Ann Inst Stat Math (2019) 71:811–835
https://doi.org/10.1007/s10463-018-0660-2

Weighted allocations, their concomitant-based estimators, and asymptotics

Nadezhda Gribkova[1] · **Ričardas Zitikis**[2]

Received: 9 October 2017 / Revised: 29 January 2018 / Published online: 7 April 2018

Abstract Various members of the class of weighted insurance premiums and risk capital allocation rules have been researched from a number of perspectives. Corresponding formulas in the case of parametric families of distributions have been derived, and they have played a pivotal role when establishing parametric statistical inference in the area. Nonparametric inference results have also been derived in special cases such as the tail conditional expectation, distortion risk measure, and several members of the class of weighted premiums. For weighted allocation rules, however, nonparametric inference results have not yet been adequately developed. In the present paper, therefore, we put forward empirical estimators for the weighted allocation rules and establish their consistency and asymptotic normality under practically sound conditions. Intricate statistical considerations rely on the theory of induced order statistics, known as concomitants.

Keywords Weighted allocation · Insurance premium · Concomitant · Consistency · Asymptotic normality

Research supported by the Natural Sciences and Engineering Research Council of Canada.

✉ Ričardas Zitikis
zitikis@stats.uwo.ca

Nadezhda Gribkova
n.gribkova@spbu.ru

1 Faculty of Mathematics and Mechanics, St. Petersburg State University, St. Petersburg, Russia 199034

2 School of Mathematical and Statistical Sciences, Western University, London, ON N6A 5B7, Canada

1 Introduction

The tail conditional expectation, exponential tilting, and various members of the class of weighted insurance premiums and the corresponding risk capital allocation rules have been extensively researched (e.g., Pflug and Römisch 2007; Rüschendorf 2013; McNeil et al. 2015; Föllmer and Schied 2016). Their formulas in the case of various parametric families of distributions have been derived (e.g., Furman and Landsman 2005, 2010; Asimit et al. 2013; Su 2016; Asimit et al. 2016; Su and Furman 2017; Ratovomirija et al. 2017; Vernic 2017; and references therein), thus facilitating parametric statistical inference in the area.

The literature also contains a number of nonparametric inference results (e.g., Brazauskas and Serfling 2003; Brazauskas 2009; Brazauskas and Kleefeld 2009; and references therein), particularly in special cases such as the value-at-risk (e.g., Maesono and Penev 2013; and references therein), the tail conditional expectation (e.g., Brazauskas et al. 2008; and references therein), and distortion risk measures (e.g., Jones and Zitikis 2007; and references therein), with results available in light- and heavy-tailed settings (e.g., Necir and Meraghni 2009; Necir et al. 2007, 2010; Rassoul 2013; Brahimi et al. 2012; and references therein).

Nonparametric statistical inference for weighted allocation rules has not yet, however, been adequately developed, and we therefore devote the current paper to this topic. In particular, in next Sect. 2 we construct empirical estimators for the weighted allocations. Main asymptotic results are presented in Sect. 3, and their illustrations are given in Sect. 4. Proofs are in Sect. 5. They are, essentially, delicate combinations of the classical groundbreaking works of Shorack (1972), Bhattacharya (1974), van Zwet (1980), and Yang (1981). Section 6 concludes the paper with additional thoughts, some inspired by questions and suggestions by the reviewers of this paper, on various paths that can be taken to further statistical inference in the area, given the multitude and diversity of applications.

2 Premiums, allocations, and their estimators

Let X be a real-valued random variable, which could, for example, be a financial or insurance risk associated with a business line of a company. Denote the cumulative distribution function (cdf) of X by F_X. When X is viewed as a stand-alone risk, then the capital needed to mitigate the risk can be calculated using (e.g., Furman and Zitikis 2008a)

$$\pi_w = \frac{\mathbf{E}[X w \circ F_X(X)]}{\mathbf{E}[w \circ F_X(X)]} \tag{1}$$

with an appropriately chosen weight function $w : [0, 1] \to [-\infty, \infty]$, where $w \circ F_X$ denotes the composition of the functions w and F_X. We of course assume that the two expectations in the definition of π_w are well defined and finite, and $\mathbf{E}[w \circ F_X(X)]$ is not zero.

The function w may or may not take infinite values and may or may not be non-decreasing, depending on the context. Throughout the paper, we always assume that w is finite on the open interval $(0, 1)$ and, at each point of $(0, 1)$, is either left continuous or right continuous. As far as we are aware of, all practically relevant weight functions satisfy these properties, with a few illustrative examples given next.

When dealing with insurance losses, researchers work with nonnegative and non-decreasing weight functions, which ensure that π_w is nonnegatively loaded, that is, the bound $\pi_w \geq \mathbf{E}[X]$ holds for all risks X under consideration. In other contexts, such as econometrics and, more specifically, measurement of income inequality, the function w can be non-increasing.

For example, $w(t) = \mathbf{1}\{t > p\}$ for any parameter $p \in (0, 1)$ is non-decreasing and leads to the insurance version of the tail conditional expectation. Another example is $w(t) = \nu(1-t)^{\nu-1}$ with parameter $\nu > 0$. If $\nu \in (0, 1]$, then π_w is the proportional hazards transform (Wang 1995, 1996). If $\nu \geq 1$, then π_w reduces to the (absolute) S-Gini index used for measuring income equality (e.g., Zitikis and Gastwirth 2002; Greselin and Zitikis 2018; and references therein).

Note that if the cdf F_X is continuous, then π_w can be written as the integral

$$\pi_w = \int_0^1 F_X^{-1}(t) w^*(t) \mathrm{d}t \tag{2}$$

of the quantile function F_X^{-1} of X, with the weight function

$$w^*(t) = \frac{w(t)}{\int_0^1 w(u)\mathrm{d}u},$$

which is a probability density function (pdf) whenever $w(t) \geq 0$ for all $t \in [0, 1]$ and $\int_0^1 w(u)\mathrm{d}u \in (0, \infty)$. This representation of π_w connects our present research with the dual utility theory (Yaari 1987; Quiggin 1993; and references therein) that has arisen as a prominent alternative to the classical utility theory of von Neumann and Morgenstern (1944).

Setting appropriate insurance premiums and allocating capital to individual business lines are usually done within the company's risk profile. That is, if Y is the risk associated with the entire company, then allocating capital to the business line whose risk is X is done by taking into account the value of Y. This viewpoint leads us to the weighted capital allocation rule (Furman and Zitikis 2008b)

$$\Pi_w = \frac{\mathbf{E}[X w \circ F_Y(Y)]}{\mathbf{E}[w \circ F_Y(Y)]}, \tag{3}$$

where F_Y denotes the cdf of Y. Obviously, setting Y to X reduces Π_w to π_w, and for this reason we concentrate on developing nonparametric statistical inference for Π_w, and then we specialize the results to π_w. For the role of π_w and Π_w in the context of the weighted insurance pricing model (WIPM), we refer to Furman and Zitikis (2017).

To construct an empirical estimator for the allocation Π_w, let (X_k, Y_k), $k = 1, 2 \ldots, n$, be independent copies of the random pair (X, Y), succinctly written as $(X_1, Y_1), \ldots, (X_n, Y_n) \overset{\text{i.i.d.}}{\sim} F_{X,Y}$. For each integer $n \geq 1$, define the empirical cdf $\widehat{F}_Y$ by

$$\widehat{F}_Y(y) = \frac{1}{n+1} \sum_{k=1}^{n} \mathbf{1}\{Y_k \le y\}, \tag{4}$$

where $\mathbf{1}\{Y_k \le y\}$ is the indicator of $Y_k \le y$: it is equal to 1 when the statement $Y_k \le y$ is true, and 0 otherwise. This is an empirical estimator of the cdf $F_Y(y)$ that slightly differs from the classical empirical cdf because of the use of $1/(n+1)$ instead of $1/n$. This adjustment, also utilized by Gribkova and Zitikis (2017), is important because $\widehat{F}_Y(y)$ defined in this way takes only values $k/(n+1)$, $k = 1, \dots, n$, which are always inside the open interval (0, 1) on which the weight function w is finite.

We are now in the position to define the empirical estimator of Π_w by the formula

$$\widetilde{\Pi}_w = \frac{\sum_{k=1}^{n} X_k w \circ \widehat{F}_Y(Y_k)}{\sum_{k=1}^{n} w \circ \widehat{F}_Y(Y_k)}, \tag{5}$$

with a tilde used instead of the usual hat on top of Π_w because we reserve the latter notation for another estimator to be introduced in a moment. Note that when $Y = X$ and thus $Y_k = X_k$ for all $k \ge 1$, the empirical allocation $\widetilde{\Pi}_w$ reduces to the estimator

$$\widetilde{\pi}_w = \frac{\sum_{k=1}^{n} X_k w \circ \widehat{F}_X(X_k)}{\sum_{k=1}^{n} w \circ \widehat{F}_X(X_k)} \tag{6}$$

of π_w, where $\widehat{F}_X$ is defined by Eq. (4) but now based on $X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} F_X$. Estimators (5) and (6) are ratio statistics, whose deep asymptotic properties have been explored by Maesono (2005, 2010).

When the underlying population cdf F_Y is continuous, the random variable $F_Y(Y)$ is uniform on (0, 1) and, therefore, the denominator in the definition of Π_w is equal to $\int_0^1 w(u)\,\mathrm{d}u$. Since we do not need to estimate the latter integral, we can therefore use the following simpler estimator

$$\widehat{\Pi}_w = \frac{\widehat{\Delta}_w}{\int_0^1 w(u)\mathrm{d}u} \tag{7}$$

of Π_w, where

$$\widehat{\Delta}_w = \frac{1}{n} \sum_{k=1}^{n} X_k w \circ \widehat{F}_Y(Y_k).$$

Note that, almost surely, we have the equation

$$\widehat{\Delta}_w = \frac{1}{n} \sum_{k=1}^{n} X_{[k:n]} w_{k,n},$$

where

$$w_{k,n} = w\left(\frac{k}{n+1}\right)$$

and $X_{[1:n]}, \ldots, X_{[n:n]}$ are the induced order statistics, known as concomitants, corresponding to the order statistics $Y_{1:n}, \ldots, Y_{n:n}$. When $Y = X$ and thus $Y_k = X_k$ for all $k \geq 1$, then the concomitants reduce to the order statistics $X_{1:n}, \ldots, X_{n:n}$. In this case, the estimator $\widehat{\Pi}_w$ reduces to the estimator $\widehat{\pi}_w$ of π_w given by the equation

$$\widehat{\pi}_w = \frac{1}{n} \sum_{k=1}^{n} X_{k:n} \frac{w_{k,n}}{\int_0^1 w(u)\mathrm{d}u}. \tag{8}$$

While $\widehat{\pi}_w$ is a linear combination of order statistics (e.g., Stigler 1974; Helmers 1982; Shorack 2017; see also Gribkova 2017, for recent references), which is a less technically demanding object, the estimators $\widetilde{\Pi}_w$, $\widehat{\Pi}_w$, and $\widehat{\Delta}_w$ are linear combinations of concomitants, which require much more sophisticated methods of analysis. In the next section, we establish conditions under which these estimators are consistent and asymptotically normal.

3 Main results

Asymptotic behavior of the above introduced estimators is influenced by the weight function w and the joint distribution of X and Y, which interact with each other in delicate ways. Hence, determining their influence on asymptotic results, such as consistency and asymptotic normality, at one stroke becomes undesirable from several points of view, most notably from the practical point of view because resulting conditions are quite unwieldy. Due to this reason, we next set out to derive asymptotic results in several complementary forms, and with easily verifiable conditions as we shall see in next Sect. 4. We start with strong consistency, using $\stackrel{\text{a.s.}}{\to}$ to denote convergence almost surely.

Theorem 1 *If the first moment* $\mathbf{E}[X]$ *is finite and the weight function* w *is continuous on* $[0, 1]$, *then* $\widetilde{\Pi}_w \stackrel{a.s.}{\to} \Pi_w$ *and thus* $\widetilde{\pi}_w \stackrel{a.s.}{\to} \pi_w$ *when* $n \to \infty$.

The theorem is attractive in the sense that it does not impose any condition on the underlying random variables, except the very minimal condition that the first moment of X is finite. We shall further elaborate on moment-type requirements at the end of this section, as well as in concluding Sect. 6.

A shortcoming of Theorem 1 is that the condition on the weight function w is very strong. For example, it is not satisfied by $w(t) = \nu(1-t)^{\nu-1}$ for any $\nu \in (0, 1)$. Furthermore, the condition is not satisfied by $w(t) = \mathbf{1}\{t > p\}$ for any $p \in (0, 1)$, and we thus cannot use the theorem to deduce strong consistency of the tail conditional expectation.

In the next two theorems, we no longer assume continuity and thus boundedness of the function w on the compact interval $[0, 1]$. Instead, we require finite higher moments of X, as well as the continuity of the cdf F_Y when dealing with $\widehat{\Pi}_w$, and the continuity of the cdf F_X when dealing with $\widehat{\pi}_w$.

For any $1 \leq p \leq \infty$, we use L_p to denote the space of all Borel measurable functions $h : [0, 1] \to \mathbb{R}$ such that $\|h\|_p := (\int_0^1 |f(t)|^p \mathrm{d}t)^{1/p} < \infty$ when $1 \leq p <$

∞ and $\|h\|_\infty := \operatorname{ess\,sup}_{t\in[0,1]} |h(t)| < \infty$ when $p = \infty$. We use ":=" when wanting to emphasize that certain equations are by definition.

The following two theorems are consequences of the strong law of large numbers for L-statistics, proved in various levels of generality by van Zwet (1980). For example, Theorem 2.1 and Corollary 2.1 of van Zwet (1980) imply the following theorem.

Theorem 2 *Let the cdf F_X be continuous. If $\mathbf{E}[|X|^p] < \infty$ and $w \in L_q$ for some $p, q \in [1, \infty]$ such that $p^{-1} + q^{-1} = 1$, then $\widetilde{\pi}_w \overset{a.s.}{\to} \pi_w$ when $n \to \infty$.*

The next theorem, which follows from Theorem 3.1 of van Zwet (1980), allows us to use different values of p and q on different subintervals of (0, 1), thus enabling different growth rates of the quantile function F_X^{-1} and the weight function w near the end points of the interval (0, 1). Following van Zwet (1980), let $0 =: a_0 < a_1 < \cdots < a_j := 1$ be points dividing the interval (0, 1) into $j \geq 1$ subintervals, which we denote by

$$A_i = (a_{i-1}, a_i), \quad i = 1, \ldots, j,$$

whose ε-neighborhoods within the interval (0, 1) are

$$B_{i,\varepsilon} = (a_{i-1} - \varepsilon, a_i + \varepsilon) \cap (0, 1).$$

Theorem 3 *Let the cdf F_X be continuous, and let $p_i, q_i \in [1, \infty]$ be such that $p_i^{-1} + q_i^{-1} = 1$. If there is $\varepsilon > 0$ such that $F_X^{-1}\mathbf{1}_{B_{i,\varepsilon}} \in L_{p_i}$ and $w\mathbf{1}_{A_i} \in L_{q_i}$ for every $i = 1, \ldots, j$, then $\widehat{\pi}_w \overset{a.s.}{\to} \pi_w$ when $n \to \infty$.*

From the practical point of view, it is (weak) consistency that really matters, which also naturally leads to the exploration of asymptotic normality, and our following research path is in this direction. It leads us to practically attractive and justifiable conditions on the weight function w as well as on the joint cdf of X and Y. Our focus now is also shifting from the simpler $\widehat{\pi}_w$ toward the more complex weighted allocation rule $\widehat{\Pi}_w$. Not surprisingly, therefore, in what follows we employ the conditional expectation function

$$g_{X|Y}(y) = \mathbf{E}[X \mid Y = y]$$

defined on the support of Y, as well as the conditional variance function

$$v^2_{X|Y}(y) = \mathbf{Var}[X \mid Y = y].$$

We note that the function $g_{X|Y} \circ F_Y^{-1}(t)$ is known in the literature as the quantile regression function of X on Y, and it has prominently manifested in the literature (e.g., Rao and Zhao 1995; Tse 2009, 2015; and the references therein). The quantile conditional-variance function $v^2_{X|Y} \circ F_Y^{-1}(t)$ is also prominently featured in these works. For a work devoted mainly to the analysis and modeling of these functions, we refer to Kamnitui et al. (2015). All these functions play a pivotal role throughout the rest of the present paper. We use $\overset{\mathbf{P}}{\to}$ to denote convergence in probability.

Theorem 4 *Assume that* $\mathbf{E}[X^2]$ *is finite and the cdf* F_Y *is continuous. If* $v^2_{X|Y} \circ F_Y^{-1} \in L_p$ *and* $w^2 \in L_q$ *for some* $p, q \in [1, \infty]$ *such that* $p^{-1} + q^{-1} = 1$*, then* $\widehat{\Pi}_w \xrightarrow{\mathbf{P}} \Pi_w$ *when* $n \to \infty$.

To appreciate the theorem from the practical perspective, we look at several special cases. First, when $p = 1$, we have $q = \infty$, which is to say that the weight function w is bounded. This covers the weight function $w(t) = \mathbf{1}\{t > p\}$ for every $p \in (0, 1)$, and also the weight function $w(t) = \nu(1-t)^{\nu-1}$ for every $\nu \geq 1$. Note also that the condition $v^2_{X|Y} \circ F_Y^{-1} \in L_1$ is equivalent to $\mathbf{E}[X^2] < \infty$, which we assume.

Second, when $p = \infty$, which implies $q = 1$, the function $v^2_{X|Y}(y)$ must be bounded and the function w^2 integrable on $(0, 1)$, that is, $w \in L_2$. The weight function $w(t) = \mathbf{1}\{t > p\}$ is always such, whereas $w(t) = \nu(1-t)^{\nu-1}$ belongs to L_2 only when $\nu > 1/2$. The latter restriction has appeared naturally in Jones and Zitikis (2003, 2007), Brahimi et al. (2011), and other insurance-related works dealing with the proportional hazards premium.

The next theorem, which is in the spirit of Theorem 3 and uses the notations introduced before it, concludes our explorations of consistency.

Theorem 5 *Assume that* $\mathbf{E}[X^2]$ *is finite and the cdf* F_Y *is continuous. If there is* $\varepsilon > 0$ *such that* $v^2_{X|Y} \circ F_Y^{-1}\mathbf{1}_{B_{i,\varepsilon}} \in L_{p_i}$ *and* $w\mathbf{1}_{A_i} \in L_{2q_i}$ *for every* $i = 1, \ldots, j$*, then* $\widehat{\Pi}_w \xrightarrow{\mathbf{P}} \Pi_w$ *when* $n \to \infty$.

We now set out to establish asymptotic normality of the estimator $\widehat{\Pi}_w$. We show, in particular, that its asymptotic variance is

$$\sigma^2_w = \left(\sigma^2_{w,1} + \sigma^2_{w,2}\right) \Big/ \left(\int_0^1 w(u)\mathrm{d}u\right)^2$$

where

$$\sigma^2_{w,1} = \int_0^1 v^2_{X|Y} \circ F_Y^{-1}(t)\, w^2(t)\, \mathrm{d}t \tag{9}$$

and

$$\sigma^2_{w,2} = \int_0^1 \int_0^1 w(s)w(t)\, (\min\{s,t\} - st)\, \mathrm{d}g_{X|Y} \circ F_Y^{-1}(s)\, \mathrm{d}g_{X|Y} \circ F_Y^{-1}(t). \tag{10}$$

We note at the outset that in the theorem that follows, we impose conditions that assure the finiteness of $\sigma^2_{w,1}$ and $\sigma^2_{w,2}$. Note also that for the variance $\sigma^2_{w,2}$ to be well defined, we need to, and thus do, assume—without explicitly saying this every time—that the quantile regression function $g_{X|Y} \circ F_Y^{-1}$ is left continuous on $(0, 1)$ and of bounded variation on $[\varepsilon, 1-\varepsilon]$ for every $0 < \varepsilon < 1/2$. Following Shorack's (1972) terminology, this means that $g_{X|Y} \circ F_Y^{-1}$ belongs to the class $\mathcal{L}$. In what follows, we also use the notation

$$\mu_{2,X|Y}(y) = \mathbf{E}[X^2 \mid Y = y],$$

and $\xrightarrow{d}$ denotes convergence in distribution.

Theorem 6 *Let F_Y be continuous. Furthermore, let the weight function w satisfy the following conditions:*

(i) *w is continuous on $(0,1)$ except possibly at a finite number of points $t_1 < \cdots < t_m$, and there is $r > 1/2$ such that for every $\varepsilon > 0$ there is a constant $c < \infty$ such that*

$$|w(u) - w(v)| \le c|u - v|^r \tag{11}$$

for all $u, v \in (t_{i-1}, t_i) \cap (\varepsilon, 1-\varepsilon)$ and every $i = 1, \ldots, m+1$, where $t_0 := 0$ and $t_{m+1} := 1$;

(ii) *there is (small) $\varepsilon > 0$ such that w is differentiable on the set $\Theta_\varepsilon := (0,\varepsilon) \cup (1-\varepsilon, 1)$, and there are $\kappa_1, \kappa_2 \in [0,1)$ such that the two bounds*

$$|w(t)| \le ct^{-\kappa_1/2}(1-t)^{-\kappa_2/2} \tag{12}$$

and

$$t(1-t)|w'(t)| \le ct^{-\kappa_1/2}(1-t)^{-\kappa_2/2} \tag{13}$$

hold for all $t \in \Theta_\varepsilon$.

If the function $g_{X|Y} \circ F_Y^{-1}$ is continuous at every point t_i of condition (i), and, for some $\delta > 0$, the bound

$$\mu_{2,X|Y} \circ F_Y^{-1}(t) \le ct^{-1+\kappa_1+\delta}(1-t)^{-1+\kappa_2+\delta} \tag{14}$$

holds for all $t \in \Theta_\varepsilon$ with the same κ_1 and κ_2 as in condition (ii), then

$$n^{1/2}(\widehat{\Pi}_w - \Pi_w) \xrightarrow{d} \mathcal{N}(0, \sigma_w^2) \tag{15}$$

when $n \to \infty$.

Note that condition (14) implies $\mathbf{E}[X^2] < \infty$. In the special case $Y = X$, we have the following corollary to Theorem 6.

Corollary 1 *Let F_X be continuous. If the weight function w satisfies conditions (i) and (ii) of Theorem 6, the quantile function F_X^{-1} is continuous at every point t_i of condition (i), and $(F_X^{-1}(t))^2$ does not exceed the right-hand side of bound (14) near the end points of the interval $(0,1)$, then*

$$n^{1/2}(\widehat{\pi}_w - \pi_w) \xrightarrow{d} \mathcal{N}(0, \sigma_w^2) \tag{16}$$

when $n \to \infty$, where

$$\sigma_w^2 = \frac{1}{\left(\int_0^1 w(u)\mathrm{d}u\right)^2} \int_0^1 \int_0^1 w(s)w(t)\,(\min\{s,t\} - st)\ \mathrm{d}F_X^{-1}(s)\,\mathrm{d}F_X^{-1}(t).$$

In a variety of forms, Corollary 1 has frequently appeared in literature. Indeed, if F_X is continuous, then $\widehat{\pi}_w$ is an L-statistic and π_w is its asymptotic mean, called L-functional. For details and references on the topic, we refer to the monographs by Helmers (1982), Serfling (1980), and Shorack (2017).

It is clear from the above results that the tails of the weight function w and the quantile conditional variance, or quantile conditional second moment, interact, and thus there is always a delicate balancing act to maintain: stronger conditions on w lead to weaker conditions on the quantile-based functions, and vice versa. There is, however, a possibility to weaken both sets of conditions at the same time, but this leads to drastically different results and hinges on other techniques of proof, as seen from the works of Necir and Meraghni (2009), and Necir et al. (2007), who tackle the proportional hazards transform; Necir et al. (2010), and Rassoul (2013), who tackle the tail conditional expectation; and Brahimi et al. (2012), who tackle the general distortion risk measure.

4 Examples

From the practical point of view, it is (weak) consistency and asymptotic normality that are of most interest, and so our aim in this section is to illustrate Theorems 5 and 6. We use three classes of weight functions that are of particular importance in econometrics, insurance, and financial engineering. The functions never take infinite values on the open interval (0, 1) and are continuous from either left- or right-hand sides at each point of (0, 1).

The basic assumptions on the joint distribution of X and Y are the finiteness of the second moment $\mathbf{E}[X^2]$ and the continuity of the cdf F_Y. Recall also our agreement to always—when considering asymptotic normality—assume that the quantile regression function $g_{X|Y} \circ F_Y^{-1}$ is left continuous on (0, 1) and has bounded variation on $[\varepsilon, 1-\varepsilon]$ for every $0 < \varepsilon < 1/2$, that is, belongs to the class $\mathcal{L}$ in the terminology of Shorack (1972).

The following three examples are of increasing complexity.

4.1 Step functions

Let $p \in (0, 1)$ be any fixed parameter, and let w be any of the following two step functions:

$$w(t) = \mathbf{1}\{t < p\} \tag{17}$$

or

$$w(t) = \mathbf{1}\{t > p\}. \tag{18}$$

Both functions are bounded on the interval [0, 1] and have discontinuities at the point p. Function (17) gives rise to the financial version of the average-value-at-risk, whereas function (18) gives rise to the insurance version of the average-value-at-risk (e.g., Pflug and Römisch 2007; Rüschendorf 2013; McNeil et al. 2015; Föllmer and Schied 2016). The duality is a consequence of the fact that insurers view losses as nonnegative random variables, whereas financial engineers view losses as negative realizations of random variables.

4.1.1 Consistency

The assumption $\mathbf{E}[X^2] < \infty$ implies $v^2_{X|Y} \circ F_Y^{-1} \in L_1(0, 1)$, and since $w^2 \in L_\infty(0, 1)$, we can therefore use Theorem 5 with $j = 1$ and $A_1 = (0, 1)$ to conclude that the estimator $\widehat{\Pi}_w$ is consistent.

4.1.2 Asymptotic normality

Condition (i) of Theorem 6 is satisfied with $m = 1$ and $t_1 = p$. In particular, bound (11) holds with $r = 1$ for all $u, v \in (0, p)$, as well as for all $u, v \in (p, 1)$. Condition (ii) of the same theorem is also satisfied, with $\kappa_1 = 0$ and $\kappa_2 = 0$.

Hence, the estimator $\widehat{\Pi}_w$ is asymptotically normal, that is, statement (15) holds, if (1) the function $g_{X|Y} \circ F_Y^{-1}$ is continuous at the point p, and (2) with $\kappa_1 = 0$, $\kappa_2 = 0$, and some $\delta > 0$, bound (14) holds near the end points of the interval $(0, 1)$. The latter condition, which is

$$\mu_{2,X|Y} \circ F_Y^{-1}(t) \le ct^{-1+\delta}(1-t)^{-1+\delta}$$

for all $t \in (0, \varepsilon) \cup (1 - \varepsilon, 1)$, is just slightly stronger than the very basic requirement $\mathbf{E}[X^2] < \infty$.

4.2 One-sided power function

Let $\nu > 0$ be any fixed parameter, and consider the one-sided power function

$$w(t) = \nu(1-t)^{\nu-1}.$$

As noted in Sect. 2, this weight function arises in insurance, as well as when measuring economic inequality. When $\nu = 1$, the function gives rise to the mean, which is viewed as the net premium by insurers, and the mean income by those working in the area of measuring economic inequality. When $\nu \in (0, 1)$, the (increasing) weight function gives rise to the proportional hazards transform (Wang 1995, 1996), and when $\nu > 1$, the (decreasing) weight function gives rise to the absolute S-Gini index (e.g., Greselin and Zitikis 2018; and references therein). This duality is natural because insurers are particularly concerned with large losses, therefore making them even larger when setting premiums with positive risk-loadings, whereas those working in the area of measurement inequality tend to emphasize the less fortunate members of the population.

4.2.1 Consistency

The one-sided power function may or may not be bounded, depending on the parameter ν value. Hence, assuming that there is $p \geq 1$ such that

$$v_{X|Y}^2 \circ F_Y^{-1} \in L_p(1/2, 1), \tag{19}$$

we need to specify those values of ν that satisfy $w \in L_{2q}(1/2, 1)$. This is equivalent to the requirement $\nu > 1 - 1/(2q)$, which, due to the relationship $p^{-1} + q^{-1} = 1$, is equivalent to

$$\nu > \frac{1}{2} + \frac{1}{2p}. \tag{20}$$

Under these assumptions, the estimator $\widehat{\Pi}_w$ is consistent.

To make the above discussion more intuitive, we can slightly strengthen conditions by requiring $\mathbf{E}[|X|^{2p} \mid Y > \text{med}(Y)] < \infty$, where $\text{med}(Y)$ denotes the median of Y. Then, if condition (20) holds, then the estimator $\widehat{\Pi}_w$ is consistent.

4.2.2 Asymptotic normality

Condition (i) of Theorem 6 is satisfied with $m = 0$. Condition (ii) of the same theorem is satisfied with $\kappa_1 = 0$ and $\kappa_2 = \max\{2(1-\nu), 0\}$. Hence, asymptotic normality of $\widehat{\Pi}_w$ follows if

$$\mu_{2,X|Y} \circ F_Y^{-1}(t) \leq c \begin{cases} t^{-1+\delta}(1-t)^{1-2\nu+\delta} & \text{when } 0 < \nu < 1, \\ t^{-1+\delta}(1-t)^{-1+\delta} & \text{when } \nu \geq 1, \end{cases} \tag{21}$$

near the end points of the interval (0, 1).

To appreciate this condition from the intuitive point of view, we assume that $\mu_{2,X|Y} \circ F_Y^{-1}(t)$ is of the order $c(1-t)^{-1/p}$ for t's in a neighborhood of 1, which roughly speaking means that the moment $\mathbf{E}[|X|^{2p} \mid Y > \text{med}(Y)]$ is finite. Then the estimator $\widehat{\Pi}_w$ is asymptotically normal, that is, statement (15) holds, whenever ν satisfies condition (20).

4.3 Two-sided power function

Consider the weight function

$$w(t) = \begin{cases} c_1 t^{\nu_1} & \text{when } \; 0 < t < t_*, \\ c_2(1-t)^{\nu_2} & \text{when } \; t_* \leq t < 1, \end{cases} \tag{22}$$

with parameters $c_1, c_2 \in (0, \infty)$, $\nu_1, \nu_2 > -1$, and $t_* \in (0, 1)$. Depending on the parameter values, function (22) may or may not be discontinuous at the point t_*, but it is always continuous on the open intervals $(0, t_*)$ and $(t_*, 1)$.

The continuous version

$$w(t) = \frac{ct^{\nu_1}}{ct^{\nu_1} + (1-t)^{-\nu_2}} \tag{23}$$

of function (22) has been used as a probability weighting function in research areas dealing with risk and uncertainty (e.g., Tversky and Kahneman 1992; Gonzalez and Wu 1999; Wakker 2010; and references therein).

Our following considerations of consistency and asymptotic normality are equally applicable to any of these two weight functions, with the only exception that in the case of function (22) we need to require continuity of $g_{X|Y} \circ F_Y^{-1}$ at the point t_*, but there is no such need when dealing with function (23) as it is continuous on the entire open interval (0, 1).

4.3.1 Consistency

Analogously to Sect. 4.2.1, we conclude that $\widehat{\Pi}_w$ is consistent whenever

1. there are $p_1 \geq 1$ and $p_2 \geq 1$ such that, with $\mathrm{med}(Y)$ denoting the median of Y, the moments $\mathbf{E}[|X|^{2p_1} \mid Y < \mathrm{med}(Y)]$ and $\mathbf{E}[|X|^{2p_2} \mid Y > \mathrm{med}(Y)]$ are finite;
2. the parameters $\nu_1 > -1$ and $\nu_2 > -1$ are such that

$$\nu_1 > -\frac{1}{2} + \frac{1}{2p_1} \quad \text{and} \quad \nu_2 > -\frac{1}{2} + \frac{1}{2p_2}. \tag{24}$$

4.3.2 Asymptotic normality

Irrespective of the parameter values, if we set $m = 1$ and $t_1 = t_*$, we see that the function w satisfies condition (11) with $r = 1$ for all $u, v \in (\varepsilon, t_*)$, as well as for all $u, v \in (t_*, 1-\varepsilon)$. This verifies condition (i) of Theorem 6. Condition (ii) of the same theorem is satisfied with $\kappa_1 = \max\{-2\nu_1, 0\}$ and $\kappa_2 = \max\{-2\nu_2, 0\}$.

If we deal with function (22), then we need to assume that $g_{X|Y} \circ F_Y^{-1}$ is continuous at the point t_*. In the case of function (23), there is no need for such an assumption on $g_{X|Y} \circ F_Y^{-1}$.

To understand condition (14), let the p_1th power of $\mu_{2,X|Y} \circ F_Y^{-1}(t)$ be integrable over $t \in (0, \varepsilon)$, and the p_2th power of the same function be integrable over $t \in (1-\varepsilon, 1)$. Roughly speaking, these assumptions mean that the moments $\mathbf{E}[|X|^{2p_1} \mid Y < \mathrm{med}(Y)]$ and $\mathbf{E}[|X|^{2p_2} \mid Y > \mathrm{med}(Y)]$ are finite.

Under these conditions, the estimator $\widehat{\Pi}_w$ is asymptotically normal, that is, statement (15) holds, provided that bounds (24) hold.

5 Proofs

Proof of Theorem 1 We have $\widetilde{\Pi}_w \overset{\text{a.s.}}{\to} \Pi_w$ provided that, when $n \to \infty$,

$$\frac{1}{n}\sum_{k=1}^{n} X_k w \circ \widehat{F}_Y(Y_k) \overset{\text{a.s.}}{\to} \mathbf{E}[Xw \circ F_Y(Y)] \tag{25}$$

and

$$\frac{1}{n}\sum_{k=1}^{n} w \circ \widehat{F}_Y(Y_k) \overset{\text{a.s.}}{\to} \mathbf{E}[w \circ F_Y(Y)]. \tag{26}$$

Statement (26) follows from statement (25) if we set X_k to 1. Hence, we only need to prove statement (25). We write

$$\frac{1}{n}\sum_{k=1}^{n} X_k w\circ\widehat{F}_Y(Y_k) = \frac{1}{n}\sum_{k=1}^{n} X_k w\circ F_Y(Y_k) + \frac{1}{n}\sum_{k=1}^{n} X_k \left(w \circ \widehat{F}_Y(Y_k) - w \circ F_Y(Y_k)\right). \tag{27}$$

The classical strong law of large numbers implies that $n^{-1}\sum_{k=1}^{n} X_k w \circ F_Y(Y_k)$ converges to $\mathbf{E}[Xw \circ F_Y(Y)]$ almost surely. Hence, we are left to prove that the second average on the right-hand side of Eq. (27) converges to 0 almost surely. This we achieve by first estimating its absolute value by

$$\left(\frac{1}{n}\sum_{k=1}^{n} |X_k|\right) \sup_{y\in\mathbb{R}} \left|w \circ \widehat{F}_Y(y) - w \circ F_Y(y)\right|. \tag{28}$$

By the strong law of large numbers, $n^{-1}\sum_{k=1}^{n} |X_k|$ converges almost surely to the (finite) mean of $|X|$, and the supremum in (28) converges to zero almost surely because of the classical Glivenko–Cantelli theorem and the uniform continuity of w, which holds because w is continuous on the compact interval [0, 1]. Hence, statement (25) holds, and so does $\widetilde{\Pi}_w \overset{\text{a.s.}}{\to} \Pi_w$. Statement $\widetilde{\pi}_w \overset{\text{a.s.}}{\to} \pi_w$ follows as a special case when $X = Y$ and $X_k = Y_k$ for all $k = 1, \ldots, n$. This completes the proof of Theorem 1. □

Proof of Theorem 4 The theorem follows from the statement

$$\widehat{\Delta}_w \overset{\mathbf{P}}{\to} \mathbf{E}[Xw \circ F_Y(Y)]. \tag{29}$$

To prove it, we write the decomposition $\widehat{\Delta}_w = (T_{n,1} + T_{n,2})/n$, where

$$T_{n,1} = \sum_{k=1}^{n} g_{X|Y}(Y_{k:n}) w_{k,n}$$

and

$$T_{n,2} = \sum_{k=1}^{n} \left(X_{[k:n]} - g_{X|Y}(Y_{k:n})\right) w_{k,n}.$$

The rest of the proof consists of two steps:

$$\frac{1}{n} T_{n,1} \overset{\mathbf{P}}{\to} \mathbf{E}[Xw \circ F_Y(Y)], \tag{30}$$

$$\frac{1}{n}T_{n,2} \xrightarrow{\mathbf{P}} 0. \tag{31}$$

In fact, statement (30) holds with convergence in probability replaced by convergence almost surely. Indeed, the strong law of large numbers for L-statistics (van Zwet 1980; Corollary 2.1) implies

$$\frac{1}{n}T_{n,1} \overset{\text{a.s.}}{\to} \int_0^1 g_{X|Y} \circ F_Y^{-1}(t)w(t)\mathrm{d}t \tag{32}$$

when $n \to \infty$. It remains to note that the integral on the right-hand side of statement (32) is equal to $\mathbf{E}[Xw \circ F_Y(Y)]$. Hence, we are left to prove statement (31), which means that, for every $\delta > 0$, we need to show

$$\mathbf{P}\left(|T_{n,2}| > n\delta\right) \to 0 \tag{33}$$

when $n \to \infty$. Recall that, conditionally on $Y_{1:n}, \ldots, Y_{n:n}$, the concomitants $X_{[1:n]}, \ldots, X_{[n:n]}$ are independent (Bhattacharya 1974; Lemma 1). Hence, with the help of Markov's inequality, we obtain

$$\begin{aligned}
\mathbf{P}\left(|T_{n,2}| > n\delta\right) &= \mathbf{E}\left[\mathbf{P}\left(\left|\sum_{k=1}^n \left(X_{[k:n]} - g_{X|Y}(Y_{k:n})\right) w_{k,n}\right| > n\delta \mid Y_1, \ldots, Y_n\right)\right] \\
&\le \frac{1}{n^2\delta^2}\sum_{k=1}^n \mathbf{E}\left[\mathbf{E}\left[\left(X_{[k:n]} - g_{X|Y}(Y_{k:n})\right)^2 \mid Y_1, \ldots, Y_n\right]\right] w_{k,n}^2 \\
&= \frac{1}{n^2\delta^2}\sum_{k=1}^n \mathbf{E}\left[v_{X|Y}^2(Y_{k:n})\right] w_{k,n}^2,
\end{aligned} \tag{34}$$

where the right-most equation follows from the fact that (Bhattacharya 1974; Lemma 1) conditionally on $Y_{1:n}, \ldots, Y_{n:n}$, the concomitants $X_{[1:n]}, \ldots, X_{[n:n]}$ follow the cdf's $F(x \mid Y_{1:n}), \ldots, F(x \mid Y_{n:n})$, respectively, where $F(x \mid y) = \mathbf{P}[X \le x \mid Y = y]$. Next we apply Hölder's inequality on the right-hand side of bound (34) and obtain

$$\begin{aligned}
\mathbf{P}\left(|T_{n,2}| > n\delta\right) &\le \frac{1}{n\delta^2}\left(\frac{1}{n}\sum_{k=1}^n \left(\mathbf{E}\left[v_{X|Y}^2(Y_{k:n})\right]\right)^p\right)^{1/p}\left(\frac{1}{n}\sum_{k=1}^n |w_{k,n}|^{2q}\right)^{1/q} \\
&\le \frac{1}{n\delta^2}\left(\frac{1}{n}\sum_{k=1}^n \mathbf{E}\left[v_{X|Y}^{2p}(Y_{k:n})\right]\right)^{1/p}\left(\frac{1}{n}\sum_{k=1}^n |w_{k,n}|^{2q}\right)^{1/q} \\
&= \frac{1}{n\delta^2}\left(\frac{1}{n}\sum_{k=1}^n \mathbf{E}\left[v_{X|Y}^{2p} \circ F_Y^{-1}(U_k)\right]\right)^{1/p}\left(\frac{1}{n}\sum_{k=1}^n |w_{k,n}|^{2q}\right)^{1/q},
\end{aligned} \tag{35}$$

where $U_1, \ldots, U_n \overset{\text{i.i.d.}}{\sim} \text{Uni}(0,1)$ are i.i.d. random variables from the uniform distribution on the interval [0, 1]. The first average on the right-hand side of Eq. (35) is equal to the integral $\int_0^1 v_{X|Y}^{2p} \circ F_Y^{-1}(t)\,\mathrm{d}t$, whereas the second average converges to $\int_0^1 |w(t)|^{2q}\mathrm{d}t$. Both integrals are finite by assumption. This completes the proof of Theorem 4. □

Proof of Theorem 5 We need to prove statement (29) under the conditions of Theorem 5. We start again with the decomposition $\widehat{\Delta}_w = (T_{n,2} + T_{n,1})/n$. Statement (32) follows by the strong law of large numbers for L-statistics (van Zwet 1980; Theorem 3.1). It remains to prove statement (31). We fix any $\delta > 0$ and write

$$\mathbf{P}\left(|T_{n,2}| > n\delta\right) = \mathbf{P}\left(\left|\sum_{k=1}^{n}\left(X_{[k:n]} - g_{X|Y} \circ F^{-1}(U_{k:n})\right) w_{k,n}\right| > n\delta\right) \\ \leq \Delta + \mathbf{P}\left(\mathcal{D}^c\right) \tag{36}$$

for any subset $\mathcal{D}$ of the sample space, where

$$\Delta := \mathbf{P}\left(\left\{\left|\sum_{k=1}^{n}\left(X_{[k:n]} - g_{X|Y} \circ F^{-1}(U_{k:n})\right) w_{k,n}\right| > n\delta\right\} \cap \mathcal{D}\right)$$

and $U_{1:n}, \ldots, U_{n:n}$ are the order statistics of $U_1, \ldots, U_n \overset{\text{i.i.d.}}{\sim} \text{Uni}(0,1)$. We next make a special choice of $\mathcal{D}$.

First, we recall the definitions of A_i and $B_{i,\varepsilon}$, given before Theorem 3. Then we define $r_i = \min\{k : k/(n+1) \in A_i\}$ and $s_i = \max\{k : k/(n+1) \in A_i\}$, and with the notation $\mathcal{D}_i = \left\{U_{r_i:n} \in B_{i,\varepsilon}\right\} \cap \left\{U_{s_i:n} \in B_{i,\varepsilon}\right\}$, we define

$$\mathcal{D} = \bigcap_{i=1}^{j} \mathcal{D}_i.$$

Since $r_i/(n+1),\ s_i/(n+1) \in A_i \subset B_{i,\varepsilon}$, Bernstein's inequality implies $\mathbf{P}\big(\mathcal{D}_i^c\big) \leq \exp\{-c_i n\}$ for some $c_i > 0$, where

$$\mathcal{D}_i^c = \left\{U_{r_i:n} \notin B_{i,\varepsilon}\right\} \cup \left\{U_{s_i:n} \notin B_{i,\varepsilon}\right\}.$$

Consequently,

$$\mathbf{P}\left(\mathcal{D}^c\right) = \mathbf{P}\left(\bigcup_{i=1}^{j} \mathcal{D}_i^c\right) \leq \sum_{i=1}^{j} e^{-c_i n}. \tag{37}$$

In view of estimate (37), from now on we restrict our attention to only the quantity Δ.

Since conditionally on $Y_{1:n}, \ldots, Y_{n:n}$, the concomitants $X_{[1:n]}, \ldots, X_{[n:n]}$ follow the cdf's $F(x \mid Y_{1:n}), \ldots, F(x \mid Y_{n:n})$, respectively, we use Markov's inequality and obtain the bound

$$\begin{aligned} \varDelta &= \mathbf{E}\left[\mathbf{1}_{\mathcal{D}}\mathbf{P}\left(\left|\sum_{k=1}^{n}\left(X_{[k:n]} - g_{X|Y}\circ F^{-1}(U_{k:n})\right)w_{k,n}\right| > n\delta \mid U_1,\ldots,U_n\right)\right] \\ &\le \frac{1}{(n\delta)^2}\mathbf{E}\left[\mathbf{1}_{\mathcal{D}}\sum_{k=1}^{n} v_{X|Y}^2\circ F_Y^{-1}(U_{k:n})w_{k,n}^2\right]. \end{aligned} \tag{38}$$

We split the sum $\sum_{k=1}^{n}$ into two sums: one is the sum over those k such that $k/(n+1) \in \{a_1,\ldots,a_j\}$, and another sum is over all the remaining k's. That is, we write

$$\sum_{k=1}^{n} = \sum_{i=1}^{j}\sum_{k:k/(n+1)=a_i} + \sum_{i=1}^{j}\sum_{k:\,k/(n+1)\in A_i}.$$

If a sum is empty, we set it to 0. For example, the sum $\sum_{k:k/(n+1)=a_i}$ is always empty when $i=j$ because $a_j=1$ and $k\in\{1,\ldots,n\}$. Hence, in order to show that $\varDelta$ converges to 0 when $n\to\infty$, we need to prove that

$$\varDelta_{1,i} := \frac{1}{n^2}\sum_{k:\,k/(n+1)=a_i}\mathbf{E}\left[\mathbf{1}_{\mathcal{D}}v_{X|Y}^2\circ F_Y^{-1}(U_{k:n})\right]w_{k,n}^2 \to 0 \tag{39}$$

for every $i=1,\ldots,j-1$, and

$$\varDelta_{2,i} := \frac{1}{n^2}\sum_{k:\,k/(n+1)\in A_i}\mathbf{E}\left[\mathbf{1}_{\mathcal{D}}v_{X|Y}^2\circ F_Y^{-1}(U_{k:n})\right]w_{k,n}^2 \to 0 \tag{40}$$

for every $i=1,\ldots,j$.

We begin with $\varDelta_{1,i}$, whose sum $\sum_{k:k/(n+1)=a_i}$ is either empty or contains only one summand. If it is not empty, then let k be the (only) integer that satisfies $k=(n+1)a_i$. As we have already noted above, the case $i=j$ can be dropped. Hence, we only consider $i\le j-1$. Under this condition, $w_{k,n}^2$ can only be one of the values $w^2(a_i)$, $i=1,\ldots,j-1$. Since the function w is bounded on every compact subinterval of $(0,1)$, we therefore conclude that $\max_{i=1,\ldots,j-1} w^2(a_i)$ is finite. This implies the first bound below, with the rest of calculations being obvious:

$$\begin{aligned} \frac{1}{n}\mathbf{E}\left[\mathbf{1}_{\mathcal{D}}v_{X|Y}^2\circ F_Y^{-1}(U_{k:n})w_{k,n}^2\right] &\le \frac{c}{n}\mathbf{E}\left[v_{X|Y}^2\circ F_Y^{-1}(U_{k:n})\right] \\ &\le \frac{c}{n}\sum_{k=1}^{n}\mathbf{E}\left[v_{X|Y}^2\circ F_Y^{-1}(U_{k:n})\right] \\ &= c\int_0^1 v_{X|Y}^2\circ F_Y^{-1}(t)\mathrm{d}t < \infty. \end{aligned} \tag{41}$$

Statement (39) follows, and we are left to prove statement (40).

With the help of Hölder's inequality, we have

$$\Delta_{2,i} \leq \frac{1}{n}\left(\frac{1}{n}\sum_{k:\,k/(n+1)\in A_i}\left(\mathbf{E}\left[\mathbf{1}_{\mathcal{D}}v^2_{X|Y}\circ F_Y^{-1}(U_{k:n})\right]\right)^{p_i}\right)^{1/p_i} \times\left(\frac{1}{n}\sum_{k:\,k/(n+1)\in A_i}|w_{k,n}|^{2q_i}\right)^{1/q_i}. \tag{42}$$

Furthermore, when $n\to\infty$,

$$\frac{1}{n}\sum_{k:\,k/(n+1)\in A_i}|w_{k,n}|^{2q_i}\to\int_{A_i}|w(t)|^{2q_i}\,\mathrm{d}t<\infty.$$

Since $\mathcal{D}=\bigcap_{i=1}^{j}\mathcal{D}_i$, we have the first bound below, with the rest of calculations being obvious:

$$\begin{aligned}
&\frac{1}{n}\sum_{k:\,k/(n+1)\in A_i}\left(\mathbf{E}\left[\mathbf{1}_{\mathcal{D}}v^2_{X|Y}\circ F_Y^{-1}(U_{k:n})\right]\right)^{p_i}\\
&\quad\leq\frac{1}{n}\sum_{k:\,k/(n+1)\in A_i}\mathbf{E}\left[\mathbf{1}_{\mathcal{D}_i}v^{2p_i}_{X|Y}\circ F_Y^{-1}(U_{k:n})\right]\\
&\quad=\frac{1}{n}\sum_{k=r_i}^{s_i}\mathbf{E}\left[\mathbf{1}_{\mathcal{D}_i}v^{2p_i}_{X|Y}\circ F_Y^{-1}(U_{k:n})\right]\\
&\quad\leq\frac{1}{n}\sum_{k=1}^{n}\mathbf{E}\left[\mathbf{1}_{B_{i,\varepsilon}}v^{2p_i}_{X|Y}\circ F_Y^{-1}(U_{k:n})\right]\\
&\quad=\int_{B_{i,\varepsilon}}v^{2p_i}_{X|Y}\circ F_Y^{-1}(t)\,\mathrm{d}t<\infty.
\end{aligned} \tag{43}$$

This implies statement (40) and completes the proof that the quantity Δ on the right-hand side of bound (36) converges to 0 when $n\to\infty$. The proof of Theorem 5 is finished. □

Proof of Theorem 6 Asymptotic normality of $\widehat{\Pi}_w$ follows if we show that, when $n\to\infty$,

$$n^{1/2}(\widehat{\Pi}_w-\Pi_{w,n})\xrightarrow{d}\mathcal{N}(0,\sigma^2_w), \tag{44}$$

$$n^{1/2}(\Pi_{w,n}-\Pi_w)\to 0, \tag{45}$$

where

$$\Pi_{w,n}:=\sum_{k=1}^{n}\frac{w_{k,n}}{\int_0^1 w(u)\mathrm{d}u}\int_{(k-1)/n}^{k/n}g_{X|Y}\circ F_Y^{-1}(t)\,\mathrm{d}t.$$

We next establish statements (44) and (45), and in this way complete the proof of Theorem 6.

We note at the outset that statements (44) and (45) require different subsets of conditions formulated in Theorem 6. In the proofs that follow, we shall specify which of them, and where, are required.

Proof of statement (44) We have $\Pi_{w,n} = \Delta_{w,n}/\int_0^1 w(u)\mathrm{d}u$, where

$$\Delta_{w,n} = \sum_{k=1}^{n} w_{k,n} \int_{(k-1)/n}^{k/n} g_{X|Y} \circ F_Y^{-1}(t)\, \mathrm{d}t.$$

Hence, statement (44) is equivalent to

$$n^{1/2}(\widehat{\Delta}_w - \Delta_{w,n}) \overset{d}{\to} \mathcal{N}(0, \sigma_{w,1}^2 + \sigma_{w,2}^2)$$

when $n \to \infty$. We write $n^{1/2}(\widehat{\Delta}_w - \Delta_{w,n}) = W_n + T_n$, where

$$W_n = n^{-1/2} \sum_{k=1}^{n} \left(X_{[k:n]} - g_{X|Y}(Y_{k:n})\right) w_{k,n} \tag{46}$$

and

$$T_n = n^{1/2} \left(\frac{1}{n} \sum_{k=1}^{n} g_{X|Y}(Y_{k:n}) w_{k,n} - \Delta_{w,n} \right). \tag{47}$$

Hence, to prove the theorem, we need to show that

$$W_n + T_n \overset{d}{\to} \mathcal{N}\left(0, \sigma_{w,1}^2 + \sigma_{w,2}^2\right), \tag{48}$$

when $n \to \infty$. We follow the approach of Yang (1981) for proving the central limit theorem for linear combinations of concomitants. Namely, we rely on the following fundamental theorem of Yang (1981), which in a somewhat different context has recently, and effectively, been utilized by Gribkova and Zitikis (2017). □

Theorem 7 (Yang 1981) *Let $(X_1, Y_1), (X_2, Y_2), \ldots$ be random pairs. Denote $\mathbf{Z}_n = ((X_1, Y_1), \ldots, (X_n, Y_n))$ and $\mathbf{Y}_n = (Y_1, \ldots, Y_n)$, and let $W_n := W_n(\mathbf{Z}_n)$ and $T_n := T_n(\mathbf{Y}_n)$ be measurable vector-valued functions of $\mathbf{Z}_n$ and $\mathbf{Y}_n$, respectively. Suppose that T_n converges in distribution to F_T, and the conditional distribution of W_n given $\mathbf{Y}_n$ converges weakly to a distribution F_W which does not depend on the Y_k's. Then $(W_n, T_n) \overset{d}{\to} F_W F_T$.*

First, we work with the quantity W_n defined by Eq. (46) and prove that its conditional distribution given $\mathbf{Y}_n$ converges to the normal distribution with the mean 0 and variance $\sigma_{w,1}^2$ for almost all sequences $(Y_m)_{m \ge 1}$, with the limiting distribution not depending on the sequence $(Y_m)_{m \ge 1}$. Next, we prove that the quantity T_n defined by Eq. (47) is asymptotically normal with the mean 0 and variance $\sigma_{w,2}^2$. Given these two results, Theorem 7 implies that the joint distribution of (W_n, T_n) converges to the product of the two aforementioned normal distributions. In turn, this implies that $W_n + T_n$ is asymptotically normal with the mean 0 and variance $\sigma_{w,1}^2 + \sigma_{w,2}^2$. Hence, the rest of the proof consists of two parts, which deal with the asymptotic normality of W_n and T_n, respectively.

Part 1 Using Bhattacharya's (1974) result already utilized in the proof of Theorem 5, we have $\mathbf{E}[W_n \mid \boldsymbol{Y}_n] = 0$ with the conditional variance $V_n^2 := \mathbf{Var}[W_n \mid \boldsymbol{Y}_n]$ expressed by

$$V_n^2 = \frac{1}{n} \sum_{k=1}^{n} v_{X|Y}^2(Y_{k:n}) w_{k,n}^2.$$

Applying Lindeberg's normal-convergence criterion, we conclude that the sequence of the (conditional) distributions of W_n/V_n is asymptotically standard normal if, for every $\varepsilon > 0$ and when $n \to \infty$,

$$\frac{1}{nV_n^2} \sum_{k=1}^{n} w_{k,n}^2 h_{\theta_{k,n}}(Y_{k:n}) \to 0 \tag{49}$$

for almost all realizations of the sequence $Y_1, Y_2, \ldots$, where

$$h_{\theta_{k,n}}(y) = \int (x - g_{X|Y}(y))^2 \mathbf{1}\{|x - g_{X|Y}(y)| \geq \theta_{k,n}\} \mathrm{d}F(x \mid y) \tag{50}$$

with the notation

$$\theta_{k,n} = \frac{\varepsilon n^{1/2} V_n}{|w_{k,n}|}.$$

(If $w_{k,n} = 0$, the corresponding summand in statement (49) vanishes, and hence $\theta_{k,n}$ can be defined arbitrarily in this case.) The strong law of large numbers for L-statistics (van Zwet 1980; Theorem 3.1) implies

$$V_n^2 \overset{\text{a.s.}}{\to} \int_0^1 v_{X|Y}^2 \circ F_Y^{-1}(t) w^2(t)\, \mathrm{d}t, \tag{51}$$

with the integral on the right-hand side being equal to $\sigma_{w,1}^2$. To verify $\theta_{k,n} \overset{\text{a.s.}}{\to} \infty$, we write the bounds

$$\begin{aligned} \theta_{k,n} &\geq \varepsilon n^{1/2} V_n / \max_{k=1,\ldots,n} |w_{k,n}| \\ &\geq \varepsilon n^{1/2} V_n / \max_{k=1,\ldots,n} \left(\frac{k(n-k)}{n^2} \right)^{-\max(\kappa_1,\kappa_2)/2} \\ &= \varepsilon n^{1/2} V_n / n^{\max(\kappa_1,\kappa_2)/2}. \end{aligned} \tag{52}$$

Since $\max(\kappa_1, \kappa_2) < 1$, we have $\theta_{k,n} \overset{\text{a.s.}}{\to} \infty$. Applying the strong law of large numbers for L-statistics (van Zwet 1980; Theorem 3.1), we have that, for every $K > 0$,

$$\frac{1}{n}\sum_{k=1}^{n} w_{k,n}^2 h_K(Y_{k:n}) \overset{\text{a.s.}}{\to} \int_0^1 w^2(t)\, h_K \circ F_Y^{-1}(t)\, \mathrm{d}t \tag{53}$$

when $n \to \infty$. (The function $h_K(y)$ is defined by Eq. (50) with K instead of $\theta_{k,n}$.) Since $\theta_{k,n} \to \infty$, statements (53) and (51) imply the Lindeberg's criterion for almost all realizations of the sequence $(Y_m)_{m\geq 1}$. Hence, the conditional distribution of W_n/V_n given $\boldsymbol{Y}_n$ converges to the standard normal distribution almost surely.

Part 2 In order to prove statement (44), it remains to show that the distribution of T_n given by (47) converges to the normal law with the mean 0 and variance $\sigma_{w,2}^2$. The latter fact is a direct consequence of a result of Shorack's (1972) Theorem 1 on asymptotic normality of linear combination of functions of order statistics.

Indeed, let $U_1, \ldots, U_n \overset{\text{i.i.d.}}{\sim} \text{Uni}(0, 1)$, and let $U_{1:n}, \ldots, U_{n:n}$ denote the corresponding order statistics. Then, with the equality holding in distribution, we have

$$T_n = n^{1/2}\left(\frac{1}{n}\sum_{k=1}^{n} g_{X|Y} \circ F_Y^{-1}(U_{k:n}) w_{k,n} - \Delta_{w,n}\right).$$

Since under the conditions of Theorem 6, the conditions of Theorem 1 by Shorack (1972) are satisfied, the aforementioned asymptotic normality of T_n holds. Statement (44) follows. □

Proof of statement (45) We start with the equations

$$\begin{aligned}\Delta_{w,n} &= \sum_{k=1}^{n} w_{k,n} \int_{(k-1)/n}^{k/n} g_{X|Y} \circ F_Y^{-1}(t)\, \mathrm{d}t \\ &= \int_0^1 g_{X|Y} \circ F_Y^{-1}(t) w_n(t)\, \mathrm{d}t, \end{aligned} \tag{54}$$

where the function $w_n : (0, 1] \to \mathbb{R}$ is defined by $w_n(t) = w_{k,n}$ when $(k-1)/n < t \leq k/n$, for all $k = 1, \ldots, n$. Next we write

$$n^{1/2}(\widehat{\Delta}_w - \Delta_{w,n}) = I_{n,1} + I_{n,2} + I_{n,3},$$

where

$$I_{n,l} = n^{1/2} \int_{D_l} (w_n(t) - w(t))\, g_{X|Y} \circ F_Y^{-1}(t)\, \mathrm{d}t$$

with the sets $D_1 = (0, \varepsilon)$, $D_2 = (\varepsilon, 1-\varepsilon)$, and $D_3 = (1-\varepsilon, 1)$, and with a sufficiently small $\varepsilon > 0$ so that we could use condition (ii) of Theorem 6. We shall prove that $I_{n,1}$

and $I_{n,2}$ converge to zero when $n \to \infty$; the treatment of $I_{n,3}$ is similar as in the case of $I_{n,1}$ and is therefore omitted.

Without loss of generality, we can, and thus do, assume $1/n < \varepsilon$, and thus $|I_{n,1}|$ does not exceed the sum $I^*_{n,1} + I^{**}_{n,1}$, where

$$I^*_{n,1} = n^{1/2} \int_0^{1/n} |w_n(t) - w(t)| \left| g_{X|Y} \circ F_Y^{-1}(t) \right| \, \mathrm{d}t,$$
$$I^{**}_{n,1} = n^{1/2} \int_{1/n}^{\varepsilon} |w_n(t) - w(t)| \left| g_{X|Y} \circ F_Y^{-1}(t) \right| \, \mathrm{d}t.$$

Using bounds (12) and (14), we obtain

$$\begin{aligned} I^*_{n,1} &\le n^{1/2} \, |w(1/n)| \int_0^{1/n} \left| g_{X|Y} \circ F_Y^{-1}(t) \right| \, \mathrm{d}t \\ &\quad + n^{1/2} \int_0^{1/n} |w(t)| \left| g_{X|Y} \circ F_Y^{-1}(t) \right| \, \mathrm{d}t \\ &\le c\, n^{-\delta/2}. \end{aligned} \tag{55}$$

Furthermore, upon recalling that $w_n(t)$ is equal to $w_{k,n}$ and therefore to $w(k/(n+1))$ for all $t \in \big((k-1)/n,\ k/n\big)$, we use bounds (13) and (14) to obtain

$$I^{**}_{n,1} \le cn^{-1/2} \sum_{k=2}^{[n\varepsilon]+1} \int_{(k-1)/n}^{k/n} \tau_k^{-\kappa_1/2-1} \, t^{-1/2+\kappa_1/2+\delta/2} \, \mathrm{d}t$$

with some $\tau_k \in \big((k-1)/n,\ k/n\big)$, where $[\cdot]$ denotes the integer part. Without loss of generality, we let $\delta > 0$ be smaller than $1 - \kappa_1$. We have

$$I^{**}_{n,1} \le cn^{-1/2} \left(n^{1/2-\delta/2} + \frac{1}{n} \sum_{k=2}^{[n\varepsilon]} \left(\frac{k}{n} \right)^{-3/2+\delta/2} \right).$$

Since

$$\frac{1}{n} \sum_{k=2}^{n} \left(\frac{k}{n} \right)^{-3/2+\delta/2} \le c \left(\frac{1}{n} \right)^{-1/2+\delta/2},$$

we conclude that

$$I^{**}_{n,1} \le c\, n^{-\delta/2}. \tag{56}$$

Bounds (55) and (56) complete the proof that $I_{n,1} \to 0$ when $n \to \infty$.

It remains to show that $I_{n,2}$ converges to 0 when $n \to \infty$. For this, we first rewrite $I_{n,2}$ as follows

$$I_{n,2} = n^{1/2} \sum_{i=1}^{m+1} \int_{t_{i-1}}^{t_i} \mathbf{1}_{\mathcal{D}_i} \big(w_n(t) - w(t) \big) g_{X|Y} \circ F_Y^{-1}(t)\, \mathrm{d}t.$$

By condition (i) of Theorem 6 and the integrability of $g_{X|Y} \circ F_Y^{-1}$ (which holds because the first moment of X is finite), the absolute value of $I_{n,2}$ does not exceed $cn^{1/2-r}$, which converges to 0 because of $r > 1/2$. This completes the proof statement (45). □

Having thus established statements (44) and (45), we have also concluded the proof of Theorem 6. □

6 Concluding remarks

Insurance losses are usually nonnegative random variables, and the arising weight functions w are therefore defined on the nonnegative real half-line and take nonnegative values. In financial engineering, both losses and profits are of interest at the same time and are thus modeled with real-valued random variables.

Since our developed statistical inference results are not sensitive to the positivity and negativity of random variables and weight functions, we have therefore accommodate these scenarios by considering real-valued random variables as well as real-valued weight functions. We have covered a large set of scenarios of interest, but of course not all of them:

- Random variables whose cdf's contain discrete components are not covered by our results. This is a limitation as insurance losses may have a mass at, for example, zero, due to the fact that some losses are not covered. Incorporating the discrete component into asymptotic results is not trivial and requires considerable effort, as seen from the pioneering article by Chernoff et al. (1967); we also refer to the monographs by Serfling (1980) and Shorack (2017).
- Insurance literature contains numerous examples when losses with infinite variances arise. We refer to a recent work by Brazauskas and Kleefeld (2016), who analyze one of such examples. They also provide a wealth of earlier references on the topic.
- In operational risk literature (Nešlehová et al. 2006), we find examples when random variable do not even have finite first moments.

These are challenging problems which, we think, should be tackled separately, as they likely require specialized techniques. In the case of infinite variances, and especially when the first moments are infinite, our proposed empirical estimators may not even be suitable, unless $w(t)$ converges to 0 fast enough when t approaches the end points of the interval (0, 1), as can be seen from, for example, the pioneering work of Necir and Meraghni (2009), as well as their subsequent prolific research, who tackle actuarial risk measures within the framework of extreme-value theory (e.g., Embrechts et al. 1997; Beirlant et al. 2004; Castillo et al. 2005; de Haan and Ferreira 2006).

Acknowledgements We are indebted to two anonymous reviewers for suggestions, insightful comments, and constructive criticism that guided our work on the revision.

References

Asimit, A. V., Vernic, R., Zitikis, R. (2013). Evaluating risk measures and capital allocations based on multi-losses driven by a heavy-tailed background risk: The multivariate Pareto-II model. *Risks*, *1*, 14–33.

Asimit, A. V., Furman, E., Vernic, R. (2016). Statistical inference for a new class of multivariate Pareto distributions. *Communications in Statistics: Simulation and Computation*, *45*, 456–471.

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J. L. (2004). *Statistics of extremes: Theory and applications*. Chichester: Wiley.

Bhattacharya, P. K. (1974). Convergence of sample paths of normalized sum of induced order statistics. *Annals of Statistics*, *2*, 1034–1039.

Brahimi, B., Meddi, F., Necir, A. (2012). Bias-corrected estimation in distortion risk premiums for heavy-tailed losses. *Afrika Statistika*, *7*, 474–490.

Brahimi, B., Meraghni, D., Necir, A., Zitikis, R. (2011). Estimating the distortion parameter of the proportional-hazard premium for heavy-tailed losses. *Insurance: Mathematics and Economics*, *49*, 325–334.

Brazauskas, V. (2009). Robust and efficient fitting of loss models: diagnostic tools and insights. *North American Actuarial Journal*, *13*, 356–369.

Brazauskas, V., Kleefeld, A. (2009). Robust and efficient fitting of the generalized Pareto distribution with actuarial applications in view. *Insurance: Mathematics and Economics*, *45*, 424–435.

Brazauskas, V., Kleefeld, A. (2016). Modeling severity and measuring tail risk of Norwegian fire claims. *North American Actuarial Journal*, *20*, 1–16.

Brazauskas, V., Jones, B. L., Puri, M. L., Zitikis, R. (2008). Estimating conditional tail expectation with actuarial applications in view. *Journal of Statistical Planning and Inference*, *138*(11), 3590–3604 (**special issue in Honor of Junjiro Ogawa: Design of experiments, multivariate analysis and statistical inference**).

Brazauskas, V., Serfling, R. (2003). Favourable estimators for fitting Pareto models: A study using goodness-of-fit measures with actual data. *ASTIN Bulletin*, *33*, 365–381.

Castillo, E., Hadi, A. S., Balakrishnan, N., Sarabia, J. M. (2005). *Extreme value and related models with applications in engineering and science*. Hoboken: Wiley.

Chernoff, H., Gastwirth, J. L., Johns, M. V. (1967). Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation. *Annals of Mathematical Statistics*, *38*, 52–72.

de Haan, L., Ferreira, A. (2006). *Extreme value theory: An introduction*. New York: Springer.

Embrechts, P., Klüppelberg, C., Mikosch, T. (1997). *Modelling extremal events: For insurance and finance*. New York: Springer.

Föllmer, H., Schied, A. (2016). *Stochastic finance: An introduction in discrete time* (4th ed.). Berlin: Walter de Gruyter.

Furman, E., Landsman, Z. (2005). Risk capital decomposition for a multivariate dependent gamma portfolio. *Insurance: Mathematics and Economics*, *37*, 635–649.

Furman, E., Landsman, Z. (2010). Multivariate Tweedie distributions and some related capital-at-risk analyses. *Insurance: Mathematics and Economics*, *46*, 351–361.

Furman, E., Zitikis, R. (2008a). Weighted premium calculation principles. *Insurance: Mathematics and Economics*, *42*, 459–465.

Furman, E., Zitikis, R. (2008b). Weighted risk capital allocations. *Insurance: Mathematics and Economics*, *43*, 263–269.

Furman, E., Zitikis, R. (2017). Beyond the Pearson correlation: Heavy-tailed risks, weighted Gini correlations, and a Gini-type weighted insurance pricing model. *ASTIN Bulletin*, *47*, 919–942.

Gonzalez, R., Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, *38*, 129–166.

Greselin, F., Zitikis, R. (2018). From the classical Gini index of income inequality to a new Zenga-type relative measure of risk: A modeller's perspective. *Econometrics*, *6*, 1–20 (**special issue on econometrics and income inequality, with Guest Editors Martin Biewen and Emmanuel Flachaire**).

Gribkova, N. V. (2017). Cramér type large deviations for trimmed L-statistics. *Probability and Mathematical Statistics*, *37*, 101–118.
Gribkova, N. V., Zitikis, R. (2017). Statistical foundations for assessing the difference between the classical and weighted-Gini betas. *Mathematical Methods of Statistics*, *26*, 267–281.
Helmers, R. (1982). *Edgeworth expansions for linear combinations of order statistics*. Amsterdam: Mathematisch Centrum.
Jones, B. L., Zitikis, R. (2003). Empirical estimation of risk measures and related quantities. *North American Actuarial Journal*, *7*, 44–54.
Jones, B. L., Zitikis, R. (2007). Risk measures, distortion parameters, and their empirical estimation. *Insurance: Mathematics and Economics*, *41*, 279–297.
Kamnitui, N., Santiwipanont, T., Sumetkijakan, S. (2015). Dependence measuring from conditional variances. *Dependence Modeling*, *3*, 98–112.
Maesono, Y. (2005). Asymptotic representation of ratio statistics and their mean squared errors. *Journal of the Japan Statistical Society*, *35*, 73–97.
Maesono, Y. (2010). Edgeworth expansion and normalizing transformation of ratio statistics and their application. *Communications in Statistics: Theory and Methods*, *39*, 1344–1358.
Maesono, Y., Penev, S. (2013). Improved confidence intervals for quantiles. *Annals of the Institute of Statistical Mathematics*, *65*, 167–189.
McNeil, A. J., Frey, R., Embrechts, P. (2015). *Quantitative risk management: Concepts, techniques and tools* (Revised ed.). Princeton, NJ: Princeton University Press.
Necir, A., Meraghni, D. (2009). Empirical estimation of the proportional hazard premium for heavy-tailed claim amounts. *Insurance: Mathematics and Economics*, *45*, 49–58.
Necir, A., Meraghni, D., Meddi, F. (2007). Statistical estimate of the proportional hazard premium of loss. *Scandinavian Actuarial Journal*, *2007*, 147–161.
Necir, A., Rassoul, A., Zitikis, R. (2010). Estimating the conditional tail expectation in the case of heady-tailed losses. *Journal of Probability and Statistics*, *2010*, 1–17.
Nešlehová, J., Embrechts, P., Chavez-Demoulin, V. (2006). Infinite-mean models and the LDA for operational risk. *Journal of Operational Risk*, *1*, 3–25.
Pflug, G. C., Römisch, W. (2007). *Modeling, measuring and managing risk*. Singapore: World Scientific.
Quiggin, J. (1993). *Generalized expected utility theory*. Dordrecht: Kluwer.
Rao, C. R., Zhao, L. C. (1995). Convergence theorems for empirical cumulative quantile regression function. *Mathematical Methods of Statistics*, *4*, 81–91.
Rassoul, A. (2013). Kernel-type estimator of the conditional tail expectation for a heavy-tailed distribution. *Insurance: Mathematics and Economics*, *53*, 698–703.
Ratovomirija, G., Tamraz, M., Vernic, R. (2017). On some multivariate Sarmanov mixed Erlang reinsurance risks: Aggregation and capital allocation. *Insurance: Mathematics and Economics*, *74*, 197–209.
Rüschendorf, L. (2013). *Mathematical risk analysis: Dependence, risk bounds, optimal allocations and portfolios*. New York: Springer.
Serfling, R. J. (1980). *Approximation theorems of mathematical statistics*. New York: Wiley.
Shorack, G. R. (1972). Functions of order statistics. *Annals of Mathematical Statistics*, *43*, 412–427.
Shorack, G. R. (2017). *Probability for statisticians* (2nd ed.). New York: Springer.
Stigler, S. M. (1974). Linear functions of order statistics with smooth weight functions. *Annals of Statistics*, *2*, 676–693.
Su, J. (2016). *Multiple risk factors dependence structures with applications to actuarial risk management*. Ph.D. thesis, York University, Ontario, Canada.
Su, J., Furman, E. (2017). A form of multivariate Pareto distribution with applications to financial risk measurement. *ASTIN Bulletin*, *47*, 331–357.
Tse, S. M. (2009). On the cumulative quantile regression process. *Mathematical Methods of Statistics*, *18*, 270–279.
Tse, S. M. (2015). The cumulative quantile regression function with censored and truncated response. *Mathematical Methods of Statistics*, *24*, 147–155.
Tversky, A., Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323.
van Zwet, W. R. (1980). A strong law for linear functions of order statistics. *Annals of Probability*, *8*, 986–990.
Vernic, R. (2017). Capital allocation for Sarmanov's class of distributions. *Methodology and Computing in Applied Probability*, *19*, 311–330.

von Neumann, J., Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.
Wakker, P. P. (2010). *Prospect theory: For risk and ambiguity*. Cambridge: Cambridge University Press.
Wang, S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics*, *17*, 43–54.
Wang, S. S. (1996). Premium calculation by transforming the layer premium density. *ASTIN Bulletin*, *26*, 71–92.
Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica*, *55*, 95–115.
Yang, S. S. (1981). Linear combinations of concomitants of order statistics with application to testing and estimation. *Annals of the Institute of Statistical Mathematics*, *33*, 463–470.
Zitikis, R., Gastwirth, J. L. (2002). Asymptotic distribution of the S-Gini index. *Australian and New Zealand Journal of Statistics*, *44*, 439–446.