METRON - *International Journal of Statistics*
2007, vol. LXV, n. 2, pp. 175-199

VYTARAS BRAZAUSKAS - BRUCE L. JONES - RIČARDAS ZITIKIS

# Robustification and performance evaluation of empirical risk measures and other vector-valued estimators

*Summary* - Actuarial axioms lead to risk measures that for loss distributions assign weighted integrals of the corresponding value-at-risk functions. Hence, constructing robust and efficient estimators for the integrals using observed losses becomes a task of practical interest. Furthermore, a number of risk measures are functionals of several such weighted integrals. Hence, in order to cover many cases of practical interest, in the present paper we consider robust estimators of vector-valued risk measures as well as other population parameters, and we also study their large-sample efficiency properties. In addition, we discuss robust estimators for location, scale, and location-scale parameters for several parametric families of interest in actuarial science, econometrics, and beyond.

*Key Words* - Risk measure; Distortion function; Empirical estimator; Robustness; Efficiency; Parametric inference; Nonparametric inference.

## 1. Introduction

Risk measures are, in general, functionals from the set of loss distributions to the extended real line. Depending on the adopted system of axioms (see, e.g., Young, 2004, for a discussion and references), researchers have arrived at various representations of risk measures, such as integrals of loss distributions (see, e.g., Wang, 1998, and references therein). This topic is extensively researched in numerous journal articles and monographs in the actuarial and financial areas.

Jones and Zitikis (2003) observed and put into good use the fact that a large class of risk measures can be expressed as weighted integrals $\int_0^1 J(u)F^{-1}(u)\mathrm{d}u$ of the quantile function $F^{-1}$ or, in other words, value-at-risk function corresponding to the underlying loss distribution $F$, as illustrated by the following example.

*Received January 2007 and revised July 2007.*

**Example 1.1.** Among numerous risk measures that appear in the actuarial literature, three very frequent ones are proportional hazards transform (PHT) risk measure, Wang's transform (WT) risk measure, and conditional tail expectation (CTE). The three risk measures can be written as the integral

$$R[X] = \int_0^1 J(u)F^{-1}(u)\mathrm{d}u$$

with the weight function $J(u)$ defined – depending on the risk measure – as follows:

$$\begin{aligned}
\text{PHT}: \quad & J(u) = \frac{1-\rho}{(1-u)^{\rho}} \quad \text{with} \quad 0 \le \rho < 1,\\
\text{WT}: \quad & J(u) = \exp\left\{\lambda\Phi^{-1}(u) - \frac{1}{2}\lambda^2\right\} \quad \text{with} \quad 0 \le \lambda < \infty,\\
\text{CTE}: \quad & J(u) = \frac{I_{[\alpha,1]}(u)}{1-\alpha} \quad \text{with} \quad 0 \le \alpha < 1,
\end{aligned}$$

where $\Phi^{-1}(t)$ is the standard normal quantile function, and $I_{[\alpha,1]}(t)$ is the indicator function taking value 1 if $t \in [\alpha, 1]$ and 0 otherwise. Note that setting the parameters $\rho$, $\lambda$, and $\alpha$ to 0 yields $J(u) \equiv 1$ in all the cases and thus the equation $R[X] = \mathbf{E}[X]$. □

In view of the representation of risk measures as integrals $\int_0^1 J(u)F^{-1}(u)\mathrm{d}u$, $L$-statistics (i.e., linear combinations of order statistics) such as $\sum_{i=1}^n c_{i,n}X_{i:n}$ with $c_{i,n} = \int_{(i-1)/n}^{i/n} J(u)\mathrm{d}u$ become natural estimators of the risk measures. Brazauskas and Kaiser (2004), and Kaiser and Brazauskas (2006) investigate robustness properties of the estimators under the assumption that their asymptotic variances are finite, which restricts the set of functionals that can be used for constructing risk measures. In the present paper we enlarge the class of risk measures by employing truncated versions of the aforementioned asymptotic variances. Furthermore, we establish joint asymptotics of a finite number of $L$-statistics, which is important as empirical risk measures can be differences, ratios, or other functionals of several $L$-statistics, as illustrated in the next example.

**Example 1.2.** The risk measure $R[X]$ (see Example 1.1) is usually 'loaded', that is, $R[X]$ is not smaller than the 'net premium' $\mathbf{E}[X]$. To measure how much larger $R[X]$ is, it is natural to use either the absolute distance $R[X]-\mathbf{E}[X]$ or the relative distance $R[X]/\mathbf{E}[X]-1$ (see, e.g., Wang, 1998). Both distances can be estimated from data by the functionals $h(x, y) = x - y$ and $h(x, y) = x/y - 1$, respectively, of the bivariate $L$-statistic $(\sum_{i=1}^n c_{i,n}X_{i:n}, \sum_{i=1}^n X_{i:n})$, thus justifying the need for establishing asymptotic properties such as consistency, asymptotic normality, etc. For other examples of actuarial and econometric uses of multivariate $L$-statistics, we refer to Brazauskas et al. (2007), Jones et al. (2006), Zitikis (2002), and references therein. □

Generally, in the present paper we aim at robust estimators for a very broad spectrum of vector-parameters such as those related to actuarial risk measures and indices of economic inequality and also parameters that are related to several parametric classes of distributions (e.g., location-scale families, gamma, beta, Pareto and Weibull distributions) of particular interest in actuarial science, econometrics, reliability engineering, and other areas (see, e.g., Kleiber and Kotz, 2003).

There are numerous methods in the statistical literature for estimating vector-parameters. They include maximum likelihood, minimum distance, least squares, moments, to name a few. These methods can usually be found as special cases of some general classes of statistics, such as $M$-, $L$-, or $R$-statistics (see, e.g., Serfling, 1980, Chapters 7–9). If one's objective is robust estimation of vector-parameters, which is the case in the present paper, then $M$-, $L$-, and $R$-statistics offer such estimators, with the class of $M$-statistics being arguably the most popular choice, which is mostly due to a close relationship between the objective function of $M$-statistic and its influence function. The latter function is an important tool for studying robustness properties of the estimators. However, the class of robust $M$-estimators has not been received too enthusiastically in practice because of several reasons: First, in order to define the estimators in terms of the underlying influence function, the practitioner needs to have a fairly deep understanding of robust statistics. Second, to compute the estimators, the practitioner needs to solve a system of nonlinear equations, which can be a challenging task from the computational point of view.

In view of these and other challenges related to practical implementation of $M$-estimators – and our main concern in the present paper is practical implementation – robust estimators based on $L$-statistics are particularly attractive: they are easy to define, typically have explicit expressions, and one can easily see, understand and thus control the actions of the estimators on data; these are particularly important features from the practical viewpoint. Moreover, as already noted above, $L$-statistics have natural uses and interpretations in actuarial and econometric sciences. For these reasons in particular, in the present paper we look at the performance of $L$-statistics via simulation studies in both parametric and non-parametric settings.

In addition to estimating actuarial risk measures and indices of economic inequality, the other goal of the present paper is to analyze asymptotically robust estimators of location, scale, and location-scale parameters. We shall also investigate efficiency properties of these estimators for a number of parametric families: Cauchy, Student's $t$, normal, and Laplace. In addition, we shall compute asymptotic relative efficiency of robust estimators with respect to a) the sample mean for estimation of location, b) modified sample standard deviation for estimation of scale, and c) sample mean and modified sample standard deviation for joint estimation of location and scale. Readers familiar with robust estimation problems in location-scale families will recognize these estimators, which are thoroughly studied and widely scattered across the literature. To achieve these goals, in Section 3 we present a theoretical basis concerning asymptotic results for multivariate trimmed $L$-statistics in various scenarios of interest. In Sections 4 and 5 we provide examples of asymptotically robust estimators of location, scale, and location-scale parameters with their efficiency properties for various symmetric location-scale families. A numerical illustration of the theoretical results is given in Section 5. A summary and concluding notes are given in Section 6.

## 2. Background and (un)trimmed risk measure values

Given $K$ populations with distribution functions $F_1, \dots, F_K$, we are interested in estimating the vector-parameter $\boldsymbol{\mu}_0 = (\mu_0(1), \dots, \mu_0(K))$, where the coordinates are the (untrimmed) 'risk measure values'

$$\mu_0(k) = \int_0^1 J_k(u)\, h_k \circ F_k^{-1}(u) \quad \mathrm{d}u$$

with $J_k : (0, 1) \to [0, \infty)$ and $h_k : (-\infty, \infty) \to (-\infty, \infty)$ specified by the researcher and with the symbol '$\circ$' denoting the composition of two functions, that is, $h_k \circ F_k^{-1}(u) = h_k(F_k^{-1}(u))$.

**Assumption 2.1.** For every $1 \le k \le K$, the function $h_k$ can be expressed as the sum $\sum_{j=1}^{m_k} h_{k,j}$ of non-decreasing and left-continuous functions $h_{k,1}, \dots, h_{k,m_k}$.

Considering just one coordinate $\mu_0(k)$ is already of interest in many applications, let alone the vector-parameter $\boldsymbol{\mu}_0$ and its functionals. As a simple example, consider the functions $J_k(u) = 1$ and $h_k(u) = 1$ for all $u$; then $\mu_0(k)$ is the mean of $F_k$. Take $h_k(u) = 1$ for all $u$; then $\mu_0(k)$ is one of the most popular actuarial risk measures $\int_0^1 J_k(u)\, F_k^{-1}(u) \quad \mathrm{d}u$. The latter integral also includes a number of well known indices of economic/income inequality. Furthermore, take $J_k(u) = 1$ for all $u$ and let $h_k(u)$ be a generic function; then $\mu_0(k)$ defines yet another well known class of actuarial risk measures $\int_0^1 h_k \circ F_k^{-1}(u) \quad \mathrm{d}u$, which distorts the quantile function $F_k^{-1}$ in a 'non-linear' fashion, unlike the earlier noted integral $\int_0^1 J_k(u)\, F_k^{-1}(u) \quad \mathrm{d}u$, which is a linear functional of $F_k^{-1}$.

Since the $K$ population distributions $F_1, \ldots, F_K$ are generally unknown, we collect data:

$$\begin{array}{ccccc} X_1(1), & \ldots, & X_{n(1)}(1) & \sim & F_1 \\ & \ddots & & & \vdots \\ X_1(k), & \ldots, & X_{n(k)}(k) & \sim & F_k \\ & \ddots & & & \vdots \\ X_1(K), & \ldots, & X_{n(K)}(K) & \sim & F_K \end{array}$$

**Assumption 2.2.** For every $k$, the random variables $X_1(k), \ldots, X_{n(k)}(k)$ are independent and identically distributed, and no additional information is assumed (unless explicitly stated otherwise) about the dependence structure between the vectors $(X_1(k), \ldots, X_{n(k)}(k))$, $k = 1, \ldots, K$.

Various dependence structures are of interest in practice. For example, we encounter situations where the vectors $(X_1(k), \ldots, X_{n(k)}(k))$, $1 \le k \le K$, are independent; this is the so-called case of $K$ independent populations. In other situations, we have coordinate-wise equalities, that is, $X_j(1) = \cdots = X_j(K)$ for every $1 \le j \le n$ with $n = n(1) = \cdots = n(K)$ and $F_1 = \cdots = F_K$. This case occurs when dealing with combinations of several statistics based on the same sample; examples include the aforementioned relative risk measures (see, e.g., Wang, 1998) and indices of economic inequality (see, e.g., Tarsitano, 2004).

Returning to the general vector-parameter $\boldsymbol{\mu}_0$, we estimate it using the multivariate $L$-statistic $\mathbf{L}_{n(1),\ldots,n(K)} = (L_{n(1)}(1), \ldots, L_{n(K)}(K))$ with the coordinates defined by

$$L_{n(k)}(k) = \frac{1}{n(k)} \sum_{i=m_n(k)+1}^{n(k)-m_n^*(k)} c_{i,n(k)}(k)\, h_k(X_{i:n(k)}(k)),$$

where the following notation has been used: The random variables $X_{1:n(k)}(k) \le X_{2:n(k)}(k) \le \cdots \le X_{n(k):n(k)}(k)$ are the order statistics of $X_1(k), X_2(k), \ldots, X_{n(k)}(k)$. The integers $m_n(k)$ and $m_n^*(k)$ are such that $0 \le m_n(k) < n(k) - m_n^*(k) \le n(k)$. The coefficients $c_{i,n(k)}(k)$ are generated by a function $J_k$ – which we assume to be non-negative and continuous on $(0,1)$ – in such a way that, for every $1 \le i \le n(k)$,

$$c_{i,n(k)}(k) = n(k) \int_{(i-1)/n(k)}^{i/n(k)} J_k(u) \, \mathrm{d}u.$$

To give a flavour of the $L$-statistics that we utilize later, an example follows.

**Example 2.1.** When estimating the location and scale parameters in Section 4, we use the following bivariate $L$-statistic $(L_{n(1)}(1), L_{n(2)}(2)) = (\widehat{\mu}_{a_1}, S^2_{a_2})$, where

$$\widehat{\mu}_{a_1} = \sum_{i=m_n(1)+1}^{n-m_n(1)} \left[\frac{1}{n-2m_n(1)}\right] X_{i:n} \tag{1}$$

and

$$S^2_{a_2} = \sum_{i=m_n(2)+1}^{n-m_n(2)} \left[\frac{1}{n}\right] (X_{i:n} - \mu)^2. \tag{2}$$

The equation $\widehat{\mu}_{a_1} = L_{n(1)}(1)$ holds with $n(1) = n$ and $m_n(1) = m_n^*(1)$ such that $m_n(1)/n \to a_1$ when $n \to \infty$, and the functions $h_1(x) = x$ and $J_1(u) \equiv 1/(1-2a_1)$. Furthermore, $S^2_{a_2} = L_{n(2)}(2)$ holds with $n(2) = n$ and $m_n(2) = m_n^*(2)$ with $m_n(2)/n \to a_2$, and the functions $h_2(x) = (x-\mu)^2$ and $J_2(u) \equiv 1$. $\square$

Our next task is to specify the centering vector

$$\boldsymbol{\mu}_{n(1),\dots,n(K)} = (\mu_{n(1)}(1), \dots, \mu_{n(K)}(K))$$

and the normalizing vector $\mathbf{b}_{n(1),\dots,n(K)} = (b_{n(1)}(1), \dots, b_{n(K)}(K))$ such that the asymptotic result

$$\mathbf{S}_{n(1),\dots,n(K)} = \left(\mathbf{L}_{n(1),\dots,n(K)} - \boldsymbol{\mu}_{n(1),\dots,n(K)}\right) \mathbf{D}\left(\mathbf{b}_{n(1),\dots,n(K)}\right) \longrightarrow_d \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}) \tag{3}$$

holds with the $K \times K$-dimensional matrix $\mathbf{D}(\mathbf{b}_{n(1),\dots,n(K)})$ whose diagonal entries are $b_{n(1)}(1), \dots, b_{n(K)}(K)$ and the off-diagonal entries are 0. The $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$ denotes a Gaussian vector with the mean $\mathbf{0} = (0, \dots, 0)$ and the covariance-variance matrix $\boldsymbol{\Sigma}$ whose entries will be specified later, depending on the trimming considered.

For the validity of statement (3), it is necessary (but not sufficient) to have the asymptotic normality of

$$\mathbf{S}_{n(1),\dots,n(K)}\mathbf{e}(k)^T = b_{n(k)}(k)\left(L_{n(k)}(k) - \mu_{n(k)}(k)\right), \tag{4}$$

where $\mathbf{e}(k) = (0, \dots, 0, 1, 0, \dots, 0)$ with 1 in the $k$th place. A detailed investigation of such one-dimensional theoretical results and references have been provided by Mason and Shorack (1990). Multivariate untrimmed $L$-statistics have been investigated by Zitikis (1993) but those results are not applicable in the context of the present paper as we now need truncation due to reasons noted above. To gain further intuition on the topic and to also introduce additional notation, we note (see, e.g., Serfling, 1980) that when the quantity

$$\sigma_0^2(k) = \iint_{[0,1)} (u \wedge v - uv)\, J_k(u) J_k(v)\ \ \mathrm{d}h_k \circ F_k^{-1}(u)\ \ \mathrm{d}h_k \circ F_k^{-1}(v) \tag{5}$$

is finite (with the notation $\iint_A \equiv \int_A \int_A$), then after centering the untrimmed $L$-statistic (i.e., when $m_n(k) = 0$ and $m_n^*(k) = 0$) with the earlier introduced 'asymptotic mean' $\mu_0(k)$, the right-hand side of equation (4) is equal to $n(k)^{-1/2} \sum_{i=1}^{n(k)} Y_i(k) + o_{\mathbf{P}}(1)$, where $Y_1(k), \dots, Y_{n(k)}(k)$ are i.i.d. random variables with zero means and variances $\sigma_0^2(k)$. We know (see, e.g., Serfling, 1980) that the latter random variables can be expressed by

$$Y_i(k) = -\int_0^1 \left(\mathbf{1}_{\{U_i(k)\le t\}} - t\right) J_k(u) \;\; \mathrm{d}h_k \circ F_k^{-1}(u), \tag{6}$$

where the random variables $U_1(k), \dots, U_{n(k)}(k)$ are independent and uniformly distributed on (0, 1). In various uses of $L$-statistics, the variance $\sigma_0^2(k)$ is finite. However, there are situations where it is of interest to go beyond finite $\sigma_0^2(k)$. For this reason, the normalizing constant

$$\iint_{[a_n(k),1-b_n(k))} (u \wedge v - uv)\, J_k(u) J_k(v) \;\; \mathrm{d}h_k \circ F_k^{-1}(u) \;\; \mathrm{d}h_k \circ F_k^{-1}(v) \tag{7}$$

with $0 < a_n(k) < 1 - b_n(k) < 1$ becomes natural as truncation assures the finiteness of the integral for larger classes of functions $F_k$ and $J_k$. This is the reason why in this paper we introduce the 'trimmed risk measure value'

$$\mu_{n(k)}(k) = \int_{m_n(k)/n(k)}^{1-m_n^*(k)/n(k)} J_k(u)\, h_k \circ F_k^{-1}(u) \;\; \mathrm{d}u.$$

Naturally, the estimator of $\mu_{n(k)}(k)$ is $L_{n(k)}(k)$ and the length of confidence intervals for $\mu_{n(k)}(k)$ depends on variance-type quantities like that in (7).

## 3. Estimating trimmed risk measure values

The asymptotic results for untrimmed $L$-statistics discussed in the previous section can easily be extended to the case in which we trim only a fixed number of smallest and largest order statistics:

$$m_n(k) = m(k) \quad \text{and} \quad m_n^*(k) = m^*(k) \quad \text{with fixed} \quad m(k),\ m^*(k) \ge 1. \tag{8}$$

With the normalizing vector $\mathbf{b}^0_{n(1),\dots,n(K)} = (b^0_{n(1)}(1), \dots, b^0_{n(K)}(K))$ whose coordinates are $b^0_{n(k)}(k) = \sqrt{n(k)}/\sigma_0(k)$, we introduce a particular form of $\mathbf{S}_{n(1),\dots,n(K)}$, defined as follows:

$$\mathbf{S}^0_{n(1),\dots,n(K)} = \left(\mathbf{L}_{n(1),\dots,n(K)} - \boldsymbol{\mu}_{n(1),\dots,n(K)}\right) \mathbf{D}\left(\mathbf{b}^0_{n(1),\dots,n(K)}\right).$$

**Assumption 3.1.** Let the weight functions $J_k(t)$ be of the form $J_k(t) = t^{\alpha_k}(1-t)^{\beta_k}$. (If desired, the class of possible $J_k$'s can be somewhat enlarged; see Condition J in Mason and Shorack, 1990, p. 114.)

**Theorem 3.1.** *Under the trimming scheme in* (8) *and Assumption* 3.1*, in addition to Assumptions* 2.1 *and* 2.2*, we have that*

$$\mathbf{S}^0_{n(1),\dots,n(K)} \longrightarrow_d \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}^{FIX}), \tag{9}$$

*where the entries of the covariance-variance matrix* $\mathbf{\Sigma}^{FIX} = [\rho^{FIX}(k,l)]_{k,l=1}^K$ *are given by*

$$\rho^{FIX}(k,l) = \frac{1}{\sigma_0(k)\sigma_0(l)} \times \\ \times \iint_{[0,1)} \Big(\mathbf{P}\left[U(k) \le u, U(l) \le v\right] - uv\Big) J_k(u) J_l(v)\ \ dh_k \circ F_k^{-1}(u)\ \ dh_l \circ F_l^{-1}(v) \tag{10}$$

*with* $\sigma_0^2(k)$ *defined by equation* (5). [*We assume that* $\sigma_0^2(k)$ *and* $\rho^{FIX}(k,l)$ *are well defined and finite.*]

The formula for $\rho^{FIX}(k,l)$ contains 'unobservable' random variables $U(k)$ and $U(l)$; both of them have uniform distributions on (0, 1). However, we need to know their joint distribution in order to calculate (or estimate) the right-hand side of equation (10). For this reason we first note that the joint distribution of $X(k)$ and $X(l)$ is equal to that of $F_k^{-1}(U(k))$ and $F_l^{-1}(U(l))$. This is useful as we shall see in the next two paragraphs.

Suppose that the random variables $X(k)$ and $X(l)$ are independent when $k \ne l$; this happens when the vectors $(X_1(k), \dots, X_{n(k)}(k))$, $k = 1, \dots, K$, are independent. In this case $U(k)$ and $U(l)$ are independent, and thus $\mathbf{P}[U(k) \le u, U(l) \le v] = uv$. This gives the equation $\rho^{FIX}(k,l) = 0$ whenever $k \ne l$. Since $\rho^{FIX}(k,k) = 1$, the matrix $\mathbf{\Sigma}^{FIX}$ is unit.

Next we deal with the situation in which $X_j(1) = \dots = X_j(K)$ for every $1 \le j \le n$, where $n = n(1) = \dots = n(K)$. This situation can be reformulated as just one set of independent and identically distributed random variables $X_1, \dots, X_n \sim F$. Hence, $U(k) \equiv U(l)$ and thus $\mathbf{P}[U(k) \le u, U(l) \le v] = u \wedge v$. This in turn gives the equation

$$\rho^{FIX}(k,l) \\ = \frac{1}{\sigma_0(k)\sigma_0(l)} \iint_{[0,1)} (u \wedge v - uv)\, J_k(u) J_l(v)\ \ \mathrm{d}h_k \circ F^{-1}(u)\ \ \mathrm{d}h_l \circ F^{-1}(v) \tag{11}$$

with

$$\sigma_0^2(k) = \iint_{[0,1)} (u \wedge v - uv)\, J_k(u) J_k(v)\ \ \mathrm{d}h_k \circ F^{-1}(u)\ \ \mathrm{d}h_k \circ F^{-1}(v).$$

In a simulation study later in this paper, we use the following two-dimensional (i.e., $K = 2$) version of the above theorem: The vector $(L_n(1), L_n(2))$ is $\mathcal{A}N(\boldsymbol{\mu}_n, n^{-1}\boldsymbol{\Xi}^{FIX})$ with $\boldsymbol{\mu}_n = (\mu_n(1), \mu_n(2))$ and

$$\boldsymbol{\Xi}^{FIX} = \begin{pmatrix} \sigma_0^2(1) & \sigma_0(1)\sigma_0(2)\rho^{FIX}(1,2) \\ \sigma_0(1)\sigma_0(2)\rho^{FIX}(1,2) & \sigma_0^2(2) \end{pmatrix},$$

where

$$L_n(k) = \sum_{i=m_n(k)+1}^{n-m_n^*(k)} \left[\int_{(i-1)/n}^{i/n} J_k(u) \ \mathrm{d}u\right] h_k(X_{i:n}), \tag{12}$$

$$\mu_n(k) = \int_{m_n(k)/n}^{1-m_n^*(k)/n} J_k(u) h_k \circ F^{-1}(u) \ \mathrm{d}u. \tag{13}$$

We next consider statistical inference under vanishing-fraction trimming. Here we trim a vanishing fraction of smallest and largest order statistics, that is, we assume that

$$m_n(k),\ m_n^*(k) \to \infty \quad \text{but} \quad \frac{m_n(k)}{n(k)},\ \frac{m_n^*(k)}{n(k)} \to 0, \tag{14}$$

when $n \to \infty$.

**Assumption 3.2.** Let the weight functions $J_k(t)$ be such that the random variable $H_k(U)$ (see notation below) is in the domain of attraction of a stable law with index $\alpha \in (0, 2]$, where the random variable $U$ is uniformly distributed on $(0, 1)$ and

$$H_k(t) = \begin{cases} \displaystyle\int_{[c,t)} J_k(u) \ \mathrm{d}h_k \circ F_k^{-1}(u) & \text{when } t > c, \\ 0 & \text{when } t = c, \\ \displaystyle -\int_{[t,c)} J_k(u) \ \mathrm{d}h_k \circ F_k^{-1}(u) & \text{when } t < c, \end{cases}$$

with a fixed continuity point $c \in (0, 1)$ of the function $h_k \circ F_k^{-1}$. (If desired, the class of weight functions $J_k$ can be somewhat increased; see condition (1.37) in Mason and Shorack, 1990, p. 117.)

**Theorem 3.2.** *Under the trimming scheme in* (14) *and Assumption* 3.2*, in addition to Assumptions 2.1 and* 2.2*, we have that*

$$\mathbf{S}_{n(1),\dots,n(K)} \longrightarrow_d \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}^{VANISH}), \tag{15}$$

*where* $\boldsymbol{\Sigma}^{VANISH} = [\rho^{VANISH}(k,l)]_{k,l=1}^{K}$ *with the entries defined by*

$$\rho^{VANISH}(k,l) = \lim_{n\to\infty} \frac{1}{\sigma_{n(k)}(k)\sigma_{n(k)}(l)} \int_{[m_n(k)/n(k),1-m_n^*(k)/n(k))} \int_{[m_n(l)/n(l),1-m_n^*(l)/n(l))} \times \Big(\mathbf{P}[U(k)\le u, U(l)\le v] - uv\Big) J_k(u)J_l(v)\ dh_k\circ F_k^{-1}(u)\ dh_l\circ F_l^{-1}(v) \tag{16}$$

*with*

$$\sigma_{n(k)}^2(k) = \iint_{[m_n(k)/n(k),1-m_n^*(k)/n(k))} (u\wedge v - uv)\, J_k(u)J_k(v)\ dh_k\circ F_k^{-1}(u)\ dh_k\circ F_k^{-1}(v).$$

[*We assume that the limit* $\rho^{VANISH}(k,l)$ *exists and is finite.*]

When $X_1,\dots,X_n \sim F$ and $K=2$, statement (15) implies the following one to be used in later sections: The vector $(L_n(1), L_n(2))$ is $\mathcal{AN}(\boldsymbol{\mu}_n, n^{-1}\boldsymbol{\Xi}_n^{VANISH})$ with $\boldsymbol{\mu}_n = (\mu_n(1), \mu_n(2))$ and

$$\boldsymbol{\Xi}_n^{VANISH} = \begin{pmatrix} \sigma_n^2(1) & \sigma_n(1)\sigma_n(2)\rho_n^{VANISH}(1,2) \\ \sigma_n(1)\sigma_n(2)\rho_n^{VANISH}(1,2) & \sigma_n^2(2) \end{pmatrix}.$$

We next consider statistical inference under fixed-fraction trimming. Here we trim a fixed fraction of smallest and largest order statistics, that is, we assume that, for some $0 < a_k < 1-b_k < 1$,

$$\frac{m_n(k)}{n(k)} \to a_k + o\Big(\frac{1}{\sqrt{n(k)}}\Big) \text{ and } \frac{m_n^*(k)}{n(k)} \to b_k + o\Big(\frac{1}{\sqrt{n(k)}}\Big). \tag{17}$$

**Assumption 3.3.** Let the weight functions $J_k(t)$ satisfy the Lipschitz condition on an open interval containing $[a_k, b_k]$, and let both $a_k$ and $b_k$ be continuity points of the function $h_k\circ F_k^{-1}$. (If desired, the class of possible functions $J_k$ and $h_k\circ F_k^{-1}$ can be somewhat enlarged; see condition (1.33) in Mason and Shorack, 1990, p. 117.)

**Theorem 3.3.** *Under the trimming scheme in* (17) *and Assumption* 3.3*, in addition to Assumptions* 2.1 *and* 2.2*, we have that*

$$\mathbf{S}_{n(1),\dots,n(K)} \longrightarrow_d \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}^{FRAC}), \tag{18}$$

*where the entries of the covariance-variance matrix* $\boldsymbol{\Sigma}^{FRAC} = [\rho^{FRAC}(k,l)]_{k,l=1}^{K}$ *are*

$$\rho^{FRAC}(k,l) = \frac{1}{\sigma_{a_k,b_k}(k)\sigma_{a_k,b_k}(l)} \int_{[a_k,1-b_k)} \int_{[a_l,1-b_l)} \times \Big(\mathbf{P}[U(k)\le u, U(l)\le v] - uv\Big) J_k(u)J_l(v)\ dh_k\circ F_k^{-1}(u)\ dh_l\circ F_l^{-1}(v) \tag{19}$$

*with*

$$\sigma^2_{a_k,b_k}(k) = \iint_{[a_k,b_k)} (u \wedge v - uv)\, J_k(u) J_k(v)\;\, dh_k \circ F_k^{-1}(u)\;\, dh_k \circ F_k^{-1}(v).$$

[*We assume that* $\sigma^2_{a_k,b_k}(k)$ *and* $\rho^{FRAC}(k,l)$ *are well defined and finite.*]

In the special case in which $X_1, \ldots, X_n \sim F$ and $K = 2$, statement (18) implies that the vector $(L_n(1), L_n(2))$ is $\mathcal{AN}(\boldsymbol{\mu}_n, n^{-1}\boldsymbol{\Xi}^{FRAC})$ with $\boldsymbol{\mu}_n = (\mu_n(1), \mu_n(2))$ and

$$\boldsymbol{\Xi}^{FRAC} = \begin{pmatrix} \sigma^2_{a_1,b_1}(1) & \sigma_{a_1,b_1}(1)\sigma_{a_2,b_2}(2)\rho^{FRAC}(1,2) \\ \sigma_{a_1,b_1}(1)\sigma_{a_2,b_2}(2)\rho^{FRAC}(1,2) & \sigma^2_{a_2,b_2}(2) \end{pmatrix}.$$

## 4. Estimating the location and scale parameters

Let $X_1, \ldots, X_n$ be i.i.d. random variables with a distribution function $F_{\mu,\sigma}$ which belongs to a location-scale family; denote the corresponding pdf by $f_{\mu,\sigma}$, and assume it is symmetric. At least one of the parameters $\mu \in (-\infty, \infty)$ and $\sigma > 0$ is unknown, and we shall use trimmed $L$-statistics to estimate them. In particular, we shall study the asymptotic relative efficiency (ARE) of the estimators:

$$\mathrm{ARE}(E_1,\ E_0) = \frac{\text{asymptotic variance of } \textit{reference} \text{ estimator } E_0}{\text{asymptotic variance of } \textit{competing} \text{ estimator } E_1}\ .$$

In the multivariate version of the ARE, the variance is replaced by the generalized variance, which is defined as the determinant of the asymptotic covariance-variance matrix, and then the ratio is raised to the power $1/K$, where $K$ is the dimension of the vector-parameter. For further details on the topic, we refer, for example, to Serfling (1980, Section 4.1).

### Estimating the location parameter when scale is known

Let an $L$-statistic $L_n(1)$ be (see Example 2.1)

$$\widehat{\mu}_{a_1} = \sum_{i=m_n(1)+1}^{n-m_n(1)} \left[\frac{1}{n - 2m_n(1)}\right] X_{i:n},$$

which is achieved by choosing $m_n(1) = m_n^*(1)$, which in turn implies $m_n(1)/n = m_n^*(1)/n$ and $m_n(1)/n \to a_1 = b_1$ when $n \to \infty$, and $h_1(x) = x$ and $J_1(u) =$

$1/(1-2a_1)$ for $0 < u < 1$. We next justify the use of the symmetric trimming, $a_1 = b_1$. Elementary calculations show that

$$\begin{aligned} \mu_n(1) &= \int_{a_1}^{1-a_1} \left[\frac{1}{1-2a_1}\right] F_{\mu,\sigma}^{-1}(u) \quad \mathrm{d}u \\ &= \frac{1}{1-2a_1}\left[\sigma \int_{F_{0,1}^{-1}(a_1)}^{F_{0,1}^{-1}(1-a_1)} z \quad \mathrm{d}F_{0,1}(z)\right] + \mu \;=\; \mu. \end{aligned} \tag{20}$$

It is clear from the second line in (20) that in order to eliminate the dependence of $\mu_n(1)$ on $\sigma$ we have to choose $a_1 = b_1$. (The elimination of $\sigma$ is not necessary when it is known but desirable when $\sigma$ is unknown.) We next compute the asymptotic variance of $\widehat{\mu}_{a_1}$, which is (see Appendix B)

$$\sigma^2_{a_1,a_1}(1) = \frac{\sigma^2}{(1-2a_1)^2}\left\{2a_1\left[F_{0,1}^{-1}(a_1)\right]^2 + \delta_2(F_{0,1},\, a_1)\right\} \tag{21}$$

with

$$\delta_2(F_{0,1},\, a_1) = 2\int_0^{F_{0,1}^{-1}(1-a_1)} s^2 \quad \mathrm{d}F_{0,1}(s). \tag{22}$$

We call $\delta_2(F_{0,1},\, a_1)$ the 'trimmed second moment' of $F_{0,1}$ as it converges to the (finite or infinite) second moment of $F_{0,1}$ when $a_1 \downarrow 0$. We summarize the above discussion by saying that $\widehat{\mu}_{a_1}$ is $\mathcal{A}N(\mu,\, n^{-1}\sigma^2\Delta_\mu)$, where

$$\Delta_\mu \equiv \Delta_\mu(F_{0,1},\, a_1) = \frac{1}{(1-2a_1)^2}\left[2a_1\left[F_{0,1}^{-1}(a_1)\right]^2 + \delta_2(F_{0,1},\, a_1)\right].$$

ESTIMATING THE SCALE PARAMETER WHEN LOCATION IS KNOWN

Consider an $L$-statistic $L_n(2)$ with such parameters (see Example 2.1) that it becomes

$$S^2_{a_2} = \sum_{i=m_n(2)+1}^{n-m_n(2)} \left[\frac{1}{n}\right] (X_{i:n} - \mu)^2.$$

Namely, $m_n(2) = m_n^*(2)$ or, equivalently, $m_n(2)/n = m_n^*(2)/n$ and $m_n(2)/n \to a_2 = b_2$. Furthermore, $h_2(x) = (x-\mu)^2$ and $J_2(u) = 1$, for $0 < u < 1$; the choices of functions $h_2$ and $J_2$ will become clear when we derive asymptotic properties of $S^2_{a_2}$.

Unlike in the case of the location estimator $\widehat{\mu}_{a_1}$, symmetric trimming in the current situation is not necessary. However, when we consider the joint estimation of $\mu$ and $\sigma$, the estimator of $\sigma$ will look similar to $S^2_{a_2}$ and in its formula $\mu$ will be replaced with $\widehat{\mu}_{a_1}$. Since robustness properties, e.g., the breakdown point, of a vector-estimator are defined by those of its weakest (i.e., least robust) coordinate-estimator, it makes sense to use similar trimming schemes for all coordinates of a vector-estimator.

Straightforward calculations give (see Appendix B) the equations

$$\mu_n(2) = \sigma^2\,\delta_2(F_{0,1},\ a_2), \tag{23}$$

and

$$\begin{aligned}&\sigma^2_{a_2,a_2}(2)\\&=\sigma^4\Big\{\delta_4(F_{0,1},a_2)-4a_2\big[F_{0,1}^{-1}(a_2)\big]^2\delta_2(F_{0,1},a_2)-\delta_2^2(F_{0,1},a_2)+2a_2(1-2a_2)\big[F_{0,1}^{-1}(a_2)\big]^4\Big\}\end{aligned} \tag{24}$$

with

$$\delta_4(F_{0,1},\ a_2) = 2\int_0^{F_{0,1}^{-1}(1-a_2)} s^4\ \ \mathrm{d}F_{0,1}(s). \tag{25}$$

We call $\delta_4(F_{0,1},\ a_2)$ the 'trimmed fourth moment' of $F_{0,1}$ as it converges to the (finite or infinite) fourth moment of $F_{0,1}$ when $a_2 \downarrow 0$. We see from equation (23) that $L_n(2)$ is not a consistent estimator of $\sigma$. However, the following modification

$$\widehat{\sigma}_{a_2} = \sqrt{L_n(2)/\delta_2(F_{0,1},\ a_2)}$$

estimates $\sigma$ consistently, which can easily be shown using the delta method (see, e.g., Serfling, 1980, p. 118). The asymptotic mean of $\widehat{\sigma}_{a_2}$ is $\mu_n^*(2) = \sqrt{\mu_n(2)\,/\delta_2\,(F_{0,1},\ a_2)} = \sigma$ and its asymptotic variance is $\sigma^{*2}_{a_2,\,a_2}(2) = \sigma^2_{a_2,a_2}(2)\big[4\sigma^2\delta_2(F_{0,1},\ a_2)\big]^{-1}$. To summarize the discussion, we have that $\widehat{\sigma}_{a_2}$ is $\mathcal{A}N(\sigma,\ n^{-1}\sigma^2\Delta_\sigma)$, where

$$\Delta_\sigma \equiv \Delta_\sigma(F_{0,1}, a_2) = 0.25\delta_2^{-2}\Big[\delta_4-\delta_2^2+2a_2(1-2a_2)\big[F_{0,1}^{-1}(a_2)\big]^4\Big]-a_2\big[F_{0,1}^{-1}(a_2)\big]^2\delta_2^{-1}.$$

JOINT ESTIMATION OF LOCATION AND SCALE PARAMETERS

The problem of joint estimation of $\mu$ and $\sigma$ has essentially been solved in the previous two sections. Indeed, we estimate $\mu$ with $\widehat{\mu}_{a_1}$ and $\sigma$ with $\widehat{\sigma}_{a_2}$, where the now unknown parameter $\mu$ in the definition of $L_n(2)$ (which is needed for $\widehat{\sigma}_{a_2}$) is replaced by its estimator $\widehat{\mu}_{a_1}$. Note that the latter estimator of $\sigma$, which we denote $\widehat{\sigma}^*_{a_2}$, has the same asymptotic distribution as $\widehat{\sigma}_{a_2}$ as seen from the decomposition

$$\widehat{\sigma}^{*2}_{a_2} = \sum_{i=m_n(2)+1}^{n-m_n(2)} \left[\frac{1}{n\delta_2}\right](X_{i:n} - \widehat{\mu}_{a_1})^2 = \widehat{\sigma}^2_{a_2} + \text{rem}$$

and Slutsky's theorem, where the remainder term

$$\text{rem} = -2\frac{1-2a_2}{\delta_2}(\widehat{\mu}_{a_1}-\mu)(\widehat{\mu}_{a_2}-\mu) + \frac{1-2a_2}{\delta_2}(\widehat{\mu}_{a_1}-\mu)^2$$

is of the order $O_{\mathbf{P}}(n^{-1})$ in view of $m_n(1) \sim a_1 n$ and $m_n(2) \sim a_2 n$ (see Example 2.1). Therefore, for asymptotic considerations, we can still rely on the results established for $\widehat{\sigma}_{a_2}$ above. This implies that for joint estimation of $\mu$ and $\sigma$ we only need to evaluate the covariance between $\widehat{\mu}_{a_1}$ and $\widehat{\sigma}_{a_2}$. Note that (see Appendix B)

$$\sigma_{a_1,a_1}(1)\,\sigma_{a_2,a_2}(2)\,\rho^{FRAC}(1,\ 2) = 0. \tag{26}$$

In summary, the vector

$$(\widehat{\mu}_{a_1}, \widehat{\sigma}^*_{a_2}) = \left(\sum_{i=m_n(1)+1}^{n-m_n(1)} \left[\frac{1}{n-2m_n(1)}\right] X_{i:n},\ \left(\sum_{i=m_n(2)+1}^{n-m_n(2)} \left[\frac{1}{n\delta_2}\right](X_{i:n}-\widehat{\mu}_{a_1})^2\right)^{1/2}\right)$$

is

$$\mathcal{A}N\left((\mu,\sigma),\ \frac{\sigma^2}{n}\begin{pmatrix}\Delta_\mu & 0\\ 0 & \Delta_\sigma\end{pmatrix}\right).$$

## 5. Numerical illustration

Here we study the ARE of the estimators $\widehat{\mu}_{a_1}$, $\widehat{\sigma}_{a_2}$, and $(\widehat{\mu}_{a_1}, \widehat{\sigma}^*_{a_2})$ for the parameters $\mu \in \mathbf{R}$, $\sigma > 0$, and $(\mu, \sigma)$ of the following symmetric location-scale families: Student's $t_\nu$ with (integer) $\nu \geq 1$ degrees of freedom, normal, and Laplace; in Table 1 (see Appendix A) we also report results for the Cauchy distribution, which is a special case of the Student's distribution when $\nu = 1$. When verifying numerical computations, it is also helpful to keep in mind that the normal distribution is, loosely speaking, the Student's distribution when $\nu = \infty$.

The ARE of the robust estimators is computed with respect to 1) the sample mean for estimating the location, 2) the modified sample standard deviation for estimating the scale, and 3) the sample mean and modified sample standard deviation for the joint estimation of location and scale. The asymptotic distributions of these reference estimators are well-known in the statistical literature (see, e.g., Serfling, 1980, Chapter 3). We list them next for convenient reference. Namely, for estimating the location when the scale is known, the sample mean

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \ \text{ is } \mathcal{A}N\left(\mu,\ \frac{\sigma^2}{n}\,\Delta^\circ_\mu\right)$$

where $\Delta_\mu^\circ \equiv \Delta_\mu^\circ(F_{0,1}) = \lim_{a_1 \to 0} \Delta_\mu(F_{0,1},\, a_1)$. For estimating the scale when the location is known, the distribution-adjusted sample standard deviation

$$S_* = \sqrt{\frac{1}{n\delta_2^\circ} \sum_{i=1}^{n} (X_i - \mu)^2} \text{ is } \mathcal{AN}\left(\sigma,\ \frac{\sigma^2}{n}\,\Delta_\sigma^\circ\right),$$

where

$$\delta_2^\circ \equiv \delta_2^\circ(F_{0,1}) = \lim_{a_2 \to 0} \delta_2(F_{0,1},\, a_2) \quad \text{and} \quad \Delta_\sigma^\circ \equiv \Delta_\sigma^\circ(F_{0,1}) = \lim_{a_2 \to 0} \Delta_\sigma(F_{0,1}, a_2).$$

For the joint estimation of the location and scale, the vector-estimator

$$(\overline{X},\, S_{**}) = \left(\frac{1}{n}\sum_{i=1}^{n} X_i,\, \sqrt{\frac{1}{n\delta_2^\circ}\sum_{i=1}^{n}(X_i - \overline{X})^2}\right)$$

is

$$\mathcal{AN}\left((\mu, \sigma),\ \frac{\sigma^2}{n}\begin{pmatrix} \Delta_\mu^\circ & 0 \\ 0 & \Delta_\sigma^\circ \end{pmatrix}\right).$$

(Note that the numerical values of the efficiency-related constants of the reference estimators can be easily found from the corresponding constants of the robust estimators by letting $a_1 \downarrow 0$ and $a_2 \downarrow 0$.) These results combined with our findings in earlier sections imply that

$$\mathrm{ARE}(\widehat{\mu}_{a_1},\, \overline{X}) = \frac{\Delta_\mu^\circ}{\Delta_\mu}, \qquad \mathrm{ARE}(\widehat{\sigma}_{a_2},\, S_*) = \frac{\Delta_\sigma^\circ}{\Delta_\sigma},$$

and

$$\mathrm{ARE}\big((\widehat{\mu}_{a_1}, \widehat{\sigma}_{a_2}^*),\, (\overline{X},\, S_{**})\big) = \sqrt{\frac{\Delta_\mu^\circ \Delta_\sigma^\circ}{\Delta_\mu \Delta_\sigma}}.$$

Hence, to calculate the ARE's, we need to evaluate $\Delta_\mu^\circ$, $\Delta_\mu$, $\Delta_\sigma^\circ$, $\Delta_\sigma$, and thus, in turn, $\delta_2^\circ$, $\delta_2$, $\delta_4^\circ$, $\delta_4$. We accomplish this next, using equations (22) and (25), and calculating their right-hand sides for the aforementioned three distributions.

- Student's $t_\nu$ distribution:

$$\delta_2(F_{0,1}, a_1) = \begin{cases} 2\log\left(\sqrt{1 + \frac{\left[F_{0,1}^{-1}(a_1)\right]^2}{2}} - \frac{F_{0,1}^{-1}(a_1)}{\sqrt{2}}\right) - 2\,(1 - 2a_1)\,, \ \nu = 2. \\ \frac{\nu}{\nu - 2}\ (1 - 2a_1) \\ \quad + 2\frac{\nu}{\nu - 2} F_{0,1}^{-1}(a_1)\, f_{0,1}(F_{0,1}^{-1}(a_1))\left(1 + \frac{\left[F_{0,1}^{-1}(a_1)\right]^2}{\nu}\right), \quad \nu \neq 2. \end{cases}$$

$$\delta_4(F_{0,1}, a_2) = \begin{cases} 4(1-2a_2)-4F_{0,1}^{-1}(a_2)\, f_{0,1}(F_{0,1}^{-1}(a_2)) \left(1+\dfrac{\left[F_{0,1}^{-1}(a_2)\right]^2}{2}\right)^2 & \\ \quad -6\log\left(\sqrt{1+\dfrac{\left[F_{0,1}^{-1}(a_2)\right]^2}{2}} - \dfrac{F_{0,1}^{-1}(a_2)}{\sqrt{2}}\right), & \nu = 2. \\ -32(1-2a_2) & \\ \quad -32F_{0,1}^{-1}(a_2)\, f_{0,1}(F_{0,1}^{-1}(a_2)) \left(1+\dfrac{\left[F_{0,1}^{-1}(a_2)\right]^2}{4}\right) & \\ \quad +24\log\left(\sqrt{1+\dfrac{\left[F_{0,1}^{-1}(a_2)\right]^2}{4}} - \dfrac{F_{0,1}^{-1}(a_2)}{2}\right), & \nu = 4. \\ \dfrac{3\nu^2}{(\nu-2)(\nu-4)}(1-2a_2) & \\ \quad +\dfrac{2\nu}{\nu-4}F_{0,1}^{-1}(a_2)\, f_{0,1}\left(F_{0,1}^{-1}(a_2)\right)\left(1+\dfrac{\left[F_{0,1}^{-1}(a_2)\right]^2}{\nu}\right) & \\ \quad \times\left(\dfrac{3\nu}{\nu-2}+\left[F_{0,1}^{-1}(a_2)\right]^2\right), & \nu \neq 2, 4. \end{cases}$$

Also, $\delta_2^\circ = \nu/(\nu-2)$ for $\nu > 2$, and $= \infty$ for $\nu \leq 2$; $\delta_4^\circ = 3\nu^2/((\nu-4)(\nu-2))$ for $\nu > 4$, and $= \infty$ for $\nu \leq 4$. Consequently, $\Delta_\mu^\circ = \nu/(\nu-2)$ for $\nu > 2$, and $= \infty$ for $\nu \leq 2$; $\Delta_\sigma^\circ = (\nu-1)/(2\nu-8)$ for $\nu > 4$, and $= \infty$ for $\nu \leq 4$.

- Normal distribution:

$$\delta_2(F_{0,1}, a_1) = \big(1-2a_1\big) + 2F_{0,1}^{-1}(a_1)\, f_{0,1}\big(F_{0,1}^{-1}(a_1)\big),$$
$$\delta_4(F_{0,1}, a_2) = 2\big[F_{0,1}^{-1}(a_2)\big]^3 f_{0,1}\big(F_{0,1}^{-1}(a_2)\big) + 3\delta_2(F_{0,1}, a_2).$$

Also, $\delta_2^\circ = 1$, $\delta_4^\circ = 3$ and $\Delta_\mu^\circ = 1$, $\Delta_\sigma^\circ = 1/2$.

- Laplace distribution:

$$\delta_2(F_{0,1}, a_1) = 2\big(1-2a_1\big) - 2a_1\left[\log^2(2a_1) - 2\log(2a_1)\right],$$
$$\delta_4(F_{0,1}, a_2) = -2a_2\log^4(2a_2) + 8a_2\log^3(2a_2) + 12\delta_2(F_{0,1}, a_2).$$

Also, $\delta_2^\circ = 2$, $\delta_4^\circ = 24$ and $\Delta_\mu^\circ = 2$, $\Delta_\sigma^\circ = 5/4$.

The formulas above provide all the necessary components for the development of computer code. Numerical illustrations are provided in Table 1 for asymptotic variances and in Figure 1 for ARE's, in the case $a_1 = a_2$. In the table and figure, the reference estimators correspond to the estimators with no trimming, that is, when $a_1 = 0$, $a_2 = 0$, and $a_1 = a_2 = 0$. Several conclusions emerge from these illustrations. First, for the Cauchy distribution even the mildest trimming provides an infinite improvement over the standard empirical estimators. Second, except for the estimation of location in the Laplace family (where the median estimator is optimal), the general pattern is that high levels of trimming lead to small efficiencies. Third, favorable efficiencies (often exceeding 1) are attained in the range of 2%–10% trimming for estimation of $\mu$, $\sigma$, and $(\mu, \sigma)$, which implies that the breakdown points of favorable procedures range between 0.02 and 0.10.

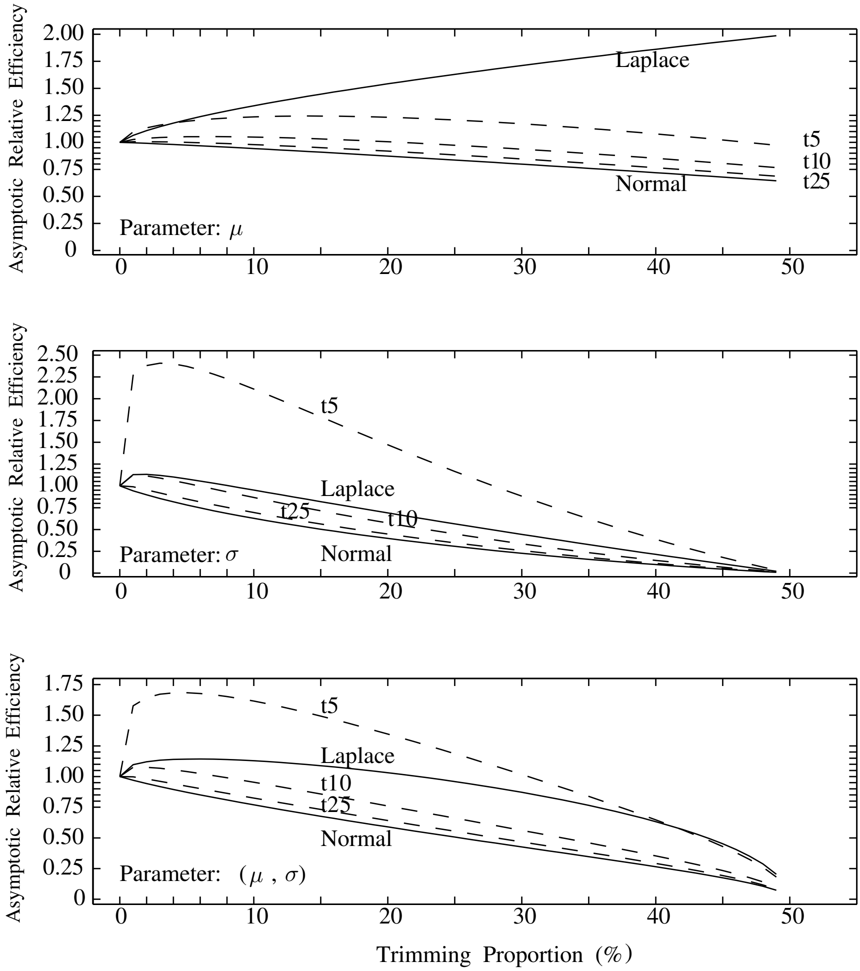


Figure 1. Asymptotic relative efficiency as functions of trimming percentages of robust estimators with respect to the estimators $\overline{X}$, $S_*$, $(\overline{X}, S_{**})$ of the parameters $\mu$, $\sigma$, $(\mu, \sigma)$, respectively.

## 6. Summary and concluding remarks

In this paper we have considered the problem of robust estimation of vector-parameters that originate in actuarial science as risk measure values. We have investigated the performance of estimators based on vectors of trimmed $L$-statistics whose weights are formed by means of weight functions $J$. Asymptotic normality and numerical performance of the statistics have been investigated in the case of trimming vanishing fraction, fixed fraction and fixed number of observations. Examples of asymptotically robust estimators (i.e, those based on trimmed $L$-statistics with fixed fractions of observations removed) of location, scale, and location-scale parameters have been provided. The asymptotic relative efficiency of the estimators has been quantified for the Cauchy, Student's $t$, normal, and Laplace distributions.

Results of the present paper suggest several interesting venues for further research. To highlight one of them, choose the simplest but prominent and popular member of the class of univariate trimmed $L$-statistics: the trimmed mean, which is well studied in the literature and often used in statistical practice. First-order asymptotic properties – including asymptotic normality – of the trimmed mean were established by Stigler (1973), whereas simple and explicit second-order approximations of Edgeworth-type (correcting for skewness and bias of a trimmed mean) have recently been derived by Gribkova and Helmers (2006). Results of the latter paper suggest a possibility for going beyond the first-order asymptotic properties considered in the present paper and investigate Edgeworth-type approximations for trimmed (multivariate) $L$-statistics.

## 7. Appendix A: numerical evaluations

Table 1: *Asymptotic variances (divided by $\sigma^2/n$) for one-dimensional and generalized asymptotic variances (divided by $\sigma^4/n^2$) for two-dimensional empirical and robust estimators of $\mu$, $\sigma$, and $(\mu, \sigma)$, for selected location-scale families.*

| **Estimator** | **Trimming proportions** | Symmetric location-scale family: Cauchy | $t_5$ | $t_{10}$ | $t_{15}$ | $t_{20}$ | $t_{25}$ | $t_{50}$ | $t_{100}$ | Normal | Laplace |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\widehat{\mu}_{a_1}$ | 0 | $\infty$ | 1.67 | 1.25 | 1.15 | 1.11 | 1.09 | 1.04 | 1.02 | 1.00 | 2.00 |
| | 0.02 | 20.90 | 1.47 | 1.20 | 1.13 | 1.10 | 1.08 | 1.04 | 1.03 | 1.01 | 1.80 |
| | 0.04 | 10.79 | 1.41 | 1.19 | 1.13 | 1.10 | 1.08 | 1.05 | 1.04 | 1.02 | 1.70 |
| | 0.06 | 7.43 | 1.38 | 1.19 | 1.13 | 1.11 | 1.09 | 1.06 | 1.05 | 1.03 | 1.62 |
| | 0.08 | 5.76 | 1.36 | 1.19 | 1.14 | 1.11 | 1.10 | 1.07 | 1.06 | 1.05 | 1.55 |
| | 0.10 | 4.77 | 1.35 | 1.19 | 1.15 | 1.12 | 1.11 | 1.08 | 1.07 | 1.06 | 1.49 |
| | 0.12 | 4.12 | 1.34 | 1.20 | 1.15 | 1.13 | 1.12 | 1.10 | 1.09 | 1.08 | 1.45 |
| | 0.14 | 3.66 | 1.34 | 1.21 | 1.17 | 1.15 | 1.14 | 1.11 | 1.10 | 1.09 | 1.40 |
| | 0.16 | 3.32 | 1.34 | 1.22 | 1.18 | 1.16 | 1.15 | 1.13 | 1.12 | 1.11 | 1.36 |
| | 0.18 | 3.07 | 1.35 | 1.23 | 1.19 | 1.18 | 1.17 | 1.15 | 1.14 | 1.13 | 1.33 |
| | 0.20 | 2.87 | 1.35 | 1.24 | 1.21 | 1.19 | 1.18 | 1.16 | 1.15 | 1.14 | 1.30 |
| $\widehat{\sigma}_{a_2}$ | 0 | $\infty$ | 2.00 | 0.75 | 0.64 | 0.59 | 0.57 | 0.53 | 0.52 | 0.50 | 1.25 |
| | 0.02 | 9.02 | 0.84 | 0.67 | 0.63 | 0.61 | 0.60 | 0.58 | 0.57 | 0.56 | 1.10 |
| | 0.04 | 4.89 | 0.83 | 0.71 | 0.67 | 0.66 | 0.65 | 0.63 | 0.62 | 0.61 | 1.13 |
| | 0.06 | 3.54 | 0.86 | 0.75 | 0.72 | 0.71 | 0.70 | 0.68 | 0.68 | 0.67 | 1.18 |
| | 0.08 | 2.90 | 0.90 | 0.81 | 0.78 | 0.77 | 0.76 | 0.74 | 0.74 | 0.73 | 1.24 |
| | 0.10 | 2.54 | 0.95 | 0.87 | 0.84 | 0.83 | 0.82 | 0.81 | 0.80 | 0.80 | 1.31 |
| | 0.12 | 2.32 | 1.01 | 0.93 | 0.91 | 0.90 | 0.89 | 0.88 | 0.88 | 0.87 | 1.39 |
| | 0.14 | 2.19 | 1.08 | 1.01 | 0.99 | 0.98 | 0.97 | 0.96 | 0.96 | 0.95 | 1.48 |
| | 0.16 | 2.12 | 1.16 | 1.09 | 1.08 | 1.07 | 1.06 | 1.05 | 1.05 | 1.04 | 1.58 |
| | 0.18 | 2.10 | 1.25 | 1.19 | 1.17 | 1.17 | 1.16 | 1.15 | 1.15 | 1.14 | 1.69 |
| | 0.20 | 2.11 | 1.36 | 1.30 | 1.29 | 1.28 | 1.28 | 1.27 | 1.26 | 1.26 | 1.81 |
| $(\widehat{\mu}_{a_1}, \widehat{\sigma}^*_{a_1})$ | 0 | $\infty$ | 3.33 | 0.94 | 0.73 | 0.66 | 0.62 | 0.55 | 0.53 | 0.50 | 2.50 |
| | 0.02 | 188.44 | 1.24 | 0.81 | 0.71 | 0.67 | 0.65 | 0.60 | 0.58 | 0.56 | 1.99 |
| | 0.04 | 52.75 | 1.18 | 0.84 | 0.76 | 0.72 | 0.70 | 0.66 | 0.64 | 0.62 | 1.93 |
| | 0.06 | 26.34 | 1.18 | 0.89 | 0.82 | 0.78 | 0.76 | 0.73 | 0.71 | 0.69 | 1.91 |
| | 0.08 | 16.71 | 1.22 | 0.96 | 0.89 | 0.85 | 0.83 | 0.80 | 0.78 | 0.76 | 1.93 |
| | 0.10 | 12.11 | 1.28 | 1.03 | 0.96 | 0.93 | 0.91 | 0.88 | 0.86 | 0.84 | 1.96 |
| | 0.12 | 9.56 | 1.35 | 1.12 | 1.05 | 1.02 | 1.00 | 0.97 | 0.95 | 0.93 | 2.01 |
| | 0.14 | 8.03 | 1.44 | 1.22 | 1.15 | 1.12 | 1.10 | 1.07 | 1.05 | 1.04 | 2.07 |
| | 0.16 | 7.06 | 1.55 | 1.33 | 1.27 | 1.24 | 1.22 | 1.19 | 1.17 | 1.15 | 2.15 |
| | 0.18 | 6.44 | 1.68 | 1.47 | 1.40 | 1.37 | 1.35 | 1.32 | 1.30 | 1.29 | 2.24 |
| | 0.20 | 6.06 | 1.84 | 1.62 | 1.56 | 1.53 | 1.51 | 1.47 | 1.46 | 1.44 | 2.35 |

## 8. Appendix B: Proofs

Proof of statement (9)

Under the trimming scheme in (8), for every $1 \le k \le K$, Theorem 1.1 in Mason and Shorack (1990, pp. 115–116) implies that $\|\mathbf{S}^0_{n(1),\ldots,n(K)} + \mathbf{S}^{FIX}_{n(1),\ldots,n(K)}\| = o_{\mathbf{P}}(1)$, where the norm $\|\cdot\|$ can be any (e.g., Euclidean) and $\mathbf{S}^{FIX}_{n(1),\ldots,n(K)} = \mathbf{T}^{FIX}_{n(1),\ldots,n(K)}\mathbf{D}(\mathbf{b}^0_{n(1),\ldots,n(K)})$ with the vector $\mathbf{T}^{FIX}_{n(1),\ldots,n(K)} = (T^{FIX}_{n(1)}(1), \ldots, T^{FIX}_{n(K)}(K))$ whose coordinates are (see equation (6))

$$T^{FIX}_{n(k)}(k) = \frac{1}{n(k)} \sum_{i=1}^{n(k)} \int_{[0,1)} \left(\mathbf{1}_{\{U_i(k)\le u\}} - u\right) J_k(u) \ \ \mathrm{d}h_k \circ F_k^{-1}(u).$$

The multivariate CLT implies that $\mathbf{S}^{FIX}_{n(1),\ldots,n(K)} \longrightarrow_d \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}^{FIX})$ with the matrix $\mathbf{\Sigma}^{FIX} = [\rho^{FIX}(k,l)]_{k,l=1}^K$ in the formulation of the theorem. This concludes the proof. □

Proof of statement (15)

Under the trimming scheme in (14), for every $1 \le k \le K$, Theorem 1.3 in Mason and Shorack (1990, p. 117) implies that $\|\mathbf{S}_{n(1),\ldots,n(K)} + \mathbf{S}^{VANISH}_{n(1),\ldots,n(K)}\| = o_{\mathbf{P}}(1)$, where $\mathbf{S}^{VANISH}_{n(1),\ldots,n(K)} = \mathbf{T}^{VANISH}_{n(1),\ldots,n(K)}\mathbf{D}(\mathbf{b}^{VANISH}_{n(1),\ldots,n(K)})$ with the coordinates of the vector $\mathbf{T}^{VANISH}_{n(1),\ldots,n(K)} = (T^{VANISH}_{n(1)}(1), \ldots, T^{VANISH}_{n(K)}(K))$ defined by

$$T^{VANISH}_{n(k)}(k) = \frac{1}{n(k)} \sum_{i=1}^{n(k)} \int_{[m_n(k)/n(k),1-m_n^*(k)/n(k))} \left(\mathbf{1}_{\{U_i(k)\le t\}} - t\right) J_k(u)\ \mathrm{d}h_k \circ F_k^{-1}(u).$$

The coordinates of $\mathbf{b}^{VANISH}_{n(1),\ldots,n(K)} = (b^{VANISH}_{n(1)}(1), \ldots, b^{VANISH}_{n(K)}(K))$ are defined by the equation

$$b^{VANISH}_{n(k)}(k) = \sqrt{n(k)}/\sigma_{n(k)}(k),$$

where $\sigma^2_{n(k)}(k)$ is defined in the formulation of the theorem. The multivariate CLT implies that $\mathbf{S}^{VANISH}_{n(1),\ldots,n(K)} \longrightarrow_d \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}^{VANISH})$ with the variance-covariance matrix $\mathbf{\Sigma}^{VANISH} = [\rho^{VANISH}(k,l)]_{k,l=1}^K$ whose entries are given in the formulation of the theorem. This concludes the proof. □

PROOF OF STATEMENT (18)

Under the trimming scheme in (17), for every $1 \le k \le K$, Theorem 1.2 in Mason and Shorack (1990, p. 117) implies that $\|\mathbf{S}_{n(1),\dots,n(K)} + \mathbf{S}^{FRAC}_{n(1),\dots,n(K)}\| = o_{\mathbf{P}}(1)$, where $\mathbf{S}^{FRAC}_{n(1),\dots,n(K)} = \mathbf{T}^{FRAC}_{n(1),\dots,n(K)} \mathbf{D}(\mathbf{b}^{FRAC}_{n(1),\dots,n(K)})$. The coordinates of $\mathbf{T}^{FRAC}_{n(1),\dots,n(K)} = (T^{FRAC}_{n(1)}(1), \dots, T^{FRAC}_{n(K)}(K))$ are

$$T^{FRAC}_{n(k)}(k) = \frac{1}{n(k)} \sum_{i=1}^{n(k)} \int_{[a_k, 1-b_k)} \big(\mathbf{1}_{\{U_i(k) \le t\}} - t\big)\, J_k(u) \quad \mathrm{d}h_k \circ F_k^{-1}(u).$$

The coordinates of $\mathbf{b}^{FRAC}_{n(1),\dots,n(K)} = (b^{FRAC}_{n(1)}(1), \dots, b^{FRAC}_{n(K)}(K))$ are $b^{FRAC}_{n(k)}(k) = \sqrt{n(k)}/\sigma_{a_k,b_k}(k)$ with $\sigma^2_{a_k,b_k}(k)$ defined in the formulation of the theorem. The multivariate CLT implies that $\mathbf{S}^{FRAC}_{n(1),\dots,n(K)} \longrightarrow_d \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}^{FRAC})$ with the variance-covariance matrix $\mathbf{\Sigma}^{FRAC} = [\rho^{FRAC}(k,l)]_{k,l=1}^{K}$ defined in the formulation of the theorem. This concludes the proof. $\square$

To proceed, we first list several well-known properties of the pdf's, cdf's, and quantile functions of symmetric location-scale families:

- For $x \in (-\infty, \infty)$: $F_{\mu,\sigma}(x) = F_{0,1}\big((x-\mu)/\sigma\big)$ and $f_{\mu,\sigma}(x) = f_{0,1}\big((x-\mu)/\sigma\big)/\sigma$.
- For $t \in (0,1)$: $F^{-1}_{\mu,\sigma}(t) = \sigma F^{-1}_{0,1}(t) + \mu$ and $F^{-1}_{0,1}(t) = -F^{-1}_{0,1}(1-t)$.
- For $z \in (-\infty, \infty)$: $F_{0,1}(-z) = 1 - F_{0,1}(z)$ and $f_{0,1}(-z) = f_{0,1}(z)$.

Using these properties, it is straightforward to establish the following two formulas, which we shall find particularly useful in the proofs below. Namely, for any $\varepsilon \ge 0$, we have that

$$\int_{-\varepsilon}^{\varepsilon} z^r \quad \mathrm{d}F_{0,1}(z) = \begin{cases} 2\int_0^{\varepsilon} z^r \quad \mathrm{d}F_{0,1}(z)\ , & \text{if } r \ge 0 \text{ is even,} \\ 0\ , & \text{if } r \ge 1 \text{ is odd,} \end{cases} \tag{27}$$

and

$$\int_{-\varepsilon}^{\varepsilon} z^r F_{0,1}(z)\ \mathrm{d}z = \begin{cases} \dfrac{\varepsilon^{r+1}}{r+1}\Big[2F_{0,1}(\varepsilon) - 1\Big] \\ \qquad -\dfrac{2}{r+1} \displaystyle\int_0^{\varepsilon} z^{r+1} \quad \mathrm{d}F_{0,1}(z), & \text{if } r \ge 1 \text{ is odd,} \\ \varepsilon^{r+1}/(r+1), & \text{if } r \ge 0 \text{ is even.} \end{cases} \tag{28}$$

PROOF OF EQUATION (21)

The proof involves two changes of variables, $s = F_{0,1}^{-1}(u)$ and $t = F_{0,1}^{-1}(v)$, an application of formula (28) with $r = 0$ and $r = 1$, and a number of simplifications:

$$\sigma^2_{a_1,a_1}(1) = \int_{a_1}^{1-a_1} \int_{a_1}^{1-a_1} (u \wedge v - uv) \left[\frac{1}{1-2a_1}\right] \left[\frac{1}{1-2a_1}\right] \, \mathrm{d}F_{\mu,\sigma}^{-1}(u) \ \mathrm{d}F_{\mu,\sigma}^{-1}(v)$$

$$= \frac{\sigma^2}{(1-2a_1)^2} \int_{F_{0,1}^{-1}(a_1)}^{F_{0,1}^{-1}(1-a_1)} \int_{F_{0,1}^{-1}(a_1)}^{F_{0,1}^{-1}(1-a_1)} \Big[F_{0,1}(s) \wedge F_{0,1}(t) - F_{0,1}(s) F_{0,1}(t)\Big] \mathrm{d}s \mathrm{d}t$$

$$= \frac{\sigma^2}{(1-2a_1)^2} \Bigg\{ \int_{F_{0,1}^{-1}(a_1)}^{F_{0,1}^{-1}(1-a_1)} \left[ \int_{F_{0,1}^{-1}(a_1)}^{t} F_{0,1}(s) \, \mathrm{d}s + \int_{t}^{F_{0,1}^{-1}(1-a_1)} F_{0,1}(t) \mathrm{d}s \right] \mathrm{d}t$$

$$- \left[ \int_{F_{0,1}^{-1}(a_1)}^{F_{0,1}^{-1}(1-a_1)} F_{0,1}(s) \, \mathrm{d}s \right]^2 \Bigg\}$$

$$= \frac{\sigma^2}{(1-2a_1)^2} \left\{ 2a_1 \left[F_{0,1}^{-1}(a_1)\right]^2 + 2 \int_0^{F_{0,1}^{-1}(1-a_1)} s^2 \, \mathrm{d}F_{0,1}(s) \right\}$$

$$= \frac{\sigma^2}{(1-2a_1)^2} \left\{ 2a_1 \left[F_{0,1}^{-1}(a_1)\right]^2 + \delta_2(F_{0,1}, \ a_1) \right\},$$

This completes the proof of equation (21). □

PROOF OF EQUATION (23)

The proof is based on the changes of variables $z = F_{0,1}^{-1}(u)$ and an application of formula (27) with $r = 2$:

$$\mu_n(2) = \int_{a_2}^{1-a_2} [1] \left( F_{\mu,\sigma}^{-1}(u) - \mu \right)^2 \, \mathrm{d}u$$

$$= \sigma^2 \int_{a_2}^{1-a_2} \left( F_{0,1}^{-1}(u) \right)^2 \, \mathrm{d}u$$

$$= \sigma^2 \left[ 2 \int_0^{F_{0,1}^{-1}(1-a_2)} z^2 \, \mathrm{d}F_{0,1}(z) \right]$$

$$= \sigma^2 \, \delta_2(F_{0,1}, \ a_2).$$

This completes the proof of equation (23). □

PROOF OF EQUATION (24)

The proof involves two changes of variables, $s = F_{0,1}^{-1}(u)$ and $t = F_{0,1}^{-1}(v)$, several applications of formula (28) with $r = 1$ and $r = 3$, and several simplifications:

$$\sigma^2_{a_2,a_2}(2) = \int_{a_2}^{1-a_2}\int_{a_2}^{1-a_2} (u \wedge v - uv)\,[1]\,[1]\,\mathrm{d}\left[F_{\mu,\sigma}^{-1}(u) - \mu\right]^2 \mathrm{d}\left[F_{\mu,\sigma}^{-1}(v) - \mu\right]^2$$

$$= 4\sigma^4 \int_{F_{0,1}^{-1}(a_2)}^{F_{0,1}^{-1}(1-a_2)}\int_{F_{0,1}^{-1}(a_2)}^{F_{0,1}^{-1}(1-a_2)} \left[F_{0,1}(s) \wedge F_{0,1}(t) - F_{0,1}(s)F_{0,1}(t)\right] s\,t\,\mathrm{d}s\,\mathrm{d}t$$

$$= 4\sigma^4 \Bigg\{ \int_{F_{0,1}^{-1}(a_2)}^{F_{0,1}^{-1}(1-a_2)} \left[\int_{F_{0,1}^{-1}(a_2)}^{t} s\,t\,F_{0,1}(s)\,\mathrm{d}s + \int_{t}^{F_{0,1}^{-1}(1-a_2)} s\,t\,F_{0,1}(t)\,\mathrm{d}s\right] \mathrm{d}t$$

$$- \left[\int_{F_{0,1}^{-1}(a_2)}^{F_{0,1}^{-1}(1-a_2)} s\,F_{0,1}(s)\,\mathrm{d}s\right]^2 \Bigg\}$$

$$= \sigma^4 \Bigg\{ 2\int_0^{F_{0,1}^{-1}(1-a_2)} z^4\,\mathrm{d}F_{0,1}(z) - 8a_2\left[F_{0,1}^{-1}(a_2)\right]^2 \int_0^{F_{0,1}^{-1}(1-a_2)} z^2\,\mathrm{d}F_{0,1}(z)$$

$$- \left[2\int_0^{F_{0,1}^{-1}(1-a_2)} z^2\,\mathrm{d}F_{0,1}(z)\right]^2 + 2a_2(1-2a_2)\left[F_{0,1}^{-1}(a_2)\right]^4 \Bigg\}$$

$$= \sigma^4 \Big\{ \delta_4(F_{0,1},\, a_2) - 4a_2\left[F_{0,1}^{-1}(a_2)\right]^2 \delta_2(F_{0,1},\, a_2)$$

$$- \;\delta_2^2(F_{0,1},\, a_2) + 2a_2(1-2a_2)\left[F_{0,1}^{-1}(a_2)\right]^4 \Big\}.$$

This completes the proof of equation (24). □

PROOF OF EQUATION (26)

The following steps are similar to those already used in earlier proof:

$$
\begin{aligned}
&\sigma_{a_1,a_1}(1)\,\sigma_{a_2,a_2}(2)\,\rho^{FRAC}(1,2)\\
&= \int_{a_2}^{1-a_2}\int_{a_1}^{1-a_1}(u\wedge v-uv)\left[\frac{1}{1-2a_1}\right][1]\;\;\mathrm{d}F_{\mu,\sigma}^{-1}(u)\;\;\mathrm{d}\left[F_{\mu,\sigma}^{-1}(v)-\mu\right]^2\\
&= \frac{2\sigma^3}{1-2a_1}\int_{F_{0,1}^{-1}(a_2)}^{F_{0,1}^{-1}(1-a_2)}\int_{F_{0,1}^{-1}(a_1)}^{F_{0,1}^{-1}(1-a_1)}\Big[F_{0,1}(s)\wedge F_{0,1}(t)-F_{0,1}(s)F_{0,1}(t)\Big]t\;\;\mathrm{d}s\;\;\mathrm{d}t\\
&= \frac{2\sigma^3}{1-2a_1}\Bigg\{\int_{F_{0,1}^{-1}(a_2)}^{F_{0,1}^{-1}(1-a_2)}\int_{F_{0,1}^{-1}(a_1)}^{F_{0,1}^{-1}(1-a_1)} t\,F_{0,1}(s)\wedge F_{0,1}(t)\;\;\mathrm{d}s\;\;\mathrm{d}t\\
&-\int_{F_{0,1}^{-1}(a_2)}^{F_{0,1}^{-1}(1-a_2)}\int_{F_{0,1}^{-1}(a_1)}^{F_{0,1}^{-1}(1-a_1)} t\,F_{0,1}(s)\,F_{0,1}(t)\;\;\mathrm{d}s\;\;\mathrm{d}t\Bigg\}\\
&= \frac{2\sigma^3}{1-2a_1}\Big\{0.5F_{0,1}^{-1}(a_1)\delta_2(F_{0,1},\,a_2)-0.5(1-2a_2)F_{0,1}^{-1}(a_1)\big[F_{0,1}^{-1}(a_2)\big]^2\\
&-\Big[0.5F_{0,1}^{-1}(a_1)\delta_2(F_{0,1},\,a_2)-0.5(1-2a_2)F_{0,1}^{-1}(a_1)\big[F_{0,1}^{-1}(a_2)\big]^2\Big]\Big\}\;=\;0.
\end{aligned}
$$

This completes the proof of equation (26). □

ACKNOWLEDGMENTS

The authors are grateful to two anonymous referees for suggestions, queries, and constructive criticism that helped to revise the paper. The second and third authors also gratefully acknowledge their research support by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

REFERENCES

BRAZAUSKAS, V., JONES, B. L., PURI, M. L., and ZITIKIS, R. (2007) Nested *L*-statistics and their use in comparing the riskiness of portfolios, *Scandinavian Actuarial Journal*, to appear.

BRAZAUSKAS, V. and KAISER, T. (2004) Discussion of “Empirical estimation of risk measures and related quantities” by B. L. Jones and R. Zitikis, *North American Actuarial Journal*, 8 (3), 114–117.

GRIBKOVA, N. V. and HELMERS, R. (2006) The empirical Edgeworth expansion for a Studentized trimmed mean, *Mathematical Methods of Statistics*, 15, 61–87.

JONES, B. L., PURI, M. L., and ZITIKIS, R. (2005) Testing hypotheses about the equality of several risk measure values with applications in insurance, *Insurance: Mathematics and Economics*, 38 (2), 253–270.

JONES, B. L. and ZITIKIS, R. (2003) Empirical estimation of risk measures and related quantities, *North American Actuarial Journal*, 7 (4), 44–54.

KAISER, T. and BRAZAUSKAS, V. (2006) Interval estimation of actuarial risk measures, *North American Actuarial Journal*, 10 (4), 249–268.

KLEIBER, C. and KOTZ, S. (2006) *Statistical Size Distributions in Economics and Actuarial Sciences*, Wiley, New York.

MASON, D. R. and SHORACK, G. R. (1990) Necessary and sufficient conditions for asymptotic normality of trimmed $L$-statistics, *Journal of Statistical Planning and Inference*, 25, 111–139.

SERFLING, R. J. (1980) *Approximation Theorems of Mathematical Statistics*, Wiley, New York.

STIGLER, S. M. (1973) The asymptotic distribution of the trimmed mean, *Annals of Statistics*, 1, 472–477.

TARSITANO, A. (2004) A new class of inequality measures based on a ratio of $L$-statistics, *Metron*, 62, 137–160.

WANG, S. (1998) An actuarial index of the right-tail risk, *North American Actuarial Journal*, 2 (2), 88–101.

YOUNG, V. R. (2004) Premium Principles. In: *Encyclopedia of Actuarial Science* (ed. J. L. Teugels and B. Sundt), Wiley, New York.

ZITIKIS, R. (1993) A Berry-Esséen bound for multivariate $L$-estimates with explicit dependence on dimension. In: *Stability Problems for Stochastic Models*, pp. 197–211, Lecture Notes in Mathematics, Vol. 1546, Springer, Berlin.

ZITIKIS, R. (2002) Large sample estimation of a family of economic inequality indices, *Pakistan Journal of Statistics*, (Special Issue in Honour of Dr. S. Ejaz Ahmad) 18, 225–248.

VYTARAS BRAZAUSKAS
Department of Mathematical Sciences
University of Wisconsin-Milwaukee
P.O. Box 413, Milwaukee
Wisconsin 53201, U.S.A.
vytaras@uwm.edu

BRUCE L.JONES
Department of Statistical
and Actuarial Sciences
University of Western Ontario
London, Ontario N6A 5B7, Canada
jones@stats.uwo.ca

RIČARDAS ZITIKIS
Department of Statistical
and Actuarial Sciences
University of Western Ontario
London, Ontario N6A 5B7, Canada
zitikis@stats.uwo.ca