} & \text{for } x \in (\inf Q_X, \sup Q_X], \\ 1 & \text{for } x > \sup Q_X, \\ 0 & \text{for } x \leq \inf Q_X. \end{cases} \quad (2)$$

Yet another description of a random variable X comes in terms of the *superquantile function* $\bar{Q}_X: (0, 1) \rightarrow (-\infty, \infty]$, with

$$\bar{Q}_X(p) = \text{expectation in the upper } p\text{-tail distribution of } X, \text{ for } p \in (0, 1).$$

The expectation here refers to the probability distribution on $[Q_X(p), \infty)$, which, in the case of $F_X(Q_X(p)) = p$, is the conditional distribution of X subject $X \geq Q_X(p)$, but which slightly modify that conditional distribution when F_X has a jump at the quantile $Q_X(p)$. In the latter case there is a probability atom at $Q_X(p)$ causing the interval $[Q_X(p), \infty)$ to have

FIGURE 1. Distribution function F_X and quantile function Q_X .



probability larger than $1 - p$ and the interval $(Q_X(p), \infty)$ to have probability smaller than $1 - p$. To take care of the discrepancy, the p -tail distribution is defined in general as having $F_X^{[p]}(x) = \max\{0, F_X(x) - p\} / (1 - p)$ as its distribution. This amounts to an appropriate splitting of a probability atom at $Q_X(p)$.

The superquantile function is equivalently given by

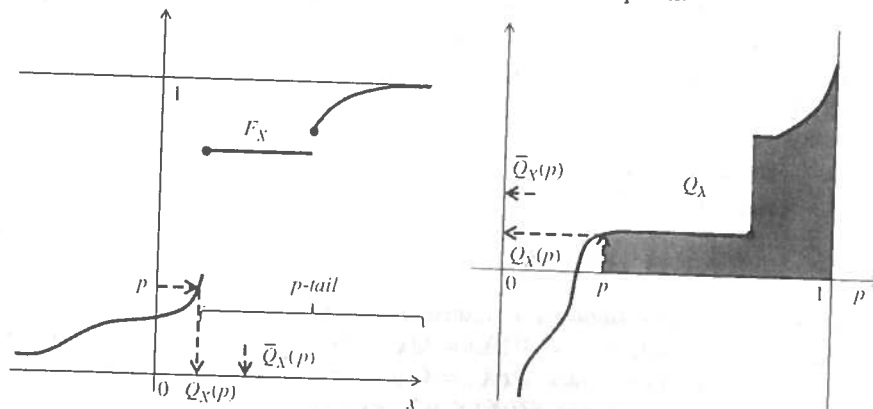
$$\bar{Q}_X(p) = \frac{1}{1-p} \int_p^1 Q_X(p') dp' \quad \text{for } p \in (0, 1). \quad (3)$$

This expression highlights a connection between quantiles and superquantiles, with the latter being an “average” of the former. We refer to Rockafellar and Royset [28] for an explanation of the equivalence between the two formulae; see also Figure 2, where the left-hand side indicates the p -tail and the right-hand side illustrates the area under the quantile function. Since Q_X is a nondecreasing, left-continuous function with at most a countable number of jumps, Q_X is Lebesgue measurable. Moreover, the integrand in (3) is bounded below by $Q_X(p)$, and therefore the integral is well defined, though the value may be infinity. For $p = 0$ and 1, it is natural to extend the definition by setting

$$\bar{Q}_X(0) = E[X] \quad \text{and} \quad \bar{Q}_X(1) = \sup X,$$

yielding a continuous function on $[0, 1]$, where the former statement requires the expectation to be well defined and the latter implies the essential supremum. Presently, it may not be

FIGURE 2. The p th quantile and the p -tail.



clear how the knowledge of \bar{Q}_X provides an equivalent characterization of X to that of F_X and Q_X , but we see in §4 that this is indeed the case.

While the definition of $\bar{Q}_X(p)$ holds for any random variable X and $p \in (0, 1)$, it is tailored to applications where high realizations of X is of particular concern. Such situations may arise when X describes “cost,” “loss,” or “damage,” and the upper-tail-centered definition of a superquantile coincides with a risk-averse decision maker’s focus on the upper tail of the distribution. Of course, a parallel development with an opposite orientation of the random variable X , focused on profits and gains, is also possible. In that case, one could define *subquantiles* as the expectation in the lower p -tail distribution of X . However, we do not pursue that topic further.

3. Superquantile in Risk Modeling and Optimization

In applications one is often faced with the need to determine whether one random variable is “adequately” small or, in comparison with another random variable, if it is “better” in some sense. This situation arises in stochastic optimization where families of parametric random variables (random functions) are compared with the goal of identifying the parameters that return the “smallest” random variable. The random variables may represent cost, loss, and damage associated some future actions, with parameters representing quantities that a decision maker can select to “shape” the probability of the various outcomes. Superquantiles, and more generally, risk measures, enable comparison across random variables.

3.1. Risk Measures

A *measure of risk* is a functional \mathcal{R} that assigns to a random variable X a value $\mathcal{R}(X)$ in $(-\infty, \infty]$ as a quantification of the risk in it. The comparison of two random variables X and Y can then be reduced to that of comparing the real numbers $\mathcal{R}(X)$ and $\mathcal{R}(Y)$. Moreover, the question of X being sufficiently small, say “adequately” no greater than 0, can then be represented by $\mathcal{R}(X) \leq 0$. The ill-posed problem of minimizing a parametric random variable subject to constraints on other parametric random variables being “adequately” small can then be interpreted as one in terms of risk measures. We refer to Krokmal et al. [19] and Rockafellar and Uryasev [31] for detailed surveys of this approach and its connection to expected utility theory.

There are several natural candidates for risk measures. The choice $\mathcal{R}(X) = E[X]$ places the focus on the “average outcome” and ignores the possible variability of X . For example, a constraint $\mathcal{R}(X) \leq 0$ then only ensures that X is no greater than zero on average. A risk-averse decision maker may steer away from this choice and consider $\mathcal{R}(X) = E[X] + \lambda\sigma(X)$ for some $\lambda > 0$, where $\sigma(X)$ denotes standard deviation. Then, the interpretation of $\mathcal{R}(X) \leq 0$ is that outcomes of X above zero can only be in the part of the distribution of X lying more than λ standard deviation above the mean, with an obvious parallel to the construction of confidence intervals in statistics. Another risk measure is to set $\mathcal{R}(X) = \sup X$ (the essential supremum of X). Although conservative, this choice may lead to infeasible demands as $\mathcal{R}(X) = \infty$ often.

Two more possibilities are based on quantiles and superquantiles, with the former being closely related to failure probability and chance constraints in stochastic optimization. By setting $\mathcal{R}(X) = Q_X(p)$, for some $p \in (0, 1)$, the relations

$$Q_X(p) \leq 0 \iff \text{prob}\{X \leq 0\} \geq p \iff \text{prob}\{X > 0\} \leq 1 - p$$

imply that $\mathcal{R}(X) \leq 0$ is equivalent to a chance constraint on X , which is frequently used in practice. In financial applications, $\mathcal{R}(X) = Q_X(p)$ is referred to as the value-at-risk of X at probability level p . The choice $\mathcal{R}(X) = \bar{Q}_X(p)$ implies, in view of the definition of a superquantile, that a requirement $\mathcal{R}(X) \leq 0$ is satisfied if and only if X is on average

below zero even in its p -tail distribution. In reliability terminology, such a requirement is equivalent to the buffered failure probability being no greater than $1 - p$ (Rockafellar and Royset [27]).

With the abundance of possible risk measures (and we refer to Rockafellar and Uryasev [31] and references therein for many more), there is a demand for guidance on what would constitute a good and useful measure of risk. There are two concepts that stand out in this regards: coherency and regularity. We discuss each in turn.

A measure of risk \mathcal{R} is *coherent* in the sense of Artzner et al. [4] (see also Delbaen [10]) if it satisfies the following axioms:

- $\mathcal{R}(C) = C$ for constant random variables $X \equiv C$,
- $\mathcal{R}(X) \leq \mathcal{R}(X')$ when $X \leq X'$ almost surely (monotonicity),
- $\mathcal{R}(X + X') \leq \mathcal{R}(X) + \mathcal{R}(X')$ (subadditivity),
- $\mathcal{R}(\lambda X) = \lambda \mathcal{R}(X)$ for $\lambda > 0$ (positive homogeneity).

For the examples above, coherency holds for the choices $\mathcal{R}(X) = E[X]$, $\mathcal{R}(X) = \sup X$, and $\mathcal{R}(X) = \bar{Q}_X(p)$, but it is absent in general for $\mathcal{R}(X) = E[X] + \lambda\sigma(X)$ with $\lambda > 0$ (because the monotonicity axiom fails) and for $\mathcal{R}(X) = Q_X(p)$ (because the subadditivity axiom fails).

A measure of risk \mathcal{R} is *regular* in the sense of Rockafellar and Uryasev [31] if it satisfies the following axioms:

- $\mathcal{R}(C) = C$ for constant random variables $X \equiv C$,
- $\mathcal{R}((1 - \tau)X + \tau X') \leq (1 - \tau)\mathcal{R}(X) + \tau\mathcal{R}(X')$ for all X, X' and $\tau \in (0, 1)$ (convexity),
- $\{X \mid \mathcal{R}(X) \leq C\}$ is closed for all $C \in \mathbb{R}$ (closedness),
- $\mathcal{R}(X) > E[X]$ for nonconstant X (aversity).

The third axiom (closedness) requires a topology on the space of random variables under consideration. Without going into technical details, we simply note that $\mathcal{R}(X) = \bar{Q}_X(p)$ is regular for $p \in (0, 1)$ when we consider the space of random variables X with $E|X| < \infty$ equipped with the L_1 -norm topology. (This fact follows by the continuity of $\bar{Q}_X(p)$ as a functional on the space of such random variables for $p \in [0, 1]$; see Theorem 3(a) below.) For the other examples above, we find that $\mathcal{R}(X) = \sup X$ and $\mathcal{R}(X) = E[X] + \lambda\sigma(X)$, with $\lambda > 0$, are regular on the space of random variables X with $E[X^2] < \infty$ equipped with the L_2 -norm topology. The choice $\mathcal{R}(X) = E[X]$ fails the aversity axiom and $\mathcal{R}(X) = Q_X(p)$ the convexity axiom. We refer to Rockafellar and Uryasev [31] for further details. The coherency and regularity axioms overlap, but are not equivalent as the above examples illustrate. However, both notions impose conditions that are practically important in modeling and implementation, with the superquantile risk measures emerging as both coherent and regular; see, for example, Artzner et al. [4] and Rockafellar and Uryasev [31] for discussions.

Although the definition of superquantiles may at first indicate difficulties in evaluating and implementing them in practice, a trade-off formula reduces the task to that of evaluating expectations and minimizing over a scalar, thereby eliminating that concern Rockafellar and Uryasev [29, 30]. Specifically,

$$\bar{Q}_X(p) = \min_x \{x + \mathcal{V}_p(X - x)\}, \quad \text{where } \mathcal{V}_p(X) = \frac{1}{1 - p} E[\max\{0, X\}], \quad (4)$$

$$Q_X(p) = \arg \min_x \{x + \mathcal{V}_p(X - x)\} \quad (\text{left endpoint, if not a singleton}). \quad (5)$$

Here, the “argmin,” consisting of the x -values for which the minimum is attained, is, in this formula, a nonempty, closed, bounded interval that typically reduces to a single point. Interestingly, the corresponding quantile of a superquantile is a byproduct of the minimization.

A consequence of the trade-off formula in stochastic optimization problems with superquantiles of parametric random variables in constraints and/or objectives is that the

trade-off formula can be substituted in each instance with an associated auxiliary variable in the overall minimization. If the random variables depend in a convex manner on the parameters, then the resulting expressions remain convex. In the absence of other complications, this may lead to a convex optimization problem for which efficient algorithms exist.

3.2. Superquantiles and the Newsvendor Problem

Whereas superquantiles are often adopted as a measure of risk when modeling a risk-averse decision maker, they also arise “naturally” in situations with trade-offs between various costs. We illustrate such a situation by considering the newsvendor problem.

A newsvendor acquires x newspapers every morning at a unit price c , sells them for b a piece, and salvages each unsold newspaper for a value a at the end of the day. Naturally, $0 \leq a < c < b$. The demand for newspapers is unknown but given by the random variable X . The newsvendor would like to determine x such that the expected cost is minimized. The newsvendor’s (random) cost, consisting of initial expense minus income from sales and salvage, is

$$cx - b \min\{X, x\} - a \max\{0, x - X\} = (c - a) \left[x + \frac{b - a}{c - a} \max\{0, X - x\} \right] + (a - b)X,$$

where the latter expression is established by simple algebra. By setting $p = (b - c)/(b - a)$, we obtain in view of (4) that the minimum expected cost is given by

$$\min_x E[cx - b \min\{X, x\} - a \max\{0, x - X\}] = (c - a)\bar{Q}_X(p) + (a - b)E[X].$$

The corresponding optimal number of newspaper is $Q_X(p)$ by (5), which of course coincides with the textbook solution for the problem. Interestingly, in this problem the probability level p of the superquantile is determined by a , b , c . These coefficients thereby impose a degree of “risk averseness” in some sense on the decision maker.

The above situation generalizes to numerous other contexts with X possibly being a parametric random variable depending on decision variables u that also must be optimized. This may lead to minimization, with respect to u , of $\bar{Q}_{X(u)}(p)$ and other terms. Again, the value of p may be determined, or at least informed, by cost parameters as in the newsvendor problem.

4. Superexpectation and Duality

Convex analysis provides deeper insight about superquantile functions and their connections with distribution and quantile functions. We start by placing F_X and Q_X in the context of monotone relations. Although these functions may have jumps (see Figure 1), the graphs obtained by filling in the vertical gaps and adding infinite vertical segments at the right and left ends of the resulting curve for Q_X , when the range of X is bounded, are *maximal monotone relations*. We let the filled-in graphs for F_X and Q_X be denoted by Γ_X and Δ_X , respectively; see Figure 3. These graphs are the reflections of each other across the line $x = p$. The relations Γ_X and Δ_X identify with subdifferentials of convex functions and therefore enable us to bring to bear the machinery of convex analysis including conjugate duality. For the remainder of this section we assume that all random variables X have $E|X| < \infty$.

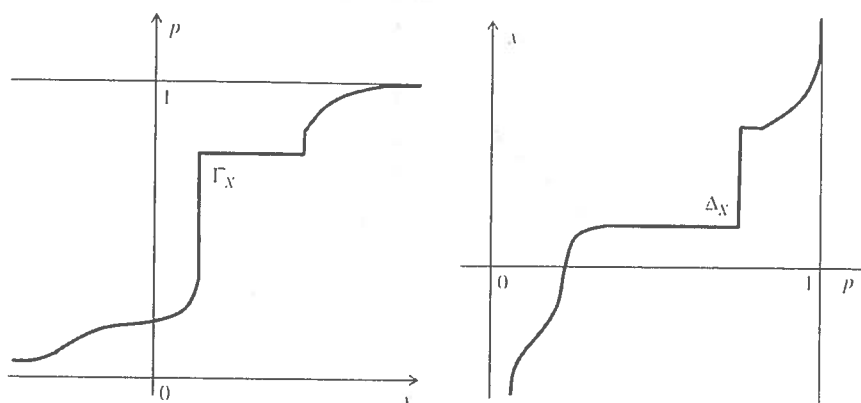
The distribution function F_X is nondecreasing and right-continuous, but we can also define a left-continuous counterpart F_X^- given for each $x \in \mathbb{R}$ by

$$F_X^-(x) = \lim_{x' \nearrow x} F_X(x').$$

Then,

$$\Gamma_X = \{(x, p) \in \mathbb{R} \times \mathbb{R} \mid F_X^-(x) \leq p \leq F_X(x)\}.$$

FIGURE 3. Relations Γ_X and Δ_X .



We proceed similarly with the nondecreasing left-continuous function Q_X , but first we need to extend it beyond $(0, 1)$ by setting

$$Q_X(1) = \lim_{p \nearrow 1} Q_X(p), \quad Q_X(p) = \infty \text{ for } p > 1, \quad Q_X(p) = -\infty \text{ for } p \leq 0,$$

which results in a nondecreasing left-continuous function on \mathbb{R} . We also define its right-continuous counterpart Q_X^+ by

$$Q_X^+(p) = \lim_{p' \searrow p} Q_X(p')$$

and find that

$$\Delta_X = \{(p, x) \in \mathbb{R} \times \mathbb{R} \mid Q_X(p) \leq x \leq Q_X^+(p)\}.$$

In view of (1) and (2), $\Delta_X = \Gamma_X^{-1}$ and $\Gamma_X = \Delta_X^{-1}$; i.e.,

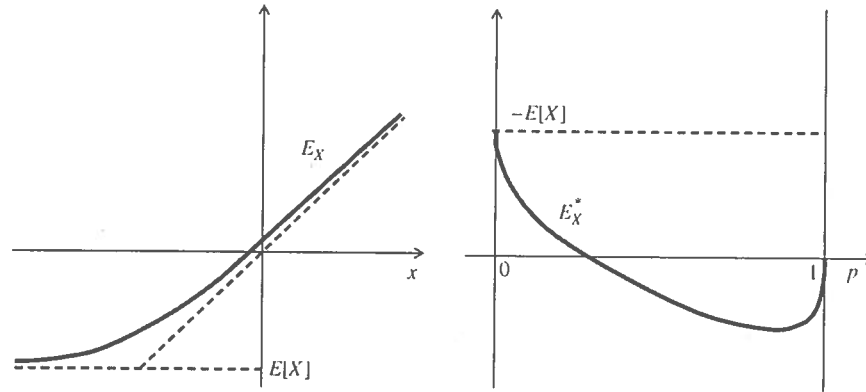
$$(x, p) \in \Gamma_X \iff (p, x) \in \Delta_X.$$

We recall from convex analysis that a pair of maximal monotone relations that are the inverses of each other are the graphs of the subdifferentials of two convex functions that are conjugate to each other; see, for example, Rockafellar and Royset [28] for details. There is flexibility in the choice of convex functions as knowledge of a subdifferentials of a function leaves the function determined only up to an additive constant. The consideration of conjugate pairs of convex functions defined in some way in terms of F_X and Q_X is due to Ogryczak and Ruszczyński [23, 24, 25]. Whereas they focus on random variables with a “profit” or “gain” orientation and the lower tail of the distribution of such variables, we here make choices more natural to the opposite orientation. A shift to a “cost” orientation is accomplished in Dentcheva and Martinez [11], which led to a development similar to that below. However, we adopt slightly different choices that coordinate well with superquantiles; see Rockafellar and Royset [28] for further connections.

We define the *superexpectation function* $E_X: \mathbb{R} \rightarrow \mathbb{R}$ associated with a random variable X by

$$E_X(x) = E[\max\{x, X\}] = \int_{-\infty}^{\infty} \max\{x, x'\} dF_X(x') = \int_0^1 \max\{x, Q_X(p)\} dp,$$

with the value $E_X(x)$ being termed the *superexpectation of X at level x* . Here, the last equality follows from a change-of-variable formula; see, for example, Billingsley [6, Theorem 16.13]. The connection between the superexpectation function E_X and Γ_X is clear from the next theorem, where we use the notation ∂E_X to denote the subdifferential of E_X and $\text{gph } \partial E_X$ the corresponding graph.

FIGURE 4. Dual Pair E_X and E_X^* .

Theorem 1. (ROCKAFELLAR AND ROYSET [28, THEOREM 1], CHARACTERIZATION OF SUPEREXPECTATIONS). *The superexpectation function E_X for a random variable X having $E|X| < \infty$ is a nondecreasing finite convex function on \mathbb{R} with the following properties:*

- (i) $\Gamma_X = \text{gph } \partial E_X$.
- (ii) $F_X(x) = \text{right-derivative of } E_X \text{ at } x$.
- (iii) $E_X(x) - x \geq 0$, $\lim_{x \rightarrow \infty} [E_X(x) - x] = 0$, and $\lim_{x \rightarrow -\infty} E_X(x) = E[X]$.
- (iv) For any random variables X_0 and X_1 having $E|X_0|, E|X_1| < \infty$,

$$E_X(x) \leq (1 - \lambda)E_{X_0}(x) + \lambda E_{X_1}(x) \quad \text{when } X = (1 - \lambda)X_0 + \lambda X_1 \text{ with } 0 < \lambda < 1.$$

Moreover, any convex function f on \mathbb{R} with the properties that

$$f(x) - x \geq 0, \quad \lim_{x \rightarrow \infty} [f(x) - x] = 0, \quad \lim_{x \rightarrow -\infty} f(x) = \text{a finite value},$$

is a superexpectation function for a random variable X having $E|X| < \infty$.

The left portion of Figure 4 illustrates the properties in Theorem 1(iii). The additional convexity property in Theorem 1(iv) is valuable for applications in stochastic optimization, which often involve random variables $X(u)$ that depend convexly on a parameter vector u .

The connection between the superexpectation function and the superquantile function emerges from the Legendre–Fenchel transform. We recall that a closed proper¹ convex function f on \mathbb{R} defines a conjugate function f^* on \mathbb{R} , through the Legendre–Fenchel transform

$$f^*(p) = \sup_x \{xp - f(x)\},$$

which is also closed, proper, and convex. The next theorem gives an expression for the conjugate function of E_X and its properties. In essence, the result is established in Ogryczak and Ruszczyński [25], and Dentcheva and Martinez [11], but we follow the notation in Rockafellar and Royset [28], which also provides an alternative proof.

Theorem 2. (ROCKAFELLAR AND ROYSET [28, THEOREM 2], DUAL OF SUPEREXPECTATIONS). *The closed proper convex function E_X^* on \mathbb{R} that is conjugate to the superexpectation function E_X for a random variable X with $E|X| < \infty$ is given by*

$$E_X^*(p) = \begin{cases} -(1-p)\bar{Q}_X(p) & \text{for } p \in (0, 1), \\ -E[X] & \text{for } p = 0, \\ 0 & \text{for } p = 1, \\ \infty & \text{for } p \notin [0, 1]. \end{cases}$$

¹ A function $f: \mathbb{R} \rightarrow [-\infty, \infty]$ is closed if it is lower semicontinuous and proper if $f(x) > -\infty$ for all $x \in \mathbb{R}$ and $f(x) < \infty$ for some $x \in \mathbb{R}$.

E_X^* has the following properties:

- (i) $\Delta_X = \text{gph } \partial E_X^*$.
- (ii) $Q_X(p) = \text{left-derivative of } E_X^* \text{ at } p$.
- (iii) E_X^* is continuous relative to $[0, 1]$ with

$$\lim_{p \nearrow 1} (1-p)\bar{Q}_X(p) = 0, \quad \lim_{p \searrow 0} \bar{Q}_X(p) = E[X].$$

Moreover, any function g on \mathbb{R} that is finite convex and continuous on $[0, 1]$ with $g(1) = 0$, but $g(p) = \infty$ for $p \notin [0, 1]$, is the conjugate of a superexpectation function for some random variable X .

The right portion of Figure 4 illustrates E_X^* . We note that the conjugate E_X^* is uniquely determined by the superquantile function \bar{Q}_X . Not only it but also E_X , F_X , and Q_X , along with Γ_X and Δ_X , can be reconstructed from knowledge of \bar{Q}_X . Moreover the following properties of a function \bar{g} on $(0, 1)$ are necessary and sufficient to have $\bar{g} = \bar{Q}_X$ for a random variable X with $E|X| < \infty$:

$$(1-p)\bar{g}(p) \text{ is concave in } p \text{ with } \lim_{p \nearrow 1} (1-p)\bar{g}(p) = 0, \\ \lim_{p \searrow 0} \bar{g}(p) = \text{a finite value.}$$

Consequently, the claim in §2 that the superquantile function of a random variable is as fundamental to a random variable as the distribution and quantile functions is justified.

We also note that since

$$\begin{aligned} \bar{Q}_X(p) &= -\frac{1}{1-p} E_X^*(p) \\ &= -\frac{1}{1-p} \sup_x \{xp - E_X(x)\} = \min_x \left\{ x + \frac{1}{1-p} E[\max\{0, X-x\}] \right\}, \end{aligned} \quad (6)$$

the trade-off formula (4) is directly recovered. In fact, this insight was the source of the discovery of that formula.

A simple example may help illustrate the concepts.

Example 1. Let X be exponentially distributed with parameter $\lambda > 0$. Then the distribution function is $F_X(x) = 1 - \exp(-\lambda x)$, the superexpectation function is

$$E_X(x) = \begin{cases} x + (1/\lambda) \exp(-\lambda x) & \text{for } x \geq 0, \\ 1/\lambda & \text{for } x < 0, \end{cases}$$

and the conjugate superexpectation function has $E_X^*(p) = (1/\lambda)(p-1)(1 - \log(1-p))$ for $p \in [0, 1)$. Quantiles and superquantiles are thus given on $(0, 1)$ by

$$Q_X(p) = -(1/\lambda) \log(1-p), \quad \bar{Q}_X(p) = (1/\lambda)[1 - \log(1-p)].$$

These insights lead to the following bounds on the superquantile function.

Theorem 3. (ROCKAFELLAR AND ROYSET [28, THEOREM 3], SUPERQUANTILE ESTIMATES). For $p \in [0, 1)$, one has

- (i) $|\bar{Q}_X(p) - \bar{Q}_Y(p)| \leq 1/(1-p)E|X-Y|$ when $E|X| < \infty$, $E|Y| < \infty$.
- (ii) $E[X] \leq \bar{Q}_X(p) \leq E[X] + (1/(\sqrt{1-p}))\sigma(X)$ when $E[X^2] < \infty$, $\sigma(X) = \text{standard deviation}$.

The above development allows for alternative characterizations of the classical notion of convergence in distribution for a sequence of random variables. Again, the superquantile function takes on a new role. We recall that a sequence of random variables X_k converges in distribution to a random variable X when $F_{X_k}(x) \rightarrow F_X(x)$ at all continuity points x of F_X .

Theorem 4. (ROCKAFELLAR AND ROYSET [28, THEOREM 4], CHARACTERIZATIONS OF CONVERGENCE IN DISTRIBUTION). *Let X, X_k be random variables with $E|X|, E|X_k| < \infty$, $k = 1, 2, \dots$. Then, the following conditions are equivalent:*

- (i) X_k converges in distribution to X .
- (ii) Γ_{X_k} converges graphically² to Γ_X .
- (iii) Δ_{X_k} converges graphically to Δ_X .
- (iv) $Q_{X_k}(p) \rightarrow Q_X(p)$ at all continuity points p of Q_X in $(0, 1)$.
- (v) $E_{X_k}(x) \rightarrow E_X(x)$ for all $x \in \mathbb{R}$.
- (vi) $\bar{Q}_{X_k}(p) \rightarrow \bar{Q}_X(p)$ for all $p \in (0, 1)$.

The everywhere pointwise convergence in (v) and (vi) can be replaced by pointwise convergence on a dense subset or uniform convergence on compact intervals of \mathbb{R} and $(0, 1)$, respectively.

It is apparent from Theorem 4(vi) that a superquantile is stable under perturbations of the underlying probability distribution. This has importance consequences for optimization problems with superquantiles of parametric random variables as objective functions and constraints. If the superquantiles remain convex and finite as functions of the parameters, then Theorem 4(vi) ensures epiconvergence of approximations obtained by replacing true probability distributions with approximating ones. Moreover, optimal solutions of problems with the approximations will tend to those of the true problems, justifying the use of approximate probability distributions in applications.

We end this section with a brief discussion of stochastic dominance, which has importance applications in modeling and stochastic optimization; see, for example, Dentcheva and Ruszczyński [12]. Typically, that topic is presented in the context of random variables with “profit” and “gain” orientation, and therefore, to be consistent with the other parts of this tutorial, we need to adopt a parallel definition for random variables with a “cost” orientation. First-order stochastic dominance of X over Y , denoted by $X \leq_1 Y$, is defined as

$$X \leq_1 Y \iff E[g(X)] \leq E[g(Y)] \text{ for continuous bounded increasing } g.$$

Second-order stochastic dominance, denoted by $X \leq_2 Y$, means that

$$X \leq_2 Y \iff E[g(X)] \leq E[g(Y)] \text{ for finite convex increasing } g.$$

The latter property is also known as “increasing convex order”; see Müller and Stoyan [22]. The connections between these notions and distribution, quantile, superexpectation, and superquantile functions follow next, where we see that the superexpectation function as well as the superquantile function characterize second-order stochastic dominance.

Theorem 5. (ROCKAFELLAR AND ROYSET [28, THEOREM 8], STOCHASTIC DOMINANCE). *First-order stochastic dominance is characterized by*

$$X \leq_1 Y \iff F_X \geq F_Y \iff Q_X \leq Q_Y.$$

Second-order stochastic dominance is characterized by

$$X \leq_2 Y \iff E_X \leq E_Y \iff \bar{Q}_X \leq \bar{Q}_Y.$$

² Graphical convergence here corresponds to the convergence of the corresponding subsets of \mathbb{R}^2 in the sense of Painlevé–Kuratowski; see Rockafellar and Wets [32, Chapter 4].

5. Superdistribution and Measures of Regret, Error, and Deviation

A further role of the superquantile function of a random variable is in the construction of a distribution function of another random variable and higher-order superquantile functions, with resulting applications in generalized regression and risk analysis. The construction results in measures of risk, regret, error, and deviation that are connected in a *risk quadrangle* described in detail below.

We start with a side-by-side graphical comparison between the superquantile function \bar{Q}_X and the quantile function Q_X , as in the right portion of Figure 5. We assume that X is a nonconstant random variable with $E[X^2] < \infty$. An immediate insight is that \bar{Q}_X is the inverse of a distribution function \bar{F}_X (see the left portion of Figure 5) in perfect analogy to the pairing of Q_X and F_X . Specifically,

$$\bar{F}_X(x) = \begin{cases} \bar{Q}_X^{-1}(x) & \text{for } \lim_{p \searrow 0} \bar{Q}_X(p) < x < \lim_{p \nearrow 1} \bar{Q}_X(p), \\ 0 & \text{for } x \leq \lim_{p \searrow 0} \bar{Q}_X(p), \\ 1 & \text{for } x \geq \lim_{p \nearrow 1} \bar{Q}_X(p). \end{cases}$$

We call \bar{F}_X the *superdistribution function* of X . Specifically, \bar{F}_X is the distribution function for an auxiliary random variable \bar{X} derived from X , and it is given by

$$\bar{X} = \bar{Q}_X(F_X(X)).$$

Consequently, $\bar{F}_X = F_{\bar{X}}$ and

$$Q_{\bar{X}}(p) = \bar{Q}_X(p) \quad \text{for } p \in (0, 1). \tag{7}$$

In view of (3) and (7), we then find that the superquantile function of \bar{X} is given by

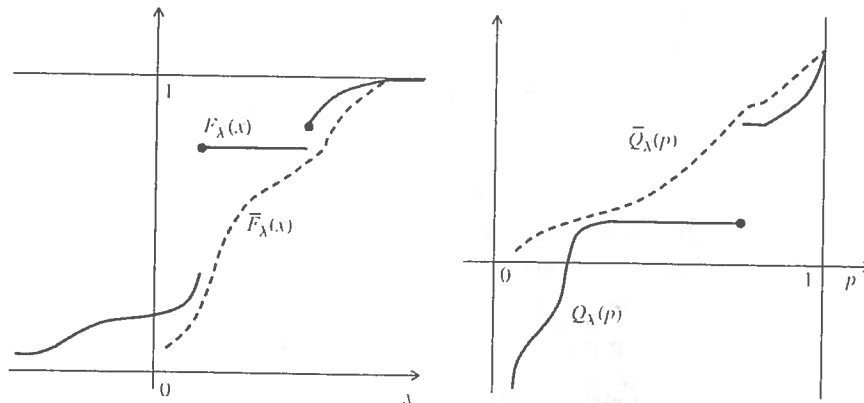
$$\bar{\bar{Q}}_X(p) = \frac{1}{1-p} \int_p^1 Q_{\bar{X}}(p') dp' = \frac{1}{1-p} \int_p^1 \bar{Q}_X(p') dp', \tag{8}$$

which we refer to as the *second-order superquantile function* of X and denote it by $\bar{\bar{Q}}_X$. Of course, this process can be repeated with \bar{X} in the role of X to generate even higher-order superquantiles.

We note that the assumption of $E[X^2] < \infty$ implies that $E|\bar{X}| < \infty$ through Theorem 3(b) so that the derivations of §4 hold with X replaced by \bar{X} . Consequently, in parallel to (6), we obtain that

$$\begin{aligned} \bar{\bar{Q}}_X(p) &= \bar{Q}_{\bar{X}}(p) = \min_x \{x + \bar{V}_p(X - x)\}, \\ \bar{Q}_X(p) &= Q_{\bar{X}}(p) = \arg \min_x \{x + \bar{V}_p(X - x)\}, \end{aligned} \tag{9}$$

FIGURE 5. Superdistribution function \bar{F}_X and superquantile function \bar{Q}_X .



where

$$\begin{aligned}\bar{V}_p(X) &= \frac{1}{1-p} E[\max\{0, \bar{X}\}] \\ &= \frac{1}{1-p} \int_{-\infty}^{\infty} \max\{0, x\} d\bar{F}_X(x) = \frac{1}{1-p} \int_0^1 \max\{0, \bar{Q}_X(p')\} dp',\end{aligned}$$

with the last equality following by a change-of-variable formula (see Billingsley [6, Theorem 16.13]) and (7).

The second-order superquantile gives a regular measure of risk, which also leads to measures of regret, error, and deviation, with further applications in generalized regression and risk modeling. Before making these claims formal, we introduce the additional concepts.

A *measure of deviation* is a functional \mathcal{D} that assigns to a random variable X a value $\mathcal{D}(X)$ in $[0, \infty]$ that quantifies its nonconstancy. It is *regular* if it is closed and convex (analogously to the second and third axioms in the regularity condition of §3) and in addition satisfies the axiom

$$\mathcal{D}(C) = 0 \text{ for constant random variables } X \equiv C, \text{ but } \mathcal{D}(X) > 0 \text{ for nonconstant } X.$$

A *measure of regret* is a functional \mathcal{V} that assigns to a random variable X a value $\mathcal{V}(X)$ in $(-\infty, \infty]$ that quantifies the perceived displeasure with the mix of possible outcomes for X . It is *regular* if it is closed and convex and in addition satisfies the following axioms:

$$\mathcal{V}(0) = 0, \text{ but } \mathcal{V}(X) > E[X] \text{ when } X \neq 0;$$

$$\text{for any sequence } X_k, \lim_{k \rightarrow \infty} \{\mathcal{V}(X_k) - E[X_k]\} = 0 \text{ implies } \lim_{k \rightarrow \infty} E[X_k] = 0.$$

A *measure of error* is a functional \mathcal{E} that assigns to a random variable X a value $\mathcal{E}(X)$ in $[0, \infty]$ that quantifies its nonzeroness. It is *regular* if it is closed and convex and in addition satisfies the following axioms:

$$\mathcal{E}(0) = 0, \text{ but } \mathcal{E}(X) > 0 \text{ when } X \neq 0;$$

$$\text{for any sequence } X_k, \lim_{k \rightarrow \infty} \mathcal{E}(X_k) = 0 \text{ implies } \lim_{k \rightarrow \infty} E[X_k] = 0.$$

We refer to Rockafellar and Uryasev [31] for examples and discussion of these quantities.

Theorem 6. *On the space of random variables X with $E[X^2] < \infty$ and the L_2 -norm topology, for any $p \in (0, 1)$, the functionals $\bar{\mathcal{R}}_p$, $\bar{\mathcal{D}}_p$, $\bar{\mathcal{V}}_p$, and $\bar{\mathcal{E}}_p$, given by*

$$\bar{\mathcal{R}}_p(X) = \bar{Q}_X(p), \quad \bar{\mathcal{D}}_p(X) = \bar{\mathcal{R}}_p(X) - E[X],$$

and

$$\bar{\mathcal{V}}_p(X) = \frac{1}{1-p} \int_0^1 \max\{0, \bar{Q}_X(p')\} dp', \quad \bar{\mathcal{E}}_p(X) = \bar{\mathcal{V}}_p(X) - E[X], \quad (10)$$

are regular measures of risk, deviation, regret, and error, respectively. Moreover, they form a risk quadrangle in the sense of Rockafellar and Uryasev [31], and thereby (9) and

$$\bar{Q}_X(p) = \arg \min_x \{x + \bar{\mathcal{V}}_p(X - x)\} = \arg \min_x \bar{\mathcal{E}}_p(X - x) \quad (11)$$

hold.

Proof. The regularity of $\bar{\mathcal{R}}_p$ and $\bar{\mathcal{D}}_p$ follows from (8) and the properties of the superquantile function. The closedness and convexity of $\bar{\mathcal{V}}_p$ and $\bar{\mathcal{E}}_p$ follow similarly. By Rockafellar et al. [33, Proposition 2], we know that

$$\begin{aligned}\bar{\mathcal{E}}_p(X) &= 0 \text{ when } X \equiv 0, \\ \bar{\mathcal{E}}_p(X) &> 0 \text{ when } X \neq 0, \text{ and} \\ \bar{\mathcal{E}}_p(X) &\geq \min\{1, p/(1-p)\} |E[X]|.\end{aligned}$$

Consequently, we find that the remaining axioms required for regularity of $\bar{\mathcal{V}}_p$ and $\bar{\mathcal{E}}_p$ are also satisfied.

As a direct consequence of the quadrangle theorem in Rockafellar and Uryasev [31], the measures of risk, deviation, error, and regret form a risk quadrangle. The trade-off formula follows by the quadrangle theorem in Rockafellar and Uryasev [31] as well. \square

The measure of risk $\bar{\mathcal{R}}_p$ is more conservative than the one based on a superquantile in the sense that $\bar{\mathcal{R}}_p(X) \geq \bar{Q}_X(p)$. Moreover, it mitigates the difficulty a decision maker may have with selecting an appropriate probability level p for the choice $\mathcal{R}(X) = \bar{Q}_X(p)$ by considering an “average” of superquantiles. The presence of integrals in the expressions for the measures of risk, deviation, regret, and error may require numerical integration, but this causes little complication in practice as such one-dimensional integration is easily carried out with high accuracy; see, for example, §6 and Rockafellar et al. [33]. The measure of error $\bar{\mathcal{E}}_p$ has in view of (11) significant implications in generalized regression as discussed next.

6. Superquantile Regression

In applications, it may be beneficial to attempt to approximate a random variable Y by means of an n -dimensional explanatory random vector X that is more accessible in some sense. This situation naturally leads to least-squares regression and related models that estimate conditional expectations. Although such models are adequate in many situations, they fall short in contexts where a decision maker is risk averse, i.e., is more concerned about upper-tail realizations of Y than average loss, and views errors asymmetrically with underestimating losses being more detrimental than overestimating. Quantile regression (see Koenker [17], Koenker and Bassett [18], Gilchrist [14], and references therein) accommodates risk averseness and an asymmetric view of errors by estimating conditional quantiles at a certain probability level such as those in the tail of the conditional distribution of Y . However, with the increasing focus on superquantiles and their desired properties as coherent and regular measures of risk, we would like to also carry out generalized regression that is consistent with superquantiles. Theorem 6 provides the framework for such a regression methodology.

We start by recalling quantile regression. One obtains a p -quantile of a random variable Y , with $E|Y| < \infty$, by computing

$$Q_Y(p) = \arg \min_y \mathcal{E}_p(Y - y), \quad (12)$$

where

$$\mathcal{E}_p(Y) = \frac{1}{1-p} E[\max\{Y, 0\}] - E[Y].$$

(These expressions are closely connected to (5); see Rockafellar et al. [33].) Here, we assume that the argmin is unique for simplicity. In general, the quantile is taken as the left-most point in the argmin set. The functional \mathcal{E}_p is a regular measure of error referred to as the (scaled) Koenker–Bassett error. Although $Q_Y(p)$ is the “best” scalar representation of Y in the sense of the Koenker–Bassett error, we need to go beyond the class of constant functions to utilize the connection with an underlying explanatory random vector X . We focus on regression functions of the form

$$f(x) = C_0 + \langle C, h(x) \rangle, \quad C_0 \in \mathbb{R}, C \in \mathbb{R}^m,$$

for a given “basis” function $h = (h_1, h_2, \dots, h_m): \mathbb{R}^n \rightarrow \mathbb{R}^m$. This class satisfies most practical needs including that of linear regression where $m = n$ and $h(x) = x$. Consequently, instead of solving (12), quantile regression centers on finding optimal solutions of the problem

$$\min_{C_0 \in \mathbb{R}, C \in \mathbb{R}^m} \mathcal{E}_p(Y - [C_0 + \langle C, h(X) \rangle]),$$

with the resulting optimal coefficients C_0 and C yielding a regression function.

Since minimizing $\bar{\mathcal{E}}_p(Y - y)$ by choice of $y \in \mathbb{R}$ returns the p -superquantile of Y by Theorem 6, a parallel development of the previous paragraph leads to *superquantile regression* by solving

$$P: \min_{C_0 \in \mathbb{R}, C \in \mathbb{R}^m} \bar{\mathcal{E}}_p(Y - [C_0 + \langle C, h(X) \rangle]).$$

The remainder of this tutorial summarizes properties of superquantile regression, gives means of assessing the goodness of fit, and discusses computational strategies for solving P .

6.1. Existence, Uniqueness, and Stability

We start by examining the existence and uniqueness of regression functions obtained from P and then proceed with studying the stability of such functions under perturbations of the distribution of (X, Y) . For notational simplicity, we let

$$Z(C_0, C) = Y - (C_0 + \langle C, h(X) \rangle)$$

be an error random variable, whose distribution depends on C_0 , C , h , and the joint distribution of (X, Y) . Moreover, for any $p \in [0, 1]$ and random variable X , we let

$$\bar{q}_p(X) = \bar{Q}_X(p).$$

We denote by $\bar{C} \subset \mathbb{R}^{m+1}$ the set of optimal solutions of P and refer to $(\bar{C}_0, \bar{C}) \in \bar{C}$ as a *regression vector*.

In view of the regression theorem in Rockafellar and Uryasev [31] (see also Theorem 3.1 in Rockafellar et al. [34]), we find that a regression vector can equivalently be determined from the measure of deviation \bar{D}_p by first solving

$$D: \min_{C \in \mathbb{R}^m} \bar{D}_p(Z_0(C)),$$

where $Z_0(C) = Y - \langle C, h(X) \rangle$, to obtain the optimal “slope” coefficients \bar{C} and then setting the “intercept” coefficient

$$\bar{C}_0 = \bar{q}_p(Z_0(\bar{C})).$$

Clearly, in comparison to P , solving D involves one less optimization variable and also a simpler objective function; see (10) and (8).

The existence and uniqueness of a regression vector are given by the next theorem.

Theorem 7. (ROCKAFELLAR ET AL. [33, THEOREM 2], EXISTENCE AND UNIQUENESS OF REGRESSION VECTOR). *If $E[Y^2]$, $E[h_i(X)^2] < \infty$, $i = 1, 2, \dots, m$, then P is a convex problem with a set of optimal solutions \bar{C} that is nonempty, closed, and convex.*

(a) \bar{C} is bounded if and only if the random vector X and the basis function h satisfy the condition that $\langle C, h(X) \rangle$ is not constant unless $C = 0$.

(b) If, in addition, for every $(C_0, C), (C'_0, C') \in \mathbb{R}^{m+1}$, with $C \neq C'$, there exists a $p_0 \in [0, 1]$ such that

$$0 \leq \bar{q}_p(Z(C_0, C) + Z(C'_0, C')) < \bar{q}_p(Z(C_0, C)) + \bar{q}_p(Z(C'_0, C')) \quad (13)$$

for all $p \in [p_0, 1]$, then \bar{C} is a singleton.

Although (13) is not always satisfied, we know that if $(h(X), Y)$ is normally distributed with a positive definite variance-covariance matrix, P has a unique solution, and therefore superquantile regression returns a unique regression vector in that case (Rockafellar et al. [33]).

We next turn to consistency and stability of the regression vector. Of course, the joint distribution of (X, Y) is rarely available in practice, and one may need to pass to an approximating empirical distribution generated by a sample. Moreover, perturbations of the “true” distribution of (X, Y) may occur due to measurement errors in the data and other factors. We consider these possibilities and let (X^ν, Y^ν) be a random vector whose joint distribution approximates that of (X, Y) in some sense. For example, (X^ν, Y^ν) may be governed by the empirical distribution generated by an independent and identically distributed sample of size

ν from (X, Y) . Presumably, as $\nu \rightarrow \infty$, the approximation of (X, Y) by (X^ν, Y^ν) improves as stated formally below. Regardless of the nature of (X^ν, Y^ν) , we define an approximate error random variable

$$Z^\nu(C_0, C) = Y^\nu - (C_0 + (C, h(X^\nu))),$$

and the corresponding approximate superquantile regression problem

$$P^\nu: \min_{C_0 \in \mathbb{R}, C \in \mathbb{R}^m} \bar{E}_p(Z^\nu(C_0, C)).$$

The next result, which utilizes Theorem 4(vi), shows that as (X^ν, Y^ν) approximates (X, Y) , a regression vector obtained from P^ν approximates one from P , which provides the justification for basing a regression analysis on P^ν .

Theorem 8. (ROCKAFELLAR ET AL. [33, THEOREM 3], STABILITY OF REGRESSION VECTOR). *Suppose that (X^ν, Y^ν) , $\nu = 1, 2, \dots$, and (X, Y) are $n + 1$ -dimensional random vectors such that (X^ν, Y^ν) converges to (X, Y) in distribution and that the basis function h is continuous except possibly on a subset $S \subset \mathbb{R}^n$ with $\text{Prob}(X \in S) = 0$. Moreover, let $E[(h_i(X))^2], E[Y^2], \sup_\nu E[(h_i(X^\nu))^2], \sup_\nu E[(Y^\nu)^2] < \infty, i = 1, 2, \dots, m$.*

If $\{(\bar{C}_0^\nu, \bar{C}^\nu)\}_{\nu=1}^\infty$ is a sequence of optimal solutions of P^ν , with $p \in (0, 1)$, then every accumulation point of that sequence is a regression vector of P .

6.2. Goodness-of-Fit Criterion

Regression modeling must be associated with means of assessing the goodness of fit of a computed regression vector. In least-squares regression, the frequently used *coefficient of determination* is given by the residual sum of squares and the total sum of squares, which in our notation takes the form

$$R^2 = 1 - \frac{E[Z(C_0, C)^2]}{\sigma^2(Y)}, \tag{14}$$

where $\sigma^2(Y)$ denotes the variance of Y . Although R^2 can not be relied on exclusively, it provides an indication of the goodness of fit that is easily extended to the present context of superquantile regression.

From Example 1' in Rockafellar and Uryasev [31], we know that the numerator in (14) is an error measure applied to $Z(C_0, C)$ and that it corresponds to the deviation measure $\sigma^2(\cdot)$. Moreover, the minimization of that error of $Z(C_0, C)$ results in the least-squares regression vector. According to Rockafellar and Uryasev [31], these error and deviation measures are in correspondence and belong to a risk quadrangle that yields the expectation as its statistic. This observation motivates us to define a coefficient of determination for superquantile regression as

$$\bar{R}_p^2(C_0, C) = 1 - \frac{\bar{E}_p(Z(C_0, C))}{\bar{D}_p(Y)}$$

at probability level $p \in (0, 1)$. As in the classical case, higher values of \bar{R}_p^2 are "better," and in fact, since P aims to minimize $\bar{E}_p(Z(C_0, C))$, the goal of superquantile regression is to find the highest possible value of \bar{R}_p^2 . Clearly, though, $\bar{R}_p^2 \leq 1$, since error and deviation measures are nonnegative.

6.3. Computational Methods for Superquantile Regression

Although it at first may appear difficult to solve P , several simplification may come into play. As discussed above, it suffices to solve D , which in fact reduces to a linear program when the distribution of (X, Y) is given by a normalized counting measure (Rockafellar et al. [33]). The use of that measure is of course the standard assumption in practice, where a set of observations of (X, Y) , each assumed equally likely to occur, is usually available. Even if another discrete measure is assumed, which is relevant when observations are "weighted"

unevenly, D is easily solved to high accuracy through numerical integration. By replacing the integral in D with a finite sum using some standard numerical integration scheme, the problem becomes one of minimizing mixed superquantiles that can be transcribed into a linear program using standard techniques. Moreover, nonsmooth optimization algorithms for unconstrained convex problems such as solvers in Portfolio Safeguard (American Optimal Decisions, Inc. [3]) are available. We refer to Rockafellar et al. [33] for further details and numerical illustrations.

Acknowledgments

This work was supported by the Air Force Office of Scientific Research [Grants FA9550-11-1-0206 and F1ATAO1194GOO1].

References

- [1] C. Acerbi and D. Tasche. On the coherence of expected shortfall. *Journal of Banking and Finance* 26(7):1487–1503, 2002.
- [2] S. Alexander, T. F. Coleman, and Y. Li. Minimizing CVaR and VaR for a portfolio of derivatives. *Journal of Banking and Finance* 30(2):583–605, 2006.
- [3] American Optimal Decisions, Inc. *Portfolio Safeguard (PSG) in Windows Shell Environment: Basic Principles*. AORDA, Gainesville, FL, 2011.
- [4] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance* 9(3):203–227, 1999.
- [5] A. Balbas, B. Balbas, and R. Balbas. CAPM and APT-like models with risk measures. *Journal of Banking and Finance* 34(6):1166–1174, 2010.
- [6] P. Billingsley. *Probability and Measure*, Anniversary Edition. John Wiley & Sons, Hoboken, NJ, 2012.
- [7] M. Carrion, A. B. Philpott, A. J. Conejo, and J. M. Arroyo. A stochastic programming approach to electric energy procurement for large consumers. *IEEE Transactions on Power Systems* 22(2):744–754, 2007.
- [8] C. W. Commander, P. M. Pardalos, V. Ryabchenko, S. Uryasev, and G. Zrazhevsky. The wireless network jamming problem. *Journal of Combinatorial Optimization* 14(4):481–498, 2007.
- [9] A. J. Conejo, R. Garcia-Bertrand, M. Carrion, A. Caballero, and A. de Andres. Optimal involvement in futures markets of a power producer. *IEEE Transactions on Power Systems* 23(2):703–711, 2008.
- [10] F. Delbaen. Coherent measures of risk on general probability spaces. P. J. Schonbucher and K. Sandmann, eds. *Advances in Finance and Stochastics: Essays in Honor of Dieter Sondermann*. Springer, Berlin, 1–37, 2002.
- [11] D. Dentcheva and G. Martinez. Two-stage stochastic optimization problems with stochastic ordering constraints on the recourse. *European Journal of Operational Research* 219(1):1–8, 2012.
- [12] D. Dentcheva and A. Ruszczyński. Stochastic optimization with dominance constraints. *SIAM Journal on Optimization* 14:548–566, 2003.
- [13] H. Föllmer and A. Schied. *Stochastic Finance: An Introduction in Discrete Time*, 2nd ed. De Gruyter Studies in Mathematics, Walter de Gruyter, Berlin, 2004.
- [14] W. Gilchrist. Regression revisited. *International Statistical Review* 76(3):401–418, 2008.
- [15] T. Haukaas, S. Allahdadian, and M. Mahsuli. Risk measures for minimization of earthquake costs. G. Deodatis, B. R. Ellingwood, and D. M. Frangopol, eds. *Proceedings of the 11th International Conference on Structural Safety and Reliability*. Balkema, Leiden, The Netherlands, 2013.
- [16] K. Kalinchenko, A. Veremyev, V. Boginski, D. E. Jeffcoat, and S. Uryasev. Robust connectivity issues in dynamic sensor networks for area surveillance under uncertainty. *Structural and Multidisciplinary Optimization* 7(2):235–248, 2011.
- [17] R. Koenker. *Quantile Regression*. Cambridge University Press, Cambridge, UK, 2005.
- [18] R. Koenker and G. Bassett Jr. Regression quantiles. *Econometrica* 46(1):33–50, 1978.

- [19] P. Krokhmal, M. Zabarankin, and S. Uryasev. Modeling and optimization of risk. *Surveys in Operations Research and Management Science* 16(2):49–66, 2011.
- [20] R. Minguez, E. Castillo, and J. L. Lara. Iterative scenario reduction technique to solve reliability based optimization problems using the buffered failure probability. G. Deodatis, B. R. Ellingwood, and D. M. Frangopol, eds. *Proceedings of the 11th International Conference on Structural Safety and Reliability*. Balkema, Leiden, The Netherlands, 2013.
- [21] A. Molyboha and M. Zabarankin. Stochastic optimization of sensor placement for diver detection. *Operations Research* 60(2):292–312, 2012.
- [22] A. Müller and D. Stoyan. *Comparison Methods for Stochastic Models and Risks*. John Wiley & Sons, Chichester, UK, 2002.
- [23] W. Ogryczak and A. Ruszczyński. From stochastic dominance to mean-risk models: Semideviations and risk measures. *European Journal of Operational Research* 116(1):33–50, 1999.
- [24] W. Ogryczak and A. Ruszczyński. On consistency of stochastic dominance and mean-semideviation models. *Mathematical Programming* 89(2):217–232, 2001.
- [25] W. Ogryczak and A. Ruszczyński. Dual stochastic dominance and related mean-risk models. *SIAM Journal on Optimization* 13(1):60–78, 2002.
- [26] R. T. Rockafellar. Coherent approaches to risk in optimization under uncertainty. T. Klastorin, ed. *Tutorials in Operations Research*. INFORMS, Hanover, MD, 38–61, 2007.
- [27] R. T. Rockafellar and J. O. Royset. On buffered failure probability in design and optimization of structures. *Reliability Engineering and System Safety* 95(5):499–510, 2010.
- [28] R. T. Rockafellar and J. O. Royset. Random variables, monotone relations, and convex analysis. *Mathematical Programming, Series B*. Forthcoming, 2013.
- [29] R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk* 2(3):21–42, 2000.
- [30] R. T. Rockafellar and S. Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance* 26(7):1443–1471, 2002.
- [31] R. T. Rockafellar and S. Uryasev. The fundamental risk quadrangle in risk management, optimization and statistical estimation. *Surveys in Operations Research and Management Science* 18(1–2):33–53, 2013.
- [32] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. Springer, Berlin, 1998.
- [33] R. T. Rockafellar, J. O. Royset, and S. I. Miranda. Superquantile regression with applications to buffered reliability, uncertainty quantification, and conditional value-at-risk. Working paper, University of Washington, Seattle, 2013.
- [34] R. T. Rockafellar, S. Uryasev, and M. Zabarankin. Risk tuning with generalized linear regression. *Mathematics of Operations Research* 33(3):712–729, 2008.
- [35] S. Sarykalin, G. Serraino, and S. Uryasev. Value-at-risk vs. conditional value-at-risk in risk management and optimization. Z.-L. Chen, S. Raghavan, eds. *Tutorials in Operations Research*. INFORMS, Hanover, MD, 270–294, 2008.
- [36] B. Tomlin. On the value of mitigation and contingency strategies for managing supply chain disruption risks. *Management Science* 52(5):639–667, 2006.
- [37] S. Uryasev, U. A. Theiler, and G. Serraino. Risk-return optimization with different risk-aggregation strategies. *Journal of Risk Finance* 11(2):129–146, 2010.
- [38] P. M. Verderame and C. A. Floudas. Multisite planning under demand and transportation time uncertainty: Robust optimization and conditional value-at-risk frameworks. *Industrial and Engineering Chemistry Research* 50(9):4959–4982, 2011.
- [39] C.-J. Wang and S. Uryasev. Efficient execution in the secondary mortgage market: A stochastic optimization model using CVaR constraints. *Journal of Risk* 10(1):41–66, 2007.
- [40] R. B. Webby, P. T. Adamson, J. Boland, P. G. Howlett, A. V. Metcalfe, and J. Piantadosi. The mekongapplications of value at risk (VaR) and conditional value at risk (CVaR) simulation to the benefits, costs, and consequences of water resources development in a large river basin. *Ecological Modelling* 201(1):89–96, 2007.
- [41] Y. Yamai and T. Yoshida. Value-at-risk versus expected shortfall: A practical perspective. *Journal of Banking and Finance* 29(4):997–1015, 2005.
- [42] G. M. Yamout, K. Hatfield, and H. E. Romeijn. Comparison of new conditional value-at-risk-based management models for optimal allocation of uncertain water supplies. *Water Resources Research* 43(7):1–13, 2007.