QATAR UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

IMPROVED INFERENCE FOR THE SCALE PARAMETER IN THE LOMAX DISTRIBUTION BASED ON ADJUSTED PROFILE LIKELIHOOD FUNCTIONS

BY

MAISOUN F. M. SEWAILEM

A Thesis Submitted to

the College of Arts and Sciences

in Partial Fulfillment of the Requirements for the Degree of

Masters of Science in Master of Science in Applied Statistics

June 2022

COMMITTEE PAGE

The members of the Committee approve the Thesis of
Maisoun F. M. Sewailem defended on 2022.

Prof. Ayman Baklizi
Thesis/Dissertation Supervisor

Dr. Saddam Abbasi
Committee Member

Dr. Faiz Elfaki
Committee Member

Approved:

Dean, College of Arts and Sciences

ABSTRACT

SEWAILEM, MAISOUN, F., Masters: June : Master of Science in Applied Statistics

Title: Improved Inference for the Scale Parameter in the Lomax Distribution Based on Adjusted Profile Likelihood Functions.

Supervisor of Thesis: Ayman, Suleiman, Baklizi.

In this thesis, we consider improving maximum likelihood inference for the scale parameter of the Lomax distribution. The improvement is based on using modification to the maximum likelihood estimator based on Barndorff-Nielsen's modified profile likelihood function. We apply this modification to obtain improved estimator for the scale parameter of the Lomax distribution in the presence of a nuisance shape parameter. Due to its complicated expressions, several approximations to Barndorff-Nielsen's modified profile log-likelihood function are used, including the modification based on the empirical covariances and the modification based on an ancillary statistic approximation. We consider complete as well as type I and type II censored data. Comparison between maximum profile likelihood estimator and modified profile likelihood estimators in terms of their biases and mean squared errors were carried out using simulation technique. We found that according to the criteria used, the point estimate of the Lomax scale parameter using the modified profile likelihood function based on empirical covariances approximation have the best performance under type I and type II censoring data. Examples based on real data are given to illustrate the methods considered in this thesis.

Keywords: Modified maximum profile likelihood method, Lomax distribution, Barndorff-Nielsen's adjustment method, Censoring, Approximate ancillary statistic.

DEDICATION

This thesis is dedicated to my father and husband, who have provided me with inspiration and strength when I felt like giving up, and who continue to provide moral, spiritual, and financial support. To my children and friends who have given me encouragement and support to complete this thesis.

Thank you for compelling me to see this journey through to the end.

ACKNOWLEDGMENTS

I'd want to give my special thanks to everyone who assisted me in completing my thesis. First and foremost, a special thanks to Professor Ayman Suleiman Baklizi for his supervision and guidance throughout all of the stages that this work went through.

I would like also to thank my father and spouse for their encouragement and motivation while I worked through the challenging process of finishing this thesis.

Finally, many thanks to my friends and colleagues for their motivation and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTODUCTION

1.1 Fundamental concepts in lifetime data analysis

The term lifetime data analysis includes two main branches, survival analysis and reliability theory. It represents a set of statistical techniques for describing and quantifying time to event data. The term 'failure' is used in survival analysis to describe the presence of the event of interest, even if the event is a ‘success'. Furthermore, the period of time necessary for failure to occur is referred to as 'survival time' (Stevenson et al [38]). Survival analysis methods rely on the survival distribution, which can be specified in two ways: the survival function and the hazard function (Moore [31]). The survival function expresses the likelihood of survival up to a point t. Specifically

$$S(t) = P(T > t), \;\; 0 < t < \infty \tag{1}$$

At time 0, this function has a value of 1 and diminishes (or stays unchanged) through time.

The hazard function, which represents the failure rate at any given time, is frequently used to define the survival function. It's the probability of a subject failing in the next little period of time divided by the length of that period, given that he or she has survived up to time t (Moore [31]). Explicitly, this could be stated as

$$h(t) = lim_{\delta \to 0} \frac{P(t < T < t + \delta | T > t)}{\delta} \tag{2}$$

There are three types of survival time models: parametric, semi-parametric, and non-parametric. Non-negative distributions, such as exponential, Weibull, gamma, and lognormal distributions, are frequently required in parametric survival analysis models. The shape of the model's hazard function will be affected by the distribution we choose.

So, we must select the one that best matches the hazard (Stevenson [37]), or we can compare different parametric models and select the best one using a criterion such as AIC function (Moore [31]). As a result, the most difficult task in running parametric models is finding an acceptable distribution. When using data to estimate the hazard function, this method is known as semi-parametric. It makes no assumptions about the distribution of failure times, but does make assumptions about how covariates affect survival experience (Stevenson [37]). Usually, it is difficult to know which parametric family to use when modeling human or animal survival, and, in many cases, none of the existing families have enough adaptability to model the distribution's true shape. As a result, the Cox proportional hazard regression model (Cox [12]) is far more common than parametric regression, because the nonparametric estimation of the hazard function provides far more flexibility than most parametric techniques (Moore [31]).

It is important to note that the presence of censorship in its data is the most important feature that distinguishes survival analysis from other statistical analyses. Censorship occurs when only a portion of an individual's survival time is known (Klein and Moeschberger[24]).

1.2 Lifetime data and Censoring

The response variable in survival data is a non-negative discrete or continuous random variable that describes the time from a well-defined origin to a well-defined event. Censoring occurs when the starting or ending events are not precisely observed (Moore [31]). Right, left, and interval censoring are all examples of censoring (Klein and Moeschberger [24]).

Right censoring is the most frequent and easiest to deal with in the analysis, and it is classified into three types: Type I, Type II, and random. It is so named because the times of failure to the right (i.e., greater than t) are missing (Lawless [25]).

In Type I censoring, the censoring time t is pre-specified. We observe r failures during the t hours of testing (where r can be any number from 0 to n). The (exact) failure times are $t_1, t_2, \dots \dots . . t_r$, and there are $(n - r)$ units that passed the entire t-hour test. t is fixed ahead of time and r is random because we don't know how many failures will occur until the test is performed. Also, we assume that when there are failures, the precise times of failure are registered. (Lee and Wang [26]).

In Type II censoring, the number of failure times r is pre-specified in advance, and the study continues until the first r failure objects occur. For example, you could test 100 units and decide that you want at least half of them to fail. Then $r = 50$, but t is unknown until the 50^{th} failure occurs (Klein and Moeschberger [24]).

Each unit in random censoring has a possible censoring time C_i and a possible lifetime t_i, which are assumed to be statistically independent random variables. We look at $y_i = \min(C_i, t_i)$ which is the lowest of the censoring and life times, as well as an indicator variable called d_i , which indicates whether the observation ended due to death or censoring (Lawless [25]).

It is worth noting that Type I censoring is most commonly used in survival analysis, whereas Type II censoring is most commonly used in reliability studies.

1.3 The Lomax distribution

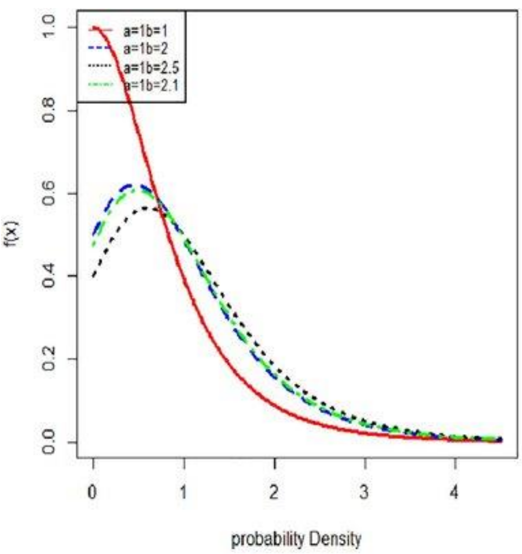
The Lomax distribution was developed to model business failure data (Lomax [29]). It is commonly referred to as the "Pareto type II" distribution (Arnold [2]). The Lomax model belongs to the declining failure rate family in the lifetime distribution context, see Chahkandi and Ganjali [8]. It has been proposed as a heavy tailed distribution by Bryson [7] to replace the Exponential, Weibull, and Gamma distributions. The Lomax distribution was designed to model business failure data, but it's also been used for life testing and reliability modeling. The Lomax distribution is critical for analyzing lifetime data sets in a variety of fields, including business, medicine, and engineering (Johnson et al. [23]). Corbellini et al. [9], Ghitany et al. [17], and Holland et al. [21] provide additional examples. Arnold [3] also includes applications and some properties of this distribution.

Suppose T is the Lomax random variable, then the pdf and cdf of the underlying Lomax lifetime distribution with two parameters β and θ are given respectively by

$$f(t,\theta,\beta) = \frac{\beta\theta}{(1+\beta t)^{\theta+1}} \quad , \quad t > 0, \beta > 0, \theta > 0 \tag{3}$$

$$F(t,\theta,\beta) = 1 - \frac{1}{(1+\beta t)^{\theta}} \quad , \quad t > 0, \beta > 0, \theta > 0 \tag{4}$$

where θ and β are the shape parameter and the scale parameter, respectively. The scale and shape of the distribution are controlled by these parameters, as shown in the diagram below.



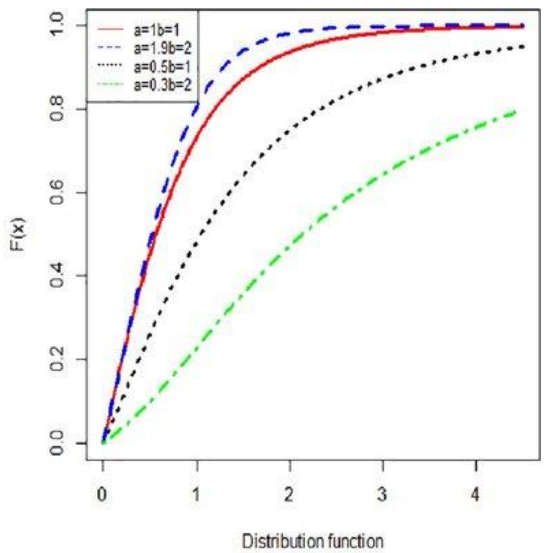


Figure 1.Graph of the pdf and cdf with a=1 and different values of b

In Figure 1 above, "a" represent the scale parameter and "b" represent the shape parameter. It is clear that the shape of the Lomax distribution is declining for any value of the shape parameter (b). In addition, as the value of the shape parameter "b" decrease, the shape of the distribution sharply decreases.

The Lomax distribution's survival function is as follows:

$$S(t) = \frac{1}{(1+\beta t)^{\theta}} \quad , \qquad t > 0, \beta > 0, \theta > 0 \tag{5}$$

The hazard function of the Lomax distribution is linked to the pdf and survival function by

$$h(t) = \frac{f(t)}{S(t)} = \frac{\beta\theta}{(1+\beta t)^{\theta+1}} \cdot (1+\beta t)^{\theta} = \frac{\beta\theta}{(1+\beta t)} \tag{6}$$

Figure 2 bellow illustrate the hazard function of the Lomax distribution.

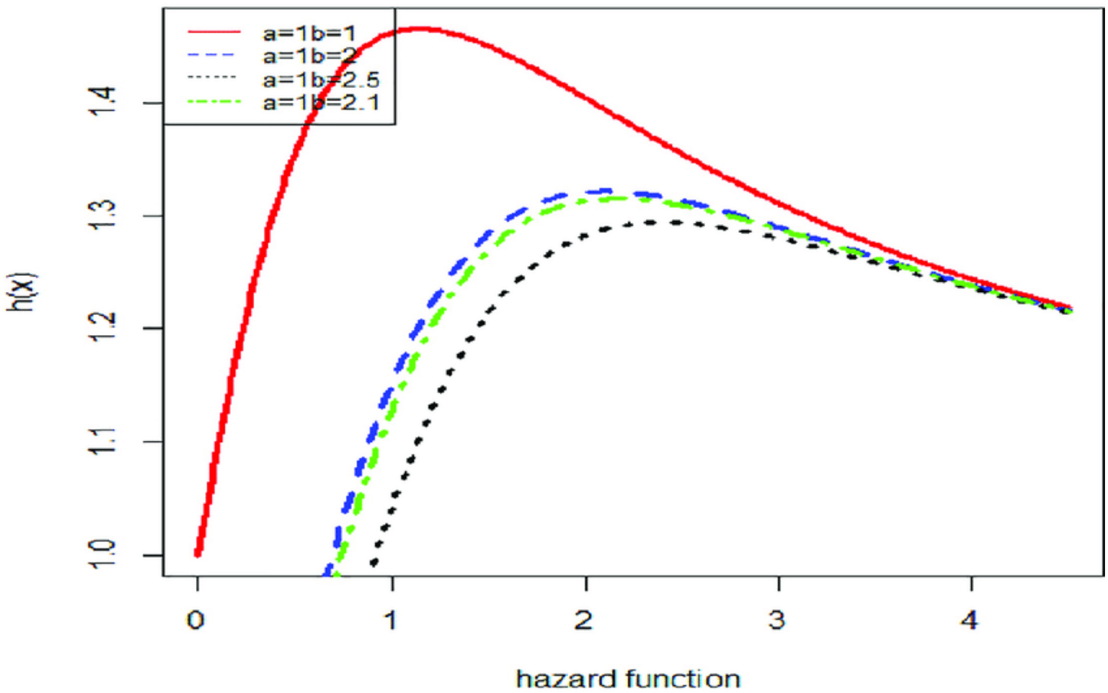


Figure 2.Graph of hazard function of the Lomax distribtion with a=1 and different values of b

The hazard function's shape is continually decreasing, as shown in Figure 2, demonstrating that the Lomax model is part of a family of models with decreasing failure rates (hazard rate).

1.4 Censoring and the likelihood functions

The likelihood function is central to a wide range of standard model-based statistical theory. It is known that the inferential practices derived from it have asymptotically ideal properties under very general regularity conditions (Lehmann and Casella [27]).

1.4.1 Complete Data

Suppose that lifetimes for individuals in some population following a specific distribution with probability function $f(t,\underline{\boldsymbol{\theta}})$,and that the lifetimes $t_1, \dots \dots, t_n$ for a random sample of n individuals are observed. Then the likelihood function according to Lawless [25] is given by

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(t_i, \underline{\boldsymbol{\theta}})$$

(7)

Which can be maximized to give an estimate $\hat{\underline{\boldsymbol{\theta}}}$.

1.4.2 Type I censoring

Assume that $t_j, (j = 1, \dots, n)$ is a random sample from a mathematical model with probability density function given by $f(t_j, \underline{\boldsymbol{\theta}})$, and survival function specified by $S(t_j, \underline{\boldsymbol{\theta}})$ where $\underline{\boldsymbol{\theta}} = (\theta_1, \dots \dots, \theta_k)^{`}$ .Let T is a censoring constant. Lawless [25] claim that, data can be denoted by n independent pairs of random variables in the form $(x_j, \delta_j), j = 1, \dots, n$, where

$$x_j = min(t_j, T)$$

and,

$$\delta_j = \begin{cases} 1 & , t_j < T \\ 0 & , t_j > T \end{cases}$$

The resulting likelihood function will be specified by

$$L(\theta, \beta) = \prod_{j=1}^{n} f(t_j; \theta, \beta)^{\delta_j} S(t_j; \theta, \beta)^{1-\delta_j}$$

(8)

1.4.3 Type II censoring

Let $t_{(j)}, (j = 1, \dots, r)$ be the smallest r order statistic from a sample of size n following a parametric distribution with pdf given by $f(t_{(j)}, \underline{\boldsymbol{\theta}})$ and survival function specified by $S(t_{(j)}, \underline{\boldsymbol{\theta}})$, where $\underline{\boldsymbol{\theta}} = (\theta_1, \dots \dots, \theta_k)^{`}$. Here, the number of failure time (say r) is static and prespecified and the time of study T is random. Therefore, Observations stopped after the r^{th} failure ($r < n$). According to Lawless [25], the likelihood function will be given by

$$L(\Theta, \beta) = \prod_{j=1}^{r} f(t_{(j)}; \Theta, \beta) \prod_{j=r+1}^{n} S(t_{(r)}; \Theta, \beta) \tag{9}$$

1.5 Literature review

1.5.1 Lomax distribution

The Lomax distribution is critical for analyzing data collected over a lifetime in business, computer science, medical and biological sciences, engineering, economics, income and wealth inequality, and Internet traffic and dependability analysis (see Johnson et al. [23]). Harris [19], for example, employs the Lomax distribution to examine income and wealth data. Furthermore, Corbellini et al. [9] used it to analyze company size and queuing statistics. According to Ghitany et al. [17], it can be used to create a lifetime distribution. It can also be used in biosciences and estimating the size distribution of computer data on servers, according to Holland et al [21].

1.5.2 Inferences from the Lomax model

In the literature, many authors have addressed Lomax model inferences from Bayesian, E-Bayesian, and maximum likelihood estimation perspectives. For instance, Okasha [32] computed estimates of the unidentified parameters in the Lomax distribution under type II censored data using E-Bayesian estimation, which is the expectation of bayes estimate, as well as estimating related survival time characteristics (hazard and reliability functions). The new technique (E-Bayesian) is compared to the equivalent Bayes and maximum likelihood procedures using Monte Carlo simulation.

The results show that the E-Bayes estimates (based on balanced squared error loss function (BSEL)) outperform the Bayes estimates when comparing the estimated risks of the estimates. Baklizi et al. [4] estimated parameters in this model with progressively censored data using likelihood and Bayesian inference. They considered point and interval estimation for the Lomax distribution's parameters. Numerical results show that point estimation appears similar between the MLE and the Bayesian estimators, while Bayesian intervals are considered better than the corresponding likelihood intervals (Wald intervals). Using the methods of Lindley [28] and Tierney and Kadane [38], Howlader and Hossain [22] investigated Bayesian survival estimation of the Pareto II distribution (Lomax distribution) based on failure censored data. Comparisons of these two methods, as well as its competitor, the maximum likelihood method, are done using Monte Carlo simulation. They conclude that the Lindley method should be used for a few samples and for big samples, any of the three approaches. Cramer and Schmiedt [13] used type-II censored competing risks data from Lomax distribution to calculate maximum likelihood estimates for the distribution parameters. The results show that the estimation that one-step censoring plans $(0, \dots \dots, 0, n-m, 0, \dots .0)$ are ideal in many situations. AL-Zahrani and Al-Sobhi [1] used Bayesian and maximum likelihood estimation to estimate the parameters according to general progressive censored data. The results reveal that in small samples, bayes estimators beat ML estimators, whereas in big samples, the estimators are nearly equal. Moreover, Abdul Wahab et al. [30] compared the maximum likelihood approach to the Bayes-approach with respect to the lifetime performance index using progressively Type-II censored competing risks data from Lomax distributions. The results show that the Bayes estimator performs slightly better than the maximum likelihood estimator. However, the change is slight, and it is debatable if the Bayes technique is genuinely advantageous in the circumstance at hand.

1.5.3 Modified profile likelihood function

Numerous modifications to the profile likelihood function have been proposed in the literature for a variety of distributions. For example, under type II data censoring, Barreto et al. [6] developed adjusted profile likelihood inference (point estimation, interval estimation, and hypothesis test) for the Birnbaum–Saunders distribution shape parameter. They considered the profile likelihood function modifications proposed by Barndorff-Nielsen [5], Cox and Reid [11], Fraser and Reid [15], Fraser et al. [16], and Severini [33]. Also, take into account bootstrap-based inference (hypothesis testing & Interval estimation). According to the results, the best performing point estimator was obtained by maximizing Barndorff-Nielsen's adjusted profile log-likelihood function. It usually had the smallest relative bias. Specifically, an approximation based on an ancillary statistic proposed by Fraser and Reid [15] and Fraser et al. [16]. The bootstrap test outperforms tests based on altered profile likelihood functions by a significant margin. The numerical findings show that the bootstrap-t confidence interval is the best method for estimating intervals. In addition, For the Weibull shape parameter, Ferrari et al. [14] give a number of possible adjusted profile likelihoods. They employ Barndorff-modified Nielsen's profile likelihood. An approximation based on population covariances, an approximation based on empirical covariances, an approximation based on an ancillary statistic, and an approximation based on orthogonal parameters. Point estimation and hypothesis testing are both taken into account. Estimation using the Cox and Reid likelihood function was significantly more precise, both without and with (types I and II) censoring, and inference based on Cox and Reid's adjusted profile likelihood outpaces not only conventional profile likelihood inference, as proposed by Yang and Xie [39], but also inference based on alternative adjusted profile likelihoods.

Overall, the available literature reviews present inference (point estimation, interval estimation, and hypothesis test) on the Lomax distribution from the perspectives of Bayesian, E-Bayesian, and maximum likelihood estimation. In addition, several modifications to the profile likelihood function for distributions other than the Lomax distribution have been proposed in the literature. As a result of the literature, we can conclude that there is a limit to inference on the Lomax distribution when using modifications to the profile likelihood function.

1.6 Problem statement

One of the most significant distributions in dependability theory is the Lomax distribution. It is a massive probability distribution utilized in business, economics, actuarial science, queueing theory, and Web traffic modeling.

There is a difficulty in conducting statistical inference for a model with a nuisance parameter, which is a common occurrence. Furthermore, the larger the nuisance parameter's dimension, the higher its potential impact on the results for the parameter of interest. Replacing the nuisance parameter with its partial maximum likelihood estimator is a simple way to eliminate the nuisance parameter's impact on inference. The likelihood function that result is then denoted to as the profile likelihood function. However, because the profile likelihood function does not approximate a genuine conditional or marginal likelihood function, the profile likelihood function is not a genuine likelihood function, which means that it is not a real likelihood function. Hence, it will not always give us accurate inference (Severini [35]). Therefore, adjustments to the profile likelihood function are needed.

In the literature, several variations to the profile likelihood function have been presented for a variety of distributions by approximating a marginal or conditional likelihood function; they're all trying to diminish the effect of the nuisance parameter on inferences about the parameter of interest. For example, Barreto et al. [6] use Barndorff-Nielsen's modified profile likelihood function and some approximation to it that approximates the marginal or conditional likelihood function for the parameter of interest, if one exists, to get adjusted profile maximum likelihood estimators for the Birnbaum-Saunders distribution shape parameter under Type II data censoring. In addition, using the Cox and Reid method of orthogonal parameters, they got a modified profile log-likelihood function that approximates the conditional density function of observation given the nuisance parameter. Ferrari et al. [14] present different adjustment profile likelihoods for the Weibull shape parameter with and without censoring (types I and II).

Therefore, due to the significance of the Lomax distribution in reliability problems and lifetime analysis, particularly in business failure, and the scarcity of studies on the statistical inference of Lomax scale parameter (interest parameter) in the existence of nuisance shape parameter based on adjusted profile likelihood function, this study is designed to address this gap in the literature.

1.7 Objectives and significant of the study

The thesis's goal is to improve inference for the scale parameter of the Lomax distribution in the presence of a nuisance shape parameter. The improvement is through the use of the Barndorff-Nielsen adjustment of the profile likelihood function. This adjustment is quite complicated and cannot be obtained exactly for the case of the Lomax distribution. Several approximations were derived and investigated. We apply these approximations to both complete and type I and type

II censored data. We obtained the maximum modified profile likelihood estimators. We investigated and compared the performance of the various approximations using their biases and mean squared errors using a simulation study.

1.7.1 Specified research goals

1. Find the maximum profile likelihood estimator (MPLE) for the scale parameter under different sampling schemes (complete sample & Type I and Type II censoring).
2. Find the maximum modified profile likelihood estimators (adjusted MPLE) for the scale parameter under different sampling schemes using different approximations to Brandorff-Nielsen’s modified profile likelihood function.
3. Using simulation, examine the performance and properties of the MPLE, and adjusted MPLE in terms of bias and MSE under various sampling strategies.
4. Use real data examples to illustrate the inference methods that we proposed in this thesis.

1.8 Overview

The following is an overview of how this thesis is organized. The profile likelihood function and its properties are covered in Chapter 2. This chapter also includes several approximations to Barndorff-Nielsen's adjustment. Additionally, adjustments for point estimation on the Lomax scale parameter are derived using various approximations to Brandorff Nielson's method with and without censoring (type I and type II) data. The numerical findings of a simulation research aiming to evaluate and compare the performance of estimators derived from profile likelihood functions and adjusted profile likelihood functions under various sample types are given in Chapter 3. In Chapter 4, numerical examples based on real-world data are provided. The thesis is concluded in Chapter 5 with a summary, recommendations, and suggestions for future research.

CHAPTER 2: PROFILE LIKELIHOOD FUNCTION FOR THE LOMAX DISTRIBUTION AND ITS MODIFICATIONS

2.1 Profile likelihood function

We consider a model parametrized by a parameter (θ, β), where β denotes the parameter of concern and θ is a nuisance parameter. Replacing θ with the restricted maximum likelihood estimator $\hat{\theta}_\beta$ is a simple approach of removing the effect of the nuisance parameter on inference. Let $L(\theta, \beta)$ be the likelihood function and let $l(\theta, \beta) = log\big(L(\theta, \beta)\big)$, where log is the natural logarithm, then $l_p(\beta) = l(\theta, \beta)|_{\theta=\hat{\theta}_\beta} = l\big(\hat{\theta}_\beta, \beta\big)$ is called the profile log-likelihood function and the maximum profile likelihood estimator of β , under this approach , is represented as $\hat{\beta}_p$. However, because $l_p(\beta)$ does not attempt to approximate a true conditional or marginal likelihood function, the profile likelihood function isn't a real likelihood function and thus lacks some of the favorable characteristics of a true likelihood function. This is because, by keeping the nuisance parameter at its point estimate, we are ignoring the uncertainty that comes with such estimation to some extent (Severini [35]).

2.2 Modified profile likelihood function

There are numerous adjustments to the profile likelihood function proposed in the literature. They're all meant to mitigate the influence of the nuisance parameter on inference about the parameter of interest. We will discuss some of them in the following subsections.

2.2.1 Barndorff-Nielsen's modified profile likelihood function

Barndorff-Nielsen [5] developed a modification that, if it exists, approximates the marginal or conditional likelihood function for the parameter of interest. He proposed a formula for calculating the approximate conditional density of the maximum likelihood method given an ancillary statistic a. He called this formula the p* equation. Several authors have utilized modified profile likelihood functions for inference including Yang and Xie [39], Barreto et al. [6], and Ferrari et al. [14]. The approach used in this thesis follows closely the approach of Ferrari et al. [14] for the Weibull shape parameters. The modified profile log-likelihood function of Barndorff-Nielsen is

$$l_{BN}(\beta) = l_p(\beta) - \log\left|\frac{\partial\hat{\theta}_\beta}{\partial\hat{\theta}}\right| - \frac{1}{2}\log\left|j_{\theta\theta}\left(\hat{\theta}_\beta,\beta\right)\right|, \tag{10}$$

where $j_{\theta\theta}\left(\hat{\theta}_\beta,\beta\right) = -\frac{\partial^2 l(\hat{\theta}_\beta,\beta)}{\partial\theta^2}$ and $\frac{\partial\,\hat{\theta}_\beta}{\partial\hat{\theta}}$ is a partial derivatives matrix of $\hat{\theta}_\beta$ with respect to $\hat{\theta}$. The most challenging part of computing the $l_{BN}(\beta)$ is in finding $\left|\frac{\partial\hat{\theta}_\beta}{\partial\hat{\theta}}\right|$. There is another equivalent modification for $l_{BN}(\beta)$ that avoid this term. It requires a sample space derivative of the log-likelihood function, as well as an ancillary statistic a such that $\left(\hat{\theta},\hat{\beta},a\right)$ is a minimal sufficient statistic, see [5].

The next three approximation approaches avoid the difficulties of evaluating the sample space derivatives emanating from this Barndorff-Nielsen's approach.

2.2.2 An approximation based on population covariances

Severini [33] presented the following approximation for Barndorff-Nielsen's modified profile likelihood function:

$$\bar{l}_{BN}(\beta) = l_p(\beta) + \frac{1}{2}\log\left|j_{\theta\theta}(\hat{\theta}_\beta,\beta)\right| - \log\left|\mathrm{I}_\theta\left(\hat{\theta}_\beta,\beta\,;\hat{\theta},\hat{\beta}\right)\right|, \tag{11}$$

where,

$$\mathrm{I}_{\theta}(\theta, \beta\,; \theta_0, \beta_0) = E_{(\theta_0,\beta_0)}\{l_{\theta}(\theta, \beta) l_{\theta}(\theta_0, \beta_0)^T\}$$

(12)

with $l_{\theta}(\theta, \beta) = \frac{\partial l(\theta,\beta)}{\partial \theta}$. Here, $\hat{\theta}_{\beta}$ is the restricted maximum likelihood estimator. The maximum likelihood estimators of θ and β, respectively, are $\hat{\theta}$ and $\hat{\beta}$. $I_{\theta}(\theta, \beta\,; \theta_0, \beta_0)$ is independent of the ancillary statistic "a" and $I_{\theta}(\theta, \beta\,; \theta_0, \beta_0)$ represents the covariance between $l_{\theta}(\theta, \beta)$ and $l_{\theta}(\theta_0, \beta_0)$ .The corresponding modified maximum profile likelihood estimator (MMPLE) is represented as $\hat{\hat{\beta}}_{BN}$.

2.2.3 An approximation based on empirical covariances

According to Severini [34], the empirical covariances approximation, presented below, is useful when calculating expected values of log likelihood derivative products is difficult. This approximation is as follows:

$$\check{l}_{BN}(\beta) = l_p(\beta) + \frac{1}{2}\log\left|j_{\theta\theta}(\hat{\theta}_{\beta}, \beta)\right| - \log\left|\check{I}_{\theta}(\hat{\theta}_{\beta}, \beta; \hat{\Theta}, \hat{\beta})\right|,$$

(13)

where,

$$\check{\mathrm{I}}_{\theta}\,(\hat{\theta}_{\beta}, \beta; \hat{\Theta}, \hat{\beta}) = \sum_{j=1}^{n} l_{\theta}^{(j)}(\hat{\theta}_{\beta}, \beta)\, l_{\theta}^{(j)}(\hat{\Theta}\,, \hat{\beta}\,)^T$$

(14)

Here, $l_{\theta}^{(j)}$ is the score function of the j^{th} observation, and the equivalent modified maximum profile likelihood estimator (MMPLE) under this approximation is represented as $\check{\hat{\beta}}_{BN}$.

2.2.4 An approximation based on an ancillary statistic

An approximation was proposed by Fraser and Reid [15] and Fraser et al. [16], which is given by

$$\tilde{l}_{BN}\ (\beta)=l_p(\beta)+\frac{1}{2}\log\left|j_{\theta\theta}(\hat{\theta}_\beta,\beta)\right|-\log\left|l_{\theta;y}(\hat{\theta}_\beta,\beta)\hat{V}_\theta\right|,$$

(15)

where,

$$l_{\theta;y}(\theta,\beta)=\frac{\partial l_\theta(\theta,\beta)}{\partial \mathcal{Y}^T}$$

(16)

Here, $\frac{\partial l(\theta,\beta)}{\partial\theta}$is the score function for, $\mathcal{Y}^T=(y_1,\dots\dots.,y_n)$ and

$$\hat{V}_\Theta=\left(-\frac{\partial F(y_1;\widehat{\Theta},\ \hat{\beta})/\partial\widehat{\Theta}}{f_1(y_1;\widehat{\Theta},\hat{\beta})},\dots\dots\dots\dots,-\frac{\partial F(y_n;\widehat{\Theta},\ \hat{\beta})/\partial\widehat{\Theta}}{f_n(y_n;\widehat{\Theta},\hat{\beta})}\right),$$

(17)

$f_j(y;\theta,\beta)$ and $F_j(y;\theta,\beta)$ are the probability density and cumulative distribution functions of y_j ,respectively, and where, $\hat{V}_\Theta$ is the approximate ancillary statistic. The corresponding modified maximum profile likelihood estimator (MMPLE) under this approximation ($\tilde{l}_{BN}$ (β)) is represented as $\hat{\tilde{\beta}}_{BN}$.

2.3 Modified profile likelihoods for the Lomax scale parameter

2.3.1 Complete Data

Let $y_1,\dots\dots.,y_n$ be distributed independently and uniformly (i.i.d.) variables created by Lomax. Using (7), the likelihood function for the $(\theta\ ,\beta)$ parameters is specified by

$$L(\theta,\beta,y_i)\ =\prod_{i=1}^{n}\beta\theta(1+\beta y_i)^{-(\theta+1)}\ =\beta^n\theta^n\prod_{i=1}^{n}(1+\beta y_i)^{-(\theta+1)}$$

(18)

Therefore, the log-likelihood function is defined as follows:

$$l(\theta,\beta) = n\log\beta + n\log\theta - (\theta+1)\sum_{i=1}^{n}\log(1+\beta y_i)$$

(19)

The log-likelihood function's first derivative with respect to θ is given by

$$\frac{\partial l(\theta,\beta)}{\partial\theta} = l_\theta(\theta,\beta) = \frac{n}{\theta} - \sum_{i=1}^{n}\log(1+\beta y_i)$$

(20)

The root of this equation in θ for a fixed value of β is

$$\hat{\theta}_\beta = \frac{n}{\sum_{i=1}^{n}\log(1+\beta y_i)}$$

(21)

This root represents the constrained MLE of θ for a given β. We get the profile log-likelihood function by substituting $\hat{\theta}_\beta$ in the log-likelihood equation.

$$l_p(\beta) = n\log\beta + n\log\hat{\theta}_\beta - (\hat{\theta}_\beta + 1)\sum_{i=1}^{n}\log(1+\beta y_i)$$

(22)

After substituting the value of $\widehat{\Theta}_\beta$, it is become

$$l_p(\beta) = n\log\beta + n\log\frac{n}{\sum_{i=1}^{n}\log(1+\beta y_i)} - \left(\frac{n}{\sum_{i=1}^{n}\log(1+\beta y_i)} + 1\right)\sum_{i=1}^{n}\log(1+\beta y_i)$$

(23)

The solution of the following equation yields the MLE of β, which is the maximum profile likelihood estimator of β.

$$\frac{\partial l_p(\beta)}{\partial\beta} = \frac{n}{\beta} - \hat{\theta}_\beta\sum_{i=1}^{n}\frac{y_i}{1+\beta y_i} - \sum_{i=1}^{n}\frac{y_i}{1+\beta y_i} = 0$$

(24)

Which is equivalent to the following equation,

$$\frac{\partial l_p(\beta)}{\partial \beta} = \frac{n}{\beta} - \left[\frac{n}{\sum_{i=1}^{n} \log(1+\beta y_i)}\right] \sum_{i=1}^{n} \frac{y_i}{1+\beta y_i} - \sum_{i=1}^{n} \frac{y_i}{1+\beta y_i} = 0$$

(25)

The MLE $\hat{\beta}$ can't be obtained analytically and we need to find it numerically by applying some iterative methods like the Newton-Raphson method or direct optimization techniques.

Calculating $j_{\theta\theta}(\hat{\theta}_\beta, \beta)$ from the observed Fisher information matrix $j(\theta, \beta) = -\left(\frac{\partial^2}{\partial \theta^2} l(\theta, \beta)\right)_{\theta=\hat{\theta}_\beta}$ which is obtained from the log-likelihood function for Lomax distribution evaluated at $(\hat{\theta}_\beta, \beta)$ we obtain

$$j_{\theta\theta}(\hat{\theta}_\beta, \beta) = -\left(-\frac{n}{(\hat{\theta}_\beta)^2}\right) = \frac{n}{\left(\frac{n}{\sum_{i=1}^{n} log(1+\beta y_i)}\right)^2} = \frac{\left(\sum_{i=1}^{n} log(1+\beta y_i)\right)^2}{n}$$

(26)

Now, we will consider some approximations to the modified profile likelihood for Lomax parameter β using Barndorff-Nielsen's Method that are described in section 3.

From (20), we obtain

$$l_\theta(\theta, \beta) = \frac{n}{\theta} - \sum_{i=1}^{n} \log(1+\beta y_i)$$

(27)

and

$$l_\theta(\theta_0, \beta_0) = \frac{n}{\theta_0} - \sum_{i=1}^{n} \log(1+\beta_0 y_i)$$

(28)

Then, using (12)

$$l_{\theta}(\theta,\beta)l_{\theta}(\theta_0,\beta_0)=\left(\frac{n}{\theta}-\sum_{j=1}^{n}\log\bigl(1+\beta y_j\bigr)\right)\left(\frac{n}{\theta_0}-\sum_{j=1}^{n}\log\bigl(1+\beta_0 y_j\bigr)\right)$$

$$=\frac{n^2}{\theta\theta_0}-\frac{n}{\theta}\sum_{i=1}^{n}\log(1+\beta_0 y_i)-\frac{n}{\theta_0}\sum_{i=1}^{n}\log(1+\beta y_i)+\sum_{i=1}^{n}\log(1+\beta y_i)\sum_{k=1}^{n}\log(1+\beta_0 y_k)$$

$$=\frac{n^2}{\theta\theta_0}-\frac{n}{\theta}\sum_{i=1}^{n}log(1+\beta_0 y_i)-\frac{n}{\theta_0}\sum_{i=1}^{n}log(1+\beta y_i)+\sum_{i=1}^{n}(\log(1+\beta y_i))(\log(1+\beta_0 y_i))+\sum_{i=1}^{n}\sum_{\substack{k=1\\k\neq i}}^{n}\bigl(log(1+\beta_0 y_i)\,(log(1+\beta y_k))\bigr)$$

(29)

Thus,

$$E_{(\theta_0,\beta_0)}\{l_{\theta}(\theta,\beta)l_{\theta}(\theta_0,\beta_0)\}=$$

$$=\frac{n^2}{\theta\theta_0}-\frac{n}{\theta}\sum_{i=1}^{n}E(\log(1+\beta_0 y_i))-\frac{n}{\theta_0}\sum_{i=1}^{n}E(\log(1+\beta y_i))$$
$$+\sum_{i=1}^{n}E(\log(1+\beta y_i))(\log(1+\beta_0 y_i))$$
$$+\sum_{i=1}^{n}\sum_{\substack{k=1\\k\neq i}}^{n}E(\log(1+\beta y_i))E(\log(1+\beta_0 y_k))$$

(30)

These expectations are very difficult to find in closed form. Thus, we can’t find the Barndorff-Nielsen modified profile likelihood function based on population covariance approximation (equation 11). Therefore, we make use of the empirical covariances approximation.

Using (14), it follows that

$$l_{\theta}^{(j)}\bigl(\beta,\hat{\theta}_{\beta}\bigr)=\frac{n}{\hat{\Theta}_{\beta}}-\log(1+\beta y_i)$$

(31)

$$l_{\theta}^{(j)}\bigl(\hat{\beta}\,,\hat{\theta}\bigr)=\frac{n}{\hat{\theta}}-\log(1+\hat{\beta} y_i)$$

(32)

Then,

$$\breve{I}_{\theta}\ (\beta, \hat{\theta}_{\beta}; \hat{\beta}, \hat{\theta}) = \sum_{i=1}^{n}[(\frac{n}{\hat{\theta}_{\beta}} - log(1+\beta y_i))\left(\frac{n}{\hat{\theta}} - log(1+\hat{\beta} y_i)\right)]$$

(33)

From (13,14,23 & 26), we obtain

$$\breve{l}_{BN}\ (\beta) = n\, log\, \beta + n\, log\, \hat{\theta}_{\beta} - (\hat{\theta}_{\beta}+1)\sum_{j=1}^{n} log(1+\beta y_j) + \frac{1}{2} log\left|\frac{n}{(\hat{\theta}_{\beta})^2}\right| - log\left|\sum_{j=1}^{n}[(\frac{n}{\hat{\theta}_{\beta}} - log(1+\beta y_j))\left(\frac{n}{\hat{\theta}} - log(1+\hat{\beta} y_j)\right)]\right|$$

(34)

where,

$$\hat{\theta}_{\beta} = \frac{n}{\sum_{i=1}^{n} log(1+\beta y_i)}$$

and where $\hat{\theta}$ and $\hat{\beta}$ are the maximum likelihood estimators of θ and β. The corresponding estimator is $\hat{\breve{\beta}}_{BN}$. The MLE $\hat{\breve{\beta}}_{BN}$ does not have a closed form expression and we need to find it numerically by applying some iterative methods to solve the likelihood equation and compute the estimate $\hat{\breve{\beta}}_{BN}$.

From (16)

$$\partial l_{\theta}(\hat{\theta}_{\beta}, \beta) = \frac{n}{\hat{\theta}_{\beta}} - \sum_{i=1}^{n} \log(1+\beta y_i)$$

(35)

Therefore,

$$l_{\theta;y}(\hat{\theta}_{\beta}, \beta) = \frac{\partial l_{\theta}(\hat{\theta}_{\beta}, \beta)}{\partial \mathcal{Y}^{T}} = -\frac{\beta}{1+\beta y_j} \quad , \quad j = 1,2 \ldots., n$$

(36)

From (17)

$$\frac{\partial F(y_i;\widehat{\Theta},\,\hat{\beta})}{\partial\widehat{\Theta}} = \frac{(1+\widehat{\beta}\ y_j)^{\widehat{\Theta}}\log(1+\hat{\beta}\,y_j)}{(1+\widehat{\beta}\,y_j\,)^{2\widehat{\Theta}}}$$

(37)

$$f_j\big(y;\widehat{\theta}\,,\hat{\beta}\big) = \frac{\hat{\beta}\hat{\theta}}{(1+\hat{\beta}y_j)^{\hat{\theta}+1}}$$

(38)

Therefore,

$$\hat{V}_\theta(j) = -\frac{(1+\hat{\beta}\ y_j)\log(1+\widehat{\beta}\ y_j)}{\hat{\beta}\widehat{\Theta}} \qquad ,\ \ j = 1,2,\ldots.n$$

(39)

Hence from (15)

$$\tilde{l}_{BN}(\beta)= n\log\beta + n\log\hat{\theta}_\beta - [\,\hat{\theta}_\beta+1]\sum_{j=1}^{n}\log\big(1+\beta y_j\big) + \frac{1}{2}log\left|\frac{\left(\sum_{j=1}^{n} log(1+\beta y_j)\right)^2}{n}\right| - log\left|\left(\sum_{j=1}^{n}\left(\frac{\beta}{1+\beta y_j}\cdot\frac{(1+\widehat{\beta}\ y_j)\,log(1+\widehat{\beta}\,y_j}{\widehat{\beta}\ \widehat{\Theta}}\right)\right)\right|$$

(40)

where,

$$\hat{\theta}_\beta = \frac{n}{\sum_{i=1}^{n} log(1+\beta y_i)}$$

The corresponding estimator is $\hat{\hat{\beta}}_{BN}$, which will be computed numerically.

2.3.2 Type II censored data

It is worth to note that this data is used mostly in reliability theory. Here, the number of failure time (Say r) is static and prespecified and the time of study T is random. As a result, following the r^{th} failure($r< n$), observations stopped. Let $y_{(1)},\ldots\ldots\ldots,y_{(r)}$ be the smallest order statistics from a Lomax distribution sample of size n. Using (9), the likelihood function is

$$L(\Theta,\beta)=\prod_{j=1}^{r}f(y_{(j)};\Theta,\beta)\prod_{j=r+1}^{n}S(y_{(r)};\Theta,\beta)=\left[S(y_{(r)};\Theta,\beta)\right]^{n-r}\prod_{j=1}^{r}f\big(y_{(j)};\Theta,\beta\big)$$

$$=\left[(1+\beta y_{(r)})^{-\Theta}\right]^{n-r}\prod_{j=1}^{r}\left[\beta\Theta\big(1+\beta y_{(j)}\big)^{-(\Theta+1)}\right]$$

$$=\left[\big(1+\beta y_{(r)}\big)^{-\Theta}\right]^{n-r}\cdot\beta^{r}\Theta^{r}\cdot\prod_{j=1}^{r}\big(1+\beta y_{(j)}\big)^{-(\Theta+1)}$$

(41)

The log-likelihood function for the (Θ,β) parameters is

$$l(\Theta,\beta)=(n-r)\,log\big(1+\beta y_{(r)}\big)^{-\Theta}+log(\beta^{r}\Theta^{r})+\sum_{j=1}^{r}[-(\Theta+1)\,log(1+\beta y_{(j)})]$$

$$=-\Theta(n-r)\,log\big(1+\beta y_{(r)}\big)+r\,log\,\beta+r\,log\,\Theta-(\Theta+1)\sum_{j=1}^{r}log(1+\beta y_{(j)})$$

$$=-\Theta(n-r)\,log\big(1+\beta y_{(r)}\big)+r\,log\,\beta+r\,log\,\Theta$$

$$-\Theta\sum_{j=1}^{r}log\big(1+\beta y_{(j)}\big)-\sum_{j=1}^{r}log(1+\beta y_{(j)})$$

(42)

Hence,

$$l(\Theta,\beta)=r\log\beta+r\log\Theta$$

$$-\sum\nolimits_{j=1}^{r}\log\big(1+\beta y_{(j)}\big)-\Theta\left[\sum_{j=1}^{r}\log(1+\beta y_{(j)})+(n-\ r)\log(1+\beta y_{(r)})\right]$$

(43)

For fixed β (parameter of interest), the constrained maximum likelihood estimator of θ can be derived as follow

$$\frac{\partial l}{\partial\Theta}=\frac{r}{\Theta}-\left[\sum_{j=1}^{r}\log\big(1+\beta y_{(j)}\big)+(n-r)\log\big(1+\beta y_{(r)}\big)\right]=0$$

$$\frac{r}{\Theta} = \sum_{j=1}^{r} \log\left(1 + \beta y_{(j)}\right) + (n - r)\log(1 + \beta y_{(r)})$$

$$\frac{\Theta}{r} = \frac{1}{\sum_{j=1}^{r} log\left(1 + \beta y_{(j)}\right) + (n - r)\, log(1 + \beta y_{(r)})}$$

Hence, the constrained maximum likelihood estimator of θ is

$$\widehat{\Theta}_{\beta} = \frac{r}{\sum_{j=1}^{r} \log\left(1 + \beta y_{(j)}\right) + (n - r)\log(1 + \beta y_{(r)})} \quad (44)$$

This root represents the restricted MLE of θ for a given β. This estimate reduces to the one given previously in Uncensored data when there is no censoring$(r = n)$. We get the profile log-likelihood function by substituting $\hat{\theta}_{\beta}$ in the log-likelihood equation.

$$\begin{aligned} l_p(\beta) = l\left(\widehat{\Theta}_{\beta}, \beta\right) & \\ &= r\, log\, \beta + r\, log\left(\widehat{\Theta}_{\beta}\right) \\ &- \sum_{j=1}^{r} log(1 + \beta y_{(j)}) - \left(\widehat{\Theta}_{\beta}\right) \\ &\cdot \left[\sum_{j=1}^{r} log(1 + \beta y_{(j)}) + (n - r)\, log(1 + \beta y_{(r)})\right] \end{aligned} \quad (45)$$

After substituting the value of $\widehat{\Theta}_{\beta}$, it is become

$$l_p(\beta) = l(\hat{\theta}_\beta, \beta)$$

$$= r \log \beta + r \log \left(\frac{r}{\sum_{j=1}^{r} \log(1+\beta y_{(j)}) + (n-r)\log(1+\beta y_{(r)})} \right)$$

$$- \sum_{j=1}^{r} \log(1+\beta y_{(j)})$$

$$- \left(\frac{r}{\sum_{j=1}^{r} \log(1+\beta y_{(j)}) + (n-r)\log(1+\beta y_{(r)})} \right)$$

$$\cdot \left[\sum_{j=1}^{r} \log(1+\beta y_{(j)}) + (n-\ r)\log(1+\beta y_{(r)}) \right]$$

(46)

The MLE of β is the root of the following equation, which is the maximum profile likelihood estimator of

$$\frac{\partial l_p(\beta)}{\partial \beta} = \frac{r}{\beta} - \sum_{j=1}^{r} \frac{y_{(j)}}{1+\beta y_{(j)}} - \hat{\theta}_\beta \sum_{j=1}^{r} \frac{y_{(j)}}{1+\beta y_{(j)}} - \frac{\hat{\theta}_\beta (n-r) y_{(r)}}{1+\beta y_{(r)}} = 0$$

(47)

Which is equivalent to the following equation,

$$\frac{\partial l_p(\beta)}{\partial \beta} = \frac{r}{\beta} - \sum_{j=1}^{r} \frac{y_{(j)}}{1+\beta y_{(j)}}$$

$$- \left(\frac{r}{\sum_{j=1}^{r} \log(1+\beta y_{(j)}) + (n-r)\log(1+\beta y_{(r)})} \right) \sum_{j=1}^{r} \frac{y_{(j)}}{1+\beta y_{(j)}}$$

$$- \frac{\left(\frac{r}{\sum_{j=1}^{r} \log(1+\beta y_{(j)}) + (n-r)\log(1+\beta y_{(r)})} \right)(n-r) y_{(r)}}{1+\beta y_{(r)}} = 0$$

(48)

The MLE $\hat{\beta}$ can’t be obtained analytically and we need to find it numerically by applying some iterative methods like the Newton-Raphson method or direct optimization techniques.

Calculating $j_{\theta\theta}(\hat{\theta}_\beta, \beta)$ from the observed Fisher information matrix $j(\theta, \beta) = -\left(\frac{\partial^2}{\partial\theta^2} l(\theta, \beta)\right)_{\theta=\hat{\theta}_\beta}$ which is obtained from the log-likelihood function for Lomax distribution evaluated at $(\hat{\theta}_\beta, \beta)$ we obtain

$$j_{\theta\theta}(\hat{\theta}_\beta, \beta) = \frac{\left(\sum_{j=1}^{r} \log(1 + \beta y_{(j)}) + (n-r)\log(1 + \beta y_{(r)})\right)^2}{r} \tag{49}$$

Now, we will consider some approximations to the modified profile likelihood for Lomax parameter β using Barndorff-Nielsen's Method that are described in section 3.

Equation (11) is not possible to derive in type II censoring because y_j are order statistics (Not iid). Therefore, we make use of the empirical covariances.

Using (14), it follows that

$$l_{\theta}^{(j)}(\beta, \hat{\theta}_\beta) = \frac{r}{\hat{\theta}_\beta} - log(1 + \beta y_{(j)}) - (n-r)\, log(1 + \beta y_{(r)}) \tag{50}$$

$$l_{\theta}^{(j)}(\hat{\beta}, \hat{\theta}) = \frac{r}{\hat{\theta}} - log(1 + \hat{\beta} y_{(j)}) - (n-r)\, log(1 + \hat{\beta} y_{(r)}) \tag{51}$$

Then,

$$\check{I}_{\theta}(\beta, \hat{\theta}_\beta; \hat{\beta}, \hat{\theta}) = \sum_{j=1}^{r}\left[\left(\frac{r}{\hat{\theta}_\beta} - \log(1 + \beta y_{(j)}) - (n-r)\log(1 + \beta y_{(r)})\right)\left(\frac{r}{\hat{\theta}} - \log(1 + \hat{\beta} y_{(j)}) - (n-r)\log(1 + \hat{\beta} y_{(r)})\right)\right] \tag{52}$$

From (13,14,46&49), we obtain

$$\breve{l}_{BN}(\beta) = r\log\beta$$

$$+ r\log[\widehat{\Theta}_\beta] - \sum_{j=1}^{r}\log(1+\beta y_{(j)}) - (\widehat{\Theta}_\beta)$$

$$\cdot\left[\sum_{j=1}^{r}\log(1+\beta y_{(j)}) + (n-\ r)\log\bigl(1+\beta y_{(r)}\bigr)\right]$$

$$+\frac{1}{2}\log\left[\frac{\left(\sum_{j=1}^{r}\log\bigl(1+\beta y_{(j)}\bigr)+(n-r)\log(1+\beta y_{(r)})\right)^2}{r}\right]$$

$$-\log\left[\sum_{j=1}^{r}\left[\left(\frac{r}{\widehat{\Theta}_\beta} - \log(1+\beta y_{(j)}) - (n-r)\log(1+\beta y_{(r)})\right)\left(\frac{r}{\hat{\widehat{\Theta}}}\right.\right.\right.$$

$$\left.\left.\left. - \log\bigl(1+\hat{\beta} y_{(j)}\bigr) - (n-r)\log(1+\hat{\beta} y_{(r)})\right)\right]\right] \tag{53}$$

where,

$$\widehat{\Theta}_\beta = \frac{r}{\sum_{j=1}^{r}\log\bigl(1+\beta y_{(j)}\bigr)+(n-r)\log(1+\beta y_{(r)})}$$

Here, $\hat{\theta}$ and $\hat{\beta}$ are the maximum likelihood estimators of θ and β under type II censoring. The corresponding estimator is $\hat{\hat{\beta}}_{BN}$. There is no closed form expression for the MLE $\hat{\hat{\beta}}_{BN}$ and we need to find it numerically by applying some iterative methods such as optimization techniques to solve the likelihood equation and compute the estimate $\hat{\hat{\beta}}_{BN}$.

We also obtain, an approximation based on ancillary statistics. Using (16) it follows that

$$\partial l_\theta\bigl(\hat{\theta}_\beta,\beta\bigr) = \frac{r}{\hat{\theta}_\beta} - \left(\sum_{j=1}^{r}\log[1+\beta y_{(j)}] + (n-r)\log[1+\beta y_{(r)}]\right)$$

(54)

Thus,

$$l_{\theta;y}(\hat{\theta}_\beta,\beta) = \frac{\partial l_\theta(\hat{\theta}_\beta,\beta)}{\partial \mathcal{Y}^T} = \begin{cases} -\left[\dfrac{\beta}{1+\beta y_{(j)}}\right] & ,j = 1,2 \dots ., r-1 \\ -\left[\dfrac{\beta}{1+\beta y_{(r)}} + \dfrac{(n-r)\beta}{1+\beta y_{(r)}}\right] & ,j = r \end{cases}$$

(55)

Using (17),

$$\frac{\partial F(y_i;\widehat{\Theta},\hat{\beta})}{\partial\widehat{\Theta}} = \frac{(1+\widehat{\beta}\ y_{(j)})^{\widehat{\Theta}}\, log(1+\hat{\beta}\, y_{(j)})}{(1+\widehat{\beta}\, y_{(j)}\,)^{2\widehat{\Theta}}} \qquad ,j = 1,2 \dots ., r$$

(56)

Now,

$$\hat{V}_\theta(j) = -\frac{\left(1+\widehat{\beta}\ y_{(j)}\right)^{\widehat{\Theta}} \log(1+\hat{\beta}\, y_{(j)})}{\left(1+\widehat{\beta}\, y_{(j)}\,\right)^{2\widehat{\Theta}}} \cdot \frac{\left(1+\hat{\beta}\, y_{(j)}\right)^{\hat{\theta}+1}}{\widehat{\beta}\,\hat{\theta}}$$

$$= \frac{\left(1+\widehat{\beta}\ y_{(j)}\right)^{2\hat{\theta}+1} \log(1+\hat{\beta}\, y_{(j)})}{\left(1+\widehat{\beta}\, y_{(j)}\,\right)^{2\widehat{\Theta}}\hat{\beta}\hat{\theta}}$$

$$= -\frac{(1+\hat{\beta}\, y_{(j)}) \log(1+\widehat{\beta}\ y_{(j)})}{\hat{\beta}\widehat{\Theta}} \qquad , \ j = 1,2, \dots . r-1$$

$$= -\frac{(1+\hat{\beta}\, y_{(r)}) \log(1+\widehat{\beta}\ y_{(r)})}{\hat{\beta}\widehat{\Theta}} \qquad , \ j = r$$

(57)

From (15),

$$
\begin{aligned}
\tilde{l}_{BN}(\beta) = r\log\beta + r\log\left(\widehat{\Theta}_{\beta}\right) \\
-\sum_{j=1}^{r}\log(1+\beta y_{(j)}) - \left(\widehat{\Theta}_{\beta}\right) \\
\cdot\left[\sum_{j=1}^{r}\log(1+\beta y_{(j)}) + (n-\ r)\log\left(1+\beta y_{(r)}\right)\right] \\
+\frac{1}{2}\log\left[\frac{\left(\sum_{j=1}^{r}\log\left(1+\beta y_{(j)}\right)+(n-r)\log(1+\beta y_{(r)})\right)^{2}}{r}\right] \\
-\log\left[\left(\sum_{j=1}^{r-1}\left(\left(\frac{\beta}{1+\beta y_{(j)}}\right)\left(\frac{(1+\hat{\beta}y_{(j)})\log(1+\hat{\beta}y_{(j)})}{\hat{\beta}\widehat{\Theta}}\right)\right)\right)\right. \\
\left.+\left(\left(\frac{\beta}{1+\beta y_{(r)}}+\frac{(n-r)\beta}{1+\beta y_{(r)}}\right)\cdot\frac{(1+\hat{\beta}y_{(r)})\log(1+\hat{\beta}y_{(r)})}{\hat{\beta}\widehat{\Theta}}\right)\right]
\end{aligned} \tag{58}
$$

where,

$$
\widehat{\Theta}_{\beta} = \frac{r}{\sum_{j=1}^{r}\log\left(1+\beta y_{(j)}\right)+(n-r)\log(1+\beta y_{(r)})}
$$

Here, $\hat{\theta}$ and $\hat{\beta}$ are the maximum likelihood estimators of θ and β under type II censoring. The corresponding estimator is $\hat{\hat{\beta}}_{BN}$. The MLE $\hat{\hat{\beta}}_{BN}$ does not have a closed form expression, so we must find it numerically by solving the likelihood equation and computing the estimate $\hat{\hat{\beta}}_{BN}$ using iterative methods such as optimization techniques.

3.2.3 Type I censored data

We'll now consider a scenario in which type I censoring is present, which is commonly utilized in survival analysis. The number of failure times (Say r) is random (unlike type II censoring, where r is fixed and predetermined), whereas the time of study

(t) is fixed and predetermined (unlike type II censoring, T is random). Let $x_1, \dots \dots, x_n$ be Lomax random variables that are independently and identically distributed (i.i.d.). Using (8), we have

$$L(\theta,\beta) = \prod_{j=1}^{n} f(x_j;\theta,\beta)^{\delta_j} S(x_j;\theta,\beta)^{1-\delta_j} = \left(\prod_{j=1}^{r} f(y_j;\theta,\beta)\right)\left(S(T;\theta,\beta)\right)^{n-r}$$

(59)

where, $y_j = x_{(j)}$ is the j^{th} order statistics, and $r = \sum_{j=1}^{n} \delta_j$ is the number of complete observations in the sample.

Now, the likelihood function for the $L(\Theta,\beta)$ is

$$L(\theta,\beta) = \prod_{j=1}^{n} f(x_j,\theta,\beta)^{\delta_j} S(x_j)^{1-\delta_j} = \left(\prod_{j=1}^{r} f(y_j,\theta,\beta)\right)\left(S(T)\right)^{n-r}$$

$$= \left[(1+\beta\, T)^{-\Theta}\right]^{n-r} \prod_{j=1}^{r} \left[\beta\Theta\left(1+\beta y_j\right)^{-(\Theta+1)}\right]$$

$$= \left[(1+\beta T)^{-\Theta}\right]^{n-r} \cdot \beta^r \Theta^r \cdot \prod_{j=1}^{r}\left(1+\beta y_j\right)^{-(\Theta+1)}$$

(60)

Then, the log likelihood function for is

$$l(\Theta,\beta) = (n-r)\, log(1+\beta\, T)^{-\Theta} + log(\beta^r \Theta^r) + \sum_{j=1}^{r}[-(\Theta+1)\, log(1+\beta y_j)]$$

$$= -\Theta(n-r)\, log(1+\beta\, T) + r\, log\, \beta + r\, log\, \Theta - (\Theta+1)\sum_{j=1}^{r} log(1+\beta y_j)$$

$$= -\Theta(n-r)\, log(1+\beta\, T) + r\, log\, \beta + r\, log\, \Theta$$

$$- \Theta \sum_{j=1}^{r} log\left(1+\beta y_j\right) - \sum_{j=1}^{r} log(1+\beta y_j)$$

(61)

After simplifying,

$$l(\Theta,\beta) = r\,log\,\beta + r\,log\,\Theta$$

$$-\sum_{j=1}^{r} log\big(1+\beta y_j\big) - \Theta\left[\sum_{j=1}^{r} log(1+\beta y_j) + (n-\ r)\,log(1+\beta\,T)\right]$$

(62)

For fixed (parameter of interest), the constrained maximum likelihood estimator of θ can be stated as

$$\frac{\partial l}{\partial \Theta} = \frac{r}{\Theta} - \left[\sum_{j=1}^{r} \log\big(1+\beta y_j\big) + (n-r)\log(1+\beta\,T)\right] = 0$$

$$\frac{r}{\Theta} = \sum_{j=1}^{r} \log\big(1+\beta y_j\big) + (n-r)\log(1+\beta\,T)$$

$$\frac{\Theta}{r} = \frac{1}{\sum_{j=1}^{r} log\big(1+\beta y_j\big) + (n-r)\,log(1+\beta\,T)}$$

Hence,

$$\widehat{\Theta}_\beta = \frac{r}{\sum_{j=1}^{r} \log\big(1+\beta y_j\big) + (n-r)\log(1+\beta\,T)}$$

(63)

This estimator corresponds to the one given previously in uncensored data when there is no censoring $(r = n)$. As a result, we get

$$l_p(\beta) = l\big(\widehat{\Theta}_\beta, \beta\big)$$

$$= r\log\beta + r\log\big(\widehat{\Theta}_\beta\big)$$

$$-\sum_{j=1}^{r} \log(1+\beta y_j) - \big(\widehat{\Theta}_\beta\big)\cdot\left[\sum_{j=1}^{r} \log(1+\beta y_j) + (n-\ r)\log(1+\beta\,T)\right]$$

(64)

After substituting the value of $\widehat{\Theta}_\beta$, it is become

$$l_p(\beta) = l(\widehat{\Theta}_\beta, \beta)$$

$$= r \log \beta + r \log \left(\frac{r}{\sum_{j=1}^{r} \log(1+\beta y_j) + (n-r)\log(1+\beta\, T)} \right)$$

$$- \sum_{j=1}^{r} \log(1+\beta y_j) - \left(\frac{r}{\sum_{j=1}^{r} \log(1+\beta y_j) + (n-r)\ln(1+\beta\, T)} \right)$$

$$\cdot \left[\sum_{j=1}^{r} \log(1+\beta y_j + (n-\ r)\log(1+\beta\, T) \right]$$

(65)

The MLE of β is the solution to the following problem equation

$$\frac{\partial l_p(\beta)}{\partial \beta} = \frac{r}{\beta} - \sum_{j=1}^{r} \frac{y_j}{1+\beta y_j} - \hat{\theta}_\beta \sum_{j=1}^{r} \frac{y_j}{1+\beta y_j} - \frac{\hat{\theta}_\beta (n-r)T}{1+\beta T} = 0$$

(66)

Which is equivalent to the following equation,

$$\frac{\partial l_p(\beta)}{\partial \beta} = \frac{r}{\beta} - \sum_{j=1}^{r} \frac{y_j}{1+\beta y_j}$$

$$- \left(\frac{r}{\sum_{j=1}^{r} log(1+\beta y_j) + (n-r)\, log(1+\beta T)}\right) \sum_{j=1}^{r} \frac{y_j}{1+\beta y_j}$$

$$- \frac{\left(\frac{r}{\sum_{j=1}^{r} log(1+\beta y_j) + (n-r)\, log(1+\beta T)} \right)(n-r)T}{1+\beta T} = 0$$

(67)

The MLE $\hat{\beta}$ can't be obtained analytically and we need to find it numerically by applying some iterative methods like the Newton-Raphson method or direct optimization techniques.

Now,

$$j_{\Theta\Theta}(\widehat{\Theta}_\beta, \beta) = \frac{\left(\sum_{j=1}^{r} \log(1+\beta y_j + (n-r)\log(1+\beta\, T)\right)^2}{r}$$

(68)

As with uncensored data and type II censoring**,** an approximation to Barndorff-Nielsen's modified profile log-likelihoods based on population covariance ($\bar{l}_{BN}(\beta)$) is not possible to derive since $y_j = x_{(j)}$ is the j^{th} order statistics (Not iid). However, we obtained the following two approximations to the modified profile log-likelihoods by Brandorff-Nielsen's method. Using (14), it follows that

$$l_{\theta}^{(j)}(\beta, \hat{\theta}_{\beta}) = \frac{r}{\hat{\theta}_{\beta}} - \log(1 + \beta y_j) - (n - r)\log(1 + \beta\, T)$$

(69)

$$l_{\theta}^{(j)}(\hat{\beta}\,, \hat{\theta}) = \frac{r}{\hat{\theta}} - \log(1 + \hat{\beta} y_j) - (n - r)\log(1 + \hat{\beta}\, T)$$

(70)

Then,

$$\breve{I}_{\theta}(\beta, \hat{\theta}_{\beta}; \hat{\beta}, \hat{\theta}) = \sum_{j=1}^{r}\left[\left(\frac{r}{\hat{\theta}_{\beta}} - log(1 + \beta y_j) - (n - r)\, log(1 + \beta\, T)\right)\left(\frac{r}{\hat{\theta}} - log(1 + \hat{\beta} y_j) - (n - r)\, log(1 + \hat{\beta}\, T)\right)\right]$$

(71)

From (13,14,65 & 68), we obtain

$$\breve{l}_{BN}(\beta) = r\ln\beta + r\log[\widehat{\Theta}_\beta] - \sum_{j=1}^{r}\log(1+\beta y_j) - (\widehat{\Theta}_\beta)$$

$$\cdot\left[\sum_{j=1}^{r}\log(1+\beta y_j) + (n-\ r)\log(1+\beta\, T)\right]$$

$$+\frac{1}{2}\log\left[\frac{\left(\sum_{j=1}^{r}\log\left(1+\beta y_j\right) + (n-r)\log(1+\beta\, T\right)^2}{r}\right]$$

$$-\log\left[\sum_{j=1}^{r}\left[\left(\frac{r}{\widehat{\Theta}_\beta} - \log(1+\beta y_j) - (n-r)\log(1+\beta\, T)\right)\left(\frac{r}{\widehat{\Theta}}\right.\right.\right.$$

$$\left.\left.\left.-\log\left(1+\hat{\beta}y_j\right) - (n-r)\log(1+\hat{\beta}\, T)\right)\right]\right]$$

(72)

where ,

$$\widehat{\Theta}_\beta = \frac{r}{\sum_{j=1}^{r} log\left(1+\beta y_j\right) + (n-r)\, log(1+\beta\, T)}$$

Here, $\hat{\theta}$ and $\hat{\beta}$ are the maximum likelihood estimators of θ and β under type I censoring. The corresponding estimator is $\hat{\hat{\beta}}_{BN}$. There is no closed form expression for the MLE $\hat{\hat{\beta}}_{BN}$ and we need to find it numerically by applying some iterative methods to solve the likelihood equation and compute the estimate $\hat{\hat{\beta}}_{BN}$.

We also obtain, an approximation based on ancillary statistics. Using (16) it follows that

$$\partial l_\theta\left(\hat{\theta}_\beta,\beta\right) = \frac{r}{\hat{\theta}_\beta} - \left(\sum_{j=1}^{r}\log(1+\beta y_j) + (n-r)\log(1+\beta T)\right)$$

(73)

Thus,

$$l_{\theta;y}\left(\hat{\theta}_{\beta},\beta\right)=\frac{\partial l_{\theta}(\hat{\theta}_{\beta},\beta)}{\partial\mathcal{Y}^{T}}=-\left[\frac{\beta}{1+\beta y_j}\right]\quad ,j=1,2,\dots,r$$

(74)

And, using (17)

$$\frac{\partial F(y_j;\widehat{\Theta},\,\hat{\beta})}{\partial\widehat{\Theta}}=\frac{(1+\widehat{\beta}\;\,y_j)^{\widehat{\Theta}}\log(1+\hat{\beta}\,y_j)}{(1+\widehat{\beta}\,y_j\,)^{2\widehat{\Theta}}}\qquad ,j=1,2\dots,r$$

(75)

Now,

$$\hat{V}_{\theta}(j)=-\frac{\left(1+\widehat{\beta}\;\,y_j\right)^{\widehat{\Theta}}\log(1+\hat{\beta}\,y_j)}{\left(1+\widehat{\beta}\,y_j\,\right)^{2\widehat{\Theta}}}\cdot\frac{\left(1+\hat{\beta}\,y_j\right)^{\hat{\theta}+1}}{\widehat{\beta}\,\hat{\theta}}=\frac{\left(1+\widehat{\beta}\;\,y_j\right)^{2\hat{\theta}+1}\log(1+\hat{\beta}\,y_j)}{\left(1+\widehat{\beta}\,y_j\,\right)^{2\widehat{\Theta}}\hat{\beta}\hat{\theta}}$$

$$=-\frac{(1+\hat{\beta}\,y_j)\log(1+\widehat{\beta}\;\,y_j)}{\hat{\beta}\widehat{\Theta}}\qquad ,\ \ j=1,2,\dots.r$$

(76)

From (15),

$$\tilde{l}_{BN}(\beta)=r\log\beta+r\log\left(\widehat{\Theta}_{\beta}\right)$$

$$-\sum_{j=1}^{r}\log(1+\beta y_j)-\left(\widehat{\Theta}_{\beta}\right)\cdot\left[\sum_{j=1}^{r}\log(1+\beta y_j+(n-\ r)\log(1+\beta T)\right]$$

$$+\frac{1}{2}\log\left[\frac{\left(\sum_{j=1}^{r}\log\left(1+\beta y_j\right)+(n-r)\log(1+\beta T)\right)^2}{r}\right]$$

$$-\log\left[\sum_{j=1}^{r}\left(\frac{\beta}{1+\beta y_j}\cdot\frac{(1+\hat{\beta}\,y_j)\log(1+\widehat{\beta}\;\,y_j)}{\hat{\beta}\widehat{\Theta}}\right)\right]$$

(77)

where,

$$\widehat{\Theta}_{\beta}=\frac{r}{\sum_{j=1}^{r}log\left(1+\beta y_j\right)+(n-r)\,log(1+\beta\,T)}$$

Here, $\hat{\theta}$ and $\hat{\beta}$ are the maximum likelihood estimators of θ and β under type I censoring.

The corresponding estimator is $\hat{\hat{\beta}}_{BN}$. The MLE $\hat{\hat{\beta}}_{BN}$ does not have a closed form expression, so we must find it numerically by solving the likelihood equation and computing the estimate $\hat{\hat{\beta}}_{BN}$ using iterative methods such as optimization techniques.

CHAPTER 3: SIMULATION STUDY

3.1 Simulations using Monte Carlo methods

Monte Carlo simulation using R programming on point estimation was performed to compare the performance of estimators of the Lomax scale parameter (the parameter of interest) discussed in chapter 2 according to the standard profile likelihood function and modified profile likelihood functions. The samples from Lomax distribution (complete, types I and II censoring) are generated using the inverse transformation technique (see Appendix A). The "optim" function in R is used to find numerically the maximum likelihood estimators of the scale parameter that maximizes the profile likelihood and modified profile likelihood functions. The maximum profile likelihood function using complete data was used to find the first starting point for the maximum profile likelihood function under types I and II censoring, whereas the initial starting point for the profile likelihood function using complete data is beta=1. The profile likelihood estimator found from the associated data is the initial starting point for the maximum modified profile likelihood functions under complete, types I, and II censoring.

Different sample sizes are considered under both complete data (no censoring) and (types I and II) censoring. The sample sizes considered for complete data are n=25,50,75, and 100. For type I censoring, we considered sample sizes n=50,75, and 100 with censoring proportions 40% and 20% and for type II censoring we considered the same sample sizes as type I censoring (n=50,75, and 100), but with failure rates of 60% and 80%. The true value of the parameter of interest is static at $1(\beta = 1)$ and the true values the nuisance parameter are $\theta = 0.8, 1.0, 1.2$. The simulation results are based on 5000 iterations. Bias and mean square errors (MSEs) are presented for all the following point estimators: $\hat{\beta}_p$, $\hat{\bar{\beta}}_{BN}$ and $\hat{\bar{\bar{\beta}}}_{BN}$ under both no and (types I and II) censoring. $\hat{\beta}_p$ denotes the profile likelihood estimator. $\hat{\bar{\beta}}_{BN}$ and $\hat{\bar{\bar{\beta}}}_{BN}$ are the modified profile likelihood estimators derived from Barndorff-Nielsen's modified profile likelihood function based on an empirical covariances and an ancillary statistic approximation, respectively.

3.2 Results and comparisons

3.2.1 Complete data

Tables 1 and 2 show the simulation results on point estimation for the Lomax scale parameter (parameter of interest) with no censoring (complete data). We reported the bias on Table 1 and mean squared errors (MSEs) on Table 2 of the following point estimators: $\hat{\beta}_p$, $\hat{\bar{\beta}}_{BN}$ and $\hat{\bar{\bar{\beta}}}_{BN}$. The results show that the estimator with the least bias is the standard profile likelihood estimator ($\hat{\beta}_p$) even it is very close to the modified profile likelihood estimator based on empirical covariance ($\hat{\bar{\beta}}_{BN}$). This figure is true for all the sample sizes and the true parameters values that we considered$(\theta,\beta) = (0.8.0,1.0),(1.0,1.0),(1.2,1.0)$.For instance, in Table 1, with $\theta = 1$ and $n = 50$, the biases are $0.02639078(\hat{\beta}_p)$, $0.02641483(\hat{\bar{\beta}}_{BN})$, $0.09311917(\hat{\bar{\bar{\beta}}}_{BN})$.

Table 1.Bias of β for Different Values of θ, β and Different Sample Sizes.

	$\boldsymbol{\theta = 0.8 , \beta = 1.0}$		
n	$\boldsymbol{\hat{\beta}_p}$	$\boldsymbol{\hat{\bar{\beta}}_{BN}}$	$\boldsymbol{\hat{\bar{\bar{\beta}}}_{BN}}$
$n = 25$	0.06944827	0.06963663	0.2068714

$n = 50$	0.0319663	0.03198901	0.09369476
$n = 75$	0.02012667	0.02013311	0.06004821
$n = 100$	0.01905495	0.01905751	0.04856389
		$\boldsymbol{\theta = 1.0\,, \beta = 1.0}$	
n	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\hat{\hat{\boldsymbol{\beta}}}}_{\boldsymbol{BN}}$
$n = 25$	0.04492102	0.04511849	0.1909607
$n = 50$	0.02639078	0.02641483	0.09311917
$n = 75$	0.01152717	0.01153404	0.05426695
$n = 100$	0.01011534	0.0101181	0.04145646
		$\boldsymbol{\theta = 1.2\,, \beta = 1.0}$	
n	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\hat{\hat{\boldsymbol{\beta}}}}_{\boldsymbol{BN}}$
$n = 25$	0.04872336	0.04893382	0.2069312
$n = 50$	0.02849368	0.02851939	0.1002826
$n = 75$	0.01684183	0.0168492	0.06267788
$n = 100$	0.01415081	0.01415379	0.05300757

From Table 2 below, we note that the estimator ($\hat{\beta}_p$) has also the smallest mean squared errors (MSEs) for all sample sizes and the true parameters values considered. It's also worth noting that the mean squared errors (MSEs) for all estimators drop as the sample size grows. Therefore, based on bias and MSEs, the best performing estimator under no censoring is the standard profile likelihood estimator ($\hat{\beta}_p$).

Table 2.Mean Squared Errors of β for Different Values of θ, β and Different Sample Sizes.

	$\boldsymbol{\theta = 0.8, \beta = 1.0}$		
n	$\widehat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\widehat{\widehat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\widehat{\widehat{\widehat{\boldsymbol{\beta}}}}_{\boldsymbol{BN}}$
$n = 25$	0.7516971	0.7519412	0.9696935
$n = 50$	0.3049	0.3049131	0.3464013
$n = 75$	0.1772604	0.1772627	0.1927288
$n = 100$	0.1288766	0.1288773	0.1375648
	$\boldsymbol{\theta = 1.0, \beta = 1.0}$		
n	$\widehat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\widehat{\widehat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\widehat{\widehat{\widehat{\boldsymbol{\beta}}}}_{\boldsymbol{BN}}$
$n = 25$	0.7483676	0.7486032	0.971693
$n = 50$	0.3125195	0.312533	0.3563715
$n = 75$	0.1795949	0.1795972	0.1954329
$n = 100$	0.1328643	0.132865	0.1413831
	$\boldsymbol{\theta = 1.2, \beta = 1.0}$		
n	$\widehat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\widehat{\widehat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\widehat{\widehat{\widehat{\boldsymbol{\beta}}}}_{\boldsymbol{BN}}$
$n = 25$	0.844925	0.8451934	1.099157
$n = 50$	0.3416982	0.3417125	0.3881213
$n = 75$	0.1940211	0.1940237	0.2118529
$n = 100$	0.149568	0.1495689	0.1600617

3.2.2 Type II censored data

Tables 3-6 show the bias and mean squared errors (MSEs) for all estimators for the Lomax scale parameter under type II censoring data that we discussed earlier in the complete data section $\left(\hat{\beta}_p\ , \hat{\hat{\beta}}_{BN} \text{ and } \hat{\bar{\beta}}_{BN}.\right)$. In Table 3, We notice that ,the modified profile maximum likelihood estimator based on empirical covariance approximation ($\hat{\hat{\beta}}_{BN}$) has the smallest bias not only compared to the standard profile likelihood estimator ($\hat{\beta}_p$), but also to the modified profile maximum likelihood estimator based on ancillary statistics approximation ($\hat{\bar{\beta}}_{BN}$) when only the sample size is 50 and failure rate of 60% across all the values of the true parameters considered $(\theta, \beta) = (0.8.0,1.0), (1.0,1.0), (1.2,1.0)$. For instance, in Table 3, with $\theta = 0.8$,$n = 50$, the biases are 0.04988537 ($\hat{\beta}_p$), -0.01721757 ($\hat{\hat{\beta}}_{BN}$), 0.2874958 ($\hat{\bar{\beta}}_{BN}$). When the sample is 75 or 100 under the same failure rate (60%) and across all the true values of parameters considered $(\theta, \beta) = (0.8.0,1.0), (1.0,1.0), (1.2,1.0)$, the standard profile likelihood estimator ($\hat{\beta}_p$) has the smallest bias than the modified profile likelihood estimators ($\hat{\hat{\beta}}_{BN}$, $\hat{\bar{\beta}}_{BN}$). For example, in Table 3, with $\theta = 0.8$,$n = 100$, the biases are 0.003011131($\hat{\beta}_p$), -0.03104611 $\left(\hat{\hat{\beta}}_{BN}\right)$, 0.1159797($\hat{\bar{\beta}}_{BN}$). Further that, we notice that, when the failure rate is 80% (censored observations is 20 %), the modified profile maximum likelihood estimator based on an empirical covariance approximation ($\hat{\hat{\beta}}_{BN}$) has the smallest bias always for sample sizes n=75 and n=100 and across all the true values of parameters considered $(\theta, \beta) = (0.8.0,1.0), (1.0,1.0), (1.2,1.0)$. When the sample size is 50, this figure is not correct. This means that the biases of estimators her depend on the sample size and the censoring percentage.

Table 3. Bias of β for Different Values of θ, β and Different Sample Sizes, Failure Rates of 60%.

$\boldsymbol{\theta = 0.8}$, $\boldsymbol{\beta = 1.0}$

n	r	$\widehat{\beta}_p$	$\widehat{\widehat{\beta}}_{BN}$	$\widehat{\widehat{\widehat{\beta}}}_{BN}$
$n = 50$	$r = 30$	0.04988537	−0.01721757	0.2874958
$n = 75$	$r = 45$	0.009476831	−0.03546534	0.1623593
$n = 100$	$r = 60$	0.003011131	−0.03104611	0.1159797
		$\theta = 1.0, \beta = 1.0$		
n	r	$\widehat{\beta}_p$	$\widehat{\widehat{\beta}}_{BN}$	$\widehat{\widehat{\widehat{\beta}}}_{BN}$
$n = 50$	$r = 30$	0.0520114	−0.02226076	0.3225419
$n = 75$	$r = 45$	0.009669833	−0.04078877	0.1861122
$n = 100$	$r = 60$	0.005528588	−0.03300539	0.1369028
		$\theta = 1.2, \beta = 1.0$		
n	r	$\widehat{\beta}_p$	$\widehat{\widehat{\beta}}_{BN}$	$\widehat{\widehat{\widehat{\beta}}}_{BN}$
$n = 50$	$r = 30$	0.07264376	−0.007274964	0.3685922
$n = 75$	$r = 45$	0.03799421	−0.0177692	0.2358413
$n = 100$	$r = 60$	0.011129	−0.03164209	0.1605773

Table 4.Bias of β for Different Values of θ, β and Different Sample Sizes ,Failure Rates of 80%.

		$\theta = 0.8, \beta = 1.0$		
n	r	$\widehat{\beta}_p$	$\widehat{\widehat{\beta}}_{BN}$	$\widehat{\widehat{\widehat{\beta}}}_{BN}$
$n = 50$	$r = 40$	0.01448543	−0.01418991	0.1325044
$n = 75$	$r = 60$	0.0223255	0.002905222	0.09885229
$n = 100$	$r = 80$	0.0021629	−0.01225346	0.05769923
		$\theta = 1.0, \beta = 1.0$		
n	r	$\widehat{\beta}_p$	$\widehat{\widehat{\beta}}_{BN}$	$\widehat{\widehat{\widehat{\beta}}}_{BN}$

$n = 50$	$r = 40$	0.004189844	−0.02738142	0.1378329
$n = 75$	$r = 60$	0.01804031	−0.003606574	0.1057426
$n = 100$	$r = 80$	0.01135549	−0.004815155	0.0755752
			$\boldsymbol{\theta = 1.2\,, \beta = 1.0}$	
$\boldsymbol{n}$	$\boldsymbol{r}$	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\hat{\hat{\boldsymbol{\beta}}}}_{\boldsymbol{BN}}$
$n = 50$	$r = 40$	−0.002989626	−0.03752231	0.1447843
$n = 75$	$r = 60$	0.01656271	−0.007219475	0.1143707
$n = 100$	$r = 80$	0.01068609	−0.007226078	0.08313256

Tables 5 and 6 contains the results of mean squared errors for all estimators under Failure Rates of 60% and 80 % ,respectively. We notice that the mean sequared errors is the least for the the modified profile maximum likelihood estimator based on empirical covariance approximation ($\hat{\hat{\beta}}_{BN}$) for all sample sizes whether the failure rate is 60% or 80% and across all the true values of the parameters considerd $(\theta, \beta) = (0.8.0,1.0), (1.0,1.0), (1.2,1.0)$. For example , in Table 5 , with $\theta = 0.8$, $n = 50$,the mean squared errors (MSEs) are $1.192677(\hat{\beta}_p)$,$1.088475(\hat{\hat{\beta}}_{BN})$, 1.611714 ($\hat{\hat{\beta}}_{BN}$) ,and in Table 6 , with $\theta = 0.8$, $n = 50$,the mean squared errors (MSEs) are $0.5318778(\hat{\beta}_p)$,$0.5097712(\hat{\hat{\beta}}_{BN})$, 0.6404126 ($\hat{\hat{\beta}}_{BN}$).Therfore , we can conclude that based on mean squared errors , the best performing estimator under type II censoring is $\hat{\hat{\beta}}_{BN}$, then $\hat{\beta}_p$ and finally $\hat{\hat{\beta}}_{BN}$. It is also worth noting that the mean squared errors (MSEs) for all estimators decrease as the sample size increases and the proportion of censored observations decreases. Furthermore, the values of estimator's bias depend on the true value of parameters, sample sizes, and the proportion of censored observations in the sample.

Table 5.MSEs of β for Different Values of θ, β and Different Sample Sizes ,Failure Rates of 60%.

		$\boldsymbol{\theta = 0.8 \, , \beta = 1.0}$		
$\boldsymbol{n}$	$\boldsymbol{r}$	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\hat{\hat{\boldsymbol{\beta}}}}_{\boldsymbol{BN}}$
$n = 50$	$r = 30$	1.192677	1.088475	1.611714
$n = 75$	$r = 45$	0.7161285	0.6757677	0.8769252
$n = 100$	$r = 60$	0.5010118	0.4802361	0.5801214
		$\boldsymbol{\theta = 1.0 \, , \beta = 1.0}$		
$\boldsymbol{n}$	$\boldsymbol{r}$	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\hat{\hat{\boldsymbol{\beta}}}}_{\boldsymbol{BN}}$
$n = 50$	$r = 30$	1.430854	1.301745	1.969685
$n = 75$	$r = 45$	0.9000115	0.8466787	1.118009
$n = 100$	$r = 60$	0.6702818	0.6407744	0.7857691
		$\boldsymbol{\theta = 1.2 \, , \beta = 1.0}$		
$\boldsymbol{n}$	$\boldsymbol{r}$	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\hat{\hat{\boldsymbol{\beta}}}}_{\boldsymbol{BN}}$
$n = 50$	$r = 30$	1.637519	1.490294	2.27276
$n = 75$	$r = 45$	1.095941	1.0286	1.374207
$n = 100$	$r = 60$	0.7987492	0.7620716	0.9470012

Table 6.MSEs of β for Different Values of θ, β and Different Sample Sizes ,Failure Rates of 80%.

		$\boldsymbol{\theta = 0.8 \, , \beta = 1.0}$		

n	r	$\hat{\beta}_p$	$\widehat{\bar{\beta}}_{BN}$	$\widehat{\tilde{\beta}}_{BN}$
$n = 50$	$r = 40$	0.5318778	0.5097712	0.6404126
$n = 75$	$r = 60$	0.338342	0.328273	0.3845593
$n = 100$	$r = 80$	0.2252858	0.2206274	0.2461143
			$\theta = 1.0$, $\beta = 1.0$	
n	r	$\hat{\beta}_p$	$\widehat{\bar{\beta}}_{BN}$	$\widehat{\tilde{\beta}}_{BN}$
$n = 50$	$r = 40$	0.6070636	0.5829056	0.7312216
$n = 75$	$r = 60$	0.3907775	0.3791078	0.4466518
$n = 100$	$r = 80$	0.2615879	0.2558273	0.2884857
			$\theta = 1.2$, $\beta = 1.0$	
n	r	$\hat{\beta}_p$	$\widehat{\bar{\beta}}_{BN}$	$\widehat{\tilde{\beta}}_{BN}$
$n = 50$	$r = 40$	0.7063077	0.6765366	0.8570163
$n = 75$	$r = 60$	0.4472451	0.4336605	0.512278
$n = 100$	$r = 80$	0.3153406	0.30831	0.3484352

3.2.3 Type I censored data

Tables 7 and 8 below shows the bias results for all estimators that we discussed earlier in this chapter, but her under type I censoring data. The modified profile maximum likelihood estimator based on an empirical covariances approximation ($\hat{\tilde{\beta}}_{BN}$) present the smallest bias comparing to the standard profile likelihood estimator ($\hat{\beta}_p$),and also, to the modified profile maximum likelihood estimator based on ancillary statistics approximation ($\hat{\bar{\beta}}_{BN}$).This figure is consistent across all the sample sizes, proportion of censored observation, and the values of the true parameters considered $(\theta, \beta) = (0.8.0,1.0), (1.0,1.0), (1.2,1.0)$. In addition, the value of biases decreases when the sample size increase and proportion of censored observations decrease. For example, in Tabel 7, with $\theta = 0.8$,$n = 50$, the biases are $0.2973097(\hat{\beta}_p)$, 0.2148689 $\left(\hat{\hat{\beta}}_{BN}\right)$, $0.4389123(\hat{\hat{\beta}}_{BN})$,and in Tabel 8 , with $\theta = 0.8$,$n = 50$, the biases are $0.1264803(\hat{\beta}_p)$, 0.09293404$\left(\hat{\hat{\beta}}_{BN}\right)$, $0.2088536(\hat{\hat{\beta}}_{BN})$.

Table 7.Bias of β for Different Values of θ, β and Different Sample Sizes ,40% Censored Data.

	$\boldsymbol{\theta = 0.8, \beta = 1.0}$		
$\boldsymbol{n}$	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$
$n = 50$	0.2973097	0.2148689	0.4389123
$n = 75$	0.1763403	0.124541	0.2627149
$n = 100$	0.1359589	0.09820738	0.198376
	$\boldsymbol{\theta = 1.0, \beta = 1.0}$		
$\boldsymbol{n}$	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$
$n = 50$	0.3226008	0.2305478	0.4798584
$n = 75$	0.1864183	0.1282869	0.2839516
$n = 100$	0.1497622	0.1066336	0.2208916
	$\boldsymbol{\theta = 1.2, \beta = 1.0}$		
$\boldsymbol{n}$	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$
$n = 50$	0.4327258	0.3283987	0.608986
$n = 75$	0.2495663	0.1840209	0.3591117
$n = 100$	0.1886426	0.1401105	0.2686519

Table 8.Bias of β for Different Values of θ, β and Different Sample Sizes ,20% Censored Data.

	$\boldsymbol{\theta = 0.8\,,\beta = 1.0}$		
$\boldsymbol{n}$	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\hat{\hat{\boldsymbol{\beta}}}}_{\boldsymbol{BN}}$
$n = 50$	0.1264803	0.09293404	0.2088536
$n = 75$	0.08561933	0.06416345	0.1380114
$n = 100$	0.06464137	0.04903077	0.1026434
	$\boldsymbol{\theta = 1.0\,,\beta = 1.0}$		
$\boldsymbol{n}$	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\hat{\hat{\boldsymbol{\beta}}}}_{\boldsymbol{BN}}$
$n = 50$	0.1749115	0.1365016	0.2680733
$n = 75$	0.1025145	0.07856541	0.1603866
$n = 100$	0.0633128	0.04597983	0.105016
	$\boldsymbol{\theta = 1.2\,,\beta = 1.0}$		
$\boldsymbol{n}$	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\hat{\hat{\boldsymbol{\beta}}}}_{\boldsymbol{BN}}$
$n = 50$	0.178889	0.1373527	0.2801677
$n = 75$	0.1295151	0.102768	0.1935463
$n = 100$	0.09464382	0.07508975	0.1411572

Regarding the mean squared errors (MSEs), as shown blow in Table 9 and 10, like type II censoring, the modified profile maximum likelihood estimator based on empirical covariance approximation ($\hat{\hat{\beta}}_{BN}$) has the least mean squared errors among all the sample sizes considered whether the data have 40 % or 20 % censored observations and across all the true values of the parameters that we considered $(\theta, \beta) = (0.8.0,1.0), (1.0,1.0), (1.2,1.0)$. Furthermore, as the sample size grows and the number of censored data drops, the bias and mean squared errors (MSEs) for all estimators decrease. Therefore, we can conclude that, like the previous case in type II censoring, the best performing estimator based on biases and mean squared errors (MSEs) under type I censoring is also $\hat{\hat{\beta}}_{BN}$, then $\hat{\beta}_p$ and finally $\hat{\hat{\beta}}_{BN}$.

Table 9.MSEs of β for Different Values of θ, β and Different Sample Sizes ,40% Censored Data.

	$\boldsymbol{\theta = 0.8\,, \beta = 1.0}$		
$\boldsymbol{n}$	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$
$n = 50$	1.723507	1.535303	2.095795
$n = 75$	0.8776371	0.8062395	1.011576
$n = 100$	0.5918235	0.555391	0.6596562
	$\boldsymbol{\theta = 1.0\,, \beta = 1.0}$		
$\boldsymbol{n}$	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$
$n = 50$	1.991764	1.769823	2.428213
$n = 75$	1.050137	0.9645024	1.21562
$n = 100$	0.7462274	0.6981894	0.8359341
	$\boldsymbol{\theta = 1.2\,, \beta = 1.0}$		
$\boldsymbol{n}$	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$
$n = 50$	2.657058	2.356114	3.229395
$n = 75$	1.492934	1.37348	1.719617
$n = 100$	0.9819477	0.9162409	1.103524

Table 10.MSEs of β for Different Values of θ, β and Different Sample Sizes ,20% Censored Data.

$\boldsymbol{n}$	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\breve{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\tilde{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$
		$\boldsymbol{\theta = 0.8 , \beta = 1.0}$	
$n = 50$	0.6092399	0.5685677	0.721311
$n = 75$	0.3682551	0.3514674	0.4121435
$n = 100$	0.2360463	0.2279355	0.2573829
		$\boldsymbol{\theta = 1.0 , \beta = 1.0}$	
$\boldsymbol{n}$	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\breve{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\tilde{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$
$n = 50$	0.769086	0.7150742	0.9132357
$n = 75$	0.4463088	0.4252308	0.5001879
$n = 100$	0.2867205	0.2769057	0.3119131
		$\boldsymbol{\theta = 1.2 , \beta = 1.0}$	
$\boldsymbol{n}$	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\breve{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\tilde{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$
$n = 50$	0.9349742	0.8713125	1.104755
$n = 75$	0.524831	0.4992046	0.590599
$n = 100$	0.3715427	0.3577615	0.4066636

CHAPTER 4: NUMERICAL EXAMPLES WITH REAL DATA

In this chapter, we utilize two real-world data sets to show how the suggested methods function and to validate how well our estimators perform in reality. For both examples, the initial guess used in the profile and modified profile maximum likelihood functions is based on the method of moments (See appendix B) for all sampling schemes (complete data, types I and II censoring).

4.1 Meteorological study

In this first numerical illustration, we consider a date set taken from a meteorological study by Simpson [26] and subsequently investigated by Giles et al. [18], Bryson [7], Helu et al. [20], and A. Baklizi et al. [4]. The Lomax distribution's applicability for this data was tested by Helu et al. [20] using The Kolmogrov– Smirnov (K–S) test, as well as the Anderson–Darling (A–D) and chi-square tests. The research was based on radar-evaluated rainfall from 52 south Florida cumulus clouds, 26 seeded clouds, and 26 control clouds. We obtained the profile and modified profile likelihood estimators for the Lomax scale parameter using the following measurements from the control group only:

26.1,26.3,87,95,373.4,0,17.3,24.4,11.5,321.2,68.5,81.2,47.3,28.6,830.1,345.5,1202.6, 36.6,4.9,4.9,41.1,29,163,244.3,147.8,21. Here $n = 26$ and there is no censoring. For type II censoring, we consider a subset of measurements of the control group and impose a failure rate of 80%. For type I censoring, we consider also a subset of measurements of the control group and impose 20% censored data. The values of the point estimates of the Lomax scale parameter gotten by maximizing the profile and the modified profile likelihoods under complete, type I and type II censoring are summarized in the following table (Tabel 11).

Table 11.Point estimation for the Lomax Scale Parameter Under Different Sampling Schems

Data Schemes	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$
Complete data	0.01270243	0.01270243	0.01457947
Type II censoring	0.01909989	0.018055	0.02299024
Type I censoring	0.01377231	0.01274749	0.01612275

From Tabel 11, under complete data, the standard profile maximum likelihood estimator ($\hat{\beta}_p$), and the modified profile maximum likelihood point estimate based on empirical covariances approximation ($\hat{\hat{\beta}}_{BN}$) are identical. Point estimate obtained using the modified profile likelihood based on an ancillary statistics approximation ($\hat{\hat{\beta}}_{BN}$), on the other hand, are numerically larger than the previous estimators. We observed that the standard profile likelihood estimator ($\hat{\beta}_p$), is the best performing estimator based on bias and MSE in the simulation study with complete data. As a result, this estimator ($\hat{\beta}_p$) is recommended to use for this type of data, sample size, and distribution based on the result of the corresponding simulation study.

From Tabel 11, the point estimate obtained under type II censoring obtained from the modified profile likelihood function based on empirical covariances approximation ($\hat{\hat{\beta}}_{BN}$) is smaller than the standard profile likelihood estimator ($\hat{\beta}_p$), and also, to the modified profile maximum likelihood estimator based on an ancillary statistics approximation ($\hat{\hat{\beta}}_{BN}$). Based on a simulation study with type II censoring for $n = 50, r = 30$ (failure rate =80%) which has roughly the close number of failure observations as this example ($r = 21$), we found that the modified profile likelihood based on empirical covariances ($\hat{\hat{\beta}}_{BN}$) is the best performing estimator based on bias and MSE. As a result, in this data set example and distribution, it is recommended that you utilize this estimator ($\hat{\hat{\beta}}_{BN}$).

From Table 11, the point estimate under type I censoring resulting from the modified profile likelihood function based on empirical covariances approximation ($\hat{\hat{\beta}}_{BN}$) has the least value, as it does for type II censoring. When we have type I censoring in this example, we can recommend using the modified profile maximum likelihood point estimate based on empirical covariances approximation ($\hat{\hat{\beta}}_{BN}$) as the corresponding simulation study with type I censoring for n=50 and censoring of 20%, giving us 30 complete observations, which is roughly close to the number of complete observations as this example (21 complete observations under 20 percent censoring proportion). Because this estimator had the best performance in the simulation research in terms of bias and MSE.

4.2 Computer file sizes

The data set in the second numerical example represents computer file sizes (in bytes) for all 269 files with the.ini extension on David Giles' PC (running Windows) on March 19, 2011 (see Giles et al. [18]). Previous work by Holland et al. [21] demonstrated the Lomax distribution's superiority against a number of other competitors for modeling such file sizes. We calculated the profile and modified profile likelihood estimators for the Lomax scale parameter using the following data:

67,67,67,67,67,67,67,67,67,67,67,20,62,113,113,67,67,67,67,67,67,67,6850,320,1727,66,66,1418,23,1192,175,8698,2470,1771,243,28529,141,685,51,197,60,248,11117,204,169,125,172,123,161,173,178,170,177,125,172,202,123,544,24684,485,294,1114,266,266,3285,16504,9540,2204,70,51,51,51,84,72,70,165,88,88,88,88,88,88,88,88,88,529,253,213,326,95,95,429,2730,3170,166,165,207,209,224,186,272708,182,6651,166,207,162,165,446,166,212,171,143,236,166,167,168,95,230,211,526,2452,1838,603,105,2281,124,69,98,148,70,807,223,79,6000,71,75,74,77,75,80,79,78,72,175,477,6

2792,2515,278,143,3718,9073,5335,65,544,2885,4311,4586,4107,4728,4862,4101,4632,4245,3803,4378,4418,4369,2939,1298,3718,9073,2,525,1405,376,4161,231,36,37,762,513508,65,67,65,436724,513508,510714,2435152,240,53478,62,55070,9605,2,1015477,10110,1931,2695,2891,1152,360124,2732,343,6877,3458,12082,53478,13223,62,62,113,113,67,67,67,67,67,67,184,181,62,206,482,348,84,84,4069,13804,11886,5013,3865,16226,14416,3529,2142,18594,11865,5809,13610,3971,6735,2409,14450,8389,4510,3203,6761,2078,5848,5171,3106,3068,6073,5832,4543,6085,5915,3951,24142,240,66,1698,65,145,67,67,67,67,67).

Here $n = 269$ and there is no censoring. For type II censoring, we consider a subset of computer file sizes and impose a failure rate of 60%. For type I censoring, we consider also a subset of computer file sizes and impose 40% censored proportion. The following table (Tabel 12) summarize the results of the point estimates for the Lomax scale parameter gotten by maximizing the profile and the modified profile likelihoods under complete, type I and type I censoring data.

Table 12.Point estimation for the Lomax Scale Parameter Under Different Sampling Schems

	$\hat{\boldsymbol{\beta}}_{\boldsymbol{p}}$	$\hat{\hat{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$	$\hat{\tilde{\boldsymbol{\beta}}}_{\boldsymbol{BN}}$
Complete data	0.007858893	0.007858893	0.007914804
Type II censoring	0.00502907	0.004980985	0.005151622
Type I censoring	0.005047209	0.005057715	0.005080545

As illustrated in Tabel 12, under complete data, the point estimates from the conventional profile and modified profile maximum likelihood based on empirical covariances are equivalent, while the modified profile maximum likelihood based on an ancillary statistic is numerically less. We can recommend using the point estimate derived from the standard profile likelihood function ($\hat{\beta}_p$) based on the corresponding simulation study with simulation indices similar to this example , because it performs the best in terms of bias and MSE.

From Tabel 12, under type II censoring, the point estimates from the conventional profile and the modified profile maximum likelihood based on empirical covariances are similar and both less than the modified profile maximum likelihood based on ancillary statistics. All the point estimates resulting from the standard profile and modified profile maximum likelihoods are nearly identical for type I censoring. The modified profile maximum likelihood point estimate based on empirical covariances approximation ($\hat{\hat{\beta}}_{BN}$) is recommended based on the corresponding simulation study for types I and II censoring with simulation indices similar to this example because it shows the best performance in terms of bias and MSE.

CHAPTER 5: CONCLUSION

5.1 Summary and conclusion

The Barndorff-Nielsen modified profile likelihood function, which is based on empirical covariances and an ancillary statistic approximation, is used to modify the standard maximum profile likelihood estimator for the Lomax scale parameter (parameter of interest) in the presence of the nuisance shape parameter. These estimators (standard profile likelihood estimator and modified profile likelihood estimators) are not available in simple closed forms, but can be obtained numerically as roots of some complicated likelihood equations. We used simulation techniques to compare the biases and mean squared errors of the maximum profile likelihood estimator and the modified profile likelihood estimator in order to find the best performing estimator for the Lomax scale parameter.

According to the criteria used (bias and MSEs) in the simulation study, the numerical results show that under type I and II censored data, the modified profile maximum likelihood estimator based on empirical covariances approximation outperforms not only the standard maximum profile likelihood estimator, but also the modified profile maximum likelihood estimator based on an ancillary statistics approximation. It has almost lowest bias and always has the lowest mean squared errors for all sample sizes considered $(n = 50,75,100)$, proportion of censored observations (40%,20%), and across all true values of the parameters that we considered $(\theta, \beta) = (0.8.0,1.0), (1.0,1.0), (1.2,1.0)$. When there is no censoring, the best performing estimator is the standard profile maximum likelihood estimator because it has the smallest bias and mean squared errors for all sample sizes, proportion of censored observations, and across all the true values of parameters that we considered.

5.2 Recommendations

Researchers should base their inferences about the Lomax scale parameter on the Barndorff-Nielsen's modified profile likelihood function, which is based on empirical covariances estimate. This is because more precise inferences regarding the scale parameter are possible. This is especially true if the data has been censored (Types I and II).

5.3 Suggestions for further research

In this thesis, we obtained adjusted profile maximum likelihood estimators for the Lomax scale parameter in the presence of a shape parameter under non-censored and censored (Types I and II) data. We considered profile likelihood function adjustments based on several approximations to Barndorff- Nielsen's modified profile log-likelihood function.

Researchers might utilize the same best estimator that we discovered in future research to find interval estimation and construct hypothesis testing procedures. They can also use other adjustments to the profile maximum likelihood function, such as Cox and Reid's [11] conditional profile likelihood or conditioning on complete and sufficient statistics, to get more precise point estimation if it is possible. Also, they could use the same profile maximum likelihood function modifications used in this thesis on other important lifetime distributions and/or using more general censoring patterns like progressive or hybrid censoring schemes.

Bootstrap inference using simulation is also suggested to examine and validate the accuracy of the values of the point estimations that we obtained from the samples of real data examples based on bias and mean squared errors. Moreover, it is recommended to use bootstrap point estimation, interval estimation, and confidence interval estimation which may give us more accurate estimation for the Lomax scale parameter.

Because prediction is such an essential topic in statistical inference, it is recommended that we use the cross-validation approach to see if the model fits the data and can be trusted in prediction. This can be done by dividing the given data into two parts and comparing part of the model prediction with real data. The first part (the largest part) is used to fit the proposed model, while the second part (the smallest part) is used to compare the predicted value from the fitted model to the real data. If the prediction and real data points are almost same, we can trust this model to predict.

REFERENCES

[1] Al-Zahrani, Bander, and Mashail Al-Sobhi. "On parameters estimation of Lomax distribution under general progressive censoring." *Journal of Quality and Reliability Engineering* 2013 (2013).

[2] Arnold B.C. (2008) Pareto and Generalized Pareto Distributions. In: Chotikapanich D. (eds) Modeling Income Distributions and Lorenz Curves. Economic Studies in Equality, Social Exclusion and Well-Being, vol 5. Springer, New York, NY.

[3] Arnold, B. C. "Pareto distributions fairland." *Maryland: International Cooperative Publishing House* (1983).

[4] Baklizi, A. (2021), Saadati Nik, A., & Asgharzadeh,A. (n.d.). Likelihood and Bayesian Inference in the Lomax Distribution Under progressive censoring. To appear.

[5] Barndorff-Nielsen, Ole. "On a formula for the distribution of the maximum likelihood estimator." *Biometrika* 70, no. 2 (1983): 343-365.

[6] Barreto, Larissa Santana, Audrey HMA Cysneiros, and Francisco Cribari-Neto. "Improved Birnbaum–Saunders inference under type II censoring." *Computational Statistics & Data Analysis* 57, no. 1 (2013): 68-81

[7] Bryson, Maurice C. "Heavy-tailed distributions: properties and tests." *Technometrics* 16, no. 1 (1974): 61-68.

[8] Chahkandi, Majid, and Mojtaba Ganjali. "On some lifetime distributions with decreasing failure rate." *Computational Statistics & Data Analysis* 53, no. 12 (2009): 4433-4440.

[9] Corbellini, Aldo, Lisa Crosato, Piero Ganugi, and Marco Mazzoli. "Fitting Pareto II distributions on firm size: Statistical methodology and economic puzzles." In *Advances in Data Analysis*, pp. 321-328. Birkhäuser Boston, 2010.

[10] Cox, David R. "Regression models and life-tables." *Journal of the Royal Statistical Society: Series B (Methodological)* 34, no. 2 (1972): 187-202.

[11] Cox, David Roxbee, and Nancy Reid. "Parameter orthogonality and approximate conditional inference." *Journal of the Royal Statistical Society: Series B (Methodological)* 49, no. 1 (1987): 1-18.

[12] Cox, David Roxbee, and Nancy Reid. "Parameter orthogonality and approximate conditional inference." *Journal of the Royal Statistical Society: Series B (Methodological)* 49, no. 1 (1987): 1-18.

[13] Cramer, Erhard, and Anja Bettina Schmiedt. "Progressively type-II censored competing risks data from Lomax distributions." *Computational Statistics & Data Analysis* 55, no. 3 (2011): 1285-1303.

[14] Ferrari, Silvia LP, Michel Ferreira Da Silva, and Francisco Cribari–neto. "Adjusted profile likelihoods for the Weibull shape parameter." *Journal of Statistical Computation and Simulation* 77, no. 7 (2007): 531-548.

[15] Fraser, D. A. S., and N. Reid. "Ancillaries and third order significance." *Utilitas Mathematica* 7 (1995): 33-53.

[16] Fraser, Donald Alexander Stuart, Nancy Reid, and J. Wu. "A simple general formula for tail probabilities for frequentist and Bayesian inference." *Biometrika* 86, no. 2 (1999): 249-264.

[17] Ghitany, M. E., F. A. Al-Awadhi, and LA1122 Alkhalfan. "Marshall–Olkin extended Lomax distribution and its application to censored data." *Communications in Statistics—Theory and Methods* 36, no. 10 (2007): 1855-1866.

[18] Giles, David E., Hui Feng, and Ryan T. Godwin. "On the bias of the maximum likelihood estimator for the two-parameter Lomax distribution." *Communications in Statistics-Theory and Methods* 42, no. 11 (2013): 1934-1950.

[19] Harris, Carl M. "The Pareto distribution as a queue service discipline." *Operations Research* 16, no. 2 (1968): 307-313.

[20] Helu, Amal, Hani Samawi, and Mohammad Z. Raqab. "Estimation on Lomax progressive censoring using the EM algorithm." *Journal of Statistical Computation and Simulation* 85, no. 5 (2015): 1035-1052.

[21] Holland, Oliver, Assen Golaup, and A. H. Aghvami. "Traffic characteristics of aggregated module downloads for mobile terminal reconfiguration." *IEE Proceedings-Communications* 153, no. 5 (2006): 683-690.

[22] Howlader, Hatem A., and Anwar M. Hossain. "Bayesian survival estimation of Pareto distribution of the second kind based on failure-censored data." *Computational statistics & data analysis* 38, no. 3 (2002): 301-314.

[23] Johnson, Norman L., Samuel Kotz, and Narayanaswamy Balakrishnan. *Continuous univariate distributions, volume 2*. Vol. 289. John wiley & sons, 1995.

[24] Klein, John P., and Melvin L. Moeschberger. *Survival analysis: techniques for censored and truncated data*. Vol. 2. New York: Springer, 2003.

[25] Lawless, Jerald F. *Statistical models and methods for lifetime data*. John Wiley & Sons, 2011.

[26] Lee, Elisa T., and John Wang. *Statistical methods for survival data analysis*. Vol. 476. John Wiley & Sons, 2003.

[27] Lehmann, Erich L., and George Casella. *Theory of point estimation*. Springer Science & Business Media, 2006.

[28] Lindley, Dennis V. "Approximate bayesian methods." *Trabajos de estadística y de investigación operativa* 31, no. 1 (1980): 223-245.

[29] Lomax, Kenneth S. "Business failures: Another example of the analysis of failure data." *Journal of the American Statistical Association* 49, no. 268 (1954): 847-852.

[30] Mahmoud, Mohamed Abdul Wahab, Rashad Mohamed El-Sagheer, Ahmed Abo-Elmagd Soliman, and Ahmed Hamed Abd Ellah. "Inferences of the lifetime performance index with Lomax distribution based on progressive type-II censored data." *Economic Quality Control* 29, no. 1 (2014): 39-51.

[31] Moore, Dirk F. *Applied survival analysis using R*. New York, NY: Springer, 2016.

[32] Okasha, Hassan M. "E-Bayesian estimation for the Lomax distribution based on type-II censored data." *Journal of the Egyptian Mathematical Society* 22, no. 3 (2014): 489-495.

[33] Severini, Thomas A. "An approximation to the modified profile likelihood function." *Biometrika* 85, no. 2 (1998): 403-411.

[34] Severini, Thomas A. "An empirical adjustment to the likelihood ratio statistic." *Biometrika* 86, no. 2 (1999): 235-247.

[35] Severini, Thomas A. *Likelihood methods in statistics*. Oxford University Press, 2000.

[36] Simpson, Joanne. "Use of the gamma distribution in single-cloud rainfall analysis." *Monthly Weather Review* 100, no. 4 (1972): 309-312.

[37] Stevenson, Mark, and I. V. A. B. S. EpiCentre. "An introduction to survival analysis." *EpiCentre, IVABS, Massey University* (2009).

[38] Tierney, Luke, and Joseph B. Kadane. "Accurate approximations for posterior moments and marginal densities." *Journal of the american statistical association* 81, no. 393 (1986): 82-86.

[39] Yang, Zhenlin, and Min Xie. "Efficient estimation of the Weibull shape parameter based on a modified profile likelihood." *Journal of Statistical Computation and Simulation* 73, no. 2 (2003): 115-123.

APPENDIX A: THE INVERSE TRANSFORMATION TECHNIQUE

Let U be a uniform random variable for obtaining a random sample from a Lomax distribution with complete data. To get random numbers like x from the Lomax distribution, solve U=F(t), where F(t) is the cumulative distribution function of the Lomax distribution (see equation (4)).

Now,

$$U = F(t)$$

$$U = 1 - \frac{1}{(1+\beta t)^{\theta}}$$

$$(1+\beta t)^{-\theta} = 1 - U$$

$$\ln(1+\beta t) = \frac{\ln(1-U)}{-\theta}$$

$$e^{\ln(1+\beta t)} = e^{\frac{1}{-\theta}\ln(1-n)}$$

$$(1+\beta t) = e^{-\frac{1}{\theta}\ln(1-U)}$$

$$\beta t = e^{-\frac{1}{\theta}\ln(1-U)} - 1$$

$$t = \frac{1}{\beta}\left(e^{-\frac{1}{\theta}\ln(1-U)} - 1\right) = \frac{1}{\beta}\left(e^{ln(1-U)^{-\frac{1}{\theta}}} - 1\right)$$

$$t = \frac{1}{\beta}\left((1-U)^{-\frac{1}{\theta}} - 1\right) \qquad (*)$$

Then we apply the equation (*) above to generate complete data from the Lomax distribution.

Let T be the censoring constant from Lomax, then the type I censored sample from Lomax distribution is gotten as follow:

$$x_j = min(t_j, T)$$

Let $t_{(j)}, (j = 1, \dots., r)$ be the smallest r order statistics from a sample of size n .Here r be the r^{th} failure ($r< n$) , then $t_{(j)}, (j = 1, \dots., r)$ represent sample from type II censoring.

APPENDIX B: METHODS OF MOMENTS

$$\frac{1}{\beta(\theta-1)} = \bar{x} \qquad (1)$$

$$\frac{\theta}{\left(\beta(\theta-1)\right)^2(\theta-2)} = s^2 \qquad (2)$$

Dividing the second equation by the square of the first equation we obtain

$$\frac{\theta}{(\theta-2)} = \frac{s^2}{\bar{x}^2} \rightarrow \theta\frac{s^2}{\bar{x}^2} - 2\frac{s^2}{\bar{x}^2} = \theta \rightarrow \hat{\theta} = \frac{2\frac{s^2}{\bar{x}^2}}{\left(\frac{s^2}{\bar{x}^2} - 1\right)} = \frac{2}{1-\frac{\bar{x}^2}{s^2}}$$

Substituting in equation (1) we obtain

$$\hat{\beta} = \frac{1}{\bar{x}\left(\frac{2}{1-\frac{\bar{x}^2}{s^2}} - 1\right)}$$

Use this as initial guess for β