Proceedings of the 2022 Winter Simulation Conference
B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C.G. Corlu, L.H. Lee, E.P. Chew, T. Roeder, and P. Lendermann, eds.

QUANTILE SENSITIVITY ESTIMATION THROUGH DELTA FAMILY METHOD

Zhenyu Cui

School of Business
Stevens Institute of Technology
1 Castle Point on Hudson
Hoboken, NJ 07030, USA

Kailin Ding

Academy of Mathematics and Systems Science,
Chinese Academy of Sciences,
55 Zhongguancun East Rd, Haidian District,
Beijing, 100190, P. R. CHINA

ABSTRACT

In this paper, we consider the estimation of the quantile sensitivity through Monte Carlo simulations. We first propose a new representation of the quantile by writing it as an expectation involving Dirac Delta payoff functions. Then we consider two alternative approximations for the Dirac Delta function, one is the Delta sequence, and the other is through orthogonal series. Then we derive quantile sensitivity estimators by combining them with the conditional expectation representation derived in (Hong 2009). Numerical examples demonstrate the accurateness and efficiency of the proposed method, and compare with the existing literature.

1 INTRODUCTION

The quantile of a random variable is an important measure of the random performance, and its equivalent, the value at risk (VaR) is an important risk measure for quantitative risk management. Except for a few special cases, the quantile is usually not available in closed-form, hence Monte Carlo simulation is often used to estimate it. The quantile sensitivity measures the sensitivity of the quantile with respect to a model parameter, and it provides information useful for hedging purposes and also in gradient-based optimization. A seminal work in quantile sensitivity estimation through Monte Carlo simulation is (Hong 2009), where the quantile sensitivity is represented as a conditional expectation. Subsequently, a kernel smoothing method is proposed in (Liu and Hong 2009) based on this representation. Representative alternative estimation methods include the use of infinitesimal perturbation analysis (IPA) in (Jiang and Fu 2015), conditional Monte Carlo method in (Fu, Hong, and Hu 2009), measure-valued differentiation approach in (Heidergott and Volk-Makarewicz 2016), and the generalized likelihood ratio (GLR) approach in (Peng, Fu, Hu, and Lei 2019), etc.

The motivation of the present paper stems from recent developments in utilizing the Delta family method in probability density estimations (Yang, Chen, and Wan 2019; Cui and Xu 2022), implied volatility calculation (Cui, Kirkby, Nguyen, and Taylor 2021) and high-dimensional stochastic control (Ma, Lu, and Cui 2022). The main idea of the Delta family method is based on smooth approximating sequences of the Dirac Delta function, which is a generalized function itself. In the literature, there are two distinct approximating sequences, with one being the *Delta sequence* first proposed in statistics, see (Walter and Blum 1979; Susarla and Walter 1981); The second class of approximating sequence is based on the distributional representation of the Dirac Delta function through projections onto orthonormal basis, see (Li and Wong 2008). In the present paper, we group these two types of approximation methods under the hood of the "Delta family method". We apply this Delta family method to the quantile sensitivity estimation problem, and in this paper we shall report initial findings and numerical evidences for the performance of the method. More specifically, we consider the first and second order sensitivities of the quantile, and also an example involving the sensitivity of the conditional value at risk.

The contributions are two fold: first, we establish a general representation of the quantile function in terms of the expectation with respect to a payoff function involving the Dirac Delta function, through a novel use of the sifting property of the Dirac Delta function and the change of variable technique. This representation is explicit, as compared to the traditional sample estimator based on interpolating order statistics of the sample. Based on this new representation, we propose two alternative converging approximations to the quantile based on the Delta family method. Second, we obtain explicit approximate expressions for the quantile sensitivity by applying the Delta family method to the conditional expectation representation of (Hong 2009), which are shown to converge to the true value. Numerical examples demonstrate the accuracy and efficiency of the new approach, and compare with existing methods in the literature.

The remainder of the paper is organized as follows: Section 2 presents the main results on a new quantile representation and the corresponding first and second order quantile sensitivities expressions. Section 3 presents numerical experiments and comparisons with existing methods in the literature. The method is shown to be accurate and efficient as compared to benchmark cases. Section 4 concludes the paper.

2 QUANTILE SENSITIVITY

We cast our setting where the parameter of interest can appear in both the performance function and the underlying input distribution, which is a very general setup incorporating many applications of interests, see also (Peng, Fu, Hu, and Heidergott 2018). We have $Y_\theta = g(X;\theta)$, where $g(\cdot;\theta)$ is the performance function, X is the input random variable having density function $f_X(\cdot;\theta)$, and $\theta \in \Theta$ with an open set $\Theta \subset \mathbb{R}$. The α quantile q_θ^α is defined as the unique root of $F(q_\theta^\alpha;\theta) = \alpha$, where F is the cumulative distribution function of Y_θ. The goal is to estimate $\frac{d}{d\theta} q_\theta^\alpha$, where θ is the parameter of interest.

We shall first establish a probabilistic representation of the quantile. From the sifting property of the Dirac Delta function, we have

$$\begin{aligned} q_\theta^\alpha = F^{-1}(\alpha;\theta) &= \int_0^1 F^{-1}(u;\theta)\delta(u-\alpha)du \\ &= \int_{\mathbb{R}} F^{-1}(F(s;\theta);\theta)\delta(F(s;\theta)-\alpha)dF(s;\theta) \\ &= \int_{\mathbb{R}} s\cdot\delta(F(s;\theta)-\alpha)f(s;\theta)ds \\ &= E[Y_\theta\cdot\delta(F(Y_\theta;\theta)-\alpha)]. \end{aligned} \tag{1}$$

Hence the quantile can be elegantly represented as the expectation in (1), and the payoff involves the Dirac Delta function, which belongs to the class of generalized functions. This *explicit* representation (1) of the quantile is new to the literature, and is in contrast with the *implicit* representation of quantile as the root of the equation $F(q_\theta^\alpha;\theta) = \alpha$. In the next step, we shall replace the Dirac Delta function by equivalent distributional representations using the tools of Delta family method, i.e. through Delta sequences or orthogonal series expansions.

As for the Delta sequences, there are many possible choices as documented in (Walter and Blum 1979), see Table 1 below.

For simplicity, we can consider the Delta sequence based on the normal density function:

$$\delta(x-a) = \lim_{\varepsilon\to 0}\frac{1}{2\sqrt{\pi\varepsilon}}e^{-\frac{(x-a)^2}{4\varepsilon}}.$$

Then combined with (1), we have the following equivalent representation of the quantile:

$$q_\theta^\alpha = \lim_{\varepsilon\to 0}\frac{1}{2\sqrt{\pi\varepsilon}}E\left[Y_\theta\cdot e^{-\frac{(F(Y_\theta,\theta)-\alpha)^2}{4\varepsilon}}\right]. \tag{2}$$

Table 1: Different types of Delta sequences.

Types	Delta sequences $\delta(x-y)$
Normal density	$\frac{1}{2\sqrt{\pi\varepsilon}}e^{-\frac{(x-y)^2}{4\varepsilon}}$
Lorentzian type	$\frac{1}{\pi}\frac{\varepsilon}{(x-y)^2+\varepsilon^2}$
Fourier integral	$\frac{1}{2\pi}\int_{-1/\varepsilon}^{1/\varepsilon} e^{i(x-y)t}dt$
Trigonometric function	$\frac{1}{2\pi}\frac{\sin\left[\left(\frac{1}{\varepsilon}+\frac{1}{2}\right)(x-y)\right]}{\sin\left(\frac{1}{2}(x-y)\right)}$

Note that the representation (2) is very similar to the Gaussian kernel approximation to the density function, with ε playing the role of *bandwidth* in the language of kernel density estimation. Since the kernel method has been previously applied in quantile sensitivity estimation in (Liu and Hong 2009), it is very important to distinguish our work from theirs when the Delta sequence approximation is utilized. The main distinction lies in that our derivation is based on a totally different representation of the quantile function as given in (1), in contrast to the sample estimator based on order statistics as utilized in (Liu and Hong 2009). More specifically, they utilize the quantile estimator $\hat{q}_\alpha^n = L_{\lceil n\alpha\rceil:n}$, where $L_{i:n}$ denotes the i-th order statistic from the n observations of L. Then their estimator for quantile sensitivity is given by equation (2) on page 513 of their paper, and recalled below:

$$\bar{V}_n = \frac{\sum_{i=1}^n K\left(\frac{\hat{q}_\alpha^n - L_i}{b_n}\right)\cdot D_i}{\sum_{i=1}^n K\left(\frac{\hat{q}_\alpha^n - L_i}{b_n}\right)},$$

where K denotes a kernel function on $\mathbb{R}$ and b_n is the bandwidth parameter. It can be seen that the main difference lies in how we approximate the quantile in the above expression, for which we apply the Delta family method to the new representation (1). Note that our second implementation based on orthogonal series expansions is totally different from the kernel approximation approach. We shall next discuss this second approach.

An alternative exact distributional representation of Dirac Delta function is as follows: let $\{g_k(y)\}_{k=0}^\infty$ be a complete orthonormal basis, then from (Li and Wong 2008), the Dirac Delta function can be represented by

$$\delta(x-a) = \sum_{k=0}^\infty g_k(x)g_k(a). \tag{3}$$

Note that the above identity holds in a "distribution" sense, and that the above representation is essentially equivalent to the completeness of the basis. There are many choices of the orthgonal basis for various situations, and more specifically, we have the following series representations of the Dirac Delta function:

$$\delta(x-a) = \sum_{k=0}^\infty \left(k+\frac{1}{2}\right)P_k(x)P_k(a),$$
$$\delta(x-a) = \frac{e^{-(x^2+a^2)/2}}{\sqrt{\pi}}\sum_{k=0}^\infty \frac{1}{2^k k!}H_k(x)H_k(a),$$
$$\delta(x-a) = e^{-(x+a)/2}\sum_{k=0}^\infty L_k(x)L_k(a),$$

and each of the above respectively corresponds to the choices of $g_k(x) = \left(k+\frac{1}{2}\right)^{1/2}P_k(x)$, $g_k(x) = \frac{1}{\pi^{1/4}2^{k/2}\sqrt{k!}}e^{-x^2/2}H_k(x)$, and $g_k(x) = e^{-x/2}L_k(x)$. Here $P_k(x)$, $H_k(x)$ and $L_k(x)$ are respectively the Legendre, Hermite and Laguerre polynomials, all of which are representative classical orthogonal polynomials.

Note that the above formulas appear respectively as formula (1.17.22), (1.17.24) and (1.17.23) on page 38 of the NIST handbook of Mathematical Functions, a definite reference on special functions. See https://dlmf.nist.gov

Combining (3) with (1), we can represent the quantile function as

$$q_\theta^\alpha = \sum_{k=0}^{\infty} g_k(\alpha) E\left[Y_\theta \cdot g_k(F(Y_\theta, \theta))\right]. \tag{4}$$

Note that the quantile is expressed as an orthogonal series expansion whose coefficients are given by expectations.

Remark 1 In numerical experiments, we have tried both Delta sequence method and the orthonormal basis method. It turns out that in order to reach the same accuracy level, the computational time of the orthonormal basis method is significantly larger than that of the Delta sequence method. We conjecture that this is due to the oscillatory behaviors at the boundaries of the involved orthogonal polynomials, i.e. Gibbs phenomenon. The efficient implementation of this alternative approach is left to future research.

With the above preparation, we shall study the quantile sensitivity using the Delta family method. We make the following standard assumptions as in (Hong 2009) using our notations:

ASSUMPTION 1. The pathwise derivative $\partial_\theta Y_\theta$ exists w.p.1 for any $\theta \in \Theta$, and there exists a function $k(X)$ with $\mathrm{E}[k(X)] < \infty$, such that

$$|Y_{\theta_2} - Y_{\theta_1}| \leqslant k(X)\,|\theta_2 - \theta_1|$$

for all $\theta_1, \theta_2 \in \Theta$.

ASSUMPTION 2. For any $\theta \in \Theta, Y_\theta$ has a continuous density $f(a;\theta)$ in a neighborhood of $t = a$ with any real number a, and $\partial_\theta F(t;\theta)$ exists and is continuous with respect to both θ and t at $t = a$.

For any $\theta \in \Theta$, let

$$h(t;\theta) = \mathrm{E}\left[\partial_\theta Y_\theta \mid Y_\theta = t\right]$$

We make the following assumption on $h(t;\theta)$.

ASSUMPTION 3. For any $\theta \in \Theta, h(t;\theta)$ is continuous at $t = a$.

From Theorem 2 of (Hong 2009), suppose that Assumptions 1-3 above are satisfied at $a = q_\theta^\alpha$, then the quantile sensitivity is given by the following conditional expectation:

$$\frac{d}{d\theta} q_\theta^\alpha = E\left[\frac{dY_\theta}{d\theta}\middle| Y_\theta = q_\theta^\alpha\right]. \tag{5}$$

ASSUMPTION 4. $E\left[\left|\frac{dY_\theta}{d\theta}\right| \cdot \mathbf{1}_{\{Y_\theta = q_\theta^\alpha\}}\right] < +\infty.$

Based on the above equation, we have the following result:

Proposition 2 Suppose that Assumptions 1-4 are satisfied at $a = q_\theta^\alpha$, then we arrive at an alternative representation of the first order quantile sensitivity utilizing the Dirac Delta function:

$$\frac{d}{d\theta} q_\theta^\alpha = \frac{E[\frac{dY_\theta}{d\theta}\delta(Y_\theta - q_\theta^\alpha)]}{E[\delta(Y_\theta - q_\theta^\alpha)]}. \tag{6}$$

Proof. Denote $f(y,x)$ as the joint density function of $\left(Y_\theta, \frac{dY_\theta}{d\theta}\right)$, then we have

$$\begin{aligned}\frac{d}{d\theta}q_\theta^\alpha = E\left[\frac{dY_\theta}{d\theta}\Big|Y_\theta = q_\theta^\alpha\right] = \frac{\int_{\mathbb{R}} x f(q_\theta^\alpha, x)dx}{\int_{\mathbb{R}} f(q_\theta^\alpha, x)dx} \\ =& \frac{\int_{\mathbb{R}} xE[\delta(Y_\theta - q_\theta^\alpha)\delta(\frac{dY_\theta}{d\theta} - x)]dx}{f(q_\theta^\alpha;\theta)} \\ =& \frac{E[\delta(Y_\theta - q_\theta^\alpha)\int_{\mathbb{R}} x\delta(\frac{dY_\theta}{d\theta} - x)dx]}{f(q_\theta^\alpha;\theta)} \\ =& \frac{E[\frac{dY_\theta}{d\theta}\delta(Y_\theta - q_\theta^\alpha)]}{f(q_\theta^\alpha;\theta)} \\ =& \frac{E[\frac{dY_\theta}{d\theta}\delta(Y_\theta - q_\theta^\alpha)]}{E[\delta(Y_\theta - q_\theta^\alpha)]},\end{aligned} \tag{7}$$

where $f(q_\theta^\alpha;\theta)$ is the density function of Y_θ at $Y_\theta = q_\theta^\alpha$. From Assumption 1, there exists a function $k(X)$, such that $\mathrm{E}\left[\left|\frac{Y_{\theta_1}-Y_{\theta_2}}{\theta_1-\theta_2}\right|\right] \leqslant \mathrm{E}[k(X)] < \infty$ for all $\theta_1,\theta_2 \in \Theta$, which is compatible with $E\left[\left|\frac{dY_\theta}{d\theta}\right|\right] < \infty$. Then from Fubini's theorem, the fourth equality in (7) holds for $\int_{\mathbb{R}} |x| f(q_\theta^\alpha, x)dx = E\left[\left|\frac{dY_\theta}{d\theta}\right| \cdot \mathbf{1}_{\{Y_\theta = q_\theta^\alpha\}}\right] < \infty$. $\square$

Next, the following result provides the second order quantile sensitivity:

Proposition 3 Suppose that Assumptions 1-4 are satisfied and suppose further $\partial_\theta f(a;\theta)$ exists and is continuous with respect to both θ and $a = q_\theta^\alpha$. Then the second order quantile sensitivity is given by

$$\begin{aligned}\frac{d^2}{d\theta^2}q_\theta^\alpha =& \frac{E[\frac{d^2Y_\theta}{d\theta^2}\delta(Y_\theta - q_\theta^\alpha)]}{E[\delta(Y_\theta - q_\theta^\alpha)]} + \frac{E[\frac{dY_\theta}{d\theta}\frac{d}{da}\delta(Y_\theta - a)|_{a=q_\theta^\alpha} \cdot \frac{d}{d\theta}q_\theta^\alpha]}{E[\delta(Y_\theta - q_\theta^\alpha)]} + \frac{d}{d\theta}q_\theta^\alpha \cdot \frac{E[\frac{dY_\theta}{d\theta}\frac{d}{da}\delta(Y_\theta - a)|_{a=q_\theta^\alpha}]}{E[\delta(Y_\theta - q_\theta^\alpha)]} \\ &- \left(\frac{d}{d\theta}q_\theta^\alpha\right)^2 \cdot \frac{E\left[\frac{d}{da}\delta(Y_\theta - a)\right]|_{a=q_\theta^\alpha}}{E\left[\delta(Y_\theta - q_\theta^\alpha)\right]}.\end{aligned} \tag{8}$$

Proof. Follow the ideas of (Hong 2009) and also the expression (6), we can carry out the following derivation:

$$\begin{aligned}\frac{d^2}{d\theta^2}q_\theta^\alpha =& \frac{E[\frac{d^2Y_\theta}{d\theta^2}\delta(Y_\theta - q_\theta^\alpha)]}{E[\delta(Y_\theta - q_\theta^\alpha)]} + \frac{E[\frac{dY_\theta}{d\theta}\frac{d}{d\theta}\delta(Y_\theta - q_\theta^\alpha)]}{E[\delta(Y_\theta - q_\theta^\alpha)]} - \frac{d}{d\theta}q_\theta^\alpha \cdot \frac{\frac{d}{d\theta}f(q_\theta^\alpha;\theta)}{f(q_\theta^\alpha;\theta)} \\ =:& \mathrm{I} + \mathrm{II} - \mathrm{III}.\end{aligned} \tag{9}$$

Next we shall calculate III. Denote $\pi(a,\theta) = E\left[(a - Y_\theta)\cdot \mathbf{1}_{\{Y_\theta \le a\}}\right]$, from the proof of Theorem 1 in (Hong 2009), there are

$$\partial_a \pi(a;\theta) = F(a;\theta),$$

for any t in the neighborhood of a, and $\partial_a\partial_t\pi(a;\theta) = \partial_t\partial_a\pi(a;\theta) = -f(a;\theta)\cdot E\left[\frac{d}{\partial\theta}Y_\theta|Y_\theta = a\right]$. Then we have

$$\begin{aligned}\frac{\partial}{\partial a}f(a;\theta) =& \partial_\theta\partial_{a^2}\pi(a;\theta) \\ =& \partial_a\left[-f(a;\theta)\cdot E[\frac{d}{\partial\theta}Y_\theta|Y_\theta = a]\right] \\ =& -\frac{d}{da}f(a;\theta)\cdot E\left[\frac{d}{\partial\theta}Y_\theta|Y_\theta = a\right] - f(a;\theta)\cdot\frac{d}{da}E\left[\frac{dY_\theta}{\partial\theta}|Y_\theta = a\right].\end{aligned} \tag{10}$$

For $a = q_\theta^\alpha$, it holds that

$$
\begin{aligned}
\frac{d}{d\theta} f(q_\theta^\alpha;\theta) =& \frac{\partial}{\partial\theta} f(a;\theta)\Big|_{a=q_\theta^\alpha} + \frac{\partial}{\partial a} f(a;\theta)|_{a=q_\theta^\alpha} \cdot \frac{d}{d\theta} q_\theta^\alpha \\
=& -\frac{\partial}{\partial a} f(a;\theta)\Big|_{a=q_\theta^\alpha} \cdot \frac{d}{d\theta} q_\theta^\alpha - f(q_\theta^\alpha;\theta)\cdot \frac{\partial}{\partial a} E\left[\frac{dY_\theta}{d\theta}|Y_\theta = a\right]\Bigg|_{a=q_\theta^\alpha} + \frac{\partial}{\partial a} f(a;\theta)\Big|_{a=q_\theta^\alpha} \cdot \frac{d}{d\theta} q_\theta^\alpha \\
=& - f(q_\theta^\alpha;\theta)\cdot \frac{\partial}{\partial a} E\left[\frac{dY_\theta}{d\theta}|Y_\theta = a\right]\Bigg|_{a=q_\theta^\alpha}. \qquad (11)
\end{aligned}
$$

Hence,

$$
\mathrm{III} = -\frac{d}{d\theta} q_\theta^\alpha \cdot \frac{\partial}{\partial a} E\left[\frac{dY_\theta}{d\theta}\Big|Y_\theta = a\right]\Bigg|_{a=q_\theta^\alpha}. \qquad (12)
$$

□

2.1 Extension to sensitivities of risk measures

Note that the value at risk (VaR) is equivalent to the quantile of the loss random variable, which is denoted also as Y_θ for notational convenience. The explicit representation in (1) also allows us to compute more general risk measures named the range value at risk (RVaR). It was proposed in (Cont, Deguest, and Scandolo 2010) as a robust risk measure, defined as

$$
\mathrm{RVaR}_{\alpha,\beta}[Y_\theta] = \frac{1}{\beta-\alpha}\int_\alpha^\beta \mathrm{VaR}_u[Y_\theta]du, \quad \forall 0<\alpha<\beta<1,
$$

and see (Embrechts, Liu, and Wang 2018) for some recent study of its theoretical properties.

From (1), we have $\mathrm{VaR}_u[Y_\theta] = E[Y_\theta \cdot \delta(F(Y_\theta;\theta)-u)]$, where $F(\cdot;\theta)$ is the distribution function of the loss random variable. After plugging it into the above expression, we have

$$
\mathrm{RVaR}_{\alpha,\beta}[Y_\theta] = \frac{1}{\beta-\alpha}\int_\alpha^\beta E[Y_\theta \cdot \delta(F(Y_\theta;\theta)-u)]du, \quad \forall\ 0<\alpha<\beta<1.
$$

We can plug in either of the two approximate representations of the quantile as given in (2) or (4), and obtain some further simplification of the final expressions for the quantile. First, we utilize the representation in (2), and there is

$$
\mathrm{RVaR}_{\alpha,\beta}[Y_\theta] = \lim_{\varepsilon\to 0}\frac{1}{2(\beta-\alpha)\sqrt{\pi\varepsilon}} E\left[Y_\theta \int_\alpha^\beta e^{-\frac{(F(Y_\theta,\theta)-u)^2}{4\varepsilon}}du\right]. \qquad (13)
$$

Recall the identity $\int_0^a e^{-t^2}dt = \frac{\sqrt{\pi}}{2}\mathrm{erf}(a)$, then we can alternatively simplify the above expression as

$$
\mathrm{RVaR}_{\alpha,\beta}[Y_\theta] = \lim_{\varepsilon\to 0}\frac{1}{2(\beta-\alpha)} E\left[Y_\theta\left(\mathrm{erf}\left(\frac{\beta-F(Y_\theta,\theta)}{2\sqrt{\varepsilon}}\right) - \mathrm{erf}\left(\frac{\alpha-F(Y_\theta,\theta)}{2\sqrt{\varepsilon}}\right)\right)\right]. \qquad (14)
$$

We further pose the following assumptions on Y_θ:

ASSUMPTION 5. $\mathrm{VaR}_\alpha[Y_\theta]$ and $\mathrm{VaR}_\beta[Y_\theta]$ are differentiable for any $\theta\in\Theta$.

ASSUMPTION 6. For any $\theta \in \Theta$, $P(Y_\theta = \mathrm{VaR}_\alpha[Y_\theta]) = 0$ and $P\left(Y_\theta = \mathrm{VaR}_\beta[Y_\theta]\right) = 0$.

Alternatively, we have the following equivalent representation

$$\begin{aligned}
\mathrm{RVaR}_{\alpha,\beta}[Y_\theta] =& \frac{1}{\beta-\alpha}\int_\alpha^\beta \mathrm{VaR}_u[Y_\theta]du \\
=& \frac{1}{\beta-\alpha}\int_\alpha^\beta F_Y^{-1}(v)dv \\
=& \frac{1}{\beta-\alpha}\int_{F_{Y_\theta}^{-1}(\alpha)}^{F_{Y_\theta}^{-1}(\beta)} u dF_{Y_\theta}(u) \\
=& \frac{1}{\beta-\alpha}E\left[Y_\theta \middle| \mathrm{VaR}_\alpha[L_\theta] \le L_\theta < \mathrm{VaR}_\beta[Y_\theta]\right] \\
=& \frac{1}{\beta-\alpha}\left(E\left[Y_\theta \mathbf{1}_{\{Y_\theta > \mathrm{VaR}_\alpha[Y_\theta]\}}\right] - E\left[Y_\theta \mathbf{1}_{\{Y_\theta > \mathrm{VaR}_\alpha[Y_\beta]\}}\right]\right).
\end{aligned} \tag{15}$$

From (15), we can similarly propose to estimate the sensitivity of the RVAR, which includes the popular risk measures VaR and CVaR as special cases.

Proposition 4 Suppose that Assumptions 5-6 are satisfied. Then the sensitivity of the RVAR can be represented as

$$\frac{d}{d\theta}\mathrm{RVaR}_{\alpha,\beta}[Y_\theta] = E\left[\frac{d}{d\theta}Y_\theta \middle| Y_\theta \ge \mathrm{VaR}_\alpha[Y_\theta]\right] - E\left[\frac{d}{d\theta}Y_\theta \middle| Y_\theta \ge \mathrm{VaR}_\beta[Y_\theta]\right]. \tag{16}$$

From Theorem 3.1 in (Hong and Liu 2009), the sensitivity of conditional value at risk (CVaR) is given by

$$\frac{d}{d\theta}CVaR_\alpha[Y_\theta] = E\left[\frac{d}{d\theta}Y_\theta | Y_\theta \ge q_\theta^\alpha\right] = \frac{1}{1-\alpha}E\left[\frac{d}{d\theta}Y_\theta \cdot \mathbf{1}_{\{Y_\theta \ge q_\theta^\alpha\}}\right]. \tag{17}$$

Note that (17) is a special case of (16).

3 NUMERICAL EXPERIMENTS

In this section, we carry out several numerical experiments to illustrate our results. All experiments are conducted in Matlab 9.3 on a personal computer with an Intel Core i5 CPU @ 2.8 GHz.

We note the following fact that $F(Y_\theta) \sim Uni(0,1)$, and assume that we have M simulated values of $Y_\theta(\omega_i), i = 1,2,\ldots,M$. If we replace the distribution function $F(\cdot)$ by the empirical cumulative distribution function (ECDF) $\hat{F}(\cdot)$, then we have that $\hat{F}_M(Y_\theta^{(m)}) = m/M$ for $m = 1,2,\ldots,M$. Here $Y_\theta^{(m)}$ is the m-th order statistic of the simulated M samples.

From (1), and combined with the above analysis, we have the following nonparametric estimates of the quantile function:

$$q_\theta^\alpha \approx \frac{1}{M}\sum_{m=1}^{M} Y_\theta^{(m)} \cdot \delta\left(\frac{m}{M} - \alpha\right), \tag{18}$$

or equivalently the quantile can be estimated as the weighted average of the order statistics of the simulated sample.

It is important to distinguish the representation (18) from the classical method for estimating the quantile by treating it as a fractional order statistic, i.e. $\hat{q}_\alpha^n = Y_{\lceil n\alpha \rceil:n}$, where $Y_{i:n}$ denotes the i-th order statistic from the n observations of Y. Intuitively, the traditional method is a *local* method that involves only one order statistics, while our proposed method is a *global* method that involves the (weighted average) of *all* the order statistics.

We can substitute in the orthogonal series representation of the Dirac Delta function, and we have

$$\begin{aligned} q_\theta^\alpha &= \sum_{k=0}^{\infty} g_k(\alpha) E\left[Y_\theta \cdot g_k(F(Y_\theta,\theta))\right] \\ &\approx \sum_{k=0}^{\infty} \frac{g_k(\alpha)}{M} \sum_{m=1}^{M} Y_\theta^{(m)} g_k\left(\frac{m}{M}\right). \end{aligned} \tag{19}$$

Remark 5 For the Delta sequences representation, $E[\delta(x-y)]$ is given as $\varepsilon \to 0$. In actual implementations, let $\varepsilon = \frac{1}{n}$ for an integer $n \in \mathbb{N}$, then $\varepsilon \to 0$ as $n \to \infty$ and we have

$$\frac{1}{2\sqrt{\pi\varepsilon}} E\left[e^{-\frac{(x-y)^2}{4\varepsilon}}\right] = \frac{\sqrt{n}}{2\sqrt{\pi}} E\left[e^{-\frac{n(x-y)^2}{4}}\right] =: A_n. \tag{20}$$

Through the use of telescopic series, we can rewrite $E[\delta(x-y)]$ in the following way:

$$E[\delta(x-y)] = \lim_{n\to\infty} A_n = \sum_{i=1}^{\infty} (A_i - A_{i-1}) \tag{21}$$

with $A_0 = 0$.

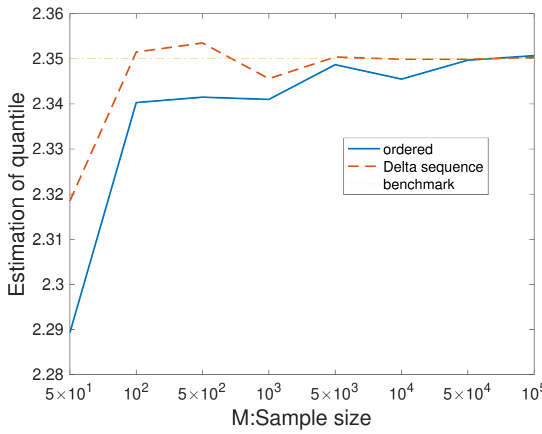
3.1 A toy example

We shall consider the same toy example as in (Peng, Fu, Hu, and Pierre 2021). Consider a simple stochastic model $Y_\theta = \theta X_1 + X_2 + U$, and X_1, X_2 are standard normal random variables with common cumulative distribution function $\Phi(\cdot)$, and U is an independent uniform random variable.

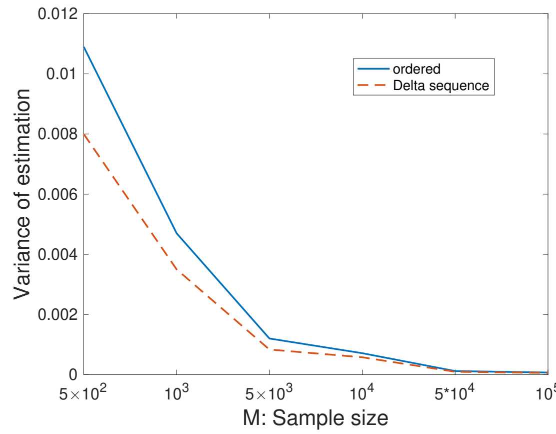
Denote q_θ^α the quantile of Y_θ. We shall verify the value of quantile sensitivity $\frac{dq_\theta^\alpha}{d\theta}$ using our method compared with the GLR method in (Peng, Fu, Hu, and Pierre 2021). In our implementation, we first compute the value of quantile given by (18), where we utilize the Gaussian delta sequence with parameter ε_1 to approximate the Dirac Delta function $\delta(\cdot)$. Next, the value of quantile sensitivity can be obtained based on the representation in (6) and the estimator of the quantile. In this step, we also choose the Delta sequences to represent the Dirac Delta function, which is more efficient than the orthogonal series expansions from pilot numerical experiments. Each result is conducted through Monte Carlo simulation and the values of quantile sensitivity are estimated through 100 independent runs. We fix $\alpha = \{0.1, 0.3, 0.5, 0.7, 0.9\}$ and $\theta = 1$, which are replicated from the parameter settings in (Peng, Fu, Hu, and Pierre 2021).

In the first step, we compare our Delta sequence method with the classical method using sample quantiles for the estimation of quantile. The comparison results are reported in Figure 1, where we set $\alpha = 0.9$ and the classical method is denoted as "ordered", and the benchmark is obtained by using the classical method with 10^8 sample sizes. We plot the estimations of quantile based on the classical ordered samples method and our Delta sequence method with respect to the sample size M in Figure 1(a), and the variances of these two estimations in Figure 1(b). It can be seen that the Delta sequence method converges faster to the exact benchmark as compared to the classical method of using sample quantile.

In the second step, we illustrate numerically the convergence results and CPU times for computing the quantile sensitivity with respect to θ for various values of ε and sample size M of the Monte Carlo simulation. Note that ε is the parameter of the Gaussian Delta sequence. We plot the convergence results in Figure 2(a), (b) using the benchmark given in (Peng, Fu, Hu, and Pierre 2021). The root mean square error (RMSE) therein is defined as $RMSE = \sqrt{\frac{1}{100}\sum_{i=1}^{100}(\hat{C}_i - C)^2}$, where $\hat{C}_i$ is the estimation through the



(a)



(b)

Figure 1: Comparisons of Delta sequence and ordered method based on 100 independently replications for the estimations of quantile. $\alpha = 0.9$.

i^{th} independent run and C is the benchmark. Figure 2(c) plots the CPU time of our method for each run with respect to varying values of $1/\varepsilon$. Here we fix $\varepsilon = 1/800$ and the CPU time includes the times of both steps of quantile computation and quantile sensitivity computation. From Figure 2, we observe that log absolute errors and RMSEs converge to 0 respectively as ε goes to 0 with a fixed M or as M goes to infinity with a fixed ε. We provide the results for all four α levels and the results show that the bias is largest when $\alpha = 0.9$ compared to other three levels, which is similar to those in (Hong 2009). To reach an accuracy of $1e^{-8}$, it approximately takes 2 seconds of CPU time when $M = 10^5$.

In Table 2, we take the GLR estimator as the benchmark and report the results of our method by implementing the Monte Carlo method with $M = 10^6$ sample sizes, and compare with the IPA estimator for the quantile sensitivity. We also provide the 95% confidence interval ("CI") of the IPA and our estimator based on 100 replications. The results show that our estimator is more accurate (with smaller variance) and the confidence intervals of our estimator are smaller than those of the IPA estimator.

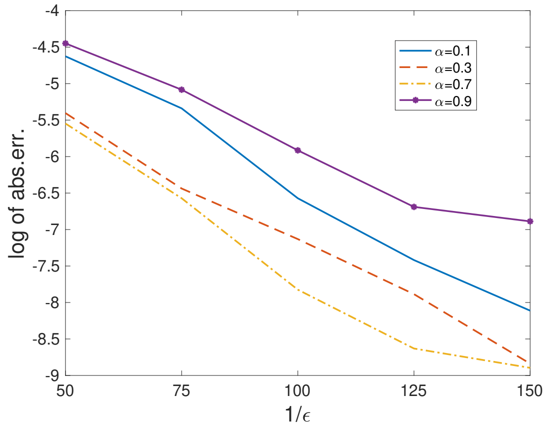
Table 2: The Delta family method in comparison with GLR method for the quantile sensitivity with $M = 10^6$ across 100 replications for Delta family and IPA method.

	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.5$	$\alpha = 0.7$	$\alpha = 0.9$
GLR	-0.888	-0.363	0	0.363	0.888
Delta family	-0.889	-0.363	0.004	0.363	0.888
(CI)	([-0.896,-0.883])	([-0.368,-0.360)	([0.001,0.005])	([0.359,0.367])	([0.883,0.894])
IPA	-0.8818	-0.360	0	0.366	0.884
(CI)	([-1.112,-0.651])	([-0.518,-0.202)	([-0.148,0.148])	([0.216,0.516])	([0.650,1.118])

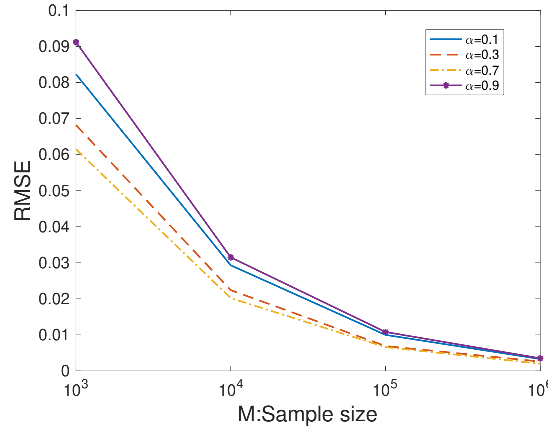
3.2 Example of sensitivity of conditional value at risk

We shall consider a loss model and the problem of estimating the sensitivity of conditional value at risk, which is considered in (Hong and Liu 2009): Denote ΔS the changes in the risk factors in the time period. Assume that $\Delta S \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ with mean $\boldsymbol{\mu} = (\mu_1, \mu_2)' = (0.01, 0.03)'$ and covariance

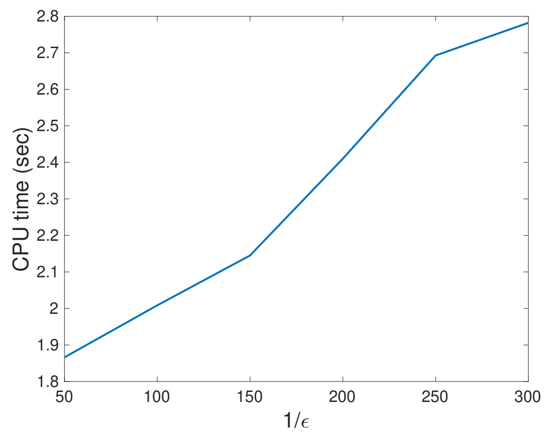
$$\Sigma = 0.02 \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix} \tag{22}$$



(a) $M = 10^5$



(b) $\varepsilon = 1/100$



(c) $M = 10^5$

Figure 2: The convergence results and the CPU times of the Delta family method for four quantile levels $\alpha = \{0.1, 0.3, 0.7, 0.9\}$.

follows a multivariate normal distribution. The loss model, which is the quadratic function of ΔS, can be given by

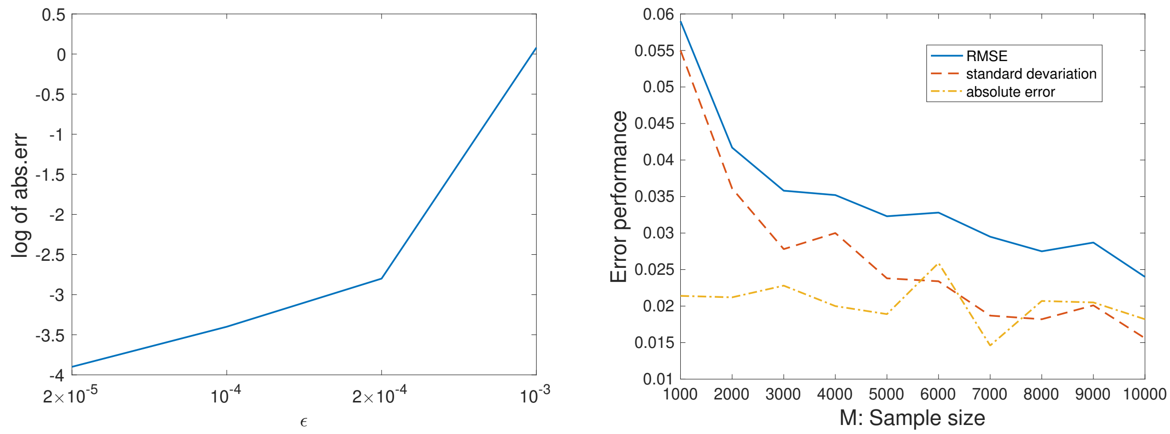
$$L = a_0 + a_1'\Delta + \Delta S' A \Delta S. \tag{23}$$

Here the parameter settings are $a_0 = 0.3$, $a_1 = (0.8, 1.5)'$ and

$$A = \begin{pmatrix} 1.2 & 0.6 \\ 0.6 & 1.5 \end{pmatrix}. \tag{24}$$

For the CVaR of L, assume $\alpha = 0.95$. The CVaR ($c^{\alpha}_{\mu_1}$) sensitivity $\frac{dc^{\alpha}_{\mu_1}}{\partial \mu_1}$ given by (17) is computed through our Delta family method combined with Monte Carlo simulations and 100 independent replications, where the value of $\frac{\partial L}{\partial \mu_1}$ is obtained by the IPA method. The convergence of numerical results is reported in Figure 3. In the left panel, we fix the sample size of Monte Carlo to be $M = 10^4$ and plot the logarithm of the absolute error with respect to different values of ε, which indicates that the results get more accurate as ε tends to zero. The absolute error is defined as $abs.err. = |\frac{dc^{\alpha}_{\mu_1}}{d\mu_1} - \frac{dc_{ben}}{d\mu_1}|$. The benchmark is given by

$\frac{dc_{ben}}{d\mu_1} = 1.7391$ which has been estimated in (Hong and Liu 2009) by using the finite difference method. The right panel of Figure 3 provides the root mean square error (RMSE), the standard deviations and the absolute errors with respect to increasing sample sizes M of Monte Carlo, where the Delta sequence parameter is fixed at $\varepsilon = 2 \times 10^{-5}$. The taken sample interval is 10^3, which is relatively small, and the randomness of the small sample interval that causes the errors to minorly fluctuate up and down. The overall trends are convergent with increasing sample size.



(a) $M = 10^4$

(b) $\varepsilon = 2 \times 10^{-5}$

Figure 3: Error performance of our method with respect to ε and sample size for the CVaR sensitivity.

4 CONCLUSION

In this paper, we propose to estimate the quantile sensitivity using the Delta family method. Two approaches, respectively the Delta sequence and orthogonal series expansions, are utilized and yield fast and accurate results as compared to benchmark in the literature. For future research, we shall explore high-dimensional sensitivity analysis based on the tensor decomposition property of the Dirac Delta function.

REFERENCES

Cont, R., R. Deguest, and G. Scandolo. 2010. "Robustness and sensitivity analysis of risk measurement procedures". *Quantitative Finance* 10(6):593–606.

Cui, Z., J. Kirkby, D. Nguyen, and S. Taylor. 2021. "A Closed-Form Model-Free Implied Volatility Formula through Delta Families". *The Journal of Derivatives* 28(4):111–127.

Cui, Z., and Y. Xu. 2022. "A new representation of the risk-neutral distribution and its applications". *Quantitative Finance* 22(5):817–834.

Embrechts, P., H. Liu, and R. Wang. 2018. "Quantile-based risk sharing". *Operations Research* 66(4):936–949.

Fu, M. C., L. J. Hong, and J.-Q. Hu. 2009. "Conditional Monte Carlo estimation of quantile sensitivities". *Management Science* 55(12):2019–2027.

Heidergott, B., and W. Volk-Makarewicz. 2016. "A measure-valued differentiation approach to sensitivities of quantiles". *Mathematics of Operations Research* 41(1):293–317.

Hong, L. J. 2009. "Estimating quantile sensitivities". *Operations research* 57(1):118–130.

Hong, L. J., and G. Liu. 2009. "Simulating sensitivities of conditional value at risk". *Management Science* 55(2):281–293.

Jiang, G., and M. C. Fu. 2015. "On estimating quantile sensitivities via infinitesimal perturbation analysis". *Operations Research* 63(2):435–441.

Li, Y., and R. Wong. 2008. "Integral and series representations of the Dirac delta function". *Communications on Pure and Applied Analysis* 7(2):229–247.

Liu, G., and L. J. Hong. 2009. "Kernel estimation of quantile sensitivities". *Naval Research Logistics* 56(6):511–525.

Ma, J., Z. Lu, and Z. Cui. 2022. "Delta family approach for the stochastic control problems of utility maximization". *Available at SSRN*.

Peng, Y., M. C. Fu, J. Hu, and L. Pierre. 2021. "Variance reduction for generalized likelihood ratio method in quantile sensitivity estimation". In *Proceedings of the 2021 Winter Simulation Conference*, edited by S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo, and M. Loper, 1–12. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Peng, Y., M. C. Fu, J.-Q. Hu, and B. Heidergott. 2018. "A new unbiased stochastic derivative estimator for discontinuous sample performances with structural parameters". *Operations Research* 66(2):487–499.

Peng, Y., M. C. Fu, J.-Q. Hu, and L. Lei. 2019. "Estimating quantile sensitivity for financial models with correlations and jumps". In *Proceedings of the 2019 Winter Simulation Conference*, edited by N. Mustafee, K.-H. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, , and Y.-J. Son, 962–973. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Susarla, V., and G. Walter. 1981. "Estimation of a multivariate density function using delta sequences". *Annals of Statistics* 9(2):347–355.

Walter, G., and J. Blum. 1979. "Probability density estimation using delta sequences". *Annals of Statistics* 7(2):328–340.

Yang, N., N. Chen, and X. Wan. 2019. "A new delta expansion for multivariate diffusions via the Itô-Taylor expansion". *Journal of Econometrics* 209(2):256–288.

AUTHOR BIOGRAPHIES

ZHENYU CUI is an Associate Professor in Financial Engineering at the School of Business at Stevens Institute of Technology. He holds a Ph.D. degree in Statistics from University of Waterloo, Canada. His research interests include financial engineering, Monte Carlo simulation, and financial econometrics. His e-mail address is zcui6@stevens.edua. His website is https://sites.google.com/site/zhenyucui86/home.

KAILIN DING is a Postdoctoral Fellow at the Academy of Mathematics and Systems Science, Chinese Academy of Sciences. She holds a Ph.D. degree in Mathematics from Nankai University. Her research interests include quantitative finance and numerical simulation. She is a member of IEEE and INFORMS. Her email address is klding@amss.ac.cn.