Insurance: Mathematics and Economics 52 (2013) 255–262

Contents lists available at SciVerse ScienceDirect

Insurance: Mathematics and Economics

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

A nonparametric approach to calculating value-at-risk

Ramon Alemany, Catalina Bolancé, Montserrat Guillén *

Department of Econometrics, Riskcenter-IREA, University of Barcelona, Av. Diagonal, 690, 08034 Barcelona, Spain

ARTICLE INFO

Article history:
Received August 2012
Received in revised form October 2012
Accepted 25 December 2012

Keywords:
Kernel estimation
Bandwidth selection
Quantile
Risk measures

ABSTRACT

A method to estimate an extreme quantile that requires no distributional assumptions is presented. The approach is based on transformed kernel estimation of the cumulative distribution function (cdf). The proposed method consists of a double transformation kernel estimation. We derive optimal bandwidth selection methods that have a direct expression for the smoothing parameter. The bandwidth can accommodate to the given quantile level. The procedure is useful for large data sets and improves quantile estimation compared to other methods in heavy tailed distributions. Implementation is straightforward and R programs are available.

1. Introduction

Risk measures and their mathematical properties have been widely studied in the literature (see, for instance, the books by McNeil et al. (2005) and Jorion (2007) or articles such as Dhaene et al. (2006) among many others). Most of those contributions and applications in risk management usually assume a parametric distribution for the loss random variable,[1] but deviations from parametric hypothesis can be critical in the extremes and produce inaccurate results (see, Kupiec, 1995). Krätschmer and Zähle (2011) investigated the error made even when the normal approximation is plugged in a general distribution-invariant risk measure. Alternatively, a suitable heavy tailed parametric distribution can be fitted (see, for example, McNeil et al., 2005; Jorion, 2007; Bolancé et al., 2012b). Extreme value theory can also be used to locate the tail of the distribution (see, Reiss and Thomas, 1997; Hill, 1975; Guillén et al., 2011).

Our approach is nonparametric as in Peng et al. (2012), Cai and Wang (2008) and Jones and Zitikis (2007). We propose a method to estimate quantiles that is based on a nonparametric estimate of the cumulative distribution function with an optimal bandwidth at the desired quantile level. Eling (2012) recently used a similar benchmark nonparametric fit to describe claims severity distributions in property-liability insurance (see, Bolancé et al., 2012b, for details) but the choice of the smoothing parameter needs further analysis. Besides, Eling (2012) was interested in the fit for the density of claims severity, not on risk measurement or quantiles.[2] We present the nonparametric estimation approach and focus on the bandwidth choice. We also carry out a simulation exercise.

A risk measure widely used to quantify the risk is the value-at-risk with level α. It is defined as follows,

$$\mathrm{VaR}_{\alpha}(X) = \inf\{x, F_X(x) \geq \alpha\} = F_X^{-1}(\alpha), \tag{1}$$

where X is a random variable with probability distribution function (pdf) f_X, and cumulative distribution function (cdf) F_X. Artzner et al. (1999) discussed other risk measures, but they stated that expected shortfall is preferred in practice due to its better properties, although value-at-risk is widely used in applications.

The VaR_{α} is used both as an internal risk management tool and as a regulatory measure of risk exposure to calculate capital adequacy requirements in financial and insurance institutions. In this paper we propose a method to estimate the VaR_{α} in extreme quantiles, based on transformed kernel estimation (TKE) of the cdf of losses. The proposed method consists of a double transformation kernel estimation (DTKE), and it works well for very extreme levels and a large sample size. It also improves quantile estimation compared to existing methods. An additional contribution is that we propose a simple expression for an optimal bandwidth parameter. Thus, we advocate that there is little advantage of assuming parametric distributions when calculating value-at-risk for heavy

* Corresponding author. Tel.: +34 934037039.
E-mail addresses: mguillen@ub.edu, mguillene@uoc.edu (M. Guillén).

[1] Standard industry models such as CreditRisk$^+$ are parametric. See, Fan and Gu (2003), and references therein for semiparametric models.

[2] Eling (2012) worked with two empirical data sets. The first dataset is US indemnity losses and the second is comprised of Danish fire losses. His work indicated that the transformation kernel (Bolancé et al., 2003) is the best and second best approach when compared with the parametric distributions in terms of the log likelihood value in his applications. The transformation kernel approach performed extremely well there and confirmed the results presented by Bolancé et al. (2008a) for auto insurance.

0167-6687/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
doi:10.1016/j.insmatheco.2012.12.008

tailed data, given that the nonparametric approach implementation is very straightforward and provides consistent results.

Some previous research has already studied nonparametric estimation of quantiles. On the one hand Azzalini (1981) suggested to estimate the cdf and then to obtain the quantile from its inverse function. On the other hand Harrell and Davis (1982) proposed an alternative quantile estimator, based on a weighted sum of sample observations. Later, Sheather and Marron (1990) analyzed the existing kernel methods for quantile estimation and proposed a smoothing parameter. None of those contributions, however, focused on highly skewed or heavy tailed distributions, which most often appear in financial and insurance risk management.

Recently, Swanepoel and Van Graan (2005) presented kernel estimation of a cdf using nonparametric transformation, i.e. a simple form of transformed kernel estimation. Instead, Bolancé et al. (2008b) used a parametric transformation, which provides good results in the estimation of conditional tail expectation. Here, we propose an improved nonparametric procedure to estimate the VaR_α in finance and insurance applications and derive an optimal expression for the bandwidth parameter.

A principal difference between our transformed kernel estimation and the fit of a heavy tailed parametric loss distribution is that we use sample information to estimate the parameters of an initial parametric model and, later, we also use the sample information to correct this initial fit. The proposed method works when losses have a heavy tailed distribution and it is easy to implement. It is very flexible, so it is comparable to the empirical distribution approach. We can affirm that the method proposed in this work smooths the shape of the empirical distribution and extrapolates its behavior when dealing with extremes, where data are very scarce or non existent.

The results of our simulation study show that our double transformed kernel estimation method can be applied to risk measurement and is specially suitable when the sample size is large. This is useful when basic parametric densities provide a poor fit in the tail. In the transformed kernel approach, no parametric form is imposed on the loss distribution, but, most importantly, this method avoids defining where the tail of the loss distribution starts in order to apply extreme value theory.

When writing this article, we decided to summarize basic nonparametric concepts that appear quite frequently elsewhere.[3] To make the presentation self-contained, we introduce kernel estimation notation in Section 2 and we present nonparametric estimation of a pdf and a cdf. We also describe nonparametric estimation of cdf in connection with estimation of value-at-risk. Section 3 introduces transformation kernel estimation of a cdf and a new result on its asymptotic properties. Double transformation kernel estimation of a cdf and the selection of the smoothing parameter are studied in Section 4. Section 5 presents a simulation study where we can confirm the properties of the methods proposed in the previous sections. The most relevant conclusions and a discussion are given in the last section. Implementation tools in *R* are available from the authors and detailed hands-on examples of transformation kernel estimation can be found in Bolancé et al. (2012b).

2. Nonparametric estimation of a cumulative distribution function

Let X be a random variable which represents a loss amount; its cdf is F_X. Let us assume that $X_i i = 1, \ldots, n$ denotes data observations from the loss random variable X. For instance, loss data may arise from historical simulation or they may have been generated in a Monte Carlo analysis. A natural nonparametric method to estimate cdf is the empirical distribution,

$$\widehat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le x), \tag{2}$$

where $I(\cdot) = 1$ if condition between parentheses is true. Then, the empirical estimator of value-at-risk is:

$$\mathrm{VaR}_\alpha(X) = \inf\left\{x, \widehat{F}_n(x) \ge \alpha\right\}. \tag{3}$$

Estimation of the empirical distribution is very simple, but it cannot extrapolate beyond the maximum observed data point. This is especially troublesome if the sample is not too large, and one may suspect that the probability of a loss larger than the maximum observed loss in the data sample is not zero.

Classical kernel estimation (CKE) of cdf F_X is obtained by integration of the classical kernel estimation of its pdf f_X. By means of a change of variable, the usual expression for the kernel estimator of a cdf is obtained:

$$\begin{aligned}\widehat{F}_X(x) &= \int_{-\infty}^{x} \widehat{f}_X(u)du = \int_{-\infty}^{x} \frac{1}{nb}\sum_{i=1}^{n} k\left(\frac{u - X_i}{b}\right) du \\ &= \frac{1}{n}\sum_{i=1}^{n}\int_{-\infty}^{\frac{x-X_i}{b}} k(t)\,dt = \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{x - X_i}{b}\right),\end{aligned} \tag{4}$$

where $k(\cdot)$ is a pdf, which is known as the kernel function. It is usually a symmetric pdf, but this does not imply that the final estimate of F_X is symmetric. Function $K(\cdot)$ is the cdf of $k(\cdot)$. Some examples of very common kernel functions are the Epanechnikov and the Gaussian kernel (see, Silverman, 1986). Parameter b is the *bandwidth* or the smoothing parameter. It controls for the smoothness of the cdf estimate. The larger b is, the smoother the resulting cdf. The classical kernel estimation of a cdf as defined in (4) is not much different to the expression of the well-known empirical distribution in (2). Indeed, in (4) one should replace $K\left(\frac{x-X_i}{b}\right)$ by $I(X_i \le x)$ in order to obtain (2). The main difference between (2) and (4) is that the empirical cdf only uses data below x to obtain the point estimate of $F_X(x)$, while the classical kernel cdf estimator uses all the data above and below x. In other words, the empirical cdf gives more weight to the observations that are smaller than x than it does to the observations that are larger than x.

In practice, to estimate VaR_α from $\widehat{F}_X(\cdot)$, we use the Newton–Raphson method to solve the equation:

$$\widehat{F}_X(x) = \alpha. \tag{5}$$

Properties of kernel cdf estimator were analyzed by Reiss (1981) and Azzalini (1981). Both point out that when $n \to \infty$, the mean squared error (MSE) of $\widehat{F}_X(x)$ can be approximated by:

$$\begin{aligned}E\left\{\widehat{F}_X(x) - F_X(x)\right\}^2 \sim\ & \frac{F_X(x)\left[1 - F_X(x)\right]}{n} \\ & - f_X(x)\frac{b}{n}\left(1 - \int_{-1}^{1} K^2(t)\,dt\right) \\ & + b^4\left(\frac{1}{2}f_X'(x)\int t^2 k(t)\,dt\right)^2 \\ =\ & \frac{F_X(x)\left[1 - F_X(x)\right]}{n} - u(x) + b^4 v(x),\end{aligned} \tag{6}$$

where as in Azzalini (1981)

$$u(x) = f_X(x)\frac{b}{n}\left(1 - \int_{-1}^{1} K^2(t)\,dt\right)$$

[3] Many recent contributions in insurance are based on nonparametric statistical methods. For instance, Lopez (2012) provided a new nonparametric estimator of the joint distribution of two lifetimes for mortality analysis and Kim (2010) studied the bias of the empirical distortion risk measure estimate.