TI 2013-082/III
Tinbergen Institute Discussion Paper

A Measure-Valued Differentiation Approach to Sensitivity Analysis of Quantiles

*Bernd Heidergott**
Warren Volk-Makarewicz

Faculty of Economics and Business Administration, VU University Amsterdam.

** Tinbergen Institute.*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at http://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: http://www.dsf.nl/

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

A Measure-Valued Differentiation Approach to Sensitivity Analysis of Quantiles

Bernd Heidergott
Tinbergen Institute, and Department of Econometrics and Operations Research,
VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam
bheidergott@feweb.vu.nl
and
Warren Volk-Makarewicz
Department of Econometrics and Operations Research, VU University Amsterdam,
De Boelelaan 1105, 1081 HV Amsterdam
wmakarewicz@feweb.vu.nl

Abstract

Quantiles play an important role in modelling quality of service in the service industry and in modelling risk in the financial industry. Recently, Hong showed in his breakthrough papers [25, 26] that efficient simulation based estimators can be obtained for quantile sensitivities by means of sample path differentiation. This has led to an intensive search for sample-path differentiation based estimators for quantile sensitivities. In this paper we present a novel approach to quantile sensitivity estimation. Our approach elaborates on the concept of measure-valued differentiation (MVD). Thereby, we overcome the main obstacle of the sample path approach which is the requirement that the sample cost have to be Lipschitz continuous with respect to the parameter of interest. Specifically, we perform a sensitivity analysis of the quantile of the value of a multi-asset option and a portfolio. In addition, we discuss application of our sensitivity estimator to the Variance-Gamma process and to queueing networks.

1 Introduction

Let Z have distribution function F. The quantile of Z (resp. F) at a level $\alpha \in (0,1)$, denoted by q_α, is defined as the largest value y such that the probability of obtaining a value $y \leq Z$ is less than or equal to α:

$$q_\alpha = \sup\{y : F(y) \leq \alpha\}.$$

Throughout this paper we will assume that F is continuous, so that Z possess a density function (pdf), denoted by f, and the quantile can be written more simply since F is now a bijection:

$$q_\alpha = \sup\{y : F(y) = \alpha\} = F^{-1}(\alpha). \tag{1}$$

Quantiles and quantile related performance measures are common in modelling quality of service (QoS). Indeed, in the call center industry, QoS is typically measured by the fraction of customer interactions that meet a predefined service level, for instance, the fraction of customers that could be helped within a pre-specified time (e.g. 90 % of customers can contact an agent in less than 10 seconds). In public transportation networks, QoS is measured by the achieved punctuality (e.g. 95 % of trains are no more delayed than 2 minutes). In risk analysis, value at risk and conditional value at risk are defined through quantiles. Finally, note that the 6-σ quality control approach in business management is another example, here it is the goal to guarantee that 99.9996 % of produced parts are with a pre-specified range of boundary values.

To improve or optimize the quantile related performance of a system, sensitivity analysis of quantiles with respect to changes in the parameters of the underlying model are essential. In this paper we are interested in estimating quantile sensitivities with the help of simulation. In Section 2 we discuss our approach and present the estimators. The contribution of the paper will be highlighted and the further analysis will be motivated. In Section 3 we provide a detailed discussion of the literature. The mathematical analysis of the estimator(s) is provided in Section 4 and Section 5. Applications are discussed in Section 6 and Section 7.

2 Quantile Sensitivity Analysis

Let $\mathbf{Z} = (Z_i : 1 \leq i \leq n)$ be a sequence of independent and identically distributed copies of Z. For $l = 1, \ldots, n$, the order statistic $Z_{l:n}$ is the l^{th} smallest random variable from the collection $\mathbf{Z}$. The order-statistic vector is given by

$$Z_{1:n} < Z_{2:n} < \ldots < Z_{l:n} < \ldots < Z_{n:n}.$$

Note that the order statistic is well-defined with probability one as the distribution of Z is assumed to be continuous. The order statistic $Z_{l:n}$ is the standard statistical estimator for the quantile. The relationship between quantiles and order statistics was first determined by [2] for i.i.d. random variables by showing that

$$\lim_{n \to \infty} Z_{\lceil \alpha n \rceil : n} = q_\alpha, \tag{2}$$

with probability one. The requirement of independence in the result has been weakened progressively, for instance, by Sen, [39], for m-dependent random variables[1]; and, [40], for ϕ-mixing random variables in which the random variables are asymptotically independent. The greatest technical weakening was introduced in [17], where an almost sure result was proved for particular classes of linear stationary processes. Apart from (2) we will use the following results from the theory of spacings of order statistics of i.i.d. data:

$$n\big(Z_{\lceil \alpha n \rceil : n} - Z_{\lceil \alpha n \rceil - 1 : n}\big) \stackrel{d}{\to} \frac{\mathcal{E}}{f(q_\alpha)}, \tag{3}$$

as $n \to \infty$, where $\stackrel{d}{\to}$ denotes convergence in distribution and $\mathcal{E}$ is an Exponential random variable with mean one; see [36].

In this paper we assume that the distribution of Z depends on some controllable distributional parameter θ, where we assume for sake of simplicity that $\theta \in \Theta = (a, b) \subset \mathbb{R}$. We express the dependency of the distribution of Z on θ through writing F_θ for the cumulative distribution function (cdf.) and f_θ for the probability density function (pdf.). Since Z is a continuous random variable, $F_\theta^{-1}(x)$ is differentiable w.r.t. to its argument x, and if F_θ is differentiable w.r.t θ, so is $F_\theta^{-1}(y)$. By definition, see (1),

$$\alpha = F_\theta\big(q_\alpha(\theta)\big)$$

and we obtain an expression for the quantile sensitivity by differentiating the previous w.r.t θ

$$0 = \partial_\theta F_\theta\big(q_\alpha(\theta)\big) + f_\theta\big(q_\alpha(\theta)\big)\, \partial_\theta q_\alpha(\theta),$$

or

$$\partial_\theta q_\alpha(\theta) = -\frac{\partial_\theta F_\theta\big(q_\alpha(\theta)\big)}{f_\theta\big(q_\alpha(\theta)\big)}, \tag{4}$$

[1] A sequence $\{X_n\}$ of random variables are called m-dependent if X_n and X_{n+k} are independent for any n provided that $k \geq m$.

where ∂_θ is a typographic shorthand for $\frac{\partial}{\partial\theta}$, which we will frequently use. Combining the order statistic limit in (2) with the limit of spacings in (3) suggests the following result

$$\lim_{m\to\infty} \mathbb{E}_\theta[-m\big(Z_{\lceil \alpha m\rceil:m} - Z_{\lceil \alpha m\rceil-1:m}\big)F'_\theta(Z_{\lceil \alpha m\rceil:m})] = q'_\alpha(\theta). \tag{5}$$

The paper is devoted to the analysis of (5), which has two main aspects: the statistical and the distributional differentiation aspect. For the *statistical analysis* we will first provide sufficient conditions for the above limit to hold, which already yields an asymptotically unbiased estimator for $q'_\alpha(\theta)$. Provided that (5) holds, taking averages over i.i.d. realizations of

$$-m\big(Z_{\lceil \alpha m\rceil:m} - Z_{\lceil \alpha m\rceil-1:m}\big)F'_\theta(Z_{\lceil \alpha m\rceil:m}), \tag{6}$$

then yields a strongly consistent estimator for $q'_\alpha(\theta)$. As we will show in this paper, confidence intervals for $q'_\alpha(\theta)$ can be established as well. The other aspect in (5) (resp. (6)) is that on how to deal with the *distributional derivative* $F'_\theta(Z_{\lceil \alpha m\rceil:m})$. In other words, efficient simulation of $F'_\theta(q_\alpha)$ has to be addressed. In this paper we will apply measure-valued differentiation (MVD) to operationalize $F'_\theta(Z_{\lceil \alpha m\rceil:m})$ for estimation. In particular, if F'_θ exists then it can under rather weak conditions be written as $F'_\theta = c_\theta(F^+_\theta - F^-_\theta)$, with c_θ a constant and $F^\pm_\theta$ distribution functions. Inserting the above difference expression for F'_θ into (5), we arrive at the estimator

$$-mc_\theta\big(Z_{\lceil \alpha m\rceil:m} - Z_{\lceil \alpha m\rceil-1:m}\big)\Big(F^+_\theta(Z_{\lceil \alpha m\rceil:m}) - F^-_\theta(Z_{\lceil \alpha m\rceil:m})\Big). \tag{7}$$

Note that the above estimator is a single-run estimator as no additional simulations apart of sampling the order statistics is required.

Denote by $D_{m,k}$ the sample average over k realizations of one of the estimators in (6) and (7). Let n denote the computational budget, i.e., n is total number of realizations of Z that can be used for estimating $\partial_\theta q_\alpha(\theta)$. While taking $m = n$ (and $k = 1$) for the estimator yields the most accurate point estimation for $\partial_\theta q_\alpha(\theta)$, no statistical assessment on the quality of the estimator can be made. Therefore, one typically splits the overall budget into parts, namely $n = mk$, where m is the number of realizations of Z assigned to the estimator and k is the number of independent replications of the estimator. Letting m and k tends to infinity simultaneously, the limit with respect to m yields that $\partial_\theta q_\alpha(\theta)$ will be approximated arbitrarily close and the limit with respect to k allows for constructing confidence intervals for $\partial_\theta q_\alpha(\theta)$. Details will be provided in the statistical analysis part of the paper.

Now suppose that F_θ is not known or computationally intractable. In this case we will split the estimation process into two parts. The first part will use i.i.d. replications of the spacing variable on the LHS of (3), which by Pyke's result yields a strongly consistent estimator for the value of the inverse density at the quantile, i.e., $1/f_\theta(q_\alpha)$. Alongside with estimating the inverse of the density we use $Z_{\lceil \alpha n\rceil:n}$ as estimator for q_α, see (2). In the second part of the procedure, we will use an MVD estimator for estimating $F'_\theta(q_\alpha)$. Typically, the distribution of Z is analytically intractable as it is distribution of some measurable mapping h of a collection of random variables X_i, $1 \le i \le I$, i.e., $Z = h(X_1, \ldots, X_I)$. In this case, we will resort to the product rule of weak differentiation [20] for deriving unbiased derivative estimator(s) for F'_θ at q_α. Generally speaking, if θ is a parameter of the distribution of X_1, denoted by $F_{1,\theta}$, and h as well as X_2 to X_I are independent of θ, then, provided that $F_{1,\theta}$ is weakly differentiable with weak derivative $(c_\theta, F^+_{1,\theta}, F^-_{1,\theta})$, it holds under fairly general conditions that

$$F'_\theta(z) = c_\theta \mathbb{E}_\theta[1\{h(X^+_1, X_2, \ldots, X_I) \le z\} - 1\{h(X^-_1, X_2, \ldots, X_I) \le z\}],$$

where $X^\pm_1$ are distributed according to $F^\pm_{1,\theta}$, and $1\{A\}$ denotes the indicator mapping on the set A, which equals one if the condition defining set A is true and is zero otherwise. Let $Z^\pm = h(X^\pm_1, X_2, \ldots, X_I)$. The overall estimator is build on the statistical data available for the nominal random variable $\mathbf{Z} = (Z_i : 1 \le i \le m)$ and weak derivative $Z^\pm$, and it is given as follows

$$\lim_{m\to\infty} c_\theta \mathbb{E}_\theta\left[-m\big(Z_{\lceil \alpha m\rceil:m} - Z_{\lceil \alpha m\rceil-1:m}\big)\big(1\{Z^+ \le Z_{\lceil \alpha m\rceil:m}\} - 1\{Z^- \le Z_{\lceil \alpha m\rceil:m}\}\big)\right] = q'_\alpha(\theta), \tag{8}$$

or, alternatively,

$$\lim_{m\to\infty} c_\theta \mathbb{E}_\theta\big[-m\big(Z_{\lceil \alpha m\rceil:m} - Z_{\lceil \alpha m\rceil-1:m}\big)\big] \times \mathbb{E}_\theta[1\{Z^+ \leq Z_{\lceil \alpha m\rceil:m}\} - 1\{Z^- \leq Z_{\lceil \alpha m\rceil:m}\}] = q'_\alpha(\theta). \tag{9}$$

The analysis of the estimators in (8) and (9) involves splitting the simulation budget. More specifically, $m \times k$ observations are used for obtaining k i.i.d. replications yielding the spacing estimator for inverse of the density together with a quantile estimator, and l i.i.d. samples of $Z^\pm$ for estimating the distributional derivative. We will first provide a full analysis of the case that F'_θ is analytically tractable in Section 4. The estimator incorporating the derivative estimation part on F'_θ will be discussed in Section 5, where also a more detailed description of the concept of measure-valued differentiation is provided. Applications to queues are discussed in Section 6 and applications to finance are presented in Section 7. A detailed literature review is provided in Section 3.

The main contributions of the paper are the following:

- We provide a quantile sensitivity estimator that is applicable for non-smooth cost functions. The estimator is independent of the particular cost function and requires no conditioning. For the most general variant of our estimator, it is only required that h is measurable.
- Our estimator is derived by measure-valued differentiation. In general, MVD has the often lamented drawback that it requires simulating different versions of the model in order to estimate the sensitivity. As we show for the case of quantile sensitivities, in case that F is analytically tractable, MVD leads to a single-run estimator. This is a novelty in the literature.
- In case that F'_θ fails to be obtainable, we will apply MVD to estimate F'_θ empirically via simulation. This can be done in a natural way as F'_θ is already a distributional derivative and can thus be easily evaluated in a simulation by means of MVD. Thereby, we extended the results known in the literature so far on sensitivity analysis of quantiles
- We provide a first application to sensitivity analysis of the cash-flow of multi-asset options and of options defined on the Variance-Gamma process.

3 Discussion of the Literature

Sensitivity analysis of stochastic models has been an area of active research since the advent of sample-path approaches such as infinitesimal perturbation analysis and the score function method. As of today there are three main approaches to sensitivity analysis: infinitesimal perturbation analysis [24, 11] (IPA) (together with it's variations such as smoothed perturbation analysis (SPA) [6] and realization probabilities [4]); the score function method [38], and measure valued differentiation [37, 22, 20]. Sensitivity analysis is well developed for performance characteristics that can be obtained as an expected value of an appropriate random variable, and the above methods are now belonging to the mainstream in applied probability and operations research, see [7, 29]. Quantiles do not fall into this category and sensitivity analysis of quantiles has been a long standing problem. The breakthrough papers by Hong [26, 25] were the first results on sensitivity analysis for quantiles. The basic idea of Hong's approach is the following. Suppose $Z_i(\theta)$ are path-wise monotone with respect to θ. For example, if θ is a location or scale parameter of the distribution of $Z_i(\theta)$, then sampling $Z_i(\theta)$ by means of the inverse cdf. yields a path-wise monotone representation. Then, the order indices $l : n$ are independent of θ. In other words, if $Z_i(\theta)$ is entry $l : n$ of the order statistic for some realization ω, then $Z_i(\theta)$ is entry $l : n$ of the order statistic for all values of θ for this particular realization ω. If $Z_i(\theta)$ is differentiable, then

$$\frac{d}{d\theta} Z_{\lceil \alpha n\rceil:n}(\theta)$$

is a natural estimator for the quantile sensitivity. Hong provides in [26, 25] sufficient conditions for the above estimator to be asymptotically unbiased and, provided that sample averages are taken over k

replications of the above estimator and k is chosen well in connection with n, strong consistency of the resulting estimator is established as well. The main assumption imposed in these papers is that $Z_i(\theta)$ is Lipschitz continuous with respect to θ. Specifically, provided that $Z(\theta) = h(X(\theta))$, the condition requires that (i) $X(\theta)$ is almost surely differentiable, (ii) $h(x)$ is (piece-wise) differentiable with respect to x, and that (iii) $|h(X(\theta_2)) - h(X(\theta_1))| \leq k(X)|\theta_2 - \theta_1|$, with $\mathbb{E}[k(X)] < \infty$. Under these conditions it can be shown that

$$\frac{d}{d\theta} Z_i(\theta) = h'(X(\theta)) \frac{d}{d\theta} X(\theta) \quad \text{and} \quad \frac{d}{d\theta} \mathbb{E}[Z_i(\theta)] = \mathbb{E}\left[h'(X(\theta)) \frac{d}{d\theta} X(\theta) \right],$$

where $h'(x)$ denote the derivative of $h(x)$ with respect to x. The main challenge for this approach is that Lipschitz continuity often fails to hold or is hard to check. See, for example, the portfolio credit risk model in [9], where an SPA-like conditioning approach is proposed in order to come up with a weaker set of assumption with respect to Lipschitz continuity. While this approach allows to deal with non-Lipschitz continuous mappings h, it requires a careful analysis of the model conditioned on each of the discontinuities of h. Apart from the fact that this becomes infeasible for complex models, it requires in mathematical terms that h is invertible. A simple example were h fails to be invertible is provided in [27].

An alternative line of research originates from attempting to directly estimate $q_\alpha(\theta)$ via (4). Indeed, again provided that conditions (i) to (iii) hold, it can be shown that

$$q'_\alpha(\theta) = \mathbb{E}\left[\frac{d}{d\theta} h(X(\theta)) \middle| h(X(\theta)) = q_\alpha(\theta) \right].$$

Unfortunately, the above expression requires conditioning on an event of probability zero. In [30], kernel estimators are used to smooth out the conditioning on $q_\alpha(\theta)$. While kernel estimators introduce bias into the estimator, it has been shown in [18] that this bias is negligible in applications, which renders the kernel estimator a very interesting alternative for sensitivity analysis of quantiles. However, it requires an IPA estimator to exist with known convergence properties. This is not always the case for models in applications. Think, for example, of the stationary waiting time in a (non)-exponential queueing network, where, except in special cases, path-wise differentiation of the stationary waiting time is not applicable (a general result is [13], which provides an IPA estimator based on regenerative analysis). Or, as a second example, consider the maintenance problem studied in [21]. No IPA estimator exists for this model and the corresponding SPA estimator suffers from computational complexity that makes it impractical even for small system size. Finally, it is worth mentioning that the application of IPA to the Variance Gamma process is still an open question. Fu proposed in [8] IPA estimators but without proof of unbiasedness. It is worth noting that kernel estimators can also be applied to smooth out discontinuities of h. An analysis of this approach has been provided in [31]. For an alternative approach to deal with discontinuities for sample path derivatives we refer to [32].

To summarize, the key obstacle in applying the sample path differentiation approach is the requirement that $h(x)$ is (piece-wise) differentiable with respect to x, and that $h(X(\theta))$ is Lipschitz continuous in the above sense. The approach proposed in this paper requires neither assumption.

4 Statistical Analysis

In this section we provide the analysis of the statistical properties of our estimator. We will first provide an analysis for the estimator in (5) in Section 4.1. In Section 4.2 we discuss possible extensions.

4.1 Main Analysis

The main assumptions required for our analysis are the following:

(A1) $\mathbf{Z} = (Z_i)_{i=1}^m$ *is a sequence of i.i.d. continuous random variables taking values in* $S \subset \mathbb{R}$ *with density function* f_θ*, cumulative distribution function (cdf.)* F_θ*, and finite second moment.*

(A2) *There exists an open neighborhood $B(\alpha)$ of $q_\alpha(\theta)$, such that*

- *the density $f_\theta(x)$ is continuous and strictly positive within this neighborhood, i.e.,*

$$\inf_{x \in B(\alpha)} f_\theta(x) > 0,$$

- *the derivative of the density $f_\theta(x)$ with respect to x is bounded on $B(\alpha)$, i.e.,*

$$\sup_{x \in B(\alpha)} \left| \frac{\partial}{\partial x} f_\theta(x) \right| < \infty .$$

(A3) *The cdf. $F_\theta(x)$ is differentiable with respect to θ, and the derivative is continuous as a mapping of x.*

Before we can prove the main result on uniform convergence of spacings, we state a preliminary technical result, which provides a bound on the difference between the quantile and an associated order statistic.

Lemma 4.1 (Serfling, [41], pp.97) *Let $\alpha \in (0,1)$ and suppose that* **(A1)** *and* **(A2)** *hold. Let l_n be a sequence of positive integers (with $1 \le l_n \le n$) such that*

$$\frac{l_n}{n} = \alpha + o\big((\ln(n))^\delta n^{-\frac{1}{2}}\big)$$

for some $\delta \ge 1/2$. Then with probability one

$$\big| Z_{l_n:n} - q_\alpha(\theta) \big| \le \frac{2}{f_\theta\big(q_\alpha(\theta)\big)} \frac{(\ln(n))^\delta}{n^{\frac{1}{2}}}, \tag{10}$$

for all n sufficiently large.

We will now turn to the proof of uniform integrability of spacings.

Lemma 4.2 (Uniform Integrability of Spacings) *Suppose that* **(A1)** *and* **(A2)** *hold, then*

(i) The spacing sequences $m(Z_{\lceil \alpha m \rceil:m} - Z_{\lceil \alpha m \rceil - 1:m})$ and $m^2(Z_{\lceil \alpha m \rceil:m} - Z_{\lceil \alpha m \rceil - 1:m})^2$ are uniformly integrable;

(ii) The sequence $m^a(Z_{\lceil \alpha m \rceil:m} - Z_{\lceil \alpha m \rceil - 1:m})^b$ tends to 0 a.s. as $m \to \infty$ for $a = 1, 2$ and $b > a$.

Proof: To simplify the notation, let

$$Y_{1,m} = Z_{\lceil \alpha m \rceil - 1:m},\ Y_{2,m} = Z_{\lceil \alpha m \rceil:m}.$$

Note that

$$Y_{2,m} - Y_{1,m} = F_\theta^{-1}(T_{2,m}) - F_\theta^{-1}(T_{1,m}),$$

where $T_{1,m} = U_{\lceil \alpha m \rceil - 1:m}$, $T_{2,m} = U_{\lceil \alpha m \rceil:m}$ are taken from uniform-[0,1] order statistics. Since F_θ is differentiable and monotone, there exist, for almost all realizations, $\xi \in (Y_{1,m}, Y_{2,m})$, such that by the Mean Value Theorem,

$$Y_{2,m} - Y_{1,m} = \partial_x F_\theta^{-1}(\xi)(T_{2,m} - T_{1,m}) = \frac{1}{f_\theta(F_\theta^{-1}(\xi))}(T_{2,m} - T_{1,m}). \tag{11}$$

Let

$$r_m = \frac{2}{f_\theta\big(q_\alpha(\theta)\big)} \left(\frac{\ln(m)}{m}\right)^{\frac{1}{2}}.$$

Note that letting $\delta = 1/2$ in Lemma 4.1, implies

$$|q_\alpha(\theta) - Y_{2,m}| \leq r_m, \quad |q_\alpha(\theta) - Y_{1,m}| \leq r_m,$$

for all but finitely many m, which gives

$$|Y_{2,m} - Y_{1,m}| \leq |q_\alpha(\theta) - Y_{2,m}| + |q_\alpha(\theta) - Y_{1,m}| \leq 2r_m.$$

By **(A2)**, $f_\theta(x) > 0$ within a neighbourhood $B(\alpha)$ of $q_\alpha(\theta)$. Letting $B_{r_m}(q_\alpha(\theta))$ be the open ball centred at $x = q_\alpha(\theta)$ with radius r_m, it then holds for m sufficiently large that $B_{r_m}(q_\alpha(\theta)) \subset B(\alpha)$, which implies the constants

$$c_m = \inf_{y \in B_{r_m}(q_\alpha(\theta))} f_\theta(y), \quad C_m = \sup_{y \in B_{r_m}(q_\alpha(\theta))} f_\theta(y)$$

are finite. Inserting these constants into (11) we arrive at

$$\frac{1}{C_m^2}\mathbb{E}[(T_{2,m} - T_{1,m})^2] \leq \mathbb{E}_\theta[(Y_{2,m} - Y_{1,m})^2] \leq \frac{1}{c_m^2}\mathbb{E}[(T_{2,m} - T_{1,m})^2]. \tag{12}$$

Note that by construction, it holds that

$$\lim_{m\to\infty} c_m = \lim_{m\to\infty} C_m = \frac{1}{f_\theta(q_\alpha(\theta))}. \tag{13}$$

Hence, taking the limit in (12) as m tends to infinity, it follows from the limit in (13) together with the fact that

$$\mathbb{E}[(T_{2,m} - T_{1,m})^2] = \frac{2}{m(m+1)}$$

for all m, that

$$\lim_{m\to\infty} \mathbb{E}_\theta[m^2(Y_{2,m} - Y_{1,m})^2] = \frac{2}{f_\theta^2(q_\alpha(\theta))}. \tag{14}$$

Noting that

$$\mathbb{E}[T_{2,m} - T_{1,m}] = \frac{1}{m+1},$$

it follows from the same line of arguments that

$$\lim_{m\to\infty} \mathbb{E}_\theta[m(Y_{2,m} - Y_{1,m})] = \frac{1}{f_\theta(q_\alpha(\theta))}. \tag{15}$$

We now turn to the proof of uniform convergence. From **(A1)** together with [36] it follows that

$$m(Y_{2,m} - Y_{1,m}) \overset{d}{\to} \frac{\mathcal{E}}{f_\theta\big(q_\alpha(\theta)\big)}, \tag{16}$$

where $\mathcal{E}$ is an exponential random variable with mean one. Applying the mapping $g(x) = x^2$ to $m(Y_{2,m} - Y_{1,m})$, the Continuity Limit Theorem yields

$$m^2(Y_{2,m} - Y_{1,m})^2 \overset{d}{\to} V := \frac{\mathcal{E}^2}{f_\theta^2\big(q_\alpha(\theta)\big)}, \tag{17}$$

with

$$\mathbb{E}[V] = \frac{2}{f_\theta^2\big(q_\alpha(\theta)\big)}. \tag{18}$$

A sufficient condition for uniform integrability of a sequence $\{X_n\}$ is that $(X_n)_{n\geq 1} \overset{d}{\to} X$ and $\lim_{n\geq 1}\mathbb{E}|X_n| = \mathbb{E}|X| < \infty$, see [41]. Hence, from (15) together with (16) follows uniform integrability of $m(Y_{2,m} - Y_{1,m})$, and from (14) together with (17) and (18) follows uniform integrability of $m^2(Y_{2,m} - Y_{1,m})^2$.

We now turn to the proof of part (ii) . For $b > a$, the sequence, $(m^a(Y_{2,m} - Y_{1,m})^b)_{m\geq 1}$ is positive with probability one with mean bounded above by

$$\mathbb{E}_\theta[m^a(Y_{2,m} - Y_{1,m})^b] \leq 2\left(\frac{\ln(m)}{m}\right)^{(1/2)(b-a)} \left(\frac{2}{f_\theta\big(q_\alpha(\theta)\big)}\right)^{b-a} \mathbb{E}_\theta[m^a(Y_{2,m} - Y_{1,m})^a],$$

which follows from Lemma 4.1. Since for $a = 1, 2$, with $\mathbb{E}_\theta[m^a(Y_{2,m} - Y_{1,m})^a] = O(1)$, this means $\mathbb{E}_\theta[m^a(Y_{2,m} - Y_{1,m})^b] \to 0$ as $m \to \infty$. Combining these results with Property H in [42], p. 185, we arrive at

$$m^a(Y_{2,m} - Y_{1,m})^b \overset{a.s.}{\to} 0,$$

which concludes the proof for part (ii) of Lemma 4.2. □

The first statistical property of our estimator we derive is asymptotic unbiasdeness. For the proof we use the fact that uniform convergence together with convergence in distribution implies convergence in the mean. For the proof we need the following property of the derivative of the cdf.:

(A4) *It holds that*

$$\sup_{x\in S} |F'_\theta(x)| < \infty.$$

As will show in Lemma 5.1 in the section on distributional differentiation, condition **(A4)** holds under rather weak conditions on F_θ. We now state the main theorem.

Theorem 4.1 *Let* **(A1)** *to* **(A4)** *hold, then*

$$\lim_{m\to\infty} \mathbb{E}[-m\big(Z_{\lceil \alpha m\rceil:m} - Z_{\lceil \alpha m\rceil-1:m}\big)F'_\theta(Z_{(\lceil \alpha m\rceil):m})] = q'_\alpha(\theta).$$

Proof: By **(A3)** it follows that

$$\mathbb{E}\left[m^2\big(Z_{\lceil \alpha m\rceil:m} - Z_{\lceil \alpha m\rceil-1:m}\big)^2\big(F'_\theta(Z_{\lceil \alpha m\rceil:m})\big)^2\right]$$

is well defined. Using the fact that $\big(F'_\theta(Z_{\lceil \alpha m\rceil:m})\big)^2$ is bounded (which follows from by **(A4)**), we may, by **(A1)** together with **(A2)**, argue like in the proof of Lemma 4.2 to show that

$$\sup_m \left|\mathbb{E}\big[m^2\big(Z_{\lceil \alpha m\rceil:m} - Z_{\lceil \alpha m\rceil-1:m}\big)^2\big(F'_\theta(Z_{\lceil \alpha m\rceil:m})\big)^2\big]\right| < \infty.$$

This implies uniform integrability of the sequence $(m\big(Z_{\lceil \alpha m\rceil:m} - Z_{\lceil \alpha m\rceil-1:m}\big)F'_\theta(Z_{\lceil \alpha m\rceil:m}))$. By **(A1)**, $F'_\theta(Z_{\lceil \alpha m\rceil:m})$ converges almost surely towards the deterministic value $F'_\theta(q_\alpha(\theta))$. From Slutsky's theorem it then follows that $m\big(Z_{\lceil \alpha m\rceil:m} - Z_{\lceil \alpha m\rceil-1:m}\big)F'_\theta(Z_{\lceil \alpha m\rceil:m})$ converges weakly. This together with the fact that this sequence is uniformly integrable, establishes convergence in the mean. □

Let

$$d_m = -m\big(Z_{\lceil \alpha m\rceil:m} - Z_{\lceil \alpha m\rceil-1:m}\big)F'_\theta(Z_{\lceil \alpha m\rceil:m})$$

and denote by $d_m(i)$, for $1 \leq i \leq k$, a realization of d_m. Then d_m is the estimator for $q'_\alpha(\theta)$ introduced in (5), and we denote by $D^\alpha_{m,k}$ the sample average over k realizations $d_m(i)$, $1 \leq i \leq k$, based on a sample of $\mathbf{Z}$ (i.e., m i.i.d. samples of Z):

$$D^\alpha_{m,k} = \frac{1}{k}\sum_{i=1}^k d_m(i).$$

The previous result on asymptotic unbiasedness, Theorem 4.1, together with the i.i.d. assumption of our samples implies that $D^\alpha_{m,k} \overset{a.s.}{\to} q'_\alpha(\theta)$ as $k, m \to \infty$ via the Strong Law of Large Numbers, [42]. This is strong consistency of our estimator.

We can delve further in our asymptotic result and determine the extent of biasedness for finite samples between our estimator and the mean value. This result is also needed for our Central Limit Theorem result, which will be provided in Theorem 4.3 later on.

Theorem 4.2 *Under assumptions* **(A1)** *to* **(A3)** *it holds that*

$$|\mathbb{E}_\theta[D^\alpha_{k.m} - q'_\alpha(\theta)]| = O\left(m^{-1}\right).$$

Proof: To simplify the notation, let $Y_{1,m} = Z_{\lceil \alpha m \rceil - 1:m}$, $Y_{2,m} = Z_{\lceil \alpha m \rceil :m}$, and

$$W_{1,m} = Y_{2,m} - Y_{1,m} \quad \text{and} \quad W_{2,m} = F^+_\theta(Y_{2,m}) - F^-_\theta(Y_{2,m}).$$

With this notation, our estimator reads

$$\mathbb{E}_\theta[D^\alpha_{k,m}] = \mathbb{E}_\theta[-mc_\theta W_{1,m} W_{2,m}]. \tag{19}$$

Now we expand $W_{1,m}$, $W_{2,m}$ via a Taylor polynomial approximation, heeding the method in [5], pp. 84.

We first deal with $W_{1,m}$. We have assumed that **(A1)** and **(A2)** hold, and we may take $s_m = 2(\ln(m)/m)^{1/2}$ in Lemma 4.1, so that for m sufficiently large there exists a neighbourhood $B_{s_m}(\alpha)$ of $q_\alpha(\theta)$ such that $B_{s_m}(\alpha) \subset B(\alpha)$ and the bound in (10) applies. By means of the mapping $F_\theta(x)$, a set $B \subset S$ corresponds to a pre-image B^{-1} on the unit interval, and we define

$$B^{-1} = \{F_\theta(x) : x \in B\} \subset [0, 1].$$

By Assumption **(A2)** it now holds that

$$c_{1,m} := \sup_{x \in (B_{s_m}(\alpha))^{-1}} \left| \frac{\partial^2}{\partial x^2} F^{-1}_\theta(x) \right| = \sup_{x \in B_{s_m}(\alpha)} \left| -f^{-3}_\theta(x) \frac{\partial}{\partial x} f_\theta(x) \right| < \infty,$$

where we have used the fact that $\partial_x F^{-1}_\theta(x) = 1/f_\theta(F^{-1}_\theta(x))$, which follows from the fact that $F_\theta(x)$ is monotone and continuous. Letting $(U_{l:m} : 1 \leq l \leq m)$ be the order statistic of i.i.d. uniform [0,1] random variables. Writing $Y_{i,m} = F^{-1}_\theta(T_{i,m})$, for $i = 1, 2$, where $T_{1,m} = U_{\lceil \alpha m \rceil - 1:m}$, $T_{2,m} = U_{\lceil \alpha m \rceil :m}$, we can expand $Y_{i,m}$ around α as follows

$$Y_{i,m} = F^{-1}_\theta(\alpha) + \frac{\partial}{\partial x} F^{-1}_\theta(\alpha)\big(T_{i,m} - \alpha\big) + \frac{1}{2} \frac{\partial^2}{\partial x^2} F^{-1}_\theta(\xi_i)\big(T_{i,m} - \alpha\big)^2,$$

with $\xi_i \in (B_{s_m}(\alpha))^{-1}$, for $i = 1, 2$. Note that $(B_{s_m}(\alpha))^{-1}$ is a decreasing sequence of sets with limit $\{\alpha\}$ as m tends to ∞, which implies that

$$\lim_{m \to \infty} c_{1,m} = \left| -f^{-3}_\theta(\alpha) \frac{\partial}{\partial x} f_\theta(\alpha) \right|.$$

Hence, for m sufficiently large

$$\begin{aligned} W_{1,m} &\approx \Big(F^{-1}_\theta(\alpha) + \partial_x F^{-1}_\theta(\alpha)\big(T_{2,m} - \alpha\big) + c_{1,m}\big(T_2 - \alpha\big)^2\Big) \\ &\quad - \Big(F^{-1}_\theta(\alpha) + \partial_x F^{-1}_\theta(\alpha)\big(T_{1,m} - \alpha\big) + c_{1,m}\big(T_1 - \alpha\big)^2\Big) \\ &= \frac{\partial}{\partial x} F^{-1}_\theta(\alpha)\big(T_{2,m} - T_{1,m}\big) + c_{1,m}\big(T_{2,m} - T_{1,m}\big)\big(T_{2,m} + T_{1,m} - 2\alpha\big). \end{aligned}$$

We now turn to the term $W_{2,m}$. We use a linear Taylor expansion with a Lagrangian remainder around $x = q_\alpha(\theta)$, providing

$$\begin{aligned} F^+_\theta(Y_{2,m}) &= F^+_\theta(q_\alpha(\theta)) + \frac{\partial}{\partial x} F^+_\theta(\eta_1)\big(Y_{2,m} - q_\alpha(\theta)\big) \\ &= F^+_\theta(q_\alpha(\theta)) + f^+_\theta(\eta_1)\big(Y_{2,m} - q_\alpha(\theta)\big), \end{aligned}$$

for $\eta_1 \in B_{s_m}(\alpha)$, and $f^+_\theta(x)$ denoting the density of $F^+_\theta(x)$. In the same vein

$$F^-_\theta(Y_{2,m}) = F^-_\theta(q_\alpha(\theta)) + f^-_\theta(\eta_2)\big(Y_{2,m} - q_\alpha(\theta)\big),$$

for $\eta_2 \in B_{s_m}(\alpha)$, and $f_\theta^-(x)$ denoting the density of $F_\theta^-(x)$. Letting

$$c_{2,m} = \sup_{x \in B_{s_m}(\alpha)} \max(f_\theta^+(x), f_\theta^-(x)) < \infty,$$

it follows from the fact that $B_{s_m}(\alpha)$ is a decreasing sequence with limit $\{q_\alpha(\theta)\}$, that $c_{2,m}$ converges to $\max(f_\theta^+(q_\alpha(\theta)), f_\theta^-(q_\alpha(\theta)))$. Hence, for m sufficiently large it holds that

$$W_{2,m} \approx F_\theta^+(q_\alpha(\theta)) - F_\theta^-(q_\alpha(\theta)) + c_{2,m}\big(Y_{2,m} - \alpha\big).$$

Now let

$$c_{3,m} = \sup_{x \in (B_{s_m}(\alpha))^{-1}} \left|\partial_x F_\theta^{-1}(x)\right| = \sup_{x \in (B_{s_m}(\alpha))^{-1}} \left|\frac{1}{f_\theta(F_\theta^{-1}(x))}\right|,$$

where we have used the fact that $\partial_x F_\theta^{-1}(x) = 1/f_\theta(F_\theta^{-1}(x))$, which follows from the fact that $F_\theta(x)$ is monotone and continuous, and note that, again by a Taylor series approximation, it holds for m sufficiently large that

$$Y_{2,m} - q_\alpha(\theta) \approx c_{3,m}(T_{2,m} - \alpha).$$

Note that by **(A2)** it holds that $c_{2,m}, c_{3,m} < \infty$. Inserting the above approximation for $Y_{2,m} - q_\alpha(\theta)$ into the approximation for W_2 we arrive at

$$W_{2,m} \approx F_\theta^+(q_\alpha(\theta)) - F_\theta^-(q_\alpha(\theta)) + c_{2,m}c_{3,m}\big(T_{2,m} - \alpha\big),$$

for m sufficiently large. The constants are only important in the sense that they have a finite limit, which they do by the Continuity Limit Theorem. From the definition for the sensitivity of the quantile, see Equation (4), it follows with $\partial_\theta F_\theta(q_\alpha(\theta)) = c_\theta(F_\theta^+(q_\alpha(\theta)) - F_\theta^-(q_\alpha(\theta)))$ that

$$q'_\alpha(\theta) = c_\theta(F_\theta^+(q_\alpha(\theta)) - F_\theta^-(q_\alpha(\theta)))\frac{\partial}{\partial x}F_\theta^{-1}(\alpha),$$

where we make use of the fact that $\partial F_\theta^{-1}(\alpha)/\partial x = f_\theta(F_\theta^{-1}(\alpha))$. Collecting the above approximations and inserting them into Equation (19) yields

$$\begin{aligned}
\mathbb{E}_\theta[D^\alpha_{m,k}] \approx\; & m q'_\alpha(\theta)\mathbb{E}[T_{2,m} - T_{1,m}] \\
& - m\big(\partial_\theta F_\theta\big(q_\alpha(\theta)\big)\big)c_{1,m}\mathbb{E}\big[\big(T_{2,m} - T_{1,m}\big)\big(T_{2,m} + T_{1,m} - 2\alpha\big)\big] \\
& - m c_\theta\big(\partial_x F_\theta^{-1}(\alpha)\big)c_{2,m}c_{3,m}\mathbb{E}\left[\big(T_{2,m} - T_{1,m}\big)\big(T_{2,m} - \alpha\big)\right] \\
& - m c_\theta c_{1,m}c_{2,m}c_{3,m}\mathbb{E}\left[\big(T_{2,m} - T_{1,m}\big)\big(T_{2,m} - \alpha\big)\big(T_{2,m} + T_{1,m} - 2\alpha\big)\right].
\end{aligned}$$

Since $\mathbb{E}[T_{2,m} - T_{1,m}] = 1/(m+1)$, it holds that

$$m q'_\alpha(\theta)\mathbb{E}[T_{2,m} - T_{1,m}] \approx q'_\alpha(\theta),$$

for m sufficiently large, which yields

$$\begin{aligned}
\mathbb{E}_\theta[D^\alpha_{m,k}] - q'_\alpha(\theta) \approx\; & -m\big(\partial_\theta F_\theta\big(q_\alpha(\theta)\big)\big)c_{1,m}\mathbb{E}[\big(T_{2,m} - T_{1,m}\big)\big(T_{2,m} + T_{1,m} - 2\alpha\big)] \\
& - m c_\theta\big(\partial_x F_\theta^{-1}(\alpha)\big)c_{2,m}c_{3,m}\mathbb{E}\left[\big(T_{2,m} - T_{1,m}\big)(T_{2,m} - \alpha)\right] \\
& - m c_\theta c_{1,m}c_{2,m}c_{3,m}\mathbb{E}\left[\big(T_{2,m} - T_{1,m}\big)\big(T_{2,m} - \alpha\big)\big(T_{2,m} + T_{1,m} - 2\alpha\big)\right].
\end{aligned} \tag{20}$$

Computing the remaining expected values of uniform order statistics, yields

$$\mathbb{E}\left[\big(T_{2,m} - T_{1,m}\big)\big(T_{2,m} - \alpha\big)\right] = \frac{\lceil \alpha m \rceil - \alpha m + 1 - \alpha}{m(m+1)},$$

and for m sufficiently large

$$\mathbb{E}\left[\big(T_{2,m} - T_{1,m}\big)\big(T_{2,m} - \alpha\big)\right] = \frac{2-\alpha}{m(m+1)}.$$

Also,

$$\mathbb{E}\left[\big(T_{2,m} - T_{1,m}\big)\big(T_{2,m} + T_{1,m} - 2\alpha\big)\right] = \frac{2(\lceil \alpha m \rceil - \alpha m - \alpha)}{m(m+1)} \approx \frac{2(1-\alpha)}{m(m+1)},$$

and

$$\begin{aligned}
&\mathbb{E}\left[\big(T_2 - T_1\big)\big(T_2 - \alpha\big)\big(T_2 + T_1 - 2\alpha\big)\right] \\
&= \frac{2\alpha(1-\alpha)m + 2(\lceil \alpha m \rceil - \alpha m)^2 + 4(1-2\alpha)(\lceil \alpha m \rceil - \alpha m) - 4\alpha(1-\alpha)}{m(m+1)(m+2)} \\
&\approx \frac{2\alpha(1-\alpha)}{m(m+1)} + \frac{2(1-2\alpha)(3-2\alpha)}{m(m+1)(m+2)}.
\end{aligned}$$

Hence, $\mathbb{E}_\theta[D^\alpha_{k,m}] - q'_\alpha(\theta)$ in (20) behaves for sufficiently large m like a sum of terms that are at most of order $O\left(m^{-1}\right)$, which proves the claim. □

Since we have finite second moments, the limiting distribution for the Central Limit Theorem in Theorem 4.3 is the standard normal distribution. The precise statement is as follows.

Theorem 4.3 (Central Limit Theorem) *Suppose that assumptions* **(A1)** *to* **(A4)** *hold and suppose that $k^{1/2}/m \to 0$ as $k, m \to \infty$. Then*

$$\frac{D^\alpha_{m,k} - q'_\alpha(\theta)}{\big(\mathrm{Var}_\theta\big(D^\alpha_{m,k}\big)\big)^{1/2}} \xrightarrow{d} \mathcal{N}(0,1) \tag{21}$$

as $k, m \to \infty$.

Proof: The left-hand side of (21) can be written as

$$\frac{D^\alpha_{m,k} - q'_\alpha(\theta)}{\big(\mathrm{Var}_\theta\big(D^\alpha_{m,k}\big)\big)^{1/2}} = \frac{D^\alpha_{m,k} - \mathbb{E}_\theta\big[D^\alpha_{m,k}\big]}{\big(\mathrm{Var}_\theta\big(D^\alpha_{m,k}\big)\big)^{1/2}} + \frac{\mathbb{E}_\theta\big[D^\alpha_{m,k}\big] - q'_\alpha(\theta)}{\big(\mathrm{Var}_\theta\big(D^\alpha_{m,k}\big)\big)^{1/2}}. \tag{22}$$

By construction $D^\alpha_{m,k}$ is the sample average of i.i.d. replications $d_m(i)$. Hence, by the Lévy Central Limit Theorem the first term on the right hand side of (22) converges in distribution to the standard normal distribution as m tends to infinity. As for the second term, the variance of the estimator satisfies

$$\mathrm{Var}_\theta\big(D^\alpha_{m,k}\big) = \frac{1}{k}\mathrm{Var}_\theta\big(d_m(1)\big).$$

By Theorem 4.2, $|\mathbb{E}_\theta[D^\alpha_{m.k}] - q_\alpha(\theta)| = O\left(m^{-1}\right)$, which means that $|\mathbb{E}_\theta[D^\alpha_{m.k}] - q_\alpha(\theta)|$ behaves asymptotically like L/m for some finite constant L. We obtain an upper bound for the second term

$$\left|\frac{\mathbb{E}_\theta[D^\alpha_{m,k}] - q'_\alpha(\theta)}{\big(\mathrm{Var}_\theta(D^\alpha_{m,k})\big)^{1/2}}\right| \le \frac{L}{\Big(\mathrm{Var}_\theta(d_m(1))\Big)^{1/2}} \frac{k^{1/2}}{m}. \tag{23}$$

Given our requirement on the relation between k and m, Equation (23) tends to zero as $k, m \to \infty$. Slutsky's Theorem provides the final result. □

In order to construct confidence intervals for $q'_\alpha(\theta)$, the variance of the estimator $D^\alpha_{k,m}$ has to be estimated as well. Let

$$S^\alpha_{m,k} = \frac{1}{k-1}\left\{\sum_{i=1}^{k}\big(d_m(i)\big)^2 - \frac{1}{k}\left(\sum_{i=1}^{k} d_m(i)\right)^2\right\}$$

denote our estimator for $\mathrm{Var}_\theta(d_m(i))$. The following lemma provides a sufficient condition for strong consistency of the above estimator.

Lemma 4.3 *Suppose assumption* **(A1)** *holds, then*

$$\lim_{k\to\infty} S^{\alpha}_{m,k} = \mathrm{Var}_\theta\big(d_m(i)\big).$$

with probability one.

Proof: The proof is an application of the Strong Law of Large Numbers (SLLN) [42]. We begin with

$$S^{\alpha}_{m,k} = \frac{k}{k-1}\left\{\frac{1}{k}\sum_{i=1}^{k}\big(d_m(i)\big)^2 - \frac{1}{k^2}\left(\sum_{i=1}^{k} d_m(i)\right)^2\right\}. \tag{24}$$

From assumption (**A1**), $\mathbb{E}_\theta[(D^{\alpha}_{m,k})^2] < \infty$ for all m. Combining this result, with the SLLN and the Continuity Limit Theorem, where we use the mapping $g(x) = x^2$, we obtain

$$\lim_{k\to\infty}\frac{1}{k}\sum_{i=1}^{k}\big(d_m(i)\big)^2 = \mathbb{E}_\theta[(D^{\alpha}_{m,k})^2]$$

and

$$\lim_{k\to\infty}\frac{1}{k^2}\left(\sum_{i=1}^{k} d_m(i)\right)^2 = \mathbb{E}^2_\theta[D^{\alpha}_{m,k}], \tag{25}$$

with probability one. Combining Equation (25) with Equation (24), and the algebra of limits for almost sure convergence, see, for example, [16], we obtain

$$\lim_{k\to\infty} S^{\alpha}_{m,k} = \lim_{k\to\infty}\left(\mathbb{E}_\theta\left[\big(D^{\alpha}_{m,k}\big)^2\right] - \mathbb{E}^2_\theta\big[D^{\alpha}_{m,k}\big]\right) = \mathrm{Var}_\theta\big(D^{\alpha}_{m,k}\big)$$

with probability one. □

With the above statistical analysis we can construct two-sided confidence interval for $q'_\alpha(\theta)$ as follows, where we assume that the conditions put forward in Theorem 4.3 hold. Let β denote the confidence level and let $t_{\beta,k}$ denote the $(1-\beta/2)$ quantile of Student's t-distribution with k degrees of freedom. Then, it holds asymptotically that

$$q'_\alpha(\theta) \in \left(D^{\alpha}_{m,k} - \frac{t_{\beta,k-1}}{k}S^{\alpha}_{m,k},\, D^{\alpha}_{m,k} + \frac{t_{\beta,k-1}}{k}S^{\alpha}_{m,k}\right)$$

with probability of at least $1-\beta$.

4.2 Extensions

There exists an almost sure version of Pyke's [36] basic spacing convergence result put forward in (3). The corresponding density estimator requires the choice of a sequence k_m such that k_m tends to infinity as m tends to infinity and k_m/m tends to zero as n tends to infinity. Given observation $\mathbf{Z} = (Z_i : 1 \le i \le m)$, and $z \in \mathbb{R}$, the density estimator is based on the difference of the order statistic values lying $2k_m$ steps apart and including z. More specifically,

$$f_{\theta,m}(z) = \begin{cases} \frac{2k_m}{m(Z_{\lceil 2k_m+j\rceil} - Z_{\lceil 1+j\rceil})} & \text{if } z \in [Z_{\lceil k_m+j\rceil}, Z_{\lceil k_m+j+1\rceil}) \text{ for } j = 0,1,\ldots,m-2k_m; \\ 0 & \text{if } z < Z_{\lceil k_m\rceil} \text{ or } z \ge Z_{\lceil m-k_m+1\rceil}\,. \end{cases} \tag{26}$$

Then it holds that

$$\lim_{m\to\infty} f_{\theta,m}(z) = f_\theta(z), \tag{27}$$

with probability one, see [44, 43]. Intuitively, the almost sure result is equivalent to a Strong Law of Large Numbers result where the expression is a sum of spacings which are approximately independent for large m. Elaborating on (27), a strongly consistent estimator for $q'_\alpha(\theta)$ can be obtained. The precise statement is presented in the following lemma.

Lemma 4.4 *Let* **(A1)** *to* **(A4)** *hold. Let $f_{\theta,m}$ be defined as in (26) and assume that*

(i) $f_\theta(x)$ is uniformly continuous as a mapping in x on $\mathbb{R}$,

(ii) the sequence $\{k_m\}$ is such that $\lim_{m\to\infty} k_m/m = 0$ and,

$$\sum_{m=1}^{\infty} e^{-ck_m} < \infty$$

for all $c > 0$, then

$$\lim_{m\to\infty} f_{\theta,m}(Z_{\lceil \alpha m\rceil:m})\, F'_\theta(Z_{(\lceil \alpha m\rceil):m}) = q'_\alpha(\theta),$$

with probability one.

Proof: Bahadur [2] showed that for i.i.d. random variables $Z_{\lceil \alpha n\rceil:n}$ tends towards q_α almost surely. This together with continuity of F'_θ implies that $F'_\theta(Z_{(\lceil \alpha m\rceil):m})$ tends to $F'_\theta(q_\alpha)$ as m tends to infinity with probability one. Note that

$$\lim_{m\to\infty} f_{\theta,m}(Z_{\lceil \alpha m\rceil:m}) \quad = \quad \lim_{m\to\infty} f_{\theta,m}(q_\alpha) + \lim_{m\to\infty} \big(f_{\theta,m}(Z_{\lceil \alpha m\rceil:m}) - f_{\theta,m}(q_\alpha) \big).$$

Provided that the conditions put forward in the lemma hold, the limit in (27) holds uniformly, that is, with probability one it holds that

$$\lim_{n\to\infty} \sup_z |f_{\theta,m}(z) - f_\theta(z)| = 0,$$

see [45]. For m sufficiently large it thus holds that $|f_{\theta,m}(Z_{\lceil \alpha m\rceil:m}) - f_{\theta,m}(q_\alpha)|$ becomes arbitrarily close to $|f_\theta(Z_{\lceil \alpha m\rceil:m}) - f_\theta(q_\alpha)|$ and, together with Lemma 4.1, we conclude that $f_{\theta,m}(Z_{\lceil \alpha m\rceil:m}) - f_{\theta,m}(q_\alpha) = o(1)$. Hence,

$$\lim_{m\to\infty} f_{\theta,m}(Z_{\lceil \alpha m\rceil:m}) = f_\theta(q_\alpha)$$

which probability one, which concludes the proof. □

Remark 4.1 *The optimal rate of convergence is given in [45], [46], and depends on the further requirements on the density function f_θ. For example, if $f_\theta(x)$ is Lipschitz continuous as a mapping of x, then the optimal sequence is $k_m = \lceil m^{4/7}(\log(m))^{3/7}\rceil$.*

While on the one side the strongly consistent estimator in Lemma 4.4 is preferable to the estimator in (6), which is based on the weak limit in (3), there is no result available on the asymptotic bias (compare with Theorem 4.2), which is a key result for establishing a central limit theorem for the estimator. In addition, uniform continuity of $f_\theta(x)$ as a mapping of x is required in Lemma 4.4 is hard to check for complex models.

5 Distributional Differentiation

For our analysis, we will work within the framework of *measure valued differentiation* (MVD), and we refer to [37, 22] for details.

5.1 Basic Definitions

Let μ_θ be a probability measure and let $F_\theta(x) = \mu_\theta((-\infty, x])$, for $x \in \mathbb{R}$, denote the corresponding cumulative distribution function, where we denote the probability density function of F_θ by f_θ. Let $v \geq 1$ be some measurable mapping that is absolutely integrable with respect to μ_θ for all θ. We denote by $\mathcal{B}_v$ the set of all measurable mappings that are bounded by cv for some finite constant c, i.e., $f \in \mathcal{B}_v$ if f is measurable and $f(x) \leq c\,v(x)$ for all $x \in \mathbb{R}$. In addition, we denote by $\mathcal{B}^b$ the set of bounded measurable mappings, i.e., for $\hat{v}(x) = 1$, $x \in \mathbb{R}$, it holds that $\mathcal{B}_{\hat{v}} = \mathcal{B}^b$.

The measure μ_θ is called $\mathcal{B}_v$-differentiable if a signed measure μ'_θ exists such that for all $h \in \mathcal{B}_v$:

$$\frac{\partial}{\partial \theta} \int_{\mathbb{R}} h(x)\mu_\theta(dx) = \int_{\mathbb{R}} h(x)\mu'_\theta(dx),$$

for examples we refer to [22]. The measure μ'_θ is called the *$\mathcal{B}_v$-derivative of μ_θ.*

Provided that the set of bounded and continuous mappings is a subset of $\mathcal{B}_v$, the derivative measure μ'_θ is uniquely defined. The derivative measure can be written in a non-unique way as the difference of two probability measures μ^+_θ and μ^-_θ with pre-factor $c_\theta > 0$:

$$\frac{\partial}{\partial \theta} \int_{\mathbb{R}} h(x)\mu_\theta(dx) = \int_{\mathbb{R}} h(x)\mu'_\theta(dx) = c_\theta \int_{\mathbb{R}} h(x)\big(\mu^+_\theta(dx) - \mu^-_\theta(dx)\big), \tag{28}$$

for all $h \in \mathcal{B}_v$, and we can write a measure-valued derivative of μ_θ as the triple $(c_\theta, \mu^+_\theta, \mu^-_\theta)$. In the same vein, we will call $(c_\theta, F^+_\theta, F^-_\theta)$ a $\mathcal{B}_v$-derivative of F_θ if $F^\pm_\theta$ is the cumulative distribution function of $\mu^\pm_\theta$. Since $\mathcal{B}^b \subset \mathcal{B}_v$, it holds that if μ_θ (resp. F_θ) is $\mathcal{B}_v$-differentiable, then μ_θ (resp. F_θ) is also $\mathcal{B}^b$-differentiable and the derivatives coincide.

Let X be distributed according to μ_θ, and let $X^\pm$ follow distribution $\mu^\pm_\theta$, then $\mathcal{B}_v$-differentiability of μ_θ (resp. of X) implies for all $h \in \mathcal{B}_v$ that

$$\frac{\partial}{\partial \theta}\mathbb{E}_\theta\big[h(X)\big] = c_\theta\big(\mathbb{E}_\theta\big[h(X^+)\big] - \mathbb{E}_\theta\big[h(X^-)\big]\big)$$

and we call (c_θ, X^+, X^-) a $\mathcal{B}_v$-derivative of X. Note that provided that X is $\mathcal{B}_v$-differentiable with differentiable density f_θ, it holds that

$$\frac{\partial}{\partial \theta} F_\theta(x) = \int_{-\infty}^{x} 1_S(x) \frac{\partial}{\partial \theta} f_\theta(x)\, dx, \tag{29}$$

where S denotes the support of X independent of θ.

For the statistical analysis we required in Theorem 4.1 that $F'_\theta(x)$ is bounded as a mapping of x. As the following lemma shows this conditions is always satisfied provided the corresponding random variable (resp. distribution) is $\mathcal{B}_v$-differentiable.

Lemma 5.1 *Let X have distribution μ_θ. If X (resp. μ_θ) is $\mathcal{B}^b$-differentiable with cdf. $F_\theta(x)$ for $v \geq 1$, then F'_θ exits and*

$$\sup_{x \in S} \big|F'_\theta(x)\big| < \infty.$$

Moreover, $F'_\theta(x)$ is continuous as a mapping in x.

Proof: Let X have distribution μ_θ with $\mathcal{B}^b$-derivative $(c_\theta, \mu^+_\theta, \mu^-_\theta)$. Note $\mathcal{B}^b$ contains all indicator mappings. From (28) it then follows that

$$\begin{aligned} F'_\theta(x) &= \int_S 1_{(-\infty,x]}\mu'_\theta(dx) \\ &= c_\theta\left(\int_S 1_{(-\infty,x]}\mu^+_\theta(dx) - \int_S 1_{(-\infty,x]}\mu^-_\theta(dx)\right) \\ &= c_\theta(F^+_\theta(x) - F^-_\theta(x)), \end{aligned}$$

and the proof follows from the fact that $F^+_\theta(x) - F^-_\theta(x) \leq 1$ for any x. The fact that $F'_\theta(x)$ is continuous as a mapping of x stems from the fact that $F_\theta(x)$ is $\mathcal{B}^b$-differentiable, which implies strong differentiability of F_θ and thereby implies continuity of F'_θ; see [20]. $\square$

Since the set of bounded continuous mappings is a subset of $\mathcal{B}_v$, there exists a unique triple representation of μ_θ (resp. X) in the sense of minimal total variation, given by the Hahn-Jordan decomposition, see [28]. For measure valued differentiation this implies that our two probability measures have support on disjoint sets. While for independent random variables the Hahn-Jordan decomposition provides the minimal variance estimator, sampling (X^+, X^-) may be difficult.

Example 5.1 *Let F_θ denote the cdf. of an exponentially distributed random variable with rate θ, i.e., $F_\theta(x) = 1 - e^{-\theta x}$, for $x \geq 0$. Let $v(x) = (1+x)^p$, for $p \in \mathbb{N}$. Then it holds that the $\mathcal{B}_v$-derivative of F_θ, denoted by F'_θ, is given by*

$$F'_\theta(x) = xe^{-\theta x}, \quad x \geq 0,$$

for any $p \in \mathbb{N}$. For details, see [22]. Denote the cdf. of Erlang-$(2,\theta)$-distribution by $E_{2,\theta}(x)$, then it holds that

$$F'_\theta(x) = \frac{1}{\theta}(F_\theta(x) - E_{2,\theta}(x)).$$

Example 5.2 *Let $\mathcal{N}(\mu,\sigma)$ denote the normal distribution with mean μ and standard deviation σ and write $\mathcal{N}(\mu,\sigma)(x)$ for the cdf. of $\mathcal{N}(\mu,\sigma)$. The density of $\mathcal{N}(\mu,\sigma)(x)$ is given by*

$$\phi_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

Then $\mathcal{N}(\mu,\sigma)$ is $\mathcal{B}_v$-differentiable with respect to μ and σ, for $v(x) = (1+x)^p$, with $p \in \mathbb{N}$; for details, see [22]. Denote the $\mathcal{B}_v$-derivative of $\mathcal{N}(\mu,\sigma)$ with respect to μ by $\mathcal{N}_{\boldsymbol{\mu}}(\mu,\sigma)$ and the $\mathcal{B}_v$-derivative of $\mathcal{N}(\mu,\sigma)$ with respect to σ by $\mathcal{N}_{\boldsymbol{\sigma}}(\mu,\sigma)$. Differentiating $\phi_{\mu,\sigma}(x)$ with respect to θ yields

$$\frac{\partial}{\partial\mu}\phi_{\mu,\sigma}(x) = \frac{1}{\sigma^3\sqrt{2\pi}}(x-\mu)e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

Integrating the above derivative of the density out yields

$$\mathcal{N}_{\boldsymbol{\mu}}(\mu,\sigma)(x) = -\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R},$$

see (29). Alternatively, we can write

$$\frac{\partial}{\partial\mu}\phi_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}}\left(\frac{(x-\mu)}{\sigma^2}e^{-\frac{(x-\mu)^2}{2\sigma^2}}1_{x\geq\mu} - \frac{(\mu-x)}{\sigma^2}e^{-\frac{(x-\mu)^2}{2\sigma^2}}1_{x\leq\mu}\right), \quad x \in \mathbb{R}.$$

Note that $xe^{-\frac{x^2}{2\sigma^2}}/\sigma^2$, for $x \geq 0$ is the density of the Rayleigh distribution the cdf. of which is given by $R_\sigma(x) = 1 - e^{-\frac{x^2}{2\sigma^2}}$, for $x \geq 0$. Hence,

$$\mathcal{N}_{\boldsymbol{\mu}}(\mu,\sigma)(x) = \frac{1}{\sigma\sqrt{2\pi}}\left(R_\sigma(x-\mu)1_{x\geq\mu} - R_\sigma(|x-\mu|)1_{x\leq\mu}\right),$$

with $R_\sigma(x-\mu)$ being the Rayleigh distribution shifted by μ. Note that a sample from $R_\sigma(x-\mu)$ can be obtained from $\mu + \sigma\sqrt{-2\ln(1-U)}$, for $U \in [0,1]$, and that a sample from $R_\sigma(|x-\mu|)$ can be obtained from $\mu - \sigma\sqrt{-2\ln(1-U)}$.

We now turn to the derivative with respect to σ. It can be shown that

$$\mathcal{N}_{\boldsymbol{\sigma}}(\mu,\sigma)(x) = \frac{1}{\sigma}\left(\mathcal{DM}(\mu,\sigma)(x) - \mathcal{N}(\mu,\sigma)(x)\right),$$

for $x \in \mathbb{R}$, where $\mathcal{DM}(\mu,\sigma)$ denotes the cdf. of the double-sided distribution with mean μ and shape parameter σ. Sampling from the Double-Maxwell distribution is discussed in [23]. In terms of minimizing the variance of the estimator it is of interest that the Double-Maxwell distribution and normal distribution *can be coupled in a simple way. If the random variable M has cdf. $M \sim \mathcal{DM}(\mu,\sigma)(x)$, then UM, with U being uniformly distributed in $[0,1]$ and independent of everything else, has cdf. $\mathcal{N}(\mu,\sigma)(x)$. See, [23] for details. The normal distribution and the Double-Maxwell distribution can be computed by means of the error function. Fortunately, the expression for $\mathcal{N}_{\boldsymbol{\sigma}}(\mu,\sigma)(x)$ is the difference between the two cdf.'s and the corresponding error function terms cancel out. Specifically, it holds*

$$\mathcal{N}_{\boldsymbol{\sigma}}(\mu,\sigma)(x) = -\frac{1}{\sqrt{2\pi}\sigma}\frac{x-\mu}{\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

Suppose that μ and σ are mappings of a common parameter θ, i.e., $\mu = \mu(\theta)$ and $\sigma = \sigma(\theta)$. Provided that $\mu(\theta)$ and $\sigma(\theta)$ are differentiable with respect to θ, applying the chain rule of differentiation yields for the $\mathcal{B}_v$-derivative of $\mathcal{N}(\mu,\sigma)(x)$ with respect to θ, denoted by $\partial\mathcal{N}(\mu,\sigma)(x)/\partial\theta$

$$\frac{\partial}{\partial\theta}\mathcal{N}(\mu,\sigma)(x) = \frac{d}{d\theta}\mu(\theta)\,\mathcal{N}_{\boldsymbol{\mu}}(\mu,\sigma)(x) + \frac{d}{d\theta}\sigma(\theta)\,\mathcal{N}_{\boldsymbol{\sigma}}(\mu,\sigma)(x), \tag{30}$$

for $x \in \mathbb{R}$.

Remark 5.1 *It is worth noting that our estimators can be interpreted (up to some limitations) in terms of the* Score Function Method. *It is well known that the score function estimator has typically larger variance than an MVD estimator whereas the score function estimator is a single run estimator avoiding the re-simulation required by standard MVD estimators. However, the advantage of the single-run representation of the Score Function Method usually does not compensate for its (typically much) larger variance and, in general, an MVD estimator is more efficient than a Score Function Method estimator.*

5.2 The Inverse-Transformation Approach and Extensions to Finite Product Measures

Weak differentiation of X carries over to that of $h(X)$. The precise statement is provided in the following lemma, which will be useful for establishing **(A3)** in applications.

Lemma 5.2 *Assume that X is $\mathcal{B}^b$-differentiable with weak derivative (c, X^+, X^-). Let h be a real-valued measurable mapping defined on the state space of X. Then $Z = h(X)$ is $\mathcal{B}^b$-differentiable with weak derivative (c, Z^+, Z^-), with $Z^+ = h(X^+)$ and $Z^- = h^-(X)$. In addition the cdf. of Z is $\mathcal{B}^b$-differentiable.*

Proof: For any $g \in \mathcal{B}^b$ it holds that $g \circ h$, with $(g \circ h)(x) = g(h(x))$ belongs to $\mathcal{B}^b$. By computation, it holds for any $g \in \mathcal{B}^b$ that

$$\begin{aligned}\frac{d}{d\theta}\mathbb{E}[g(Z)] = \frac{d}{d\theta}\mathbb{E}[(g\circ h)(X)] &= c(\mathbb{E}[(g\circ h)(X^+)] - \mathbb{E}[(g\circ h)(X^-)]) \\ &= c(\mathbb{E}[g(h(X^+))] - \mathbb{E}[g(h(X^-))]),\end{aligned}$$

which shows that $(c, h(X^+), h(X^-))$ is an instance of an $\mathcal{B}^b$-derivative of $Z = h(X)$. Differentiability of the cdf. of Z follows from Lemma 5.1, which concludes the proof. □

Often the stochastic model under consideration has not a single random variable as input but a collection $X = (X_1, \ldots, X_J)$ of independent real-valued random variables. The aggregated model is then given by $Z = h(X_1, \ldots, X_J)$, with h being some measurable mapping from $\mathbb{R}^J$ onto $\mathbb{R}$. Let F denote the cdf. of the aggregate random variable Z and let F_i denote the cdf. of X_i and let μ_i denote the associate measure, for $1 \le i \le J$, i.e., $\mu_i((-\infty, z]) = F_i(z)$ for $z \in \mathbb{R}$. In what follows we assume without loss of generality [2] that cdf.s are defined on $\mathbb{R}$. Applying the product rule of weak differentiation leads to an unbiased sensitivity estimator in this case. The precise statement is provided in the following lemma.

[2] If a cdf. F (resp. measure μ) has support $[0,\infty)$, then we extend F (resp. μ) to $\mathbb{R}$ in the obvious way be setting $F(z) = 0$, $z \le 0$ (resp. $\mu(A) = \mu(A \cap [0,\infty))$.

Lemma 5.3 *For $1 \le i \le J$, suppose that X_i are $\mathcal{B}^b$-differentiable w.r.t. θ and let (c_i, X_i^+, X_i^-) denote a version of the $\mathcal{B}^b$-derivative. Let h be a measurable mapping from $\mathbb{R}^J$ onto $\mathbb{R}$. Then $Z = h(X_1, \ldots, X_J)$ is $\mathcal{B}^b$-differentiable with $\mathcal{B}^b$-derivative (c, Z^+, Z^-), which is given by*

$$Z^+ = h(X_1, \ldots, X_{\rho-1}, X_\rho^+, X_{\rho+1}, \ldots, X_J), \;\; Z^- = h(X_1, \ldots, X_{\rho-1}, X_\rho^-, X_{\rho+1}, \ldots, X_J),$$

where ρ is uniformly distributed on $\{1, \ldots, J\}$ independent of everything else, and $c = Jc_\rho$.

Proof: Applying Theorem 6.1 in [20] is readily follows that Z is $\mathcal{B}^b$-differentiable. The particular form of the weak derivative stems from the randomization principle, see Corollary 4 in [19]. □

The estimator in Lemma 5.3 requires re-simulation, which might render the estimator inefficient. In the following we will show that in case that h is invertible, then we can rewrite these estimator in single run form by elaborating on the distributional derivatives. More specifically, for $1 \le j \le J$, let $X_1 = x_1, \ldots, X_{j-1} = x_{j-1}, \cdot, X_{j+1} = x_{j+1}, \ldots, X_J = x_J$ be given, and define the inverse image of $h(x_1, \ldots, x_{j-1}, \cdot, x_{j+1}, \ldots, x_J)$ as follows

$$\{x \in \mathbb{R} : h(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_J) \le z\} = h^{I,j}_{(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_J)}(z).$$

Hence, for $j = 1$,

$$\begin{aligned}\mathbb{P}(Z \le z | X_2 = x_2, \ldots, X_j = x_J) &= \mathbb{P}\Big(X_1 \in h^{I,1}_{(x_2,\ldots,x_J)}(z) \Big| X_2 = x_2, \ldots, X_j = x_J\Big) \\ &= \mathbb{E}\Big[\mu_1(h^{I,1}_{(x_2,\ldots,x_J)}(z)) \Big| X_2 = x_2, \ldots, X_j = x_J\Big].\end{aligned}$$

We call the inverse image *simple* if $h^{I,i}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_J)(z)$ yields sets of type $(-\infty, x]$ for any value of z and $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_J$. Note that this implies that $h^{I,i}$ is measurable. If the inverse image is simple, we let $h^{-1,i}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_J)(z) = x$ if $(-\infty, x] = h^{I,i}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_J)(z)$, and the above RHS can be written as

$$= \mathbb{E}\Big[F_1\Big(h^{-1,1}_{(x_2,\ldots,x_J)}(z)\Big) \Big| X_2 = x_2, \ldots, X_j = x_J\Big]. \tag{31}$$

As has been noted in [27], the inverse of h may not be available in closed form or the evaluation of it may be numerically infeasible, see the portfolio example in [27].

In the next lemma we will present the single run version of the estimator in Lemma 5.3.

Lemma 5.4 *Let $F_{i,\theta}$ denote the cdf. of X_i, for $1 \le j \le J$. If h is simple, then it holds under the conditions put forward in Lemma 5.3 that*

$$\partial_\theta F(z) = J\mathbb{E}_\theta\left[\partial_\theta F_{\rho,\theta}\left(h^{-1,\rho}_{(X_1,\ldots,X_{\rho-1},X_{\rho+1},\ldots,X_J)}(z)\right)\right],$$

for ρ being uniformly distributed on $\{1, \ldots, J\}$ independent of everything else.

Proof: By Lemma 5.3 it holds that

$$\begin{aligned}\partial_\theta F_\theta(z) &= Jc_\rho \mathbb{E}\left[1_{h(X_1,\ldots,X_{\rho-1},X_\rho^+,X_{\rho+1},\ldots,X_J)\le z} - 1_{h(X_1,\ldots,X_{\rho-1},X_\rho^-,X_{\rho+1},\ldots,X_J)\le z}\right] \\ &= Jc_\rho \mathbb{E}\left[\mathbb{E}\left[1_{h^{-1,\rho}_{(X_1,\ldots,X_{\rho-1},X_{\rho+1},\ldots,X_J)}(z)} - 1_{h^{-1,\rho}_{(X_1,\ldots,X_{\rho-1},X_{\rho+1},\ldots,X_J)}(z)} \middle| X_1, \ldots, X_{\rho-1}, X_{\rho+1}, \ldots, X_J\right]\right] \\ &= J\mathbb{E}\left[\partial_\theta F_{\rho,\theta}(h^{-1,\rho}_{(X_1,\ldots,X_{\rho-1},X_{\rho+1},\ldots,X_J)}(z))\right],\end{aligned}$$

which proves the claim. □

Remark 5.2 *The Finite Difference method is a general propose approach to gradient estimation. While easy to use it has the main drawback of producing biased estimates. Finding satisfying settings for the Finite Difference method in practice is often rather difficult and Finite Differences is not considered as a competitive method in the gradient estimation literature. For a detailed discussion on the Finite Difference method and some analytical results available on this method we refer to Section 1 of Chapter VII of [1].*

5.3 Incorporating Distributional Derivative Estimators

If F'_θ is not available in a closed form, we resort to estimating F'_θ through re-simulation. We first address the estimator put forward in (8). Let

$$\mathrm{distr}_l^\alpha(z) = \frac{1}{l}\sum_{i=1}^{l}\big(1\{Z^+(i) \le z\} \;-\; 1\{Z^-(i) \le z\}\big),$$

be the unbiased estimator of $F'_\theta(z)$ based on l i.i.d. replications of $Z^\pm(i)$, see Lemma 5.3. Furthermore, for m, k let

$$\mathrm{dens}_{m,k}^\alpha = -\frac{m}{k}\sum_{i=1}^{k}\big(Z_{\lceil \alpha m\rceil:m}(i) - Z_{\lceil \alpha m\rceil-1:m}(i)\big)$$

be an unbiased estimator for $1/f_\theta(q_\alpha(\theta))$. The following lemma provides by means of re-simulation an extension of Theorem 4.2 to the case that F'_θ is not obtainable in a closed-form.

Lemma 5.5 *Under assumptions* **(A1)** *to* **(A3)** *it holds that*

$$\left|\mathbb{E}_\theta\left[\mathrm{dens}_m^\alpha - \frac{1}{f_\theta(q_\alpha(\theta))}\right]\right| = \mathrm{O}\big(m^{-1}\big)\ .$$

Proof: Replacing F'_θ by one in the proof of Theorem 4.2, the result readily follows. □

Lemma 5.6 *Under assumptions* **(A1)** *to* **(A3)** *it holds that*

$$\big|\mathbb{E}_\theta[\mathrm{dens}_m^\alpha \times \mathrm{distr}_l^\alpha(Z_{\lceil \alpha m\rceil:m}) - q'_\alpha(\theta)]\big| = \mathrm{O}\big(m^{-1}\big)\,.$$

Proof: By definition,

$$\begin{aligned}
\mathbb{E}_\theta[D_{m,k}^\alpha - q'_\alpha(\theta)] &= \mathbb{E}_\theta[\mathrm{dens}_m^\alpha \times F'_\theta(Z_{\lceil \alpha m\rceil:m}) - q'_\alpha(\theta)]\\
&= \mathbb{E}_\theta[\mathrm{dens}_m^\alpha \times \mathrm{distr}_l^\alpha(Z_{\lceil \alpha m\rceil:m}) + \mathrm{dens}_m^\alpha \times (F'_\theta(Z_{\lceil \alpha m\rceil:m}) - \mathrm{distr}_l^\alpha(Z_{\lceil \alpha m\rceil:m})) - q'_\alpha(\theta)]\\
&= \mathbb{E}_\theta[\mathrm{dens}_m^\alpha \times \mathrm{distr}_l^\alpha(Z_{\lceil \alpha m\rceil:m}) - q'_\alpha(\theta)]\\
&\quad + \mathbb{E}_\theta[\mathrm{dens}_m^\alpha \times (F'_\theta(Z_{\lceil \alpha m\rceil:m}) - \mathrm{distr}_l^\alpha(Z_{\lceil \alpha m\rceil:m}))].
\end{aligned}$$

Hence,

$$\begin{aligned}
&\big|\mathbb{E}_\theta[\mathrm{dens}_m^\alpha \times \mathrm{distr}_l^\alpha(Z_{\lceil \alpha m\rceil:m}) - q'_\alpha(\theta)]\big|\\
&\le \big|\mathbb{E}_\theta[D_{m,k}^\alpha - q'_\alpha(\theta)]\big| + \big|\mathbb{E}_\theta[\mathrm{dens}_m^\alpha \times (F'_\theta(Z_{\lceil \alpha m\rceil:m}) - \mathrm{distr}_l^\alpha(Z_{\lceil \alpha m\rceil:m}))]\big|\,.
\end{aligned}$$

We have computed the rate of convergence of $\mathbb{E}_\theta[D_{m,k}^\alpha - q'_\alpha(\theta)]$ in Theorem 4.2. Since $\mathrm{distr}_l^\alpha(z)$ is an unbiased estimator for $F'_\theta(z)$, there is no bias introduced by switching from $F'_\theta(Z_{\lceil \alpha m\rceil:m})$ to $\mathrm{distr}_l^\alpha(Z_{\lceil \alpha m\rceil:m})$, i.e., $\mathbb{E}_\theta[\mathrm{dens}_m^\alpha \times (F'_\theta(Z_{\lceil \alpha m\rceil:m}) - \mathrm{distr}_l^\alpha(Z_{\lceil \alpha m\rceil:m}))] = 0$. □

Following the same line of argument as for the proof of the CLT for estimator $D_{m,k}^\alpha$, see Theorem 4.3, we arrive at the following result:

Lemma 5.7 (Central Limit Theorem) *Suppose that Assumptions* **(A1)** *to* **(A4)** *hold and suppose that $k^{1/2}/m \to 0$ as $k, m \to \infty$. If $Z^\pm$ have finite second moments, then*

$$\frac{\mathrm{dens}_m^\alpha \times \mathrm{distr}_l^\alpha(Z_{\lceil \alpha m\rceil:m}) - q'_\alpha(\theta)}{\big(\mathrm{Var}_\theta\big(\mathrm{dens}_m^\alpha \times \mathrm{distr}_l^\alpha(Z_{\lceil \alpha m\rceil:m})\big)\big)^{1/2}} \overset{d}{\to} \mathcal{N}(0,1)$$

as $k, m, l \to \infty$.

Assuming that $Z = h(X_1, \ldots, X_J)$, where X_i has cdf. $F_{i,\theta}$, we can alternatively estimate the distributional derivative term through

$$\frac{J}{l} \sum_{i=1}^{l} F'_{\rho,\theta}\left(h^{-1}_{(X_1(i),\ldots,X_{\rho-1}(i),X_{\rho+1}(l),\ldots,X_J(i))}(z) \right)$$

see (31) and Lemma 5.4, where $(X_1(i), \ldots, X_J(i))$, for $1 \leq i \leq l$, are i.i.d. copies of $(X_1, \ldots, X_J)$.

We conclude this section with a discussion of the estimator in (9). For this estimator, a confidence interval can be obtained by choosing α_1, α_2, such that $\alpha = 1 - (1 - \alpha_1)(1 - \alpha_2)$, where α denotes the level of the overall confidence interval, and estimating confidence intervals for $\mathbb{E}[\mathrm{dens}^{\alpha}_m]$ of level α_1 and for $\mathbb{E}[\mathrm{distr}^{\alpha}_l(Z_{\lceil \alpha m \rceil:m})]$ of level α_2, separately.

It is worth noting that in estimating a quantile sensitivity one has to balance bias with variance. As k and m are related through $k^{1/2}/m \to 0$ as $k, m \to \infty$, see Lemma 5.7, the choice is between $n = km$ and l. Increasing n will reduce bias (and to some extent variance) and increasing l will only reduce variance. As argued in Section 6 of Chapter III in [1], the relation between n and l should be such that the standard deviation dominates the bias. As the standard deviation is of order $1/\sqrt{kl}$ and bias of order $1/m$, one should therefore choose k, l such that $m > \sqrt{kl}$.

Remark 5.3 *For the estimator put forward in Lemma 4.4, let*

$$\mathrm{dens}^{\alpha}_m = \frac{2k_m}{m(Z_{\lceil \alpha m \rceil + k_m:m} - Z_{\lceil \alpha m \rceil - k_m + 1:m})}$$

be the estimator for $f_\theta(Z_{\lceil \alpha m \rceil:m})$ as provided in (27). Then the overall estimator becomes $\mathrm{dens}^{\alpha}_m \mathrm{distr}^{\alpha}_l(Z_{\lceil \alpha m \rceil:m})$, where the bias is reduced by letting m tend to infinity and variance is reduced by letting l tend to infinity. Unfortunately, no analytical results are available on the rate the bias is reduced through the choice of m, so no guidelines for choosing m in relation to l are available.

6 Application to Queues

Consider a tandem queue with Poisson λ arrival stream consisting of two queues with infinite buffer capacity. Jobs arrive from the outside to server 1 and are being served with exponential service rate μ_1. From server 1 they continue to server 2 where they are served with exponential service rate μ_2. The jobs leave the system once the service at station 2 is completed. For the sake of simplicity, we let $\mu = \mu_1 = \mu_2$, and we assume that the tandem queue is stable, i.e., $\lambda < \mu$. By Burke's theorem [3], it is known that the departure process at the first server is a Poisson λ process and with this result it can be shown that the distribution of the sojourn time of a job, i.e., the total time elapsed between entering and leaving the system, is the convolution of two exponential distributions with rate $\mu - \lambda$, i.e., an Erlang distribution with shape 2 and rate $\mu - \lambda$.

Let q_α denote the α-quantile of the stationary sojourn time. In the following we apply our estimator to estimate the sensitivity of q_α with respect to, for example, μ. Note that, even though the distribution of the stationary sojourn time can be obtained in a closed form, the α-quantile cannot be expressed in a closed form and has to be solved numerically. Let F^D denote the cdf. of the stationary sojourn time, denoted by Z, and since this distribution is of Erlang type it can be obtained in a closed for as follows

$$F^D(z) = 1 - e^{-(\mu-\lambda)z} - (\mu - \lambda) z e^{-(\mu-\lambda)z}, \; z \geq 0. \tag{32}$$

Following the line of thought of Example 5.1, it follows that Z is $\mathcal{B}^b$-differentiable and that

$$\frac{\partial}{\partial \mu} F^D(z) \;\; = \;\; (\mu - \lambda)^2 z e^{-(\mu-\lambda)z}, \; z \geq 0.$$

Inserting the above expression into (6) yields

$$-m(\mu - \lambda)^2 (Z_{\lceil \alpha m \rceil:m} - Z_{\lceil \alpha m \rceil - 1:m}) Z_{\lceil \alpha m \rceil:m} e^{-(\mu-\lambda) Z_{\lceil \alpha m \rceil:m}}$$

as an estimator, where the order statistic is obtained from $\mathbf{Z}$ being an i.i.d. sample of m stationary sojourn times.

Next, consider the excess sojourn time $D1_{D\geq d}$ for some fixed value d. For example, in service systems waiting is often only considered to be relevant if it exceeds some pre-specified threshold value d. We are again interested in the sensitivity with regard to the common service time parameter μ. Let $h(x) = x1_{x\geq d}$ and note that $h(x) = y$ implies $x = y$ for $y > d$. Hence, under the reasonable assumption that $q_\alpha > d$, we can apply estimator (6), which yields

$$-m(\mu-\lambda)^2(Z_{\lceil\alpha m\rceil:m} - Z_{\lceil\alpha m\rceil-1:m})Z_{\lceil\alpha m\rceil:m}e^{-(\mu-\lambda)Z_{\lceil\alpha m\rceil:m}}1_{Z_{\lceil\alpha m\rceil:m}>d}$$

as an estimator, where the order statistic is obtained from $\mathbf{Z}$ being an i.i.d. sample of m stationary sojourn times. Note that $h(x)$ fails to be Lipschitz continuous and, therefore, IPA cannot be applied.

In the following we show that assumptions **(A1)** to **(A4)** hold for the above examples. Since the distribution of the stationary sojourn time in the tandem queue is explicitly known and has finite second moment, sampling $\mathbf{Z}$ from this distribution is feasible and condition **(A1)** is thus satisfied. We now turn to assumption **(A2)**. Since the support of the sojourn time distribution is $\mathbb{R}^+$, $f_\theta(x) > 0$ on every ball $B(\alpha)$, and the first part of **(A2)** is satisfied. Let $f^D(x)$ denote the density of the stationary sojourn time, which is given by

$$f^D(x) = (\mu-\lambda)^2 x e^{-(\mu-\lambda)x},$$

for $x \geq 0$. The remaining conditions for **(A2)** follow from

$$\frac{\partial}{\partial\mu} f^D(x) = \left(2(\mu-\lambda)x - (\mu-\lambda)^2 x^2\right) e^{-(\mu-\lambda)x} \tag{33}$$

which is bounded since $x^2 e^{-(\mu-\lambda)x}$ is bounded on $\mathbb{R}^+$.

We have already argued that the stationary waiting is $\mathcal{B}^b$-differentiable. Therefore, **(A3)** and **(A4)** follow from applying Lemma 5.1. For the excess waiting time, we resort to Lemma 5.2 for showing that $h(Z) = Z_{Z\geq d}$ is $\mathcal{B}^b$-differentiable.

It is worth noting that the above line of argument can be extended to more general feed forward exponential queueing networks.

7 Application to Option Pricing

In this section we will discuss sensitivities of financial options. Section 7.1 provides the basic stock price models. Plain vanilla options are analyzed in Section 7.2. Section 7.3 is devoted to rainbow options. Eventually, we discuss sensitivity analysis of the value at risk of a portfolio in Section 7.4.

7.1 The Underlying Financial Models

In the following we will introduce the Black-Scholes-Merton model and Variance Gamma process model for a stock price.

7.1.1 The Black-Scholes-Merton Model

The famous Black-Scholes-Merton (BSM) Model, [35], is a simplified model of a financial market. The BSM model is composed of one stock, having value $S(t)$ at time $t \geq 0$, that pays no dividends and a bond, having value e^{rt} at time $t \geq 0$, where $r \geq 0$. The BSM model is complete which means that every contingent claim can be replicated. For hedging purposes the price of the stock is determined under the risk-neutral or equivalent martingale measure, and the value of the stock price becomes

$$S(t) = S(0)e^{\left(r-\frac{\sigma^2}{2}\right)t+\sigma\sqrt{t}W},$$

with W being a standard normal random variable. Alternatively, let $X_{a(t),b(t)}$ denote a normal random variable with mean

$$a(t) = \ln(S(0)) + \left(r - \frac{\sigma^2}{2}\right) t$$

and standard deviation

$$b(t) = \sigma\sqrt{t},$$

then

$$S(t) = e^{X_{a(t),b(t)}},$$

for $t \geq 0$. Note that the above presentation is a simplified version of the actual BSM model. Indeed, the standard approach for obtaining $S(t)$ is to model the time evolution of the stock price by means of a stochastic differential equation, the solution of which is given by a geometric Brownian motion, say, $\{\hat{S}(t) : t \geq 0\}$. Specifically, $\{\hat{S}(t) : t \geq 0\}$ is constructed in such a way that $\{e^{-rt}\hat{S}(t) : t \geq 0\}$ becomes a martingale. The above definition of $S(t)$ is such that the marginal distribution of $\{\hat{S}(t) : t \geq 0\}$ at time t is equal to the distribution of $S(t)$.

While the BSM model is widely used, it fails to represent the market properly and more accurate models that build on the BSM model have been developed in the literature. In the subsequent section we will discuss one of these extensions which is the Variance-Gamma (VG) process, [33], [34]. The VG process is obtained from substituting the time parameter of a Brownian motion with a Gamma process as subordinator. In other words, the value of a VG-process at time t is obtained by evaluating the value of a Brownian motion at time $\tau(t)$, where $\tau(t)$ is a Gamma process, i.e., the increments of $\tau(t)$ are independently Gamma distributed. The VG process has been introduced, though not a complete model, as it better mimics the return of an actual stock as it allows for skewness, leptokurticity, and its sample path appears as a series of jumps, which reflects the change in stock value due to successive trades on the stock.

7.1.2 The Variance-Gamma Process

In defining the VG process, we first give a precise definition of the subordinator process $\{\tau(t) : t \geq 0\}$. Let $\gamma(a, b)$ denote the Gamma distribution with shape parameter a and scale parameter b. Recall that for $a \in \mathbb{N}$, $\gamma(a, b)$ represents the distribution of the sum of a independent exponential mean b random variables. Let ν denote the variance-rate of the gamma time change, then $\tau(t)-\tau(s) = \tau(t-s)$, $0 \leq s < t$, is $\gamma(t/\nu, 1/\nu)$ distributed, i.e., the mean value of $\tau(t)$ is t and the variance of $\tau(t)$ is given by νt.

The VG-process is determined through choosing appropriate values for its parameters (r, σ, ν, κ), where r is, like in the BSM model, the interest rate on a risk free bond, σ denotes in analogy with the BSM model the implied volatility, ν is the parameter determining the Gamma process and models the kurtosis of the stock price process, and κ is an artificial parameter that allows to introduce asymmetry in the model. With mean $\mathbb{E}[\tau(t)] = t$, the Gamma process represents a business 'time' of which each event represents a 'trade' during the trading time of the stock. The Gamma process is composed of a countable number of very small positive jumps over any given time interval. More specifically, $\tau(t) < t$ represents trading of the stock in a market that is relatively quiet, and $\tau(t) > t$ represents trading of the stock in a market that is relatively active.

As the VG-process has more parameters than the BSM model it encompasses more features of the stock price process. It is worth noting that efficient statistical estimators for determining the actual values of (r, σ, ν, κ) exist in the literature, which makes the VG-process interesting from a practical point of view. Like for the BSM model we may simplify the presentation of the VG process as we are only interested in the marginal distribution of the process at some time point t. The following construction of the VG-process follows [8].

Let $\{W(t) : t \geq 0\}$ denote the Wiener process, then

$$X(t) = \kappa\tau(t) + \sigma W(\tau(t))$$

yields a VG-process. The price of the asset at time t in the VG-model is given by

$$S(t) = S(0)e^{(r+\omega)t+X(t)},$$

where

$$\omega = \frac{1}{\nu}\ln\left(1 - \kappa\nu - \frac{\sigma^2\nu}{2}\right), \tag{34}$$

is chosen such $\{S(t) : t \geq 0\}$ becomes a martingale, i.e., $\mathbb{E}[e^{-rt}S(t)|S(0)] = S(0)$. Like for the BSM model we are only interested in the marginal distribution of $\{S(t) : t \geq 0\}$ at time t. Therefore, we fix t and set

$$\alpha(y) := \alpha_t(y) = \ln(S(0)) + (r+\omega)t + \kappa y \text{ and } \beta(y) := \beta_t(y) = \sigma\sqrt{y}$$

and we let $Y(y)$ be a normal random variable with mean $\alpha(y)$ and standard deviation $\beta(y)$. Then, the price of the asset at time t in the VG-model is in distribution equal to

$$S(t) = e^{Y(\tau(t))}.$$

In the following we will derive the quantile sensitivity with respect to the initial price $S(0)$ and the implied volatility σ. Note that these parameters are only present in the normal distribution and do not affect the gamma process. We can write the density function of $S(t)$ in the VG model at time $t > 0$, denoted by $\phi^{\mathrm{VG}}(\cdot, t)$, as

$$\phi^{\mathrm{VG}}(x,t) = \int_0^\infty \phi_{\alpha(y),\beta(y)}(x)\frac{y^{\frac{t}{\nu}-1}e^{-\frac{y}{\nu}}}{\Gamma\left(\frac{t}{\nu}\right)\nu^{\frac{t}{\nu}}}\,dy$$

where $\phi_{\mu,\sigma}(x)$ denotes the density function of the normal distribution with mean μ and standard deviation σ. By Fubini's Theorem, the cdf. of $S(t)$ in the VG-model at time $t > 0$ is obtained as

$$\mathcal{VG}(x,t) = \int_0^\infty \mathcal{N}(\alpha(y),\beta(y))(x)\frac{y^{\frac{t}{\nu}-1}e^{-\frac{y}{\nu}}}{\Gamma\left(\frac{t}{\nu}\right)\nu^{\frac{t}{\nu}}}\,dy.$$

For a parameter θ present either in the mean or the standard deviation of the normal cdf., the Measure Valued Derivative of the VG model can be written as follows

$$\frac{\partial}{\partial\theta}\mathcal{VG}(x,t) = \int_0^\infty \left(\frac{d}{d\theta}\alpha(y)\mathcal{N}_{\boldsymbol{\mu}}(\alpha(y),\beta(y))(x) + \frac{d}{d\theta}\beta(y)\mathcal{N}_{\boldsymbol{\sigma}}(\alpha(y),\beta(y))(x)\right)\frac{y^{\frac{t}{\nu}-1}e^{-\frac{y}{\nu}}}{\Gamma\left(\frac{t}{\nu}\right)\nu^{\frac{t}{\nu}}}\,dy, \tag{35}$$

where $\mathcal{N}_{\boldsymbol{\mu}}(\cdot,\cdot)$ denotes the weak derivative of the normal distribution with respect to the mean and $\mathcal{N}_{\boldsymbol{\sigma}}(\cdot,\cdot)$ denotes the weak derivative of the normal distribution with respect to the standard deviation, see Example 5.2 for details. Note that interchanging of derivative and integral in (35) is permitted as the density function of a gamma distribution is B_v differentiable with $v(x) = 1$, since the normal distribution is a bounded function.

7.2 A Plain Vanilla Option

In this section we determine quantile sensitivities of a vanilla call option with respect to the initial price as well as to the implied volatility.

Let r denote the interest rate, then the value of a call option at time t is given by

$$H_1(S(t)) = e^{-rt}\max(S(t)-K,0) \;=\; h_1(X_{a(t),b(t)}),$$

with

$$h_1(x) = e^{-rt}\max(e^x - K, 0), \tag{36}$$

where $K > 0$ denotes the strike price, i.e., the price at which the stock can be purchased within the contract. The pay-off mapping $H_1(x)$ is continuous throughout $\mathbb{R}$ and differentiable on $\mathbb{R}\setminus\{K\}$. As $S(t) \neq K$ with probability one, the conditions for IPA are met and the IPA quantile estimator is applicable.

7.2.1 The α-quantile Delta

Suppose now that we are interested in the sensitivity of the α-quantile of $H_1(S(t))$ with respect to the initial stock price $S(0)$, and consider first the Black-Scholes-Merton model. In other words, we are interested in the "α-quantile Delta." To simplify the notation, we suppress the argument "t" in $a(t)$ and $b(t)$. In light of the representation $H_1(S(t)) = h_1(X_{a,b})$, the main task in applying our estimator is to compute the derivative of the cdf. of $X_{a,b}$ with respect to $S(0)$. We apply (30) in Example 5.2 to $\theta = S(0)$, which yields

$$\frac{\partial}{\partial S(0)}\mathcal{N}(a,b)(x) = -\frac{1}{S(0)}\frac{1}{\sqrt{2\pi}b}e^{-\frac{(x-a)^2}{2b^2}},$$

for $x \in \mathbb{R}$. Note that the inverse of $h_1(x)$ for $h_1(x) > 0$ is given by

$$h_1^{-1}(x) = \ln(xe^{rt} + K), \quad x > 0.$$

Inserting the expression for $\partial\mathcal{N}(a,b)(x)/\partial S(0)$ for F_θ' into the estimator (6) yields

$$m\big(Z_{(\lceil\alpha m\rceil):m} - Z_{(\lceil\alpha m\rceil)-1:m}\big)\frac{1}{S(0)}\frac{1}{\sqrt{2\pi}b}e^{-\frac{(h_1^{-1}(Z_{\lceil\alpha m\rceil:m})-a)^2}{2b^2}}.$$

For the VG-model, using integral result 3.471.9 in Gradshteyn & Rhyzik, [15], the partial derivative w.r.t. $S(0)$ has the form

$$\begin{aligned}\frac{\partial}{\partial S(0)}\mathcal{VG}(x,t) &= -\frac{1}{S(0)}\int_0^\infty \frac{1}{\sqrt{2\pi}\beta(y)}e^{-\frac{(x-\alpha(y))^2}{2(\beta(y))^2}}\frac{y^{\frac{t}{\nu}-1}e^{-\frac{y}{\nu}}}{\Gamma\left(\frac{t}{\nu}\right)v^{\frac{t}{\nu}}}\,dy \\ &= -\frac{2}{\sqrt{2\pi}\sigma S(0)\Gamma(\frac{t}{\nu})\nu^{\frac{t}{\nu}}}e^{\frac{\kappa}{\sigma^2}\eta(x)}\left(\frac{\eta^2(x)}{\kappa^2+\frac{2\sigma^2}{\nu}}\right)^{\frac{t}{2\nu}-\frac{1}{4}}K_{\frac{t}{\nu}-\frac{1}{2}}\left(\frac{1}{\sigma^2}\sqrt{\eta^2(x)\left(\kappa^2+\frac{2\sigma^2}{\nu}\right)}\right),\end{aligned}$$

with

$$\eta(x) = x - \ln(S(0)) - (r+w)t, \tag{37}$$

where ω is defined in (34), and K_n is the modified Bessel function of the second kind with order n. The consequent estimator in the form of Equation (6) is

$$\begin{aligned}&\frac{2}{\sqrt{2\pi}\sigma S(0)\Gamma(\frac{t}{\nu})v^{\frac{t}{\nu}}}m\big(Z_{(\lceil\alpha m\rceil):m} - Z_{(\lceil\alpha m\rceil)-1:m}\big) \\ &\cdot e^{\frac{\kappa}{\sigma^2}\eta(h_1^{-1}(Z_{\lceil\alpha m\rceil:m}))}\left(\frac{\eta^2(h_1^{-1}(Z_{\lceil\alpha m\rceil:m}))}{\kappa^2+\frac{2\sigma^2}{\nu}}\right)^{\frac{t}{2\nu}-\frac{1}{4}}K_{\frac{t}{\nu}-\frac{1}{2}}\left(\frac{1}{\sigma^2}\sqrt{\eta^2(h_1^{-1}(Z_{\lceil\alpha m\rceil:m}))\left(\kappa^2+\frac{2\sigma^2}{\nu}\right)}\right).\end{aligned}$$

Provided that conditions **(A1)** to **(A4)** hold, we have strongly consistent estimator for $\partial q_\alpha/\partial S(0)$ and a CLT, where the order statistic is obtained from $\mathbf{Z}$ which is a sample of m i.i.d. copies of $H_1(S(t)) = h_1(X_{a(t),b(t)})$, respectively $h_1(Y_{\alpha(t),\beta(t)})$.

In the following we show that the conditions for applying our quantile sensitivity estimator to the call option for both pricing models are satisfied. For assumption **(A1)**, the moment generating function (mgf) is finite for the normal distribution and sampling $\mathbf{Z}$ as an i.i.d. vector satisfies **(A1)**. For the VG-model, a similar line of argument holds for $S(t)$, though we require the more stringent constraint $2(\kappa\nu+\sigma^2\nu) < 1$ for the second moment to be finite. The proof that **(A2)** is satisfied for the BSM and the VG-process is postponed to the Appendix. We now turn to (**A3**) and **(A4)**. As h_1 is a measurable mapping, $\mathcal{B}^b$-differentiability of normal and variance gamma distribution carries over to $\mathcal{B}^b$-differentiability of the option value. Hence, **(A3)** and **(A4)** then follow from applying Lemma 5.1. This assumption also holds for the VG-process for the same reason.

7.2.2 The α-quantile Vega

Let us consider the same situation but with the parameter being the implied volatility, σ. In parlance, this is considered the α-quantile Vega. Letting $\theta = \sigma$, for the BSM model (where we simplify the notation by suppressing the argument "t" in $a(t)$ and $b(t)$), it follows from Example 5.2 that

$$\begin{aligned}\frac{\partial}{\partial\sigma}\mathcal{N}(a,b)(x) &= -\sigma t \mathcal{N}_{\boldsymbol{\mu}}(a,b)(x) + \sqrt{t}\mathcal{N}_{\boldsymbol{\sigma}}(a,b)(x) \\ &= \sqrt{t}\left(1 - \frac{(x-a)}{b^2}\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2b^2}},\end{aligned} \tag{38}$$

for $x \in \mathbb{R}$. Letting $F'_\theta(x) = \partial\mathcal{N}(a,b)(x)/\partial\sigma$, as provided in (38), would then yield the estimator in (6). We will however not choose (6) as our estimator but we rather will illustrate the application of the general estimator put forward in (8). Inserting the representation for the derivatives obtained in Example 5.2 into (38) yields

$$\begin{aligned}\frac{\partial}{\partial\sigma}\mathcal{N}(a,b)(x) &= -\frac{t}{b\sqrt{2\pi}}\left(R_b(x-a)1_{x\geq a} - R_b(|x-a|)1_{x\leq a}\right) \\ &\quad + \frac{\sqrt{t}}{b}\left(\mathcal{DM}(a,b)(x) - \mathcal{N}(a,b)(x)\right) \\ &= (c_1+c_2)\Bigg(\frac{c_1}{c_1+c_2}R_b(|x-a|)1_{x\leq a} + \frac{c_2}{c_1+c_2}\mathcal{DM}(a,b)(x) \\ &\quad - \frac{c_1}{c_1+c_2}R_b(x-a)1_{x\geq a} - \frac{c_2}{c_1+c_2}\mathcal{N}(a,b)(x)\Bigg)\end{aligned}$$

for $x \in \mathbb{R}$, where

$$c_1 = \frac{t}{\sqrt{2\pi}} \quad \text{and} \quad c_2 = \sqrt{t}.$$

We now denote $R_1 \sim R_b(|x-a|)1_{x\geq a}$, $R_2 \sim R_b(|x-a|)1_{x\leq a}$, being the random variables associated with the Rayleigh cumulative distribution functions and $M \sim \mathcal{DM}(a,b)(x)$ being the Double-Maxwell random variable. We denote X^+ according to

$$X^+ = \frac{c_1}{c_1+c_2}R_2 + \frac{c_2}{c_1+c_2}M$$

i.e., with probability $c_1/(c_1+c_2)$ let X^+ be distributed according to a shifted Rayleigh distribution and with probability $c_2/(c_1+c_2)$ let X^+ follow a Double-Maxwell distribution. In the same vein, let X^- be defined as

$$X^- = \frac{c_1}{c_1+c_2}R_1 + \frac{c_2}{c_1+c_2}X$$

i.e., with probability $c_1/(c_1+c_2)$ let X^+ be distributed according to a shifted negative Rayleigh distribution and with probability $c_2/(c_1+c_2)$ let X^- follow a normal distribution. With this in mind, the estimator in Equation (8) becomes

$$-m(c_1+c_2)\big(Z_{\lceil\alpha m\rceil:m} - Z_{\lceil\alpha m\rceil-1:m}\big)\Big(1_{h_2(X^+)\leq Z_{\lceil\alpha m\rceil:m}} - 1_{h_2(X^-)\leq Z_{\lceil\alpha m\rceil:m}}\Big).$$

For the VG-model, the estimator via Equation (6) is analogous where the measure-valued derivative random variables are also mean-variance mixtures with a Gamma distributed mixing distribution. There is also the additional complication with the pre-factor depending on τ. This will be discussed shortly. We can obtain a closed form expression for $\partial\mathcal{VG}(x)/\partial\sigma$ following Equation (35) with $\eta(x)$ defined as in

(37). After some calculation, with the result of the same integral, 3.471.9 in Gradshteyn & Rhyzik, [15], we get

$$\begin{aligned}
\frac{\partial}{\partial\sigma}\mathcal{VG}(x,t) &= \int_0^\infty \frac{1}{\sqrt{2\pi}\beta(y)}\left(\frac{\sigma t}{1-\kappa\nu-\frac{\sigma^2}{2}\nu} - \frac{\sqrt{y}(x-\alpha(y))}{\beta(y)}\right)e^{-\frac{1}{2}\left(\frac{x-\alpha(t)}{\beta(t)}\right)^2}\frac{y^{\frac{t}{\nu}-1}e^{-\frac{y}{\nu}}}{\Gamma\left(\frac{t}{\nu}\right)\nu^{\frac{t}{\nu}}}\,dy \\
&= \frac{2e^{\frac{\kappa}{\sigma^2}\eta(x)}}{\sqrt{2\pi}\sigma\Gamma\left(\frac{t}{\nu}\right)\nu^{\frac{t}{\nu}}}\left(\left(\sigma t e^{-\nu\omega} - \frac{\eta(x)}{\sigma}\right)\left(\frac{\eta^2(x)}{\kappa^2+\frac{2\sigma^2}{\nu}}\right)^{\frac{t}{2\nu}-\frac{1}{4}} K_{\frac{t}{\nu}-\frac{1}{2}}\left(\frac{1}{\sigma^2}\sqrt{\eta^2(x)\left(\kappa^2+\frac{2\sigma^2}{\nu}\right)}\right)\right. \\
&\quad \left. + \frac{\kappa}{\sigma}\left(\frac{\eta^2(x)}{\kappa^2+\frac{2\sigma^2}{\nu}}\right)^{\frac{t}{2\nu}+\frac{1}{4}} K_{\frac{t}{\nu}+\frac{1}{2}}\left(\frac{1}{\sigma^2}\sqrt{\eta^2(x)\left(\kappa^2+\frac{2\sigma^2}{\nu}\right)}\right)\right).
\end{aligned}$$

The resulting estimator from Equation (6), where $h_1^{-1}(x)$ for $x>0$ defined previously, has the formula

$$\begin{aligned}
&-\frac{2}{\sqrt{2\pi}\sigma\Gamma\left(\frac{t}{\nu}\right)\nu^{\frac{t}{\nu}}} m\big(Z_{(\lceil\alpha m\rceil):m} - Z_{(\lceil\alpha m\rceil)-1:m}\big) \\
&\cdot\left(\left(\sigma t e^{-\nu\omega} - \frac{\eta(h_1^{-1}(Z_{\lceil\alpha m\rceil:m}))}{\sigma}\right)\left(\frac{\eta^2(h_1^{-1}(Z_{\lceil\alpha m\rceil:m}))}{\kappa^2+\frac{2\sigma^2}{\nu}}\right)^{\frac{t}{2\nu}-\frac{1}{4}} K_{\frac{t}{\nu}-\frac{1}{2}}\left(\frac{1}{\sigma^2}\sqrt{\eta^2(h_1^{-1}(Z_{\lceil\alpha m\rceil:m}))\left(\kappa^2+\frac{2\sigma^2}{\nu}\right)}\right)\right. \\
&\left. + \frac{\kappa}{\sigma}\left(\frac{\eta^2(h_1^{-1}(Z_{\lceil\alpha m\rceil:m}))}{\kappa^2+\frac{2\sigma^2}{\nu}}\right)^{\frac{t}{2\nu}+\frac{1}{4}} K_{\frac{t}{\nu}+\frac{1}{2}}\left(\frac{1}{\sigma^2}\sqrt{\eta^2(h_1^{-1}(Z_{\lceil\alpha m\rceil:m}))\left(\kappa^2+\frac{2\sigma^2}{\nu}\right)}\right)\right).
\end{aligned}$$

As for the BSM model, for the general estimator, differentiation w.r.t. σ of $\mathcal{VG}(\cdot,t)$ is split up in relation to the mean and standard deviation obtaining similar conditional MVD distributions. By (35),

$$\begin{aligned}
\mathcal{N}_{\boldsymbol{\sigma}}(\alpha(y),\beta(y))(x) &= -\frac{\sigma t}{1-\kappa\nu-\frac{\sigma^2}{2}\nu}\mathcal{N}_{\boldsymbol{\mu}}(\alpha(y),\beta(y))(x) + \sqrt{y}\mathcal{N}_{\boldsymbol{\sigma}}(\alpha(y),\beta(y))(x) \\
&= \frac{\sigma t}{\sqrt{2\pi}\beta(y)}e^{-\nu\omega}\left(R_{\beta(y)}(|x-\alpha(y)|)1_{x\le\alpha(y)} - R_{\beta(y)}(x-\alpha(y))1_{x\ge\alpha(y)}\right) \\
&\quad + \frac{1}{\sigma}\left(\mathcal{DM}(\alpha(y),\beta(y))(x) - \mathcal{N}(\alpha(y),\beta(y))(x)\right) \\
&= (d_1(y)+d_2)\left(\frac{d_1(y)}{d_1(y)+d_2}R_{\beta(y)}(|x-\alpha(y)|)1_{x\le\alpha(y)} + \frac{d_2}{d_1(y)+d_2}\mathcal{DM}(\alpha(y),\beta(y))(x)\right. \\
&\quad \left. - \frac{d_1(y)}{d_1+d_2(y)}R_{\beta(y)}(x-\alpha(y))1_{x\ge\alpha(y)} - \frac{d_2}{d_1(y)+d_2}\mathcal{N}(\alpha(y),\beta(y))(x)\right)
\end{aligned}$$

for $x\in\mathbb{R}$, where

$$d_1(y) = \frac{\sigma t}{\sqrt{2\pi}\beta(y)}\frac{1}{1-\kappa\nu-\frac{\sigma^2\nu}{2}} \quad \text{and} \quad d_2 = \frac{1}{\sigma},$$

which is a Gamma random variable dependent pre-factor. Consequently,

$$\frac{\partial}{\partial\sigma}\mathcal{VG}(x,t) = \int_0^\infty \mathcal{N}_{\boldsymbol{\sigma}}(\alpha(y),\beta(y))(x)\frac{y^{\frac{t}{\nu}-1}e^{-\frac{y}{\nu}}}{\Gamma\left(\frac{t}{\nu}\right)\nu^{\frac{t}{\nu}}}\,dy.$$

We denote $\tau\sim\gamma(t/\nu,1/\nu)$ as the mixing random variable, and $S_1\sim R_1(x)1_{x>0}$, $S_2\sim R_1(|x|)1_{x<0}$, and $L\sim\mathcal{DM}(0,1)(x)$ as the standardized Rayleigh and Double-Maxwell random variables. We now denote $R_1=\alpha(\tau)+\beta(\tau)S_1$, $R_2=\alpha(\tau)+\beta(\tau)S_2$ as the Rayleigh mean-variance mixtures and $M=\alpha(\tau)+\beta(\tau)L$ respectively as the Double-Maxwell variant. As for the BSM model, we define Y^+ as

$$Y^+ = \frac{d_1(\tau)}{d_1(\tau)+d_2}R_2 + \frac{d_2}{d_1+d_2}M \tag{39}$$

that is with probability $d_1(\tau)/(d_1(\tau)+d_2)$, depending on the Gamma distributed random variable, Y^+ is a shifted Rayleigh distribution on the negative half-line and with probability $d_2(\tau)/(d_1+d_2(\tau))$, Y^+ is distributed with as Double-Maxwell distribution. In addition, Y^- is defined via

$$Y^- = \frac{d_1(\tau)}{d_1(\tau)+d_2} R_1 + \frac{d_2}{d_1(\tau)+d_2} X, \tag{40}$$

i.e. as a shifted Rayleigh distribution on the postive half-line with contingent probability $d_1(\tau)/(d_1(\tau)+d_2)$, and as a normal random variable with probability $d_2/(d_1(\tau)+d_2)$, the general estimator, Equation (8), has the expression

$$-m(d_1(\tau(t))+d_2)\left(Z_{\lceil \alpha m\rceil:m} - Z_{\lceil \alpha m\rceil-1:m}\right)\left(1_{h_1(Y^+)\le Z_{\lceil \alpha m\rceil:m}} - 1_{h_1(Y^-)\le Z_{\lceil \alpha m\rceil:m}}\right).$$

The order statistic within both estimators are obtained from $\mathbf{Z}$ which is a sample of m i.i.d. copies of either $H_1(S(t)) = h_1(X_{a(t),b(t)})$ or $H_1(S(t)) = h_1(Y_{\alpha(t),\beta(t)})$, which shows that **(A1)** holds. Assumptions **(A3)** to **(A4)** are identical from Section 7.2, where we use the fact that $h_1(\cdot)$ is an element of $\mathcal{B}_{1+e^x}$. For **(A2)** we refer to the Appendix.

Remark 7.1 *The above analysis shows that Y^+ and Y^- constructed in (39) and (40) is a weak derivative of the Variance Gamma process. This is an important result in its own right, as this is the first unbiased estimator for the derivative of the Variance Gamma process with respect to σ.*

7.3 Rainbow Options

Rainbow Options are contingent financial claims that are based on multiple stocks. Let $S(t) = (S_1(t), S_2(t))$, with $t \ge 0$, denote the joint price vector of two assets, where the price of an individual stock is depicted by the Black-Scholes-Merton model

$$S_i(t) = S_i(0) e^{\left(r - \frac{\sigma_i^2}{2}\right)t + \sigma_i \sqrt{t} X_i} := e^{X_{a_i(t),b_i(t)}},$$

for $i = 1, 2$, and $X = (X_1, X_2)$, a standard normal random vector. For the two assets, we define

$$a_i(t) = \ln S(0) + (r - \sigma_i^2/2)t \quad \text{and} \quad b_i(t) = \sigma_i \sqrt{t},$$

following the notation of the BSM model in the previous section. We note here for the normal random variables

$$X_{a_i(t),b_i(t)} = a_i(t) + b_i(t) X_i,$$

and for the moment we will just consider the standard random vector X which has the Lebesgue density

$$\phi(x,\Sigma) = \frac{1}{2\pi\sqrt{|\Sigma|}} e^{-\frac{1}{2} x^T (\Sigma)^{-1} x/2}, \tag{41}$$

for $x = (x_1, x_2)$. The covariance matrix Σ is given via

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

with $\rho \in [-1, 1]$ being the correlation of X_1 and X_2. Specifically,

$$|\Sigma| = 1 - \rho^2 \quad \text{and} \quad \Sigma^{-1} = \frac{1}{1-\rho^2} \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}.$$

For $x = (x_1, x_2)^T$, it holds

$$x^T (\Sigma)^{-1} x = (x_1^2 - 2\rho x_1 x_2 + x_2^2)/(1-\rho^2),$$

and we may therefore write the Lebesgue density in (41) as a product of two univariate normal densities with X_2 conditioned on X_1:

$$\begin{aligned}\phi(x,\Sigma) &= \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(x_1^2-2\rho x_1 x_2+x_2^2)} \\ &= \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} e^{-\frac{1}{2}\frac{(x_2-\rho x_1)^2}{1-\rho^2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_1^2}.\end{aligned}$$

In the following we consider the spread option and consider the sensitivity of the quantile w.r.t. the parameter ρ. Let

$$H_2(S(t)) = \max(S_2(t) - S_1(t) - K, 0),$$

for $K > 0$. For $x = (x_1, x_2)$ we set

$$h_2(x) = \max(e^{x_2} - e^{x_1} - K, 0),$$

then

$$H_2(S(t)) = h_2(X_1, X_2).$$

Since the parameters $a_i(t), b_i(t)$, for $i = 1, 2$, are independent of ρ, the eventual final measure-valued derivative will just be an affine transformation of $X_2^{\pm}$.

Provided above, the conditional distribution to X_2 given $X_1 = y$ is $\mathcal{N}(\rho x, \sqrt{1-\rho^2})$. Evoking Example 5.2 it follows that

$$\frac{\partial}{\partial\rho}\mathcal{N}(\rho y, \sqrt{1-\rho^2}) = y\mathcal{N}_{\boldsymbol{\mu}}(\rho y, \sqrt{1-\rho^2}) - \frac{\rho}{\sqrt{1-\rho^2}}\mathcal{N}_{\boldsymbol{\sigma}}(\rho y, \sqrt{1-\rho^2}).$$

Consider a digital spread option that entitles the holder a profit of $S_1(t) - S_2(t)$ provided that $S_1(t) - S_2(t) \geq K$, for some $K > 0$, and otherwise nothing. Hence, $H_2(S(t)) = (S_1(t) - S_2(t))1_{S_2(t)-S_1(t)\geq K} = h_2(X_1, X_2)$. Provided that $X_1 = y$, $h_2(y, \cdot)$ becomes invertible and the conditional version of the estimator in (6) applies.

In the following we show how to represent the distributional derivatives in terms of random variables so as to obtain an estimator in the form of Equation (8), respectively (9). Rearranging the positive and negative parts, we arrive at

$$\begin{aligned}\frac{\partial}{\partial\rho}\mathcal{N}(\rho y, \sqrt{1-\rho^2})(x) &= \frac{y}{\sqrt{2\pi}\sqrt{1-\rho^2}}\left(R_{\sqrt{1-\rho^2}}(x-\rho y)1_{x\geq\rho y} - R_{\sqrt{1-\rho^2}}(|x-\rho y|)1_{x<\rho y}\right) \\ &= -\frac{\rho}{1-\rho^2}\left(\mathcal{DM}(\rho y, \sqrt{1-\rho^2})(x) - \mathcal{N}(\rho y, \sqrt{1-\rho^2})(x)\right) \\ &= (c_1(y)+c_2)\Bigg(\frac{c_1(y)}{c_1(y)+c_2}R_{\sqrt{1-\rho^2}}(y-\rho x)1_{y\geq\rho x} + \frac{c_2}{c_1(y)+c_2}\mathcal{N}(\rho x, \sqrt{1-\rho^2})(y) \\ &\quad - \frac{c_1(y)}{c_1(y)+c_2}R_{\sqrt{1-\rho^2}}(|y-\rho x|)1_{y<\rho x} - \frac{c_2}{c_1(y)+c_2}\mathcal{DM}(\rho x, \sqrt{1-\rho^2})(y)\Bigg)\end{aligned}$$

where

$$c_1(y) = \frac{y}{\sqrt{2\pi}\sqrt{1-\rho^2}} \quad \text{and} \quad c_2 = \frac{\rho}{1-\rho^2}.$$

Accordingly, we can a define Rayleigh distribution random variables $S_1 \sim R_{\sqrt{1-\rho^2}}(x)1_{x\geq 0}$ and $S_2 \sim R_{\sqrt{1-\rho^2}}(|x|)1_{x<0}$ in which we can write the X_1-dependent shift Rayleigh random variables by $R_1 = \rho X_1 + S_1$ and $R_2 = \rho X_1 + S_2$. Similarly for the conditional Double-Maxwell random variable with $L \sim \mathcal{DM}(0, \sqrt{1-\rho^2})(x)$, $M = \rho X_1 + L$. The final form of the positive component X_2^+ is

$$X_2^+ = \frac{c_1(X_1)}{c_1(X_1)+c_2}R_1 + \frac{c_2}{c_1(X_1)+c_2}X_2$$

which is a shifted Rayleigh random variable with normally distributed probability $c_1(X_1)/(c_1(X_1)+c_2)$ and the original, conditioned normal random variable with probability $c_2/(c_1(X_1)+c_2)$. For the random variable X_2^-, its expression is similar

$$X_2^- = \frac{c_1(X_1)}{c_1(X_1)+c_2} R_2 + \frac{c_2}{c_1(X_1)+c_2} M,$$

a shifted Rayleigh random variable with normally distributed probability $c_1(X_1)/(c_1(X_1)+c_2)$ or otherwise a conditioned Double-Maxwell distributed random variable.

The sampling scheme for the derivative of the joint distribution of (X_1, X_2) is as follows. First, a standard normal random variable is sampled, yielding $X_1 = x$. Then, given this value, the positive part and the negative part of X_2, denoted by X_2^+ and X_2^-, are sampled according to the above distributions. Eventually, $X^{\pm}_{a_2(t),b_2(t)}$ can be obtained from

$$X^{\pm}_{a_2(t),b_2(t)} = a_2(t) + b_2(t) X_2^{\pm},$$

and the general estimator in (8) becomes

$$-\, m(c_1(X_1)+c_2)\left(1_{h_2(X_1,X_2^+)\le Z_{\lceil \alpha m\rceil:m}} - 1_{h_2(X_1,X_2^-)\le Z_{\lceil \alpha m\rceil:m}}\right),$$

where $Z_{\lceil \alpha m\rceil:m}$ is obtained from an independent i.i.d. sample of $h_3(X_1, X_2)$, and the pre-factor gives the normalizing constant originating from the mixture interpretation of the weak derivative. The fact that conditions **(A1)**, **(A2)** and **(A4)** hold follows readily from arguments provided in Section 7.2. For **(A3)** note that we cannot apply the product rule of weak differentiation to (X_1, X_2) as the components are not independent. However, since $1_{h_2(x,y)\le z} \in \mathcal{B}^b$, it follows from the dominated convergence theorem that

$$\partial_\rho \int\int 1_{h_2(x,y)\le z} \mathcal{N}(\rho x, \sqrt{1-\rho^2})(dy)\mathcal{N}(dx) = \int \left(\int 1_{h_2(x,y)\le z} \partial_\rho \mathcal{N}(\rho x, \sqrt{1-\rho^2})(dy)\right)\mathcal{N}(dx).$$

for any z, which implies that $h(X_1, X_2)$ is $\mathcal{B}^b$-differentiable w.r.t. ρ. Condition **(A3)** then follows from Lemma 5.1.

Sampling the derivative information via a conditional distribution may not be the numerically most efficient way and directly differentiating the multi-variate normal distribution seems more natural. Unfortunately, the latter approach leads to rather complex expression for the weak derivatives and efficient sampling from these expressions is still an open problem.

7.4 The VAR of a Portfolio

We consider a finite set of assets $j = 1, \ldots, J$, which can be any kind of financial assets such as bonds, stocks, futures, and options. Following the model provided in [10], a portfolio, at a certain time, is given by a vector $x = (x_, \ldots, x_J)$ of positions representing the amount of money allocated to each of these assets. The uncertain value of the financial assets in the next period are given by Y_j with c.d.f. $F_{j,\theta}$, $1 \le j \le J$. In the following we will assume that the portfolio consists of long equity positions, and that the risk factors relate to the stock prices at the end of the next period. In this case the value of the portfolio is linear with respect to the portfolio positions as well as with respect to the risk factors, and we let

$$V := V(x_1, \ldots, x_J; Y_1, \ldots, Y_J) = \sum_{j=1}^{J} x_j Y_j$$

denote the value of the portfolio at the end of the next period, and denote the c.d.f. of V by F. Let $q_\alpha(x_1, \ldots, x_J)$ denote the α-quantile of the portfolio value, and let $\mathbf{V}(x_1, \ldots, x_J)$ denote the expected value of the uncertain portfolio value. Then the value at risk (VAR) of the portfolio is given by

$$V^*(x_1, \ldots, x_J) = \mathbf{V}(x_1, \ldots, x_J) - q_\alpha(x_1, \ldots, x_J),$$

which describes the maximal amount by which the portfolio can fall short of its expected value with probability α. See [10] for more details.

Assume that θ is a distributional parameter of the underlying stocks. Then, the sensitivity of V^* with respect to θ will allow us to assess the robustness of the VAR with respect to small changes in θ. We will consider the simpler case that Y_j, $1 \le j \le J$, are given by stock prices. We focus on $\partial_\theta q_\alpha(x_1, \dots, x_J)$. Let $\mathbf{Z} = (Z_i : 1 \le i \le n)$ be an i.i.d. sample of $V(x_1, \dots, x_J; Y_1, \dots, Y_J)$. Assume that the price Y_i is given by e^{W_i}, where W_i represent an appropriate random variable resulting in a BSM or VG-process model of the stock price; see Section 7.1 for details. Fix $x_1, \dots, x_J$, and let

$$h(w_1, \dots, w_J) = \sum_{j=1}^{J} x_j e^{w_j}.$$

and, for $1 \le j \le J$, let

$$h^{-1}_{w_1,\dots,w_{j-1},w_{j+1},\dots,w_J}(z) = \ln \left(\frac{1}{x_j} \left(z - \sum_{i=1, i \ne j}^{J} x_i e^{w_i} \right) \right)$$

provided that $z \ge \sum_{i=1,i\ne j}^{J} x_i e^{w_i}$ and otherwise zero. Then, by Lemma 5.4 it holds that

$$\partial_\theta \mathbb{P}(V \le z) = \partial_\theta F(z) = J \mathbb{E}_\theta \left[\partial F_{\rho,\theta} \left(h^{-1}_{W_1,\dots,W_{j-1},W_{j+1},\dots,W_J}(z) \right) \right].$$

Importance sampling can be used in order to increase the probability of a non-zero derivative contribution. Hence, for θ either the initial state $S(0)$ or σ in either the BSM model or the VG-process, the sensitivity of the quantile of the portfolio can be estimated by combining the above distributional estimator with the inverse density estimator, see Section 5.3, where the particular form of $\partial F_{\rho,\theta}$ can be found in Section 7.1. The above result can be extended to a portfolio of options in a straightforward way.

Remark 7.2 *Let (c_θ, W_j^+, W_j^-) denote the $\mathcal{B}^b$-derivative of W_j, for $1 \le j \le J$. Following Lemma 5.3, an alternative estimator for $\partial_\theta \mathbb{P}(V \le z)$ can be obtained from*

$$c_\theta J \mathbb{E} \left[1_{\left(x_\rho e^{W_\rho^+} + \sum_{j=1, j\ne\rho}^{J} x_j e^{W_j} \right) \le z} - 1_{\left(x_\rho + e^{W_\rho^-} + \sum_{j=1, j \ne \rho}^{J} x_j e^{W_j} \right) \le z} \right],$$

with ρ uniformly distributed on $\{1, \dots, J\}$ and independent of everything else.

Remark 7.3 *Using the product rule of weak differentiation for general spaces (rather than $\mathcal{B}^b$), an unbiased estimator for $\mathbf{V}(x_1, \dots, x_J)$ of the type presented in Lemma 5.3 can be obtained as well.*

Acknowledgement

The authors are grateful to two anonymous referees whose comments and suggestions helped improving the paper.

Conclusion

In this paper we provided expressions for quantile sensitivities that can be used for gradient estimation by means of Monte Carlo simulation. The examples illustrate the flexibility of our measure-valued differentiation and statistical spacing theory based approach. Specifically, multivariate problems and models containing the Variance-Gamma process can be dealt with as well. Future research will be on the extension of our approach to time-dependent systems such as queueing networks.

References

[1] Asmussen, S., Glynn, P., 2007. *Stochastic Simulation.* Springer, New York NY.

[2] Bahadur, R., 1966. *A Note on Quantiles in Large Samples.* The Annals of Mathematical Statistics, 37 (3), 577-580.

[3] Burke, P., 1956. The output of a queueing system. *Operations Research*, 4 (6), 699-704.

[4] Cao, X. R., 2007. *Stochastic Learning and Optimization: A Sensitivity-Based Approach.* Springer, New York NY.

[5] David, H. A., Nagaraja, H. N., 2003. *Order Statistics.* 3rd ed., Wiley, Hoboken NJ.

[6] Fu, M. C. and Hu, J. Q., 1997. *Conditional Monte Carlo: Gradient Estimation and Optimization Applications.* Kluwer Academic Publishers, Boston MA.

[7] Fu, M. C., 2006. Gradient estimation. In *Handbook on Operations Research and Management Science: Simulation*, S. Henderson and B. Nelson, Eds. Vol. 13. Elsevier, 575–616.

[8] Fu, M. C., 2007. Variance-Gamma and Monte Carlo. In *Advances in Mathematical Finance*, M. C. Fu, R. Jarrow, J. Yen, and R. Elliott Eds., 21-25, Birkhäuser, Boston MA.

[9] Fu, M. C., Hu, J. Q., Hong, L. J., 2009. Conditional Monte Carlo Estimation of Quantile Sensitivities. *Management Science*, 55 (12), 2019-2027.

[10] Gaivoronski, A., Pflug, G., 2005. Value-at-risk in portfolio optimization: properties and computational approach. *Journal of Risk*, 7 (2), 1-31.

[11] Glasserman, P. 1991. *Gradient Estimation via Perturbation Analysis.* Kluwer Academic, Boston, MA.

[12] Glasserman, P., 2004. *Monte Carlo Methods in Financial Engineering.* Springer, New York.

[13] Glasserman, P., 1993. Regenerative derivatives for regenerative sequences, *Advances in Applied Probability*, 25 (1), 116-139.

[14] Ghosh M., Sukhatme, S., 1981. On Bahadur's representation of quantiles in nonregular cases. *Communications in Statistics - Theory and Methods*, 10 (3), 269-282.

[15] Gradstheyn, I. S., Rhyzik, I. M., 1980. *Table of Integrals, Series, and Products*, Academic Press, New York, NY.

[16] Grimmett, G. R., Stirzaker, D. R., 2001. *Probability and Random Processes*, 3rd ed., Oxford University Press, Oxford, UK.

[17] Hesse, C. H., 1990. A Bahadur-Type Representation for Empirical Quantiles of a Large Class of Stationary, Possibly Infinite, Linear Processes. *The Annals of Statistics.* 18 (3), 1188-1202.

[18] Liu, G., Hong, L. J., 2011. Kernel estimation of the Greeks for options with discounted payoffs *Operations Research*, 59 (1), 96-108.

[19] Heidergott, B., 2001. A differential calculus for random matrices with applications to (max,+)-linear stochastic systems. *Mathematics of Operations Research*, 26 (4), 679-699.

[20] Heidergott, B., Leahu H., 2010. Weak differentiability of product measures. *Mathematics of Operations Research*, 35 (1), 27-51.

[21] Heidergott, B., Fahrenhorst-Yuan, T. , 2010. Gradient estimation for a multi-component maintenance system using measure-valued differentiation, *Operations Research*, 58 (3), 706-718.

[22] Heidergott, B., Vázquez-Abad, F., Pflug, G., Farenhorst-Yuan, T., 2010. Gradient estimation for discrete-event systems by measure-valued differentiation. *Transactions on Modelling and Computer Simulation*, 20 (1), Article No. 5, (28 pages).

[23] Heidergott, B., Vázquez-Abad, F., Volk-Makarewicz, W., 2008. Sensitivity estimation for Gaussian systems, *European Journal of Operational Research*, 187, 193-207.

[24] Ho, Y. C. and Cao, X. R. 1991. *Perturbation Analysis of Discrete Event Dynamic Systems.* Kluwer Academic, Boston, MA.

[25] Hong, L. J., 2009. Estimating Quantile Sensitivities. *Operations Research.* 57 (1), 118-130.

[26] Hong, L. J., Liu, G., 2008. Simulating Sensitivities of Conditional Value at Risk, *Management Science*, 55 (2), 281-293.

[27] Hong, L. J., Liu, G., 2010. Pathwise Estimation of probability sensitivities through terminating or terminating or steady-state simulations, *Operations Research*, 58 (2), 357-370.

[28] Kallenberg, O., 2001. *Foundations of Modern Probability*, 2nd Ed., Springer, New York, NY.

[29] Kroese, D., Taimre, T., Botev, Z., 2011. *Handbook of Monte Carlo Methods*, Wiley Series in Probability, New York, NY.

[30] Liu, G., Hong, L. J., 2009. Kernel estimation of quantile sensitivities. *Naval Research Logistics*, 56 (6), 511-525.

[31] Hong, L. J., Liu, G., 2011. Kernel estimation of the Greeks for options with discounted payoffs, *Operations Research*, 59 (1), 96-108.

[32] Lyuu, Y.-D., Teng, H.-W., 2011. Unbiased and efficient Greeks of rainbow options. *Finance and Stochastics*, 15 (1), 141-181.

[33] Madan, D. B., Milne, F., 1991. Option Pricing with VG Martingale Components, Mathematical Finance 1 (4), 39-55.

[34] Madan, D. B., Carr, P. P., Chang, E. C., 1998. The Variance Gamma Process and Option Pricing. *European Financial Review*, 2, 79-105.

[35] Nielsen, L. 1999. *Pricing and Hedging of Derivative Securities*, Oxford University Press, Oxford, UK.

[36] Pyke, R., 1965. Spacings, *Journal of the Royal Statistical Society. Series B (Methodological)*, 27 (3), 395-449.

[37] Pflug, G., 1996. *Optimization of Stochastic Models.* Kluwer Academic, Norwell, MA.

[38] Rubinstein, R. , Shapiro, A. 1993. *Discrete Event Systems: Sensitivity Analysis and Optimization by the Score Function Method.* Wiley, Chichester, NY.

[39] Sen, P., 1968. Asymptotic Normality of Sample Quantiles for m-Dependent Processes. *The Annals of Mathematical Statistics*, 39 (5), 1724-1730.

[40] Sen, P., 1972. On the Bahadur representation of Sample Quantiles for Sequences of ϕ-Mixing Random Variables. *Journal of Multivariate Analysis.* 2 (1), 77-95.

[41] Serfling, R., 1980. *Approximation Theorems of Mathematical Statistics.* Wiley, New York, NY.

[42] Shiryaev, A., 1984. *Probability, 2nd ed.* Springer, New York, NY.

[43] Yu, J., 1986. Consistency of a k-nearest neighbor probability density function estimator. *Acta Mathematica Scientia*, 6, 467-477.

[44] Yu, J., 1993. Consistency of an alternative nearest neighbor probability density function estimator. *Statistical Research Report 1993-6*, Umeå University.

[45] Yu, J., 1993. Uniform convergence rates for a nearest neighbor estimator for a class of probability density functions. *Statistical Research Report 1993-11*, Umeå University.

[46] Yu, J., 1997. Uniform convergence rates for a nearest neighbor density estimator under dependence assumptions. *Communications in Statistics - Theory and Methods*, 26(3): 601-616.

Appendix

A.1 Condition (A2) for the α-quantile Delta

We now show that **(A2)** is satisfied for the plain vanilla option in Section 7.2. The support of $\mathcal{N}(a,b^2)$ is the entire real line, as well as for $\mathcal{VG}(\cdot,t)$ with $t>0$. For the normal density function

$$\frac{\partial}{\partial x}\phi_{a,b}(x) = -\frac{1}{\sqrt{2\pi}b}\frac{x-a}{b^2}e^{-\frac{1}{2}\frac{(x-a)^2}{b^2}},$$

which yields

$$\sup_{x\in\mathbb{R}}\left|\frac{\partial}{\partial x}\phi_{a,b}(x)\right| = \frac{1}{\sqrt{2\pi}b^2}e^{-\frac{1}{2}}.$$

Also,

$$\frac{\partial}{\partial S(0)}\phi_{a,b}(x) = \frac{1}{\sqrt{2\pi}S(0)b}\frac{x-a}{b^2}e^{-\frac{1}{2}\frac{(x-a)^2}{b^2}},$$

which yields

$$\sup_{x\in\mathbb{R}}\left|\frac{\partial}{\partial S(0)}\phi_{a,b}(x)\right| = \frac{1}{\sqrt{2\pi}S(0)b^2}e^{-\frac{1}{2}},$$

which establishes **(A2)** for the BSM model.

For the VG-process, we need to show that the supremum of the derivative of the density over a neighbourhood $B(\alpha)$ of the quantile is finite for a sufficiently large value of m. Note that for $t>0$ it holds that

$$\begin{aligned}\frac{\partial}{\partial x}\phi^{\mathrm{VG}}(x,t) &= \int_0^\infty \frac{\partial}{\partial x}\phi_{\alpha(y),\beta(y)}(x)\frac{y^{\frac{t}{\nu}-1}e^{-\frac{y}{\nu}}}{\Gamma\left(\frac{t}{\nu}\right)\nu^{\frac{t}{\nu}}}\,dy\\ &= \int_0^\infty -\frac{1}{\sqrt{2\pi}\beta(y)}\frac{x-\alpha(y)}{\beta^2(y)}e^{-\frac{1}{2}\frac{(x-\alpha(y))^2}{\beta^2(y)}}\frac{y^{\frac{t}{\nu}-1}e^{-\frac{y}{\nu}}}{\Gamma\left(\frac{t}{\nu}\right)\nu^{\frac{t}{\nu}}}\,dy\end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial S(0)}\phi^{\mathrm{VG}}(x,t) &= \int_0^\infty \frac{1}{\sqrt{2\pi}\beta(y)}\frac{x-\alpha(y)}{\beta^2(y)S(0)}e^{-\frac{1}{2}\frac{(x-\alpha(y))^2}{\beta^2(y)}}\frac{y^{\frac{t}{\nu}-1}e^{-\frac{y}{\nu}}}{\Gamma\left(\frac{t}{\nu}\right)\nu^{\frac{t}{\nu}}}\,dy\\ &= -\frac{1}{S(0)}\phi^{\mathrm{VG}}(x,t).\end{aligned}$$

By determining the supremum by a naïve approach, using Fubini's theorem and using the global supremum of the normal distribution will leads to the expectation $\mathbb{E}[\tau^{-1/2}]$ for the Gamma distribution, $\gamma(t/\nu,1/\nu)$. This only leads to a finite value if $\nu>t/2$ for fixed t.

For the derivation, since the two derivatives are related, we will only look at the case for the derivative w.r.t. x. Recall that $\eta(x) = x - \ln S(0) - (r+\omega)t$, see (37), such that for a fixed value of $\tau(t) = y$,

$$x - \alpha(y) = \eta(x) - \kappa y, \tag{42}$$

from which it follows that $x - \alpha(y) > 0$ is equivalent to $y < \eta(x)/\kappa$. The region in which we determine the local supremum of $\phi^{\mathrm{VG}}(x,t)$ is $B_{r_m}(q_\alpha) \subset B(\alpha)$ as defined in Lemma 4.2. For m sufficiently large, it holds that

$$q_\alpha - r_m \leq x \leq q_\alpha + r_m.$$

Subtracting $\ln(S(0)) + (r+\omega)t$ from the above row of inequalities leads to

$$\eta(q_\alpha) - r_m \leq \eta(x) \leq \eta(q_\alpha) + r_m. \tag{43}$$

Before turning to the actual proof, we define, for notational convenience,

$$\zeta_\xi(x) = e^{\frac{\kappa}{\sigma^2}\eta(x)} \left(\frac{\eta^2(x)}{\kappa^2 + \frac{2\sigma^2}{\nu}} \right)^{\frac{\xi}{2}} K_\xi \left(\frac{1}{\sigma^2} \sqrt{\eta^2(x) \left(\kappa^2 + \frac{2\sigma^2}{\nu} \right)} \right)$$

and for the derivation, given parameters $\nu, \beta > 0, \gamma > 0$, we require the integral result 3.471.9 in [15], given below

$$\int_0^\infty x^{\nu-1} e^{-\frac{\beta}{x} - \gamma x}\, dx = 2 \left(\frac{\beta}{\gamma} \right)^{\frac{\nu}{2}} K_\nu(2\sqrt{\beta\gamma}). \tag{44}$$

Now we can provide the suprema. Writing g_ν as the Lebesgue density for the gamma function, noting $x - \alpha(y) = \eta(x) - \kappa y$, see (42), it follows that

$$\begin{aligned} \left| \frac{\partial}{\partial x} \phi^{\mathrm{VG}}(x) \right| &\leq \int_0^\infty \frac{1}{\sqrt{2\pi}\beta(y)} \frac{|x - \alpha y|}{\beta^2(y)} e^{-\frac{1}{2}\frac{(x-\alpha(y))^2}{\beta^2(y)}} g_\nu(y)\, dy \\ &\leq \frac{1}{\sqrt{2\pi}\sigma^3 \Gamma\left(\frac{t}{\nu}\right) \nu^{\frac{t}{\nu}}} \int_0^\infty |\eta(x) - \kappa y| e^{-\frac{1}{2}\frac{(\eta(x)-\kappa y)^2}{\sigma^2 y}} e^{-\frac{y}{\nu}}\, dy. \end{aligned} \tag{45}$$

A crucial step for computing the expression in (45) is the computation of bounds for

$$|\eta(x) - \kappa y| \quad \text{and} \quad e^{-\frac{1}{2}\frac{(\eta(x)-\kappa y)^2}{\sigma^2 y}},$$

where we distinguish between the cases when both $\eta(q_\alpha) > 0$ and $\kappa > 0$ are positive or negative; when one of $\eta(q_\alpha)$ and κ is positive and the other negative, and $\kappa = 0$. We will also require for m to be sufficiently large such that if

$$\eta(q_\alpha) - \kappa y > 0 \quad \text{this implies} \quad \eta(q_\alpha) - r_m - \kappa y \;>\; 0,$$

and if

$$\eta(q_\alpha) - \kappa y < 0 \quad \text{this implies} \quad \eta(q_\alpha) + r_m - \kappa y \;<\; 0.$$

First, we consider the case $\eta(q_\alpha) > 0$, and $\kappa > 0$. When both parameters are negative, the upper bound is similar. Within the region $B_{r_m}(q_\alpha)$ it holds that

$$\eta(x) + \kappa y \leq \eta(q_\alpha) + r_m + \kappa y,$$

see (43). To maximize the exponential term, we need to split up the integral into the regions $\eta(q_\alpha) - \kappa y < 0$ and $\eta(q_\alpha) - \kappa y > 0$. In the first interval, $\eta(x) - \kappa y \leq \eta(q_\alpha) + r_m - \kappa y$, and, for m sufficiently large, $\eta(q_\alpha) - \kappa y < 0$ implies $\eta(q_\alpha) + r_m - \kappa y < 0$, from which it follows $(\eta(x) - \kappa y)^2 > (\eta(q_\alpha) + r_m - \kappa y)^2$.

For the case $\eta(q_\alpha) - \kappa y > 0$ we obtain by similar arguments $(\eta(x) - \kappa y)^2 > (\eta(q_\alpha) - r_m - \kappa y)^2$. We now obtain

$$\begin{aligned}
&\sup_{x \in B_{r_m}(q_\alpha)} \left| \frac{\partial}{\partial x} \phi^{\mathrm{VG}}(x) \right| \\
&\leq \frac{1}{\sqrt{2\pi}\sigma^3 \Gamma\left(\frac{t}{\nu}\right) \nu^{\frac{t}{\nu}}} \left(\int_0^{\frac{\eta(q_\alpha)}{\kappa}} (\eta(q_\alpha) + r_m + \kappa y) e^{-\frac{1}{2}\frac{(\eta(q_\alpha) - r_m - \kappa y)^2}{\sigma^2 y}} e^{-\frac{y}{\nu}}\, dy \right. \\
&+ \left. \int_{\frac{\eta(q_\alpha)}{\kappa}}^\infty (\eta(q_\alpha) + r_m + \kappa y) e^{-\frac{1}{2}\frac{(\eta(q_\alpha) + r_m - \kappa y)^2}{\sigma^2 y}} e^{-\frac{y}{\nu}}\, dy \right) \\
&\leq \frac{1}{\sqrt{2\pi}\sigma^3 \Gamma\left(\frac{t}{\nu}\right) \nu^{\frac{t}{\nu}}} \left(\eta(q_\alpha + r_m) \left(e^{\frac{\kappa}{\sigma^2}\eta(q_\alpha + r_m)} \int_0^\infty y^{\frac{t}{\nu} - \frac{5}{2}} e^{-\frac{1}{2}\frac{\eta(q_\alpha + r_m))}{\sigma^2 y}} e^{-\left(\frac{1}{2}\frac{\kappa^2}{\sigma^2} + \frac{1}{\nu}\right) y}\, dy \right. \right. \\
&+ \left. e^{\frac{\kappa}{\sigma^2}\eta(q_\alpha - r_m)} \int_0^\infty y^{\frac{t}{\nu} - \frac{5}{2}} e^{-\frac{1}{2}\frac{\eta(q_\alpha - r_m)}{\sigma^2 y}} e^{-\left(\frac{1}{2}\frac{\kappa^2}{\sigma^2} + \frac{1}{\nu}\right) y}\, dy \right) \\
&+ \left. \kappa \left(e^{\frac{\kappa}{\sigma^2}\eta(q_\alpha + r_m)} \int_0^\infty y^{\frac{t}{\nu} - \frac{3}{2}} e^{-\frac{1}{2}\frac{\eta(q_\alpha + r_m)}{\sigma^2 y}} e^{-\left(\frac{1}{2}\frac{\kappa^2}{\sigma^2} + \frac{1}{\nu}\right) y}\, dy + e^{\frac{\kappa}{\sigma^2}\eta(q_\alpha - r_m)} \int_0^\infty y^{\frac{t}{\nu} - \frac{3}{2}} e^{-\frac{1}{2}\frac{\eta(q_\alpha - r_m)}{\sigma^2 y}} e^{-\left(\frac{1}{2}\frac{\kappa^2}{\sigma^2} + \frac{1}{\nu}\right) y}\, dy \right) \right) \\
&= \frac{2}{\sqrt{2\pi}\sigma^3 \Gamma\left(\frac{t}{\nu}\right) \nu^{\frac{t}{\nu}}} \left(\eta(q_\alpha + r_m) \left(\zeta_{\frac{t}{\nu} - \frac{3}{2}}(q_\alpha + r_m) + \zeta_{\frac{t}{\nu} - \frac{3}{2}}(q_\alpha - r_m) \right) + \kappa \left(\zeta_{\frac{t}{\nu} - \frac{1}{2}}(q_\alpha + r_m) + \zeta_{\frac{t}{\nu} - \frac{1}{2}}(q_\alpha - r_m) \right) \right).
\end{aligned}$$

Next, we discuss the case $\eta(q_\alpha) < 0$ and $\kappa > 0$, i.e., the case that $\eta(q_\alpha) - \kappa y < 0$ for all $y > 0$. For m sufficiently large, $\eta(q_\alpha) - \kappa y < 0$ implies that $\eta(x) - \kappa y < 0$ for all $y > 0$, see (43). As a result $\eta(x) - \kappa y \leq \kappa y - \eta(x) < \kappa y + |\eta(q_\alpha)| + r_m$ within $B_{r_m}(q_\alpha)$. Additionally, for the exponent, $(\eta(x) - \kappa y)^2 > (\eta(q_\alpha) + r_m - \kappa y)^2$. Putting this altogether, we have the supremum

$$\begin{aligned}
\sup_{x \in B_{r_m}(q_\alpha)} \left| \frac{\partial}{\partial x} \phi^{\mathrm{VG}}(x) \right| &\leq \frac{1}{\sqrt{2\pi}\sigma^3 \Gamma\left(\frac{t}{\nu}\right) \nu^{\frac{t}{\nu}}} \int_0^\infty y^{\frac{t}{\nu} - \frac{5}{2}} (\kappa y + |\eta(q_\alpha)| + r_m) e^{-\frac{1}{2}\frac{(\eta(q_\alpha) + r_m - \kappa y)^2}{\sigma^2 y}} e^{-\frac{y}{\nu}}\, dy \\
&= \frac{e^{\frac{\kappa}{\sigma^2}\eta(q_\alpha + r_m)}}{\sqrt{2\pi}\sigma^3 \Gamma\left(\frac{t}{\nu}\right) \nu^{\frac{t}{\nu}}} \left((|\eta(q_\alpha)| + r_m) \int_0^\infty y^{\frac{t}{\nu} - \frac{5}{2}} e^{-\frac{1}{2}\frac{\eta^2(q_\alpha + r_m)}{\sigma^2}\frac{1}{y}} e^{-\left(\frac{1}{2}\frac{\kappa^2}{\sigma^2} + \frac{1}{\nu}\right) y}\, dy \right. \\
&\quad + \left. \kappa \int_0^\infty y^{\frac{t}{\nu} - \frac{3}{2}} e^{-\frac{1}{2}\frac{\eta^2(q_\alpha + r_m)}{\sigma^2}\frac{1}{y}} e^{-\left(\frac{1}{2}\frac{\kappa^2}{\sigma^2} + \frac{1}{\nu}\right) y}\, dy \right)
\end{aligned}$$

and we arrive at

$$\sup_{x \in B_{r_m}(q_\alpha)} \left| \frac{\partial}{\partial x} \phi^{VG}_{\alpha(t),\beta(t)}(x) \right| = \frac{2}{\sqrt{2\pi}\sigma^3 \Gamma\left(\frac{t}{\nu}\right) \nu^{\frac{t}{\nu}}} \left((|\eta(q_\alpha)| + r_m) \zeta_{\frac{t}{\nu} - \frac{3}{2}}(q_\alpha + r_m) + \kappa \zeta_{\frac{t}{\nu} - \frac{1}{2}}(q_\alpha + r_m) \right).$$

Finally, we discuss that case $\kappa = 0$. When $\kappa = 0$, we obtain a simpler expression. We note that within the open ball $B_{r_m}(q_\alpha)$, accounting for the sign of $\eta(q_\alpha)$, $\eta(x) \leq |\eta(q_\alpha)| + r_m$ and for the exponential term, $\eta^2(x) \geq (|\eta(q_\alpha)| - r_m)^2$. We again have made the assumption that if $\eta(q_\alpha) > 0$, m is chosen sufficiently large such that $\eta(q_\alpha) - r_m > 0$. A like assumption holds if $\eta(q_\alpha) < 0$. The supremum now has the form:

$$\sup_{x \in B_{r_m}(q_\alpha)} \left| \frac{\partial}{\partial x} \phi^{\mathrm{VG}}(x) \right| \leq \frac{|\eta(q_\alpha)| + r_m}{\sqrt{2\pi}\sigma^3 \Gamma\left(\frac{t}{\nu}\right) \nu^{\frac{t}{\nu}}} \int_0^\infty y^{\frac{t}{\nu} - \frac{5}{2}} e^{-\frac{1}{2}\frac{(|\eta(q_\alpha)| - r_m)^2}{\sigma^2}\frac{1}{y}} e^{-\frac{y}{\nu}}\, dy$$

and after some rearrangement of the constants and the use of the same integral expression

$$= \frac{2^{\frac{5}{4}}(|\eta(q_\alpha)| + r_m)}{\sigma^{\frac{3}{2}} \Gamma\left(\frac{t}{\nu}\right) (2\nu\sigma^2)^{\frac{t}{2\nu}}} (|\eta(q_\alpha)| - r_m)^{\frac{t}{\nu}} K_{\frac{t}{\nu} - \frac{3}{2}} \left(\sqrt{\frac{2}{\nu}} \frac{|\eta(q_\alpha)| - r_m}{\sigma} \right).$$

Note that while $\eta(q_\alpha) = 0$ is possible this event has probability zero. The case $\kappa = 0$, corresponds to the symmetric Variance Gamma process.

A.2 Condition (A2) for the α-quantile Vega

As for the α-quantile vega, determining the supremum of the derivative is standard. Ensuring that each term in the normal derivative is positive, we have the expression

$$\left|\frac{\partial}{\partial\sigma}\phi^{\mathrm{VG}}(x)\right| = \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma\beta(y)}\left(\frac{(x-\alpha(y))^2}{\beta^2(y)}+1\right)e^{-\frac{1}{2}\frac{(x-\alpha(y))^2}{\beta^2(y)}}g_\nu(y)\,dy + \int_0^\infty \frac{1}{\sqrt{2\pi}\beta(y)}\frac{|\eta(x)-\kappa y|}{\beta^2(y)}\sigma t e^{-\nu\omega}e^{-\frac{1}{2}\frac{(x-\alpha(y))^2}{\beta^2(y)}}g_\nu(y)\,dy$$

where $e^{-\nu\omega}=(1-\kappa\nu-\sigma^2\nu/2)^{-1}$. Though the techniques to determine a bound for the supremum are of the same type as in the previous example, the process is more involved. We will only consider here the more difficult case $\eta(q_\alpha)/\kappa>0$ where both quantities are positive. The supremum for the partial derivative w.r.t σ has the form

$$\begin{aligned}&\sup_{x\in B_{r_m}(q_\alpha)}\left|\frac{\partial}{\partial\sigma}\phi^{\mathrm{VG}}(x)\right|\\ &=\frac{2}{\sqrt{2\pi}\sigma^2\Gamma\left(\frac{t}{\nu}\right)\nu^{\frac{t}{\nu}}}\Bigg(\left(\eta(q_\alpha)+r_m\right)\left(\frac{\eta(q_\alpha+r_m)}{\sigma^2}+te^{-\nu\omega}\right)\left(\zeta_{\frac{t}{\nu}-\frac{3}{2}}(q_\alpha+r_m)+\zeta_{\frac{t}{\nu}-\frac{3}{2}}(q_\alpha-r_m)\right)\\ &+\left(\frac{2\kappa\eta(q_\alpha+r_m)}{\sigma^2}+1\right)\zeta_{\frac{t}{\nu}-\frac{1}{2}}(q_\alpha+r_m)+\zeta_{\frac{t}{\nu}-\frac{1}{2}}(q_\alpha-r_m)+\frac{\kappa^2}{\sigma^2}\left(\zeta_{\frac{t}{\nu}+\frac{1}{2}}(q_\alpha+r_m)+\zeta_{\frac{t}{\nu}+\frac{1}{2}}(q_\alpha-r_m)\right)\Bigg).\end{aligned}$$

In the following we provide some details on how to obtain the above bound. For the derivation, within the integrand, as before $|\eta(x)-\kappa y|\le\eta(q_\alpha)+r_m+\kappa y$ in the interval $B_{r_m}(q_\alpha)$. We do also need to split both integrals into the regions $\eta(q_\alpha)-\kappa y>0$ and $\eta(q_\alpha)-\kappa y<0$. For the first region, in $B_{r_m}(q_\alpha)$

$$(\eta(q_\alpha)-r_m-\kappa y)^2\le(\eta(x)-\kappa y)^2 \;\le\; (\eta(q_\alpha)+r_m-\kappa y)^2.$$

The lower bound is for the exponent and the upper bound is for the quadratic term. For $\eta(q_\alpha)-\kappa y<0$

$$(\eta(q_\alpha)+r_m-\kappa y)^2\le(\eta(x)-\kappa y)^2 \;\le\; (\eta(q_\alpha)-r_m-\kappa y)^2.$$

Inserting these bounds into the expression $|\frac{\partial}{\partial\sigma}\phi^{\mathrm{VG}}(x)|$, we get our first bound of the supremum

$$\begin{aligned}&\sup_{x\in B_{r_m}(q_\alpha)}\left|\frac{\partial}{\partial\sigma}\phi^{\mathrm{VG}}(x)\right|\\ &\le\int_0^{\frac{\eta(q_\alpha)}{\kappa}}\frac{1}{\sqrt{2\pi y}\sigma^2}e^{-\frac{1}{2}\left(\frac{\eta(q_\alpha)-r_m-\kappa y}{\sigma\sqrt{y}}\right)^2}g_\nu(y)\,dy+\int_{\frac{\eta(q_\alpha)}{\kappa}}^\infty\frac{1}{\sqrt{2\pi y}\sigma^2}e^{-\frac{1}{2}\left(\frac{\eta(q_\alpha)+r_m-\kappa y}{\sigma\sqrt{y}}\right)^2}g_\nu(y)\,dy\\ &+\int_0^{\frac{\eta(q_\alpha)}{\kappa}}\frac{1}{\sqrt{2\pi}y^{\frac{3}{2}}\sigma^4}(\eta(q_\alpha)+r_m-\kappa y)^2e^{-\frac{1}{2}\left(\frac{\eta(q_\alpha)-r_m-\kappa y}{\sigma\sqrt{y}}\right)^2}g_\nu(y)\,dy\\ &+\int_{\frac{\eta(q_\alpha)}{\kappa}}^\infty\frac{1}{\sqrt{2\pi}y^{\frac{3}{2}}\sigma^4}(\eta(q_\alpha)-r_m-\kappa y)^2e^{-\frac{1}{2}\left(\frac{\eta(q_\alpha)+r_m-\kappa y}{\sigma\sqrt{y}}\right)^2}g_\nu(y)\,dy\\ &+\sigma^2te^{-\nu\omega}(\eta(q_\alpha)+r_m)\Bigg(\int_0^{\frac{\eta(q_\alpha)}{\kappa}}\frac{1}{\sqrt{2\pi}y^{\frac{3}{2}}\sigma^4}e^{-\frac{1}{2}\left(\frac{\eta(q_\alpha)-r_m-\kappa y}{\sigma\sqrt{y}}\right)^2}g_\nu(y)\,dy\\ &\qquad+\int_{\frac{\eta(q_\alpha)}{\kappa}}^\infty\frac{1}{\sqrt{2\pi}y^{\frac{3}{2}}\sigma^4}e^{-\frac{1}{2}\left(\frac{\eta(q_\alpha)+r_m-\kappa y}{\sigma\sqrt{y}}\right)^2}g_\nu(y)\,dy\Bigg).\end{aligned}$$

For the third integral, knowing $\eta(q_\alpha)-\kappa y>0$,

$$(\eta(q_\alpha)+r_m-\kappa y)^2\le\eta^2(q_\alpha+r_m)+\kappa^2y^2$$

since $\eta(q_\alpha) + r_m > 0$, where we use the fact that $\eta(x+y) = \eta(x) + y$, see (37). For the fourth integral, where $\eta(q_\alpha) - \kappa y < 0$, $\eta(q_\alpha) - r_m$ may be a negative value and so

$$(\eta(q_\alpha) - r_m - \kappa y)^2 < (\eta(q_\alpha) + r_m + \kappa y)^2 = \eta^2(q_\alpha + r_m) + 2\kappa\eta(q_\alpha + r_m)y + \kappa^2 y^2,$$

separating the components. Extending the intervals for each of the integrals to $[0, \infty)$, and rearranging the integrals, we bound above again by

$$\begin{aligned}
&\leq \frac{1}{\sqrt{2\pi}\sigma^2\Gamma\left(\frac{t}{\nu}\right)\nu^{\frac{t}{\nu}}} \left(\eta(q_\alpha + r_m)\left(\frac{\eta(q_\alpha + r_m)}{\sigma^2} + te^{-\nu\omega}\right)\left(e^{\frac{\kappa}{\sigma^2}\eta(q_\alpha + r_m)} \int_0^\infty y^{\frac{t}{\nu} - \frac{5}{2}} e^{-\frac{1}{2}\frac{\eta^2(q_\alpha + r_m)}{\sigma^2}\frac{1}{y}} e^{-\left(\frac{1}{2}\frac{\kappa^2}{\sigma^2} + \frac{1}{\nu}\right)y}\, dy\right.\right. \\
&+ \left. e^{\frac{\kappa}{\sigma^2}\eta(q_\alpha - r_m)} \int_0^\infty y^{\frac{t}{\nu} - \frac{5}{2}} e^{-\frac{1}{2}\frac{\eta^2(q_\alpha - r_m)}{\sigma^2}\frac{1}{y}} e^{-\left(\frac{1}{2}\frac{\kappa^2}{\sigma^2} + \frac{1}{\nu}\right)y}\, dy\right) \\
&+ \left(\frac{2\kappa\eta(q_\alpha + r_m)}{\sigma^2} + 1\right) e^{\frac{\kappa}{\sigma^2}\eta(q_\alpha + r_m)} \int_0^\infty y^{\frac{t}{\nu} - \frac{3}{2}} e^{-\frac{1}{2}\frac{\eta^2(q_\alpha + r_m)}{\sigma^2}\frac{1}{y}} e^{-\left(\frac{1}{2}\frac{\kappa^2}{\sigma^2} + \frac{1}{\nu}\right)y}\, dy \\
&+ e^{\frac{\kappa}{\sigma^2}\eta(q_\alpha - r_m)} \int_0^\infty y^{\frac{t}{\nu} - \frac{3}{2}} e^{-\frac{1}{2}\frac{\eta^2(q_\alpha - r_m)}{\sigma^2}\frac{1}{y}} e^{-\left(\frac{1}{2}\frac{\kappa^2}{\sigma^2} + \frac{1}{\nu}\right)y}\, dy \\
&+ \frac{\kappa^2}{\sigma^2}\left(e^{\frac{\kappa}{\sigma^2}\eta(q_\alpha + r_m)} \int_0^\infty y^{\frac{t}{\nu} - \frac{1}{2}} e^{-\frac{1}{2}\frac{\eta^2(q_\alpha + r_m)}{\sigma^2}\frac{1}{y}} e^{-\left(\frac{1}{2}\frac{\kappa^2}{\sigma^2} + \frac{1}{\nu}\right)y}\, dy\right. \\
&+ \left.\left. e^{\frac{\kappa}{\sigma^2}\eta(q_\alpha - r_m)} \int_0^\infty y^{\frac{t}{\nu} - \frac{1}{2}} e^{-\frac{1}{2}\frac{\eta^2(q_\alpha - r_m)}{\sigma^2}\frac{1}{y}} e^{-\left(\frac{1}{2}\frac{\kappa^2}{\sigma^2} + \frac{1}{\nu}\right)y}\, dy\right)\right).
\end{aligned}$$

From Equation (44) and the definition of $\zeta_{(\cdot)}$, the final result is obtained.