Proceedings of the 2018 Winter Simulation Conference
M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, eds.

MONTE CARLO ESTIMATION OF ECONOMIC CAPITAL

Zachary T. Kaplan
Yajuan Li
Marvin K. Nakayama

Computer Science Department
New Jersey Institute of Technology
Newark, NJ 07102, USA

ABSTRACT

Economic capital (EC) is a risk measure that has been used by financial firms to help determine capital levels to hold to protect (with high probability) against large unexpected losses of credit portfolios. Given a stochastic model for a portfolio's loss over a given time horizon, the EC is defined as the difference between a quantile and the mean of the loss distribution. We describe Monte Carlo methods for estimating the EC. We apply measure-specific importance sampling to separately estimate the two components of the EC, which can lead to much smaller variance than when estimating both terms simultaneously. We provide Bahadur-type representations for our estimators of the EC, which we further exploit to establish central limit theorems and asymptotically valid confidence intervals. We present numerical results for a simple model to demonstrate the effectiveness of our approaches.

1 INTRODUCTION

Risk management aims to protect a financial institution from future uncertainties. For example, risk managers employ various *risk measures* (e.g., Section 2.3 of McNeil et al. 2015) to help determine appropriate levels of capital needed to be able to absorb (with high probability) large unexpected losses in credit portfolios comprising loans, bonds, and other financial instruments subject to default. One such risk measure is a *quantile*, also known as a *percentile* or *value-at-risk* (VaR). For a specified constant $0 < p < 1$, e.g., $p = 0.999$, the p-quantile is the constant ξ such that there is probability p that the loss Y over a fixed time horizon, e.g., one year, is less than ξ. Another example is the *expected shortfall*, also called the *conditional tail expectation* or *conditional value-at-risk*, which, for continuous loss distributions, is the expected loss given that it exceeds ξ.

This paper focuses on another risk measure, *economic capital* (EC), defined as the difference between a quantile and the mean loss; see p. 5 of Klaassen and van Eeghen (2009), Section 2.4 of Lütkebohmert (2009), and the Moody's Analytics white paper by Levy et al. (2013). (The EC is alternatively called the *relative* or *mean-adjusted VaR*; see p. 108 of Jorion 2007; and p. 300 of McNeil et al. 2015.) For example, Deutsche Bank (2017) appears to use EC to help determine capital levels: p. 60 states, "In line with our economic capital framework, economic capital for credit risk is set at a level to absorb with a probability of 99.9% very severe aggregate unexpected losses within one year. Our economic capital for credit risk is derived from the loss distribution of a portfolio via Monte Carlo Simulation of correlated rating migrations."

Because of the rarity of extreme losses, Monte Carlo simulation with *simple random sampling* (SRS) can produce noisy estimates of the EC. This motivates the use of *variance-reduction techniques* (VRTs), such as *importance sampling* (IS); e.g., see Chapters V and VI of Asmussen and Glynn (2007) and Chapter 4 of Glasserman (2004) for overviews of these methods. Glasserman and Li (2005) devise IS schemes for estimating tail probabilities of multi-factor credit-risk models with a Gaussian copula to capture

978-1-5386-6572-5/18/$31.00

dependencies of default events across obligors (e.g., corporations to which a bank has extended credit). Bassamboo et al. (2008) design IS methods to handle non-Gaussian forms of dependencies.

While IS can be effective in reducing the variance of estimators of tail probabilities and extreme quantiles of losses, the same approach may be detrimental when further used to estimate the mean loss in the EC. An IS scheme that works well in estimating an extreme quantile will typically sample more from the corresponding tail of the distribution and less around the mean, degrading the mean's estimator. This suggests employing different simulation techniques to separately estimate the two measures comprising the EC. Specifically, we utilize IS to estimate the quantile and independently apply SRS to estimate the mean. Goyal et al. (1987) call this approach *measure-specific importance sampling* (MSIS), which they adopt to separately estimate the numerator and denominator in a ratio formula, where only the denominator corresponds to a rare event. Through a simple example, we demonstrate numerically the benefits when $p \approx 1$ of applying MSIS to estimate the EC rather than estimating both the quantile and mean simultaneously. For the models in Glasserman and Li (2005) and Bassamboo et al. (2008), computing the mean loss may not require simulation because of their models' tractability. But more complicated stochastic models may preclude analytically solving for the mean loss, thereby motivating the use of Monte Carlo to estimate it.

The rest of our paper proceeds as follows. Section 2 describes the mathematical framework that we adopt. Section 3 develops the SRS estimator of the EC, proves a Bahadur-type asymptotic representation (Bahadur 1966) for the estimator, and also establishes a *central limit theorem* (CLT). To account for the sampling error in our EC estimator, we also provide two asymptotically valid methods for constructing *confidence intervals* (CIs) for the EC: one based on batching and the other on sectioning. Section 4 employs IS to estimate the EC, and proves that the resulting estimator satisfies a Bahadur-type representation and a CLT. Section 5 gives the details for using MSIS to estimate the EC. We also develop an asymptotically valid CI for η when applying MSIS and sectioning. Section 6 presents numerical results for a simple model. Section 7 has some concluding remarks. Our paper contains several theorems, none of which have appeared previously in the literature to the best of our knowledge, but space limitations necessitate deferring their formal proofs to a forthcoming follow-up paper.

2 MATHEMATICAL FRAMEWORK

Let Y be a random variable denoting the loss of a credit-portfolio model over a given time horizon, and let F be the *cumulative distribution function* (CDF) of Y. We assume that F is unknown or intractable, but we have a simulation model that generates observations of $Y \sim F$. Let $\mu = E[Y]$ be the mean of $Y \sim F$. For a CDF H and constant $0 < q < 1$, we define the *q-quantile* of H as $H^{-1}(q) = \inf\{y : H(y) \geq q\}$; e.g., the median is the 0.5-quantile, also known as the 50th percentile. Then we define the *economic capital* as

$$\eta = \xi - \mu, \tag{1}$$

where $\xi = F^{-1}(p)$ for a given $0 < p < 1$, e.g., $p = 0.999$. (As ξ and η depend on p, we should instead write them as ξ_p and η_p, but we often omit the subscript p to simplify notation.) The goal is to use simulation to estimate η and provide an asymptotically valid confidence interval for η to measure the sampling error.

We will sometimes (but not always) assume that the loss Y has the form

$$Y = c(\mathbf{X}) \tag{2}$$

for a given function $c : \Re^d \to \Re$ with $d \geq 1$, and random vector $\mathbf{X} = (X_1, X_2, \ldots, X_d)$ having a specified joint CDF G, where G can allow the components $X_1, X_2, \ldots, X_d$ of $\mathbf{X}$ to be dependent and non-identically distributed. Let G_j be the marginal CDF of X_j. If we further assume that

$$G(\mathbf{x}) = \prod_{j=1}^{d} G_j(x_j) \quad \text{for all } \mathbf{x} = (x_1, x_2, \ldots, x_d) \in \Re^d, \tag{3}$$

then $X_1, X_2, \ldots, X_d$ are independent. In the case that X_j is a continuous (resp., discrete) random variable, then let g_j be the density (resp., probability mass function) of G_j. We can think of the function c in (2) as a computer code, which can be quite involved, that transforms an input $\mathbf{X} \sim G$ into a loss $Y \sim F$.

For example, as in Glasserman and Li (2005), Bassamboo et al. (2008), and Lütkebohmert (2009), we may consider a multi-factor credit-risk model for which the loss Y has the form in (2), with mutually independent components in $\mathbf{X}$. Specifically, suppose there are $m \geq 1$ obligors, and we induce dependence among the default events across obligors as follows. Let $X_1, \ldots, X_r$ be independent and identically distributed (i.i.d.) $N(0,1)$ random variables representing the *systematic risk factors*, which model global, country, and sector factors that impact all obligors, where $N(q, s^2)$ represents a normal random variable with mean q and variance s^2. For each $k = 1, 2, \ldots, m$, let X_{r+k} be another independent $N(0,1)$ random variable denoting the *idiosyncratic risk* associated with obligor k. Define constants $a_{k,j}$, $k = 1, 2, \ldots, m$, $j = 1, 2, \ldots, r$, as the *loading factors*, satisfying $\sum_{j=1}^{r} a_{k,j}^2 \leq 1$ for each obligor k. Let $b_k = [1 - \sum_{j=1}^{r} a_{k,j}^2]^{1/2}$, so $(\sum_{j=1}^{r} a_{k,j} X_j) + b_k X_{r+k}$ is $N(0,1)$ for each obligor k. Let X_{r+m+1} be a positive random variable representing a *common shock* affecting all obligors. For each obligor $k = 1, 2, \ldots, m$, let t_k be a given constant, and we assume that obligor k defaults if and only if

$$\frac{(\sum_{j=1}^{r} a_{k,j} X_j) + b_k X_{r+k}}{X_{r+m+1}} > t_k.$$

Glasserman and Li (2005) and Bassamboo et al. (2008) assume for simplicity that when obligor k defaults, the loss resulting from that default is a constant c_k, but they note that their results can also permit the loss to be stochastic (under appropriate conditions). For obligor k, let $X_{r+m+1+k}$ be another random variable, and define the loss from obligor k defaulting as $v_k(X_1, \ldots, X_{r+m+1}, X_{r+m+1+k})$ for a given function $v_k : \Re^{r+m+2} \to \Re_+$. Thus, the loss can depend on the systematic and idiosyncratic risk factors and common shock, as in Farinelli and Shkolnikov (2012). Finally, the function c in (2) for the total portfolio loss is

$$c(\mathbf{X}) = \sum_{k=1}^{m} v_k(X_1, \ldots, X_{r+m+1}, X_{r+m+1+k}) \, I\left(\frac{(\sum_{j=1}^{r} a_{k,j} X_j) + b_k X_{r+k}}{X_{r+m+1}} > t_k \right),$$

where $I(\cdot)$ denotes the indicator function, which equals 1 (resp., 0) if its argument is true (resp., false). Thus, the vector $\mathbf{X}$ in (2) has dimension $d = r + 2m + 1$ in this example.

3 SIMPLE RANDOM SAMPLING

We start by applying SRS to estimate η, and the results throughout this section do not require that the loss $Y \sim F$ has the form in (2). Let $Y_1, Y_2, \ldots, Y_n$ be a *random sample* of size n from F; i.e., $Y_1, Y_2, \ldots, Y_n$ are i.i.d. random variables, each with CDF F. In the special case when Y has the form in (2), we generate $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ as i.i.d. copies of $\mathbf{X} \sim G$, and let $Y_i = c(\mathbf{X}_i)$ for each $i = 1, 2, \ldots, n$. In general, define

$$\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^{n} Y_i \tag{4}$$

as the sample mean. Let $\hat{F}_n$ be the *empirical CDF*, i.e.,

$$\hat{F}_n(y) = \frac{1}{n} \sum_{i=1}^{n} I(Y_i \leq y). \tag{5}$$

Let $\hat{\xi}_n = \hat{F}_n^{-1}(p)$ be an SRS p-quantile estimator. The SRS estimator of the EC $\eta = \xi - \mu$ is then

$$\hat{\eta}_n = \hat{\xi}_n - \hat{\mu}_n. \tag{6}$$

We can equivalently compute $\hat{\xi}_n$ through order statistics: let $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(n)}$ be the sorted values of $Y_1, Y_2, \ldots, Y_n$, and $\hat{\xi}_n = Y_{(\lceil np \rceil)}$, where $\lceil \cdot \rceil$ denotes the ceiling (i.e., round-up) function. For simplicity, we do not consider other quantile estimators (Hyndman and Fan 1996), e.g., based on kernel methods or by using an interpolated version of $\hat{F}_n$.

3.1 Asymptotic Properties of SRS EC Estimator

Although the estimator $\hat{\mu}_n$ in (4) of the mean is a sample average, the p-quantile estimator $\hat{\xi}_n = \hat{F}_n^{-1}(p)$ is *not*, so the EC estimator $\hat{\eta}_n$ in (6) is also not simply a sample average. This complicates the analysis of $\hat{\eta}_n$. However, Bahadur (1966) shows that when the sample size n is large, $\hat{\xi}_n$ is well approximated by a sample average of i.i.d. quantities, and we will do the same for $\hat{\eta}_n$. Specifically, let f be the derivative (when it exists) of the CDF F. Also, let $\Rightarrow$ denote convergence in distribution (e.g., Chapter 5 of Billingsley 1995). Then (see Section 2.5 of Serfling 1980) if $f(\xi) > 0$, the p-quantile estimator satisfies

$$\hat{\xi}_n = \xi - \frac{1}{f(\xi)}\left[\hat{F}_n(\xi) - p\right] + R_n, \tag{7}$$

with

$$\sqrt{n}R_n \Rightarrow 0 \quad \text{as } n \to \infty. \tag{8}$$

If in addition F is twice differentiable at ξ, then for either choice of sign below,

$$\limsup_{n\to\infty} \pm \frac{n^{3/4}R_n}{(\log\log n)^{3/4}} = \frac{2^{5/4}[p(1-p)]^{1/4}}{3^{3/4}} \quad \text{with probability 1.} \tag{9}$$

Note that (9) implies (8), and we call (7) with (8) (resp., (9)) a *weak* (resp., *strong*) *Bahadur representation* for $\hat{\xi}_n$. The key point of (7)–(9) is that they enable us to analyze the asymptotics of $\hat{\xi}_n$ through the simpler $\hat{F}_n(\xi)$, which is a sample average of i.i.d. terms by (5). The following result establishes that the SRS estimator $\hat{\eta}_n$ of the EC η has similar *Bahadur-type representations*.

Theorem 1 Suppose $Y_1, Y_2, \ldots$ are i.i.d. with CDF F, where F is differentiable at ξ with $f(\xi) > 0$. Then

$$\hat{\eta}_n = \eta - \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{f(\xi)}[I(Y_i \le \xi) - p] + [Y_i - \mu]\right) + R_n \tag{10}$$

with R_n from (7), so (8) holds. If in addition F is twice differentiable at ξ, then (9) further holds.

Theorem 1 provides useful insight into the asymptotic behavior of $\hat{\eta}_n$. As noted before, $\hat{\eta}_n$ is not a sample mean, complicating its analysis. But (10), which follows from (4)–(9), shows that when n is large, $\hat{\eta}_n$ can be represented as the sum of η, the sample mean of the $W_i \equiv -[I(Y_i \le \xi) - p]/f(\xi) - [Y_i - \mu]$, and remainder term R_n. Because R_n vanishes faster than $1/\sqrt{n}$ (by (8) or (9)), we can then determine the $n^{1/2}$-asymptotics (e.g., CLT) of $\hat{\eta}_n$ through the sample mean of W_i, $i = 1, 2, \ldots, n$, which are i.i.d.

Let $\sigma^2 = \mathrm{Var}[Y]$, the variance of Y. We next establish a CLT for $\hat{\eta}_n$, which follows from Theorem 1.

Theorem 2 Suppose $Y_1, Y_2, \ldots$ are i.i.d. with CDF F and $f(\xi) > 0$. If $0 < \sigma^2 < \infty$, then

$$\sqrt{n}\,[\hat{\eta}_n - \eta] \Rightarrow N(0, \psi^2) \quad \text{as } n \to \infty,$$

where

$$\psi^2 = \frac{p(1-p)}{f^2(\xi)} + \sigma^2 + \frac{2}{f(\xi)}\mathrm{Cov}[I(Y \le \xi), Y]. \tag{11}$$

We next develop approaches to construct asymptotically valid confidence intervals for η based on the CLT in Theorem 2 and the Bahadur-type representation in Theorem 1. Although it is possible to consistently estimate the asymptotic variance ψ^2 in (11), we instead consider methods that avoid this issue.

3.2 Confidence Interval Using Batching

We first consider *batching* (e.g., p. 202 of Glasserman 2004) to construct a CI for η. The method builds $b \ge 2$ independent estimates of η, and then constructs a CI from the sample mean and sample variance of the b estimates of η. Numerical studies in Nakayama (2014) of batching CIs for a quantile for different b suggest that setting $b = 10$ may be reasonable in practice. We next provide the details of batching for η.

Let n be the overall sample size, and let $m = n/b$ be the batch size, which we assume is integer-valued. We will establish the asymptotic validity of the batching CI as the overall sample size n grows large with b fixed, which means that the batch size m grows large. For $j = 1, 2, \ldots, b$, we take the jth batch to consist of the outputs $Y_{j,i} \equiv Y_{(j-1)m+i}$, $i = 1, 2, \ldots, m$. For batch j, compute the sample mean $\hat{\mu}_{j,m} = (1/m)\sum_{i=1}^{m} Y_{j,i}$, the empirical CDF $\hat{F}_{j,m}$ with $\hat{F}_{j,m}(y) = (1/m)\sum_{i=1}^{m} I(Y_{j,i} \le y)$, and the p-quantile estimator $\hat{\xi}_{j,m} = \hat{F}_{j,m}^{-1}(p)$. Then, the estimator of the EC η from batch j is $\hat{\eta}_{j,m} = \hat{\xi}_{j,m} - \hat{\mu}_{j,m}$. The b estimators $\hat{\eta}_{j,m}$, $j = 1, 2, \ldots, b$, across the batches are i.i.d., and we compute their sample mean and sample variance as

$$\bar{\eta}_{b,m} = \frac{1}{b}\sum_{j=1}^{b} \hat{\eta}_{j,m} \quad \text{and} \quad S_{b,m}^2 = \frac{1}{b-1}\sum_{j=1}^{b} [\hat{\eta}_{j,m} - \bar{\eta}_{b,m}]^2, \tag{12}$$

respectively. For each batch j, the estimator $\hat{\eta}_{j,m}$ satisfies a CLT as the batch size $m \to \infty$ by Theorem 2, so each $\hat{\eta}_{j,m}$ is approximately normal when the batch size m is large. Moreover, batches are independent, so $\hat{\eta}_{j,m}$, $j = 1, 2, \ldots, b$, are independent. For large m, we have that $[\bar{\eta}_{b,m} - \eta]/[S_{b,m}^2/b]^{1/2}$ has approximately a Student t distribution with $b-1$ degrees of freedom. Let H_{b-1} be the CDF of a Student t random variable with $b-1$ degrees of freedom, and let $\tau_{b-1,q} = H_{b-1}^{-1}(q)$ be the q-quantile of H_{b-1}. We then define the two-sided, β-level *batching CI* for η as

$$I_{b,m} = (\bar{\eta}_{b,m} \pm \tau_\beta S_{b,m}/\sqrt{b}), \tag{13}$$

where $\tau_\beta \equiv \tau_{b-1,1-(1-\beta)/2}$. For example, for $b = 10$ batches and $\beta = 0.95$, the 95% CI $I_{b,m}$ uses $\tau_\beta = 2.262$. The following theorem establishes the asymptotic validity of the SRS batching CI.

Theorem 3 Under the assumptions of Theorem 2, we have $\lim_{m\to\infty} P(\eta \in I_{b,m}) = \beta$ for any fixed $b \ge 2$.

3.3 Confidence Interval Using Sectioning

Although Theorem 3 establishes that the batching CI $I_{b,m}$ in (13) is asymptotically valid as the batch size $m \to \infty$ for a fixed number $b \ge 2$ of batches, the CI may have poor coverage when the overall sample size $n = bm$ is not very large. To understand why, we examine the bias of the batching point estimator $\bar{\eta}_{b,m}$ from (12). Note that

$$\begin{aligned} \text{Bias}[\bar{\eta}_{b,m}] = E[\bar{\eta}_{b,m}] - \eta &= \frac{1}{b}\sum_{j=1}^{b} E[\hat{\eta}_{j,m}] - \eta = E[\hat{\eta}_{j,m}] - \eta \\ &= E[\hat{\xi}_{j,m}] - E[\hat{\mu}_{j,m}] - (\xi - \mu) = E[\hat{\xi}_{j,m}] - \xi \end{aligned}$$

because $\hat{\mu}_{j,m}$ is unbiased. While the bias of $\hat{\xi}_{j,m}$ converges to 0 as $m \to \infty$, the bias is nonzero in general for fixed m. The bias of the batching point estimator $\bar{\eta}_{b,m}$ is determined by the bias of $\hat{\xi}_{j,m}$, which depends on the batch size $m = n/b$. But because $m < n$, the bias of $\bar{\eta}_{b,m}$ can be significant when the overall sample size n is not very large. Thus, the batching CI $I_{b,m}$ may be poorly centered on average, so the coverage can suffer for small n; i.e., $P(\eta \in I_{b,m})$ may differ significantly from β for small n.

To address this issue, we may instead apply *sectioning*, which was originally proposed in Section III.5a of Asmussen and Glynn (2007). The basic idea is to modify the batching CI $I_{b,m}$ by replacing the batching point estimator $\bar{\eta}_{b,m}$ throughout with the overall estimator $\hat{\eta}_n$ from (6) based on the total sample size n. Because the bias of the overall point estimator $\hat{\eta}_n$ is based on overall sample size n rather than the batch size $m = n/b$, the bias of $\hat{\eta}_n$ can be smaller than that of $\bar{\eta}_{b,m}$, and this modification can lead to a CI with improved coverage when n is small since the CI may now be better centered on average. Specifically, let

$$S_{b,m}'^2 = \frac{1}{b-1}\sum_{j=1}^{b} [\hat{\eta}_{j,m} - \hat{\eta}_n]^2, \tag{14}$$

which is similar to $S^2_{b,m}$ in (12) but where we replace the batching point estimator $\bar{\eta}_{b,m}$ with the overall point estimator $\hat{\eta}_n$. Then we define the two-sided, β-level *sectioning CI* for η as

$$I'_{b,m} = (\hat{\eta}_n \pm \tau_\beta S'_{b,m}/\sqrt{b}), \tag{15}$$

which is centered at the overall point estimator with sample size $n = bm$. The following result establishes the asymptotic validity of the SRS sectioning CI $I'_{b,m}$.

Theorem 4 Under the assumptions of Theorem 2, we have $\lim_{m\to\infty} P(\eta \in I'_{b,m}) = \beta$ for any fixed $b \geq 2$.

4 IMPORTANCE SAMPLING

SRS may produce an estimator of the EC $\eta = \xi - \mu$ with large variance when $p \approx 1$ because ξ is then an extreme quantile, so now we consider applying variance reduction. We focus on importance sampling, but other VRTs can also be employed. To use IS, we assume throughout this section that Y has the form in (2).

In the setting of (2), we write the mean μ in (1) as $\mu = E_G[c(\mathbf{X})]$, where E_G denotes the expectation operator when the $\Re^d$-valued random vector $\mathbf{X} \sim G$. Let $\tilde{G}$ be another joint distribution on $\Re^d$ such that the measure m_G corresponding to G is absolutely continuous (p. 422 of Billingsley 1995) with respect to the measure $m_{\tilde{G}}$ corresponding to $\tilde{G}$; i.e., $m_G(A) = 0$ for every measurable set $A \subseteq \Re^d$ for which $m_{\tilde{G}}(A) = 0$. Then we can express the mean μ as

$$\mu = E_G[c(\mathbf{X})] = \int_{\Re^d} c(\mathbf{x})\, dG(\mathbf{x}) = \int_{\Re^d} c(\mathbf{x}) \frac{dG(\mathbf{x})}{d\tilde{G}(\mathbf{x})}\, d\tilde{G}(\mathbf{x}) = E_{\tilde{G}}[c(\mathbf{X})L(\mathbf{X})], \tag{16}$$

where $E_{\tilde{G}}$ denotes expectation when $\mathbf{X} \sim \tilde{G}$, and

$$L(\mathbf{x}) = \frac{dG(\mathbf{x})}{d\tilde{G}(\mathbf{x})} \tag{17}$$

is the *likelihood ratio*. By (16), we can obtain an unbiased estimator of μ by generating i.i.d. $\mathbf{X}_i \sim \tilde{G}$, $i = 1, 2, \ldots, n$, and then forming the *IS estimator* of μ as

$$\tilde{\mu}_n = \frac{1}{n} \sum_{i=1}^{n} c(\mathbf{X}_i) L(\mathbf{X}_i). \tag{18}$$

Suppose that the original joint CDF G satisfies (3), so a random vector with CDF G has independent components. Let $\tilde{G}_j$ be the marginal CDF of X_j when $\mathbf{X} = (X_1, X_2, \ldots, X_d) \sim \tilde{G}$. Also, suppose that

$$\tilde{G}(\mathbf{x}) = \prod_{j=1}^{d} \tilde{G}_j(x_j) \qquad \text{for all } \mathbf{x} = (x_1, x_2, \ldots, x_d) \in \Re^d, \tag{19}$$

so $\mathbf{X} \sim \tilde{G}$ has independent components. Further suppose that each G_j (resp., $\tilde{G}_j$) has density or probability mass function g_j (resp., $\tilde{g}_j$). Then the likelihood ratio in (17) becomes

$$L(\mathbf{x}) = \prod_{j=1}^{d} \frac{g_j(x_j)}{\tilde{g}_j(x_j)}. \tag{20}$$

To estimate the p-quantile ξ using IS, we will follow an approach developed by Glynn (1996): first apply IS to estimate the CDF F, and then invert the estimated CDF to obtain the IS quantile estimator. Specifically, let P_G denote the probability measure when $\mathbf{X} \sim G$, and write

$$\begin{aligned} F(y) &= 1 - P(Y > y) = 1 - P_G(c(\mathbf{X}) > y) = 1 - E_G[I(c(\mathbf{X}) > y)] = 1 - \int_{\Re^d} I(c(\mathbf{x}) > y)\, dG(\mathbf{x}) \\ &= 1 - \int_{\Re^d} I(c(\mathbf{x}) > y) \frac{dG(\mathbf{x})}{d\tilde{G}(\mathbf{x})}\, d\tilde{G}(\mathbf{x}) = 1 - E_{\tilde{G}}[I(c(\mathbf{X}) > y)L(\mathbf{X})]. \end{aligned} \tag{21}$$

Thus, by (21), we can obtain an unbiased estimator of $F(y)$ for each y as

$$\tilde{F}_n(y) = 1 - \frac{1}{n}\sum_{i=1}^{n} I(c(\mathbf{X}_i) > y)L(\mathbf{X}_i), \tag{22}$$

where $\mathbf{X}_i \sim \tilde{G}$, $i = 1,2,\ldots,n$, are the same as in (18). We call $\tilde{F}_n(y)$ the *IS estimator* of $F(y)$.

The corresponding IS estimator $\tilde{\xi}_n$ of the p-quantile $F^{-1}(p)$ inverts (22), i.e.,

$$\tilde{\xi}_n = \tilde{F}_n^{-1}(p), \tag{23}$$

which can be computed as follows. Let $Y_i = c(\mathbf{X}_i)$, and let $Y_{1:n} \le Y_{2:n} \le \cdots \le Y_{n:n}$ be the sorted values of $Y_1, Y_2, \ldots, Y_n$. Also, let $\mathbf{X}_{i::n}$ be the $\mathbf{X}_j$ corresponding to $Y_{i:n}$. Then we have that $\tilde{\xi}_n = Y_{i_p:n}$, where i_p is the greatest integer for which $\sum_{\ell=i_p}^{n} L(\mathbf{X}_{\ell::n}) \ge n(1-p)$. Chu and Nakayama (2012) establish a weak Bahadur representation for the quantile estimator obtained through a combination of IS and stratified sampling, and $\tilde{\xi}_n$ in (23) is a special case of IS only; i.e., their Theorem 4.2 shows that if

$$\text{there exist constants } \varepsilon > 0 \text{ and } \lambda > 0 \text{ such that } E_{\tilde{G}}[I(c(\mathbf{X}) > \xi - \lambda)L^{2+\varepsilon}(\mathbf{X})] < \infty, \tag{24}$$

then

$$\tilde{\xi}_n = \xi - \frac{1}{f(\xi)}[\tilde{F}_n(\xi) - p] + \tilde{R}_n$$

with $\tilde{R}_n$ satisfying

$$\sqrt{n}\tilde{R}_n \Rightarrow 0 \quad \text{as } n \to \infty.$$

We can obtain an alternative quantile estimator by first writing $F(y) = E_G[I(c(\mathbf{X}) \le y)] = E_{\tilde{G}}[I(c(\mathbf{X}) \le y)L(\mathbf{X})]$ and using i.i.d. $\mathbf{X}_i \sim \tilde{G}$, $i = 1,2,\ldots,n$, to form an unbiased estimator of $F(y)$ as

$$\tilde{F}'_n(y) = \frac{1}{n}\sum_{i=1}^{n} I(c(\mathbf{X}_i) \le y)L(\mathbf{X}_i). \tag{25}$$

This leads to another p-quantile estimator $\tilde{\xi}'_n = \tilde{F}_n'^{-1}(p)$. Theorem 4.1 of Chu and Nakayama (2012) (resp., Sun and Hong 2010) establishes a weak (resp., strong) Bahadur representation for $\tilde{\xi}'_n$. But it turns out that when trying to estimate the p-quantile for $p \approx 1$ using IS, Glynn (1996) shows for a simple example that the p-quantile estimator $\tilde{\xi}_n$ in (23) based on (22) has smaller asymptotic variance in its CLT than the estimator $\tilde{\xi}'_n$ obtained by inverting (25). (In contrast, when $p \approx 0$, the p-quantile estimator $\tilde{\xi}'_n$ can have smaller asymptotic variance than $\tilde{\xi}_n$.)

We define the IS estimator of the EC as

$$\tilde{\eta}_n = \tilde{\xi}_n - \tilde{\mu}_n, \tag{26}$$

where both $\tilde{\xi}_n$ and $\tilde{\mu}_n$ are constructed using the same sample $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$, with each $\mathbf{X}_i \sim \tilde{G}$. The following results show that $\tilde{\eta}_n$ has a Bahadur-type representation and obeys a CLT.

Theorem 5 Suppose that $Y \sim F$ has the form in (2), and assume that $f(\xi) > 0$. Suppose that $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ are i.i.d. with CDF $\tilde{G}$, where the measure m_G corresponding to G is absolutely continuous with respect to the measure $m_{\tilde{G}}$ corresponding to $\tilde{G}$. Also suppose that (24) holds for $L(\mathbf{x})$ in (17). Then

$$\tilde{\eta}_n = \eta - \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{f(\xi)}[(1 - I(c(\mathbf{X}_i) > \xi)L(\mathbf{X}_i)) - p] + [c(\mathbf{X}_i)L(\mathbf{X}_i) - \mu]\right) + \tilde{R}_n$$

with

$$\sqrt{n}\tilde{R}_n \Rightarrow 0 \quad \text{as } n \to \infty.$$

Theorem 6 In addition to the assumptions of Theorem 5, also assume $\zeta^2 \equiv \mathrm{Var}_{\tilde{G}}[c(\mathbf{X})L(\mathbf{X})] < \infty$. Then $\sqrt{n}\,[\tilde{\eta}_n - \eta] \Rightarrow N(0, \kappa^2)$ as $n \to \infty$, where

$$\kappa^2 = \frac{\upsilon^2}{f^2(\xi)} + \zeta^2 + \frac{2\gamma}{f(\xi)}, \tag{27}$$

with

$$\upsilon^2 = E_{\tilde{G}}[I(c(\mathbf{X}) > \xi)L^2(\mathbf{X})] - (1-p)^2, \tag{28}$$
$$\gamma = -\mathrm{Cov}_{\tilde{G}}[I(c(\mathbf{X}) > \xi)L(\mathbf{X}), c(\mathbf{X})L(\mathbf{X})].$$

5 MEASURE-SPECIFIC IMPORTANCE SAMPLING

If we estimate both ξ and μ using the same IS distribution $\tilde{G}$ to sample $\mathbf{X}$, as in Section 4, the resulting estimator of $\eta = \xi - \mu$ may have large variance. Many VRTs are designed to handle ξ or μ individually but not both simultaneously. The problem is that ξ is a property of the tail of F, whereas μ typically measures its central tendency. A VRT that is constructed to efficiently analyze the tail of F may not work well to estimate its mean, and vice versa. It may be difficult to design a single VRT that does well to simultaneously estimate ξ and μ.

Thus, we instead use IS to estimate ξ only and independently estimate μ via SRS. This approach is known as *measure-specific importance sampling*, which Goyal et al. (1992) apply to estimate a ratio of means by estimating one (rare-event) mean using IS and independently estimating the other (non-rare) mean without IS. More generally, we could also apply one VRT to estimate ξ and another VRT to estimate μ, where each VRT may be something other than IS, but we do not pursue that here.

We now provide the details of MSIS. Fixing an overall sample size n, we allocate a fraction $0 < \delta < 1$ of the sample size to estimate ξ via IS, and we use the rest of the sample size to estimate μ using SRS. Let $n_1 = \delta n$ be the sample size for estimating ξ with IS, and $n_2 = (1-\delta)n$ be the SRS sample size to estimate μ. Here, we are assuming that both n_1 and n_2 are integer-valued; if not, let $n_1 = \lfloor \delta n \rfloor$ and $n_2 = n - n_1$, where $\lfloor \cdot \rfloor$ denotes the floor function. Let $\tilde{F}_{n,\delta}(y) = \tilde{F}_{n_1}(y)$, for $\tilde{F}_{n_1}(y)$ in (22) with n_1 replacing n, be the IS CDF estimator based on a sample size n_1, and let $\tilde{\xi}_{n,\delta} = \tilde{F}_{n,\delta}^{-1}(p)$ be the corresponding p-quantile estimator. Let $\hat{\mu}_{n,\delta} = \hat{\mu}_{n_2}$, for $\hat{\mu}_{n_2}$ in (4) with n_2 replacing n, be the SRS estimator of μ based on the sample size n_2. The MSIS estimator of η is then

$$\tilde{\eta}_{n,\delta} = \tilde{\xi}_{n,\delta} - \hat{\mu}_{n,\delta}. \tag{29}$$

The following result establishes a weak Bahadur-type representation for the MSIS estimator $\tilde{\eta}_{n,\delta}$ of η.

Theorem 7 Suppose that $Y \sim F$ has the form in (2), that $f(\xi) > 0$, and that (24) holds. Then for any fixed $0 < \delta < 1$, the MSIS estimator of η satisfies

$$\tilde{\eta}_{n,\delta} = \eta - \frac{1}{f(\xi)}[\tilde{F}_{n,\delta}(\xi) - p] - (\hat{\mu}_{n,\delta} - \mu) + \tilde{R}_{n,\delta}$$

with

$$\sqrt{n}\tilde{R}_{n,\delta} \Rightarrow 0 \quad \text{as } n \to \infty.$$

A consequence of the Bahadur-type representation for $\tilde{\eta}_{n,\delta}$ in Theorem 7 is that the MSIS estimator of η also satisfies a CLT.

Theorem 8 In addition to the assumptions of Theorem 7, further suppose that $\sigma^2 < \infty$. Then for any $0 < \delta < 1$, the MSIS estimator of η satisfies

$$\sqrt{n}\left[\tilde{\eta}_{n,\delta} - \eta\right] \Rightarrow N(0, \kappa_\delta^2) \quad \text{as } n \to \infty,$$

where, for υ^2 from (28),

$$\kappa_\delta^2 = \frac{\upsilon^2}{\delta f^2(\xi)} + \frac{\sigma^2}{1-\delta}. \tag{30}$$

Note that (30) does not contain a covariance term because MSIS independently estimates ξ and μ. In contrast, (11) and (27) include a covariance term because ξ and μ are estimated from the same sample.

5.1 Optimal Sampling Allocation

The asymptotic variance κ_δ^2 in (30) depends on the sampling-allocation parameter δ specified by the user. The optimal choice of δ to minimize κ_δ^2 can be easily found by setting the derivative of κ_δ^2 with respect to δ to 0, and solving. This leads to the optimal δ as $\delta^* = [\upsilon/f(\xi)]/[\sigma + (\upsilon/f(\xi))]$ to minimize the asymptotic variance.

In practice, the value of the optimal δ^* is unknown because σ^2, υ^2 and $f(\xi)$ are all unknown. But one could apply a two-stage procedure. In the first stage, fix a sampling allocation δ_0, e.g., $\delta_0 = 1/2$, and use a small sample size n' to estimate σ^2, υ^2 and $f(\xi)$, where $f(\xi)$ could be estimated, e.g., via a finite difference (Chu and Nakayama 2012). Then employ these estimates to approximate the optimal sampling allocation δ^*, which is then used in the second stage with some overall sample size $n'' \gg n'$.

5.2 Confidence Interval Using Sectioning

We now develop a CI for η when using MSIS and sectioning. First divide the overall sample size n into sample sizes $n_1 = \delta n$ for estimating ξ via IS and $n_2 = (1-\delta)n$ for estimating μ with SRS. Then fix $b \geq 2$ as the number of batches (e.g., $b = 10$), and let $m_1 = n_1/b$ and $m_2 = n_2/b$ be the batch sizes for MSIS, with $m = m_1 + m_2 = n/b$ as the overall batch size. We assume that m_1 and m_2 are integer valued; otherwise, let $m_1 = \lfloor n_1/b \rfloor$ and $m_2 = \lfloor n_2/b \rfloor$. For each batch $j = 1,2,\ldots,b$, let $\tilde{\xi}_{j,m,\delta}$ be the IS estimator of ξ from batch j, and $\tilde{\xi}_{j,m,\delta}$, $j = 1,2,\ldots,b$, are independent. Similarly, let $\hat{\mu}_{j,m,\delta}$ be the SRS estimator of μ from batch j, with $\hat{\mu}_{j,m_2,\delta}$, $j = 1,2,\ldots,b$, independent. From batch j, we define the MSIS estimator of η as $\tilde{\eta}_{j,m,\delta} = \tilde{\xi}_{j,m,\delta} - \hat{\mu}_{j,m,\delta}$, and $\tilde{\eta}_{j,m,\delta}$, $j = 1,2,\ldots,b$, are independent. Analogous to the SRS batching CI in (13), the MSIS batching CI uses the sample mean and sample variance of the $\tilde{\eta}_{j,m,\delta}$, $j = 1,2,\ldots,b$.

To construct the MSIS sectioning CI for η, we compute $\tilde{S}'^2_{b,m,\delta} = \frac{1}{b-1}\sum_{j=1}^{b}[\tilde{\eta}_{j,m,\delta} - \tilde{\eta}_{n,\delta}]^2$, which uses the overall MSIS estimator $\tilde{\eta}_{n,\delta}$ of η from (29) with sample size $n = bm$. (This is analogous to (14), which we used in the SRS sectioning CI $I'_{b,m}$ in (15).) We then obtain

$$\tilde{I}_{b,m,\delta} = \left(\tilde{\eta}_{n,\delta} \pm \tau_\beta \tilde{S}'_{b,m,\delta}/\sqrt{b}\right)$$

as the two-sided β-level CI for η using MSIS with sectioning, where we note the CI is centered at the overall MSIS estimator $\tilde{\eta}_{n,\delta}$. The following result shows that the MSIS sectioning CI $\tilde{I}_{b,n,\delta}$ for η is asymptotically valid. (Nakayama 2014 proves the asymptotic validity of the IS sectioning CI for just the quantile ξ.)

Theorem 9 Under the assumptions of Theorem 8, $\lim_{m\to\infty} P(\eta \in \tilde{I}_{b,m,\delta}) = \beta$ for fixed $b \geq 2$ and $0 < \delta < 1$.

As in Section 3.3, the overall MSIS estimator $\tilde{\eta}_{n,\delta}$ in (29) often has lower bias than the MSIS batching point estimator $(1/b)\sum_{j=1}^{b}\tilde{\eta}_{j,m,\delta}$. Thus, the sectioning CI can achieve better coverage than the batching CI when the overall sample size n is not large because the former can be better centered on average.

6 NUMERICAL EXPERIMENTS

We next provide numerical results for a simple model to demonstrate the benefits when $p \approx 1$ of estimating the EC $\eta_p \equiv \xi_p - \mu$ via MSIS rather than using either SRS or IS to estimate both the p-quantile ξ_p and μ. Specifically, we assume that Y has the form in (2), with $\mathbf{X} = (X_1, X_2, \ldots, X_d)$ a vector of $d = 10$ i.i.d. $N(0,1)$ random variables, where we define the function $c : \Re^d \to \Re$ in (2) as the sum $c(\mathbf{x}) = \sum_{j=1}^{d} x_j$. Thus, we have that $Y \sim N(0,d)$; i.e., the CDF F of Y is $F(y) = \Phi(y/\sqrt{d})$ with density $f(y) = \phi(y/\sqrt{d})/\sqrt{d}$, where Φ is the CDF of $N(0,1)$ and ϕ its density. The mean of F is $\mu = 0$, its variance is $\sigma^2 = d$, and the p-quantile is $\xi_p = F^{-1}(p) = \mu + \sigma\Phi^{-1}(p) = \sqrt{d}\Phi^{-1}(p)$. Hence, in (16), we have that $dG(\mathbf{x}) = \prod_{j=1}^{d} \phi(x_j)dx_j$, and the EC is $\eta_p = \xi_p - \mu = \sqrt{d}\Phi^{-1}(p)$.

To design an importance sampler that can be effective for estimating ξ_p, we choose the IS distribution for $\mathbf{X}$ so that the sum $Y = \sum_{j=1}^{d} X_j$ has mean ξ_p. We can do this by specifying the IS joint CDF $\tilde{G}$ to satisfy (19) (i.e., vector $\mathbf{X} \sim \tilde{G}$ has independent components), where each marginal $\tilde{G}_j$ is the CDF of $N(\nu_p, 1)$, with $\nu_p = \xi_p/d$. Thus, we have that $\tilde{G}_j(x_j) = \Phi(x_j - \nu_p)$, which has density $\tilde{g}_j(x_j) = \phi(x_j - \nu_p)$. In this case, the likelihood ratio in (20) becomes

$$L(\mathbf{x}) = \prod_{j=1}^{d} \frac{\phi(x_j)}{\phi(x_j - \nu_p)} = \prod_{j=1}^{d} \frac{(1/\sqrt{2\pi})e^{-x_j^2/2}}{(1/\sqrt{2\pi})e^{-(x_j-\nu_p)^2/2}} = \exp\left(\frac{d}{2}\nu_p^2 - \nu_p \sum_{j=1}^{d} x_j\right). \tag{31}$$

Then, we use the IS estimator of $F(y)$ to be (22), where the likelihood ratio is (31) and each $\mathbf{X}_i \sim \tilde{G}$.

Although we designed the IS joint CDF $\tilde{G}$ to be appropriate for estimating the p-quantile ξ_p, we can also use the same $\tilde{G}$ to estimate the mean μ via IS. The corresponding IS estimator of μ is given by (18), where we use (31) for the likelihood ratio.

Because of the tractability of our simple model, we are able to derive analytical expressions for the asymptotic variances in the CLTs of estimators of the EC $\eta_p = \xi_p - \mu$ based on various simulation methods. Table 1 gives the exact values of the asymptotic variances, which we evaluated numerically, for different values of p. (Deutsche Bank 2017, p. 43, appears to report the EC for $p = 0.999$, whereas $p = 0.9998$ was used the previous year.) The second (resp., third) column of Table 1 corresponds to estimating both ξ_p and μ from a single sample using SRS as in (6) (resp., using IS as in (26)), so the asymptotic variance is given by (11) (resp., (27)). The last three columns of Table 1 estimate ξ_p and μ using independent samples, with a proportion δ (resp., $1-\delta$) of the overall sample size allocated to estimate ξ_p (resp., μ). We set $\delta = 1/2$ in our calculations. In the second row, the notation "m1+m2" indicates that simulation method m1 (resp., m2) is employed to estimate ξ_p (resp., μ). Thus, the asymptotic variance for the column labeled "SRS+SRS" (resp., "IS+IS") is $p(1-p)/[\delta f^2(\xi)] + \sigma^2/(1-\delta)$ (resp., $\upsilon^2/[\delta f^2(\xi)] + \zeta^2/(1-\delta)$). The last column ("IS+SRS") of Table 1 corresponds to MSIS (Section 5), and (30) gives its asymptotic variance.

Table 1: We numerically computed the asymptotic variances of estimators of the EC $\eta_p = \xi_p - \mu$ based on various simulation methods for different p. "Single Sample" denotes estimating both ξ_p and μ from the same sample, using either SRS or IS. "Measure-Specific Sampling" corresponds to estimating ξ_p and μ independently, where the notation "m1+m2" in the second row denotes that method m1 (resp., m2) is used to estimate ξ_p (resp., μ), so MSIS corresponds to the last column.

	Single Sample		Measure-Specific Sampling		
p	SRS	IS	SRS+SRS	IS+IS	IS+SRS
0.9	1.92e+01	1.37e+02	7.84e+01	2.84e+02	3.09e+01
0.99	1.29e+02	1.44e+04	2.99e+02	2.87e+04	2.75e+01
0.999	8.71e+02	1.48e+06	1.78e+03	2.96e+06	2.61e+01
0.9998	3.47e+03	3.75e+07	6.98e+03	7.50e+07	2.56e+01

Table 1 shows that for $p = 0.9$, single-sample (SS) SRS has the lowest asymptotic variance of all the methods considered. But for more extreme p, MSIS outperforms all other methods, with a variance-reduction factor of $3.47\text{e+}03/2.56\text{e+}01 \approx 136$ compared to SS SRS for $p = 0.9998$. Thus, to achieve about the same width confidence interval as MSIS, SS SRS would need a sample size that is about 136-fold larger. Also, for $p = 0.9998$, the asymptotic variance of SS IS is a factor of $3.75\text{e+}07/2.56\text{e+}01 \approx 1.47\text{e+}06$ larger than MSIS, demonstrating the enormous benefit of MSIS by separately estimating ξ_p and μ.

To compare the performance of the batching and sectioning CIs when using MSIS, we ran coverage experiments when simulating our model. Table 2 lists the estimated coverage and average half width

(AHW) over $r = 10^4$ independent replications for various sample sizes n with $b = 10$ batches. The table also gives the sample variance of the batching and sectioning (i.e., overall) point estimators of the EC.

Table 2: We ran 10^4 independent replications to estimate the coverage and average half width of sectioning and batching confidence intervals with nominal confidence level $\beta = 0.95$ for the EC η_p estimated using MSIS for different sample sizes n for $p = 0.999$. We also give the sample variances of the point estimators of the EC across the 10^4 replications.

	Batching			Sectioning		
n	Coverage	AHW	Sample Variance	Coverage	AHW	Sample Variance
40	0.9154	2.162	9.864e–01	0.9790	2.269	6.948e–01
100	0.9096	1.294	3.589e–01	0.9698	1.351	2.714e–01
400	0.9411	0.576	6.853e–02	0.9568	0.584	6.480e–02

Table 2 shows that both the sectioning and batching CIs approach the nominal coverage $\beta = 0.95$ as n increases, which agrees with Theorem 9 for sectioning. However, at $n = 100$, the coverage for sectioning appears to be closer to nominal than for batching. As explained in Sections 3.3 and 5.2, sectioning can achieve better coverage than batching because the former centers its CI at a point estimator that often has lower bias. It is also interesting to note that for each n, the sectioning point estimator has lower sample variance than the batching one.

7 CONCLUDING REMARKS

The economic capital is a risk measure that can be used to help determine the amount of capital needed to protect (with high probability) against large unexpected losses of a credit portfolio. The EC is the difference between the p-quantile ξ and the mean μ of the loss distribution, where p is often chosen with $p \approx 1$, and estimating both ξ and μ using a single simulation can produce an EC estimator with large variance. We examined instead using independent simulations to estimate separately the two components, where we applied importance sampling for estimating ξ and simple random sampling for μ. Our numerical results show that when $p \approx 1$, measure-specific importance sampling can greatly reduce the asymptotic variance compared to estimating both ξ and μ with the same simulation method. We are currently investigating applying alternative VRTs to estimate the EC, as well as experimenting with other models.

ACKNOWLEDGMENTS

This work has been supported in part by the National Science Foundation under Grant No. CMMI-1537322. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Asmussen, S., and P. Glynn. 2007. *Stochastic Simulation: Algorithms and Analysis*. New York: Springer.

Bahadur, R. R. 1966. "A Note on Quantiles in Large Samples". *Annals of Mathematical Statistics* 37(3):577–580.

Bassamboo, A., S. Juneja, and A. Zeevi. 2008. "Portfolio Credit Risk with Extremal Dependence: Asymptotic Analysis and Efficient Simulation". *Operations Research* 56(3):593–606.

Billingsley, P. 1995. *Probability and Measure*. 3rd ed. New York: John Wiley and Sons.

Chu, F., and M. K. Nakayama. 2012. "Confidence Intervals for Quantiles When Applying Variance-Reduction Techniques". *ACM Transactions On Modeling and Computer Simulation* 22(2):10:1–10:25.

Deutsche Bank 2017. "Annual Report 2017". Technical report, Deutsche Bank, Frankfurt am Main, Germany.

Farinelli, S., and M. Shkolnikov. 2012. "Two Models of Stochastic Loss Given Default". *Journal of Credit Risk* 8(2):3–20.

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

Glasserman, P., and J. Li. 2005. "Importance Sampling for Portfolio Credit Risk". *Management Science* 51(11):1643–1656.

Glynn, P. W. 1996. "Importance Sampling for Monte Carlo Estimation of Quantiles". In *Mathematical Methods in Stochastic Simulation and Experimental Design: Proceedings of the 2nd St. Petersburg Workshop on Simulation*, edited by S. M. Ermakov and V. B. Melas, 180–185. St. Petersburg, Russia: Publishing House of St. Petersburg Univ.

Goyal, A., P. Heidelberger, and P. Shahabuddin. 1987. "Measure Specific Dynamic Importance Sampling for Availability Simulations". In *Proceedings of the 1987 Winter Simulation Conference*, edited by A. Thesen et al., 351–357. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Goyal, A., P. Shahabuddin, P. Heidelberger, V. Nicola, and P. W. Glynn. 1992. "A Unified Framework for Simulating Markovian Models of Highly Dependable Systems". *IEEE Transactions on Computers* C-41(1):36–51.

Hyndman, R. J., and Y. Fan. 1996. "Sample Quantiles in Statistical Packages". *American Statistician* 50(4):361–365.

Jorion, P. 2007. *Value at Risk: The New Benchmark for Managing Financial Risk*. 3rd ed. New York: McGraw-Hill.

Klaassen, P., and I. van Eeghen. 2009. *Economic Capital: How It Works, and What Every Manager Needs to Know*. Burlington, MA: Elsevier.

Levy, A., A. Kaplin, Q. Meng, and J. Zhang. 2013, August. "A Unified Approach to Accounting for Regulatory and Economic Capital". Technical Report 2013-01-08, Moody's Analytics.

Lütkebohmert, E. 2009. *Concentration Risk in Credit Portfolios*. Berlin: Springer.

McNeil, A. J., R. Frey, and P. Embrechts. 2015. *Quantitative Risk Management: Concepts, Techniques, Tools*. Revised ed. Princeton, New Jersey: Princeton University Press.

Nakayama, M. K. 2014. "Confidence Intervals Using Sectioning for Quantiles When Applying Variance-Reduction Techniques". *ACM Transactions on Modeling and Computer Simulation* 24(4):19:1–19:21.

Serfling, R. J. 1980. *Approximation Theorems of Mathematical Statistics*. New York: John Wiley and Sons.

Sun, L., and L. J. Hong. 2010. "Asymptotic Representations for Importance-Sampling Estimators of Value-at-Risk and Conditional Value-at-Risk". *Operations Research Letters* 38(4):246–251.

AUTHOR BIOGRAPHIES

ZACHARY KAPLAN is an undergraduate student of computer science and applied mathematics at the New Jersey Institute of Technology and a software engineering intern at Google. His email is ztk4@njit.edu.

YAJUAN LI is a Ph.D. student of computer science at the New Jersey Institute of Technology. She received an undergraduate degree in Information Management and Information Systems at Beijing Forestry University. Her email address is yl935@njit.edu.

MARVIN K. NAKAYAMA is a professor in the Department of Computer Science at the New Jersey Institute of Technology. He received an M.S. and Ph.D. in operations research from Stanford University and a B.A. in mathematics-computer science from U.C. San Diego. He is a recipient of a CAREER Award from the National Science Foundation, and a paper he co-authored received the Best Theoretical Paper Award for the 2014 Winter Simulation Conference. He is an associate editor for *ACM Transactions on Modeling and Computer Simulation*, and served as the simulation area editor for the *INFORMS Journal on Computing* from 2007–2016. His research interests include simulation, modeling, statistics, risk analysis, and energy. His email address is marvin@njit.edu.