Insurance: Mathematics and Economics 76 (2017) 28–47

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

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

Haezendonck–Goovaerts risk measure with a heavy tailed loss

Qing Liu [a], Liang Peng [b,*], Xing Wang [b]

[a] School of Statistics and Research Center of Applied Statistics, Jiangxi University of Finance and Economics, Nanchang, Jiangxi 330013, PR China
[b] Department of Risk Management and Insurance, Georgia State University, Atlanta, GA 30303, USA

ARTICLE INFO

Article history:
Received October 2016
Received in revised form June 2017
Accepted 12 June 2017
Available online 8 July 2017

Keywords:
Empirical process
Fixed level
Haezendonck–Goovaerts risk measure
Heavy tail
Intermediate level

ABSTRACT

Recently Haezendonck–Goovaerts (H–G) risk measure has received much attention in (re)insurance and portfolio management. Some nonparametric inferences have been proposed in the literature. When the loss variable does not have enough moments, which depends on the involved Young function, the nonparametric estimator in Ahn and Shyamalkumar (2014) has a nonnormal limit, which challenges interval estimation. Motivated by the fact that many loss variables in insurance and finance could have a heavier tail such as an infinite variance, this paper proposes a new estimator which estimates the tail by extreme value theory and the middle part nonparametrically. It turns out that the proposed new estimator always has a normal limit regardless of the tail heaviness of the loss variable. Hence an interval with asymptotically correct confidence level can be obtained easily either by the normal approximation method via estimating the asymptotic variance or by a bootstrap method. A simulation study and real data analysis confirm the effectiveness of the proposed new inference procedure for estimating the H–G risk measure.

1. Introduction

Risk management generally involves risk identification, risk quantification, and risk prediction. Measuring a risk and quantifying its uncertainty is an important task. Recently a so-called Haezendonck–Goovaerts (H–G) risk measure has received much attention in actuarial science with applications to optimal portfolio management and optimal reinsurance policy; see Bellini and Gianin (2008a, b), Cheung and Lo (2013), Zhu et al. (2013), and references therein.

Let $\psi : [0, \infty] \to [0, \infty]$ be a convex function satisfying $\psi(0) = 0, \psi(1) = 1$ and $\psi(\infty) = \infty$, i.e., ψ is a so-called normalized Young function. For a number $q \in (0, 1)$ and each $\beta > 0$, let $\alpha = \alpha(\beta)$ be a solution to

$$E\left\{\psi\left(\frac{(X-\beta)_+}{\alpha}\right)\right\} = 1-q, \tag{1}$$

where $x_+ = \max(x, 0)$. Then, Haezendonck and Goovaerts (1982) proposed the so-called H–G risk measure at level q as

$$\theta = \inf_{\beta>0}\{\beta + \alpha(\beta)\}. \tag{2}$$

Some important properties and connections with other risk measures are given in Goovaerts et al. (2012). For example, if $\psi(x) = x$, then $\alpha(\beta) = \frac{1}{1-q}E\{(X-\beta)_+\}$ and $\theta = \frac{1}{1-q}E\{(X - F^{-}(q))_+\}$, where $F(x) = P(X \le x)$ and $F^{-}(x)$ denotes the inverse function of $F(x)$. Hence, in this case, the H–G risk measure equals the expected shortfall.

In order to employ this risk measure in practice, an efficient statistical inference is needed. Ahn and Shyamalkumar (2014) first proposed a nonparametric estimation and derived its asymptotic limit, which may be nonnormal when the loss variable has no enough moments, which depends on the involved Young function ψ. When the limit is normal, Peng et al. (2015) developed an empirical likelihood method to effectively construct an interval when the H–G risk measure is defined at a fixed level. Further, Wang and Peng (2016) showed that this empirical likelihood method is still valid for an intermediate level, which leads to a unified interval estimator of the H–G risk measure at either a fixed level or an intermediate level. We refer to Owen (2001) for an overview of empirical likelihood methods, which has been shown to be quite effective in interval estimation and hypothesis test. Properties of the H–G risk measure at an extreme level are available in Tang and Yang (2012), Tang and Yang (2014) and Mao and Hu (2012).

To better understand the inference issue, we formulate the H–G risk measure as a solution to the following estimating equations. Suppose $X, X_1, \ldots, X_n$ are independent and identically distributed random variables with distribution function $F(x)$. Since the H–G risk measure is equivalent to solving the following estimating

* Corresponding author.
E-mail address: lpeng@gsu.edu (L. Peng).

http://dx.doi.org/10.1016/j.insmatheco.2017.06.004
0167-6687/

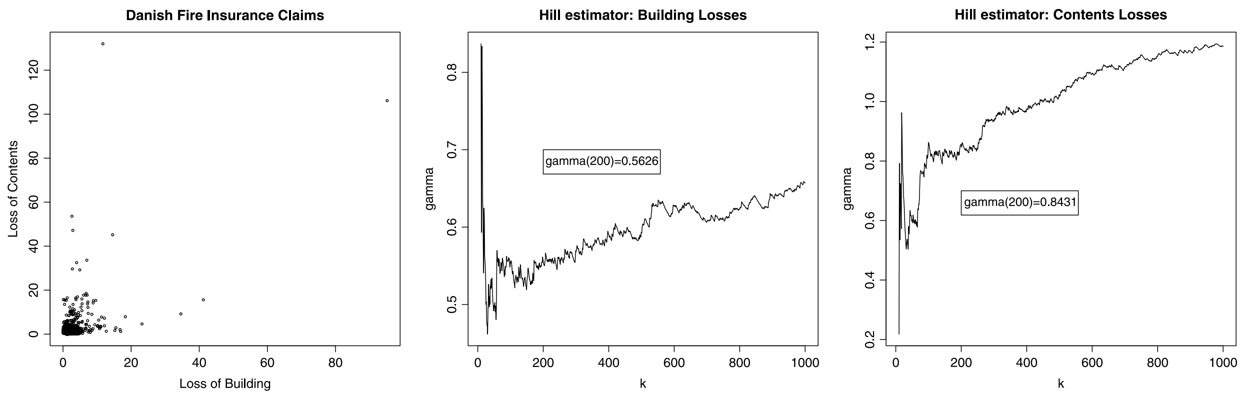

Fig. 1. Left panel: Danish fire losses to building and contents; Middle panel: Hill estimate for losses to building; Right panel: Hill estimate for losses to contents.

equations

$$\begin{cases} E\left\{\psi\left(\frac{X_i-\beta}{\theta-\beta}\right)I(X_i>\beta)\right\}=1-q, \\ E\left\{\psi'\left(\frac{X_i-\beta}{\theta-\beta}\right)(X_i-\theta)I(X_i>\beta)\right\}=0 \end{cases} \tag{3}$$

for some β and $\theta>\beta$ under some conditions (see Tang and Yang 2014), one can estimate β and θ by solving

$$\begin{cases} \frac{1}{n}\sum_{i=1}^{n}\psi\left(\frac{X_i-\beta}{\theta-\beta}\right)I(X_i>\beta)=1-q, \\ \frac{1}{n}\sum_{i=1}^{n}\psi'\left(\frac{X_i-\beta}{\theta-\beta}\right)(X_i-\theta)I(X_i>\beta)=0, \end{cases} \tag{4}$$

which will result in a nonnormal limit when either

$$E\left\{\psi\left(\frac{X_i-\beta_0}{\theta_0-\beta_0}\right)I(X_i>\beta_0)\right\}^2=\infty \quad \text{or} \\ E\left\{\psi'\left(\frac{X_i-\beta_0}{\theta_0-\beta_0}\right)(X_i-\theta_0)I(X_i>\beta_0)\right\}^2=\infty, \tag{5}$$

where θ_0 and β_0 denote the true values of θ and β, respectively. This makes interval estimation nontrivial since one has to employ different methods to separately deal with the cases of having a normal limit and a nonnormal limit.

Practically it is often observed that loss data in insurance have a heavy tailed distribution and even have an infinite variance, which implies that (5) holds quite frequently. Particularly this paper is motivated by analyzing the Danish fire loss data (see left panel in Fig. 1), which consists of losses to building and losses to contents. The data were collected at the Copenhagen Reinsurance Company and comprise 2167 fire losses over the period 1980 to 1990. By assuming that

$$\lim_{t\to\infty}\frac{1-F(tx)}{1-F(t)}=x^{-1/\gamma} \quad \text{for } x>0, \tag{6}$$

i.e., $1-F$ has a heavy tail with tail index $1/\gamma$, γ can be estimated by the well-known Hill estimator

$$\hat{\gamma}(k)=\frac{1}{k}\sum_{i=1}^{k}\log\frac{X_{n,n-i+1}}{X_{n,n-k}}, \tag{7}$$

where $X_{n,1}\le\cdots\le X_{n,n}$ denote the order statistics of $X_1,\ldots,X_n$, $k=k(n)\to\infty$ and $k/n\to 0$ as $n\to\infty$; see Hill (1975) for details. Note that (6) implies that $EX_+^d<\infty$ for $d<1/\gamma$ and $EX_+^d=\infty$ for $d>1/\gamma$. Moreover (6) holds for many commonly used loss distributions in insurance such as Pareto distribution, inverse gamma distribution, student t distribution, Cauchy distribution, Burr distribution, Log-gamma distribution, etc. The middle and right panels in Fig. 1 show that γ is between 0.5 and 1, which implies that $EX_+<\infty$ but $EX_+^2=\infty$. Therefore, when $\psi(x)=\psi_r(x)=x^r$ with some $r>1$, the nonparametric estimator of the H–G risk measure based on (4) has a nonnormal limit, which makes interval estimation nontrivial and it generally requires a subsample bootstrap method.

Motivated by the idea of estimating the mean of a heavy-tailed distribution in Peng (2001) and Peng (2004) and the expected shortfall of a heavy-tailed loss variable in Necir and Meraghni (2009), this paper proposes to separately estimate the expectations in (3) by two parts: semi-parametric estimation for the tail and nonparametric estimation for the middle part. It turns out the proposed new estimator will always have a normal limit regardless of the tail heaviness of X. Hence interval estimation can be done by using either the normal approximation method via estimating the asymptotic variance or a bootstrap method. In the simulation study and data analysis below, we simply employ the naive bootstrap method, i.e., resample directly from original data, and a comparison study shows that a blind application of methods without considering a nonnormal limit would forecast risk inaccurately.

We organize this paper as follows. Section 2 presents the new methodologies and main results for estimating the H–G risk measure at both a fixed level and an intermediate level. A simulation study is given in Section 3. Analysis of the Danish fire loss data is presented in Section 4. All proofs are put in Section 5.

2. Methodologies and main results

Throughout we assume $X, X_1,\ldots,X_n$ are independent and identically distributed random variables with distribution function F satisfying (6), and

$$\begin{cases} \psi(x) \text{ is a normalized Young function with } \psi'(0)<\infty \text{ and} \\ \text{continuous second derivatives on } (0,\infty), \text{ and satisfies} \\ \lim_{x\to\infty}\frac{\psi''(x)}{r(r-1)x^{r-2}}=d_0>0 \text{ for some } r>1. \end{cases} \tag{8}$$

Since we want to estimate the tail semiparametrically, it is necessary to specify an approximation rate in (6) as usual in the context of extreme value theory for controlling the bias of an employed tail probability estimator. Put $\bar{F}(x)=1-F(x)$ and let $\bar{F}^-(t)$ denote the inverse function of $\bar{F}(t)$. Then it is known that (6) is equivalent to

$$\lim_{t\to 0}\frac{\bar{F}^-(tx)}{\bar{F}^-(t)}=x^{-\gamma} \quad \text{for } x>0.$$

Hence we assume there exists a function $A(t)\to 0$ with a constant sign near zero such that

$$\lim_{t\to 0}\frac{\frac{\bar{F}^-(tx)}{\bar{F}^-(t)}-x^{-\gamma}}{A(t)}=x^{-\gamma}\frac{x^{\rho}-1}{\rho} \tag{9}$$