Variance Reduction Method for a Least-Squares Monte Carlo Approach to the Calculation of Risk Measures

Hongjun Ha*

Department of Mathematics; Saint Joseph's University

5600 City Avenue, Philadelphia, PA 19131. USA

September 2018

Abstract

Estimation of risk measures at a risk horizon is a challenging problem due to difficulty in obtaining a distribution of a loss random variable. A simplified approach to an approximation of the loss random variable is the usage of the quadratic form of state variables that drive uncertainty. However, if the risk horizon is a relatively extended period, this technique may deliver a meager projection of a possible event. A natural extension is to approximate the loss random variable through higher degrees of polynomials and apply it to a generation of samples to form the distribution of losses. Like other simulation methods for the risk management, a required number of simulations is also significant for estimating a probability of a massive loss or capital requirements if we rely on the functional approximated loss distribution, which is the culprit of expensive computational costs. In this paper, we provide a simple and efficient algorithm to approximate the loss random variable defined at the future time and calculate a probability of a large loss based on the Least-Squares Monte Carlo method (LSM) and the Importance Sampling (IS). We test the algorithm with a variable annuity contract and verify that a significant variance reduction is achieved with a relatively small computational budget. In particular, the algorithm can be exceptionally efficient for estimating a probability of occurring Black swan events.

Keywords: loss distribution, probability of large loss, least-squares algorithm, basis functions, importance sampling.

*Corresponding author. Phone: +1-610-660-1562. E-mail addresses: hha@sju.edu (H. Ha).

1 Introduction

The experiences of the financial crisis make financial institutions recognize the importance of risk management. One of the fundamental components of management operations is to measure risk quantitatively. Famous risk measure is to calculate a probability of significant loss, Value-at-Risk or Expected Shortfall. For these risk measurements, it is required to have a loss distribution or at least a reasonable approximation of the loss distribution. However, financial institutions have difficulty calculating risk measures since there are many factors which the companies should consider when estimating a loss distribution. Therefore, it is essential to have a sound numerical method to approximate for measuring risk with acceptable computational cost. The most famous method is the usage of the Monte Carlo method to approximate a loss distribution.

The näive Monte Carlo method, however, may not be feasible due to the well-known nested structure of simulation scheme in many practical situations. A company should generate sample paths of specific state variables until a risk horizon. At each realization of state variables, it should generate a large number of simulations again to approximate values of losses. To avoid this troublesome, one can consider the usage of quadratic approximation, so-called Delta-Gamma approximation. In Glasserman et al. (2000), the authors assume that a loss random variable may be approximated by polynomials of the second degree and associated delta and gamma coefficients are known in advance. They provide the algorithm of the importance and stratified sampling to reduce a variance under the Gaussian setting. In particular, they show that the exponential tilting gives the optimal way to choose an alternative density function for the importance sampling. The crucial point to justify the Delta-Gamma approximation is that the risk horizon (Δt in Glasserman et al. (2000)), on which a loss random variable is defined, should not be a long period.

In insurance regulation such as Solvency II, however, an insurance company should report capital requirements with an assumption that the risk horizon is one year. So, we easily expect that the Delta-Gamma approximation does not produce a viable approximation for a loss random variable at a long risk horizon. Therefore, it is natural to extend the Delta-Gamma approximation to a more significant approximation by considering more polynomials with higher degrees. If one considers an approximation based on a more complex functional form, there are two issues to be addressed: i) how can we find coefficients associated to functions or estimate them? and ii) how can we use the approximation to find a risk measure with a small effort? The first question is answered by the Least-Squares Monte Carlo simulation (LSM) in Ha and Bauer (2018). The coefficients are estimated by the least-squares regression using data generated from the Monte Carlo simulation. After obtaining the coefficients, we use the approximation to calculate a risk measure. In this paper, we focus on estimating a probability of large loss, that is, we are interested mainly in estimating $P[L_\tau > x]$, where L_τ is the loss random variable at the risk horizon τ. If one uses this approximation to assess a risk measure, a number of simulations is usually significant. This simple exploitation of the approximated loss random variable causes computational cost impeding efficient risk management. To have sound results with an affordable cost, we employ the Importance Sampling (IS). The main task in the IS is to select an alternative probability measure $\tilde{P}$ to obtain a good approximation of $P[L_\tau > x]$ with a small effort. The hint to choose the best alternative density comes from that one minimize the second raw moment of the product of $1_{\{L>x\}}$ and likelihood ratio, where 1 is the indicator function. This idea is adapted in quantitative finance literatures (see e.g., Glasserman et al. (2000); Capriotti (2008)). In Glasserman et al. (2000), the optimization is achieved by the exponential tilting, and in Capriotti (2008), the author estimates the second raw moment by the Monte Carlo estimator and minimizes it by the least-squares method with sample realizations, so-called the least-squares importance sampling.

In this paper, we approximate the loss random variable by the LSM and employ the IS technique to improve the convergence of the probability of a large loss. For the importance sampling, we minimize an estimated first raw moment of the product of $1_{\{L>x\}}$ and the likelihood ratio under the physical measure. Our numerical tests show that a significant variance reduction is possible if one relies on the IS. In particular, if x is extremely large and optimization is implemented at x, the LSM-IS method is exceptionally efficient for estimating small probabilities.

The remainder of the paper is organized as follows: Section 2 introduces the model setting for approximating a loss random variable and the importance sampling. Section 3 presents more concrete explanations about the algorithm under the Gaussian setting. Section 4 provides our numerical results. And, finally, Section 5 concludes.

2 The LSM and IS Approach

In this paper, we employ two simulation approaches to estimate a probability of large loss. The first is the LSM method to approximate a loss random variable at the risk horizon. On the other hand, we discuss how to select an alternative probability measure $\tilde{\mathbb{P}}$ for the IS, which generates right sample paths to reduce the variance of an estimator.

2.1 LSM Approach

We refer Ha and Bauer (2018) for the specific model setting. The longest maturity of a company's liability is given by $T < \infty$. Let $(\Omega, \mathcal{F}, \mathbf{F} = (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a complete filtered probability space, where $\mathbb{P}$ denotes the physical measure. We assume that a regular Markov process $Y_t \in \mathbb{R}^d$ drives all relevant risk factors. In what follows, all random variables considered are square-integrable in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. A risk-neutral probability measure (equivalent martingale measure) $\mathbb{Q}$ equivalent to $\mathbb{P}$ exists under which payment streams can be valued as expected discounted cash flows with respect to a given numéraire process $(N_t)_{t\in[0,T]}$. Let L_τ be a loss random variable at the risk horizon τ, $0 < \tau < T$:

$$L_\tau = -(A_\tau(Y_\tau) - V_\tau(Y_\tau)),$$

where $A_\tau(Y_\tau)$ is a price of the asset the company has at time τ, and $V_\tau(Y_\tau)$ is the market value of liabilities of the company at time τ, respectively.

The primary problem is to estimate a probability that the company's loss is larger than x at time τ. In the probabilistic statement, we focus on calculating

$$P[L_\tau > x]. \tag{1}$$

The first step to calculate (1) is to obtain a distribution of L_τ. In general, it is not feasible to recover the distribution of L_τ analytically because of complex profiles of payoffs, multiple risk factors, and illiquidity of liabilities. The only feasible method is the Monte Carlo method to approximate L_τ. However, a simple application of the Monte Carlo method to the valuation problem at the future time causes another unwelcome problem, so-called the nested Monte Carlo simulation implementation; The company generates a large number of simulation for Y_τ under $\mathbb{P}$ and generates a large number of simulations at each realization of Y_τ untile time T under $\mathbb{Q}$, which is impracticable in many cases. The need for a doable algorithm with small efforts arises.

The LSM approach relies on two sequential approximation. L_τ is a function of random variable Y_τ in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. It is important to note that the function space $L^2(\Omega, \mathcal{F}, \mathbb{P})$ is separable. Therefore, there is an at least one set of complete orthonormal functions, $\{\varphi_j\}_{j\geq 1}$, such that

$$f(Y_\tau) = \sum_{j=1}^{\infty} < f, \varphi_j > \varphi_j(Y_\tau),$$

where f is any function in $L^2(\Omega, \mathcal{F}, \mathbb{P})$ and $< \cdot, \cdot >$ is a scalar product with respect to $\mathbb{P}$. That is, $\{\varphi_j\}_{j\geq 1}$ is the basis for $L^2(\Omega, \mathcal{F}, \mathbb{P})$. If one uses enough number of orthonormal basis functions, it is possible to approximate the loss random variable L_τ suitably. More precisely, L_τ is replaced with a finite linear combination by a set of M basis functions:

$$L_\tau \approx \tilde{L}_\tau = \sum_{j=1}^{M} \alpha_j \varphi_j(Y_\tau), \tag{2}$$

where $\alpha_j =< L_\tau, \varphi_j >$ is the jth coefficients for the jth basis function $\varphi_j(\cdot)$. The most famous basis functions are polynomials such as Hermite polynomials, Legendre polynomials, Laguerre polynomials, power monomials, etc. In Glasserman et al. (2000), the loss random variable is approximated by the Delta-Gamma approximation. If we map the Delta-Gamma approximation to the idea in (2), the basis functions are monomials (not orthonormalized) of the second degree, and M depends on d. The justification of the Delta-Gamma approximation is valid when τ is small. The necessity of large M due to a higher degree is an important issue in the insurance industry. In financial literature, normal values of τ are one day or one week. In insurance regulations such as Solvency II, the risk horizon is usually one-year, and it is easy to expect that the second-order approximation is not enough to capture the complicated dynamics of L_τ. So, a natural extension from the Delta-Gamma approximation is to consider basis functions of higher degree, which can represent the complexity of non-linear behaviors of $L\tau$ at time τ. In this paper, we consider an approximation problem for L_τ when relying on an arbitrary M. From now on, we assume that the first basis function is constant, that is, $\varphi_1(Y_\tau) = 1$. If we use the functional approximation in (2), we have the following basic characteristics:

Theorem 2.1. *The (unconditional) expectation and variance of $\tilde{L}_\tau$ are α_1 and $\sigma^2_{\tilde{L}_\tau} = \sum_{j=2}^{M} \alpha_j^2$, respectively.*

Proof of Theorem 2.1. Note that the set of orthonormal basis functions include 1 as the first basis function. Moreover, the orthogonality gives

$$\mathbb{E}^{\mathbb{P}}[\tilde{L}_\tau] = \mathbb{E}^{\mathbb{P}}\left[\sum_{j=1}^{M}\alpha_j\varphi_j(Y_\tau)\right] = \alpha_1 + \sum_{j=2}^{M}\alpha_j\mathbb{E}^{\mathbb{P}}\left[1\times\varphi_j(Y_\tau)\right] = \alpha_1.$$

The variance of L_τ is obtained from orthonormality of basis functions:

$$\begin{aligned} Var^{\mathbb{P}}[\tilde{L}_\tau] &= Var^{\mathbb{P}}\left[\sum_{j=1}^{M}\alpha_j\varphi_j(Y_\tau)\right] = Var^{\mathbb{P}}\left[\sum_{j=2}^{M}\alpha_j\varphi_j(Y_\tau)\right] \\ &= \sum_{j=2}^{M}\alpha_j^2\mathbb{E}^{\mathbb{P}}\left[(\varphi_j(Y_\tau))^2\right] = \sum_{j=2}^{M}\alpha_j^2. \end{aligned}$$

□

In practice, it is unclear how to choose a good x when calculating $Pr[L_\tau > x]$. Based on Theorem 2.1, one can guess a decent x for risk management purpose, for instance, $x = a_1 + \rho\sigma_{\tilde{L}_\tau}$, where $\rho > 0$.

The second approximation of the LSM algorithm is the employment of the least-squares regression to estimate $\alpha = (\alpha_1, ..., \alpha_M)'$. For the convenience, suppose that the assets and liabilities generate cash flows at a discrete time $\tau+1, ...,$ and T. The LSM approach generate N sample paths of $Y_t^{(i)}$ $0 < t \le \tau$,$i = 1, ..., N$ under the physical measure and generate one sample path of $Y_t^{(i)}$ $\tau < t \le T$ under the risk neutral measure at each $Y_\tau^{(i)}$. Based on the realization of $Y_t^{(i)}$, $\tau < t \le T$, we compute a noisy estimator $\hat{V}_\tau^{(i)}$ for $L_\tau(Y_\tau^{(i)})$ using the numéraire process. We costruct a design matrix D:

$$D = \begin{bmatrix} \varphi_1(Y_\tau^{(1)}) & \dots & \varphi_M(Y_\tau^{(1)}) \\ \vdots & \ddots & \vdots \\ \varphi_1(Y_\tau^{(N)}) & \dots & \varphi_M(Y_\tau^{(N)}) \end{bmatrix}.$$

The least-squares estimator $\hat{\alpha}$ is obtained by regressing $\hat{V}_\tau$ on D:

$$\hat{\alpha} = (D'D)^{-1}D'\hat{V}_\tau \qquad (3)$$

where $\hat{V}_\tau = (\hat{V}_\tau^{(1)}, ..., \hat{V}_\tau^{(N)})'$. The needed assumptions for $\hat{\alpha}$ to be the Best Linear Unbiased Estimator (BLUE) are satisfied with the nature of Monte Carlo simulation and the assumption on the considered function space:

Lemma 2.1. *The least-sqaures regression estimator for α in (2) produces the BLUE.*

Proof of Lemma 2.1. The linearity in α is clearly satisfied in (2), and randomness of sample is automatically satisfied from the nature of the Monte Carlo simulation. Orthogonality of basis functions rules out the perfect collinearity. The omitted basis functions form the error:

$$L_\tau(y) = \sum_{j=1}^{M}\alpha_j\varphi_j(y) + u,$$

where $u = \sum_{j=M+1}^{\infty} \alpha_j \varphi_j(y)$. Note that

$$\mathbb{E}^{\mathbb{P}}[u|\varphi_j(y), j = 1, ..., M] = 0,$$

since non constant basis functions are orthogonal with 1. We know that

$$Var^{\mathbb{P}}[u|\varphi_j(y), j = 1, ..., M] = \sum_{j=M+1}^{\infty} \alpha_j^2 = \sigma_u^2,$$

which is a finite constant since basis functions are orthornormal and L_τ is square-integrable. Therefore, the least-squares regression for α produces the BLUE. □

The final approximation $\hat{L}_\tau$ for L_τ is given by

$$L_\tau \approx \tilde{L}\tau \approx \hat{L}_\tau = \sum_{j=1}^{M} \hat{\alpha}_j \varphi_j(Y_\tau). \tag{4}$$

The detailed algorithm to obtain (2) is provided in Algorithm 1

Result: Functional Approximation for L_τ

1. Set an orthonormal basis function $\{\varphi_j\}_{\jmath \geq 1}$ and fix M and N.;
2. For $i = 1, ..., N$,
   (a) Generate $Y_t^{(i)}$, $0 < t \leq \tau$ under $\mathbb{P}$.
   (b) For each i, generate one sample path of $Y_t^{(i)}$, under $\mathbb{Q}$.
   (c) Cacluate $\hat{V}_\tau^{(i)}$.
   (d) Construct the matrix D.
3. Run the least-squares regression $\hat{V}_\tau$ on D and save the estimated coefficient as $\hat{\alpha}$.

Algorithm 1: Approximation of L_τ through the LSM

If a good approximation for L_τ is available from Algorithm 1, we obtain the distribution of $\hat{L}_\tau$ easily by plugging $Y_\tau^{(i)}$ into (4). This distribution, in turn, serves as the approximated distribution for L_τ.

It is not necessary to generate a high volume of Y_τ to estimate α. However, a lot of sample paths of Y_τ is still required to estimate $Pr[L_\tau > x] \approx Pr[\hat{L}_\tau > x]$ for large x. In Section 2.2, we discuss the efficient algorithm to obtain the probability of large loss via the IS technique.

2.2 Combination of Important Sampling and the LSM

In this Section, we assume that the reliable approximation is obtained through Algorithm 1 and focus on estimating $Pr[\hat{L}_\tau > x]$ via the IS technique. The basic idea of the IS is to draw

Y_τ from a new distribution which reduce a variance of an estimator. Note that

$$
\begin{aligned}
Pr[\hat{L}_\tau > x] &= \mathbb{E}^{\mathbb{P}}\left[1_{\{\hat{L}_\tau > x\}}\right] = \int 1_{\{\hat{L}_\tau(y) > x\}} d\mathbb{P}(y), \\
&= \int 1_{\{\hat{L}_\tau(y) > x\}} \frac{d\mathbb{P}(y)}{d\tilde{\mathbb{P}}_\theta(y)} d\tilde{\mathbb{P}}_\theta(y) = \mathbb{E}^{\tilde{\mathbb{P}}_\theta}\left[1_{\{\hat{L}_\tau > x\}} l_\theta\right], \qquad (5)
\end{aligned}
$$

where $\tilde{\mathbb{P}}_\theta$ is an importance sampling measure, θ is a parameter vector for $\tilde{\mathbb{P}}_\theta$ and $l_\theta = \frac{d\mathbb{P}(y)}{d\tilde{\mathbb{P}}_\theta(y)}$ is the likelihood ratio. The last expectation in (5) shows that the probability can be approximated via the Monte Carlo simulation where the state variable Y_τ follows the distribution $\tilde{\mathbb{P}}_\theta$.

Analytical expression for the considered probability is too seldom found in the practice due to difficulty of obtaining a distribution function of $\hat{L}_\tau$. The Monte Carlo simulation again plays a important role in estimating a probability again:

$$
\mathbb{E}^{\tilde{\mathbb{P}}_\theta}\left[1_{\{\hat{L}_\tau > x\}} l_\theta\right] \approx \frac{1}{N_{IS}} \sum_{i=1}^{N_{IS}} 1_{\{\hat{L}_\tau(Y_\tau^{(i)}) > x\}}, \qquad (6)
$$

where N_{IS} is a number of simulations for the IS and $Y_\tau^{(i)}$ is a realized sample drawn from $\tilde{\mathbb{P}}_\theta$. The problem at this stage is what distribution $\tilde{\mathbb{P}}_\theta$ is incorporated with (6), which reduces a variability. If we fix a family of distributions, the problem amounts to finding θ. From now on, we assume that a family of distribution $\tilde{\mathbb{P}}_\theta$ is fixed, and we are interested in determining θ for the efficiency of the IS. In Glasserman et al. (2000), the authors provide the method which minimizes the second raw moment of $1_{\{\hat{L}_\tau > x\}} l_\theta$ asymptotically by minimizing the upper limit of $\mathbb{E}^{\mathbb{P}}\left[1_{\{\hat{L}_\tau > x\}} l_\theta\right]$ under the normality of Y_τ. It is required that one should know the probability distribution $\hat{L}_\tau$ to find a good θ analytically, which is not feasible if the approximation depends on the complex basis function. In Capriotti (2008), the author numerically minimizes the second raw moment of $1_{\{\hat{L}_\tau > x\}} l_\theta$ using the Monte Carlo simulation and the least-squares algorithm.

In this paper, we search θ by employing Capriotti (2008)'s idea. Note that the second raw moment, $m_2(\theta)$, of $1_{\{\hat{L}_\tau > x\}} l_\theta$ under $\tilde{\mathbb{P}}_\theta$ is

$$
m_2(\theta) = \mathbb{E}^{\tilde{\mathbb{P}}_\theta}\left[1_{\{\hat{L}_\tau > x\}} l_\theta^2\right] = \mathbb{E}^{\mathbb{P}}\left[1_{\{\hat{L}_\tau > x\}} l_\theta\right]. \qquad (7)
$$

The last expectation is approximated by the Monte Carlo simulation:

$$
\begin{aligned}
m_2(\theta) &= \mathbb{E}^{\mathbb{P}}\left[1_{\{\hat{L}_\tau > x\}} l_\theta\right] \\
&\approx \frac{1}{N_O} \sum_{k=1}^{N_O} 1_{\{\hat{L}_\tau(Y_\tau^{(i)}) > x\}} l_\theta(Y_\tau^{(i)}), \qquad (8)
\end{aligned}
$$

where N_O is a number of simulations and $Y_\tau^{(i)}$ is a realized sample of Y_τ under $\mathbb{P}$. The optimal θ is determined by minimizing (8):

$$
\theta^* = \text{argmin}_\theta \frac{1}{N_O} \sum_{k=1}^{N_O} 1_{\{\hat{L}_\tau(Y_\tau^{(i)}) > x\}} l_\theta(Y_\tau^{(i)}). \qquad (9)
$$

Usually, the optimization is done numerically. The final estimate for $Pr[\hat{L}_\tau > x]$ is obtained through Algorithm 2.

Result: Approximation for $Pr[\hat{L}_\tau > x]$

1. Part 1
   (a) Generate N_O sample paths of Y_τ under $\mathbb{P}$.
   (b) Calculate $\hat{m}_2(\theta) = N_O^{-1} \sum_{k=1}^{N_O} 1_{\{\hat{L}_\tau(Y_\tau^{(i)})>x\}} l_\theta(Y_\tau^{(i)})$.
   (c) Find θ^* by minimizing $\hat{m}_2(\theta)$.
2. Part 2
   (a) Generate N_{IS} samples of Y_τ under $\tilde{\mathbb{P}}_{\theta^*}$.
   (b) Calculate $N_{IS}^{-1} \sum_{i=1}^{N_{IS}} 1_{\{\hat{L}_\tau(Y_\tau^{(i)})>x\}} l_{\theta^*}(Y_\tau^{(i)})$ and return it.

Algorithm 2: Approximation of $Pr[\hat{L}_\tau > x]$ through the IS

Note that Algorithm 1 and Part 1 of Algorithm 2 can be implemented simultaneously by setting $N_O = N$. That is, the sample paths of Y_τ under Algorithm 1 can be used for Part 1 of Algorithm 2, and this enhances computational efficiency. A required number of simulations for Algorithm 1 and Part 1 of Algorithm 2 is not formidable. Moreover, a required number of simulation for Part 2 of Algorithm 2 should not be significant. Therefore, the combination of Algorithm 1 and 2 can deliver the efficient method for estimating a probability of large loss. In Section 3, we provide concrete application of the Algorithms under the Gaussian setting.

3 Gaussian Setting

Suppose that the state variable Y_t, $t > 0$ is normally distributed with a mean vector μ_t and a variance-covariance matrix Σ_t. So, all cash-flows generated from assets and liabilities at each time t in the future are functions of the multivariate normal random variable. Under this setting, the function $L^2(\mathbb{R}^d, \mathcal{B}^d, \mathbb{P})$, where $\mathcal{B}^d$ is a Borel sigma-algebra on $\mathbb{R}^d$, is separable and admits a set of complete orthonormal basis functions for the space. In particular, the complete set of orthonormal basis functions is products of a univariate normalized Hermite polynomials with respect the multivariate standard normal distribution (Khare and Zhou, 2009; Rahman1, 2017). That is, L_τ is expressed by

$$L_\tau(y) = \sum_{j=1}^{\infty} < L_\tau, h_j^{|n|} > h_j^{|n|}(y), \tag{10}$$

$$h_j^{|n|}(y) = \Pi_{s=1}^{d} h_{n_s}(z_s(y)), \quad |n| = \sum_{s=1}^{d} n_s, \tag{11}$$

$$z(y) = (z_1(y), ..., z_d(y))' = \Sigma_\tau^{-\frac{1}{2}}(y - \mu_\tau), \tag{12}$$

where $n_s \in \mathbb{N}_0$, $\mathbb{N}_0$ is the set of non-negative integer and h_k is a univariate normalizaed Hermite polynomial with degree k with the following recursion:

$$h_0(x) = 1,\ h_1(x) = x,\ h_k(x) = \frac{1}{\sqrt{k}}\left(x\, h_{k-1}(x) - \sqrt{k-1}h_{k-2}(x)\right), \quad j = 2, 3, ...$$

If we use M basis function for the approximation,

$$L_\tau(y) \approx \tilde{L}_\tau(y) = \sum_{j=1}^{M} \alpha_j h_j^{|n|}(y), \tag{13}$$

where $\alpha_j =< L_\tau, h_j^{|n|} >$ and a scalar product is computed under $N(0, I_d)$. In Ha and Bauer (2018), the authors provide the systematic way to arrange basis functions in an optimal manner to minimize the functional error of (13) based on the spectral value decomposition (SVD) of the compact operator if there is one cash flow at time T. In this paper, we assume that the choice of M does not cause a notable cost to a certain degree and the M is large enough to capture a non-linear dynamics of L_τ.

Remark: In Ha and Bauer (2018), the authors provide the systematic method to arrange the multivariate Hermite polynomials optimally under a fixed M. In general, there does not exit how to choose an appropriate M for a functional approximation. One simple rule which one can consider is to take advantage of Theorem 2.1. Let $\hat{\sigma}_{\hat{L}_\tau,M}$ be an estimated standard deviation of $\sigma_{\hat{L}_\tau,M}$, which is easily obtained from the usual least-squares regression. This estimate quantifies the functional error from omitting basis functions over M. If we increase one more basis function under the same realization of Y_τ, we have $\sigma_{\hat{L}_\tau,M+1}$ and calculate

$$c(M, M+1) = \frac{\sigma_{\hat{L}_\tau,M} - \sigma_{\hat{L}_\tau,M+1}}{\sigma_{\hat{L}_\tau,M}}.$$

If $c(M, M+1) < \epsilon$, where ϵ is a small positive number, we stop increasing M.

Next, we move to the problem on finding θ^*. We only focus on an another multivatiate normal density function for the IS, that is, $\theta = (\mu, \Sigma)$. We want to recover (μ^*, Σ^*) such that

$$(\mu^*, \Sigma^*) = \text{argmin}_\theta \frac{1}{N_O} \sum_{k=1}^{N_O} 1_{\{\hat{L}_\tau(Z_i) > x\}} l_{(\mu,\Sigma)}(Z_i), \tag{14}$$

where $l_{(\mu,\Sigma)}(Z) = f(Z|0,I_d)/f(Z|\mu,\Sigma)$, $Z \sim N(0,I_d)$, and $f(Z|\mu,\Sigma)$ is a density function of d-dimensional normal random variable with a mean vector μ and a variance-cpvariance matrix Σ. Detailed algorithm for estimating $Pr[\hat{L}_\tau > x]$ is provided in Algorithm 3.

Result: Approximation for $Pr[\hat{L}_\tau > x]$ under the Gaussian setting

1. Part 1
   (a) Set M and N.
   (b) Generate sample paths of $Y_t^{(i)}$ $0 < t \leq \tau$ $i = 1, ..., N$ under $\mathbb{P}$.
   (c) For each $Y_\tau^{(i)}$, generate one sample path of $Y_t^{(i)}$, $\tau < t \leq T$ under $\mathbb{Q}$.
   (d) Obtain $V_\tau^{(i)}$, $i = 1, ..., N$.
   (e) Obtain $\hat{L}_\tau$ via Hermite basis functions and the least-squares regression.
   (f) Search (μ^*, Σ^*) by minimizing $N^{-1}\sum_{k=1}^{N} 1_{\{\hat{L}_\tau(Z_i)>x\}} l_{(\mu,\Sigma)}(Z_i)$, where $Z_i = \Sigma^{-\frac{1}{2}}(Y_\tau^{(i)} - \mu_\tau)$
2. Part 2
   (a) Set x.
   (b) Generate N_{IS} normal random numbers under $N(\mu^*, \Sigma^*)$.
   (c) Calculate $N_{IS}^{-1}\sum_{i=1}^{N_{IS}} 1_{\{\hat{L}_\tau(Z_i)>x\}} l_{(\mu^*,\Sigma^*)}(Z_i)$ and return it.

Algorithm 3: Approximation of $Pr[\hat{L}_\tau > x]$ through the Gaussian setting

4 Numerical Illustration

We consider a Variable Annuity (VA) plus guaranteed minimum income benefit (GMIB) in Ha and Bauer (2018). At maturity T, the policyholder has the right to choose between a lump sum payment amounting to the current account value or a guaranteed annuity payment b determined as a guaranteed rate applied to a guaranteed amount. GMIBs are famous riders for VA contracts: Between 2011 and 2013, roughly 15% of the more than \$150 billion worth of Variable Annuities sold in the US contained a GMIB (LIMRA). Importantly, GMIBs are subject to a variety of risk factors, including a fund (investment) risk, mortality risk, and – as long-term contracts – interest rate risk. Consequently, we consider its risk and valuation in a multivariate Markov setting for these three risk factors.

4.1 Model and Payoff of the GMIB

As in the previous section, we consider a large portfolio of GMIBs with policyholder age x, policy maturity T, and a fixed guaranteed amount – so that the guaranteed annuity payment

b is fixed at time zero.[1] The payoff of the VA plus GMIB at T in case of survival is given by:

$$\max \{S_T,\ b\, a_{x+T}(T)\}, \tag{15}$$

where S_T is the underlying account value which evolves according to a reference asset net various fees (which we ignore for simplicity).

We consider a three-dimensional state process Y_t governing financial and biometric risks:

$$Y_t = (q_t,\, r_t,\, \mu_{x+t})',$$

where q_t denotes the log-price of the risky asset at time t, r_t is the short rate, and μ_{x+t} is the force of mortality of an $(x+t)$-aged person at time t. We assume Y_t satisfies the following stochastic differential equations under $\mathbb{P}$:

$$dq_t = \left(m - \frac{1}{2}\sigma_S^2\right) dt + \sigma_S\, dW_t^S, \tag{16}$$

$$dr_t = \alpha(\gamma - r_t)\, dt + \sigma_r\, dW_t^r, \tag{17}$$

$$d\mu_{x+t} = \kappa\mu_{x+t}\, dt + \psi\, dW_t^\mu, \tag{18}$$

where m is the instantaneous rate of return of the risk asset, σ_S is the asset volatility, κ is an instantaneous rate of increment of mortality (Gompertz exponent), ψ is the volatility of mortality, and W_t^S, W_t^r, and W_t^μ are standard Brownian motions under $\mathbb{P}$ with $dW_t^S\, dW_t^r = \rho_{12}\, dt$, $dW_t^S\, dW_t^\mu = dW_t^r\, dW_t^\mu = 0$, i.e. we assume independence of financial and biometric risks. Note that the solutions to the above stochastic differential equations at time t are Normal distributed so that we can use Hermite polynomials using the approach in Section 3.

The dynamics of Y_t under the risk-neutral measure $\mathbb{Q}$ are given by:

$$dq_t = \left(r_t - \frac{1}{2}\sigma_S^2\right) dt + \sigma_S\, d\tilde{W}_t^S,$$

$$dr_t = \alpha(\bar{\gamma} - r_t)\, dt + \sigma_r\, d\tilde{W}_t^r,$$

$$d\mu_{x+t} = \kappa\mu_{x+t}\, dt + \psi\, d\tilde{W}_t^\mu,$$

where $\tilde{W}_t^S$, $\tilde{W}_t^r$ and $\tilde{W}_t^\mu$ are standard Brownian motions under $\mathbb{Q}$ with the same correlation coefficients. Here, for simplicity and without loss of generality, we assume that there is no risk premium for mortality risk. Since the force of mortality is stochastic, the k-year survival probability ${}_kp_{x+t}$ is given by:

$${}_kp_{x+t} = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^k \mu_{x+t+s}\, ds} | Y_t\right],$$

and the at time-t-value of the VA plus GMIB contract is:

$$V(t) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s\, ds} e^{-\int_0^{T-t} \mu_{x+t+y}\, dy} \max\{e^{q_T},\ b\, a_{x+T}(T)\}\, |Y_t\right]. \tag{19}$$

[1] Some contract variants include path-dependent features such as ratchet guarantees (**?**).

Since it is not possible to obtain an analytical expression for the GMIB, particularly when considering additional features such as step ups or ratchets, it is necessary to rely on numerical methods for valuation and estimating risk capital. To directly apply our LSM framework, we adjust the presentation by changing the numéraire to a pure endowment with maturity T and maturity value one. The price of GMIB at time k using the pure endowment as the numéraire is:

$$V(t) = {}_{T-t}E_{x+t}\, \mathbb{E}^{\mathbb{Q}_E}\left[\max\left\{e^{q_T},\ b\, a_{x+T}(T)\right\}|Y_t\right], \tag{20}$$

where $\tau \leq t \leq T$, ${}_{T-t}E_{x+t}$ is the price of the pure endowment contract at time t, and $\mathbb{Q}_E$ is the risk-neutral measure using the pure endowment contract as the numéraire.

Under our assumption of independence between financial and biometric risk, we obtain:

$$\begin{aligned}
{}_{T-t}E_{x+t} &= \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s\, ds} e^{-\int_0^{T-t} \mu_{x+t+y}\, dy} |Y_t\right] = p(t,T)\ \times\ {}_{T-t}p_{x+t} \\
&= A_r(t,T) \exp\left(-r_t\, B_r(t,T)\right)\ A_\mu(t,T) \exp\left(-\mu_{x+t}\, B_\mu(t,T)\right)
\end{aligned}$$

since (r_t) and (μ_t) are affine with

$$\begin{aligned}
B_r(t,T) &= \frac{1-e^{-\alpha(T-t)}}{\alpha}, \quad A_r(t,T) = \exp\left\{\left(\bar{\gamma} - \frac{\sigma_r^2}{2\alpha^2}\right)\left(B_r(t,T) - T + t\right) - \frac{\sigma_r^2}{4\alpha} B_r^2(t,T)\right\}, \\
B_\mu(t,T) &= \frac{1-e^{\kappa(T-t)}}{\kappa}, \quad A_\mu(t,T) = \exp\left\{\frac{\psi^2}{2\kappa^2}\left(B_m u(t,T) + T - t\right) + \frac{\psi^2}{4\kappa} B_\mu(t,T)^2\right\}.
\end{aligned}$$

Thus, applying Itô's formula, the dynamics of the pure endowment price are:

$$\begin{aligned}
d_{T-t}E_{x+t} &= dp \times {}_{T-t}p_{x+t} + p \times d({}_{T-t}p_{x+t}) \\
&= {}_{T-t}E_{x+t}\left[(r_t + \mu_{x+t})dt - \sigma_r B_r(t,T) d\tilde{W}_t^t - \psi B_\mu(t,T) d\tilde{W}_t^\mu\right],
\end{aligned}$$

and from Brigo and Mercurio (2006), the new dynamics of Y_t under $\mathbb{Q}_E$ for $\tau \leq t \leq T$ become:

$$dq_t = \left(r_t - \frac{1}{2}\sigma_S^2 - \rho_{12}\, \sigma_s\, \sigma_r B_r(t,T)\right) dt + \sigma_S\, dZ_t^S, \tag{21}$$

$$dr_t = \alpha(\bar{\gamma} - \sigma_r^2 B_r(t,T)/\alpha - r_t)\, dt + \sigma_r\, dZ_t^r, \tag{22}$$

$$d\mu_{x+t} = \kappa(\mu_{x+t} - \psi^2 B_\mu(t,T)/\kappa)\, dt + \psi\, dZ_t^\mu, \tag{23}$$

where Z_t^S, Z_t^r, and Z_t^μ are standard Brownian motions under $\mathbb{Q}_E$ with $dZ_t^S dZ_t^r = \rho_{12} dt$, $dZ_t^S dZ_t^\mu = 0$, and $dZ_t^r dZ_t^\mu = 0$.

We ignore the asset side in the calculation of the risk capital for the VA plus GMIB contract, and estimate the risk measure $Pr[L_\tau > x]$ via the LSM and IS algorithm. In particular, the cash flow functional in the current setting is occurred at time T only with $\hat{V}_\tau$ without superscript

$$\hat{V}_\tau = {}_{T-\tau}E_{x+\tau} \max\{e^{q_T}, b a_{x+T}(T)\}.$$

Nest, we fix M and determine the order of basis functions, we apply Algorithm 1 and 2 to estimation of a probability of a large loss.

x	LSM	LSM-IS	VR
120	0.0738	0.0730	11.29
125	0.0386	0.0383	12.99
130	0.0189	0.0191	11.50
135	0.0094	0.0093	9.08
140	0.0041	0.0044	6.14
145	0.0020	0.0019	5.94

Table 1: Estimation of $Pr[\hat{L}_\tau > x]$. For probabilities via the LSM and optimal parameter estimation, the number of simulations is fixed by $N = 200,000$. The number of simulations for the IS is fixed by $N_{IS} = 100,000$. (μ^*, Σ^*) is estimated at $x = 120$.

4.2 Numerical Results

As in the previous application, we set the model parameters using representative values. The initial price of the risky asset is one hundred – so $q_0 = 4.605$ – and for the risky asset parameters we assume $m = 0.05$ (instantaneous rate of return) and $\sigma_S = 20\%$ (asset volatility). The initial interest rate is assumed to be $r_0 = 2\%$, $\alpha = 20\%$ (speed of mean reversion), $\gamma = 2.5\%$ (mean reversion level), $\sigma_r = 1\%$ (interest rate volatility), $\lambda = 2\%$ (market price of risk), and $\rho_{12} = -30\%$ (correlation between asset and interest rate). For the mortality rate, $x = 55$ (age of the policyholder), $\mu_{55} = 1\%$ (initial value of mortality), $\kappa = 10\%$ (instantaneous rate of increment), and $\psi = 0.03\%$ (mortality volatility) are assumed. For the insurance contract, we let the maturity $T = 15$, and the guaranteed annuity payout $b = 16.81$[2] per year. We set the risk horizon $\tau = 1$ as in the previous application.

We consider Hermite polynomials ararangement expained in Ha and Bauer (2018) to minimize a functional error. Separately, one can determine the number of basis functions with the preliminary remark, but it may be inefficient. We implement the LSM-IS approximation to the probability estimation problem with $M = 10$. To assess a efficiency of LSM-IS algorithm, we define a relative variance ratio:

$$\text{VR} = \frac{\text{Var of LSM estimator}}{\text{Var of LSM-IS estimator}}.$$

In Table 1, we verify that the LSM and LSM-IS methods produce similar results for the probabilities. However, the LSM-IS performs better than the LSM concerning the efficiency. In particular, although the number of simulations for the IS is smaller than the number of simulation for the LSM to estimate probabilities, the relative efficiency from the LSM-IS outperforms over the LSM. The estimated parameter at $x = 120$ works on a wide range of values.

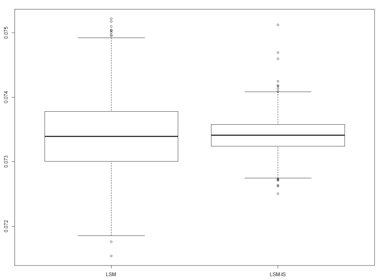
[2]Here, b is determined by:

$$b = \frac{S_0(1+m_g)^T}{a^*_{x+T}} = \frac{S_0(1+m_g)^T}{\left(\sum_{k=1}^{\infty} {}_{T+k}p_x\, p(0,T+k)\right) / \left({}_T p_x\, p(0,T)\right)},$$

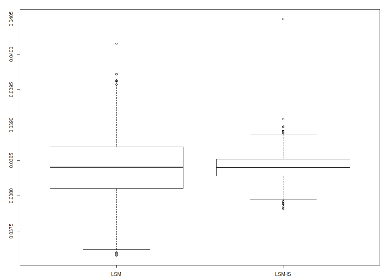
where m_g is the guaranteed rate of return and a^*_{x+T} is actuarial present value based on forward rates. We set $m_g = 2\%$. With above parameter values, the probability that $S_T > b\, a_{x+T}(T)$ is estimated 51.42% via Monte Carlo simulation.

The efficiency of the LSM-IS algorithm can be verifiable through graphical analysis. Figure 1 shows the distributions of $\widehat{Pr}[\hat{L}_\tau > x]$ at various x. All distributions generated from the LSM-IS are less volatile compared with the distributions of probabilities from the LSM.

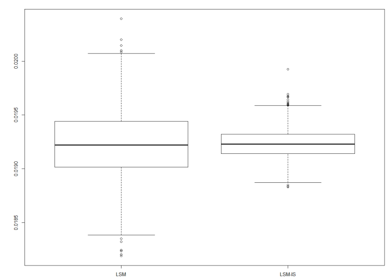
The LSM-IS algorithm delivers enormous improvement when x is large and (μ, Σ) is optimized at x. As seen in Figure 2, the ratio of two variances of distributions is large in Figure 2a, and the LSM-IS algorithm produces very efficient results around x in Figure 2b. It seems to be advisable to use the LSM-IS when we care about the substantial value of x.



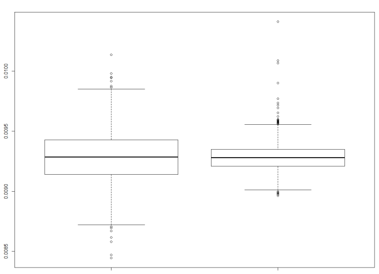
(a) $x = 120$, $\kappa_x = 4.83$



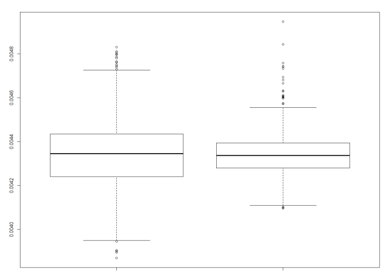
(b) $x = 125$, $\kappa_x = 5.68$



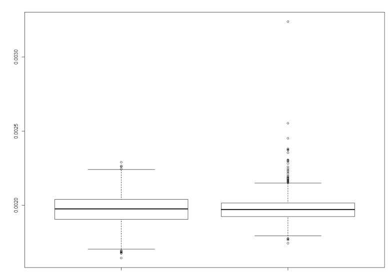
(c) $x = 130$, $\kappa_x = 5.75$



(d) $x = 135$, $\kappa_x = 3.67$

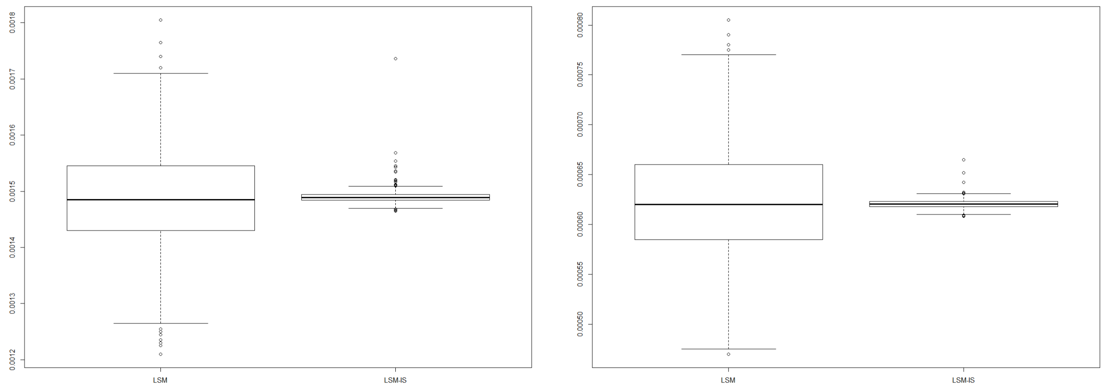


(e) $x = 140$, $\kappa_x = 2.94$



(f) $x = 145$, $\kappa_x = 1.52$

Figure 1: Box-and-whisker diagrams for $Pr[\hat{L}_\tau > x]$ based on two-thousand implementations of the LSM and the LSM-IS algorithms. The number of simulations for the LSM and optimization is $N = 200,000$, and the number of simulation for the IS is $N_{IS} = 100,000$. (μ^*, Σ^*) is estimated at $x = 120$. We report the ratio, κ_x, of the variance of distribution from the LSM probabilities to the variance of distribution from the LSM-IS probabilities at each x.



(a) $x = 145$, $\kappa_x = 73.30$ (b) $x = 150$, $\kappa_x = 193.93$

Figure 2: Box-and-whisker diagrams for $Pr[\hat{L}_\tau > x]$ based on two-thousand implementations of the LSM and the LSM-IS algorithms. The number of simulations for the LSM and optimization is $N = 200,000$, and the number of simulation for the IS is $N_{IS} = 100,000$. (μ^*, Σ^*) is estimated at $x = 145$. We report the ratio, κ_x, of the variance of distribution from the LSM probabilities to the variance of distribution from the LSM-IS probabilities at each x.

5 Conclusion

We propose a simple algorithm combining the LSM and IS for estimating risk measures in an arbitrary degree approximation for L_τ, which delivers reliable results with a relatively small computational effort. In particular, we show how to choose an IS density numerically, which is essential when one employs a higher order approximation for L_τ, where τ is a long period. Moreover, the desirable properties of the approximation can be derived if a set of orthonormal basis functions is incorporated.

Our numerical illustrations show that the algorithm can deliver superior results with small computational costs comparing with a basic Monte Carlo simulation. Specifically, the LSM-IS algorithm is exceptionally efficient for calculating a probability that the loss exceeds a high threshold. Therefore, we expect that the algorithm serves as the primary tool for calculating a probability of a substantial loss. Moreover, it can be used for calculating other risk measures such as Value-at-Risk by some trials for searching x which gives a desired significant level.

In this paper, we focus on the Gaussian setting where all quantities such as basis functions and an alternative density function are easily obtainable. In the further researches, it is necessary to consider a multivariate distribution which has a different tail behavior from the normal distribution for a more realistic description of the market situations.

References

Amemiya, T., 1985. Advanced econometrics. Harvard university press, Cambridge.

Brigo, D., Mercurio, F., 2006. Interest rate models – theory and practice. Springer.

Capriotti, L., 2008. Least-Squares Importance Sampling for Monte Carlo security pricing. Quantitative Finance, 5: 485-497.

Glasserman, P., Heiderlberger, P., Shahabuddin, P., 2000. Variance Reduction Techniques for Estimating Value-at-Risk. Management Science, 46: 1139-1164.

Ha, H., Bauer, D., 2018. A Least-Squares Monte Carlo Approach to the Calculation of Capital requirements. Working paper.

Khare, K., Zhou, H., 2009. Rate of Convergence of Some Multivariate Markov Chains with Polynomial Eigenfunctions. The Annals of Applied Probability, 19: 737-777.

Rahman1, S., 2017. Wiener-Hermite Polynomial Expansion for Multivariate Gaussian Probability Measures. Working paper.