

Model Efficiency and Uncertainty in Quantile Estimation of Loss Severity Distributions

VYTARAS BRAZAUSKAS¹

University of Wisconsin-Milwaukee

SAHADEB UPRETEE²

University of Wisconsin-Milwaukee

To appear in *Risks*

(*Submitted:* January 29, 2019 *Revised:* April 9, 2019; April 18, 2019 *Accepted:* April 26, 2019)

Abstract. Quantiles of probability distributions play a central role in the definition of risk measures (e.g., value-at-risk, conditional tail expectation) which in turn are used to capture the riskiness of the distribution tail. Estimates of risk measures are needed in many practical situations such as in pricing of extreme events, developing reserve estimates, designing risk transfer strategies, and allocating capital. In this paper, we present the empirical nonparametric and two types of parametric estimators of quantiles at various levels. For parametric estimation, we employ the maximum likelihood and percentile-matching approaches. Asymptotic distributions of all the estimators under consideration are derived when data are left-truncated and right-censored, which is a typical loss variable modification in insurance. Then, we construct relative efficiency curves, REC, for all the parametric estimators. Specific examples of such curves are provided for exponential and single-parameter Pareto distributions for a few data truncation and censoring cases. Additionally, using simulated data we examine how wrong quantile estimates can be when one makes incorrect modeling assumptions. The numerical analysis is also supplemented with standard model diagnostics and validation (e.g., quantile-quantile plots, goodness-of-fit tests, information criteria) and presents an example of when those methods can mislead the decision maker. These findings pave the way for further work on RECs with potential for them being developed into an effective diagnostic tool in this context.

Keywords & Phrases. Data truncation and censoring; empirical estimator; maximum likelihood; model uncertainty; percentile matching; quantile estimation.

¹ CORRESPONDING AUTHOR: Vytautas Brazauskas, Ph.D., ASA, is a Professor in the Department of Mathematical Sciences, University of Wisconsin-Milwaukee, P.O. Box 413, Milwaukee, WI 53201, USA. *e-mail:* vytaras@uwm.edu

² Sahadeb Uprettee is a PhD candidate and Graduate Teaching Assistant in the Department of Mathematical Sciences, University of Wisconsin-Milwaukee, P.O. Box 413, Milwaukee, WI 53201, USA. *e-mail:* supreteee@uwm.edu

1 Introduction

Quantiles of probability distributions play a central role in the definition of risk measures (e.g., value-at-risk, conditional tail expectation) which in turn are used to capture the riskiness of the distribution tail. Estimates of risk measures are needed in many practical situations such as in pricing of extreme events, developing reserve estimates, designing risk transfer strategies, and allocating capital. When solving such problems, the first highly consequential task is to find point estimates of quantiles and to assess their variability. In this context, the empirical nonparametric approach is the simplest one to use (see Jones and Zitikis, 2003), but it lacks efficiency due to the scarcity of sample data in the tails. On the other hand, parametric estimators can significantly improve quantile estimators' efficiency (see Brazauskas and Kaiser, 2004, and Kaiser and Brazauskas, 2006). Moreover, the parametric approach can accommodate truncation and censoring that are common features of insurance loss data. Of course, the main drawback of parametric estimators is that they are sensitive to initial modeling assumptions, which creates model uncertainty¹

There is a growing number of studies on various aspects of model risk in modeling, measuring and pricing risks. Cairns (2000) was the first author to systematically study model risk in insurance. He discussed different sources of model risk, including parameter uncertainty and model uncertainty, and presented methods to treat these uncertainties coherently. Hartman *et al.* (2017) focused on parameter uncertainty and analyzed its impact in different sectors of insurance practice, namely, life insurance, health insurance, and property/casualty insurance. They also gave a comprehensive review of the literature concerning parameter uncertainty. A recent article by Hong *et al.* (2018) shows how typical claim predictions change when the model is uncertain. In particular, they illustrate such effects by using standard model selection tools such as Akaike Information Criterion to determine the “best” regression subset of covariates, and then apply the selected model for claim prediction. Bignozzi *et al.* (2015) and Samanthi *et al.* (2017) are two recent examples of theoretical and practical investigations, respectively, of the effects of the data dependence assumption on subsequent risk measuring. Also, an extensive simulation study involving estimation of upper quantiles of lognormal, log-logistic, and log-

¹ Note that some authors use ‘model risk’ instead of ‘model uncertainty’ to describe the same phenomenon.

double exponential distributions under model and parameter uncertainty was conducted by Modarres *et al.* (2002). Their overall conclusion was that when modeling is done by assuming one of the three families and treating the other two as possible misspecification, the least severe effect on upper quantile estimates occurs when the lognormal distribution is assumed.

Further, there is even more interest in this topic in the financial risk management literature. Model uncertainty within the risk aggregation problems has been recently studied by Embrechts *et al.* (2015) and Cambou and Filipović (2017), and for value-at-risk estimation by Alexander and Sarabia (2012). Cont *et al.* (2010) and Glasserman and Xu (2014) linked financial risk measurement procedures, model risk, and robustness. The first paper suggests to use the classical robust statistics techniques for managing model risk, while the second pursues model distance and entropy based techniques to derive the worst-case risk measurements (relative to measurements from a baseline model). Finally, Aggarwal *et al.* (2016) and Black *et al.* (2018) provide comprehensive accounts on model risk identification, measurement, and management in practice. These authors develop a model risk framework, identify distinct model cultures within an organization, review common methods and challenges for quantifying model risk, and discuss difficulties that arise in mapping model errors to actual financial impact.

The implied conclusion in many academic and practice oriented papers on model risk is that it can be reduced or mitigated by using all or a combination of the following: performing model validation, fitting multiple models, and applying various stress tests or sensitivity analysis. This idea was in part adopted in the case studies of Brazauskas and Kleefeld (2016), which were based on well-known (real) reinsurance data. What was discovered by these authors, however, is that fitting multiple models and using extensive model validation for each of them may not be sufficient if data are left-truncated. That is, they used quantile-quantile plots, Kolmogorov-Smirnov (KS) and Anderson-Darling (AD) tests, Akaike and Bayesian information criteria (AIC and BIC) and had concluded that six different models are acceptable for each of the 12 data sets analyzed. However, when all those models were used to estimate the 90% and 95% quantiles (value-at-risk measures) for ground-up loss, for some data sets they resulted in similar estimates, which would be expected, while for others they were far apart, which is counterintuitive. Moreover, using left-truncated operational risk data, Yu and Brazauskas (2017) have shown that even shifted parametric models (which might seem like a plausible option but

nonetheless incorrectly account for data truncation) can pass those standard model validation tests. Next, due to the presence of deductibles and policy limits in insurance contracts, data truncation and censoring are unavoidable modifications of the loss severity variable. This suggests that quantile and, more generally, risk measure estimation requires careful thinking and analysis.

In this paper, we present the empirical nonparametric, maximum likelihood, and percentile-matching estimators of ground-up loss distribution quantiles (at various levels). Asymptotic distributions of these estimators are derived when data are left-truncated and right-censored. Relative efficiency curves, REC, for all the estimators are then constructed, and plots of such curves are provided for exponential and single-parameter Pareto distributions. Then, we generate a sample of 50 observations from a left-truncated and right-censored Pareto I model and using that data set investigate how biased quantile estimates can be when one makes incorrect distributional assumptions or relies on a wrong modeling approach. The numerical analysis is also supplemented with standard model diagnostics and validation (e.g., quantile-quantile plots and KS and AD tests) and demonstrates how those methods can mislead the decision maker. In addition, we examine the information provided by RECs and conclude that such curves have strong potential for being developed into an effective diagnostic tool in this context.

The rest of the paper is organized as follows. In Section 2, nonparametric and parametric quantile estimators are defined and their asymptotic distributions are specified when the underlying random variable is left-truncated and right-censored. The next section presents two illustrative examples of RECs for exponential and single-parameter Pareto distributions. Specifically, RECs of maximum likelihood, percentile matching, and empirical estimators of quantiles of these distributions are plotted. Section 4 studies the effects of distribution choice and modeling approach on estimates of quantiles. Concluding remarks are offered in Section 5. Finally, the appendix provides two asymptotic theorems of mathematical statistics and a detailed description of how to construct RECs. These results are essential to analytic derivations in the paper, and we recommend the reader to review them first.

2 Quantile Estimation

Insurance contracts have coverage modifications that need to be taken into account when modeling the underlying loss severity variable. In this section, we specify the estimators of quantiles of the ground-up distribution and derive their asymptotic distributions when the loss variable is affected by left truncation (due to deductible) and right censoring (due to policy limit). We consider three types of estimators: empirical (Section 2.1), maximum likelihood, MLE (Section 2.2), and percentile matching, PM (Section 2.3).

To present the estimators and their properties, let us start with notation and assumptions. Suppose we observe n continuous *independent identically distributed (i.i.d.)* random variables X_1^*, \dots, X_n^* , where each X^* is equal to the ground-up variable X , if X exceeds threshold t ($t \geq 0$) but is capped at upper limit u ($u > t$). That is, X^* is a mixed discrete-continuous random variable that satisfies the following conditional event relationship:

$$X^* \stackrel{d}{=} \min\{X, u\} \mid X > t,$$

where $\stackrel{d}{=}$ denotes “equal in distribution.” Also, let us denote the probability density function (pdf), cumulative distribution function (cdf), and quantile function (qf) of X as f , F , and F^{-1} , respectively. Then, the cdf F_* , pdf f_* , qf F_*^{-1} of X^* are related to F , f , F^{-1} and given by:

$$F_*(x^* | t; u) = \begin{cases} 0, & x^* \leq t; \\ \frac{F(x^*) - F(t)}{1 - F(t)}, & t < x^* < u; \\ 1, & x^* \geq u, \end{cases} \quad (2.1)$$

$$f_*(x^* | t; u) = \begin{cases} \frac{f(x^*)}{1 - F(t)}, & t < x^* < u; \\ \frac{1 - F(u^-)}{1 - F(t)}, & x^* = u; \\ 0, & \text{elsewhere,} \end{cases} \quad (2.2)$$

$$F_*^{-1}(s | t; u) = \begin{cases} F^{-1}(s + (1 - s)F(t)), & 0 \leq s < \frac{F(u) - F(t)}{1 - F(t)}; \\ u, & \frac{F(u) - F(t)}{1 - F(t)} \leq s \leq 1. \end{cases} \quad (2.3)$$

Note that we are interested in estimating the p th quantile of X (i.e., $F^{-1}(p)$) based on the observed data $X_1^* = x_1^*, \dots, X_n^* = x_n^*$. Thus, Theorems A.1 and A.2 in Appendix A.1 and the REC construction of Appendix A.2 have to be applied to functions (2.1)–(2.3), not F , f , F^{-1} .

2.1 Empirical Approach

As mentioned earlier, the empirical approach is restricted to the range of observed data. Indeed, based on x_1^*, \dots, x_n^* , the empirical estimator $\widehat{F}_{\text{EMP}}(t) = 0$. Thus, it cannot take full advantage of formulas (2.1)–(2.3), and yields a biased estimator that works within a limited range of quantile levels. In this case, the $F^{-1}(p)$ estimator is $\widehat{F}_{\text{EMP}}^{-1}(p) = x_{[\lceil np \rceil]}^* < u$, and as follows from Theorem A.1,

$$\widehat{F}_{\text{EMP}}^{-1}(p) \text{ is } \mathcal{AN} \left(F_*^{-1}(p), \frac{1}{n} \frac{p(1-p)}{f_*^2(F_*^{-1}(p))} \right), \quad 0 < p < \frac{F(u) - F(t)}{1 - F(t)}. \quad (2.4)$$

To see that this estimator is positively biased, i.e., any (estimable) quantile of the observable variable X^* is never below the corresponding quantile of the unobservable variable X (which is what we want to estimate), notice that for the mean parameter in (2.4), we have

$$F_*^{-1}(p) = F^{-1}(p + (1-p)F(t)) \geq F^{-1}(p), \quad 0 < p < \frac{F(u) - F(t)}{1 - F(t)},$$

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

2.2 MLE Approach

Parametric methods use the observed data x_1^*, \dots, x_n^* and fully recognize its distributional properties. The MLE approach is one of the most common estimation techniques. It takes into account (2.1)–(2.3) and finds parameter estimates by maximizing the following log-likelihood function:

$$\begin{aligned} \log \mathcal{L}(\boldsymbol{\theta} | x_1^*, \dots, x_n^*) &= \log \left[\prod_{i=1}^n f_*(x_i^* | t; u) \right] = \log \left[\prod_{i=1}^n \left[\frac{f(x_i^*)}{1 - F(t)} \right]^{\mathbf{1}_{\{t < x_i^* < u\}}} \left[\frac{1 - F(u^-)}{1 - F(t)} \right]^{\mathbf{1}_{\{x_i^* = u\}}} \right] \\ &= \sum_{i=1}^n \log [f(x_i^*)] \mathbf{1}_{\{t < x_i^* < u\}} - n \log [1 - F(t)] + \log [1 - F(u^-)] \sum_{i=1}^n \mathbf{1}_{\{x_i^* = u\}}, \end{aligned} \quad (2.5)$$

where $\mathbf{1}\{ \}$ denotes the indicator function.

Once parameter MLEs, $\widehat{\theta}_1, \dots, \widehat{\theta}_k$, are available, the p th quantile estimate is found by plugging those MLE values into the parametric expression of $F^{-1}(p) = h(\theta_1, \dots, \theta_k)$. Let us denote this estimator as $\widehat{F}_{\text{MLE}}^{-1}(p) = h(\widehat{\theta}_1, \dots, \widehat{\theta}_k)$. Then, as follows from the MLE's asymptotic distribution and the delta method,

$$\widehat{F}_{\text{MLE}}^{-1}(p) \text{ is } \mathcal{AN} \left(F^{-1}(p), \frac{1}{n} \mathbf{d}_\theta \mathbf{I}_\theta^{-1} \mathbf{d}'_\theta \right), \quad 0 < p < 1, \quad (2.6)$$

where $\mathbf{d}_\theta = \left(\partial h / \partial \hat{\theta}_1, \dots, \partial h / \partial \hat{\theta}_k \right) \Big|_{(\theta_1, \dots, \theta_k)}$, and the entries of \mathbf{I}_θ are given by (A.7) with g replaced by (2.2). Note that (2.6) is defined for $0 < p < 1$, while (2.4) for $0 < p < \frac{F(u) - F(t)}{1 - F(t)} \leq 1$.

2.3 PM Approach

A popular alternative to the MLE approach for estimation of loss distribution parameters is percentile matching. To estimate k unknown parameters with the PM method and using the ordered data $x_{(1)}^* \leq \dots \leq x_{(n)}^*$, one has to solve the following system of equations with respect to $\theta_1, \dots, \theta_k$:

$$F_*^{-1}(p_1) = x_{([\!|np_1|])}^*, \quad F_*^{-1}(p_2) = x_{([\!|np_2|])}^*, \quad \dots, \quad F_*^{-1}(p_k) = x_{([\!|np_k|])}^*,$$

where $p_1 < \dots < p_k < \frac{F(u) - F(t)}{1 - F(t)}$ and $x_{([\!|np_k|])}^* < u$. Once parameter PMs, $\tilde{\theta}_1, \dots, \tilde{\theta}_k$, are available, the p th quantile estimate is found by plugging those PM values into $F^{-1}(p) = h(\theta_1, \dots, \theta_k)$. Let us denote this estimator as $\hat{F}_{\text{PM}}^{-1}(p) = h(\tilde{\theta}_1, \dots, \tilde{\theta}_k)$. Then, as follows from Theorem A.2 and the delta method,

$$\hat{F}_{\text{PM}}^{-1}(p) \text{ is } \mathcal{AN} \left(F^{-1}(p), \frac{1}{n} \mathbf{d}_\theta \mathbf{D}_\theta^* \boldsymbol{\Sigma}_\theta (\mathbf{D}_\theta^*)' \mathbf{d}'_\theta \right), \quad 0 < p < 1, \quad (2.7)$$

where $\mathbf{d}_\theta = \left(\partial h / \partial \tilde{\theta}_1, \dots, \partial h / \partial \tilde{\theta}_k \right) \Big|_{(\theta_1, \dots, \theta_k)}$ and \mathbf{D}_θ^* is specified in Theorem A.2. The entries of $\boldsymbol{\Sigma}_\theta$ are given by (A.1) with g and G^{-1} replaced by expressions (2.2) and (2.3), respectively. Note that (2.7) is defined for $0 < p < 1$, while (2.4) for $0 < p < \frac{F(u) - F(t)}{1 - F(t)} \leq 1$.

3 RECs for Exponential and Pareto Models

In this section, we provide examples of RECs for exponential and single-parameter Pareto distributions under several data-truncation and censoring scenarios. For each model, we choose the (biased) empirical estimator of $F^{-1}(p)$ as the benchmark estimator. Then, using formulas (2.4), (2.6), and (2.7), we evaluate asymptotic relative efficiencies, ARE_p 's, for the MLE and PM estimators with respect to the empirical estimator, as well as ARE_p of PM with respect to MLE. The three definitions of ARE_p 's are given by equations (A.8)–(A.10).

3.1 Exponential Distribution

Let X_1, X_2, \dots be i.i.d. exponentially distributed random variables with cdf $F(x) = 1 - e^{-(x-x_0)/\theta}$, $x \geq x_0$, pdf $f(x) = (1/\theta)e^{-(x-x_0)/\theta}$, $x > x_0$, and qf $F^{-1}(s) = x_0 - \theta \log(1 - s)$, $0 \leq s \leq 1$, and where

$x_0 \geq 0$ is known and $\theta > 0$ is an unknown scale parameter. According to the model setup of Section 2, however, the X_i 's are unobservable. The data are generated by variables X_1^*, \dots, X_n^* which are i.i.d. with cdf, pdf, and qf given by (2.1), (2.2), and (2.3), respectively. This implies that when X_i 's are exponentially distributed, we have

$$F_*(x^* | t; u) = \begin{cases} 0, & x^* \leq t; \\ 1 - e^{-(x^*-t)/\theta}, & t < x^* < u; \\ 1, & x^* \geq u, \end{cases}$$

$$f_*(x^* | t; u) = \begin{cases} (1/\theta)e^{-(x^*-t)/\theta}, & t < x^* < u; \\ e^{-(u-t)/\theta}, & x^* = u; \\ 0, & \text{elsewhere,} \end{cases}$$

$$F_*^{-1}(s | t; u) = \begin{cases} -\theta \log(1-s) + t, & 0 \leq s < 1 - e^{-(u-t)/\theta}; \\ u, & 1 - e^{-(u-t)/\theta} \leq s \leq 1. \end{cases}$$

Now, for the empirical estimator $\widehat{F}_{\text{EMP}}^{-1}(p) = x_{([\!np])}^*$, the asymptotic result (2.4) becomes

$$\widehat{F}_{\text{EMP}}^{-1}(p) \text{ is } \mathcal{AN} \left(-\theta \log(1-p) + t, \frac{\theta^2 p}{n(1-p)} \right), \quad 0 < p < 1 - e^{-(u-t)/\theta}. \quad (3.1)$$

The statement (3.1) shows that the asymptotic bias of $\widehat{F}_{\text{EMP}}^{-1}(p)$ is $t - x_0$.

Further, MLE of θ is found by maximizing the log-likelihood (2.5) which in this case is

$$\log \mathcal{L}(\theta | x_1^*, \dots, x_n^*) = -\log \theta \sum_{i=1}^n \mathbf{1}\{t < x_i^* < u\} - \frac{1}{\theta} \sum_{i=1}^n \left[(x_i^* - t) \mathbf{1}\{t < x_i^* < u\} + (u - t) \mathbf{1}\{x_i^* = u\} \right].$$

It yields a closed-form solution for θ :

$$\widehat{\theta}_{\text{MLE}} = \frac{\sum_{i=1}^n \left[(x_i^* - t) \mathbf{1}\{t < x_i^* < u\} + (u - t) \mathbf{1}\{x_i^* = u\} \right]}{\sum_{i=1}^n \mathbf{1}\{t < x_i^* < u\}}.$$

This in turn implies that $\widehat{F}_{\text{MLE}}^{-1}(p) = x_0 - \widehat{\theta}_{\text{MLE}} \log(1-p)$, and the asymptotic result (2.6) becomes

$$\widehat{F}_{\text{MLE}}^{-1}(p) \text{ is } \mathcal{AN} \left(x_0 - \theta \log(1-p), \frac{\theta^2 \log^2(1-p)}{n(1 - e^{-(u-t)/\theta})} \right), \quad 0 < p < 1. \quad (3.2)$$

Furthermore, since for the exponential distribution there is only one unknown parameter θ , its PM estimator is derived by solving a single equation, $F_*^{-1}(p_1) = x_{([\!np_1])}^*$. Note that p_1 has to be chosen

from the range $0 < p_1 < 1 - e^{-(u-t)/\theta}$ (equivalently, $x_{([np_1])}^* < u$). In this case, the resulting estimator is also explicit and given by

$$\hat{\theta}_{\text{PM}} = \frac{t - x_{([np_1])}^*}{\log(1 - p_1)}.$$

Subsequently, $\hat{F}_{\text{PM}}^{-1}(p) = x_0 - \hat{\theta}_{\text{PM}} \log(1 - p)$, and the asymptotic result (2.7) becomes

$$\hat{F}_{\text{PM}}^{-1}(p) \text{ is } \mathcal{AN} \left(x_0 - \theta \log(1 - p), \frac{\theta^2}{n} \frac{p_1}{1 - p_1} \left[\frac{\log(1 - p)}{\log(1 - p_1)} \right]^2 \right), \quad 0 < p < 1. \quad (3.3)$$

Finally, we have everything in place for computation of ARE_p . Since $\hat{F}_{\text{EMP}}^{-1}(p)$ is our benchmark estimator which is biased, formulas (A.8) and (A.9) will be modified by replacing estimators' variances with their mean-square errors (MSE). The MSE ratios based on (3.1)–(3.3) are:

$$\text{ARE} \left(\hat{F}_{\text{MLE}}^{-1}(p), \hat{F}_{\text{EMP}}^{-1}(p) \right) = \frac{\frac{\theta^2}{n} \frac{p}{1-p} + (t - x_0)^2}{\frac{\theta^2}{n} \frac{\log^2(1-p)}{1 - e^{-(u-t)/\theta}}}, \quad 0 < p < 1 - e^{-(u-t)/\theta}, \quad (3.4)$$

$$\text{ARE} \left(\hat{F}_{\text{PM}}^{-1}(p), \hat{F}_{\text{EMP}}^{-1}(p) \right) = \frac{\frac{\theta^2}{n} \frac{p}{1-p} + (t - x_0)^2}{\frac{\theta^2}{n} \frac{\log^2(1-p)}{\log^2(1-p_1)} \frac{p_1}{1-p_1}}, \quad 0 < p < 1 - e^{-(u-t)/\theta}, \quad (3.5)$$

$$\text{ARE} \left(\hat{F}_{\text{PM}}^{-1}(p), \hat{F}_{\text{MLE}}^{-1}(p) \right) = \frac{\frac{\theta^2}{n} \frac{\log^2(1-p)}{1 - e^{-(u-t)/\theta}}}{\frac{\theta^2}{n} \frac{\log^2(1-p)}{\log^2(1-p_1)} \frac{p_1}{1-p_1}} = \frac{(1 - p_1) \log^2(1 - p_1)}{p_1 (1 - e^{-(u-t)/\theta})}, \quad 0 < p < 1. \quad (3.6)$$

Note that for $p \geq 1 - e^{-(u-t)/\theta}$, the ratios (3.4) and (3.5) are infinite because $\hat{F}_{\text{EMP}}^{-1}(p)$ is undefined. Also, in (3.6), the probability level p_1 has to be chosen from the range $0 < p_1 < 1 - e^{-(u-t)/\theta}$.

In Figure 3.1, RECs of quantile estimators of the exponential($x_0 = 100, \theta$) distribution are plotted for the left-truncation level $t = 500$ and right-censoring at $u = 2500$. In the first column of plots, the distribution is lighter tailed ($\theta = 250$) with $F(t) = 0.7981$, $F(u) = 0.9999$, and $F_*(u | t; u) = 0.9995$. In the second column of plots, the distribution has a heavier tail ($\theta = 500$) with $F(t) = 0.5507$, $F(u) = 0.9918$, and $F_*(u | t; u) = 0.9817$. Due to the high bias of the empirical estimator (which goes to ∞ as $p \rightarrow 0$), the vertical axes are plotted on the logarithmic scale to minimize visual distortions. Comparison of plots across the rows reveals a couple of patterns: first, in the top row it is clearly visible that a combination of heavier tail and a slightly smaller percentage of actually observed data $F_*(u | t; u)$ shifts all curves significantly upward (especially for small p); second, as is evident from all plots, the efficiency of PM estimators increases monotonically for $0 < p_1 < 0.80$ and then starts to

decrease for $0.80 < p_1 < 1$ (i.e., the curves PM_4 are above those of PM_3 which are above PM_2 , etc., but PM_6 are below PM_5). Thus the $p_1 \approx 0.80$ level is optimal for PM estimation. This fact is in agreement with the complete sample optimality result (see discussion in Section 3.1 of Brazauskas, 2009).

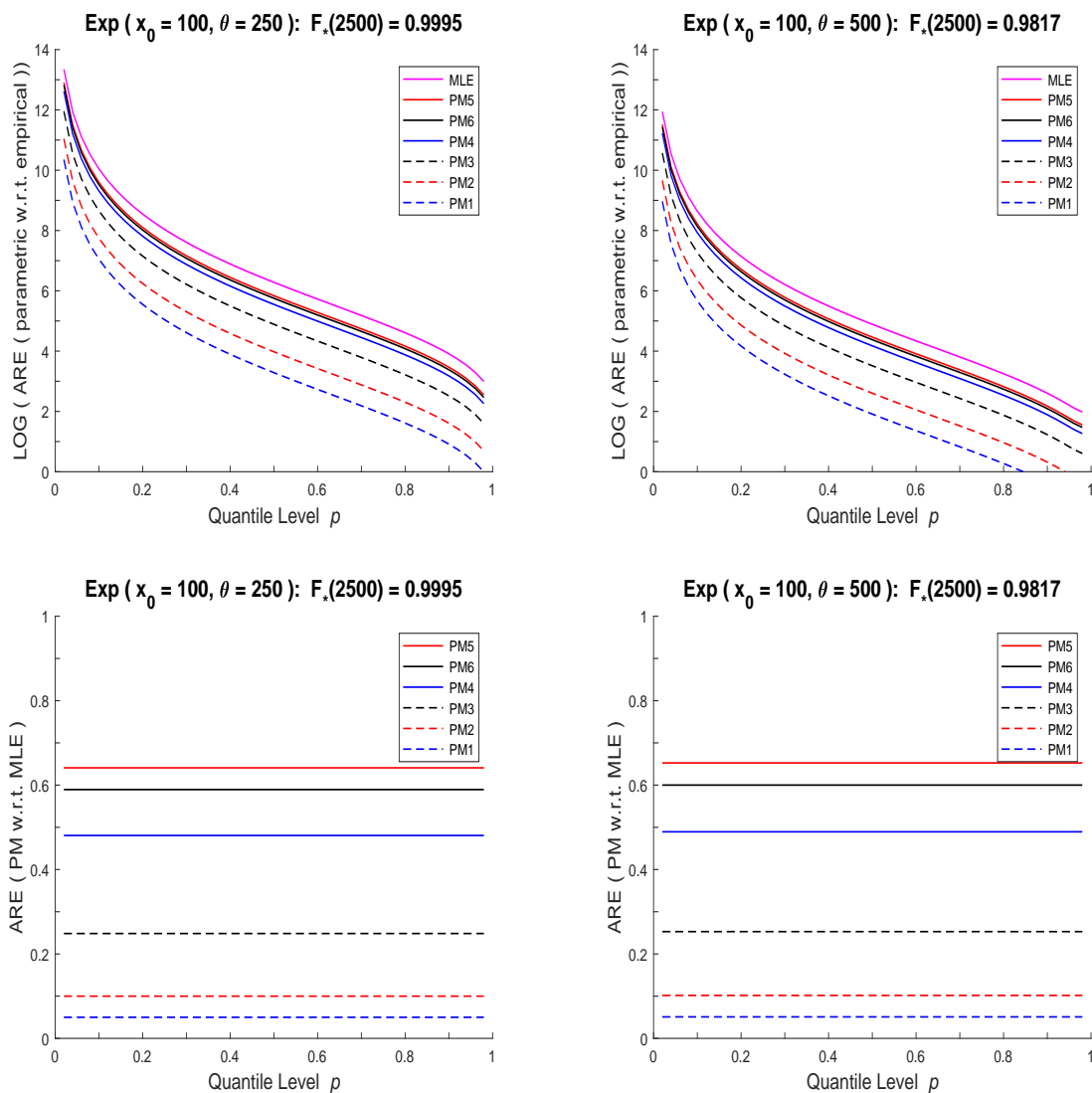


FIGURE 3.1. RECs of quantile estimators of the exponential($x_0 = 100, \theta$) distribution for $t = 500, u = 2500, n = 100$, and $\theta = 250$ (left column), $\theta = 500$ (right column). Level $p_1 = 0.05$ (PM_1), 0.10 (PM_2), 0.25 (PM_3), 0.50 (PM_4), 0.75 (PM_5), 0.90 (PM_6). *Top row*: Plots of formulas (3.4) and (3.5). *Bottom row*: Plots of formula (3.6).

3.2 Pareto Distribution

Let X_1, X_2, \dots be i.i.d. random variables distributed according to a single-parameter Pareto distribution with cdf $F(x) = 1 - (x_0/x)^\alpha$, $x > x_0$, pdf $f(x) = (\alpha/x_0)(x_0/x)^{\alpha+1}$, $x > x_0$, and qf $F^{-1}(s) = x_0(1-s)^{-1/\alpha}$, $0 \leq s \leq 1$. Here $x_0 > 0$ is known and $\alpha > 0$ is an unknown shape parameter, thus justifying the single-parameter characterization. As before, X_i 's are unobservable and the data are generated by variables X_1^*, \dots, X_n^* which are i.i.d. with cdf, pdf, and qf given by (2.1), (2.2), and (2.3), respectively. This implies that when X_i 's are Pareto distributed, we have

$$F_*(x^* | t; u) = \begin{cases} 0, & x^* \leq t; \\ 1 - (t/x^*)^\alpha, & t < x^* < u; \\ 1, & x^* \geq u, \end{cases}$$

$$f_*(x^* | t; u) = \begin{cases} (\alpha/t)(t/x^*)^{\alpha+1}, & t < x^* < u; \\ (t/u)^\alpha, & x^* = u; \\ 0, & \text{elsewhere,} \end{cases}$$

$$F_*^{-1}(s | t; u) = \begin{cases} t(1-s)^{-1/\alpha}, & 0 \leq s < 1 - (t/u)^\alpha; \\ u, & 1 - (t/u)^\alpha \leq s \leq 1. \end{cases}$$

Next, for the empirical estimator $\hat{F}_{\text{EMP}}^{-1}(p) = x_{([\!n p])}^*$, the asymptotic result (2.4) becomes

$$\hat{F}_{\text{EMP}}^{-1}(p) \text{ is } \mathcal{AN} \left(t(1-p)^{-1/\alpha}, \frac{(t/\alpha)^2}{n} \frac{p}{(1-p)^{1+2/\alpha}} \right), \quad 0 < p < 1 - (t/u)^\alpha. \quad (3.7)$$

As evident from the statement (3.7) and the fact that $t(1-p)^{-1/\alpha} \geq x_0(1-p)^{-1/\alpha}$ (since $t \geq x_0$), this estimator is asymptotically (positively) biased.

Further, MLE of α is found by maximizing the log-likelihood (2.5) which in this case is

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

It yields a closed-form solution for α :

$$\hat{\alpha}_{\text{MLE}} = \frac{\sum_{i=1}^n \mathbf{1}\{t < x_i^* < u\}}{\sum_{i=1}^n \left[\log(x_i^*/t) \mathbf{1}\{t < x_i^* < u\} + \log(u/t) \mathbf{1}\{x_i^* = u\} \right]}.$$

This in turn implies that $\widehat{F}_{\text{MLE}}^{-1}(p) = x_0(1-p)^{-1/\widehat{\alpha}_{\text{MLE}}}$, and the asymptotic result (2.6) becomes

$$\widehat{F}_{\text{MLE}}^{-1}(p) \text{ is } \mathcal{AN} \left(x_0(1-p)^{-1/\alpha}, \frac{1}{n} \frac{\log^2(1-p)}{(1-p)^{2/\alpha}} \frac{(x_0/\alpha)^2}{1-(t/u)^\alpha} \right), \quad 0 < p < 1. \quad (3.8)$$

Furthermore, similar to the exponential distribution case, PM estimator of α is derived by solving a single equation, $F_*^{-1}(p_1) = x_{(\lceil np_1 \rceil)}^*$, where $x_{(\lceil np_1 \rceil)}^* < u$. The resulting estimator is given by

$$\widehat{\alpha}_{\text{PM}} = \frac{\log(1-p_1)}{\log(t/x_{(\lceil np_1 \rceil)}^*)}.$$

Subsequently, $\widehat{F}_{\text{PM}}^{-1}(p) = x_0(1-p)^{-1/\widehat{\alpha}_{\text{PM}}}$, and the asymptotic result (2.7) becomes

$$\widehat{F}_{\text{PM}}^{-1}(p) \text{ is } \mathcal{AN} \left(x_0(1-p)^{-1/\alpha}, \frac{(x_0/\alpha)^2}{n} \frac{p_1 \log^2(1-p)}{(1-p_1) \log^2(1-p_1)} (1-p)^{-2/\alpha} \right), \quad 0 < p < 1. \quad (3.9)$$

Note that p_1 has to be chosen from the range $0 < p_1 < 1 - (t/u)^\alpha$ (equivalently, $x_{(\lceil np_1 \rceil)}^* < u$).

Finally, for computation of ARE_p , formulas (A.8) and (A.9) are modified the same way as in Section 3.1. The MSE ratios based on (3.7)–(3.9) are:

$$\begin{aligned} \text{ARE} \left(\widehat{F}_{\text{MLE}}^{-1}(p), \widehat{F}_{\text{EMP}}^{-1}(p) \right) &= \frac{\frac{(t/\alpha)^2}{n} \frac{p}{(1-p)^{1+2/\alpha}} + [t(1-p)^{-1/\alpha} - x_0(1-p)^{-1/\alpha}]^2}{\frac{1}{n} \frac{\log^2(1-p)}{(1-p)^{2/\alpha}} \frac{(x_0/\alpha)^2}{1-(t/u)^\alpha}} \\ &= \frac{p/(1-p) + n\alpha^2(1-x_0/t)^2}{(x_0/t)^2 \log^2(1-p)/(1-(t/u)^\alpha)}, \quad 0 < p < 1 - (t/u)^\alpha, \quad (3.10) \end{aligned}$$

$$\begin{aligned} \text{ARE} \left(\widehat{F}_{\text{PM}}^{-1}(p), \widehat{F}_{\text{EMP}}^{-1}(p) \right) &= \frac{\frac{(t/\alpha)^2}{n} \frac{p}{(1-p)^{1+2/\alpha}} + [t(1-p)^{-1/\alpha} - x_0(1-p)^{-1/\alpha}]^2}{\frac{(x_0/\alpha)^2}{n} \frac{p_1 \log^2(1-p)}{(1-p_1) \log^2(1-p_1)} (1-p)^{-2/\alpha}} \\ &= \frac{(t/x_0)^2 [p/(1-p) + n\alpha^2(1-x_0/t)^2]}{(p_1/(1-p_1)) (\log(1-p)/\log(1-p_1))^2}, \quad 0 < p < 1 - (t/u)^\alpha, \quad (3.11) \end{aligned}$$

$$\begin{aligned} \text{ARE} \left(\widehat{F}_{\text{PM}}^{-1}(p), \widehat{F}_{\text{MLE}}^{-1}(p) \right) &= \frac{\frac{1}{n} \frac{\log^2(1-p)}{(1-p)^{2/\alpha}} \frac{(x_0/\alpha)^2}{1-(t/u)^\alpha}}{\frac{(x_0/\alpha)^2}{n} \frac{p_1 \log^2(1-p)}{(1-p_1) \log^2(1-p_1)} (1-p)^{-2/\alpha}} \\ &= \frac{(1-p_1) \log^2(1-p_1)}{p_1(1-(t/u)^\alpha)}, \quad 0 < p < 1. \quad (3.12) \end{aligned}$$

For $p \geq 1 - (t/u)^\alpha$, the ratios (3.10) and (3.11) are infinite. In (3.12), the probability level p_1 has to be chosen from the range $0 < p_1 < 1 - (t/u)^\alpha$.

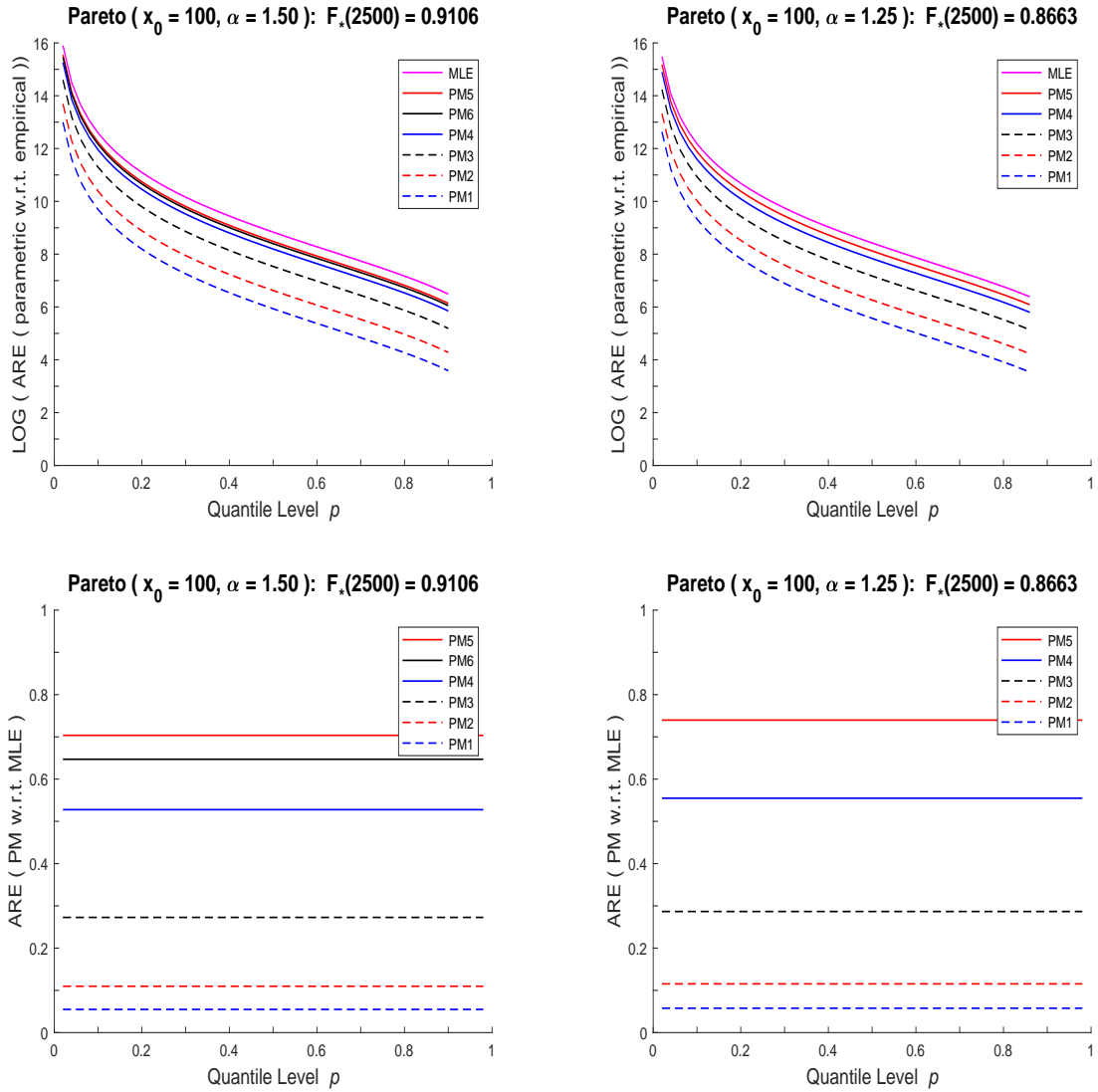


FIGURE 3.2. RECs of quantile estimators of the Pareto ($x_0 = 100, \alpha$) distribution for $t = 500$, $u = 2500$, $n = 100$, and $\alpha = 1.50$ (left column), $\alpha = 1.25$ (right column). Level $p_1 = 0.05$ (PM₁), 0.10 (PM₂), 0.25 (PM₃), 0.50 (PM₄), 0.75 (PM₅), 0.90 (PM₆). *Top row*: Plots of formulas (3.10) and (3.11). *Bottom row*: Plots of formula (3.12).

In Figure 3.2, RECs of quantile estimators of the Pareto ($x_0 = 100, \alpha$) distribution are plotted for the left-truncation level $t = 500$ and right-censoring at $u = 2500$. In the first column of plots, the distribution is heavy tailed ($\alpha = 1.50$) with $F(t) = 0.9106$, $F(u) = 0.9920$, and $F_*(u|t;u) = 0.9106$. In the second column of plots, the distribution has even heavier tail ($\alpha = 1.25$) with $F(t) = 0.8663$, $F(u) = 0.9821$, and $F_*(u|t;u) = 0.8663$. Due to the high bias of the empirical estimator (which goes to ∞ as $p \rightarrow 0$), the vertical axes are plotted on the logarithmic scale to minimize visual distortions.

Comparison of plots shows the same ordering of PM curves as those under the exponential distribution assumption. The choice of $p_1 \approx 0.80$ is also optimal for PM estimation. A change from heavy to even heavier tail and a decrease in the percentage of actually observed data $F_*(u|t;u)$ results in less pronounced shifts of the Pareto-based REC curves; but they are much higher than the exponential RECs. Thus, since both models are truncated and censored at the identical $t = 500$ and $u = 2500$, this suggests that the significant differences in the REC curves between the distributions can be used to construct a model selection method. This idea will be further discussed in Section 4.

4 Evaluation of Model Uncertainty

In this section, using simulated data we demonstrate how model uncertainty can emerge in a surprising way and examine how wrong quantile estimates can be when one makes incorrect modeling assumptions. In particular, we generate $n = 50$ observations from the exponential distribution of Section 3.1 (with $x_0 = 100$, $\theta = 500$, $t = 500$, $u = 2500$), fit the exponential model using MLE and PM ($p_1 = 0.80$) estimators to it, and perform standard model diagnostics (e.g., quantile-quantile plots) and validation (e.g., Kolmogorov-Smirnov and Anderson-Darling tests). As expected, the exponential distribution is not rejected by any of the tests. Then, using the same data we repeat the exercise by assuming a Pareto distribution, and find that it also passes all the tests. In both cases, we additionally compute AIC and BIC values, which under the *incorrect* Pareto assumption are better than the ones under the *correct* exponential assumption. Next, to make sure that this conclusion was not random, we simulate $n = 50$ observations from the Pareto distribution of Section 3.2 (with $x_0 = 100$, $\alpha = 1.50$, $t = 500$, $u = 2500$), fit and validate both models, and find yet again that both distributional assumptions are acceptable. This exercise shows that standard model diagnostic methods can mislead the decision maker, which would be no big deal if quantile estimates based on incorrect modeling assumptions were close to the true values of quantiles, but that's not the case. For completeness, we include the empirical estimates of quantiles although it is known they are incorrect. Below we provide the details of the described exercises so the interested reader can reproduce the results.

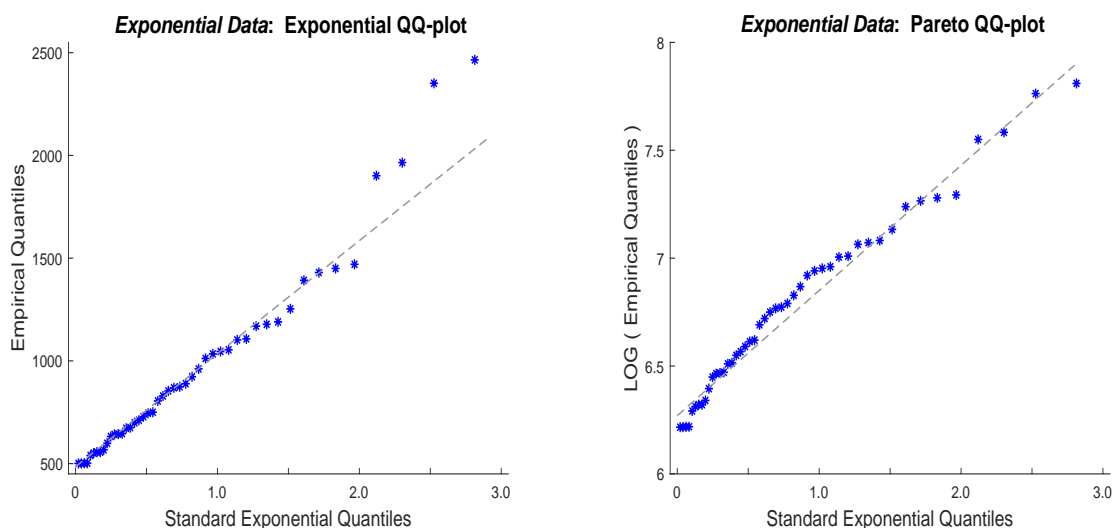
The data sets were simulated using R with a seed of 200 (it is used to initialize the random number generator). They are presented in Table 4.1, where censored observations are *italicized*.

TABLE 4.1. Left-truncated (at $t = 500$) and right-censored (at $u = 2500$) data simulated from the exponential ($x_0 = 100, \theta = 500$) and Pareto ($x_0 = 100, \alpha = 1.50$) distributions.

Exponential Data: 501, 501, 502, 502, 540, 551, 556, 556, 567, 599, 632, 642, 644, 646, 672, 675, 699, 711, 728, 745, 750, 805, 829, 854, 869, 874, 889, 923, 961, 1012, 1034, 1046, 1054, 1102, 1107, 1169, 1178, 1190, 1253, 1392, 1430, 1450, 1470, 1901, 1965, 2351, 2465, 2500, 2500, 2500.

Pareto Data: 516, 526, 535, 542, 550, 570, 593, 603, 605, 608, 609, 661, 674, 688, 694, 728, 734, 751, 751, 768, 778, 782, 786, 797, 825, 836, 836, 847, 940, 962, 968, 1034, 1080, 1115, 1118, 1120, 1134, 1137, 1175, 1213, 1224, 1271, 1379, 1725, 1861, 2000, 2500, 2500, 2500, 2500.

In Figure 4.1, the quantile-quantile plots (QQ-plots) are provided. The plots are parameter free. That is, since the exponential and Pareto distributions are location-scale and log-location-scale families, respectively, their QQ-plots can be constructed without first estimating model parameters. Note also that only actual data can be used in these plots (i.e., no observations $u = 2500$). As is evident from Figure 4.1, the points in all graphs form a (roughly) straight line; thus both distributions are acceptable for both data sets.



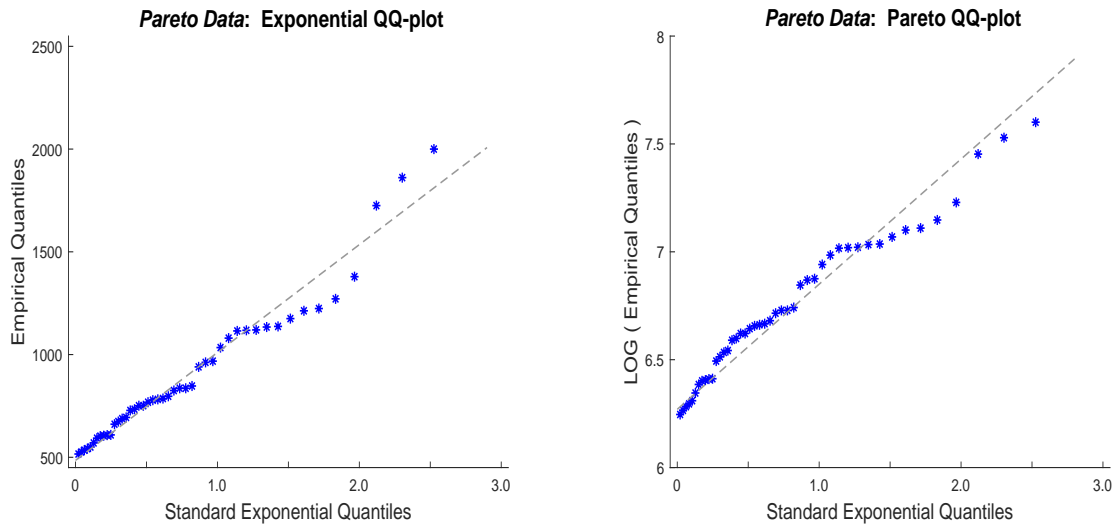


FIGURE 4.1. Exponential and Pareto quantile-quantile plots for the data sets of Table 4.1. The dashed line represents the “best” fit line. LEFT column: $y = 485 + 550x$ (top) and $y = 485 + 525x$ (bottom). RIGHT column: $y = 6.27 + 0.58x$ (top and bottom).

To formally evaluate the appropriateness of the fitted model to data, we perform KS and AD goodness-of-fit tests. The models are fitted using two parameter estimation methods, MLE and PM ($p_1 = 0.80$), to check the sensitivity of overall conclusions to model fitting procedures. The values of the test statistics along with the corresponding p -values are reported in Table 4.2. (The p -values are computed using parametric bootstrap with 1000 simulation runs. For a brief description of the parametric bootstrap procedure, see, for example, Section 20.4.5 of Klugman *et al.* 2012). We can see that except for one isolated case (Pareto data, Pareto model, PM estimation) the p -values are above 0.10 for both distributions, all parameter estimation methods, and both tests. Thus, the fitted exponential and Pareto models are acceptable for both data sets. In addition, the table contains AIC and BIC values, which can be used as model selection tools. Based on these metrics (smaller values are better), the Pareto model would be chosen for both data sets. Of course, the decision to accept Pareto when data came from an exponential distribution is incorrect.

TABLE 4.2. Parameter estimates, goodness-of-fit measures, and information criteria for the exponential and Pareto models fitted to the data sets of Table 4.1.

Assumed Model	Parameter Estimates	Goodness-of-Fit Measures		Information Criteria	
		Kolmogorov-Smirnov (p -value*)	Anderson-Darling (p -value*)	AIC	BIC
<i>Exponential Data</i>					
Exponential	$\hat{\theta}_{\text{MLE}} = 595.57$	0.077 (0.914)	1.099 (0.317)	696.62	698.53
	$\hat{\theta}_{\text{PM}} = 554.23$	0.076 (0.637)	0.942 (0.374)	696.86	698.78
Pareto	$\hat{\alpha}_{\text{MLE}} = 1.491$	0.095 (0.371)	0.898 (0.344)	679.29	681.20
	$\hat{\alpha}_{\text{PM}} = 1.572$	0.109 (0.538)	1.112 (0.262)	682.92	684.83
<i>Pareto Data</i>					
Exponential	$\hat{\theta}_{\text{MLE}} = 579.33$	0.109 (0.573)	0.564 (0.670)	695.99	697.90
	$\hat{\theta}_{\text{PM}} = 443.01$	0.102 (0.610)	1.006 (0.329)	696.12	698.03
Pareto	$\hat{\alpha}_{\text{MLE}} = 1.487$	0.128 (0.354)	1.025 (0.294)	678.29	680.20
	$\hat{\alpha}_{\text{PM}} = 1.816$	0.195 (0.000)	2.525 (0.000)	681.44	683.35

* The p -values are computed using parametric bootstrap with 1000 simulation runs.

Next, to see whether it really matters which model we select at this stage of the analysis, we have to examine the true probability models that generated data and check how much off target our upper quantile estimates are. For the data sets of Table 4.1, the underlying distributions are exponential ($x_0 = 100, \theta = 500$) and Pareto ($x_0 = 100, \alpha = 1.50$), with the quantile functions given by:

$$F^{-1}(p) = 100 - 500 \log(1 - p) \quad (\text{exponential}), \quad F^{-1}(p) = 100(1 - p)^{-1/1.50} \quad (\text{Pareto}).$$

Thus, the true values of the 90%, 95% and 99% quantiles (estimation targets) are:

$$\text{Exponential: } F^{-1}(0.90) = 1251, \quad F^{-1}(0.95) = 1598, \quad F^{-1}(0.99) = 2403.$$

$$\text{Pareto: } F^{-1}(0.90) = 464, \quad F^{-1}(0.95) = 737, \quad F^{-1}(0.99) = 2154.$$

The quantile estimation results are summarized in Table 4.3. There, we clearly see that the parametric estimates of the quantiles based on the correctly identified model are fairly close to their targets, but those based on the incorrect model are significantly off their targets. Also, the empirical estimates are way off target ($\hat{F}^{-1}(0.99) = 2500$ for the exponential data set is a lucky coincidence, not a rule).

TABLE 4.3. Parametric and empirical estimates of the 90%, 95% and 99% quantiles for the exponential and Pareto data sets of Table 4.1.

Data Set	Estimation Methodology		Quantile Estimates		
	Assumed Model	Estimator	$\hat{F}^{-1}(0.90)$	$\hat{F}^{-1}(0.95)$	$\hat{F}^{-1}(0.99)$
Exponential	Exponential	MLE	1471	1884	2843
		PM	1387	1774	2674
	Pareto	MLE	468	746	2194
		PM	436	679	1902
	Empirical	— — —	2004	2484	2500
Pareto	Exponential	MLE	1434	1836	2768
		PM	1123	1431	2146
	Pareto	MLE	471	750	2216
		PM	356	522	1269
	Empirical	— — —	1875	2500	2500

Finally, in Table 4.4 we present estimated RECs, given by (3.4) and (3.10), at selected quantile levels. The curves are estimated using the MLE values from Table 4.2 and show how many times the parametric approach is more efficient than the empirical one in estimating a quantile. Note that as was seen in Figures 3.1 and 3.2, RECs based on PM estimators have the same shapes as those of MLE, just rescaled by a constant (smaller than one). Thus PM based conclusions would not change from those of MLE and one method of analysis will be sufficient. What stands out from these computations is the vast differences between the corresponding exponential and Pareto RECs, when they are estimated using the same data set (especially for small p). We conjecture that with some additional work one can develop an effective diagnostic tool to differentiate between the models.

TABLE 4.4. Estimated Pareto and exponential RECs (MLE, empirical) at selected quantile levels. Model parameter estimates are from Table 4.2, based on the data of Table 4.1.

Quantile Level p	Exponential Data		Pareto Data	
	Exponential Model	Pareto Model	Exponential Model	Pareto Model
0.05	8293	615077	8792	611390
0.10	1971	145899	2089	145025
0.25	267	19631	283	19513
0.50	47	3413	50	3393
0.75	13	877	14	872
0.90	6	344	6	342
0.95	4	228	5	227

5 Concluding Remarks

The relative efficiency curves, REC, were introduced as a practical tool for comparison of two competing statistical procedures, when data are complete. In this paper, we have redesigned and extended such curves to the left-truncated and right-censored data scenarios that are common in insurance analytics. Our illustrations have focused on the parametric (MLE and PM) and empirical nonparametric approaches for estimation of quantiles that are key inputs for further risk analysis (e.g., contract pricing, risk measurement, capital allocation). Further, we have developed specific examples of RECs for exponential and single-parameter Pareto distributions under a few data truncation and censoring scenarios. Then, using simulated exponential and Pareto data we have examined how wrong quantile estimates can be when incorrect modeling assumptions are made. The numerical analysis involved application of standard model diagnostics and validation (e.g., QQ-plots, KS and AD tests, AIC and BIC criteria) and has demonstrated how those methods can mislead the decision maker. Finally, the newly developed RECs have been applied to study the discrepancies between the quality of quantile estimates of the fitted exponential and Pareto distributions. Our conclusion is that RECs have strong potential for being developed into an effective diagnostic tool in this context.

Author Contributions

Both authors have contributed equally to various aspects of the paper, including methodology, analysis, simulations, writing, and reviewing.

Funding

This research received no external funding.

Acknowledgements

The authors are very appreciative of valuable suggestions, insightful queries, and useful comments provided by three anonymous referees, which helped to significantly improve the paper.

Conflicts of Interest

The authors declare no conflict of interest.

Appendix

In the appendix, we provide some theoretical results that are key to analytic derivations in the paper. Specifically, in Appendix A.1, the asymptotic normality theorems for sample quantiles and percentile-matching (PM) estimators of model parameters are presented. The construction of the relative efficiency curves (REC) is described in Appendix A.2. Note that more detailed presentations of parts of this material are available in Brazauskas (2009) and Yu and Brazauskas (2017).

Suppose we have a sample of *independent and identically distributed (i.i.d.)* continuous random variables, X_1, \dots, 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 $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$. Further, let $X_{(1)} \leq \dots \leq X_{(n)}$ denote the ordered sample values. Also, throughout the paper the notation \mathcal{AN} is used for “asymptotically normal.”

A.1 Asymptotic Theorems

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 A.1 [ASYMPTOTIC NORMALITY OF SAMPLE QUANTILES]

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

$$\sigma_{ij}^2 = \frac{p_i(1-p_j)}{g(G^{-1}(p_i))g(G^{-1}(p_j))}. \quad (\text{A.1})$$

In the univariate case ($k = 1$), the sample quantile

$$X_{(\lceil np \rceil)} \text{ is } \mathcal{AN} \left(G^{-1}(p), \frac{1}{n} \frac{p(1-p)}{g^2(G^{-1}(p))} \right). \quad (\text{A.2})$$

The main drawback of statistical inference based on the empirical nonparametric approach is that it is restricted to the range of observed data. For the problems encountered with claim severity data, this is a major limitation. Therefore, a more appropriate alternative is to estimate distribution quantiles parametrically, which first requires estimates of the model parameters and then those values are applied to $G^{-1}(p)$. The most common technique for parameter estimation is MLE. Its asymptotic distribution is well known and available, for example, in Section 4.2 of Serfling (1980).

Percentile matching is a popular alternative to the MLE approach for estimation of loss distribution parameters (see Klugman *et al.*, 2012, Section 13.1). If the distribution has k unknown parameters, $(\theta_1, \dots, \theta_k)$, PM estimators are found by matching $G^{-1}(p_i)$ with $X_{(\lceil np_i \rceil)}$, $i = 1, \dots, k$, and then solving the resulting system of equations with respect to $\theta_1, \dots, \theta_k$. Assuming the system of equations has a unique solution, it is clear that PM estimators of $\theta_1, \dots, \theta_k$ are functions of $X_{(\lceil np_1 \rceil)}, \dots, X_{(\lceil np_k \rceil)}$. Let us denote these estimators as $\tilde{\theta}_i = h_i^*(X_{(\lceil np_1 \rceil)}, \dots, X_{(\lceil np_k \rceil)})$, $i = 1, \dots, k$. Their joint asymptotic normality follows, with some modifications, from Theorem A.1 and the delta method (see, e.g., Section 3.3 of Serfling, 1980).

Theorem A.2 [ASYMPTOTIC NORMALITY OF PMS]

Let $\tilde{\boldsymbol{\theta}} = (\tilde{\theta}_1, \dots, \tilde{\theta}_k)$ denote the PM estimator of parameter $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$. Then,

$$(\tilde{\theta}_1, \dots, \tilde{\theta}_k) \text{ is } \mathcal{AN} \left((\theta_1, \dots, \theta_k), \frac{1}{n} \mathbf{D}_{\boldsymbol{\theta}}^* \boldsymbol{\Sigma}_{\boldsymbol{\theta}} (\mathbf{D}_{\boldsymbol{\theta}}^*)' \right), \quad (\text{A.3})$$

where the entries of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ are given by (A.1) and $\mathbf{D}_{\boldsymbol{\theta}}^* = [d_{ij}^*]_{k \times k}$ is the Jacobian of the transformations h_1^*, \dots, h_k^* evaluated at $(\theta_1, \dots, \theta_k)$, that is, $d_{ij}^* = \partial h_i^* / \partial X_{(\lceil np_j \rceil)} \Big|_{(\theta_1, \dots, \theta_k)}$.

A.2 Relative Efficiency Curves

For complete data, the relative efficiency curve, REC, was introduced by Brazauskas (2009). It is constructed by using asymptotic properties of quantile estimators. Suppose two asymptotically normal estimators of a fixed quantile of the underlying distribution are available. Plotting the ratio of their variances versus the probability level of quantile yields an REC. Such a curve provides information about the accuracy of one estimator relative to another when both are designed to estimate the same (fixed but arbitrary) quantile of the distribution. If one or both estimators are biased, REC is constructed by replacing their variances with the mean-square errors.

Next, for a fixed probability level p , $0 < p < 1$, consider the empirical nonparametric and parametric estimators of the population quantile $G_{\boldsymbol{\theta}}^{-1}(p)$. Then, as follows from (A.2), the empirical estimator $\hat{G}_{\text{EMP}}^{-1}(p) = X_{(\lceil np \rceil)}$ satisfies:

$$\hat{G}_{\text{EMP}}^{-1}(p) \text{ is } \mathcal{AN} \left(G_{\boldsymbol{\theta}}^{-1}(p), \frac{1}{n} \frac{p(1-p)}{g_{\boldsymbol{\theta}}^2(G_{\boldsymbol{\theta}}^{-1}(p))} \right). \quad (\text{A.4})$$

For MLE and PM estimators, we use their asymptotic distributions in conjunction with the delta method (where $G_{\boldsymbol{\theta}}^{-1}(p)$ is viewed as a function of $\boldsymbol{\theta}$, say, $h(\boldsymbol{\theta})$) and arrive at the following results:

$$\hat{G}_{\text{MLE}}^{-1}(p) \text{ is } \mathcal{AN} \left(G_{\boldsymbol{\theta}}^{-1}(p), \frac{1}{n} \mathbf{d}_{\boldsymbol{\theta}} \mathbf{I}_{\boldsymbol{\theta}}^{-1} \mathbf{d}_{\boldsymbol{\theta}}' \right) \quad (\text{A.5})$$

and

$$\hat{G}_{\text{PM}}^{-1}(p) \text{ is } \mathcal{AN} \left(G_{\boldsymbol{\theta}}^{-1}(p), \frac{1}{n} \mathbf{d}_{\boldsymbol{\theta}} \mathbf{D}_{\boldsymbol{\theta}}^* \boldsymbol{\Sigma}_{\boldsymbol{\theta}} (\mathbf{D}_{\boldsymbol{\theta}}^*)' \mathbf{d}_{\boldsymbol{\theta}}' \right). \quad (\text{A.6})$$

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

$$I_{ij} = \mathbf{E} \left[\frac{\partial \log g(X)}{\partial \theta_i} \cdot \frac{\partial \log g(X)}{\partial \theta_j} \right], \quad (\text{A.7})$$

matrices \mathbf{D}_θ^* and Σ_θ are same as those specified in (A.3), and $\mathbf{d}_\theta = \left(\partial h / \partial \hat{\theta}_1, \dots, \partial h / \partial \hat{\theta}_k \right) \Big|_{(\theta_1, \dots, \theta_k)}$.

Note that the asymptotic variances of $\widehat{G}_{\text{MLE}}^{-1}(p)$ and $\widehat{G}_{\text{PM}}^{-1}(p)$ also depend on p .

Now, the asymptotic relative efficiency, ARE, of $\widehat{G}_{\text{EMP}}^{-1}(p)$ relative to $\widehat{G}_{\text{MLE}}^{-1}(p)$ is

$$\text{ARE}_p := \text{ARE} \left(\widehat{G}_{\text{EMP}}^{-1}(p), \widehat{G}_{\text{MLE}}^{-1}(p) \right) = \frac{g_\theta^2(G_\theta^{-1}(p))}{p(1-p)} \mathbf{d}_\theta \mathbf{I}_\theta^{-1} \mathbf{d}_\theta' \quad \text{for } 0 < p < 1, \quad (\text{A.8})$$

and relative to $\widehat{G}_{\text{PM}}^{-1}(p)$ it is

$$\text{ARE}_p := \text{ARE} \left(\widehat{G}_{\text{EMP}}^{-1}(p), \widehat{G}_{\text{PM}}^{-1}(p) \right) = \frac{g_\theta^2(G_\theta^{-1}(p))}{p(1-p)} \mathbf{d}_\theta \mathbf{D}_\theta^* \Sigma_\theta (\mathbf{D}_\theta^*)' \mathbf{d}_\theta' \quad \text{for } 0 < p < 1. \quad (\text{A.9})$$

For comparison of $\widehat{G}_{\text{PM}}^{-1}(p)$ relative to $\widehat{G}_{\text{MLE}}^{-1}(p)$, we have

$$\text{ARE}_p := \text{ARE} \left(\widehat{G}_{\text{PM}}^{-1}(p), \widehat{G}_{\text{MLE}}^{-1}(p) \right) = \frac{\mathbf{d}_\theta \mathbf{I}_\theta^{-1} \mathbf{d}_\theta'}{\mathbf{d}_\theta \mathbf{D}_\theta^* \Sigma_\theta (\mathbf{D}_\theta^*)' \mathbf{d}_\theta'} \quad \text{for } 0 < p < 1. \quad (\text{A.10})$$

Plotting the points (p, ARE_p) yields corresponding relative efficiency curves, where ARE_p is defined by (A.8), (A.9), or (A.10).

References

- [1] Aggarwal, Ankur, Michael B. Beck, Matthew Cann, Tim Ford, Dan Georgescu, Nirav Morjaria, Andrew D. Smith, Yvonne Taylor, Andreas Tsanakas, Louise Witts, and Ivy Ye. (2016). Model risk – daring to open up the black box. *British Actuarial Journal*, **21**(2), 229–296.
- [2] Alexander, Carol, and Jose M. Sarabia. (2012). Quantile uncertainty and value-at-risk model risk. *Risk Analysis*, **32**(8), 1293–1308.
- [3] Bignozzi, Valeria, Giovanni Puccetti, and Ludger Rüschemdorf. (2015). Reducing model risk via positive and negative dependence assumptions. *Insurance: Mathematics and Economics*, **61**, 17–26.
- [4] Black, Rob, Andreas Tsanakas, Andrew D. Smith, Michael B. Beck, Iain D. MacLugash, Jasvir Grewal, Louise Witts, Nirav Morjaria, R.J. Green, and Zhixin Lim. (2018). Model risk: illuminating the black box. *British Actuarial Journal*, **23**(2), 1–58.
- [5] Brazauskas, Vytautas. (2009). Quantile estimation and the statistical relative efficiency curve. *Metron*, **LXVII**(3), 289–301.
- [6] Brazauskas, Vytautas, and Thomas Kaiser. (2004). Discussion of “Empirical estimation of risk measures and related quantities” by Jones and Zitikis. *North American Actuarial Journal*, **8**(3), 114–117.
- [7] Brazauskas, Vytautas, and Andreas Kleefeld. (2016). Modeling severity and measuring tail risk of Norwegian fire claims. *North American Actuarial Journal*, **20**(1), 1–16.

- [8] Cairns, Andrew J. (2000). A discussion of parameter and model uncertainty in insurance. *Insurance: Mathematics and Economics*, **27**(3), 313–330.
- [9] Cambou, Mathieu, and Damir Filipović. (2017). Model uncertainty and scenario aggregation. *Mathematical Finance*, **27**(2), 534–567.
- [10] Cont, Rama, Romain Deguest, and Giacomo Scandolo. (2010). Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance*, **10**(6), 593–606.
- [11] Embrechts, Paul, Bin Wang, and Ruodu Wang. (2015). Aggregation-robustness and model uncertainty of regulatory risk measures. *Finance and Stochastics*, **19**(4), 763–790.
- [12] Glasserman, Paul, and Xingbo Xu. (2014). Robust risk measurement and model risk. *Quantitative Finance*, **14**(1), 1–30.
- [13] Hartman, Brian, Robert Richardson, and Rylan Bateman. (2017). *Parameter Uncertainty*. Casualty Actuarial Society, Canadian Institute of Actuaries, Society of Actuaries. Technical Report.
- [14] Hong, Liang, Todd Kuffner, and Ryan Martin. (2018). On prediction of future insurance claims when the model is uncertain. *Variance*, **12**(1), 90–99.
- [15] Jones, Bruce L., and Ricardas Zitikis. (2003). Empirical estimation of risk measures and related quantities (with discussion). *North American Actuarial Journal*, **7**(4), 44–54. Discussion: **8**(3), 114–117; Reply: **8**(3), 117–118.
- [16] Kaiser, Thomas, and Vytautas Brazauskas. (2006). Interval estimation of actuarial risk measures. *North American Actuarial Journal*, **10**(4), 249–268.
- [17] Klugman, Stuart A., Harry H. Panjer, and Gordon E. Willmot. (2012). *Loss Models: From Data to Decisions*, 4th edition. Wiley, New York.
- [18] Modarres, Reza, Tapan K. Nayak, and Joseph L. Gastwirth. (2002). Estimation of upper quantiles under model and parameter uncertainty. *Computational Statistics and Data Analysis*, **39**(4), 529–554.
- [19] Samanthi, Ranadeera, Wei Wei, and Vytautas Brazauskas. (2017). Comparing the riskiness of dependent portfolios via nested L -statistics. *Annals of Actuarial Science*, **11**(2), 237–252.
- [20] Serfling, Robert J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.
- [21] Yu, Daoping, and Vytautas Brazauskas. (2017). Model uncertainty in operational risk modeling due to data truncation: A single risk case. *Risks*, **5**(49), 17 pages, doi:10.3390/risks5030049.