Annals of Operations Research
https://doi.org/10.1007/s10479-019-03373-1

S.I.: RECENT DEVELOPMENTS IN FINANCIAL MODELING AND RISK MANAGEMENT

Calculating CVaR and bPOE for common probability distributions with application to portfolio optimization and density estimation

Matthew Norton[1] · **Valentyn Khokhlov**[2] · **Stan Uryasev**[3]

Abstract

Conditional value-at-risk (CVaR) and value-at-risk, also called the superquantile and quantile, are frequently used to characterize the tails of probability distributions and are popular measures of risk in applications where the distribution represents the magnitude of a potential loss. buffered probability of exceedance (bPOE) is a recently introduced characterization of the tail which is the inverse of CVaR, much like the CDF is the inverse of the quantile. These quantities can prove very useful as the basis for a variety of risk-averse parametric engineering approaches. Their use, however, is often made difficult by the lack of well-known closed-form equations for calculating these quantities for commonly used probability distributions. In this paper, we derive formulas for the superquantile and bPOE for a variety of common univariate probability distributions. Besides providing a useful collection within a single reference, we use these formulas to incorporate the superquantile and bPOE into parametric procedures. In particular, we consider two: portfolio optimization and density estimation. First, when portfolio returns are assumed to follow particular distribution families, we show that finding the optimal portfolio via minimization of bPOE has advantages over superquantile minimization. We show that, given a fixed threshold, a single portfolio is the minimal bPOE portfolio for an entire class of distributions simultaneously. Second, we apply our formulas to parametric density estimation and propose the method of superquantiles (MOS), a simple variation of the method of moments where moments are replaced by superquantiles at different confidence levels. With the freedom to select various combinations of confidence levels, MOS allows the user to focus the fitting procedure on different portions of the distribution, such as the tail when fitting heavy-tailed asymmetric data.

Keywords Conditional value-at-risk · Buffered probability of exceedance · Superquantile · Density estimation · Portfolio optimization

✉ Matthew Norton
mnorton@nps.edu

Extended author information available on the last page of the article

1 Introduction

When faced with randomness and uncertainty, some of the most popular techniques for dealing with such randomness are parametric in nature. Given a real valued random variable X, analysis can be greatly simplified if one assumes that X belongs to a specific parametric family of distributions. For example, the Method of Moments (MM) is one of the simplest and most widely used methods for parametric density estimation. These techniques, however, often require that certain characteristics of the distribution family be representable by a simple, ideally closed-form, expression. For example, traditional MM uses closed-form expressions for the moments of the parametric distribution family. Similarly, the Matching of Quantiles (MOQ) procedure (see e.g., Sgouropoulos et al. 2015; Karian and Dudewicz 1999) uses expressions for the quantile function. In portfolio optimization, the availability of simple expressions for the mean and variance of portfolio returns yields a tractable Markowitz portfolio optimization problem.[1] For a variety of problems, application of a parametric method relies upon the availability of a closed-form expression for a specific characteristic of the parametric family of interest.

Luckily, for a variety of distributions, closed-from expressions are available for commonly utilized characteristics. These include characteristics such as the moments, the quantile, and the CDF. Over the past 2 decades, however, new fundamental characteristics like the superquantile have emerged from the field of quantitative risk management with important applications across engineering fields like financial, civil, and environmental engineering (see e.g., Rockafellar and Royset 2010; Rockafellar and Uryasev 2000; Davis and Uryasev 2016). Furthermore, closed-form expressions for these characteristics, for a large variety of common parametric distribution families, have not been widely disseminated. While emerging from specific engineering applications, some of these characteristics are very general and can be viewed as fundamental aspects of a random variable just like the mean or quantile. Thus, utilization of these characteristics within parametric methods is a natural consideration. To facilitate their use, however, we must develop closed-form expressions.[2]

We focus on developing these expressions for the superquantile and Buffered Probability of Exceedance (bPOE) for a variety of distribution families. Developments in financial risk theory over the last 2 decades have heavily emphasized measurement of tail risk. After Artzner et al. (1999) introduced the concept of a *coherent* risk measure, Rockafellar and Uryasev (2000) introduced the superquantile, also called Conditional Value at Risk (CVaR) in the financial literature. This began to be considered a preferable characterization of tail risk compared to the quantile, or Value-at-Risk (VaR). While some closed-form expression are available to use the superquantile within parametric procedures, see e.g., Rockafellar and Uryasev (2000), Landsman and Valdez (2003), Andreev et al. (2005), the variety of distributions discussed within each of these sources is limited.

We illustrate that for a variety of common distributions, straightforward techniques such as integration of the quantile function obtain a closed-form expression for the superquantile that is easy to use within subsequent parametric methods. We attempt to include a variety, providing superquantile formulas for the Exponential, Pareto/Generalized Pareto (GPD), Laplace, Normal, LogNormal, Logistic, LogLogistic, Generalized Student-t, Weibull, and Generalized Extreme Value (GEV) distributions. These provide examples varying from the exponentially tailed (Exponential, Pareto/GPD, Laplace), to the symmetric (Normal, Laplace,

[1] See Sect. 4 for specifics.

[2] When closed-form expressions are not available, we look to provide simple calculation methods that might still be utilized within parametric methods.

Logistic, Student-t), to the asymmetric heavier tailed (Weibull, LogLogistic, GEV) distributions. While some of these formulas may exist elsewhere, and we make every attempt to cite sources accordingly, we hope that this paper serves as a good resource for practitioners in search of superquantile formulas. We note that we were not able to derive closed form solutions for the popularly used Gamma distribution. This is not too surprising, particularly since there is also no known closed form quantile function. Nevertheless, we point out that the Weibull, which does have closed form superquantile formula, is a good approximation if the Gamma is of particular interest.

While the superquantile has risen in popularity over the past decade, a related characteristic called Buffered Probability of Exceedance (bPOE) has recently been introduced, first by Rockafellar and Royset (2010) in the context of Buffered Failure Probability and then generalized by Mafusalov and Uryasev (2018). This concept has grown in popularity within the risk management community with application in finance, logistics, analysis of natural disasters, statistics, stochastic programming, and machine learning (Shang et al. 2018; Uryasev 2014; Davis and Uryasev 2016; Mafusalov et al. 2018; Norton et al. 2017; Norton and Uryasev 2016). Specifically, bPOE is the inverse of the superquantile in the same way that the CDF is the inverse of the quantile. However, much like the superquantile when compared against the quantile, bPOE has many mathematically advantageous properties over the traditionally used Probability of Exceedance (POE). Direct optimization often reduces to convex or linear programming, it can be calculated via a one dimensional convex optimization problem, and it provides a risk-averse probabilistic assessment of the risk of experiencing outcomes larger than some fixed upper threshold. Thus, the second aim of this paper is to provide closed-form expressions for bPOE and, when unable to do so, show that calculation of bPOE is still simple, reducing to a one-dimensional convex optimization problem or a one-dimensional root finding problem. For the parametric portfolio application, in particular, we will see that when closed-form bPOE is unavailable and the superquantile is available, finding the optimal bPOE portfolio is no more difficult, computationally, than finding the optimal superquantile (CVaR) portfolio.

Motivating us to derive closed-form expressions (or simple calculation formulas) for the superquantile and bPOE for common distributions is the inclusion of these risk averse, tail measurements within parametric methods. In particular, we explore the use of the superquantile and bPOE within parametric portfolio optimization and density estimation. First, we consider parametric portfolio optimization, where returns are assumed to follow a specific distribution and, using these assumptions, a tractable portfolio optimization problem is formulated and solved. We begin by narrowing our choices of distribution to only those that both fit the pattern of portfolio returns and generate tractable portfolio optimization problems. Then, we consider two companion problems, solving for portfolios that minimize the superquantile (CVaR) of the distribution of potential losses (i.e. the average of the worst-case $100(1-\alpha)\%$ scenarios) and portfolios that minimize bPOE of the loss distribution (i.e. the *buffered* probability that losses will exceed a fixed upper threshold x). In comparing these problems, we discover that bPOE optimization can often be highly preferable to superquantile (CVaR) optimization in the parametric context. Specifically, for fixed α, the portfolio that minimizes the superquantile depends upon the distributional assumption (i.e., even if α is fixed, changing the assumed parametric distribution for returns will change the contents of the optimal portfolio). However, for fixed threshold x, the portfolio that minimizes bPOE does not depend upon the distributional assumption (at least for the specific class of distributions we consider, which includes the Logistic, Laplace, Normal, Student-t, and GEV). In other words, no matter which of these distributions we choose, we will always achieve the same optimal portfolio for fixed value of threshold x. Thus, bPOE-based portfolio optimization can provide additional consistency with respect to parameter choices, eliminating one source of additional variability for the decision maker.

Finally, we consider parametric density estimation, proposing a variant of MM where moments are replaced by superquantiles. This can also be seen as a natural variation of the MOQ procedure where quantiles are replaced by superquantiles. Made possible by the closed-form superquantile expressions, we show that this framework allows one to flexibly perform density estimation, allowing the user to focus the fitting procedure on specific portions of the distribution. For example, we illustrate by fitting a Weibull with additional emphasis put onto estimating the right tail. Compared against traditional MM and maximum likelihood (ML), we see that we get a better fit for such asymmetric, heavy tailed situations.

1.1 Organization of paper

We first provide a brief introduction to superquantiles and bPOE in Sect. 1.2. In Sect. 2, we give formulas for both the superquantile and bPOE for the Exponential, Pareto, Generalized Pareto, and Laplace distributions. Along the way, we highlight some simple relationships between POE, bPOE, the quantile, and the superquantile. In Sect. 3, we treat distributions for which a closed-form superquantile formula exists, but where we are unable to derive a simple closed-form bPOE formula. In order of appearance, we consider the Normal, Log-Normal, Logistic, Generalized Student-t, Weibull, LogLogistic, and Generalized Extreme Value Distribution. However, we point out for these cases that because a formula for the superquantile is known, bPOE can be solved for via a simple root finding problem. We also illustrate for some cases that the one-dimensional convex optimization formula for bPOE can also be used in these cases. In Sects. 1 and 2, all proofs can be found in the "Appendix". In Sect. 4, we illustrate the use of these formulas in portfolio optimization and parametric distribution approximation. Note that most of the referred to formulas for PDFs, CDFs, and quantile functions are readily availability on Wikipedia or statistics reference books such as Everitt (2006).

1.2 Background and notation

When working with optimization of tail probabilities, one frequently works with constraints or objectives involving *probability of exceedance* (POE), $p_x(X) = P(X > x)$, or its associated quantile $q_\alpha(X) = \min\{x \mid P(X \leq x) \geq \alpha\}$, where $\alpha \in [0, 1]$ is a probability level. The quantile is a popular measure of tail risk in financial engineering, but when included in optimization problems via constraints or objectives, is quite difficult to treat with continuous (linear or non-linear) optimization techniques.

A significant advancement was made in Rockafellar and Uryasev (2000, 2002) in the development of a replacement called the superquantile or CVaR. The superquantile is a measure of uncertainty similar to the quantile, but with superior mathematical properties. Formally, the superquantile (CVaR) for a continuously distributed X is defined as,

$$\bar{q}_\alpha(X) = E\left[X|X > q_\alpha(X)\right] = \frac{1}{1-\alpha}\int_{q_\alpha(X)}^{\infty} xf(x)dx = \frac{1}{1-\alpha}\int_{\alpha}^{1} q_p(X)dp. \qquad (1)$$

Similar to $q_\alpha(X)$, the superquantile can be used to assess the tail of the distribution. The superquantile, though, is far easier to handle in optimization contexts due to its convexity w.r.t. random variables and continuity w.r.t. α. It also has the important property that it considers the magnitude of events within the tail. Therefore, in situations where a distribution may

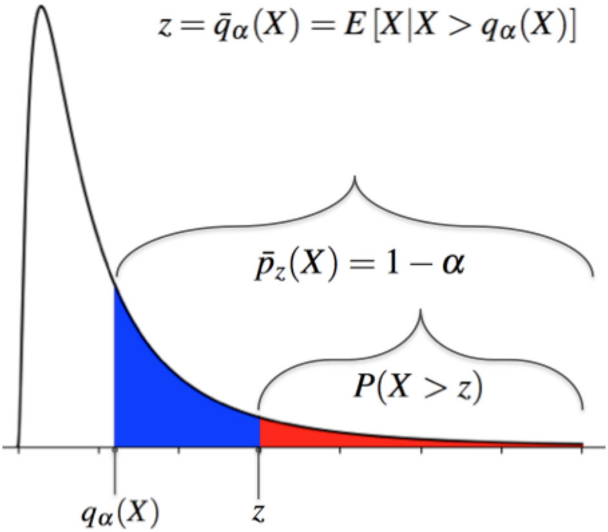

Fig. 1 Shown is the probability density function (PDF) of $X \sim$ Lognormal($\sigma = 1, \mu = 0$). Given threshold $z \in \mathbb{R}$, POE equals $P(X > z)$ the cumulative density in red. For the same threshold z, bPOE equals $\bar{p}_z(X)$ the combined cumulative density in red and blue. By definition, the expectation of the worst-case $1 - \alpha = \bar{p}_z(X)$ outcomes equals $z = \bar{q}_\alpha(X)$. These worst-case outcomes are those that are larger than the quantile $q_\alpha(X)$. (Color figure online)

have a heavy tail, the superquantile accounts for magnitudes of low-probability large-loss tail events while the quantile does not account for this information.

The notion of *buffered probability* was originally introduced by Rockafellar and Royset (2010) in the context of the design and optimization of structures as the Buffered Probability of Failure (bPOF). Working to extend this concept, bPOE was developed as the inverse of the superquantile by Mafusalov and Uryasev (2018) in the same way that POE is the inverse of the quantile. Specifically, for continuously distributed X, bPOE at threshold x is defined in the following way, where $\sup X$ denotes the essential supremum of random variable X and threshold $x \in [E[X], \sup X]$;

$$\bar{p}_x(X) = \{1 - \alpha \mid \bar{q}_\alpha(X) = x\}. \tag{2}$$

In words, bPOE calculates one minus the probability level at which the superquantile, the tail expectation, equals the threshold x. Roughly speaking, bPOE calculates the proportion of worst-case outcomes which average to x. Figure 1 presents an illustration of bPOE for a Lognormal distributed random variable X. We note that there exist two slightly different variants of bPOE, called Upper and Lower bPOE which are identical for continuous random variables. For this paper, we utilize only continuous random variables. For the interested reader, details regarding the difference between Upper and Lower bPOE can be found in Mafusalov and Uryasev (2018).

Similar to the superquantile, bPOE is a more robust measure of tail risk, as it considers not only the probability that events/losses will exceed the threshold x, but also the magnitude of these potential events. Also, much like the superquantile, bPOE can be represented as the unique minimum of a one-dimensional convex optimization problem with the formulas given by Norton and Uryasev (2016), Mafusalov and Uryasev (2018) as follows, where $[\cdot]^+ = \max\{\cdot, 0\}$.

$$\bar{p}_x(X) = \min_{a \geq 0} E[a(X-x)+1]^+ = \min_{\gamma < x} \frac{E[X-\gamma]^+}{x-\gamma}, \quad (3)$$

$$\bar{q}_\alpha(X) = \min_{\gamma} \gamma + \frac{E[X-\gamma]^+}{1-\alpha}. \quad (4)$$

Note that formulas (3) and (4) are valid for general real valued random variables, not only continuously distributed random variables. It is also useful to note that the argmin of both (3) and (4) gives the quantile. For the bPOE calculation formula, we have that the argmin is $\gamma^* = q_\alpha(X)$ where $\alpha = 1 - \bar{p}_x(X)$ and $a^* = \frac{1}{x-\gamma^*}$ for the other representation. For the superquantile calculation formula, we have that the argmin is $\gamma^* = q_\alpha(X)$ where α was the desired probability level for calculating the superquantile.

The bPOE concept is also closely related to the concept of a superdistribution function $\bar{F}(x)$, introduced by Rockafellar and Royset (2014). For the CDF, we have that POE equals $P(X > x) = 1 - F(x)$ and we have the inverse CDF given by $F^{-1}(\alpha) = q_\alpha(X)$. The superdistribution function $\bar{F}(x)$ is motivated by the inverse relation $\bar{F}^{-1}(\alpha) = \bar{q}_\alpha(X)$. Thus, bPOE equals $1 - \bar{F}(x)$. The superdistribution function of a random variable X can also be understood as the CDF of an auxiliary random variable $\bar{X} = \bar{q}_u(X)$ where $u \sim U(0,1)$ is a uniformly distributed random variable. In this case, $\bar{F}_X(x) = F_{\bar{X}}(x)$ where the subscript indicates that it is the distribution function associated with a particular random variable.

2 Distributions with closed form superquantile and bPOE

In this section, we derive closed-form expressions for both the superquantile and bPOE for the Exponential, Pareto, Generalized Pareto, and Laplace distributions. For these distributions, we see that they exhibit a common property, where the formula for POE is identical to bPOE up to a constant. The Laplace distribution presents an interesting case in which only the right tail exhibits this reproducing property. Along the way, for completeness, we also highlight relationships between the expressions for bPOE, POE, the superquantile, and the quantile. All proofs can be found in the "Appendix".

2.1 Exponential

For this section, we have Exponential random variable $X \sim Exp(\lambda)$. Recall that the Exponential parameter has range $\lambda > 0$ with $E[X] = \sigma(X) = \frac{1}{\lambda}$, and that the Exponential CDF, PDF, and quantile are given by,

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}, \quad f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}, \quad q_\alpha(X) = \frac{-\ln(1-\alpha)}{\lambda}$$

Proposition 1 *Let $X \sim Exp(\lambda)$. Then,*

$$\bar{q}_\alpha(X) = \frac{-\ln(1-\alpha)+1}{\lambda} \quad \text{and} \quad \bar{p}_x(X) = e^{1-\lambda x}.$$

Next, we relate bPOE and POE as well as the superquantile and quantile.

Corollary 1 *Let $X \sim Exp(\lambda)$, with mean $\mu = \frac{1}{\lambda}$. Then, $\bar{p}_x(X) = P(X > x - \mu)$ and $\bar{q}_\alpha(X) = q_\alpha(X) + \mu$.*

2.2 Pareto

Assume $X \sim Pareto(a, x_m)$. Recall that Pareto parameters have range $a > 0$, $x_m > 0$ with

$$E[X] = \begin{cases} \infty, & a \in (0,1), \\ \frac{a x_m}{a-1}, & a > 1 \end{cases} \quad \text{and} \quad \sigma^2(X) = \begin{cases} \infty, & a \in (0,2], \\ \frac{a x_m^2}{(a-1)^2(a-2)}, & a > 2 \end{cases},$$

and that the Pareto CDF, PDF, and quantile are given by

$$F(x) = \begin{cases} 1 - \left(\frac{x_m}{x}\right)^a & x \geq x_m, \\ 0 & x < x_m. \end{cases}, \quad f(x) = \begin{cases} \frac{a x_m^a}{x^{a+1}} & x \geq x_m, \\ 0 & x < x_m. \end{cases}, \quad q_\alpha(X) = \frac{x_m}{(1-\alpha)^{\frac{1}{a}}}.$$

Proposition 2 *Assume $X \sim Pareto(a, x_m)$ with $a > 1$. Then, for $\alpha \in [0,1]$ and $x \geq E[X]$,*

$$\bar{q}_\alpha(X) = \frac{x_m a}{(1-\alpha)^{\frac{1}{a}}(a-1)}, \quad \bar{p}_x(X) = \left(\frac{x_m a}{x(a-1)}\right)^a.$$

Note that if $a \in (0,1]$, then $E[X] = \infty$ implying that $\bar{q}_\alpha(X) = \infty$ and $\bar{p}_x(X) = 1$ for all $\alpha \in [0,1]$ and $x \in \mathbb{R}^n$.

Corollary 2 *Relating bPOE and POE, as well as the quantile and superquantile, we can say that,*

$$\bar{p}_x(X) = P\left(X > \frac{x(a-1)}{a}\right) = P(X > x)\left(\frac{a}{a-1}\right)^a \quad \textit{and} \quad \bar{q}_\alpha(X) = q_\alpha(X)\frac{a}{a-1}.$$

2.3 Generalized Pareto distribution (GPD)

Assume $X \sim GPD(\mu, s, \xi)$. Recall that GPD parameters have range $\mu \in \mathbb{R}$, $s > 0$, $\xi \in \mathbb{R}$ with $E[X] = \mu + \frac{s}{1-\xi}$ if $\xi < 1$ and $\sigma^2(X) = \frac{s^2}{(1-\xi)^2(1-2\xi)}$ if $\xi < .5$, and that the GPD CDF and PDF are given by,

$$F(x) = \begin{cases} 1 - \left(1 + \frac{\xi(x-\mu)}{s}\right)^{-1/\xi} & \text{for } \xi \neq 0, \\ 1 - \exp\left(-\frac{x-\mu}{s}\right) & \text{for } \xi = 0. \end{cases}, \quad f(x) = \frac{1}{s}\left(1 + \frac{\xi(x-\mu)}{s}\right)^{\left(-\frac{1}{\xi}-1\right)}.$$

for $x \geq \mu$ when $\xi \geq 0$ and $\mu \leq x \leq \mu - \frac{s}{\xi}$ when $\xi < 0$. Furthermore, the quantiles are given by,

$$q_\alpha(X) = \begin{cases} \mu + \frac{s\left((1-\alpha)^{-\xi} - 1\right)}{\xi} & \text{for } \xi \neq 0, \\ \mu - s\ln(1-\alpha) & \text{for } \xi = 0. \end{cases}$$

Proposition 3 *Assume $X \sim GPD(\mu, s, \xi)$ with $-1 < \xi < 1$. Then,*

$$\bar{q}_\alpha(X) = \begin{cases} \mu + s\left[\frac{(1-\alpha)^{-\xi}}{1-\xi} + \frac{(1-\alpha)^{-\xi} - 1}{\xi}\right] & \textit{for } \xi \neq 0, \\ \mu + s[1 - \ln(1-\alpha)] & \textit{for } \xi = 0. \end{cases},$$

$$\bar{p}_x(X) = \begin{cases} \frac{\left(1 + \frac{\xi(x-\mu)}{s}\right)^{-\frac{1}{\xi}}}{(1-\xi)^{\frac{1}{\xi}}} & \textit{for } \xi \neq 0, \\ e^{1-\left(\frac{x-\mu}{s}\right)} & \textit{for } \xi = 0. \end{cases}.$$

2.4 Laplace

Assume $X \sim Laplace(\mu, b)$. Recall that Laplace parameters have range $\mu \in \mathbb{R}, b > 0$ with $E[X] = \mu$ and $\sigma^2(X) = 2b^2$, and that the Laplace CDF, PDF, and quantile function are given by,

$$F(x) = \begin{cases} 1 - \frac{1}{2}e^{-\frac{x-\mu}{b}} & x \geq \mu, \\ \frac{1}{2}e^{\frac{x-\mu}{b}} & x < \mu. \end{cases}, \quad f(x) = \frac{1}{2b}e^{\frac{-|x-\mu|}{b}},$$

$$q_\alpha(X) = \mu - b\, sign(\alpha - .5)\ \ln(1 - 2|\alpha - .5|).$$

Proposition 4 *If $X \sim Laplace(\mu, b)$, then*

$$\bar{q}_\alpha(X) = \begin{cases} \mu + b\left(\frac{\alpha}{1-\alpha}\right)(1 - \ln(2\alpha)) & \alpha < .5, \\ \mu + b\,(1 - \ln(2(1-\alpha))) & \alpha \geq .5. \end{cases},$$

$$\bar{p}_x(X) = \begin{cases} \frac{1}{2}e^{1-\left(\frac{x-\mu}{b}\right)} & x \geq \mu + b, \\ 1 + \frac{z}{\mathcal{W}(-2e^{-z-1}z)} & x < \mu + b. \end{cases}$$

where $z = \frac{x-\mu}{b}$ and $\mathcal{W}$ is the Lambert-$\mathcal{W}$ function.[3]

3 Distributions with closed form superquantile

In this section, we derive closed-form expressions for the superquantile of the Normal, Lognormal, Logistic, Student-t, Weibull, LogLogistic, and GEV distributions. The Normal, Logistic, and Student-t provide us with examples of symmetric distributions with varying tail heaviness. The Lognormal, Weibull, LogLogistic, and GEV provide us with examples of asymmetric distributions that have heavy right tails. In particular, we will utilize the Weibull formula for density estimation in Sect. 5. All proofs can be found in the "Appendix".

For these distributions, we are not able to reduce calculation of bPOE to closed-form. However, we highlight for the case of the Normal and Logistic distributions that bPOE can be calculated by solving a one-dimensional convex optimization problem or one-dimensional root finding problem. In general, we note that for continuous X, bPOE at x equals $1 - \alpha$ where α solves $\bar{q}_\alpha(X) = x$. Thus, if the superquantile is known in closed-form, this reduces to a simple one-dimensional root finding problem in α.

3.1 Normal

Let $X \sim \mathcal{N}(0, 1)$ be a standard normal random variable. Recall that

$$F(x) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{x}{\sqrt{2}}\right)\right], \quad f(x) = \frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}, \quad q_\alpha(X) = \sqrt{2}\text{erf}^{-1}(2\alpha - 1),$$

where $\text{erf}(\cdot)$ is the error function with $\text{erf}^{-1}(\cdot)$ denoting its inverse.

We show that the superquantile can be calculated by utilizing the quantile function and PDF. We also show that bPOE can be calculated in two ways: by solving a simple root finding problem involving only the PDF and CDF or by solving a convex optimization problem with gradients calculated with the error function.

[3] Also called the product logarithm or omega function.

Proposition 5 *If* $X \sim \mathcal{N}(\mu, \sigma)$, *then*

$$\bar{q}_\alpha(X) = \mu + \sigma \frac{f\left(q_\alpha\left(\frac{X-\mu}{\sigma}\right)\right)}{1-\alpha}.$$

Proposition 6 *If* $X \sim \mathcal{N}(0, 1)$, *then*

$$\bar{p}_x(X) = \min_{\gamma < x} \frac{f(\gamma) - \gamma(1 - F(\gamma))}{x - \gamma}.$$

Furthermore, if $\gamma \in argmin$, *then* γ *equals the quantile of* X *at probability level* $1 - \bar{p}_x(X)$.

Proposition 7 *Let* $X \sim \mathcal{N}(0, 1)$ *with* $x \in \mathbb{R}$ *given. If* γ *is the solution to the equation*

$$\frac{f(\gamma)}{1 - F(\gamma)} = x$$

then $\bar{p}_x(X) = \frac{f(\gamma)-\gamma(1-F(\gamma))}{x-\gamma}$. *Additionally, we will have that* $q_\alpha(X) = \gamma$ *and* $\bar{q}_\alpha(X) = x$ *at probability level* $\alpha = 1 - \bar{p}_x(X)$.

The following proposition provides the gradient calculation for solving the bPOE minimization problem.

Proposition 8 *For* $X \sim \mathcal{N}(0, 1)$, *we have that the bPOE minimization formula has the following integral representation,*

$$\begin{aligned}
\bar{p}_x(X) &= \min_{\gamma < x} \frac{f(\gamma) - \gamma(1 - F(\gamma))}{x - \gamma} \\
&= \min_{\gamma < x} \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{u e^{\frac{-(\gamma+u)^2}{2}}}{x - \gamma} du \\
&= \min_{\gamma < x} \frac{e^{\frac{-\gamma^2}{2}} - \gamma\sqrt{\frac{\pi}{2}}\,\mathrm{erfc}\left(\frac{\gamma}{\sqrt{2}}\right)}{\sqrt{2\pi}(x - \gamma)}
\end{aligned}$$

Furthermore, the function $g(u, \gamma; x) = \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{u e^{\frac{-(\gamma+u)^2}{2}}}{x-\gamma} du$ *is convex w.r.t.* γ *over the range* $\gamma \in (-\infty, x)$. *Additionally,* g *has gradient given by,*

$$\begin{aligned}
\frac{\partial g}{\partial \gamma} &= \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{\partial}{\partial \gamma} \left(\frac{u e^{\frac{-(\gamma+u)^2}{2}}}{x - \gamma} \right) du \\
&= \frac{e^{\frac{-\gamma^2}{2}} - x\sqrt{\frac{\pi}{2}}\,\mathrm{erfc}\left(\frac{\gamma}{\sqrt{2}}\right)}{\sqrt{2\pi}(x - \gamma)^2}
\end{aligned}$$

where $\mathrm{erfc}(\cdot)$ *denotes the complementary error function.*

3.2 Lognormal

Assume $X \sim Lognormal(\mu, s)$. Recall that Lognormal parameters have range $\mu \in \mathbb{R}$, $s > 0$, with $E[X] = e^{\mu + \frac{s^2}{2}}$ and $\sigma^2(X) = (e^{s^2} - 1)e^{2\mu + s^2}$ and that the Lognormal CDF, PDF, and quantile function are given by,

$$F(x) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{\ln x - \mu}{s\sqrt{2}}\right)\right], \quad f(x) = \frac{1}{xs\sqrt{2\pi}} e^{\frac{-(\ln x - \mu)^2}{2s^2}},$$
$$q_\alpha(X) = e^{\mu + s\sqrt{2}\text{erf}^{-1}(2\alpha - 1)}.$$

Proposition 9 *If $X \sim Lognormal(\mu, s)$, then*

$$\bar{q}_\alpha(X) = \frac{1}{2} e^{\mu + \frac{s^2}{2}} \frac{\left[1 + \text{erf}\left(\frac{s}{\sqrt{2}} - \text{erf}^{-1}(2\alpha - 1)\right)\right]}{1 - \alpha}.$$

3.3 Logistic

Assume $X \sim Logistic(\mu, s)$. Recall that Logistic parameters have range $\mu \in \mathbb{R}$, $s > 0$, with $E[X] = \mu$ and $\sigma^2(X) = \frac{s^2\pi^2}{3}$ and that the Logistic CDF, PDF, and quantile function are given by,

$$F(x) = \frac{1}{1 + e^{-\frac{x-\mu}{s}}}, \quad f(x) = \frac{e^{-\frac{x-\mu}{s}}}{s\left(1 + e^{-\frac{x-\mu}{s}}\right)^2}, \quad q_\alpha(X) = \mu + s\ln\left(\frac{\alpha}{1-\alpha}\right).$$

Here, we derive a closed-form expression for the superquantile for the logistic distribution and derive a simple root finding problem for calculating bPOE via (2). We also find that these quantities have a correspondence with the binary entropy function.

Proposition 10 *If $X \sim Logistic(\mu, s)$, then*

$$\bar{q}_\alpha(X) = \mu + \frac{sH(\alpha)}{1 - \alpha}$$

where $H(\alpha)$ is the binary entropy function $H(\alpha) = -\alpha\ln(\alpha) - (1-\alpha)\ln(1-\alpha)$. Furthermore, for any $x \geq \mu$, if α solves the equation,

$$\frac{H(\alpha)}{1-\alpha} = \frac{x - \mu}{s},$$

then $\bar{p}_x(X) = 1 - \alpha$. Additionally, $\bar{p}_x(X) = 1 - \alpha$ if α is the solution to the transformed system,

$$(1-\alpha)\alpha^{\frac{\alpha}{1-\alpha}} = e^{-\left(\frac{x-\mu}{s}\right)}.$$

Note that both functions $\frac{H(\alpha)}{1-\alpha}$ and $(1-\alpha)\alpha^{\frac{\alpha}{1-\alpha}}$ are one-dimensional, convex, and monotonic over the range $\alpha \in [0, 1]$, and thus unique solutions exist and can easily be found via root finding methods.

We can also utilize the minimization formula to calculate bPOE. Calculating bPOE in this way has the added benefit of simultaneously calculating the quantile $q_{1-\bar{p}_x(X)}(X)$.

Proposition 11 *If $X \sim Logistic(\mu, s)$, then*

$$\bar{p}_x(X) = \min_{\gamma < x} \frac{s\ln\left(1 + e^{-\left(\frac{\gamma-\mu}{s}\right)}\right)}{x - \gamma},$$

which is a convex optimization problem over the range $\gamma \in (-\infty, x)$. Furthermore, the minimum occurs at γ such that,

$$\frac{s \ln\left(1 + e^{-\left(\frac{\gamma-\mu}{s}\right)}\right)}{x-\gamma} = 1 - F(\gamma).$$

3.4 Student-*t*

Assume $X \sim Student\text{-}t(\nu, s, \mu)$. Recall that Student-$t$ parameters have range $\nu > 0, s > 0$, $\mu > 0$ with $E[X] = \mu$ and $\sigma^2(X) = \frac{s^2\nu}{\nu-2}$, and that the Student-$t$ CDF and PDF are given by,

$$F(x) = 1 - \frac{1}{2}\mathcal{I}_{\nu(x)}\left(\frac{\nu}{2}, \frac{1}{2}\right), \quad f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\nu\pi}s}\left(1 + \frac{(x-\mu)^2}{\nu s^2}\right)^{\frac{-(\nu+1)}{2}},$$

where $\nu(x) = \frac{\nu}{\frac{x-\mu}{s}+\nu}$, $\mathcal{I}_t(a, b)$ is the regularized incomplete Beta function, and $\Gamma(a)$ is the Gamma function. Note that a general closed-form expression for $q_\alpha(X)$ is not known but is a readily available function within common spreadsheet and statistical packages.

Proposition 12 *If $X \sim Student\text{-}t(\nu, s, \mu)$, then*

$$\bar{q}_\alpha(X) = \mu + s\left(\frac{\nu + T^{-1}(\alpha)^2}{(\nu-1)(1-\alpha)}\right)\tau(T^{-1}(\alpha))$$

where $T^{-1}(\alpha)$ is the inverse of the standardized Student-t CDF and $\tau(x)$ is standardized Student-t PDF.

3.5 Weibull

Assume $X \sim Weibull(\lambda, k)$. Recall that Weibull parameters have range $\lambda > 0, k > 0$ with $E[X] = \lambda\Gamma(1 + \frac{1}{k})$ and $\sigma^2(X) = \lambda^2\left[\Gamma(1 + \frac{2}{k}) - \Gamma(1 + \frac{1}{k})^2\right]$, and that the Weibull CDF, PDF, and quantile function are given by

$$F(x) = 1 - e^{-(x/\lambda)^k}, \quad f(x) = \begin{cases} \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} & x \geq 0, \\ 0 & x < 0, \end{cases},$$
$$q_\alpha(X) = \lambda(-\ln(1-\alpha))^{1/k}.$$

where $\Gamma(a) = \int_0^\infty p^{a-1}e^{-p}dp$ is the gamma function.

Proposition 13 *If $X \sim Weibull(\lambda, k)$, then*

$$\bar{q}_\alpha(X) = \frac{\lambda}{1-\alpha}\Gamma_U\left(1 + \frac{1}{k}, -\ln(1-\alpha)\right)$$

where $\Gamma_U(a, b) = \int_b^\infty p^{a-1}e^{-p}dp$ is the upper incomplete gamma function.

3.6 Log–logistic

Assume $X \sim LogLogistic(a, b)$. Recall that Log-Logistic parameters have range $a > 0$, $b > 0$ with $E[X] = a\frac{\pi}{b} csc\left(\frac{\pi}{b}\right)$ when $b > 1$ and $\sigma^2(X) = a^2\left(\frac{2\pi}{b} csc\left(\frac{2\pi}{b}\right) - (\frac{\pi}{b} csc\left(\frac{\pi}{b}\right))^2\right)$ when $b > 2$, and that the Log-Logistic CDF, PDF, and quantile function are given by,

$$F(x) = \frac{1}{1 + \left(\frac{x}{a}\right)^{-b}}, \quad f(x) = \frac{(b/a)(x/a)^{b-1}}{\left(1 + (x/a)^b\right)^2}, \quad q_\alpha(X) = a\left(\frac{\alpha}{1-\alpha}\right)^{\frac{1}{b}},$$

where $csc(\cdot)$ is the cosecant function.

Proposition 14 *If $X \sim LogLogistic(a, b)$, then*

$$\bar{q}_\alpha(X) = \frac{a}{1-\alpha}\left(\frac{\pi}{b} csc\left(\frac{\pi}{b}\right) - B_\alpha\left(\frac{1}{b} + 1, 1 - \frac{1}{b}\right)\right)$$

where $B_y(A_1, A_2) = \int_0^y p^{A_1-1}(1-p)^{A_2-1}dp$ is the incomplete beta function.

3.7 Generalized extreme value distribution

Assume X follows a Generalized Extreme Value (GEV) distribution, which we denote as $X \sim GEV(\mu, s, \xi)$. Recall that GEV parameters have range $\mu \in \mathbb{R}$, $s > 0$, $\xi \in \mathbb{R}$ with

$$E[X] = \begin{cases} \mu + s(g_1 - 1)/\xi & \text{if } \xi \neq 0, \xi < 1, \\ \mu + s\,y & \text{if } \xi = 0, \\ \infty & \text{if } \xi \geq 1, \end{cases} \quad \text{and}$$

$$\sigma^2(X) = \begin{cases} s^2\,(g_2 - g_1^2)/\xi^2 & \text{if } \xi \neq 0, \xi < \frac{1}{2}, \\ s^2\,\frac{\pi^2}{6} & \text{if } \xi = 0, \\ \infty & \text{if } \xi \geq \frac{1}{2}, \end{cases}$$ where $g_k = \Gamma(1 - k\xi)$ and y is the Euler-Mascheroni constant.

Additionally, recall that the GEV has CDF, PDF, and quantile function given by,

$$F(x) = \begin{cases} e^{-\left(1+\frac{\xi(x-\mu)}{s}\right)^{\frac{-1}{\xi}}} & \xi \neq 0, \\ e^{-e^{-\left(\frac{x-\mu}{s}\right)}} & \xi = 0 \end{cases}, .$$

$$f(x) = \begin{cases} \frac{1}{s}\left(1 + \frac{\xi(x-\mu)}{s}\right)^{\frac{-1}{\xi}-1} e^{-\left(1+\frac{\xi(x-\mu)}{s}\right)^{\frac{-1}{\xi}}} & \xi \neq 0, \\ \frac{1}{s}e^{-\left(\frac{x-\mu}{s}\right)}e^{-e^{-\left(\frac{x-\mu}{s}\right)}} & \xi = 0, \end{cases}$$

$$q_\alpha(X) = \begin{cases} \mu + \frac{s}{\xi}\left((\ln(\frac{1}{\alpha})^{-\xi} - 1\right) & \xi \neq 0, \\ \mu - s\ln(-\ln(\alpha)) & \xi = 0. \end{cases}$$

Proposition 15 *If $X \sim GEV(\mu, s, \xi)$, then*

$$\bar{q}_\alpha(X) = \begin{cases} \mu + \frac{s}{\xi(1-\alpha)}\left[\Gamma_L(1-\xi, \ln(\frac{1}{\alpha})) - (1-\alpha)\right] & \xi \neq 0, \\ \mu + \frac{s}{(1-\alpha)}\left(y + \alpha\ln(-\ln(\alpha)) - \text{li}(\alpha)\right) & \xi = 0, \end{cases}$$

where $\Gamma_L(a,b) = \int_0^b p^{a-1}e^{-p}dp$ *is the lower incomplete gamma function,* $\mathrm{li}(x) = \int_0^\alpha \frac{1}{\ln p}dp$ *is the logarithmic integral function, and* y *is the Euler-Mascheroni constant.*

4 Portfolio optimization

A common parametric approach to portfolio optimization is to assume that portfolio returns follow some specified distribution. With this assumption, particularly when taking a risk averse approach, closed-form representations of the superquantile and bPOE for the specified distribution allow one to formulate a tractable portfolio optimization problem. In this section, we show that our derived formulas for the superquantile and bPOE reveal important properties about portfolio optimization problems formulated with particular distributional assumptions placed upon portfolio returns.

Portfolio optimization with the superquantile is common, so we begin by simply pointing out which of the closed-form superquantile formulas yield tractable portfolio optimization problems. Portfolio optimization with bPOE, however, is not common and we show that it can be advantageous compared to the superquantile approach. In particular, superquantile optimization requires that one sets the probability level α. One can then observe that for fixed α, the optimal superquantile portfolio may change based upon the distribution utilized to model returns. We show that if portfolio returns are assumed to follow a Laplace, Logistic, Normal, or Student-t distribution, the minimal bPOE portfolios for fixed threshold x are the same regardless of the distribution chosen, meaning that there exists a single portfolio that is x-bPOE optimal for multiple choices of distribution.

Note that in this section we will be dealing with asset returns $\mathcal{R}$, as it is typical for financial problems, and the loss is the opposite of return: $X = -\mathcal{R}$, and $q_\alpha(X) = -q_{1-\alpha}(\mathcal{R})$.

The portfolio optimization problem consists of finding a vector of asset weights $w \in \mathbb{R}^n$ for a set of n assets with unknown random returns $R = [R_1, R_2, ..., R_n]$ that solves the following optimization problem,

$$\begin{aligned} \max_{w\in\mathbb{R}^n} \quad & L(w,R) \\ s.t. \quad & g_i(w,R) \le 0, i = 1, ..., I \\ & h_j(w,R) = 0, j = 1, ..., J \\ & w^T 1 = 1 \\ & l \le w \le u \end{aligned} \tag{5}$$

where $L(w, R)$ is some function to be maximized,[4] functions $g_j(w, R)$ and $h_i(w, R)$ enforce inequality and equality constraints, and vectors l, u enforce upper and lower bounds on the individual asset weights. A simple example is the standard Markowitz optimization problem where we maximize the expected utility, which is a weighted combination of the expected return and its variance via a positive trade-off parameter $\lambda \ge 0$:

$$\begin{aligned} \max_{w\in\mathbb{R}^n} \quad & w^T\eta - \lambda w^T \Sigma w \\ s.t. \quad & w^T 1 = 1 \\ & l \le w \le u \end{aligned} \tag{6}$$

[4] Or minimized if we consider the negative.

An important aspect of the random portfolio return $w^T R$ which can be seen within the Markowitz problem and will be used later on in this section is the fact that the expectation $E[w^T R]$ and variance $\sigma^2(w^T R)$ are given by $w^T \eta$ and $w^T \Sigma w$ respectively, where $\eta \in \mathbb{R}^n$ is the vector of expected returns for the n assets and $\Sigma \in \mathbb{R}^{n\times n}$ is the covariance matrix for the n assets. This allows us to represent the expected value and variance of the portfolio return in terms of w, and consequently to formulate an optimization problem with decision vector w.

4.1 Superquantile and bPOE optimization with qualified distributions

As we are dealing with asset returns, and not losses, we need to define the superquantile using that notation. The superquantile is the expected loss above the quantile (conditional expected value of losses in the right tail), so in terms of returns it would be the conditional expected value of returns in the left tail, which can be described by the left superquantile:

$$\tilde{q}_{1-\alpha}(\mathcal{R}) = \frac{1}{1-\alpha}\int_0^{1-\alpha} q_p(\mathcal{R})dp.$$

We can use the closed-form superquantile formulas derived in the previous sections for the right superquantile $\bar{q}_\alpha(R)$ to calculate the left superquantile $\tilde{q}_\alpha(\mathcal{R})$, as

$$\alpha\tilde{q}_\alpha(\mathcal{R}) + (1-\alpha)\bar{q}_\alpha(\mathcal{R}) = \int_0^1 q_p(\mathcal{R})dp = E[\mathcal{R}],$$

so

$$\tilde{q}_{1-\alpha}(\mathcal{R}) = \frac{1}{1-\alpha}(E[\mathcal{R}] - \alpha\bar{q}_{1-\alpha}(\mathcal{R})).$$

Since $-\tilde{q}_{1-\alpha}(\mathcal{R}) = \bar{q}_\alpha(X)$, bPOE is given as

$$\bar{p}_x(X) = \{1-\alpha | \bar{q}_\alpha(X) = x\} = \bar{p}_x(-\mathcal{R}) = \{1-\alpha | \tilde{q}_{1-\alpha}(\mathcal{R}) = -x\}.$$

4.1.1 Qualified distributions for portfolio optimization

The superquantile or bPOE portfolio optimization problem has objective or constraints formulated with either $\tilde{q}_{1-\alpha}(w^T R)$ or $\bar{p}_x(w^T R)$. To formulate such a problem using a given distribution, we begin by defining a set of *qualified* distributions which we will consider. These qualified distributions satisfy the following requirement that the distributions admit a superquantile/bPOE formula in terms of decision variable w:

Definition 1 (*Qualified Distribution*) A *qualified* distribution $\mathcal{D}$ satisfies the following condition:
(C1) $w^T R \sim \mathcal{D} \implies \tilde{q}_{1-\alpha}(w^T R) = w^T\eta - \sqrt{w^T \Sigma w}\zeta(\alpha)$, where $\zeta(\alpha)$ is a function depending upon α, η is the vector of the expected asset returns, and Σ is the covariance matrix for asset returns.

Why should we enforce this precondition? Condition (C1) guarantees that the superquantile can be expressed in terms of w. This is necessary to express the superquantile optimization problem. For example, if we assume that $w^T R \sim Logistic(\mu, s)$, we need to be able to express μ and s in terms of w. Since $\mu = E[w^T R] = w^T\eta$ and $w^T \Sigma w = \sigma^2(w^T R) = \frac{s^2\pi^2}{3}$, we have

$$\begin{aligned}\tilde{q}_{1-\alpha}(\mathcal{R}) &= \frac{1}{1-\alpha}(E[\mathcal{R}] - \alpha\bar{q}_{1-\alpha}(\mathcal{R})) \\ &= \frac{1}{1-\alpha}\left(\mu - \alpha\left[\mu + \frac{s}{\alpha}(-(1-\alpha)ln(1-\alpha) - \alpha ln(\alpha))\right]\right) \\ &= \mu - \frac{s}{1-\alpha}(-\alpha ln(\alpha) - (1-\alpha)ln(1-\alpha)) \\ &= w^T\eta - \sqrt{w^T\Sigma w}\frac{\sqrt{3}(-\alpha\ln(\alpha) - (1-\alpha)\ln(1-\alpha))}{\pi(1-\alpha)},\end{aligned}$$

which satisfies (C1). Other examples that satisfy this condition are Laplace, Normal, Exponential, Student-t, Pareto, GPD, and GEV. Note that for Student-t, we assume that parameter ν is fixed and the same for all assets, i.e. $\Theta = \{\nu\}$, and for GPD/GEV distributions $\Theta = \{\xi\}$.

Additionally, we note that in the portfolio optimization context we would typically like to have our distribution satisfy the following informally stated conditions for the distribution of returns $w^T R \sim \mathcal{D}$:

(C2) The statistical parameters of the distribution $\mathcal{D}$ must be consistent with the descriptive statistics of real-life asset returns.

(C3) The shape of the PDF for the given distribution $\mathcal{D}$ must conform to the shape of the empirical distribution of typical real-life asset returns.

Condition (C2) and (C3) are simple sanity checks on our model assumptions in a portfolio context. For example, for exponential distribution $E[\mathcal{R}] = \frac{1}{\lambda} = \sigma(\mathcal{R})$, but for the real-life asset returns the sample mean is not generally equal to the sample standard deviation. So, Exponential and Pareto distributions make no sense in portfolio optimization problems even if they satisfy (C1). As for (C3), a distribution is not practical if there is obvious discrepancy between the shape of its PDF and the shape of the empirical PDF observed using real-life asset returns. The latter is generally bell-shaped or, more likely, inverse-V shaped, and is never shaped like the PDF of an Exponential, Pareto/GPD, or Weibull for $k < 1$.

This leaves us with a set of four elliptical distributions which satisfy all three conditions: Logistic, Laplace, Normal, and Student-t, as well as the nonelliptical GEV distribution. For the latter, with $\xi \neq 0$ the left superquantile can be expressed as

$$\tilde{q}_{1-\alpha}(\mathcal{R}) = w^T\nu - \sqrt{w^T\Sigma w}\frac{\alpha\Gamma(1-\xi) - \Gamma_U(1-\xi, ln(\frac{1}{1-\alpha}))}{(1-\alpha)\sqrt{g_2 - g_1^2}},$$

where $\Gamma_U(a, b) = \int_b^\infty p^{a-1}e^{-p}dp$ is the upper incomplete Gamma function, $g_k = \Gamma(1 - k\xi)$.

4.1.2 Superquantile and bPOE optimization

An alternative to the Markowitz problem is to find the portfolio with minimal superquantile (7) or bPOE (8).

$$\begin{aligned} \min_{w\in\mathbb{R}^n} \quad & -\tilde{q}_{1-\alpha}(w^T R) \\ s.t. \quad & w^T 1 = 1 \\ & l \leq w \leq u \end{aligned} \tag{7}$$

$$\begin{aligned} \min_{w \in \mathbb{R}^n} \quad & \bar{p}_x(-w^T R) \\ s.t. \quad & w^T 1 = 1 \\ & l \le w \le u \end{aligned} \tag{8}$$

For qualified distributions, however, these problems can be greatly simplified. First, we see that (7) reduces to (9):

$$\begin{aligned} \max_{w \in \mathbb{R}^n} \quad & w^T \eta - \sqrt{w^T \Sigma w} \zeta(\alpha) \\ s.t. \quad & w^T 1 = 1 \\ & l \le w \le u \end{aligned} \tag{9}$$

A well-known result, which can be seen by simple analysis of optimality conditions, is that (9) and (6) are equivalent. This is to say that any optimal solution obtained by solving (9) with some α can be obtained by solving (6) with some λ. Thus, the superquantile optimal portfolio is also mean−variance optimal in a Markowitz sense.

Now, for bPOE we see that the picture is actually much simpler. Specifically, we have the following proposition.

Proposition 16 *If we assume that $w^T R \sim \mathcal{D}$ and we have that $\mathcal{D}$ is a qualified distribution, then* (8) *reduces to* (10).

$$\begin{aligned} \max_{w \in \mathbb{R}^n} \quad & \frac{w^T \eta + x}{\sqrt{w^T \Sigma w}} \\ s.t. \quad & w^T 1 = 1 \\ & l \le w \le u \end{aligned} \tag{10}$$

Proof Since $\bar{p}_x(X) = \bar{p}_x(-R) = \{1 - \alpha | \tilde{q}_{1-\alpha}(R) = -x\}$ and $\tilde{q}_{1-\alpha}(w^T R) = w^T \eta - \sqrt{w^T \Sigma w} \zeta(\alpha)$ for qualified distributions, the problem (8) can be rewritten as:

$$\begin{aligned} \min_{w \in \mathbb{R}^n, \alpha} \quad & 1 - \alpha \\ s.t. \quad & - w^T \eta + \sqrt{w^T \Sigma w} \zeta(\alpha) = x \\ & w^T 1 = 1 \\ & l \le w \le u \end{aligned} \tag{11}$$

which can further be written as:

$$\begin{aligned} \max_{w \in \mathbb{R}^n} \quad & \alpha \\ s.t. \quad & \zeta(\alpha) = \frac{w^T \eta + x}{\sqrt{w^T \Sigma w}} \\ & w^T 1 = 1 \\ & l \le w \le u \end{aligned} \tag{12}$$

Next, note that by definition of the superquantile (both left and right), we know that $\zeta(\alpha)$ must be an increasing function w.r.t. $\alpha \in [0, 1]$. Finally, since $\zeta(\alpha)$ is an increasing function w.r.t. α, we see that we can formulate the maximization as (10) without changing the argmax. □

This proposition has the important implication for portfolio theory that the optimal bPOE portfolio is simultaneously optimal for all qualified distributions at the same threshold x. This means that bPOE minimization is advantageous compared to superquantile minimization, particularly when one wants to asses how the risk associated with the *optimal* portfolio is affected by distributional assumptions. For example, imagine you want to know how distributional assumptions affect the risk of the optimal portfolio. If we choose to solve for superquantile optimal portfolios at fixed level $\alpha = .95$ and want to see how the value of the optimal superquantile value changes as distributions change, we would need to re-solve the optimization problem for every distribution choice because changing the distribution may alter the optimal portfolio composition. However, if we solve the bPOE optimization for fixed x, we need to only solve a single optimization problem to yield a portfolio that will be optimal for all distribution choices. Critically, while the same portfolio will be optimal for all distributions, the value of bPOE at x for the *optimal* portfolio will still depend upon the distribution choice! Thus, in one solve we can see how bPOE at x (the risk) will change for the optimal portfolio as distributional assumptions change. Therefore, it is much easier to see how the risk (bPOE) associated with the optimal portfolio changes when only distributional assumptions change.

Additionally, if the optimal portfolio has bPOE at x for a particular distribution equaling $1 - \alpha$, we know that it would be an optimal portfolio for the left superquantile maximization problem with parameter $1 - \alpha$.[5] This helps the user not only assess the sensitivity to distributional choices, but also the relationship with choices of $1 - \alpha$ for the corresponding superquantile minimization problems.

4.1.3 Numerical demonstration

In this example, we consider a global equity portfolio that consists of 6 market portfolios—U.S., Japan, U.K., Germany, France, and Switzerland, represented by the corresponding MSCI indices – MXUS, MXJP, MXGB, MXDE, MXFR, MXCH. Parameters of returns for portfolio components are provided in Table 1 (source: Capital IQ sample of monthly returns from April 1987 to April 1996, annualized).

We first find the minimal superquantile (CVaR) portfolios with $\alpha = \{.95, .99\}$, denoted as CVaR 99% and CVaR 95% portfolios. This problem was solved using a non-linear programming algorithm and the results are provided in Table 2. The respective values of λ are also provided, which allows deriving the same portfolios using the standard MVO solver that uses a quadratic programming algorithm.

Table 2 shows that superquantile optimal portfolios are not the same as the global minimum variance portfolio (min. risk portfolio), but are quite close. Distributional assumptions play their role in the optimal portfolio's composition, with the Student-t distribution rendering the most conservative allocation for CVaR 99%. Thus, as noted before, if one wanted to assess the affect that distributional choices have on optimal CVaR 95% portfolios, four optimization problems would need to be solved; one for each distribution choice.

We can note from (9) that if portfolio return is constrained from below, unless this constraint is very close to the return of the global minimum variance portfolio, it results in the superquantile optimization being essentially the same as the variance minimization. If the risk is constrained from above, that superquantile optimization is the same as return maximization.

[5] If this was not true, it would imply that there exists a portfolio with smaller bPOE and hence the current portfolio is not bPOE optimal. A formal proof can be found in Mafusalov and Uryasev (2018).

Table 1 Portfolio return data

Asset ticker	Expected return (%)	SD (%)	Correlations					
			MXUS	MXJP	MXGB	MXDE	MXFR	MXCH
MXUS	10.25	13.79	1	0.190041	0.639133	0.481857	0.499406	0.605384
MXJP	6.90	26.05	0.190041	1	0.450337	0.251601	0.378753	0.373964
MXGB	8.81	19.16	0.639133	0.450337	1	0.579918	0.584215	0.654687
MXDE	9.15	20.31	0.481857	0.251601	0.579918	1	0.753072	0.628426
MXFR	8.83	20.40	0.499406	0.378753	0.584215	0.753072	1	0.580626
MXCH	13.85	17.45	0.605384	0.373964	0.654687	0.628426	0.580626	1

Table 2 Optimal superquantile (CVaR) portfolios

Asset ticker	Min risk portfolio (%)	CVaR 99% optimal portfolios				CVaR 95% optimal portfolios			
		Normal	t (df = 3)	Laplace	Logistic	Normal	t (df = 3)	Laplace	Logistic
MXUS	70.99	65.80%	67.59%	67.03%	66.53%	64.23%	64.78%	65.05%	64.64%
MXJP	13.98	9.61%	11.11%	10.64%	10.21%	8.28%	8.74%	8.97%	8.62%
MXGB	0.00	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
MXDE	9.24	2.87%	5.07%	4.37%	3.76%	0.95%	1.61%	1.94%	1.44%
MXFR	0.00	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
MXCH	5.79	21.72%	16.22%	17.96%	19.50%	26.54%	24.87%	24.04%	25.30%
Return	9.89	10.68%	10.40%	10.49%	10.57%	10.91%	10.83%	10.79%	10.85%
SD	12.86	13.01%	12.93%	12.95%	12.97%	13.11%	13.08%	13.06%	13.09%
λ		20.48	31.28	26.82	23.80	15.73	17.11	17.88	16.73

Using the same set of assets, we also solved the bPOE optimization problem (8) with thresholds $x = 0.16$ and $x = 0.25$ (i.e. losses exceeding 16% and 25%), $l = 0$, and $u = 1$.[6] Table 3 shows the results: the minimal bPOE achieved, the optimal portfolio composition and parameters, and CVaR[7] for the optimal portfolios for all distributions. All the optimal solutions were generated by the non-linear programming algorithm for different distributions, but the results support the conclusion that the optimal portfolio composition does not depend on the distribution (small discrepancies are due to the optimization algorithm accuracy). One can see, however, that the value of bPOE does indeed change as distribution assumptions change. For instance, we see that the optimal portfolio at threshold $x = .16$ has larger bPOE for Laplace assumptions compared to Student-t assumptions. For $x = .25$, however, we see this trend reversed. Overall, with only two solves, one for each threshold choice $x = .16, .25$, we are able to asses the changes in risk that occur for optimal portfolios across 4 different distribution choices. This is compared to eight solves required for the same analysis of optimal CVaR portfolios with $\alpha = .95, .99$.

As noted in Sect. 4.1.2, the value of bPOE for a bPOE optimal portfolio also tells us the level of α required for it to be an optimal CVaR portfolio under the same distributional assumption. We can see this reflected by comparing Tables 2 and 3. If we look at the optimal bPOE value in Table 3 for the normal distribution, we see that bPOE equals 5.13% and .8% respectively for threshold choices $x = 16\%$ and $x = 25\%$. This implies that these portfolios would be optimal if we minimized the (right) superquantile (of losses) at probability level $\alpha = .9487$ and $\alpha = .992$. Looking at Table 2, we see that indeed this is approximately the case. Specifically, we see that an almost identical portfolio composition is achieved when the superquantile was minimized with $\alpha = .95$ and $\alpha = .99$ and the distribution was assumed to be normal.

5 Parametric density estimation with superquantiles

One of the motivations for providing closed-form superquantile formulas is so that they can be used within common parametric estimation frameworks. The Exponential, Parteo/GPD, Laplace, Normal, Lognormal, Logistic, Student-t, Weibull, LogLogistic, and GEV represent a wide range of distributions that can now be utilized within these parametric procedures, but with superquantiles incorporated into the fitting criteria. We illustrate this idea by proposing a simple variation of the Method of Moments (MM), which we call the Method of superquantiles (MOS), where superquantiles at varying levels of α take the place of moments. Our numerical example utilizes a heavy tailed Weibull to illustrate MOS, since it is particularly well-suited for asymmetric heavy-tailed data. However, any of the listed distributions could be used as well.

5.1 Method of superquantiles

The MM is a well known tool for estimating the parameters of a distribution when moments are available in parametric form and desired moments are either assumed to be known or are measured from empirical observations. It looks for the distribution $f_\Theta(x)$, parameterized by Θ, with moments equal to some known moments or, if moments are unknown, empirical

[6] Note that bPOE thresholds refer to loss thresholds instead of return thresholds. See Sect. 4.1.

[7] Right superquantile of loss distribution.

Table 3 Optimal bPOE portfolios

Assumed distribution	Normal (%)	t (df = 3) (%)	Laplace (%)	Logistic (%)	Normal (%)	t (df = 3) (%)	Laplace (%)	Logistic (%)
bPOE threshold, x	16				25			
bPOE value, $\bar{p}_x(X)$	5.13	6.21	7.46	6.36	0.80	2.93	2.81	1.86
Asset tiker	bPOE-optimal portfolio composition							
MXUS	64.20	64.19	64.20	64.20	65.95	65.95	65.95	65.95
MXJP	8.26	8.27	8.25	8.25	9.73	9.73	9.73	9.73
MXGB	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MXDE	0.90	0.91	0.90	0.90	3.05	3.05	3.06	3.05
MXFR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MXCH	26.64	26.63	26.64	26.65	21.27	21.27	21.27	21.27
Return	10.92	10.92	10.92	10.92	10.65	10.65	10.65	10.65
SD	13.12	13.12	13.12	13.12	13.00	13.00	13.00	13.00
Test distribution	CVaR for the test distributions							
normal	16.00	14.93	13.87	14.79	25.00	18.95	19.16	21.16
t (df = 3)	18.14	16.00	14.05	15.74	46.31	25.00	25.56	31.46
Laplace	19.48	17.70	16.00	17.48	36.62	24.61	25.00	28.79
logistic	17.61	16.18	14.81	16.00	31.14	21.71	22.01	25.00

moments. With n moments used, the problem reduces to solving a set of n equations w.r.t. the set of parameters Θ of the distribution family.

This method, though, can be generalized where moments are replaced by other distributional characteristics, such as the superquantile and quantile. We utilize superquantiles. This method provides flexibility through the choices of different α, allowing the user to focus the fitting procedure on particular portions of the distribution. This flexibility is advantageous compared to other methods such as MM or maximum likelihood (ML), since these fitting methods treat each portion of the distribution equally. When fitting the tail is important, for example, and there are many samples around the mean with few samples in the tail, it can be desirable to focus the fitting procedure on carefully fitting the tail samples. As will be shown, one can focus MOS by choice of α. One will see that this procedure is similar to fitting with Probability Weighted Moments (PWM),[8] but where MOS is much more straightforward with superquantiles far easier to interpret than PWMs.

We formulate the following problem, where $\hat{\bar{q}}_\alpha(X)$ denotes either a known superquantile or an empirical estimate[9] from a sample of X and $\bar{q}_\alpha(X_{f_\Theta})$ denotes parameterized formulas for the superquantile when X has density function f_Θ with the set of parameters Θ:

Method of Superquantiles Fix $\alpha_1, ..., \alpha_k \in [0, 1]$ and choose a parametric distribution family f_Θ with parameters Θ. Solve for Θ such that,

$$\bar{q}_{\alpha_i}(X_{f_\Theta}) = \hat{\bar{q}}_{\alpha_i}(X) \quad \text{for all } i = 1, ..., k,$$

which is a system of k equations in $|\Theta|$ unknowns.

This problem, however, may not have a solution. For example, if $k = 2$ and the parametric family only has a single parameter (i.e. $|\Theta| = 1$). In this case, one can solve the following surrogate Least Squares minimization problem:

LS Method of Superquantiles (LS-MOS) Fix $\alpha_1, ..., \alpha_k \in [0, 1]$ and choose a parametric distribution family f_Θ with parameters Θ. Choose weights $c_1, ..., c_k > 0$ and solve for,

$$\Theta \in \operatorname*{argmin}_{\Theta} \sum_i c_i \left(\bar{q}_{\alpha_i}(X_{f_\Theta}) - \hat{\bar{q}}_{\alpha_i}(X)\right)^2.$$

This procedure finds the distribution which has superquantiles that are *close* to the empirical superquantiles. The freedom to select α_i as well as c_i provides the user with much flexibility as to which portion of the distribution should match more exactly the empirical superquantiles.

5.1.1 Example customization: conservative tail fitting

When sample size is small and the tail of the distribution at hand is long, it is likely that the tail will be difficult to characterize from empirical data since few observations will be observed in the tail (with high probability). The proposed method of superquantiles can easily, however, be made more conservative based upon empirical data in an intuitive way. For example, one could have the following condition where ϵ_i is a pre-specified constant such that $0 < \epsilon_i \leq \alpha_i$:

$$\bar{q}_{\alpha_i - \epsilon_i}(X_{f_\Theta}) = \hat{\bar{q}}_{\alpha_i}(X).$$

[8] Also sometimes called L-moments.

[9] A simple empirical estimate from a sample of N observations $S = \{X_1, \ldots, X_N\}$ can be obtained by sorting S and calculating the average of the largest $(1 - \alpha)N$ observations. More precise estimates can be obtained by weighted averages; see Proposition 8 of Rockafellar and Uryasev (2002) for details.

Table 4 Parameter values

Method	Fit on S_1		Fit on S_2		k = 1	λ = .5
	k = 1.4	λ = .5	k = 1.4	λ = .5		
LS1	1.63	.52	1.27	.46	.92	.45
LS2	**1.41**	**.48**	**1.42**	**.49**	**1.02**	**.49**
MM	1.64	.52	1.29	.46	.95	.45
ML	1.71	.53	1.21	.45	.86	.43

Bold indicates the estimated parameters that are closest to the true value

Or, for the least squares variant, one can fit the problem,

$$\min_{\Theta} \sum_i c_i \left(\bar{q}_{\alpha_i - \epsilon_i}(X_{f_\Theta}) - \hat{\bar{q}}_{\alpha_i}(X)\right)^2$$

Notice that these conditions are effectively making the assumption that the empirical superquantile has underestimated the true tail expectation, which is often the case with heavy tailed distributions.

5.1.2 Example: Weibull distribution fitting

We illustrate the basic method on fitting a Weibull distribution, with $\Theta = (\lambda, k)$, from a small sample of 50 observations. We took two independent samples, denoted S_1, S_2, of size 50 from a Weibull with $\lambda = .5, k = 1.4$. We then estimated the Weibull parameters using MM, ML, and the LS-MOS. The MM was solved using the first two moments. The LS-MOS was solved twice. It was first solved with $\alpha_1 = .15, \alpha_2 = .75, c_1 = c_2 = 1$, a choice which was made to mimic the behavior of MM and ML, where the fit emphasizes most of the observed data. To put more emphasis on the tail observations, it was also solved with $\alpha_1 = .5, \alpha_2 = .75, \alpha_3 = .95, c_1 = c_2 = c_3 = 1$. We denote these solutions as LS1, LS2 respectively. The ML solution is available in closed-form and we solved MM, LS1, and LS2 using Scipy's[10] optimization library.[11] Parameters obtained from each fitting procedure are shown in Table 4.

Looking at Fig. 2 for S_1 and Fig. 3 for S_2, we see that the LS1 fit is, indeed, much like the MM a ML fit for both data sets. However, we see that the LS2 fit is the best in both cases. This assessment is both visual and can be seen by looking at the fit parameter values in Table 4, which are much closer to the true parameter values. The ML, MM, and LS1 methods have put too much emphasis on the observations around the mode. The LS2 fit, however, has put appropriate emphasis on the less frequent observations in the tail.

It is also important to notice how the differences in S_1 and S_2 have affected the fit from each method. Looking at the differences between S_1 and S_2, we can see that the samples differ in the observed density in the lower portion of the range. This is directly reflected in the fits given by MM, ML, and LS1. Compared to their fits on S_1, they are more heavily favoring the left side of the distribution. The LS2 fit, however, is robust to these differences between data sets and, by focusing on the tail, has remained mostly unchanged from the fit on S_1. This is the intended effect from the selection of larger values for α in LS2.

10 www.scipy.org.

11 Specifically, we used the *leastsq* function which implements MINPACK's *lmdif* routine. This routine requires function values and calculates the Jacobian by a forward-difference approximation.

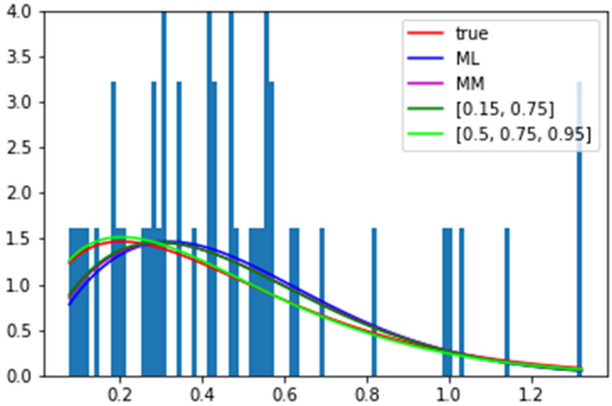

Fig. 2 Fits using sample S_1. PDFs displayed with normalized histogram of S_1 sample given in background

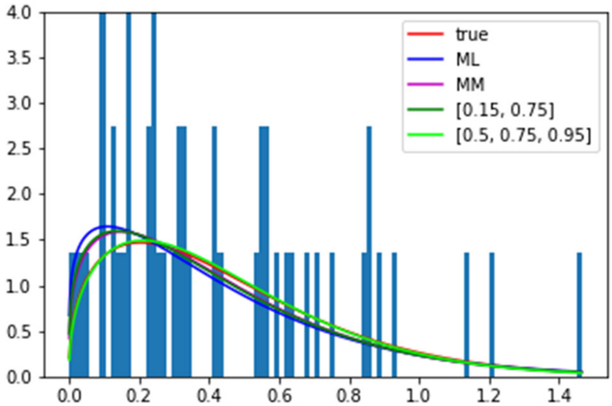

Fig. 3 Fits using sample S_2. PDFs displayed with normalized histogram of S_2 sample given in background

We duplicated this procedure on a heavier tailed Weibull. We took 50 samples of a Weibull with true parameters $k = 1$, $\lambda = .5$ and fit MM, ML, LS1, and LS2 using the empirical data. Figures 4 and 5 highlight different aspects of the resulting fits with Table 4 also showing the fit parameter values for each method. We see that LS2 clearly provides the best fit, with Fig. 5 in particular showing that MM, ML, and LS1 underestimate the tail densities. Figure 5 shows a zoomed in view of the tail of the distribution. MM, ML, and LS1 put more emphasis on fitting the observations around the mode. As intended, however, LS2 focuses more on fitting the right tail observations and arrives at a better fit.[12]

5.1.3 Constrained likelihood and entropy maximization

While we focused primarily on a variant of the method of moments, the formulas provided for superquantiles and bPOE can be used in other parametric procedures. For example, one could consider a constrained variant of the maximum likelihood or maximum entropy

[12] The fit from LS2 curves toward zero because it is the only fit with $k > 1$.

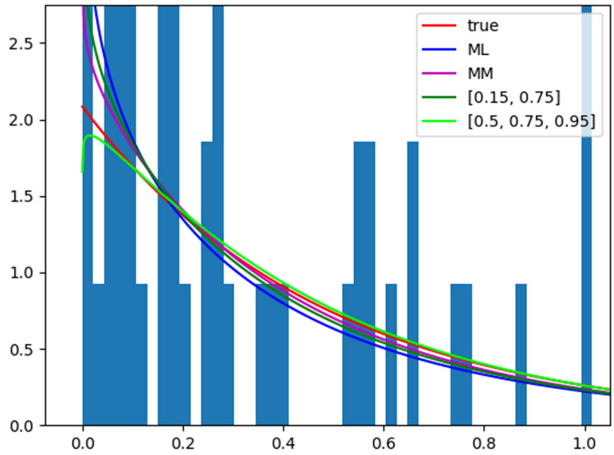

Fig. 4 Left side of distribution for fits on sample from Weibull with true $k = 1, \lambda = .5$

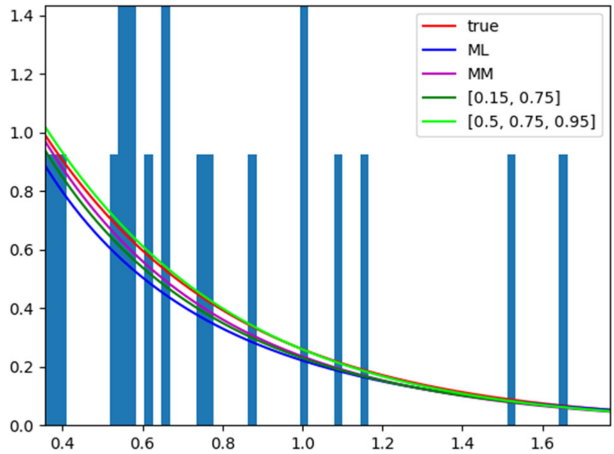

Fig. 5 Right side of distribution (zoomed in) for fits on sample from Weibull with true $k = 1, \lambda = .5$

method, where superquantile constraints are introduced. Letting $H(f_\Theta)$ denote the entropy of the random variable with density function f_Θ, and y_i denote an observation, constrained entropy maximization and maximum likelihood can be set up as follows:

$$ML: \max_{f_\Theta \in \mathcal{F}} \sum_i \log(f_\Theta(y_i)) \ , \quad ME: \max_{f_\Theta \in \mathcal{F}} H(f_\Theta)$$

where $\mathcal{F} = \{f_\Theta \mid \bar{q}_{\alpha_i}(X_{f_\Theta}) \leq \hat{\bar{q}}_{\alpha_i}(X) \ \ \forall i = 1, ..., k\}$.

While we leave full exploration of this framework for future work, this simple formulation illustrates another potential use for the provided superquantile and bPOE formulas within traditional parametric frameworks.

6 Conclusion

In this paper, we first derived closed-form formulas for the superquantile and bPOE, then utilized them within parametric portfolio optimization and density estimation problems. We were able to derive superquantile formulas for a variety of distributions, including ones with exponential tails (Exponential, Pareto/GPD, Laplace), symmetric distributions (Normal, Laplace, Logistic, Student-t), and asymmetric distributions with heavy tails (Lognormal, Weibull, LogLogistic, GEV). With bPOE formulas, while we had less success deriving truly closed-form solutions, we saw that it can still be calculated by solving a one-dimensional convex optimization problem or one-dimensional root finding problem.

We then utilized these formulas to develop two parametric procedures, one in portfolio optimization and one in density estimation. We first found that formulas for Normal, Laplace, Student-t, Logistic, and GEV are all distributions which yield tractable superquantile and bPOE portfolio optimization problems. Furthermore, we found that bPOE-optimal portfolios are more robust to changing distributional assumptions compared to superquantile-optimal portfolios. Specifically, bPOE optimal portfolios are optimal, simultaneously, for an entire class of distributions. Finally, we presented a variation on the method of moments where moments are replaced by superquantiles. This parametric procedure is made possible by our closed-form formulas and we illustrate its use in the case of heavy tailed asymmetric data, where additional emphasis on fitting the tail via superquantile conditions can be highly desirable. We find that this method makes it easy to direct the focus of the fitting procedure on tail samples.

Funding This research was supported by the Naval Postgraduate School's Research Initiation Program.

A Proofs

Proposition 1

Proof First, note that $q_\alpha(X) = \frac{-\ln(1-\alpha)}{\lambda}$ for exponential RVs with rate parameter λ. We then have,

$$
\begin{aligned}
\bar{q}_\alpha(X) &= \frac{1}{1-\alpha}\int_\alpha^1 q_p(X)dp \\
&= \frac{-1}{\lambda(1-\alpha)}\int_\alpha^1 \ln(1-p)dp \\
&= \frac{-1}{\lambda(1-\alpha)}\int_{1-\alpha}^0 -\ln(y)dy = \frac{-1}{\lambda(1-\alpha)}\int_0^{1-\alpha} \ln(y)dy.
\end{aligned}
$$

Since $\int \ln(y)dy = y\ln(y) - y + C$, we have

$$
\begin{aligned}
\bar{q}_\alpha(X) &= \frac{-1}{\lambda(1-\alpha)}\int_0^{1-\alpha} \ln(y)dy \\
&= \frac{-1}{\lambda(1-\alpha)}\left[(1-\alpha)\ln(1-\alpha) - (1-\alpha)\right] = \frac{-\ln(1-\alpha)+1}{\lambda}.
\end{aligned}
$$

We can then see that

$$
\begin{aligned}
\bar{p}_x(X) &= \{1-\alpha \mid \bar{q}_\alpha(X) = x\} \\
&= \{1-\alpha \mid \frac{-\ln(1-\alpha)+1}{\lambda} = x\}
\end{aligned}
$$

$$= \{1-\alpha | \ln(1-\alpha) = 1-\lambda x\}$$
$$= \{1-\alpha | e^{\ln(1-\alpha)} = e^{1-\lambda x}\} = \{1-\alpha | 1-\alpha = e^{1-\lambda x}\} = e^{1-\lambda x}.$$

□

Corollary 1

Proof We know that X, being exponential, has CDF given by $P(X \geq x) = 1 - e^{-\lambda x}$. From Proposition 1, we know that

$$\bar{p}_x(X) = e^{(1-\lambda x)} = e^{-\lambda(\frac{-1}{\lambda}+x)}.$$

Then, since $\mu = \frac{1}{\lambda}$, it follows that $\bar{p}_x(X) = e^{-\lambda(x-\mu)} = 1 - P(X \leq x-\mu) = P(X > x-\mu)$. The equality for CVaR follows easily from Proposition 1 since $q_\alpha(X) = \frac{-\ln(1-\alpha)}{\lambda}$.

□

Proposition 2

Proof First, note that the conditional distribution of a Pareto, conditioned on the event that the random value is larger than some γ, is simply another Pareto with parameters a, γ. This implies that $E[X|X > \gamma] = \frac{a\gamma}{a-1}$ if $a \geq 1$; otherwise the expectation is ∞. Also, $1 - F(\gamma) = \left(\frac{x_m}{\gamma}\right)^a$. Since,

$$E[X-\gamma]^+ = (E[X|X > \gamma] - \gamma)(1 - F(\gamma)),$$

we will have that,

$$E[X-\gamma]^+ = \left(\frac{a\gamma}{a-1} - \gamma\right)\left(\frac{x_m}{\gamma}\right)^a.$$

This gives us bPOE formula,

$$\bar{p}_x(X) = \min_{x_m \leq \gamma < x} \frac{\left(\frac{a\gamma}{a-1} - \gamma\right) x_m^a}{\gamma^a (x-\gamma)}$$
$$= \min_{x_m \leq \gamma < x} \frac{\left(\frac{a}{a-1} - 1\right) x_m^a}{\gamma^{a-1}(x-\gamma)} = \left(\max_{x_m \leq \gamma < x} \frac{\gamma^{a-1}(x-\gamma)(a-1)}{x_m^a}\right)^{-1}$$

Since $a > 1$, the maximization objective is concave over the range $\gamma \in (0, \infty)$ which contains the range $[x_m, x)$, so we just need to take the gradient of function $g(\gamma) = \frac{\gamma^{a-1}(x-\gamma)(a-1)}{x_m^a}$ and set it to zero to find the optimal γ as follows:

$$\frac{\partial g}{\partial \gamma} = \frac{x(a-1)^2\gamma^{a-2} - (a-1)a\gamma^{a-1}}{x_m^a} = 0 \implies x(a-1)^2\gamma^{a-2} = (a-1)a\gamma^{a-1}$$
$$\implies \frac{x(a-1)}{a} = \gamma$$

Plugging this value of γ into the objective of our bPOE formula yields,

$$\bar{p}_x(X) = \frac{\left(\frac{\frac{ax(a-1)}{a}}{a-1} - \frac{x(a-1)}{a}\right) x_m^a}{\left(\frac{x(a-1)}{a}\right)^a \left(x - \frac{x(a-1)}{a}\right)}$$
$$= \left(\frac{x_m a}{x(a-1)}\right)^a$$

CVaR is then equal to the value of x which solves the equation $1-\alpha = \bar{p}_x(X)$ or,

$$1 - \alpha = \left(\frac{x_m a}{x(a-1)}\right)^a,$$

which has solution,

$$\bar{q}_\alpha(X) = \frac{x_m a}{(1-\alpha)^{\frac{1}{a}}(a-1)}.$$

□

Corollary 2

Proof It follows by simply comparing the formulas from Proposition 1 and the CDF and quantile formulas for a Pareto RV. □

Proposition 3

Proof For these results, we rely on the fact that if $X \sim GPD(\mu, s, \xi)$, then $X - \gamma | X > \gamma \sim GPD(0, s + \xi(\gamma - \mu), \xi)$, meaning that the excess distribution of a GPD random variable is also GPD. Now, note also that if $\xi < 1$, then $E[X] = \mu + \frac{s}{1-\xi}$. This gives us,

$$E[X - \gamma | X > \gamma] = E[GPD(0, s + \xi(\gamma - \mu), \xi)] = \frac{s + \xi(\gamma - \mu)}{1 - \xi}$$

which further implies that,

$$\bar{q}_\alpha(X) = E[X - q_\alpha(X) | X > q_\alpha(X)] + q_\alpha(X)$$
$$= \frac{s + \xi(q_\alpha(X) - \mu)}{1 - \xi} + q_\alpha(X).$$

Plugging in the values of the quantile functions yields the final formulas. Using the formulas we just found for $\bar{q}_\alpha(X)$, it is straightforward to solve for $\bar{p}_x(X)$ which equals $1-\alpha$ such that α solves the equation $x = \bar{q}_\alpha(X)$. □

Proposition 4

Proof To get the superquantile, we begin with the integral representation:

$$\bar{q}_\alpha(X) = \frac{1}{1-\alpha} \int_\alpha^1 q_p(X) dp$$
$$= \frac{1}{1-\alpha} \int_\alpha^1 \mu - b\, sign(p - .5) \ln(1 - 2|p - .5|) dp$$

$$
\begin{aligned}
&= \mu - \frac{b}{1-\alpha} \int_{\alpha}^{1} sign(p - .5) \ln(1 - 2|p - .5|) dp \\
&= \mu - \frac{b}{1-\alpha} \left(\int_{\min\{\alpha,.5\}}^{.5} - \ln(2p) dp + \int_{\max\{\alpha,.5\}}^{1} \ln(2(1-p)) dp \right).
\end{aligned}
$$

To evaluate the integral, we use simple substitution as well as the identity $\int \ln(y)dy = y\ln(y) - y + C$. After simplifying, we see that with $\alpha < .5$ the integral evaluates to,

$$
\bar{q}_\alpha(X) = \mu + b\left(\frac{\alpha}{1-\alpha}\right)(1 - \ln(2\alpha)).
$$

Similarly, we find that with $\alpha \geq .5$ the integral evaluates to,

$$
\bar{q}_\alpha(X) = \mu + b\left(1 - \ln\left(2(1-\alpha)\right)\right).
$$

For bPOE, first assume that threshold $x \geq \mu + b$. Using our formula for CVaR, we see that $\bar{q}_{.5}(X) = \mu + b$. Thus, $x \geq \mu + b$ implies that $1 - \bar{p}_x(X) \geq .5$ implying that

$$
\begin{aligned}
\bar{p}_x(X) &= \{1 - \alpha | \bar{q}_\alpha(X) = x, \alpha \geq .5\} \\
&= \{1 - \alpha | \mu + b\left(1 - \ln\left(2(1-\alpha)\right)\right) = x\} \\
&= \frac{1}{2} e^{1 - \left(\frac{x-\mu}{b}\right)}.
\end{aligned}
$$

Assume contrarily that $x < \mu + b$. Since $\bar{q}_{.5}(X) = \mu + b$, we have that $1 - \bar{p}_x(X) < .5$ which implies that

$$
\begin{aligned}
\bar{p}_x(X) &= \{1 - \alpha | \bar{q}_\alpha(X) = x, \alpha < .5\} \\
&= \left\{1 - \alpha | \mu + b\left(\frac{\alpha}{1-\alpha}\right)(1 - \ln(2\alpha)) = x\right\}.
\end{aligned}
$$

Letting $z = \frac{x-\mu}{b}$, we must now find α which solves the equation $\left(\frac{\alpha}{1-\alpha}\right)(1 - \ln(2\alpha)) = z$. We do so as follows:

$$
\begin{aligned}
\left(\frac{\alpha}{1-\alpha}\right)(1 - \ln(2\alpha)) = z \implies & \frac{-z}{\alpha} = \frac{(\ln(2\alpha) - 1)}{1-\alpha} \\
\implies & e^{\frac{-z}{\alpha}} = e^{\frac{(\ln(2\alpha)-1)}{1-\alpha}} = \left(\frac{2\alpha}{e}\right)^{\frac{1}{1-\alpha}} \\
\implies & e^{\frac{-z(1-\alpha)}{\alpha}} = \left(\frac{2\alpha}{e}\right) \\
\implies & \frac{-z}{\alpha} e^{-z(\frac{1}{\alpha}-1)} = -2ze^{-1} \\
\implies & \frac{-z}{\alpha} e^{\frac{-z}{\alpha}} = -2ze^{-z-1} \\
\implies & \frac{-z}{\alpha} = \mathcal{W}(-2ze^{-z-1}).
\end{aligned}
$$

where the final step follows from the definition of the Lambert-$\mathcal{W}$ function which is given by the relation $xe^x = y \iff \mathcal{W}(y) = x$. Thus, $\frac{-z}{\alpha} = \mathcal{W}(-2ze^{-z-1}) \implies \bar{p}_x(X) = 1 - \alpha = 1 + \frac{z}{\mathcal{W}(-2e^{-z-1}z)}$. □

Proposition 5

Proof It is well known that if $X \sim \mathcal{N}(0, 1)$, then the conditional expectation is given by the inverse Mills Ratio, $E[X|X > \gamma] = \frac{f(\gamma)}{1-F(\gamma)}$. It follows then that $\bar{q}_\alpha(X) = E[X|X > q_\alpha(X)] = \frac{f(q_\alpha(X))}{1-F(q_\alpha(X))} = \frac{f(q_\alpha(X))}{1-\alpha}$. □

Proposition 6

Proof Note that for a standard normal random variable, the tail expectation beyond any threshold γ is given by the inverse Mills Ratio,

$$E[X|X > \gamma] = \frac{f(\gamma)}{1 - F(\gamma)}.$$

Note also that for any threshold γ and any random variable we have,

$$E[X - \gamma]^+ = (E[X|X > \gamma] - \gamma)(1 - F(\gamma)).$$

Using the Mills ratio gives us,

$$E[X - \gamma]^+ = \left(\frac{f(\gamma)}{1 - F(\gamma)} - \gamma\right)(1 - F(\gamma)) = f(\gamma) - \gamma(1 - F(\gamma)).$$

Plugging this result into the minimization formula for bPOE yields the final formula. □

Proposition 7

Proof This follows from the fact that $\bar{q}_\alpha(X) = E[X|X > q_\alpha(X)] = \frac{f(q_\alpha(X))}{1-F(q_\alpha(X))}$ and the optimization formula of bPOE given in the previous proposition for normally distributed variables. □

Proposition 8

Proof To derive the integral representation, simply plug in the formula for $E[X - \gamma]^+$, then utilize the definition of the PDF and CDF. The gradient calculation is a standard calculus exercise. □

Proposition 9

Proof We simply evaluate the integral of the quantile function as follows.

$$\begin{aligned}
\bar{q}_\alpha(X) &= \frac{1}{1 - \alpha} \int_\alpha^1 q_p(X) dp \\
&= \frac{1}{1 - \alpha} \int_\alpha^1 e^{\mu + s\sqrt{2}\text{erf}^{-1}(2p-1)} dp \\
&= \frac{e^\mu}{1 - \alpha} \int_\alpha^1 e^{s\sqrt{2}\text{erf}^{-1}(2p-1)} dp \\
&= \frac{e^\mu}{1 - \alpha} \left[-\frac{1}{2} e^{\frac{s^2}{2}} \left(1 + \text{erf}\left(\frac{s}{\sqrt{2}} - \text{erf}^{-1}(2p - 1)\right)\right)\right]_{p=\alpha}^1
\end{aligned}$$

$$
\begin{aligned}
&= \frac{e^{\mu}}{1-\alpha}\left[\frac{1}{2}e^{\frac{s^2}{2}} + \frac{1}{2}e^{\frac{s^2}{2}}\left(1 + \text{erf}\left(\frac{s}{\sqrt{2}} - \text{erf}^{-1}(2p-1)\right)\right)\right] \\
&= \frac{1}{2}e^{\mu+\frac{s^2}{2}}\frac{\left[1 + \text{erf}\left(\frac{s}{\sqrt{2}} - \text{erf}^{-1}(2\alpha-1)\right)\right]}{1-\alpha}.
\end{aligned}
$$

□

Proposition 10

Proof To obtain the superquantile, we have

$$
\begin{aligned}
\bar{q}_\alpha(X) &= \frac{1}{1-\alpha}\int_\alpha^1 q_p(X)dp \\
&= \frac{1}{1-\alpha}\int_\alpha^1 \mu + s\ln\left(\frac{\alpha}{1-\alpha}\right)dp \\
&= \mu + \frac{s}{1-\alpha}\int_\alpha^1 \ln(p) - \ln(1-p)dp \\
&= \mu + \frac{s}{1-\alpha}\left(\int_\alpha^1 \ln(p)dp + \int_\alpha^1 -\ln(1-p)dp\right)
\end{aligned}
$$

Utilizing simple substitution as well as the identity $\int \ln(y)dy = y\ln(y) - y + C$, we get

$$
\begin{aligned}
\bar{q}_\alpha(X) &= \mu + \frac{s}{1-\alpha}(-1 - \alpha\ln\alpha + \alpha - (1-\alpha)\ln(1-\alpha) + (1-\alpha)) \\
&= \mu + \frac{s}{1-\alpha}(-\alpha\ln\alpha - (1-\alpha)\ln(1-\alpha)) \\
&= \mu + \frac{s}{1-\alpha}H(\alpha).
\end{aligned}
$$

To get bPOE, we simply follow the bPOE definition, needing to find α which solves $\mu + \frac{s}{1-\alpha}H(\alpha) = x$. The transformed system arises from combining logarithms within the superquantile formula and applying exponential transformations. □

Proposition 11

Proof This follows from the fact that $E[X-\gamma]^+ = \int_\gamma^\infty (1-F(t))dt$. Evaluating this integral for $X \sim Logistic(\mu, s)$ yields, $E[X-\gamma]^+ = s\ln(1+e^{-(\frac{\gamma-\mu}{s})})$ which can then be plugged into the minimization formula for bPOE. The second part of the proposition follows from the fact that the gradient of the objective function w.r.t. γ is given by,

$$
\frac{s\ln(1+e^{-(\frac{\gamma-\mu}{s})})}{(x-\gamma)^2} - \frac{e^{-(\frac{\gamma-\mu}{s})}}{(x-\gamma)\left(1+e^{-(\frac{\gamma-\mu}{s})}\right)}.
$$

Setting this gradient to zero and simplifying yields the stated optimality condition. □

Proposition 12

Proof Since there is no closed-form expression for the quantile, we utilize the representation of the superquantile given by $\frac{1}{1-\alpha}\int_{q_\alpha(X)}^{\infty} tf(t)dt$. To evaluate this integral, we first take the derivative of the PDF, giving

$$\frac{df(x)}{dx} = \frac{-f(x)(x-\mu)(\nu+1)}{\nu s^2 + (x-\mu)}.$$

Rearranging yields,

$$xf(x)dx = \frac{-\nu s^2 df(x)}{(\nu+1)} - \frac{(x-\mu)^2 df(x)}{(\nu+1)} + \mu f(x)dx.$$

We can then integrate both sides,

$$\int xf(x)dx = \frac{-\nu s^2 f(x)}{(\nu+1)} - \frac{1}{(\nu+1)}\int (x-\mu)^2 df(x) + \mu F(x).$$

Integrating by parts gives us the following form of the middle term;

$$\int (x-\mu)^2 df(x) = (x-\mu)^2 f(x) - 2\int xf(x)dx + 2\mu F(x).$$

Then, finally, after substituting this new expression for the middle term and simplifying, we get

$$\int xf(x)dx = -\frac{(\nu s^2 + (x-\mu)^2)}{(\nu-1)} f(x) + \mu F(x).$$

Taking the definite integral yields,

$$\int_{q_\alpha(X)}^{\infty} xf(x)dx = \left(-\lim_{x\to\infty}\frac{(\nu s^2 + (x-\mu)^2)}{(\nu-1)} f(x) + \lim_{x\to\infty}\mu F(x)\right) - \left(-\frac{(\nu s^2 + (q_\alpha(X)-\mu)^2)}{(\nu-1)} f(q_\alpha(X)) + \mu F(q_\alpha(X))\right).$$

It is easy to see that the second limit goes to one and, after applying l'Hôpital's rule where necessary, that the first limit goes to zero. This leaves

$$\begin{aligned}\int_{q_\alpha(X)}^{\infty} xf(x)dx &= \mu - \left(-\frac{(\nu s^2 + (q_\alpha(X)-\mu)^2)}{(\nu-1)} f(q_\alpha(X)) + \mu F(q_\alpha(X))\right)\\ &= \mu(1-\alpha) + \left(\frac{\nu s^2 + (q_\alpha(X)-\mu)^2}{(\nu-1)}\right) f(q_\alpha(X))\\ &= \mu(1-\alpha) + s\left(\frac{\nu + T^{-1}(\alpha)^2}{(\nu-1)}\right)\tau(T^{-1}(\alpha)),\end{aligned}$$

where the final step comes from writing the non-standardized quantile $q_\alpha(X)$ and PDF $f(x)$ in their standardized form. Then, finally, dividing by $1-\alpha$ yields the formula

$$\bar{q}_\alpha(X) = \frac{1}{1-\alpha}\int_{q_\alpha(X)}^{\infty} xf(x)dx = \mu + s\left(\frac{\nu + T^{-1}(\alpha)^2}{(\nu-1)(1-\alpha)}\right)\tau(T^{-1}(\alpha)).$$

□

Proposition 13

Proof To calculate the superquantile, we utilize the integral representation (1), which is

$$\begin{aligned}\bar{q}_\alpha(X) &= \frac{1}{1-\alpha}\int_\alpha^1 q_p(X)dp \\ &= \frac{1}{1-\alpha}\int_\alpha^1 \lambda(-\ln(1-p))^{1/k}dp.\end{aligned}$$

To put this integral into the form of the upper incomplete gamma function, make the change of variable $y = -\ln(1-p)$. This gives $e^y = \frac{1}{1-p}$ and $dp = (1-p)dy = e^{-y}dp$ with new lower limit of integration $-\ln(1-\alpha)$ and upper limit of integration ∞. Applying to the integral yields

$$\begin{aligned}\bar{q}_\alpha(X) &= \frac{\lambda}{1-\alpha}\int_{-\ln(1-\alpha)}^\infty y^{1/k}e^{-y}dy \\ &= \frac{\lambda}{1-\alpha}\Gamma_U\left(1+\frac{1}{k}, -\ln(1-\alpha)\right).\end{aligned}$$

□

Proposition 14

Proof To calculate the superquantile, we utilize the integral representation as follows:

$$\begin{aligned}\bar{q}_\alpha(X) &= \frac{1}{1-\alpha}\int_\alpha^1 q_p(X)dp \\ &= \frac{1}{1-\alpha}\left(\int_0^1 q_p(X)dp - \int_0^\alpha q_p(X)dp\right) \\ &= \frac{1}{1-\alpha}\left(E[X] - \int_0^\alpha q_p(X)dp\right) \\ &= \frac{1}{1-\alpha}\left(E[X] - a\int_0^\alpha \left(\frac{p}{1-p}\right)^{\frac{1}{b}} dp\right).\end{aligned}$$

Now, note first that for $X \sim LogLogistic(a, b)$, we have $E[X] = a\frac{\pi}{b}csc\left(\frac{\pi}{b}\right)$. Next, for the incomplete beta function, letting $A_1 = \frac{1}{b} + 1$ and $A_2 = 1 - \frac{1}{b}$, we can see that

$$B_\alpha\left(\frac{1}{b}+1, 1-\frac{1}{b}\right) = \int_0^\alpha p^{\frac{1}{b}}(1-p)^{-\frac{1}{b}}dp.$$

Using these two facts, we have,

$$\begin{aligned}\bar{q}_\alpha(X) &= \frac{1}{1-\alpha}\left(E[X] - a\int_0^\alpha \left(\frac{p}{1-p}\right)^{\frac{1}{b}} dp\right) \\ &= \frac{1}{1-\alpha}\left(a\frac{\pi}{b}csc\left(\frac{\pi}{b}\right) - aB_\alpha\left(\frac{1}{b}+1, 1-\frac{1}{b}\right)\right) \\ &= \frac{a}{1-\alpha}\left(\frac{\pi}{b}csc\left(\frac{\pi}{b}\right) - B_\alpha\left(\frac{1}{b}+1, 1-\frac{1}{b}\right)\right).\end{aligned}$$

□

Proposition 15

Proof Assume we have $\xi = 0$. Then, we have

$$\begin{aligned}\bar{q}_\alpha(X) &= \frac{1}{1-\alpha}\int_\alpha^1 \mu - s\ln(-\ln(p))dp \\ &= \mu - \frac{s}{1-\alpha}\left(\int_0^1 \ln(-\ln(p))dp - \int_0^\alpha \ln(-\ln(p))dp\right) \\ &= \mu - \frac{s}{1-\alpha}\left(y - \int_0^\alpha \ln(-\ln(p))dp\right) \\ &= \mu - \frac{s}{1-\alpha}\left(y + \alpha\ln(-\ln(\alpha)) - \mathrm{li}(\alpha)\right).\end{aligned}$$

Assume now that $\xi \neq 0$. Then, we have that,

$$\begin{aligned}\bar{q}_\alpha(X) &= \frac{1}{1-\alpha}\int_\alpha^1 \mu + \frac{s}{\xi}\left((\ln(\frac{1}{p})^{-\xi} - 1\right)dp \\ &= \mu - \frac{s}{\xi(1-\alpha)}\int_\alpha^1 \left((\ln(\frac{1}{p})^{-\xi} - 1\right)dp \\ &= \mu + \frac{s}{\xi(1-\alpha)}\left[\Gamma_L(1-\xi, \ln(\frac{1}{\alpha})) - (1-\alpha)\right].\end{aligned}$$

□

References

Andreev, A., Kanto, A., & Malo, P. (2005). *On closed-form calculation of CVaR*. Helsinki School of Economics working paper W-389.

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

Davis, J. R., & Uryasev, S. (2016). Analysis of tropical storm damage using buffered probability of exceedance. *Natural Hazards*, *83*(1), 465–483.

Everitt, B. S. (2006). *The Cambridge dictionary of statistics*. Cambridge: Cambridge University Press.

Karian, Z. A., & Dudewicz, E. J. (1999). Fitting the generalized lambda distribution to data: A method based on percentiles. *Communications in Statistics: Simulation and Computation*, *28*(3), 793–819.

Landsman, Z. M., & Valdez, E. A. (2003). Tail conditional expectation for elliptical distributions. *North American Actuarial Journal*, *7*(4), 55–71.

Mafusalov, A., Shapiro, A., & Uryasev, S. (2018). Estimation and asymptotics for buffered probability of exceedance. *European Journal of Operational Research*, *270*(3), 826–836.

Mafusalov, A., & Uryasev, S. (2018). Buffered probability of exceedance: Mathematical properties and optimization. *SIAM Journal on Optimization*, *28*(2), 1077–1103.

Norton, M., Mafusalov, A., & Uryasev, S. (2017). Soft margin support vector classification as buffered probability minimization. *The Journal of Machine Learning Research*, *18*(1), 2285–2327.

Norton, M., & Uryasev, S. (2016). Maximization of AUC and buffered AUC in binary classification. *Mathematical Programming*, *174*(1–2), 575–612.

Rockafellar, R., & Royset, J. (2010). On buffered failure probability in design and optimization of structures. *Reliability Engineering & System Safety*, *95*, 499–510.

Rockafellar, R., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *The Journal of Risk*, *2*(3), 21–41.

Rockafellar, R. T., & Royset, J. O. (2014). Random variables, monotone relations, and convex analysis. *Mathematical Programming*, *148*(1–2), 297–331.

Rockafellar, R. T., & Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, *26*(7), 1443–1471.
Sgouropoulos, N., Yao, Q., & Yastremiz, C. (2015). Matching a distribution by matching quantiles estimation. *Journal of the American Statistical Association*, *110*(510), 742–759.
Shang, D., Kuzmenko, V., & Uryasev, S. (2018). Cash flow matching with risks controlled by buffered probability of exceedance and conditional value-at-risk. *Annals of Operations Research*, *260*(1–2), 501–514.
Uryasev, S. (2014). *Buffered probability of exceedance and buffered service level: Definitions and properties*. Department of Industrial and Systems Engineering, University of Florida, research report 3.

Affiliations

Matthew Norton[1] · Valentyn Khokhlov[2] · Stan Uryasev[3]

Valentyn Khokhlov
vkhokhlov.embals2016@london.edu

Stan Uryasev
uryasev@ufl.edu

1 Operations Research Department, Naval Postgraduate School, Monterey, CA, USA

2 CFA Society, Kyiv, Ukraine

3 Department of Industrial and Systems Engineering, Risk Management and Financial Engineering Laboratory, University of Florida, Gainesville, FL, USA