Proceedings of the 2009 Winter Simulation Conference
M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, eds.

A GENERAL FRAMEWORK OF IMPORTANCE SAMPLING FOR VALUE-AT-RISK AND CONDITIONAL VALUE-AT-RISK

Lihua Sun
L. Jeff Hong

Department of Industrial Engineering and Logistics Management
The Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong, China

ABSTRACT

Value-at-risk (VaR) and conditional value-at-risk (CVaR) are important risk measures. Importance sampling (IS) is often used to estimate them. We derive the asymptotic representations for IS estimators of VaR and CVaR. Based on these representations, we are able to give simple conditions under which the IS estimators have smaller asymptotic variances than the ordinal estimators. We show that the exponential twisting can yield an IS distribution that satisfies the conditions for both the IS estimators of VaR and CVaR. Therefore, we may be able to estimate VaR and CVaR accurately at the same time.

1 INTRODUCTION

Value-at-risk (VaR) and conditional value-at-risk (CVaR) are two widely used risk measures. They play important roles in risk measurement, portfolio management and regulatory control of financial institutions. The α-VaR, defined as the α quantile of a portfolio value L. The Basel Accord II has incorporated the concept of VaR and encourages banks to use VaR for daily risk management. It has the advantage of being flexible and conceptually simple. However, VaR also has some shortcomings. It provides no information about the amount of loss that investors might suffer beyond the VaR. Moreover, it has also been criticized for not being a coherent risk measure because it is not subadditive (Artzner et al. 1999). The α-CVaR, defined as the average of $\beta-$VaR of L for $0<\beta<\alpha$, has a long history of being used in insurance industry. It provides information on the potential large losses that an investor can suffer. Rockafellar and Uryasev (2000) showed that CVaR always satisfies the subadditivity and is therefore a coherent risk measure. Heyde et. al (2007), however, argued that requiring subadditivity may lead to risk measures that are not robust to the underlying models and data. They proposed replacing subadditivity axiom by the comonotonic subadditivity axiom and found that both VaR and CVaR satisfy the new axiom and VaR is more robust.

Because VaR and CVaR both have advantages and disadvantages, it is often difficult to decide which one is better. Risk managers may consider both of them at the same time to complement each other. There are typically three approaches to estimate VaR and CVaR: the variance-covariance approach, the historical simulation approach and the Monte Carlo simulation approach. Among the three, the Monte Carlo simulation approach is frequently used, since it is very general and can be applied to a wide range of risk models. However, the Monte Carlo simulation approach is often time consuming. In risk management, α is typically close to 0. A large number of replications are needed to obtain accurate estimation of the tail behavior. Therefore, variance reduction techniques are often used to increase the efficiency of the estimation. Among all variance reduction techniques, importance sampling (IS) is a natural choice, because it can allocate more samples to the tail of the distribution that is most relevant to the estimation of VaR and CVaR. In this paper, we study the asymptotic properties of the IS estimators of VaR and CVaR and propose simple schemes for choosing IS distributions.

Because IS estimators of both VaR and CVaR are rather complicated compared to the typical sample means, to analyze their asymptotic properties, we use the method of asymptotic representations. Bahadur (1966) used this method to analyze the asymptotic properties of the ordinary estimator of VaR (quantile) by showing that the estimator can be approximated by a sample mean except for a high-order term. Then, the consistency and asymptotic normality of the estimator can be derived easily. In this paper, we derive the asymptotic representations for the IS estimators of both VaR and CVaR, and use them to prove the consistency and asymptotic normality of both estimators. From the asymptotic normality, we give simple conditions on the IS distributions under which the IS scheme is guaranteed to work asymptotically. A good feature of the conditions is that they are the same for both VaR and CVaR. Therefore, one can estimate the VaR and CVaR simultaneously using the same IS distribution. This feature will greatly help those risk managers who consider both risk measures and use them to complement each others.

We then consider how to choose a good IS distribution for estimating both VaR and CVaR. We consider this problem under a special family of change of measure that is known as exponential twisting and show how to choose good IS distributions from the family. Interestingly, the distributions that we choose for VaR and CVaR are again the same. We show that the chosen IS distribution guarantees to reduce the asymptotic variances for both VaR and CVaR under mild conditions.

The literature on using IS to estimate VaR is growing rapidly. For example, Glasserman et al. (2000, 2002) used IS to estimate the VaR of a portfolio loss for both light-tail and heavy-tail situations; Glasserman and Li (2005) applied IS to estimate the VaR of portfolio credit risk; Glasserman and Juneja (2008) used IS to estimate the VaR of a sum of i.i.d random variables. To the best of our knowledge, however, there is no published work on using IS to estimate CVaR. Our paper differs from others for several reasons. We consider the general theory of using IS to estimate VaR and CVaR, without focusing on specific applications. We show that the variances of VaR and CVaR can be reduced by using the same IS distribution. Therefore, risk managers who use both VaR and CVaR can obtain accurate estimates of both without running two separate simulation experiments.

The rest of the paper is organized as follows. Section 2 reviews the IS estimators of VaR and CVaR. In section 3, we develop asymptotic representations for the IS estimators for both VaR and CVaR. From these representations, we can easily prove the consistency and asymptotic normality of the IS estimators. In section 4, we consider the exponential twisting to choose an IS distribution, followed by the conclusions in section 5.

2 IMPORTANCE SAMPLING FOR VAR AND CVAR

Let L be a real-valued random variable with a cumulative distribution function (c.d.f.) $F(\cdot)$, and let v_α and c_α denote the α-VaR and α-CVaR of L, respectively, for any $0<\alpha<1$. Then,

$$v_\alpha = F^{-1}(\alpha) = \inf\{x : F(x) \geq \alpha\}, \tag{1}$$

$$c_\alpha = v_\alpha - \frac{1}{\alpha}\mathrm{E}[v_\alpha - L]^+, \tag{2}$$

where $x^+ = \max\{x,0\}$. Note that v_α is also the α-quantile of L, and c_α is the average of L conditioning that $L \leq v_\alpha$ if L has a positive density at v_α. Under the definitions, we are interested in the left tail of the distribution of L. Therefore, α is often close to 0. Sometimes, v_α and c_α are defined for the right tail of a random variable (e.g., Hong and Liu (2009)). We may convert the right tail to the left tail by adding a negative sign to the random variable.

Conventional Monte Carlo estimation of v_α and v_α involves generating n independent and identically distributed (i.i.d) random observations of L, denoted as $L_1,\ldots,L_n$, and estimating them by

$$\tilde{v}_\alpha^n = \tilde{F}_n^{-1}(\alpha) = \inf\{x : \tilde{F}_n(x) \geq \alpha\}, \tag{3}$$

$$\tilde{c}_\alpha^n = \tilde{v}_\alpha^n - \frac{1}{n\alpha}\sum_{i=1}^n [\tilde{v}_\alpha^n - L_i]^+, \tag{4}$$

respectively, where

$$\tilde{F}_n(x) = \frac{1}{n}\sum_{i=1}^n I\{L_i \leq x\} \tag{5}$$

is the empirical distribution of L constructed from $L_1,\ldots,L_n$ and $I\{\cdot\}$ is the indicator function. Note that $\tilde{F}_n(x)$ is an unbiased and consistent estimator of $F(x)$. By Serfling (1980), $\tilde{F}_n^{-1}(\alpha) = L_{\lceil n\alpha\rceil:n}$, where $L_{i:n}$ is the ith order statistic of the L. Serfling (1980) showed that $\tilde{v}_\alpha^n$ and $\tilde{c}_\alpha^n$ are consistent estimators of v_α and c_α, respectively, as $n \to \infty$.

As α is typically chosen close to 0, e.g., $\alpha = 0.01$ or 0.05, to estimate v_α and c_α accurately using the conventional Monte Carlo method, the sample size n often needs to be very large. Therefore, it is important to apply some variance reduction techniques to improve the estimation efficiency. Importance sampling (IS) is intuitively appealing in estimating v_α and c_α especially when α is close to 0, because it may allocate more samples to the left tail of the distribution that are important in estimating v_α and c_α.

Now we introduce the IS estimators of v_α and c_α. Suppose we choose an IS distribution function G for which the probability measure associated to F is absolutely continuous with respect to that associated with G, i.e., $F(dx)=0$ if

$G(dx) = 0$ for any $x \in \Re$. Let $\mathscr{L}(x) = \frac{F(dx)}{G(dx)}$, then $\mathscr{L}$ is called the *likelihood ratio* (LR) function. Then for any $x \in \Re$, we may estimate $F(x)$ by

$$\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} I\{L_i \le x\}\mathscr{L}(L_i). \tag{6}$$

Note that $\hat{F}_n(x)$ is an unbiased and consistent estimator of $F(x)$ as $\tilde{F}_n(x)$ that is defined in Equation (5). Let $\hat{v}_\alpha^n$ and $\hat{c}_\alpha^n$ denote the IS estimators of v_α and c_α. Similar to Equations (3) and (4), we define

$$\begin{aligned} \hat{v}_\alpha^n &= \hat{F}_n^{-1}(\alpha) = \inf\{x : \hat{F}_n(x) \ge \alpha\}, \\ \hat{c}_\alpha^n &= \hat{v}_\alpha^n - \frac{1}{n\alpha}\sum_{i=1}^{n} [\hat{v}_\alpha^n - L_i]^+ \mathscr{L}(L_i). \end{aligned}$$

In the rest of this paper, we analyze the asymptotic properties of $\hat{v}_\alpha^n$ and $\hat{c}_\alpha^n$, and discuss how to select G to reduce the variances of $\hat{v}_\alpha^n$ and $\hat{c}_\alpha^n$. Throughout the rest of the paper, we make the following assumptions.

Assumption 1. *There exists an $\varepsilon > 0$ such that when $x \in (v_\alpha - \varepsilon, v_\alpha + \varepsilon)$ L has a density $f(x) > 0$ and $f(x)$ is differentiable*
.

Assumption 1 requires that L has a density in a neighborhood of v_α. It also implies that $F(v_\alpha) = \alpha$ and $c_\alpha = \mathrm{E}[L|L \ge v_\alpha]$ (Hong and Liu 2009).

Assumption 2. *There exists a constant $C > 0$ such that $\mathscr{L}(x) \le C$ for all $x < v_\alpha + \varepsilon$.*

To estimate v_α and c_α, the samples in the set $\{L < v_\alpha + \varepsilon\}$ are most useful. Then, any reasonable IS distribution should allocate more samples to the region, i.e., $\mathscr{L}(x) \le 1$ for all $x < v_\alpha + \varepsilon$. Therefore, Assumption 2 is a very weak assumption on the IS distribution.

3 ASYMPTOTIC REPRESENTATIONS OF THE IS ESTIMATORS

A complicated estimator can often be represented as the sum of several terms, where the asymptotic behaviors of all the terms are clear. Then, the representation is called an *asymptotic representation* of the estimator. Based on the asymptotic representation of an estimator, many asymptotic properties of the estimator, e.g., consistency and asymptotic normality, can be studied easily.

A famous example is the asymptotic representation of the VaR estimator $\tilde{v}_\alpha^n$ (also known as Bahadur representation of the quantile estimator). By Bahadur (1966), under Assumption 1,

$$\tilde{v}_\alpha^n = v_\alpha + \frac{1}{f(v_\alpha)}\left(\alpha - \frac{1}{n}\sum_{i=1}^{n} I\{L_i \le v_\alpha\}\right) + R_n, \tag{7}$$

where $R_n = O(n^{-3/4}(\log n)^{3/4})$ with probability 1 (w.p.1). The statement $Y_n = O(g(n))$ w.p.1 means that there exist a set Ω_0 with $P(\Omega_0) = 1$ such that for any $\omega \in \Omega_0$ we can find out a constant $C(\omega)$ satisfies $|Y_n(\omega)| \le C(\omega)g(n)$ for all n large enough. Given the representation, many asymptotic properties of $\tilde{v}_\alpha^n$ can be analyzed easily. For instance, we may use it to prove the strong consistency and asymptotic normality of $\tilde{v}_\alpha^n$. By the strong law of large numbers (Durrett 2005), $\frac{1}{n}\sum_{i=1}^{n} I\{L_i \le v_\alpha\} \to F(v_\alpha)$ w.p.1 as $n \to \infty$. Furthermore, because $F(v_\alpha) = \alpha$ by Assumption 1, it is clear that $\tilde{v}_\alpha^n \to v_\alpha$ w.p.1 as $n \to \infty$. Thus, $\tilde{v}_\alpha^n$ is a strongly consistent estimator of v_α. Similarly, by the central limit theorem (Durrett 2005),

$$\sqrt{n}\left(\alpha - \frac{1}{n}\sum_{i=1}^{n} I\{L_i \le v_\alpha\}\right) \Rightarrow \sqrt{\alpha(1-\alpha)}N(0,1)$$

as $n \to \infty$, where "$\Rightarrow$" means "convergence in distribution" and $N(0,1)$ is standard normal random variable. Then,

$$\sqrt{n}(\tilde{v}_\alpha^n - v_\alpha) \Rightarrow \frac{\sqrt{\alpha(1-\alpha)}}{f(v_\alpha)} N(0,1) \tag{8}$$

as $n \to \infty$. Therefore, $\tilde{v}_\alpha^n$ is asymptotically normally distributed.

In the rest of this section, we develop asymptotic representations of the IS estimators $\hat{v}_\alpha^n$ and $\hat{c}_\alpha^n$, and use them to analyze the consistency and asymptotic normality of $\hat{v}_\alpha^n$ and $\hat{c}_\alpha^n$.

3.1 Asymptotic Representation of $\hat{v}_\alpha^n$

We first consider the asymptotic representation of $\hat{v}_\alpha^n$. Note that by a second order Taylor expansion, when $|\hat{v}_\alpha^n - v_\alpha| < \varepsilon$,

$$F(\hat{v}_\alpha^n) - F(v_\alpha) = f(v_\alpha)(\hat{v}_\alpha^n - v_\alpha) - A_{1,n},$$

where $A_{1,n} = O(\hat{v}_\alpha^n - v_\alpha)^2$. Then, we have

$$\hat{v}_\alpha^n = v_\alpha + \frac{F(\hat{v}_\alpha^n) - F(v_\alpha)}{f(v_\alpha)} + \frac{1}{f(v_\alpha)} A_{1,n}. \tag{9}$$

Let $A_{2,n} = F(\hat{v}_\alpha^n) + \hat{F}_n(v_\alpha) - \hat{F}_n(\hat{v}_\alpha^n) - F(v_\alpha)$ and $A_{3,n} = \hat{F}_n(\hat{v}_\alpha^n) - F(v_\alpha)$. It is easy to see that

$$F(\hat{v}_\alpha^n) - F(v_\alpha) = F(v_\alpha) - \hat{F}_n(v_\alpha) + A_{2,n} + A_{3,n}.$$

Then by Equation (9), we have

$$\hat{v}_\alpha = v_\alpha + \frac{F(v_\alpha) - \hat{F}_n(v_\alpha)}{f(v_\alpha)} + \frac{1}{f(v_\alpha)}(A_{1,n} + A_{2,n} + A_{3,n}). \tag{10}$$

In the following lemma, we provide the orders of $A_{1,n}$, $A_{2,n}$ and $A_{3,n}$.

Lemma 3.1. *Suppose that Assumptions 1 and 2 are satisfied. Then for any $\alpha \in (0,1)$, $A_{1,n} = O\left(n^{-1}\log n\right)$, $A_{2,n} = O\left(n^{-\frac{3}{4}}(\log n)^{-\frac{3}{4}}\right)$ and $A_{3,n} = O\left(n^{-1}\right)$ w.p.1.*

Let $\mathscr{L}_i$ denote $\mathscr{L}(L_i)$ for all $i = 1, \ldots, n$. By Lemma 3.1, we can prove the following theorem on the asymptotic expansion of $\hat{v}_\alpha^n$.

Theorem 3.1. *Suppose that Assumptions 1 and 2 are satisfied. Then for any $\alpha \in (0,1)$,*

$$\hat{v}_\alpha^n = v_\alpha + \frac{1}{f(v_\alpha)} \left(\alpha - \frac{1}{n} \sum_{i=1}^n I\{L_i \le v_\alpha\} \mathscr{L}_i \right) + A_n,$$

where $A_n = O\left(n^{-\frac{3}{4}}(\log n)^{-\frac{3}{4}}\right)$ w.p.1.

Proof. By Assumption 1, $F(v_\alpha) = \alpha$. Then by Equation (6),

$$F(v_\alpha) - \hat{F}_n(v_\alpha) = \alpha - \frac{1}{n} \sum_{i=1}^n I\{L_i \le v_\alpha\} \mathscr{L}_i.$$

Let $A_n = \frac{1}{f(v_\alpha)}(A_{1,n} + A_{2,n} + A_{3,n})$. Since $f(v_\alpha) > 0$ by Assumption 1, and by Lemma 3.1, we have $A_n = O\left(n^{-\frac{3}{4}}(\log n)^{-\frac{3}{4}}\right)$ w.p.1. Therefore, the conclusion of the theorem follows directly from Equation (10). □

Because $\tilde{v}_\alpha^n$ is a special case of $\hat{v}_\alpha$ where $\mathscr{L}(x)=1$ for all $x\in\Re$, the Bahadur representation of Equation (7) may be viewed as a special case of Theorem 3.1.

Let E_G and Var_G denote the expectation and variance under the IS distribution G. From Theorem 3.1, it is also straight-forward to prove the following corollary on the strong consistency and asymptotic normality of $\hat{v}_\alpha^n$.

Corollary 3.1. *Suppose that Assumptions 1 and 2 are satisfied. Then for any $\alpha\in(0,1)$, $\hat{v}_\alpha^n\to v_\alpha$ w.p.1 and*

$$\sqrt{n}(\hat{v}_\alpha^n-v_\alpha)\Rightarrow\frac{\sqrt{\mathrm{Var}_G\left[I\{L\le v_\alpha\}\mathscr{L}(L)\right]}}{f(v_\alpha)}N(0,1)$$

as $n\to\infty$.

Remark 3.1. *The conclusions of Corollary 3.1 have also been proved by Glynn (1996) under Assumption 1 and the assumption that $E_G[\mathscr{L}^3(L)]<\infty$.*

Note that

$$\mathrm{Var}_G\left[I\{L\le v_\alpha\}\mathscr{L}(L)\right]=\mathrm{E}_G\left[I\{L\le v_\alpha\}\mathscr{L}^2(L)\right]-\mathrm{E}_G^2\left[I\{L\le v_\alpha\}\mathscr{L}(L)\right]=\mathrm{E}\left[I\{L\le v_\alpha\}\mathscr{L}(L)\right]-\alpha^2. \tag{11}$$

If the IS distribution allocates more samples to the left tail of the distribution of L, i.e., $\mathscr{L}(x)<1$ for all $x\le v_\alpha$, then by Equation (11), $\mathrm{Var}_G\left[I\{L\le v_\alpha\}\mathscr{L}(L)\right]<\alpha(1-\alpha)$. Compared to Equation (8), the IS estimator $\hat{v}_\alpha^n$ has a smaller asymptotic variance than the conventional estimator $\tilde{v}_\alpha^n$. This explains why IS can improve the efficiency of VaR estimation when the IS distribution is selected appropriately.

Theorem 3.1 not only can be applied to obtain the strong consistency and asymptotic normality of $\hat{v}_\alpha^n$, it can also be applied to facilitate another asymptotic analysis related to $\hat{v}_\alpha$. For instance, it helps in developing the asymptotic representation of $\hat{c}_\alpha^n$ as shown in Section 3.2.

3.2 Asymptotic Representation of $\hat{c}_\alpha^n$

We now consider the asymptotic representation of $\hat{c}_\alpha^n$. Note that

$$\begin{aligned}\hat{c}_\alpha^n &= \hat{v}_\alpha^n-\frac{1}{n\alpha}\sum_{i=1}^n(\hat{v}_\alpha^n-L_i)^+\mathscr{L}_i\\ &= v_\alpha-\frac{1}{n\alpha}\sum_{i=1}^n(v_\alpha-L_i)^+\mathscr{L}_i+(\hat{v}_\alpha^n-v_\alpha)-\frac{1}{n\alpha}\sum_{i=1}^n\left[(\hat{v}_\alpha^n-L_i)^+-(v_\alpha-L_i)\right]\mathscr{L}_i. \end{aligned}\tag{12}$$

Furthermore, note that

$$\begin{aligned}&\frac{1}{n\alpha}\sum_{i=1}^n\left[(\hat{v}_\alpha^n-L_i)^+-(v_\alpha-L_i)\right]\mathscr{L}_i=\frac{1}{n\alpha}\sum_{i=1}^n\left[(\hat{v}_\alpha^n-L_i)I\{L_i\le\hat{v}_\alpha^n\}-(v_\alpha-L_i)I\{L_i\le v_\alpha\}\right]\mathscr{L}_i\\ &= \frac{1}{n\alpha}\sum_{i=1}^n\left[(\hat{v}_\alpha^n-v_\alpha)I\{L_i\le\hat{v}_\alpha^n\}\right]\mathscr{L}_i+\frac{1}{n\alpha}\sum_{i=1}^n(v_\alpha-L_i)\left[I\{L_i\le\hat{v}_\alpha^n\}-I\{L_i\le v_\alpha\}\right]\mathscr{L}_i. \end{aligned}\tag{13}$$

Since

$$\frac{1}{n\alpha}\sum_{i=1}^n\left[(\hat{v}_\alpha^n-v_\alpha)I\{L_i\le\hat{v}_\alpha^n\}\right]\mathscr{L}_i=\frac{1}{\alpha}(\hat{v}_\alpha^n-v_\alpha)\hat{F}_n(\hat{v}_\alpha^n)$$

and

$$\left|\frac{1}{n\alpha}\sum_{i=1}^n(v_\alpha-L_i)\left[I\{L_i\le\hat{v}_\alpha^n\}-I\{L_i\le v_\alpha\}\right]\mathscr{L}_i\right|\le\frac{1}{\alpha}\left|\hat{v}_\alpha^n-v_\alpha\right|\left|\hat{F}_n(\hat{v}_\alpha^n)-\hat{F}_n(v_\alpha)\right|,$$

then by Equations (12) and (13),

$$\hat{c}_{\alpha}^{n} = v_{\alpha} - \frac{1}{n\alpha}\sum_{i=1}^{n}(v_{\alpha} - L_i)^{+}\mathscr{L}_i + B_n, \tag{14}$$

where

$$|B_n| \leq \frac{1}{\alpha}\left|\hat{v}_{\alpha}^{n} - v_{\alpha}\right|\left(|\alpha - \hat{F}_n(\hat{v}_{\alpha}^{n})| + |\hat{F}_n(\hat{v}_{\alpha}^{n}) - \hat{F}_n(v_{\alpha})|\right) \leq \frac{1}{\alpha}\left|\hat{v}_{\alpha}^{n} - v_{\alpha}\right|\left(2|\hat{F}_n(\hat{v}_{\alpha}^{n}) - F(v_{\alpha})| + |\hat{F}_n(v_{\alpha}) - F(v_{\alpha})|\right).$$

In the following lemma, we prove the order of B_n.

Lemma 3.2. *Suppose that Assumptions 1 and 2 are satisfied. Then for any $\alpha \in (0,1)$, $B_n = O\left(n^{-1}\log n\right)$ w.p.1.*

Then, we have the following theorem on the asymptotic expansion of $\hat{c}_{\alpha}^{n}$. Note that the conclusion of the theorem follows directly from Equation (14) and Lemma 3.2. Therefore, we omit the proof.

Theorem 3.2. *Suppose that Assumptions 1 and 2 are satisfied. Then for any $\alpha \in (0,1)$,*

$$\hat{c}_{\alpha}^{n} = c_{\alpha} + \left(\frac{1}{n}\sum_{i=1}^{n}\left[v_{\alpha} - \frac{1}{\alpha}(v_{\alpha} - L_i)^{+}\mathscr{L}_i\right] - c_{\alpha}\right) + B_n,$$

where $B_n = O\left(n^{-1}\log n\right)$ w.p.1.

Note that $\mathrm{E}_G\left[v_{\alpha} - \frac{1}{\alpha}(v_{\alpha} - L)^{+}\mathscr{L}\right] = c_{\alpha}$. Then by the strong law of large numbers and the central limit theorem, it is easy to prove the following corollary on the strong consistency and asymptotic normality of $\hat{c}_{\alpha}^{n}$.

Corollary 3.2. *Suppose that Assumptions 1 and 2 are satisfied. Then for any $\alpha \in (0,1)$, $\hat{c}_{\alpha}^{n} \to c_{\alpha}$ w.p.1 and*

$$\sqrt{n}(\hat{c}_{\alpha}^{n} - c_{\alpha}) \Rightarrow \frac{\sqrt{\mathrm{Var}_G\left[(v_{\alpha} - L)^{+}\mathscr{L}(L)\right]}}{\alpha}N(0,1)$$

as $n \to \infty$.

Furthermore, we may set $\mathscr{L}(x) = 1$ for all $x \in \Re$. Then, the conclusions of Theorem 3.2 and Corollary 3.2 apply to $\tilde{c}_{\alpha}$, the conventional Monte Carlo estimator of c_{α}. We summarize the results in the following corollary.

Corollary 3.3. *Suppose that Assumption 1 is satisfied. Then for any $\alpha \in (0,1)$,*

$$\tilde{c}_{\alpha}^{n} = c_{\alpha} + \left(\frac{1}{n}\sum_{i=1}^{n}\left[v_{\alpha} - \frac{1}{\alpha}(v_{\alpha} - L_i)^{+}\right] - c_{\alpha}\right) + C_n$$

where $C_n = O\left(n^{-1}\log n\right)$ w.p.1, $\tilde{c}_{\alpha}^{n} \to c_{\alpha}$ w.p.1, and

$$\sqrt{n}(\tilde{c}_{\alpha}^{n} - c_{\alpha}) \Rightarrow \frac{\sqrt{\mathrm{Var}\left[(v_{\alpha} - L)^{+}\right]}}{\alpha}N(0,1)$$

as $n \to \infty$.

Remark 3.2. *The strong consistency and asymptotic normality of $\tilde{c}_{\alpha}^{n}$ have also been studied by a number of papers in the literature, including Trindade et al. (2007), using different methods.*

If the IS distribution allocates more samples to the left tail of the distribution of L, i.e., $\mathscr{L}(x) < 1$ for all $x \leq v_{\alpha}$, then it is easy to show that $\mathrm{Var}_G\left[(v_{\alpha} - L)^{+}\mathscr{L}(L)\right] < \mathrm{Var}\left[(v_{\alpha} - L)^{+}\right]$. Therefore, by Corollaries 3.2 and 3.3, the IS estimator $\hat{c}_{\alpha}^{n}$ has a smaller asymptotic variance than the conventional estimator $\tilde{c}_{\alpha}^{n}$. This explains why IS can improve the efficiency of CVaR estimation when the IS distribution is selected appropriately.

4 SELECTING IS DISTRIBUTIONS

By Theorems 3.1 and 3.2, it is clear that the variances of $\hat{v}_\alpha^n$ and $\hat{c}_\alpha^n$ are denominated by

$$\mathrm{Var}_G\left[\frac{1}{f(v_\alpha)}\left(\alpha-\frac{1}{n}\sum_{i=1}^n I\{L_i\le v_\alpha\}\mathscr{L}_i\right)\right] \text{ and } \mathrm{Var}_G\left[\frac{1}{n}\sum_{i=1}^n\left[v_\alpha-\frac{1}{\alpha}(v_\alpha-L_i)^+\mathscr{L}_i\right]\right],$$

respectively. Note that

$$\begin{aligned}
&\mathrm{Var}_G\left[\frac{1}{f(v_\alpha)}\left(\alpha-\frac{1}{n}\sum_{i=1}^n I\{L_i\le v_\alpha\}\mathscr{L}_i\right)\right]\\
&= \frac{1}{f^2(v_\alpha)}\mathrm{Var}_G\left[I\{L\le v_\alpha\}\mathscr{L}\right]=\frac{1}{f^2(v_\alpha)}\left\{\mathrm{E}_G\left[I\{L\le v_\alpha\}\mathscr{L}^2\right]-\alpha^2\right\},\\
&\mathrm{Var}_G\left[\frac{1}{n}\sum_{i=1}^n\left(v_\alpha-\frac{1}{\alpha}(v_\alpha-L_i)^+\mathscr{L}_i\right)\right]\\
&= \frac{1}{\alpha^2}\mathrm{Var}_G\left[(v_\alpha-L)^+\mathscr{L}\right]=\frac{1}{\alpha^2}\left\{\mathrm{E}_G\left[(v_\alpha-L)^2I\{L\le v_\alpha\}\mathscr{L}^2\right]-E^2\left[(v_\alpha-L)^+\right]\right\}.
\end{aligned}$$

Therefore, we only need to find IS distributions that reduce $\mathrm{E}_G\left[I\{L\le v_\alpha\}\mathscr{L}^2\right]$ and $\mathrm{E}_G\left[(v_\alpha-L)^2I\{L\le v_\alpha\}\mathscr{L}^2\right]$.

To find a good IS distribution function G, we restrict our considerations to the technique of "exponential twisting", where we set $\mathscr{L}=e^{-\theta L+\phi(\theta)}$ and $\phi(\theta)=\log \mathrm{E}e^{\theta L}$, and our goal is to choose a parameter θ that can reduce $\mathrm{E}_G\left[I\{L\le v_\alpha\}\mathscr{L}^2\right]$ and $\mathrm{E}_G\left[(v_\alpha-L)^2I\{L\le v_\alpha\}\mathscr{L}^2\right]$.

We consider $\theta<0$. Then, $\mathscr{L}<e^{-\theta(v_\alpha+\varepsilon)+\phi(\theta)}$ for $L\in(-\infty,v_\alpha+\varepsilon)$. Therefore, Assumption 2 is satisfied. Note that

$$\mathrm{E}_G\left[I\{L\le v_\alpha\}\mathscr{L}^2\right]=\mathrm{E}\left[I\{L\le v_\alpha\}\mathscr{L}\right]\le \mathrm{E}\left[I\{L\le v_\alpha\}\right]e^{-v_\alpha\theta+\phi(\theta)} \tag{15}$$

$$\mathrm{E}_G\left[(v_\alpha-L)^2I\{L\le v_\alpha\}\mathscr{L}^2\right]=\mathrm{E}\left[(v_\alpha-L)^2I\{L\le v_\alpha\}\mathscr{L}\right]\le \mathrm{E}\left[(v_\alpha-L)^2I\{L\le v_\alpha\}\right]e^{-v_\alpha\theta+\phi(\theta)}. \tag{16}$$

Therefore, we propose to choose θ to minimize the upper bounds in Equations (15) and (16), i.e., minimizing $-v_\alpha\theta+\phi(\theta)$. Note that $\phi(\theta)$ is called the cumulant-generating function and it is convex. Therefore, the optimal θ, denoted as θ^*, satisfies $\phi'(\theta^*)=v_\alpha$. Then, we have the following theorem.

Theorem 4.1. *Suppose that Suppose that Assumption 1 is satisfied and $v_\alpha<\mathrm{E}[L]$. If the IS distribution G satisfies $\mathscr{L}=F(L)/G(L)=e^{-\theta^*L+\phi(\theta^*)}$, then*

$$\begin{aligned}
\mathrm{Var}_G\left[I\{L\le v_\alpha\}\mathscr{L}(L)\right] &< \mathrm{Var}\left[I\{L\le v_\alpha\}\right],\\
\mathrm{Var}_G\left[(v_\alpha-L)^+\mathscr{L}(L)\right] &< \mathrm{Var}\left[(v_\alpha-L)^+\right].
\end{aligned}$$

Note that this θ^* applies to both the IS estimators of VaR and CVaR. Therefore, this allows risk managers to apply the IS distribution and estimate these two risk measures at the same time. By Theorem 4.1, we show that the exponential twisting with θ^* guarantees to reduce the asymptotic variances of both $\hat{v}_\alpha^n$ and $\hat{c}_\alpha^n$.

5 CONCLUSIONS

In this paper, we develop the asymptotic representations for the IS estimators of both VaR and CVaR, and derive the consistency and asymptotic normality for both estimators. We show that there are simple conditions for choosing IS distributions that guarantee to reduce the asymptotic variances of the estimators. Furthermore, we show that the technique of exponential twisting guarantees to find an IS distribution that reduces the asymptotic variances of both estimators at the same time.

REFERENCES

Artzner, P., F. Delbaen, J. -M. Eber and D. Heath. 1999. Coherent measures of risk. *Mathematical Finance* 9(3): 203-228.

Bahadur, R. R. 1966. A note on quantiles in large samples. *Annals of Mathematical Statistics* 37: 577-580.

Delbaen, F. 2001. Coherent risk measures. Lecture Notes, Scuola Normale Superiore di Pisa.

Durret, R. 2005. *Probability: Theory and Examples*, 3rd ed. Duxury Press, Belmont, MA.

Glasserman, P. 2004. *Monte Carlo Methods in Financial Engineering*. Springer, New York.

Glasserman, P., H. Heidelberger and P. Shahabuddin. 2000. Variance reduction techniques for estimating value-at-risk. *Management Science* 46: 1349-1364.

Glasserman, P. and S. Juneja. 2008. Uniformly efficient importance sampling for the tail distribution of sums of random variables. *Mathematics of Operations Research* 33: 35-50.

Glasserman, P. and J. Li. 2005. Importance sampling or portfolio credit risk. *Management Science* 51:1643-1656.

Glynn, P. W. 1996. Importance sampling for Monte Carlo estimation of quantiles. *Proceedings of Second International Workshop on Mathematical Methods in Stochastic Simulation and Experimental Design*. St. Petersburg, FL, 180-185.

Heyde, C. C., S. G. Kou and X. H. Peng. 2007. What is a good external risk measure: Bridging the gaps between robustness, subadditivity, and insurance risk measures. Working paper, Department of Industrial Engineering and Operations research, Columbia University, New York.

Hong, L. J. and G. Liu. 2009. Simulating sensitivities of conditional value at risk. *Management Science* 55: 281-293.

Rockafellar, R. T. and S. Uryasev. 2000. Optimization of conditional value-at-risk. *Journal of Risk* 272-281.

Serfling, R. J. 1980. *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

Trindade, A. A., S. Uryasev, A. Shapiro and G. Zrazhevsky. 2007. Financial prediction with constrained tail risk. *Journal of Banking and Finance* 31: 3524-3538.

AUTHOR BIOGRAPHIES

LIHUA SUN is a Ph.D. student in the Department of Industrial Engineering and Logistics Management at The Hong Kong University of Science and Technology. Her research interests include simulation methodologies and financial engineering. Her e-mail address is <sunlihua@ust.hk>.

L. JEFF HONG is an associate professor in the Department of Industrial Engineering and Logistics Management at The Hong Kong University of Science and Technology. His research interests include Monte Carlo method, financial engineering, stochastic optimization. He is currently an associate editor of *Operations Research*, *Naval Research Logistics* and *ACM Transactions on Modeling and Computer Simulation*.His email address is <hongl@ust.hk>.