Scandinavian Actuarial Journal, 2007, 3, 162–179

Original Article

Nested *L*-statistics and their use in comparing the riskiness of portfolios

VYTARAS BRAZAUSKAS*, BRUCE L. JONES†, MADAN L. PURI‡ and RIČARDAS ZITIKIS†

*Department of Mathematical Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA
†Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Ontario N6A 5B7, Canada
‡Department of Mathematics, Indiana University, Bloomington, IN 47405, USA

(*Accepted 10 April 2007*)

Inspired by the problem of testing hypotheses about the equality of several risk measure values, we find that the 'nested *L*-statistic'—a notion introduced herein—is natural and particularly convenient. Indeed, the test statistic that we explore in this paper is a nested *L*-statistic. We discuss large-sample properties of the statistic, investigate its performance using a simulation study, and consider an example involving the comparison of risk measure values where the risks of interest are those associated with tornado damage in different time periods and different regions.

Keywords: Asymptotic distribution; conditional tail expectation; hypothesis testing; insurance losses; proportional hazards transform; risk measure

1. Introduction

In valuing a portfolio of risks, one often uses a *risk measure* which captures the riskiness associated with the portfolio. Formally, a risk measure is a functional mapping from the set of all distribution functions to the set of extended real numbers. There exist infinitely many such functional mappings, and the choice of a suitable risk measure is a subjective decision that depends on one's intuition regarding the nature of the risk and one's preference.

In constructing nonparametric estimates of risk measure values, $R = R[F]$, based on a sample of size n, where F is the distribution function of the underlying risk, it is often appropriate to replace F by its empirical estimator $\hat{F}$, so that $\hat{R} = R[\hat{F}]$. Sometimes a modification of the functional R is appropriate and useful. For example, we can often increase robustness by suitably choosing the functional R to depend on the sample size. There are, of course, situations where it is appropriate to use parametric estimators.

Corresponding author: Bruce L. Jones, Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Ontario, Canada, N6A 5B7. E-mail: jones@stats.uwo.ca

Scandinavian Actuarial Journal
ISSN 0346-1238 print/ISSN 1651-2030 online
http://www.tandf.no/saj
DOI: 10.1080/03461230701390287

However, in this paper, we shall concentrate on the nonparametric approach, which is justified in applications such as automobile insurance where data sets are sufficiently large and the underlying distributions are unknown.

In many cases, nonparametric estimators that are constructed as described above are L-statistics. That is, they are linear combinations of the ordered sample outcomes, which may, for example, be asset returns or insurance losses. Jones and Zitikis (2003) employed this representation in studying estimators of risk measures and their asymptotic properties. As we shall see below the well-known Gini index (cf. Eq. (1.1) below), which has been used for measuring income inequality for almost a century, can also be estimated by an L-statistic (cf. Eq. (1.2) below).

Portfolios of interest often contain natural groupings of risks. For example, a portfolio of automobile insurance policies may include policies from several different geographic regions, and a portfolio of assets may comprise several classes of assets. In other situations, temporal groupings may be of interest. When groups of risks arise, it is important to understand how the risk measure value varies across groups. In an insurance context, this is essential in establishing appropriate rating classes. In an investment context, differences among risk measure values are important in portfolio selection. It is therefore of practical importance to study hypothesis tests associated with the relative values of the risk measure for several groups.

Relationships among risk measure values were explored by Wang and Young (1998) and Wirch and Hardy (2000). Jones and Zitikis (2005) considered nonparametric and parametric tests for the order of two risk measure values. Nonparametric tests of hypotheses about the equality of several risk measures were provided by Jones, Puri and Zitikis (2006). These authors considered null hypotheses involving both known and unknown (but equal) risk measure values. In the latter case, alternative hypotheses involving ordered risk measure values were investigated. In the present paper, we also consider nonparametric tests about the equality of several risk measure values. However, the alternative hypothesis involves no order specification. The test statistic is based on the Gini index (cf. Eq. (1.1) below) applied to risk measure estimators, and therefore results in a nested L-statistic. We next define the problem formally.

Let $R_1, \ldots, R_k$ be risk measure values corresponding to k populations with distribution functions $F_1, \ldots, F_k$, respectively. The populations can be dependent or independent. In the current paper we concentrate on the independent case, but for those who may wish to apply the herein developed method for dependent populations, we offer a hint in Note 2.3.

The risk measure values $R_1, \ldots, R_k$ are obtained using a risk measure, R, which is common to all k populations. The risk measure could, for example, be the mean, a distorted mean, or some other functional. There may be many hypotheses of interest. In this paper, however, we focus our interest on whether or not the k risk measures are all equal. That is, we wish to test the hypothesis

$$H_0 : R_1 = \cdots = R_k, \quad \text{against}$$

$$H_1 : \exists\ (i,j) \text{ such that } R_i \neq R_j.$$

With the help of the parameter

$$\gamma := \frac{1}{k^2} \sum_{1 \le i,j \le k} |R_i - R_j|, \tag{1.1}$$

which is the Gini index (Gini, 1914) of the risk measure values $R_1, \ldots, R_k$, we can reformulate the hypotheses as follows:

$$H_0 : \gamma = 0,$$
$$H_1 : \gamma > 0.$$

It is interesting to note that the aforementioned parameter γ can be expressed as (David, 1970)

$$\begin{aligned}\gamma &= \frac{1}{k^2} \sum_{i=1}^{k} (4i - 2(k+1)) R_{i:k} \\ &= \sum_{i=1}^{k} \left(\int_{(i-1)/k}^{i/k} K(u) du \right) R_{i:k}, \end{aligned} \tag{1.2}$$

where

$$K(u) := 4u - 2, \tag{1.3}$$

and $R_{1:k}, \ldots, R_{k:k}$ are the k ordered risk measure values. We shall use the representation (1.2) later in the paper.

To test the hypothesis H_0, we construct an empirical estimator for γ and then establish its asymptotic distribution when the sample size increases. A natural way to introduce such an estimator is

$$\hat{\gamma} := \frac{1}{k^2} \sum_{1 \le i,j \le k} |\hat{R}_i - \hat{R}_j|,$$

where $\hat{R}_i$ denotes an empirical estimator of R_i. We construct the estimator $\hat{R}_i$ as follows. First, we assume that, for all $1 \le i \le k$, the risk measure values R_i can be expressed by the equation $R_i = R[F_i]$, where $R[\cdot]$ is a functional defined on a set of distribution functions F by the equation

$$R[F] = \int_0^1 F^{-1}(u) J(u) du, \tag{1.4}$$

where the function J is such that the integral in Eq. (1.4) is finite for the set of distribution functions F under consideration. Many risk measures can be expressed in this form, and we therefore do not consider representation (1.4) unnecessarily restrictive (cf. Jones and Zitikis, 2003, Brazauskas *et al.*, 2007, and references therein). The risk measure $R[F]$ is also called 'spectral risk measure' and the function J a 'risk aversion function'; we refer to Acerbi (2002) for more detail on the subject as well as for further references. The following are examples of such risk measures.

Example 1.1 (MEAN). With $J(u) = 1$ for all $0 \le u \le 1$, Eq. (1.4) gives the mean of the distribution F. Instead of the traditional μ or μ_F, we shall frequently use the notation MEAN[F], which makes the presentation of risk measures consistent throughout the paper.

Example 1.2 (PHT, proportional hazards transform). Let $r > 0$ be a fixed number, chosen by an actuary and called the *distortion level*. With the function $J(u) = r(1-u)^{r-1}$ for all $0 \le u \le 1$, Eq. (1.4) gives the PH transform (cf. Wang, 1995; Jones and Zitikis, 2003). The values of r of practical interest are $0 < r \le 1$. We shall frequently denote the PHT risk measure value by PHT[F], or by PHT[F, r] if we need to specify the distortion level r.

Example 1.3 (CTE, conditional tail expectation). Let $0 \le t < 1$ be a fixed number, chosen by an actuary, and called the *threshold level*. With the function $J(u) = 0$ for all $0 \le u < t$, and $J(u) = 1/(1-t)$ for all $t \le u \le 1$, Eq. (1.4) gives the 100t% CTE. We shall frequently denote the CTE risk measure value by CTE[F] or by CTE[F, t] if we need to specify the threshold level t. Finally, we note that the CTE risk measure is also called Tail Conditional Expectation (TCE), Conditional Value-at-Risk (CVaR), Expected Shortfall (ES); we refer to Artzner *et al.* (1999), Acerbi (2002), Acerbi and Tasche (2002) for extensive discussions of this risk measure and further references on the subject.

Note 1.1. The risk functional R in Eq. (1.4) can be further generalized into the following one:

$$R[F; \Psi] = \int_0^1 F^{-1}(u) d\Psi(u),$$

where $\Psi(u)$ is a function of bounded variation. We first note that if $\Psi(u)$ is differentiable with the first derivative $J(u)$, then $R[F; \Psi]$ equals $R[F]$ given by Eq. (1.4). If, however, $\Psi(u)$ is the indicator function, that is, $\Psi(u) = 0$ for all $u \in (0, \alpha)$ and $\Psi(u) = 1$ for all $u \in [\alpha, 1]$, where $\alpha \in (0,1)$ is a fixed parameter, then $R[F; \Psi] = F^{-1}(\alpha)$, which is the α-th quantile or, in other words, Value-at-Risk (VaR) (cf. Artzner *et al.*, 1999). The latter risk measure (i.e. VaR) is not covered by the results of the present paper, as assumptions (A1)–(A3) to be formulated in the next section are not satisfied for indicator functions $\Psi(u)$. This is natural as the smoothing character of the integration operator disappears in this case, and thus quite different mathematical forces start acting. (This is convincingly seen by comparing the classical results on estimating the mean $\mu = \int_0^1 F^{-1}(u) du$ and the quantile $F^{-1}(\alpha)$.)

Now we are in the position to define the estimator $\hat{R}_i$ by the equation

$$\hat{R}_i := \int_0^1 \hat{F}_i^{-1}(u) J(u) du,$$

where $\hat{F}_i$ is the empirical distribution function based on the sample $X_1(i), \ldots, X_{n_i}(i)$ of size n_i drawn from the population with the distribution function F_i, and $\hat{F}_i^{-1}$ is the corresponding quantile function. We assume that, for every fixed $1 \le i \le k$, the random variables $X_1(i), \ldots, X_{n_i}(i)$ are independent and identically distributed. As to whether the vectors

$$\boldsymbol{X}(i) := (X_1(i), \ldots, X_{n_i}(i)), \quad 1 \le i \le k,$$

are independent or not depends on the problem considered. Our focus in this paper is on independent vectors.

We now reformulate the above quantities so that the main object of the present paper – nested L-statistics – would emerge in a most natural way. First we write the expression

$$\hat{R}_i = \sum_{m=1}^{n_i} \left(\int_{(m-1)/n_i}^{m/n_i} J(u)du \right) X_{m:n_i}(i), \tag{1.5}$$

where $X_{1:n_i}(i) \le \cdots \le X_{n_i:n_i}(i)$ are the ordered observations from the population i. Hence, $\hat{R}_i$ is a linear combination of order statistics or, in other words, an L-statistic, based on the random variables $X_1(i), \ldots, X_{n_i}(i)$.

Following Eq. (1.2) for γ, we introduce an estimator for γ by the equation

$$\hat{\gamma} := \sum_{i=1}^{k} \left(\int_{(i-1)/k}^{i/k} K(u)du \right) \hat{R}_{i:k}. \tag{1.6}$$

Since $\hat{R}_1, \ldots, \hat{R}_k$ are L-statistics, as we have shown above, the estimator $\hat{\gamma}$ can naturally be called a 'nested L-statistic'. With the notation

$$L[\psi, \boldsymbol{Y}] := L[\psi, Y_1, \ldots, Y_k] := \sum_{i=1}^{k} \left(\int_{(i-1)/k}^{i/k} \psi(u)du \right) Y_{i:k},$$

where $\boldsymbol{Y} := (Y_1, \ldots, Y_k)$, we express $\hat{\gamma}$ as follows

$$\hat{\gamma} = L[K,\ L[J, \boldsymbol{X}(1)], \ldots, L[J, \boldsymbol{X}(k)]]. \tag{1.7}$$

The right-hand side of Eq. (1.7) justifies the herein introduced name 'nested L-statistic' for the estimator $\hat{\gamma}$. In the special case $J(u) = 1$ for all $0 < u < 1$, we have the equality $\hat{\gamma} = L[K, \bar{\boldsymbol{X}}(1), \ldots, \bar{\boldsymbol{X}}(k)]$, where $\bar{\boldsymbol{X}}(i)$ is the sample mean of the random variables $X_1(i), \ldots, X_{n_i}(i)$, which are the coordinates of $\boldsymbol{X}(i)$.

The rest of the paper is organized as follows. The next section, Section 2, presents large-sample statistical inferential results for the nested L-statistics introduced above. In Section 3 we describe a simulation study which illustrates the performance of our theoretical results. An example of practical relevance is explored in Section 4. Concluding remarks are offered in Section 5.

2. Asymptotic results

Statistical inference dealing with the problem stated in the previous section depends on two accomplishments concerning the test statistic

$$T := \frac{\hat{\gamma}}{\sqrt{\sum_{m=1}^{k} n_m^{-1}}}.$$

They are:

- Establishing the asymptotic distribution of T under the null hypothesis H_0, so that critical values would be possible to obtain.
- Showing that, under the alternative H_1, the test statistic T converges to infinity when the sample sizes n_m increase, in which case the asymptotic power is 1.

Using the notation $D_i := \hat{R}_i - R_i$, we obtain under the null hypothesis H_0 that

$$\begin{aligned}\hat{\gamma} &= \frac{1}{k^2} \sum_{1 \le i,j \le k} |D_i - D_j| \\ &= \frac{1}{k^2} \sum_{i=1}^{k} (4i - 2(k+1)) D_{i:k},\end{aligned} \tag{2.1}$$

where $D_{1:k} \le \cdots \le D_{k:k}$ are the ordered values of $D_1, \ldots, D_k$. Continuing with Eq. (2.1), we have

$$T = \frac{1}{k^2} \sum_{i=1}^{k} (4i - 2(k+1)) \Delta_{i:k}, \tag{2.2}$$

where $\Delta_{1:k} \le \cdots \le \Delta_{k:k}$ are the ordered values of

$$\Delta_i := \Theta_i(n_1, \ldots, n_k) \sqrt{n_i} (\hat{R}_i - R_i), \quad i = 1, \ldots, k,$$

with the notation $\Theta_i(n_1, \ldots, n_k) = \sqrt{n_i^{-1} / \Sigma_{m=1}^{k} n_m^{-1}}$. It is now natural to make an assumption about the 'comparability of sample sizes,' which requires that there are constants $\theta_i \in (0, 1)$ such that $\Sigma_{i=1}^{k} \theta_i^2 = 1$ and the sample sizes n_i tend to infinity in such a way that, for every $1 \le i \le k$,

$$\Theta_i(n_1, \ldots, n_k) \to \theta_i,$$

when $\min_i(n_i) \to \infty$. With the function $\mathcal{T}_k : \mathbf{R}^k \to \mathbf{R}^k$ defined by

$$\mathcal{T}_k(x_1, \ldots, x_k) := \frac{1}{k^2} \sum_{i=1}^{k} (4i - 2(k+1)) x_{i:k},$$

we rewrite Eq. (2.2) as follows:

$$T = \mathcal{T}_k(\Delta_1, \ldots, \Delta_k). \tag{2.3}$$

In view of the asymptotic result, when $n_i \to \infty$,

$$\sqrt{n_i}(\hat{R}_i - R_i) \to_d \sigma_i G_i, \tag{2.4}$$

whose validity is discussed in the next paragraph, we have that $\Delta_i \to_d \theta_i \sigma_i G_i$. Hence, with the help of Eq. (2.3) we obtain that, when $\min_i(n_i) \to \infty$,

$$T \to_d \mathcal{T}_k(\theta_1 \sigma_1 G_1, \ldots, \theta_k \sigma_k G_k), \tag{2.5}$$

where, under the assumption of independent populations, $G_1, \ldots, G_k$ are *independent* standard normal random variables. Given statement (2.5), we can test hypotheses about, or construct confidence intervals for, the parameter γ. To make this result readily applicable in practice, we shall later discuss a bootstrap technique for calculating critical values of the test.

The validity of statement (2.4) is thoroughly discussed in the literature (cf. Serfling, 1980, Chapter 8; Shorack and Wellner, 1986, Chapter 19). We know from these references that the result holds under the following assumptions:

(A1) The weight function J is continuous on the interval (0, 1), except possibly at a finite number of points at which F_i^{-1} is continuous (this condition can, if desired, be relaxed; cf. Assumption 2 on page 664 in Shorack and Wellner, 1986).

(A2) There exist α, $\beta > 1/2$ and $c < \infty$ such that $|J(t)| \leq c\, t^{\alpha-1}(1-t)^{\beta-1}$ on the interval (0,1).

(A3) The moment $\mathbf{E}[|X_1(i)|^{\gamma}]$ is finite for some γ such that $\gamma > 1/(\alpha - 1/2)$ and $\gamma > 1/(\beta - 1/2)$.

In the case of the PHT risk measure, the above three assumptions (A1)–(A3) hold provided that $r \in (1/2, 1]$ and the moment $\mathbf{E}[X^p]$ is finite for some $p > 1/(r - 1/2)$. When $r = 1$, we have the MEAN risk measure and need to assume that $\mathbf{E}[X^p] < \infty$ for some $p > 2$, even though the classical CLT requires only $p = 2$. However, this would be an improvement of no particular significance in the context of the present paper. As to the CTE risk measure, we need to assume that the quantile function $F^{-1}(t)$ is continuous at the point $t = \alpha$ and that the moment $\mathbf{E}[X^p]$ is finite for some $p > 2$. We have relegated more detailed thoughts on moment conditions to Note 4 at the end of the current section.

Now we shall prove that under the alternative H_1 the test statistic T converges to infinity when $\min_i(n_i) \to \infty$. We proceed as follows:

$$T = \frac{1}{k^2\sqrt{\sum_{m=1}^{k} n_m^{-1}}} \sum_{1 \leq i,j \leq k} |(D_i - D_j) + (R_i - R_j)|$$
$$\geq -\frac{1}{k^2\sqrt{\sum_{m=1}^{k} n_m^{-1}}} \sum_{1 \leq i,j \leq k} |D_i - D_j| + \frac{1}{k^2\sqrt{\sum_{m=1}^{k} n_m^{-1}}} \sum_{1 \leq i,j \leq k} |R_i - R_j| \qquad (2.6)$$

We already know that the first summand (with 'minus' in front of it) on the right-hand side of bound (2.6) has a non-degenerate distribution. The second summand converges to infinity, since $\sqrt{\Sigma_{m=1}^{k} n_m^{-1}}$ converges to 0 and the sum $\Sigma_{1 \leq i,j \leq k}|R_i - R_j|$ is strictly positive under the currently assumed H_1.

We suggest the following bootstrap approximation for calculating critical values of the test. For every $1 \leq i \leq k$, we sample with replacement from the set of random variables $X_1(i), \ldots, X_{n_i}(i)$ and obtain n_i new ones, which we denote by $X_1^*(i), \ldots, X_{n_i}^*(i)$. Using the latter variables, for every $1 \leq i \leq k$ we calculate the risk measure value $\hat{R}_i^*$, which is defined using formula (1.5) with $X_{m:n_i}(i)$ replaced by $X_{m:n_i}^*(i)$. Denote

$$\hat{\gamma}^* := \frac{1}{k^2} \sum_{i=1}^{k} (4i - 2(k+1)) D_{i:k}^*,$$

where $D^*_{1:k} \le \cdots \le D^*_{k:k}$ are the ordered values of $D^*_i := \hat{R}^*_i - \hat{R}_i$, $i = 1, \ldots, k$. With the above notation, we define the bootstrapped version of the test statistic T as follows:

$$T^* := \frac{\hat{\gamma}^*}{\sqrt{\sum_{m=1}^{k} n_m^{-1}}}.$$

We repeat the above resampling procedure B times and in this way obtain B values of T^*. Then we calculate the $100(1-\alpha)$ percentile of these values of T^*. Denote the percentile by $x_\alpha[T^*]$, which is, by definition, the critical value of the test. Specifically, the decision rule is as follows: we reject the null hypothesis H_0 in favour of the alternative H_1 if the value of the test statistic T exceeds $x_\alpha[T^*]$; otherwise, we retain H_0.

Note 2.1. From the practical (and pragmatic) point of view, it makes sense to reformulate the above decision rule in terms of $\hat{\gamma}$ and $\hat{\gamma}^*$. Namely, after resampling B times we obtain B values of $\hat{\gamma}^*$. Then we calculate the $100(1-\alpha)$ percentile of these values; denote the percentile by $x_\alpha[\hat{\gamma}^*]$. The decision rule is as follows: we reject the null hypothesis H_0 in favour of the alternative H_1 if the value of $\hat{\gamma}$ exceeds $x_\alpha[\hat{\gamma}^*]$; otherwise, we retain H_0. We adopt this approach for testing H_0 against H_1 in Section 4 below.

Note 2.2. The standard deviation σ_i in statement (2.4) is given by the equation (cf., e.g., Serfling, 1980; Helmers, 1982; Shorack and Wellner, 1986)

$$\sigma_i^2 := \int\int (F_i(x \wedge y) - F_i(x)F_i(y))J(F_i(x))J(F_i(y))dxdy,$$

where $x \wedge y$ denotes the minimum of x and y. A non-parametric estimator of the variance σ_i^2 can easily be constructed by replacing the theoretical distribution functions F_i by their empirical counterparts, which leads to the estimator

$$s_i^2 := \sum_{1 \le k \le n_i - 1} \sum_{1 \le m \le n_i - 1} (X_{k+1:n_i}(i) - X_{k:n_i}(i))(X_{m+1:n_i}(i) - X_{m:n_i}(i)) \times \left(\frac{k \wedge m}{n_i} - \frac{k}{n_i}\frac{m}{n_i}\right) J\left(\frac{k}{n_i}\right) J\left(\frac{m}{n_i}\right).$$

Note 2.3. It is well known (cf., e.g., Shorack and Wellner, 1986, Chapter 19) that, under the three assumptions (A1)–(A3), we have the representation

$$\sqrt{n_i}\,(\hat{R}_i - R_i) = \frac{1}{\sqrt{n_i}} \sum_{m=1}^{n_i} A_i(X_m(i)) + o_{\mathbf{P}}(1) \tag{2.7}$$

when $n_i \to \infty$, where

$$A_i(y) := -\int_{-\infty}^{\infty} (1\{y \le x\} - F_i(x))J(F_i(x))dx.$$

We easily check that the variance of $A_i(X_m(i))$ is equal to the above noted σ_i^2. From representation (2.7) we therefore arrive at statement (2.4). In an obvious way, representation (2.7) can also be used for deriving the asymptotic distribution of the test statistic T when the vectors $\boldsymbol{X}(i) = (X_1(i), \ldots, X_{n_i}(i))$, $1 \le i \le k$, are dependent. In this case, the limiting distribution is of the same form as that on the right-hand side of statement (2.5), but now with the standard normal random variables $G_1, \ldots, G_k$ being *dependent*, with the entries of the covariance matrix being equal to $\sigma_i^{-1}\sigma_j^{-1}\mathbf{cov}(A_i(X_1(i)), A_j(X_1(j)))$.

Note 2.4. The asymptotic results and, consequently, statistical tests discussed in this section require at least two finite moments, if not more, as in the case of the PHT risk measure, which depends on the distortion level r. The moment requirements are natural since the very nature of asymptotically normal distributions (that we aim at) does not allow the observations to move towards infinity with too large a probability. If it is not reasonable to assume that the first two moments are finite, then we suggest using a parametric model with a heavy-tail and developing the corresponding parametric inferential tools. This, however, does not remove the problem, as mle's are still desired to be asymptotically normal, for which a moment-type assumption is needed. Employing asymptotically stable distributions, instead of asymptotically normal used above, provides a further venue for relaxing moment assumptions.

3. Simulation study

In this section we augment the asymptotic results of Section 2 with finite-sample performance investigations. We have two objectives. The first one is to see how large the sample size $n = \min\{n_1, \ldots, n_k\}$ is needed for the proposed (asymptotic) test to attain the nominal level of significance, and the second is to estimate the power of the test against selected types of alternatives, for various n. Using Monte Carlo simulations, we generate three portfolios (of insurance losses) that are either equally risky (H_0 setting) or unequally risky (H_1 setting) according to a fixed risk measure, then perform the test and compute its proportion of rejections. (Such a proportion estimates the nominal level of significance under H_0 and the power of the test under H_1.) Specifically, for generation of insurance portfolios, we follow the similar work done by Brazauskas and Kaiser (2004), Brazauskas *et al.* (2007), and Kaiser and Brazauskas (2006), and choose the following three parametric families:

- *Exponential* with the cdf

$$F_1(x) = 1 - e^{-(x-x_0)/\theta}, \quad x > x_0, \;\; \theta > 0.$$

- *Pareto* with the cdf

$$F_2(x) = 1 - (x_0/x)^{\beta}, \quad x > x_0, \;\; \beta > 0.$$

- *Lognormal* with the cdf

$$F_3(x) = \Phi(\log(x - x_0) - \mu), \quad x > x_0, \; -\infty < \mu < \infty,$$

where $\Phi(\cdot)$ denotes the standard normal cdf.

The parameter x_0 in the above distributions can be interpreted as a deductible or a retention level. (Note that, due to x_0, the distributions F_1, F_2, and F_3 have the same support.) The remaining parameters θ, β, and μ are selected so that the families F_1, F_2, F_3 follow the hypothesized scenario with respect to a fixed risk measure. In particular, under H_0, they are equally risky and therefore satisfy the equation

$$R[F_1] \;=\; R[F_2] \;=\; R[F_3], \tag{3.1}$$

where $R[\cdot]$ represents any of the three risk measures presented in Examples 1.1–1.3 (i.e., MEAN, PHT, CTE). Evaluation of these measures for the distributions F_1, F_2, F_3 yields the following expressions of Eq. (3.1).

- For the MEAN measure:

$$x_0 + \theta = \frac{x_0 \beta}{\beta - 1} = x_0 + e^{\mu + 0.5}, \tag{3.2}$$

- For the PHT measure:

$$x_0 + \frac{\theta}{r} = x_0 + \frac{x_0}{\beta r - 1} = x_0 + C_r \, e^{\mu}, \tag{3.3}$$

where, for fixed r, the integral $C_r = \int_{-\infty}^{\infty} [1 - \Phi(z)]^r e^z \; dz$ is found numerically. For example, as reported by Brazauskas and Kaiser (2004), $C_{0.55} = 3.896$, $C_{0.70} = 2.665$, $C_{0.85} = 2.030$, $C_{0.95} = 1.758$. Note that when $r = 1$, then the PHT measure becomes the MEAN.

- For the CTE measure:

$$x_0 - \theta(\log(1 - t) - 1) = \frac{x_0 \beta}{\beta - 1}(1 - t)^{-1/\beta} = x_0 + \frac{1}{1 - t} e^{\mu + 0.5} \; \Phi(1 - \Phi^{-1}(t)). \tag{3.4}$$

Note that when $t = 0$, then the CTE measure becomes the MEAN.

Under H_1, the families are unequally risky, and we consider two types of alternatives:

- Two portfolios are equally risky but the third one differs; that is,

$$R[F_1^{\star}] = c_{\star} R[F_1], \quad R[F_2^{\star}] = R[F_2], \quad R[F_3^{\star}] = R[F_3], \tag{3.5}$$

where $c_{\star} \neq 1$ and $R[F_1] = R[F_2] = R[F_3]$.

- Relative riskiness of all three portfolios is equally-spaced; that is,

$$R[F_1^{\star\star}] = c_{\star\star} R[F_1], \quad R[F_2^{\star\star}] = R[F_2], \quad R[F_3^{\star\star}] = c_{\star\star}^2 R[F_3], \tag{3.6}$$

where $c_{\star\star} > 1$ and $R[F_1] = R[F_2] = R[F_3]$.

Note that the choices $c_{\star} = 1$ and $c_{\star\star} = 1$ reduce the H_1 scenario to that of H_0.

We use the following design for the simulation study. For a fixed risk measure and a fixed scenario of riskiness, we generate three *independent* samples of size $n\,(= n_1 = n_2 = n_3)$

from the distributions F_1, F_2, and F_3, respectively. These samples are then resampled according to the bootstrap method described in the previous section, an α-level test is performed, and its decision—reject the null hypothesis or not—is recorded. This procedure is repeated 5000 times, for each of the three risk measures and for each of the hypothesized scenarios. Using the recorded 5000 decisions for the tests based on the MEAN, PHT, and CTE measures, respectively, we estimate the proportion, say $\hat{p}$, of test's rejections. The standard error of such an estimate is then evaluated using the formula $\sqrt{\hat{p}}(1-\hat{p})/5000$. The study was performed for the following choices of simulation parameters:

- *Level of significance*: $\alpha=0.01$, 0.05, 0.10.
- *Sample size*: $n=25$, 50, 100, 200.
- *Number of bootstrap samples*: $B=1000$.
- *Measure-related parameters*:
  - *Distortion level* (for PHT measure): $r=0.85$.
  - *Threshold level* (for CTE measure): $t=0.75$.
- *Distribution-related parameters, under H_0* (derived from Eqs (3.2)–(3.4)):
  - $1.222=\text{MEAN}[F_1]=\text{MEAN}[F_2]=\text{MEAN}[F_3]$,

    where
    * F_1 is exponential with $x_0=1$, $\theta=0.222$,
    * F_2 is Pareto with $x_0=1$, $\beta=5.5$,
    * F_3 is lognormal with $x_0=1$, $\mu=-2.004$, $\sigma=1$.
  - $1.272=\text{PHT}[F_1,\ r=0.85]=\text{PHT}[F_2,\ r=0.85]=\text{PHT}[F_3,\ r=0.85]$,

    where
    * F_1 is exponential with $x_0=1$, $\theta=0.231$
    * F_2 is Pareto with $x_0=1$, $\beta=5.5$,
    * F_3 is lognormal with $x_0=1$, $\mu=-2.010$, $\sigma=1$.
  - $2.107=\text{CTE}[F_1,\ t=0.75]=\text{CTE}[F_2,\ t=0.75]=\text{CTE}[F_3,\ t=0.75]$,

    where
    * F_1 is exponential with $x_0=1$, $\theta=0.277$,
    * F_2 is Pareto with $x_0=1$, $\beta=5.5$,
    * F_3 is lognormal with $x_0=1$, $\mu=-2.044$, $\sigma=1$.

- *Distribution-related parameters, under H_1* (derived from Eqs (3.5)–(3.6)):
  - H_1 specified by Eq. (3.5): $F_2^\star=F_2$, $F_3^\star=F_3$ and
    * For MEAN: $F_1^\star$ is exponential with $x_0=1$, $\theta^\star=x_0(c_\star-1)+c_\star\theta$,
    * For PHT: $F_1^\star$ is exponential with $x_0=1$, $\theta^\star=x_0 r(c_\star-1)+c_\star\theta$,
    * For CTE: $F_1^\star$ is exponential with $x_0=1$, $\theta^\star=\dfrac{x_0(c_\star-1)}{1-\log(1-t)}+c_\star\theta$,

    where θ as under H_0 and $c_\star=0.85$, 0.90, 0.95, 1.05, 1.10, 1.15, 1.25, 1.50, 2.00.
  - H_1 specified by Eq. (3.6): $F_2^{\star\star}=F_2$ and
    * For MEAN: $F_1^{\star\star}$ is exponential with $x_0=1$, $\theta^{\star\star}=x_0(c_{\star\star}-1)+c_{\star\star}\theta$, and $F_3^{\star\star}$ is lognormal with $x_0=1$, $\mu^{\star\star}=\log(x_0(c_{\star\star}^2-1)+c_{\star\star}^2e^{\mu+0.5})-0.5$, $\sigma=1$,

* For PHT: $F_1^{\star\star}$ is exponential with $x_0 = 1$, $\theta^{\star\star} = x_0 r(c_{\star\star} - 1) + c_{\star\star}\theta$, and $F_3^{\star\star}$ is lognormal with $x_0 = 1$, $\mu^{\star\star} = \log\Big(\frac{x_0(c_{\star\star}^2 - 1)}{C_r} + c_{\star\star}^2 e^{\mu}\Big)$, $\sigma = 1$,

* For CTE: $F_1^{\star\star}$ is exponential with $x_0 = 1$, $\theta^{\star\star} = \frac{x_0(c_{\star\star} - 1)}{1 - \log(1 - t)} + c_{\star\star}\theta$, and $F_3^{\star\star}$ is lognormal with $x_0 = 1$, $\mu^{\star\star} = \log\Big(\frac{(1 - t)x_0(c_{\star\star}^2 - 1)}{\Phi(1 - \Phi^{-1}(t))} + c_{\star\star}^2 e^{\mu + 0.5}\Big) - 0.5$, $\sigma = 1$,

where θ and μ as under H and $c_{\star\star} = 1.05 : (0.05) : 1.25,\ \ 1.50 : (0.50) : 3.00$.

Our simulation results are summarized in Table 1 with probabilities of type *I* error, as well as in Tables A.1 and A.2 with estimated power function values. Specifically, we notice in Table 1 that the convergence of estimated probability of the type *I* error depends on the underlying risk measure. For instance, if the risk measure is 'light' (such as the MEAN), then the nominal level of significance is attained even with sample sizes as small as $n = 25$. However, for 'heavier' risk measures (such as the PHT and CTE), the sample size $n = 100$ may still be not large enough. In summary, for the risk measures considered in our study, the results for $n = 200$ are satisfactory at all levels of α.

Tables A.1 and A.2 provide power estimates against the two types of alternatives described above, for various choices of α and n. (As a quick reference, we repeated, though with lesser accuracy, the entries of Table 1 in these two tables; they correspond to the cases $c_{\star} = 1$ and $c_{\star\star} = 1$.) Similar to the type *I* error investigations, we notice that the power of the test depends on 'heaviness' of the underlying risk measure. That is, all things being equal, the test is more powerful for the 'light' measure than for the 'heavy' one. Other patterns of estimated power function values agree with the general behavior of any reasonable statistical test. Specifically, for a fixed alternative, i.e., fixed $c_{\star}$ or $c_{\star\star}$, and fixed n, the power increases (decreases) as α increases (decreases); for fixed alternative and α, the power changes in unison with n; and for fixed α and n, the test becomes more powerful as $c_{\star}$ ($c_{\star\star}$) moves further away from $c_{\star} = 1$ ($c_{\star\star} = 1$), i.e., when data go further into the alternative. Also, comparison of the two types of alternatives reveals that the test is more powerful against the second type of alternatives, which can be anticipated because under the second scenario the differences in portfolio riskiness are more pronounced.

Table 1. Estimated probabilities of the type *I* error of the tests based on the MEAN, PHT, CTE measures, for selected n and α. Standard errors for the entries are presented in parentheses.

α	*Risk measure*	$n = 25$	$n = 50$	$n = 100$	$n = 200$
0.01	MEAN	0.009 (.001)	0.009 (.001)	0.008 (.001)	0.009 (.001)
	PHT [$r = 0.85$]	0.014 (.002)	0.011 (.001)	0.014 (.002)	0.012 (.002)
	CTE [$t = 0.75$]	0.022 (.002)	0.011 (.001)	0.014 (.002)	0.011 (.002)
0.05	MEAN	0.051 (.003)	0.053 (.003)	0.049 (.003)	0.052 (.003)
	PHT [$r = 0.85$]	0.072 (.004)	0.067 (.004)	0.067 (.004)	0.059 (.003)
	CTE [$t = 0.75$]	0.083 (.004)	0.063 (.003)	0.060 (.003)	0.051 (.003)
0.10	MEAN	0.113 (.004)	0.114 (.004)	0.104 (.004)	0.104 (.004)
	PHT [$r = 0.85$]	0.144 (.005)	0.136 (.005)	0.134 (.005)	0.119 (.005)
	CTE [$t = 0.75$]	0.153 (.005)	0.123 (.005)	0.111 (.004)	0.105 (.004)

Finally, we conclude that, except for some extreme situations (related to CTE and small α), the test will successfully detect, with the probability substantially above 0.50, the differences in portfolio riskiness of at least 15% (corresponding to $c_{\star} \leq 0.85$ or $c_{\star} \geq 1.15$, and $c_{\star\star} \geq 1.15$) for portfolios of $n \geq 100$ insurance losses.

4. Example

To illustrate the methods presented in this paper, we consider data from Brooks and Doswell (2000), providing the damage from 137 major tornadoes in the United States from 1890 to 1999. The possibility of tornado damage creates an important risk for insurers. The need to understand this and other catastrophic risks is stated very well by Brooks and Doswell: "A major challenge to preparedness and recovery is maintaining a level of readiness during the gaps. Catastrophic events pose significant threats to the insurance and reinsurance industries. Thus, accurate estimates of the threat are important for long-term planning both in the private and public sectors."

Brooks and Doswell (2000) argue that, in order to compare tornado losses over time, it is appropriate to adjust for inflation and wealth. We therefore normalized the damage amounts by dividing each by the nominal GDP per capita in the year of occurrence and scaling the resulting values so that the smallest is 100. The GDP estimates were obtained from Johnston and Williamson (2005). The resulting damage amounts are shown in Table A.3 along with the year and census region of occurrence. The test described in this paper was used to investigate whether or not the normalized damage amounts differ by time period and by census region.

Nonparametric estimates of risk measure values were obtained for each of three time periods using each of the following three different risk measures: the MEAN, the PHT with $r = 0.85$, and the CTE with $t = 0.75$. Specifically, the estimates were obtained using Eq. (1.5) with the function J given by

$$J(u) = 1, \quad 0 \leq u \leq 1 \quad (\text{MEAN}),$$

$$J(u) = 0.85\,(1-u)^{-0.15}, \quad 0 \leq u \leq 1 \quad (\text{PHT}),$$

$$J(u) = \begin{cases} 0, & 0 \leq u < 0.75 \\ 4, & 0.75 \leq u \leq 1 \end{cases} \quad (\text{CTE}),$$

respectively. The resulting estimates of $\hat{R}_i$ (corresponding to the time periods 1890–1929, 1930–1969, 1970–1999), and $\hat{\gamma}$ are shown in Table 2.

To test the hypothesis that the risk measure values are equal for the three time periods, 10,000 bootstrap samples were generated, and for each bootstrap sample, the value of $\hat{\gamma}^*$ was computed. The critical values at 5 and 10% levels of significance where calculated as the empirical 95th and 90th percentiles of the 10,000 values of $\hat{\gamma}^*$. These critical values are shown for each of the three risk measures in Table 2. We see that, using all three risk measures, we are unable to reject the hypothesis of equal risk measures at the 10% level, since the values of $\hat{\gamma}$ do not exceed the corresponding values of $x_{0.1}[\hat{\gamma}^*]$. This is consistent with the conclusion of Brooks and Doswell (2000), who state "We find nothing to suggest

Table 2. Estimates for analysis of Tornado damage by time period.

	MEAN	PHT	CTE
$\hat{R}_{1890\text{–}1929}$	7119.66	9531.28	23548.7
$\hat{R}_{1930\text{–}1969}$	7244.21	8615.25	18067.3
$\hat{R}_{1970\text{–}1999}$	11692.60	13885.00	30832.1
$\hat{\gamma}$	2032.41	2342.10	5673.25
$x_{0.05}[\hat{\gamma}^*]$	2864.78	3445.64	9528.50
$x_{0.1}[\hat{\gamma}^*]$	2477.16	3009.17	8215.26

that damage from individual tornadoes has increased through time, except as a result of the increasing cost of goods and accumulation of wealth of the U.S."

We next perform a similar analysis by census region. Since almost all tornadoes occurred in the Midwest (region 2) or the South (region 3), only these two regions were included in the analysis. We consider the same three risk measures and summarize our results in Table 3. Comparing the $\hat{\gamma}$ values to the corresponding critical values, we see that for the MEAN and the PHT, we reject the null hypothesis of equal risk measure values at the 5% level. However for the CTE, $\hat{\gamma}$ is less than the 5% critical value (but greater than the 10% critical value). So, we have rather weak evidence against the null hypothesis that the CTE value is the same for the two regions. This reflects our greater uncertainty about the CTE values for the two regions.

The latter analysis brings out the important point that, depending on the choice of risk measure, conclusions may differ as to whether or not to reject the null hypothesis of equality of the risk measure values. However, there is no inconsistency here, as two distributions may indeed produce the same risk measure value under one risk measure and different risk measure values under another risk measure.

5. Concluding remarks

In this paper, we have considered testing hypotheses about the equality of risk measures using a nested L-statistic, which appears to be a natural construct in this context. We have investigated asymptotic properties of the test statistic and assessed its performance using a simulation study. We have applied the herein developed approach to a practical example.

There are several avenues for further research that go beyond the scope of the present paper, but are certainly of interest. One of them involves robustification of the herein

Table 3. Estimates for analysis of Tornado damage by region.

	MEAN	PHT	CTE
$\hat{R}_2$	12287.30	14819.00	31314.50
$\hat{R}_3$	5786.50	7381.12	16883.80
$\hat{\gamma}$	3250.38	3718.95	7215.35
$x_{0.05}[\hat{\gamma}^*]$	2336.75	2888.30	7750.51
$x_{0.1}[\hat{\gamma}^*]$	1952.77	2432.66	6469.71

proposed test statistic by using trimming techniques. Another avenue involves relaxing the presently used assumption of independence within the sub-populations. Yet another problem would be to investigate asymptotic properties of the nested L-statistic when the number k of sub-populations increases. There may, of course, be many other generalizations that arise in practice and require modifications to the approach presented in this paper.

Acknowledgements

The authors wish to thank an anonymous referee whose helpful comments have improved the paper. Research of the second and fourth authors was partially supported by individual Discovery Research Grants from the Natural Sciences and Engineering Research Council (NSERC) of Canada. We also gratefully acknowledge the hospitality of the Department of Mathematics at Indiana University in Bloomington, where all four of us met for a week in February, 2006, and started working on the project.

References

Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance* **26**, 1487–1503.

Acerbi, C. & Tasche, D. (2002). On the coherance of expected shortfall. *Journal of Banking & Finance* **26**, 1487–1503.

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

Brazauskas, V., Jones, B. L., Puri, M. L. & Zitikis, R. (2007). Estimating conditional tail expectations with actuarial applications in view. *Journal of Statistical Planning Inference* (in press).

Brazauskas, V. & 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**, 114–117.

Brooks, H. E. & Doswell, C. A. III (2000). *Normalized Damage from Major Tornadoes in the United States: 1890–1999*. http://www.nssl.noaa.gov/users/brooks/public_html/damage/tdam1.html (accessed 18 December 2006).

David, H. A. (1970). *Order statistics*. New York: Wiley.

Gini, C. (1914). On the measurement of concentration and variability of characters. *Metron* **LXIII** (2005), 3–38 (translation of the original C. Gini (1914) article by Fulvio De Santis).

Helmers, R. (1982). *Edgeworth Expansions for Linear Combinations of Order Statistics*. Amsterdam: Mathematisch Centrum.

Johnston, L. D. & Williamson, S. H. (2005). *The Annual Real and Nominal GDP for the United States, 1790 – Present*. Economic History Services, October 2005, http://www.eh.net/hmit/gdp/

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

Jones, B. L. & Zitikis, R. (2005). Testing for the order of risk measures: an application of L-statistics in actuarial science. *Metron* **LXIII** (2), 193–211.

Jones, B. L., Puri, M. L. & Zitikis, R. (2006). Testing hypotheses about the equality of several risk measure values with applications in insurance. *Insurance Mathematics and Economics* **38**, 253–270.

Kaiser, T. & Brazauskas, V. (2006). Interval estimation of actuarial risk measures. *North American Actuarial Journal* **10** (4), 249–268.

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

Shorack, G. R. & Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*. New York: Wiley.

Wang, S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance Mathmatics and Economics* **17** (1), 43–54.

Wang, S. S. & Young, V. R. (1998). Ordering risks: expected utility theory versus Yaari's dual theory of risk. *Insurance Mathematics and Economics* **22** (2), 145–161.

Wirch, J. L. & Hardy, M. L. (2000). Proper ordering of risk measures, *AFIR Congress Proceedings*. Tromso, Norway.

Appendix A

Table A.1. The first type of alternatives. Estimated power function values of the tests based on the MEAN, PHT, CTE measures, for selected n and α.

Risk measure	α	n	$c_\star$									
			0.85	0.90	0.95	1.00	1.05	1.10	1.15	1.25	1.50	2.00
MEAN	0.01	25	0.74	0.18	0.02	0.01	0.04	0.13	0.27	0.58	0.92	0.99
		50	0.96	0.54	0.04	0.01	0.08	0.30	0.59	0.91	1.00	1.00
		100	1.00	0.94	0.13	0.01	0.17	0.62	0.91	1.00	1.00	1.00
		200	1.00	1.00	0.44	0.01	0.37	0.91	1.00	1.00	1.00	1.00
	0.05	25	0.97	0.54	0.11	0.05	0.14	0.33	0.54	0.83	0.99	1.00
		50	1.00	0.89	0.19	0.05	0.22	0.54	1.00	0.98	1.00	1.00
		100	1.00	1.00	0.41	0.05	0.38	0.82	0.97	1.00	1.00	1.00
		200	1.00	1.00	0.77	0.05	0.60	0.97	1.00	1.00	1.00	1.00
	0.10	25	1.00	0.74	0.22	0.11	0.24	0.46	0.66	0.90	0.99	1.00
		50	1.00	0.96	0.33	0.11	0.33	0.67	0.88	0.99	1.00	1.00
		100	1.00	1.00	0.59	0.10	0.50	0.89	0.99	1.00	1.00	1.00
		200	1.00	1.00	0.88	0.10	0.71	0.99	1.00	1.00	1.00	1.00
PHT [$r=0.85$]	0.01	25	0.55	0.11	0.02	0.01	0.05	0.15	0.28	0.56	0.91	0.99
		50	0.89	0.29	0.03	0.01	0.08	0.28	0.53	0.85	0.99	1.00
		100	0.99	0.73	0.07	0.01	0.14	0.51	0.81	0.98	1.00	1.00
		200	1.00	0.98	0.20	0.01	0.28	0.80	0.97	1.00	1.00	1.00
	0.05	25	0.90	0.36	0.09	0.07	0.16	0.32	0.51	0.79	0.98	1.00
		50	0.99	0.66	0.13	0.07	0.23	0.50	0.74	0.94	1.00	1.00
		100	1.00	0.95	0.26	0.07	0.31	0.71	0.92	0.99	1.00	1.00
		200	1.00	1.00	0.50	0.06	0.48	0.91	0.99	1.00	1.00	1.00
	0.10	25	0.97	0.55	0.19	0.14	0.26	0.44	0.63	0.86	0.99	1.00
		50	1.00	0.82	0.24	0.14	0.34	0.62	0.83	0.97	1.00	1.00
		100	1.00	0.98	0.41	0.13	0.44	0.80	0.95	1.00	1.00	1.00
		200	1.00	1.00	0.67	0.12	0.60	0.95	1.00	1.00	1.00	1.00
CTE [$t=0.75$]	0.01	25	0.07	0.03	0.02	0.02	0.04	0.07	0.12	0.23	0.53	0.84
		50	0.09	0.03	0.01	0.01	0.03	0.07	0.15	0.36	0.79	0.98
		100	0.27	0.06	0.01	0.01	0.04	0.13	0.28	0.63	0.97	1.00
		200	0.74	0.20	0.02	0.01	0.07	0.26	0.55	0.91	1.00	1.00
	0.05	25	0.21	0.11	0.08	0.08	0.12	0.18	0.26	0.42	0.74	0.96
		50	0.33	0.13	0.06	0.06	0.10	0.19	0.32	0.59	0.93	1.00
		100	0.68	0.25	0.07	0.06	0.13	0.29	0.50	0.82	0.99	1.00
		200	0.96	0.52	0.11	0.05	0.18	0.47	0.75	0.98	1.00	1.00
	0.10	25	0.35	0.20	0.16	0.15	0.21	0.27	0.36	0.54	0.83	0.98
		50	0.53	0.25	0.12	0.12	0.18	0.30	0.44	0.70	0.96	1.00
		100	0.85	0.42	0.15	0.11	0.21	0.41	0.62	0.89	1.00	1.00
		200	0.99	0.71	0.22	0.11	0.28	0.59	0.84	0.99	1.00	1.00

NOTE: Standard errors for all entries *do not exceed* $\sqrt{0.5(1-0.5)/5000}=0.007$.

Table A.2. The second type of alternatives. Estimated power function values of the tests based on the MEAN, PHT, CTE measures, for selected n and α.

Risk measure	α	n	$c_{\star\star}$									
			1.00	1.05	1.10	1.15	1.20	1.25	1.50	2.00	2.50	3.00
MEAN	0.01	25	0.01	0.04	0.19	0.40	0.58	0.71	0.92	0.95	0.96	0.96
		50	0.01	0.12	0.53	0.82	0.93	0.96	0.99	0.99	0.99	0.99
		100	0.01	0.37	0.91	0.99	1.00	1.00	1.00	1.00	1.00	1.00
		200	0.01	0.76	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	0.05	25	0.05	0.19	0.48	0.74	0.99	0.95	0.99	0.99	1.00	1.00
		50	0.05	0.35	0.82	0.97	0.99	1.00	1.00	1.00	1.00	1.00
		100	0.05	0.64	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		200	0.05	0.91	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	0.10	25	0.11	0.32	0.65	0.87	0.96	0.99	1.00	1.00	1.00	1.00
		50	0.11	0.51	0.90	0.99	1.00	1.00	1.00	1.00	1.00	1.00
		100	0.10	0.76	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		200	0.10	0.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
PHT [$r = 0.85$]	0.01	25	0.01	0.05	0.17	0.35	0.53	0.68	0.93	0.96	0.97	0.97
		50	0.01	0.10	0.41	0.71	0.89	0.95	0.99	0.99	0.99	0.99
		100	0.01	0.23	0.77	0.96	0.99	0.99	1.00	1.00	1.00	1.00
		200	0.01	0.52	0.98	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	0.05	25	0.07	0.19	0.44	0.67	0.83	0.92	0.99	1.00	1.00	1.00
		50	0.07	0.30	0.70	0.91	0.98	0.99	1.00	1.00	1.00	1.00
		100	0.07	0.47	0.92	0.99	1.00	1.00	1.00	1.00	1.00	1.00
		200	0.06	0.75	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	0.10	25	0.14	0.30	0.58	0.80	0.92	0.96	1.00	1.00	1.00	1.00
		50	0.14	0.43	0.80	0.96	0.99	1.00	1.00	1.00	1.00	1.00
		100	0.13	0.61	0.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		200	0.12	0.84	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
CTE [$t = 0.75$]	0.01	25	0.02	0.03	0.06	0.11	0.17	0.25	0.58	0.80	0.84	0.85
		50	0.01	0.02	0.08	0.17	0.31	0.47	0.84	0.94	0.95	0.95
		100	0.01	0.04	0.19	0.46	0.69	0.85	0.98	0.99	0.99	0.99
		200	0.01	0.09	0.49	0.86	0.97	0.99	1.00	1.00	1.00	1.00
	0.05	25	0.08	0.10	0.19	0.28	0.40	0.51	0.85	0.96	0.97	0.97
		50	0.06	0.11	0.25	0.45	0.63	0.78	0.98	0.99	1.00	1.00
		100	0.06	0.15	0.45	0.75	0.91	0.97	1.00	1.00	1.00	1.00
		200	0.05	0.28	0.77	0.97	1.00	1.00	1.00	1.00	1.00	1.00
	0.10	25	0.15	0.18	0.31	0.43	0.56	0.68	0.96	1.00	1.00	1.00
		50	0.12	0.20	0.40	0.61	0.78	0.89	1.00	1.00	1.00	1.00
		100	0.11	0.27	0.60	0.85	0.96	0.99	1.00	1.00	1.00	1.00
		200	0.11	0.41	0.86	0.99	1.00	1.00	1.00	1.00	1.00	1.00

NOTE: Standard errors for all entries *do not exceed* $\sqrt{0.5(1-0.5)/5000} = 0.007$.

Table A.3. Normalized Damage from Major Tornadoes in the United States: 1890–1999 with Year and Census Region (1 Northeast, 2 Midwest, 3 South) of Occurrence.

Year	Region	Damage	Year	Region	Damage	Year	Region	Damage
1890	3	18809	1932	3	2561	1967	2	7850
1893	3	130	1932	3	1601	1968	2	9891
1893	2	2606	1932	3	640	1968	2	9891
1896	3	1380	1933	3	1005	1968	2	6923
1896	2	2071	1933	3	670	1968	3	1648
1896	2	82827	1933	3	335	1969	3	923
1898	3	12135	1936	3	6902	1969	2	4617
1899	2	1153	1936	3	29910	1970	3	39901
1899	2	1730	1942	3	747	1971	3	2209
1900	3	1106	1942	3	3111	1973	3	20470
1902	3	987	1942	3	622	1973	3	5520
1903	3	4658	1942	3	622	1974	2	21401
1904	3	479	1944	3	5458	1974	3	6420
1905	2	437	1944	1	13646	1974	2	10700
1905	3	1312	1945	3	1340	1974	3	3210
1908	3	2200	1947	3	7073	1974	3	3638
1908	3	2200	1947	3	707	1974	2	7490
1908	3	2200	1948	2	2939	1974	3	3210
1908	3	440	1949	3	1088	1975	3	11081
1909	3	2536	1952	3	2289	1975	2	49470
1909	3	423	1952	3	458	1975	2	3958
1909	3	423	1952	3	1962	1978	3	6262
1909	3	423	1953	3	9412	1978	3	14563
1913	2	18669	1953	3	25727	1979	3	52731
1917	2	3905	1953	2	11922	1979	3	3559
1917	2	5207	1953	1	32629	1979	1	26366
1917	3	2603	1953	3	15687	1980	2	6129
1918	2	4094	1955	3	4767	1980	2	17160
1919	2	7031	1956	2	5754	1982	3	5354
1920	3	3626	1957	2	22242	1982	2	10709
1920	3	363	1957	2	8341	1984	2	3609
1920	3	363	1959	2	5236	1985	1	8483
1921	3	222	1964	3	6501	1986	3	6477
1921	3	4439	1965	2	14172	1987	3	100
1924	3	1975	1965	2	8098	1988	3	5550
1924	2	23696	1965	2	10123	1989	3	6772
1925	2	30787	1965	2	6074	1990	2	10680
1925	3	385	1965	2	4049	1991	2	3809
1927	3	2249	1965	2	6074	1993	3	5875
1927	3	3935	1965	2	4859	1993	3	2761
1927	2	41224	1966	3	6724	1994	3	2799
1929	3	1591	1966	3	11207	1994	3	2799
1930	3	2031	1966	2	37356	1995	3	5416
1930	3	203	1966	2	4483	1998	3	2373
1932	3	960	1967	2	5352	1999	3	45302
1932	3	4802	1967	2	10704			