Insurance: Mathematics and Economics 74 (2017) 99–108

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

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

Bootstrap consistency and bias correction in the nonparametric estimation of risk measures of collective risks

Alexandra Lauer, Henryk Zähle *

Department of Mathematics, Saarland University, Germany

ARTICLE INFO

Article history:
Received September 2016
Received in revised form February 2017
Accepted 4 March 2017
Available online 10 March 2017

Keywords:
Law-invariant risk measure
Collective risk
Nonparametric estimation
Bootstrap consistency
Weighted exchangeable bootstrap
Bootstrap-based bias correction
Value at Risk
Average Value at Risk

ABSTRACT

We consider two nonparametric estimators for the risk measure of the sum of n i.i.d. individual insurance risks divided by n, where the number of historical single claims that are used for the statistical estimation is of order n. This framework matches the situation that nonlife insurance companies are faced with within the scope of premium calculation. Indeed, the risk measure of the collective risk divided by n can be seen as a suitable premium for each of the individual risks. For both estimators asymptotic normality has been obtained recently. Here we derive almost sure bootstrap consistency for both estimators, where we allow for the weighted exchangeable bootstrap and rather general law-invariant risk measures. Both estimators are subject to a relevant negative bias for small to moderate n. For one of them we investigate by means of numerical experiments the benefit of a bootstrap-based bias correction. The numerical experiments are performed for the Value at Risk and the Average Value at Risk, and the results are comparable to those of Kim and Hardy (2007) who did analogous experiments for classical nonparametric plug-in estimators. For the other estimator the benefit of a bootstrap-based bias correction can be ruled out by theoretical arguments.

1. Introduction

Let (X_i) be a sequence of nonnegative i.i.d. random variables on a common probability space with distribution μ. In the context of actuarial theory, the random variable $S_n := \sum_{i=1}^n X_i$ can be seen as the total claim of a homogeneous insurance collective consisting of n risks. This corresponds to the individual risk model, and one should therefore assume that μ has large point mass at 0. The distribution of S_n is given by the n-fold convolution μ^{*n} of μ. A central task in insurance practice is the specification of the premium $\mathcal{R}_\rho(\mu^{*n})$ for the collective risk S_n, where $\mathcal{R}_\rho$ is the statistical functional associated with any suitable law-invariant risk measure ρ. Note that

$$\mathcal{R}_n := \frac{1}{n}\mathcal{R}_\rho(\mu^{*n}) \tag{1}$$

can be seen as a suitable premium for each of the individual risks $X_1, \ldots, X_n$, where it is important to note that most often $\frac{1}{n}\mathcal{R}_\rho(\mu^{*n})$ is essentially smaller than $\mathcal{R}_\rho(\mu)$.

In Krätschmer and Zähle (2011) and Lauer and Zähle (2015) the nonparametric estimators

$$^{\mathrm{NA}}\widehat{\mathcal{R}}_n := \frac{1}{n}\mathcal{R}_\rho(\mathcal{N}_{n\widehat{m}_{u_n}, n\widehat{s}^2_{u_n}}) \quad \text{and} \quad ^{\mathrm{CE}}\widehat{\mathcal{R}}_n := \frac{1}{n}\mathcal{R}_\rho(\widehat{\mu}^{*n}_{u_n}) \tag{2}$$

for the individual premium $\mathcal{R}_n$ based on observed historical single claims $Y_1, \ldots, Y_{u_n}$ were studied, where the Y_i are assumed to be i.i.d. random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with distribution μ. In (2) and in the rest of the paper, we use $\widehat{m}_{u_n}$, $\widehat{s}^2_{u_n}$, and $\widehat{\mu}^{*n}_{u_n}$ to denote the empirical mean $\frac{1}{u_n}\sum_{i=1}^{u_n} Y_i$, the empirical variance $\frac{1}{u_n}\sum_{i=1}^{u_n}(Y_i - \widehat{m}_{u_n})^2$, and the n-fold convolution of the empirical probability measure $\widehat{\mu}_{u_n} := \frac{1}{u_n}\sum_{i=1}^{u_n}\delta_{Y_i}$ of the observed historical single claims, respectively. Recall that $\widehat{m}_{u_n}$, $\widehat{s}^2_{u_n}$, and $\widehat{\mu}_{u_n}$ are the standard nonparametric estimators for the mean m of μ, the variance s^2 of μ, and μ itself, respectively. Moreover $\mathcal{N}_{m,s^2}$ stands for the normal distribution with mean m and variance s^2. The former estimator in (2) is motivated by the central limit theorem, and the latter one by the Glivenko–Cantelli theorem. For computational aspects of the latter estimator in (2) see Lauer and Zähle (2015), Appendix A.

On the one hand, it was shown in Krätschmer and Zähle (2011) and Lauer and Zähle (2015) that for many risk measures ρ and under

$$\lim_{n\to\infty} u_n/n = c \quad \text{for some constant } c \in (0,\infty) \tag{3}$$

and some mild assumptions on μ the estimators in (2) are strongly consistent in the sense that

$$^{\mathrm{NA}}\widehat{\mathcal{R}}_n - \mathcal{R}_n \longrightarrow 0 \quad \mathbb{P}\text{-a.s.} \qquad \text{and}$$
$$^{\mathrm{CE}}\widehat{\mathcal{R}}_n - \mathcal{R}_n \longrightarrow 0 \quad \mathbb{P}\text{-a.s.} \tag{4}$$

* Corresponding author.
E-mail addresses: lauer@math.uni-sb.de (A. Lauer), zaehle@math.uni-sb.de (H. Zähle).

http://dx.doi.org/10.1016/j.insmatheco.2017.03.001

and asymptotically normal in the sense that

$$\mathbb{P}\circ\left\{\sqrt{u_n}({}^{\mathsf{NA}}\widehat{\mathcal{R}}_n-\mathcal{R}_n)\right\}^{-1}\stackrel{\mathsf{w}}{\longrightarrow}\mathcal{N}_{0,s^2}\quad\text{and}$$
$$\mathbb{P}\circ\left\{\sqrt{u_n}({}^{\mathsf{CE}}\widehat{\mathcal{R}}_n-\mathcal{R}_n)\right\}^{-1}\stackrel{\mathsf{w}}{\longrightarrow}\mathcal{N}_{0,s^2}. \tag{5}$$

Here and in the rest of the paper $\stackrel{\mathsf{w}}{\longrightarrow}$ refers to weak convergence. Condition (3) is motivated by the fact that the premium is typically estimated on the basis of the historical claims of the same collective from the last year or from the last few years. Note that this condition is somehow nonstandard, because in the literature on asymptotic statistical inference for convolutions it is usually assumed that the number of summands n is fixed and the number of observations u tends to infinity; see, for instance, Pitts (1994).

On the other hand, in Lauer and Zähle (2015) it was demonstrated by means of Monte Carlo simulations that the estimators in (2) are subject to a negative bias for finite sample size n. In particular when the conditional single claim distribution $\mu_{>0}[\,\cdot\,]:=\mu[\,\cdot\cap(0,\infty)]/\mu[(0,\infty)]$ is "heavy-tailed" the bias can be considerable. A conventional method for correcting the bias of an estimator is based on "the" bootstrap. We will recall the fundamental idea of this method in Section 2. For background see also Section 10.6 in Efron and Tibshirani (1994).

In the paper at hand, we address the question whether the biases of the estimators ${}^{\mathsf{NA}}\widehat{\mathcal{R}}_n$ and ${}^{\mathsf{CE}}\widehat{\mathcal{R}}_n$ for the individual premium $\mathcal{R}_n$ can be reduced by means of the bootstrap technique to be explained in Section 2. For the estimator $\mathcal{R}_\rho(\widehat{\mu}_n)$ of $\mathcal{R}_\rho(\mu)$ analogous investigations have been done by Kim and Hardy (2007) for the Value at Risk and the Average Value at Risk, and by Kim (2010) for more general distortion risk measures. Ahn and Shyamalkumar (2010) provided some asymptotic analysis for the Average Value at Risk in this context. Part (iii) of Remark 3.3 indicates that the bootstrap approach for reducing the bias is *not* expedient for the estimator ${}^{\mathsf{NA}}\widehat{\mathcal{R}}_n$. On the other hand, the bootstrap approach can be (slightly) useful for ${}^{\mathsf{CE}}\widehat{\mathcal{R}}_n$. In our numerical examples for ${}^{\mathsf{CE}}\widehat{\mathcal{R}}_n$ with ρ the Value at Risk and the Average Value at Risk (see Section 4), we obtain results that are qualitatively comparable to the numerical results of Kim and Hardy (2007) and Kim (2010). Whereas for the Value at Risk an application of the bootstrap-based method of Section 2 seems not useful, for the Average Value at Risk we can observe that on average a small to moderate reduction of the bias goes along with a small increase of the variance (and of the mean squared error).

In the framework of Ahn and Shyamalkumar (2010), Kim and Hardy (2007) and Kim (2010) the plug-in estimator $\mathcal{R}_\rho(\widehat{\mu}_n)$ for a distortion risk measure $\mathcal{R}_\rho(\mu)$ is an L-statistic, and thus bootstrap consistency is known from the literature. According to Gribkova (2016), Theorem 7 of Gribkova (2002) applies, at least for distortion functions that are piecewise differentiable with Hölder continuous derivative. For the Average Value at Risk functional, see also Corollary 4.2 in Beutner and Zähle (2016). Moreover, for L-statistics even the exact bootstrap mean can be calculated explicitly (Hutson and Ernst, 2000). In our setting, where the individual premium $\mathcal{R}_n=\mathcal{R}_\rho(\mu^{*n})/n$ is estimated by ${}^{\mathsf{CE}}\widehat{\mathcal{R}}_n$, bootstrap results seem not to exist so far (up to the best of our knowledge). For this reason we will derive in Section 3 a result on bootstrap consistency for this estimator to give a mathematical justification for the use of the bootstrap-based method of Section 2. Theorem 3.2 is the theoretical contribution of our article. Its proof will be carried out in Section 5. Although the method of Section 2 seems not to be appropriate for the estimator ${}^{\mathsf{NA}}\widehat{\mathcal{R}}_n$ (see part (iii) of Remark 3.3), in Theorem 3.2 we also establish bootstrap consistency for this estimator. In Section 4 we will present the results of some numerical experiments.

2. Bootstrap-based bias correction

As mentioned in the Section 1 the estimators defined in (2) have a negative bias w.r.t. $\mathcal{R}_n$. As a countermeasure one can try to "estimate" the bias and subtract it from the original estimator. The "estimation" of the bias can sometimes be done by means of bootstrap methods. The idea of the bootstrap was introduced by Efron in 1979 in his seminal paper (Efron, 1979). Since then many variants of the bootstrap have been discussed in the literature; for background and details one may refer to Davison and Hinkley (1997), Efron and Tibshirani (1994), Lahiri (2003) and Shao and Tu (1995) among others.

To explain the bootstrap-based method for correcting the bias more precisely, let $\widehat{\mathcal{R}}_n$ be an estimator for a real-valued characteristic $\mathcal{R}_n$, $n\in\mathbb{N}$, where $\mathcal{R}_n$ may or may not be defined by (1). In any case assume that $\widehat{\mathcal{R}}_n$ is given by a statistical functional evaluated at a (random) probability measure which is uniquely determined by observed data $Y_1,\ldots,Y_{u_n}$, where the latter are given by the first u_n terms of a sequence (Y_i) of i.i.d. random variables defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$. For illustrations of such estimators see (2). Assume that $\widehat{\mathcal{R}}_n$ is biased, i.e. that

$$\mathrm{Bias}(\widehat{\mathcal{R}}_n):=\mathbb{E}[\widehat{\mathcal{R}}_n-\mathcal{R}_n] \tag{6}$$

differs from 0 for finite sample size n. Further assume that

$$\mathbb{P}\circ\left\{\sqrt{u_n}(\widehat{\mathcal{R}}_n-\mathcal{R}_n)\right\}^{-1}\stackrel{\mathsf{w}}{\longrightarrow}\mathcal{N}_{0,s^2} \tag{7}$$

holds for some $s^2\in(0,\infty)$. See (5) for an illustration of condition (7). Now extend the original probability space $(\Omega,\mathcal{F},\mathbb{P})$ to the product $(\overline{\Omega},\overline{\mathcal{F}},\overline{\mathbb{P}}):=(\Omega\times\Omega',\mathcal{F}\otimes\mathcal{F}',\mathbb{P}\otimes\mathbb{P}')$ with any other probability space $(\Omega',\mathcal{F}',\mathbb{P}')$, and assume that the result ω' of $(\Omega',\mathcal{F}',\mathbb{P}')$ and the original sample $Y_1(\omega),\ldots,Y_{u_n}(\omega)$ specify a new (random) probability measure. The latter is plugged in the underlying statistical functional to obtain a "bootstrap version" of $\widehat{\mathcal{R}}_n$, denoted by $\widehat{\mathcal{R}}_n^{\mathsf{B}}$. Note that $\widehat{\mathcal{R}}_n^{\mathsf{B}}$ depends on ω and ω', that is, it is defined on the probability space $(\overline{\Omega},\overline{\mathcal{F}},\overline{\mathbb{P}})$. Also note that, up to some measurability issues, the mapping $\omega'\mapsto\widehat{\mathcal{R}}_n^{\mathsf{B}}(\omega,\omega')$ can be seen as a random variable on $(\Omega',\mathcal{F}',\mathbb{P}')$ for any fixed ω. For illustrations of $\widehat{\mathcal{R}}_n^{\mathsf{B}}$ see (12) and (14). In fact $\widehat{\mathcal{R}}_n^{\mathsf{B}}$ should be called (almost sure) bootstrap version of $\widehat{\mathcal{R}}_n$ only if

$$\mathbb{P}'\circ\left\{\sqrt{u_n}\big(\widehat{\mathcal{R}}_n^{\mathsf{B}}(\omega,\,\cdot\,)-\widehat{\mathcal{R}}_n(\omega)\big)\right\}^{-1}\stackrel{\mathsf{w}}{\longrightarrow}\mathcal{N}_{0,s^2}\qquad\mathbb{P}\text{-a.e. }\omega. \tag{8}$$

The left-hand side of (8) is often referred to as the conditional distribution of $\sqrt{u_n}(\widehat{\mathcal{R}}_n^{\mathsf{B}}-\widehat{\mathcal{R}}_n)$ given the observation $Y_1,\ldots,Y_{u_n}$. For a justification of this interpretation see, for instance, the discussion at the end of Section 2 in Beutner and Zähle (submitted for publication).

Whenever (7) and (8) can be shown, we have

$$\mathbb{P}\circ\left\{\widehat{\mathcal{R}}_n-\mathcal{R}_n\right\}^{-1}\approx\mathcal{N}_{0,s^2/u_n}$$

and

$$\mathbb{P}'\circ\left\{\widehat{\mathcal{R}}_n^{\mathsf{B}}(\omega,\,\cdot\,)-\widehat{\mathcal{R}}_n(\omega)\right\}^{-1}\approx\mathcal{N}_{0,s^2/u_n}\qquad\mathbb{P}\text{-a.e. }\omega$$

for "large" n. That is, informally,

$$\mathbb{P}\circ\left\{\widehat{\mathcal{R}}_n-\mathcal{R}_n\right\}^{-1}\approx\mathbb{P}'\circ\left\{\widehat{\mathcal{R}}_n^{\mathsf{B}}(\omega,\,\cdot\,)-\widehat{\mathcal{R}}_n(\omega)\right\}^{-1}\qquad\mathbb{P}\text{-a.e. }\omega \tag{9}$$

for "large" n. Sometimes it turns out that the two laws in (9) are not only "close" but even have a similar skewness so that the means of these two laws are close to each other. In this case the mean of the law on the right-hand side of (9) is a reasonable approximation of $\mathrm{Bias}(\widehat{\mathcal{R}}_n)$ defined in (6). Though the law on the right-hand side of (9) can be seldom specified explicitly, it can be numerically approximated through

$$\frac{1}{L}\sum_{\ell=1}^{L}\delta_{\widehat{\mathcal{R}}_n^{\mathsf{B},\ell}(\omega,\,\cdot\,)-\widehat{\mathcal{R}}_n(\omega)}\qquad\text{with }L\gg n$$