Journal of Probability and Statistical Science
24(2), 1-15Aug. 2026

Estimation VaR and CVaR by Gradient and Langevin MC Methods

Reza Habibi
Iran Banking Institute, Central Bank of Iran, Iran

ABSTRACT

Value at Risk (VaR) is a key risk management metric used to quantify potential losses in financial portfolios over a specified time horizon under normal market conditions. Traditional methods for estimating VaR, such as historical simulation and parametric approaches, can be limited by their reliance on strong assumptions about the underlying return distributions and their inability to accurately capture the behavior of extreme market events. In this paper, we introduce the Gradient Monte Carlo (GMC) and Langevin MC methods as robust alternatives for VaR estimation. By leveraging the principles of MC simulation combined with gradient-based optimization techniques, GMC efficiently approximates the distribution of portfolio returns, enabling accurate and reliable estimation of VaR while LMC provides a rigorous theoretical framework for the application of GMC and LMC in risk assessment and demonstrate their advantages over conventional methods through extensive empirical validation. Our results indicate that GMC and LMC not only enhance computational efficiency but also improve the accuracy of VaR and CVaR estimates, particularly in tail regions where conventional approaches may fail.

Keywords: CVaR, GMC, LMC, Monte Carlo simulation, Risk measure, VaR

1. Introduction

Financial risk management is the process of identifying and assessing potential risks that may impact a bank's overall financial ability. In cases where potential risks are identified, professionals assess whether they have to accept or mitigate the risks. The risks can be qualitative or quantitative. For analyzing qualitative risks, professional judgment is required, whereas, for quantitative risks, the analysis is based on numerical or historical data. Finance managers will use available financial tools to ensure the business is safe from any risks or, in financial terms, to hedge the business against them. The main goals of financial risk management include minimizing financial losses, maintaining financial stability, and ensuring compliance with relevant regulations, see Trindade *et al*. (2017). VaR and CVaR are both risk measures used in banking risk management to quantify the potential loss in value of a portfolio over a defined period. While VaR provides an estimate of potential losses up to a certain period, the CVaR offers a more comprehensive view by accounting for the expected losses beyond the VaR threshold. Let L be the loss of portfolio with distribution function $F_L := F$. Considering the confidence level$1-\alpha$, the VaR measure v_α is defined as percentile F, *i.e.*,

$$F(v_\alpha) = 1-\alpha.$$

The truncated distribution of L given event $[L > v_\alpha]$ is given by

$$P(L \le l | L > v_\alpha) = 1 - \frac{1-F(l)}{\alpha}, \quad l \ge v_\alpha.$$

The expectation of this truncated distribution is given by CVaR. *i.e.*, c_α, given by

$$c_\alpha = v_\alpha + \frac{1}{\alpha} E(L - v_\alpha)_+,$$

where $a_+ = \max(x, 0)$, see Hong and Liu (2021). Based on these formulas, Hong and Liu (2021) using MC simulation and Sun and Hong (2019) using importance sampling based MC estimated c_α and studied the related asymptotic distribution of estimates. Bardou *et al*. (2020)

Received June 20, 2026, in final form August 2026.
Reza Habibi (corresponding author) is affiliated with Iran Banking Institute, Central Bank of Iran, Iran.
Email: habibi1356@yahoo.com

computed VaR and CVaR using stochastic approximation and stochastic gradient algorithm. They mention that this method produces estimates which are more close to actual VaR and CVaR with respect to traditional methods such as the inversion of distribution function and quantile estimation. By adapting the Bayesian learning setting, Barbu and Zhu (2025) studied the Hamiltonian and Langevin MC methods as optimization algorithms to estimate v_α and c_α. They also provide a comprehensive review on these types of MC method while Baker *et al.* (2019) gives rich *R* software package $sgmcmc$ to apply these methods.

In the current paper, we extend their work. Let $c(v)$ be the CVaR allocated to VaR v, then

$$c(v) = E(L|L > v) = v + E(L - v|L - v > 0).$$

For every real number a, then

$$a_+ = \frac{a+|a|}{2},$$

thus,

$$c(v) = v + \frac{E(L-v)_+}{P(L>v)} = v + \frac{E\left((L-v)+|L-v|\right)}{2P(L>v)}.$$

Using the Knight identity, see Maanan *et al.* (2025), *i.e.*,

$$|x - y| - |x| = -y\{I(x > 0) - I(x < 0)\} - \int_0^y (2I(x \le s) - 1)ds,$$

which holds for $x \neq 0$, where $I(A)$ stands for the indicator function of event *A*, it is seen that

$$c(v) = v + \frac{E((L-v)+(L-v)+2\int_0^v I(L\le s)ds)}{2P(L>v)} =$$

$$= v + \frac{E\left((L-v)+\int_0^v I(L\le s)ds)\right)}{P(L>v)} =$$

$$= v + \frac{E(L)-v+\int_0^v P(L\le s)ds)}{P(L>v)}.$$

Notice that

$$\int_0^v P(L \le s)ds) - v = -\int_v^\infty P(L > s)ds, \quad \text{and}$$

since L is a positive variable, then

$$E(L) = \int_0^\infty P(L > s)ds.$$

Let

$$G(s) = P(L > s)$$

be the survive function related to *L*, then

$$c(v) = v + \frac{\int_v^\infty G(s)ds}{G(v)}.$$

By taking derivative from both sides, it is seen that

$$c'(v) = \frac{-G\prime(v)}{G(v)} \times \frac{\int_v^\infty G(s)ds}{G(v)},$$

where c' is derivative of c with respect to v. Then, the following differential equation is derived for $c(v)$, *i.e.*,

$$c'(v) = \lambda(v)(c(v) - v),$$

where

$$\lambda(v) = \frac{-G\prime(v)}{G(v)},$$

is the hazard function.

Remark 1. Using integration by parts, an alternative formula is given by

$$c(v) = -\frac{\int_v^\infty sdG(s)}{G(v)}.$$

Some examples. Here, some examples with closed form CVaR are given.

Example 1. When L has exponential distribution with hazard parameter λ, and

$$G(s) = e^{-s/\lambda}, s > 0,$$

it is seen that

$$\lambda(v) = \lambda \text{ (constant)},$$
$$v_\alpha = -\lambda \log(\alpha),$$

and

$$c(v) = v + \lambda,$$

Example 2. When L has Pareto distribution with shape and scale positive parameters a, b, $c(v_\alpha) = \lambda(1 - \log(\alpha))$. respectively, with

$$G(s) = (\frac{b}{v})^a,$$

then

$$v_\alpha = b\alpha^{-1/a},$$
$$c(v) = v + \frac{v}{a-1}(\max(v, b))^{1-a},$$
$$c(v_\alpha) = v_\alpha\left(1 + \frac{{v_\alpha}^{1-a}}{a-1}\right).$$

Example 3. For special case, when L has Weibull distribution with shape parameter 2 and scale parameter $a > 0$, then

$$G(s) = e^{-(\frac{s}{a})^2},$$

then

$$\frac{\int_v^\infty G(s)ds}{G(v)} = \frac{\int_v^\infty \frac{1}{\sqrt{2\pi}a} e^{-(\frac{s}{a})^2} ds}{\frac{1}{\sqrt{2\pi}a} e^{-(\frac{v}{a})^2}} = \frac{1-N(v,0,a)}{n(v,0,a)},$$

where $N(v, 0, a)$, and $n(v, 0, a)$ are cumulative distribution and density functions of a normal distribution with zero mean and standard deviation $a > 0$.

In the next section, recursive estimates of VaR and CVaR are proposed. Limiting distributions and variance reduction results are also proposed in section 2. Section 2 also studies the Langevin MC method and also studies the translated gamma approximation.

Sensitivity analysis as well as randomized step size issues are proposed in section 3.

2 Recursive relations

To find recursive relations, let

$$c_n = c(v_n), \lambda_n = \lambda(v_n)$$

and

$$\delta c_{n+1} = c_{n+1} - c_n.$$

Then, the discrete recursive relation for c_n is given by

$$\delta c_{n+1} = \lambda_n(c_n - v_n)\delta v_{n+1}.$$

Here, recursive estimation for VaR v_n is proposed. Notice that $G(v) = \alpha$ and

$$c(v_\alpha) = v_\alpha + \frac{\int_{v_\alpha}^\infty G(s)ds}{\alpha} =$$
$$= E\frac{\left\{\int_{v_\alpha}^\infty I(L>s)ds + \alpha v_\alpha\right\}}{\alpha}.$$

Following Bardou *et al*. (2020), it is enough to minimize the

$$E\left\{\int_{v_\alpha}^\infty I(L > s)ds + \alpha v_\alpha\right\},$$

with respect to v_α. Using the stochastic gradient method, the following recursive estimate of v_n is obtained:

$$\delta v_{n+1} = -\gamma_n\{I(L_n \le v_n) - \alpha\}.$$

where γ_n is a sequence of positive real numbers such that
$\sum_{n=1}^{\infty}\gamma_n = \infty$ and $\sum_{n=1}^{\infty}\gamma_n^2 < \infty$ (say, $\gamma_n = \frac{1}{n}$).

The following proposition summarizes the above discussion and the proof is omitted.

Proposition 1. Let $\lambda(v)$ be the hazard function allocated to loss variable *L*, computed at value at risk *v*, then (*a*)-(*c*) are correct.

(a) The conditional value at risk $c(v)$ satisfies the following differential equation as

$$c'(v) = \lambda(v)(c(v) - v).$$

(b) The discrete format of $c(v)$ is given by

$$\delta c_{n+1} = \lambda_n(c_n - v_n)\delta v_{n+1}.$$

(c) The recursive estimates of VaR is given by

$$\delta v_{n+1} = -\gamma_n\{I(L_n \le v_n) - \alpha\}.$$

Remark 2. One can notice that using the Hong and Liu (2021) percentile estimator and the Bardou *et al*. (2020) method, the same recursive for v_n is derived. To this end, first, notice that the criteria

$$F(v_\alpha) = 1 - \alpha$$

of Hong and Liu (2021) implies that

$$E\{H(L_n; v_n)\} = 0,$$

where

$$H(L_n; v_n) = I(L_n \le v_\alpha) - \alpha.$$

Second, notice that function H_1 of Bardou *et al*. (2020) is a function of

$$H(L_n; v_n).$$

Remark 3. Following Hong and Liu (2021), the estimation of v_α is the minimizer of

$$v + \frac{\sum_{i=1}^{n}\{I(L_i \le v) - \alpha\}}{n(1-\alpha)} = v + \frac{F_n(v) - \alpha}{1-\alpha}.$$

2.1 Asymptotic results

Here, asymptotic results are proposed. Results are compared with asymptotic behaviors quantile estimations.

VaR asymptotic. Let

$$Z_{n+1} = I(L_{n+1} \le v_n) - \alpha.$$

The mean and variance of Z_n are zero and $\alpha(1-\alpha)$, respectively. By solving the recursive relation of v_n, it is seen that

$$v_n = v_\alpha - \sum_{k=1}^{n-1}\gamma_k Z_k.$$

Using the central limit theorem, it is seen that

$$\frac{v_n - v_\alpha}{\sqrt{\sum_{k=1}^{n-1}\gamma_k^2}}$$

converges in distribution (denoted by $\to^d$) to

$$\sqrt{\alpha(1-\alpha)}N(0,1)$$

distribution.

Remark 4. Let $\hat{v}_n$ be the quantile estimation of VaR, then using the Bahadur representation, see Trindade *et al*. (2017), it is seen that

$$\hat{v}_n = v_\alpha - \frac{1}{f(v_\alpha)}\sum_{k=1}^{n}\{I(L_n \le v_\alpha) - \alpha\} + O_{a.s}(a_n),$$

where

$$a_n = \left(\frac{\log(n)}{n}\right)^{0.75}$$

and $f(\cdot)$ is the density of *L*. Here,

$$A_n = O_{a.s}(a_n),$$

if A_n/a_n is bounded with probability of one. Following Hong and Liu (2021), then

$$\sqrt{n}(\hat{v}_n - v_\alpha) \to^d \frac{\sqrt{\alpha(1-\alpha)}}{f(v_\alpha)} N(0,1).$$

Consistencies of v_n and $\hat{v}_n$ are obvious. The ratio of rate of convergence of v_n and $\hat{v}_n$ to v_α is

$$\frac{1}{\sqrt{n\sum_{k=1}^{n-1}\gamma_k^2}}.$$

The ratio of variances of v_n and $\hat{v}_n$ is

$$\sqrt{f(v_\alpha)}.$$

CVaR asymptotic. As $n \to \infty$, then $v_n \to^p v_\alpha$ with the rate of convergence of

$$\frac{1}{\sqrt{\sum_{k=1}^{n-1}\gamma_k^2}}.$$

This result is true for c_n, that is $c_n \to^p c_\alpha$, Notation $\to^p$ stands for convergence in probability. *Remark 5*. Let $\hat{c}_n$ be the CVaR related to $\hat{v}_n$. Following Hong and Liu (2021), then

$$\sqrt{n}(\hat{c}_n - c_\alpha) \to^d \vartheta N(0,1),$$

where

$$\vartheta = \frac{1}{(1-\alpha)^2} var(L - v_\alpha)_+,$$

Notation $\to^d$ stands for convergence in distribution.

Continuous time asymptotic. Here, it is assumed that

$$\gamma_k = \gamma_{kn} = \gamma\left(\frac{k}{n}\right),$$

for some function $\gamma(.)$. Then,

$$n^{-0.5}\sum_{k=1}^{n-1}\gamma_k Z_k \to^d \sqrt{\alpha(1-\alpha)}\int_0^1 \gamma(t)dB_t,$$

where B_t is the standard Brownian motion on [0,1]. Thus,

$$n^{-0.5}(v_n - v_\alpha) \to^d \sqrt{\alpha(1-\alpha)}\int_0^1 \gamma(t)dB_t.$$

In the continuous time, it is seen that

$$\begin{cases} dv_t = v_\alpha dt + \sqrt{\alpha(1-\alpha)}\gamma(t)dB_t \\ dc_t = \lambda_t(c_t - \lambda_t)dt. \end{cases}$$

Proposition 2. The scaled VaR

$$\frac{v_n - v_\alpha}{\sqrt{\sum_{k=1}^{n-1}\gamma_k^2}}$$

is distributed as

$$\sqrt{\alpha(1-\alpha)}N(0,1)$$

and in the continuous time setting then

$$\begin{cases} dv_t = v_\alpha dt + \sqrt{\alpha(1-\alpha)}\gamma(t)dB_t \\ dc_t = \lambda_t(c_t - \lambda_t)dt. \end{cases}$$

2.2 Variance reduction

The importance sampling is an effective method among variance reduction methods such as stratified sampling and control variable in estimating risk measures with rare event aspects such as VaR and CVaR, see Trindade *et al*. (2017). To simulate loss *L* under another distribution *H* defined on $(0,\infty)$, let

$$\Delta(l) = \frac{F(l)}{H(l)}$$

be the likelihood ratio. It is easy to see that

$$F(l) = E_H\{I(L \le l)\Delta(L)\},$$

where E_H is the expectation under H distribution. Let v_n^{IS}, c_n^{IS} be the VaR and CVaR using importance sampling methods. The following proposition summarizes the results. *Proposition 3*. The recursive relations of v_n^{IS}, c_n^{IS} are proposing by replacing

$$I(L \leq l)$$

with

$$I(L \leq l)\Delta(L)$$

in propositions 1 and 2.

2.3 Randomized γ_n

Previously, it was seen that

$$v_n = v_\alpha - \sum_{k=1}^{n} \gamma_k Z_k,$$

where

$$Z_k = I(L_k \leq v_{k-1}) - \alpha.$$

Here, γ_n may be random variables. For example, following Bardou *et al*. (2020), let

$$\gamma_n = \frac{\Delta_n}{\sum_{k=1}^{n} \Delta_k},$$

where

$$\Delta_n = \log(S_n) - \log(S_{n-1}),$$

where S_n is the market value of portfolio computed at some time $t_n > 0$. If

$$\sum_{j=1}^{n} \gamma_j = 1,$$

then

$$v_n = v_\alpha - E(z_{\Gamma_n} | Z_1, \dots, Z_n),$$

where Γ_n is a random variable which gets values $k = 1, \dots, n$ with probabilities γ_k's, respectively. Here, the MC estimate reduces to the bootstrap estimate of expectation. Motivated by Sun and Hong (2019), let γ_n's be random variables. For example let

$$\gamma_n = \frac{U_n}{\sum_{k=1}^{n} U_k},$$

where U_n's are independent all come from standard uniform $U(0,1)$ distribution. Then,

$$v_n \rightarrow^{as.} v_\alpha.$$

Notation $\rightarrow^{as.}$ stands for almost sure convergence.

2.4 Sensitivity analysis

The v_α is the minimizing point of $c(v)$. That is

$$c'(v_\alpha) = 0.$$

Equivalently,

$$c(v_\alpha) = v_\alpha.$$

That is v_α is the root of

$$D(v) = c(v) - v.$$

Thus, v_n satisfies the Newton-Raphson iterative scheme, *i. e.*,

$$v_{n+1} = v_n - \frac{D(v_n)}{D'(v_n)} = v_n - \frac{c(v_n) - v_n}{c'(v_n) - 1}.$$

One can see that

$$\frac{1}{v_{n+1} - v_n} = \lambda_n - \frac{1}{c_n - v_n}.$$

As λ_n gets large, then the rate of convergence of $v_{n+1} - v_n$ to zero is more.

3. Other types of MC

In this section, other types of MC simulations are proposed.

3.1 Langevin MC

Using the Langevin MC method, similar recursive relations for VaR is obtainable. Notice that

$$c(v) = v + \frac{\int_v^\infty G(s)ds}{G(v)},$$
$$c'(v) = \frac{-G\prime(v)}{G(v)} \times \frac{\int_v^\infty G(s)ds}{G(v)}.$$

Then, the Langevin MC implies that

$$dv = -\varphi c'(v)dt + \sqrt{2}dB_t,$$

where φ is the time step factor and B_t is the standard Brownian motion on [0,1]. The discrete samples are taken by

$$\delta v_{n+1} = -\varphi c'(v_n) + \sqrt{2\varphi}Z_{n+1},$$

where Z_{n+1} is a sequence of 8ndependent samples from standard normal distribution.

3.2 Translated gamma

The translated gamma approximation (see Kaas *et al*., 2020), approximate G(s) by a survival function of a shifted gamma distributed random variable $Z + z_0$, wherez_0 is a constant and Z has gamma distribution with shape and scale parameters ϑ_1, ϑ_2, respectively. The idea is to approximate distribution of L with $Z + z_0$ by matching the first three moments. According to the Kaas *et al*. (2020), it is seen that

$$\vartheta_1 = \frac{4}{\gamma_L^2}, \vartheta_2 = \frac{2}{\gamma_L \sigma_L},$$
$$z_0 = \mu_L - \frac{2\sigma_L}{\gamma_L},$$

where μ_L, σ_L and γ_L are mean, standard deviation and skew of loss L which are strictly positive and estimated using Monte Carlo simulation. Thus,

$$G(s) = P(Z > s - z_0),$$

and

$$\int_v^\infty G(s)ds = \int_{v-z_0}^\infty P(Z > s)ds =$$
$$= \int_{v-z_0}^\infty P\left(Z^* > \frac{2s}{\vartheta_2}\right)ds,$$

where

$$Z^* = \frac{2Z}{\vartheta_2}$$

has chi-squared distribution with $2\vartheta_1$ degrees of freedom. Thus, the above integral equals to

$$= \frac{\vartheta_2}{2}\int_{\frac{2(v-z_0)}{\vartheta_2}}^\infty P(Z^* > s)ds.$$

Beside this,

$$G(s) = P(Z^* > \frac{2(v-z_0)}{\vartheta_2}).$$

Notice that $G(v_\alpha) = \alpha$. Therefore,

$$P\left(Z^* > \frac{2(v_\alpha - z_0)}{\vartheta_2}\right) = \alpha.$$

Thus,

$$v_\alpha = \frac{\vartheta_2}{2}\chi^2{}_{2\vartheta_1,\alpha} + z_0.$$

Here, $\chi^2{}_{2\vartheta_1,\alpha}$ is $(1-\alpha)$-th quantile of chi-squared distribution with $2\vartheta_1$ degrees of freedom.

Remark 6. Notice that

$$Z^{**} = Z/\vartheta_2$$

has the gamma distribution with shape parameter ϑ_1/ϑ_2 and scale parameter 1. That is,

$$G(s) = \vartheta_2 \int_{\frac{v-z_0}{\vartheta_2}}^\infty P(Z^{**} > s)ds,$$

When ϑ_1/ϑ_2 is close to natural number m, then

$$P(Z^{**} > s) = P(N_s < m),$$

where N_s has Poisson distribution with rate s. Therefore,

$$\int_v^{\infty} G(s)ds = \vartheta_2 \int_{\frac{v-z_0}{\vartheta_2}}^{\infty} P(N_s < m)ds,$$

$$G(v) = P\left(Z^{**} > \frac{v-z_0}{\vartheta_2}\right) = P\left(N_{\frac{v-z_0}{\vartheta_2}} < m\right).$$

References

[1]. Barbu, A. and Zhu, S. C. (2025). *Monte Carlo methods*. Springer. USA.

[2]. Bardou, O., Frikha, N., and Pages, G. (2020). *Computing VaR and CVaR using stochastic approximation and adaptive unconstrained importance sampling*. *arXiv:0812.3381v5*

[3]. Baker, J., Fearnhead, P., Fox, E. B., and Nemeth, C. (2019). $sgmcmc$: An R package for stochastic gradient Markov Chain Monte Carlo. *Journal of Statistical Software* **91**, 1-27.

[4]. Hong, L. J. and Liu, G. (2021). Monte Carlo estimation of Value at Risk and conditional Value at Risk and their sensitivities. *Proceedings of the 2021 Winter Simulation Conference*, 95-107.

[5]. Kaas, R., Goovaerts, M., Dhaene, J., and Denuit, M. (2020). *Modern actuarial risk theory usingR* . Springer. USA.

[6]. Maanan, S., Dermoune, A., and El-Ghini, A. (2025). Multi-scale asymptotic normality in percentile regression: Hilbert matrices and polynomial designs. *arXiv:2503.15041v3*.

[7]. Sun, L. and Hong, L. J. (2019). A general framework of importance sampling for Value at Risk and Conditional Value at Risk. *Proceedings of the 2011 Winter Simulation Conference*, 415-422.

[8]. Trindade, A. A., Uryasev, S., Shapiro, A. and Zrazhevsky, G. (2017). Financial prediction with constrained tail risk. *Journal of Banking and Finance* **31**, 3524–3538.