Bootstrapping the Expected Shortfall

Shuxia Sun[1], Fuxia Cheng[2]

[1]Wright State University, Dayton, OH, USA
[2]Illinois State University, Normal, IL, USA
Email: shuxia.sun@wright.edu, fcheng@ilstu.edu

Abstract

The expected shortfall is a popular risk measure in financial risk management. It is defined as the conditional expected loss given that the loss is greater than a given high quantile. We derive the asymptotic properties of the blocking bootstrap estimators for the expected shortfall of a stationary process under strong mixing conditions.

Keywords

High Quantile, Risk Measure, Moving Block Bootstrap, Nonparametric Estimation, Strong Mixing Sample Quantile

1. Introduction

The expected shortfall is a risk measure which has been mostly used among actuaries and insurance companies. The expected shortfall on a portfolio of financial assets is the conditional expected loss given that the loss is greater than a high quantile named as value at risk (VaR). While the expected shortfall and VaR are two popular risk measures, the expected shortfall is becoming increasingly important due to its better properties. For example, the expected shortfall satisfies the sub-additive property, whereas the VaR does not. Sub-additivity, which is one of the four requirements of coherent risk measures, implies in the context of risk management that the total risk on a portfolio should not be greater than the sum of the individual risks. See [1] and [2] for details.

Let $\{X_t\}_{t \in \mathbb{Z}}$ be a sequence of stationary random variables with common distribution function F that describes the negative profit and loss distribution, where $\mathbb{Z} \equiv \{0, \pm 1, \pm 2, \cdots\}$ denotes the set of all integers. For a given positive value p close to zero, the $(1-p)$ confidence level VaR, denoted by v_p, is defined as the $(1-p)$-th high quantile of the loss function F. That is

$$v_p = F^{-1}(1-p) = \inf\{t : F(t) \geq 1-p\}. \tag{1.1}$$

VaR is defined as the loss of a financial position over a time horizon that would be exceeded with small probability p. See [3] [4] [5] for more discussions on VaR. The expected shortfall associated with a confidence level $(1-p)$ on a portfolio, denoted by μ_p, is a conditional expectation defined as follows

$$\mu_p = E\left(X_t \mid X_t > v_p\right).$$

As can be seen from the definitions above, the expected shortfall gives the potential size of the loss that exceeds it, whereas VaR does not. Moreover, as a risk measure, in addition to being coherent, the expected shortfall gives weights to all quantiles greater than the high quantile v_p, whereas the VaR gives all its weight to a single quantile v_p and tells nothing about the size of the loss that exceeds it. Thus, as a risk measure, the expected shortfall is more applicable and produces better incentive for traders than VaR.

In the past, the estimation of expected shortfall has been mainly developed for the identically independently distributed (IID) observations based on the extreme value theory in a parametric or semi-parametric framework. We refer the reader to [6] and [7] for details. However, many empirical studies showed that financial data are often weakly dependent with heavy tails. It is also challenging to build a parametric model that can capture the tail behavior for calculating risk measures since the data are generally sparse in the tail part of the loss distribution. As a result, nonparametric based methods can play an important role in diverse problems in risk managements.

For the weak dependence case, under suitable mixing conditions, [8] first proposed a nonparametric kernel estimator for the expected shortfall in the context of portfolio allocation and derived its asymptotic properties. Reference [9] compared the performance of the sample estimator and the kernel smoothed estimator of the expected shortfall and showed that extra smoothing does not result in more accurate approximation for the expected shortfall.

Although properties of expected shortfall are well studied in the literature, no work seems to be available on the properties of bootstrap approximations for the expected shortfall. Our main contribution in this research is to provide a theoretical foundation to the practical applications of the moving block bootstrap for the expected shortfall.

In this paper, we investigate the asymptotic properties of nonparametric block bootstrap methods for estimating the sampling distribution and the asymptotic variance of expected shortfall under weak dependence setting.

The rest of the paper is organized as follows. In Section 2, we introduce some background material, including a brief description of the moving block bootstrap (MBB) method. In Section 3, we state the main results of this paper. Technical details and proofs will be presented in Section 4.

2. Background

We first define the sample estimator of expected shortfall. For a sample $X_1, \cdots, X_n, n \geq 1$, of stationary random variables with common distribution

function F, let F_n denote the corresponding empirical distribution function, putting mass $1/n$ on each X_i, *i.e.*,

$$F_n(x) = n^{-1} \sum_{i=1}^{n} I(X_i \leq x), \quad x \in \mathbb{R}, \tag{2.1}$$

where $I(\cdot)$ denotes the indicator function which is defined as follows

$$I(S) \equiv \begin{cases} 1, & \text{if the statement } S \text{ is true} \\ 0, & \text{if the statement } S \text{ if false} \end{cases}$$

Then, $\hat{v}_n = F_n^{-1}(1-p)$ is the sample estimator of VaR at confidence level $(1-p)$. The sample estimator of the expected shortfall, denoted by $\hat{\mu}_n$, can be defined as

$$\hat{\mu}_n = \frac{\sum_{i=1}^{n} X_i I(X_i \geq \hat{v}_n)}{\sum_{i=1}^{n} I(X_t \geq \hat{v}_n)} = \frac{1}{\lfloor np \rfloor + 1} \sum_{i=1}^{n} X_i I(X_i \geq \hat{v}_n),$$

where $\lfloor x \rfloor$ denotes the largest integer not exceeding x for $x \in \mathbb{R}$, *i.e.*,

$$\lfloor x \rfloor = \max \{ n \in \mathbb{Z} : n \leq x \}$$

Then, the normalized expected shortfall, denoted by Z_n, is defined as below:

$$Z_n \equiv \sqrt{n} p (\hat{\mu}_n - \mu_p).$$

Our goal in this paper is to investigate the asymptotic properties of bootstrap approximations to the distribution and the variance of the normalized expected shortfall, Z_n.

Reference [10] first introduced the bootstrap method into the statistical world. The bootstrap is a flexible method that can be applied to a variety of problems. It can be used to approximate the quantities of interest such as distribution, bias, variance, significance level in a nonparametric framework.

However, Efron's bootstrap method fails when the data are not independent [11] [12]. Block bootstrap methods for dependent data have been put forward by several authors, notably by [13]-[20]. See [21] and references therein for a detailed account of results on bootstrap methods (for smooth functions of the data) in the dependent case.

It is worth mentioning that although there has been a considerable amount of work on properties of block bootstrap methods for *smooth* functionals of weakly dependent data, not many theoretical results seem to be available on properties of the moving block bootstrap (MBB) and other block bootstrap methods in case of nonsmooth functionals. Reference [22] first showed that blocking bootstrap methods provide a valid approximation to the distribution and its asymptotic variance of a non-smooth function of data, the normalized sample quantile. Recently, due to its applications in financial times series data analysis, quantile based methods (for nonsmooth functions of data) are becoming increasingly attractive, such as expected shortfall [8], quantile hedging [23], risk management [24], among others.

In this paper, we investigate the blocking bootstrap approximation to the expected shortfall based on time series data. For definiteness and conciseness, we shall exclusively concentrate on the MBB method that was independently proposed by [15] and [16].

For the sake of integrity, we briefly describe the MBB method for estimating the sampling distributions of statistics based on weakly dependent observations. Let $X_1, X_2, \cdots, X_n$ be a sample from the stationary process $\{X_i\}_{i \in \mathbb{Z}}$. For ℓ, a positive integer between 1 and n, we define the overlapping blocks of size ℓ as:

$$B_i = (X_i, \cdots, X_{i+\ell-1}), \quad i = 1, \cdots, N = n - \ell + 1.$$

Let $B_1^*, \cdots, B_b^*$ be a random sample of blocks from $\{B_1, \cdots, B_N\}$, where $b = \lfloor n/\ell \rfloor$, that is, $B_1^*, \cdots, B_b^*$ are independently and identically distributed as Uniform$\{B_1, \cdots, B_N\}$.

The observations in the resampled block B_i^* are denoted by $X_{(i-1)\ell+1}^*, \cdots, X_{i\ell}^*, 1 \le i \le b$. The MBB sample consists of $X_1^*, \cdots, X_\ell^*, \cdots, X_{n_1}^*$, where $n_1 = b\ell$. Let

$$T_n \equiv t_n(X_1, \cdots, X_n; \theta) \tag{2.2}$$

be a random quantity of interest that is a function of the random variables $\{X_1, \cdots, X_n\}$ and of some unknown population parameter θ. Then, the MBB version of T_n is defined as

$$T_n^* = t_{n_1}(X_1^*, \cdots, X_{n_1}^*; \hat{\theta}_n), \tag{2.3}$$

where $\hat{\theta}_n$ is a sensible estimator of θ based on $\{X_1, \cdots, X_n\}$. The MBB estimator of the distribution of T_n can be defined as the conditional distribution of T_n^*, given $\mathbf{X} \equiv \{\cdots, X_1, X_2, \cdots\}$. Note that Efron's bootstrap method is a special case of the MBB method with block length $\ell = 1$.

An alternative definition of the MBB version of T_n of (2.2) is given by resampling $\lceil n/\ell \rceil$ blocks from $\{B_1, \cdots, B_N\}$, and using the first n out of the $\lceil n/\ell \rceil \cdot \ell$-many resampled values. However, the difference between the two versions is asymptotically negligible. To simplify the proofs of the main results, here we shall use the version given by (2.3) based on *b complete* resampled blocks.

Throughout this paper, we use P_*, E_*, and Var_* to denote, respectively, the conditional probability, the conditional expectation, and the conditional variance, given $\mathbf{X}$.

Now, we are in a position to define the MBB version of the normalized expected shortfall, Z_n, for a given $p \in (0,1)$. Let F_n^* denote the MBB empirical distribution function, *i.e.*, $F_n^*(x) = n_1^{-1} \sum_{i=1}^{n_1} I(X_i^* \le x), x \in \mathbb{R}$. Then, the MBB version of the sample VaR, $\hat{v}_n = F_n^{-1}(1-p)$, is defined as $v_n^* \equiv F_n^{*-1}(1-p)$. Similarly, the MBB version of the expected shortfall can be defined as below

$$\mu_n^* = \frac{\sum_{i=1}^n X_i^* I(X_i^* \ge \tilde{v}_n)}{\sum_{i=1}^{n_1} I(X_i^* \ge \tilde{v}_n)} = \frac{1}{\lfloor n_1 p \rfloor + 1} \sum_{i=1}^n X_i^* I(X_i^* \ge \tilde{v}_n),$$

where $\tilde{v}_n = \tilde{F}_n^{-1}(p)$, and $\tilde{F}_n(\cdot) = E_* F_n^*$. The MBB version of the normalized expected shortfall, $Z_n = \sqrt{n}\, p\left(\hat{\mu}_n - \mu_p\right)$, is given by

$$Z_n^* \equiv \sqrt{n_1}\, p\left(\mu_n^* - \tilde{\mu}_n\right), \quad \tilde{\mu}_n = E_* \mu_n^*. \qquad (2.4)$$

Note that in the definition of the MBB version of Z_n, we center μ_n^* by $\tilde{\mu}_n$. As in the case of the sample mean [25], and the sample quantile [22], this helps center constant for the MBB sample expected shortfall.

Let

$$H_n(x) = P(Z_n \leq x), \quad x \in \mathbb{R}, \qquad (2.5)$$

denote the distribution function of Z_n. Then, the MBB estimator of $H_n(x)$ is given by the conditional distribution of Z_n^*, *i.e.*, by

$$\hat{H}_n(x) = P_*(Z_n^* \leq x), \quad x \in \mathbb{R}. \qquad (2.6)$$

We conclude this section with an introduction of some standard dependence condition on the X_i's. Suppose that the random variables $\{X_i\}_{i \in \mathbb{Z}}$ are defined on a common probability space $(\Omega, \mathcal{F}, P)$. Let $\mathcal{F}_m^n = \sigma\langle X_i : m \leq i \leq n, i \in \mathbb{Z}\rangle$ be the σ-field generated by random variables $X_m, \cdots, X_n$, $-\infty \leq m \leq n \leq \infty$. For $n \geq 1$, we define

$$\alpha(n) = \sup_{m \in \mathbb{Z}} \sup_{A \in \mathcal{F}_{-\infty}^m, B \in \mathcal{F}_{m+n}^\infty} \left| P(A \cap B) - P(A)P(B) \right|.$$

The sequence $\{X_i\}_{i \in \mathbb{Z}}$ is called *strongly mixing* or *α-mixing* if $\alpha(n) \to 0$ as $n \to \infty$. Strong mixing is a fairly non-restrictive dependence assumption. Empirical studies have showed that many log financial returns are strong mixing with exponential decay coefficients.

As a convention, we assume throughout this paper that unless otherwise specified, limits are taken as $n \to \infty$. Next we state the main results of the paper.

3. Main Results

The first result asserts consistency of the MBB approximation for the distribution function of Z_n.

Theorem 1. *Assume that $0 < p < 1$ with p close to zero and that F has a positive and continuous derivative f in a neighborhood $\mathcal{N}_p$ of v_p with $f(v_p) > 0$. Also, suppose that $\alpha(n) \leq C\rho^n$ for some $C \in (0,\infty)$ and $\rho \in (0,1)$, and that $\ell^{-1} = o(1)$ and $\ell = o\left(n^{1/4}(\log\log n)^{-1/4}\right)$. Then, under the moment condition*

$$E|X_1|^{4+\delta} < \infty, \quad \text{for some } \delta > 0, \qquad (3.1)$$

$$\sup_{x \in \mathbb{R}} \left| P_*\left(\sqrt{n_1}\, p\left(\mu_n^* - \tilde{\mu}_n\right) \leq x\right) - P\left(\sqrt{n}\, p\left(\hat{\mu}_n - \mu_p\right) \leq x\right) \right| = o_p(1). \qquad (3.2)$$

Theorem 1 shows that the MBB method provides a valid approximation to the distribution of the centered and scaled sample expected shortfall Z_n under geometric mixing and under a mild condition on the block length ℓ.

The next result proves the consistency of the MBB variance estimators under

the same conditions as Theorem 1.

Theorem 2. *Under the conditions of Theorem* 1,

$$\hat{\sigma}_n^2 = Var_* \left(\sqrt{n_1} p \left(\mu_n^* - \tilde{\mu}_n \right) \right) \to^p \sigma_\infty^2,$$

where

$$\sigma_\infty^2 = \sum_{i \in \mathbb{Z}} Cov \left(X_1 I \left(X_1 > v_p \right), X_{1+i} I \left(X_{i+1} > v_p \right) \right). \tag{3.3}$$

Theorem 2 shows that under some mild conditions, the MBB estimator $\hat{\sigma}_n^2$ of the asymptotic variance of the centered and scaled expected shortfall at confidence level $(1-p)$ converges in probability.

Remark 1. *Note that in addition to the regularity conditions, we require the moment condition* (3.1) *for both main results. The consistency of the MBB approximation to the distributions of some quantities, as in the cases of sample mean and sample quantile, does not require moment condition although consistency of the MBB variance estimator in general needs some moment condition. The moment condition of Theorem* 1 *may be relaxed.*

4. Proofs

We now introduce some basic notation. Let C, $C(\cdot)$ denote generic constants in $(0, \infty)$ that depend on their arguments (if any) but not on the variables n and x. For real numbers x and y, write $x \vee y = \max\{x, y\}$, $x \wedge y = \min\{x, y\}$.

For any real number x, $i = 1, \cdots, n$, we introduce the following

$$Y_i(x) = X_i I(X_i \geq x), \quad Y_i^*(x) = X_i^* I(X_i^* \geq x)$$

$$U_i(x) = \frac{1}{\ell} \sum_{j=1}^{\ell} Y_{(i-1)\ell+j}(x), \quad U_i^*(x) = \frac{1}{\ell} \sum_{j=1}^{\ell} Y_{(i-1)\ell+j}^*(x)$$

Note that $U_i(\cdot)$ and $U_i^*(\cdot)$, $i = 1, \cdots, b$, are block averages. Then,

$$\hat{\mu}_n(x) = \frac{1}{\lfloor np \rfloor + 1} \sum_{i=1}^{n} Y_i(x), \quad \mu_n^*(x) = \frac{1}{\lfloor n_1 p \rfloor + 1} \sum_{i=1}^{n_1} Y_i^*(x)$$

which implies that

$$\hat{\mu}_n = \hat{\mu}_n(\hat{v}_n), \quad \mu_n^* = \mu_n^*(\tilde{v}_n).$$

For a random variable X and a real number q, we define

$$\|X\|_q \equiv \left(E|X|^q \right)^{\frac{1}{q}}, \quad q \in [1, \infty).$$

Recall that unless otherwise indicated, limits are taken by letting n tend to infinity. The first lemma is Theorem 1 of [9]. It states the asymptotic normality of the centered and scaled sample estimator of expected shortfall, $Z_n = \sqrt{n} p (\hat{\mu}_n - \mu_p)$, for a given $p \in (0,1)$ close to zero. We include this result here for the sake of completeness.

Lemma 1. *Suppose that F is differentiable at v_p with a positive derivative $f(v_p) > 0$ and that $\alpha(n) \leq C \rho^n$ for some $C > 0$ and $0 < \rho < 1$. Then,*

$$Z_n = \sqrt{n}\, p\left(\hat{\mu}_n - \mu_p\right) \to^d N\left(0, \sigma_\infty^2\right), \tag{4.1}$$

where σ_∞^2 is as defined in (3.3).

The next lemma is a consistency result of the MBB variance estimator of σ_∞^2, the asymptotic variance of the normalized expected shortfall.

Lemma 2. *Under conditions of Theorem* 2, *we have*

$$\ell Var_* U_1^*\left(\tilde{v}_n\right) \to^p \sigma_\infty^2, \quad \text{as } n \to \infty.$$

Proof: By Lemma 5.4 (iii) of [22], for $\ell = o\left(n^{1/2}\right)$, we have,

$$\left|\tilde{v}_n - v_p\right| \le C_1 n^{-1/2}\left(\log\log n\right)^{1/2}, \text{ a.s.} \tag{4.2}$$

where $C_1 > 0$ is a constant. Let

$$x_{n1} = v_p - C_1 n^{-1/2}\left(\log\log n\right)^{1/2}, \quad x_{n2} = v_p + C_1 n^{-1/2}\left(\log\log n\right)^{1/2},$$

and

$$R_{i,n} = \frac{1}{\ell}\sum_{j=1}^{\ell}\left\{X_{(i-1)\ell+j}\left[I\left(X_{(i-1)\ell+j} \ge \tilde{v}_n\right) - I\left(X_{(i-1)\ell+j} \ge v_p\right)\right]\right\}, \quad 1 \le i \le N. \tag{4.3}$$

Then, for $r = 1, 2$, $j = 1, \cdots, \ell$,

$$\begin{aligned}
&E\left|X_{(i-1)\ell+j}\left[I\left(X_{(i-1)\ell+j} \ge \tilde{v}_n\right) - I\left(X_{(i-1)\ell+j} \ge v_p\right)\right]\right|^{2r+\delta} \\
&= E\left|X_1\left[I\left(X_1 \ge \tilde{v}_n\right) - I\left(X_1 \ge v_p\right)\right]\right|^{2r+\delta} \\
&\le E\left[\left|X_1\right|^{2r+\delta} I\left(x_{n1} \le X_1 \le x_{n2}\right)\right] \\
&= \int_{x_{n1}}^{x_{n2}} |x|^{2r+\delta} f(x)\,dx \\
&= O\left(n^{-1/2}\left(\log\log n\right)^{1/2}\right),
\end{aligned} \tag{4.4}$$

since $f(x)$ is continuous (and hence bounded) in a neighborhood $\mathcal{N}_p$ of v_p. Because the mixing coefficient $\alpha(\cdot)$ decays exponentially, it is easily seen that, for $1 \le k \le 2r$,

$$\Delta(k, r) = 1 + \sum_{j=1}^{\ell} j^k \left[\alpha(j)\right]^{\delta/(2r+\delta)} < \infty$$

Hence, applying Lemma 3.2 of Lahiri (2003) with $r = 1, 2$, and $1 \le k \le 2r$, we obtain

$$\begin{aligned}
E\left|R_{i,n}\right|^k &\le \ell^{-k} E\left(\sum_{j=1}^{\ell}\left|X_{(i-1)\ell+j}\left[I\left(X_{(i-1)\ell+j} \ge \tilde{v}_n\right) - I\left(X_{(i-1)\ell+j} \ge v_p\right)\right]\right|\right)^k \\
&\le \ell^{-k} \cdot C(r)\Delta(k, r)\left\|X_1\left(I\left(x_{n1} \le X_1 \le x_{n2}\right)\right)\right\|_{2r+\delta}^k \ell^{k/2} \\
&= O\left(\ell^{-k/2} n^{-k/(4r+2\delta)}\left(\log\log n\right)^{k/(4r+2\delta)}\right).
\end{aligned} \tag{4.5}$$

Thus,

$$E\left(\frac{1}{N}\sum_{i=1}^{N}\left|R_{i,n}\right|\right) = E\left|R_{1,n}\right| = O\left(\ell^{-1/2} n^{-1/(4+2\delta)}\left(\log\log n\right)^{1/(4+2\delta)}\right),$$

and

$$E\left(\frac{1}{N}\sum_{i=1}^{N}\left|R_{i,n}\right|^2\right) = E\left|R_{1,n}\right|^2 = O\left(\ell^{-1}n^{-1/(2+\delta)}\left(\log\log n\right)^{1/(2+\delta)}\right).$$

which together with the Markov's inequality lead to

$$\frac{1}{N}\sum_{i=1}^{N}\left|R_{i,n}\right| = O_p\left(\ell^{-1/2}n^{-1/(4+2\delta)}\left(\log\log n\right)^{1/(4+2\delta)}\right), \qquad (4.6)$$

and

$$\frac{1}{N}\sum_{i=1}^{N}\left|R_{i,n}\right|^2 = O_p\left(\ell^{-1}n^{-1/(2+\delta)}\left(\log\log n\right)^{1/(2+\delta)}\right). \qquad (4.7)$$

Define,

$$\bar{U}_N = \frac{1}{N}\sum_{i=1}^{N}U_i\left(v_p\right), \quad \bar{R}_N = \frac{1}{N}\sum_{i=1}^{N}R_{i,n}. \qquad (4.8)$$

Now, we investigate $Var_*\left[U_1^*\left(\tilde{v}_n\right)\right]$.

$$\begin{aligned}
Var_*\left[U_1^*\left(\tilde{v}_n\right)\right] &= \frac{1}{N}\sum_{i=1}^{N}\left[U_i\left(\tilde{v}_n\right) - \frac{1}{N}\sum_{i=1}^{N}U_i\left(\tilde{v}_n\right)\right]^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}\left[\left(U_i\left(v_p\right) + R_{i,n}\right) - \frac{1}{N}\sum_{i=1}^{N}\left(U_i\left(v_p\right) + R_{i,n}\right)\right]^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}\left[\left(U_i\left(v_p\right) - \bar{U}_N\right) + \left(R_{i,n} - \bar{R}_N\right)\right]^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}\left[U_i\left(v_p\right) - \bar{U}_N\right]^2 + \frac{1}{N}\sum_{i=1}^{N}\left[R_{i,n} - \bar{R}_N\right]^2 \\
&\quad + \frac{2}{N}\sum_{i=1}^{N}\left[U_i\left(v_p\right) - \bar{U}_N\right]\left[R_{i,n} - \bar{R}_N\right] \\
&= Var_*\left[U_1^*\left(v_p\right)\right] + \frac{1}{N}\sum_{i=1}^{N}\left[R_{i,n} - \bar{R}_N\right]^2 \\
&\quad + \frac{2}{N}\sum_{i=1}^{N}\left[U_i\left(v_p\right) - \bar{U}_N\right]\left[R_{i,n} - \bar{R}_N\right].
\end{aligned} \qquad (4.9)$$

Theorem 3.1 of Lahiri (2003) implies that

$$\ell Var_*\left[U_1^*\left(v_p\right)\right] \to^p \sigma_\infty^2. \qquad (4.10)$$

Next, we show that

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^{N}\left[R_{i,n} - \bar{R}_N\right]^2 \\
&= O_p\left(\ell^{-1}n^{-1/(2+\delta)}\left(\log\log n\right)^{1/(2+\delta)}\right),
\end{aligned} \qquad (4.11)$$

and

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^{N}\left[U_i\left(v_p\right) - \bar{U}_N\right]\left[R_{i,n} - \bar{R}_N\right] \\
&= O_p\left(\ell^{-1/2}n^{-(1+\delta)/(4+2\delta)}\left(\log\log n\right)^{1/(4+2\delta)}\right).
\end{aligned} \qquad (4.12)$$

Using (4.6), (4.7), we get

$$\frac{1}{N}\sum_{i=1}^{N}\left[R_{i,n}-\bar{R}_{N}\right]^{2} \leq \frac{2}{N}\sum_{i=1}^{N}\left(\left|R_{i,n}\right|^{2}+\left|\bar{R}_{N}\right|^{2}\right)$$

$$\leq \frac{2}{N}\sum_{i=1}^{N}\left|R_{i,n}\right|^{2}+2\left(\frac{1}{N}\sum_{i=1}^{N}\left|R_{i,n}\right|\right)^{2}$$

$$= O_{p}\left(\ell^{-1}n^{-1/(2+\delta)}\left(\log\log n\right)^{1/(2+\delta)}\right)$$

$$+ O_{p}\left(\ell^{-1}n^{-2/(4+2\delta)}\left(\log\log n\right)^{2/(4+2\delta)}\right)$$

$$= O_{p}\left(\ell^{-1}n^{-1/(2+\delta)}\left(\log\log n\right)^{1/(2+\delta)}\right).$$

Hence, Equation (4.11) is proved. By (4.10), (4.11), and Cauchy-Schwartz inequality, we have,

$$\frac{1}{N}\sum_{i=1}^{N}\left[U_{i}\left(v_{p}\right)-\bar{U}_{N}\right]\left[R_{i,n}-\bar{R}_{N}\right]$$

$$\leq \frac{1}{N}\left(\sum_{i=1}^{N}\left[U_{i}\left(v_{p}\right)-\bar{U}_{N}\right]^{2}\right)^{1/2}\left(\sum_{i=1}^{N}\left[R_{i,n}-\bar{R}_{N}\right]^{2}\right)^{1/2}$$

$$= \frac{1}{\sqrt{N}}\left(Var_{*}\left[U_{1}^{*}\left(v_{p}\right)\right]\right)^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left|R_{i,n}\right|^{2}\right)^{1/2}$$

$$= N^{-1/2}\cdot O_{p}\left(1\right)\cdot O_{p}\left(\ell^{-1/2}n^{-1/(4+2\delta)}\left(\log\log n\right)^{1/(4+2\delta)}\right)$$

$$= O_{p}\left(\ell^{-1/2}n^{-(1+\delta)/(4+2\delta)}\left(\log\log n\right)^{1/(4+2\delta)}\right).$$

We complete the proof of (4.12). Combining (4.9)-(4.12), we obtain

$$\ell Var_{*}\left[U_{1}^{*}\left(\tilde{v}_{n}\right)\right] = \sigma_{\infty}^{2}+\ell\cdot O_{p}\left(\ell^{-1}n^{-1/(2+\delta)}\left(\log\log n\right)^{1/(2+\delta)}\right)$$

$$+ \ell\cdot O_{p}\left(\ell^{-1/2}n^{-(1+\delta)/(4+2\delta)}\left(\log\log n\right)^{1/(4+2\delta)}\right)$$

$$= \sigma_{\infty}^{2}+O_{p}\left(\ell^{1/2}n^{-(1+\delta)/(4+2\delta)}\left(\log\log n\right)^{1/(4+2\delta)}\right)$$

$$= \sigma_{\infty}^{2}+o_{p}\left(1\right)$$

Here we used the condition on the block length $\ell = o\left(n^{1/4}\left(\log\log n\right)^{-1/4}\right)$. We complete the proof of Lemma 2. □

Lemma 3 below gives a convergence result of the third moment of the MBB block average.

Lemma 3. *Under the conditions of Theorem* 2,

$$\frac{n\sqrt{n}}{b^{2}}E_{*}\left|U_{1}^{*}\left(\tilde{v}_{n}\right)-E_{*}U_{1}^{*}\left(\tilde{v}_{n}\right)\right|^{3} \to^{p} 0.$$

Proof: It can be verified by using (4.5) with $k=3, r=2$,

$$E\left(\frac{1}{N}\sum_{i=1}^{N}\left|R_{i,n}\right|^{3}\right) = E\left|R_{1,n}\right|^{3} = O\left(\ell^{-3/2}n^{-3/(8+2\delta)}\left(\log\log n\right)^{3/(8+2\delta)}\right),$$

which leads to

$$\frac{1}{N}\sum_{i=1}^{N}\left|R_{i,n}\right|^{3} = O_{p}\left(\ell^{-3/2}n^{-3/(8+2\delta)}\left(\log\log n\right)^{3/(8+2\delta)}\right). \quad (4.13)$$

Using (4.6), (4.13), and the fact that

$$\left(x + y \right)^{m} \le 2^{m} \left(x^{m} + y^{m} \right), \quad \forall m, x, y > 0 \tag{4.14}$$

we obtain,

$$\frac{1}{N} \sum_{i=1}^{N} \left| R_{i,n} - \bar{R}_{N} \right|^{3}$$

$$\le \frac{1}{N} \sum_{i=1}^{N} 2^{3} \left(\left| R_{i,n} \right|^{3} + \left| \bar{R}_{N} \right|^{3} \right)$$

$$\le \frac{8}{N} \sum_{i=1}^{N} \left| R_{i,n} \right|^{3} + 8 \left(\frac{1}{N} \sum_{i=1}^{N} \left| R_{i,n} \right| \right)^{3}$$

$$= O_{p} \left(\ell^{-3/2} n^{-3/(8+2\delta)} \left(\log \log n \right)^{3/(8+2\delta)} \right).$$

Therefore,

$$\frac{1}{N} \sum_{i=1}^{N} \left| R_{i,n} - \bar{R}_{N} \right|^{3} = O_{p} \left(\ell^{-3/2} n^{-3/(8+2\delta)} \left(\log \log n \right)^{3/(8+2\delta)} \right). \tag{4.15}$$

Note that for $1 \le j \le \ell$, $r = 1, 2$, and $1 \le k \le 2r$,

$$\left\| Y_{j} \left(v_{p} \right) \right\|_{2r+\delta} \le \left\| X_{1} \right\|_{2r+\delta} < \infty$$

and

$$\Delta \left(k, r \right) = 1 + \sum_{j=1}^{\ell} j^{k} \left[\alpha \left(j \right) \right]^{\delta/(2r+\delta)} < \infty$$

Then, Lemma 3.2 of Lahiri (2003) implies,

$$E \left| \sum_{j=1}^{\ell} Y_{j} \left(v_{p} \right) \right|^{k} \le \Delta \left(k, r \right) \left\| Y_{1} \left(v_{p} \right) \right\|_{2r+\delta}^{k} \ell^{k/2} \le C \left(k, r \right) \left\| X_{1} \right\|_{2r+\delta}^{k} \ell^{k/2} = O \left(\ell^{k/2} \right) \tag{4.16}$$

which implies

$$E \left(\frac{1}{N} \sum_{i=1}^{N} \left| U_{i} \left(v_{p} \right) \right|^{3} \right) = E \left| U_{1} \left(v_{p} \right) \right|^{3} = E \left(\frac{1}{\ell} \sum_{j=1}^{\ell} Y_{j} \left(v_{p} \right) \right)^{3} = O \left(\ell^{-3/2} \right)$$

and

$$E \left(\frac{1}{N} \sum_{i=1}^{N} \left| U_{i} \left(v_{p} \right) \right| \right) = E \left| U_{1} \left(v_{p} \right) \right| = E \left(\frac{1}{\ell} \sum_{j=1}^{\ell} Y_{j} \left(v_{p} \right) \right) = O \left(\ell^{-1/2} \right)$$

Therefore,

$$\frac{1}{N} \sum_{i=1}^{N} \left| U_{i} \left(v_{p} \right) \right|^{3} = O_{p} \left(\ell^{-3/2} \right), \quad \frac{1}{N} \sum_{i=1}^{N} \left| U_{i} \left(v_{p} \right) \right| = O_{p} \left(\ell^{-1/2} \right). \tag{4.17}$$

Finally, combining (4.14), (4.15), and (4.17) gives,

$$E_{*} \left| U_{1}^{*} \left(\tilde{v}_{n} \right) - E_{*} U_{1}^{*} \left(\tilde{v}_{n} \right) \right|^{3}$$

$$= \frac{1}{N} \sum_{i=1}^{N} \left| U_{i} \left(\tilde{v}_{n} \right) - \frac{1}{N} \sum_{i=1}^{N} U_{i} \left(\tilde{v}_{n} \right) \right|^{3}$$

$$= \frac{1}{N} \sum_{i=1}^{N} \left| \left(U_{i} \left(v_{p} \right) - \bar{U}_{N} \right) + \left(R_{i,n} - \bar{R}_{N} \right) \right|^{3}$$

$$\le \frac{1}{N} \sum_{i=1}^{N} 2^{3} \left(\left| U_{i} \left(v_{p} \right) - \bar{U}_{N} \right|^{3} + \left| R_{i,n} - \bar{R}_{N} \right|^{3} \right)$$

$$= \frac{8}{N} \sum_{i=1}^{N} \left| U_i(\nu_p) - \bar{U}_N \right|^3 + \frac{8}{N} \sum_{i=1}^{N} \left| R_{i,n} - \bar{R}_N \right|^3$$

$$= \frac{64}{N} \sum_{i=1}^{N} \left| U_i(\nu_p) \right|^3 + 64 \left| \bar{U}_N \right|^3 + \frac{8}{N} \sum_{i=1}^{N} \left| R_{i,n} - \bar{R}_N \right|^3 \qquad (4.18)$$

$$= O_p \left(\ell^{-3/2} + \ell^{-3/2} n^{-3/(8+2\delta)} \left(\log \log n \right)^{3/(8+2\delta)} \right).$$

Thus, we obtain,

$$\frac{n\sqrt{n}}{b^2} E_* \left| U_1^*(\tilde{\nu}_n) - E_* U_1^*(\tilde{\nu}_n) \right|^3$$

$$= \ell^2 n^{-1/2} \cdot O_p \left(\ell^{-3/2} + \ell^{-3/2} n^{-3/(4+2\delta)} \left(\log \log n \right)^{3/(4+2\delta)} \right)$$

$$= o_p(1)$$

Lemma 3 is proved. □

Proof of Theorem 1: By the definitions of μ_n^*, $\tilde{\mu}_n$, Y_i^*, U_i^*, $n_1 = b\ell$, and the fact that $U_1^*, U_2^*, \cdots, U_b^*$ are IID, we get

$$\mu_n^* - \tilde{\mu}_n = \frac{\sum_{i=1}^{n} X_i^* I\left(X_i^* \geq \tilde{\nu}_n \right)}{\sum_{i=1}^{n_1} I\left(X_i^* \geq \tilde{\nu}_n \right)} - E_* \left(\frac{\sum_{i=1}^{n} X_i^* I\left(X_i^* \geq \tilde{\nu}_n \right)}{\sum_{i=1}^{n_1} I\left(X_i^* \geq \tilde{\nu}_n \right)} \right)$$

$$= \frac{n_1}{\lfloor n_1 p \rfloor + 1} \left\{ \frac{1}{n_1} \sum_{i=1}^{n_1} Y_i^*(\tilde{\nu}_n) - E_* \left(\frac{1}{n_1} \sum_{i=1}^{n_1} Y_i^*(\tilde{\nu}_n) \right) \right\}$$

$$= \frac{n_1}{\lfloor n_1 p \rfloor + 1} \left\{ \frac{1}{b} \sum_{i=1}^{b} U_i^*(\tilde{\nu}_n) - E_* \left(\frac{1}{b} \sum_{i=1}^{b} U_i^*(\tilde{\nu}_n) \right) \right\} \qquad (4.19)$$

$$= \frac{n_1}{\lfloor n_1 p \rfloor + 1} \left(\frac{1}{b} \sum_{i=1}^{b} \left[U_i^*(\tilde{\nu}_n) - E_* U_i^*(\tilde{\nu}_n) \right] \right)$$

Then, for any $y \in \mathbb{R}$,

$$P_* \left(\sqrt{n_1} p \left(\mu_n^* - \tilde{\mu}_n \right) \leq y \right)$$

$$= P_* \left(\frac{1}{n_1} \sum_{i=1}^{n_1} \left[Y_i^*(\tilde{\nu}_n) - \tilde{\mu}_n \right] \leq \frac{\lfloor n_1 p \rfloor + 1}{p n_1 \sqrt{n_1}} y \right)$$

$$= P_* \left(\frac{1}{b} \sum_{i=1}^{b} \left[U_i^*(\tilde{\nu}_n) - E_* U_i^*(\tilde{\nu}_n) \right] \leq \frac{\lfloor n_1 p \rfloor + 1}{p n_1 \sqrt{n_1}} y \right),$$

which together with the Berry-Esseen Theorem gives

$$\left| P_* \left(\sqrt{n_1} p \left(\mu_n^* - \tilde{\mu}_n \right) \leq y \right) - \Psi \left(\frac{\sqrt{b}}{\sqrt{Var_* U_1^*(\tilde{\nu}_n)}} \frac{\lfloor n_1 p \rfloor + 1}{p n_1 \sqrt{n_1}} y \right) \right|$$

$$\leq \frac{3 E_* \left| U_1^*(\tilde{\nu}_n) - E_* U_1^*(\tilde{\nu}_n) \right|^3}{\sqrt{b} \left(Var_* U_1^*(\tilde{\nu}_n) \right)^{3/2}} \qquad (4.20)$$

Lemma 2 and Lemma 3 imply, respectively,

$$\frac{\sqrt{b}}{\sqrt{Var_* U_1^*(\tilde{\nu}_n)}} \frac{\lfloor n_1 p \rfloor + 1}{p n_1 \sqrt{n_1}} y = \frac{\lfloor n_1 p \rfloor + 1}{p n_1} \frac{y}{\sqrt{\ell Var_* U_1^*(\tilde{\nu}_n)}} \to^p \frac{y}{\sigma_\infty}, \qquad (4.21)$$

and

$$\frac{E_*\left|U_1^*\left(\tilde{v}_n\right)-E_*U_1^*\left(\tilde{v}_n\right)\right|^3}{\sqrt{b}\left(Var_*U_1^*\left(\tilde{v}_n\right)\right)^{3/2}}=\frac{\frac{n\sqrt{n}}{b^2}E_*\left|U_1^*\left(\tilde{v}_n\right)-E_*U_1^*\left(\tilde{v}_n\right)\right|^3}{\left(\ell Var_*U_1^*\left(\tilde{v}_n\right)\right)^{3/2}}\to^p 0 \qquad (4.22)$$

Then, by (4.20)-(4.22), we have

$$\left|P_*\left(\sqrt{n_1}\,p\left(\mu_n^*-\tilde{\mu}_n\right)\le y\right)-\Psi\left(\frac{y}{\sigma_\infty}\right)\right|\to^p 0,$$

uniformly in $y\in\mathbb{R}$. This together with Lemma 1 yields

$$\sup_{y\in\mathbb{R}}\left|P_*\left(\sqrt{n_1}\,p\left(\mu_n^*-\tilde{\mu}_n\right)\le y\right)-P\left(\sqrt{n}\,p\left(\hat{\mu}_n-\mu_p\right)\le y\right)\right|\to^p 0.$$

That is, $Z_n^*=\sqrt{n_1}\,p\left(\mu_n^*-\tilde{\mu}_n\right)$ converges in probability to $Z_n=\sqrt{n}\,p\left(\hat{\mu}_n-\mu_p\right)$. We complete the proof of Theorem 1. □

Proof of Theorem 2: Theorem 2 is a consequence of Lemma 2 and Equation (4.19).

$$Var_*\left(Z_n^*\right)=Var_*\left(\sqrt{n_1}\,p\mu_n^*\right)$$

$$=Var_*\left(\frac{n_1\sqrt{n_1}\,p}{\lfloor n_1 p\rfloor+1}\cdot\frac{1}{n_1}\sum_{i=1}^{n_1}Y_i^*\right)$$

$$=\left(\frac{n_1\sqrt{n_1}\,p}{\lfloor n_1 p\rfloor+1}\right)^2 Var_*\left(\frac{1}{b}\sum_{i=1}^{b}U_i^*\left(\tilde{v}_n\right)\right)$$

$$=\left(\frac{n_1 p}{\lfloor n_1 p\rfloor+1}\right)^2\left(\ell Var_*\left(U_1^*\left(\tilde{v}_n\right)\right)\right)$$

$$\to\sigma_\infty^2,$$

since $U_1^*,U_2^*,\cdots,U_b^*$ are IID, which implies that

$$Var_*\left(\frac{1}{b}\sum_{i=1}^{b}U_i^*\left(\tilde{v}_n\right)\right)=\frac{1}{b}Var_*\left(U_1^*\left(\tilde{v}_n\right)\right)=\frac{1}{n_1}\left(\ell Var_*\left(U_1^*\left(\tilde{v}_n\right)\right)\right).$$

The proof of Theorem 2 is completed. □

5. Conclusions

In this paper, we establish the asymptotic properties of the blocking bootstrap estimators of the expected shortfall. We prove that the MBB method provides a valid approximation to the distribution of the centered and scaled sample estimator of the expected shortfall and show that under mild regularity conditions, the MBB variance estimator is consistent.

As in many situations where the block bootstrap methodology is involved, the performance of the block bootstrap distribution function and variance estimators of Z_n critically depends on the block size ℓ. Although there have been some theoretical works on the choice of the optimal block length in the literature, the optimal block lengths and/or the optimal rates of convergence in Theorems 1 and 2 are unknown at this stage.

It is also of interest to conduct both empirical studies and simulations to investigate the performance of the MBB estimators and compare the obtained results with those in [9].

Acknowledgements

References

[1] Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999) Coherent Measures of Risk. *Mathematical Finance*, **9**, 203-228. https://doi.org/10.1111/1467-9965.00068

[2] Föllmer, H. and Schied, A. (2002) Convex Measures of Risk and Trading Constraints. *Finance and Stochastic*, **6**, 429-447. https://doi.org/10.1007/s007800200072

[3] Duffie, D. and Pan, J. (1997) An Overview of Value at Risk. *Journal of Derivative*, **4**, 7-49. https://doi.org/10.3905/jod.1997.407971

[4] Jorion, P. (2001) Value at Risk. 2nd Edition, McGraw-Hill, New York.

[5] Chen, S.X. and Tang, C.Y. (2005) Nonparametric Inference of Value at Risk for Dependent Financial Returns. *Journal of Financial Econometrics*, **3**, 227-255. https://doi.org/10.1093/jjfinec/nbi012

[6] Embrechts, P., Klueppelberg, C. and Mikosch, T. (1997) Modeling Extremal Events for Insurance and Finance. Springer, Berlin. https://doi.org/10.1007/978-3-642-33483-2

[7] McNeil, A. (1997) Estimating the Tails of Loss Severity Distributions Using Extreme Value Theory. *ASTIN Bulletin*, **27**, 117-137. https://doi.org/10.2143/AST.27.1.563210

[8] Scaillet, O. (2004) Nonparametric Estimation and Sensitivity Analysis of Expected Shortfall. *Mathematical Finance*, **14**, 115-129. https://doi.org/10.1111/j.0960-1627.2004.00184.x

[9] Chen, S.X. (2008) Nonparametric Estimation of Expected Shortfall. *Journal of Financial Economics*, 87-107.

[10] Efron, B. (1979) Bootstrap Methods: Another Look at the Jackknife. *Annals of Statistics*, **7**, 1-26. https://doi.org/10.1214/aos/1176344552

[11] Bickel, P.J. and Freedman, D.A. (1981) Some Asymptotic Theory for the Bootstrap. *Annals of Statistics*, **9**, 1196-1217. https://doi.org/10.1214/aos/1176345637

[12] Singh, K. (1981) On Asymptotic Accuracy of Efron's Bootstrap. *Annals of Statistics*, **9**, 1187-1195. https://doi.org/10.1214/aos/1176345636

[13] Hall, P. (1985) Resampling a Coverage Pattern. *Stochastic Processes and Their Applications*, **20**, 231-246. https://doi.org/10.1016/0304-4149(85)90212-1

[14] Carlstein, E. (1986) The Use of Subseries Methods for Estimating the Variance of a General Statistic from a Stationary Time Series. *Annals of Statistics*, **14**, 1171-1179. https://doi.org/10.1214/aos/1176350057

[15] Künsch, H.R. (1989) The Jackknife and the Bootstrap for General Stationary Observations. *Annals of Statistics*, **17**, 1217-1261. https://doi.org/10.1214/aos/1176347265

[16] Liu, R.Y. and Singh, K. (1992) Moving Blocks Jackknife and Bootstrap Capture Weak Convergence. In: Lepage, R. and Billard, L., Eds., *Exploring the Limits of*

Bootstrap, Wiley, New York, 225-248.

[17] Politis, D. and Romano, J.P. (1992) A Circular Block Resampling Procedure for Stationary Data. In: Lepage, R. and Billard, L., Eds., *Exploring the Limits of Bootstrap*, Wiley, New York, 263-270.

[18] Politis, D. and Romano, J.P. (1994) Stationary Bootstrap. *Journal of the American Statistical Association*, **89**, 1303-1313.
https://doi.org/10.1080/01621459.1994.10476870

[19] Politis, D., Romano, J.P. and Wolf, M. (1997) Subsampling for Heteroskedastic Time Series. *Journal of Econometrics*, **81**, 281-317.
https://doi.org/10.1016/S0304-4076(97)86569-4

[20] Paparoditis, E. and Politis, D.N. (2001) Tapered Block Bootstrap. *Biometrika*, **88**, 1105-1119. https://doi.org/10.1093/biomet/88.4.1105

[21] Lahiri, S.N. (2003) Resampling Methods for Dependent Data. Springer-Verlag, New York. https://doi.org/10.1007/978-1-4757-3803-2

[22] Sun, S. and Lahiri, S.N. (2006) Bootstrapping the Sample Quantile Based on Weakly Dependent Observations. *Sankhyā*, **68**, 130-166.

[23] Föllmer, H. and Leukert, P. (1999) Quantile Hedging. *Finance and Stochastic*, **3**, 251-273. https://doi.org/10.1007/s007800050062

[24] Melnikov, S. and Romaniuk, Y. (2006) Evaluating the Performance of Gompertz, Makeham and Lee-Carter Mortality Models for Risk Management with Unit-Linked Contracts. *Insurance: Mathematics and Economics*, **39**, 310-329.
https://doi.org/10.1016/j.insmatheco.2006.02.012

[25] Lahiri, S.N. (1992) Edgeworth Correction by Moving Block Bootstrap for Stationary and Non Stationary Data. In: Lepage, R. and Billard, L., Eds., *Exploring the Limits of Bootstrap*, Wiley, New York, 225-248.