Submitted to *INFORMS Journal on Computing*
manuscript (Please, provide the mansucript number!)

Computing Sensitivities for Distortion Risk Measures

Peter W. Glynn
Department of Management Science and Engineering, Stanford University, Stanford, California 94305 USA, glynn@stanford.edu

Yijie Peng
Department of Management Science and Information Systems, Guanghua School of Management, Peking University, Beijing, 100871, China, pengyijie@pku.edu.cn

Michael C. Fu
The Robert H. Smith School of Business, Institute for Systems Research, University of Maryland, College Park, Maryland 20742-1815, USA, mfu@rhsmith.umd.edu

Jian-Qiang Hu
Department of Management Science, School of Management, Fudan University, Shanghai, 200433, China, hujq@fudan.edu.cn

Distortion risk measure, defined by an integral of a distorted tail probability, has been widely used in behavioral economics and risk management as an alternative to expected utility. The sensitivity of the distortion risk measure is a functional of certain distribution sensitivities. We propose a new sensitivity estimator for the distortion risk measure that uses the generalized likelihood ratio estimators in Peng et al. (2020) for distribution sensitivities as input and establish a central limit theorem for the new estimator. The proposed estimator can handle discontinuous sample paths and distortion functions.

Key words: sensitivity analysis; distortion risk measure; asymptotic analysis; functional limit theory

1. Introduction

Sensitivity analysis (or stochastic derivative estimation) is an important area in stochastic optimization (see Asmussen and Glynn 2007). The classic problem in sensitivity analysis is to estimate the derivative of an expectation of a random performance. Although finite difference (FD) methods are always implementable, they require simulation of multiple sample paths and face a bias-variance trade-off, whereas direct derivative estimators are unbiased and only need a single sample path. Two of the most popular direct derivative estimators are the likelihood ratio (LR) method (Glynn 1990, Rubinstein and Shapiro 1993) and infinitesimal perturbation analysis (IPA) (Ho and Cao 1991, Glasserman 1991); see Fu (2015) for recent review.

An important class of risk measures that is widely used in finance and economics are *distortion risk measures* that take the form

$$\vartheta = \int_0^\infty w(P(Z > z))dz, \tag{1}$$

where the "risk" Z is described by a non-negative random variable (rv). (When Z is of mixed sign, one can apply separate distortion risk measures to the positive and negative parts of Z.) In (1), the function w is called the *distortion function* and satisfies $w(0) = 0$ and $w(1) = 1$. The use of distortion risk measures is supported by cumulative prospect theory, and is an alternative to basing decisions on the foundation of expected utility theory; see Kahneman and Tversky (1979) and Tversky and Kahneman (1992).

Suppose $\theta \in \mathbb{R}^r$ is a (continuous) decision variable, and that the decision involves an objective based on the distortion risk measure (1). If the distribution of Z depends on θ, we can write the distortion risk measure as

$$\vartheta(\theta) = \int_0^\infty w(1 - F(z;\theta))dz, \tag{2}$$

where $F(\cdot;\theta)$ is the cumulative distribution function of Z under the decision parameter θ. In such a setting, many numerical optimization algorithms will then require the efficient computation of the gradient $\nabla\vartheta(\theta)$. In many such settings, no closed form for $F(\cdot;\theta)$ will exist. Rather, one may often have only a simulation algorithm capable of generating a sequence $(Z_i : 1 \leq i \leq n)$ in which the marginal distribution of the Z_i's is $F(\cdot;\theta)$. In such a setting, we are led naturally to the consideration of simulation-based estimators for $\nabla\vartheta(\theta)$.

This paper is concerned with simulation-based computation of $\nabla\vartheta(\theta)$. We note that the risk measures VaR ("value-at-risk") and CVaR ("conditional value-at-risk") are special cases of (1), where $w(y) = \mathbf{1}\{y > 1 - \alpha\}$ and $w(y) = \min\{y/(1-\alpha), 1\}$, respectively. As a result, this paper can be viewed as a generalization of the existing literature on computing gradients (often known as "sensitivities") for VaR and CVaR; see Hong (2009) and Hong and Liu (2009). Moreover, Artzner et al. (1999) and Kusuoka (2001) showed that in general, a coherent risk measure defined by some axioms can be represented as a distortion risk measure under mild regularity conditions. Given our interest in developing a theory that covers VaR and CVaR, the theory will need to accommodate situations in which the distortion function need not be continuously differentiable nor even continuous.

An earlier work by Gourieroux and Liu (2006) studied the sensitivity analysis for distortion risk measures with respect to the parameters in the distortion function and established a central limit theorem for the estimator. However, their method cannot estimate the sensitivities with respect to the parameters in the underlying stochastic models, which is the main focus of our work. The estimator in Gourieroux and Liu (2006) can be viewed as a functional of the empirical distribution function, whereas our estimator is a functional of the distribution sensitivity estimators. The establishment of the central limit theorem for our method entails more sophisticated analysis.

Cao and Wan (2014) provided a sensitivity estimator for distortion risk measures, which is a Riemann integral of the IPA-based quantile sensitivity estimators in Hong (2009). Cao and Wan (2014) assumed continuity in both the distortion function and sample paths in the underlying stochastic model and did not study the asymptotic property of the proposed estimator. For many examples of this work, e.g., the payoff of a barrier option, the sample paths of the the underlying stochastic models are discontinuous with respect to structural parameters (parameters directly appearing in an output function of the stochastic model rather than through the input distributions), e.g., the initial price of the underlying asset. In this case, the pathwise derivative in the IPA-based quantile sensitivity estimator does not exist. In this paper, the continuity assumption in both the distortion function and sample paths in the underlying stochastic model is relaxed, and we offer a new estimator which can be represented by either a Riemann integral or Lebesgue-Stieltjes integral of the generalized likelihood ratio (GLR) estimators for distribution sensitivities in Peng et al. (2020). The GLR method, first studied in Peng et al. (2018), allows for the existence of structural parameters and can handle discontinuous sample paths in the underlying stochastic model, which cannot be handled by either IPA or LR.

In estimating the derivative of an expectation, the theoretical focus is on establishing unbiasedness of the stochastic derivative estimator by justifying the interchange between derivative and expectation. A key challenge has been handing discontinuities in underlying stochastic models with respect to structural parameters, which arise in a wide variety of applications including production/inventory management and financial engineering. The GLR method recently proposed in Peng et al. (2018) is capable of handing a large scope of discontinuities in a general framework. Furthermore, Peng et al. (2020) provide GLR estimators for any order of distribution sensitivities (derivatives of the distribution function of the random performance with respect to both the argument and parameters in underlying stochastic models), which is an expectation of an indicator of the random performance. For complex simulation-based stochastic models, distribution sensitivity estimation plays a central role in quantile sensitivity estimation pioneered by Hong (2009), the statistical inference problem studied in Peng et al. (2020), and sensitivity estimation for distortion risk measures.

Establishing a central limit theorem for estimators is central for both theory and practice, because it provides a basis for hypothesis testing and constructing a confidence interval or region. For stochastic derivative estimators for expectations, this may often be straightforward. However, asymptotic analysis becomes more challenging for estimating sensitivities of quantile and CVaR (Hong 2009, Hong and Liu 2009, Liu and Hong 2009, Fu et al. 2009, Jiang and Fu 2015, Heidergott and Volk-Makarewicz 2016, Lei et al. 2018). Asymptotic analysis is simpler for the batched estimators (Hong 2009, Jiang and Fu 2015, Heidergott and Volk-Makarewicz 2016), but the convergence rates for these estimators are generally worse than their non-batched counterparts (Liu and Hong 2009, Liu and Hong 2009, Fu et al. 2009, Lei et al. 2018). Recently, Peng et al. (2017) and Glynn et al. (2020) provide a unified treatment for establishing a central limit theorem for various types of non-batched quantile sensitivity estimators, which can be viewed as a ratio of the estimates for two distribution sensitivities evaluated at a quantile estimate.

The asymptotic analysis for estimating derivatives of distortion risk measures, which can be viewed as a functional of the empirical processes of the distribution function and distribution sensitivities, is even more complicated than that in quantile sensitivity estimation. We establish a central limit theorem for the proposed sensitivity estimators of the distortion risk measures using a single batch of samples. To the best of our knowledge, our work is the first to establish a functional type limit theory in sensitivity analysis. Our results are established for an integral with the tail of the distribution or quantile truncated. Estimation of quantile and the distribution sensitivities becomes more challenging in the tails of the distribution, because samples become rare. In general, more samples are needed to reach certain accuracy for the proposed sensitivity estimators of the distortion risk measures as the truncation size of the integral become smaller. Our results can be extended straightforwardly for a distribution with compact support or $w(\cdot)$ is zero in neighborhoods containing 0 and 1, e.g., the distortion function of VaR. In many practical problems, the distributions of the random performance indeed have bounded support. For example, investors may have an exit plan when the gain or loss reaches certain boundaries, and the contract of a barrier option would specify a knock-out price to avoid unlimited potential loss in exercising the option. Derivative estimation for the expected payoff of barrier options was studied in Wang et al. (2012) and Peng et al. (2018), whereas in this work, we apply the proposed estimator to estimate sensitivities of VaR and CVaR for the payoffs of barrier options. Since VaR and CVaR provide more information about the tail of a distribution, our proposed method offers a new tool for better hedging tail risk of financial derivatives.

We adopt the convention that all vectors (and vector-valued functions) are written as column vectors, with the exception of gradients (which will be written as row vectors). The rest of the paper is organized as follows. Section 2 provides the general theory for the central limit theorems of two types of estimators. The general theory is applied to gradient estimation for distortion risk measures in Section 3. How to construct a confidence region is discussed in Section 4. Applications are given in Section 5.

2. The General Theory

When the distortion function w is continuously differentiable, we expect that

$$\nabla\vartheta(\theta) = -\int_0^\infty w'(1-F(z;\theta))\nabla_\theta F(z;\theta)dz, \tag{3}$$

assuming that the limit interchange is valid. The interchange of gradient and integration is typically justified by the condition for applying the dominated convergence theorem, e.g., $\int_0^\infty \sup_{\theta\in\Theta}\|w'(1-F(z;\theta))\nabla_\theta F(z;\theta)\|dz<\infty$. Suppose that we have available a simulatable random element (Z,Γ), where Z is a scalar rv and $\Gamma=(\Gamma(x):x\geq 0)$ is an $\mathbb{R}^k$-valued stochastic process for which

$$P(Z\leq z)=F(z;\theta) \tag{4}$$

and

$$E[\Gamma(z)] = \nabla_\theta F(z;\theta)^T \tag{5}$$

for $z\geq 0$. The stochastic process $\Gamma=(\Gamma(x):x\geq 0)$ can be obtained by the GLR method for distribution sensitivities in Peng et al. (2020). We defer the detailed form of the GLR estimators and conditions for justifying the unbiasedness to Section 3. Note that (5) is a statement that $\nabla_\theta F(\cdot;\theta)$ can be estimated consistently by simulating independent and identically distributed (iid) copies of $\Gamma(\cdot)$. In particular, if we simulate iid copies $((Z_n,\Gamma_n(\cdot)):n\geq 1)$ of $(Z,\Gamma(\cdot))$, we can then consistently estimate $F(\cdot,\theta)$ and $E[\Gamma(\cdot)]$ via

$$F_n(z)=\frac{1}{n}\sum_{i=1}^n \mathbf{1}\{Z_i\leq z\}$$

and

$$\overline{\Gamma}_n(z)=\frac{1}{n}\sum_{i=1}^n \Gamma_i(z)$$

for $z\geq 0$ (provided that $\Gamma(z)$ is appropriately integrable), and the gradient $\nabla\vartheta(\theta)$ can be estimated via

$$\delta_{n1} = -\int_0^\infty w'(1-F_n(z))\overline{\Gamma}_n(z)dz. \tag{6}$$

To study the rate of convergence of δ_{n1} to $\nabla\vartheta(\theta)$, and to develop large-sample confidence regions for $\nabla\vartheta(\theta)$, we now discuss the central limit theorem (CLT) for δ_{n1}. If we set

$$X(z)=(\mathbf{1}\{Z\leq z\},\Gamma(z)),$$
$$\overline{X}_n(z)=\left(F_n(z),\overline{\Gamma}_n(z)\right),$$
$$\varphi(x_1,x_2)=-w'(1-x_1)x_2,$$

and

$$\nu(dz) = dz,$$

for $z \geq 0$, we note that δ_{n1} can be represented as a special case of the rv

$$\beta_n = \int_{[a,b]} \varphi\left(\overline{X}_n(z)\right)\nu(dz), \tag{7}$$

where $a = 0$ and $b = \infty$, and $\varphi : \mathbb{R}^{k+1} \to \mathbb{R}^k$. We now provide a general CLT for β_n, and explore sufficient conditions specific to the estimator δ_{n1} in Section 3.

Recall that $\overline{X}_n(z)$ is $\mathbb{R}^{k+1}$-valued, and that $\varphi(\cdot)$ is $\mathbb{R}^k$-valued, with its i'th component given by $\varphi_i(\cdot)$, $1 \leq i \leq k$. We assume that $E[X(\cdot)]$ is continuous on $[a,b]$ and that $\varphi(\cdot)$ is continuously differentiable, with $k \times (k+1)$ Jacobian matrix $(J\varphi)(\cdot)$, on a compact subset $\mathbb{K}_\epsilon$ containing $\{x : \|x - E[X(z)]\| < \epsilon$ for some $z \in [a,b]\}$ for some $\epsilon > 0$. Denote $D_{\mathbb{R}}[a,b]$ as the space comprising of $\mathbb{R}$-valued functions that are right-continuous and have left limits on $[a,b]$.

Theorem 1 *Assume that $-\infty < a < b < \infty$, ν is a finite measure, and that $(X(z) : z \in [a,b]) \in D_{\mathbb{R}^{k+1}}[a,b]$, the space of right-continuous functions with left limits. Suppose that there exists a continuous path Gaussian process $(G(z) : z \in [a,b])$ for which*

$$n^{1/2}\left(\overline{X}_n(\cdot) - E[X(\cdot)]\right) \Rightarrow G(\cdot) \tag{8}$$

as $n \to \infty$, where $\Rightarrow$ denotes weak convergence (in $D_{\mathbb{R}^{k+1}}[a,b]$). Then,

$$n^{1/2}\left(\beta_n - \int_{[a,b]} \varphi\left(E[X(z)]\right)\nu(dz)\right) \Rightarrow \int_{[a,b]} (J\varphi)\left(E[X(z)]\right)G(z)\nu(dz) \tag{9}$$

as $n \to \infty$, where $\Rightarrow$ denotes weak convergence (in $\mathbb{R}^k$).

Proof. We note that $D_{\mathbb{R}^{k+1}}[a,b]$ is a separable metric space under the (standard) J_1 topology; see p. 381 of Whitt (2002). We may therefore apply the Skorohod representation theorem; see p. 102 of Ethier and Kurtz (2009). Accordingly we may replace the weak convergence in (8) with almost sure (a.s.) convergence in the J_1 metric. Since the limit $G(\cdot)$ is continuous, convergence in the J_1 metric implies that the a.s. convergence in (8) corresponds to uniform convergence of the left-hand side to $G(\cdot)$; see p. 112 of Billingsley (1968). Consequently, $\overline{X}_n(z)$ converges a.s. to $E[X(z)]$ uniformly in $z \in [a,b]$. It follows that $\overline{X}_n(\cdot) \in \mathbb{K}_\epsilon$ for n sufficiently large, so that Taylor's theorem implies that

$$n^{1/2}\left(\varphi_i\left(\overline{X}_n(z)\right) - \varphi_i\left(E[X(z)]\right)\right) = n^{1/2}\nabla\varphi_i\left(\xi_{ni}(z)\right)\left(\overline{X}_n(z) - E[X(z)]\right),$$

where $\xi_{ni}(z)$ lies on the line segment connecting $\overline{X}_n(z)$ and $E[X(z)]$. Since $\overline{X}_n(\cdot)$ converges uniformly a.s. to $E[X(\cdot)]$, the same is true of $\xi_{ni}(\cdot)$. Hence,

$$n^{1/2}\nabla\varphi_i\left(\xi_{ni}(z)\right)\left(\overline{X}_n(z)-E[X(z)]\right)\to\nabla\varphi_i(E[X(z)])G(z)$$

uniformly a.s. in $z\in[a,b]$, for $1\le i\le k$. Because $\nu(\cdot)$ is a finite measure, the bounded convergence theorem implies that

$$\int_{[a,b]} n^{1/2}\nabla\varphi_i\left(\xi_{ni}(z)\right)\left(\overline{X}_n(z)-E[X(z)]\right)\nu(dz)\to\int_{[a,b]}\nabla\varphi_i(E[X(z)])G(z)\nu(dz)$$

for $1\le i\le k$, yielding

$$\int_{[a,b]} n^{1/2}\left(\varphi\left(\overline{X}_n(z)\right)-\varphi(E[X(z)])\right)\nu(dz)\to\int_{[a,b]}(J\varphi)(E[X(z)])G(z)\nu(dz)$$

a.s. as $n\to\infty$. This immediately yields the theorem. □

Noting that $\overline{X}_n(\cdot)$ is an average of iid copies of $X(\cdot)$, it is obvious that if $E[||X(z)||^2]<\infty$ for $z\in[a,b]$, the CLT for iid (finite-dimensional) random vectors guarantees that the finite-dimensional distributions of $n^{1/2}\left(\overline{X}_n(\cdot)-E[X(\cdot)]\right)$ must converge weakly to those of $G(\cdot)$, so that the Gaussian process $G(\cdot)$ of Theorem 1 must be such that

$$\begin{aligned}&E[G(z_1)]=0\\&E[G(z_1)G(z_2)^T]=E[X(z_1)X(z_2)^T]-E[X(z_1)]E[X(z_2)]^T\end{aligned}$$

for $z_1,z_2\in[a,b]$. As for the limit object in (9), one can approximate it via a Riemann-Stieltjes approximation. Such an approximation is clearly Gaussian, so that by taking limits, we conclude that the right-hand side of (9) is Gaussian with mean zero and covariance matrix

$$\int_{[a,b]}\int_{[a,b]}(J\varphi)(E[X(y)])\left[E[X(y)X(z)^T]-E[X(y)]E[X(z)]^T\right](J\varphi)^T(E[X(z)])\nu(dy)\nu(dz).$$

We turn next to sufficient conditions ensuring the validity of (8). In particular, we will now discuss how to improve the just-discussed finite-dimensional weak convergence to weak convergence in $D_{\mathbb{R}^{k+1}}[a,b]$. For this purpose, we can follow two different approaches.

Approach 1: We can apply CLT's for averages of iid $D_{\mathbb{R}}[a,b]$-valued random variables to each component of $n^{1/2}\left(\overline{X}_n(\cdot)-E[X(\cdot)]\right)$; see, for example, Theorem 1 of Bloznelis and Paulauskas (1994). We claim that this implies that (8) holds.

To see this, note that the weak convergence of each component implies that each component is tight in $D_{\mathbb{R}}[a,b]$. Consequently, $\left(n^{1/2}\left(\overline{X}_n(\cdot)-E[X(\cdot)]\right):n\ge1\right)$ is tight in the product space $D_{\mathbb{R}}[a,b] \times \cdots \times D_{\mathbb{R}}[a,b]$ ($k+1$ times); see p. 390 of Whitt (2002). Because the finite-dimensional distributions of $n^{1/2}\left(\overline{X}_n(\cdot) - E[X(\cdot)]\right)$ converge weakly to those of $G(\cdot)$, it follows that $n^{1/2}\left(\overline{X}_n(\cdot) - E[X(\cdot)]\right)$ converges weakly to $G(\cdot)$ in the product topology on the product space.

We now apply the Skorohod representation theorem, so that we can replace the weak convergence by a.s. convergence in the product topology. Since $G(\cdot)$ is continuous, it follows that each component of $n^{1/2}\left(\overline{X}_n(\cdot) - E[X(\cdot)]\right)$ converges uniformly a.s. to the corresponding component of $G(\cdot)$. Consequently, $n^{1/2}\left(\overline{X}_n(\cdot) - E[X(\cdot)]\right)$ converges uniformly a.s. to $G(\cdot)$, thereby implying (8).

Approach 2: We can apply empirical process theory to the collection of rv's $\left(\overline{X}_n(z) : z \in [a,b]\right)$. In particular, suppose that $\mathcal{G} = \left\{\overline{X}_n(z) : z \in [a,b]\right\}$ is a P-Donsker class; see p. 81 of van der Vaart and Wellner (1996). Then,

$$n^{1/2}\left(\overline{X}_n(\cdot) - E[X(\cdot)]\right) \Rightarrow G(\cdot)$$

in $\ell^\infty(\mathcal{G})$ which is a collection of bounded functions from $\mathcal{G}$ to $\mathbb{R}$, equipped with the sup (infinity) norm. The weak convergence in $\ell^\infty(\mathcal{G})$ to a continuous limit clearly implies weak convergence in the Skorohod J_1 metric because the uniform metric is finer than the J_1 metric; see p 150 of Billingsley (1968). Note that we have not assumed so far that $G(\cdot)$ is continuous. If $G(\cdot)$ is now assumed to be continuous, then weak convergence in $\ell^\infty(\mathcal{G})$ implies uniform convergence in $z \in [a,b]$ to a $D_{\mathbb{R}^k}[a,b]$-valued random process, so that (8) holds.

With the weak convergence in $\ell^\infty(\mathcal{G})$, there is a systematic approach to establish CLT for a functional of a random element, given that the functional is Hadamard differentiable; see p. 296 of Van der Vaart (2000). Estimator (7) can be viewed as a functional $\mathcal{F}(\cdot)$ of the stochastic process $\overline{X}_n(\cdot)$, where

$$\mathcal{F}(\eta) = \int_{[a,b]} \varphi\left(\eta(z)\right) \nu(dz),$$

where $\eta(\cdot)$ is a $D_{\mathbb{R}^{k+1}}[a,b]$-valued function. By applying the mean value theorem to the integration function, it is straightforward to show that

$$\begin{aligned}
\mathcal{F}'_\eta(\tau) &= \lim_{\epsilon \to 0} \frac{1}{\epsilon}\left[\mathcal{F}(\eta + \epsilon\tau_\epsilon) - \mathcal{F}(\eta)\right] \\
&= \lim_{\epsilon \to 0} \frac{1}{\epsilon}\left[\int_{[a,b]} \varphi\left(\eta(z) + \epsilon\tau_\epsilon(z)\right) \nu(dz) - \int_{[a,b]} \varphi\left(\eta(z)\right) \nu(dz)\right] \\
&= \int_{[a,b]} \left(J\varphi\right)(\eta(z))\tau(z)\nu(dz),
\end{aligned}$$

where $\tau_\epsilon(\cdot)$ is a $D_{\mathbb{R}^{k+1}}[a,b]$-valued function converging to a $C_{\mathbb{R}^{k+1}}[a,b]$-valued function $\tau(\cdot)$ in the uniform norm of $\ell^\infty(\mathcal{G})$ as $\epsilon \to 0$. By the functional delta method (see p. 297 of Van der Vaart 2000),

$$n^{1/2}\left[\mathcal{F}\left(\eta + \frac{1}{n^{1/2}}\overline{G}_n\right) - \mathcal{F}(\eta)\right] \Rightarrow \mathcal{F}'_\eta(G), \tag{10}$$

where $\eta(\cdot) = E[X(\cdot)]$ and $\overline{G}_n(\cdot) = n^{1/2}(\overline{X}_n(\cdot) - \eta(\cdot))$. Moreover, note that $\mathcal{F}'_\eta(G)$ is the same as the right hand of (9). Therefore, the CLT (10) is essentially the result of Theorem 1.

Our theory above gives a general framework for the analysis of gradient estimators in which the distortion function is continuously differentiable. However, as pointed out in the Introduction, this does not cover the distortions that arise in the setting of VaR or CVaR. To handle such distortion functions, we use the fact that we can re-write (2) as

$$-\int_{[0,1]} F^{-1}(y;\theta)\tilde{w}(dy),$$

where $\tilde{w}(y) = w(1-y)$, when $w(\cdot)$ is non-decreasing and left-continuous; see Dhaene et al. (2012). Assuming that $F^{-1}(z,\cdot)$ is suitably smooth and that the gradient and integral can be interchanged, we arrive at the expression

$$\nabla\vartheta(\theta) = -\int_{[0,1]} \nabla_\theta F^{-1}(y;\theta)\tilde{w}(dy).$$

The implicit function theorem then reveals that if $F(\cdot;\theta)$ has a positive continuous density $f(\cdot;\theta)$,

$$\nabla_\theta F^{-1}(y;\theta) = -\frac{\nabla_\theta F(F^{-1}(y;\theta);\theta)}{f(F^{-1}(y;\theta);\theta)},$$

which is the foundation for calculating the IPA estimator (Suri and Zazanis 1988). For input random variables, the distribution is known, but for output random variables, the numerator and denominator usually do not have analytical forms and must be estimated by simulation, by the two distribution sensitivity estimates in this work. Assume that we have a simulatable pair $(Z,(\chi(z): z \geq 0))$ and a deterministic function $\kappa: \mathbb{R}^{k+1} \to \mathbb{R}^k$ for which

$$\kappa(E[\chi(z)]) = -\frac{\nabla_\theta F(z;\theta)^T}{f(z;\theta)}$$

and $P(Z \leq z) = F(z;\theta)$ for $z \geq 0$. Similarly, the stochastic process $(\chi(z): z \geq 0)$ can be obtained by the GLR method for distribution sensitivities in Peng et al. (2020), and the detailed discussion is deferred to Section 3, where the specific form of κ can also be found. By generating iid copies $((Z_i,(\chi_i(z): z \geq 0)): i \geq 1)$ of $(Z,(\chi(z): z \geq 0))$, we can therefore estimate $\nabla\vartheta(\theta)$ via the estimator

$$\delta_{n2} = \int_{[0,1]} \kappa\left(\overline{\chi}_n(F_n^{-1}(y))\right)\tilde{w}(dy), \tag{11}$$

where

$$\overline{\chi}_n(z) = \frac{1}{n}\sum_{i=1}^{n}\chi_i(z)$$

and $F_n(\cdot)$ is again just the empirical cumulative distribution function of $Z_1, Z_2, \ldots, Z_n$ given by $F_n(z) = n^{-1}\sum_{i=1}^{n}\mathbf{1}\{Z_i \le z\}$.

As with the estimator δ_{n1}, we will now discuss a CLT for a more general estimator γ_n, for which δ_{n2} will be a special case. Suppose that Y is a scalar rv with cumulative distribution function $H(\cdot)$, and that $W = (W(z) : z \in [c,d])$ is a $D_{\mathbb{R}^{k+1}}[c,d]$-valued rv. Let $((Y_i, W_i) : i \ge 1)$ be a sequence of iid copies of (Y, W), and set

$$H_n(y) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}\{Y_i \le y\}$$
$$\overline{W}_n(z) = \frac{1}{n}\sum_{i=1}^{n}W_i(z)$$

for $y \in \mathbb{R}$ and $z \in [c,d]$. For $0 < u < 1$, put $H_n^{-1}(u) = \inf\{y : H_n(y) \ge u\}$, and for $0 \le a < b \le 1$, set

$$\gamma_n = \int_{[a,b]} \phi\left(\overline{W}_n(H_n^{-1}(u))\right)\nu(du) \tag{12}$$

for some finite measure $\nu(\cdot)$, for some deterministic $\phi : \mathbb{R}^{k+1} \to \mathbb{R}^k$. Clearly, δ_{n2} is a special case of γ_n.

We assume that $\omega(\cdot) = E[W(\cdot)]$ is continuously differentiable on $[c,d]$ with derivative $\omega'(\cdot)$ and that $H(\cdot)$ has a continuous positive density $h(\cdot)$ on $[a,b]$. We also require the existence of $\epsilon > 0$ for which $c \le H^{-1}(a) - \epsilon < H^{-1}(b) + \epsilon \le d$ and that $\phi(\cdot)$ is continuously differentiable, with $k \times (k+1)$ Jacobian matrix $(J\phi)(\cdot)$, on a compact set $\mathbb{K}_\epsilon$ containing $\{x \in \mathbb{R}^{k+1} : ||x - \omega(z)|| < \epsilon$ for some $z \in [c,d]\}$.

Theorem 2 *Assume that $0 < a < b < 1$, and that there exists an $\mathbb{R}^{k+2}$-valued continuous path Gaussian process $G = ((G_1(z), G_2(z)) : z \in [c,d])$ such that*

$$n^{1/2}\left(\overline{W}_n(\cdot) - \omega(\cdot), H_n(\cdot) - H(\cdot)\right) \Rightarrow (G_1(\cdot), G_2(\cdot)) \tag{13}$$

as $n \to \infty$, where $\Rightarrow$ denotes weak convergence (in $D_{\mathbb{R}^{k+2}}[c,d]$). Then,

$$\begin{aligned} & n^{1/2}\left(\gamma_n - \int_{[a,b]}\phi(\omega(H^{-1}(u)))\nu(du)\right) \\ & \Rightarrow \int_{[a,b]}(J\phi)(\omega(H^{-1}(u)))\left[G_1(H^{-1}(u)) - \omega'(H^{-1}(u))\frac{G_2(H^{-1}(u))}{h(H^{-1}(u))}\right]\nu(du) \end{aligned} \tag{14}$$

as $n \to \infty$, where $\Rightarrow$ denotes weak convergence (in $\mathbb{R}^k$).

Proof. Because $0<a<b<1$, we may, without loss of generality, assume that $-\infty<c<d<\infty$. As in the proof of Theorem 1, we apply the Skorohod representation theorem to replace weak convergence by a.s. convergence.

Because $n^{1/2}(H_n(\cdot)-H(\cdot))$ converges uniformly to $G_2(\cdot)$ on $[c,d]$ and $c\le H^{-1}(a)-\epsilon<H^{-1}(b)+\epsilon\le d$, it follows that $H_n^{-1}(\cdot)\to H^{-1}(\cdot)$ uniformly on $[a,b]$. Furthermore,

$$n^{1/2}\left(H_n(H_n^{-1}(u))-H(H_n^{-1}(u))\right)=G_2(H_n^{-1}(u))+o(1) \tag{15}$$

uniformly in $u\in[a,b]$, where $o(b_n)$ is a term for which $o(b_n)/b_n\to 0$ as $n\to\infty$. Since $G_2(\cdot)$ is continuous,

$$G_2(H_n^{-1}(u))=G_2(H^{-1}(u))+o(1), \tag{16}$$

uniformly in $u\in[a,b]$. Also, because $H(\cdot)$ is a continuous distribution, the jumps in $H_n(\cdot)$ are all of size n^{-1}, so $H_n(H_n^{-1}(u))=u+o(n^{-1/2})$ uniformly in $u\in[a,b]$. Consequently, (15) and (16) imply that

$$n^{1/2}\left(H(H_n^{-1}(u))-u\right)=-G_2(H^{-1}(u))+o(1), \tag{17}$$

uniformly in $u\in[a,b]$. But Taylor's theorem implies that

$$\begin{aligned}n^{1/2}\left(H(H_n^{-1}(u))-u\right)&=n^{1/2}\left(H(H_n^{-1}(u))-H(H^{-1}(u))\right)\\&=h(H^{-1}(u))n^{1/2}(H_n^{-1}(u)-H^{-1}(u))+o(1)\end{aligned}$$

uniformly in $u\in[a,b]$, so that (17) yields

$$n^{1/2}(H_n^{-1}(u)-H^{-1}(u))=-\frac{G_2(H^{-1}(u))}{h(H^{-1}(u))}+o(1) \tag{18}$$

uniformly in $u\in[a,b]$.

As in the proof of Theorem 1, $\left(\overline{W}_n(H_n^{-1}(u)):u\in[a,b]\right)\in K_\epsilon$ for n sufficiently large, so that

$$n^{1/2}\left(\phi\left(\overline{W}_n(H_n^{-1}(u))\right)-\phi(\omega(H^{-1}(u)))\right)=(J\phi)(\omega(H^{-1}(u)))n^{1/2}\left[\overline{W}_n(H_n^{-1}(u))-\omega(H^{-1}(u))\right]+o(1)$$

uniformly in $u\in[a,b]$. But

$$\begin{aligned}n^{1/2}\left[\overline{W}_n(H_n^{-1}(u))-\omega(H^{-1}(u))\right]&=n^{1/2}\left[\overline{W}_n(H_n^{-1}(u))-\omega(H_n^{-1}(u))\right]+n^{1/2}\left[\omega(H_n^{-1}(u))-\omega(H^{-1}(u))\right]\\&=G_1(H_n^{-1}(u))+\omega'(H^{-1}(u))n^{1/2}\left[H_n^{-1}(u)-H^{-1}(u)\right]+o(1)\\&=G_1(H^{-1}(u))+\omega'(H^{-1}(u))n^{1/2}\left[H_n^{-1}(u)-H^{-1}(u)\right]+o(1) \quad (19)\end{aligned}$$

uniformly in $u\in[a,b]$, since $G_1(\cdot)$ is continuous.

It follows from (18) and (19) that

$$n^{1/2}\left(\phi\left(\overline{W}_n(H_n^{-1}(u))\right)-\phi(\omega(H^{-1}(u)))\right)\to(J\phi)(\omega(H^{-1}(u)))\left[G_1(H^{-1}(u))-\omega'(H^{-1}(u))\frac{G_2(H^{-1}(u))}{h(H^{-1}(u))}\right]$$

uniformly in $u\in[a,b]$. The bounded convergence theorem then proves that

$$\begin{aligned}&\int_{[a,b]} n^{1/2}\left(\phi\left(\overline{W}_n(H_n^{-1}(u))\right)-\phi(\omega(H^{-1}(u)))\right)\nu(du)\\ &\to\int_{[a,b]}(J\phi)(\omega(H^{-1}(u)))\left[G_1(H^{-1}(u))-\omega'(H^{-1}(u))\frac{G_2(H^{-1}(u))}{h(H^{-1}(u))}\right]\nu(du)\end{aligned}$$

as $n\to\infty$, establishing the theorem. □

As with Theorem 1, the limit appearing in (14) must be a mean zero $\mathbb{R}^k$-valued Gaussian rv with covariance matrix

$$\int_{[a,b]}\int_{[a,b]}(J\phi)(\omega(H^{-1}(u_1)))\left[E[V(u_1)V(u_2)^T]-E[V(u_1)]E[V(u_2)]\right](J\phi)^T(\omega(H^{-1}(u_2)))\nu(du_1)\nu(du_2),$$

where

$$V(u)=W(H^{-1}(u))-\omega'(H^{-1}(u))\frac{\mathbf{1}\{Y\le H^{-1}(u)\}}{h(H^{-1}(u))}.$$

Also, as with Theorem 1, we have two different approaches to establishing the key hypothesis (13), one based on existing CLT's for $D_{\mathbb{R}}[c,d]$-valued rv's and another based on application of empirical process theory.

With the weak convergence in ℓ^∞, we can also potentially establish Theorem 2 via the functional delta method. The inverse map can be viewed a functional $\mathcal{Q}(\cdot)$ of a function. From p. 307 of Van der Vaart (2000), the Hadamard derivative is

$$\mathcal{Q}_H'(\tau_2)(u)=-\frac{\tau_2(\mathcal{Q}(H)(u))}{h(\mathcal{Q}(H)(u))},\qquad u\in[a,b],\tag{20}$$

where $\tau_2(\cdot)$ is a $C_{\mathbb{R}}[a,b]$-valued function. Estimator (12) can be viewed as a functional $\mathcal{H}(\cdot)$ of $\Pi_n(\cdot)=(\overline{W}_n(\cdot),H_n(\cdot))$, where

$$\mathcal{H}(\pi)=\int_{[a,b]}\phi\left(\omega(\mathcal{Q}(H)(u))\right)\nu(du),$$

where $\pi(\cdot)=(\omega(\cdot),H(\cdot))$. By applying the mean value theorem to the integration function,

$$\begin{aligned}\mathcal{H}_\pi'(\zeta)&=\lim_{\epsilon\to0}\frac{1}{\epsilon}\left[\mathcal{H}(\pi+\epsilon\zeta_\epsilon)-\mathcal{H}(\pi)\right]\\ &=\lim_{\epsilon\to0}\frac{1}{\epsilon}\left[\int_{[a,b]}\phi\left((\omega+\epsilon\tau_{1,\epsilon})(\mathcal{Q}(H+\epsilon\tau_{1,\epsilon})(u))\right)\nu(du)-\int_{[a,b]}\phi\left(\omega(\mathcal{Q}(H)(u))\right)\nu(du)\right]\\ &=\int_{[a,b]}(J\phi)(\omega(\mathcal{Q}(H)(u)))\left[\tau_1(\mathcal{Q}(H)(u))-\omega'(\mathcal{Q}(H)(u))\frac{\tau_2(\mathcal{Q}(H)(u))}{h(\mathcal{Q}(H)(u))}\right]\nu(du),\end{aligned}$$

where $\zeta_\epsilon(\cdot) = (\tau_{1,\epsilon}(\cdot), \tau_{2,\epsilon}(\cdot))$ is a $D_{\mathbb{R}^{k+2}}[a,b]$-valued function converging to a $C_{\mathbb{R}^{k+2}}[a,b]$-valued function $\zeta(\cdot) = (\tau_1(\cdot), \tau_2(\cdot))$ in the uniform norm as $\epsilon \to 0$. By the functional delta method,

$$n^{1/2}\left[\mathcal{H}\left(\pi + \frac{1}{n^{1/2}}\overline{G}_n\right) - \mathcal{H}(\pi)\right] \Rightarrow \mathcal{H}'_\pi(G), \tag{21}$$

where $\overline{G}_n(\cdot) = n^{1/2}\left(\overline{W}_n(\cdot) - \omega(\cdot), H_n(\cdot) - H(\cdot)\right)$. The CLT (21) implies Theorem 2.

3. Application to Gradient Estimation for Distortion Risk Measures

We now specialize the general theory of Section 2 to distortion risk measure gradients. Specifically, the simulatable stochastic processes $\Gamma(\cdot)$ and $\chi(\cdot)$ assumed in Section 2 will be given by applying GLR estimators for distribution sensitivities in Peng et al. (2020), and conditions to justify unbiasedness of the GLR estimators and weak convergence of simulatable stochastic processes (8) and (13) in Theorems 1 and 2 will be provided. In the setting of δ_{n1}, Theorem 1 requires modifying the estimator (slightly) to

$$\delta_{n1}(b) = -\int_0^b w'(1 - F_n(z))\overline{\Gamma}_n(z)dz \tag{22}$$

for $b < \infty$. When the dependence of Z on θ arises purely through the distribution of an input rv in with density $q(\cdot;\theta)$ known and support independent of θ, then $F(z;\theta) = E[\mathbf{1}\{Z \le z\}\lambda(\Upsilon;\theta)]$, where $\Upsilon \in \mathbb{R}^m$ is a rv with the density $q(\cdot;\theta_0)$, for some likelihood ratio

$$\lambda(\upsilon;\theta) = \frac{q(\upsilon;\theta)}{q(\upsilon;\theta_0)},$$

where $\upsilon = (\upsilon_1, \dots, \upsilon_m)$, and $\Gamma(z) = \mathbf{1}\{Z \le z\}\nabla_\theta\lambda(\Upsilon;\theta)^T$. However, in many applications

$$F(z;\theta) = E[\mathbf{1}\{g(\Upsilon;\theta) \le z\}\lambda(\Upsilon;\theta)] \tag{23}$$

for some function $g : \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}$, so that the dependence of $Z = g(\Upsilon;\theta)$ on θ is affected not only by θ in the distribution of input rv, but also the structural dependence on θ through the function g. In this setting, one can appeal to the GLR gradient estimators. Suppose that the following technical conditions on q and g hold:

(C.1) Inverse function $g^{-1}(\cdot, \upsilon_{-1};\theta)$ of g w.r.t. the first argument exists for all $\upsilon_{-1} = (\upsilon_2, \dots, \upsilon_m)$.

(C.2) There exists $\epsilon > 0$ such that $|(\partial g(\upsilon;\theta)/\partial\upsilon_1)^{-1}| > \epsilon$.

(C.3) Moment condition:

$$E\left[\left|\frac{\partial^2 g}{\partial\upsilon_1^2}(\Upsilon;\theta)\right|\right] < \infty\ .$$

(C.4) Condition on density:

$$\lim_{\upsilon_1 \to \pm\infty} q_1(\upsilon_1;\theta) = 0, \quad q(\upsilon;\theta) < \infty,$$

where q_1 is the marginal density of the first coordinate in Υ.

For a simple example $g(\upsilon;\theta)=\upsilon_1+\theta\upsilon_2$, its inverse function w.r.t. the respect to the first argument is $g^{-1}(\cdot,\upsilon_{-1};\theta)=\cdot-\theta\upsilon_2$. Conditions (C.1) and (C.2) can justify that g^{-1} is globally Lipchitz continuous with respect to the argument. When g is a linear function of υ, $\partial g(\upsilon;\theta)/\partial\upsilon_1$ is a constant so condition (C.2) holds, and $\partial^2 g(\upsilon;\theta)/\partial\upsilon_1^2$ is zero so condition (C.3) above holds. Condition (C.4) holds for most distributions supported on the whole space, e.g., the exponential family distributions. In Peng et al. (2020), it is shown that when conditions (C.1)-(C.4) are in force, we have

$$\Gamma(z)=(\mathbf{1}\{Z\le z\}A_1,\dots,\mathbf{1}\{Z\le z\}A_k)^T,$$

where, for $1\le j\le k$,

$$\begin{aligned}A_j(\Upsilon)=&\frac{\partial\lambda}{\partial\theta_j}(\Upsilon;\theta)-\lambda(\Upsilon;\theta)\left(\frac{\partial g}{\partial\upsilon_1}(\Upsilon;\theta)\right)^{-1}\\&\times\left[\frac{\partial g}{\partial\theta_j}(\Upsilon;\theta)\left(\frac{\partial\log q}{\partial\upsilon_1}(\Upsilon;\theta)-\frac{\partial^2 g}{\partial\upsilon_1^2}(\Upsilon;\theta)\left(\frac{\partial g}{\partial\upsilon_1}(\Upsilon;\theta)\right)^{-1}\right)+\frac{\partial^2 g}{\partial\theta_j\partial\upsilon_1}(\Upsilon;\theta)\right],\end{aligned}\tag{24}$$

in which case $E[\Gamma(z)]=\nabla_\theta F(z;\theta)^T$. When applying Theorem 1, we now take $\varphi:\mathbb{R}^{k+1}\to\mathbb{R}^k$ so that $\varphi(x_0,x_1,\dots,x_k)=(w'(1-x_0)x_1,\dots,w'(1-x_0)x_k)^T$ and

$$\Gamma(z)=(\mathbf{1}\{Z\le z\},\mathbf{1}\{Z\le z\}A_1,\dots,\mathbf{1}\{Z\le z\}A_k)^T.$$

To validate the hypotheses of Theorem 1, we now require that w be twice differentiable and $E[A_j^2]<\infty$ for $1\le j\le k$, and turn to verification of (8), either via Approach 1 or Approach 2.

Verification via Approach 1: We apply Theorem 1 of Bloznelis and Paulauskas (1994) to the process $\mathbf{1}\{Z\le\cdot\}A_j$ (where $A_0\equiv 1$), and note that if $E[A_j^2\mathbf{1}\{Z\le\cdot\}]$ is continuously differentiable, then $E[(\mathbf{1}\{Z\le z_2\}A_j-\mathbf{1}\{Z\le z_1\}A_j)^2]=E[A_j^2\mathbf{1}\{Z\le z_2\}]-E[A_j^2\mathbf{1}\{Z\le z_1\}]$ for $z_1\le z_2$. The differentiability then implies this latter quantity is of order z_2-z_1 as $z_2\downarrow z_1$, so that (1.5) of their Theorem 1 follows. (Hypothesis (1.4) is immediate in this setting.)

Verification via Approach 2: Put $K_l=\mathbf{1}\{Z\le z_l\}A_1$ and note that, without loss of generality, we may assume $A_1\ge 0$ (for otherwise we can work with the positive and negative parts of A_1). Note that if $0=z_0<z_1<\cdots<z_t=\infty$, then the K_l's bracket the family $\mathcal{C}_1=\{\mathbf{1}\{Z\le z\}A_1:z\ge 0\}$, in the sense that $K_l\le\mathbf{1}\{Z\le z\}A_1\le K_{l+1}$ for $z\in[z_l,z_{l+1}]$. By choosing $t=\lceil E[A_1^2]/\epsilon^2\rceil+1$ and appropriately selecting the z_l's, we can guarantee that

$$\sqrt{E[(K_{l+1}-K_l)^2]}=\sqrt{E[A_1^2\mathbf{1}\{z_l<Z\le z_{l+1}\}]}\le\epsilon.$$

The ϵ-bracket number $N_{[]}(\epsilon, \mathcal{C}_1, \mathbb{L}^2)$ is the minimal number of ε-brackets needed to cover $\mathcal{C}_1$, Consequently, $N_{[]}(\epsilon, \mathcal{C}_1, \mathbb{L}^2) \leq \lceil E[A_1^2]/\epsilon^2 \rceil + 1$, so that

$$\int_0^\infty \sqrt{\log N_{[]}(\epsilon, \mathcal{C}_1, \mathbb{L}^2)} d\epsilon < \infty.$$

It follows that $\mathcal{C}_1$ is P-Donsker; see p. 85 of van der Vaart and Wellner (1996). Similarly, $\mathcal{C}_j = \{\mathbf{1}\{Z \leq z\} A_j : z \geq 0\}$ is P-Donsker for $0 \leq j \leq k$, so that $\mathcal{C} = \{\mathbf{1}\{Z \leq z\} A_j : z \geq 0, 0 \leq j \leq k\}$ is P-Donsker; see Theorem 2.10.3 of van der Vaart and Wellner (1996). Let Υ_i, $i = 1, \ldots, n$, be iid copies of Υ such that $Z_i = g(\Upsilon_i; \theta)$, $i = 1, \ldots, n$. Hence,

$$n^{1/2} \left(\frac{1}{n} \sum_{i=1}^n \mathbf{1}\{Z_i \leq \cdot\} A_j(\Upsilon_i) - E[\mathbf{1}\{Z \leq \cdot\} A_j] : 0 \leq j \leq k \right)^T \Rightarrow G(\cdot)$$

in $\ell^\infty(\mathcal{C})$, yielding (8).

We note that the A_j's defined in (24) involve partial derivatives of g with respect to the first argument. Under hypotheses (C.1)-(C.4), similar estimators can be defined in terms of partial derivatives of g taken with respect to x_2 (or, indeed, any υ_j, for $1 \leq j \leq m$). Since all m such estimators are unbiased for $\nabla_\theta F(z;\theta)$, any linear combination of the m estimators for which the coefficients sum to 1 is also unbiased as an estimator for $\nabla_\theta F(z;\theta)$. Of course, these more general estimators also can be analyzed as special cases of Theorem 1. Summarizing the discussion above, we have the following proposition.

Proposition 1 *If $E[A_j^2] < \infty$, $j = 1, \ldots, j$, then assumption (8) in Theorem 1 holds.*

We turn next to the analysis of estimators for $\nabla \vartheta(\theta)$ when the distortion function w is non-smooth. As for the estimator δ_{n1}, we (slightly) modify the estimator to take the form

$$\delta_{n2}(\epsilon) = \int_{[\epsilon, 1-\epsilon]} \kappa \left(\overline{\chi}_n (F_n^{-1}(y)) \right) \tilde{w}(dy) \tag{25}$$

for $\epsilon > 0$. In this case, we note that

$$\kappa(x_0, x_1, \ldots, x_k) = (x_1/x_0, \ldots, x_k/x_0)^T$$

and

$$\chi(z) = (\mathbf{1}\{Z \leq z\} B, \mathbf{1}\{Z \leq z\} A_1, \ldots, \mathbf{1}\{Z \leq z\} A_k)^T,$$

where A_j is as before, and

$$B(\Upsilon) = \lambda(\Upsilon;\theta) \left(\frac{\partial g}{\partial \upsilon_1}(\Upsilon;\theta) \right)^{-1} \left(\frac{\partial \log q}{\partial \upsilon_1}(\Upsilon;\theta) - \frac{\partial^2 g}{\partial \upsilon_1^2}(\Upsilon;\theta) \left(\frac{\partial g}{\partial \upsilon_1}(\Upsilon;\theta) \right)^{-1} \right).$$

Recall that $(E[\mathbf{1}\{Z \leq z\}A_1], \ldots, E[\mathbf{1}\{Z \leq z\}A_k]) = \nabla_\theta F(z;\theta)$. As shown in Peng et al. (2020) and Peng et al. (2018),

$$E[\mathbf{1}\{Z \leq z\}B] = f(z;\theta)$$

when $f(z;\theta)$ is positive and continuously differentiable for $z \in [c,d]$, and g is appropriately smooth. As for (13), this can be addressed via either Approach 1 or Approach 2, precisely as for Theorem 1, in the presence of the additional hypothesis $E[B^2] < \infty$. Since the mapping ϕ has the differentiability required, Theorem 2 therefore can be applied to $\delta_{n2}(\epsilon)$. Along the way, further estimators of $\nabla\vartheta(\theta)$ can be obtained by taking linear combinations of estimators of $\nabla_\theta F(z;\theta)$ obtained from (24) by using partial derivatives of g taken with respect to $\upsilon_2, \upsilon_3, \ldots, \upsilon_m$. Summarizing the discussion above, we have the following proposition.

Proposition 2 *If $E[B^2] < \infty$, then assumption (13) in Theorem 2 holds.*

For a distribution with compact support or $w(\cdot)$ is zero in neighborhoods containing 0 and 1, e.g., the distortion function of VaR, central limit theorems in Theorems 1 and 2 can be established analogously in these special cases. We conclude this section with a brief discussion of the technical difficulties associated with rigorous proof of Theorems 1 and 2 when $b = \infty$ in $\delta_{n1}(b)$ or $\epsilon = 0$ in $\delta_{n2}(\epsilon)$. To obtain a rigorous proof of the CLT for $\delta_{n1}(b)$ with $b = \infty$, it is no longer sufficient to establish weak convergence in $D_{\mathbb{R}^k}[0,\infty)$. The problem is that the topology of uniform convergence on compact subintervals of $[0,\infty)$ associated with a continuous limit processes is not sufficient to guarantee that the integral over $[0,\infty)$ converges to that of the limit process. In particular, more control over the behavior of $\varphi(\overline{X}_n(z))$ as $z \to \infty$ would be needed, so that some version of the dominated convergence theorem could be applied. Similar issues arise with $\delta_{n2}(\epsilon)$. In particular, $h(H^{-1}(u)) \to 0$ as $u \to 1$ (and often also as $u \to 0$), so that there are singularities that need to be controlled as $\epsilon \downarrow 0$; see (14).

4. Large-Sample Confidence Regions

Central limit theorems provide a basis for hypothesis testing and constructing confidence intervals or regions. In this section, we briefly discuss the construction of large-sample confidence regions for the estimators described in Theorems 1 and 2. For estimator (11), $\omega'(\cdot)$ in (14) involves higher order distribution sensitivities. Peng et al. (2020) provided the GLR estimator for any distribution sensitivity, which is a multiplication of the indicator and a weight function like $\chi(z)$. The easiest approach is to use the (multivariate) method of batch means; see Munoz and Glynn (2001). This method avoids the need to compute $(J\varphi)(\cdot)$ or $(J\phi)(\cdot)$, or to estimate the density $h(\cdot)$ and its derivatives appearing in the limit distribution for Theorem 2, as would occur if one constructs the confidence regions via consistent estimation of the covariance matrices appearing in the limit distributions in Theorems 1 and 2.

For $p \geq k+1$, we assume that our total sample size $n = lp$ for some (large) integer l. In addition to the estimators β_n and γ_n, we also compute the p "batch means" estimators

$$\beta_{ni} = \int_{[a,b]} \varphi\left(\overline{X}_{ni}(z)\right) \nu(dz)$$

and

$$\gamma_{ni} = \int_{[a,b]} \phi\left(\overline{W}_{ni}(H_{ni}^{-1}(u))\right) \nu(du)$$

for $1 \leq i \leq p$, where

$$\overline{X}_{ni}(z) = \frac{1}{l} \sum_{j=l(i-1)+1}^{li} X_j(z),$$

$$\overline{W}_{ni}(z) = \frac{1}{l} \sum_{j=l(i-1)+1}^{li} W_j(z),$$

and

$$H_{ni}(y) = \frac{1}{l} \sum_{j=l(i-1)+1}^{li} \mathbf{1}\{Y_j \leq y\}.$$

We now compute sample covariance matrices for the β_{ni}'s and γ_{ni}'s, namely

$$S_{p1}(n) = \frac{1}{p-1} \sum_{i=1}^{p} [\beta_{ni} - \beta_n][\beta_{ni} - \beta_n]^T,$$
$$S_{p2}(n) = \frac{1}{p-1} \sum_{i=1}^{p} [\gamma_{ni} - \gamma_n][\gamma_{ni} - \gamma_n]^T.$$

Our next theorem provides large-sample confidence regions for the asymptotic means of β_n and γ_n. Let $F_{(k,p-k,\alpha)}$ be the $(1-\alpha)$-quantile of an F distribution with $(k, p-k)$ degrees of freedom.

Proposition 3 *Assume that the limit rv's appearing in Theorems 1 and 2 have non-singular covariance matrices.*

(a) Under the conditions of Theorem 1,

$$P\left(\int_{[a,b]} \varphi(E[X(z)])\nu(dz) \in R_1(n)\right) \to 1-\alpha$$

as $n \to \infty$, where

$$R_1(n) = \left\{x \in \mathbb{R}^k : p(\beta_n - x)^T S_{p1}(n)^{-1}(\beta_n - x) \leq \frac{k(p-1)}{p-k} F_{(k,p-k,\alpha)}\right\}.$$

(b) Under the conditions of Theorem 2,

$$P\left(\int_{[a,b]} \phi(\omega(H^{-1}(u)))\nu(du) \in R_2(n)\right) \to 1-\alpha$$

as $n \to \infty$*, where*

$$R_2(n) = \left\{ x \in \mathbb{R}^k : p(\gamma_n - x)^T S_{p2}(n)^{-1}(\gamma_n - x) \le \frac{k(p-1)}{p-k} F_{(k,p-k,\alpha)} \right\}.$$

Proof. These results are almost immediate, given the proofs of Theorems 1 and 2. For example, the proof of Theorem 1 establishes that

$$\begin{aligned} &\beta_{ni} - \int_{[a,b]} \varphi(E[X(z)])\nu(dz) \\ &= \int_{[a,b]} (J\varphi)(E[X(z)]) \left(\overline{X}_{ni}(z) - E[X(z)]\right) \nu(dz) + o(l^{-1/2}) \end{aligned}$$

for $1 \le i \le p$, from which it is evident that

$$\begin{aligned} &\beta_n - \int_{[a,b]} \varphi(E[X(z)])\nu(dz) \\ &= \frac{1}{p}\sum_{i=1}^{p} \left[\beta_{ni} - \int_{[a,b]} \varphi(E[X(z)])\nu(dz)\right] + o(l^{-1/2}) \end{aligned}$$

as $l \to \infty$. Consequently,

$$\begin{aligned} &l^{1/2}\left(\beta_{n1} - \int_{[a,b]} \varphi(E[X(z)])\nu(dz), \ldots, \beta_{np} - \int_{[a,b]} \varphi(E[X(z)])\nu(dz), \beta_n - \int_{[a,b]} \varphi(E[X(z)])\nu(dz)\right) \\ &\Rightarrow \left(N_1, N_2, \ldots, N_p, \frac{1}{p}\sum_{i=1}^{p} N_i\right) \end{aligned}$$

as $l \to \infty$, where $N_1, N_2, \ldots, N_p$ are iid mean zero Gaussian random vectors sharing the covariance matrix of the limit rv appearing in Theorem 1. Part (a) then follows as in Munoz and Glynn (2001); see p. 416. A similar argument works for part (b). □

5. Applications

We can approximate the Riemann-Stieltjes integral in (7) or (12) by a Riemann-Stieltjes summation on a partition. For a given partition, the asymptotic analysis for the corresponding Riemann-Stieltjes summations of (7) and (12) can be established by the results in Peng et al. (2017) and Glynn et al. (2020), which are dependent on the partition points, so they are weaker than the results established in Section 2.

We extend the stochastic model $g(\Upsilon;\theta)$ to a more general setting allowing a discontinuous sample path as follows:

$$\begin{aligned} &\sum_{j=1}^{m} \Lambda_j \mathbf{1}\{\Lambda_j \in [a_j, b_j]\} \prod_{l=1}^{j-1} \mathbf{1}\{\Lambda_l \notin (a_l, b_l)\} + c_1 \mathbf{1}\{\Lambda_m < a_m\} \prod_{l=1}^{m-1} \mathbf{1}\{\Lambda_l \notin (a_l, b_l)\} \\ &\quad + c_2 \mathbf{1}\{\Lambda_m > b_m\} \prod_{l=1}^{m-1} \mathbf{1}\{\Lambda_l \notin (a_l, b_l)\}, \end{aligned} \tag{26}$$

where

$$\Lambda_j = g_j(\Upsilon;\theta), \quad j = 1,\dots,m.$$

In the special case $a_1 = 0$ and $b_1 = \infty$, stochastic model (26) reduces to the stochastic model in Section 3. The general formula for the sensitivity estimator of the distortion risk measure under stochastic model (26) is given in the online appendix, which basically applies a general version of the GLR estimator in Peng et al. (2018). Numerical experiments for testing the performance of the proposed method can be found in the online appendix.

5.1. Assets with Exiting Boundaries

We first deal with a relatively simple example:

$$Z = a\mathbf{1}\{\bar{\Lambda} < a\} + \bar{\Lambda}\mathbf{1}\{a \le \bar{\Lambda} \le b\} + b\mathbf{1}\{\bar{\Lambda} > b\},$$

where $\bar{\Lambda} = \exp(\Upsilon(1) + \theta\Upsilon(2))$, and $\Upsilon(1)$ and $\Upsilon(2)$ are independent standard normal random variables. This stochastic model can be viewed as the payoff of an investment on an asset with exiting boundaries a and b. The distribution function of Z can be given by the following form: for $z < a$, $F(z;\theta) = 0$, for $z \ge b$, $F(z;\theta) = 1$, and for $a \le z < b$,

$$F(z;\theta) = E\left[\mathbf{1}\{Z \le z\}\right] = E\left[\mathbf{1}\{\Lambda \le \log z\}\right] ,$$

where $\Lambda = \Upsilon(1) + \theta\Upsilon(2)$. After rewriting, the GLR estimators can be applied to estimate the sensitivities with respect to θ and z. The weight functions in the GLR estimators for the sensitivities with respect to θ and z are

$$A(\Upsilon) = \Upsilon(1)\Upsilon(2), \quad B(\Upsilon) = -\Upsilon(1)/z .$$

The stochastic model of Λ is a linear function of $\Upsilon(1)$ and $\Upsilon(2)$, so it falls into the special case discussed after conditions (C.1)-(C.4), which can be checked in this example. Moreover, the second moment conditions for A and B in Propositions 1 and 2 hold. Therefore, all assumptions in Theorems 1 and 2 are satisfied in this example.

We generalize the simple example to the model that can be viewed as the payoff of an investment on two dependent assets with exiting boundaries:

$$Z = a\mathbf{1}\{\Lambda < a\} + \Lambda\mathbf{1}\{a \le \Lambda \le b\} + b\mathbf{1}\{\Lambda > b\},$$

where

$$\Lambda = e^{\Upsilon(1)} + e^{\Upsilon(2)},$$

and $(\Upsilon(1), \Upsilon(2))$ follows a bivariate normal distribution with the joint density

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\theta^2}}\exp\left(-\frac{1}{2(1-\theta^2)}\left[\frac{(y_1-\mu_1)^2}{\sigma_1^2}+\frac{(y_2-\mu_2)^2}{\sigma_2^2}-\frac{2\theta(y_1-\mu_1)(y_2-\mu_2)}{\sigma_1\sigma_2}\right]\right).$$

The weight functions in the GLR estimators for the sensitivities with respect to θ and z are

$$A(\Upsilon)=\frac{1}{1-\theta^2}\left[\theta-\frac{\theta}{(1-\theta^2)}\left(\frac{(\Upsilon(1)-\mu_1)^2}{\sigma_1^2}+\frac{(\Upsilon(2)-\mu_2)^2}{\sigma_2^2}-\frac{2\theta(\Upsilon(1)-\mu_1)(\Upsilon(2)-\mu_2)}{\sigma_1\sigma_2}\right)+\frac{(\Upsilon(1)-\mu_1)(\Upsilon(2)-\mu_2)}{\sigma_1\sigma_2}\right],$$

$$B(\Upsilon)=-\left[\frac{1}{1-\theta^2}\left(\frac{\Upsilon(1)-\mu_1}{\sigma_1^2}-\frac{\theta(\Upsilon(2)-\mu_2)}{\sigma_1\sigma_2}\right)+1\right]e^{-\Upsilon(1)}.$$

5.2. Barrier Option

We consider financial options that give the owner a right to buy or sell a particular underlying asset at strike price K. There are various types of options including vanilla option whose payoff only depends on the underlying asset price at expiration and exotic options whose payoffs are path dependent. A knockout barrier option is worthless if the path of the underlying asset exceeds a barrier L. The event when the barrier option stays "alive" is $\{\max_{j=1,\ldots,m} S_{t_j} < L\}$, where $S_t \in \mathbb{R}$ is the underlying asset price at time t. For a European barrier option, the payoff is

$$Z=\bar{\Lambda}_m\prod_{j=1}^{m-1}\mathbf{1}\{\bar{\Lambda}_j\le b\}\mathbf{1}\{0<\bar{\Lambda}_m<b\},$$

where $b=e^{-rm\Delta}(L-K)$, Δ is the step size of the discrete monitoring points $t_j=j\Delta$. Suppose $S_t=S_0\exp\{(r-\sigma^2/2)t+\sigma B_t+\sum_{j=1}^{N_t}J_j\}$ follows a geometric jump-diffusion process, where $\{N_t\}$ is a counting process, and J_i, $i\in\mathbb{Z}^+$, are the jump sizes. Then we have

$$\bar{\Lambda}_j=e^{-rm\Delta}\left\{S_0\exp\left(j\left(r-\frac{\sigma^2}{2}\right)\Delta+\sigma\sqrt{\Delta}\sum_{l=1}^{j}\Upsilon(l)+\sum_{i=1}^{N(j)}J_{j,i}\right)-K\right\},\quad j=1,\ldots,m,$$

where S_0 is the initial underlying asset price, r is the interest rate, σ is the implied volatility, and $\Upsilon(j)=(B_{j\Delta}-B_{(j-1)\Delta})/\sqrt{\Delta}$, $j=1,\ldots,m$, which are i.i.d. with standard normal distribution, $N(j)=N_{j\Delta}-N_{j(\Delta-1)}$, and $J_{j,i}$ is the ith jump in the jth period. A vanilla European option is the special case when $L=\infty$, and we can substitute $\bar{\Lambda}_m$ with $(\sum_{j=1}^{m}\bar{\Lambda}_j)/m$ in a vanilla European option to obtain the payoff of an Asian option. American option allows the owner to exercise before the expiration. For an American option exercisable at the specified discrete monitoring points, the payoff is given by

$$Z=\sum_{j=1}^{m}\bar{\Lambda}_j\prod_{l=1}^{j-1}\mathbf{1}\{\bar{\Lambda}_l<b_l\}\mathbf{1}\{\bar{\Lambda}_j\ge b_j\},$$

where $b_j = e^{-rj\Delta}(L_j - K)$, and L_j is the exercise threshold at time t_j. Obviously, the payoffs of the options described above are special cases of stochastic model (26). The sample paths of the payoffs of the barrier option and American option are discontinuous with respect to parameters $S_0, L, r, \sigma, L_j,\ j = 1, \ldots, m$, because the sample path of the underlying asset price could pass the barrier or exercise boundary before expiration with the parameters perturbed.

We show the analytical form of the GLR estimator for the distribution sensitivity of the barrier option with respect to $\theta = S_0$. The distribution function of Z can be given by the following form: for $z > b$, $F(z;\theta) = 1$, and for $0 \le z \le b$,

$$F(z;\theta) = E\left[\prod_{l=1}^{m-1} \mathbf{1}\{\bar{\Lambda}_l \le b\}\mathbf{1}\{\bar{\Lambda}_m \le z\}\right] + \sum_{j=1}^{m} E\left[\prod_{l=1}^{j-1} \mathbf{1}\{\bar{\Lambda}_l \le b\}\mathbf{1}\{\bar{\Lambda}_j > b\}\right].$$

We can also represent the distribution function by

$$F(z;\theta) = E\left[\prod_{l=1}^{m-1} \mathbf{1}\{\Lambda_l \le \log L\}\mathbf{1}\{\Lambda_m \le \log(e^{rm\Delta}z + K)\}\right] + \sum_{j=1}^{m} E\left[\prod_{l=1}^{j-1} \mathbf{1}\{\Lambda_l \le \log L\}\mathbf{1}\{\Lambda_j > \log L\}\right],$$

where

$$\Lambda_j = \log\theta + j\left(r - \frac{\sigma^2}{2}\right)\Delta + \sigma\sqrt{\Delta}\sum_{l=1}^{j}\Upsilon(l) + \sum_{i=1}^{N(j)} J_{j,i}, \quad j = 1, \ldots, m.$$

Using the formulas in the online appendix, the estimators for $\partial F(z;\theta)/\partial\theta$ and $f(z;\theta)$ are given respectively by

$$\left\{\prod_{j=1}^{m-1} \mathbf{1}\{\Lambda_j \le \log L\}\mathbf{1}\{\Lambda_m \le \log(e^{rm\Delta}z + K)\} + \sum_{i=1}^{m}\prod_{j=1}^{i-1} \mathbf{1}\{\Lambda_j \le \log L\}\mathbf{1}\{\Lambda_i > \log L\}\right\}\frac{\Upsilon(1)}{\theta\sigma\sqrt{\Delta}},$$

$$\prod_{j=1}^{m-1} \mathbf{1}\{\Lambda_j \le \log L\}\mathbf{1}\{\Lambda_m \le \log(e^{rm\Delta}z + K)\}\frac{-\Upsilon(m)}{\sigma\sqrt{\Delta}(e^{rm\Delta}z + K)}.$$

6. Conclusion

This paper proposes a new sensitivity estimator of the distortion risk measure. The stochastic model in our paper covers the case where the sample path is discontinuous with respect to the parameter, which IPA cannot handle, and by using the quantile representation, the new method can deal with the non-smooth distortion functions that cover some important risk measures such as VaR and CVaR, which have been studied separately in literature. In addition, the new estimator is proved to be asymptotically normal, using the functional limit theory. The stochastic uniform convergence offered by the empirical process technique could be a powerful tool to address the asymptotic analyses for various types of problems in stochastic optimization (e.g., Lim and Glynn 2012).

A technical difficulty in this work is to find a more general condition for establishing central limit theorems of Theorem 1 and 2 without truncating the tail of the distribution or quantile. In general, to obtain such results requires imposing additional conditions on the distortion function and the distribution of the output performance due to the nature of a difficulty that the estimations of quantile and the distribution sensitivities become poorly behaved as they get closer to the tail, because samples rarely appear around the tail. A potential way to achieve the results could be proving that the variances of estimators (22) and (25) are bounded uniformly for b and ϵ under certain extra conditions. We leave for future work proofs of Theorems 1 and 2 that cover $\delta_{n1}(\infty)$ or $\delta_{n2}(0)$ in more general scenarios.

In this work, we do not take into consideration the computational effort on evaluating the estimation at the partition points of the integral. The proposed estimator in this paper would have a subcanonical convergence rate in terms of the total computational budget that could be quantified as a product of the sample size and the number of partition points. Other numerical integration techniques such as Simpson's rule can be applied to achieve the canonical convergence rate under appropriate regularity conditions (see Andradóttir and Glynn 2016). Furthermore, asymptotic analysis for the setting where the input rvs are not iid can be considered.

Acknowledgments.

This work was supported in part by the National Science Foundation of China (NSFC) under Grants 71901003, 71720107003, 71690232, 91846301, and 71790615, by the National Science Foundation (NSF) under Grant CMMI-1434419, and by the Air Force of Scientific Research (AFOSR) under Grant FA9550-15-10050.

References

Andradóttir, Sigrún, Peter W Glynn. 2016. Computing Bayesian means using simulation. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* **26** 10.

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, David Heath. 1999. Coherent measures of risk. *Mathematical Finance* **9** 203–228.

Asmussen, Søren, Peter W Glynn. 2007. *Stochastic Simulation: Algorithms and Analysis*, vol. 57. Springer Science & Business Media.

Billingsley, Patrick. 1968. *Convergence of probability measures*. John Wiley & Sons.

Bloznelis, M., V. Paulauskas. 1994. A note on the central limit theorem for stochastically continuous processes. *Stochastic Processes and their Applications* **53** 351–361.

Cao, Xi-Ren, Xiangwei Wan. 2014. Sensitivity analysis of nonlinear behavior with distorted probability. *Mathematical Finance* .

Dhaene, Jan, Alexander Kukush, Daniël Linders, Qihe Tang. 2012. Remarks on quantiles and distortion risk measures. *European Actuarial Journal* **2** 319–328.

Ethier, Stewart N, Thomas G Kurtz. 2009. *Markov processes: characterization and convergence*, vol. 282. John Wiley & Sons.

Fu, Michael C. 2015. Stochastic gradient estimation. *Fu, Michael C. (ed.), Chapter 5 in Handbooks of Simulation Optimization*, vol. 216. Springer, 105–147.

Fu, Michael C., L. Jeff Hong, Jian-Qiang Hu. 2009. Conditional Monte Carlo estimation of quantile sensitivities. *Management Science* **55** 2019–2027.

Glasserman, Paul. 1991. *Gradient Estimation via Perturbation Analysis*. Springer.

Glynn, Peter W. 1990. Likelihood ratio gradient estimation for stochastic systems. *Communications of the ACM* **33** 75–84.

Glynn, W Peter, Lin Fan, Michael C Fu, Jianqiang Hu, Yijie Peng. 2020. Central limit theorems for estimated functions at estimated points. *Operations Research, forthcoming* .

Gourieroux, Christian, Wei Liu. 2006. Sensitivity analysis of distortion risk measures. `https://econpapers.repec.org/paper/crswpaper/2006-33.htm.`

Heidergott, Bernd, Warren Volk-Makarewicz. 2016. A measure-valued differentiation approach to sensitivity analysis of quantiles. *Mathematics of Operations Research* **41** 293–317.

Ho, Yu-Chi, Xi-Ren Cao. 1991. *Discrete Event Dynamic Systems and Perturbation Analysis*. Kluwer Academic Publishers, Boston, MA.

Hong, L Jeff. 2009. Estimating quantile sensitivities. *Operations Research* **57** 118–130.

Hong, L Jeff, Guangwu Liu. 2009. Simulating sensitivities of conditional value at risk. *Management Science* **55** 281–293.

Jiang, Guangxin, Michael C Fu. 2015. Technical note — On estimating quantile sensitivities via infinitesimal perturbation analysis. *Operations Research* **63** 435–441.

Kahneman, Daniel, Amos Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society* 263–291.

Kusuoka, Shigeo. 2001. On law invariant coherent risk measures. *Advances in Mathematical Economics*. Springer, 83–95.

Lei, Lei, Yijie Peng, Michael C. Fu, Jian-Qiang Hu. 2018. Applications of generalized likelihood ratio method to distribution sensitivities and steady-state simulation. *Journal of Discrete Event Dynamic Systems* **28** 109–125.

Lim, Eunji, Peter W Glynn. 2012. Consistency of multidimensional convex regression. *Operations Research* **60** 196–208.

Liu, Guangwu, Liu Jeff Hong. 2009. Kernel estimation of quantile sensitivities. *Naval Research Logistics* **56** 511–525.

Munoz, David F., Peter W. Glynn. 2001. Standardized time series for steady-state simulation output analysis. *Operations Research* **49** 413–422.

Peng, Yijie, Michael C. Fu, Peter W. Glynn, Jian-Qiang Hu. 2017. On the asymptotic analysis of quantile sensitivity estimation by Monte Carlo simulation. *Proceedings of 2017 Winter Simulation Conference* 2336–2347.

Peng, Yijie, Michael C Fu, Bernd Heidergott, Henry Lam. 2020. Maximum likelihood estimation by Monte Carlo simulation: Data driven model calibration. *Operations Research, forthcoming* .

Peng, Yijie, Michael C Fu, Jian-Qiang Hu, Bernd Heidergott. 2018. A new unbiased stochastic derivative estimator for discontinuous sample performances with structural parameters. *Operations Research* **66** 487–499.

Rubinstein, Reuven Y, Alexander Shapiro. 1993. *Discrete Event Systems: Sensitivity Analysis and Stochastic Optimization by the Score Function Method*, vol. 346. Wiley New York.

Suri, Rajan, Michael A. Zazanis. 1988. Perturbation analysis gives strongly consistent sensitivity estimates for the M/G/1 queue. *Management Science* **34** 39–64.

Tversky, Amos, Daniel Kahneman. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* **5** 297–323.

Van der Vaart, Aad W. 2000. *Asymptotic Statistics*. Cambridge University Press.

van der Vaart, Aad W., Jon A. Wellner. 1996. *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer.

Wang, Yongqiang, Michael C Fu, Steven I Marcus. 2012. A new stochastic derivative estimator for discontinuous payoff functions with application to financial derivatives. *Operations Research* **60** 447–460.

Whitt, Ward. 2002. *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Applications to Queues*. Springer.