Journal of Statistical Planning and Inference 138 (2008) 3590–3604

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

Estimating conditional tail expectation with actuarial applications in view

Vytaras Brazauskas[a], Bruce L. Jones[b], Madan L. Puri[c], Ričardas Zitikis[b,*]

[a]Department of Mathematical Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA
[b]Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Ontario, Canada N6A 5B7
[c]Department of Mathematics, Indiana University, Bloomington, IN 47405, USA

ARTICLE INFO

Article history:
Received 7 April 2005
Accepted 4 November 2005
Available online 16 March 2008

MSC:
62F12
62F25
62G09
62G15
62G20
62G30
62P05
62P25

Keywords:
Conditional tail expectation
Quantile function
Consistency
Asymptotic normality
Confidence intervals
Vervaat process

ABSTRACT

We develop statistical inferential tools for estimating and comparing conditional tail expectation (CTE) functions, which are of considerable interest in actuarial science. In particular, we construct estimators for the CTE functions, develop the necessary asymptotic theory for the estimators, and then use the theory for constructing confidence intervals and bands for the functions. Both parametric and non-parametric approaches are explored. Simulation studies illustrate the performance of estimators in various situations. Results are obtained under minimal assumptions, and the general Vervaat process plays a crucial role in achieving these goals.

1. Introduction

In recent years, risk measures have become important tools in finance and actuarial science. One often wishes to place a value on a random financial quantity, such as an insurance liability, in such a way that this value allows for the variability of the quantity. This is accomplished using a risk measure, which can be defined as a mapping from the set of possible outcomes of the financial quantity to the real numbers. For simplicity in our presentation, we refer to the financial quantity as a "loss'' denoted by the random variable X. There are certain properties that a risk measure should have in order to be sensible. In fact, the term "coherent'' risk measure is reserved for risk measures that satisfy a specific set of properties (cf. Artzner, 1999). The conditional tail expectation (CTE) risk measure (also known as Tail-VaR or expected shortfall), which is the subject of this paper, is an example of a coherent risk measure (cf. Acerbi and Tasche, 2002; Artzner, 1999; Tasche, 2002; Wirch and Hardy, 1999).

The CTE risk measure is the conditional expectation of the loss random variable X given that X exceeds a specified quantile. The CTE has become increasingly popular due to its simplicity as well as its coherence. In fact, use of the CTE for determining liabilities

* Corresponding author. Tel.: +1 519 661 2111; fax: +1 519 661 3813.
E-mail addresses: vytaras@uwm.edu (V. Brazauskas), jones@stats.uwo.ca (B.L. Jones), puri@indiana.edu (M.L. Puri), zitikis@stats.uwo.ca (R. Zitikis).

0378-3758/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2005.11.011

associated with variable life insurance and annuity products with guarantees is required in Canada (cf. Canadian Institute of Actuaries, 2002) and recommended in the United States (cf. American Academy of Actuaries, 2002).

Application of the CTE in a multivariate context to elliptical distributions was considered by Landsman and Valdez (2003). Hardy and Wirch (2004) recognize the fact that there is normally no allowance for risk measures to evolve over time by introducing a dynamic risk measure called the iterated CTE. We do not consider such generalizations in this paper.

Manistre and Hancock (2005) present an empirical estimator of the CTE as well as an estimator of its variance. These authors deal with the situation in which the CTE is estimated based on simulated outcomes that are generated under a specified set of assumptions. In this case, the value of the CTE is, in principle, known but rather difficult to determine. In the present paper we address the case in which the CTE must be estimated from data.

Formally, the CTE risk measure, or function, can be defined as follows. Given a loss variable X (which is a real-valued random variable) with finite mean $\mathbf{E}[X]$, let F_X denote its distribution function. Next, let F_X^{-1} be the left-continuous inverse of F_X called the quantile function in the statistical literature. That is, for every $t \in [0,1]$, we have

$$F_X^{-1}(t) = \inf\{x : F_X(x) \geqslant t\}.$$

With the above notations, the CTE function is defined by

$$\mathrm{CTE}_X(t) = \mathbf{E}[X|X > F_X^{-1}(t)].$$

If F_X is a continuous function, then we have that $F_X(F_X^{-1}(t)) = t$ for every $t \in [0,1]$. Hence, $F_X^{-1}(t)$ is the point such that $t \times 100\%$ of losses are on or below the point. This can equivalently be reformulated by saying that there are $(1-t) \times 100\%$ of losses above the point $F_X^{-1}(t)$. To see how CTE_X looks in the continuous case, we first note that in this case the random variable $U = F_X(X)$ is uniformly distributed on the interval $[0,1]$. We therefore have the equality:

$$\mathrm{CTE}_X(t) = \frac{1}{\mathbf{P}\{U > t\}} \mathbf{E}[F_X^{-1}(U) I_{\{U>t\}}]. \tag{1.1}$$

Denote the right-hand side of (1.1) by $C_X(t)$. The following equality holds:

$$C_X(t) = \frac{1}{1-t} \int_t^1 F_X^{-1}(u)\, \mathrm{d}u. \tag{1.2}$$

The function C_X is well defined for all distributions. Equality (1.1) shows that the functions CTE_X and C_X coincide on the whole interval $[0,1]$ when the distribution function F_X is continuous. Note, however, that when F_X is not continuous, then the two curves CTE_X and C_X might not coincide (cf. Appendix at the end of the paper). Nevertheless, both CTE_X and C_X are called CTE in the literature.

Throughout the paper we concentrate on estimating C_X. We do not assume the continuity of F_X unless we need to do so. Specifically, when constructing consistent estimators of C_X, we do not assume the continuity of F_X, but in the case of confidence intervals we need and thus do assume it.

An empirical estimator of the function C_X is defined as follows. Let $X_1, \ldots, X_n$ be independent and identically distributed random variables. Denote the corresponding empirical distribution function by $F_{X,n}$ and the quantile function by $F_{X,n}^{-1}$. The empirical estimator of C_X is defined by

$$C_{X,n}(t) = \frac{1}{1-t} \int_t^1 F_{X,n}^{-1}(u)\, \mathrm{d}u.$$

In following sections we investigate point-wise and uniform over the interval $[0,1]$ consistency of $C_{X,n}$, as well as its asymptotic distribution. Based on these results, we then derive point-wise and simultaneous confidence intervals for the function C_X.

We conclude this section with brief discussions of two topics that are closely related to our paper. The first topic concerns the mean residual life (MRL) function, which is a somewhat *less* complicated mathematical object than the CTE. The second topic concerns the absolute concentration curve (ACC), which is a somewhat *more* complicated object than the CTE. Definitions and further details about these functions follow.

There might be situations when instead of the quantile $F_X^{-1}(t)$ one would wish to use a specified value x that does not depend on the population distribution. This problem is directly linked to the MRL function

$$\mathrm{MRL}_X(x) = \mathbf{E}[X - x|X > x].$$

There is an extensive literature on estimating the MRL function point-wise and simultaneously over its domain of definition, usually $[0,\infty)$ but possibly the whole real line. For example, we refer to Csörgö et al. (1986), Csörgö and Zitikis (1996) and Kochar et al. (2000) and references therein. We conclude this paragraph with the note that estimating MRL and CTE functions are two closely related but obviously different tasks.

A more general curve than the CTE function $\mathrm{CTE}_X(t)$ is the dual to the ACC. The latter is defined by

$$\mathrm{ACC}_{Y,X}(t)=\mathbf{E}[Y|X\leqslant F_X^{-1}(t)],$$

where Y is a random variable with finite mean $\mathbf{E}[Y]$ and dependent on X. Note that when $Y{=}X$, then $\mathrm{ACC}_{Y,X}(t)$ equals $\mathbf{E}[X|X\leqslant F_X^{-1}(t)]$. The latter function (in t) is the dual to the above defined CTE function $\mathbf{E}[X|X>F_X^{-1}(t)]$.

The ACC was introduced by Shalit and Yitzhaki (1994). For subsequent developments in the area we refer, for example, to Mayshar and Yitzhaki (1995), Chow (2001), Seiler (2001), Shalit and Yitzhaki (2003), Schechtman and Yitzhaki (2004), and references therein.

2. Consistency

Since $C_{X,n}(t)$ is a function of t, its consistency might, for example, mean convergence of $C_{X,n}(t)$ to $C_X(t)$ when $n\to\infty$ either at a given fixed point $t\in[0,1]$ or uniformly over all points $t\in[0,1]$. Furthermore, convergence can, for example, be almost surely or in probability. The latter mode of convergence is more natural from the applications point of view, but the former one is stronger and thus implies the latter one. Since our considerations would be almost identical for any of the two modes of convergence, we therefore restrict ourselves to convergence almost surely, that is, to strong consistency of $C_{X,n}(t)$. Naturally, we start with the simpler case of point-wise consistency.

Theorem 2.1. *Assuming that the first moment* $\mathbf{E}[X]$ *is finite, we have for every* $t\in[0,1]$ *that* $C_{X,n}(t)$ *converges to* $C_X(t)$ *almost surely and thus in probability. In other words,* $C_{X,n}(t)$ *is a strongly (and thus weakly) consistent estimator of* $C_X(t)$ *for every fixed* $t\in[0,1]$.

The proof of Theorem 2.1 as well as those of subsequent theorems are given in the Appendix at the end of the paper.

Uniform over the interval $[0,1]$ consistency does not hold, which is seen from the following considerations:

$$\begin{aligned}\sup_{0<t<1}|C_{X,n}(t)-C_X(t)| &\geqslant \sup_{1-1/n<t<1}|C_{X,n}(t)-C_X(t)|\\ &=\sup_{1-1/n<t<1}|X_{n:n}-C_X(t)|\\ &\geqslant -\sup_{1-1/n<t<1}X_{n:n}+\sup_{1-1/n<t<1}C_X(t)\end{aligned}\tag{2.1}$$

which is infinite since the order statistic is finite almost surely and the supremum is infinite for distributions whose right-end of the support is infinite. Thus, in this case the right-hand side of (2.1) does not converge to 0 (neither almost surely nor in probability), and thus, in turn, the left-hand side does not converge either. We therefore conclude that $C_{X,n}$ is not a uniformly consistent estimator of C_X over the whole interval $[0,1]$.

The reason why the uniform consistency fails is that the denominators $1-t$ in the definitions of the theoretical and empirical CTE functions emphasize the corresponding numerators so much when t approaches 1 that the difference between $C_{X,n}(t)$ and $C_X(t)$ becomes large. This, in turn, suggests a way to fix the problem. Namely, let $q:(0,1)\to[0,\infty)$ be a function such that $q(t)\to 0$ when $t\to 1$. We expect that if we multiply $|C_{X,n}(t)-C_X(t)|$ by $q(t)$, then the resulting quantity will converge to 0 uniformly over the interval $[0,1]$. The following theorem describes the class of weight functions q that can be used for the purpose. Certainly, the class depends on the upper tail of the distribution function F_X, which we control using a moment condition.

Theorem 2.2. *Let the rth moment* $\mathbf{E}[|X|^r]$ *be finite for some* $r\geqslant 1$, *and let the weight function* $q:(0,1)\to[0,\infty)$ *be such that*

$$\sup_{0<t<1}\frac{q(t)}{(1-t)^{1/r}}<\infty.\tag{2.2}$$

Then

$$\sup_{0<t<1}q(t)|C_{X,n}(t)-C_X(t)|\to_{\text{a.s.}}0.\tag{2.3}$$

One might be interested to know whether assumption (2.2) can be relaxed. Note in this regard that the assumption allows us to use the weight function $q(t)=(1-t)^{1/r-\varepsilon}$ with $\varepsilon=0$, but not with $\varepsilon>0$ no matter how small it is. The following theorem shows that these notes are also applicable in the case of statement (2.3) itself, thus proving the optimality of assumptions in Theorem 2.2.

Theorem 2.3. *For any (small)* $\varepsilon>0$, *the supremum in* (2.3) *with the weight function* $q(t)=(1-t)^{1/r-\varepsilon}$ *does not converge to* 0 *in probability and thus, in turn, almost surely*.

3. Confidence intervals and bands

In the previous section we discussed when it is appropriate to use $C_{X,n}(t)$ for estimating $C_X(t)$ either at any fixed points $t \in [0, 1]$ or uniformly over all t. The next step is to construct point-wise and simultaneous confidence intervals for $C_X(t)$. We call them confidence intervals and confidence bands, respectively. The following two paragraphs justify this research in the actuarial context.

We start with the note that in many cases one may be interested in a confidence interval for the CTE at a given value of t. For example, the Canadian Institute of Actuaries Task Force on Segregated Fund Investment Guarantees (cf. Canadian Institute of Actuaries, 2002) recommends using $t = 0.90$ in determining the balance sheet liability for variable products with investment guarantees.

In other instances, one may wish to explore the impact of changing the value of t, say from t_1 to t_2 for certain $0 < t_1 < t_2 < 1$, on the resulting CTE estimates. One might then estimate the CTE for values of t in the interval $[t_1, t_2]$. To properly reflect uncertainty about the estimates, it is appropriate to construct a confidence band for the values of the CTE function.

We start with (point-wise) confidence intervals, whose construction is based on the following asymptotic result.

Theorem 3.1. *Assume that the second moment* $\mathbf{E}[X^2]$ *is finite. Let* $t \in [0, 1]$ *be fixed, and let the distribution function* F_X *be continuous at the point* $F_X^{-1}(t)$. *Then*

$$\sqrt{n}(C_{X,n}(t) - C_X(t)) \to_d \mathcal{N}(0, \sigma_X^2(t)), \tag{3.1}$$

where $\mathcal{N}(0, \sigma_X^2(t))$ *denotes a centered normal random variable with the variance*

$$\sigma_X^2(t) = \frac{1}{(1-t)^2} \int_{F_X^{-1}(t)}^{\infty} \int_{F_X^{-1}(t)}^{\infty} (F_X(x \wedge y) - F_X(x)F_X(y))\, dx\, dy. \tag{3.2}$$

In particular, statement (3.1) *holds for any fixed* $t \in [0, 1]$ *if the distribution function* F_X *is continuous everywhere on the real line.*

Quantities like $\sigma_X^2(t)$ in (3.2) appear naturally and frequently when investigating L- and R-statistics. For results and discussions on the subject, we refer, for example, to Puri and Sen (1971), Serfling (1980) and Shorack and Wellner (1986).

Using statement (3.1), we derive the following $(1-\alpha)100\%$ level asymptotic confidence interval for the CTE $C_X(t)$:

$$C_{X,n}(t) \pm \frac{z_{\alpha/2}\sigma_X(t)}{\sqrt{n}},$$

where $z_{\alpha/2}$ is the $(1-\alpha/2) \times 100\%$ percentile of the standard normal distribution. The standard deviation $\sigma_X(t)$ is unknown and thus needs to be estimated empirically. This we do next.

Under the assumptions of Theorem 3.1, we can easily check that

$$\sigma_{X,n}^2(t) = \frac{1}{(1-t)^2} \int_{F_{X,n}^{-1}(t)}^{\infty} \int_{F_{X,n}^{-1}(t)}^{\infty} (F_{X,n}(x \wedge y) - F_{X,n}(x)F_{X,n}(y))\, dx\, dy \tag{3.3}$$

is a consistent estimator of $\sigma_X^2(t)$. Hence, we have the following $(1-\alpha)100\%$ level asymptotic confidence interval for the CTE $C_X(t)$:

$$C_{X,n}(t) \pm \frac{z_{\alpha/2}\sigma_{X,n}(t)}{\sqrt{n}}.$$

We note in passing that formula (3.3) can be rewritten in a more computationally convenient form (for details, see Eq. (A.27) in Appendix):

$$\sigma_{X,n}^2(t) = \frac{1}{(1-t)^2} \sum_{nt \leqslant j \leqslant n-1} \sum_{nt \leqslant k \leqslant n-1} \left(\frac{j \wedge k}{n} - \frac{j}{n}\frac{k}{n} \right) (X_{j+1:n} - X_{j:n})(X_{k+1:n} - X_{k:n}). \tag{3.4}$$

In view of statement (3.1), we can also use a bootstrap approximation to construct confidence intervals. To this end, we sample with replacement from $X_1, \ldots, X_n$ and obtain a new sample $X_1^*, \ldots, X_n^*$. Using the latter sample, we then obtain the corresponding CTE, which we denote by $C_{X,n}^*(t)$. Hence, we obtain a value of the quantity (note the absolute value)

$$\sqrt{n}|C_{X,n}^*(t) - C_{X,n}(t)|. \tag{3.5}$$

We repeat the above simulation procedure M times and obtain M values of quantity (3.5). Then we define x^* as the smallest number x such that the proportion of those M values of quantity (3.5) that are at or below x is at least $1-\alpha$. We arrive at the following asymptotic $(1-\alpha)100\%$ level confidence interval for the CTE $C_X(t)$:

$$C_{X,n}(t) \pm \frac{x^*}{\sqrt{n}}.$$

Summarizing the discussion above, we have constructed point-wise confidence intervals for the CTE $C_X(t)$ at any fixed point $t \in [0, 1]$. These confidence intervals, however, do not imply simultaneous confidence intervals for the function C_X over the interval $[0, 1]$. A stronger asymptotic result than that in Theorem 3.1 is needed, and we formulate it as Theorem 3.2.

Before formulating the theorem, we find it instructive to give a few introductory notes on the assumptions to be imposed. First, recall that in Theorem 3.1 we required continuity of F_X at a fixed point. Due to its uniform character, Theorem 3.2 will require continuity of F_X at every point. Second, as we can guess from the above result concerning uniform consistency, we shall need to employ weight functions in the theorem that follows. Third, certain moment assumptions will be imposed to control the behavior of the aforementioned weight functions, depending on the tail behavior of the underlying distribution function F_X. Now we are ready to formulate our next theorem.

Theorem 3.2. *Let the distribution function F_X be continuous. Assume that the moment $\mathbf{E}[|X|^{2+\varepsilon}]$ is finite for some $\varepsilon > 0$, and let the moment $\mathbf{E}[X_+^r]$ be also finite for some $r > 2$, where $X_+ = \max(X, 0)$. If the weight function $q : (0, 1) \to [0, \infty)$ is such that, for some $v > 0$,*

$$\sup_{0<t<1} \frac{q(t)}{(1-t)^{1/r+1/2+v}} < \infty, \tag{3.6}$$

then we have that

$$\sqrt{n} \sup_{0<t<1} q(t)|C_{X,n}(t) - C_X(t)| \to_{\mathrm{d}} \sup_{0<t<1} \frac{q(t)}{1-t} \left| \int_{F_X^{-1}(t)}^{\infty} \mathscr{B}(F_X(x))\, \mathrm{d}x \right|, \tag{3.7}$$

where $\mathscr{B}$ is the Brownian bridge on the interval $[0, 1]$.

From (3.7) we obtain the following $(1-\alpha)100\%$ level asymptotic confidence band for the CTE function C_X:

$$C_{X,n}(t) \pm \frac{s_\alpha}{q(t)\sqrt{n}} \quad \text{for all } t \in [0,1] \text{ such that } q(t) > 0,$$

where s_α is the $(1-\alpha)$th quantile of the distribution function of the limiting random variable on the right-hand side of (3.7). Obviously, s_α depends on both the distribution function F_X and the weight function q.

Since s_α depends on the (unknown) distribution function F_X, it needs to be estimated empirically. This can be done using a bootstrap approximation analogous to the one described in the two paragraphs below Eq. (3.4) but now with

$$\sqrt{n} \sup_{0<t<1} q(t)|C^*_{X,n}(t) - C_{X,n}(t)| \tag{3.8}$$

instead of quantity (3.5). We obtain the asymptotic $(1-\alpha)100\%$ level confidence band for the CTE function C_X:

$$C_{X,n}(t) \pm \frac{x^*}{q(t)\sqrt{n}} \quad \text{for all } t \in [0,1] \text{ such that } q(t) > 0.$$

Note that if a confidence band for C_X is desired, for example, over an interval $[t_1, t_2] \subseteq [0, 1]$, then it is natural to choose q so that $q(t) = 0$ for all t's outside the interval $[t_1, t_2]$.

4. Parametric confidence intervals

In this section we derive parametric confidence intervals for the CTE function at any given $t \in [0, 1]$. Just like non-parametric confidence intervals above, the parametric ones constructed in this section are also asymptotic. They are constructed using asymptotic theory for the maximum likelihood estimators (MLE) and also the delta method.

Let us focus on *interval* estimation of the expected maximum loss $C_X(t)$ in the $(1-t)100\%$ worst cases, based on a sample $X_1, \ldots, X_n$ having distribution function F_X. First we note that empirical (i.e., non-parametric) confidence intervals for $C_X(t)$ are readily available (see Section 3). Below we shall construct *parametric* confidence intervals for $C_X(t)$ for three parametric distributions:

- *Exponential* with the cdf given by $F_{X_1}(x) = 1 - \mathrm{e}^{-(x-x_0)/\theta}$, for $x > x_0$ and $\theta > 0$.
- *Pareto* with the cdf given by $F_{X_2}(x) = 1 - (x_0/x)^\gamma$, for $x > x_0$ and $\gamma > 0$.
- *Lognormal* with the cdf given by $F_{X_3}(x) = \Phi(\log(x - x_0) - \mu)$, for $x > x_0$ and $-\infty < \mu < \infty$, where $\Phi(\cdot)$ denotes the cdf of the standard normal distribution.

Application of these distributions covers a wide spectrum of areas ranging from actuarial science, economics, finance to telecommunications and extreme value theory. A more detailed discussion of these applications along with various techniques for estimation of parameters of these families can be found in Brazauskas and Serfling (2000), Serfling (2002) and Brazauskas and Kaiser (2004). Practical aspects of fitting Pareto models to real loss data are investigated in Brazauskas and Serfling (2003).

The parameter x_0 in the above distributions can be interpreted as a deductible or a retention level and, thus, assumed to be known. (Note that, due to x_0, the distribution functions F_{X_1}, F_{X_2}, and F_{X_3} have the same support.) The parameters θ, γ, and μ are unknown but can be estimated from the data using, for example, the maximum likelihood approach. Straightforward derivations yield the following formulas for the MLE's of the parameters θ, γ, and μ, respectively:

$$\widehat{\theta} = \frac{1}{n}\sum_{i=1}^{n}(X_i - x_0), \quad \widehat{\gamma} = \left(\frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{X_i}{x_0}\right)\right)^{-1}, \quad \widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\log(X_i - x_0). \tag{4.1}$$

It is well known from asymptotic theory for the maximum likelihood procedures that the MLE's $\widehat{\theta}$, $\widehat{\gamma}$, and $\widehat{\mu}$ are each asymptotically normal with, respectively, the means θ, γ, and μ, and the variances θ^2/n, γ^2/n, and $1/n$.

The CTE measures for the exponential, Pareto, and lognormal distributions are found by computing integral (1.2) with F_{X_1}, F_{X_2}, and F_{X_3}, respectively. This yields the formulas, respectively:

$$C_{X_1}(t) = x_0 - \theta(\log(1-t) - 1), \quad C_{X_2}(t) = \frac{x_0\gamma}{\gamma - 1}(1-t)^{-1/\gamma},$$
$$C_{X_3}(t) = x_0 + \frac{1}{1-t}\, e^{\mu+0.5}\Phi(1 - \Phi^{-1}(t)).$$

The corresponding empirical estimators of CTE's are then found by replacing unknown parameters with their respective MLE's. That is, we have

$$\widehat{C}_{X_1}(t) = x_0 - \widehat{\theta}(\log(1-t) - 1), \quad \widehat{C}_{X_2}(t) = \frac{x_0\widehat{\gamma}}{\widehat{\gamma} - 1}(1-t)^{-1/\widehat{\gamma}},$$
$$\widehat{C}_{X_3}(t) = x_0 + \frac{1}{1-t}\, e^{\widehat{\mu}+0.5}\Phi(1 - \Phi^{-1}(t)).$$

In view of the discussion above, in order to get parametric confidence intervals for the CTE measures, one has to apply the classical theory of transformations for asymptotically normal statistics (see, e.g., Serfling, 1980, Chapter 3). Application of this approach—also known as the delta method—implies the following statement, for $k = 1, 2, 3$,

$$\sqrt{n}(\widehat{C}_{X_k}(t) - C_{X_k}(t)) \to_d \mathcal{N}(0, Q^2_{X_k}), \tag{4.2}$$

where the asymptotic variances are given by formulas:

$$Q^2_{X_1} = (1 - \log(1-t))^2\theta^2,$$
$$Q^2_{X_2} = \frac{x_0^2(1-t)^{-2/\gamma}}{(\gamma-1)^2}\left(\frac{\log(1-t)}{\gamma} - \frac{1}{\gamma - 1}\right)^2\gamma^2,$$
$$Q^2_{X_3} = \left(\frac{1}{1-t}e^{\mu+0.5}\Phi(1 - \Phi^{-1}(t))\right)^2.$$

We can now formulate the $100(1-\alpha)\%$ parametric confidence intervals for the CTE measures $C_{X_1}(t)$, $C_{X_2}(t)$, and $C_{X_3}(t)$, respectively:

$$x_0 - \widehat{\theta}(\log(1-t) - 1)(1 \pm z_{\alpha/2}\sqrt{1/n}),$$
$$\frac{x_0\widehat{\gamma}}{\widehat{\gamma}-1}(1-t)^{-1/\widehat{\gamma}}(1 \pm z_{\alpha/2}(\log(1-t)/\widehat{\gamma} - 1/(\widehat{\gamma} - 1))\sqrt{1/n}),$$
$$x_0 + \frac{1}{1-t}e^{\widehat{\mu}+0.5}\Phi(1 - \Phi^{-1}(t))(1 \pm z_{\alpha/2}\sqrt{1/n}),$$

where $z_{\alpha/2}$ is the $(1 - \alpha/2)$th quantile of the standard normal distribution.

5. A simulation study

In this section we examine—via Monte Carlo simulations—finite-sample performance of the proposed (asymptotic) confidence intervals for the CTE measure, with particular emphasis on comparisons between empirical (i.e., non-parametric) and parametric confidence intervals. Data are generated by three similar shape parametric families that have equal CTE measures and their "CTE-riskiness'' is either "mild'' or "severe''.

We used the following design for the Monte Carlo simulation study. Ten thousand samples of size n were generated from a distribution F_X. For each sample, a $(1-\alpha)100\%$ confidence interval for the CTE measure was constructed using the empirical and parametric approaches. Then, based on these 10,000 intervals for each approach, the average length of the interval and the proportion of times the interval covered the true value of the CTE measure was evaluated. This procedure was repeated 10 times, and the means and standard errors of the average length and the proportion of coverage were recorded. The study was performed for the following (specific) choices of simulation parameters:

- *Sample sizes*: $n=20, 100, 250$.
- *Confidence level*: $1-\alpha=0.95$.
- *Threshold levels*: $t=0.95, 0.80$.
- *Target quantities*: We choose to work with the exponential (F_{X_1}), Pareto (F_{X_2}), and lognormal (F_{X_3}) distribution functions (cf. the previous section for formulas). The parameters x_0, γ, θ, and μ are chosen so that all three distributions are equally risky, i.e., they have identical CTE values, that is, $C_{X_1}(t)=C_{X_2}(t)=C_{X_3}(t)$, which can equivalently be written as the equalities

$$x_0-\theta(\log(1-t)-1)=\frac{x_0\gamma}{\gamma-1}(1-t)^{-1/\gamma}$$
$$=x_0+\frac{1}{1-t}e^{\mu+0.5}\Phi(1-\Phi^{-1}(t)).$$

Since x_0 is known, without loss of generality we can and thus do assume that $x_0=1$. Next, taking into consideration that the three CTE values are equal by our choice, we select values of the parameters θ, γ, and μ in such a way that they represent the following two scenarios:

- *Mild* riskiness:

$$\gamma=10,$$
$$\theta=\frac{1-(10/9)(1-t)^{-1/10}}{\log(1-t)-1},$$
$$\mu=\log\left(\frac{(1-t)((10/9)(1-t)^{-1/10}-1)}{\Phi(1-\Phi^{-1}(t))}\right)-0.5.$$

Hence, the values

$$C_{X_1}(0.95)=C_{X_2}(0.95)=C_{X_3}(0.95)=1.499,$$
$$C_{X_1}(0.80)=C_{X_2}(0.80)=C_{X_3}(0.80)=1.305.$$

- *Severe* riskiness:

$$\gamma=3,$$
$$\theta=\frac{1-1.5(1-t)^{-1/3}}{\log(1-t)-1},$$
$$\mu=\log\left(\frac{(1-t)(1.5(1-t)^{-1/3}-1)}{\Phi(1-\Phi^{-1}(t))}\right)-0.5.$$

Hence, the values

$$C_{X_1}(0.95)=C_{X_2}(0.95)=C_{X_3}(0.95)=4.072,$$
$$C_{X_1}(0.80)=C_{X_2}(0.80)=C_{X_3}(0.80)=2.565.$$

Our simulation results are summarized in Table 1. The following information is presented there: length and proportion of coverage of 95% empirical and parametric confidence intervals for $C_X(t)$, for selected t and sample size n, when losses are generated from exponential, Pareto, and lognormal distributions. Parameters of the distributions are chosen so that, under a fixed scenario of riskiness, mild or severe, all three families are equally risky according to the $C_X(t)$ criterion. Standard errors for the entries are presented in parentheses. A discussion of our findings presented in Table 1 is given in the next paragraph.

Coverage proportions of the empirical intervals are quite low in small samples (0.52–0.61 for $n=20$) for $t=0.95$ and for both scenarios of riskiness of underlying distributions. These proportions increase as sample size gets larger (0.76–0.85 for $n=100$; 0.82–0.89 for $n=250$) or when t is less extreme (0.74–0.83 for $t=0.80$ and $n=20$). Overall, coverages of the empirical intervals get reasonably close to the intended 95% confidence level for $n\geqslant 250$. Parametric intervals, on the other hand, perform very well

Table 1

t	Method of estimation	Sample size	Exponential		Pareto		Lognormal	
			Length	Coverage	Length	Coverage	Length	Coverage
Riskiness of distributions: mild [true values: $C_X(0.80) = 1.305$, $C_X(0.95) = 1.499$]								
0.95	Empirical	20	0.48 (0.004)	0.61 (0.004)	0.55 (0.007)	0.59 (0.004)	0.69 (0.009)	0.54 (0.006)
		100	0.29 (0.001)	0.85 (0.003)	0.35 (0.001)	0.83 (0.003)	0.48 (0.005)	0.79 (0.003)
		250	0.18 (0.000)	0.89 (0.003)	0.23 (0.001)	0.87 (0.002)	0.32 (0.002)	0.85 (0.002)
	Parametric	20	0.44 (0.001)	0.93 (0.003)	0.55 (0.001)	0.92 (0.003)	0.45 (0.001)	0.94 (0.003)
		100	0.20 (0.000)	0.94 (0.002)	0.24 (0.000)	0.94 (0.003)	0.20 (0.000)	0.95 (0.002)
		250	0.12 (0.000)	0.95 (0.002)	0.15 (0.000)	0.95 (0.002)	0.12 (0.000)	0.95 (0.003)
0.80	Empirical	20	0.29 (0.002)	0.82 (0.004)	0.33 (0.002)	0.81 (0.003)	0.40 (0.004)	0.76 (0.004)
		100	0.13 (0.000)	0.92 (0.003)	0.16 (0.000)	0.90 (0.004)	0.21 (0.001)	0.88 (0.003)
		250	0.09 (0.000)	0.93 (0.002)	0.10 (0.000)	0.93 (0.003)	0.14 (0.001)	0.91 (0.004)
	Parametric	20	0.27 (0.001)	0.92 (0.002)	0.32 (0.001)	0.92 (0.002)	0.27 (0.001)	0.94 (0.002)
		100	0.12 (0.000)	0.94 (0.002)	0.14 (0.000)	0.94 (0.002)	0.12 (0.000)	0.95 (0.001)
		250	0.08 (0.000)	0.95 (0.002)	0.09 (0.000)	0.95 (0.002)	0.08 (0.000)	0.95 (0.002)
Riskiness of distributions: severe [true values: $C_X(0.80) = 2.565$, $C_X(0.95) = 4.072$]								
0.95	Empirical	20	2.93 (0.022)	0.61 (0.004)	4.66 (0.067)	0.52 (0.004)	4.23 (0.057)	0.54 (0.005)
		100	1.78 (0.007)	0.84 (0.003)	3.46 (0.040)	0.76 (0.004)	2.99 (0.021)	0.79 (0.002)
		250	1.14 (0.002)	0.89 (0.003)	2.35 (0.016)	0.82 (0.003)	1.98 (0.012)	0.84 (0.003)
	Parametric	20	2.69 (0.007)	0.93 (0.001)	6.40 (0.034)	0.89 (0.004)	2.76 (0.008)	0.94 (0.002)
		100	1.20 (0.001)	0.95 (0.002)	2.47 (0.008)	0.94 (0.002)	1.21 (0.001)	0.95 (0.003)
		250	0.76 (0.000)	0.95 (0.002)	1.53 (0.003)	0.94 (0.002)	0.76 (0.001)	0.95 (0.003)
0.80	Empirical	20	1.46 (0.010)	0.83 (0.005)	2.30 (0.020)	0.74 (0.004)	2.05 (0.018)	0.76 (0.003)
		100	0.69 (0.002)	0.92 (0.003)	1.21 (0.008)	0.86 (0.004)	1.05 (0.004)	0.88 (0.003)
		250	0.44 (0.001)	0.93 (0.003)	0.82 (0.005)	0.89 (0.004)	0.69 (0.002)	0.91 (0.002)
	Parametric	20	1.37 (0.003)	0.92 (0.002)	2.65 (0.008)	0.90 (0.004)	1.41 (0.003)	0.94 (0.002)
		100	0.61 (0.001)	0.95 (0.003)	1.07 (0.002)	0.94 (0.002)	0.62 (0.001)	0.95 (0.002)
		250	0.39 (0.000)	0.95 (0.002)	0.67 (0.001)	0.94 (0.002)	0.39 (0.000)	0.95 (0.003)

with respect to the coverage criterion having coverage proportions of at least 0.89 (=0.89 for severe riskiness, at Pareto, $t = 0.95$, $n = 20$) for both scenarios of riskiness and for all sample sizes that we considered. Further, except for a few cases (e.g., severe riskiness, at Pareto, $n = 20$), they also dominate empirical counterparts with respect to the length criterion. Superior overall performance of parametric intervals should not come as a surprise because they have an advantage of "knowing" the underlying distribution, i.e., they are equipped with additional information. Empirical procedures, however, are more flexible and perform similarly under all distributional scenarios. Finally, for both methods of estimation, the impact of specific distribution or type of riskiness on coverage proportions is minimal but is significant on the interval length, sometimes even dramatic (e.g., for $t = 0.95$, parametric approach, $n = 20$, at Pareto, the interval length increases more than 10 times at severe riskiness (6.40) compared to mild riskiness (0.55), whereas the true value of $C_X(0.95)$ changes from 4.072 to 1.499).

6. Comparing CTE's when samples are independent

There are situations when it is of interest to compare the CTE functions C_X and C_Y that correspond to loss variables X and Y, respectively. Here is an example.

In an insurance context, companies may be interested in comparing the CTE risk measure values corresponding to different policy characteristics. For example, an automobile insurer may offer policyholders a choice of deductible. It is well known that losses for those who choose a less expensive high deductible policy do not have the same distribution (before applying the deductible) as losses for those who choose a more expensive low deductible policy.

Just like in previous sections, we estimate C_X using $C_{X,n}$ based on independent and identically distributed random variables $X_1, \ldots, X_n$, and estimate C_Y using $C_{Y,m}$ based on independent and identically distributed random variables $Y_1, \ldots, Y_m$. In practice, there might be many possible scenarios of interest concerning dependence between X's and Y's. In this section we consider the case when the two samples are independent. An example where such a situation arises is as follows.

The insurer may wish to compare the CTE for policies with different deductibles. This can be done by collecting data on claims arising from policies with the different deductibles. These samples can be assumed to be independent.

Testing hypotheses about the equality $C_X(t) = C_Y(t)$ or, say, dominance $C_X(t) \leqslant C_Y(t)$ either at a fixed point $t \in [0, 1]$ or over a region of the interval $[0, 1]$ are, naturally, based on the asymptotic behavior of the appropriately normalized difference $C_{X,n}(t) - C_{Y,m}(t)$. In the case when the X's and Y's are independent, the asymptotic behavior of the aforementioned difference can be obtained using the corresponding one-sample results already established above. Namely, assuming that the ratio $n/(n+m)$ converges to a constant $\eta \in [0, 1]$ when both n and m tend to infinity, we have under the conditions of Theorem 3.1 for both F_X and F_Y that the following statement holds:

$$\Theta_n(t) = \sqrt{\frac{nm}{n+m}}((C_{X,n}(t) - C_{Y,m}(t)) - (C_X(t) - C_Y(t))) \to_{\mathrm{d}} \mathcal{N}(0, (1-\eta)\sigma_X^2(t) + \eta\sigma_Y^2(t)). \tag{6.1}$$

Using statement (6.1), we can construct confidence intervals for the difference $C_X(t) - C_Y(t)$, test hypotheses about whether, say, $C_X(t)$ is equal to or above/below $C_Y(t)$. For example, we write the following $(1-\alpha)100\%$ level asymptotic confidence interval for the difference $C_X(t) - C_Y(t)$:

$$C_{X,n}(t) - C_{Y,m}(t) \pm z_{\alpha/2}\sqrt{\frac{\sigma^2_{X,n}(t)}{n} + \frac{\sigma^2_{Y,m}(t)}{m}}.$$

Instead of using the asymptotic variance and their empirical estimators, we can use bootstrap methods instead. As an example, suppose we are interested in having (two-sided) confidence intervals for the difference $C_X(t) - C_Y(t)$. We generate independent random variables $X_1^*, \dots, X_n^*$ from the distribution $F_{X,n}$, as well as independent random variables $Y_1^*, \dots, Y_m^*$ from the distribution $F_{Y,m}$. From these new samples we then calculate $C^*_{X,n}(t)$ and $C^*_{Y,m}(t)$, respectively. We obtain a numerical value of the quantity

$$\sqrt{\frac{nm}{n+m}}|(C^*_{X,n}(t) - C^*_{Y,m}(t)) - (C_{X,n}(t) - C_{Y,m}(t))|. \tag{6.2}$$

Next we proceed as in the two paragraphs below Eq. (3.4) but now with quantity (6.2) instead of (3.5). We obtain the asymptotic $(1-\alpha)100\%$ level confidence interval for the difference $C_X(t) - C_Y(t)$:

$$C_{X,n}(t) - C_{Y,m}(t) \pm x^*\sqrt{\frac{n+m}{nm}}.$$

The above discussion concerned the case when $t \in [0,1]$ was fixed. However, similar considerations are also applicable for comparing CTE functions C_X and C_Y over regions of the interval $[0,1]$ as well. The only major difference is that now we need to use considerations in the proof of Theorem 3.2 to obtain, for example, convergence in distribution of quantities such as $\sup_t q(t)\Theta_n(t)$, $\sup_t q(t)|\Theta_n(t)|$, or other ones depending on the problem of interest. For more detail on choosing appropriate functionals, we refer, for example, to Horváth et al. (2006), Schechtman et al. (2008), and references therein.

The asymptotic distributions of the aforementioned quantities are not distribution free. Hence, a bootstrap approximation can be used. Below we present an illustration of how it can be done when we want, for example, to determine whether or not two CTE functions, say C_X and C_Y, coincide on a certain region, say $[t_1, t_2]$, of their domain of definition $[0,1]$.

We sample with replacement from $X_1, \dots, X_n$ and $Y_1, \dots, Y_m$, and obtain $C^*_{X,n}(t)$ and $C^*_{Y,m}(t)$, respectively, and so in turn the quantity

$$\sqrt{\frac{nm}{n+m}}\sup_{0<t<1} q(t)|(C^*_{X,n}(t) - C^*_{Y,m}(t)) - (C_{X,n}(t) - C_{Y,m}(t))|, \tag{6.3}$$

where the weight function q vanishes outside the region of interest $[t_1, t_2]$. Next we follow the two paragraphs below Eq. (3.4) but now with quantity (6.3) instead of (3.5) and arrive at the asymptotic $(1-\alpha)100\%$ level confidence band for the difference $C_X - C_Y$:

$$C_{X,n}(t) - C_{Y,n}(t) \pm \frac{x^*}{q(t)}\sqrt{\frac{n+m}{nm}} \quad \text{for all } t \in [0,1] \text{ such that } q(t) > 0.$$

Various other tests about relationships between C_X and C_Y can be developed along the lines above. We omit the details and refer to the already noted papers by Horváth et al. (2006) and Schechtman et al. (2008) for further hints and references.

7. Comparing CTE's when observations are paired

In this section we deal with two CTE functions when observations are independent and identically distributed bivariate random vectors $(X_1, Y_1), \dots, (X_n, Y_n)$. This situation may arise, for example, in the following context.

An insurer selling variable products with guarantees may wish to compare the CTE risk measure values associated with products with different policy designs, which may reflect different guaranteed returns. This can be done by collecting data on losses in several years for each policy design. In comparing the resulting risk measure estimates, however, it is important to recognize that the losses in a given year for the different policy designs will be dependent if both are linked to the performance of the same investment fund.

Obtaining asymptotic results for two empirical CTE functions in the case of paired observations is a more complex problem than that in the case of independent samples. In particular, statement (6.1) does not hold in the paired case, nor therefore its uniform (over a range of t values) version based on Theorem 3.2. To rectify the situation, we shall now look at the proofs of Theorems 3.1 and 3.2 and see what changes have to be made there in order to derive the limiting distribution of

$$\Delta_n(t) = \sqrt{n}((C_{X,n}(t) - C_{Y,n}(t)) - (C_X(t) - C_Y(t)))$$

either at any fixed $t \in [0,1]$ or uniformly over a range of t values.

Under the assumptions of Theorem 3.1 for both X's and Y's, we have representation (A.16) for X's and a similar one for Y's. Hence, up to a negligible remainder term, we have that $\Delta_n(t)$ is the arithmetic mean of the random variables $H(X_i;t) - H(Y_i;t)$. Hence, the asymptotic distribution of $\Delta_n(t)$ is normal with mean zero and variance $\sigma_X^2(t) + \sigma_Y^2(t) - 2\sigma_{X,Y}(t)$, where

$$\sigma_{X,Y}(t) = \frac{1}{(1-t)^2}\int_{F_X^{-1}(t)}^{\infty}\int_{F_Y^{-1}(t)}^{\infty}(F_{X,Y}(x,y) - F_X(x)F_Y(y))\,dx\,dy.$$

Based on this asymptotic result, we can now derive, for example, confidence intervals for the difference $C_X(t) - C_Y(t)$, test hypotheses about whether any of the two quantities $C_X(t)$ or $C_Y(t)$ dominates another one. These considerations proceed along the above lines where we discussed the case of two independent samples. The main difference now is that we have to empirically estimate the covariance $\sigma_{X,Y}(t)$. This we do next.

An estimator for $\sigma_{X,Y}(t)$ can be constructed by replacing F_X, F_Y, and $F_{X,Y}(x,y)$ by their empirical counterparts $F_{X,n}$, $F_{Y,n}$, and $F_{X,Y,n}(x,y)$, respectively. We obtain the formula (cf., Section 8 for details)

$$\sigma_{X,Y,n}(t) = \frac{1}{(1-t)^2}\sum_{nt\leqslant j\leqslant n-1}\sum_{nt\leqslant k\leqslant n-1}\left(\frac{\kappa_n(j,k)}{n} - \frac{j}{n}\frac{k}{n}\right)(X_{j+1:n} - X_{j:n})(Y_{k+1:n} - Y_{k:n}) \tag{7.1}$$

with the notation

$$\kappa_n(j,k) = \sum_{i=1}^{j}\mathbf{I}\{Y_{i:n}^{\mathrm{IND}} \leqslant Y_{k:n}\},$$

where $Y_{1:n}\leqslant\cdots\leqslant Y_{n:n}$ denote the order statistics of $Y_1,\ldots,Y_n$, and $Y_{1:n}^{\mathrm{IND}},\ldots,Y_{n:n}^{\mathrm{IND}}$ denote the *induced* order statistics by $X_1,\ldots,X_n$. The right-hand side of equality (7.1) gives a computationally convenient formula for $\sigma_{X,Y,n}(t)$.

The above considerations lead, for example, to the following $(1-\alpha)100\%$ level asymptotic confidence interval for the difference $C_X(t) - C_Y(t)$:

$$C_{X,n}(t) - C_{Y,n}(t) \pm \frac{z_{\alpha/2}}{\sqrt{n}}\sqrt{\sigma_{X,n}^2(t) + \sigma_{Y,n}^2(t) - 2\sigma_{X,Y,n}(t)}.$$

One can also construct bootstrap based confidence intervals for the difference $C_{X,n}(t) - C_{Y,m}(t)$. This can be done by using simple random sampling from the pairs $(X_1,Y_1),\ldots,(X_n,Y_n)$, and thus getting n new ones $(X_1^*,Y_1^*),\ldots,(X_n^*,Y_n^*)$. From the first coordinates of the new pairs we construct $C_{X,n}^*(t)$, and then $C_{Y,n}^*(t)$ from the second coordinates. This gives a value for the quantity

$$\sqrt{n}|(C_{X,n}^*(t) - C_{Y,n}^*(t)) - (C_{X,n}(t) - C_{Y,n}(t))|. \tag{7.2}$$

Following the two paragraphs below Eq. (3.4) but using quantity (7.2) instead of (3.5), we arrive at the asymptotic $(1-\alpha)100\%$ level confidence interval for the difference $C_X(t) - C_Y(t)$:

$$C_{X,n}(t) - C_{Y,n}(t) \pm \frac{x^*}{\sqrt{n}}.$$

With obvious modifications, similar considerations to those in the last few paragraphs of Section 6 hold in the case of comparing two CTE functions when observations are paired. We omit further details to avoid repetition.

8. Parametric comparison of CTE's

When all the X's and Y's are independent, then, just like in the above discussed non-parametric case, statistical inferential theory for comparing $C_X(t)$ and $C_Y(t)$ can be derived from the corresponding "univariate'' statements, cf. (4.2). Namely, we obtain the following asymptotic statement:

$$\frac{(\widehat{C}_X(t) - \widehat{C}_Y(t)) - (C_X(t) - C_Y(t))}{\sqrt{Q_X^2/n + Q_Y^2/m}} \to_d \mathrm{N}(0,1) \tag{8.1}$$

when n and m converge to infinity in such a way that the ratio $n/(n+m)$ converges to a constant $\eta\in(0,1)$. If X and Y in statement (8.1) are random variables from the set $\{X_1,X_2,X_3\}$ (i.e., exponential, Pareto, and lognormal, respectively), then expressions for Q_X^2 and Q_Y^2 can be found below statement (4.2). In turn, their parametric estimators can be obtained by replacing θ, γ and μ by the corresponding MLE $\widehat{\theta}$, $\widehat{\gamma}$, and $\widehat{\mu}$ in the formulas for Q_X^2 and Q_Y^2. Denote the estimators by $\widehat{Q}_X^2$ and $\widehat{Q}_Y^2$, respectively. Hence, for example, we can now formulate the following $100(1-\alpha)\%$ asymptotic confidence interval for $C_X(t) - C_Y(t)$:

$$\widehat{C}_X(t) - \widehat{C}_Y(t) \pm z_{\alpha/2}\sqrt{\frac{\widehat{Q}_X^2}{n} + \frac{\widehat{Q}_Y^2}{m}}. \tag{8.2}$$

When the samples are paired, then the individual asymptotic results for the empirical measures $\widehat{C}_X(t)$ and $\widehat{C}_Y(t)$ may not imply the corresponding results for their differences. Indeed, with n denoting the number of observed pairs, we write the following asymptotic result:

$$\frac{\sqrt{n}((\widehat{C}_X(t)-\widehat{C}_Y(t))-(C_X(t)-C_Y(t)))}{\sqrt{Q_X^2+Q_Y^2-2Q_{X,Y}}}\to_{\mathrm{d}}\mathrm{N}(0,1), \tag{8.3}$$

where the denominator on the left-hand side of (8.3) is the asymptotic standard deviation of $\sqrt{n}(\widehat{C}_X(t)-\widehat{C}_Y(t))$.

Depending on X and Y, the formulas for Q_X^2 and Q_Y^2 can be found below statement (4.2). Thus, we only need to find a formula for $Q_{X,Y}$ depending on the parametric distribution of X and Y. To illustrate how this problem can be solved, we now assume that X is exponential and Y is lognormal. We have the equalities:

$$\begin{aligned}
&\sqrt{n}((\widehat{C}_X(t)-\widehat{C}_Y(t))-(C_X(t)-C_Y(t)))\\
&\quad=-\sqrt{n}\left((\widehat{\theta}-\theta)(\log(1-t)-1)+\frac{1}{1-t}(\mathrm{e}^{\widehat{\mu}+0.5}-\mathrm{e}^{\mu+0.5})\Phi(1-\Phi^{-1}(t))\right)\\
&\quad=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{((X_i-x_0)-\theta)(1-\log(1-t))\right.\\
&\qquad-\left.(\log(Y_i-x_0)-\mu)\frac{1}{1-t}\mathrm{e}^{\mu+0.5}\Phi(1-\Phi^{-1}(t))\right\}+o_{\mathrm{P}}(1).
\end{aligned} \tag{8.4}$$

From (8.4), we see that

$$Q_{X,Y}=\frac{1-\log(1-t)}{1-t}\mathrm{e}^{\mu+0.5}\Phi(1-\Phi^{-1}(t))\,\mathrm{Cov}(X-x_0,\log(Y-x_0)). \tag{8.5}$$

The joint distribution of X and Y can be specified by choosing a suitable copula. Copulas might involve unknown parameters, but they can be estimated (along with θ and γ) using the maximum likelihood estimation technique. We refer to Frees and Valdez (1998) for a discussion on copulas, their use in an actuarial context, as well as for estimation of copula parameters.

Acknowledgments

The authors sincerely thank the referee and Professor Takeshi Hayakawa for valuable suggestions that helped us to improve the manuscript. The second and the fourth authors gratefully acknowledge the hospitality of the Department of Mathematics at Indiana University in Bloomington, where they worked on the project during several visits at the University. The research of the second and the fourth authors was partially supported by individual Discovery Research Grants from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

Appendix A. Proofs

Proof that CTE_X**.** may differ from C_X when F_X is not continuousFix values of $X_1,\ldots,X_n$, and let X^* denote a random variable with the distribution function $F_{X,n}$. Then

$$\begin{aligned}
\mathrm{CTE}_{X^*}(t)&=\mathbf{E}[X^*|X^*>F_{X,n}^{-1}(t)]\\
&=\frac{1}{1-F_{X,n}(F_{X,n}^{-1}(t))}\int_{(F_{X,n}^{-1}(t),\infty)}x\,\mathrm{d}F_{X,n}(x).
\end{aligned} \tag{A.1}$$

Subdivide the interval $(0,1]$ into n subintervals $((k-1)/n,k/n]$, $k=1,\ldots,n$. When $t\in((k-1)/n,k/n]$, then we have that $F_{X,n}^{-1}(t)=X_{k:n}$, the kth order statistic of $X_1,\ldots,X_n$. Clearly, $F_{X,n}(F_{X,n}^{-1}(t))=k/n$ and $\int_{(F_{X,n}^{-1}(t),\infty)}x\,\mathrm{d}F_{X,n}(x)=(1/n)\sum_{i=k+1}^{n}X_{i:n}$. Hence, for all $t\in((k-1)/n,k/n]$ we have the equation

$$\mathrm{CTE}_{X^*}(t)=\frac{1}{n-k}\sum_{i=k+1}^{n}X_{i:n}. \tag{A.2}$$

Furthermore, for the same values of $t\in((k-1)/n,k/n]$ we have that

$$C_{X^*}(t)=\frac{1}{1-t}\left(\frac{k}{n}-t\right)X_{k:n}+\frac{1}{n-nt}\sum_{i=k+1}^{n}X_{i:n}. \tag{A.3}$$

From (A.2) and (A.3) we see that the equality $\mathrm{CTE}_{X^*}(t)=C_{X^*}(t)$ holds only when $t=k/n$. □

Proof of Theorem 2.1. The statement of Theorem 2.1 is equivalent to the convergence almost surely of $\int_t^1 F_n^{-1}(u)\,du$ to $\int_t^1 F^{-1}(u)\,du$. This latter convergence follows (even uniformly over all $t \in [0,1]$) if the statement

$$\int_0^1 |F_{X,n}^{-1}(u) - F_X^{-1}(u)|\,du \to_{a.s.} 0 \tag{A.4}$$

holds. The integral on the left-hand side of (A.4) is the L_1-distance between the quantile functions $F_{X,n}^{-1}$ and F_X^{-1}. Using another terminology, the integral is the Dobrushin's distance between the distribution functions $F_{X,n}$ and F_X. Hence, assuming that the random variables $X_1, X_2, \ldots$ are independent and identically distributed, we know (cf., e.g., Shorack and Wellner, 1986, p. 65) that statement (A.4) is true if and only if the following two statements $F_{X,n} \Rightarrow F_X$ (weak convergence) and $\int |x|\,dF_{X,n}(x) \to \int |x|\,dF_X(x)$ hold almost surely. The first statement follows from the classical Glivenko–Cantelli theorem, which says that the supremum distance between $F_{X,n}$ and F_X converges almost surely to 0. The second statement can be written as $n^{-1}\sum_{i=1}^n |X_i| \to_{a.s.} \mathbf{E}[|X|]$, which holds by the strong law of large numbers whenever the moment $\mathbf{E}[X]$ is finite. This finishes the proof of Theorem 2.1. □

Proof of Theorem 2.2. Using Hölder's inequality, we have the bounds

$$\begin{aligned}\sup_{0<t<1} q(t)|C_{X,n}(t) - C_X(t)| &\leqslant \sup_{0<t<1} \frac{q(t)}{1-t}\int_t^1 |F_{X,n}^{-1}(u) - F_X^{-1}(u)|\,du \\ &\leqslant \sup_{0<t<1} \frac{q(t)}{(1-t)^{1/r}}\left(\int_t^1 |F_{X,n}^{-1}(u) - F_X^{-1}(u)|^r\,du\right)^{1/r} \\ &\leqslant \sup_{0<t<1} \frac{q(t)}{(1-t)^{1/r}}\left(\int_0^1 |F_{n,X}^{-1}(u) - F_X^{-1}(u)|^r\,du\right)^{1/r}.\end{aligned} \tag{A.5}$$

The integral on the left-hand side of (A.5) is the L_r-distance between the quantile functions $F_{X,n}^{-1}$ and F_X^{-1}. We know (cf., e.g., Shorack and Wellner, 1986, p. 65) that

$$\int_0^1 |F_{n,X}^{-1}(u) - F_X^{-1}(u)|^r\,du \to_{a.s.} 0$$

if and only if $F_{X,n} \Rightarrow F_X$ (holds due to the Glivenko–Cantelli theorem) and $\int |x|^r\,dF_{X,n}(x) \to_{a.s.} \int |x|^r\,dF_X(x)$. The latter statement holds due to the strong law of large numbers and the assumption $\mathbf{E}[|X|^r] < \infty$. □

Proof of Theorem 2.3. We start as follows:

$$\begin{aligned}\sup_{0<t<1} q(t)|C_{X,n}(t) - C_X(t)| &\geqslant \sup_{1-1/n<t<1} q(t)|C_{X,n}(t) - C_X(t)| \\ &= \sup_{1-1/n<t<1} q(t)|X_{n:n} - C_X(t)| \\ &\geqslant \sup_{1-1/n<t<1} q(t)X_{n:n} - \sup_{1-1/n<t<1} q(t)C_X(t) \\ &= \Delta_n,\end{aligned} \tag{A.6}$$

where

$$\Delta_n = \frac{1}{n^{1/r-\varepsilon}}X_{n:n} - \sup_{1-1/n<t<1}\frac{(1-t)^{1/r-\varepsilon}}{1-t}\int_t^1 F_X^{-1}(u)\,du.$$

We shall now find a distribution function such that $\mathbf{E}[|X|^r] < \infty$ but the right-hand side of (A.6) does not converge to 0 in probability. Let $F_X(x) = 1 - 1/x^{1/(1/r-\varepsilon)}$ for all $x \geqslant 1$, with the same $\varepsilon > 0$ as above. The quantile $F_X^{-1}(u)$ in this case is $1/(1-u)^{1/r-\varepsilon}$, the rth moment is finite, and the supremum on the right-hand side of (A.6) equals $1/(1-1/r+\varepsilon)$. Furthermore, since the distribution function F_X is continuous, we have that $X_{n:n} = F_X^{-1}(U_{n:n})$, where $U_{n:n} = F_X(X_{n:n})$ is the largest uniform on $[0,1]$ order statistic. Hence,

$$\Delta_n = \frac{1}{(n(1-U_{n:n}))^{1/r-\varepsilon}} - \frac{1}{1-1/r+\varepsilon}. \tag{A.7}$$

It is an easy exercise to show that, for any $\delta > 0$,

$$\lim_{n\to\infty} \mathbf{P}\{\Delta_n \geqslant \delta\} > 0. \tag{A.8}$$

This concludes the proof of the theorem. □

Proof of Theorem 3.1. We start the proof of Theorem 3.1 with the representation

$$C_{X,n}(t) - C_X(t) = \frac{1}{1-t}\int_t^1 (F_{X,n}^{-1}(u) - F_X^{-1}(u))\,\mathrm{d}u. \tag{A.9}$$

Our next step is to extract a sum of random variables from the right-hand side of (A.9). To understand how to do this, we shall now look at the integral

$$\int_t^1 (F_{X,n}^{-1}(u) - F_X^{-1}(u))\,\mathrm{d}u. \tag{A.10}$$

When $t=0$, the integral is the difference between the empirical mean $n^{-1}\sum_{i=1}^n X_i$ and the theoretical one $\mathbf{E}[X]$. This difference is certainly a sum of centered i.i.d. random variables, which is a desired representation. Note also that the difference can be written as the integral

$$-\int_{-\infty}^{\infty} (F_{X,n}(x) - F_X(x))\,\mathrm{d}x, \tag{A.11}$$

which is again a sum of centered i.i.d. random variables. The latter representation gives a most important clue on how to extract a sum of i.i.d. random variables from the integral in (A.10). This we rigorously accomplish next.

We start with the equation

$$\int_t^1 (F_{X,n}^{-1}(u) - F_X^{-1}(u))\,\mathrm{d}u = -\int_{F_X^{-1}(t)}^{\infty} (F_{X,n}(x) - F_X(x))\,\mathrm{d}x + R_{X,n}(t), \tag{A.12}$$

where the remainder term $R_{X,n}(t)$ is defined by Eq. (A.12) itself. Note that $R_{X,n}(t)$ equals 0 when $t=0$, which follows from the already noted equality of the integrals in (A.10) and (A.11). This equality, in turn, implies the following representation:

$$R_{X,n}(t) = -V_{X,n}, \tag{A.13}$$

where

$$V_{X,n}(t) = \int_0^t (F_{X,n}^{-1}(u) - F_X^{-1}(u))\,\mathrm{d}u + \int_0^{F_X^{-1}(t)} (F_{X,n}(x) - F_X(x))\,\mathrm{d}x.$$

At first sight it is difficult to see why one should prefer representation (A.13) to the original definition of $R_{X,n}(t)$ given by Eq. (A.12). The reason is as follows. The process $V_{X,n}$ has been thoroughly investigated in the literature and is known as the (general) Vervaat process (cf., e.g., the survey papers by Zitikis, 1998, and Davydov and Zitikis, 2004). Among many facts about the process, we know, for example, that $V_{X,n}(t)$ is non-negative for all $t \in [0,1]$ and satisfies the following bound:

$$V_{X,n}(t) \leqslant -(F_{X,n}(F_X^{-1}(t)) - t)(F_{X,n}^{-1}(t) - F_X^{-1}(t)) \tag{A.14}$$

for any distribution function F_X. If, however, we know that the distribution function F_X is continuous at the point $F_X^{-1}(t)$, then we have the equality $t = F_X(F_X^{-1}(t))$ and thus, in turn, the following bounds:

$$\begin{aligned} |V_{X,n}(t)| &\leqslant |F_{X,n}(F_X^{-1}(t)) - F_X(F_X^{-1}(t))||F_{X,n}^{-1}(t) - F_X^{-1}(t)| \\ &\leqslant \sup_{x\in\mathbf{R}} |F_{X,n}(x) - F_X(x)||F_{X,n}^{-1}(t) - F_X^{-1}(t)|. \end{aligned} \tag{A.15}$$

By the classical Kolmogorov–Smirnov theorem (cf., e.g., Shorack and Wellner, 1986), the supremum on the right-hand side of (A.15) is of the order $\mathcal{O}_{\mathbf{P}}(n^{-1/2})$. Thus, $\sqrt{n}|V_{X,n}(t)| = o_{\mathbf{P}}(1)$ whenever $F_n^{-1}(t) \to_{\mathbf{P}} F^{-1}(t)$. The latter convergence holds if the distribution function F_X is continuous at the point $F_X^{-1}(t)$. Hence, for every fixed $t \in [0,1]$, we have that

$$\begin{aligned} \sqrt{n}(C_{X,n}(t) - C_X(t)) &= -\frac{1}{1-t}\int_{F^{-1}(t)}^{\infty} \sqrt{n}(F_n(x) - F(x))\,\mathrm{d}x - \frac{1}{1-t}\sqrt{n}V_{X,n}(t) \\ &= -\frac{1}{1-t}\int_{F^{-1}(t)}^{\infty} \sqrt{n}(F_n(x) - F(x))\,\mathrm{d}x + o_{\mathbf{P}}(1) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^n H(X_i;t) + o_{\mathbf{P}}(1), \end{aligned} \tag{A.16}$$

where

$$H(X_i;t) = -\frac{1}{1-t}\int_{F^{-1}(t)}^{\infty} (\mathbf{I}\{X_i \leqslant x\} - F(x))\,\mathrm{d}x.$$

For every fixed $t \in [0,1]$, the random variables $H(X_i;t)$, $1 \leqslant i \leqslant n$, are centered, i.i.d., and have variances $\sigma_X^2(t)$. The variance $\sigma_X^2(t)$ is finite for every $t \in [0,1]$ if the second moment of X is finite. The latter holds by assumption. This completes the proof of Theorem 3.1. □

Proof of Theorem 3.2. With $V_{X,n}$ denoting the Vervaat process, we have the equality

$$q(t)(C_{X,n}(t)-C_X(t)) = -\frac{q(t)}{1-t}\int_{F^{-1}(t)}^{\infty}(F_n(x)-F(x))\,\mathrm{d}x - \frac{q(t)}{1-t}V_{X,n}(t). \tag{A.17}$$

We shall now prove that

$$\sqrt{n}\sup_{0<t<1}\frac{q(t)}{1-t}V_{X,n}(t)\to_{\mathbf{P}}0. \tag{A.18}$$

Using the first bound in (A.15), we have that $|V_{X,n}(t)|$ does not exceed the product of $t^{1/2-\nu}(1-t)^{1/2-\nu}|F_{X,n}^{-1}(t)-F_X^{-1}(t)|$ and

$$\sup_{0<u<1}\frac{|F_{X,n}(F_X^{-1}(u))-F_X(F_X^{-1}(u))|}{u^{1/2-\nu}(1-u)^{1/2-\nu}}. \tag{A.19}$$

By the weighted Kolmogorov–Smirnov theorem (cf., e.g., Shorack and Wellner, 1986), quantity (A.19) is of the order $\mathcal{O}_{\mathbf{P}}(n^{-1/2})$ for any $\nu>0$. Hence, statement (A.18) holds provided that

$$\sup_{0<t<1}\frac{q(t)t^{1/2-\nu}}{(1-t)^{1/2+\nu}}|F_n^{-1}(t)-F^{-1}(t)|\to_{\mathbf{P}}0. \tag{A.20}$$

In view of assumption (3.6), statement (A.20) follows from the statement

$$\sup_{0<t<1} t^{1/r_1}(1-t)^{1/r}|F_n^{-1}(t)-F^{-1}(t)|\to_{\mathbf{P}}0. \tag{A.21}$$

Since, by assumption, $\mathbf{E}[X_-^{r_1}]<\infty$ with some $r_1>2$ and $\mathbf{E}[X_+^{r}]<\infty$, we have (cf., e.g., Shorack and Wellner, 1986) that statement (A.21) holds. This proves statement (A.18).

From equality (A.17) and statement (A.18) we have that

$$\sqrt{n}\sup_{0<t<1}q(t)|C_{X,n}(t)-C_X(t)| = \sqrt{n}\sup_{0<t<1}\frac{q(t)}{1-t}\left|\int_{F^{-1}(t)}^{\infty}(F_n(x)-F(x))\,\mathrm{d}x\right| + o_{\mathbf{P}}(1). \tag{A.22}$$

We need to establish convergence in distribution of the main term on the right-hand side of (A.22). For this, we rewrite the term as follows:

$$\sup_{0<t<1}\frac{q(t)}{1-t}\left|\int_{F^{-1}(t)}^{\infty}\left[\frac{\sqrt{n}(F_n(x)-F(x))}{(1-F(x))^{1/2-\delta}}\right](1-F(x))^{1/2-\delta}\,\mathrm{d}x\right|, \tag{A.23}$$

where $\delta>0$ will be specified below. The process with respect to $-\infty<x<\infty$ in the brackets $[\cdot]$ above converges in the appropriate functional space to the process $\mathscr{B}(F(x))/(1-F(x))^{1/2-\delta}$ (cf., e.g., Shorack and Wellner, 1986). Hence, if the quantity

$$\sup_{0<t<1}\frac{q(t)}{1-t}\int_{F^{-1}(t)}^{\infty}(1-F(x))^{1/2-\delta}\,\mathrm{d}x \tag{A.24}$$

is finite for some $\delta>0$, then by the continuous mapping theorem (cf., e.g., Shorack and Wellner, 1986) we have that quantity (A.23) converges in distribution to the right-hand side of (3.7). Hence, in order to finish the proof of Theorem 3.2, we only need to verify that the quantity in (A.24) is finite for some $\delta>0$.

If we replace the supremum in quantity (A.24) by the supremum over $0<t\leqslant\frac{1}{2}$, then the resulting quantity will be finite for sufficiently small $\delta>0$, due to the assumption that $\mathbf{E}[|X|^{2+\varepsilon}]<\infty$ for some $\varepsilon>0$. Thus, we are left to check that quantity (A.24) is also finite if we replace the supremum there by the supremum over $\frac{1}{2}\leqslant t<1$. We do this as follows. First, using (3.6), we have that

$$\begin{aligned}&\sup_{1/2\leqslant t<1}\frac{q(t)}{1-t}\int_{F^{-1}(t)}^{\infty}(1-F(x))^{1/2-\delta}\,\mathrm{d}x\\ &\quad\leqslant c\sup_{1/2\leqslant t<1}\frac{1}{(1-t)^{1-(1/r+1/2+\nu)}}\int_{F^{-1}(t)}^{\infty}(1-F(x))^{1/2-\delta}\,\mathrm{d}x.\end{aligned} \tag{A.25}$$

Now we choose $v>0$ sufficiently small so that $1/r+1/2+v<1$ would hold (such a choice is possible since $r>2$). Continuing with (A.25), we have

$$\sup_{1/2\leqslant t<1}\frac{1}{(1-t)^{1-(1/r+1/2+v)}}\int_{F^{-1}(t)}^{\infty}(1-F(x))^{1/2-\delta}\,\mathrm{d}x$$
$$\leqslant \sup_{1/2\leqslant t<1}\int_{F^{-1}(t)}^{\infty}\frac{(1-F(x))^{1/2-\delta}}{(1-F(x))^{1-(1/r+1/2+v)}}\,\mathrm{d}x$$
$$\leqslant \int_{F^{-1}(1/2)}^{\infty}(1-F(x))^{1/r+v-\delta}\,\mathrm{d}x. \tag{A.26}$$

Choose now $\delta=v/2$. Since $\mathbf{E}[X_+^r]<\infty$ by assumption, the right most integral in (A.26) is finite. This finishes the proof of Theorem 3.2. □

Proof of Formula (7.1). The following equalities are straightforward:

$$\sigma_{X,Y,n}(t)=\frac{1}{(1-t)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(F_{X,Y,n}(x,y)-F_{X,n}(x)F_{Y,n}(y))\mathbf{I}\{F_{X,n}(x)\geqslant t\}\mathbf{I}\{F_{Y,n}(x)\geqslant t\}\,\mathrm{d}x\,\mathrm{d}y$$
$$=\frac{1}{(1-t)^2}\sum_{nt\leqslant j\leqslant n-1}\;\sum_{nt\leqslant k\leqslant n-1}\int_{X_{j:n}}^{X_{j+1:n}}\int_{Y_{k:n}}^{Y_{k+1:n}}\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{I}\{X_i\leqslant x,Y_i\leqslant y\}-\frac{j}{n}\frac{k}{n}\right)\mathrm{d}x\,\mathrm{d}y$$
$$=\frac{1}{(1-t)^2}\sum_{nt\leqslant j\leqslant n-1}\;\sum_{nt\leqslant k\leqslant n-1}\left(\frac{\kappa_n(j,k)}{n}-\frac{j}{n}\frac{k}{n}\right)(X_{j+1:n}-X_{j:n})(Y_{k+1:n}-Y_{k:n}). \tag{A.27}$$

This completes the proof of formula (7.1). □

References

Acerbi, C., Tasche, D., 2002. On the coherence of expected shortfall. J. Banking Finance 26, 1487–1503.
American Academy of Actuaries, 2002. Recommended approach for setting regulatory risk-based capital requirements for variable products with guarantees (excluding index guarantees). In: Presented by the American Academy of Actuaries' Life Capital Adequacy Subcommittee to the National Association of Insurance Commissioners' Life Risk-Based Capital Working Group, San Diego, CA, December 2002 ⟨http://www.actuary.org/pdf/life/rbc_16dec02.pdf⟩.
Artzner, P., 1999. Application of coherent risk measures to capital requirements in insurance. North Amer. Actuarial J. 3, 11–29.
Brazauskas, V., Kaiser, T., 2004. Discussion on: "Empirical estimation of risk measures and related quantities" [North Amer. Actuarial J. 7(4) (2003) 44–54] by B.L. Jones, R. Zitikis. With a reply by Jones and Zitikis. North Amer. Actuarial J. 8(3), 114–118.
Brazauskas, V., Serfling, R., 2000. Robust and efficient estimation of the tail index of a single-parameter Pareto distribution (with discussion). North Amer. Actuarial J. 4 (4), 12–27 (Discussion by M. Bilodeau in North Amer. Actuarial J. 5(3), 123–126, and with the authors reply in North Amer. Actuarial J. 5(3), 126–128).
Brazauskas, V., Serfling, R., 2003. Favorable estimators for fitting Pareto models: a study using goodness-of-fit measures with actual data. ASTIN Bull. 33 (2), 365–381.
Canadian Institute of Actuaries, 2002. Report of the CIA task force on segregated fund investment guarantees ⟨http://www.actuaries.ca/publications/2002/202012e.pdf⟩.
Chow, K.V., 2001. Marginal conditional stochastic dominance, statistical inference and measuring portfolio performance. J. Financial Res. 24, 289–307.
Csörgö, M., Zitikis, R., 1996. Mean residual life processes. Ann. Statist. 24, 1717–1739.
Csörgö, M., Csörgö, S., Horváth, L., 1986. An asymptotic theory for empirical reliability and concentration processes. Lecture Notes in Statistics, vol. 33. Springer, New York.
Davydov, Y., Zitikis, R., 2004. Convex rearrangements of random elements. In: Horváth, L., Szyszkowicz, B. (Eds.), Asymptotic Methods in Stochastics, Fields Institute Communications. American Mathematical Society, Providence, RI.
Frees, E.W., Valdez, E.A., 1998. Understanding relationships using copulas. North Amer. Actuarial J. 2 (1), 1–25.
Hardy, M.R., Wirch, J.L., 2004. The iterated CTE: a dynamic risk measure. North Amer. Actuarial J. 8, 62–75.
Horváth, L., Kokoszka, P., Zitikis, R., 2006. Testing for stochastic dominance using the weighted McFadden type statistic. J. Econometrics 133, 191–205.
Kochar, S.C., Mukerjee, H., Samaniego, F.J., 2000. Estimation of a monotone mean residual life. Ann. Statist. 28, 905–921.
Landsman, Z.M., Valdez, E.A., 2003. Tail conditional expectations for elliptical distributions. North Amer. Actuarial J. 7, 55–71.
Manistre, B.J., Hancock, G.H., 2005. Variance of the CTE estimator. North Amer. Actuarial J. 9 (2), 129–156.
Mayshar, J., Yitzhaki, S., 1995. Dalton-improving indirect tax reform. Amer. Econom. Rev. 85, 793–807.
Puri, M.L., Sen, P.K., 1971. Nonparametric Methods in Multivariate Analysis. Wiley, New York.
Schechtman, E., Yitzhaki, S., 2004. The Gini instrumental variables, or "the double IV estimator". Metron LXII, 1–27.
Schechtman, E., Shelef, A., Yitzhaki, S., Zitikis, R., 2008. Testing hypotheses about absolute concentration curves and marginal conditional stochastic dominance. Econometric Theory, 24 (4), to appear.
Seiler, E.J., 2001. A nonparametric test for marginal conditional stochastic dominance. Appl. Financial Econom. 11, 173–177.
Serfling, R., 2002. Efficient and robust fitting of lognormal distributions. North Amer. Actuarial J. 6 (4), 95–109.
Serfling, R.J., 1980. Approximation Theorems of Mathematical Statistics. Wiley, New York.
Shalit, H., Yitzhaki, S., 1994. Marginal conditional stochastic dominance. Management Sci. 40, 670–684.
Shalit, H., Yitzhaki, S., 2003. An asset allocation puzzle: comment. Amer. Econom. Rev. 93, 1002–1008.
Shorack, G.R., Wellner, J.A., 1986. Empirical Processes with Applications to Statistics. Wiley, New York.
Tasche, D., 2002. Expected shortfall and beyond. J. Banking Finance 26, 1519–1533.
Wirch, J.L., Hardy, M.R., 1999. A synthesis of risk measures for capital adequacy. Insurance Math. Econom. 25, 337–347.
Zitikis, R., 1998. The Vervaat process. In: Szyszkowicz, B. (Ed.), Asymptotic Methods in Probability and Statistics. North-Holland, Amsterdam, pp. 667–694.