ESTIMATING THE VARIANCE OF BOOTSTRAPPED RISK MEASURES

BY

JOSEPH H.T. KIM AND MARY R. HARDY

ABSTRACT

In Kim and Hardy (2007) the exact bootstrap was used to estimate certain risk measures including Value at Risk and the Conditional Tail Expectation. In this paper we continue this work by deriving the influence function of the exact-bootstrapped quantile risk measure. We can use the influence function to estimate the variance of the exact-bootstrap risk measure. We then extend the result to the L-estimator class, which includes the conditional tail expectation risk measure. The resulting formula provides an alternative way to estimate the variance of the bootstrapped risk measures, or the whole L-estimator class in an analytic form. A simulation study shows that this new method is comparable to the ordinary resampling-based bootstrap method, with the advantages of an analytic approach.

KEYWORDS

Exact bootstrap, L-estimator, influence function, nonparametric delta method, variance estimation, distortion risk measure.

1. INTRODUCTION

Risk measures have become an important tool in financial risk management for actuaries and other risk managers. The risk measure is often used for setting economic or regulatory capital standards for complex portfolios that are not amenable to analytic approaches. In such cases, the risk measure is generally estimated from a Monte Carlo simulation of the appropriate liability distribution.

In actuarial applications, generating scenarios can be very costly, computationally. We are often, therefore, estimating risk measures with relatively small samples. Furthermore, where the risk measure is applied to determine a solvency capital requirement, the risk measure may fall in the far tail of the loss distribution, adding to the uncertainty.

In an earlier paper (Kim and Hardy (2007)) we demonstrated the usefulness of the ordinary and (in particular) the exact bootstrap in reducing the bias

Astin Bulletin 39(1), 199-223. doi: 10.2143/AST.39.1.2038062

in certain risk measures, including the conditional tail expectation (CTE). The exact bootstrap offers an analytic form for a bootstrap estimator of a distribution statistic, thus eliminating the bootstrap resampling uncertainty, and in addition reducing the computational burden compared with the ordinary bootstrap.

The focus of this paper is the measurement of uncertainty associated with estimated risk measures. For some standard risk measures estimated using Monte Carlo samples, we have some tools available. Manistre and Hancock (2005) tackled this problem for the conditional tail expectation, using an influence function, or non-parametric delta approach. The ordinary bootstrap was used in Kim and Hardy (2007) to estimate the variance of the standard Monte Carlo estimates of the risk measures. However the underlying assumptions are changed where we have used the bootstrap methodology to determine the risk measure estimation.

Jones and Zitikis (2003) extended the non-parametric delta approach to the family of distortion risk measures, but little is known about the influence function of a bootstrapped distortion risk measure.

In this paper, we explore two non-parametric methods to estimate the variance of the *bootstrapped* distortion risk measure; in this sense this paper is a sequel of Kim and Hardy (2007). The first method is the bootstrap itself. We will examine the exact bootstrap method by Hutson and Ernst (2000) that provides an analytic bootstrap solution, thus eliminating the resampling error. The second method is the non-parametric delta method. In this paper we derive the influence function of the bootstrapped distortion risk measure in an analytic form and thus provide an alternative way of estimating its variance. The resulting formula requires only the analytic form of the risk measure and not the form of its influence function. This means that estimating the variance of the exact bootstrapped distortion risk measure is actually more straightforward than the original Monte Carlo or empirical estimate, as less information is required. Consequently the computation algorithm is generally simpler than the delta method for the empirical risk measure. The developments in this paper have other applications in statistics because the distortion risk measure is essentially the same as the L-estimator class; the variance of the bootstrapped L-estimators can also be computed with no additional difficulty.

This paper is organized as follows: In Section 2 we briefly review distortion risk measures. In Section 3 we present a brief review of the bootstrap and the delta methods as non-parametric variance estimation tools. While the bootstrap is straightforward in estimating the variance of any estimate, there has been no discussion on its relative efficiency compared to the nonparametric delta method counterpart in the actuarial context. In Section 4 we derive the influence function of the bootstrapped quantile and extend this to the L-estimator class. Also a qualitative discussion on statistical aspects of the result follows in this section. In Section 5 we illustrate the methods using a simulation study. Section 6 concludes the paper.

2. The distortion risk measure

Expressed as a functional mapping a random variable to a real value, the distortion risk measure (DRM) $t_g(F)$ is defined for a distribution function $F(x)$ by

$$\begin{aligned} t_g(F) &= -\int_{-\infty}^{0}\left[1 - g\left(\bar{F}(x)\right)\right]dx + \int_{0}^{+\infty} g\left(\bar{F}(x)\right)dx \\ &= \int_0^1 F^{-1}(u)\, g'(1-u)\, du, \end{aligned} \tag{1}$$

where g is an increasing function defined on [0,1] with $g(0) = 0$ and $g(1) = 1$, and $\bar{F}(x) = 1 - F(x)$. The DRM satisfies translation invariance, positive homogeneity, monotonicity, and additivity for comonotonic losses. If g is concave, the risk measure is *coherent*, in the sense of Artzner *et al.* (1999); see Wirch and Hardy (2000) and Dhaene *et al*. (2006).

Examples of the DRM include the Value-at-Risk (VaR) measure, the Conditional Tail Expectation (CTE), and the Proportional Hazards Transform (PHT) measure. All of these except the VaR are coherent because the corresponding g's are concave.

Turning to the actual estimation of $t_g(F)$ from a sample, an obvious choice is the empirical risk measure $t_g(\hat{F})$ where $\hat{F}$ is the empirical distribution function. That is

$$t_g(\hat{F}) = c_1 X_{(1)} + c_2 X_{(2)} + \ldots + c_n X_{(n)} = \mathbf{c}'\mathbf{X}_{:\mathbf{n}} \tag{2}$$

where $\mathbf{c} = (c_1, c_2, \ldots, c_n)'$, $\mathbf{X}_{:\mathbf{n}} = (X_{(1)}, X_{(2)}, \ldots, X_{(n)})'$, $X_{(1)} \leq \ldots \leq X_{(n)}$, and

$$c_i = \int_{(i-n)/n}^{i/n} g'(1-u)\, du = g\left(\frac{n-i+1}{n}\right) - g\left(\frac{n-i}{n}\right), \quad i = 1, 2, \ldots, n.$$

The variability of a given risk measure can be measured by a confidence interval (see, e.g., Kaiser and Brazauskas (2007)) or a variance estimate, provided that it exists. For estimating the variance of any general statistic of a distribution, $t(\hat{F})$ say, we can use non-parametric methods such as the bootstrap or the nonparametric delta method. For the latter method Jones and Zitikis (2003) identified that the DRM in (1) is equivalent to the L-estimator class whose standard expression is given by

$$\int_0^1 F^{-1}(u)\, J(u)\, du, \tag{3}$$

where $J(u)$ is commonly called the score function defined on [0,1] in the statistics literature. Thus, by setting $J(u) = g'(1-u)$, they have an access to known statistical results, such as the asymptotic variance of the empirical estimate of DRM.

Different choices, however, are possible to estimate $t_g(F)$. Most notably, Kim and Hardy (2007) investigated the bias of the Value-at-Risk (VaR) and the Conditional Tail Expectation (CTE), using the exact bootstrap by Hutson and Ernst (2000), and proposed a guideline on how to use the bootstrap without obtaining a compromising increase in variance. Their simulations show that the guideline often favors the exact bootstrapped risk measure over the empirical one for the CTE case, in terms of the mean squared error, but the variance for the *bootstrapped* risk measures is not considered in their paper. While it is true that bootstrap estimate and the empirical estimate converge, they can be substantially different for finite samples and there are many practical situations where increasing the sample size is constrained, as discussed in Kim and Hardy (2007).

3. Non-parametric variance estimation

This section provides a brief review of the two well-known non-parametric methods with the variance estimation application. We assume that F is continuous and the estimated variance makes sense in an asymptotic manner throughout this paper; see Jones and Zitikis (2003) for conditions of the variance existence of the L-estimator class.

3.1. The bootstrap method

The core idea of the nonparametric bootstrap is to repeatedly resample from the original sample with replacement. Suppose that we have a sample from an unknown distribution F and are interested in parameter $\theta = \theta(F)$ whose empirical estimate is $\hat{\theta} = \theta(\hat{F})$, where $\hat{F}$ is the empirical distribution function (e.d.f.). Now a series of the resamples – each resample is of the same size as the original sample – is drawn from $\hat{F}$ with replacement to produce $\hat{F}_i^*, \ldots, \hat{F}_R^*$, assuming R repetitions; we call $\hat{F}_i^*$ the i-th resample or bootstrap sample. The corresponding estimates $\hat{\theta}_1^*, \ldots, \hat{\theta}_R^*$, where $\hat{\theta}_i^* = \theta(\hat{F}_i^*)$, are then used for statistical inferences. Since the e.d.f. is treated as if it was the population distribution function any inference can be possible to make, though its accuracy may not be satisfactory. For example, the bootstrap mean and variance estimates of θ then are

$$\hat{\theta}^* = R^{-1} \sum_{i=1}^{R} \hat{\theta}_i^*$$

and

$$(R-1)^{-1} \sum_{i=1}^{R} (\hat{\theta}_i^* - \hat{\theta}^*)^2,$$

respectively. As a nonparametric inference tool the bootstrap is a widely-used in many areas; see, e.g., Efron and Tibshirani (1993), Shao and Tu (1995), Hall (1992), and Davidson and Hinkley (1997) for a comprehensive treatment for this subject. Although the bootstrap estimate is subject to the resampling simulation error due to a finite R, sometimes it is possible to analytically evaluate at $R = \infty$ with no simulation, in which case the bootstrap is called exact in the sense that the simulation error is eliminated. For the L-estimator class Hutson and Ernst (2000) derived the exact bootstrap (EB) mean and variance estimate.

Theorem 3.1 (Hutson and Ernst (2000)) *The EB estimate of* $E(X_{(r)}|F)$, $1 \leq r \leq n$ *is*

$$E(X_{(r)}|\hat{F}) = \sum_{j=1}^{n} w_{j(r)} X_{(j)},$$

where

$$w_{j(r)} = r\binom{n}{r}\left[B\left(\frac{j}{n}; r, n-r+1\right) - B\left(\frac{j-1}{n}; r, n-r+1\right)\right],$$

and

$$B(x; a, b) = \int_0^x t^{a-1}(1-t)^{b-1}\,dt.$$

This result immediately extends to the L-estimator class. Following the notation in (2), the EB estimate of the DRM is

$$E(t(\hat{F})|\hat{F}) = E(\mathbf{c}'\mathrm{X}_{:n}|\hat{F}) = \mathbf{c}'\mathbf{w}'\mathrm{X}_{:n}, \tag{4}$$

where the matrix $\mathbf{w} = \{w_{i(j)}\}_{i,j=1}^{n}$ comes from the EB weights for each element of $\mathbf{X_{:n}}$. Hutson and Ernst (2000) also gives the analytic formula for EB variance estimate or the EB covariance matrix

$$\Sigma_{\mathbf{:n}} = \left\{Cov(X_{(r)}, X_{(s)}\,|\,\hat{F})\right\}_{r,s=1}^{n}, \text{ as follows} \tag{5}$$

Theorem 3.2 (Hutson and Ernst (2000)) *The EB covariance matrix* $\Sigma_{\mathbf{:n}}$ *is obtained by*

$$Var(X_{(r)}\,|\,\hat{F}) = \sum_{j=1}^{n} w_{j(r)}(X_{(j)} - \hat{\mu}_{(r)})^2$$

and, for $1 \leq r \leq n$ *and* $r < s$,

$$\begin{aligned} Cov(X_{(r)}, X_{(s)}\,|\,\hat{F}) = & \sum_{j=2}^{n}\sum_{j=1}^{j-1} v_{ij(rs)}(X_{(i)} - \hat{\mu}_{(r)})\,(X_{(j)} - \hat{\mu}_{(s)}) \\ & + \sum_{j=1}^{n} v_{jj(rs)}(X_{(j)} - \hat{\mu}_{(r)})\,(X_{(j)} - \hat{\mu}_{(s)}) \end{aligned}$$

where $\hat{\mu}_{(r)} = E(X_{(r)} | \hat{F})$, *the weights are*

$$v_{ij(rs)} = \int_{(j-n)/n}^{j/n} \int_{(i-n)/n}^{i/n} f_{rs:n}(u_r, u_s)\, du_r\, du_s,$$

$$v_{jj(rs)} = \int_{(j-n)/n}^{j/n} \int_{(j-n)/n}^{u_s} f_{rs:n}(u_r, u_s)\, du_r\, du_s,$$

and $f_{rs:n}(u_r, u_s) = {}_nC_{rs} u_r^{r-1}(u_s - u_r)^{s-r-1}(1-u_s)^{n-s}$ *is the joint distribution of two uniform order statistics* $U_{r:n}$ *and* $U_{s:n}$ *with* ${}_nC_{rs} = n!/(r-1)!(s-r-1)!(n-s)!$.

Thus, the EB variance of the empirical DRM $\mathbf{c}'\mathrm{X}_{:\mathbf{n}}$ is given by

$$Var(\mathbf{c}'\mathrm{X}_{:\mathbf{n}} | \hat{F}) = \mathbf{c}'\boldsymbol{\Sigma}_{:\mathbf{n}}\mathbf{c}. \tag{6}$$

Now we turn to the bootstrapped (EB) risk measure, which is our quantity of interest. As Kim and Hardy (2007) reported, sometimes the EB DRM $\mathbf{c}'\mathbf{w}'\mathrm{X}_{:\mathbf{n}}$, or the bias-corrected EB DRM $\mathbf{c}'(\mathbf{2I}-\mathbf{w}')\mathrm{X}_{:\mathbf{n}}$, are preferred to the empirical risk measures in terms of the mean squared error. The EB variance estimates of these bootstrapped quantities therefore are

$$Var(\mathbf{c}'\mathbf{w}'\mathrm{X}_{:\mathbf{n}} | \hat{F}) = \mathbf{c}'\mathbf{w}'\boldsymbol{\Sigma}_{:\mathbf{n}}\mathbf{w}\mathbf{c}, \tag{7}$$

and

$$Var(\mathbf{c}'(\mathbf{2I}-\mathbf{w}')\mathrm{X}_{:\mathbf{n}} | \hat{F}) = \mathbf{c}'(\mathbf{2I}-\mathbf{w}')\boldsymbol{\Sigma}_{:\mathbf{n}}(\mathbf{2I}-\mathbf{w})\mathbf{c},$$

respectively.

While this exercise is straightforward and useful, a closer look prompts a computational issue on $\boldsymbol{\Sigma}_{:\mathbf{n}}$ which Hutson and Ernst (2000) did not discuss. As seen in Theorem 3.2, $\boldsymbol{\Sigma}_{:\mathbf{n}}$ is a $n \times n$ matrix but computing each element of this matrix involves another $n \times n$ matrix. The total number of computations is of order $O(n^4)$; the computational burden increases exponentially as the sample size gets larger[1]. Furthermore, because the bootstrap weights are functions of the sample size n, one should recalculate $\boldsymbol{\Sigma}_{:\mathbf{n}}$ whenever the sample size changes. For these reasons we recommend replacing $\boldsymbol{\Sigma}_{:\mathbf{n}}$ with the ordinary bootstrap (OB) version $\hat{\boldsymbol{\Sigma}}_{:\mathbf{n}}$ based on R bootstrap samples, $\hat{F}_1^*, \ldots, \hat{F}_R^*$. It is known that R needs to be bigger for the second moment than for the mean to avoid bias, as discussed in Section 5; see, e.g., Booth and Sarkar (1998).

[1] The number of computations is about 1.28×10^{10} for $n = 400$.

3.2. Nonparametric delta method

An alternative way to estimate the variance of an estimator is using the nonparametric delta method (or just delta method in short) which employs the influence function of the estimator. Estimating variance through the nonparametric delta method is well known and can be found in standard texts such as Staudte and Sheather (1990) or Hampel *et al*. (1986). Consider the von Mises expansion of any statistical functional $t(G)$ at F. The first order approximation is

$$t(G) \approx t(F) + \int L_t(x|F)\, dG(x).$$

Here L_t, the first derivative of t at F, is called the influence function (IF). The IF is a function of x given F and t, and defined by

$$L_t(x|F) = \lim_{\varepsilon \to 0} \frac{t\left[(1-\varepsilon)F + \varepsilon H_x\right] - t(F)}{\varepsilon}.$$

where H_x is the c.d.f. of a degenerate random variable at x, commonly referred to the heaviside function. The IF measures the relative influence on $t(F)$ of a very small amount of contamination at x and also can be used to estimate the variance of t. For our purpose of estimating the variance of the risk measure from the sample, we set $G = \hat{F}$, which is a choice that makes the approximation reasonably accurate as n increases, so the approximation becomes

$$t(\hat{F}) \approx t(F) + \int L_t(x|F)\, d\hat{F}_x = t(F) + \frac{1}{n}\sum_{i=1}^{n} L_t(x_i|F)$$

Now by applying the central limit theorem, $T = t(\hat{F})$ has asymptotic normality:

$$t(\hat{F}) - t(F) \xrightarrow{d} N\big(0, v_L(F)\big), \tag{8}$$

as $n \longrightarrow \infty$, where $v_L(F) = n^{-1} Var(L_t(X|F)) = n^{-1}\int L_t^2(x|F)\, dF(x)$. Assuming no information on F in practice, we estimate this variance using the sample version:

$$v_L(\hat{F}) \equiv \frac{1}{n^2}\sum_{j=1}^{n} L_t^2\left(x_j \middle| \hat{F}\right), \tag{9}$$

where x_j is the j-th observation of the sample.

If we restrict our attention to the L-estimator, defined by $\int_0^1 F^{-1}(u) J(u)\, du$, of which the score function J is bounded and continuous, the IF and the asymptotic variance is known by (see, e.g., Appendix B in Staudte and Sheather (1990)

$$L_t(x|F) = \int_{-\infty}^{x} J(F(y))\,dy - \int_{-\infty}^{\infty} (1 - F(y))\,J(F(y))\,dy \tag{10}$$

and

$$v_L(F) = n^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J(F(y))\,J(F(z))\,[F(\min(y,z)) - F(y)F(z)]\,dydz, \tag{11}$$

with its sample estimate

$$v_L(\hat{F}) = n^{-1}\sum_{j=1}^{n-1}\sum_{i=1}^{n-1} J\left(\frac{j}{n}\right)J\left(\frac{i}{n}\right)\left[\min\left(\frac{i}{n},\frac{j}{n}\right) - \frac{i}{n}\frac{j}{n}\right]\left(X_{(i+1)} - X_{(i)}\right)\left(X_{(j+1)} - X_{(j)}\right) \tag{12}$$

This is the form used in Jones and Zitikis (2003) for the DRM variance estimation. For most DRMs such as the CTE and PHT-measure, the variance estimate of the empirical estimates can be computed by this formula; see Gourieroux and Liu (2006) for variance estimates of different risk measures. However, this result does not apply to the VaR, the quantile risk measure, because its score function is a discontinuous step function. We will return to this point in the next section.

We also note that the delta method introduced in this section is meant to be used for the variance estimation of the empirical DRM estimator, not of the bootstrapped or bias-corrected counterpart, because the latter ones are not empirical estimates of continuous functionals even though all these estimators may be asymptotically equivalent.

Some examples follow to show the IF of different DRMs.

Example 3.3. (Conditional Tail Expectation) *The CTE at a confidence level of* α *is defined by* $t(F) = (1-\alpha)^{-1}\int_{Q_\alpha}^{\infty} xdF$ *and its corresponding score function is* $J(u) = (1-\alpha)^{-1} I_{(\alpha,1]}(u)$*, where* $I_A(u)$ *is the characteristic function that equals one if* $u \in A$*; zero otherwise. The IF and the asymptotic variance are given by*

$$L_t(x|F) = \begin{cases} \dfrac{x - \alpha Q_\alpha}{1-\alpha} - t(F) & \text{if } Q_\alpha < x \\ Q_\alpha - t(F) & \text{if } Q_\alpha \geq x \end{cases}$$

and

$$v_L(F) = \frac{Var(X|X > Q_\alpha) + \alpha\left(E[X|X > Q_\alpha] - Q_\alpha\right)^2}{n(1-\alpha)}.$$

For derivation of the IF of the CTE, see Appendix.

Example 3.4 (Proportional Hazard Transform) *The PHT measure with a parameter* β, $0 < \beta \leq 1$, *is defined by* $t(F) = \int_0^1 (1 - F(x))^{\beta} dx$. *Its score function is then* $J(u) = \beta(1 - u)^{\beta - 1}$ *and the IF is given by*

$$L_t(x|F) = \beta\left[\int_{-\infty}^{x} (1 - F(y))^{\beta - 1} dy - \int_{-\infty}^{\infty} (1 - F(y))^{\beta} dy\right].$$

The asymptotic variance is given by formula (11) with the specified score function and no further simplification seems possible in this case.

Example 3.5 (Wang Tansform) *The WT measure with a parameter* λ *is defined by* $t(F) = \int_0^1 \Phi(\Phi^{-1}(1 - F(x)) + \lambda) dx$, *where* Φ *is the standard normal distribution function. Its score function is then* $J(u) = \exp(-\lambda\Phi^{-1}(1 - u) - \lambda^2/2)$ *and the IF is given by*

$$L_t(x|F) = e^{-\lambda^2/2}\left[\int_{-\infty}^{x} e^{-\lambda\Phi^{-1}(1 - F(y))} dy - \int_{-\infty}^{\infty} (1 - F(y)) e^{-\lambda\Phi^{-1}(1 - F(y))} dy\right].$$

Again the asymptotic variance is given by formula (11).

3.3. Understanding the behavior of IF

In addition to its usefulness for asymptotic variance estimation, the IF of a given functional $t(F)$ also helps to describe the behavior of the given functional. In Figure 1 typical graphs of three risk measures' IFs are illustrated under the standard normal distribution, along with the quantile risk measure (also known as Value at Risk or VaR) that is discussed in the next section more closely. Depending on the parameter value and the underlying distribution F, the value of each IF may change but its shape remains similar. Heuristically speaking, the IF in each graph shows how an additional loss (or contamination) at x affects the value of the corresponding functional.

In Figure 1 IFs of PHT and WT measures show an increasing influence as the new observation gets larger. The influence will be negative for a smaller x, meaning that these risk measures will decrease in this range; the risk measures will increase for a larger x. This aspect is similar to the ordinary mean functional, whose IF is $L_t(x|F) = x - E(X)$, a linear function, but these measures put increasingly more weights to large losses to produce conservative numbers in, say, capital amount setting. The graph thus intuitively shows the impact of the *distortion* on different DRMs. Since IFs are not bounded, a single extreme value of x can make the risk measures arbitrarily large, like the ordinary mean functional, but at a greater pace.

Let us now turn to the two tail risk measures: VaR and CTE. The IF of VaR is a bounded step function and therefore the impact of one observation, however big or small, is limited; the impact actually is constant and rather

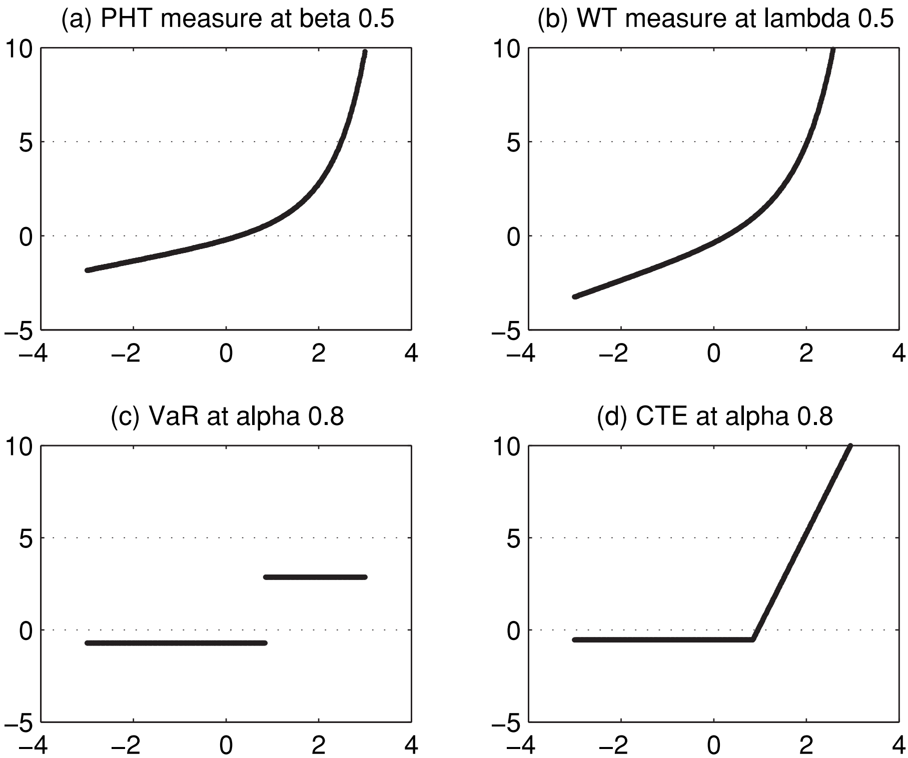

FIGURE 1: IF of different risk measures under the standard normal.

abrupt depending on whether the observation falls right or left side of the quantile. Thus change in x causes no change in the VaR and as long as x stays on one of either side of the quantile threshold. The graph intuitively proves the well-known argument that the VaR cannot account for the magnitude of extreme loss and consequently which is a significant disadvantage for an insurance risk measure.

On the other hand, the CTE does reflect the magnitude of loss once the loss lies beyond the quantile threshold as shown by an increasing linear function; if the newly-added loss is below the quantile the impact will be negative but constant and small. The unbounded IF of the CTE again indicates that the CTE value can get indefinitely large as a single jumbo loss x increases. The slope of the IF of the CTE above the threshold is also linear in x, like that of the mean, but is steeper because the coefficient $(1-\alpha)^{-1}$ is larger than 1 for $0 < \alpha < 1$, indicating the impact of a extreme loss will be bigger on the CTE than on the mean.

We also comment that neither of these tail measures account for losses below the quantile, whereas the PHT and WT measures are affected by losses from all parts of the distribution, making the latter two measures potentially advantageous in certain applications. This perhaps motivates the usage of certain risk measures in different purposes; the VaR and CTE are commonly used in economic capital management in practice, as they are tail-oriented, whereas the PHT and WT are more associated with the centre of the distribution and are more commonly used for pricing.

4. Variance of exact bootstrapped (EB) risk measures

In this section we show that the IF of the bootstrapped L-estimator (or DRM) is available in an analytic form. In particular, the influence function of the bootstrapped quantile is first derived and is later extended to the whole L-estimator class, or DRM.

4.1. EB for the quantile risk measure

Let us start with the quantile case. The IF formula in (10) is not applicable here because the score function is not smooth; it is a single mass at the quantile itself. The IF of the standard quantile, defined by $t(F) = F^{-1}(\alpha)$, is given by

$$L_t(x|F) = \begin{cases} \dfrac{\alpha - 1}{f\left(F^{-1}(\alpha)\right)}, & \textit{if } x < F^{-1}(\alpha) \\ \dfrac{\alpha}{f\left(F^{-1}(\alpha)\right)}, & \textit{if } x > F^{-1}(\alpha) \end{cases} \tag{13}$$

with the asymptotic variance

$$v_L(F) = n^{-1}\frac{\alpha(1-\alpha)}{f\left(F^{-1}(\alpha)\right)^2}.$$

Thus one needs to estimate $f(F^{-1}(\alpha))$, the value of the density at the quantile, to estimate the variance; in fact, if the score function consists of discrete masses the corresponding IF and variance involve the unknown density f. This task however might be impractical because estimating the value of density at a large α from the given sample is not easy. Although there are nonparametric tools available allowing us to estimate the density, such as kernel density estimation, this may not produce satisfactory values for tail regions, especially when the given sample is subject to skewness or (and) excess kurtosis, which is often true for financial and actuarial data. In practice therefore, using the nonparametric delta method to estimate the variance of VaR, a tail quantile risk measure, can be problematic. This is a long-standing problem in many statistical applications.

Here we consider the IF of the EB quantile instead. A closer look at the EB formula given in Theorem 3.1 sheds a new light. We first note that the

EB estimate of the r-th order statistic is given by a linear combination of all the order statistics, again belonging to the L-estimator class, and the weight coefficients $w_{i(r)}$, $i = 1, 2, \ldots, n$, form a Beta density function with parameter pair $(r, n-r+1)$ if appended together.

This observation suggests that the EB estimate

$$E(X_{(r)} | \hat{F}) = \sum_{j=1}^{n} w_{j(r)} X_{(j)} = \int_0^1 \hat{F}_n^{-1}(u) J_q(u)\, du,$$

where

$$J_q(u) = \frac{u^{r-1}(1-u)^{n-r}}{B(r, n-r+1)}, \tag{14}$$

can be thought as the plug-in estimate of a *functional*

$$\int_0^1 F^{-1}(u) J_q(u)\, du. \tag{15}$$

Note that the score function $J_q(u)$ is a density of Beta distribution which is continuous, bounded, and differentiable for a given n. We will call functional (15) the EB α-th quantile ($\alpha = r/n$). Using the beta distribution's first two moments, it is easy to see that $J_q(u)$ converges to a point mass at α in distribution as $n \to \infty$, recovering the original quantile. The following result leads us to a simple method to estimate the variance of the EB quantile for any fixed sample size n.

Lemma 4.1. *The IF of the EB α–th quantile defined in (15) is given by*

$$\frac{1}{B(n\alpha, n(1-\alpha)+1)} \times \left[\int_{-\infty}^{x} F(y)^{n\alpha-1}(1-F(y))^{n(1-\alpha)} dy - \int_{-\infty}^{\infty} F(y)^{n\alpha-1}(1-F(y))^{n(1-\alpha)+1} dy \right] \tag{16}$$

Proof: See Appendix. □

The IF of the EB quantile looks quite different from that of the standard quantile in (13). It is a function of the sample size n and, more importantly, does not involve the density f, leading to straightforward computation of variance estimate with a sample.

Consequently the variance estimate of the EB α-th quantile is then, from (12), given by

$$\frac{1}{n} \sum_{j=1}^{n-1} \sum_{i=1}^{n-1} J_q\left(\frac{j}{n}\right) J_q\left(\frac{i}{n}\right) \left[\min\left(\frac{i}{n}, \frac{j}{n}\right) - \frac{i}{n}\frac{j}{n} \right] \left(X_{(i+1)} - X_{(i)}\right)\left(X_{(j+1)} - X_{(j)}\right), \tag{17}$$

where

$$J_q\left(\frac{i}{n}\right) = \frac{\left(\frac{i}{n}\right)^{r-1}\left(1-\frac{i}{n}\right)^{n-r}}{B(r, n-r+1)}. \tag{18}$$

For actual implementation however $J_q(\frac{i}{n})$ needs to be approximated because it often produces 0 for large n and r in mathematical software. Among others we choose

$$J_q\left(\frac{i}{n}\right) = \frac{\left(\frac{i}{n}\right)^{r-1}\left(1-\frac{i}{n}\right)^{n-r}}{B(r, n-r+1)} \approx \frac{n\int_{\frac{i-1}{n}}^{\frac{i}{n}} u^{r-1}(1-u)^{n-r}\,du}{B(r, n-r+1)} = nw_{i(r)}.$$

This choice is advantageous because $w_{j(r)}$ is the (j,r)th element of the EB weight matrix $\mathbf{w}$, which is already in our hands. Using this approximation the variance estimate (17) can be presented in a most computationally-convenient form:

$$\sum_{j=1}^{n-1}\sum_{i=1}^{n-1} w_{j(r)}\,w_{i(r)}\left[\min(i,j) - \frac{ij}{n}\right]\left(X_{(i+1)} - X_{(i)}\right)\left(X_{(j+1)} - X_{(j)}\right). \tag{19}$$

4.2. Extension to EB L-estimator

To extend our result to the L-estimator class, let us reconsider the empirical DRM defined in (2). That is

$$t_g(\hat{F}) = c_1X_{(1)} + c_2X_{(2)} + \ldots + c_nX_{(n)},$$

where $c_i = g\left(\frac{n-i+1}{n}\right) - g\left(\frac{n-i}{n}\right), \quad i = 1, 2, \ldots, n.$

Following a similar argument to the quantile case, the EB estimate is given by

$$\begin{aligned} E\left(\sum_{k=1}^{n} c_k X_{(k)} \middle| \hat{F}\right) &= \sum_{k=1}^{n} c_k\, E\left(X_{(k)} \middle| \hat{F}\right) \\ &= \sum_{k=1}^{n} c_k \sum_{j=1}^{n} w_{j(k)}\, X_{(j)} \\ &= \sum_{k=1}^{n} c_k \int_0^1 \hat{F}_n^{-1}(u)\, J_q(u)\, du \\ &= \int_0^1 \hat{F}_n^{-1}(u) \sum_{k=1}^{n} c_k\, J_q(u)\, du, \end{aligned}$$

which can be thought as the plug-in estimate of a functional

$$\int_0^1 F^{-1}(u) J_L(u)\, du,$$

where

$$J_L(u) = \sum_{k=1}^{n} c_k J_q(u) = \sum_{k=1}^{n} \frac{c_k u^{k-1}(1-u)^{n-k}}{B(k, n-k+1)}. \tag{20}$$

The following is an extended result of Lemma 4.1 and covers the L-estimator class.

Corollary 4.2. *Consider an L-estimator* $\mathbf{c}'\mathbf{X}_{:\mathbf{n}} = c_1 X_{(1)} + c_2 X_{(2)} + \ldots + c_n X_{(n)}$. *Then the IF of the bootstrapped L-estimator,* $\mathbf{c}'\mathbf{w}'\mathbf{X}_{:\mathbf{n}}$, *is given by*

$$\sum_{k=1}^{n} \frac{c_k}{B(k, n-k+1)} \times$$
$$\left[\int_{-\infty}^{x} F(y)^{k-1} (1 - F(y))^{n-k}\, dy - \int_{-\infty}^{\infty} F(y)^{k-1} (1 - F(y))^{n-k+1}\, dy \right]$$

Proof: See Appendix. □

The asymptotic variance is given by (11) with $J(u)$ replaced by $J_L(u)$, with its sample version given by

$$\frac{1}{n} \sum_{j=1}^{n-1} \sum_{i=1}^{n-1} J_L\left(\frac{j}{n}\right) J_L\left(\frac{i}{n}\right) \left[\min\left(\frac{i}{n}, \frac{j}{n}\right) - \frac{i}{n}\frac{j}{n} \right] \left(X_{(i+1)} - X_{(i)}\right)\left(X_{(j+1)} - X_{(j)}\right),$$

Again we need an approximation for J_L:

$$J_L\left(\frac{i}{n}\right) = \sum_{k=1}^{n} \frac{c_k \left(\frac{i}{n}\right)^{k-1} \left(1 - \frac{i}{n}\right)^{n-k}}{B(k, n-k+1)} \approx n \sum_{k=1}^{n} \frac{c_k \int_{\frac{i-1}{n}}^{\frac{i}{n}} u^{k-1}(1-u)^{n-k}\, du}{B(k, n-k+1)} \tag{21}$$

$$= n \sum_{k=1}^{n} c_k w_{i(k)} \tag{22}$$

$$= n \mathbf{w}^{\mathbf{i}} \mathbf{c}, \quad i = 1, 2, \ldots, n-1, \tag{23}$$

where $\mathbf{w}^{\mathbf{i}}$ is the i-th row vector of the EB weight matrix $\mathbf{w} = \{w_{i(j)}\}_1^n$ and $\mathbf{c}$ is the weight vector. Thus, a most computationally-convenient form of the variance estimate of a bootstrapped DRM is

$$\sum_{j=1}^{n-1}\sum_{i=1}^{n-1}\left(\mathbf{w}^{\mathbf{i}}\mathbf{c}\right)\left(\mathbf{w}^{\mathbf{j}}\mathbf{c}\right)\left[\min(i,j)-\frac{ij}{n}\right]\left(X_{(i+1)}-X_{(i)}\right)\left(X_{(j+1)}-X_{(j)}\right). \quad (24)$$

If the bias-corrected estimator is required, one needs to change the score function to

$$J_{L_{BC}}(u) = 2\sum_{k=1}^{n} c_k - \sum_{k=1}^{n}\frac{c_k u^{k-1}(1-u)^{n-k}}{B(k,n-k+1)},$$

and repeat the argument in the same fashion.

As examples, we examine the same risk measures as in the previous section to illustrate the impact of the bootstrapping on the IFs for the selected risk measures.

Example 4.3. (CTE) *Write the weight vector for the empirical CTE estimate as* $\mathbf{c}=(c_1,\ldots,c_n)'=(n(1-\alpha))^{-1}(0,\ldots,0,1,\ldots,1)'$, *where* $c_1,\ldots,c_{n\alpha}$ *are all equal to* 0, *and* $c_{n\alpha+1},\ldots,c_n$ *are all equal to* $(n(1-\alpha))^{-1}$, *where we assume* $n\alpha$ *is an integer. The IF of the bootstrapped (EB) CTE is given by*

$$\sum_{k=n\alpha+1}^{n}\frac{(n(1-\alpha))^{-1}}{B(k,n-r+1)}\times$$
$$\left[\int_{-\infty}^{x}F(y)^{k-1}(1-F(y))^{n-k}\,dy-\int_{-\infty}^{\infty}F(y)^{k-1}(1-F(y))^{n-k+1}\,dy\right]$$

The variance estimate of the EB CTE can be computed from (24).

Example 4.4. (PHT and WT) *The weight vector for the empirical risk measure is* $\mathbf{c}=(c_1,\ldots,c_n)'$, *where* $c_i=\left(\frac{n-i+1}{n}\right)^{\beta}-\left(\frac{n-i}{n}\right)^{\beta}$ *for the PHT measure and* $c_i=\Phi\left(\Phi^{-1}\left(\frac{n-i+1}{n}\right)+\lambda\right)-\Phi\left(\Phi^{-1}\left(\frac{n-i}{n}\right)+\lambda\right)$ *for the WT measure. The IF of the corresponding EB risk measures and their variance estimate can again be computed from the above Corollary and formula (24), respectively.*

As the reader might have noticed, the variance estimation using the delta method generally requires less effort for the EB estimates than for the original estimate for a given risk measure since the IF of a bootstrapped risk measure merely requires the knowledge of the distortion function $g(u)$, whereas for the empirical risk measure we need its derivative $J(u)=g'(1-u)$ for the IF.

To examine the general impact of bootstrapping on the DRM estimation, let us look at the IF graphs of the bootstrapped risk measure. In Figure 2 we compare two IFs of the same four risk measures as in the previous section. For both the PHT and WT measure we see that the variation of the bootstrapped IF has reduced from the original IFs through flattened pattern in graph. This

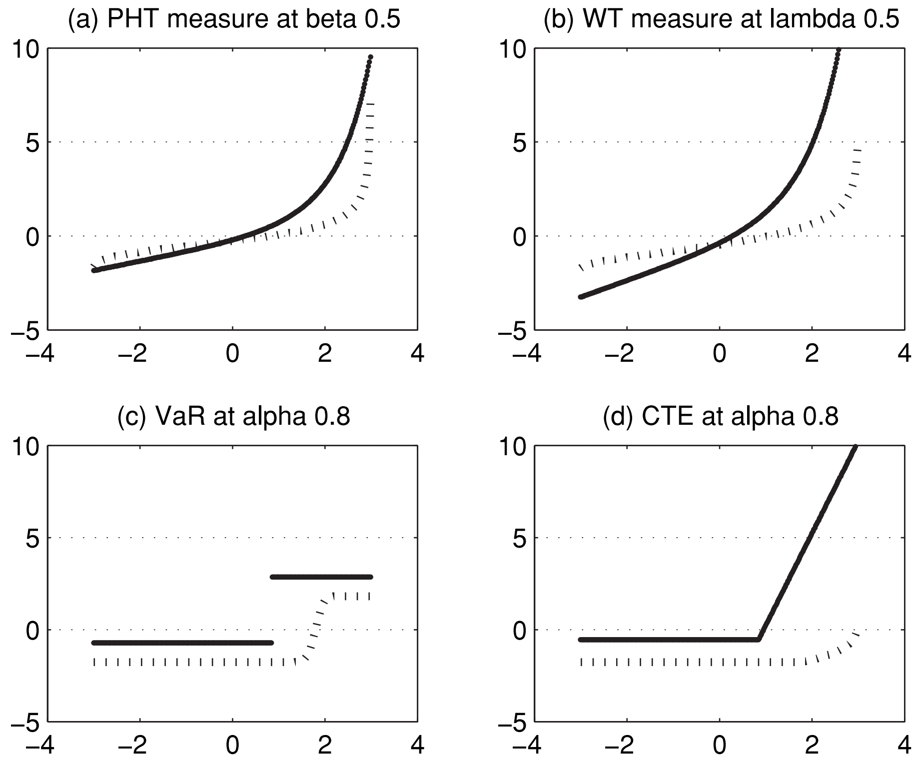

FIGURE 2: Comparison of two IFs for different risk measures under the standard normal for sample size 300: original (solid) vs. bootstrapped (dotted).

shows that the bootstrapping would make these two centre-oriented risk measures more stable. For the VaR case bootstrapping makes the quantile IF continuous and differentiable by taking average of all order statistics in the sample. For the bootstrapped CTE case the sharp edge at the quantile has now been smoothed out, but at a larger value of x, making the IF differentiable everywhere.

Focusing on positive range of x, the bootstrapped IFs are positioned below the original IFs for all four risk measures. This can again be interpreted as a reduced impact of extreme losses on risk measures. This seems consistent with the mechanism of the bootstrap for the L-estimator because the resampling procedure necessarily dilutes the magnitude of extreme losses by including smaller-valued losses in a form of weighted average of all order statistics. Our last comment is that the most statistics literature focuses on finding a robust functional which is often linked to bounded and continuous IF (see, e.g., Staudte and Sheather (1990)) in an attempt to find, say, the location parameter, but many actuarial risk measures, except for the VaR, do not fall in this category because they are designed to effectively capture the unlikely loss events that are well beyond the centre of the distribution.

4.3. Further remarks

We have derived the IF of the bootstrapped quantile and extended it to the L-estimator (or equivalently the bootstrapped DRM). There are several comments on this analytic development.

- The key idea of the nonparametric delta method in (9) states that $L_t(x|F)$ can be estimated by $L_t(x|\hat{F})$, based on the convergence of $\hat{F}$ to F. Since $\hat{F}^{EB}$ also converges to the true F, estimating $L_t(x|F)$ using $L_t(x|\hat{F}^{EB})$ can be likewise justified. This gives an alternative way to estimate the variance of a statistic where applying the nonparametric delta method is difficult for some reason. In the same line, as long as the score function J is bounded and continuous the variance estimate of the EB DRM in (24) and that of the empirical DRM in (12) are asymptotically equivalent.
- It is known that both the IF approach and the bootstrap approach produce similar variance estimates for an arbitrary statistic; see, for example, Section 2.7.4 of Davison and Hinkley (1997). This means that as sample size increases one would expect the two numbers from both methods to get close. This point will be illustrated in the numerical examples of the following section.
- The resulting formula can also be used in the practical guideline proposed in Kim and Hardy (2007) where the authors used the ordinary bootstrap to compute the variance of bootstrapped CTE.

5. Numerical example

For the simulation study we investigate bootstrapped risk measures for small sample sizes, say $n \leq 1000$, using different nonparametric methods. The sample size of less than 1000 is common in actuarial loss modelling and in operational or credit risk modelling, due to the expense of generating additional scenarios.

For the simulation study we consider the same three parametric models used in Kim and Hardy (2007). The first model is a 10-year European naked put option with the price return based on the Lognormal (LN) distribution. The initial price of the asset is set at \$100 and strike price is \$180, and the risk free rate is assumed 0.5% monthly effective. The LN parameters of the P-measure are $\mu = 0.00947$ and $\sigma = 0.04167$ which are derived from the monthly S&P 500. The second model is the identical put option as the first one except that the price return follows the Regime Switching Log-Normal distribution with two regimes (RSLN2). The parameters are derived from the same S&P data: $\mu_1 = 0.0127$, $\mu_2 = -0.0162$, $\sigma_1 = 0.0351$, $\sigma_2 = 0.0691$, $p_{12} = 0.0468$, and $p_{21} = 0.3232$. The final model is a Pareto distribution with parameters $\beta = 10$ and $\xi = 0.2$ whose distribution function is

$$F(x) = 1 - \left(\frac{\beta}{\beta + \xi x}\right)^{1/\xi}, \ x > 0. \tag{25}$$

Note that the Pareto has fatter tail than the other two models. See Kim and Hardy (2007) for the further description of each model.

For each model we consider four bootstrapped (EB) risk measures discussed in this paper, but in two separate groups. The first group includes the CTE and the VaR; in particular, we compare the EB estimate of the 95% CTE and the 97.5% VaR, which would produce the same number for the uniform distribution defined in the unit interval. The second group consists of PHT measure with $\beta = 0.8$ and the WT measure with $\lambda = 0.1976$, which again produce the same number under the uniform distribution. This grouping is due to the fact that the values produced by the former two tail risk measures will generally be too large to be matched by the latter two centre-oriented risk measures. For example, under a uniform distribution, a VaR at 97.5% can be matched by a PHT measure with $\beta = 0.02564$, but this choice of β is too small and fails to be asymptotically normal, preventing the use of the delta method; see, e.g., Jones and Zitikis (2003).

The empirical estimates of these DRM are given by

$$\widehat{CTE}_\alpha = \frac{1}{n(1-\alpha)} \sum_{j=n\alpha+1}^{n} X_{(j)},$$

$$\widehat{VaR}_\alpha = X_{(n\alpha)},$$

$$\widehat{PHT}_\lambda = \sum_{j=1}^{n} c_j X_{(j)}, \text{ with } c_i = \left(\frac{n-i+1}{n}\right)^\beta - \left(\frac{n-i}{n}\right)^\beta,$$

and

$$\widehat{WT}_\beta = \sum_{j=1}^{n} c_j X_{(j)}, \text{ with } c_i = \Phi\left(\Phi^{-1}\left(\frac{n-i+1}{n}\right) + \lambda\right) - \Phi\left(\Phi^{-1}\left(\frac{n-i}{n}\right) + \lambda\right).$$

Their EB counterparts used in our simulation study are computed by (4).

Since our focus is on the nonparametric variance estimation of the bootstrapped DRM, the influence function (IF) based nonparametric delta method developed in the previous section is compared with the variance estimate using the ordinary bootstrap (OB) simulation for each sample to obtain the estimated covariance matrix $\hat{\boldsymbol{\Sigma}}_{:n}$. We report the results for sample sizes 400 and 1,000, respectively, by repeating 5,000 sets of simulations. For the OB simulation, we used $R = 1{,}000$ for each sample of size 400, and $R = 2{,}500$ for a sample of size 1,000; the resampling size has been set at 2.5 times the sample size ($R = 2.5 \times n$) to avoid the bias in the variance estimate[2].

[2] A separate simulation shows that for $n = 1{,}000$ the bias remains unacceptably big even with $R = 1{,}000$, confirming the finding of Booth and Sarkar (1998).

Table 1 and 2 show the simulation results on the bias, standard deviation (std), and root mean squared error (rmse), along with standard errors (s.e.). All numbers are expressed in percentage of the true variance (true val.) of each risk measure under three different models.

TABLE 1

VARIANCE ESTIMATION OF PHT AND WT BOOTSTRAPPED RISK MEASURES FOR DIFFERENT SAMPLE SIZES (n = 400 AND 1000) AND MODELS (LN PUT, RSLN PUT, AND PARETO): COMPARISON OF THE ORDINARY BOOTSTRAP (OB) AND THE DELTA METHOD (IF)

Risk measure: PHT with β = 0.8						
Model	***n***	**Method**	**True val.**	**Bias in % (s.e.)**	**Std in % (s.e.)**	**rmse in %**
LN	400	OB	0.33	–3.23 (0.29)	20.3 (0.03)	20.55
		IF	0.33	–5.69 (0.27)	19.42 (0.03)	20.24
	1000	OB	0.13	–1.33 (0.18)	12.98 (0)	13.05
		IF	0.13	–2.58 (0.18)	12.59 (0)	12.85
RSLN	400	OB	0.57	–1.15 (0.24)	16.8 (0.03)	16.84
		IF	0.57	–2.99 (0.22)	15.9 (0.03)	16.17
	1000	OB	0.23	–1.18 (0.15)	10.36 (0)	10.42
		IF	0.23	–2.14 (0.14)	9.94 (0)	10.17
Pareto	400	OB	1.59	–4.04 (1.53)	108.28 (3.73)	108.35
		IF	1.59	–7.53 (1.43)	100.92 (3.24)	101.2
	1000	OB	0.68	–5.58 (0.99)	69.96 (0.66)	70.18
		IF	0.68	–8.37 (0.93)	65.6 (0.58)	66.13
Risk measure: WT with λ = 0.1976						
Model	***n***	**Method**	**True val.**	**Bias in % (s.e.)**	**Std in % (s.e.)**	**rmse in %**
LN	400	OB	0.26	–2.01 (0.31)	21.69 (0.02)	21.79
		IF	0.26	–2.74 (0.3)	21.16 (0.02)	21.34
	1000	OB	0.11	–0.62 (0.2)	13.83 (0)	13.84
		IF	0.11	–0.89 (0.19)	13.57 (0)	13.6
RSLN	400	OB	0.49	–0.23 (0.25)	18.01 (0.03)	18.01
		IF	0.49	–0.73 (0.24)	17.29 (0.03)	17.31
	1000	OB	0.2	–0.68 (0.16)	11.11 (0)	11.13
		IF	0.2	–0.93 (0.15)	10.79 (0)	10.83
Pareto	400	OB	1.12	0.17 (1.01)	71.41 (1.14)	71.41
		IF	1.12	–0.51 (0.99)	70.07 (1.1)	70.07
	1000	OB	0.45	–1.2 (0.57)	40.06 (0.15)	40.08
		IF	0.45	–1.69 (0.56)	39.63 (0.14)	39.66

TABLE 2

VARIANCE ESTIMATION OF CTE AND VAR BOOTSTRAPPED RISK MEASURES FOR DIFFERENT SAMPLE SIZES (n = 400 AND 1000) AND MODELS (LN PUT, RSLN PUT, AND PARETO): COMPARISON OF THE ORDINARY BOOTSTRAP (OB) AND THE DELTA METHOD (IF)

	Risk measure: CTE with α = 0.95					
Model	***n***	**Method**	**True val.**	**Bias in % (s.e.)**	**Std in % (s.e.)**	**rmse in %**
LN	400	OB	13.28	–1.87 (0.44)	31.14 (2.58)	31.2
		IF	13.28	–6.23 (0.44)	31.18 (2.58)	31.8
	1000	OB	5.36	–0.45 (0.29)	20.71 (0.46)	20.72
		IF	5.36	–2.15 (0.29)	20.79 (0.46)	20.9
RSLN	400	OB	14.34	–0.73 (0.45)	31.77 (2.9)	31.78
		IF	14.34	–5.15 (0.45)	31.73 (2.89)	32.15
	1000	OB	5.82	–0.54 (0.3)	20.89 (0.51)	20.9
		IF	5.82	–2.34 (0.3)	20.95 (0.51)	21.08
Pareto	400	OB	65.72	1.56 (2.26)	160.08 (336.83)	160.09
		IF	65.72	1.4 (2.25)	159.1 (332.69)	159.1
	1000	OB	26.77	–0.95 (1.15)	81.42 (35.49)	81.43
		IF	26.77	–0.96 (1.16)	82.01 (36.01)	82.02

	Risk measure: VaR with α = 0.975					
Model	***n***	**Method**	**True val.**	**Bias in % (s.e.)**	**Std in % (s.e.)**	**rmse in %**
LN	400	OB	15.24	5.17 (0.65)	45.9 (6.42)	46.19
		IF	15.24	8.14 (0.83)	58.92 (10.58)	59.48
	1000	OB	6.5	3.7 (0.54)	38.43 (1.92)	38.61
		IF	6.5	5.45 (0.68)	47.91 (2.98)	48.22
RSLN	400	OB	16.69	6.52 (0.68)	47.84 (7.64)	48.28
		IF	16.69	9.97 (0.87)	61.66 (12.69)	62.47
	1000	OB	7.15	4.25 (0.54)	37.85 (2.05)	38.08
		IF	7.15	5.96 (0.67)	47.32 (3.2)	47.69
Pareto	400	OB	36.05	15.46 (1.06)	74.62 (40.15)	76.21
		IF	36.05	21.37 (1.27)	89.6 (57.88)	92.11
	1000	OB	15.3	7.84 (0.72)	50.94 (7.94)	51.54
		IF	15.3	10.97 (0.86)	60.91 (11.35)	61.89

From this simulation study, we have several comments.

- First of all, the two different methods of variance estimation produce very similar values for the PH, the WT and the CTE, but not for VaR, where the influence function produces higher mean squared errors. We conjecture this is due to the fact that empirical quantiles defined on a single or several fixed numbers of order statistics (including the VaR used in this section) are not smooth enough. Heuristically speaking, for any consistent quantile estimator, its bootstrapped IF gets closer as n increases to the original quantile IF, a discontinuous step function. This indicates that the delta method is less satisfactory for a functional depending on local properties even after bootstrapping.

- For the other three risk measures, the OB and IF approaches are barely different. In the CTE case, the OB is marginally (but insignificantly) better in each case, and in the PHT and WT case the IF approach is marginally better in each model.
- Table 3 shows the EB risk measures values (not their variances) for each model. This table indicates that all four measures get larger as the underlying model becomes thicker in tail. Focusing on group 1, the PHT is bigger than and WT for all models (note that they would be the same under a uniform), supporting that its influence puts more weight in the right tail region than the WT would as suggested in Figure 2. Their variances in Table 1 have the same message: the variance of the PHT is larger than that of the WT, for each model. In group 2 of 3, we see that the CTE is consistently bigger than the VaR counterpart across all models (again, they would be the same under a uniform loss distribution); this is true for all continuous loss random variable with unbounded support from above. The difference between the two gets larger as the underlying model's tail gets thicker, as shown in the Pareto case. Table 2 however reveals an interesting observation. The CTE is more stable for the first two models and less stable for the Pareto model. However, in the Pareto case the CTE measure is much further into the tail than the other measures, so greater uncertainty is to be expected.

TABLE 3

EB RISK MEASURE VALUES FOR EACH MODEL WITH N = 1000.

Group		LN put	RSLN2 put	Pareto loss
1	PHT at $\beta = 0.8$	3.78	6.00	16.51
	WT at $\lambda = 0.1976$	3.08	5.08	15.30
2	CTE at $\alpha = 0.95$	30.92	42.61	63.43
	VaR at $\alpha = 0.975$	28.50	40.17	54.32

- In all cases the MSEs of these variance estimators seem large even for samples of size 1,000, especially for the Pareto example, indicating that both methods may be less than satisfactory for fat-tailed distributions in practice. This is particularly the case when the risk measure is heavily weighted in the tail region, which is the case for the PHT in the first group, and the CTE in the second group.

- Since both nonparametric methods presented here rely on asymptotic arguments, we generally cannot consider one over the other without a theoretical justification, and as far as we know there is none. Considering that the OB for variance needs more bootstrap resamplings than for mean, we believe that the IF method can sometimes be advantageous because it does not involve any simulation.

6. Concluding remarks

In this paper we derived the influence function (IF) of the EB quantile estimate and showed it exists in an analytic form where no density function needs to be computed. The result directly extends to the whole L-estimator class. Based on this finding, we conducted a simulation study to estimate the variance of the bootstrapped risk measures and compares it against the ordinary resampling method. The result shows that these two methods would produce comparable results when the corresponding risk measure is smooth.

We also illustrated the impact of the bootstrap on the IF of the distortion risk measure (DRM) using graphs of selected examples. The graphs suggests that bootstrapped DRM would lead to a more stable result than the empirical counterpart, which is consistent with our insight.

Since the proposed method computes the variance of the bootstrapped risk measure analytically, it is faster than the ordinary bootstrap and suitable for applications where variance estimation needs to be repeated. Also, coupled with the practical guideline on the bootstrap usage by Kim and Hardy (2007) this provides a package to estimate the mean squared error of given empirical risk measure non-parametrically.

7. Acknowledgements

Joseph Kim acknowledges the support by the Ph.D. Grant of the Society of Actuaries/Casualty Actuarial Society. The authors also acknowledge the support of the Natural Sciences and Engineering Research Council of Canada, and thank an anonymous referee for detailed comments that substantially improved the presentation.

References

Artzner, P., Delbaen, F., Eber, J.M. and Heath, D. (1999) Coherent measure of risk. *Mathematical Finance*, 203-228.

Booth, J.G. and Sarkar, S. (1998) Monte carlo approximation of bootstrap variances. *The American Statistician*, **52(4)**, 354-357.

Davison, A.C. and Hinkley, D.V. (1997) B*ootstrap methods and their application*. Cambridge University Press, New York.

Dhaene, J., Vanduffel, S., Tang, Q., Goovaerts, M.J., Kaas, R. and Vyncke, D. (2006) Risk measures and comonotonicity: A review. *Stochastic Models*, **22(4)**, 573-606.

EFRON, B. (1992) Jackknife-after-bootstrap standard errors and influence functions. *Journal of the Royal Statistical Society. Series B*, **54(1)**, 83-127.

EFRON, B. and TIBSHIRANI, R.J. (1993) *An introduction to the bootstrap*. Chapman & Hall, New York.

GOURIEROUX, C. and LIU, W. (2006) *Sensitivity analysis of distortion risk measures*. Technical report, University of Toronto.

HALL, P. (1992) *The Bootstap and Edgeworth Expansion*. Springer Series in Statistics. Springer-Verlag, New York.

HAMPEL, F.R., RONCHETTI, E.M., ROUSSEEUW, P.J. and STAHEL, W.A. (1986) *Robust Statistics: The Approach Based on Influence Function*. John Wiley & Sons, New York.

HUTSON, A.D. and ERNST, M.D. (2000) The exact bootstrap mean and variance of an L-estimator. *Journal of Royal Statistical Society: Series B*, **62**, 89-94.

JONES, B.L. and ZITIKIS, R. (2003) Empirical estimation of risk measures and related quantities. *North American Actuarial Journal*, **7(4)**.

KAISER, T. and BRAZAUSKAS, V. (2007) Interval estimation of actuarial risk measures. *North American Actuarial Journal*, **10(4)**.

KIM, J.H.T. and HARDY, M.R. (2007) Quantifying and correcting the bias in estimated risk measures. *ASTIN Bulletin*, **37(2)**, 365-386.

MANISTRE, B.J. and HANCOCK, G.H. (2005) Variance of the CTE estimator. *North American Actuarial Journal*, **9(2)**, 129-156.

SHAO, J. and TU, D. (1995) *The jackknife and bootstrap.* Springer Series in Statistics. Springer, New York.

STAUDTE, R.G. and SHEATHER, S.J. (1990) *Robust estimation and testing*. John Wiley & Sons, New York.

WANG, S.S. (1996) Premium calculation by transforming the layer premium density. *ASTIN Bulletin*, **26(1)**, 71-92.

WIRCH, J.L. and HARDY, M.R. (2000) *Distortion risk measures: Coherence and stochastic dominance.* Woking paper.

MARY R. HARDY
Department of Statistics and Actuarial Science
University of Waterloo
200 University Avenue West
Waterloo, Ontario, Canada N2L 3G1
E-Mail: mrhardy@math.uwaterloo.ca

JOSEPH H.T. KIM (corresponding author)
Department of Statistics and Actuarial Science
University of Waterloo
200 University Avenue West
Waterloo, Ontario, Canada N2L 3G1
E-Mail: jhtkim@math.uwaterloo.ca

APPENDIX

IF of the CTE in Example 3.3

The IF of the CTE was already stated in Manistre and Hancock (2005) but without proof. The derivation is essentially adapted from the IF of the trimmed mean as shown in, e.g., Section 3.2.2 of Staudte and Sheather (1990). From its definition the CTE for continuous F is

$$t(F) = \int_{F^{-1}(\alpha)}^{\infty} \frac{y}{1-\alpha}\, dF(y) = \int_{F^{-1}(\alpha)}^{\infty} \frac{y}{1 - F\left(F^{-1}(\alpha)\right)}\, dF(y). \qquad (26)$$

Therefore by defining $F_\varepsilon(y) = (1-\varepsilon)F(y) + \varepsilon H(y-x)$, with $H(y-x) = H_x(y)$ being the heaviside function,

$$\begin{aligned}
t(F_\varepsilon) &= \int_{F_\varepsilon^{-1}(\alpha)}^{\infty} \frac{y}{1 - F_\varepsilon\left(F_\varepsilon^{-1}(\alpha)\right)}\, dF_\varepsilon(y) \\
&= \int_{F_\varepsilon^{-1}(\alpha)}^{\infty} \frac{y}{1-\alpha}\, dF_\varepsilon(y) \\
&= \int_{F_\varepsilon^{-1}(\alpha)}^{\infty} \frac{y}{1-\alpha}\, d\left((1-\varepsilon)F(y) + \varepsilon H(y-x)\right) \\
&= \frac{1-\varepsilon}{1-\alpha} \int_{F_\varepsilon^{-1}(\alpha)}^{\infty} y dF(y) + \frac{\varepsilon}{1-\alpha} \int_{F_\varepsilon^{-1}(\alpha)}^{\infty} y dH(y-x) \\
&= \frac{1-\varepsilon}{1-\alpha} \int_{F_\varepsilon^{-1}(\alpha)}^{\infty} y dF(y) + \varepsilon \frac{x}{1-\alpha} I\left(F_\varepsilon^{-1}(\alpha) < x\right).
\end{aligned}$$

Differentiating with respect to ε gives

$$\begin{aligned}
\frac{\partial}{\partial \varepsilon} t(F_\varepsilon(y)) = {} & \frac{-1}{1-\alpha} \int_{F_\varepsilon^{-1}(\alpha)}^{\infty} y dF(y) - \frac{1-\varepsilon}{1-\alpha} F_\varepsilon^{-1}(\alpha)\, f\left(F_\varepsilon^{-1}(\alpha)\right) \cdot \frac{\partial}{\partial \varepsilon} F_\varepsilon^{-1}(\alpha) \\
& + \frac{x}{1-\alpha} I\left(F_\varepsilon^{-1}(\alpha) < x\right).
\end{aligned}$$

By setting $\varepsilon = 0$ we have the IF of the CTE

$$\begin{aligned}
L_t(x|F) &= \left. \frac{\partial}{\partial \varepsilon} t(F_\varepsilon(y)) \right|_{\varepsilon=0} \\
&= \frac{-1}{1-\alpha} \int_{F^{-1}(\alpha)}^{\infty} y dF(y) - \frac{1}{1-\alpha} F^{-1}(\alpha)\, f\left(F^{-1}(\alpha)\right) \cdot \left. \frac{\partial}{\partial \varepsilon} F_\varepsilon^{-1}(\alpha) \right|_{\varepsilon=0} \\
&\quad + \frac{x}{1-\alpha} I\left(F^{-1}(\alpha) < x\right),
\end{aligned}$$

where $\frac{\partial}{\partial\varepsilon} F_\varepsilon^{-1}(\alpha)\Big|_{\varepsilon=0}$ is the IF of the quantile. Finally plugging the quantile IF given in (13) into the above equation completes the proof.

Proof of Lemma 4.1

The α–th EB quantile, where $\alpha = r/n$, is defined by

$$t(F) = \int_0^1 J_q(u)\, F^{-1}(u)\, du, \tag{27}$$

where $J_q(u) = u^{r-1}(1-u)^{n-r}/B(r, n-r+1)$. Then from (10) the IF of $t(F)$ is

$$\int_{-\infty}^{x} J_q(F(y))\, dy - \int_{-\infty}^{\infty} (1 - F(y))\, J_q(F(y))\, dy.$$

Replacing $J_q(u) = u^{r-1}(1-u)^{n-r}/B(r, n-r+1)$ leads to

$$\frac{1}{B(r, n-r+1)}\left[\int_{-\infty}^{x} F(y)^{r-1}(1-F(y))^{n-r}\, dy - \int_{-\infty}^{\infty} F(y)^{r-1}(1-F(y))^{n-r+1}\, dy\right]$$

Finally setting $r = n\alpha$ completes the proof.

Since quantiles are not uniquely defined for the empirical distribution function due to its discreteness, other possible α-th quantiles, such as $r = (n+1)\alpha$, may also be used equally. In any case, we note that it is easy to verify that the expected value of the above IF is zero using differentiation. We emphasize that the IF is a function of x.

Proof of Corollary 4.2

First note that

$$\mathbf{c}'\mathbf{w}'\mathbf{X}_{:\mathbf{n}} = \sum_{k=1}^{n} c_k \mathbf{w}_\mathbf{k}' \mathbf{X}_{:\mathbf{n}}.$$

Each term $\mathbf{w}_\mathbf{k}'\mathbf{X}_{:\mathbf{n}}$ on the right side is the bootstrapped quantile. The result follows immediately from (16).