North American Actuarial Journal

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uaaj20

Computing and Estimating Distortion Risk Measures: How to Handle Analytically Intractable Cases?

Sahadeb Upretee & Vytaras Brazauskas

To cite this article: Sahadeb Upretee & Vytaras Brazauskas (2022): Computing and Estimating Distortion Risk Measures: How to Handle Analytically Intractable Cases?, North American Actuarial Journal, DOI: 10.1080/10920277.2022.2137201

To link to this article: https://doi.org/10.1080/10920277.2022.2137201

Published online: 30 Nov 2022.

North American Actuarial Journal, 0(0), 1–21, 2022

ISSN: 1092-0277 print / 2325-0453 online
DOI: 10.1080/10920277.2022.2137201

Routledge
Taylor & Francis Group

Check for updates

Computing and Estimating Distortion Risk Measures: How to Handle Analytically Intractable Cases?

Sahadeb Upretee[1] **and Vytaras Brazauskas**[2]
[1]*Department of Mathematics, Central Washington University, Ellensburg, Washington*
[2]*Department of Mathematical Sciences, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin*

In insurance data analytics and actuarial practice, *distortion risk measures* are used to capture the riskiness of the distribution tail. Point and interval estimates of the risk measures are then employed to price extreme events, to develop reserves, to design risk transfer strategies, and to allocate capital. Often the computation of those estimates relies on Monte Carlo simulations, which, depending upon the complexity of the problem, can be very costly in terms of required expertise and computational time. In this article, we study analytic and numerical evaluation of distortion risk measures, with the expectation that the proposed formulas or inequalities will reduce the computational burden. Specifically, we consider several distortion risk measures—*value-at-risk* (VaR), *conditional tail expectation* (CTE), *proportional hazards transform* (PHT), *Wang transform* (WT), and *Gini shortfall* (GS)—and evaluate them when the loss severity variable follows shifted exponential, Pareto I, and shifted lognormal distributions (all chosen to have the same support), which exhibit common distributional shapes of insurance losses. For these choices of risk measures and loss models, only the VaR and CTE measures always possess explicit formulas. For PHT, WT, and GS, there are cases when the analytic treatment of the measure is not feasible. In the latter situations, conditions under which the measure is finite are studied rigorously. In particular, we prove several theorems that specify two-sided bounds for the analytically intractable cases. The quality of the bounds is further investigated by comparing them with numerically evaluated risk measures. Finally, a simulation study involving application of those bounds in statistical estimation of the risk measures is also provided.

1. INTRODUCTION

Insurance is a data-rich business that is built on identifying, measuring, and providing protection against extreme or unexpected outcomes. Assessment of the riskiness of the probability distribution tail is an essential task in insurance data analytics, which is accomplished via risk measures. The outcomes of such exercise are then employed in various areas of actuarial practice: asset management, financial valuation and reporting, planning and analysis, product development, designing risk transfer strategies, pricing and reserve calculation, and risk management.

A great number of risk measures belong to the class of *distortion risk measures,* which are defined as integrals of the transformed (or "distorted") survival function of the underlying risk or loss variable. Examples of commonly used distortion risk measures include: Value-at-Risk (VaR), conditional tail expectation (CTE), proportional hazards transform (PHT), Wang transform (WT), and Gini shortfall (GS). There is a large literature on risk measures and their application to contract pricing, capital allocation, and risk management. For a quick introduction into these topics, see Albrecht (2004), Tapiero (2004), and Young (2004). Of course, VaR and CTE have been around for a long time. The other measures are more recent. A series of groundbreaking contributions made by Wang (1995, 2000, 2002) offers practical insights in the development of distortion risk measures, in particular, PHT and WT. The newest proposal among these measures, GS, was introduced by Furman, Wang, and Zitikis (2017). It builds upon and improves the risk measurement ideas studied by Furman and Landsman (2006).

Further, the current literature also provides a variety of proposals regarding the practical aspects of risk assessment exercises. For example, Balbás, Garrido, and Mayoral (2009) offered guidance on which risk measures should be used in practice. To avoid inadequate portfolio selections, these authors proposed considering two properties for risk measures (in conjunction with the well-known coherence properties): completeness and adaptability. The first property ensures that the risk measure uses all of the information of the loss distribution and the second forces the measure to use this information adequately.

Address correspondence to Vytaras Brazauskas, Department of Mathematical Sciences, University of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, WI 53201. E-mail: vytaras@uwm.edu

Among the above-mentioned risk measures, only PHT and WT satisfy these two properties. On the other hand, CTE is the only risk measure that satisfies the system of economic axioms proposed by Wang and Zitikis (2021), which paves the way for constructing Pareto-optimal insurance contracts (see Wang, Wang, and Zitikis, 2021).

Taking a different path, Belles-Sampera, Guillén, and Santolino (2014) proposed a new risk measure, called GlueVaR, that linearly combines VaR and two CTE's at different confidence levels. In addition, these authors defined a new risk measure axiom—tail subadditivity—and demonstrated that GlueVaR satisfies it. The paper also offers useful insights about the relationship between GlueVaR parameters and risk appetite, as well as GlueVaR's applicability to risk management of non-financial sector risks (e.g., health, safety, environmental, adversarial, and catastropic risks including terrorism). See also Belles-Sampera, Guillén, and Santolino (2016) on how to approximate GlueVaR measures for highly skewed loss distributions.

Furthermore, analytic and approximate calculations of risk measures are crucial when one deals with complex insurance products such as variable annuity guaranteed benefits (see Feng and Volkmer, 2012, 2014; Feng, Jing, and Dhaene, 2017). Using advanced probability and mathematical analysis techniques, these authors have reported impressive savings in the computing time of analytic methods over their simulation-based competitors. The risk measures used in those papers were VaR and CTE, and the dynamics of equity prices were driven by geometric Brownian motion (equivalently, lognormal distribution). As examples in Section 2.7 show, these are the "good" scenarios that are amenable to exact analytic calculation. The change of risk measure or distribution might result in a more challenging task.

Motivated by the practical issues discussed above, in this article we focus on computation of VaR, CTE, PHT, WT, and GS. We evaluate these risk measures when the loss variable follows shifted exponential, Pareto I, and shifted lognormal distributions (typical shapes for insurance losses). These right-skewed distributions are parameterized so that all three have the same support. If analytic treatment is feasible, explicit formulas of the risk measures are specified. To clarify what is meant by "explicit": if a risk measure formula can be expressed in terms of a known function, such as the distribution or quantile function of the standard normal variable, which are available in all standard software packages, then the formula will be declared explicit. In other cases, when the risk measures do not have such expressions, we rigorously study their convergence and specify conditions (i.e., restrictions on model parameters) under which they are finite. Specifically, we formulate and prove several theorems that establish a class of lower and upper bounds for each analytically intractable risk measure formula. The rationale for having such bounds is to eventually construct risk measure approximations that have explicit formulas and could effectively be used in large-scale simulation studies, where evaluation of the risk measure needs to be automatic and fast.

The remainder of the article is organized as follows. In Section 2, the class of distortion risk measures is introduced and five specific risk measures—VaR, CTE, PHT, WT, and GS—are defined and discussed from the perspective of risk measure coherence. Then, the measures are evaluated for three severity models and the cases resulting in explicit formulas are summarized. In Section 3, we focus on analytically intractable cases and rigorously examine them by establishing several theorems. Further, in Section 4, we parameterize the severity distributions in such a way that they have similar right tail. Then, we explore the behavior of the analytic bounds of the PHT and WT measures. Examples involving statistical estimation of PHT, WT, and GS with the help of new bounds are also provided. Furthermore, proofs of the theorems are provided in Section 5. The main results of the article are summarized and future research directions are discussed in Section 6. Finally, in Appendix A, we present key probabilistic functions of typical severity models: the shifted exponential, Pareto I, and shifted lognormal distributions. In addition, two lemmas used in the proofs of Section 5 are provided and proved in Appendix B.

2. PRELIMINARIES

In this section, we first present a broad class of risk measures (distortion risk measures) and briefly discuss the concept of risk measure coherence. Then, we specify several popular distortion risk measures: VaR, CTE, PHT, WT, and GS. Finally, Table 1 summarizing the cases where these risk measures have explicit formulas (for the parametric models of Appendix A) is provided.

2.1. Distortion Risk Measures

A distortion risk measure is defined as the expectation of loss with respect to distorted probabilities. In particular, for a continuous random variable $X \geq 0$ with cumulative distribution function (cdf) F, a risk measure R is defined as

$$R[F] = \int_0^\infty g(1-F(x))\,dx, \tag{2.1}$$

TABLE 1
Formulas of VaR, CTE, PHT, WT, and GS Risk Measures for Shifted Exponential, Pareto I, and Shifted Lognormal Distributions

Risk measure	Exponential: $\mathcal{E}xp\ (x_0, \theta)$	Pareto I: $\mathcal{P}a\ I\ (x_0, \alpha)$	Lognormal: $\mathcal{LN}\ (x_0, \mu, \sigma)$
$\mathrm{VaR}[F, \beta]$	$x_0 - \theta \log(\beta)$	$x_0 \beta^{-\frac{1}{\alpha}}$	$x_0 + e^{\mu + \sigma \Phi^{-1}(1-\beta)}$
$\mathrm{CTE}[F, \beta]$	$x_0 - \theta \log(\beta) + \theta$	$x_0 \beta^{-\frac{1}{\alpha}} \cdot \frac{\alpha}{\alpha - 1}, \quad \alpha > 1$	$x_0 + \frac{1}{\beta}\ e^{\mu + \frac{\sigma^2}{2}} \cdot \Phi(\sigma - \Phi^{-1}(1-\beta))$
$\mathrm{PHT}[F, r]$	$x_0 + \frac{\theta}{r}$	$x_0 \cdot \frac{r\alpha}{r\alpha - 1}, \quad \alpha > \frac{1}{r}$	No explicit formula
$\mathrm{WT}[F, \lambda]$	No explicit formula	No explicit formula	$x_0 + e^{\mu + \lambda\sigma + \frac{\sigma^2}{2}}$
$\mathrm{GS}[F, \beta, \delta]$	$x_0 - \theta \log(\beta) + \theta(1 + \delta)$	$x_0 \beta^{-\frac{1}{\alpha}} \cdot \frac{\alpha}{\alpha - 1} \cdot \frac{2(\alpha+\delta)-1}{2\alpha - 1}, \quad \alpha > 1$	No explicit formula

Note: For the entries under $\mathcal{P}a\ I\ (x_0, \alpha)$, if the restriction on α specified in the formula is not satisfied, then the corresponding risk measure is infinite.

where the distortion function $g : [0, 1] \to [0, 1]$ is an increasing function with $g(0) = 0$ and $g(1) = 1$. Moreover, if g is differentiable, then integration by parts in (2.1) leads to

$$R[F] \;=\; \int_0^1 F^{-1}(u)\psi(u)\, du, \tag{2.2}$$

where $\psi(u) = g'(1-u)$ and F^{-1} is the quantile function of variable X.

A number of authors studied the question of what a "good" risk measure is and what properties it should satisfy (see, for example, discussion by Albrecht, 2004). Among multiple axiomatic systems, the one proposed by Artzner et al. (1999) has become quite influential. It advocates the use of coherent measures, which are defined as follows. For loss variables X_1 and X_2, a mapping of random variables to real numbers, $\varrho[\cdot]$, is called a *coherent risk measure* if it satisfies the following four axioms:

1. Translation invariance*:* $\varrho[X_1 + a] = \varrho[X_1] + a$, where a is a real-valued constant.
2. Positive homogeneity*:* $\varrho[bX_1] = b\varrho[X_1]$, where b is a positive constant.
3. Subadditivity*:* $\varrho[X_1 + X_2] \le \varrho[X_1] + \varrho[X_2]$.
4. Monotonicity*:* If $\mathbf{P}\{X_1 \le X_2\} = 1$, then $\varrho[X_1] \le \varrho[X_2]$.

These properties have intuitively appealing interpretations. The first one says that if a risk-free amount of capital (e.g., cash) is added to or subtracted from a portfolio of risks, then the overall riskiness of the portfolio should be shifted by that amount. The second property applies to rescaling of risk (e.g., assets affected by inflation or currency exchange) and states that the risk measure should be affected by the same scale factor as the risk itself. Subadditivity is also known as the portfolio diversification property: if two portfolios are combined into one, their overall riskiness should not exceed the total riskiness of individual portfolios. The fourth property means that stochastically larger risk (or portfolio) should be riskier than stochastically smaller one.

2.2. Value-at-Risk

The VaR measure on a portfolio of risks (i.e., potential losses) is the maximum loss one might expect over a given period of time, at a given level of confidence (say, β). In mathematical terms, this measure is defined as the $(1-\beta)$-level quantile of the distribution function F:

$$\mathrm{VaR}[F, \beta] \;=\; F^{-1}(1-\beta). \tag{2.3}$$

Note that VaR can be expressed as distortion risk measure, defined by (2.1), by choosing $g(u) = 0$ for $0 \le u < \beta$, and $= 1$ for $\beta \le u \le 1$. These choices correspond to $g(1-F(x)) = 0$ for $0 \le 1-F(x) < \beta$, and $= 1$ for $\beta \le 1-F(x) \le 1$ or, equivalently, $g(1-F(x)) = 0$ for $F^{-1}(1-\beta) < x < \infty$, and $= 1$ for $0 \le x \le F^{-1}(1-\beta)$. Now expression (2.3) follows easily from (2.1):

$$\mathrm{VaR}[F,\beta] \;=\; \int_0^{F^{-1}(1-\beta)} 1\,dx \;+\; \int_{F^{-1}(1-\beta)}^{\infty} 0\,dx \;=\; F^{-1}(1-\beta).$$

This risk measure, however, is *not coherent* because it does not satisfy the subadditivity property (it does satisfy the other three properties, though). Despite this axiomatic drawback, the VaR measure remains popular among practitioners (especially in the banking industry), which is mainly due to its computational simplicity and straightforward interpretation.

2.3. Conditional Tail Expectation

The CTE measure (also known as tail-VaR, tail conditional expectation or expected shortfall) is the conditional expectation of a loss variable given that it exceeds a specified quantile, $\mathrm{VaR}[F,\beta]$. It measures the expected maximum loss in the $100\beta\%$ worst cases, over a given period of time:

$$\mathrm{CTE}[F,\beta] = F^{-1}(1-\beta) + \frac{1}{\beta}\int_{F^{-1}(1-\beta)}^{\infty} [1-F(x)]\,dx \tag{2.4}$$

$$= \frac{1}{\beta}\int_{1-\beta}^{1} F^{-1}(u)\,du. \tag{2.5}$$

It is clear from (2.4) that this measure can be expressed as (2.1) by choosing $g(t) = t/\beta$ for $0 \le t < \beta$, and $= 1$ for $\beta \le t \le 1$. Alternatively, expression (2.5) follows from (2.2) with $\psi(u) = 0$ for $0 \le u \le 1-\beta$, and $= 1/\beta$ for $1-\beta < u \le 1$. Further, CTE is a coherent risk measure and it answers the often asked "what if" question. Indeed, comparing (2.4) with (2.3), we see that there is a direct relationship between VaR and CTE. That is, in case an extreme (low probability, high impact) event happens, VaR tells us only the lower bound of possible losses; CTE, on the other hand, provides an estimate of expected loss if the extreme event occurs. Thus, CTE is more informative. Moreover, according to Wang and Zitikis (2021), CTE rewards portfolio diversification and penalizes risk concentration. These properties are not shared by any other risk measure.

2.4. Proportional Hazards Transform

The PHT measure was introduced by Wang (1995) as a new insurance premium principle, where additional risk loadings are proportional to the hazard rates (hence the name of the measure). This premium principle is scale invariant and additive for layers and enjoys some optimality properties in reinsurance sharing arrangements. The PHT measure is defined by the distortion function $g(s) = s^r$ or, equivalently, by the weight function $\psi(s) = r(1-s)^{r-1}$:

$$\mathrm{PHT}[F,r] \;=\; \int_0^{\infty} [1-F(x)]^r\,dx \;=\; r\int_0^1 F^{-1}(u)(1-u)^{r-1}\,du, \tag{2.6}$$

where constant r $(0 < r \le 1)$ represents the degree of distortion and F^{-1} is the quantile function (qf) of X. Note that $\mathrm{PHT}[F,1]$ is the expected value of X, and $\mathrm{PHT}[F,1/2] - \mathrm{PHT}[F,1]$ is the right-tailed deviation of X. Small r corresponds to high distortion, but in most practical situations r varies between 1/2 and 1. Moreover, PHT is a coherent risk measure.

2.5. Wang Transform

The WT measure was introduced by Wang (2000, 2002) as a tool for pricing both liabilities (insurance losses) and asset returns (gains). It is a very effective measure for finance models driven by normal or lognormal random variables. For example, for normally distributed asset returns, the WT measure recovers two well-known results: the capital asset pricing model and the Black-Scholes formula. Here our focus will be on insurance losses. In this context, the WT measure is defined by the distortion function $g(t) = \Phi(\Phi^{-1}(t) + \lambda)$ or, equivalently, by the weight function $\psi(u) = e^{\lambda\Phi^{-1}(u) - \lambda^2/2}$:

$$\text{WT}[F,\lambda] \;=\int_0^\infty \Phi\big(\Phi^{-1}(1-F(x))+\lambda\big)\,dx \;=\int_0^1 F^{-1}(u)\;e^{\lambda\Phi^{-1}(u)-\lambda^2/2}\,du, \tag{2.7}$$

where Φ and Φ^{-1} denote the cdf and qf of the standard normal random variable, respectively. Parameter λ ($-\infty<\lambda<\infty$) reflects the level of systematic risk and is called the *market price of risk* or *risk aversion index.* Although in theory λ can be any real number, in applications its typical range is from –1 to 1. Also, WT is a coherent risk measure.

2.6. Gini Shortfall

The GS measure was introduced by Furman, Wang, and Zitikis (2017) with the motivation to capture both the expectation and variability of X beyond an extreme quantile (i.e., beyond the VaR). Antecedents of this idea were studied by Furman and Landsman (2006), who proposed to supplement CTE with the tail standard deviation of X. The main shortcoming of such risk measures is that they require finite second moments of the underlying loss variables. The GS measure, on the other hand, replaces tail standard deviation with the tail Gini index, which captures tail variability of the loss variable X and requires only the first moment to be finite. Formally, the GS measure is defined as

$$\begin{aligned}\text{GS}[F,\beta,\delta] &= F^{-1}(1-\beta)+\frac{1}{\beta}\int_{F^{-1}(1-\beta)}^\infty [1-F(x)]\,dx\\ &\quad+\;\frac{2\delta}{\beta^2}\int_{F^{-1}(1-\beta)}^\infty [1-F(x)][\beta-1+F(x)]\,dx\end{aligned} \tag{2.8}$$

$$=\frac{1}{\beta^2}\int_{1-\beta}^1 F^{-1}(u)\;\big(\beta+4\delta(u-1+\beta/2)\big)\,du, \tag{2.9}$$

where $0<\beta<1$ is the risk appetite (actually, Furman, Wang, and Zitikis [2017], instead of parameter β used $1-\beta$, and called it the *prudence level)* and $\delta\ge 0$ is the *loading parameter*. As is evident from (2.8), the distortion function for GS is $g(t)=t/\beta+2\delta(t/\beta)(1-t/\beta)$ for $0\le t<\beta$, and $=1$ for $\beta\le t\le 1$. Alternatively, (2.9) follows from (2.2) with $\psi(u)=\beta^{-2}(\beta+4\delta(u-1+\beta/2))\;\mathbf{1}\{1-\beta\le u\le 1\}$, where $\mathbf{1}\{\}$ denotes the indicator function. Comparing (2.8) with (2.4), we see that GS is essentially CTE with an extra term for tail variability. Finally, as it was proven in theorem 4.1 of Furman, Wang, and Zitikis (2017), GS is a coherent risk measure if and only if $0\le\delta\le 1/2$.

2.7. Explicit Formulas

When claim severities follow shifted exponential, Pareto I, and shifted lognormal distributions (see Appendix A, for specific parameterizations of those models), explicit formulas for some of the risk measures are known or can be easily derived. In Table 1, we summarize such cases and identify the cases for which the selected risk measures do not have explicit formulas.

3. THEORETICAL FOUNDATIONS

The "no explicit formula" cases listed in Table 1 emerge because for those risk measures the defining formulas involve analytically intractable integrals. Thus, we repeat the question from the title of the article: how to handle analytically intractable cases? Our proposal is to establish two-sided bounds for those integrals and then use the bounds to construct explicit-formula approximations. As numerical illustrations in Section 4.3 exemplify, the approximations designed this way are easy to implement and work without human intervention, which is essential when conducting large-scale computer simulations.

In this section, we formulate four theorems that show that the relevant integrals are bounded. Then we evaluate the quality of the bounds by comparing them with direct numerical evaluations. Proofs of the theorems will be provided in Section 5.

TABLE 2
Numerical Evaluations of $C_{\text{PHT}}(r,\sigma)$ for Selected r and σ

	σ							
r	1/10	1/5	1/4	1/2	1	2	4	5
0.55	1.069	1.161	1.216	1.625	3.896	77.453	5.7×10^6	2.3×10^{10}
0.65	1.050	1.116	1.157	1.455	2.979	36.422	4.7×10^5	5.2×10^8
0.75	1.034	1.082	1.111	1.332	2.412	20.386	7.1×10^4	3.1×10^7
0.85	1.021	1.054	1.075	1.239	2.030	12.813	1.6×10^4	3.4×10^6
0.95	1.010	1.030	1.045	1.165	1.758	8.739	5.0×10^3	5.8×10^5
1	1.005	1.020	1.032	1.133	1.649	7.389	3.0×10^3	2.7×10^5

3.1. PHT of Shifted Lognormal

If $X \sim \mathcal{LN}\ (x_0, \mu, \sigma)$, then its PHT measure is found by integrating (2.6) as follows:

$$\begin{aligned} \text{PHT}[F,r] &= r\int_0^1 F^{-1}(u)(1-u)^{r-1}\,du \;=\; r\int_0^1 \left[x_0 + e^{\mu+\sigma\Phi^{-1}(u)}\right](1-u)^{r-1}\,du \\ &= x_0 + e^{\mu}\sigma\int_{-\infty}^{\infty}(1-\Phi(z))^r e^{\sigma z}\,dz \;=:\; x_0 + e^{\mu}\;C_{\text{PHT}}(r,\sigma), \end{aligned}$$

where $0<r\le 1$ is the degree of distortion. Note that for fixed r and σ, the integral $C_{\text{PHT}}(r,\sigma) = \sigma\int_{-\infty}^{\infty}(1-\Phi(z))^r\ e^{\sigma z}\,dz$ is finite and can be evaluated numerically. Theorem 1 establishes a class of lower and upper bounds for $C_{\text{PHT}}(r,\sigma)$.

Theorem 1. *For $0<r\le 1$ and $\sigma>0$, define $C_{\text{PHT}}(r,\sigma) = \sigma\int_{-\infty}^{\infty}(1-\Phi(z))^r\ e^{\sigma z}\,dz$, where Φ is the cdf of the standard normal distribution. Then the double inequality*

$$e^{\sigma x}\left[(1-\Phi(x))^r - (1-\Phi(x))\right] \;+\; e^{\sigma^2/2}\;\Phi(\sigma - x) \;\le\; C_{\text{PHT}}(r,\sigma) \;<\; e^{\sigma x} \;+\; K_x(r,\sigma),$$

where $K_x(r,\sigma) = \sigma x^{-r} r^{-1/2}(2\pi)^{(1-r)/2} e^{\sigma^2/(2r)}\Phi((\sigma - rx)/\sqrt{r})$, holds for every $x > 0$.

Proof. The proof is provided in Section 5. □

In Table 2, we provide numerical evaluations of $C_{\text{PHT}}(r,\sigma)$ for typical ranges of r and σ. Note that the lower and upper bounds established in Theorem 1 work well, although more work is needed to identify optimal value of x. A few illustrations for $x = (1/2)\times(\sigma/r)$: for $r=0.55$ and $\sigma=1$, we have $C_{\text{PHT}}(r,\sigma)\approx 3.896$ and $1.40 < 3.896 < 6.48$; for $r=0.75$ and $\sigma=2$, we have $C_{\text{PHT}}(r,\sigma)\approx 20.386$ and $6.60 < 20.386 < 43.91$; for $r=0.95$ and $\sigma=4$, we have $C_{\text{PHT}}(r,\sigma)\approx 5.0\times 10^3$ and $2.9\times 10^3 < 5.0\times 10^3 < 1.4\times 10^4$.

3.2. WT of Shifted Exponential and of Pareto I

If $X \sim \mathcal{Exp}\ (x_0, \theta)$, then its WT measure is found by integrating (2.7) as follows:

$$\begin{aligned} \text{WT}[F,\lambda] &= \int_0^1 F^{-1}(u)\ e^{\lambda\Phi^{-1}(u)-\lambda^2/2}\,du \;=\; \int_0^1 [x_0 - \theta\log(1-u)]\ e^{\lambda\Phi^{-1}(u)-\lambda^2/2}\,du \\ &= x_0 + \theta\int_{-\infty}^{\infty}\Phi(z+\lambda)\ \frac{\varphi(z)}{\Phi(z)}\,dz \;=:\; x_0 + \theta\ C_{\text{WT}}(\lambda), \end{aligned}$$

where φ is the probability density function (pdf) of the standard normal random variable and $-\infty<\lambda<\infty$ is the risk aversion index. Note that for fixed λ, the integral $C_{\text{WT}}(\lambda) = \int_{-\infty}^{\infty}\Phi(z+\lambda)\ \frac{\varphi(z)}{\Phi(z)}\,dz$ is finite and can be evaluated numerically. Theorem 2 establishes a class of lower and upper bounds for $C_{\text{WT}}(\lambda)$.

Theorem 2. *For any real constant λ, define $C_{\mathrm{WT}}(\lambda) = \int_{-\infty}^{\infty} \Phi(z+\lambda)\ \frac{\varphi(z)}{\Phi(z)}\,dz$, where Φ and φ are the cdf and pdf of the standard normal distribution, respectively. Then*

a. *For $\lambda \le 0$, the double inequality*

$$-\Phi(x+\lambda)\log[\Phi(x)] \ \le\ C_{\mathrm{WT}}(\lambda) \ \le\ \Phi(x+\lambda) - \log[\Phi(x)]$$

holds for every $x < 0$.

b. *For $\lambda > 0$, the double inequality*

$$-\Phi(x+\lambda)\log[\Phi(x)] \ \le\ C_{\mathrm{WT}}(\lambda) \ <\ \frac{x}{x+\lambda}\ \Phi(x+\lambda) - \log[\Phi(x)]$$

holds for every $x < -\lambda$.

Proof. The proof is provided in Section 5. □

If $X \sim \mathcal{P}a\ I\ (x_0, \alpha)$, then its WT measure is found by integrating (2.7) as follows:

$$\begin{aligned}\mathrm{WT}[F,\lambda] &= \int_0^1 F^{-1}(u)\ e^{\lambda\Phi^{-1}(u)-\lambda^2/2}\,du \ =\ \int_0^1 x_0(1-u)^{-1/\alpha}\ e^{\lambda\Phi^{-1}(u)-\lambda^2/2}\,du \\ &= x_0 + \frac{x_0}{\alpha}\int_{-\infty}^{\infty}\Phi(z+\lambda)\ \frac{\varphi(z)}{[\Phi(z)]^{1/\alpha+1}}\,dz \ =:\ x_0 + \frac{x_0}{\alpha}\ C_{\mathrm{WT}}(\lambda,\alpha),\end{aligned}$$

where $-\infty < \lambda < \infty$ is the risk aversion index. Note that for fixed λ and $\alpha > 1$, the integral $C_{\mathrm{WT}}(\lambda,\alpha) = \int_{-\infty}^{\infty}\Phi(z+\lambda)\ \frac{\varphi(z)}{[\Phi(z)]^{1/\alpha+1}}\,dz$ is finite and can be evaluated numerically. Theorem 3 establishes a class of lower and upper bounds for $C_{\mathrm{WT}}(\lambda,\alpha)$.

Theorem 3. *For any real constant λ and $\alpha > 1$, define $C_{\mathrm{WT}}(\lambda,\alpha) = \int_{-\infty}^{\infty}\Phi(z+\lambda)\ \frac{\varphi(z)}{[\Phi(z)]^{1/\alpha+1}}\,dz$, where Φ and φ are the cdf and pdf of the standard normal distribution, respectively. Then*

a. *For $\lambda \le 0$, the double inequality*

$$\Phi(x+\lambda)\ c_x(\alpha) \ \le\ C_{\mathrm{WT}}(\lambda,\alpha) \ \le\ c_x(\alpha) + \frac{\alpha}{\alpha-1}\ e^{-\lambda x-\lambda^2/2}\ [\Phi(x)]^{1-1/\alpha},$$

where $c_x(\alpha) = \alpha([\Phi(x)]^{-1/\alpha} - 1)$, holds for every $x < 0$.

b. *For $\lambda > 0$, the double inequality*

$$\Phi(x+\lambda)\ c_x(\alpha) \ \le\ C_{\mathrm{WT}}(\lambda,\alpha) \ <\ c_x(\alpha) + \frac{\alpha}{\alpha-1}\ \frac{x}{x+\lambda}\left[e^{-\lambda x-\lambda^2/2}\ [\Phi(x)]^{1-1/\alpha} + C_x(\alpha,\lambda)\right],$$

where $C_x(\alpha,\lambda) = \lambda\ \sqrt{\alpha(2\pi)^{1/\alpha}/(\alpha-1)}\ (-x)^{1/\alpha-1}\ e^{-(\lambda^2/2)/(\alpha-1)}\ \Phi\left(\frac{x+\lambda\alpha/(\alpha-1)}{\sqrt{\alpha/(\alpha-1)}}\right)$ holds for every $x < -\lambda$.

Proof. The proof is provided in Section 5. □

In Table 3, we provide numerical approximations of $C_{\mathrm{WT}}(\lambda)$ and $C_{\mathrm{WT}}(\lambda,\alpha)$ for typical ranges of λ and α. Note that the lower and upper bounds established in Theorems 2 and 3 are reasonably tight, although more work is needed to identify an optimal value of x. A few illustrations for $C_{\mathrm{WT}}(\lambda)$: for $\lambda = -0.5$, we have $C_{\mathrm{WT}}(\lambda) \approx 0.619$ and $0.21 < 0.619 < 1.14$ (when $x = \lambda/2$); for $\lambda = 0.25$, we have $C_{\mathrm{WT}}(\lambda) \approx 1.245$ and $0.47 < 1.245 < 1.58$ (when $x = -2\lambda$); for $\lambda = 1$, we have $C_{\mathrm{WT}}(\lambda) \approx 2.232$ and $0.60 < 2.232 < 3.94$ (when $x = -2\lambda$). Likewise, for $C_{\mathrm{WT}}(\lambda,\alpha)$: for $\lambda = -0.5$ and $\alpha = 2.5$, we have $C_{\mathrm{WT}}(\lambda,\alpha) \approx 0.886$ and $0.25 < 0.886 < 1.85$ (when $x = \lambda/2$); for $\lambda = -1$ and $\alpha = 4$, we have $C_{\mathrm{WT}}(\lambda,\alpha) \approx 0.416$ and $0.09 < 0.416 < 1.57$ (when $x = \lambda/2$); for $\lambda = 0.5$ and $\alpha = 1.25$, we have $C_{\mathrm{WT}}(\lambda,\alpha) \approx 20.965$ and $1.30 < 20.965 < 24.86$ (when $x = -2\lambda$).

TABLE 3
Numerical Evaluations of $C_{\mathrm{WT}}(\lambda)$ and $C_{\mathrm{WT}}(\lambda,\alpha)$ for Selected λ and α

		$C_{\mathrm{WT}}(\lambda,\alpha)$ for α								
λ	$C_{\mathrm{WT}}(\lambda)$	1.1	1.25	1.5	1.75	2	2.5	3	4	5
–1	0.359	0.806	0.681	0.582	0.531	0.499	0.461	0.440	0.416	0.403
–0.5	0.619	2.389	1.692	1.281	1.101	0.999	0.886	0.825	0.760	0.727
–0.25	0.792	4.719	2.820	1.938	1.595	1.412	1.217	1.116	1.011	0.958
0	1.000	11.000	5.000	3.000	2.333	2.000	1.667	1.500	1.333	1.250
0.25	1.245	33.003	9.663	4.799	3.468	2.857	2.283	2.009	1.745	1.616
0.5	1.530	141.659	20.965	8.020	5.272	4.132	3.135	2.686	2.270	2.074
1	2.232	11,090.602	158.182	26.874	13.403	9.143	6.035	4.824	3.803	3.355

3.3. GS of Shifted Lognormal

If $X \sim \mathcal{LN}\ (x_0, \mu, \sigma)$, then its GS measure is found by integrating (2.9) as follows:

$$\begin{aligned}
\mathrm{GS}[F,\beta,\delta] &= \frac{1}{\beta^2}\int_{1-\beta}^{1} F^{-1}(u)\ (\beta + 4\delta(u-1+\beta/2))\,du \\
&= \frac{1}{\beta^2}\int_{1-\beta}^{1} \left[x_0 + e^{\mu+\sigma\Phi^{-1}(u)}\right]\ (\beta(1+2\delta)-4\delta(1-u))\,du \\
&= x_0 + \beta^{-2}e^{\mu+\sigma^2/2}(\beta(1+2\delta)-4\delta)\Phi(\sigma-\Phi^{-1}(1-\beta)) \\
&+\ 4\delta\beta^{-2}e^{\mu+\sigma^2/2}\int_{\Phi^{-1}(1-\beta)}^{\infty}\Phi(z)\ \varphi(z-\sigma)\,dz \\
&=: x_0 + \beta^{-2}e^{\mu+\sigma^2/2}([\beta(1+2\delta)-4\delta]\ \Phi\big(\sigma-\Phi^{-1}(1-\beta)\big) + 4\delta\ C_{\mathrm{GS}}(\beta,\sigma)),
\end{aligned}$$

where $0<\beta<1$ is the risk appetite and $0\le\delta\le 1/2$ is the loading parameter (restricted to the interval $[0;1/2]$ to make GS coherent). Note that for fixed β and σ, the integral $C_{\mathrm{GS}}(\beta,\sigma) = \int_{\Phi^{-1}(1-\beta)}^{\infty}\Phi(z)\ \varphi(z-\sigma)\,dz$ is finite and can be evaluated numerically. Theorem 4 establishes a lower and upper bound for $C_{\mathrm{GS}}(\beta,\sigma)$.

Theorem 4. *For $0<\beta<1$ and $\sigma>0$, define $C_{\mathrm{GS}}(\beta,\sigma) = \int_{\Phi^{-1}(1-\beta)}^{\infty}\Phi(z)\ \varphi(z-\sigma)\,dz$, where Φ, φ, and Φ^{-1} denote the cdf, pdf, and qf of the standard normal distribution, respectively. Then the following two-sided inequality holds:*

$$(1-\beta)\ \Phi\big(\sigma-\Phi^{-1}(1-\beta)\big)\ \le\ C_{\mathrm{GS}}(\beta,\sigma)\ \le\ \Phi\big(\sigma-\Phi^{-1}(1-\beta)\big).$$

Proof. The proof is provided in Section 5. □

In Table 4, we provide numerical evaluations of $C_{\mathrm{GS}}(\beta,\sigma)$ for typical ranges of β and σ. Note that the lower and upper bounds established in Theorem 4 are reasonably tight. For instance: for $\beta=0.20$ and $\sigma=1/10$, we have $C_{\mathrm{GS}}(\beta,\sigma)\approx 0.207$ and $0.183<0.207<0.229$; for $\beta=0.01$ and $\sigma=1$, we have $C_{\mathrm{GS}}(\beta,\sigma)\approx 0.092$ and $0.091<0.092\le 0.092$; for $\beta=0.10$ and $\sigma=5$, we have $C_{\mathrm{GS}}(\beta,\sigma)\approx 1.000$ and $0.900<1.000\le 1.000$.

4. NUMERICAL ILLUSTRATIONS

In this section, we develop several illustrations involving risk measurement exercises. In particular, in Section 4.1, we select shifted exponential (F_1), Pareto I (F_2), and shifted lognormal (F_3) distributions in such a way that they all have identical supports and similar right tail. Then, using the specific distributions of Section 4.1, we plot the two-sided bounds of Theorems 1–3, and explore potential estimates of C_{PHT} and C_{WT}. Finally, in Section 4.3, a simulation study involving application of those bounds in statistical estimation of the risk measures is presented.

TABLE 4
Numerical Evaluations of $C_{GS}(\beta, \sigma)$ for Selected β and σ

	σ							
β	1/10	1/5	1/4	1/2	1	2	4	5
0.01	0.013	0.017	0.019	0.034	0.092	0.371	0.952	0.996
0.05	0.060	0.073	0.080	0.123	0.255	0.631	0.989	0.999
0.10	0.113	0.133	0.144	0.208	0.375	0.747	0.995	1.000
0.15	0.162	0.187	0.201	0.277	0.459	0.807	0.997	1.000
0.20	0.207	0.236	0.251	0.335	0.523	0.844	0.997	1.000
0.25	0.248	0.280	0.297	0.385	0.573	0.868	0.997	1.000

4.1. PDFs with Similar Tails

To construct three pdfs with similar tails, we select their parameters so that F_1, F_2, and F_3 are equally risky according to some risk measures. More specifically, we first completely specify Pareto I: $F_2 = \mathcal{P}a\ I\ (x_0 = 1, \alpha = 2)$. This distribution has a heavy right tail, with all its k-moments, when $k \geq 2$, being infinite. Then the other distributions are chosen to be equally risky according to VaR at β_1 and β_2; that is, $\mathrm{VaR}[F_1, \beta_1] = \mathrm{VaR}[F_2, \beta_1] = \mathrm{VaR}[F_3, \beta_1]$ and $\mathrm{VaR}[F_1, \beta_2] = \mathrm{VaR}[F_3, \beta_2]$. After solving these three equations with respect to θ, σ, and μ, we have the following expressions:

$$\theta = \frac{1-\beta_1^{-1/2}}{\log(\beta_1)}, \qquad \sigma = \frac{\log[\log(\beta_2)/\log(\beta_1)]}{\Phi^{-1}(1-\beta_2) - \Phi^{-1}(1-\beta_1)}, \qquad \mu = \log(\beta_1^{-1/2} - 1) - \sigma\ \Phi^{-1}(1-\beta_1). \tag{4.1}$$

Choosing $\beta_1 = 0.10$ and $\beta_2 = 0.05$ in (4.1) yields $\theta = 0.9391$, $\sigma = 0.7243$, $\mu = -0.1571$. In Figure 1, we plot big picture and tail view of pdfs of $F_1 = \mathcal{E}xp\ (x_0 = 1, \theta = 0.9391)$, $F_2 = \mathcal{P}a\ I\ (x_0 = 1, \alpha = 2)$, and $F_3 = \mathcal{LN}\ (x_0 = 1, \mu = -0.1571, \sigma = 0.7243)$ along with several risk measure values.

Visual assessment of the graphs suggests that F_1, F_2, and F_3 have similar right tail, and according to VaR at $\beta_1 = 0.10$, all three distributions are indeed equally risky: $\mathrm{VaR}[F_1, 0.10] = \mathrm{VaR}[F_2, 0.10] = \mathrm{VaR}[F_3, 0.10] = 3.16$. Working with real data and using traditional model validation tools, one would likely conclude that these models are similarly risky. A more comprehensive look into the tails of these distributions, however, reveals a different picture: whereas the shifted exponential and log-normal distributions remain similarly risky, the Pareto I distribution drifts away from the other two. Indeed: $\mathrm{VaR}[F_2, 0.05] = 4.47$ is 17% larger than $\mathrm{VaR}[F_1, 0.05] = \mathrm{VaR}[F_3, 0.05] = 3.81$, $\mathrm{CTE}[F_2, 0.10] = 6.32$ is 50% larger than $\mathrm{CTE}[F_3, 0.10] = 4.21$, and $\mathrm{GS}[F_2, 0.10] = 8.43$ is 77% larger than $\mathrm{GS}[F_3, 0.10, 0.50] = 4.77$.

4.2. PHT and WT Bounds

Using F_1, F_2, and F_3 as specified in Section 4.1, we now explore how effective the two-sided bounds for the constants C_{PHT} and C_{WT} are. Note that for completely known distributions, these constants (as defined in Theorems 1–3) can be approximated via direct numerical integration. However, because integration involves infinite intervals, we first have to plot the relevant curve over various intervals until the finite interval with practically all area under the curve is found. Of course, this approach works on a case-by-case basis but is impractical in simulations where such evaluations would have to be done automatically and thousands of times. Therefore, the bounds of Theorems 1–3 become essential. Also, we propose to use the average of the upper and lower bounds as an initial estimate of the corresponding constant. In Figure 2, we plot the lower and upper bounds, the average of the bounds, and the target value of $C_{PHT}(r, \sigma = 0.7243)$ for $r = 0.55, 0.75, 0.95$. (The target values are found by direct numerical integration.)

As is evident from the figure, the average of the bounds approximates the target value very well when x is chosen to minimize the length of the interval [*lower bound*; *upper bound*]. Solving such a problem in general is a complicated task, but choosing $x = 1$ is satisfactory for typical values of r (and $\sigma = 0.7243$), with the relative error being less than 10%: average = 2.20 (target = 2.29) for $r = 0.55$; average = 1.75 (target = 1.67) for $r = 0.75$; average = 1.51 (target = 1.36) for $r = 0.95$.

For constants $C_{WT}(\lambda)$ and $C_{WT}(\lambda, \alpha)$, the bounds and the domain of x depend on the sign of λ. For $\lambda < 0$, $x < 0$, and for $\lambda > 0$, $x < -\lambda$. In Figure 3, we plot the lower and upper bounds, the average of the bounds, and the target value of $C_{WT}(\lambda)$ for selected values of λ.

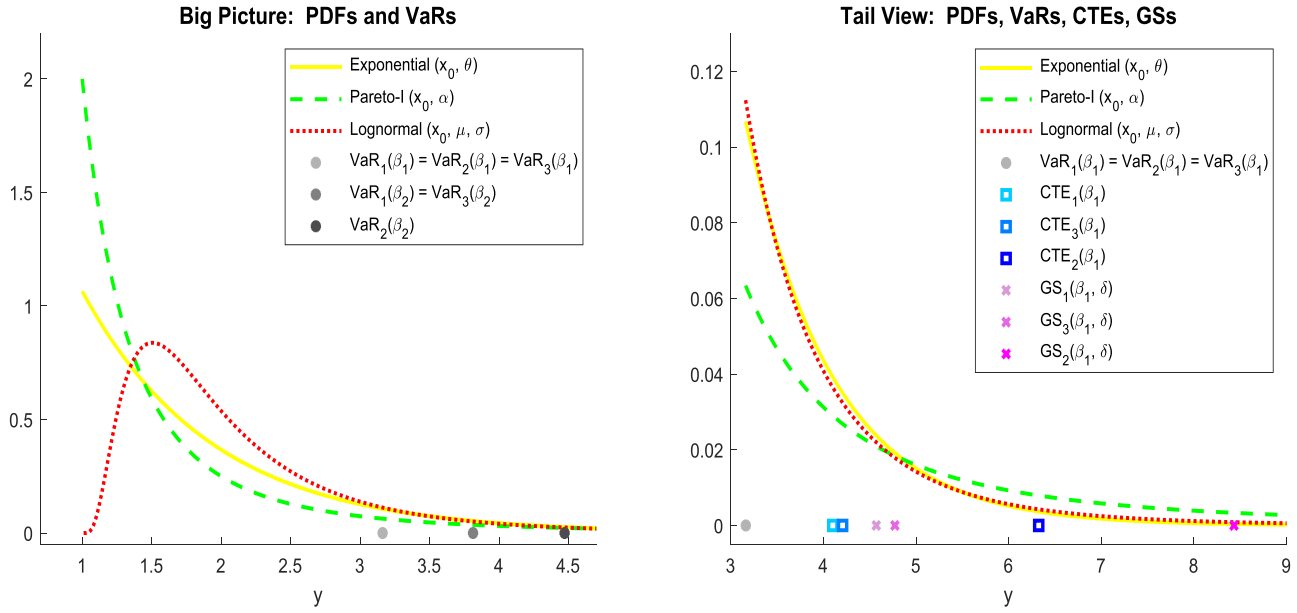

FIGURE 1. (a) Big-Picture and (b) Tail-View graphs of pdf's of $\mathcal{E}xp$ $(x_0 = 1, \theta = 0.9391)$, $\mathcal{P}a\ I$ $(x_0 = 1, \alpha = 2)$, and $\mathcal{LN}$ $(x_0 = 1, \mu = -0.1571, \sigma = 0.7243)$, with Marked Values of VaR at $\beta_1 = 0.10$ and $\beta_2 = 0.05$, CTE at $\beta_1 = 0.10$, and GS at $\beta_1 = 0.10$, $\delta = 0.50$.

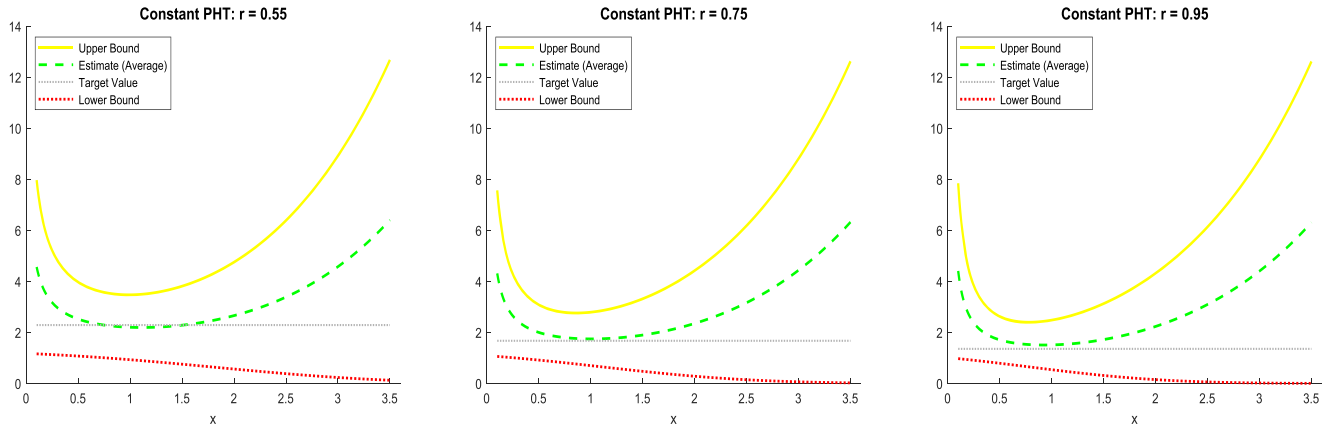

FIGURE 2. Lower and Upper Bounds, Estimate (Average of the Bounds), and the Target Value of Constant $C_{\text{PHT}}(r, \sigma = 0.7243)$ for $r = 0.55, 0.75, 0.95$.

It is clear from the figure that the average-based estimate performs reasonably for x near 0 when $\lambda < 0$ and for x slightly below $-\lambda$ when $\lambda > 0$. For example, choosing $x = -\lambda - 0.5$ for $\lambda > 0$ leads to the relative error less than 4%: average = 1.20 (target = 1.24) for $\lambda = 0.25$; average = 1.51 (target = 1.53) for $\lambda = 0.5$; average = 2.23 (target = 2.23) for $\lambda = 1$. Likewise, choosing $x = -0.25$ for $\lambda < 0$ leads to the relative error less than 37%: average = 0.75 (target = 0.79) for $\lambda = -0.25$; average = 0.67 (target = 0.62) for $\lambda = -0.5$; average = 0.56 (target = 0.36) for $\lambda = -1$.

In Figure 4, we plot the lower and upper bounds, the average of the bounds, and the target value of $C_{\text{WT}}(\lambda, \alpha = 2)$ for selected values of λ. The patterns we observe in this figure are similar to those of Figure 3. For example, choosing $x = -\lambda - 0.5$ for $\lambda > 0$ leads to the relative error less than 16%: average = 2.68 (target = 2.86) for $\lambda = 0.25$; average = 4.12 (target = 4.13) for $\lambda = 0.5$; average = 7.98 (target = 9.14) for $\lambda = 1$. Likewise, choosing $x = -0.25$ for $\lambda < 0$ leads to the relative error less than 48%: average = 1.33 (target = 1.41) for $\lambda = -0.25$; average = 1.20 (target = 1.00) for $\lambda = -0.5$; average = 0.94 (target = 0.50) for $\lambda = -1$.

Finally, one conclusion is evident from Figures 3–4: the average-based estimates of $C_{\text{WT}}(\lambda)$ and $C_{\text{WT}}(\lambda, \alpha = 2)$ perform worst when λ is "extreme" (i.e., $\lambda = \pm 1$). Because λ reflects the user's risk appetite and is chosen subjectively, a practical

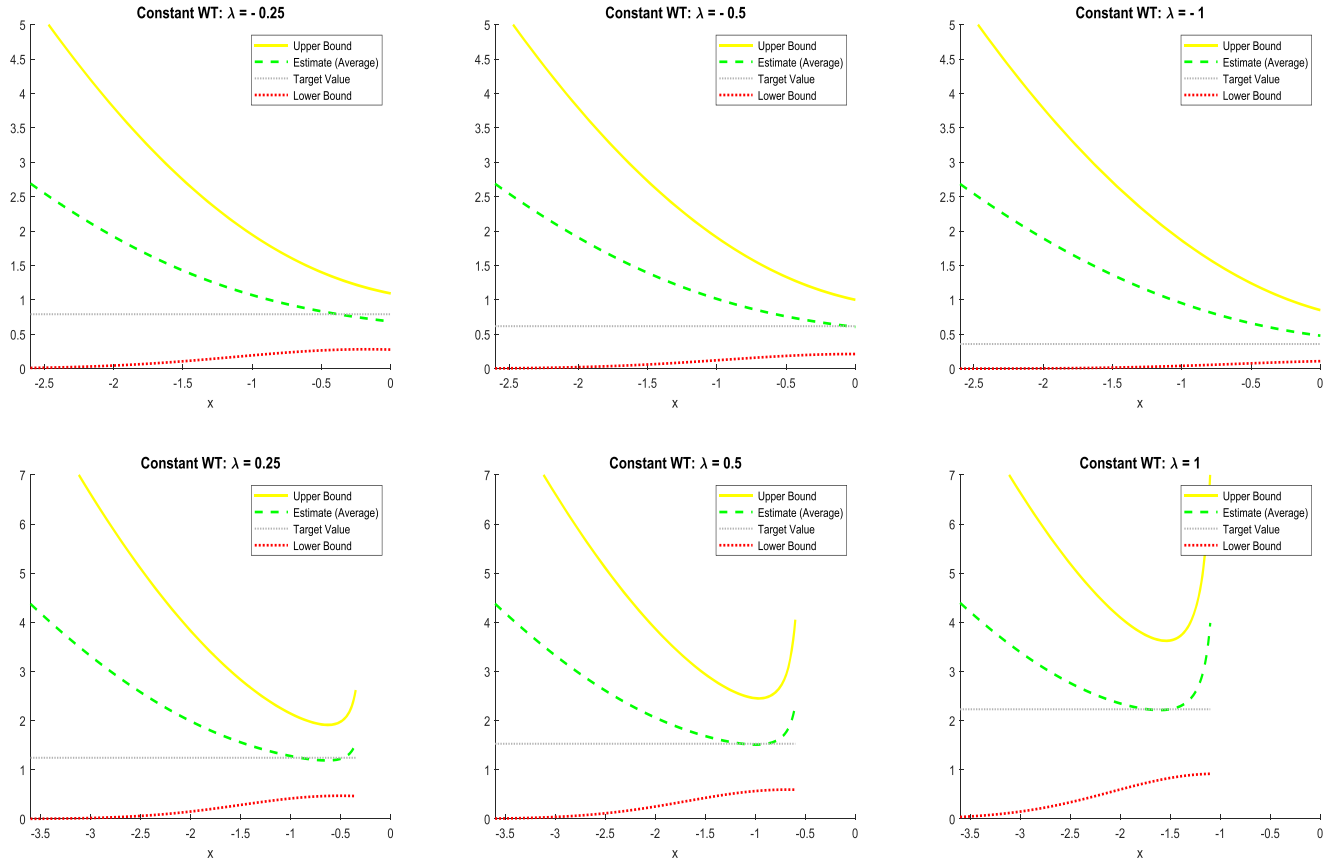

FIGURE 3. Lower and Upper Bounds, Estimate (Average of the Bounds), and the Target Value of Constant $C_{\mathrm{WT}}(\lambda)$ for (a) $\lambda = -1, -0.5, -0.25$ and (b) $\lambda = 0.25, 0.5, 1$.

remedy to resolve this issue is to employ WT measures with less extreme values of the distortion parameter. For example, consider limiting λ to the interval $[-0.5; 0.5]$.

4.3. Risk Measure Estimation

In this section, we use Monte Carlo simulations to investigate the effectiveness of risk measure approximations in statistical estimation. In particular, synthetic datasets are generated according to the distributions (F) of Section 4.1, and then for each dataset the PHT$[F, r]$ measure is estimated using nonparametric and parametric methods. At the end, the simulated PHT estimates are summarized using few standard statistics: mean, median, lower and upper quartiles (q25, q75), and root mean square error ($\sqrt{\mathrm{mse}}$). The estimates are also compared with the theoretical targets. Note that the PHT measure has explicit formulas for the exponential and Pareto I models (see Table 1), but for the lognormal distribution it has to be approximated by the average of the lower and upper bounds of Theorem 1. Below we specify the key components of the simulation study design:

- *Study parameters, estimator characteristics, and risk measure*: $M = 10^5$ (number of Monte Carlo samples), $n = 100$ (size of each sample), $a = b = 0.05$ (trimming/winsorizing proportions of MTM and MWM estimators; see below), PHT$[F, r = 0.75]$ (risk measure to be estimated from data).
- *Simulated data*: $X_1, \ldots, X_n$ are independent and identically distributed (i.i.d.) realizations of random variable X that is distributed according to F. The choice of F includes: $F_1 = \mathcal{E}xp\ (x_0 = 1, \theta = 0.9391)$, $F_2 = \mathcal{P}a\ I\ (x_0 = 1, \alpha = 2)$, and $F_3 = \mathcal{LN}\ (x_0 = 1, \mu = -0.1571, \sigma = 0.7243)$.
- *Nonparametric estimator of* PHT: The research area of nonparametric (empirical) estimation of distortion risk measures was initiated by Jones and Zitikis (2003). According to the results of section 3 in that paper, the empirical estimator of PHT is found by replacing F with the empirical cdf $\hat{F}_n$ in the formula of PHT$[F, r]$, which is (2.6), and is given by

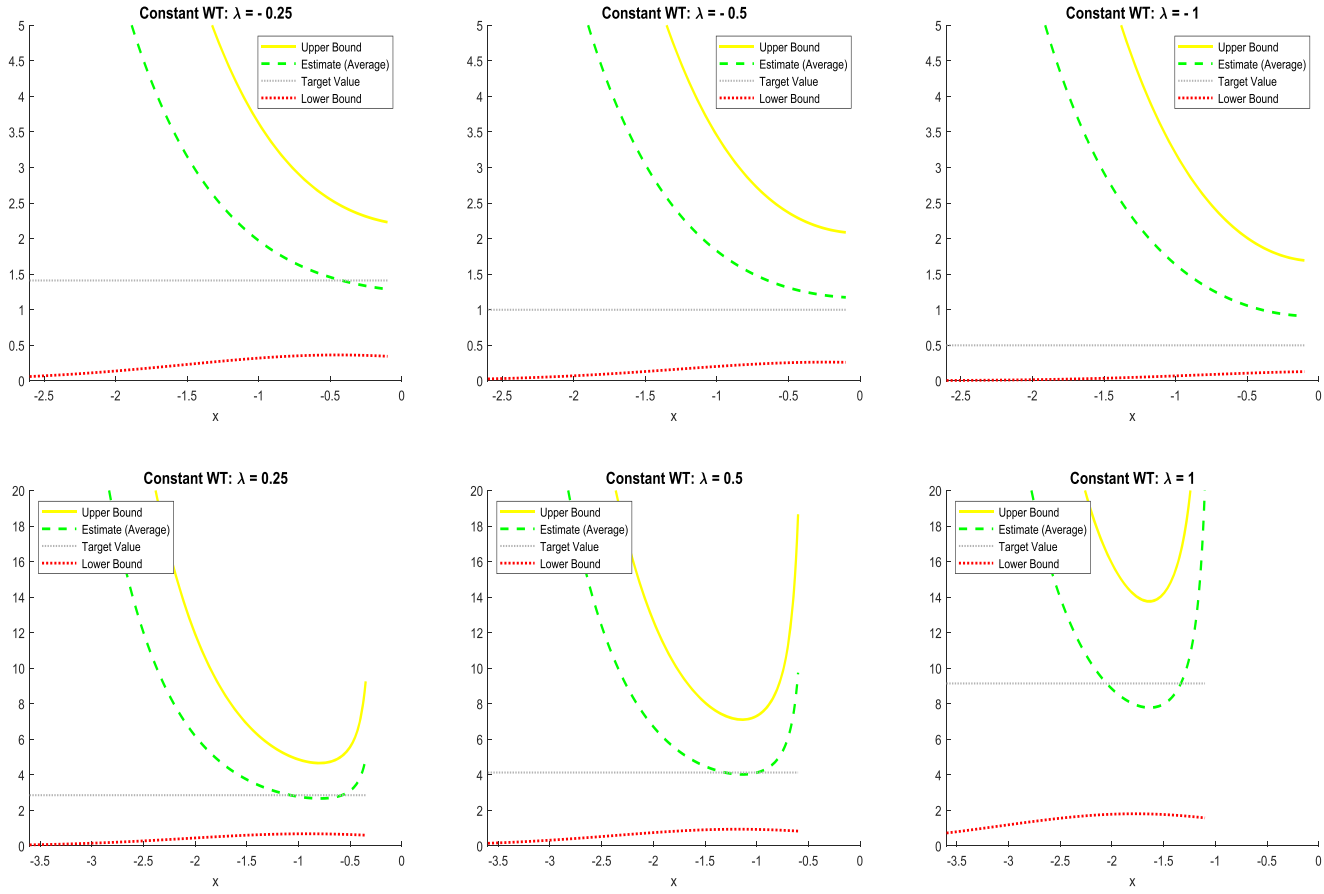

FIGURE 4. Lower and Upper Bounds, Estimate (Average of the Bounds), and the Target Value of Constant $C_{\mathrm{WT}}(\lambda, \alpha = 2)$ for (a) $\lambda = -1, -0.5, -0.25$ and (b) $\lambda = 0.25, 0.5, 1$.

$$\widehat{\mathrm{PHT}}_{\mathrm{EMP}} \equiv \mathrm{PHT}\left[\hat{F}_n, r\right] = \sum_{i=1}^{n} X_{(i)}\left((1-(i-1)/n)^r - (1-i/n)^r\right),$$

where $X_{(1)} \leq \cdots \leq X_{(n)}$ denote the ordered values of $X_1, ..., X_n$.

- *Parametric estimators of* PHT: Parametric estimators of PHT are found by first estimating the parameters of F and then plugging them in the (explicit or approximate) formula of PHT. We consider three classes of parameter estimators: *maximum likelihood estimator* (MLE), *method of trimmed moments* (MTM), and *method of winsorized moments* (MWM). For the distributions selected in this study, the formulas of MLEs are well known, and the MTM and MWM estimators were developed by Brazauskas, Jones, and Zitikis (2009) and Zhao, Brazauskas, and Ghorai (2018), respectively. Table 5 provides MLE, MTM, and MWM formulas for the unknown parameters of F_1, F_2, and F_3. (In all of the cases under consideration, parameter x_0 ($x_0 = 1$) is treated as a known threshold for the smallest possible loss and hence does not require estimation.)
- *Exponential distribution*, $\mathcal{E}xp\ (x_0, \theta)$:

$$\hat{\theta}_{\mathrm{MLE}} = \bar{X}_*, \qquad \hat{\theta}_{\mathrm{MTM}} = \frac{\hat{T}_1(a,b)}{I_t(a,b)}, \qquad \hat{\theta}_{\mathrm{MWM}} = \frac{\hat{W}_1(a,b)}{I_w(a,b)},$$

where $\bar{X}_* = n^{-1}\sum_{i=1}^{n}(X_i - x_0)$, $I_t(a,b) = (1-a-b)^{-1}\int_a^{1-b}(-\log(1-u))\,du$, $\hat{T}_1(a,b) = (n-[na]-[nb])^{-1}$ $\sum_{i=[na]+1}^{n-[nb]}(X_{(i)} - x_0)$, $I_w(a,b) = -a\log(1-a) + (1-a-b)I_t(a,b) - b\log(b)$, $\hat{W}_1(a,b) = n^{-1}([na](X_{([na]+1)} - x_0) +$ $\sum_{i=[na]+1}^{n-[nb]}(X_{(i)} - x_0) + [nb](X_{(n-[nb])} - x_0))$, and $[\cdot]$ denotes the greatest integer part. Note that $\bar{X}_* = \hat{T}_1(0,0) = \hat{W}_1(0,0)$ and $1 = I_t(0,0) = \lim_{b\to 0} I_w(0,b)$ resulting in $\hat{\theta}_{\mathrm{MLE}} = \hat{\theta}_{\mathrm{MTM}} = \hat{\theta}_{\mathrm{MWM}}$.

TABLE 5
Theoretical Targets (PHT Values) of the Exponential, Pareto I, and Lognormal Losses and Summary Statistics of the PHT Estimates based on EMP, MTM, MWM, and MLE

Distribution of losses	Theoretical target	Method of estimation	Summary Statistics				
			Mean	Median	q25	q75	$\sqrt{\text{mse}}$
Exponential (F_1)	2.252	EMP	2.233	2.229	2.146	2.316	0.129
		MTM	2.259	2.255	2.168	2.346	0.131
		MWM	2.253	2.249	2.164	2.337	0.128
		MLE	2.253	2.249	2.166	2.335	0.125
Pareto I (F_2)	3.000	EMP	2.652	2.467	2.235	2.806	1.194
		MTM	3.209	3.015	2.646	3.526	1.056
		MWM	3.162	2.986	2.629	3.469	0.899
		MLE	3.150	2.985	2.639	3.456	0.826
Lognormal (F_3)	2.430	EMP	2.401	2.389	2.302	2.487	0.146
		MTM	2.507	2.495	2.381	2.618	0.195
		MWM	2.490	2.479	2.370	2.597	0.181
		MLE	2.490	2.480	2.377	2.592	0.173

- *Pareto I distribution*, $\mathcal{P}a\ I\ (x_0, \alpha)$:

$$\hat{\alpha}_{\text{MLE}} = \frac{n}{\sum_{i=1}^{n} \log(X_i/x_0)}, \qquad \hat{\alpha}_{\text{MTM}} = \frac{I_t(a,b)}{\hat{T}_2(a,b)}, \qquad \hat{\alpha}_{\text{MWM}} = \frac{I_w(a,b)}{\hat{W}_2(a,b)},$$

  where $\hat{T}_2(a,b) = (n-[na]-[nb])^{-1} \sum_{i=[na]+1}^{n-[nb]} \log(X_{(i)}/x_0)$, $I_t(a,b)$ and $I_w(a,b)$ are as before, and $\hat{W}_2(a,b) = n^{-1}([na]\log(X_{([na]+1)}/x_0) + \sum_{i=[na]+1}^{n-[nb]} \log(X_{(i)}/x_0) + [nb]\log(X_{(n-[nb])}/x_0))$. Note that the choice $a = b = 0$ also results in $\hat{\alpha}_{\text{MLE}} = \hat{\alpha}_{\text{MTM}} = \hat{\alpha}_{\text{MWM}}$.
- *Lognormal distribution*, $\mathcal{LN}\ (x_0, \mu, \sigma)$:

$$\hat{\mu}_{\text{MLE}} = n^{-1}\sum_{i=1}^{n} \log(X_i - x_0), \qquad \hat{\sigma}_{\text{MLE}} = \left(n^{-1}\sum_{i=1}^{n} (\log(X_i - x_0) - \hat{\mu}_{\text{MLE}})^2 \right)^{1/2};$$

$$\hat{\mu}_{\text{MTM}} = \hat{T}_{3,1}(a,b) - I_{t,1}(a,b)\cdot\hat{\sigma}_{\text{MTM}}, \qquad \hat{\sigma}_{\text{MTM}} = \left(\frac{\hat{T}_{3,2}(a,b) - \hat{T}_{3,1}^2(a,b)}{I_{t,2}(a,b) - I_{t,1}^2(a,b)} \right)^{1/2};$$

$$\hat{\mu}_{\text{MWM}} = \hat{W}_{3,1}(a,b) - I_{w,1}(a,b)\cdot\hat{\sigma}_{\text{MWM}}, \qquad \hat{\sigma}_{\text{MWM}} = \left(\frac{\hat{W}_{3,2}(a,b) - \hat{W}_{3,1}^2(a,b)}{I_{w,2}(a,b) - I_{w,1}^2(a,b)} \right)^{1/2}.$$

  Here for $k=1,\ 2$: $\hat{T}_{3,k}(a,b) = (n-[na]-[nb])^{-1} \sum_{i=[na]+1}^{n-[nb]} \log^k(X_{(i)}-x_0)$; $I_{t,k}(a,b) = (1-a-b)^{-1} \int_a^{1-b} (\Phi^{-1}(u))^k\, du$; $I_{w,k}(a,b) = -a(\Phi^{-1}(a))^k + (1-a-b)I_{t,k}(a,b) - b(\Phi^{-1}(1-b))^k$; $\hat{W}_{3,k}(a,b) = n^{-1}([na]\log^k(X_{([na]+1)}-x_0) + \sum_{i=[na]+1}^{n-[nb]} \log^k(X_{(i)}-x_0) + [nb]\log^k(X_{(n-[nb])}-x_0))$. As before, the choice $a = b = 0$ yields $\hat{\mu}_{\text{MLE}} = \hat{\mu}_{\text{MTM}} = \hat{\mu}_{\text{MWM}}$ and $\hat{\sigma}_{\text{MLE}} = \hat{\sigma}_{\text{MTM}} = \hat{\sigma}_{\text{MWM}}$.
- *Theoretical targets*: $\text{PHT}[F_1, r = 0.75] = 2.252$ (exponential losses), $\text{PHT}[F_2, r = 0.75] = 3.000$ (Pareto I losses), $\text{PHT}[F_3, r = 0.75] = 2.430$ (lognormal losses).

Results of the simulation study are sumarized in Table 5. Several conclusions emerge from the table. First, the means of all estimates are fairly close—within one $\sqrt{\text{mse}}$—to their respective targets (e.g., for F_1 and MTM: target = 2.252, mean = 2.259, $\sqrt{\text{mse}} = 0.131$). Looking from a different perspective, one can see that almost all medians, except for that of EMP in the Pareto I model, are close to the target value, which falls within the interval [q25, q75]. Second, the EMP estimator performs poorly; that is,

has large bias ($-0.348 = 2.652-3.000$) and significantly inflated root mean square error ($\sqrt{\text{mse}} = 1.194$), for the Pareto I model, which generates heavy-tailed losses. Even its upper quartile is below the target value ($\text{q75} = 2.806$). In this case, the parametric methods offer much more accurate performance. Third, for the lognormal distribution, EMP has the smallest bias ($= -0.029$) and root mean square error ($\sqrt{\text{mse}} = 0.146$) because MTM, MWM, and MLE of PHT are approximated using the bounds of Theorem 1. Nonetheless, their overall performance is comparable to that based on the explicit formula estimators under the exponential and Pareto I models. Finally, we should note that estimation based on the proposed approximation formulas had no significant effect on computing times. (The entire simulation study takes about 40 s to run on a basic laptop computer).

5. PROOFS

5.1. Proof of Theorem 1

Fix $x > 0$ and split the range of integration into $(-\infty; x)$ and $(x; \infty)$:

$$\begin{aligned} C_{\text{PHT}}(r,\sigma) &= \sigma \int_{-\infty}^{x} (1-\Phi(z))^r \; e^{\sigma z}\, dz \;\; + \sigma \int_{x}^{\infty} (1-\Phi(z))^r \; e^{\sigma z}\, dz \\ &=: I_{1,x}(r,\sigma) \;\; + \;\; I_{2,x}(r,\sigma). \end{aligned}$$

The lower and upper bounds for $I_{1,x}(r,\sigma)$ follow by noticing that $(1-\Phi(x))^r \le (1-\Phi(z))^r \le 1$ for $z \le x$. That is,

$$I_{1,x}(r,\sigma) \;\; \ge \;\; \sigma \;\; (1-\Phi(x))^r \int_{-\infty}^{x} e^{\sigma z}\, dz \;\; = \;\; e^{\sigma x} \;\; (1-\Phi(x))^r$$

and

$$I_{1,x}(r,\sigma) \;\; \le \;\; \sigma \int_{-\infty}^{x} e^{\sigma z}\, dz \;\; = \;\; e^{\sigma x}.$$

Therefore,

$$e^{\sigma x} \;\; (1-\Phi(x))^r \;\; \le \;\; I_{1,x}(r,\sigma) \;\; \le \;\; e^{\sigma x}. \tag{5.1}$$

To establish the lower bound for the term $I_{2,x}(r,\sigma)$, note that $(1-\Phi(z))^r$ is decreasing in r and thus $(1-\Phi(z))^r \ge 1-\Phi(z)$ for $0 < r \le 1$. Now, first use integration by parts, then the fact that $\lim_{z\to\infty}(1-\Phi(z)) \; e^{\sigma z} = 0$, and finish with straighforward integration:

$$\begin{aligned} I_{2,x}(r,\sigma) \ge \sigma \int_{x}^{\infty} (1-\Phi(z)) \; e^{\sigma z}\, dz \;\; &= \;\; -e^{\sigma x}(1-\Phi(x)) + \int_{x}^{\infty} e^{\sigma z}\varphi(z)\, dz \\ &= -e^{\sigma x}(1-\Phi(x)) \;\; + \;\; e^{\sigma^2/2} \int_{x}^{\infty} \varphi(z-\sigma)\, dz \\ &= -e^{\sigma x}(1-\Phi(x)) \;\; + \;\; e^{\sigma^2/2}\Phi(\sigma-x). \end{aligned}$$

For the upper bound of $I_{2,x}(r,\sigma)$, we first apply Lemma B.1(a) (see Appendix B) and then $z^{-r} \le x^{-r}$ for $z \ge x$:

$$I_{2,x}(r,\sigma) \;\; < \;\; \sigma \int_{x}^{\infty} \left(\frac{1}{z}\varphi(z)\right)^r \; e^{\sigma z}\, dz \;\; \le \;\; \sigma x^{-r} \int_{x}^{\infty} (\varphi(z))^r \; e^{\sigma z}\, dz.$$

And the remaining steps are straightforward (but messy) integration:

$$\begin{aligned} \sigma x^{-r} \int_{x}^{\infty} (\varphi(z))^r \; e^{\sigma z}\, dz &= \sigma x^{-r} r^{-1/2} (2\pi)^{(1-r)/2} e^{\sigma^2/(2r)} \int_{x}^{\infty} \varphi\left(\frac{z-\sigma r^{-1}}{\sqrt{r^{-1}}}\right) dz \\ &= \sigma x^{-r} r^{-1/2} (2\pi)^{(1-r)/2} e^{\sigma^2/(2r)} \Phi((\sigma - rx)/\sqrt{r}) \;\; = \;\; K_x(r,\sigma). \end{aligned}$$

Therefore,

$$-e^{\sigma x}(1-\Phi(x)) \;+\; e^{\sigma^2/2}\;\Phi(\sigma-x) \;\le\; I_{2,x}(r,\sigma) \;<\; K_x(r,\sigma). \tag{5.2}$$

Now, adding (5.1) and (5.2) yields the statement of the theorem. □

5.2. Proof of Theorem 2

Fix $x < 0$ and split the range of integration into $(-\infty; x)$ and $(x; \infty)$:

$$C_{\mathrm{WT}}(\lambda) \;=\; \int_{-\infty}^{x} \Phi(z+\lambda)\;\frac{\varphi(z)}{\Phi(z)}\,dz \;+\int_{x}^{\infty} \Phi(z+\lambda)\;\frac{\varphi(z)}{\Phi(z)}\,dz \;=:\; I_{1,x}(\lambda)+I_{2,x}(\lambda).$$

Start with the term $I_{2,x}(\lambda),$ which, after integration by parts, becomes

$$I_{2,x}(\lambda) \;=\; -\Phi(x+\lambda)\log[\Phi(x)]-\int_{x}^{\infty} \log[\Phi(z)]\;\varphi(z+\lambda)\,dz.$$

Notice that for $z \ge x,$ we have $\log[\Phi(x)] \le \log[\Phi(z)] \le 0$ and therefore

$$-\Phi(x+\lambda)\log[\Phi(x)] \;\le\; I_{2,x}(\lambda) \;\le\; -\log[\Phi(x)]. \tag{5.3}$$

Next, it is clear from its definition that $I_{1,x}(\lambda) \ge 0,$ but to find an upper bound for it, the sign of λ has to be taken into consideration. Thus, for $\lambda \le 0,$ Lemma B.2(a) (see Appendix B) leads to

$$0 \;\le\; I_{1,x}(\lambda) \;\le\; \int_{-\infty}^{x} \left[e^{-\lambda z-\lambda^2/2}\;\Phi(z)\right]\frac{\varphi(z)}{\Phi(z)}\,dz \;=\; \int_{-\infty}^{x} \varphi(z+\lambda)\,dz \;=\; \Phi(x+\lambda). \tag{5.4}$$

Adding (5.3) and (5.4) proves the double inequality in (a).

For $\lambda > 0,$ we apply Lemma B.2(b) (note that the upper bound is valid for $x < -\lambda$) and then $z/(z+\lambda) \le x/(x+\lambda)$ for $z \le x < -\lambda$:

$$\begin{aligned} 0 \;&\le\; I_{1,x}(\lambda) < \textstyle\int_{-\infty}^{x} \left[\frac{z}{z+\lambda}\;\; e^{-\lambda z-\lambda^2/2}\;\Phi(z)\right]\frac{\varphi(z)}{\Phi(z)}\,dz \\ &\le \frac{x}{x+\lambda}\int_{-\infty}^{x} \varphi(z+\lambda)\,dz \;=\; \frac{x}{x+\lambda}\;\Phi(x+\lambda). \end{aligned} \tag{5.5}$$

Adding (5.3) and (5.5) proves the double inequality in (b). □

5.3. Proof of Theorem 3

Fix $x < 0$ and split the range of integration into $(-\infty; x)$ and $(x; \infty)$:

$$\begin{aligned} C_{\mathrm{WT}}(\lambda,\alpha) &= \int_{-\infty}^{x} \Phi(z+\lambda)\;\frac{\varphi(z)}{[\Phi(z)]^{1/\alpha+1}}\,dz \;+\int_{x}^{\infty} \Phi(z+\lambda)\;\frac{\varphi(z)}{[\Phi(z)]^{1/\alpha+1}}\,dz \\ &=: I_{1,x}(\lambda,\alpha) \;+\; I_{2,x}(\lambda,\alpha). \end{aligned}$$

Start with the term $I_{2,x}(\lambda,\alpha),$ which, after integration by parts, becomes

$$I_{2,x}(\lambda,\alpha) \;=\; \alpha\left[\Phi(x+\lambda)[\Phi(x)]^{-1/\alpha}-1+\int_{x}^{\infty} [\Phi(z)]^{-1/\alpha}\;\varphi(z+\lambda)\,dz\right].$$

Note that for $z \geq x$, we have $1 \leq [\Phi(z)]^{-1/\alpha} \leq [\Phi(x)]^{-1/\alpha}$ and, after straighforward simplifications,

$$\Phi(x+\lambda)\ c_x(\alpha) \ \leq \ I_{2,x}(\lambda,\alpha) \ \leq \ c_x(\alpha). \tag{5.6}$$

Next, it is clear from its definition that $I_{1,x}(\lambda,\alpha) \geq 0$, but to find an upper bound for it, the sign of λ has to be taken into account. Thus, for $\lambda \leq 0$, first Lemma B.2(a) and then integration by parts (with the condition $\alpha > 1$) lead to

$$\begin{aligned} 0 \ \leq \ I_{1,x}(\lambda,\alpha) &\leq \int_{-\infty}^{x} \left[e^{-\lambda z-\lambda^2/2}\ \Phi(z)\right] \frac{\varphi(z)}{[\Phi(z)]^{1/\alpha+1}}\, dz \\ &= \frac{\alpha\ e^{-\lambda^2/2}}{\alpha-1}\left[e^{-\lambda x}[\Phi(x)]^{1-1/\alpha} + \lambda\int_{-\infty}^{x}\left[\Phi(z)\right]^{1-1/\alpha}\ e^{-\lambda z}\, dz\right]. \end{aligned}$$

For $z \leq x$, we have $0 \leq [\Phi(z)]^{1-1/\alpha} \leq [\Phi(x)]^{1-1/\alpha}$, and because $\lambda \leq 0$,

$$0 \ \leq \ I_{1,x}(\lambda,\alpha) \ \leq \ \frac{\alpha}{\alpha-1}\ e^{-\lambda x-\lambda^2/2}\ [\Phi(x)]^{1-1/\alpha}. \tag{5.7}$$

Adding (5.6) and (5.7) proves the double inequality in (a).

For $\lambda > 0$, we apply Lemma B.2(b) (note that the upper bound is valid for $x < -\lambda$), $z/(z+\lambda) \leq x/(x+\lambda)$ for $z \leq x < -\lambda$, and then integration by parts:

$$\begin{aligned} 0 \ \leq \ I_{1,x}(\lambda,\alpha) &< \int_{-\infty}^{x}\left[\frac{z}{z+\lambda}\ e^{-\lambda z-\lambda^2/2}\ \Phi(z)\right]\frac{\varphi(z)}{[\Phi(z)]^{1/\alpha+1}}\, dz \\ &\leq \frac{x\ e^{-\lambda^2/2}}{x+\lambda}\int_{-\infty}^{x} e^{-\lambda z}\ \frac{\varphi(z)}{[\Phi(z)]^{1/\alpha}}\, dz \\ &= \frac{\alpha}{\alpha-1}\ \frac{x\ e^{-\lambda^2/2}}{x+\lambda}\left[e^{-\lambda x}[\Phi(x)]^{1-1/\alpha} + \lambda\int_{-\infty}^{x}\left[\Phi(z)\right]^{1-1/\alpha}\ e^{-\lambda z}\, dz\right], \end{aligned} \tag{5.8}$$

where the last step involves the following computation (assuming $\alpha > 1$):

$$\begin{aligned} \lim_{z\to-\infty}\frac{[\Phi(z)]^{1-1/\alpha}}{e^{\lambda z}} &= \left(\lim_{z\to-\infty}\frac{\Phi(z)}{e^{\lambda z(\alpha/(\alpha-1))}}\right)^{1-1/\alpha} = \left(\lim_{z\to-\infty}\frac{\varphi(z)}{(\lambda\alpha/(\alpha-1))e^{\lambda z(\alpha/(\alpha-1))}}\right)^{1-1/\alpha} \\ &= \text{constant} \times \left(\lim_{z\to-\infty}\varphi(z+\lambda\alpha/(\alpha-1))\right)^{1-1/\alpha} = \ 0. \end{aligned}$$

Further, we continue (5.8) by first applying Lemma B.1(b), then $-z^{-1} \leq -x^{-1}$ for $z \leq x < -\lambda$, and finishing with straightforward (but messy) integration:

$$\begin{aligned} 0 \ \leq \ I_{1,x}(\lambda,\alpha) &< \frac{\alpha}{\alpha-1}\ \frac{x\ e^{-\lambda^2/2}}{x+\lambda}\left[e^{-\lambda x}[\Phi(x)]^{1-1/\alpha} + \lambda\int_{-\infty}^{x}\left[\frac{\varphi(z)}{-z}\right]^{1-1/\alpha} e^{-\lambda z}\, dz\right] \\ &\leq \frac{\alpha}{\alpha-1}\ \frac{x\ e^{-\lambda^2/2}}{x+\lambda}\left[e^{-\lambda x}[\Phi(x)]^{1-1/\alpha} + \lambda(-x)^{1-1/\alpha}\int_{-\infty}^{x}[\varphi(z)]^{1-1/\alpha}\ e^{-\lambda z}\, dz\right] \\ &= \frac{\alpha}{\alpha-1}\ \frac{x}{x+\lambda}\left[e^{-\lambda x-\lambda^2/2}\ [\Phi(x)]^{1-1/\alpha} + C_x(\alpha,\lambda)\right], \end{aligned} \tag{5.9}$$

where $C_x(\alpha,\lambda) = \lambda\ \sqrt{\alpha(2\pi)^{1/\alpha}/(\alpha-1)}\ (-x)^{1/\alpha-1}\ e^{-(\lambda^2/2)/(\alpha-1)}\ \Phi\left(\frac{x+\lambda\alpha/(\alpha-1)}{\sqrt{\alpha/(\alpha-1)}}\right)$. Adding (5.6) and (5.9) proves the double inequality in (b). □

5.4. Proof of Theorem 4

Notice that $\Phi^{-1}(1-\beta) \le z \le \infty$ yields $1-\beta \le \Phi(z) \le 1$. This leads to

$$(1-\beta)\int_{\Phi^{-1}(1-\beta)}^{\infty} \varphi(z-\sigma)\,dz \;\le\; C_{\mathrm{GS}}(\beta,\sigma) \;\le\; \int_{\Phi^{-1}(1-\beta)}^{\infty} \varphi(z-\sigma)\,dz.$$

Because $\int_{\Phi^{-1}(1-\beta)}^{\infty} \varphi(z-\sigma)\,dz = \Phi(\sigma - \Phi^{-1}(1-\beta))$, the statement of the theorem follows. □

6. CONCLUDING REMARKS

In this article, we have considered five commonly used distortion risk measures—VAR, CTE, PHT, WT, and GS—and three right-skewed distributions with identical support: shifted exponential, Pareto I, and shifted lognormal. Analytic and numerical evaluations of these risk measures are the focus of the article. When analytic treatment is feasible, explicit formulas of the risk measures have been presented. When it is not, lower and upper bounds for analytically intractable integrals that appear in the formulas of PHT (for shifted lognormal loss), WT (for shifted exponential and Pareto I losses), and GS (for shifted lognormal loss) risk measures have been established. In addition, the integrals have been evaluated numerically and the quality of the analytic bounds has been investigated.

The research presented in this article invites follow-up studies in several directions. First, within the classes of lower and upper bounds for the integrals of Section 3 (see Theorems 1–3), one could identify the optimal split point of the integration range. The smallest difference between the bounds could be used as an optimality criterion. Alternatively, a search for tighter bounds based on different inequalities for the standard normal distribution tails could be pursued. The ultimate goal of such improvements is to accurately approximate those integrals so that their computation within large-scale simulation studies becomes automatic. Second, it is certainly of interest to expand the list of loss severity models to other popular probability distributions such as gamma, Weibull, generalized Pareto, as well as other types of Pareto (see Klugman, Panjer, and Willmot [2012], appendix A, for an inventory of continuous distributions used in actuarial studies). Or, to evaluate these risk measures for recently proposed spliced and folded distributions, see Scollnik (2007, 2014), Scollnik and Sun (2012), Brazauskas and Kleefeld (2011, 2014, 2016), Nadarajah and Bakar (2015). We believe the techniques established in Section 5 would be useful for computing and estimating various risk measures under these other distributions as well.

In addition, to capture the riskiness of aggregate losses, approximations based on central limit theorems would lead to a normal distribution. Thus, the risk measure formulas for normal distributions are also needed. However, one can anticipate that due to the direct relationship between the lognormal and normal distributions, similar formulas for both distributions should emerge. Also, in this context, typical generalizations of the normal distribution include Student's t and elliptical distributions. The GS measures for these and related skew distributions were derived by Furman, Wang, and Zitikis (2017).

ACKNOWLEDGMENTS

The authors are very appreciative of valuable insights and useful comments provided by two anonymous referees, which helped to substantially improve the article. Also, much of this work was completed while the first author was a Ph.D. student in the Department of Mathematical Sciences at the University of Wisconsin–Milwaukee.

ORCID

Vytaras Brazauskas http://orcid.org/0000-0003-2161-053X

REFERENCES

Albrecht, P. 2004. Risk measures. In *Encyclopedia of actuarial science*, ed. B. Sundt and J. Teugels, vol. 3, 1493–501. London: Wiley.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath, D. 1999. Coherent measures of risk. *Mathematical Finance* 9 (3):203–28.

Balbás, A., J. Garrido, and S. Mayoral. 2009. Properties of distortion risk measures. *Methodology and Computing in Applied Probability* 11 (3):385–99.

Belles-Sampera, J., M. Guillén, and M. Santolino. 2014. Beyond Value-at-Risk: GlueVaR distortion risk measures. *Risk Analysis* 34 (1):121–34.

Belles-Sampera, J., M. Guillén, and M. Santolino. 2016. The use of flexible quantile-based measures in risk assessment. *Communications in Statistics: Theory and Methods* 45 (6):1670–81.

Brazauskas, V., B. Jones, and R. Zitikis. 2009. Robust fitting of claim severity distributions and the method of trimmed moments. *Journal of Statistical Planning and Inference* 139 (6):2028–43.

Brazauskas, V., and A. Kleefeld. 2011. Folded and log-folded-t distributions as models for insurance loss data. *Scandinavian Actuarial Journal* 2011 (1): 59–79.
Brazauskas, V., and A. Kleefeld. 2014. Authors' reply to "Letter to the Editor: Regarding folded models and the paper by Brazauskas and Kleefeld (2011)" by Scollnik. *Scandinavian Actuarial Journal* 2014 (8):753–57.
Brazauskas, V., and A. Kleefeld. 2016. Modeling severity and measuring tail risk of Norwegian fire claims. *North American Actuarial Journal* 20 (1):1–16.
Brazauskas, V., and T. Kaiser. 2004. Discussion of "Empirical estimation of risk measures and related quantities" by Jones and Zitikis. *North American Actuarial Journal*, 8 (3):114–17.
Feller, W. 1968. *An introduction to probability theory and its applications*. Vol. 1, 3rd ed. New York: Wiley.
Feng, R., and H. W. Volkmer. 2012. Analytical calculation of risk measures for variable annuity guaranteed benefits. *Insurance: Mathematics and Economics* 51 (3):636–48.
Feng, R., and H. W. Volkmer. 2014. Spectral methods for the calculation of risk measures for variable annuity guaranteed benefits. *ASTIN Bulletin* 44 (3): 653–81.
Feng, R., X. Jing, and J. Dhaene. 2017. Comonotonic approximations of risk measures for variable annuity guaranteed benefits with dynamic policyholder behavior. *Journal of Computational and Applied Mathematics* 311:272–92.
Furman, E., and Z. Landsman. 2006. Tail variance premium with applications for elliptical portfolio of risks. *ASTIN Bulletin* 36 (2):433–62.
Furman, E., R. Wang, and R. Zitikis. 2017. Gini-type measures of risk and variability: Gini shortfall, capital allocations, and heavy-tailed risks. *Journal of Banking and Finance* 83:70–84.
Johnson, N. L., S. Kotz, and N. Balakrishnan. 1994. *Continuous univariate distributions*. Vol. 1, 2nd ed. New York: Wiley.
Jones, B. L., and R. Zitikis. 2003. Empirical estimation of risk measures and related quantities (with discussion). *North American Actuarial Journal* 7 (4): 44–54.
Jones, B., and R. Zitikis. 2004. Authors' reply to Discussion of "Empirical estimation of risk measures and related quantities" by Brazauskas and Kaiser. *North American Actuarial Journal*, 8 (3):117–18.
Klugman, S. A., H. H. Panjer, and G. E. Willmot. 2012. *Loss models: From data to decisions*. 4th ed. New York: Wiley.
Nadarajah, S., and S. A. A. Bakar. 2015. New folded models for the log-transformed Norwegian fire claim data. *Communications in statistics: Theory and methods* 44 (20):4408–40.
Scollnik, D. P. M. 2007. On composite lognormal-Pareto models. *Scandinavian Actuarial Journal* 2007 (1):20–33.
Scollnik, D. P. M. 2014. Letter to the Editor: Regarding folded models and the paper by Brazauskas and Kleefeld (2011). *Scandinavian Actuarial Journal* 2014 (3):278–81.
Scollnik, D. P. M., and C. Sun. 2012. Modeling with Weibull-Pareto models. *North American Actuarial Journal* 16 (2):260–72.
Tapiero, C. S. 2004. Risk management: An interdisciplinary framework. In *Encyclopedia of actuarial science*, ed. B. Sundt and J. Teugels, vol. 3, 1483–93. London: Wiley.
Wang, Q., R. Wang, and R. Zitikis. 2021. Risk measures induced by efficient insurance contracts. *Insurance: Mathematics and Economics* 103:56–65.
Wang, R., and R. Zitikis. 2021. An axiomatic foundation for the expected shortfall. *Management Science* 67 (3):1413–29.
Wang, S. 1995. Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics* 17 (1):43–54.
Wang, S. 2000. A class of distortion operators for pricing financial and insurance risks. *Journal of Risk and Insurance* 67 (1):15–36.
Wang, S. 2002. A universal framework for pricing financial and insurance risks. *ASTIN Bulletin* 32 (2):213–34.
Young, V. R. 2004. Premium principles. In *Encyclopedia of actuarial science*, ed. B. Sundt and J. Teugels, vol. 3, 1322–31. London: Wiley.
Zhao, Q., V. Brazauskas, and J. Ghorai. 2018. Robust and efficient fitting of severity models and the method of winsorized moments. *ASTIN Bulletin* 48 (1): 275–309.

Discussions on this article can be submitted until October 1, 2023. The authors reserve the right to reply to any discussion. Please see the Instructions for Authors found online at http://www.tandfonline.com/uaaj for submission instructions.

APPENDIX A. PARAMETRIC MODELS

In this appendix, we present probability density function (pdf), cumulative distribution function (cdf), and quantile function (qf) for three continuous non-negative random variables X: shifted exponential, Pareto I, and shifted lognormal. Note that X represents so-called ground up losses, which are of great interest in product design (e.g., for specifying insurance contract parameters or for choosing loss retention levels in reinsurance) as well as for other business decisions.

There are many probability distributions used to model claim severity and new ones being actively developed. The proposed models have varying numbers of parameters (and thus varying levels of flexibility) and different degrees of tail heaviness. In this article, the models chosen to illustrate the risk measuring concepts and theoretical results are typical and mathematically tractable, plus they are designed to have identical supports.

A.1. Shifted Exponential Distribution

Let us assume that random variable X is distributed according to a shifted exponential distribution with a location (shift) parameter $x_0 > 0$ and scale parameter $\theta > 0$. We will denote this fact as $X \sim \mathcal{E}xp\ (x_0, \theta)$. As is well known (see Jonhson, Kotz, and Balakrishnan, 1994, ch. 19), the pdf, cdf, and qf of X are:

$$\begin{aligned} pdf: \quad f(x) &= \theta^{-1} e^{-(x-x_0)/\theta}, \qquad x \geq x_0, \\ cdf: \quad F(x) &= 1-e^{-(x-x_0)/\theta}, \qquad x \geq x_0, \\ qf: \quad F^{-1}(p) &= x_0 - \theta \log(1-p), \qquad 0 \leq p \leq 1. \end{aligned}$$

We assume x_0 is a known parameter representing the smallest possible loss.

A.2. Pareto I Distribution

Let random variable X be distributed according to a Pareto I distribution with a scale parameter $x_0 > 0$ and shape parameter $\alpha > 0$. We will denote this fact as $X \sim \mathcal{P}a\ I\ (x_0, \alpha)$. As is well known (see Jonhson, Kotz, and Balakrishnan, 1994, ch. 20), the pdf, cdf, and qf of X are

$$\begin{aligned} pdf: \quad f(x) &= (\alpha/x_0)(x_0/x)^{\alpha+1}, \qquad x \geq x_0, \\ cdf: \quad F(x) &= 1-(x_0/x)^{\alpha}, \qquad x \geq x_0, \\ qf: \quad F^{-1}(p) &= x_0(1-p)^{-1/\alpha}, \qquad 0 \leq p \leq 1. \end{aligned}$$

As in Section A.1, x_0 is assumed to be a known parameter representing the smallest possible loss.

A.3. Shifted Lognormal Distribution

Suppose random variable X is distributed according to a shifted lognormal distribution with a location (shift) parameter $x_0 > 0$, log-location $-\infty < \mu < \infty$, and log-scale parameter $\sigma > 0$. We will denote this fact as $X \sim \mathcal{LN}(x_0, \mu, \sigma)$. As is well known (see Jonhson, Kotz, and Balakrishnan, 1994, ch. 14), X is related to a normal random variable, and its pdf, cdf, and qf are

$$\begin{aligned} pdf: \quad f(x) &= (\sigma(x-x_0))^{-1}\ \varphi\left(\frac{\log(x-x_0)-\mu}{\sigma}\right), \qquad x \geq x_0, \\ cdf: \quad F(x) &= \Phi\left(\frac{\log(x-x_0)-\mu}{\sigma}\right), \qquad x \geq x_0, \\ qf: \quad F^{-1}(p) &= x_0 + \exp\{\mu + \sigma\Phi^{-1}(p)\}, \qquad 0 \leq p \leq 1. \end{aligned}$$

Here Φ, φ, Φ^{-1} denote the cdf, pdf, and qf of the standard normal distribution, respectively. Also, similar to Sections A.1 and A.2, x_0 is a known parameter representing the smallest possible loss.

APPENDIX B. LEMMAS

To establish some of the inequalities of Section 3, it is convenient to have an approximation of the normal distribution tails as $x \to \pm\infty$. The following lemma proves simple relationships between cdf Φ and pdf φ in the upper and lower tails.

Lemma B.1. (Normal distribution tails—I)

a. *As* $x \to \infty$,

$$1-\Phi(x) \approx x^{-1}\varphi(x);$$

more precisely, the double inequality

$$(x^{-1} - x^{-3})\varphi(x) < 1-\Phi(x) < x^{-1}\varphi(x)$$

holds for every x > 0.

As $x \to -\infty$,

$$\Phi(x) \approx -x^{-1}\varphi(x);$$

more precisely, the double inequality

$$(x^{-3} - x^{-1})\varphi(x) < \Phi(x) < -x^{-1}\varphi(x)$$

holds for every x $<$ 0.

Proof. Part (a) is stated and proved in Feller (1968, pp. 175 and 179).

For part (b), we start by noticing that $1-t^{-2} < 1 < 1 + 3t^{-4}$; equivalently,

$$(1 - t^{-2})\varphi(t) < \varphi(t) < (1 + 3t^{-4})\varphi(t), \qquad t < 0.$$

Integrating each term of the inequality from $-\infty$ to x leads to

$$2\Phi(x) + x^{-1}\varphi(x) < \Phi(x) < 2\Phi(x) + (x^{-1} - x^{-3})\varphi(x).$$

Straightforward simplifications yield the stated inequality. □

Inequalities involving standard normal cdf at different points are used in proving Theorems 1–4.

Lemma B.2. (Normal distribution tails—II)

a. *For $\lambda \leq 0$, the inequality*

$$\Phi(z + \lambda) \leq e^{-\lambda z - \lambda^2/2}\ \Phi(z)$$

holds for every z $<$ 0.

b. *For $\lambda > 0$, the inequality*

$$\Phi(z + \lambda) < \frac{z}{z + \lambda}\ e^{-\lambda z - \lambda^2/2}\ \Phi(z)$$

holds for every $z < -\lambda$.

Proof. The standard normal cdf $\Phi(z + \lambda)$ can be rewritten as follows:

$$\begin{aligned} \Phi(z + \lambda) = \int_{-\infty}^{z+\lambda} \varphi(t)\, dt &= \int_{-\infty}^{z} \varphi(v + \lambda)\, dv \\ &= e^{-\lambda^2/2} \int_{-\infty}^{z} \varphi(v)\ e^{-\lambda v}\, dv. \end{aligned} \tag{B.1}$$

In part (a), $\lambda \leq 0$ implies that $e^{-\lambda v}$ is a nondecreasing function. Thus, for $v \leq z$, we have $e^{-\lambda v} \leq e^{-\lambda z}$. Combining this inequality with (B.1), we get

$$\Phi(z + \lambda) \leq e^{-\lambda^2/2} e^{-\lambda z} \int_{-\infty}^{z} \varphi(v)\, dv = e^{-\lambda z - \lambda^2/2}\ \Phi(z).$$

For part (b), first use integration by parts in (6.1), then apply the upper bound from Lemma B.1(b), notice that $-v^{-1} \leq -z^{-1}$ when $v \leq z < 0$, and complete the integration:

$$\begin{aligned}
\Phi(z+\lambda) &= e^{-\lambda^2/2}\left[\Phi(z)\ e^{-\lambda z} + \lambda\int_{-\infty}^{z}\Phi(v)\ e^{-\lambda v}\,dv\right] \\
&< e^{-\lambda^2/2}\left[\Phi(z)\ e^{-\lambda z} + \lambda\int_{-\infty}^{z}(-v^{-1})\ \varphi(v)\ e^{-\lambda v}\,dv\right] \\
&\le e^{-\lambda z-\lambda^2/2}\ \Phi(z) - \frac{\lambda}{z}\ e^{-\lambda^2/2}\int_{-\infty}^{z}\varphi(v)\ e^{-\lambda v}\,dv \\
&= e^{-\lambda z-\lambda^2/2}\ \Phi(z) - \frac{\lambda}{z}\ \int_{-\infty}^{z}\varphi(v+\lambda)\,dv \\
&= e^{-\lambda z-\lambda^2/2}\ \Phi(z) - \frac{\lambda}{z}\ \Phi(z+\lambda).
\end{aligned}$$

Finally, we rearrange this inequality into

$$\Phi(z+\lambda)\ <\ \frac{z}{z+\lambda}\ e^{-\lambda z-\lambda^2/2}\ \Phi(z)$$

and notice that the rearrangement is valid when $z/(z+\lambda)>0$ or $z<-\lambda<0$. □