Proceedings of the 2024 Winter Simulation Conference
H. Lam, E. Azar, D. Batur, S. Gao, W. Xie, S. R. Hunter, and M. D. Rossetti, eds.

NESTED SIMULATION FOR VALUE-AT-RISK WITH PRECISION TOLERANCE

Xianyu Kuang, Guangwu Liu, and Qianwen Zhu

Department of Management Science, City University of Hong Kong, Kowloon Tong, Hong Kong

ABSTRACT

This paper investigates nested simulation for estimating Value-at-Risk (VaR), a widely adopted risk measure in practice. We formulate a mathematical framework tailored to a practical setting where a risk estimate is measured up to a certain precision level, and any risk estimate within certain range specified by the precision tolerance level is deemed acceptable. Within this framework, we propose a rounded estimator of VaR that explicitly accounts for the pre-specified precision tolerance level, and demonstrate that its error can decay exponentially fast as the sampling budget increases. An important implication of our theoretical result is that a finite inner sample size may suffice for nested simulation in our setting, leading to a budget allocation rule that deviates substantially from the standard nested simulation procedure. Numerical examples confirm the theoretical findings, showcasing the consistent performance of the proposed rounded estimator.

1 INTRODUCTION

In portfolio risk measurement, the loss of a portfolio $L(X)$ is often represented as a real-valued function of certain random risk factors $X \in \mathbb{R}^d$. A risk measure is typically defined as a functional that maps the probability distribution of $L(X)$ to a real number. In practical applications, $L(X)$ takes the form of a conditional expectation $\mathbb{E}[Y|X]$, for which the analytical expression is unknown. Consequently, following the generation of multiple outer-level samples of X, the estimation of $L(X)$ requires the simulation of inner-level samples of Y conditioned on each outcome of X. This two-level simulation procedure is referred to as nested simulation (e.g., Gordy, 2010).

In the literature, the efficacy of a nested simulation procedure is typically characterized by the convergence rate of its mean squared error (MSE) relative to the total simulation budget, denoted as $\Gamma = cmn$ (cf. Gordy and Juneja 2010 and Zhang et al. 2022), where m and n represent the inner- and outer-level sample sizes, respectively, and c is a positive constant representing the unit sampling cost for an inner observation. Notably, Gordy and Juneja (2010) showed that the MSE of the standard nested estimator diminishes at a rate no faster than $\Gamma^{-2/3}$, and this optimal rate can be achieved by setting $n \sim \Gamma^{2/3}$ and $m \sim \Gamma^{1/3}$ for a wide array of risk measures, where "$\sim$" means equivalence in order. To further enhance this rate, leveraging special problem structures is imperative. Broadie et al. (2011) recommended allocating varying inner-level sample sizes to distinct outer-level samples and devised a sequential allocation procedure to estimate the probability of large losses, achieving a rate of $\Gamma^{-4/5+\varepsilon}$ for any positive ε.

Another branch of literature on portfolio risk measurement centers on fitting the loss function $L(X)$ using statistical learning methods. Broadie et al. (2015) employed the least-squares method (LSM) to approximate the loss function and demonstrated that the MSE converges at a rate of Γ^{-1} to a non-diminishing bias, inherent in the LSM method due to the imperfect selection of basis functions. Hong et al. (2017) utilized a non-parametric approach known as the kernel smoothing method to fit the loss function, but it is susceptible to the curse of dimensionality, a common pitfall for most non-parametric methods. Liu and Staum (2010) applied stochastic kriging to estimate the expected shortfall. Furthermore, Zhang et al. (2017) and Feng et al. (2022) develop a likelihood ratio based method and demonstrated that the rate of convergence of its MSE achieves the rate Γ^{-1}, identical to the standard rate of non-nested simulation where explicit expression of $L(\cdot)$ is known and directly computable.

In this paper, our focus is on the widely used risk measure Value-at-Risk (VaR), which represents the quantile of the probability distribution of the portfolio loss at a given confidence level. We consider a practical setting where a risk estimate is measured up to a certain precision level, and any risk estimate within a range specified by the precision tolerance level is deemed acceptable. For instance, risk managers may be concerned about whether VaR is at 1000 USD or 1001 USD, but the decimal values in the estimates may be considered as acceptable errors. In this example, risk estimates within a range specified by a precision tolerance level of 1 are considered indifferent to the risk managers. To address this precision tolerance issue, we develop a mathematical framework by defining a lattice set that represents the rounding of risk estimates. More specifically, we work within a setting where the risk manager is only interested in estimates of VaR that fall into this set. The VaR value that falls into this set is referred to as the target VaR, and our objective is to develop simulation procedures that produce VaR estimates as close as possible to this target VaR.

Tailored to this framework, we propose a rounded estimator for VaR for any user-specified precision tolerance level, which is essentially the sample quantile of the rounded version of the estimate of $L(X)$ based on inner observations. We demonstrate that the error of the proposed rounded estimator can decay exponentially fast as the sampling budget increases, implying that a finite inner sample size may suffice for nested simulation in our setting, leading to a budget allocation rule that deviates substantially from the standard nested simulation procedure. The accelerated convergence rate is achieved mainly due to two factors. Firstly, with the introduction of the precision tolerance level, a finite inner sample size would be sufficient to yield a negligible error in the rounded version of the inner estimate $L_m(X_i)$ for each i. Secondly, for the rounded inner estimates that follow a discrete distribution, the convergence rate can be exponentially fast in the sample size.

Our work is related to the literature on estimating quantiles of discrete probability distributions in a non-nested setting (i.e., the functional $L(\cdot)$ is directly computable), which has attracted the interest of researchers over an extended period (cf. Feldman and Tucker 1966, Wang and Hutson 2011, Ma et al. 2011, Chen and Lazar 2010). Our work distinguishes itself by working with the nested setting and introducing a novel performance measure, the modified mean squared error (MSE), which is essential for guiding budget allocation decisions. Specifically, our findings demonstrate that once the inner-level sample size surpasses a certain threshold, further increases result in negligible improvements. This suggests the adoption of a "one-and-a-half level" simulation, which generates only finite inner samples as the total sample budget tends toward infinity.

The rest of this paper is organized as follows. We formulate our problem in Section 2. Its analysis is provided in Section 3. Numerical experiments are presented in Section 4, followed by conclusions in Section 5.

2 PROBLEM FORMULATION

Suppose a real-valued random variable W, the Value-at-Risk (VaR) of W with a confidence level $\alpha \in (0,1)$ is defined as its α-quantile, i.e.

$$v_\alpha(W) \triangleq \inf\{y \in \mathbb{R} : \mathbb{P}(W \leq y) \geq \alpha\}. \tag{1}$$

In portfolio risk measurement, one is often interested in VaR of the loss of a portfolio over a pre-specified risk horizon τ, which depends on a set of risk factors. In many cases, especially when the portfolio includes derivative assets such as options, the portfolio loss $L(X)$ can often be expressed as a conditional expectation, i.e., $L(X) := \mathbb{E}[Y|X]$, based on classical derivative pricing theory. Here, $X \in \mathbb{R}^d$ denotes the d-dimensional risk factor up to τ, and $Y \in \mathbb{R}$ denotes the loss of the portfolio at a maturity date T discounted back to τ, where $T > \tau$. In this case, we are interested in estimating $v_\alpha(L(X))$ (abbreviated as v_α when no confusion arises), under the general setting that the functional form $L(\cdot)$ is not known explicitly and its estimation requires Monte Carlo simulation.

This problem setting is referred to as nested estimation, highlighting the feature that both the functional form of $L(\cdot)$ and the resulting v_α need to be estimated. A widely used approach to nested estimation is through nested simulation (also referred to as two-level simulation). In a standard two-level simulation procedure, n independent and identically distributed (i.i.d.) observations of X, denoted by $\{X_1,\ldots,X_n\}$ are simulated, referred to as outer-level risk scenarios. Then, for any $i \in \{1,\ldots,n\}$, conditional on each X_i, n i.i.d. observations of Y, denoted by $\{Y_1(X_i),\ldots,Y_m(X_i)\}$, are simulated, referred to as inner-level observations. It is obvious that $L(X_i)$ can be estimated by the sample mean of the inner-level observations, and we denote this estimator by

$$L_m(X_i) = \frac{1}{m}\sum_{j=1}^{m} Y_j(X_i).$$

Suppose that $\{X_{1:n},\ldots,X_{n:n}\}$ is a permutation of $\{X_1,\ldots,X_n\}$ such that

$$L_m(X_{1:n}) \le L_m(X_{2:n}) \le L_m(X_{n:n}),$$

then the standard nested simulation suggests taking

$$\widehat{v}_{n,m} = L_m(X_{\lceil n\alpha\rceil:n}) \tag{2}$$

as the estimator of v_α. Gordy and Juneja (2010) showed that under suitable regulatory conditions, the asymptotic bias and variance of this estimator are of order m^{-1} and n^{-1}, respectively. Their result suggests that in order to minimize the asymptotic mean squared error (MSE), the asymptotic optimal inner and outer sample sizes should be of order $\Gamma^{-1/3}$ and $\Gamma^{-2/3}$, respectively, where $\Gamma = cmn$ denotes the total sample budget with c being the unit sampling cost for an inner-level observation. It is well known that for the standard nested simulation estimator, the resulting convergence rate of MSE is of order $\Gamma^{-2/3}$.

The motivation of this paper is to address the precision tolerance issue in the context of nested simulation, which arises in risk measurement practices. More specifically, a risk metric of practical concern is typically measured up to a certain precision level. For instance, if the scale of the risk metric is in the range of a few thousands and the required precision is $\pm 1\%$, it is unnecessary to measure the decimal values accurately. In addition to practical considerations, the inherent presence of model errors in risk models may also render a metric measured with very high precisions meaningless and/or unnecessary. In what follows, we develop a mathematical framework to address this precision tolerance issue.

Let $\Delta > 0$ be a user-specified precision tolerance level, expressed in absolute terms rather than as a percentage, because specifying the precision tolerance as a percentage can be challenging since the true value of the risk metric is unknown. By specifying a precision tolerance level Δ, the risk manager considers estimates with an error within $\Delta/2$ to be acceptable. For example, when $\Delta = 1$ and the true (but unknown) risk metric is 1000.2, any resulting risk estimates within the range [999.7, 1000.7] are deemed satisfactory. To establish a mathematical framework, we define a lattice set $\mathbb{Z}^\Delta = \{\Delta i; i \in \mathbb{Z}\}$ that represents the rounding of risk estimates, and work within a setting where the risk manager is only interested in risk estimates that fall into this set. Referring to the aforementioned example, the target risk estimate under this setting is 1000. Hence, our objective is to develop simulation procedures that produce a risk estimate as close as possible to this target value.

To put forward a formal mathematical framework, we define

$$Q_\alpha^\Delta = \arg\min_{y\in\mathbb{Z}^\Delta} |y - v_\alpha|.$$

Note that in general, Q_α^Δ is a set with possibly one or two elements. In particular, if $v_\alpha \neq (k+1/2)\Delta$ for any $k \in \mathbb{Z}$, this set has only one element, which is our target VaR. If $v_\alpha = (k+1/2)\Delta$ for some $k \in \mathbb{Z}$, $Q_\alpha^\Delta = \{k\Delta, (k+1)\Delta\}$, values within which are indifferent to the risk manager, as both values in this set deviate from v_α by an amount no more than $\Delta/2$. We thus refer to Q_α^Δ as the *indifference set*. Within this

framework, the target VaR of interest belong to the indifference set, a slight modification is needed on performance measures of an estimator. More specifically, we define the modified MSE of any estimator $\widehat{v}$ using

$$\mathrm{MSE}^{\Delta}(\widehat{v}) = \mathbb{E}\left[\mathrm{dist}^2(\widehat{v}, Q_\alpha^{\Delta})\right],$$

where the distance between a real value a and a set A is given by $\mathrm{dist}(a,A) = \inf_{b\in A}|a-b|$, for any $a \in \mathbb{R}$ and $A \subset \mathbb{R}$.

With a pre-specified precision tolerance level Δ, our goal is to study estimators for the target VaR of interest, rendering its modified MSE as small as possible. To further simplify notation, we define the rounding operator for a real value $y \in \mathbb{R}$ as

$$[y]_\Delta := k\Delta, \text{ if } y \in I_k \triangleq [(k-1/2)\Delta, (k+1/2)\Delta], \text{ for some } k \in \mathbb{Z}.$$

Intuitively, a straightforward estimator of α-VaR is $[\widehat{v}_{n,m}]_\Delta$, where $\widehat{v}_{n,m}$ is given by Equation (2). To analyze this estimator, we note that it is equivalent to the order statistic of the rounded version of $L(X)$, which follows a discrete probability distribution. Let $L^{\Delta}(X) = [L(X)]_\Delta$ and $L_m^{\Delta}(X_i) = [L_m(X_i)]_\Delta$. Note that after rounding, the order of $\{L_m(X_{1:n}), \ldots, L_m(X_{n:n})\}$ is preserved, i.e. $L_m^{\Delta}(X_{1:n}) \le L_m^{\Delta}(X_{2:n}) \le L_m^{\Delta}(X_{n:n})$. Then the order statistic of $\{L_m^{\Delta}(X_{1:n}), \ldots, L_m^{\Delta}(X_{n:n})\}$ is given by $L_m^{\Delta}(X_{\lceil n\alpha\rceil:n})$. It can be easily verified that the VaR estimator $[\widehat{v}_{n,m}]_\Delta$ is equal to the rounded version of $L_m(X_{\lceil n\alpha\rceil:n})$, i.e.,

$$[\widehat{v}_{n,m}]_\Delta = L_m^{\Delta}(X_{\lceil n\alpha\rceil:n}).$$

We thus refer to this estimator as a *rounded estimator* with a precision tolerance level Δ, and write it as $\widehat{v}_{n,m}^{\Delta} = [\widehat{v}_{n,m}]_\Delta$ for easy of presentation in the analysis.

3 THEORETICAL RESULT

To facilitate analysis for the rounded estimator with a precision tolerance level Δ, we make the following regularity assumptions.

Assumption 1 There exists a positive constant $\sigma_u > 0$ such that $\mathrm{Var}(Y|X) \le \sigma_u^2$, w.p.1. The density function of $L(X)$, denoted by $f_L(y)$, exists, and is continuous and positive at $y = v_\alpha$.

Assumption 2

1. The density function $g_m(y,z)$ of $(L(X), \sqrt{m}(L_m(X) - L(X)))$ and its partial derivatives
$$\frac{\partial}{\partial y} g_m(y,z) \text{ and } \frac{\partial^2}{\partial y^2} g_m(y,z)$$
exist for any $m \ge 1$ and any (y,z).
2. For any $m \ge 1$, there exist nonnegative functions $p_{0,m}(z)$, $p_{1,m}(z)$ and $p_{2,m}(z)$ such that
$$\sup_y g_m(y,z) \le p_{0,m}(z), \quad \sup_y \frac{\partial}{\partial y} g_m(y,z) \le p_{1,m}(z), \quad \sup_y \frac{\partial^2}{\partial y^2} g_m(y,z) \le p_{2,m}(z),$$
for any z. In addition $\sup_m \int_{-\infty}^{\infty} |z|^r p_{i,m}(z) dz < \infty$ for $i = 0,1,2$ and $0 \le r \le 4$.

Assumption 1 imposes a requirement on the bounded conditional variance of Y given X, a condition typically employed in the nested simulation literature. Assumption 2 consists of a set of regularity conditions on the smoothness of the density function of $(L(X), \sqrt{m}(L_m(X) - L(X)))$, and the same assumption has been made in the literature, e.g., Gordy and Juneja (2010).

To simplify the presentation, we introduce further notations. Denote the cumulative distribution function (CDF) of $L(X)$ by $F_L(y)$, and the CDF and the probability density function (PDF) of $L_m(X)$ by $F_m(y)$ and $f_m(y)$, respectively. Define $\bar{F}(L(t)) = \mathbb{P}(L(X) \ge t)$ and $v_{\alpha m} = v_\alpha(L_m(X))$. When Assumption 2 holds, Lemma 1 of Gordy and Juneja (2010) showed that as $m \to \infty$, $f_m(v_{\alpha m})$ converges to $f_L(v_\alpha)$, which is positive as implied by Assumption 1. Here, to facilitate the analysis, we further make the following assumption that imposes bounds on $f_m(v_{\alpha m})$ uniformly over m.

Assumption 3 There exists positive constants p_u and p_l, such that for any $m \geq 1$, $p_l \leq f_m(v_{\alpha m}) \leq p_u$.

With these assumptions, our main theoretical result is summarized in the following theorem.

Theorem 1 Suppose that Assumptions 1-3 hold. Then, there exist an integer $m_0 \in \mathbb{N}_+$ and a constant $C_m > 0$ that may depend on m, such that for any $m \geq m_0$ and $C \in (0, C_m)$,

$$\mathrm{MSE}^{\Delta}(\widehat{v}^{\Delta}_{n,m}) = O\left(\exp\{-Cn\}\right), \tag{3}$$

where the notation $b_n = O(a_n)$ means that $\limsup_n b_n/a_n \leq c$ for some positive constant c.

Theorem 1 shows that the modified MSE decays to zero at an exponential rate in the outer sample size when the inner sample size is sufficiently large (but finite). This suggests that with a precision tolerance level, the proposed rounded estimator converges at a much faster rate compared to the standard nested simulation estimator. The accelerated convergence rate is achieved mainly due to two factors. Firstly, with the introduction of the precision tolerance level, a finite inner sample size would be sufficient to yield a negligible error in the rounded version of the inner estimate $L_m(X_i)$ for each i. Secondly, for the rounded inner estimates that follow a discrete distribution, the convergence rate can be exponentially fast in the sample size.

In an ideal case when the threshold value m_0 is known, one can simply set $n = \Gamma/m_0$, where Γ is the total sampling budget. In this scenario, the modified MSE decays exponentially fast to zero as Γ goes to infinity. However, it should be noted that in practical applications, m_0 is usually unknown. In such cases, we are interested in estimating m_0 using appropriate heuristic procedures, which shall be further discussed in Section 4.

It is important to note that a large m does not necessarily imply a faster convergence rate, as C_m is not necessarily increasing in m. Our preliminary numerical study suggests that for m's that are greater than m_0, it is possible that a larger m may lead to a slower rate of convergence.

In the remainder of this section, we provide a sketch of the proof for Theorem 1, which explains why a finite inner sample size may lead to an exponential rate of convergence in the rounded estimator with a precision tolerance level, while a more complete and detailed proof will be offered in the appendix.

Proof Sketch

Without loss of generality, throughout the proof we assume that $v_\alpha \in [-\Delta/2, \Delta/2)$. Denote $v^{\Delta}_{\alpha m} = v_\alpha(L^{\Delta}_m(X))$. A critical observation is that for $k \in \mathbb{Z}$, $v^{\Delta}_{\alpha m} > k\Delta$ if $\mathbb{P}\left(L^{\Delta}_m(X) \leq k\Delta\right) < \alpha$ and $v^{\Delta}_{\alpha m} < k\Delta$ if $\mathbb{P}\left(L^{\Delta}_m(X) \geq k\Delta\right) < 1-\alpha$.

Suppose that we have $k^{\mathrm{L}}_m, k^{\mathrm{U}}_m \in \mathbb{Z}$ (to be specified momentarily in the case-by-case analysis below) such that $v^{\Delta}_{\alpha m} \in [k^{\mathrm{L}}_m\Delta, k^{\mathrm{U}}_m\Delta]$. Define

$$b_m = \max\{\mathrm{dist}\left(k^{\mathrm{L}}_m\Delta, Q^{\Delta}_\alpha\right), \mathrm{dist}\left(k^{\mathrm{U}}_m\Delta, Q^{\Delta}_\alpha\right)\}.$$

Then, the modified MSE of $\widehat{v}^{\Delta}_{n,m}$ can be bounded as follows:

$$\begin{aligned}\mathrm{MSE}^{\Delta}(\widehat{v}^{\Delta}_{n,m}) \leq b_m^2 &+ \mathbb{E}\left[\mathrm{dist}^2\left(L^{\Delta}_m(X_{\lceil n\alpha\rceil;n}), Q^{\Delta}_\alpha\right)\mathbb{I}_{\{L^{\Delta}_m(X_{\lceil n\alpha\rceil;n}) \leq k^{\mathrm{L}}_m\Delta - \Delta\}}\right] \\ &+ \mathbb{E}\left[\mathrm{dist}^2\left(L^{\Delta}_m(X_{\lceil n\alpha\rceil;n}), Q^{\Delta}_\alpha\right)\mathbb{I}_{\{L^{\Delta}_m(X_{\lceil n\alpha\rceil;n}) \geq k^{\mathrm{U}}_m\Delta + \Delta\}}\right].\end{aligned} \tag{4}$$

The key to establishing the result in Theorem 1 includes two parts. The first part is that $b_m = 0$ for $m \geq m_0$, while the second part is that the second and third terms on the right-hand-side of Equation (4) converge to zero exponentially fast. The intuition of the latter follows from a fact that both $\mathbb{P}\left(L^{\Delta}_m(X_{\lceil n\alpha\rceil;n}) \geq k^{\mathrm{U}}_m\Delta + k\Delta\right)$ and $\mathbb{P}\left(L^{\Delta}_m(X_{\lceil n\alpha\rceil;n}) \leq k^{\mathrm{L}}_m\Delta - k\Delta\right)$ decay to zero exponentially fast in n for any $k \geq 1$, which shall be shown in detail in the proof in the appendix. The main purpose in this sketch is to provide a proof on the first part, by considering two separate cases.

Case 1: $v_\alpha \in (-\Delta/2, \Delta/2)$, implying that $Q_\alpha^\Delta = \{0\}$. In this case, there exist t_l and t_u such that $-\Delta/2 < t_l < v_\alpha < t_u < \Delta/2$. Then $F_L(t_l) < \alpha$ and $\bar{F}_L(t_u) < 1-\alpha$ by Assumption 1. Hence,

$$\begin{aligned}\mathbb{P}\left(L_m^\Delta(X) \le -\Delta\right) &= \mathbb{P}\left(L_m(X) \le -\Delta/2\right) = \mathbb{P}\left(L_m(X) \le -\Delta/2, L(X) \le t_l\right) + \mathbb{P}\left(L_m(X) \le -\Delta/2, L(X) \ge t_l\right) \\ &\le F_L(t_l) + \mathbb{P}\left(L_m(X) - L(X) \ge -\Delta/2 - t_l\right) \le F_L(t_l) + \frac{\sigma_u^2}{m(\Delta/2+t_l)^2},\end{aligned} \tag{5}$$

where the last inequality follows from Chebyshev's inequality.

Similarly, we have

$$\mathbb{P}\left(L_m^\Delta(X) \ge \Delta\right) \le \bar{F}_L(t_u) + \frac{\sigma_u^2}{m(\Delta/2-t_u)^2}. \tag{6}$$

Let

$$m_0 = \left(\frac{\sigma_u^2}{(\alpha - F_L(t_l))\,(0.5\Delta + t_l)^2}\right) \bigvee \left(\frac{\sigma_u^2}{(F_L(t_u) - \alpha)\,(0.5\Delta - t_u)^2}\right) + 1, \tag{7}$$

where $a \vee b = \max(a,b)$. Then, it can be observed that for any $m \ge m_0$, $\mathbb{P}\left(L_m^\Delta(X) \le -\Delta\right) < \alpha$ and $\mathbb{P}\left(L_m^\Delta(X) \ge \Delta\right) < 1-\alpha$, which implies that $k_m^{\mathrm{L}} = k_m^{\mathrm{U}} = 0$. By definition, it can be verified that $b_m = 0$ for any $m \ge m_0$.

Case 2: $v_\alpha = -\Delta/2$, implying $Q_\alpha^\Delta = \{-\Delta, 0\}$. Similar to Case 1, choose any t_l and t_u with $-3\Delta/2 < t_l < -\Delta/2 < t_u < \Delta/2$. By Assumption 1, we have $F_L(t_l) < \alpha$ and $\bar{F}_L(t_u) < 1-\alpha$, and

$$\mathbb{P}\left(L_m^\Delta(X) \le -2\Delta\right) \le \bar{F}_L(t_l) + \frac{\sigma_u^2}{m(3\Delta/2+t_l)^2}, \quad \mathbb{P}\left(L_m^\Delta(X) \ge \Delta\right) \le \bar{F}_L(t_u) + \frac{\sigma_u^2}{m(\Delta/2-t_u)^2}.$$

By letting

$$m_0 = \left(-\frac{\sigma_u^2}{(\alpha - F_L(t_l))\,(3\Delta/2 + t_l)^2}\right) \bigvee \left(-\frac{\sigma_u^2}{(F_L(t_u) - \alpha)\,(\Delta/2 - t_u)^2}\right) + 1,$$

we have $k_m^{\mathrm{L}} = -1$ and $k_m^{\mathrm{U}} = 0$, leading to $b_m = 0$ for any $m \ge m_0$.

4 NUMERICAL EXPERIMENTS

We consider two examples to examine the performance of the proposed rounded estimator. First, we start with a simple example where the marginal distribution of $L(X)$ and the conditional distribution of Y are normally distributed. For this case, an explicit formula for the optimal value of m_0 can be derived, based on which we suggest a heuristic procedure to determine the value of m_0 using pilot samples during practical implementation.

Furthermore, we explore a more realistic example related to portfolio risk measurement. In this case, the optimal value of m_0 is not known explicitly. We employ the heuristic procedure to set m_0, as proposed for the first example, and evaluate the performance of the rounded estimator accordingly.

4.1 A Simple Example

Suppose $L(X)$ follows a normal distribution with mean 0 and variance σ_1^2, i.e., $L(X) \sim N(0, \sigma_1^2)$, and $Y(X) = L(X) + Z$, where $Z \sim N(0, \sigma_2^2)$, and X and Z are independent. In this case, it can be easily verified that $L_m(X) = \frac{1}{m}\sum_{i=1}^m Y_i(X) \sim N(0, \sigma_3^2)$, where $\sigma_3^2 = \sigma_1^2 + \frac{1}{m}\sigma_2^2$. Then $v_\alpha = \sigma_1 z_\alpha$ and $v_{\alpha m} = \sigma_3 z_\alpha$, where z_α is the α-VaR for a standard normal distribution.

Suppose $v_\alpha \in I_p = ((p-1/2)\Delta, (p+1/2)\Delta]$ for a certain $p \in \mathbb{Z}$. To ensure $v_{\alpha m}^\Delta = v_\alpha^\Delta = p\Delta$, it is expected that $v_{\alpha m}$ falls in I_p, which is equivalent to

$$m \ge m_0 = \inf\{m \in \mathbb{N} : v_{\alpha m} \in I_p\} = \left\lceil \frac{\sigma_2^2 z_\alpha^2}{(p+1/2)^2\Delta^2 - \sigma_1^2 z_\alpha^2} \right\rceil. \tag{8}$$

Despite the fact that the right-hand-side of Equation (8) involves unknown parameters σ_1^2, σ_2^2 and p, it provides a heuristic guideline on how to estimate these parameters and thus m_0. In particular, with generated data $\{X_i, Y_j(X_i), i = 1,\ldots,n, j = 1,\ldots,m\}$, σ_2^2 and σ_3^2 can be estimated by $\hat{\sigma}_2^2 = \frac{1}{n(m-1)}\sum_{i=1}^{n}\sum_{j=1}^{m}[Y_j(X_i) - L_m(X_i)]^2$, $\hat{\sigma}_3^2 = \frac{1}{n-1}\sum_{i=1}^{n}[L_m(X_i) - \bar{L}_m(X)]^2$, respectively, where $\bar{L}_m(X) = \frac{1}{n}\sum_{i=1}^{n} L_m(X_i)$. Then, an estimator of σ_1^2 is given by $\hat{\sigma}_1^2 = \hat{\sigma}_3^2 - \frac{1}{m}\hat{\sigma}_2^2$. Furthermore, v_α and p can be estimated by $\hat{v}_\alpha = L_m(X_{\lceil n\alpha\rceil,n})$ and $\hat{p} = [\frac{\hat{v}_\alpha}{\Delta}]$, respectively, which then naturally leads to an estimator of m_0, denoted by $\hat{m}_0$.

Based on the formula for $\hat{m}_0$ and the requirement that $m \geq m_0$, we propose the following heuristic procedure to determine the inner sample size m. The heuristic procedure works as follows. It first sets $m' = (\frac{\Gamma}{10})^{1/3}$ and $n' = (\frac{\Gamma}{10})^{2/3}$ following the asymptotic optimal inner and outer sample size results in Gordy and Juneja (2010), and then generates data $\{X_i, Y_j(X_i), i = 1,\ldots,n', j = 1,\ldots,m'\}$ to obtain $\hat{m}_0$. To avoid using too small m_0 that is often subject to estimation error, we suggest setting $m = 2\hat{m}_0$, which implies a budget allocation rule with $m = 2\hat{m}_0$ and $n = \lceil\frac{\Gamma}{m}\rceil$.

Considering the reuse of data, the budget allocation rule can be more precisely described from an implementation perspective as follows. If $2\hat{m}_0 > m'$ and $\left(\Gamma - \frac{\Gamma}{10}\right) \geq n'(2\hat{m}_0 - m')$, then $m = 2\hat{m}_0$ and $n = \lceil\frac{\Gamma}{m}\rceil$, which indicates that each of the existing n' scenarios requires additional $(m - m')$ inner samples, and $(n - n')$ new scenarios are generated, each with m inner samples. If $2\hat{m}_0 > m'$ and $\left(\Gamma - \frac{\Gamma}{10}\right) < n'(2\hat{m}_0 - m')$, then $n = n', m = \lceil\frac{\Gamma}{n}\rceil (< 2\hat{m}_0)$, that is, each of the existing n' scenarios generates additional $m - m'$ inner samples, and no new scenario is simulated. If $2\hat{m}_0 \leq m'$, then $m = m' (\geq 2\hat{m}_0)$ and $n = \lceil\frac{\Gamma}{m}\rceil$, which means that $(n - n')$ new scenarios are generated, each with m' inner samples. Through this implementation, we increase the inner sample size to $2\hat{m}_0$ as much as possible, based on the previously generated data, while ensuring that each scenario corresponds to the same size of inner samples.

For the rounded estimator, an appropriate accuracy level Δ needs to be specified in advance. Generally, the order of magnitude of the true value can be inferred from historical records or other relevant sources, after which the accuracy level should be set appropriately. In our numerical experiments, Δ is set between 1% and 5% of the true value. To evaluate the performance of our procedure, we compute the true value of v_α as a benchmark, which is then used to measure the modified MSE based on $1,000$ macro replications.

In this example, the parameters are set as $\sigma_1 = 1$, $\sigma_2 = 1$, and $\alpha = 0.95$. The true value can be directly calculated using $v_\alpha = \sigma_1 z_\alpha$. The precision tolerance level is set to be $\Delta = 0.05$. In the experiments, we normalize the unit cost of sampling an inner observation of Y as 1. The modified MSEs for different sampling budgets are then summarized in Figure 1, with respect to different inner sample sizes m. In particular, inner sample sizes $m = 10$ and $m = 100$ are set as benchmarks. Moreover, explicit formula in Equation (8) suggests $m_0 = 28$, and thus $m = 2\hat{m}_0 = 56$ is considered an another benchmark. For each replication, our procedure employs the aforementioned heuristic method in determining the value of m, yielding an average value of m approximately equal to 118 across $1,000$ macro replications.

From Figure 1, it can be seen that when m is sufficiently large, say, greater than or equal to 56, the modified MSE decays rapidly as the sampling budget is sufficiently large (i.e., close to 10^7), which is consistent with the theoretical result in Theorem 1. However, when m is too small, e.g., $m = 10$, the rate of convergence of the modified MSE may be much slower. From this figure, it can also be observed that the performance of the rounded estimator with the heuristic procedure is comparable to the estimator with the known value of $m = 56$, suggesting the heuristic procedure in determining m works quite well. Moreover, when the sample size is large enough, the modified MSE vanishes to 0 because the probability of the estimator not belonging to the indifference set is extremely close to 0 and we are unable to observe such cases in a limited number of replications.

4.2 Portfolio Risk Measurement

We consider a more realistic risk measurement example for a portfolio that is comprised of options written on several underlying assets. More specifically, let $X(t) = (X_1(t),\ldots,X_d(t))^\top$ denote the price dynamics of d underlying assets, which are governed by a multidimensional geometric Brownian motion, characterized

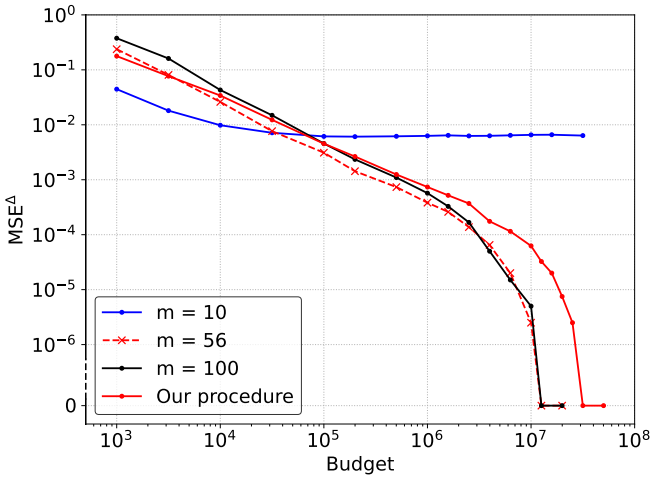

Figure 1: MSE^{Δ} for normal example.

by drifts μ_i', volatility σ_i and correlations ρ_{ij}, $i,j = 1,\ldots,d$. In other words,

$$\frac{\mathrm{d}X_i(t)}{X_i(t)} = \mu_i' \mathrm{d}t + \sum_{j=1}^{i} A_{ij} \mathrm{d}B_j(t), \quad i = 1,\ldots,d,$$

where $\{B_j(t), j = 1,\ldots,d\}$ are independent standard Brownian motions, $\Sigma = (\Sigma_{ij})$ with $\Sigma_{ij} = \sigma_i \sigma_j \rho_{ij}$ and the lower triangle matrix A satisfies $\Sigma = AA^{\top}$.

Using Itô's formula, the price of the ith underlying asset at time t can be written as

$$X_i(t) = X_i(0) \exp\left\{ \left(\mu_i' - \frac{1}{2}\sigma_i^2 \right) t + \sum_{j=1}^{i} A_{ij} B_j(t) \right\}, \quad i = 1,\ldots,d.$$

Note that μ' is chosen to be returns of the underlying asset μ under the real-world probability measure during the time horizon $[0,\tau]$, and the risk-free interest rate r under the risk neutral probability measure during $(\tau,T]$, with τ and T denoting the risk horizon and the maturity of the options, respectively. For simplicity, we assume that all underlying assets have the same initial price, drift, and volatility, and the correlation between any two underlying assest is the same. Hence we can omit subscripts in these notations. The parameters are set as $d = 4$, $X(0) = 100$, $\mu = 0.08$, $r = 0.05$, $\sigma = 15\%$, and $\rho = 0.3$.

We assume that all the options have a common maturity date T. The portfolio consists of twenty European call options, with five options written on each of the underlying assets with strike prices $K_1 = 90$, $K_2 = 95$, $K_3 = 100$, $K_4 = 105$, and $K_5 = 110$. We denote the current time by 0 and are interested in measuring the portfolio risk at a given risk horizon $\tau < T$. We set $T = 1/12$ year, that is, 1 month, and $\tau = 1/52$ year, that is, 1 week. The current value of the portfolio, denoted by $V(0)$, can be calculated explicitly using the Black-Scholes formula. The value of the portfolio at time τ, referred to as $V(\tau)$, can be written as

$$V(\tau) = \mathbb{E}\left[\sum_{k=1}^{4} \sum_{l=1}^{5} \mathrm{e}^{-r(T-\tau)} \left(X_k(T) - K_l \right)^+ \middle| X(\tau) \right].$$

Then the portfolio loss at time τ can be represented as

$$L(X(\tau)) = \mathbb{E}\left[V(0) - \sum_{k=1}^{4} \sum_{l=1}^{5} \mathrm{e}^{-r(T-\tau)} \left(X_k(T) - K_l \right)^+ \middle| X(\tau) \right].$$

To obtain the true value of the α-VaR of $L(X(\tau))$ as a benchmark, we generate a very large amount (10^8) observations of $X(\tau)$ and evaluate the corresponding values of $L(X(\tau))$ using the explicit Black-Schole formula, and then estimate α-VaR with very high accuracy. This benchmark value is then used to estimate the modified MSE of the proposed rounded estimator. In our experiments, $\alpha = 95\%$, $\Delta = 1$ and $v_\alpha = 22.627$.

The modified MSEs with respect to different sampling budgets are plotted in Figure 2. We have similar observations as in the simple example, implying that the heuristic procedure in determining m works reasonably well despite the fact that both the outer-level scenarios and the inner-level observations do not follow normal distributions. When m is small, e.g., $m = 50$, the rate of convergence of the modified MSE may be relatively slow. When m is sufficiently large, e.g., $m = 100$, the modified MSE decays at a much faster speed as the sampling budget is sufficiently large, which is consistent with the theoretical result in Theorem 1. In addition, the modified MSE also decays to zero after a large sample size as seen in the simple example.

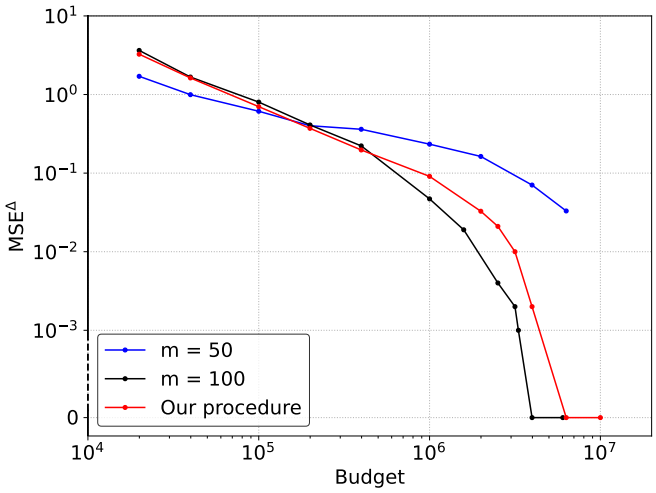

Figure 2: MSE^Δ for portfolio risk measurement example.

5 CONCLUSIONS

In this paper, we have formulated a mathematical framework that considers a pre-specified precision tolerance level, and proposed a rounded estimator of VaR that explicitly accounts for this precision tolerance level. Our theoretical analysis has demonstrated that the error of the proposed rounded estimator can decay exponentially fast as the sampling budget increases. Notably, our results suggest that a finite inner sample size may be sufficient for nested simulation within this practical setting, leading to a budget allocation rule that differs considerably from the standard nested simulation procedure. The numerical examples provided have validated our theoretical findings.

ACKNOWLEDGMENTS

The research of the second author is supported partially by the National Natural Science Foundation of China and the Research Grants Council of Hong Kong (RGC-HK), under the RGC-HK General Research Fund Project 11508620, and NSFC/RGC-HK Joint Research Scheme under project N_CityU 105/21.

A PROOF OF THEOREM 1

In this section, we shall show that the second and third terms in Equation (4) are of order $\exp\{-Cn\}$ for some $C \in (0, C_m)$. To this end, we will use the following lemma, which is the extension of the powerful Bahadur theorem to the nested simulation setting.

Lemma 1 Under Assumption 2, suppose s_n is a sequence of integer with

$$\frac{s_n}{n} = \alpha + o_n\left(\frac{(\log n)^q}{\sqrt{n}}\right),$$

for some $q \geq \frac{1}{2}$. Then

$$L_m(X_{s_n:n}) = v_{\alpha m} + \frac{1}{n}\sum_{i=1}^{n} Z_m(i) + \frac{s_n/n - \alpha}{f_m(v_{\alpha m})} + A_{nm}, \tag{9}$$

where $Z_m(i) = \dfrac{\alpha - \mathbb{I}_{\{L_m(X_i) \leq v_{\alpha m}\}}}{f_m(v_{\alpha m})}$ and $|A_{nm}| \leq a_n := c_0 \cdot n^{-3/4}(\log n)^{(1+q)/2}$ for some constant c_0 w.p.1.

The proof of Lemma 1 is skipped due to page limit. Essentially, it follow from the integration of Lemma 1 in Gordy and Juneja (2010), the standard proof techniques for the classic Bahadur representation (see, e.g., Theorem 2.5.2 in Serfling 2009) for each $m \geq 1$, and our construction of a bound on the residuals A_{nm} uniformly over m.

In addition to Lemma 1, we also need the following elementary result.

Lemma 2 Suppose X is a non-negative random variable with $\bar{F}(x) = \mathbb{P}(X \geq x)$. Then for any $t \geq 0$,

$$\mathbb{E}\left[X\mathbb{I}_{\{X \geq t\}}\right] = \int_t^\infty \bar{F}(x)dx + t\bar{F}(t).$$

Particularly if X takes value on $\mathbb{N}$, then for $t \in \mathbb{N}$,

$$\mathbb{E}\left[X^2\mathbb{I}_{\{X \geq t\}}\right] = \sum_{i=t}^{\infty}((i+1)^2 - i^2)\mathbb{P}(X \geq i) + t^2\bar{F}(t) = \sum_{i=t}^{\infty}(2i+1)\mathbb{P}(X \geq i) + t^2\bar{F}(t)$$

Proof.

$$\mathbb{E}\left[X\mathbb{I}_{\{X \geq t\}}\right] = -\int_t^\infty x d\bar{F}(x) = \int_t^\infty \bar{F}(x)dx - x\bar{F}(x)\Big|_t^\infty = \int_t^\infty \bar{F}(x)dx + t\bar{F}(t).$$

□

Now we proceed to the proof of Theorem 1. As suggested in the proof sketch provided in the main text, throughput the proof we assume that $v_\alpha \in [-\Delta/2, \Delta/2)$. The analysis proceeds separately for the cases $v_\alpha \in (-\Delta/2, \Delta/2)$ and $v_\alpha = -\Delta/2$, corresponding to $Q_\alpha^\Delta = \{0\}$ and $Q_\alpha^\Delta = \{-\Delta, 0\}$, respectively. We mainly present the proof for Case 1, while that for Case 2 follows in a similar manner and is thus omitted.

Recall that the analysis in Section 3 shows $b_m = 0$ for any $m \geq m_0$. Then, for any such m, Equation ((4)) becomes

$$\mathrm{MSE}^\Delta(\widehat{v}_{n,m}^\Delta) \leq \mathbb{E}\left[\left(L_m^\Delta(X_{\lceil n\alpha\rceil;n})\right)^2 \mathbb{I}_{\{L_m^\Delta(X_{\lceil n\alpha\rceil;n}) \leq -\Delta\}}\right] + \mathbb{E}\left[\left(L_m^\Delta(X_{\lceil n\alpha\rceil;n})\right)^2 \mathbb{I}_{\{L_m^\Delta(X_{\lceil n\alpha\rceil;n}) \geq \Delta\}}\right]. \tag{10}$$

Note that in addition to $v_{\alpha m}^\Delta = 0$, it holds that $v_{\alpha m} \in (-\Delta/2, \Delta/2)$, since $\mathbb{P}(L_m(X) \leq -\Delta) < \alpha$ and $\mathbb{P}(L_m(X) \geq \Delta) < 1 - \alpha$. Denote $d_m = \min\{v_{\alpha m} + \Delta/2, v_{\alpha m} - \Delta/2\} > 0$. By Lemma 1, for $j \geq 1$ and

sufficiently large n such that $a_n < d_m$,

$$\mathbb{P}\left(L_m^{\Delta}(X_{\lceil n\alpha\rceil;n}) \le -j\Delta\right) = \mathbb{P}\left(L_m(X_{\lceil n\alpha\rceil;n}) \le -(j-1/2)\Delta\right) = \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} Z_m(i) \le -(j-1/2)\Delta - v_{\alpha m} + a_n\right)$$
$$\le \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} Z_m(i) \le -(j-1)\Delta - d_m + a_n\right) \le \exp\left\{-2np_l^2((j-1)\Delta + d_m - a_n)^2\right\}, \tag{11}$$

where Z_m and a_n were defined in Lemma 1, p_l is a lower bound for $f_m(v_{\alpha m})$ by Assumption 3, and the last inequality follows from Hoeffding's inequality.

Similarly, it can be shown that

$$\mathbb{P}\left(L_m^{\Delta}(X_{\lceil n\alpha\rceil;n}) \ge j\Delta\right) \le \exp\left\{-2np_l^2((j-1)\Delta + d_m - a_n)^2\right\}.$$

For notational ease, denote $e_n(j) = \exp\left\{-2np_l^2((j-1)\Delta + d_m - a_n)^2\right\}$, we need the following lemma that characterize the order of $\sum_{i=1}^{\infty} e_n(i)$ and $\sum_{i=1}^{\infty} ie_n(i)$.

Lemma 3 For any constant $C \in (0, 2p_l^2 d_m^2)$,

$$\sum_{i=1}^{\infty} e_n(i) = O\left(\exp\{-Cn\}\right), \text{ and } \sum_{i=1}^{\infty} ie_n(i) = O\left(\exp\{-Cn\}\right) \tag{12}$$

Proof. For sufficiently large n such that $a_n < d_m$, because $\{e_n(i), i \ge 1\}$ is a decreasing sequence, we have,

$$\begin{aligned}
\sum_{i=1}^{\infty} e_n(i) &= \sum_{i=1}^{\infty} \exp\left\{-2np_l^2((i-1)\Delta + d_m - a_n)^2\right\} \\
&\le e_n(1) + \int_0^{\infty} \exp\left\{-2np_l^2\left(t\Delta + d_m - a_n\right)^2\right\} dt \\
&= e_n(1) + \frac{\sqrt{2\pi(4np_l^2)^{-1}}}{\Delta} \int_0^{\infty} \frac{1}{\sqrt{2\pi(4np_l^2)^{-1}}} \exp\left\{-\frac{(t + d_m - a_n))^2}{2(4np_l^2)^{-1}}\right\} dt \\
&= e_n(1) + \Delta^{-1}\sqrt{\frac{\pi}{2np_l^2}}\,\mathbb{P}\left(\mathcal{N}\left(0, (4np_l^2)^{-1}\right) \ge d_m - a_n\right) \\
&\le e_n(1) + \Delta\sqrt{\frac{\pi}{2np_l^2}} \exp\{-2np_l^2(d_m - a_n)^2\} \\
&= O\left(\exp\{-2np_l^2(d_m - a_n)^2\}\right) = O\left(\exp\{-Cn\}\right).
\end{aligned}$$

Similarly, for sufficiently large n, $\{ie_n(i), i \ge 1\}$ is decreasing, and thus

$$\begin{aligned}
\sum_{i=1}^{\infty} ie_n(i) &= \sum_{i=1}^{\infty} j\exp\left\{-2np_l^2((i-1)\Delta + d_m - a_n)^2\right\} \\
&\le e_n(1) + \int_0^{\infty} t\exp\left\{-2np_l^2\left(t\Delta + d_m - a_n\right)^2\right\} dt \\
&\le e_n(1) + \frac{1}{\Delta}\int_{d_m - a_n}^{\infty} t\exp\left\{-2np_l^2 t^2\right\} dt \\
&\qquad - 2(d_m - a_n)\int_0^{\infty} \exp\left\{-2np_l^2\left(t\Delta + d_m - a_n\right)^2\right\} dt \\
&= O\left(\exp\{-Cn\}\right).
\end{aligned}$$

□

Then, by Lemma 2, the first term on the right-hand-side of Equation (10) is bounded by

$$\mathbb{E}\left[\left(L_m^{\Delta}(X_{\lceil n\alpha\rceil;n})\right)^2 \mathbb{I}_{\{L_m^{\Delta}(X_{\lceil n\alpha\rceil;n})\le-\Delta\}}\right] = \Delta^2\mathbb{P}(L_m^{\Delta}(X_{\lceil n\alpha\rceil;n}) \le -\Delta) + \Delta^2\sum_{j=1}^{\infty}(2j+1)\mathbb{P}(L_m^{\Delta}(X_{\lceil n\alpha\rceil;n}) \le -j\Delta)$$
$$\le \Delta^2(e_n(1) + \sum_{j=1}^{\infty}(2j+1)e_n(j)) = O\left(\exp\{-Cn\}\right),$$

for any $C < C_m := 2p_l^2 d_m^2$.

The bound for the second term on the right-hand-side of Equation (10) follows in a similar manner and is thus omitted, which then completes the proof.

REFERENCES

Broadie, M., Y. Du, and C. C. Moallemi. 2011. "Efficient Risk Estimation via Nested Sequential Simulation". *Management Science* 57(6):1172–1194.

Broadie, M., Y. Du, and C. C. Moallemi. 2015. "Risk Estimation via Regression". *Operations Research* 63(5):1077–1097.

Chen, J. and N. A. Lazar. 2010. "Quantile Estimation for Discrete Data via Empirical Likelihood". *Journal of Nonparametric Statistics* 22(2):237–255.

Feldman, D. and H. G. Tucker. 1966. "Estimation of Non-unique Quantiles". *The Annals of Mathematical Statistics* 37(2):451–457.

Feng, B., G. Liu, and K. Zhang. 2022. "Portfolio Risk Measurement via Stochastic Mesh with Average Weight". In *2022 Winter Simulation Conference (WSC)*, 903–914 https://doi.org/10.1109/WSC57314.2022.10015494.

Gordy, M. B. and S. Juneja. 2010. "Nested Simulation in Portfolio Risk Measurement". *Management Science* 56(10):1833–1848.

Hong, L. J., S. Juneja, and G. Liu. 2017. "Kernel Smoothing for Nested Estimation with Application to Portfolio Risk Measurement". *Operations Research* 65(3):657–673.

Liu, M. and J. Staum. 2010. "Stochastic Kriging for Efficient Nested Simulation of Expected Shortfall". *Journal of Risk* 12(3):3.

Ma, Y., M. G. Genton, and E. Parzen. 2011. "Asymptotic Properties of Sample Quantiles of Discrete Distributions". *Annals of the Institute of Statistical Mathematics* 63:227–243.

Serfling, R. J. 2009. *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons.

Wang, D. and A. D. Hutson. 2011. "A Fractional Order Statistic Towards Defining a Smooth Quantile Function for Discrete Data". *Journal of statistical planning and inference* 141(9):3142–3150.

Zhang, K., G. Liu, and S. Wang. 2017. "Portfolio Risk Measurement via Stochastic Mesh". In *2017 Winter Simulation Conference (WSC)*, 1796–1811 https://doi.org/10.1109/WSC.2017.8247917.

Zhang, K., G. Liu, and S. Wang. 2022. "Bootstrap-Based Budget Allocation for Nested Simulation". *Operations Research* 70(2):1128–1142.

AUTHOR BIOGRAPHIES

XIANYU KUANG earned his Ph.D. in the Department of Management Sciences, College of Business at City University of Hong Kong, and is currently a Research Assistant in the same department. His research interests include stochastic simulation, machine learning, financial engineering and risk management. His email address is xykuang4-c@my.cityu.edu.hk.

GUANGWU LIU is a Professor in the Department of Management Sciences, College of Business at City University of Hong Kong. His research interests include stochastic simulation, business analytics, financial engineering, and risk management. He serves as an Associate Editor of the Asia-Pacific Journal of Operational Research, and Naval Research Logistics. His email address is guanliu@cityu.edu.hk. His website is https://www.cb.cityu.edu.hk/staff/guanliu/.

QIANWEN ZHU is a Ph.D. candidate in the Department of Management Sciences, College of Business at City University of Hong Kong. Her research focuses on stochastic simulation and machine learning, and their applications. Her email address is qianwen.zhu@my.cityu.edu.hk.