Theoretical Results on the Haezendonck-Goovaerts Risk Measures: the Design and Analysis of a Computational Algorithm

Jae Youn Ahn\textsuperscript{a,}\textsuperscript{*}, Nariankadu D. Shyamalkumar\textsuperscript{b}

\textsuperscript{a}Jae Youn Ahn, PhD, Assistant Professor, Department of Statistics, Ewha Womans University, 11-1 Daehyun-Dong, Seodaemun-Gu, Seoul 120-750, Korea. jaeyahn@ewha.ac.kr

\textsuperscript{b}Nariankadu D. Shyamalkumar, ASA, PhD, Associate Professor, Dept. of Statistics and Actuarial Science, The University of Iowa, 241 Schaeffer Hall, Iowa City, IA 52242, USA. shyamal-kumar@uiowa.edu

Abstract

The Haezendonck-Goovaerts class of risk measures (H-G risk measures) corresponds to a premium principle that is a multiplicative analog of the zero utility principle, and is thus of significant academic interest. The H-G risk measures often do not have a closed form representation and hence to compute its value we need a computational algorithm. In this article we concern ourselves with the design of such an algorithm which can also give both a lower and an upper bound for the true value. Such an algorithm can hence produce an approximation to the true value within any given absolute error tolerance. A computational algorithm is also importantly needed when one conducts a simulation study to understand the properties of the empirical non-parametric estimator of the H-G risk measure. Our work on the algorithm presented here was motivated by the needs of the simulation experiments reported in Ahn and Shyamalkumar (2014). Towards arriving at such an algorithm, we also derive results relating to the H-G risk measures which contribute further to the general understanding of these risk measures.

Key words: Orlicz Premium, Tail value-at-Risk (T-VaR), Orlicz Quantile, Young Function

Classification: 62P05: 91B82: 65Y99

\textsuperscript{*}Corresponding Author
1. Introduction

A premium principle induced by an Orlicz norm was presented in Haezendonck and Goovaerts (1982) (see also Goovaerts et al. (2003)). Based on this premium principle, Goovaerts et al. (2004) introduced a class of risk measures which is now referred to as the Haezendonck-Goovaerts risk measures. The most prominent member of this class, and in fact its minimal member, is the Tail Value-at-Risk (T-VaR) - T-VaR arguably the most popular risk measure in global insurance regulation. This general class of risk measures has received much attention recently, mainly due to it arising from a well motivated premium principle. Bellini and Gianin (2008a) present an alternate formulation of these risk measures which makes them coherent in the sense of Artzner et al. (1997, 1999). For other development and properties of this class of risk measures see Goovaerts et al. (2004), Goovaerts et al. (2012), Bellini and Gianin (2008a), and Bellini and Gianin (2012). For asymptotic expansions of the Haezendonck-Goovaerts risk measures see Tang and Yang (2012), Tang and Yang (2014), and Mao and Hu (2012). Statistical estimation of Orlicz quantiles and Haezendonck-Goovaerts risk measures is considered in Bellini and Gianin (2008b) and Ahn and Shyamalkumar (2014).

The H-G risk measures do not have a closed form representation and hence to compute its value we need an algorithm. In this article we concern ourselves with the design of such an algorithm which can also give both a lower and an upper bound for the true value. An algorithm of this kind can hence produce an approximation to the true value within any given absolute error tolerance. A computational algorithm is also importantly needed when one conducts a simulation study to understand the properties of the empirical non-parametric estimator of the H-G risk measure. The lack of an algorithm with explicit absolute error bounds was clearly noticed in Bellini and Gianin (2008b), and this impeded upon the quality of their simulation results, see Bellini and Gianin (2008b). Our work on the algorithm presented here was motivated by the needs of the simulation experiments reported in Ahn and Shyamalkumar (2014). Towards arriving at such an algorithm, we also derive results relating to the H-G risk measures which contribute further to the general understanding of these risk measures.

1.1. Risk Measures in Insurance and H-G Risk measures

Towards defining these risk measures, let $F(\cdot)$ denote the distribution function of the loss random variable $X$ associated with a certain insurance/financial portfolio, and let $\alpha$ be a real number in $(0, 1)$. The $\alpha$-level VaR for the distribution $F(\cdot)$, which we denote by $q_\alpha(F)$, is defined by

$$q_\alpha(F) := \inf\{x : F(x) \geq \alpha\}.$$  

By $q_\alpha^+(F)$ we denote $\inf\{x : F(x) > \alpha\}$. Note that $q_\alpha(F) \leq q_\alpha^+(F)$ with equality if and only if $F(\cdot)$ is strictly increasing at $q_\alpha(F)$. In a sense the set of $\alpha$ level quantiles of $F(\cdot)$ equals $[q_\alpha(F), q_\alpha^+(F)]$.

The $\alpha$-level T-VaR for $F(\cdot)$, which we denote by $\text{T-VaR}_\alpha(F)$, is defined by

$$\text{T-VaR}_\alpha(F) := \frac{1}{1-\alpha} \int_\alpha^1 q_s(F) \, ds, \quad \alpha \in (0, 1).$$

T-VaR\(^1\) has gained a prominent place in the North American insurance regulatory world, very much like VaR has for the trading desks and the banking sector at large.

\(^1\)A slight variant of T-VaR is called the Conditional Tail Expectation (CTE). CTE agrees with the T-VaR in all cases except when the corresponding quantile is a mass point.
We now begin by laying out the notation for our study of the class of H-G risk measures. A non-negative, strictly increasing, convex function $\Psi(\cdot)$ on $\mathbb{R}^+$ with $\Psi(0) = 0$ and $\Psi(1) = 1$ is called a normalized Young function (see Rao and Ren (1991) for details). In the following we will work with the extensions of such functions to the whole of $\mathbb{R}$ satisfying $\Psi(x) = 0$ for $x < 0$. For convenience we will simply refer to such extensions simply as Young functions. The class of H-G risk measures is indexed by the class of Young function, and for each Haezendonck-Goovaerts risk measure there exits a class of random variables for which it is well defined. A subset of this class of random variables is denoted by $\mathfrak{X}_\Psi$, and is defined by

$$
\mathfrak{X}_\Psi := \left\{ X \middle| \Pr (X \leq 0) = 1 \text{ or } \exists s_\infty \geq 0 \text{ such that } \mathbb{E} \left[ \Psi \left( \frac{X}{s} \right) \right] < \infty \text{ for } s > 0 \iff s > s_\infty \right\}. \tag{1}
$$

We refer to $\mathfrak{X}_\Psi$ as the Orlicz space, without explicitly noting its dependence on the underlying Young function.

In Bellini and Gianin (2008a), for convenience, the random variables are restricted to $L^\infty$ (the space of essentially bounded random variables) a subset of $\mathfrak{X}_\Psi$. We allow $s_\infty$ to be greater than 0 in (1), unlike in Goovaerts et al. (2004), to accommodate situations like those arising in Example 5 of Ahn and Shyamalkumar (2014). A useful property of $\mathfrak{X}_\Psi$ is that for $X \in \mathfrak{X}_\Psi$, $c \in \mathbb{R}$ and $s > 0$ we have

$$
\mathbb{E} \left[ \Psi \left( \frac{X}{s} \right) \right] < \infty \iff \mathbb{E} \left[ \Psi \left( \frac{X - c}{s} \right) \right] < \infty. \tag{2}
$$

We will find it convenient to define $s_\infty(\cdot)$ as

$$
s_\infty(X) := \inf \left\{ s > 0 \middle| \mathbb{E} \left[ \Psi \left( \frac{X}{s} \right) \right] < \infty \right\}. 
$$

Note that (2) is equivalent to the statement that for $X \in \mathfrak{X}_\Psi$, $s_\infty(X) = s_\infty(X - c)$, for all $c \in \mathbb{R}$.

For $X \in \mathfrak{X}_\Psi$, we define the Orlicz premium for $X - x$ corresponding to $\Psi(\cdot)$ and at level $\alpha \in [0, 1)$, denoted by $H_X(x)$, as the unique positive solution of the equation

$$
\mathbb{E} \left[ \Psi \left( \frac{X - x}{H_X(x)} \right) \right] = 1 - \alpha \quad \text{for } \Pr (X - x > 0) > 0, \tag{3}
$$

with $H_X(x) := 0$ for $x$ satisfying $\Pr (X - x \leq 0) = 1$ (see Haezendonck and Goovaerts (1982), Goovaerts et al. (2004), and Bellini and Gianin (2008a)). For $X \in \mathfrak{X}_\Psi$, we define the Haezendonck-Goovaerts risk measure corresponding to $\Psi(\cdot)$ and at level $\alpha \in [0, 1)$ by

$$
\pi_X := \inf_{x \in \mathbb{R}} (H_X(x) + x), \tag{3}
$$

and for convenience we define $\pi_X(\cdot)$ as

$$
\pi_X(x) := H_X(x) + x, \quad x \in \mathbb{R}. \tag{4}
$$

For $X \in L^\infty$, Proposition 16 of Bellini and Gianin (2008a) shows that the above infimum is attained for $\alpha \in (0, 1)$; their argument easily extends to $\mathfrak{X}_\Psi$. Moreover, examples exists where this infimum is not attained when $\alpha = 0$, see Ahn and Shyamalkumar (2014). For this reason, and also since $\alpha$ values close to one are those of interest in risk management, in the following we will restrict our attention to $\alpha$ values in $(0, 1)$.
For $X \in \mathbb{X}_\Psi$, we define the Orlicz quantile as a set

$$\mathcal{I}_X := \arg \min_{x \in \mathbb{R}} \pi_X (x).$$

In Bellini and Gianin (2008a) it is shown that $\pi_X (\cdot)$ is a convex function for $X \in L^\infty$, and this result also easily extends to $X \in \mathbb{X}_\Psi$. This, in particular, implies that $\mathcal{I}_X$ is a closed interval. We find it convenient to interchangeably use $\pi_F$ and $\mathcal{I}_F$ for $\pi_X$ and $\mathcal{I}_X$, respectively; note that the definition of the Haezendonck-Goovaerts risk measure permits such use.

We denote by $\Psi^{-1} (\cdot)$, $\Psi'_{+} (\cdot)$, and $\Psi'_{-} (\cdot)$ the inverse, right side derivative, and left side derivative of $\Psi(\cdot)$, respectively. For $X \in \mathbb{X}_\Psi$, by $H'_{X}^{+} (\cdot)$ and $H'_{X}^{-} (\cdot)$ we denote the right and left side derivatives of $H_X (\cdot)$, respectively. Similarly, for $X \in \mathbb{X}_\Psi$, by $\pi'_{X}^{+} (\cdot)$ and $\pi'_{X}^{-} (\cdot)$ we denote the right and left side derivatives of $\pi_X (\cdot)$, respectively. Finally, for $x \in \mathbb{R}$, $(x)_+$ or $x_+$ equals $x$ for non-negative $x$, and equals zero for negative $x$.

In section two, we derive a probability inequality which will be key to deriving the main result in the following section. In section three, we derive bounds for $\mathcal{I}_F$ which play a key role in the design of our algorithm. In section four we present the algorithm and establish results that prove its well functioning, and give an example to show its performance. In section five, we compare its performance to the theoretical approximations derived under certain assumptions on the underlying distribution and Young function.

### 2. A Probabilistic Inequality

In this section we derive a probabilistic inequality that is potentially of broader interest in actuarial applications. This will be used to derive an upper bound for the Orlicz quantiles in Theorem 2. In the following, for $\Psi(\cdot)$ a Young function, and $B$ a Borel set in $\mathbb{R}$, we define $\Psi_B' (\cdot)$ by

$$\Psi_B' (x) = \begin{cases} \Psi'_{+} (x), & x \in B, \\ \Psi'_{-} (x), & \text{otherwise}. \end{cases}$$

**Theorem 1.** Let $X \in \mathbb{X}_\Psi$ be such that $s_\infty (X) < 1$ and $\mathbb{E} [\Psi (X)] \geq 1$. Then for any Borel subset $B$ of $\mathbb{R}$ we have

$$\mathbb{E} \left[ X \Psi_B' (X) \right] \geq \mathbb{E} \left[ \Psi_B' (X) \right].$$

Moreover, if we have $\mathbb{E} [\Psi (X)] > 1$ then we have strict inequality in (5) as well.

**Proof.** Using that $s_\infty (X) < 1$, Lemma 3 and convexity of $\Psi(\cdot)$ we have

$$\mathbb{E} \left[ X \Psi'_{+} (X) I_{\{X \in B\}} \right] - \mathbb{E} \left[ \Psi'_{+} (X) I_{\{X \in B\}} \right]
= \mathbb{E} \left[ (X - 1) \Psi'_{+} (X) I_{\{X \geq 1 \text{ and } X \in B\}} \right] - \mathbb{E} \left[ (1 - X) \Psi'_{+} (X) I_{\{X < 1 \text{ and } X \in B\}} \right]
\geq \mathbb{E} \left[ (\Psi(X) - 1) I_{\{X \geq 1 \text{ and } X \in B\}} \right] - \mathbb{E} \left[ (1 - X) \Psi'_{+} (X) I_{\{X < 1 \text{ and } X \in B\}} \right].$$

Similarly, we have

$$\mathbb{E} \left[ X \Psi'_{-} (X) I_{\{X \in B^C\}} \right] - \mathbb{E} \left[ \Psi'_{-} (X) I_{\{X \in B^C\}} \right]
\geq \mathbb{E} \left[ (\Psi(X) - 1) I_{\{X \geq 1 \text{ and } X \in B^C\}} \right] - \mathbb{E} \left[ (1 - X) \Psi'_{-} (X) I_{\{X \leq 1 \text{ and } X \in B^C\}} \right].$$

4
Now note that
\[ \mathbb{E} [\Psi(X)] \geq 1 \Leftrightarrow \mathbb{E} [(\Psi(X) - 1)I_{X \geq 1}] \geq \mathbb{E} [(1 - \Psi(X))I_{X < 1}] . \]  
(8)

Combining (6), (7), (8) and convexity of \( \Psi(\cdot) \) we have

\[
\mathbb{E} \left[ X \Psi'^+(X)I_{X \in B} + X \Psi'^-(X)I_{X \in B^c} \right] - \mathbb{E} \left[ (1 - \Psi(X))I_{X < 1} \right] \\
- \mathbb{E} \left[ (1 - X)\Psi'^+(X)I_{X < 1} \text{ and } X \in B \right] + \mathbb{E} \left[ (1 - X)\Psi'^-(X)I_{X < 1} \text{ and } X \in B^c \right] \\
\geq \mathbb{E} \left[ (1 - X)\Psi'^+(X)I_{X < 1} \text{ and } X \in B \right] + \mathbb{E} \left[ (1 - X)\Psi'^-(X)I_{X < 1} \text{ and } X \in B^c \right] \\
- \mathbb{E} \left[ (1 - X)\Psi'^+(X)I_{X < 1} \text{ and } X \in B \right] + \mathbb{E} \left[ (1 - X)\Psi'^-(X)I_{X < 1} \text{ and } X \in B^c \right] = 0.
\]

In the case that \( \mathbb{E} [\Psi(X)] > 1 \) we have strict inequalities in (8); this results in,

\[
\mathbb{E} \left[ X \Psi'^+(X)I_{X \in B} + X \Psi'^-(X)I_{X \in B^c} \right] > \mathbb{E} \left[ (1 - \Psi(X))I_{X < 1} \right] \\
- \mathbb{E} \left[ (1 - X)\Psi'^+(X)I_{X < 1} \text{ and } X \in B \right] + \mathbb{E} \left[ (1 - X)\Psi'^-(X)I_{X < 1} \text{ and } X \in B^c \right].
\]

Hence the proof.

\[ \square \]

3. Bounds for Orlicz Quantiles

By the definition of H-G risk measures (see (3)), ess sup(\( X \)) is a natural upper bound of \( \mathcal{I}_X \). However, this bound is of little use when ess sup(\( X \)) = \( \infty \), which is the common case in insurance problems. In this section, we show that \( q^\alpha_x \) is an upper bound for \( \mathcal{I}_X \), and provide a lower bound for \( \mathcal{I}_X \). Before we study such bounds, we need to establish the following basic properties of the Orlicz premium \( H_X(\cdot) \).

**Lemma 1.** For \( X \in \mathcal{X}_q \) and \( \alpha \in (0, 1) \), we have the following:

i. \( \pi_X(\cdot) \) and \( H_X(\cdot) \) are convex functions.

ii. \( \lim_{x \to -\infty} \frac{x}{H_X(x)} = -\Psi^{-1}(1 - \alpha) \)

iii. For \( x \in \mathbb{R} \), we have

\[
1 - \frac{1}{\Psi^{-1}(1 - \alpha)} \leq \pi_X^\alpha(x) \leq 1.
\]

Furthermore, we have

\[
\lim_{x \to -\infty} \pi_X^\alpha(x) = 1 - \frac{1}{\Psi^{-1}(1 - \alpha)} < 0, \quad \text{and} \quad \lim_{x \to \infty} \pi_X^\alpha(x) = 1.
\]

iv. There exists \( x \in \mathbb{R} \) such that \( \pi_X(x) = \pi_X^\alpha \).

**Proof.** Part i is easy to show, and the result in the case of \( X \in L^\infty \) has been observed in Bellini and Gianin (2008a). For part ii, note that for \( x < a < b \leq \text{ess sup}(X) \) we have from part i that

\[
\frac{H_X(a) - H_X(x)}{a - x} \leq \frac{H_X(b) - H_X(a)}{b - a} < 0,
\]

(9)

5
which implies that
\[
\limsup_{x \to -\infty} \frac{-x}{H_X(x)} \leq \frac{b - a}{H_X(a) - H_X(b)}.
\]
The above along with the definition of \(H_X(\cdot)\) and the dominated convergence theorem (DCT) implies part ii. Towards proving part iii we note that for \(x < a < b \leq \text{ess sup}(X)\), similar to (9), we have
\[
H_X^- (x) \leq \frac{H_X(a) - H_X(x)}{a - x} \leq H_X^+ (b) .
\] (10)
Inequality (10) and part ii now implies part iii. Now part iv follows easily from part iii. □

By part i. of Lemma 1, for the derivation of the bounds for \(\mathcal{I}_X\), it suffices to analyze the derivatives of \(H_X(\cdot)\). In the following lemma we show that the reciprocal of \(H_X^+(x)\) is the expectation of \((X - x)/H_X(x)\), but with respect to a weighted distribution with weights depending on either the right or left derivative of \(\Psi(\cdot)\) - the choice of the derivative being implicitly determined.

**Lemma 2.** For \(X \in \mathbb{X}_\Psi\), \(\alpha \in (0, 1)\) and \(x < \text{ess sup}(X)\), we have
\[
H_X^+(x) = -\frac{\mathbb{E} \left[ \Psi_B \left( \frac{X - x}{H_X(x)} \right) \right]}{\mathbb{E} \left[ \left( \frac{X - x}{H_X(x)} \right) \Psi_B \left( \frac{X - x}{H_X(x)} \right) \right]},
\] (11)
where \(B = (-1/H_X^+(x), \infty)\).

**Proof.** For the proof, we find it convenient to define for \(X \in \mathbb{X}_\Psi\) a function \(\beta(\cdot)\) on \((-\infty, \text{ess sup}(X))\) given by
\[
\beta(x) := x - \frac{H_X(x)}{H_X^+(x)}.
\] (12)
Defining \(A_x(y) := \frac{y - x}{H_X(x)}\,\), and using convexity of \(\Psi(\cdot)\) we have
\[
\lim_{\epsilon \to 0^+} \frac{\Psi(A_{x+\epsilon}(y)) - \Psi(A_x(y))}{\epsilon} = \begin{cases} 
-\Psi^-(A_x(y)) \left(1 + A_x(y)H_X^+(x)\right), & y < \beta(x), \\
0, & y = \beta(x), \\
-\Psi^+(A_x(y)) \left(1 + A_x(y)H_X^+(x)\right), & y > \beta(x).
\end{cases}
\] (13)
Then we have using convexity of \(\Psi(\cdot)\), Lemma 3, and the dominated convergence theorem that
\[
\begin{align*}
0 &= \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \left( \mathbb{E} \left[ \Psi \left( \frac{X - (x + \epsilon)}{H_X(x + \epsilon)} \right) \right] \right) - \mathbb{E} \left[ \Psi \left( \frac{X - x}{H_X(x)} \right) \right] \\
&= \mathbb{E} \left[ \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \left( \Psi(A_{x+\epsilon}(X)) - \Psi(A_x(X)) \right) \right] \\
&= \frac{-\mathbb{E} \left[ (1 + A_x(X)H_X^+(x)) \left(\Psi^+(A_x(X))I_{[X > \beta(x)]} + \Psi^-(A_x(X))I_{[X \leq \beta(x)]}\right) \right]}{H_X(x)}.
\end{align*}
\]
Rearranging the last expression completes the proof. □
While the left and right derivatives of \( H_X(\cdot) \) are the same when \( \Psi(\cdot) \) is differentiable, the rather complicated expression in (11) is essential for general Young functions as the following example shows.

**Example 1.** Let \( F(\cdot) \) and \( \Psi(\cdot) \) be given by

\[
F(x) = \begin{cases} 
0, & x < 0, \\
\frac{26}{30}, & 0 \leq x < 1, \\
\frac{29}{30}, & 1 \leq x < 2, \\
1, & x \geq 2,
\end{cases} \quad \text{and} \quad \Psi(x) = \begin{cases} 
0, & x < 0, \\
2x - 1, & 0 \leq x < 1, \\
3x - 3, & x \geq 2.
\end{cases}
\]

For \( \alpha = 4/5 \) it is easily checked that \( H_X^{\ell} (0) = -2/3 \) from the fact that,

\[
H_X(x) = -\frac{2x}{3} + 1, \quad \text{for } x \in [0,1).
\]

One can confirm that in this example we have

\[
\frac{1}{H_X^{\ell} (0)} = -\frac{2 \cdot 1/30 \Psi^{\ell}(2) + 1 \cdot 3/30 \Psi^{s}(1) + 0 \cdot 26/30 \Psi^{r}(0)}{\frac{26}{30} \Psi^{s}(1) + 3/30 \Psi^{r}(1) + 26/30 \Psi^{r}(0)},
\]

which agrees with (11).

The following theorem provides tight bounds for \( I_X \).

**Theorem 2.** For \( X \in \mathbb{X}_q \) and \( \alpha \in (0,1) \), we have the following:

i. \( \sup \{ I_X \} \leq q_\alpha^+(F) \) and \( \inf \{ I_X \} \leq q_\alpha(F) \).

ii. For \( \forall x \in \mathbb{R} \), we have

\[
E \left[ X \right] - \Psi^{-1}(1 - \alpha) \left( \frac{\pi_X(x) - E [X]}{1 - \Psi^{-1}(1 - \alpha)} \right) \leq \inf \{ I_X \}
\]

(14)

**Proof.** Since by definition \( \pi_X(x) = x \) for \( x \geq \text{ess sup}(X) \), it follows that

\[
\sup \{ I_X \} \leq \text{ess sup}(X).
\]

Hence in establishing the first (resp., second) inequality of part i we assume, without loss of generality, that \( q_\alpha^+(F) < \text{ess sup}(X) \) (resp., \( q_\alpha(F) < \text{ess sup}(X) \)). Since \( q_\alpha^+(F) < \text{ess sup}(X) \), we can choose \( \delta > 0 \) such that \( q_\alpha^+(F) + \delta < \text{ess sup}(X) \). Then for \( x \in [q_\alpha(F), q_\alpha^+(F) + \delta] \) we have,

\[
E \left[ \Psi \left( \frac{X - x}{H_X(x)} \right) \middle| X > x \right] = \frac{E \left[ \Psi \left( \frac{X - x}{H_X(x)} \right) \right]}{\Pr (X > x)} \geq \frac{1 - \alpha}{\Pr (X > x)} \geq 1,
\]

(15)

with strict inequality in the case that \( q_\alpha^+(F) < x \leq q_\alpha^+(F) + \delta \). Now (15), along with Theorem 1 and
Lemma 2, implies that for \( x \in [q_\alpha(F), q_\alpha^+(F) + \delta] \),
\[
\frac{\partial}{\partial x} \pi_X(x) = 1 + H_X^\alpha(x) = 1 - \frac{E \left[ X \cdot \Psi^+ \left( \frac{X - x}{H_X(x)} \right) I_{[X > \beta(x)]} \right] - E \left[ \Psi^+ \left( \frac{X - x}{H_X(x)} \right) I_{[X \leq \beta(x)]} \right]}{E \left[ \Psi^+ \left( \frac{X - x}{H_X(x)} \right) I_{[X > \beta(x)]} \right] + E \left[ \Psi^- \left( \frac{X - x}{H_X(x)} \right) I_{[X \leq \beta(x)]} \right]},
\]
\[
\geq 0,
\]
again with strict inequality in the case that \( q_\alpha^+(F) < x \leq q_\alpha^+(F) + \delta \). Since \( \pi_X(\cdot) \) is convex, this completes the proof of part i. Towards proving part ii we note that for \( x < \text{ess sup}(X) \), using Jensen’s inequality, we have
\[
E[X] - \Psi^{-1}(1-\alpha) \left( \frac{\pi_X(x) - E[X]}{1 - \Psi^{-1}(1-\alpha)} \right) \leq x, \quad \forall x \geq \text{ess sup}(X), \tag{17}
\]
This combined with the fact that \( \Psi(\cdot) \) is non-decreasing, and that
\[
E[X] \leq \text{ess sup}(X) \leq \pi_X(x), \quad \forall x \geq \text{ess sup}(X),
\]
implies that
\[
\pi_X(x) = \pi_X(x^*), \quad \forall x^* \in \mathcal{I}_X,
\]
it follows from (17) that for all \( x \in \mathbb{R} \) and \( x^* \in \mathcal{I}_X \) we have
\[
E[X] - \Psi^{-1}(1-\alpha) \left( \frac{\pi_X(x) - E[X]}{1 - \Psi^{-1}(1-\alpha)} \right) \leq x^*.
\]
Hence the proof. \( \square \)

**Remark 1.** In the case of the T-VaR, i.e. when \( \Psi(x) = (x)_+ \), it is known that \( \mathcal{I}_X = [q_\alpha(F), q_\alpha^+(F)] \), see Rockafellar and Uryasev (2002). Hence, infimum of Orlicz quantile can be interpreted as a generalized quantile, and the upper bound in Theorem 2 is hence a tight upper bound.

**Remark 2.** The best possible lower bound provided by Theorem 2 is clearly
\[
\mathcal{L}_X := E[X] - \Psi^{-1}(1-\alpha) \left( \frac{\pi_X - E[X]}{1 - \Psi^{-1}(1-\alpha)} \right).
\]
In the case of T-VaR, generally the following equality holds:
\[
\mathcal{L}_X = q_\alpha^+(F) - \frac{1}{\alpha} E \left[ (q_\alpha^+(F) - X)_+ \right].
\]
Continuing with the case of T-VaR, we see that for a distribution function \( F(\cdot) \) with
\[
0 = F(q_\alpha(F)) - F(q_\alpha(F)) = \alpha,
\]
we have
\[
\mathcal{L}_X = q_\alpha(F).
\]
Hence, the lower bound is tight as well.

While \( \sup \{ \mathcal{I}_X \} \leq \text{ess sup}(X) \), we end this section by the following example which demonstrates that \( \inf \{ \mathcal{I}_X \} \) can be less than \( \text{ess inf}(X) \).

**Example 2.** Let \( F(\cdot) \) be a Bernoulli distribution with \( q := F(0) > \alpha \), and \( p := 1 - q \). Since \( q_\alpha(F) = 0 \), it follows from Theorem 2 that \( \sup \{ \mathcal{I}_X \} \leq 0 \). This is also easily confirmed directly by observing that
\[
H_F(x) = \frac{1 - x}{\Psi^{-1}(\frac{1-\alpha}{p})}, \quad 0 \leq x \leq 1,
\]
which in turn implies that
\[
\partial_z \pi_F(z) \bigg|_{z=x} = 1 - \frac{1}{\Psi^{-1}(\frac{1-\alpha}{p})} > 0, \quad 0 \leq x < 1.
\]

The choice \( \Psi(x) = x_+ \) leads to, as shown in Figure 1 (a) (for \( p = 2.5\% \) and \( \alpha = 95\% \)), \( \inf \{ \mathcal{I}_X \} = 0 \); hence (19) does not imply \( \sup \{ \mathcal{I}_X \} < \text{ess inf}(F) \). We note that if we instead require \( \Psi(\cdot) \) to be differentiable (which implies that \( \Psi'(0) = 0 \)) then
\[
\partial \pi_F(z) \bigg|_{z=0} = 1 - \frac{1}{\Psi^{-1}(\frac{1-\alpha}{p})} > 0,
\]
implies that \( \sup \{ \mathcal{I}_X \} < \text{ess inf}(F) = 0 \). An example of such an Young function is \( \Psi(x) = (x_+)^2 \); Figure 1 (b) provides the graph of \( \pi_F(\cdot) \) for \( p = 2.5\% \) and \( \alpha = 95\% \).

4. The Algorithm

For the design of an algorithm to compute the H-G risk measures we need the following results; toward presenting these we begin by adding to our notation for given \( X \in \mathcal{X}_\Psi \). By \( \Lambda(\cdot, \cdot) \) we denote the function on \( \mathbb{R} \times \mathbb{R}_+^+ \) defined by
\[
\Lambda(x, s) := \mathbb{E} \left[ \Psi \left( \frac{X - x}{s} \right) \right] - (1 - \alpha)
\]
Also, let \( \Lambda^1(\cdot, \cdot) \) and \( \Lambda^2(\cdot, \cdot) \) be defined by
\[
\Lambda^1(x, s) := \frac{\partial}{\partial z} \mathbb{E} \left[ \Psi \left( \frac{X - z}{s} \right) \right] \bigg|_{z=x} \quad \text{and} \quad \Lambda^2(x, s) := \frac{\partial}{\partial z} \mathbb{E} \left[ \Psi \left( \frac{X - x}{z} \right) \right] \bigg|_{z=s}.
\]
We summarize the properties of \( \Lambda(\cdot, \cdot) \) in Lemma 4 of the Appendix. Note that for \( x < \text{ess sup}(X) \), Lemma 4 implies that \( H_X(x) \) is the unique non-negative solution of \( \Lambda(x, \cdot) = 0 \). For \( x < \text{ess sup}(X) \) we
define functions $\mathcal{N}_x(\cdot)$ and $\mathcal{N}_x^n(\cdot)$ on $(s_\infty(X), \infty)$ defined by

$$\mathcal{N}_x(h) := h - \frac{\Lambda(x, h)}{\Lambda^2 - (x, h)},$$  \hspace{1cm} (20)

and

$$\mathcal{N}_x^n(h) := \mathcal{N}_x(\mathcal{N}_x^{n-1}(h)) \quad \text{for} \quad n \in \mathbb{N}, \quad \text{where} \quad \mathcal{N}_x^0(h) := h,$$

respectively. By convention, we define $\mathcal{N}_x(0) := 0$. The following proposition provides the properties of $\mathcal{N}_x(\cdot)$ which are needed for our algorithm. Note that $\mathcal{N}_x(\cdot)$ defines the Newton-Raphson steps towards computing $H_X(x)$.

**Proposition 1.** For $X \in \mathbb{X}_\Psi$, $\alpha \in (0, 1)$ and $x < \text{ess. sup}(X)$. Then we have the following:

i. $H_X(x) \geq \mathbb{E}[(X - x)_+] / \Psi^{-1}(1 - \alpha) > 0$.

ii. $h < \mathcal{N}_x(h) < \mathcal{N}_x(h') \leq H_X(x)$, for $s_\infty(X) < h < h' \leq H_X(x)$.

iii. $\mathcal{N}_x(\cdot)$ is left continuous on $(s_\infty(X), \infty)$.

iv. $\lim_{n \to \infty} \mathcal{N}_x^n(h') = H_X(x)$, for $s_\infty(X) < h' \leq H_X(x)$.

**Proof.** Part i follows easily from Jensen’s inequality. In the following $x, h$ and $h'$ denote real numbers satisfying $x < \text{ess. sup}(X)$ and $s_\infty(X) < h < h' \leq H_X(x)$. Towards establishing part ii we note that parts a and c of Lemma 4 imply that

$$\Lambda(x, h) > 0, \ \Lambda^2 - (x, h) < 0 \quad \text{and} \quad \frac{\Lambda(x, h') - \Lambda(x, h)}{h' - h} \geq \Lambda^2 - (x, h).$$
These imply that
\[ h < \mathcal{N}_x(h) = h' - (h' - h) - \frac{\Lambda(x, h)}{\Lambda^2(x, h)} \]
\[ \leq h' - \frac{\Lambda(x, h')}{\Lambda^2(x, h')} \]
\[ < h' - \frac{\Lambda(x, h')}{\Lambda^2(x, h')} = \mathcal{N}_x(h'). \]

Part ii now follows by observing that \( H_X(x) \) is a fixed point of \( \mathcal{N}_x(\cdot) \). Part iii follows from parts a and c of Lemma 4. For the proof of part iv we note that part ii implies that \( \mathcal{N}_x^n(h') \) is a monotone non-decreasing sequence bounded above by \( H_X(x) \), and by part iii \( \lim_{n \to \infty} \mathcal{N}_x^n(h) \) is a fixed point of \( \mathcal{N}_x(\cdot) \). The proof is now completed by observing that Lemma 4 implies that \( H_X(x) \) is the unique fixed point of \( \mathcal{N}_x(\cdot) \).

The following corollary provides a method that we use to compare \( \pi_X(x) \) and \( \pi_X(y) \) for any \( x, y \in \mathbb{R} \) with \( \pi_X(x) \neq \pi_X(y) \). We skip its proof as it easily follows from the above proposition.

**Corollary 1.** Let \( X \in \mathbb{R}_\Psi \), \( \alpha \in (0, 1) \), \( x < \text{ess} \sup(X) \) and \( s_\infty(X) < h \leq H_X(x) \). If we assume \( \pi_X(x) > \pi_X(y) \), then there exists \( N \in \mathbb{N} \) such that
\[ x + \mathcal{N}_x^n(h) > \pi_X(y), \] (21)
for any \( n \geq N \). Furthermore (21) implies \( \pi_X(x) > \pi_X(y) \).

Our algorithm is a grid search algorithm starting from the right end point and proceeding left in unit steps. In the following proposition we provide a method for constructing a grid satisfying certain properties that are key for the well functioning of our algorithm.

**Proposition 2.** Let \( X \in \mathbb{R}_\Psi \) and \( \alpha \in (0, 1) \) be such that \( q_\alpha^+(F) < \text{ess} \sup(X) \). Then for any given \( \epsilon > 0 \), in finitely many calculation, one can construct an explicitly defined grid \( x_1 < \ldots < x_J \) (with \( J > 3 \)) satisfying the following:

i. \( \mathcal{I}_X \subseteq (x_1, x_J) \), and \( x_J < \text{ess} \sup(X) \).

ii. \( \max_{2 \leq j \leq J} |x_j - x_{j-1}| \leq \epsilon \).

iii. \( \{\pi_X(x_j)\}_{1 \leq j \leq J} \) is an U-shaped sequence.

iv. \( \min_{1 \leq j \leq J} \pi_X(x_j) - \epsilon \leq \pi_X \leq \min_{1 \leq j \leq J} \pi_X(x_j) \).

**Proof.** We choose \( \epsilon' \) in \((0, \epsilon)\) such that \( q_\alpha^+(F) + \epsilon' < \text{ess} \sup(X) \). In the case that
\[ \mathbb{E} \left[ (X - (q_\alpha^+(F) + \epsilon'))_+ \right] > s_\infty(X)\Psi^{-1}(1 - \alpha), \]
we define
\[ h := \frac{\mathbb{E} \left[ (X - (q_\alpha^+(F) + \epsilon'))_+ \right]}{\Psi^{-1}(1 - \alpha)}; \]

11
else we choose a \( h \) in \((s_\infty(X), H_X (q_{a\alpha}^+(F) + \epsilon'))\) in finitely many steps by employing a trial and error method. Since \( h \in (s_\infty(X), H_X (q_{a\alpha}^+(F))) \), by Proposition 1

\[
\lim_{m \to \infty} \mathcal{N}^m_{q_{a\alpha}} (F)^{+\epsilon'} (h) = H_X (q_{a\alpha}^+(F) + \epsilon'),
\]

which in turn, along with part \( i \) of Theorem 2, implies that for some integer \( m > 0 \)

\[
E \left[ \Psi \left( \frac{X - q_{a\alpha}^+(F)}{\hat{\pi} - q_{a\alpha}^+(F)} \right) \right] < 1 - \alpha,
\]

(22)

where \( \hat{\pi} := \mathcal{N}_X (h) + q_{a\alpha}^+(F) + \epsilon' \). Since (22) implies that \( \pi_X (q_{a\alpha}^+(F)) < \hat{\pi} \), by Theorem 2. ii, we have

\[
\hat{L} := E [X] - \Psi^{-1} (1 - \alpha) \left( \frac{\hat{\pi} - E [X]}{1 - \Psi^{-1} (1 - \alpha)} \right) \leq \inf \{ I_X \}
\]

Now let \( J \) be an integer such that \( q_{a\alpha}^+(F) - \hat{L} \leq (J - 3) \epsilon \). Then the grid \( x_1 < x_2 < \ldots < x_J \) defined below satisfies parts \( i \) and \( ii \):

\[
x_j := \hat{L} + \frac{q_{a\alpha}^+(F) - \hat{L}}{J - 3} (j - 2), \quad 1 \leq j \leq J - 1; \quad x_J := q_{a\alpha}^+(F) + \epsilon'.
\]

Part iii follows easily from Theorem 2 and Lemma 1. Finally we can easily prove that any grid \( \{ x_1, \ldots, x_J \} \) satisfies part \( i \), \( ii \) and \( iii \) also satisfies part \( iv \) by Lemma 1.

Validity of our algorithm described in this section primarily derives from three assumptions that are usually satisfied in practice. First, we assume that \( s_\infty(X) = 0 \). This mild assumption\(^2\) has been used before in the literature (see Goovaerts et al. (2004)). Second, we assume that \( q_{a\alpha}^+(F) < \operatorname{ess. sup}(X) \) which is clearly satisfied when \( \operatorname{ess. sup}(X) = \infty \) - a common situation in practice. Also, in the case that \( \operatorname{ess. sup}(X) = \infty \), for sufficiently large sample size \( n \), and for \( F_n \) the empirical distribution function from a random sample of size \( n \), \( q_{a\alpha}^+(F_n) < \operatorname{ess. sup}(F_n) \). This assumption implies that for any \( \epsilon > 0 \) we can by Proposition 2 construct a grid \( x_1 < x_2 < \ldots < x_{J-1} < x_J \), for some \( J > 3 \), satisfying properties listed therein. We assume that this grid further satisfies

\[
\pi_X (x_{j-1}) \neq \pi_X (x_j), \quad \text{for} \quad j = 2, \ldots, J.
\]

(23)

In the case when \( I_X \) is a singleton, the set of such grids that do not satisfy (23) is of Lebesgue measure zero.

We summarize the first two assumptions for the algorithm below.

**Assumption SA:**

SA1. \( s_\infty(X) = 0 \).

SA2. \( q_{a\alpha}^+(F) < \operatorname{ess. sup}(X) \).

Finally, by \( G_\epsilon := \{ x_1, \ldots, x_J \} \) we denote the grid in Proposition 2 that satisfies (23). By standard grid we refer to grids such as \( G_\epsilon \).

---

\(^2\)Example 5 in Ahn and Shyamalkumar (2014) is an example where this assumption fails to hold.
4.1. Description of the Algorithm

Now we explain the algorithm which works under Assumption SA and under the assumption of the existence of a standard grid, \( G_\varepsilon \). Flowchart of the algorithm is given in Figure 2. We first define the initial value, \( h_0(\cdot) \), of Newton-Raphson steps at each grid point as

\[
h_0(x_j) = \frac{\mathcal{E} \left[ (X - x_j)_+ \right]}{\Psi^{-1}(1 - \alpha)}, \quad \text{for} \quad x_j \in G_\varepsilon. \tag{24}
\]

Assumption SA and Proposition 1 imply that

\[
0 < h_0(x_j) \leq H_X(x_j), \quad \text{for any} \quad x_j \in G_\varepsilon. \tag{25}
\]

Now we explain how the algorithm determines whether

\[
\pi_X(x_j) > \pi_X(x_{j+1}),
\]

or

\[
\pi_X(x_j) < \pi_X(x_{j+1}).
\]

Note that we have

\[
\pi_X(x_j) \neq \pi_X(x_{j+1}) \quad \text{for any} \quad x_j, x_{j+1} \in G_\varepsilon,
\]

by the assumption on the grid. First consider the case, \( \pi_X(x_j) > \pi_X(x_{j+1}) \); inequalities (25) and Corollary 1 guarantee the existence of \( N \in \mathbb{N}_0 \) such that

\[
x_j + N_{x_j}^n (h_0(x_j)) > \pi_X(x_{j+1}), \tag{26}
\]

for all \( n \geq N \). We note that inequality (26) is equivalent to two simultaneous inequalities

\[
\mathbb{E} \left[ \psi \left( \frac{X - x_{j+1}}{N_{x_j}^n (h_0(x_j)) + (x_j - x_{j+1})} \right) \right] < 1 - \alpha \quad \text{and} \quad N_{x_j}^n (h_0(x_j)) + (x_j - x_{j+1}) > 0, \tag{27}
\]

which can be determined in Step C1 of the algorithm in Figure 2. In the case of

\[
\pi_X(x_j) < \pi_X(x_{j+1}),
\]

we can use procedure similar to the one described above using instead \( N_{x_{j+1}}^n (h_0(x_{j+1})) \) for \( n \in \mathbb{N}_0 \). By repeating these procedures on the grid, \( G_\varepsilon \), the algorithm can find an \( x_j \in \{x_2, \cdots, x_{J-1}\} \) such that

\[
\pi_X(x_{j-1}) > \pi_X(x_j) \quad \text{and} \quad \pi_X(x_j) < \pi_X(x_{j+1}). \tag{28}
\]

Hence \( \pi_X(x_j) \) is the minimum of \( \{\pi_X(x_1), \cdots, \pi_X(x_J)\} \). Furthermore (28) and part i of Lemma 1 imply that \( I_X \subseteq [x_{j-1}, x_{j+1}] \).

Finally, we explain how the algorithm constructs an approximation for \( \pi_X \). In the algorithm, (28) is identified through the inequalities

\[
x_{j-1} + N_{x_{j-1}}^{n_{j-1}} (h_0(x_{j-1})) > \pi_X(x_j) \quad \text{and} \quad \pi_X(x_j) < x_{j+1} + N_{x_{j+1}}^{n_{j+1}} (h_0(x_{j+1})), \tag{29}
\]

for some \( n_{j-1}, n_{j+1} \in \mathbb{N} \). Using Newton-Raphson steps in (29), the algorithm constructs the following
approximation for \( \pi_X \):

\[
\hat{\pi}_X := x_{j+1} + N_{n_{j+1}}^{x_{j+1}}(h_0(x_{j+1})) - \frac{x_{j+1} - x_{j-1}}{2}.
\]  (30)

Finally, (29), Lemma 1 and convexity of \( \pi_X(\cdot) \) imply that output of the algorithm, \( \hat{\pi}_X \), satisfies

\[
\hat{\pi}_X \in [\pi_X - \epsilon, \pi_X + \epsilon].
\]  (31)

The following example demonstrates the performance of the algorithm.

**Example 3.** Let \( F(\cdot) \) denote the exponential distribution with mean 1 and let \( F_n(\cdot) \) denote the empirical distribution function corresponding to a random sample of size \( n \). Also, let \( \alpha = 0.95 \) and \( \Psi(\cdot) \) be defined by

\[
\Psi(x) = \begin{cases} 
0, & x < 0, \\
\frac{x^2 + x}{2}, & \text{otherwise}.
\end{cases}
\]

Since \( \Psi(\cdot) \) is a quadratic polynomial on the positive reals, we can easily calculate exact values of the Orlicz quantile and the Haezendonck-Goovaerts risk; these are given below:

\[
I_F = \left\{ -\log \left( 0.05 \left( \sqrt{153} - 11 \right) \right) \right\}, \quad \text{where} \quad -\log \left( \left( 0.05 \sqrt{153} - 11 \right) \right) \approx 2.681420
\]  (32)

and

\[
\pi_F = -\log \left( 0.05 \left( \sqrt{153} - 11 \right) \right) + \frac{\sqrt{153} - 11}{5 - \sqrt{17}} \approx 4.242973.
\]  (33)

First, we use the algorithm with a grid containing 1,000 points (corresponding to an error tolerance of about 0.002) which results in lower and upper bounds for \( I_F \) given by \( x_{852} \approx 2.6794 \) and \( x_{854} \approx 2.6838 \), respectively. Furthermore, the algorithm also concludes that \( \pi_F \) belongs to the interval \([4.238642, 4.242975]\), along with an approximate value for \( \pi_F \) of 4.2408. The exact values in (32) and (33) confirm the validity of these conclusions. Figure 3 plots the number of Newton-Raphson iterations employed by the algorithm to compute \( \pi_F(x_j) \), for \( j = 853, \ldots, 1000 \). From this figure it is seen that most of the grid points require a single iteration with a few grid points at both the ends being the exceptions.

Second, we employ the algorithm for \( F_n(\cdot) \), with sample size \( n = 1,000 \) with 1,000 grid points, and this is repeated for 1,000 runs. The number of Newton-Raphson iterations for each of these 1,000 runs is similar to that shown in Figure 3, i.e. most grid points require a single iteration with a few of the grid points at both ends requiring more iterations. We note that only grid points \( x_{999} \) and \( x_{1000} \) require more than two Newton-Raphson iterations. Also, the number of iterations at \( x_{999} \) and \( x_{1000} \) are always the same. The mean number of iterations at a grid point is 1.024987, and the average number of grid points visited by the algorithm is approximately 147. We report a summary of the number of iterations in Tables 1 and 2. \( \Box \)

In the above example we observe that for most values of \( j \), except for a few grid points at both ends, a single Newton-Raphson step is sufficient for the comparison of \( \pi_{F_n}(x_{j-1}) \) and \( \pi_{F_n}(x_j) \). While such a good performance of the algorithm is partly because the comparison of \( \pi_{F_n}(x_{j-1}) \) and \( \pi_{F_n}(x_j) \) does not require exact calculation of these values, we also emphasize that part of the reason is also that the Newton-Raphson steps generally converge fast.
<table>
<thead>
<tr>
<th># of Grid Points</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td># of Samples</td>
<td>408</td>
<td>0</td>
<td>359</td>
<td>0</td>
<td>231</td>
<td>0</td>
<td>2</td>
<td>1,000</td>
</tr>
</tbody>
</table>

Table 1: Distribution of No. of Grid Points Excluding \( x_{999} \) and \( x_{1000} \) that Require at Least 2 Newton-Raphson Steps

<table>
<thead>
<tr>
<th># of Newton-Raphson Steps</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td># of Samples</td>
<td>0</td>
<td>989</td>
<td>11</td>
<td>0</td>
<td>1,000</td>
</tr>
</tbody>
</table>

Table 2: Distribution of Newton-Raphson Steps at \( x_{1000} \)

5. Comparisons with Analytical Approximations of the H-G Risk Measures

Recently, there has been some interest in deriving analytical approximations for H-G risk measures as the parameter \( \alpha \) tends to 1, see Tang and Yang (2012), Tang and Yang (2014) and Mao and Hu (2012). While analytical approximations are theoretically very insightful unlike computational algorithms, as far as their use for computing approximations of the H-G risk measure is concerned, they fall short in two ways in comparison with our algorithm. First, there are no error bounds unlike our algorithm, and the guarantee is only that they are asymptotically equivalent. Second, these approximations come at the expense of assumptions on both the Young function and the distribution of the loss variable.

In the following example, we compare the approximations of the H-G risk measure using results of Tang and Yang (2012, 2014), with those derived using our algorithm. While even though the ratio of the exact values and asymptotic values of H-G risk measures goes to 1, there is no guarantee on the absolute difference as seen in Figure 4 (a). On the other hand, as claimed in (31), approximate value of the H-G risk measure can be computed using the algorithm for any given absolute error tolerance.

**Example 4.** For \( \gamma > 2, \theta > 0 \), consider a Pareto distribution \( F \) with

\[
F(x) = \begin{cases} 
0; & x \leq 0, \\
1 - \left( \frac{\theta}{x + \theta} \right)^\gamma; & x > 0,
\end{cases}
\]

and the Young function \( \psi \) defined as

\[
\psi(x) = \begin{cases} 
0, & x < 0, \\
x^2, & x \geq 0.
\end{cases}
\]

Tang and Yang (2012) provides an approximation of \( \pi_X \) as

\[
\pi_X := \frac{\gamma(\gamma - 2)^{2/\gamma}}{2^{1/\gamma}} (B(\gamma - 2, 2))^{1/\gamma} q_\alpha(F),
\]

where the beta function \( B(\cdot, \cdot) \) is defined as

\[
B(a, b) = \int_0^1 x^{a-1}(1 - x)^{b-1} \, dx, \quad a, b > 0.
\]
Figure 4 shows the performances of $\pi_X$ and $\hat{\pi}_X$, where $\hat{\pi}_X$ is calculated with $\epsilon \approx 0.14$. In this example, we have $\theta = 10$ and $\gamma = 3$. While both the analytical and computational approximations exhibit asymptotically equivalence to the exact value, behavior 1 as $\alpha$ increases as shown in Figure 4 (a), the relative ratios in Figure 4 (b) show that the performance of $\hat{\pi}_X$ is superior than $\hat{\pi}_X$ as an approximation of $\pi_X$.

**Appendix**

**Lemma 3.** For $X \in \mathcal{X}_\Psi$ and $s > s_\infty(X)$ we have

$$\mathbb{E} \left[ \frac{X}{s} \partial_x \Psi \left( \frac{X}{s} \right) \right] < \infty.$$  

**Proof.** For any $s'$ such that $s > s' > s_\infty(X)$, we have using convexity of $\Psi(\cdot)$ that

$$0 \leq \left( \frac{x}{s} \right) \Psi' \left( \frac{x}{s} \right) \leq s' \left( \frac{\Psi \left( \frac{x}{s'} \right) - \Psi \left( \frac{x}{s} \right)}{s - s'} \right), \forall x \geq 0.$$  

The above and the fact that $\Psi(x) = 0$ for $x < 0$ completes the proof. \hfill \Box

**Lemma 4.** For $X \in \mathcal{X}_\Psi$, $x < \text{ess sup}(X)$, and $s > s_\infty(X)$ we have the following:

a. $\Lambda(x, \cdot)$ and $\Lambda(\cdot, s)$ are strictly decreasing continuous functions on $(s_\infty(X), \infty)$ and $(-\infty, \text{ess. sup}(X))$, respectively. Moreover, for $x \in \mathbb{R}$, $\Lambda(x, \cdot)$ is convex on $(s_\infty(X), \infty)$, and for $s > s_\infty(X)$, $\Lambda(\cdot, s)$ is convex on $\mathbb{R}$.

b. $\Lambda^{1+}(x, s) = -\frac{1}{s} \mathbb{E} \left[ \Psi' \left( \frac{X-x}{s} \right) \right] < 0$. Moreover, $\Lambda^{1+}(\cdot, s)$ and $\Lambda^{1+}(x, \cdot)$ are right continuous functions on $\mathbb{R}$ and $(s_\infty(X), \infty)$, respectively.

c. $\Lambda^{2-}(x, s) = -\frac{1}{s} \mathbb{E} \left[ \frac{(X-x)}{s} \Psi' \left( \frac{X-x}{s} \right) \right] < 0$. Moreover, $\Lambda^{2-}(\cdot, s)$ and $\Lambda^{2-}(x, \cdot)$ are left continuous functions on $\mathbb{R}$ and $(s_\infty(X), \infty)$, respectively.

**Proof.** The first assertion is straightforward. The other two parts follow easily as applications of DCT by using convexity of $\Psi(\cdot)$, (2), Lemma 3, and Theorem 24.1 in Rockafellar (1997) (for the parts involving one-sided continuity). \hfill \Box

**Acknowledgments**

For Jae Youn Ahn, this work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean Government (2013R1A1A0763062). For Nariankadu D. Shyamalkumar, this work was partially supported by a Society of Actuaries Center of Actuarial Excellence (CAE) Research Grant.

---

3For the details of the calculation, see Ahn (2012).
References


S1. Assign $\hat{\pi}_X(x_J); j = J$

N1. Compute $\hat{\pi}_X(x_j)$

C1. Is $\pi_X(x_{j-1}) < \hat{\pi}_X(x_j)$ ?

C2. Is $\hat{\pi}_X(x_{j-1}) > \pi_X(x_j)$ ?

N2. Compute $\hat{\pi}_X(x_{j-1})$

N3. Compute $\hat{\pi}_X(x_j)$

S2. Stop: $\pi_X \in [\hat{\pi}_X(x_{j+1}) - (x_{j+1} - x_{j-1}), \hat{\pi}_X(x_{j+1})]$

Figure 2: Flowchart for the Algorithm
Figure 3: No. of Newton-Raphson Iterations at Grid Points $x_j$, $j = 853, \ldots, 1000$

Figure 4: Comparison between $\hat{\pi}_X^*$ and Asymptotics of $\pi_X$

(a) Comparison of the convergence

(b) Comparison of the convergence ratio