Insurance: Mathematics and Economics 55 (2014) 78–90

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

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

Asymptotic theory for the empirical Haezendonck–Goovaerts risk measure

Jae Youn Ahn[a,*], Nariankadu D. Shyamalkumar[b,**]

[a] Department of Statistics, Ewha Womans University, 11-1 Daehyun-Dong, Seodaemun-Gu, Seoul 120-750, Republic of Korea

[b] Department of Statistics and Actuarial Science, The University of Iowa, 241 Schaeffer Hall, Iowa City, IA 52242, United States

HIGHLIGHTS

- We study the non-parametric estimation of the Haezendonck Risk Measure via the empirical Haezendonck risk measure.
- We provide a strong consistency result for the empirical Haezendonck risk measure.
- We also provide a weak convergence result for the empirical Haezendonck risk measure.

ARTICLE INFO

Article history:
Received August 2011
Received in revised form
October 2013
Accepted 25 December 2013

IME Classifications:
10.050: IM10
10.130: IM 30

Keywords:
Orlicz premium
Tail value-at-Risk (T-VaR)
Conditional tail expectation (CTE)
Empirical CTE

ABSTRACT

Haezendonck–Goovaerts risk measures is a recently introduced class of risk measures which includes, as its minimal member, the Tail Value-at-Risk (T-VaR)—T-VaR arguably the most popular risk measure in global insurance regulation. In applications often one has to estimate the risk measure given a random sample from an *unknown* distribution. The distribution could either be truly unknown or could be the distribution of a complex function of economic and idiosyncratic variables with the complexity of the function rendering indeterminable its distribution. Hence statistical procedures for the estimation of Haezendonck–Goovaerts risk measures are a key requirement for their use in practice. A natural estimator of the Haezendonck–Goovaerts risk measure is the Haezendonck–Goovaerts risk measure of the empirical distribution, but its statistical properties have not yet been explored in detail. The main goal of this article is to both establish the strong consistency of this estimator and to derive weak convergence limits for this estimator. We also conduct a simulation study to lend insight into the sample sizes required for these asymptotic limits to take hold.

1. Introduction

Haezendonck–Goovaerts risk measures is a class of risk measures which was recently introduced in Goovaerts et al. (2004). It is based on the premium calculation principle induced by an Orlicz norm, as presented in Haezendonck and Goovaerts (1982) (see also Goovaerts et al., 2003), in the sense that it is the translation-equivariant minimal Orlicz risk measure. This class of risk measures was further studied in Bellini and Gianin (2008a), and they present an alternate formulation of these risk measures which makes them *coherent* in the sense of Artzner et al. (1997, 1999). It is worth mentioning that the Haezendonck–Goovaerts risk measures preserve the convex order; see Goovaerts et al. (2004), Bellini and Gianin (2008a), Nam et al. (2011), Ahn and Shyamalkumar (2011b) and Tang and Yang (2012) for further properties of these risk measures. The most prominent member of this class, and in fact its minimal member, is the Tail Value-at-Risk (T-VaR)—T-VaR arguably the most popular risk measure in the global insurance regulation.

In applications often one has to estimate the risk measure given a random sample from an *unknown* distribution. The distribution could either be truly unknown or could be the distribution of a complex function of economic and idiosyncratic variables with the complexity of the function rendering indeterminable its distribution. For example, under the US statutory (NAIC) regulations, for variable annuities both valuation under AG43 (VACARVM) and capital calculations under C3 Phase II require the estimation of T-VaR using data from stochastic simulations. The data used to calculate T-VaR for the portfolio is i.i.d. in nature from an unknown distribution. Soon, such calculations will be required for life insurance products as well. It is noteworthy that insurance companies are spending valuable resources to conduct these exercises efficiently. Our interest in the broader problem of developing statistical inference procedures for the estimation of risk measures is driven by

* Corresponding author.
** Corresponding author. Tel.: +1 319 335 1980; fax: +1 319 335 3017.
E-mail addresses: jaeyahn@ewha.ac.kr (J.Y. Ahn), shyamal-kumar@uiowa.edu (N.D. Shyamalkumar).

0167-6687/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.insmatheco.2013.12.003

such applications. The above, in particular, implies that good statistical procedures for the estimation of Haezendonck–Goovaerts risk measures are a key requirement for its use in practice.

An interesting alternate formulation of the problem is that of Krätschmer and Zähle (2011) where they study estimating the risk measure evaluated at the distribution of a sum of large number of identically distributed but possibly weakly dependent variables. In this case, one expects that this distribution will be close to normal, and they study the quality of estimating the risk via the normal approximation. This setup does not cover applications of the type mentioned above as while the output from a single simulation run is the sum of a large number of random variables, these random variables are both non-identically distributed and exhibit strong dependence. Hence, the output from a single simulation run is far from normal.

While the above references study properties of the Haezendonck–Goovaerts risk measures, only Bellini and Gianin (2008b) deals with its statistical estimation. Note that the natural non-parametric estimator for the Haezendonck–Goovaerts risk measure is its empirical analog, the Haezendonck–Goovaerts risk measure of the empirical distribution. In Bellini and Gianin (2008b) the authors conduct a simulation study of this estimation procedure, and also use it to estimate the efficient frontier when the risk is measured by a Haezendonck–Goovaerts risk measure. While this study suggests, in some cases, a normal asymptotic limit for this estimator, neither the consistency nor the weak convergence of this estimator has been established. This then is the main goal of this article; we provide a strong consistency and a weak convergence result for this non-parametric estimator with the latter also covering situations with a non-normal limit. The difficulty in establishing asymptotic results arises in good part from the lack of a convenient closed form expression for the Haezendonck–Goovaerts risk measure of the empirical distribution function.

We note that significant work has been done for the estimation of another class of risk measures referred to as distortion risk measures (introduced in Wang, 1996), see Beutner and Zähle (2010), Jones and Zitikis (2003) and references therein. This class of risk measures includes coherent risk measures such as T-VaR, and incoherent (albeit popular) risk measures such as VaR as well. Study of the weak limits of estimators of such risk measures, unlike the Haezendonck–Goovaerts risk measures, is greatly facilitated by the existence of closed form expressions for the risk measure, and that plug-in estimators are linear functions of order statistics. For a discussion of distortion risk measures and a connection between a subset of Haezendonck–Goovaerts risk measures and distortion risk measures, see Krätschmer and Zähle (2011). The results in this paper are established for an arbitrary Haezendonck–Goovaerts risk measure under near ideal conditions.

After some definitions and establishing the notation we provide an example which demonstrates this inherent nature of our problem, and another that demonstrates that non-normal asymptotic weak limits occur even in non-pathological situations.

A non-negative, strictly increasing, convex function $\Psi(\cdot)$ on $\mathbb{R}^+$ with $\Psi(0) = 0$ and $\Psi(1) = 1$ is called a normalized Young function (see Rao and Ren, 1991 for details). In the following we will work with the extensions of such functions to the whole of $\mathbb{R}$ satisfying $\Psi(x) = 0$ for $x < 0$. For convenience we will refer to such extensions simply as Young function. The class of Haezendonck–Goovaerts risk measures is indexed by the class of Young function, and for each Haezendonck–Goovaerts risk measure there exists a class of random variables for which it is well defined. A subset of this class of random variables is denoted by $\mathbb{X}_\Psi$, and is defined by

$$\mathbb{X}_\Psi := \left\{ X \,\middle|\, \Pr(X \le 0) = 1 \text{ or } \exists s_\infty \ge 0 \text{ such that } \times \mathbb{E}\left(\Psi\left(\frac{X}{s}\right)\right) < \infty, \text{ for } s > 0 \Leftrightarrow s > s_\infty \right\}. \tag{1}$$

In Bellini and Gianin (2008a), for convenience, the random variables were restricted to L^∞, the space of essentially bounded random variables, a subset of $\mathbb{X}_\Psi$; we allow s_∞ to be greater than 0 in (1), unlike in Goovaerts et al. (2004), to accommodate situations like those in Example 5 where $\Psi(\cdot)$ is exponential and X is an exponential random variable.

The Orlicz premium principle corresponding to $\Psi(\cdot)$, and at level $\alpha \in [0, 1)$, is denoted by $H_\Psi^\alpha(\cdot)$, and for $X \in \mathbb{X}_\Psi$, $H_\Psi^\alpha(X)$ is defined as the unique positive solution of the equation

$$\mathbb{E}\left(\Psi\left(\frac{X}{H_\Psi^\alpha(X)}\right)\right) = 1 - \alpha, \quad \text{for } \Pr(X > 0) > 0,$$

with $H_\Psi^\alpha(X) := 0$ if $\Pr(X \le 0) = 1$ (see Haezendonck and Goovaerts, 1982, Goovaerts et al., 2004, and Bellini and Gianin, 2008a). For $X \in \mathbb{X}_\Psi$, following Bellini and Gianin (2008a) and Goovaerts et al. (2004), we define the Haezendonck–Goovaerts risk measure at level $\alpha \in [0, 1)$, denoted by $\pi_\Psi^\alpha(X)$, as

$$\pi_\Psi^\alpha(X) := \inf_{x \in \mathbb{R}} \left(H_\Psi^\alpha(X - x) + x\right). \tag{2}$$

For $X \in L^\infty$, Proposition 16 of Bellini and Gianin (2008a) shows that the above infimum is attained for $\alpha \in (0, 1)$; their argument is easily extended to $\mathbb{X}_\Psi$. Moreover, examples exist where this infimum is not attained when $\alpha = 0$. Along the lines of Example 15 of Bellini and Gianin (2008a), one such example is when $\Psi(\cdot)$ and $F(\cdot)$ are defined as

$$\Psi(x) = \begin{cases} 0, & x < 0; \\ x^{2k}, & \text{otherwise}, \end{cases} \quad \text{where } k \ge 1, \quad \text{and}$$

$$F(x) = \begin{cases} 0, & x < -1; \\ \frac{1}{2}, & -1 \le x < 1; \\ 1, & \text{otherwise}. \end{cases}$$

For this reason, and also since for risk management purposes it is only the high values of α that are of interest, in the following we will restrict our attention to $\alpha \in (0, 1)$ when working with Haezendonck–Goovaerts risk measures.

For convenience we define $\pi_\Psi^\alpha(X; \cdot)$ as

$$\pi_\Psi^\alpha(X; x) := \left(H_\Psi^\alpha(X - x) + x\right), \quad x \in \mathbb{R}. \tag{3}$$

and denote by $\mathcal{I}_\Psi^\alpha(F)$, where F denotes the distribution of X, the set of minimizers of $\pi_\Psi^\alpha(X, \cdot)$. In Bellini and Gianin (2008a) it is shown that $\pi_\Psi^\alpha(X; \cdot)$ is a convex function for $X \in L^\infty$, and we note that this result too can be easily extended to $X \in \mathbb{X}_\Psi$. This extension in particular implies that $\mathcal{I}_\Psi^\alpha(F)$ is a closed interval.

From now on $X, X_1, X_2, \ldots$ will denote a sequence of independently and identically distributed random variables on our underlying probability space $(\Omega, \mathcal{F}, P)$ with $X \in \mathbb{X}_\Psi$. Also, $F(\cdot)$ we will denote their common distribution function. Note that by the definition of the Haezendonck–Goovaerts risk measure, we could define $H_\Psi^\alpha(F)$, $\pi_\Psi^\alpha(F)$ and $\pi_\Psi^\alpha(F; \cdot)$ to equal $H_\Psi^\alpha(X)$, $\pi_\Psi^\alpha(X)$ and $\pi_\Psi^\alpha(X; \cdot)$, respectively. By $F_n(\cdot)$ we will denote the empirical distribution function of the random sample of size n consisting of $X_1, \ldots, X_n$, *i.e.*

$$F_n(x) := \frac{1}{n} \sum_{i=1}^{n} \mathrm{I}_{X_i \le x}, \quad x \in \mathbb{R}.$$

We denote by $\mathbb{E}_n(g(Y))$ the expectation of $g(Y)$ with $Y \sim F_n$. As for such Y we have $Y \in L^\infty$ $(\subseteq \mathbb{X}_\Psi)$, $H_\Psi^\alpha(F_n)$ and $\pi_\Psi^\alpha(F_n)$ are both well defined, and moreover are easily seen to be a random variable defined on $(\Omega, \mathcal{F}, P)$. Note that $H_\Psi^\alpha(F_n)$ and $\pi_\Psi^\alpha(F_n)$ are natural