August 2020

Estimating Distortion Risk Measures Under Truncated and Censored Data Scenarios

Sahadeb Upretee
University of Wisconsin-Milwaukee

ESTIMATING DISTORTION RISK MEASURES UNDER TRUNCATED AND CENSORED DATA SCENARIOS

by

Sahadeb Upretee

A Dissertation Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
in Mathematics

at

The University of Wisconsin-Milwaukee
August 2020

ABSTRACT

ESTIMATING DISTORTION RISK MEASURES UNDER TRUNCATED AND CENSORED DATA SCENARIOS

by

Sahadeb Upretee

The University of Wisconsin-Milwaukee, 2020
Under the Supervision of Professor Vytaras Brazauskas

In insurance data analytics and actuarial practice, a broad class of risk measures – *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. When solving such problems, the main statistical challenge is to choose an appropriate estimate of a risk measure and to assess its variability. In this context, the empirical nonparametric approach is the simplest one to use, but it lacks efficiency due to the scarcity of data in the tails. On the other hand, parametric estimators, although prone to model misspecification, can improve estimators' efficiency significantly. Moreover, they can easily accommodate data truncation and censoring that are common features of insurance loss data.

The first objective of this dissertation is to derive the asymptotic distributions of empirical and parametric estimators of distortion risk measures under the truncated and censored data scenarios. For parametric estimation, we use maximum likelihood (ML) and percentile matching (PM) procedures. The risk measures we consider include: *value-at-risk* (VaR), *conditional tail expectation* (CTE), *proportional hazards transform* (PHT), *Wang transform* (WT), and *Gini shortfall* (GS). Conditions under which these measures are finite are studied rigorously. The ML and PM estimators of the risk measures are derived for three severity models (with identical support): shifted exponential, Pareto I, and shifted lognormal. Their asymptotic properties are established and compared with those of the empirical estimators. Then, the second objective of the dissertation

is to cross-validate and augment the theoretical results using simulations. Finally, the third objective is to provide a few numerical examples involving applications of the new estimators to actual reinsurance data.

To

Father: Devi Prasad Upretee,

Mother: Bhadrakumari Upretee,

Father-in-law: Krishna Prasad Adhikari,

Mother-in-law: Ramkumari Adhikari,

Wife: Suna Sharma,

Son: Saharsha Upretee

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to start off by expressing my sincere gratitude to my advisor Professor Vytaras Brazauskas for his patient guidance through all these years. The quality of this paper would have been unattainable without his immense contribution. Professor Vytaras' guidance, expert advice, and suggestions were invaluable and can hardly be quantified.

I feel obliged for the advice and constructive comments from my dissertation research committee members Professors Daniel Gervini, Chao Zhu, Wei Wei, and David Spade not only for serving on the dissertation committee but also for supporting and mentoring me throughout the development of this dissertation.

I am extremely in awe of Professor Yang Ho for mentoring me during my study and sharpening my theoretical knowledge. His insightful comments on my research work tremendously increased my quality of content and presentation.

A special thankyou goes to my wife, Suna Sharma, who has encouraged me to go above and beyond and constantly kept me focused and motivated. I cannot convey by words how appreciative I am for her serenity, perpetual inspiration and emotional support.

I would also like to thank my parents, siblings, and my son, Saharsha Upretee, without whom it would be impossible for me to be where I am now.

My sincere gratitude goes to faculty members in the Department of Mathematical Sciences for their advice and administrative support in ensuring I completed this dissertation in time.

Last but not the least, I would like to remember all my friends who directly and indirectly helped me while writing my dissertation.

Sahadeb Upretee

Milwaukee, Wisconsin

August 2020

Chapter 1

Introduction

1.1 Motivation

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 most common distortion risk measures include: *value-at-risk* (VaR), *conditional tail expectation* (CTE), *proportional hazards transform* (PHT), *Wang transform* (WT), and *Gini shortfall* (GS). In this dissertation, these five risk measures will be studied extensively.

Point and interval estimation of and hypothesis testing based on risk measures are very important practical problems used by insurance companies for making business decisions. They also play a significant role within the regulatory frameworks of the financial sector (e.g., *Solvency II* – the European Union insurance regulation; *ORSA*

– the U.S. insurance solvency framework; *Basel III* – regulation for the inernational banking sector). The main statistical challenge in solving various business problems is to choose an appropriate estimate of a risk measure and to assess its variability. The empirical nonparametric approach is used often because it is simple to implement and easy to understand. This approach, however, lacks efficiency due to the scarcity of data in the tails. On the other hand, parametric estimators, although prone to model misspecification, can improve estimators' efficiency significantly. Moreover, they can easily accommodate data truncation and censoring that are common features of insurance loss data. In this dissertation, we will focus on estimation of distortion risk measures under truncated and censored data scenarios.

1.2 Literature Review

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, the reader may be referred to Albrecht (2004), Tapiero (2004), and Young (2004). The development of practice-inspired risk measures was substantially advanced by a series of ground-breaking contributions made by Wang (1995, 1998a,b, 2000, 2002). However, statistical inference problems did not receive much attention or rigorous treatment until the appearance of the paper on empirical estimation of risk measures by Jones and Zitikis (2003). This paper gave momentum to the literature on risk measure estimation and testing when the underlying risk is continuous and data are fully observed (see Brazauskas and Kaiser, 2004; Jones and Zitikis, 2005, 2007; Jones *et al.*, 2006; Kaiser and Brazauskas, 2006; Brazauskas *et al.*, 2007, 2008; Necir *et al.*, 2007, 2010; Necir and Meraghni, 2009; Samanthi *et al.*, 2017).

The primary goal of this dissertation is to develop statistical inferential tools for distortion risk measures when the loss variable is continuous but data are only partially observed. In particular, we deal with typical insurance data scenarios when loss variable is affected by left truncation (due to deductibles) or right censoring (due to policy limits).

While several authors have investigated such data modifications for the ratemaking and portfolio risk retention problems (see Frees, 2017; Lee, 2017), extensions of these techniques to the class of distortion risk measures has not been considered in the actuarial literature.

1.3 Plan of the Thesis

Focusing on distortion risk measures in general and VaR, CTE, PHT, WT, and GS in particular, we present risk measure formulas when the loss variable follows *shifted exponential*, *Pareto I*, and *shifted lognormal* distributions. In some of these cases, the risk measures do not have closed form expressions, thus we specify conditions (i.e., restrictions on model parameters) under which the risk measure is finite. Further, empirical and parametric (based on maximum likelihood and percentile matching) estimation of the risk measures is treated and asymptotically normal distributions of the estimators are established. Finally, theoretical results are cross-validated by performing a small-scale simulation study and then complemented with real data illustrations.

The rest of the dissertation is organized as follows. In Chapter 2, we introduce a left-truncated and right-censored loss variable and define its probability density function, cumulative distribution function, and quantile function. Then, specific expressions of these functions are derived when the (ground up) loss variable follows the shifted exponential, Pareto I, and shifted lognormal distribution.

In Chapter 3, the class of distortion risk measures is introduced, the five specific risk measures (VaR, CTE, PHT, WT, and GS) are defined and dicussed from the perspective of risk measure *coherence*. Then, their formulas are derived for shifted exponential, Pareto I, and shifted lognormal distributions. Moreover, when the risk measure has no closed form expression (which happens in several cases), a class of lower and upper bounds is established by first proving a few probability inequalities involving the tails of the standard normal distribution. The rationale for having such bounds is to eventually construct risk measure approximations that could further be used in simulations.

In Chapter 4, we first present empirical and parametric estimators for distortion risk measures (the general case) and then specify asymptotically normal distributions of those estimators. Having the general results ready, we work out a series of analytic examples for the chosen risk measures and loss severity distributions.

Chapter 5 is devoted to numerical illustrations. First, we perform a small-scale simulation study to cross-validate the asymptotic results of Chapter 4. Second, we fit Pareto I and lognormal distributions to the well-known Norwegian fire claims data and evaluate the quality of model fits via quantile-quantile plots. Then, the fitted models and selected risk measures are used to estimate the upper-tail riskiness of these claims.

Finally, in Chapter 6, the main results of the dissertation are summarized and future research directions are discussed.

A note on notation: throughout the dissertation we will use 'log' to denote the natural logarithm.

Chapter 2

Loss Data and Models

In this chapter, we first introduce notation of probability density function (pdf), cumulative distribution function (cdf), and quantile function (qf) for continuous non-negative random variable X. Then, to account for typical transformations of insurance loss data, we modify the pdf, cdf, and qf of X when this variable is left-truncated and right-censored. The chapter ends with presentation of the pdf, cdf, and qf of *shifted exponential*, *Pareto I*, and *shifted lognormal* models under left truncation and right censoring of data. Note that these are typical and mathematically tractable loss severity distributions, plus they are designed to have identical supports.

2.1 Truncated and Censored Data

Suppose random variables

$$X_1,\, X_2, \ldots, X_N \tag{2.1}$$

are *independent and identically distributed* (*i.i.d.*) and have the pdf $f(x)$, cdf $F(x)$, and qf $F^{-1}(s)$. Since random variables corresponding to insurance loss are non-negative, the support of $f(x)$ is the set $\{x : x \geq 0\}$. Note that $X_1, \ldots, X_N$ represent so-called "ground up" losses. They 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.

In insurance practice, however, the underlying loss variable gets transformed due to coverage modifications such as deductibles and upper policy limits. Specifically, if the insurance contract has ordinary deductible d and policy limit u ($u > d$), then we actually observe a random sample of *mixed discrete-continuous* variables, $X_1^*, \ldots, X_n^*$, that satisfy the following conditional event relationship:

$$X_i^* \stackrel{d}{=} \left\{ \min\big(X_i, u\big) \,\big|\, X_i > d \right\}, \qquad i = 1, \ldots, n, \tag{2.2}$$

where $\stackrel{d}{=}$ denotes "equal in distribution." Also, the mixed pdf/pmf f_*, cdf F_*, qf F_*^{-1} of X^* are related to f, F, F^{-1} and given by:

$$f_*(x) = \begin{cases} \frac{f(x)}{1-F(d)}, & d < x < u, \\ \frac{1-F(u)}{1-F(d)}, & x = u, \\ 0, & \text{otherwise}, \end{cases} \tag{2.3}$$

$$F_*(x) = \begin{cases} 0, & x \leq d, \\ \frac{F(x)-F(d)}{1-F(d)}, & d < x < u, \\ 1, & x \geq u, \end{cases} \tag{2.4}$$

$$F_*^{-1}(p) = \begin{cases} F^{-1}\big(p + \big(1-p\big)F(d)\big), & 0 \leq p < \frac{F(u)-F(d)}{1-F(d)}, \\ u, & \frac{F(u)-F(d)}{1-F(d)} \leq p \leq 1. \end{cases} \tag{2.5}$$

Three special cases follow from equations (2.3)–(2.5), which we list in the following notes.

Note 2.1. [Left-Truncated Variable]

If $u \to \infty$, then X^* in (2.2) becomes a left-truncated variable at d, with cdf F_*, mixed pdf/pmf f_*, and qf F_*^{-1} given by

$$F_*(x) = \frac{F(x) - F(d)}{1 - F(d)} \mathbf{1}\{x > d\}, \qquad f_*(x) = \frac{f(x)}{1 - F(d)} \mathbf{1}\{x > d\},$$

$$F_*^{-1}(p) = F^{-1}\left(p + \big(1 - p\big)F(d)\right), \qquad 0 \leq p \leq 1,$$

where the indicator function $\mathbf{1}\{x > d\} = 1$ if $x > d$, and $= 0$, otherwise. □

Note 2.2. [Right-Censored Variable]
If $d = 0$, then X^* in (2.2) becomes a right-censored variable at u, with cdf F_*, mixed pdf/pmf f_*, and qf F_*^{-1} given by

$$F_*(x) = F(x)\,\mathbf{1}\{0 < x < u\} + \mathbf{1}\{x \geq u\}, \qquad f_*(x) = f(x)\,\mathbf{1}\{0 < x < u\} + \big[1 - F(u)\big]\,\mathbf{1}\{x = u\},$$

$$F_*^{-1}(p) = F^{-1}(p)\,\mathbf{1}\{0 \leq p < F(u)\} + u\,\mathbf{1}\{F(u) \leq p \leq 1\},$$

where $\mathbf{1}\{\}$ is the indicator function. □

Note 2.3. [Ground-Up Loss]
When $d = 0$ and $u \to \infty$, then X^* in (2.2) becomes a ground-up loss variable X (which is non-negative). In such a case, its respective cdf, mixed pdf/pmf, and qf are given by

$$F_*(x) = F(x), \qquad f_*(x) = f(x), \qquad F_*^{-1}(p) = F^{-1}(p),$$

where $0 \leq p \leq 1$, and $F(x) > 0$ and $f(x) > 0$ when $x > 0$. □

2.2 Severity Distributions

There are many probability distributions used to model claim severity, and new ones being actively developed. The proposed models have varying number of parameters (and thus varying levels of flexibility) and different degrees of tail heaviness. In this dissertation, we choose to illustrate the risk measuring concepts and theoretical results with three standard and mathematically tractable distributions that share the same support: shifted exponential, Pareto I, and shifted lognormal.

2.2.1 Shifted Exponential Distribution

Suppose 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 *et al.*, 1994, Chapter 19), the pdf, cdf, and qf of X are:

$$
\begin{array}{lrcll}
\text{PDF:} & f(x) & = & \theta^{-1} e^{-(x-x_0)/\theta}, & x \geq x_0, \\
\text{CDF:} & F(x) & = & 1 - e^{-(x-x_0)/\theta}, & x \geq x_0, \\
\text{QF:} & F^{-1}(p) & = & x_0 - \theta \log(1-p), & 0 \leq p \leq 1.
\end{array}
$$

We assume x_0 is a known parameter representing the smallest possible loss (e.g., one dollar).

For $X \sim \mathcal{E}xp\,(x_0, \theta)$, we have $[F(x) - F(d)]/[1 - F(d)] = 1 - e^{-(x-d)/\theta}$ for $d < x < u$ (note that $d > x_0$). Substitution of this expression in (2.4) yields the cdf of the left-truncated and right-censored variable X^* (defined by (2.2)):

$$
F_*(x) = \begin{cases} 0, & x \leq d, \\ 1 - e^{-(x-d)/\theta}, & d < x < u, \\ 1, & x \geq u. \end{cases} \tag{2.6}
$$

Further, $f(x)/[1 - F(d)] = \theta^{-1} e^{-(x-d)/\theta}$ for $d < x < u$ and $[1 - F(u)]/[1 - F(d)] = e^{-(u-d)/\theta}$ for $x = u$; substitution of these expressions in (2.3) yields the mixed pdf/pmf of X^*:

$$
f_*(x) = \begin{cases} \theta^{-1} e^{-(x-d)/\theta}, & d < x < u, \\ e^{-(u-d)/\theta}, & x = u, \\ 0, & \text{otherwise}. \end{cases} \tag{2.7}
$$

Finally, $F^{-1}\big(p + \big(1-p\big)F(d)\big) = -\theta \log(1-p) + d$ for $0 < p < 1 - e^{-(u-d)/\theta}$; substitution of this expression in (2.5) yields the qf of X^*:

$$F_*^{-1}(p) = \begin{cases} -\theta \log(1-p) + d, & 0 \le p < 1 - e^{-(u-d)/\theta}, \\ u, & 1 - e^{-(u-d)/\theta} \le p \le 1. \end{cases} \tag{2.8}$$

2.2.2 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 *et al.*, 1994, Chapter 20), the pdf, cdf, and qf of X are:

$$\begin{array}{lrcll} \text{PDF:} & f(x) & = & (\alpha/x_0)\,(x_0/x)^{\alpha+1}\,, & x \ge x_0, \\ \text{CDF:} & F(x) & = & 1 - (x_0/x)^{\alpha}\,, & x \ge x_0, \\ \text{QF:} & F^{-1}(p) & = & x_0(1-p)^{-1/\alpha}, & 0 \le p \le 1. \end{array}$$

As in Section 2.2.1, x_0 is assumed to be a known parameter representing the smallest possible loss.

For $X \sim \mathcal{P}a\, I\,(x_0, \alpha)$, we have $[F(x) - F(d)]/[1 - F(d)] = 1 - (d/x)^{\alpha}$ for $d < x < u$ (note that $d > x_0$). Substitution of this expression in (2.4) yields the cdf of the left-truncated and right-censored variable X^* (defined by (2.2)):

$$F_*(x) = \begin{cases} 0, & x \le d, \\ 1 - (d/x)^{\alpha}\,, & d < x < u, \\ 1, & x \ge u. \end{cases} \tag{2.9}$$

Further, $f(x)/[1 - F(d)] = (\alpha/d)\,(d/x)^{\alpha+1}$ for $d < x < u$ and $[1 - F(u)]/[1 - F(d)] = (d/u)^{\alpha}$ for $x = u$; substitution of these expressions in (2.3) yields the mixed pdf/pmf of

X^*:

$$f_*(x) = \begin{cases} (\alpha/d)\,(d/x)^{\alpha+1}\,, & d < x < u, \\ (d/u)^{\alpha}, & x = u, \\ 0, & \text{otherwise.} \end{cases} \tag{2.10}$$

Finally, $F^{-1}\big(p + (1-p)F(d)\big) = d(1-p)^{-1/\alpha}$ for $0 \leq p < 1 - (d/u)^{\alpha}$; substitution of this expression in (2.5) yields the qf of X^*:

$$F_*^{-1}(p) = \begin{cases} d(1-p)^{-1/\alpha}, & 0 \leq p < 1 - (d/u)^{\alpha}\,, \\ u, & 1 - (d/u)^{\alpha} \leq p \leq 1. \end{cases} \tag{2.11}$$

2.2.3 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 *et al.*, 1994, Chapter 14), X is related to a normal random variable, and its pdf, cdf, and qf are:

$$\begin{array}{rrcll} \text{PDF:} & f(x) & = & (\sigma(x-x_0))^{-1}\,\varphi\left(\dfrac{\log(x-x_0)-\mu}{\sigma}\right), & x \geq x_0, \\ \text{CDF:} & F(x) & = & \Phi\left(\dfrac{\log(x-x_0)-\mu}{\sigma}\right), & x \geq x_0, \\ \text{QF:} & F^{-1}(p) & = & x_0 + \exp\{\mu + \sigma\Phi^{-1}(p)\}, & 0 \leq p \leq 1. \end{array}$$

Here Φ, φ, Φ^{-1} denote the cdf, pdf, qf of the standard normal distribution, respectively. Also, similar to Sections 2.2.1-2.2.2, x_0 is a known parameter representing the smallest possible loss.

Next, let us first introduce the following abbreviations:

$$c_d := \frac{\log(d-x_0)-\mu}{\sigma}, \qquad c_x := \frac{\log(x-x_0)-\mu}{\sigma}, \qquad c_u := \frac{\log(u-x_0)-\mu}{\sigma}.$$

Now, for $X \sim \mathcal{LN}(x_0, \mu, \sigma)$, we have $[F(x)-F(d)]/[1-F(d)] = [\Phi(c_x)-\Phi(c_d)]/[1-\Phi(c_d)]$ for $d < x < u$ (note that $d > x_0$), where c_x and c_d are defined above. Substitution of these expressions in (2.4) yields the cdf of X^* (defined by (2.2)):

$$F_*(x) = \begin{cases} 0, & x \leq d, \\ \dfrac{\Phi(c_x) - \Phi(c_d)}{1 - \Phi(c_d)}, & d < x < u, \\ 1, & x \geq u. \end{cases} \tag{2.12}$$

Further, $f(x)/[1 - F(d)] = (\sigma(x - x_0))^{-1} \varphi(c_x)/[1 - \Phi(c_d)]$ for $d < x < u$ and $[1 - F(u)]/[1 - F(d)] = [1 - \Phi(c_u)]/[1 - \Phi(c_d)]$ for $x = u$. Substitution of these expressions in (2.3) yields the mixed pdf/pmf of X^*:

$$f_*(x) = \begin{cases} \dfrac{(\sigma(x - x_0))^{-1} \varphi(c_x)}{1 - \Phi(c_d)}, & d < x < u, \\ \dfrac{1 - \Phi(c_u)}{1 - \Phi(c_d)}, & x = u, \\ 0, & \text{otherwise}. \end{cases} \tag{2.13}$$

Finally, $F^{-1}\big(p + (1-p)F(d)\big) = x_0 + \exp\left\{\mu + \sigma\Phi^{-1}\big(p + (1-p)\Phi(c_d)\big)\right\}$ for $0 \leq p < p_u$, where $p_u = [\Phi(c_u) - \Phi(c_d)]/[1 - \Phi(c_d)]$. Substitution of these expressions in (2.5) yields the qf of X^*:

$$F_*^{-1}(p) = \begin{cases} x_0 + \exp\left\{\mu + \sigma\Phi^{-1}\Big(p + (1-p)\Phi(c_d)\Big)\right\}, & 0 \leq p < p_u, \\ u, & p_u \leq p \leq 1. \end{cases} \tag{2.14}$$

Chapter 3

Distortion Risk Measures

In this chapter, we present a broad class of risk measures – distortion risk measures – and briefly discuss the concept of risk measure coherence. Then, we define several popular distortion measures: *value-at-risk* (VaR), *conditional tail expectation* (CTE), *proportional hazards transform* (PHT), *Wang transform* (WT), and *Gini shortfall* (GS). Finally, specific formulas of these measures are derived when claim severities follow shifted exponential, Pareto I, and shifted lognormal distributions. In several instances the risk measure formulas involve integrals that are analitically intractable. In those cases, we prove that the integrals are bounded and then evaluate them numerically.

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 cdf F, a risk measure R is defined as

$$R[F] \;=\; \int_0^\infty g(1 - F(x))\, dx, \tag{3.1}$$

where the *distortion function* $g\colon [0,1] \to [0,1]$ is an increasing function with $g(0) = 0$ and $g(1) = 1$. Moreover, if g is differentiable, then integration by parts in (3.1) leads to

$$R[F] \;=\; \int_0^1 F^{-1}(u)\psi(u)\, du, \tag{3.2}$$

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. *Scale invariance*: $\varrho[bX_1] = b\varrho[X_1]$, where b is a positive constant.

3. *Subadditivity*: $\varrho[X_1 + X_2] \leq \varrho[X_1] + \varrho[X_2]$.

4. *Monotonicity*: If $\mathbf{P}\{X_1 \leq X_2\} = 1$, then $\varrho[X_1] \leq \varrho[X_2]$.

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.

3.1 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:

$$\mathrm{VaR}[F, \beta] \; = \; F^{-1}(1 - \beta). \tag{3.3}$$

Note that VaR can be expressed as distortion risk measure, defined by (3.1), by choosing $g(u) = 0$ for $0 \leq u < \beta$, and $= 1$ for $\beta \leq u \leq 1$. These choices correspond to $g(1-F(x)) = 0$ for $0 \leq 1-F(x) < \beta$, and $= 1$ for $\beta \leq 1-F(x) \leq 1$, or equivalently $g(1-F(x)) = 0$ for $F^{-1}(1-\beta) < x < \infty$, and $= 1$ for $0 \leq x \leq F^{-1}(1-\beta)$. Now expression (3.3) follows easily from (3.1):

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

This risk measure, however, is *not coherent* as it does not satisfy the subadditivity property (it does satisfy the other three properties though). To see that, let us consider the standard uniform random variable $U \sim \mathit{Uniform}\,(0,1)$ and define two loss variables: $X_1 = 100 \cdot \mathbf{1}\big\{U \leq 0.09\big\}$ and $X_2 = 100 \cdot \mathbf{1}\big\{U > 0.91\big\}$, where $\mathbf{1}\big\{\big\}$ denotes the indicator function. Let us also denote cdf's of X_1 and X_2 as F_{X_1} and F_{X_2}, respectively, and the cdf of their sum as $F_{X_1+X_2}$. Clearly, the chance of zero loss is 91% for both variables; thus $\mathrm{VaR}[F_{X_1}, 0.10] = \mathrm{VaR}[F_{X_2}, 0.10] = 0$. On the other hand, the chance of zero loss for their sum is 82%. Thus, $\mathrm{VaR}[F_{X_1+X_2}, 0.10] = 100$, which implies that $\mathrm{VaR}[F_{X_1+X_2}, 0.10] \;\nleq\; \mathrm{VaR}[F_{X_1}, 0.10] + \mathrm{VaR}[F_{X_2}, 0.10]$.

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. The following examples present VaR for the severity distributions of Section 2.2.

Example 3.1. [VaR of Shifted Exponential]
If $X \sim \mathcal{E}xp\,(x_0, \theta)$, then the VaR measure of X is its qf (defined in Section 2.2.1):

$$\mathrm{VaR}[F,\beta] \;=\; F^{-1}(1-\beta) \;=\; x_0 - \theta \log(\beta),$$

where β $(0 < \beta < 1)$ represents the confidence level; also known as the "risk appetite".

□

Example 3.2. [VaR of PARETO I]

If $X \sim \mathcal{P}a\, I\,(x_0, \alpha)$, then the VaR measure of X is its qf (defined in Section 2.2.2):

$$\mathrm{VaR}[F, \beta] \;=\; F^{-1}(1-\beta) \;=\; x_0 \beta^{-1/\alpha},$$

where β $(0 < \beta < 1)$ represents the confidence level or risk appetite. □

Example 3.3. [VaR of SHIFTED LOGNORMAL]

If $X \sim \mathcal{LN}\,(x_0, \mu, \sigma)$, then the VaR measure of X is its qf (defined in Section 2.2.3):

$$\mathrm{VaR}[F, \beta] \;=\; F^{-1}(1-\beta) \;=\; x_0 + \exp\{\mu + \sigma \Phi^{-1}(1-\beta)\},$$

where β $(0 < \beta < 1)$ represents the risk appetite. □

3.2 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:

$$\begin{aligned} \mathrm{CTE}[F, \beta] \;&=\; F^{-1}(1-\beta) + \frac{1}{\beta} \int_{F^{-1}(1-\beta)}^{\infty} \big[1 - F(x)\big]\, dx \qquad (3.4) \\ &=\; \frac{1}{\beta} \int_{1-\beta}^{1} F^{-1}(u)\, du. \qquad (3.5) \end{aligned}$$

It is clear from (3.4) that this measure can be expressed as (3.1) by choosing $g(t) = t/\beta$ for $0 \le t < \beta$, and $= 1$ for $\beta \le t \le 1$. Alternatively, expression (3.5) follows from (3.2) 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 (3.4) with (3.3) 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.

The following examples present CTE for the severity distributions of Section 2.2.

Example 3.4. [CTE of Shifted Exponential]
If $X \sim \mathcal{E}xp\,(x_0, \theta)$, then its CTE measure is found by integrating the qf of X (defined in Section 2.2.1) over the interval $[1-\beta; 1]$ and then dividing it by β:

$$\begin{aligned} \mathrm{CTE}[F, \beta] &= \frac{1}{\beta}\int_{1-\beta}^{1} F^{-1}(u)\,du \;=\; \frac{1}{\beta}\int_{1-\beta}^{1} \big[x_0 - \theta\log(1-u)\big]\,du \\ &= x_0 - \theta(\log(\beta) - 1), \end{aligned}$$

where β $(0 < \beta < 1)$ is the risk appetite. (For more integration details, see Appendix A.) $\square$

Example 3.5. [CTE of Pareto I]
If $X \sim \mathcal{P}a\,I\,(x_0, \alpha)$, then its CTE measure is found by integrating the qf of X (defined in Section 2.2.2) over the interval $[1-\beta; 1]$ and then dividing it by β:

$$\begin{aligned} \mathrm{CTE}[F, \beta] &= \frac{1}{\beta}\int_{1-\beta}^{1} F^{-1}(u)\,du \;=\; \frac{1}{\beta}\int_{1-\beta}^{1} \big[x_0(1-u)^{-1/\alpha}\big]\,du \\ &= x_0\beta^{-1/\alpha}\alpha(\alpha-1)^{-1}, \qquad \alpha > 1, \end{aligned}$$

where β $(0 < \beta < 1)$ is the risk appetite. (For more integration details, see Appendix A.) Note that CTE is infinite when Pareto distribution has very heavy upper tail (i.e., when $\alpha \leq 1$). $\square$

Example 3.6. [CTE of Shifted Lognormal]
If $X \sim \mathcal{LN}\,(x_0, \mu, \sigma)$, then its CTE measure is found by integrating the qf of X (defined in Section 2.2.3) over the interval $[1-\beta; 1]$ and then dividing it by β:

$$\begin{aligned} \mathrm{CTE}[F, \beta] &= \frac{1}{\beta}\int_{1-\beta}^{1} F^{-1}(u)\,du \;=\; \frac{1}{\beta}\int_{1-\beta}^{1} \big[x_0 + e^{\mu+\sigma\Phi^{-1}(u)}\big]\,du \\ &= x_0 + \frac{1}{\beta}\,e^{\mu+\sigma^2/2}\,\Phi\left(\sigma - \Phi^{-1}(1-\beta)\right), \end{aligned}$$

where β $(0 < \beta < 1)$ is the risk appetite. (For more integration details, see Appendix A.) $\square$

3.3 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, 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 $\psi(s) = r(1-s)^{r-1}$:

$$\text{PHT}[F, r] = \int_0^\infty \big[1 - F(x)\big]^r \, dx = r \int_0^1 F^{-1}(u)(1-u)^{r-1} \, du, \tag{3.6}$$

where constant r $(0 < r \leq 1)$ represents the degree of distortion and F^{-1} is the qf of X. Note that $\text{PHT}[F, 1]$ is the expected value of X, and $\text{PHT}[F, 1/2] - \text{PHT}[F, 1]$ is the right-tail 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 and is justified by utility theory (see Wang, 1998a,b).

The following examples present PHT for the severity distributions of Section 2.2.

Example 3.7. [PHT of Shifted Exponential]
If $X \sim \mathcal{E}xp\,(x_0, \theta)$, then its PHT measure is found by integrating (3.6) as follows:

$$\begin{aligned} \text{PHT}[F, r] &= r \int_0^1 F^{-1}(u)(1-u)^{r-1} \, du = r \int_0^1 \big[x_0 - \theta \log(1-u)\big](1-u)^{r-1} \, du \\ &= x_0 + \theta / r, \end{aligned}$$

where $0 < r \leq 1$ is the degree of distortion. (For more integration details, see Appendix A.) $\square$

Example 3.8. [PHT of PARETO I]

If $X \sim \mathcal{P}a\,I\,(x_0, \alpha)$, then its PHT measure is found by integrating (3.6) as follows:

$$\begin{aligned} \text{PHT}[F, r] &= r \int_0^1 F^{-1}(u)(1-u)^{r-1}\,du \;=\; r \int_0^1 \left[x_0(1-u)^{-1/\alpha}\right](1-u)^{r-1}\,du \\ &= x_0 + \frac{x_0}{r\alpha - 1}, \qquad \alpha > 1/r, \end{aligned}$$

where $0 < r \leq 1$ is the degree of distortion. (For more integration details, see Appendix A.) Note that PHT is infinite when $\alpha \leq 1/r$. □

Example 3.9. [PHT of SHIFTED LOGNORMAL]

If $X \sim \mathcal{LN}\,(x_0, \mu, \sigma)$, then its PHT measure is found by integrating (3.6) 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 $0 < r \leq 1$ is the degree of distortion. (For more integration details, see Appendix A.) 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 3.1 establishes a class of lower and upper bounds for $C_{\text{PHT}}(r, \sigma)$. □

Theorem 3.1. *For $0 < r \leq 1$ 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*

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

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

Proof: 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 $\big(1-\Phi(x)\big)^r \le \big(1-\Phi(z)\big)^r \le 1$ for $z \le x$. That is,

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

and

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

Therefore,

$$e^{\sigma x}\,\big(1-\Phi(x)\big)^r \;\le\; I_{1,x}(r,\sigma) \;\le\; e^{\sigma x}. \tag{3.7}$$

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.5(a) 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 lengthy) integration:

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

Therefore,

$$-e^{\sigma x}(1-\Phi(x)) \;+\; e^{\sigma^2/2}\,\Phi(\sigma - x) \;\le\; I_{2,x}(r,\sigma) \;<\; K_x(r,\sigma). \tag{3.8}$$

Now, adding (3.7) and (3.8) yields the statement of the theorem. □

In Table 3.1, 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 3.1 work well, although more work is needed in identifying 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 $2.9\times 10^3 < 5.0\times 10^3 < 1.4\times 10^4$.

Table 3.1: 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

3.4 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 $g(t) = \Phi\left(\Phi^{-1}(t) + \lambda\right)$ or, equivalently, by the weight function $\psi(u) = e^{\lambda\Phi^{-1}(u) - \lambda^2/2}$:

$$\mathrm{WT}[F, \lambda] \;=\; \int_0^{\infty} \Phi\left(\Phi^{-1}(1 - F(x)) + \lambda\right)\, dx \;=\; \int_0^1 F^{-1}(u)\, e^{\lambda\Phi^{-1}(u) - \lambda^2/2}\, du, \qquad (3.9)$$

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.

The following examples present WT for the severity distributions of Section 2.2.

Example 3.10. [WT of SHIFTED EXPONENTIAL]
If $X \sim \mathcal{E}xp\,(x_0, \theta)$, then its WT measure is found by integrating (3.9) as follows:

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

where φ is the pdf of the standard normal random variable and $-\infty < \lambda < \infty$ is the risk aversion index. (For more integration details, see Appendix A.) Note that for fixed λ, the integral $C_{\mathrm{WT}}(\lambda) = \displaystyle\int_{-\infty}^{\infty} \Phi(z + \lambda)\, \frac{\varphi(z)}{\Phi(z)}\, dz$ is finite and can be evaluated numerically. Theorem 3.2 establishes a class of lower and upper bounds for $C_{\mathrm{WT}}(\lambda)$. □

Theorem 3.2. *For any real constant λ, define $C_{\mathrm{WT}}(\lambda) = \displaystyle\int_{-\infty}^{\infty} \Phi(z + \lambda)\, \frac{\varphi(z)}{\Phi(z)}\, dz$, where Φ and φ are the cdf and pdf of the standard normal distribution, respectively. Then:*

(a) For $\lambda \leq 0$, the double inequality

$$-\Phi(x + \lambda) \log\left[\Phi(x)\right] \;\leq\; C_{\mathrm{WT}}(\lambda) \;\leq\; \Phi(x + \lambda) - \log\left[\Phi(x)\right]$$

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: 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

$$-\Phi(x+\lambda)\log[\Phi(x)] \;\le\; I_{2,x}(\lambda) \;\le\; -\log[\Phi(x)]\,. \tag{3.10}$$

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.6(a) leads to

$$0 \;\le\; I_{1,x}(\lambda) \;\le\; \int_{-\infty}^{x} \left[e^{-\lambda z-\lambda^2/2}\,\Phi(z)\right]\frac{\varphi(z)}{\Phi(z)}\,dz \;=\; \int_{-\infty}^{x} \varphi(z+\lambda)\,dz \;=\; \Phi(x+\lambda). \tag{3.11}$$

Adding (3.10) and (3.11) proves the double inequality in (*a*).

For $\lambda > 0$, we apply Lemma B.6(b) (note that the upper bound is valid for $x < -\lambda$) and then $z/(z+\lambda) \le x/(x+\lambda)$ for $z \le x < -\lambda$:

$$\begin{aligned} 0 \;\le\; I_{1,x}(\lambda) \;&<\; \int_{-\infty}^{x} \left[\frac{z}{z+\lambda}\,e^{-\lambda z-\lambda^2/2}\,\Phi(z)\right]\frac{\varphi(z)}{\Phi(z)}\,dz \\ &\le\; \frac{x}{x+\lambda}\int_{-\infty}^{x} \varphi(z+\lambda)\,dz \;=\; \frac{x}{x+\lambda}\,\Phi(x+\lambda). \end{aligned} \tag{3.12}$$

Adding (3.10) and (3.12) proves the double inequality in (*b*). □

Example 3.11. [WT of PARETO I]

If $X \sim \mathcal{P}a\,I\,(x_0, \alpha)$, then its WT measure is found by integrating (3.9) as follows:

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

where $-\infty < \lambda < \infty$ is the risk aversion index. (For more integration details, see Appendix A.) Note that for fixed λ and $\alpha > 1$, the integral $C_{\mathrm{WT}}(\lambda,\alpha) = \int_{-\infty}^{\infty} \Phi(z+\lambda)\, \frac{\varphi(z)}{\big[\Phi(z)\big]^{1/\alpha+1}}\, dz$ is finite and can be evaluated numerically. Theorem 3.3 establishes a class of lower and upper bounds for $C_{\mathrm{WT}}(\lambda,\alpha)$. $\square$

Theorem 3.3. *For any real constant λ and $\alpha > 1$, define $C_{\mathrm{WT}}(\lambda,\alpha) = \int_{-\infty}^{\infty} \Phi(z+\lambda)\, \frac{\varphi(z)}{\big[\Phi(z)\big]^{1/\alpha+1}}\, dz$, where Φ and φ are the cdf and pdf of the standard normal distribution, respectively. Then:*

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

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

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

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

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

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

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

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

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,

$$\Phi(x+\lambda)\, c_x(\alpha) \;\leq\; I_{2,x}(\lambda, \alpha) \;\leq\; c_x(\alpha). \tag{3.13}$$

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.6(a) and then integration by parts (with the condition $\alpha > 1$) lead to

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

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

$$0 \;\leq\; I_{1,x}(\lambda, \alpha) \;\leq\; \frac{\alpha}{\alpha - 1}\, e^{-\lambda x - \lambda^2/2} \big[\Phi(x)\big]^{1-1/\alpha}. \tag{3.14}$$

Adding (3.13) and (3.14) proves the double inequality in (a).

For $\lambda > 0$, we apply Lemma B.6(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:

$$\begin{aligned} 0 \;\leq\; I_{1,x}(\lambda, \alpha) \;&<\; \int_{-\infty}^x \left[\frac{z}{z+\lambda}\, e^{-\lambda z - \lambda^2/2}\, \Phi(z) \right] \frac{\varphi(z)}{\big[\Phi(z)\big]^{1/\alpha+1}}\, dz \\ &\leq\; \frac{x\, e^{-\lambda^2/2}}{x+\lambda} \int_{-\infty}^x e^{-\lambda z}\, \frac{\varphi(z)}{\big[\Phi(z)\big]^{1/\alpha}}\, dz \\ &=\; \frac{\alpha}{\alpha - 1}\, \frac{x\, e^{-\lambda^2/2}}{x+\lambda} \left[e^{-\lambda x} \big[\Phi(x)\big]^{1-1/\alpha} + \lambda \int_{-\infty}^x \big[\Phi(z)\big]^{1-1/\alpha} e^{-\lambda z}\, dz \right], \end{aligned} \tag{3.15}$$

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

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

Further, we continue (3.15) by first applying Lemma B.5(b), then $-z^{-1} \le -x^{-1}$ for $z \le x < -\lambda$, and finishing with straightforward (but messy) integration:

$$\begin{aligned} 0 \le I_{1,x}(\lambda,\alpha) &< \frac{\alpha}{\alpha-1}\,\frac{x\,e^{-\lambda^2/2}}{x+\lambda}\left[e^{-\lambda x}\big[\Phi(x)\big]^{1-1/\alpha} + \lambda\int_{-\infty}^{x}\left[\frac{\varphi(z)}{-z}\right]^{1-1/\alpha} e^{-\lambda z}\,dz\right] \\ &\le \frac{\alpha}{\alpha-1}\,\frac{x\,e^{-\lambda^2/2}}{x+\lambda}\left[e^{-\lambda x}\big[\Phi(x)\big]^{1-1/\alpha} + \lambda(-x)^{1-1/\alpha}\int_{-\infty}^{x}\left[\varphi(z)\right]^{1-1/\alpha} e^{-\lambda z}\,dz\right] \\ &= \frac{\alpha}{\alpha-1}\,\frac{x}{x+\lambda}\left[e^{-\lambda x-\lambda^2/2}\big[\Phi(x)\big]^{1-1/\alpha} + C_x(\alpha,\lambda)\right], \qquad (3.16)\end{aligned}$$

where $C_x(\alpha,\lambda) = \lambda\sqrt{\alpha(2\pi)^{1/\alpha}/(\alpha-1)}\,(-x)^{1/\alpha-1}\,e^{-(\lambda^2/2)/(\alpha-1)}\,\Phi\left(\dfrac{x+\lambda\alpha/(\alpha-1)}{\sqrt{\alpha/(\alpha-1)}}\right)$.
Adding (3.13) and (3.16) proves the double inequality in (b). $\square$

In Table 3.2, we provide numerical approximations of $C_{\text{WT}}(\lambda)$ and $C_{\text{WT}}(\lambda,\alpha)$ for typical ranges of λ and α. Note that the lower and upper bounds established in Theorems 3.2 and 3.3 are reasonably tight, although more work is needed on identifying optimal value of x. A few illustrations for $C_{\text{WT}}(\lambda)$: for $\lambda = -0.5$, we have $C_{\text{WT}}(\lambda) \approx 0.619$ and $0.21 < 0.619 < 1.14$ (when $x = \lambda/2$); for $\lambda = 0.25$, we have $C_{\text{WT}}(\lambda) \approx 1.245$ and $0.47 < 1.245 < 1.58$ (when $x = -2\lambda$); for $\lambda = 1$, we have $C_{\text{WT}}(\lambda) \approx 2.232$ and $0.60 < 2.232 < 3.94$ (when $x = -2\lambda$). Likewise, for $C_{\text{WT}}(\lambda,\alpha)$: for $\lambda = -0.5$ and $\alpha = 2.5$, we have $C_{\text{WT}}(\lambda,\alpha) \approx 0.886$ and $0.25 < 0.886 < 1.85$ (when $x = \lambda/2$); for $\lambda = -1$ and $\alpha = 4$, we have $C_{\text{WT}}(\lambda,\alpha) \approx 0.416$ and $0.09 < 0.416 < 1.57$ (when $x = \lambda/2$); for $\lambda = 0.5$ and $\alpha = 1.25$, we have $C_{\text{WT}}(\lambda,\alpha) \approx 20.965$ and $1.30 < 20.965 < 24.86$ (when $x = -2\lambda$).

Table 3.2: Numerical evaluations of $C_{\text{WT}}(\lambda)$ and $C_{\text{WT}}(\lambda, \alpha)$ for selected λ and α.

		$C_{\text{WT}}(\lambda,\alpha)$ for α								
λ	$C_{\text{WT}}(\lambda)$	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	11090.602	158.182	26.874	13.403	9.143	6.035	4.824	3.803	3.355

Example 3.12. [WT of SHIFTED LOGNORMAL]

If $X \sim \mathcal{LN}(x_0, \mu, \sigma)$, then its WT measure is found by integrating (3.9) as follows:

$$\begin{aligned} \text{WT}[F,\lambda] &= \int_0^1 F^{-1}(u)\, e^{\lambda\Phi^{-1}(u)-\lambda^2/2}\, du = \int_0^1 \left[x_0 + e^{\mu+\sigma\Phi^{-1}(u)}\right] e^{\lambda\Phi^{-1}(u)-\lambda^2/2}\, du \\ &= x_0 + \exp\left\{\mu + \lambda\sigma + \sigma^2/2\right\}, \end{aligned}$$

where $-\infty < \lambda < \infty$ is the risk aversion index. (For more integration details, see Appendix A.) □

3.5 Gini Shortfall

The GS measure was introduced by Furman *et al*. (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

$$\begin{aligned} \text{GS}[F,\beta,\delta] &= F^{-1}(1-\beta) + \frac{1}{\beta}\int_{F^{-1}(1-\beta)}^{\infty} \big[1-F(x)\big]\,dx \\ &\quad + \frac{2\delta}{\beta^2}\int_{F^{-1}(1-\beta)}^{\infty} \big[1-F(x)\big]\big[\beta-1+F(x)\big]\,dx \qquad (3.17) \\ &= \frac{1}{\beta^2}\int_{1-\beta}^{1} F^{-1}(u)\,\big(\beta+4\delta(u-1+\beta/2)\big)\,du, \qquad (3.18) \end{aligned}$$

where $0<\beta<1$ is the risk appetite (actually, Furman *et al*., 2017, instead of parameter β used $1-\beta$ and called it the *prudence level*) and $\delta \geq 0$ is the *loading parameter*. As is evident from (3.17), the distortion function for GS is $g(t) = t/\beta + 2\delta(t/\beta)(1-t/\beta)$ for $0 \leq t < \beta$, and $= 1$ for $\beta \leq t \leq 1$. Alternatively, (3.18) follows from (3.2) with $\psi(u) = \beta^{-2}\big(\beta + 4\delta(u-1+\beta/2)\big)\,\mathbf{1}\big\{1-\beta \leq u \leq 1\big\}$, where $\mathbf{1}\{\}$ denotes the indicator function. Comparing (3.17) with (3.4) 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 *et al*. (2017), GS is a coherent risk measure if and only if $0 \leq \delta \leq 1/2$.

The following examples present GS for the severity distributions of Section 2.2.

Example 3.13. [GS of SHIFTED EXPONENTIAL]

If $X \sim \mathcal{E}xp\,(x_0, \theta)$, then its GS measure is found by integrating (3.18) as follows:

$$\begin{aligned}
\text{GS}[F, \beta, \delta] &= \frac{1}{\beta^2} \int_{1-\beta}^{1} F^{-1}(u) \big(\beta + 4\delta(u - 1 + \beta/2)\big)\, du \\
&= \frac{1}{\beta^2} \int_{1-\beta}^{1} \big[x_0 - \theta \log(1-u)\big] \big(\beta(1+2\delta) - 4\delta(1-u)\big)\, du \\
&= x_0 - \theta\big[\log(\beta) - 1 - \delta\big],
\end{aligned}$$

where $0 < \beta < 1$ is the risk appetite and $0 \leq \delta \leq 1/2$ is the loading parameter (restricted to the interval $[0; 1/2]$ to make GS coherent). For more integration details, see Appendix A. □

Example 3.14. [GS of PARETO I]

If $X \sim \mathcal{P}a\,I\,(x_0, \alpha)$, then its WT measure is found by integrating (3.18) as follows:

$$\begin{aligned}
\text{GS}[F, \beta, \delta] &= \frac{1}{\beta^2} \int_{1-\beta}^{1} F^{-1}(u) \big(\beta + 4\delta(u - 1 + \beta/2)\big)\, du \\
&= \frac{1}{\beta^2} \int_{1-\beta}^{1} \big[x_0(1-u)^{-1/\alpha}\big] \big(\beta(1+2\delta) - 4\delta(1-u)\big)\, du \\
&= \alpha x_0 \beta^{-1/\alpha}(2\delta/(2\alpha-1)+1)(\alpha-1)^{-1}, \qquad \alpha > 1,
\end{aligned}$$

where $0 < \beta < 1$ is the risk appetite and $0 \leq \delta \leq 1/2$ is the loading parameter (restricted to the interval $[0; 1/2]$ to make GS coherent). For more integration details, see Appendix A. □

Example 3.15. [GS of Shifted Lognormal]

If $X \sim \mathcal{LN}(x_0, \mu, \sigma)$, then its GS measure is found by integrating (3.18) as follows:

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

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). For more integration details, see Appendix A. Note that for fixed β and σ, the integral $C_{\text{GS}}(\beta, \sigma) = \int_{\Phi^{-1}(1-\beta)}^{\infty} \Phi(z)\, \varphi(z - \sigma)\, dz$ is finite and can be evaluated numerically. Theorem 3.4 establishes a lower and upper bound for $C_{\text{GS}}(\beta, \sigma)$. □

Theorem 3.4. *For $0 < \beta < 1$ and $\sigma > 0$, define $C_{\text{GS}}(\beta, \sigma) = \int_{\Phi^{-1}(1-\beta)}^{\infty} \Phi(z)\, \varphi(z - \sigma)\, dz$, 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:*

$$(1 - \beta)\, \Phi\left(\sigma - \Phi^{-1}(1 - \beta)\right) \ \le\ C_{\text{GS}}(\beta, \sigma) \ \le\ \Phi\left(\sigma - \Phi^{-1}(1 - \beta)\right).$$

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

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

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

□

In Table 3.3, we provide numerical evaluations of $C_{\text{GS}}(\beta, \sigma)$ for typical ranges of β and σ. Note that the lower and upper bounds established in Theorem 3.4 are reasonably tight. For instance: for $\beta = 0.20$ and $\sigma = 1/10$, we have $C_{\text{GS}}(\beta, \sigma) \approx 0.207$ and $0.183 < 0.207 < 0.229$; for $\beta = 0.01$ and $\sigma = 1$, we have $C_{\text{GS}}(\beta, \sigma) \approx 0.092$ and $0.091 < 0.092 \leq 0.092$; for $\beta = 0.10$ and $\sigma = 5$, we have $C_{\text{GS}}(\beta, \sigma) \approx 1.000$ and $0.900 < 1.000 \leq 1.000$.

Table 3.3: Numerical evaluations of $C_{\text{GS}}(\beta, \sigma)$ 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

Chapter 4

Risk Measure Estimation

In this chapter, we tackle the problem of estimating the riskiness of the ground-up variable X when only its left-truncated and right-censored version X^* is observed. We start, in Section 4.1, with the (simple but incorrect) empirical estimator of the distortion risk measures; the objective is to show how biased this estimator can be and how it can mislead a decision maker. Then, in Section 4.2, two parametric – maximum likelihood (ML) and percentile matching (PM) – estimators are formulated and their asymptotic distributions are established. A series of analytic examples are worked out in Section 4.3, where the ML and PM estimators of VaR, CTE, PHT, WT, and GS are derived and their asymptotically normal distributions are specified for shifted exponential, Pareto I, and shifted lognormal severity distributions.

To put it in a few words, the problem we are interested in solving is to estimate $R[F]$ based on the observed data $X_1^* = x_1^*, \ldots, X_n^* = x_n^*$ which have common cdf F_*. Thus, for estimation of model parameters and for empirical estimation of risk measures, the asymptotic theorems of Appendix B have to be applied to functions F_*, f_*, F_*^{-1}, which are defined by (2.3)–(2.5), not F, f, F^{-1}. However, for estimation of risk measures, the parameter estimators have to be applied to risk measures based on F, f, F^{-1}, which were specified in Examples 3.1–3.15.

4.1 Empirical Approach

Let us start by noting that the empirical approach is restricted to the range of observed data. Indeed, based on $x_1^*, \ldots, x_n^*$, the empirical estimator $\widehat{F}_{\text{EMP}}(d) = n^{-1} \sum_{i=1}^{n} \mathbf{1}\{x_i^* \leq d\} = 0$. Thus, it cannot take full advantage of formulas (2.3)–(2.5), and yields a biased estimator of qf which in turn propagates the error through the integral that defines a distortion risk measure. More specifically, $R[F]$ is estimated by replacing F with $\widehat{F}_{\text{EMP}}$ in (3.2):

$$\begin{aligned} \widehat{R[F]}_{\text{EMP}} \;=\; R[\widehat{F}_{\text{EMP}}] \;&=\; \int_0^1 \widehat{F}^{-1}_{\text{EMP}}(t)\,\psi(t)\,dt \;=\; \sum_{i=1}^{n} x_{(i)}^* \left[\int_{(i-1)/n}^{i/n} \psi(t)\,dt\right] \\ &=\; \sum_{i=1}^{n} x_{(i)}^* \left[g\left(1 - \frac{i-1}{n}\right) - g\left(1 - \frac{i}{n}\right)\right], \end{aligned} \tag{4.1}$$

where $\psi(t) = g'(1-t)$ and $x_{(1)}^* \leq \cdots \leq x_{(n)}^*$ denotes the ordered values of $x_1^*, \ldots, x_n^*$. Note that $R[\widehat{F}_{\text{EMP}}]$ as defined in equation (4.1) is an L-statistic; asymptotic theory for such statistics is well known (see, e.g., Serfling, 1980, Chapter 7). Note also that $\widehat{F}^{-1}_{\text{EMP}}(t)$ does not converge to $F^{-1}(t)$, which is our target quantity, rather it converges to $F_*^{-1}(t)$. Thus, as follows from Theorem 3.2 of Jones and Zitikis (2003), with some modifications due to data truncation and censoring,

$$R[\widehat{F}_{\text{EMP}}] \;\sim\; \mathcal{AN}\left(R[F_*],\, \frac{1}{n}\,Q(\psi,\psi)\right), \tag{4.2}$$

where $\mathcal{AN}$ stands for 'asymptotically normal' and

$$Q(\psi,\psi) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Big[\min\{F_*(x), F_*(y)\} - F_*(x)F_*(y)\Big]\psi(F_*(x))\psi(F_*(y))\,dx\,dy.$$

Further, to use this result in practice, one needs an estimator for $Q(\psi,\psi)$. Jones and Zitikis (2003) proposed the following *strongly consistent* estimator (modified for data truncation and censoring):

$$\widehat{Q}(\psi,\psi) \;=\; \sum_{i=1}^{n-1}\sum_{j=1}^{n-1} c_n(i,j)\psi(i/n)\psi(j/n)\big[x^*_{(i+1)} - x^*_{(i)}\big]\big[x^*_{(j+1)} - x^*_{(j)}\big]$$

with $c_n(i,j) = \min\{i/n, j/n\} - (i/n)(j/n)$.

Finally, to see that $R[\widehat{F}_{\text{EMP}}]$ is biased, note that for the mean parameter in (4.2), we have

$$R[F_*] \;=\; \int_0^1 F_*^{-1}(t)\psi(t)\,dt \;=\; \int_0^1 F^{-1}\big(t+(1-t)F(d)\big)\psi(t)\,dt \;\geq\; \int_0^1 F^{-1}(t)\psi(t)\,dt \;=\; R[F],$$

with the inequality being strict unless $F(d) = 0$. The inequality holds because F^{-1} is strictly increasing (loss severities are non-negative absolutely continuous random variables) and $(1-t)F(d) \geq 0$.

Below is a summary of specific formulas of $R[\widehat{F}_{\text{EMP}}]$ and simplified expressions of $R[F_*]$ and $Q(\psi,\psi)$ for VaR, CTE, GS, PHT, and WT. These formulas will be used in simulations of Chapter 5.

- VALUE-AT-RISK

  For $0 < 1-\beta < \frac{F(u)-F(d)}{1-F(d)} \leq 1$, the empirical estimator

$$\text{VaR}[\widehat{F}_{\text{EMP}},\beta] = x^*_{(\lceil n(1-\beta)\rceil)}$$

  satisfies

$$\text{VaR}[\widehat{F}_{\text{EMP}},\beta] \;\sim\; \mathcal{AN}\left(\text{VaR}[F_*,\beta],\, \frac{1}{n}\,Q_{\text{VAR}}\right), \tag{4.3}$$

  where $\text{VaR}[F_*,\beta] = F^{-1}\big((1-\beta)+\beta F(d)\big)$ and $Q_{\text{VAR}} = \frac{\beta(1-\beta)\,[1-F(d)]^2}{f^2\left(F^{-1}\left((1-\beta)+\beta F(d)\right)\right)}$. Otherwise, that is for $0 \leq \frac{F(u)-F(d)}{1-F(d)} \leq 1-\beta \leq 1$, parameters in (4.3) become $\text{VaR}[F_*,\beta] = u$ and $Q_{\text{VAR}} = 0$.

- **Conditional Tail Expectation**

  For $0 < 1-\beta < \frac{F(u)-F(d)}{1-F(d)} \leq 1$, the empirical estimator

$$\text{CTE}[\widehat{F}_{\text{EMP}}, \beta] = \frac{1}{\lceil n\beta \rceil} \sum_{i=n-\lceil n\beta \rceil+1}^{n} x^*_{(i)}$$

  satisfies

$$\text{CTE}[\widehat{F}_{\text{EMP}}, \beta] \;\sim\; \mathcal{AN}\left(\text{CTE}[F_*, \beta],\, \frac{1}{n}\, Q_{\text{CTE}}\right), \tag{4.4}$$

  where

$$\text{CTE}[F_*, \beta] \;=\; \int_{1-\beta}^{\frac{F(u)-F(d)}{1-F(d)}} F^{-1}\big(t + (1-t)F(d)\big)\,\frac{1}{\beta}\,dt \;+\; \frac{u}{\beta}\frac{1-F(u)}{1-F(d)}$$

  and

$$Q_{\text{CTE}} \;=\; \int_{F^{-1}\big((1-\beta)+\beta F(d)\big)}^{u} \Big[Q^{(1)}_{\text{CTE}}(y) + Q^{(2)}_{\text{CTE}}(y)\Big]\,\frac{1}{\beta}\,dy,$$

  with

$$Q^{(1)}_{\text{CTE}}(y) \;=\; \left(1 - \frac{F(y)-F(d)}{1-F(d)}\right) \int_{F^{-1}\big((1-\beta)+\beta F(d)\big)}^{y} \frac{F(x)-F(d)}{1-F(d)}\,\frac{1}{\beta}\,dx$$

  and

$$Q^{(2)}_{\text{CTE}}(y) \;=\; \frac{F(y)-F(d)}{1-F(d)} \int_{y}^{u} \left(1 - \frac{F(x)-F(d)}{1-F(d)}\right) \frac{1}{\beta}\,dx.$$

  Here functions $Q^{(1)}_{\text{CTE}}(y)$, $Q^{(2)}_{\text{CTE}}(y)$ and subsequently Q_{CTE} can be evaluated numerically. Also, for $0 \leq \frac{F(u)-F(d)}{1-F(d)} \leq 1-\beta \leq 1$, parameters in (4.4) become $\text{CTE}[F_*, \beta] = u$ and $Q_{\text{CTE}} = 0$.

- **Proportional Hazards Transform**

  The empirical estimator

$$\text{PHT}[\widehat{F}_{\text{EMP}}, r] = \sum_{i=1}^{n} \Big[(1-(i-1)/n)^r - (1-i/n)^r\Big]\, x^*_{(i)}$$

satisfies

$$\text{PHT}[\widehat{F}_{\text{EMP}}, r] \;\sim\; \mathcal{AN}\left(\text{PHT}[F_*, r],\, \frac{1}{n}\, Q_{\text{PHT}}\right), \tag{4.5}$$

where

$$\text{PHT}[F_*, r] \;=\; \int_0^{\frac{F(u)-F(d)}{1-F(d)}} F^{-1}\big(t + (1-t)F(d)\big)\, r(1-t)^{r-1}\, dt \;+\; u\left(\frac{1-F(u)}{1-F(d)}\right)^r$$

and

$$Q_{\text{PHT}} \;=\; \int_d^u \left[Q^{(1)}_{\text{PHT}}(y) + Q^{(2)}_{\text{PHT}}(y)\right]\, r\left(1 - \frac{F(y)-F(d)}{1-F(d)}\right)^{r-1}\, dy,$$

with

$$Q^{(1)}_{\text{PHT}}(y) \;=\; \left(1 - \frac{F(y)-F(d)}{1-F(d)}\right) \int_d^y \frac{F(x)-F(d)}{1-F(d)}\, r\left(1 - \frac{F(x)-F(d)}{1-F(d)}\right)^{r-1}\, dx$$

and

$$Q^{(2)}_{\text{PHT}}(y) \;=\; \frac{F(y)-F(d)}{1-F(d)} \int_y^u \left(1 - \frac{F(x)-F(d)}{1-F(d)}\right) r\left(1 - \frac{F(x)-F(d)}{1-F(d)}\right)^{r-1}\, dx.$$

Here functions $Q^{(1)}_{\text{PHT}}(y)$, $Q^{(2)}_{\text{PHT}}(y)$ and subsequently Q_{PHT} can be evaluated numerically.

- Wang Transform

  The empirical estimator

$$\text{WT}[\widehat{F}_{\text{EMP}}, \lambda] = \sum_{i=1}^{n} \left[\Phi\left(\Phi^{-1}(i/n) - \lambda\right) - \Phi\left(\Phi^{-1}((i-1)/n) - \lambda\right)\right] x^*_{(i)},$$

satisfies

$$\text{WT}[\widehat{F}_{\text{EMP}}, \lambda] \;\sim\; \mathcal{AN}\left(\text{WT}[F_*, \lambda],\, \frac{1}{n}\, Q_{\text{WT}}\right), \tag{4.6}$$

where

$$\textsc{wt}[F_*, \lambda] \;=\; \int_0^{\frac{F(u)-F(d)}{1-F(d)}} F^{-1}\big(t+(1-t)F(d)\big)\, e^{\lambda\Phi^{-1}(t)-\lambda^2/2}\, dt$$

$$+\; u\left[1-\Phi\left(\Phi^{-1}\left(\frac{F(u)-F(d)}{1-F(d)}\right)-\lambda\right)\right]$$

and

$$Q_{\textsc{wt}} \;=\; \int_d^u \left[Q^{(1)}_{\textsc{wt}}(y)+Q^{(2)}_{\textsc{wt}}(y)\right] e^{\lambda\Phi^{-1}\left(\frac{F(y)-F(d)}{1-F(d)}\right)-\lambda^2/2}\, dy,$$

with

$$Q^{(1)}_{\textsc{wt}}(y) \;=\; \left(1-\frac{F(y)-F(d)}{1-F(d)}\right)\int_d^y \frac{F(x)-F(d)}{1-F(d)}\, e^{\lambda\Phi^{-1}\left(\frac{F(x)-F(d)}{1-F(d)}\right)-\lambda^2/2}\, dx$$

and

$$Q^{(2)}_{\textsc{wt}}(y) \;=\; \frac{F(y)-F(d)}{1-F(d)}\int_y^u \left(1-\frac{F(x)-F(d)}{1-F(d)}\right) e^{\lambda\Phi^{-1}\left(\frac{F(x)-F(d)}{1-F(d)}\right)-\lambda^2/2}\, dx.$$

Here functions $Q^{(1)}_{\textsc{wt}}(y)$, $Q^{(2)}_{\textsc{wt}}(y)$ and subsequently $Q_{\textsc{wt}}$ can be evaluated numerically.

- Gini Shortfall

  For $0 < 1-\beta < \frac{F(u)-F(d)}{1-F(d)} \leq 1$, the empirical estimator

$$\textsc{gs}[\widehat{F}_{\textsc{emp}}, \beta, \delta] = \frac{1}{\lceil n\beta\rceil^2}\sum_{i=n-\lceil n\beta\rceil+1}^{n}\Big[\lceil n\beta\rceil + 2\delta\big(\lceil n\beta\rceil - 2(n-i)-1\big)\Big]x^*_{(i)},$$

  satisfies

$$\textsc{gs}[\widehat{F}_{\textsc{emp}}, \beta, \delta] \;\sim\; \mathcal{AN}\left(\textsc{gs}[F_*, \beta, \delta],\, \frac{1}{n}\, Q_{\textsc{gs}}\right), \tag{4.7}$$

  where

$$\textsc{gs}[F_*, \beta, \delta] \;=\; \int_{1-\beta}^{\frac{F(u)-F(d)}{1-F(d)}} F^{-1}\big(t+(1-t)F(d)\big)\, \beta^{-2}\big(\beta(1+2\delta)-4\delta(1-t)\big)\, dt$$

$$+\; \frac{u}{\beta}\frac{1-F(u)}{1-F(d)} \;+\; \frac{2u\delta}{\beta}\frac{1-F(u)}{1-F(d)}\left(1-\frac{1}{\beta}\frac{1-F(u)}{1-F(d)}\right)$$

and

$$Q_{\text{GS}} = \int_{F^{-1}\left((1-\beta)+\beta F(d)\right)}^{u} \left[Q_{\text{GS}}^{(1)}(y) + Q_{\text{GS}}^{(2)}(y)\right] \widetilde{\psi}(y)\, dy$$

with

$$Q_{\text{GS}}^{(1)}(y) = \left(1 - \frac{F(y)-F(d)}{1-F(d)}\right) \int_{F^{-1}\left((1-\beta)+\beta F(d)\right)}^{y} \frac{F(x)-F(d)}{1-F(d)}\, \widetilde{\psi}(x)\, dx$$

and

$$Q_{\text{GS}}^{(2)}(y) = \frac{F(y)-F(d)}{1-F(d)} \int_{y}^{u} \left(1 - \frac{F(x)-F(d)}{1-F(d)}\right) \widetilde{\psi}(x)\, dx.$$

Here $\widetilde{\psi}(\cdot) = \beta^{-2}\left(\beta(1+2\delta) - 4\delta\left(1 - \frac{F(\cdot)-F(d)}{1-F(d)}\right)\right)$. Also, functions $Q_{\text{GS}}^{(1)}(y)$, $Q_{\text{GS}}^{(2)}(y)$ and subsequently Q_{GS} can be evaluated numerically. For $0 \leq \frac{F(u)-F(d)}{1-F(d)} \leq 1-\beta \leq 1$, parameters in (4.7) become $\text{GS}[F_*, \beta, \delta] = u$ and $Q_{\text{GS}} = 0$.

4.2 Parametric Approaches

Parametric methods use the observed data $x_1^*, \ldots, x_n^*$ and fully recognize its distributional properties.

4.2.1 ML Estimation

The ML approach takes into account (2.3)–(2.5) and finds parameter estimates by maximizing the following log-likelihood function:

$$\log \mathcal{L}\big(\boldsymbol{\theta} \,|\, x_1^*, \ldots, x_n^*\big) = \log\left[\prod_{i=1}^{n} f_*(x_i^*)\right] = \log\left[\prod_{i=1}^{n}\left[\frac{f(x_i^*)}{1-F(d)}\right]^{\mathbf{1}\{d<x_i^*<u\}}\left[\frac{1-F(u)}{1-F(d)}\right]^{\mathbf{1}\{x_i^*=u\}}\right]$$

$$= \sum_{i=1}^{n} \log\left[f(x_i^*)\right]\mathbf{1}\{d < x_i^* < u\} - n\log\left[1-F(d)\right] + \log\left[1-F(u)\right]\sum_{i=1}^{n}\mathbf{1}\{x_i^* = u\}. \tag{4.8}$$

Once parameter ML estimators, $\widehat{\boldsymbol{\theta}} = \big(\widehat{\theta}_1, \ldots, \widehat{\theta}_k\big)$, are available, the risk measure estimate is found by plugging those ML values into the parametric expression of $R[F] = h(\theta_1, \ldots, \theta_k)$. (Note that it is not $R[F_*]$.) Let us denote this estimator as $R[\widehat{F}_{\text{ML}}] =$

$h(\widehat{\theta}_1, \ldots, \widehat{\theta}_k)$. Then, as follows from the ML's asymptotic distribution and the delta method (both theorems are provided in Appendix B),

$$R[\widehat{F}_{\text{ML}}] \;\sim\; \mathcal{AN}\left(R[F],\, \frac{1}{n}\, \mathbf{d}_{\boldsymbol{\theta}} \mathbf{I}_{\boldsymbol{\theta}}^{-1} \mathbf{d}_{\boldsymbol{\theta}}'\right), \tag{4.9}$$

where $\mathbf{d}_{\boldsymbol{\theta}} = \left(\partial h / \partial \widehat{\theta}_1, \ldots, \partial h / \partial \widehat{\theta}_k\right)\Big|_{(\theta_1, \ldots, \theta_k)}$, and the entries of the Fisher information matrix $\mathbf{I}_{\boldsymbol{\theta}}$ are given by (B.3) with g replaced by (2.3).

4.2.2 PM Estimation

A popular alternative to the ML approach for estimation of loss model parameters is percentile matching, abbreviated PM (see Klugman *et al.*, 2012, Section 13.1). To estimate k unknown parameters with the PM method and using the ordered data $x^*_{(1)} \leq \cdots \leq x^*_{(n)}$, one has to solve the following system of equations with respect to $\theta_1, \ldots, \theta_k$:

$$F_*^{-1}(p_1) = x^*_{(\lceil np_1 \rceil)}, \;\; F_*^{-1}(p_2) = x^*_{(\lceil np_2 \rceil)}, \ldots, \;\; F_*^{-1}(p_k) = x^*_{(\lceil np_k \rceil)},$$

where $p_1 < \cdots < p_k < \frac{F(u) - F(d)}{1 - F(d)}$ and $x^*_{(\lceil np_k \rceil)} < u$; here $\lceil \cdot \rceil$ denotes the "rounding up" operation. Once parameter PMs, $\widetilde{\theta}_1, \ldots, \widetilde{\theta}_k$, are available, the risk measure estimate is found by plugging those PM values into $R[F] = h(\theta_1, \ldots, \theta_k)$. Let us denote this estimator as $R[\widehat{F}_{\text{PM}}] = h(\widetilde{\theta}_1, \ldots, \widetilde{\theta}_k)$. Then, as follows from Theorem B.4 and the delta method,

$$R[\widehat{F}_{\text{PM}}] \;\sim\; \mathcal{AN}\left(R[F],\, \frac{1}{n}\, \mathbf{d}_{\boldsymbol{\theta}} \mathbf{D}^*_{\boldsymbol{\theta}} \boldsymbol{\Sigma}_{\boldsymbol{\theta}} (\mathbf{D}^*_{\boldsymbol{\theta}})' \mathbf{d}_{\boldsymbol{\theta}}'\right), \tag{4.10}$$

where $\mathbf{d}_{\boldsymbol{\theta}} = \left(\partial h / \partial \widetilde{\theta}_1, \ldots, \partial h / \partial \widetilde{\theta}_k\right)\Big|_{(\theta_1, \ldots, \theta_k)}$ and $\mathbf{D}^*_{\boldsymbol{\theta}}$ is specified in Theorem B.4. The entries of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ are given by (B.1) with g and G^{-1} replaced by expressions (2.3) and (2.5), respectively.

4.3 Analytic Examples

4.3.1 Shifted Exponential Distribution

If $x_1^*, \ldots, x_n^*$ is a realization of variables (2.2) with pdf (2.7) and cdf (2.6), then the log-likelihood function (4.8) becomes

$$\begin{aligned}\log \mathcal{L}\big(\theta \,\big|\, x_1^*, \ldots, x_n^*\big) \quad = \quad & -\log\theta \sum_{i=1}^{n} \mathbf{1}\{d < x_i^* < u\} \\ & -\frac{1}{\theta}\sum_{i=1}^{n}\Big[(x_i^* - d)\mathbf{1}\{d < x_i^* < u\} + (u-d)\mathbf{1}\{x_i^* = u\}\Big].\end{aligned}$$

Straightforward maximization of $\log \mathcal{L}$ yields an explicit formula of the ML estimator of θ:

$$\widehat{\theta}_{\mathrm{ML}} \;=\; \frac{\sum_{i=1}^{n}\Big[(x_i^* - d)\mathbf{1}\{d < x_i^* < u\} + (u-d)\mathbf{1}\{x_i^* = u\}\Big]}{\sum_{i=1}^{n}\mathbf{1}\{d < x_i^* < u\}} \tag{4.11}$$

The asymptotic distribution of $\widehat{\theta}_{\mathrm{ML}}$ follows from Theorem B.2. In this case, the Fisher information matrix has a single entry (more computational details are available in Appendix A):

$$\begin{aligned} I_{11} \;&=\; -\mathbf{E}\Big[\theta^{-2}\mathbf{1}\{d < X^* < u\} - 2\theta^{-3}\big[(X^* - d)\mathbf{1}\{d < X^* < u\} + (u-d)\mathbf{1}\{X^* = u\}\big]\Big] \\ &=\; \theta^{-2}\big[1 - e^{-(u-d)/\theta}\big].\end{aligned}$$

Hence, the estimator $\widehat{\theta}_{\mathrm{ML}}$, defined by (4.11), has the following asymptotic distribution:

$$\widehat{\theta}_{\mathrm{ML}} \;\sim\; \mathcal{AN}\left(\theta, \, \frac{\theta^2}{n}\,\frac{1}{1 - e^{-(u-d)/\theta}}\right). \tag{4.12}$$

Next, since for the exponential distribution there is only one unknown parameter θ, its PM estimator is derived by solving a single equation, $F_*^{-1}(p_1) = x^*_{(\lceil np_1 \rceil)}$. Note that p_1 has to be chosen from the range $0 < p_1 < 1 - e^{-(u-d)/\theta}$ (equivalently, $x^*_{(\lceil np_1 \rceil)} < u$). In this case, the resulting estimator is also explicit and given by

$$\widehat{\theta}_{\text{PM}} \;=\; \frac{d - x^*_{(\lceil np_1 \rceil)}}{\log(1-p_1)}. \tag{4.13}$$

The asymptotic distribution of $\widehat{\theta}_{\text{PM}}$ follows from Theorem B.4, where

$$\mathbf{D}^*_{\boldsymbol{\theta}} \boldsymbol{\Sigma}_{\boldsymbol{\theta}} (\mathbf{D}^*_{\boldsymbol{\theta}})' \;=\; \frac{-1}{\log(1-p_1)} \cdot \frac{p_1(1-p_1)}{\theta^{-2}(1-p_1)^2} \cdot \frac{-1}{\log(1-p_1)} \;=\; \theta^2 \, \frac{p_1}{(1-p_1)\log^2(1-p_1)}.$$

Hence, the estimator $\widehat{\theta}_{\text{PM}}$, defined by (4.13), has the following asymptotic distribution:

$$\widehat{\theta}_{\text{PM}} \;\sim\; \mathcal{AN}\left(\theta, \, \frac{\theta^2}{n} \, \frac{p_1}{(1-p_1)\log^2(1-p_1)}\right). \tag{4.14}$$

In Examples 4.1–4.5, we use (4.11), (4.13), and (4.1) to estimate VaR, CTE, PHT, WT, and GS. Asymptotic distributions of these estimators, (4.9) and (4.10), follow by applying the delta method to (4.12) and (4.14). For the empirical estimator (4.1) the asymptotic normality is specified by (4.2).

Example 4.1. [Value-at-Risk]
Parametric and empirical estimators of VaR and their asymptotic distributions for the shifted exponential model were derived by Brazauskas and Upretee (2019, Section 3.1). Here, using the notation of this dissertation, we present those results. As derived in Example 3.1, the target parameter is $\text{VaR}[F, \beta] = x_0 - \theta \log(\beta)$. Then, its ML and PM estimators are

$$\text{VaR}[\widehat{F}_{\text{ML}}, \beta] = x_0 - \widehat{\theta}_{\text{ML}} \log(\beta) \quad \text{and} \quad \text{VaR}[\widehat{F}_{\text{PM}}, \beta] = x_0 - \widehat{\theta}_{\text{PM}} \log(\beta).$$

Therefore, the asymptotic distributions (4.9) and (4.10) become

$$\text{VaR}[\widehat{F}_{\text{ML}}, \beta] \;\sim\; \mathcal{AN}\left(x_0 - \theta \log(\beta), \, \frac{\theta^2}{n} \, \frac{\log^2(\beta)}{1 - e^{-(u-d)/\theta}}\right)$$

and

$$\text{VaR}[\widehat{F}_{\text{PM}}, \beta] \;\sim\; \mathcal{AN}\left(x_0 - \theta \log(\beta),\, \frac{\theta^2}{n} \frac{p_1 \log^2(\beta)}{(1-p_1)\log^2(1-p_1)} \right),$$

where $0 < \beta < 1$ and $0 < p_1 < 1 - e^{-(u-d)/\theta}$. □

Example 4.2. [Conditional Tail Expectation]

More detailed derivations of parametric and empirical estimators of CTE and their asymptotic distributions for the shifted exponential model are provided in Appendix A. Here we present a quick overview of those results. As derived in Example 3.4, the target parameter is $\text{CTE}[F, \beta] = x_0 - \theta(\log(\beta) - 1)$. Then, its ML and PM estimators are

$$\text{CTE}[\widehat{F}_{\text{ML}}, \beta] = x_0 - \widehat{\theta}_{\text{ML}}(\log(\beta) - 1) \quad \text{and} \quad \text{CTE}[\widehat{F}_{\text{PM}}, \beta] = x_0 - \widehat{\theta}_{\text{PM}}(\log(\beta) - 1).$$

Therefore, the asymptotic distributions (4.9) and (4.10) become

$$\text{CTE}[\widehat{F}_{\text{ML}}, \beta] \;\sim\; \mathcal{AN}\left(x_0 - \theta(\log(\beta) - 1),\, \frac{\theta^2}{n} \frac{(\log(\beta) - 1)^2}{1 - e^{-(u-d)/\theta}} \right)$$

and

$$\text{CTE}[\widehat{F}_{\text{PM}}, \beta] \;\sim\; \mathcal{AN}\left(x_0 - \theta(\log(\beta) - 1),\, \frac{\theta^2}{n} \frac{p_1(\log(\beta) - 1)^2}{(1-p_1)\log^2(1-p_1)} \right),$$

where $0 < \beta < 1$ and $0 < p_1 < 1 - e^{-(u-d)/\theta}$. □

Example 4.3. [Proportional Hazards Transform]

More detailed derivations of parametric and empirical estimators of PHT and their asymptotic distributions for the shifted exponential model are provided in Appendix A. Here we present a quick overview of those results. As derived in Example 3.7, the target parameter is $\text{PHT}[F, r] = x_0 + \theta/r$. Then, its ML and PM estimators are

$$\text{PHT}[\widehat{F}_{\text{ML}}, r] = x_0 + \widehat{\theta}_{\text{ML}}/r \quad \text{and} \quad \text{PHT}[\widehat{F}_{\text{PM}}, r] = x_0 + \widehat{\theta}_{\text{PM}}/r.$$

Therefore, the asymptotic distributions (4.9) and (4.10) become

$$\text{PHT}[\widehat{F}_{\text{ML}}, r] \;\sim\; \mathcal{AN}\left(x_0 + \frac{\theta}{r},\, \frac{\theta^2}{n} \frac{r^{-2}}{1 - e^{-(u-d)/\theta}} \right)$$

and

$$\text{PHT}[\widehat{F}_{\text{PM}}, r] \;\sim\; \mathcal{AN}\left(x_0 + \frac{\theta}{r},\; \frac{\theta^2}{n}\, \frac{r^{-2} p_1}{(1-p_1)\log^2(1-p_1)}\right),$$

where $0 < r \leq 1$ and $0 < p_1 < 1 - e^{-(u-d)/\theta}$. □

Example 4.4. [Wang Transform]

More detailed derivations of parametric and empirical estimators of WT and their asymptotic distributions for the shifted exponential model are provided in Appendix A. Here we present a quick overview of those results. As derived in Example 3.10, the target parameter is $\text{WT}[F, \lambda] = x_0 + \theta\, C_{\text{WT}}(\lambda)$. Then, its ML and PM estimators are

$$\text{WT}[\widehat{F}_{\text{ML}}, \lambda] = x_0 + \widehat{\theta}_{\text{ML}} C_{\text{WT}}(\lambda) \quad \text{and} \quad \text{WT}[\widehat{F}_{\text{PM}}, \lambda] = x_0 + \widehat{\theta}_{\text{PM}} C_{\text{WT}}(\lambda).$$

Therefore, the asymptotic distributions (4.9) and (4.10) become

$$\text{WT}[\widehat{F}_{\text{ML}}, \lambda] \;\sim\; \mathcal{AN}\left(x_0 + \theta C_{\text{WT}}(\lambda),\; \frac{\theta^2}{n}\, \frac{C^2_{\text{WT}}(\lambda)}{1 - e^{-(u-d)/\theta}}\right)$$

and

$$\text{WT}[\widehat{F}_{\text{PM}}, \lambda] \;\sim\; \mathcal{AN}\left(x_0 + \theta C_{\text{WT}}(\lambda),\; \frac{\theta^2}{n}\, \frac{p_1\, C^2_{\text{WT}}(\lambda)}{(1-p_1)\log^2(1-p_1)}\right),$$

where $-\infty < \lambda < \infty$, $C_{\text{WT}}(\lambda) = \displaystyle\int_{-\infty}^{\infty} \Phi(z+\lambda)\, \frac{\varphi(z)}{\Phi(z)}\, dz$, and $0 < p_1 < 1 - e^{-(u-d)/\theta}$. □

Example 4.5. [Gini Shortfall]

More detailed derivations of parametric and empirical estimators of GS and their asymptotic distributions for the shifted exponential model are provided in Appendix A. Here we present a quick overview of those results. As derived in Example 3.13, the target parameter is $\text{GS}[F, \beta, \delta] = x_0 - \theta\big[\log(\beta) - 1 - \delta\big]$. Then, its ML and PM estimators are

$$\text{GS}[\widehat{F}_{\text{ML}}, \beta, \delta] = x_0 - \widehat{\theta}_{\text{ML}}\big[\log(\beta) - 1 - \delta\big] \quad \text{and} \quad \text{GS}[\widehat{F}_{\text{PM}}, \beta, \delta] = x_0 - \widehat{\theta}_{\text{PM}}\big[\log(\beta) - 1 - \delta\big].$$

Therefore, the asymptotic distributions (4.9) and (4.10) become

$$\mathrm{GS}[\widehat{F}_{\mathrm{ML}},\beta,\delta] \;\sim\; \mathcal{AN}\left(x_0-\theta(\log(\beta)-1-\delta),\,\frac{\theta^2}{n}\,\frac{(\log(\beta)-1-\delta)^2}{1-e^{-(u-d)/\theta}}\right)$$

and

$$\mathrm{GS}[\widehat{F}_{\mathrm{PM}},\beta,\delta] \;\sim\; \mathcal{AN}\left(x_0-\theta(\log(\beta)-1-\delta),\,\frac{\theta^2}{n}\,\frac{p_1(\log(\beta)-1-\delta)^2}{(1-p_1)\log^2(1-p_1)}\right),$$

where $0<\beta<1$, $0\le\delta\le 1/2$, and $0<p_1<1-e^{-(u-d)/\theta}$. $\square$

4.3.2 Pareto I Distribution

If $x_1^*,\ldots,x_n^*$ is a realization of variables (2.2) with pdf (2.10) and cdf (2.9), then the log-likelihood function (4.8) becomes

$$\begin{aligned}\log\mathcal{L}\big(\alpha\,\big|\,x_1^*,\ldots,x_n^*\big) \;&=\; \sum_{i=1}^{n}\big[\log(\alpha/x_0)-(\alpha+1)\log(x_i^*/x_0)\big]\mathbf{1}\big\{d<x_i^*<u\big\}\\ &\quad-\alpha n\log(x_0/d)+\alpha\log(x_0/u)\sum_{i=1}^{n}\mathbf{1}\big\{x_i^*=u\big\}.\end{aligned}$$

Straightforward maximization of $\log\mathcal{L}$ yields an explicit formula of the ML estimator of α:

$$\widehat{\alpha}_{\mathrm{ML}} \;=\; \frac{\sum_{i=1}^{n}\mathbf{1}\big\{d<x_i^*<u\big\}}{\sum_{i=1}^{n}\log(x_i^*/d)\,\mathbf{1}\big\{d<x_i^*<u\big\}+\log(u/d)\sum_{i=1}^{n}\mathbf{1}\big\{x_i^*=u\big\}} \tag{4.15}$$

The asymptotic distribution of $\widehat{\alpha}_{\mathrm{ML}}$ follows from Theorem B.2. In this case, the Fisher information matrix has a single entry (more computational details are available in Appendix A):

$$I_{11} \;=\; -\mathbf{E}\left[-\alpha^{-2}\mathbf{1}\big\{d<X^*<u\big\}\right] \;=\; \alpha^{-2}\left[1-(d/u)^{\alpha}\right].$$

Hence, the estimator $\widehat{\alpha}_{\mathrm{ML}}$, defined by (4.15), has the following asymptotic distribution:

$$\widehat{\alpha}_{\mathrm{ML}} \;\sim\; \mathcal{AN}\left(\alpha,\,\frac{\alpha^2}{n}\,\frac{1}{1-(d/u)^{\alpha}}\right). \tag{4.16}$$

Next, since for the Pareto I distribution there is only one unknown parameter α, its PM estimator is derived by solving a single equation, $F_*^{-1}(p_1) = x^*_{(\lceil np_1 \rceil)}$. Note that p_1 has to be chosen from the range $0 < p_1 < 1 - (d/u)^\alpha$ (equivalently, $x^*_{(\lceil np_1 \rceil)} < u$). In this case, the resulting estimator is also explicit and given by

$$\widehat{\alpha}_{\text{PM}} = \frac{\log(1-p_1)}{\log\left(d/x^*_{(\lceil np_1 \rceil)}\right)}. \tag{4.17}$$

The asymptotic distribution of $\widehat{\alpha}_{\text{PM}}$ follows from Theorem B.4, where

$$\mathbf{D}^*_{\boldsymbol{\alpha}} \boldsymbol{\Sigma}_{\boldsymbol{\alpha}} (\mathbf{D}^*_{\boldsymbol{\alpha}})' = \frac{\alpha^2(1-p_1)^{1/\alpha}}{d\log(1-p_1)} \cdot \frac{d^2 p_1}{\alpha^2(1-p_1)^{1+2/\alpha}} \cdot \frac{\alpha^2(1-p_1)^{1/\alpha}}{d\log(1-p_1)} = \alpha^2 \frac{p_1}{(1-p_1)\log^2(1-p_1)}.$$

Hence, the estimator $\widehat{\alpha}_{\text{PM}}$, defined by (4.17), has the following asymptotic distribution:

$$\widehat{\alpha}_{\text{PM}} \sim \mathcal{AN}\left(\alpha, \frac{\alpha^2}{n} \frac{p_1}{(1-p_1)\log^2(1-p_1)}\right). \tag{4.18}$$

In Examples 4.6–4.10, we use (4.15), (4.17), and (4.1) to estimate VaR, CTE, PHT, WT, and GS. Asymptotic distributions of these estimators, (4.9) and (4.10), follow by applying the delta method to (4.16) and (4.18). For the empirical estimator (4.1) the asymptotic normality is specified by (4.2).

Example 4.6. [Value-at-Risk]
Parametric and empirical estimators of VaR and their asymptotic distributions for the Pareto I model were derived by Brazauskas and Upretee (2019, Section 3.2). Here, using the notation of this dissertation, we present those results. As derived in Example 3.2, the target parameter is $\text{VaR}[F,\beta] = x_0 \beta^{-1/\alpha}$. Then, its ML and PM estimators are

$$\text{VaR}[\widehat{F}_{\text{ML}}, \beta] = x_0 \beta^{-1/\widehat{\alpha}_{\text{ML}}} \quad \text{and} \quad \text{VaR}[\widehat{F}_{\text{PM}}, \beta] = x_0 \beta^{-1/\widehat{\alpha}_{\text{PM}}}.$$

Therefore, the asymptotic distributions (4.9) and (4.10) become

$$\text{VaR}[\widehat{F}_{\text{ML}}, \beta] \sim \mathcal{AN}\left(\text{VaR}[F,\beta], \frac{\alpha^{-2}}{n} \frac{\log^2(\beta)\,\text{VaR}^2[F,\beta]}{1-(d/u)^\alpha}\right)$$

and

$$\mathrm{VaR}[\widehat{F}_{\mathrm{PM}}, \beta] \;\sim\; \mathcal{AN}\left(\mathrm{VaR}[F,\beta],\; \frac{\alpha^{-2}}{n}\, \frac{p_1 \log^2(\beta)\, \mathrm{VaR}^2[F,\beta]}{(1-p_1)\log^2(1-p_1)}\right),$$

where $\mathrm{VaR}[F,\beta] = x_0\beta^{-1/\alpha}$, $0 < \beta < 1$, and $0 < p_1 < 1 - (d/u)^{\alpha}$. □

Example 4.7. [Conditional Tail Expectation]

More detailed derivations of parametric and empirical estimators of CTE and their asymptotic distributions for the Pareto I model are provided in Appendix A. Here we present a quick overview of those results. As derived in Example 3.5, the target parameter is $\mathrm{CTE}[F,\beta] = x_0\beta^{-1/\alpha}\alpha(\alpha-1)^{-1}$, which is finite for $\alpha > 1$. Then, its ML and PM estimators are

$$\mathrm{CTE}[\widehat{F}_{\mathrm{ML}}, \beta] = x_0\beta^{-1/\widehat{\alpha}_{\mathrm{ML}}}\widehat{\alpha}_{\mathrm{ML}}(\widehat{\alpha}_{\mathrm{ML}}-1)^{-1} \quad \text{and} \quad \mathrm{CTE}[\widehat{F}_{\mathrm{PM}}, \beta] = x_0\beta^{-1/\widehat{\alpha}_{\mathrm{PM}}}\widehat{\alpha}_{\mathrm{PM}}(\widehat{\alpha}_{\mathrm{PM}}-1)^{-1}.$$

Therefore, the asymptotic distributions (4.9) and (4.10) become

$$\mathrm{CTE}[\widehat{F}_{\mathrm{ML}}, \beta] \;\sim\; \mathcal{AN}\left(\mathrm{CTE}[F,\beta],\; \frac{\alpha^{-2}}{n}\, \frac{D^2_{\mathrm{CTE}}(\beta,\alpha)\, \mathrm{CTE}^2[F,\beta]}{1-(d/u)^{\alpha}}\right)$$

and

$$\mathrm{CTE}[\widehat{F}_{\mathrm{PM}}, \beta] \;\sim\; \mathcal{AN}\left(\mathrm{CTE}[F,\beta],\; \frac{\alpha^{-2}}{n}\, \frac{p_1 D^2_{\mathrm{CTE}}(\beta,\alpha)\, \mathrm{CTE}^2[F,\beta]}{(1-p_1)\log^2(1-p_1)}\right),$$

where $0 < \beta < 1$ and $0 < p_1 < 1-(d/u)^{\alpha}$. Note that $\mathrm{CTE}[F,\beta] = x_0\beta^{-1/\alpha}\alpha(\alpha-1)^{-1}$ and $D_{\mathrm{CTE}}(\beta,\alpha) = \log(\beta) - \alpha(\alpha-1)^{-1}$ are finite for $\alpha > 1$. □

Example 4.8. [Proportional Hazards Transform]

More detailed derivations of parametric and empirical estimators of PHT and their asymptotic distributions for the Pareto I model are provided in Appendix A. Here we present a quick overview of those results. As derived in Example 3.8, the target parameter is $\mathrm{PHT}[F,r] = x_0 + x_0/(r\alpha-1)$, which is finite for $\alpha > 1/r$. Then, its ML and PM estimators are

$$\mathrm{PHT}[\widehat{F}_{\mathrm{ML}}, r] = x_0 + \frac{x_0}{r\widehat{\alpha}_{\mathrm{ML}}-1} \quad \text{and} \quad \mathrm{PHT}[\widehat{F}_{\mathrm{PM}}, r] = x_0 + \frac{x_0}{r\widehat{\alpha}_{\mathrm{PM}}-1}.$$

Therefore, the asymptotic distributions (4.9) and (4.10) become

$$\text{PHT}[\widehat{F}_{\text{ML}}, r] \;\sim\; \mathcal{AN}\left(\text{PHT}[F, r],\; \frac{1}{n}\,\frac{(r\alpha-1)^{-2}\text{PHT}^2[F,r]}{1-(d/u)^{\alpha}}\right)$$

and

$$\text{PHT}[\widehat{F}_{\text{PM}}, r] \;\sim\; \mathcal{AN}\left(\text{PHT}[F, r],\; \frac{1}{n}\,\frac{p_1(r\alpha-1)^{-2}\text{PHT}^2[F,r]}{(1-p_1)\log^2(1-p_1)}\right),$$

where $0 < r \leq 1$ and $0 < p_1 < 1-(d/u)^{\alpha}$. Note that $\text{PHT}[F, r] = x_0 + x_0/(r\alpha - 1)$ is finite for $\alpha > 1/r$. $\square$

Example 4.9. [Wang Transform]

More detailed derivations of parametric and empirical estimators of WT and their asymptotic distributions for the Pareto I model are provided in Appendix A. Here we present a quick overview of those results. As derived in Example 3.11, the target parameter is $\text{WT}[F, \lambda] = x_0 + \frac{x_0}{\alpha}\, C_{\text{WT}}(\lambda, \alpha)$, where $C_{\text{WT}}(\lambda, \alpha) = \int_{-\infty}^{\infty} \Phi(z+\lambda)\, \frac{\varphi(z)}{[\Phi(z)]^{1/\alpha+1}}\; dz$ is finite for $\alpha > 1$. Then, its ML and PM estimators are

$$\text{WT}[\widehat{F}_{\text{ML}}, \lambda] = x_0 + \frac{x_0}{\widehat{\alpha}_{\text{ML}}}\, C_{\text{WT}}(\lambda, \widehat{\alpha}_{\text{ML}}) \quad \text{and} \quad \text{WT}[\widehat{F}_{\text{PM}}, \lambda] = x_0 + \frac{x_0}{\widehat{\alpha}_{\text{PM}}}\, C_{\text{WT}}(\lambda, \widehat{\alpha}_{\text{PM}}).$$

Therefore, the asymptotic distributions (4.9) and (4.10) become

$$\text{WT}[\widehat{F}_{\text{ML}}, \lambda] \;\sim\; \mathcal{AN}\left(\text{WT}[F, \lambda],\; \frac{\alpha^{-2}}{n}\,\frac{x_0^2\, D_{\text{WT}}^2(\lambda,\alpha)}{1-(d/u)^{\alpha}}\right)$$

and

$$\text{WT}[\widehat{F}_{\text{PM}}, \lambda] \;\sim\; \mathcal{AN}\left(\text{WT}[F, \lambda],\; \frac{\alpha^{-2}}{n}\,\frac{p_1\, x_0^2\, D_{\text{WT}}^2(\lambda,\alpha)}{(1-p_1)\log^2(1-p_1)}\right),$$

where $-\infty < \lambda < \infty$ and $0 < p_1 < 1-(d/u)^{\alpha}$. Note that $\text{WT}[F, \lambda] = x_0 + \frac{x_0}{\alpha}\, C_{\text{WT}}(\lambda, \alpha)$ and $D_{\text{WT}}(\lambda, \alpha) = -C_{\text{WT}}(\lambda, \alpha) + \alpha^{-1}\int_{-\infty}^{\infty} \Phi(z+\lambda)\, \frac{\varphi(z)}{[\Phi(z)]^{1/\alpha+1}}\, \log\left[\Phi(z)\right]\; dz$ are finite for $\alpha > 1$. $\square$

Example 4.10. [Gini Shortfall]

More detailed derivations of parametric and empirical estimators of PHT and their asymptotic distributions for the Pareto I model are provided in Appendix A. Here we present a quick overview of those results. As derived in Example 3.14, the target parameter is

$$\text{GS}[F,\beta,\delta] = x_0\beta^{-1/\alpha}\alpha(\alpha-1)^{-1}\left(1+2\delta(2\alpha-1)^{-1}\right),$$

which is finite for $\alpha > 1$. Then, its ML and PM estimators are

$$\text{GS}[\widehat{F}_{\text{ML}},\beta,\delta] = x_0\beta^{-1/\widehat{\alpha}_{\text{ML}}}\widehat{\alpha}_{\text{ML}}(\widehat{\alpha}_{\text{ML}}-1)^{-1}\left(1+2\delta(2\widehat{\alpha}_{\text{ML}}-1)^{-1}\right)$$

and

$$\text{GS}[\widehat{F}_{\text{PM}},\beta,\delta] = x_0\beta^{-1/\widehat{\alpha}_{\text{PM}}}\widehat{\alpha}_{\text{PM}}(\widehat{\alpha}_{\text{PM}}-1)^{-1}\left(1+2\delta(2\widehat{\alpha}_{\text{PM}}-1)^{-1}\right).$$

Therefore, the asymptotic distributions (4.9) and (4.10) become

$$\text{GS}[\widehat{F}_{\text{ML}},\beta,\delta] \;\sim\; \mathcal{AN}\left(\text{GS}[F,\beta,\delta],\ \frac{\alpha^{-2}}{n}\,\frac{D^2_{\text{GS}}(\beta,\delta,\alpha)\,\text{GS}^2[F,\beta,\delta]}{1-(d/u)^{\alpha}}\right)$$

and

$$\text{GS}[\widehat{F}_{\text{PM}},\beta,\delta] \;\sim\; \mathcal{AN}\left(\text{GS}[F,\beta,\delta],\ \frac{\alpha^{-2}}{n}\,\frac{p_1 D^2_{\text{GS}}(\beta,\delta,\alpha)\,\text{GS}^2[F,\beta,\delta]}{(1-p_1)\log^2(1-p_1)}\right)$$

where $0 < \beta < 1$, $0 \le \delta \le 1/2$, and $0 < p_1 < 1-(d/u)^{\alpha}$. Note that $\text{GS}[F,\beta,\delta] = x_0\beta^{-1/\alpha}\alpha(\alpha-1)^{-1}\left(1+2\delta(2\alpha-1)^{-1}\right)$ and $D_{\text{GS}}(\beta,\delta,\alpha) = \log(\beta)-\alpha(\alpha-1)^{-1}-4\alpha^2\delta(2\alpha-1)^{-1}(2\delta+2\alpha-1)^{-1}$ are both finite for $\alpha > 1$. □

4.3.3 Shifted Lognormal Distribution

The shifted lognormal distribution plays a major role in modeling claim severity data, but it has two unknown parameters, which makes methodological derivations considerably more complicated. Thus, in this section we will develop risk measure estimators using only the ML approach, leaving PM estimation for future investigations.

Let us start by recalling the abbreviations introduced in Section 2.2.3:

$$c_d := \frac{\log(d-x_0)-\mu}{\sigma}, \qquad c_{x_i^*} := \frac{\log(x_i^*-x_0)-\mu}{\sigma}, \qquad c_u := \frac{\log(u-x_0)-\mu}{\sigma}.$$

Using this notation, if $x_1^*, \ldots, x_n^*$ is a realization of variables (2.2) with pdf (2.13) and cdf (2.12), then the log-likelihood function (4.8) becomes

$$\begin{aligned}\log \mathcal{L}\big((\mu,\sigma)\,\big|\,x_1^*,\ldots,x_n^*\big) &= \sum_{i=1}^{n} \big[-\log(\sigma) - \log(x_i^* - x_0) + \log(\varphi(c_{x_i^*}))\big]\,\mathbf{1}\big\{d < x_i^* < u\big\} \\ &\quad - n\log\left[1-\Phi(c_d)\right] + \log\left[1-\Phi(c_u)\right]\sum_{i=1}^{n}\mathbf{1}\big\{x_i^* = u\big\}.\end{aligned}$$

Differentiation of $\log\mathcal{L}$ with respect to μ and σ, along with some straightforward simplifications, yields the following system of equations:

$$\begin{cases} \displaystyle\sum_{i=1}^{n} c_{x_i^*}\mathbf{1}\big\{d < x_i^* < u\big\} - \frac{n\,\varphi(c_d)}{1-\Phi(c_d)} + \frac{\varphi(c_u)}{1-\Phi(c_u)}\sum_{i=1}^{n}\mathbf{1}\big\{x_i^* = u\big\} = 0 \\ \displaystyle\sum_{i=1}^{n} \big(c_{x_i^*}^2 - 1\big)\mathbf{1}\big\{d < x_i^* < u\big\} - \frac{n\,\varphi(c_d)\,c_d}{1-\Phi(c_d)} + \frac{\varphi(c_u)\,c_u}{1-\Phi(c_u)}\sum_{i=1}^{n}\mathbf{1}\big\{x_i^* = u\big\} = 0 \end{cases} \tag{4.19}$$

The system of equations (4.19) has to be solved numerically. Assuming solution exists, it will be denoted $(\widehat{\mu}_{\mathrm{ML}}, \widehat{\sigma}_{\mathrm{ML}})$; its asymptotic distribution is given by Theorem B.2. In this case, the Fisher information matrix has the following entries (more details are provided in Appendix A):

$$\begin{aligned} I_{11} &= -\mathbf{E}\left[\frac{\partial^2 \log f_*(X^*)}{\partial\mu^2}\right] \\ &= \sigma^{-2}\left[\frac{\Phi(c_u)-\Phi(c_d)}{1-\Phi(c_d)} + \frac{[1-\Phi(c_d)]\varphi(c_d)c_d - \varphi^2(c_d)}{[1-\Phi(c_d)]^2} - \frac{[1-\Phi(c_u)]\varphi(c_u)c_u - \varphi^2(c_u)}{[1-\Phi(c_u)][1-\Phi(c_d)]}\right] \\ I_{12} &= I_{21} = -\mathbf{E}\left[\frac{\partial^2 \log f_*(X^*)}{\partial\mu\,\partial\sigma}\right] \\ &= \sigma^{-2}\left[\frac{[1-\Phi(c_d)]\varphi(c_d)(c_d^2+1) - \varphi^2(c_d)c_d}{[1-\Phi(c_d)]^2} - \frac{[1-\Phi(c_u)]\varphi(c_u)(c_u^2+1) - \varphi^2(c_u)c_u}{[1-\Phi(c_d)][1-\Phi(c_u)]}\right] \\ I_{22} &= -\mathbf{E}\left[\frac{\partial^2 \log f_*(X^*)}{\partial\sigma^2}\right] \\ &= \sigma^{-2}\left[\frac{2\left[\Phi(c_u)-\Phi(c_d)\right]}{1-\Phi(c_d)} + \frac{[1-\Phi(c_d)]\varphi(c_d)c_d(c_d^2+1) - \varphi^2(c_d)c_d^2}{[1-\Phi(c_d)]^2}\right. \\ &\qquad \left. - \frac{[1-\Phi(c_u)]\varphi(c_u)c_u(c_u^2+1) - \varphi^2(c_u)c_u^2}{[1-\Phi(c_d)][1-\Phi(c_u)]}\right] \end{aligned}$$

Hence, the estimator $(\widehat{\mu}_{\mathrm{ML}}, \widehat{\sigma}_{\mathrm{ML}})$, found by solving (4.19), has the following asymptotic distribution:

$$(\widehat{\mu}_{\mathrm{ML}}, \widehat{\sigma}_{\mathrm{ML}}) \;\sim\; \mathcal{AN}\left((\mu, \sigma),\, \frac{1}{n} \begin{bmatrix} \sigma_{11}^* & \sigma_{12}^* \\ \sigma_{21}^* & \sigma_{22}^* \end{bmatrix} \right), \tag{4.20}$$

where $\sigma_{11}^* = I_{22}\left(I_{11}I_{22} - I_{12}^2\right)^{-1}$, $\sigma_{12}^* = \sigma_{21}^* = -I_{12}\left(I_{11}I_{22} - I_{12}^2\right)^{-1}$, $\sigma_{22}^* = I_{11}\left(I_{11}I_{22} - I_{12}^2\right)^{-1}$, with the terms I_{ij}, $i, j = 1, 2$, specified above.

ML estimators of VaR, CTE, PHT, WT, and GS are computed by plugging in $(\widehat{\mu}_{\mathrm{ML}}, \widehat{\sigma}_{\mathrm{ML}})$, found by solving (4.19), into the parametric expression of $R[F]$. Asymptotic distributions of such estimators are derived by applying the delta method to (4.20):

$$R[\widehat{F}_{\mathrm{ML}}] \;\sim\; \mathcal{AN}\left(R[F],\, \frac{1}{n}\,(d_1, d_2) \begin{bmatrix} \sigma_{11}^* & \sigma_{12}^* \\ \sigma_{21}^* & \sigma_{22}^* \end{bmatrix} (d_1, d_2)' \right), \tag{4.21}$$

where the partial derivatives of the risk measure $R[F]$ with respect to μ and σ, denoted $d_1 = \partial R[F]/\partial\mu$ and $d_2 = \partial R[F]/\partial\sigma$, have the following expressions:

- $\mathrm{VaR}[F, \beta] = x_0 + \exp\{\mu + \sigma\Phi^{-1}(1-\beta)\}$:

$$\begin{aligned} d_1 &= \frac{\partial\,\mathrm{VaR}[F,\beta]}{\partial\mu} &=& \;\mathrm{VaR}[F,\beta] - x_0, \\ d_2 &= \frac{\partial\,\mathrm{VaR}[F,\beta]}{\partial\sigma} &=& \;\big(\mathrm{VaR}[F,\beta] - x_0\big)\,\Phi^{-1}(1-\beta). \end{aligned}$$

- $\mathrm{CTE}[F, \beta] = x_0 + \beta^{-1}\exp\{\mu + \sigma^2/2\}\,\Phi\left(\sigma - \Phi^{-1}(1-\beta)\right)$:

$$\begin{aligned} d_1 &= \frac{\partial\,\mathrm{CTE}[F,\beta]}{\partial\mu} &=& \;\mathrm{CTE}[F,\beta] - x_0, \\ d_2 &= \frac{\partial\,\mathrm{CTE}[F,\beta]}{\partial\sigma} &=& \;\sigma\big(\mathrm{CTE}[F,\beta] - x_0\big) + e^{\mu} D_{\mathrm{CTE}}(\beta,\sigma), \end{aligned}$$

  where $D_{\mathrm{CTE}}(\beta,\sigma) = \beta^{-1} e^{\sigma^2/2}\,\varphi\left(\sigma - \Phi^{-1}(1-\beta)\right)$.

- $\textsc{pht}[F,r] = x_0 + e^{\mu}\, C_{\textsc{pht}}(r,\sigma) = x_0 + e^{\mu}\,\sigma \int_{-\infty}^{\infty} \big[1-\Phi(z)\big]^r\, e^{\sigma z}\, dz$:

$$\begin{aligned} d_1 &= \frac{\partial\, \textsc{pht}[F,r]}{\partial \mu} &=& \ \textsc{pht}[F,\beta] - x_0, \\ d_2 &= \frac{\partial\, \textsc{pht}[F,r]}{\partial \sigma} &=& \ \sigma^{-1}\big(\textsc{pht}[F,r] - x_0\big) + e^{\mu} D_{\textsc{pht}}(r,\sigma), \end{aligned}$$

where $D_{\textsc{pht}}(r,\sigma) = \sigma \int_{-\infty}^{\infty} \big[1-\Phi(z)\big]^r z\, e^{\sigma z}\, dz$ can be evaluated numerically.

- $\textsc{wt}[F,\lambda] = x_0 + \exp\{\mu + \lambda\sigma + \sigma^2/2\}$:

$$\begin{aligned} d_1 &= \frac{\partial\, \textsc{wt}[F,\lambda]}{\partial \mu} &=& \ \textsc{wt}[F,\lambda] - x_0, \\ d_2 &= \frac{\partial\, \textsc{wt}[F,\lambda]}{\partial \sigma} &=& \ (\lambda+\sigma)\big(\textsc{wt}[F,\lambda] - x_0\big). \end{aligned}$$

- $\textsc{gs}[F,\beta,\delta] = x_0 + \beta^{-2} e^{\mu+\sigma^2/2}\Big([\beta(1+2\delta) - 4\delta]\ \Phi\left(\sigma - \Phi^{-1}(1-\beta)\right) + 4\delta\, C_{\textsc{gs}}(\beta,\sigma)\Big)$, where $C_{\textsc{gs}}(\beta,\sigma) = \int_{\Phi^{-1}(1-\beta)}^{\infty} \Phi(z)\, \varphi(z-\sigma)\, dz$:

$$\begin{aligned} d_1 &= \frac{\partial\, \textsc{gs}[F,\beta,\delta]}{\partial \mu} &=& \ \textsc{gs}[F,\beta,\delta] - x_0, \\ d_2 &= \frac{\partial\, \textsc{gs}[F,\beta,\delta]}{\partial \sigma} &=& \ \sigma\big(\textsc{gs}[F,\beta,\delta] - x_0\big) + e^{\mu}\left((1-2\delta) D_{\textsc{gs}}(\beta,\sigma) + \delta\,\sqrt{4/\pi}\, D^*_{\textsc{gs}}(\beta,\sigma)\right), \end{aligned}$$

where $D_{\textsc{gs}}(\beta,\sigma) = D_{\textsc{cte}}(\beta,\sigma)$ and $D^*_{\textsc{gs}}(\beta,\sigma) = \beta^{-2} e^{\sigma^2/4}\, \Phi\left(\frac{\sigma/2 - \Phi^{-1}(1-\beta)}{\sqrt{1/2}}\right)$.

Chapter 5

Numerical Illustrations

In this chapter, we supplement theoretical studies on distortion risk measure estimation with numerical illustrations. In Section 5.1, a simulation study is performed for selected risk measures and severity distributions, with the primary objective to cross-validate the asymptotic distributions of Section 4.3. In Section 5.2, we fit Pareto I and lognormal distributions to the well-known Norwegian fire claims data, evaluate the fits, and estimate the upper-tail riskiness of these claims for the year 1986.

5.1 Simulated Data

A Monte Carlo simulation study was conducted to verify and augment the asymptotic properties derived in Section 4.2. The study was performed for the following choices of simulation parameters:

- *Severity distributions*:

  $F_E = \mathcal{E}xp\,(x_0 = 10^3,\, \theta = 10^3)$; $F_P = \mathcal{P}a\,I\,(x_0 = 10^3, \alpha = 2.0)$.

- *Risk measures and their values* (measured in 1000's):

  $\mathrm{VaR}[F_E, \beta = 0.10] = 3.30$; $\mathrm{VaR}[F_P, \beta = 0.10] = 3.16$.

  $\textsc{cte}[F_E, \beta = 0.10] = 4.30$; $\textsc{cte}[F_P, \beta = 0.10] = 6.32$.

  $\textsc{gs}[F_E, \beta = 0.10, \delta = 0.25] = 4.55$; $\textsc{gs}[F_P, \beta = 0.10, \delta = 0.25] = 7.38$.

$\text{PHT}[F_E, r = 0.75] = 2.33$; $\text{PHT}[F_P, r = 0.75] = 3.00$.

$\text{WT}[F_E, \lambda = 0.50] = 2.53$; $\text{WT}[F_P, \lambda = 0.50] = 3.07$.

- *Risk measure estimators*: EMP; ML; PM (with $p_1 = 0.80$).

- *Truncation and censoring thresholds*:

  $d = 4 \cdot 10^3$ (corresponds to the 95.0% data truncation under $\mathcal{E}xp\,(x_0 = 10^3, \theta = 10^3)$ and 93.8% under $\mathcal{P}a\,I\,(x_0 = 10^3, \alpha = 2.0)$);

  $u = 14 \cdot 10^3$ (corresponds to the 0.0045% data censoring under $\mathcal{E}xp\,(d = 4 \cdot 10^3, \theta = 10^3)$ and 8.2% under $\mathcal{P}a\,I\,(d = 4 \cdot 10^3, \alpha = 2.0)$).

- *Sample size*: $n = 50$, 100, 500.

From a specified left-truncated and right-censored severity distribution, we generate 100,000 samples of a specified length n. For each sample, we estimate the risk measures VaR, CTE, PHT, WT, and GS according to their formulas derived in Examples 3.1–3.15. Then, based on those 100,000 risk measure estimates, we compute their mean and standard deviation. Simulation findings are summarized in Table 5.1, where the columns $n \to \infty$ correspond to the asymptotic mean and standard deviation of the estimator, which were derived in Chapter 4 and are included here as reference point.

Several conclusions emerge from the table. First, all finite sample estimates converge to the theoretical large-sample counterparts and thus validate the asymptotic distributions derived in Section 4.3. Second, empirical estimators are overestimating their targets, which was shown theoretically and is now checked via simulations. Third, the known ordering among some risk measures has also been confirmed. For example, for a fixed loss model and risk appetite parameter, the VaR measure is less than CTE which in turn is less than GS. Finally, the study parameters were chosen so that both loss models have the same mean (it is 2000), but Pareto I model has heavier right tail and thus its risk measure estimates are larger than those of the shifted exponential distribution (except for VaR).

Table 5.1: Means and standard deviations of ML, PM, and EMP estimators of various risk measures for $F_E = \mathcal{E}xp\,(x_0 = 10^3,\, \theta = 10^3)$ and $F_P = \mathcal{P}a\,I\,(x_0 = 10^3, \alpha = 2.0)$ distributions.

Risk Measure	Method of Estimation	Mean				Std. Deviation $(\times n^{1/2})$			
		$n = 50$	$n = 100$	$n = 500$	$n \to \infty$	$n = 50$	$n = 100$	$n = 500$	$n \to \infty$
VaR$[F_E]$	ML	3.31	3.30	3.30	3.30	2.32	2.30	2.31	2.30
	PM	3.25	3.27	3.30	3.30	2.78	2.80	2.87	2.86
	EMP	6.22	6.25	6.29	6.30	2.89	2.92	2.98	3.00
CTE$[F_E]$	ML	4.31	4.30	4.30	4.30	3.33	3.30	3.31	3.30
	PM	4.23	4.26	4.29	4.30	3.98	4.02	4.12	4.10
	EMP	7.22	7.26	7.30	7.25	4.29	4.32	4.32	4.35
GS$[F_E]$	ML	4.56	4.55	4.55	4.55	3.59	3.55	3.56	3.55
	PM	4.47	4.50	4.54	4.55	4.29	4.32	4.43	4.41
	EMP	7.42	7.48	7.54	7.47	4.79	4.89	4.93	4.97
PHT$[F_E]$	ML	2.34	2.33	2.33	2.33	1.35	1.33	1.34	1.33
	PM	2.30	2.31	2.33	2.33	1.61	1.62	1.66	1.66
	EMP	5.30	5.31	5.33	5.30	1.35	1.36	1.38	1.39
WT$[F_E]$	ML	2.53	2.53	2.53	2.53	1.54	1.53	1.53	1.53
	PM	2.50	2.51	2.53	2.53	1.85	1.86	1.91	1.90
	EMP	5.50	5.51	5.53	5.50	1.58	1.58	1.60	1.60
VaR$[F_P]$	ML	3.22	3.19	3.17	3.16	4.09	3.96	3.85	3.80
	PM	3.13	3.14	3.16	3.16	4.55	4.56	4.55	4.52
	EMP	11.86	12.22	12.59	12.65	12.33	13.83	17.52	18.97
CTE$[F_P]$	ML	6.84	6.56	6.38	6.32	18.90	16.18	14.65	14.20
	PM	6.56	6.42	6.36	6.32	20.91	18.92	17.32	16.91
	EMP	13.46	13.63	13.82	13.87	5.37	4.82	3.98	3.80
GS$[F_P]$	ML	8.08	7.70	7.45	7.38	24.78	20.83	18.65	18.03
	PM	7.74	7.53	7.43	7.38	27.41	24.41	22.07	21.47
	EMP	13.60	13.75	13.89	13.93	4.38	3.60	2.56	2.26
PHT$[F_P]$	ML	3.87	3.18	3.03	3.00	206.26	9.20	6.67	6.26
	PM	3.73	3.20	3.03	3.00	111.34	31.60	7.93	7.46
	EMP	7.69	7.71	7.72	7.72	3.31	3.34	3.34	3.32
WT$[F_P]$	ML	3.28	3.17	3.09	3.07	7.44	6.52	5.71	5.50
	PM	3.16	3.12	3.08	3.07	7.88	7.70	6.76	6.54
	EMP	8.31	8.34	8.36	8.35	3.85	3.89	3.89	3.87

NOTE: The entries for $n < \infty$ are based on 100,000 simulated samples. All entries are measured in 1000's.

5.2 Norwegian Fire Claims

In this section, we fit left-truncated Pareto I and lognormal distributions to the well-studied Norwegian fire claims data (see Nadarajah and Bakar, 2015; Brazauskas and Kleefeld, 2016), which are available at the following website:

`http://lstat.kuleuven.be/Wiley` (in Chapter 1, file NORWEGIANFIRE.TXT).

The data represent the total damage done by fires in Norway for the years 1972 through 1992; only damages in excess of a reinsurance priority of 500,000 Norwegian krones (NOK) are available. We will analyze the data set for the year 1986, which, as shown in Section 5.1 of Brazauskas and Kleefeld (2016), exhibited unusual sensitivity to the choice of the underlying loss model when used in VaR-measure calculations. The data set has $n = 647$ claims, with the three largest observations being 87, 98, and 188 (measured in millions of NOK). A summary of these data is provided in Table 5.2.

Table 5.2: Summary of *Norwegian Fire Claims* data for the year 1986.

Severity (millions NOK)	[0.5; 1.0)	[1.0; 2.0)	[2.0; 5.0)	[5.0; 10.0)	[10.0; 20.0)	[20.0; ∞)
Relative Frequency	0.507	0.312	0.119	0.032	0.017	0.012

Since no information is given below 500,000 and there is no policy limit, the random variable that generated data (using our notation, X^* defined by (2.2)) is left-truncated at $d = 500,000$ but not censored, i.e., $u = \infty$. Moreover, as is evident from Table 5.2, the data are right-skewed and heavy-tailed suggesting that Pareto I or lognormal might be appropriate models in this case.

In Figure 5.1, we present plots of the fitted-versus-observed quantiles for the Pareto I and lognormal models (both fitted using the ML approach; see equations (4.15) and (4.19)). In order to avoid visual distortions due to large spacings between the most extreme observations, both axes are measured on the logarithmic scale. That is, the points plotted in those graphs are the following pairs:

$$\left(\log\left(\widehat{F}^{-1}\big[u_i + \widehat{F}(d)(1-u_i)\big]\right),\ \log\left(x^*_{(i)}\right)\right), \quad i = 1, \ldots, n,$$

where $\widehat{F}(d)$ is the estimated parametric cdf evaluated at $d = 500,000$, $\widehat{F}^{-1}$ is the estimated parametric qf, $x^*_{(1)} < \cdots < x^*_{(n)}$ denote the ordered claim severities, $u_i = (i - 0.5)/n$ is the quantile level, and $n = 647$ is the sample size.

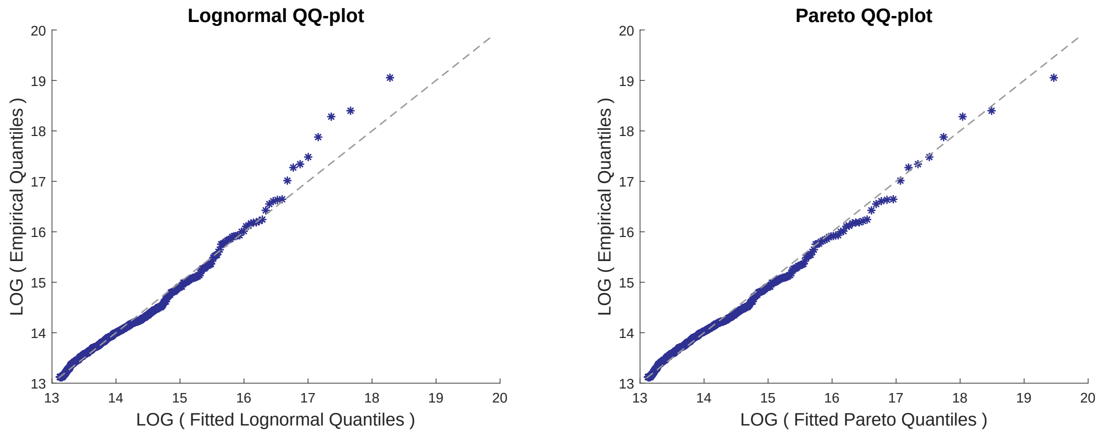

Figure 5.1: Quantile-quantile plots for *Norwegian Fire Claims* data based on $\mathcal{LN}\,(x_0 = 10^5, \widehat{\mu}_{\mathrm{ML}} = 9.7524, \widehat{\sigma}_{\mathrm{ML}} = 2.2174)$ and $\mathcal{P}a\,I\,(x_0 = 10^5, \widehat{\alpha}_{\mathrm{ML}} = 1.1270)$ models. Data and both models are left-truncated at $d = 500,000$.

Clearly, Pareto-estimated quantiles fall almost perfectly on the 45° line against the empirical quantiles. On the other hand, lognormal QQ-plot does not look as good, but note that there is only 8 observations clearly above the 45° line, which corresponds to about 1.2% of data (8 out of 647 observations). Of course, from a risk measuring perspective these 8 points are the worst because they correspond to the largest losses. Also, having top observations above the 45° line indicates that the lognormal model underestimates the right tail of the data.

In Table 5.3, we report point estimates and 90% confidence intervals of selected risk measures for Norwegian Fire Claims data. The risk measure estimates are computed using the fitted

$$\mathcal{P}a\,I\,(x_0 = 10^5, \widehat{\alpha}_{\mathrm{ML}} = 1.1270) \qquad \text{and} \qquad \mathcal{LN}\,(x_0 = 10^5, \widehat{\mu}_{\mathrm{ML}} = 9.7524, \widehat{\sigma}_{\mathrm{ML}} = 2.2174)$$

distributions. The corresponding confidence intervals are constructed using the asymptotic distributions of Examples 4.6-4.10 and equation (4.21).

Table 5.3: Point and interval estimates of selected risk measures for *Norwegian Fire Claims* data, based on $\mathcal{P}a\,I\,(x_0 = 10^5, \widehat{\alpha} = 1.1270)$ and $\mathcal{LN}\,(x_0 = 10^5, \widehat{\mu} = 9.7524, \widehat{\sigma} = 2.2174)$ models.

Risk Measure $R[F]$	$\mathcal{LN}$ distribution $R[\widehat{F}]$ (millions NOK)	$\mathcal{LN}$ distribution 90% CI (millions NOK)	$\mathcal{P}a\,I$ distribution $R[\widehat{F}]$ (millions NOK)	$\mathcal{P}a\,I$ distribution 90% CI (millions NOK)
$\mathrm{VaR}[F, \beta = 0.10]$	0.395	$[-0.139; 0.929]$	0.771	$[0.670; 0.873]$
$\mathrm{CTE}[F, \beta = 0.10]$	1.759	$[-0.070; 3.587]$	6.846	$[2.455; 11.237]$
$\mathrm{GS}[F, \beta = 0.10, \delta = 0.25]$	2.276	$[0.015; 4.536]$	9.576	$[3.117; 16.034]$
$\mathrm{PHT}[F, r = 0.95]$	0.332	$[0.066; 0.598]$	1.515	$[0.128; 2.903]$
$\mathrm{WT}[F, \lambda = 0.25]$	0.450	$[0.052; 0.848]$	2.149	$[0.329; 3.970]$

As was expected from Figure 5.1, the risk measure estimates based on the lognormal model are substantially below the corresponding estimates based on the Pareto model. Except for VaR, the lognormal confidence intervals are much shorter than those of Pareto, but the lower bound of its VaR and CTE intervals is negative and thus not informative. On the other hand, Pareto based intervals make sense but they are quite wide. This is surprising given that they are constructed using 647 observations.

Chapter 6

Concluding Remarks

6.1 Summary

In this dissertation, estimation of distortion risk measures under truncated and censored data scenarios has been studied for shifted exponential, Pareto I, and shifted lognormal loss variables. We considered five commonly used risk measures: value-at-risk (VaR), conditional tail expectation (CTE), proportional hazards transform (PHT), Wang transform (WT), and Gini shortfall (GS). We have constructed empirical (EMP), maximum likelihood (ML) and percentile matching (PM) type estimators for these risk measures and investigated their properties theoretically as well as via simulations. The estimators have also been applied to risk measurement exercises involving actual reinsurance data.

As the first contribution of the dissertation, we have derived several inequalities that established 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. Then the integrals were evaluated numerically.

The second contribution is development of the EMP, ML, and PM estimators of the five risk measures. Asymptotically normal distributions of these estimators have been derived, and their small-sample properties have been explored using Monte Carlo simulations. The simulation study revealed convergence of sample estimates to the true quantities as the sample size increased.

Finally, the third contribution is numerical illustrations based on the well-studied *Norwegian Fire Claims* data (for the year 1986). In particular, the newly developed tools have been used to evaluate the upper-tail riskiness of these claims. We have computed point estimates and constructed (asymptotic) 90% confidence intervals for VaR, CTE, PHT, WT, and GS risk measures.

6.2 Future Work

The research presented in this dissertation invites follow-up studies in several directions.

First, within the classes of lower and upper bounds for the integrals of Chapter 3 (see Theorems 3.1–3.4), 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 models to other popular probability distributions such as gamma, Weibull, generalized Pareto (as well as other types of Pareto), folded, and spliced models. In addition, to capture the riskiness of aggregate losses, the risk measure formulas for normal distribution is also needed.

Third, the PM estimators have been introduced and developed as an alternative to the ML estimators, with the expectation that they may simplify computations. And for some distributions they are computationally simpler. Focusing on the implementation of estimators, PM and other estimators could be pursued to estimate distortion risk measures. It would be especially interesting to develop robust risk measuring procedures.

Bibliography

[1] Albrecht, P. (2004). Risk measures. In *Encyclopedia of Actuarial Science* (B. Sundt and J. Teugels, eds.), volume **3**, 1493–1501; Wiley, London.

[2] Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, **9**(3), 203–228.

[3] Brazauskas, V., Jones, B., Puri, M., and Zitikis, R. (2007). Nested L-statistics and their use in comparing the riskiness of portfolios. *Scandinavian Actuarial Journal*, **107**(3), 162–179.

[4] Brazauskas, V., Jones, B., Puri, M., and Zitikis, R. (2008). Estimating conditional tail expectations with actuarial applications in view. *Journal of Statistical Planning and Inference*, **138**(11), 3590–3604.

[5] Brazauskas, V. and Kaiser, T. (2004). Discussion of "Empirical estimation of risk measures and related quantities" by Jones and Zitikis. *North American Actuarial Journal*, **8**(3), 114–117.

[6] Brazauskas, V. and Kleefeld, A. (2016). Modeling severity and measuring tail risk of Norwegian fire claims. *North American Actuarial Journal*, **20**(1), 1–16.

[7] Brazauskas, V. and Upretee, S. (2019). Model efficiency and uncertainty in quantile estimation of loss severity distributions. *Risks*, **7**(2), 16 pages, `doi:10.3390/risks7020055`.

[8] Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*, volume **1**, 3rd edition. Wiley, New York.

[9] Frees, E. (2017). Insurance portfolio risk retention. *North American Actuarial Journal*, **21**(4), 526–551.

[10] Furman, E. and Landsman, Z. (2006). Tail variance premium with applications for elliptical portfolio of risks. *ASTIN Bulletin*, **36**(2), 433–462.

[11] Furman, E., Wang, R., and Zitikis, R. (2017). Gini-type measures of risk and variability: Gini shortfall, capital allocations, and heavy-tailed risks. *Journal of Banking and Finance*, **83**, 70–84.

[12] Johnson, N.L., Kotz, S., and Balakrishnan, N. (1994). *Continuous Univariate Distributions*, volume **1**, 2nd edition. Wiley, New York.

[13] Jones, B.L., Puri, M.L., and Zitikis, R. (2006). Testing hypotheses about the equality of several risk measure values with applications in insurance. *Insurance: Mathematics and Economics*, **38**(2), 253–270.

[14] Jones, B.L. and Zitikis, R. (2003). Empirical estimation of risk measures and related quantities. *North American Actuarial Journal*, **7**(4), 44–54.

[15] Jones, B.L. and Zitikis, R. (2005). Testing for the order of risk measures: an application of L-statistics in actuarial science. *Metron*, **63**(2), 193–211.

[16] Jones, B.L. and Zitikis, R. (2007). Risk measures, distortion parameters, and their empirical estimation. *Insurance: Mathematics and Economics*, **41**(2), 279–297.

[17] Kaiser, T. and Brazauskas, V. (2006). Interval estimation of actuarial risk measures. *North American Actuarial Journal*, **10**(4), 249–268.

[18] Klugman, S.A., Panjer, H.H., and Willmot, G.E. (2012). *Loss Models: From Data to Decisions*, 4th edition. Wiley, New York.

[19] Lee, G.Y. (2017). General insurance deductible ratemaking. *North American Actuarial Journal*, **21**(4), 620–638.

[20] Nadarajah, S. and Bakar, S.A.A. (2015). New folded models for the log-transformed Norwegian fire claim data. *Communications in Statistics: Theory and Methods*, **44**(20), 4408–4440.

[21] Necir, A. and Meraghni, D. (2009). Empirical estimation of the proportional hazard premium for heavy-tailed claim amounts. *Insurance: Mathematics and Economics*, **45**(1), 49–58.

[22] Necir, A., Meraghni, D., and Meddi, F. (2007). Statistical estimate of the proportional hazard premium of loss. *Scandinavian Actuarial Journal*, **2007**(3), 147–161.

[23] Necir, A., Rassoul, A., and Zitikis, R. (2010). Estimating the conditional tail expectation in the case of heavy-tailed losses. *Journal of Probability and Statistics*, 1–17. `doi:10.1155/2010/596839`

[24] Samanthi, R., Wei, W., and Brazauskas, V. (2017). Comparing the riskiness of dependent portfolios via nested L-statistics. *Annals of Actuarial Science*, **11**(2), 237–252.

[25] Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

[26] Tapiero, C.S. (2004). Risk management: An interdisciplinary framework. In *Encyclopedia of Actuarial Science* (B. Sundt and J. Teugels, eds.), volume **3**, 1483–1493; Wiley, London.

[27] Wang, S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics*, **17**(1), 43–54.

[28] Wang, S. (1998a). An actuarial index of the right-tail risk. *North American Actuarial Journal*, **2**(2), 88–101.

[29] Wang, S. (1998b). Implementation of proportional hazards transforms in ratemaking. *Proceedings of the Casualty Actuarial Society*, **LXXXV**(1-2), 940–979.

[30] Wang, S. (2000). A class of distortion operators for pricing financial and insurance risks. *Journal of Risk and Insurance*, **67**(1), 15–36.

[31] Wang, S. (2002). A universal framework for pricing financial and insurance risks. *ASTIN Bulletin*, **32**(2), 213–234.

[32] Young, V.R. (2004). Premium principles. In *Encyclopedia of Actuarial Science* (B. Sundt and J. Teugels, eds.), volume **3**, 1322–1331; Wiley, London.

Appendix A

Derivations

In this appendix, we provide detailed derivations of risk measure formulas of Chapter 3 and show specific steps in deriving asymptotic properties of the ML and PM estimators of $R[F]$ (as presented in Chapter 4). The derivations involve either (sometimes tricky) integration or differentiation.

A.1 Derivations of Chapter 3

Example 3.4. [CTE of Shifted Exponential]

Since $\int_{1-\beta}^{1} \log(1-u)\, du = \beta(\log(\beta) - 1)$, the following steps are easily verified:

$$
\begin{aligned}
\textsc{cte}[F,\beta] &= \frac{1}{\beta}\int_{1-\beta}^{1} F^{-1}(u)\, du \;=\; \frac{1}{\beta}\int_{1-\beta}^{1} \big[x_0 - \theta\log(1-u)\big]\, du \\
&= x_0 - \frac{\theta}{\beta}\int_{1-\beta}^{1} \log(1-u)\, du \;=\; x_0 - \theta(\log(\beta) - 1).
\end{aligned}
$$

Example 3.5. [CTE of Pareto I]

For $\alpha > 1$, we have $\int_{1-\beta}^{1} (1-u)^{-1/\alpha}\, du = \alpha(\alpha-1)^{-1}\beta^{-1/\alpha+1}$; therefore

$$
\textsc{cte}[F,\beta] \;=\; \frac{1}{\beta}\int_{1-\beta}^{1} F^{-1}(u)\, du \;=\; \frac{1}{\beta}\int_{1-\beta}^{1} \big[x_0(1-u)^{-1/\alpha}\big]\, du \;=\; x_0\beta^{-1/\alpha}\alpha(\alpha-1)^{-1}.
$$

Example 3.6. [CTE of Shifted Lognormal]

Using substitution $v = \Phi^{-1}(u)$, we have

$$\begin{aligned}\int_{1-\beta}^{1} e^{\sigma\Phi^{-1}(u)}\,du &= \int_{\Phi^{-1}(1-\beta)}^{\infty} e^{\sigma v}\varphi(v)\,dv = e^{\sigma^2/2}\int_{\Phi^{-1}(1-\beta)}^{\infty}\varphi(v-\sigma)\,dv \\ &= e^{\sigma^2/2}\big[1-\Phi\big(\Phi^{-1}(1-\beta)-\sigma\big)\big] = e^{\sigma^2/2}\,\Phi\left(\sigma-\Phi^{-1}(1-\beta)\right).\end{aligned}$$

Now the following steps are easily verified:

$$\begin{aligned}\text{CTE}[F,\beta] &= \frac{1}{\beta}\int_{1-\beta}^{1} F^{-1}(u)\,du = \frac{1}{\beta}\int_{1-\beta}^{1}\big[x_0+e^{\mu+\sigma\Phi^{-1}(u)}\big]\,du \\ &= x_0+\frac{e^{\mu}}{\beta}\int_{1-\beta}^{1} e^{\sigma\Phi^{-1}(u)}\,du = x_0+\frac{1}{\beta}\,e^{\mu+\sigma^2/2}\,\Phi\left(\sigma-\Phi^{-1}(1-\beta)\right).\end{aligned}$$

Example 3.7. [PHT of Shifted Exponential]

Since $\int_0^1(1-u)^{r-1}\,du = 1/r$ and $\int_0^1(1-u)^{r-1}\log(1-u)\,du = -1/r^2$, the following steps are easily verified:

$$\begin{aligned}\text{PHT}[F,r] &= r\int_0^1 F^{-1}(u)(1-u)^{r-1}\,du = r\int_0^1\big[x_0-\theta\log(1-u)\big](1-u)^{r-1}\,du \\ &= rx_0\int_0^1(1-u)^{r-1}\,du - r\theta\int_0^1(1-u)^{r-1}\log(1-u)\,du = x_0+\theta/r.\end{aligned}$$

Example 3.8. [PHT of Pareto I]

For $\alpha > 1/r$, we have $\int_0^1(1-u)^{r-1-1/\alpha}\,du = \alpha(r\alpha-1)^{-1}$; therefore

$$\begin{aligned}\text{PHT}[F,r] &= r\int_0^1 F^{-1}(u)(1-u)^{r-1}\,du = r\int_0^1\big[x_0(1-u)^{-1/\alpha}\big](1-u)^{r-1}\,du \\ &= rx_0\alpha(r\alpha-1)^{-1} = x_0+\frac{x_0}{r\alpha-1}.\end{aligned}$$

Example 3.9. [PHT of Shifted Lognormal]

Substitution $z = \Phi^{-1}(u)$, integration by parts and $\lim_{z\to\infty} e^{\sigma z}(1-\Phi(z))^r = 0$ for $r > 0$, lead to

$$\begin{aligned}
r\int_0^1 (1-u)^{r-1} e^{\sigma\Phi^{-1}(u)}\,du &= r\int_{-\infty}^{\infty}(1-\Phi(z))^{r-1}e^{\sigma z}\varphi(z)\,dz \\
&= -\left[e^{\sigma z}(1-\Phi(z))^r\Big|_{-\infty}^{\infty} - \sigma\int_{-\infty}^{\infty}(1-\Phi(z))^r e^{\sigma z}\,dz\right] \\
&= \sigma\int_{-\infty}^{\infty}(1-\Phi(z))^r e^{\sigma z}\,dz.
\end{aligned}$$

Now using this result and $\int_0^1 (1-u)^{r-1}\,du = 1/r$ (see Example 3.7), we have

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

Example 3.10. [WT of Shifted Exponential]

Using substitution $v = \Phi^{-1}(u)$, we have

$$\int_0^1 e^{\lambda\Phi^{-1}(u)-\lambda^2/2}\,du \;=\; \int_{-\infty}^{\infty} e^{\lambda v-\lambda^2/2}\,\varphi(v)\,dv \;=\; \int_{-\infty}^{\infty}\varphi(v-\lambda)\,dv \;=\; 1.$$

Using substitution $z = -\Phi^{-1}(u)$ and integration by parts, we have

$$\begin{aligned}
\int_0^1 \log(1-u)\,e^{\lambda\Phi^{-1}(u)-\lambda^2/2}\,du &= \int_{-\infty}^{\infty}\log(1-\Phi(-z))\,e^{-\lambda z-\lambda^2/2}\,\varphi(-z)\,dz \\
&= \int_{-\infty}^{\infty}\log(\Phi(z))\,\varphi(z+\lambda)\,dz \;=\; -\int_{-\infty}^{\infty}\Phi(z+\lambda)\,\frac{\varphi(z)}{\Phi(z)}\,dz.
\end{aligned}$$

Now the following steps are easily verified:

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

Example 3.11. [WT of PARETO I]

Using substitution $z = -\Phi^{-1}(u)$ and integration by parts, we have

$$\begin{aligned}
\int_0^1 (1-u)^{-1/\alpha}\, e^{\lambda\Phi^{-1}(u)-\lambda^2/2}\, du &= \int_{-\infty}^{\infty} (1-\Phi(-z))^{-1/\alpha}\, e^{-\lambda z-\lambda^2/2}\, \varphi(-z)\, dz \\
&= \int_{-\infty}^{\infty} \big[\Phi(z)\big]^{-1/\alpha}\, \varphi(z+\lambda)\, dz \\
&= \big[\Phi(z)\big]^{-1/\alpha}\Phi(z+\lambda)\Big|_{-\infty}^{\infty} \;+\; \frac{1}{\alpha}\int_{-\infty}^{\infty} \Phi(z+\lambda)\, \frac{\varphi(z)}{\big[\Phi(z)\big]^{1/\alpha+1}}\, dz \\
&= 1 + \frac{1}{\alpha}\int_{-\infty}^{\infty} \Phi(z+\lambda)\, \frac{\varphi(z)}{\big[\Phi(z)\big]^{1/\alpha+1}}\, dz,
\end{aligned}$$

where in the last step we assume that $\alpha > 1$, which ensures that $\lim_{z\to-\infty} \big[\Phi(z)\big]^{-1/\alpha}\Phi(z+\lambda) = 0$. This limit can be found by first applying the inequalities of Lemma B.6.

- For $\lambda \le 0$ and $z < 0$, we have: $0 \;\le\; \big[\Phi(z)\big]^{-1/\alpha}\Phi(z+\lambda) \;\le\; e^{-\lambda z-\lambda^2/2}\big[\Phi(z)\big]^{1-1/\alpha}$.
- For $\lambda > 0$ and $z < -\lambda$, we have $0 \;\le\; \big[\Phi(z)\big]^{-1/\alpha}\Phi(z+\lambda) \;<\; \dfrac{z}{z+\lambda}e^{-\lambda z-\lambda^2/2}\big[\Phi(z)\big]^{1-1/\alpha}$.

Then taking the limit of these inequalities as $z \to -\infty$ yields the result.

Now the following steps are easily verified:

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

Example 3.12. [WT of SHIFTED LOGNORMAL]

As was shown in derivations of Example 3.10, $\int_0^1 e^{\lambda\Phi^{-1}(u)-\lambda^2/2}\,du = 1$. In addition, using substitution $z = \Phi^{-1}(u)$ and noticing that the resulting integral is the moment generating function of the standard normal distribution (evaluated at $\sigma + \lambda$), we have

$$\begin{aligned}
\mathrm{WT}[F,\lambda] &= \int_0^1 F^{-1}(u)\, e^{\lambda\Phi^{-1}(u)-\lambda^2/2}\,du \;=\; \int_0^1 \left[x_0 + e^{\mu+\sigma\Phi^{-1}(u)}\right] e^{\lambda\Phi^{-1}(u)-\lambda^2/2}\,du \\
&= x_0 + e^{\mu-\lambda^2/2}\int_{-\infty}^{\infty} e^{(\sigma+\lambda)z}\,\varphi(z)\,dz \;=\; x_0 + e^{\mu-\lambda^2/2}\,e^{(\sigma+\lambda)^2/2} \\
&= x_0 + \exp\left\{\mu + \lambda\sigma + \sigma^2/2\right\}.
\end{aligned}$$

Example 3.13. [GS of SHIFTED EXPONENTIAL]

Note that $\int_{1-\beta}^1 du = \beta$, $\int_{1-\beta}^1 (1-u)\,du = \beta^2/2$, $\int_{1-\beta}^1 \log(1-u)\,du = \beta(\log(\beta)-1)$, and

$$\int_{1-\beta}^1 (1-u)\log(1-u)\,du \;=\; -\frac{1}{2}\left[(1-u)^2\log(1-u)\Big|_{1-\beta}^1 + \int_{1-\beta}^1 (1-u)\,du\right] \;=\; \frac{\beta^2}{2}\log(\beta) - \frac{\beta^2}{4}.$$

Now using these integrals and some straightforward simplifications, we have

$$\begin{aligned}
\mathrm{GS}[F,\beta,\delta] &= \frac{1}{\beta^2}\int_{1-\beta}^1 F^{-1}(u)\,\big(\beta + 4\delta(u - 1 + \beta/2)\big)\,du \\
&= \frac{1}{\beta^2}\int_{1-\beta}^1 \left[x_0 - \theta\log(1-u)\right]\big(\beta(1+2\delta) - 4\delta(1-u)\big)\,du \\
&= \beta^{-2}\Bigg[x_0\beta(1+2\delta)\int_{1-\beta}^1 du - 4x_0\delta\int_{1-\beta}^1 (1-u)\,du - \beta(1+2\delta)\theta\int_{1-\beta}^1 \log(1-u)\,du \\
&\qquad + 4\delta\theta\int_{1-\beta}^1 (1-u)\log(1-u)\,du\Bigg] \;=\; x_0 - \theta\big[\log(\beta) - 1 - \delta\big].
\end{aligned}$$

Example 3.14. [GS of PARETO I]

For $\alpha > 1$, the following formulas hold: $\int_{1-\beta}^1 (1-u)^{-1/\alpha}\,du = \alpha\beta^{1-1/\alpha}(\alpha-1)^{-1}$ and

$\int_{1-\beta}^{1}(1-u)^{-1/\alpha+1}\ du = \alpha\beta^{2-1/\alpha}(2\alpha-1)^{-1}$. These formulas, with some simplifications, lead to

$$\begin{aligned}
\text{GS}[F,\beta,\delta] &= \frac{1}{\beta^2}\int_{1-\beta}^{1} F^{-1}(u)\left(\beta+4\delta(u-1+\beta/2)\right)du \\
&= \frac{1}{\beta^2}\int_{1-\beta}^{1}\left[x_0(1-u)^{-1/\alpha}\right]\left(\beta(1+2\delta)-4\delta(1-u)\right)du
\end{aligned}$$

$$\begin{aligned}
&= x_0\beta^{-2}\left[\beta(1+2\delta)\int_{1-\beta}^{1}(1-u)^{-1/\alpha}\ du - 4\delta\int_{1-\beta}^{1}(1-u)^{-1/\alpha+1}\ du\right] \\
&= \alpha x_0\beta^{-1/\alpha}(2\delta/(2\alpha-1)+1)(\alpha-1)^{-1}.
\end{aligned}$$

Example 3.15. [GS of Shifted Lognormal]

Note that $\int_{1-\beta}^{1}\ du=\beta$, $\int_{1-\beta}^{1} u\ du = \beta-\beta^2/2$, and, as was shown in Example 3.6, $\int_{1-\beta}^{1} e^{\sigma\Phi^{-1}(u)}\ du = e^{\sigma^2/2}\,\Phi(\sigma-\Phi^{-1}(1-\beta))$. Also, using substitution $z=\Phi^{-1}(u)$, we have

$$\int_{1-\beta}^{1} u\, e^{\sigma\Phi^{-1}(u)}\ du \;=\; \int_{\Phi^{-1}(1-\beta)}^{\infty}\Phi(z)\, e^{\sigma z}\varphi(z)\ dz \;=\; e^{\sigma^2/2}\int_{\Phi^{-1}(1-\beta)}^{\infty}\Phi(z)\,\varphi(z-\sigma)\ dz.$$

Now the following steps are justified:

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

A.2 Derivations of Chapter 4

A.2.1 Section 4.3.1

Fisher Information for Shifted Exponential

Recall the log-likelihood function of $X^* \stackrel{d}{=} \big\{X \,\big|\, X > d\big\}$, where $X \sim \mathcal{E}xp\,(x_0, \theta)$:

$$\begin{aligned}\log \mathcal{L}\big(\theta \,\big|\, x_1^*, \ldots, x_n^*\big) &= -\log\theta \sum_{i=1}^{n} \mathbf{1}\{d < x_i^* < u\} \\ &\quad - \frac{1}{\theta}\sum_{i=1}^{n}\Big[(x_i^* - d)\mathbf{1}\{d < x_i^* < u\} + (u-d)\mathbf{1}\{x_i^* = u\}\Big].\end{aligned}$$

For $n = 1$, its first and second derivatives with respect to θ are:

$$\begin{aligned}\frac{\partial \log \mathcal{L}}{\partial \theta} &= \left(-\frac{1}{\theta} + \frac{x^* - d}{\theta^2}\right)\mathbf{1}\{d < x^* < u\} + \frac{u-d}{\theta^2}\,\mathbf{1}\{x^* = u\}, \\ \frac{\partial^2 \log \mathcal{L}}{\partial \theta^2} &= \left(\frac{1}{\theta^2} - \frac{2(x^* - d)}{\theta^3}\right)\mathbf{1}\{d < x^* < u\} - \frac{2(u-d)}{\theta^3}\,\mathbf{1}\{x^* = u\}.\end{aligned}$$

Note that

$$\mathbf{E}\Big[\mathbf{1}\{d < X^* < u\}\Big] = \mathbf{P}\{d < X^* < u\} = \frac{F(u) - F(d)}{1 - F(d)}, \tag{A.1}$$

$$\mathbf{E}\Big[\mathbf{1}\{X^* = u\}\Big] = \mathbf{P}\{X^* = u\} = \frac{1 - F(u)}{1 - F(d)}, \tag{A.2}$$

$$\begin{aligned}\mathbf{E}\Big[(X^* - d)\mathbf{1}\{d < X^* < u\}\Big] &= \int_d^u (y-d)\, dF_*(y) = -\int_d^u (y-d)\, d\big[1 - F_*(y)\big] \\ &= \int_d^u \big[1 - F_*(y)\big]\, dy - (u-d)\big[1 - F_*(u)\big] \\ &= \int_d^u \frac{1 - F(y)}{1 - F(d)}\, dy - (u-d)\,\frac{1 - F(u)}{1 - F(d)}. \qquad \text{(A.3)}\end{aligned}$$

Application of (A.1)–(A.3) to $X \sim \mathcal{E}xp\,(x_0, \theta)$, with $F(y) = 1 - e^{-(y-x_0)/\theta}$, yields

$$\mathbf{E}\Big[\mathbf{1}\{d < X^* < u\}\Big] \;=\; 1 - e^{-(u-d)/\theta}, \qquad \mathbf{E}\Big[\mathbf{1}\{X^* = u\}\Big] \;=\; e^{-(u-d)/\theta},$$

and

$$\mathbf{E}\Big[(X^* - d)\mathbf{1}\{d < X^* < u\}\Big] \;=\; \theta\left[1 - e^{-(u-d)/\theta}\right] - (u-d)\,e^{-(u-d)/\theta}.$$

Now I_{11} is found as follows:

$$\begin{aligned} I_{11} \;&=\; -\mathbf{E}\left[\frac{\partial^2 \log f_*(X^*)}{\partial \theta^2}\right] \;=\; -\,\theta^{-2}\,\mathbf{E}\Big[\mathbf{1}\{d < X^* < u\}\Big] \\ &\quad + 2\theta^{-3}\,\mathbf{E}\Big[(X^* - d)\mathbf{1}\{d < X^* < u\}\Big] + 2(u-d)\theta^{-3}\,\mathbf{E}\Big[\mathbf{1}\{X^* = u\}\Big] \\ &=\; \theta^{-2}\left[1 - e^{-(u-d)/\theta}\right]. \end{aligned}$$

Asymptotic Distributions of ML and PM Estimators of $\boldsymbol{\theta}$

$$\widehat{\theta}_{\mathrm{ML}} \;\sim\; \mathcal{AN}\left(\theta,\, \frac{\theta^2}{n}\,\frac{1}{1 - e^{-(u-d)/\theta}}\right),$$

$$\widehat{\theta}_{\mathrm{PM}} \;\sim\; \mathcal{AN}\left(\theta,\, \frac{\theta^2}{n}\,\frac{p_1}{(1-p_1)\log^2(1-p_1)}\right).$$

Asymptotic Distributions of ML and PM Estimators of $\boldsymbol{R}[\boldsymbol{F}]$

$$R[\widehat{F}_{\mathrm{ML}}] \;\sim\; \mathcal{AN}\left(R[F],\, \frac{\theta^2}{n}\,\frac{1}{1 - e^{-(u-d)/\theta}}\; d_1^2\right),$$

$$R[\widehat{F}_{\mathrm{PM}}] \;\sim\; \mathcal{AN}\left(R[F],\, \frac{\theta^2}{n}\,\frac{p_1}{(1-p_1)\log^2(1-p_1)}\; d_1^2\right),$$

where the partial derivative with respect to θ, denoted $d_1 = \partial R[F]/\partial\theta$, has the following expression:

- **Example 4.1**: For $\mathrm{VaR}[F, \beta] = x_0 - \theta \log(\beta)$,

$$d_1 = \frac{\partial \mathrm{VaR}[F, \beta]}{\partial \theta} = -\log(\beta).$$

- **Example 4.2**: For $\textsc{cte}[F, \beta] = x_0 - \theta(\log(\beta) - 1)$,

$$d_1 = \frac{\partial \textsc{cte}[F, \beta]}{\partial \theta} = -(\log(\beta) - 1).$$

- **Example 4.3**: For $\textsc{pht}[F, r] = x_0 + \theta/r$,

$$d_1 = \frac{\partial \textsc{pht}[F, r]}{\partial \theta} = r^{-1}.$$

- **Example 4.4**: For $\textsc{wt}[F, \lambda] = x_0 + \theta\, C_{\textsc{wt}}(\lambda) = x_0 + \theta \int_{-\infty}^{\infty} \Phi(z + \lambda) \frac{\varphi(z)}{\Phi(z)}\, dz$,

$$d_1 = \frac{\partial \textsc{wt}[F, \lambda]}{\partial \theta} = C_{\textsc{wt}}(\lambda).$$

- **Example 4.5**: For $\textsc{gs}[F, \beta, \delta] = x_0 - \theta(\log(\beta) - 1 - \delta)$,

$$d_1 = \frac{\partial \textsc{gs}[F, \beta, \delta]}{\partial \theta} = -(\log(\beta) - 1 - \delta).$$

A.2.2 Section 4.3.2

Fisher Information for Pareto I

Recall the log-likelihood function of $X^* \stackrel{d}{=} \{X \mid X > d\}$, where $X \sim \mathcal{P}a\, I\,(x_0, \alpha)$:

$$
\begin{aligned}
\log \mathcal{L}\big(\alpha \,\big|\, x_1^*, \ldots, x_n^*\big) \;&=\; \sum_{i=1}^{n} \big[\log(\alpha/x_0) - (\alpha+1)\log(x_i^*/x_0)\big]\,\mathbf{1}\big\{d < x_i^* < u\big\} \\
&\quad - \alpha n \log(x_0/d) + \alpha \log(x_0/u) \sum_{i=1}^{n} \mathbf{1}\big\{x_i^* = u\big\}.
\end{aligned}
$$

For $n = 1$, its first and second derivatives with respect to α are:

$$
\begin{aligned}
\frac{\partial \log \mathcal{L}}{\partial \alpha} \;&=\; \big(1/\alpha - \log(x^*/x_0)\big)\,\mathbf{1}\{d < x^* < u\} - \log(x_0/d) + \log(x_0/u)\;\mathbf{1}\big\{x^* = u\big\}, \\
\frac{\partial^2 \log \mathcal{L}}{\partial \alpha^2} \;&=\; \big(-1/\alpha^2\big)\,\mathbf{1}\{d < x^* < u\}.
\end{aligned}
$$

Application of (A.1) to $X \sim \mathcal{P}a\, I\,(x_0, \alpha)$, with $F(y) = 1 - (x_0/y)^\alpha$, yields

$$
\mathbf{E}\Big[\mathbf{1}\{d < X^* < u\}\Big] \;=\; 1 - (d/u)^\alpha.
$$

Now I_{11} is given by:

$$
I_{11} \;=\; -\,\mathbf{E}\left[\frac{\partial^2 \log f_*(X^*)}{\partial \alpha^2}\right] \;=\; \alpha^{-2}\,\mathbf{E}\Big[\mathbf{1}\{d < X^* < u\}\Big] \;=\; \alpha^{-2}\big[1 - (d/u)^\alpha\big].
$$

Asymptotic Distributions of ML and PM Estimators of α

$$
\widehat{\alpha}_{\mathrm{ML}} \;\sim\; \mathcal{AN}\left(\alpha,\, \frac{\alpha^2}{n}\,\frac{1}{1 - (d/u)^\alpha}\right),
$$

$$
\widehat{\alpha}_{\mathrm{PM}} \;\sim\; \mathcal{AN}\left(\alpha,\, \frac{\alpha^2}{n}\,\frac{p_1}{(1 - p_1)\log^2(1 - p_1)}\right).
$$

Asymptotic Distributions of ML and PM Estimators of $\boldsymbol{R}[\boldsymbol{F}]$

$$R[\widehat{F}_{\mathrm{ML}}] \;\sim\; \mathcal{AN}\left(R[F],\, \frac{\alpha^2}{n}\, \frac{1}{1-(d/u)^{\alpha}}\, d_1^2\right),$$

$$R[\widehat{F}_{\mathrm{PM}}] \;\sim\; \mathcal{AN}\left(R[F],\, \frac{\alpha^2}{n}\, \frac{p_1}{(1-p_1)\log^2(1-p_1)}\, d_1^2\right),$$

where the partial derivative with respect to α, denoted $d_1 = \partial R[F]/\partial\alpha$, has the following expression:

- **Example 4.6**: For $\mathrm{VaR}[F,\beta] = x_0\beta^{-1/\alpha}$,

$$d_1 \;=\; \frac{\partial \mathrm{VaR}[F,\beta]}{\partial\alpha} \;=\; \alpha^{-2}\,\mathrm{VaR}[F,\beta]\,\log(\beta).$$

- **Example 4.7**: For $\textsc{cte}[F,\beta] = x_0\beta^{-1/\alpha}\alpha(\alpha-1)^{-1}$,

$$d_1 \;=\; \frac{\partial \textsc{cte}[F,\beta]}{\partial\alpha} \;=\; \alpha^{-2}\,\textsc{cte}[F,\beta]\,D_{\textsc{cte}}(\beta,\alpha),$$

  where $D_{\textsc{cte}}(\beta,\alpha) = \log(\beta) - \alpha(\alpha-1)^{-1}$.

- **Example 4.8**: For $\textsc{pht}[F,r] = x_0 + x_0/(r\alpha-1)$,

$$d_1 \;=\; \frac{\partial \textsc{pht}[F,r]}{\partial\alpha} \;=\; -\,\textsc{pht}[F,r]\,\alpha^{-1}\,(r\alpha-1)^{-1}.$$

- **Example 4.9**: For $\textsc{wt}[F,\lambda] = x_0 + \frac{x_0}{\alpha}\,C_{\textsc{wt}}(\lambda,\alpha) = x_0 + \frac{x_0}{\alpha}\int_{-\infty}^{\infty}\Phi(z+\lambda)\,\frac{\varphi(z)}{[\Phi(z)]^{1/\alpha+1}}\,dz$,

$$d_1 \;=\; \frac{\partial \textsc{wt}[F,\lambda]}{\partial\alpha} \;=\; x_0\,\alpha^{-2}\,D_{\textsc{wt}}(\lambda,\alpha),$$

  where $D_{\textsc{wt}}(\lambda,\alpha) = -C_{\textsc{wt}}(\lambda,\alpha) + \alpha^{-1}\int_{-\infty}^{\infty}\Phi(z+\lambda)\,\frac{\varphi(z)}{[\Phi(z)]^{1/\alpha+1}}\,\log\left[\Phi(z)\right]\,dz.$

- **Example 4.10**: For $\textsc{gs}[F,\beta,\delta] = x_0\beta^{-1/\alpha}\alpha(\alpha-1)^{-1}\left(1+2\delta(2\alpha-1)^{-1}\right)$,

$$d_1 \;=\; \frac{\partial \textsc{gs}[F,\beta,\delta]}{\partial\alpha} \;=\; \alpha^{-2}\,\textsc{gs}[F,\beta,\delta]\,D_{\textsc{gs}}(\beta,\delta,\alpha),$$

where $D_{\textsc{gs}}(\beta,\delta,\alpha) = \log(\beta) - \alpha(\alpha-1)^{-1} - 4\alpha^2\delta(2\alpha-1)^{-1}(2\delta+2\alpha-1)^{-1}$.

A.2.3 Section 4.3.3

Fisher Information for Shifted Lognormal

Recall the notation

$$c_d \;=\; \frac{\log(d-x_0)-\mu}{\sigma}, \qquad c_{x_i^*} \;=\; \frac{\log(x_i^*-x_0)-\mu}{\sigma}, \qquad c_u \;=\; \frac{\log(u-x_0)-\mu}{\sigma},$$

and the log-likelihood function of $X^* \stackrel{d}{=} \big\{X \,\big|\, X>d\big\}$, where $X\sim\mathcal{LN}\,(x_0,\mu,\sigma)$:

$$\begin{aligned}\log\mathcal{L}\big((\mu,\sigma)\,\big|\,x_1^*,\ldots,x_n^*\big) \;&=\; \sum_{i=1}^{n}\big[-\log(\sigma)-\log(x_i^*-x_0)+\log(\varphi(c_{x_i^*}))\big]\mathbf{1}\big\{d<x_i^*<u\big\}\\ &\quad -\,n\log\left[1-\Phi(c_d)\right]+\log\left[1-\Phi(c_u)\right]\sum_{i=1}^{n}\mathbf{1}\big\{x_i^*=u\big\}.\end{aligned}$$

For $n=1$, its first partial derivatives with respect to μ and σ are:

$$\begin{aligned}\frac{\partial\log\mathcal{L}}{\partial\mu} \;&=\; \sigma^{-1}\left\{c_{x^*}\mathbf{1}\{d<x^*<u\} - \frac{\varphi(c_d)}{1-\Phi(c_d)} + \frac{\varphi(c_u)}{1-\Phi(c_u)}\mathbf{1}\big\{x^*=u\big\}\right\},\\ \frac{\partial\log\mathcal{L}}{\partial\sigma} \;&=\; \sigma^{-1}\left\{(c_{x^*}^2-1)\mathbf{1}\{d<x^*<u\} - \frac{\varphi(c_d)}{1-\Phi(c_d)} + \frac{\varphi(c_u)}{1-\Phi(c_u)}\mathbf{1}\big\{x^*=u\big\}\right\}.\end{aligned}$$

And its second partial derivatives are:

$$
\begin{aligned}
\frac{\partial^2 \log \mathcal{L}}{\partial \mu^2} &= \sigma^{-2}\Bigg\{ -\mathbf{1}\{d < x^* < u\} - \frac{(1-\Phi(c_d))\varphi(c_d)c_d - \varphi^2(c_d)}{(1-\Phi(c_d))^2} \\
&\quad + \frac{(1-\Phi(c_u))\varphi(c_u)c_u - \varphi^2(c_u)}{(1-\Phi(c_u))^2}\,\mathbf{1}\{x^* = u\}\Bigg\}, \\
\frac{\partial^2 \log \mathcal{L}}{\partial \mu \partial \sigma} &= \sigma^{-2}\Bigg\{ -2c_{x^*}\mathbf{1}\{d < x^* < u\} - \frac{(1-\Phi(c_d))\varphi(c_d)(c_d^2-1) - \varphi^2(c_d)c_d}{(1-\Phi(c_d))^2} \\
&\quad + \frac{(1-\Phi(c_u))\varphi(c_u)(c_u^2-1) - \varphi^2(c_u)c_u}{(1-\Phi(c_u))^2}\,\mathbf{1}\{x^* = u\}\Bigg\}, \\
\frac{\partial^2 \log \mathcal{L}}{\partial \sigma^2} &= \sigma^{-2}\Bigg\{ -\frac{\partial \log \mathcal{L}}{\partial \sigma} - 2c_{x^*}^2\mathbf{1}\{d < x^* < u\} - \frac{(1-\Phi(c_d))\varphi(c_d)c_d(c_d^2-1) - \varphi^2(c_d)c_d^2}{(1-\Phi(c_d))^2} \\
&\quad + \frac{(1-\Phi(c_u))\varphi(c_u)c_u(c_u^2-1) - \varphi^2(c_u)c_u^2}{(1-\Phi(c_u))^2}\,\mathbf{1}\{x^* = u\}\Bigg\}.
\end{aligned}
$$

Application of (A.1)–(A.2) to $X \sim \mathcal{LN}(x_0, \mu, \sigma)$, with $F(y) = \Phi(c_y)$, yields

$$
\mathbf{E}\Big[\mathbf{1}\{d < X^* < u\}\Big] = \frac{\Phi(c_u) - \Phi(c_d)}{1-\Phi(c_d)} \qquad \text{and} \qquad \mathbf{E}\Big[\mathbf{1}\{X^* = u\}\Big] = \frac{1-\Phi(c_u)}{1-\Phi(c_d)}.
$$

Moreover, $\mathbf{E}\left[\frac{\partial \log \mathcal{L}}{\partial \sigma}\right] = \mathbf{E}\left[\frac{\partial \log(f_*(X^*))}{\partial \sigma}\right] = 0$, and computations analogous to (A.3) yield

$$
\begin{aligned}
\mathbf{E}\Big[c_{X^*}\mathbf{1}\{d < X^* < u\}\Big] &= \int_d^u c_y\, f_*(y)\, dy \;=\; \frac{1}{1-\Phi(c_d)}\int_{c_d}^{c_u} z\,\varphi(z)\, dz \\
&= \frac{\varphi(c_d) - \varphi(c_u)}{1-\Phi(c_d)}, \\
\mathbf{E}\Big[c_{X^*}^2\mathbf{1}\{d < X^* < u\}\Big] &= \int_d^u c_y^2\, f_*(y)\, dy \;=\; \frac{1}{1-\Phi(c_d)}\int_{c_d}^{c_u} z^2\,\varphi(z)\, dz \\
&= \frac{c_d\varphi(c_d) - c_u\varphi(c_u) + \Phi(c_u) - \Phi(c_d)}{1-\Phi(c_d)}.
\end{aligned}
$$

Putting all this together, with some straighforward simplifications, we find that the Fisher information matrix has the following entries:

$$
\begin{aligned}
I_{11} &= -\mathbf{E}\left[\frac{\partial^2 \log f_*(X^*)}{\partial \mu^2}\right] \\
&= \sigma^{-2}\left[\frac{\Phi(c_u)-\Phi(c_d)}{1-\Phi(c_d)} + \frac{[1-\Phi(c_d)]\varphi(c_d)c_d - \varphi^2(c_d)}{[1-\Phi(c_d)]^2} - \frac{[1-\Phi(c_u)]\varphi(c_u)c_u - \varphi^2(c_u)}{[1-\Phi(c_u)][1-\Phi(c_d)]}\right] \\
I_{12} &= I_{21} = -\mathbf{E}\left[\frac{\partial^2 \log f_*(X^*)}{\partial \mu\, \partial \sigma}\right] \\
&= \sigma^{-2}\left[\frac{[1-\Phi(c_d)]\varphi(c_d)(c_d^2+1) - \varphi^2(c_d)c_d}{[1-\Phi(c_d)]^2} - \frac{[1-\Phi(c_u)]\varphi(c_u)(c_u^2+1) - \varphi^2(c_u)c_u}{[1-\Phi(c_d)][1-\Phi(c_u)]}\right] \\
I_{22} &= -\mathbf{E}\left[\frac{\partial^2 \log f_*(X^*)}{\partial \sigma^2}\right] \\
&= \sigma^{-2}\left[\frac{2\,[\Phi(c_u)-\Phi(c_d)]}{1-\Phi(c_d)} + \frac{[1-\Phi(c_d)]\varphi(c_d)c_d(c_d^2+1) - \varphi^2(c_d)c_d^2}{[1-\Phi(c_d)]^2}\right. \\
&\qquad \left. - \frac{[1-\Phi(c_u)]\varphi(c_u)c_u(c_u^2+1) - \varphi^2(c_u)c_u^2}{[1-\Phi(c_d)][1-\Phi(c_u)]}\right]
\end{aligned}
$$

Appendix B

Asymptotic Theorems

Here we provide some theoretical results that are repeatedly used in Chapter 4. Specifically, the asymptotic normality theorems for sample quantiles, the maximum likelihood (ML) and percentile-matching (PM) estimators of model parameters are presented. Transformations of asymptotically normal vectors are handled by using the delta method which is also provided in this section. In addition, we extend the well-known inequality for the normal distribution upper tail to the lower tail. We also prove two additional inequalities for the standard normal distribution tails.

Suppose we have a sample of *independent and identically distributed* (*i.i.d.*) continuous random variables, $X_1, \ldots, X_n$, with the cumulative distribution function (cdf) G, probability density function (pdf) g, and quantile function (qf) G^{-1}. Let the cdf, pdf, and qf be given in a parametric form, and suppose that they are indexed by a k-dimensional parameter $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$. We will assume that g satisfies all the regularity conditions that usually accompany theorems such as the ones formulated in this appendix. (For more details, see, e.g., Serfling, 1980, Sections 2.3.3 and 4.2.2.) Further, $X_{(1)} \leq \cdots \leq X_{(n)}$ denotes the order statistics of $X_1, \ldots, X_n$. Also, throughout Appendix B and the dissertation the abbreviation $\mathcal{AN}$ stands for "asymptotically normal."

The empirical estimator of a population quantile, say $G^{-1}(p)$, is the corresponding sample quantile $X_{(\lceil np \rceil)}$, where $\lceil \cdot \rceil$ denotes the "rounding up" operation. We start with the asymptotic normality result for sample quantiles. Complete technical details are available in Section 2.3.3 of Serfling (1980).

Theorem B.1 [Asymptotic Normality of Sample Quantiles]
Let $0 < p_1 < \cdots < p_k < 1$, with $k > 1$, and suppose that pdf g is continuous, as discussed above. Then the k-variate vector of sample quantiles $\left(X_{(\lceil np_1 \rceil)}, \ldots, X_{(\lceil np_k \rceil)}\right)$ is $\mathcal{AN}$ with the mean vector $\left(G^{-1}(p_1), \ldots, G^{-1}(p_k)\right)$ and the covariance-variance matrix $\frac{1}{n}\left[\sigma_{ij}^2\right]_{i,j=1}^{k}$ with

$$\sigma_{ij}^2 = \frac{p_i(1-p_j)}{g(G^{-1}(p_i))g(G^{-1}(p_j))}\,. \tag{B.1}$$

In the univariate case ($k = 1$), the sample quantile $X_{(\lceil np \rceil)}$ satisfies:

$$X_{(\lceil np \rceil)} \;\sim\; \mathcal{AN}\left(G^{-1}(p),\, \frac{1}{n}\frac{p(1-p)}{g^2(G^{-1}(p))}\right). \tag{B.2}$$

The following theorem summarizes asymptotic distribution of the ML estimators. Description of the method, proofs and complete technical details are available in Section 4.2 of Serfling (1980).

Theorem B.2 [Asymptotic Normality of MLs]
Let $\widehat{\boldsymbol{\theta}} = \left(\widehat{\theta}_1, \ldots, \widehat{\theta}_k\right)$ denote the ML of parameter $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$. Then, under the regularity conditions mentioned above,

$$\left(\widehat{\theta}_1, \ldots, \widehat{\theta}_k\right) \;\sim\; \mathcal{AN}\left((\theta_1, \ldots, \theta_k),\, \frac{1}{n}\mathbf{I}_{\boldsymbol{\theta}}^{\mathbf{-1}}\right),$$

where $\mathbf{I}_{\boldsymbol{\theta}} = \left[I_{ij}\right]_{i,j=1}^{k}$ *is the Fisher information matrix, with the entries given by*

$$I_{ij} = \mathbf{E}\left[\frac{\partial \log g(X)}{\partial \theta_i} \cdot \frac{\partial \log g(X)}{\partial \theta_j}\right]. \tag{B.3}$$

In the univariate case $(k = 1)$*,*

$$\widehat{\theta} \;\sim\; \mathcal{AN}\left(\theta,\, \frac{1}{n}\, \frac{1}{\mathbf{E}\Big[\Big(\frac{\partial \log g(X)}{\partial \theta}\Big)^2\Big]}\right). \tag{B.4}$$

The *delta method* is a technical tool for establishing asymptotic normality of smoothly transformed asymptotically normal random vectors. Here we will present it as a direct application to Theorem B.2. For the general theorem and complete technical details, see Serfling (1980, Section 3.3).

Theorem B.3 [The Delta Method]
Suppose that $\widehat{\boldsymbol{\theta}} = \left(\widehat{\theta}_1, \ldots, \widehat{\theta}_k\right)$ *is* $\mathcal{AN}$ *with the parameters specified in Theorem B.2. Let the real-valued functions* $h_1(\boldsymbol{\theta}) = h_1\left(\theta_1, \ldots, \theta_k\right), \ldots, h_m(\boldsymbol{\theta}) = h_m\left(\theta_1, \ldots, \theta_k\right)$ *represent* m *different risk measures, tail probabilities or other transformations of model parameters. Then, under some smoothness conditions on functions* $h_1, \ldots, h_m$*, the vector of ML-based estimators*

$$\left(h_1\left(\widehat{\boldsymbol{\theta}}\right), \ldots, h_m\left(\widehat{\boldsymbol{\theta}}\right)\right) \;\sim\; \mathcal{AN}\Big(\left(h_1\left(\boldsymbol{\theta}\right), \ldots, h_m\left(\boldsymbol{\theta}\right)\right),\, \frac{1}{n}\,\mathbf{D}_{\boldsymbol{\theta}}\mathbf{I}_{\boldsymbol{\theta}}^{-1}\mathbf{D}_{\boldsymbol{\theta}}'\Big), \tag{B.5}$$

where $\mathbf{D}_{\boldsymbol{\theta}} = [d_{ij}]_{m\times k}$ *is the Jacobian of the transformations* $h_1, \ldots, h_m$ *evaluated at* $(\theta_1, \ldots, \theta_k)$*, that is,* $d_{ij} = \partial h_i / \partial \widehat{\theta}_j \Big|_{(\theta_1, \ldots, \theta_k)}$*. In the univariate case* $(m = 1)$*, the para-*

metric estimator

$$h\left(\widehat{\theta}_1,\ldots,\widehat{\theta}_k\right) \;\sim\; \mathcal{AN}\Big(h\left(\theta_1,\ldots,\theta_k\right),\, \frac{1}{n}\,\mathbf{d}_{\boldsymbol{\theta}}\mathbf{I}_{\boldsymbol{\theta}}^{-\mathbf{1}}\mathbf{d}_{\boldsymbol{\theta}}'\Big), \tag{B.6}$$

where $\mathbf{d}_{\boldsymbol{\theta}} = \left(\partial h/\partial\widehat{\theta}_1,\ldots,\partial h/\partial\widehat{\theta}_k\right)\Big|_{(\theta_1,\ldots,\theta_k)}$.

If the probability distribution has k unknown parameters, $(\theta_1,\ldots,\theta_k)$, PM estimators are found by matching $G^{-1}(p_i)$ with $X_{(\lceil np_i\rceil)}$, $i=1,\ldots,k$, and then solving the resulting system of equations with respect to $\theta_1,\ldots,\theta_k$. Assuming the system of equations has a unique solution, it is clear that PM estimators of $\theta_1,\ldots,\theta_k$ will be functions of $X_{(\lceil np_1\rceil)},\ldots,X_{(\lceil np_k\rceil)}$. Let us denote these estimators as $\widetilde{\theta}_i = h_i^*\left(X_{(\lceil np_1\rceil)},\ldots,X_{(\lceil np_k\rceil)}\right)$, $i=1,\ldots,k$. Their joint asymptotic normality follows from Theorems B.1 and B.3 (with obvious modifications to (B.5)).

Theorem B.4 [Asymptotic Normality of PMs]

Let $\widetilde{\boldsymbol{\theta}} = \left(\widetilde{\theta}_1, \ldots, \widetilde{\theta}_k\right)$ *denote the PM estimator of parameter* $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$*. Then,*

$$\left(\widetilde{\theta}_1, \ldots, \widetilde{\theta}_k\right) \;\sim\; \mathcal{AN}\left((\theta_1, \ldots, \theta_k),\, \frac{1}{n}\,\mathbf{D}^*_{\boldsymbol{\theta}}\boldsymbol{\Sigma}_{\boldsymbol{\theta}}(\mathbf{D}^*_{\boldsymbol{\theta}})'\right), \tag{B.7}$$

where the entries of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ *are given by (B.1) and* $\mathbf{D}^*_{\boldsymbol{\theta}} = [d^*_{ij}]_{k\times k}$ *is the Jacobian of the transformations* $h^*_1, \ldots, h^*_k$ *evaluated at* $(\theta_1, \ldots, \theta_k)$*, that is,* $d^*_{ij} = \partial h^*_i / \partial X_{(\lceil np_j \rceil)}\Big|_{(\theta_1,\ldots,\theta_k)}$.

For establishing some of the inequalities of Chapter 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.5. [Normal Distribution Tails - I]

(a) *As* $x \to \infty$*,*

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

more precisely, the double inequality

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

holds for every $x > 0$.

(b) *As* $x \to -\infty$*,*

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

more precisely, the double inequality

$$\left(x^{-3} - x^{-1}\right)\varphi(x) \;<\; \Phi(x) \;<\; -\,x^{-1}\varphi(x)$$

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,

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

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

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

Straightforward simplifications yield the stated inequality. □

Inequalities involving standard normal cdf at different points are used in Chapter 3 derivations.

Lemma B.6. [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

$$\begin{aligned}\Phi(z+\lambda) &= \int_{-\infty}^{z+\lambda} \varphi(t)\,dt \;=\int_{-\infty}^{z} \varphi(v+\lambda)\,dv \\ &= e^{-\lambda^2/2}\int_{-\infty}^{z} \varphi(v)\,e^{-\lambda v}\,dv. \end{aligned} \tag{B.8}$$

In part (a), $\lambda \le 0$ implies $e^{-\lambda v}$ is a non-decreasing function. Thus, for $v \le z$, we have $e^{-\lambda v} \le e^{-\lambda z}$. Combining this inequality with (B.8), we get

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

For part (b), first use integration by parts in (B.8), then apply the upper bound from Lemma B.5(b), notice that $-v^{-1} \le -z^{-1}$ when $v \le z < 0$, 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 $z < -\lambda < 0$. $\square$

CURRICULUM VITAE

Full Name: Sahadeb Upretee
Place of Birth: Bhumlu Kolanti, Kavrepalanchok, Nepal

EDUCATION

2015–2020 **Ph.D., Mathematics**, *University of Wisconsin-Milwaukee*.
Concentrations: Actuarial Science; Statistics.
Dissertation Title: Estimating Distortion Risk Measures Under Truncated and Censored Data Scenarios. Advisor: Professor V. Brazauskas.

2008–2011 **M.Phil., Mathematics**, *Kathmandu University*, Nepal.
Thesis Title: Cryptography: Its Origin and Advancement. Advisor: Professor P. R. Adhikari.

2003–2005 **M.Ed., Mathematics Education**, *Tribhuvan University*, Nepal.
Thesis Title: Mathematics Classroom Management from Multicultural Perspectives.
Advisor: Professor L. Sharma.

2000–2003 **B.Ed., Mathematics Education**, *Tribhuvan University*, Nepal.

RESEARCH INTERESTS

Actuarial Science; Loss Data Analytics; Model Uncertainty; Risk Measures.

ACTUARIAL EDUCATION

Society of Actuaries, Schaumburg, IL.
Professional Examinations Passed: P; FM; IFM; STAM; SRM (transition credit).
FAP Modules Passed: Modules 1 - 5, Interim Assessment.
VEE Credits Earned: Applied Statistics, Corporate Finance, and Economics.

EXPERIENCE

2015–2020 **Graduate Teaching Assistant**, *University of Wisconsin-Milwaukee*.
Graduate Course Taught: Probability and Statistical Inference
Undergraduate Courses Taught: Financial Mathematics (SOA Exam FM preparation); Probability (SOA Exam P preparation); Calculus and Analytic Geometry I, II, III; Algebraic Literacy II; Intermediate Algebra; Foundations of Elementary Mathematics.

2013–2015 **Assistant Professor**, *Kathmandu University School of Management*, Nepal.
Graduate Courses Taught: Optimization Techniques for Managerial Decisions - MBA
Undergraduate Courses Taught: Financial Mathematics; Calculus I,II; Logic; Quantitative Techniques.

2008–2013 **Lecturer**, *Kathmandu University*, Nepal.
Graduate Courses Taught: Analysis; Abstract Algebra; Functional Analysis; Mathematical Statistics; Multivariate Calculus; Mathematical Modeling; Topology.

PROFESSIONAL SOCIETY MEMBERSHIPS

2015–present Member of American Mathematical Society (AMS).

2009–present Member of Nepal Mathematical Council (NMC).

2008–present Member of Nepal Mathematical Society (NMS).

SELECTED PUBLICATIONS

Brazauskas, V. and Upretee, S. (2019). Model efficiency and uncertainty in quantile estimation of loss severity distributions. *Risks* **7**(2), 16 pages, doi: 10.3390/risks7020055.

Upretee, S. and Adhikari, P. R. (2010). RSA cryptosystem. α - β *Journal of Mathematics Education*. **1**, 54 - 56.

Adhikari, P. R. and Upretee, S. (2009). Substitution private key cryptosystem. *Epsilon-Delta Yearly Mathematical Magazine*, **5**, 6 - 10.

SELECTED TALKS AND PRESENTATIONS

2019 Model Efficiency and Uncertainty in Quantile Estimation of Loss Severity Distributions. *54th Actuarial Research Conference*, Indianapolis, IN.

2019 Model Efficiency and Uncertainty in Quantile Estimation of Loss Severity Distributions. *Simon Conference for Young Researchers in Risk Management and Insurance*, East Lansing, MI.

SELECTED PROFESSIONAL ACTIVITIES

Referee for: *Empirical Economics (2017, 2018).*

2009–2011 *Subject Committee Chair*, Mathematics Education, Kathmandu University, Nepal.

2013 *A subject matter expert* , Mid-Western University, Nepal.
To develop curriculum for the BS in mathematics program.

A subject matter expert, Teacher Selection Commission, Nepal.

SELECTED HONORS AND AWARDS

2019 **Actuarial Research Conference Travel Grant**, *Society of Actuaries*, Schaumburg, IL.

2019 **Graduate Student Travel Award**, *University of Wisconsin-Milwaukee.*

2015–2020 **Chancellor's Graduate Student Award**, *University of Wisconsin-Milwaukee.*

2015–2018 **Research Excellence Awards**, *Departmental of Mathematical Sciences, University of Wisconsin-Milwaukee.*

2006 **Gold Medal (Nepal Bidhya Bhusan Kha)**, *Tribhuvan University, Nepal.*
Achieving highest GPA among approximately 2000 students in Institute of Education, awarded by President of Nepal.

2006 **Himamsu Dahal Memorial Award**, *Tribhuvan University, Nepal.*
Achieving highest GPA among approximately 250 students in department of mathematics education; awarded by Prime Minister of Nepal.

PROGRAMMING SKILLS

Software Packages: MATLAB, R, SPSS, SQL, Minitab, Maple, LaTeX.

Teaching Tools: MyMathLab, ALEKS, GATEWAY, MS Word, MS Excel, Power-Point.

Programming Language: C++