

Maximum Likelihood Estimation of Parameters of Lomax Distribution Based on Progressive Type-II Hybrid Censoring Scheme.

Mwendwa Peace Mwende

B Ed(Sc)

I56/CTY/24050/2013

Department of Statistics and Actuarial Science

A Project Submitted in Partial Fulfillment of the Requirements for the Award of the Degree of Master of Science (**Statistics**) in the School of Pure and Applied Science of Kenyatta University.

©2018

Declaration

This research project is my original work and has not been presented elsewhere for a degree award.

Signature..... Date.....

Mwendwa Peace Mwende

I confirm that the work reported in this project was carried out by the candidate under my supervision.

Signature.....Date.....

Prof.Leo Odongo,
Kenyatta University.

Dedication

I dedicate this work to my Husband Peter Stanley Kariuki, my children Janice and Yisrael and my parents Sarah and Julius Mwendwa for their incredible support. May the Almighty God bless them all.

Acknowledgement

For this work to have attained the current status, it has taken not only my effort but also support of other personalities whom I wish to acknowledge here.

First and foremost, I wish to convey my sincere appreciation and recognition to my supervisor, Professor Leo Odongo who read the entire project work and made invaluable comments, corrections and suggestions that saw this project to its completion.

Peter Stanley, My Husband is a constant inspiration and source of encouragement. Peter, your belief in me is immense and it is a privilege to be walking down life's path together with you, thank you very much. On the same note, I would like to acknowledge my wonderful children Janice and Yisrael, my parents and sisters who remind me every day of the abundant blessing that is life itself.

My other special appreciation also goes to my good Lecturers Dr.C. Ouma, Dr.E. Njenga, Dr.J. Kahiri, Dr.G. Orwa, Dr.B. Muema, Mr.A. Ruto and Dr.K. Ananda for their valuable input and assistance throughout my entire study.

I would like also to thank my wonderful friends Charles Karuga, George, Kanogu and Mike for their encouragement, tireless effort and time they gave to me to ensure that this work is a success.

And above all, My gratitude and Thanks goes to Almighty God for giving me gift of life and good health over the entire period of my study. I believe without Him, I would not be the same person I am today.

Contents

Declaration	i
Dedication	ii
Acknowledgment	iii
List of Tables	vi
List of Abbreviations	vii
Abstract	viii
1 INTRODUCTION	1
1.1 Background of the Study	1
1.2 Terminologies	2
1.3 Statement of the Problem	5
1.4 Justification of the Study	7
1.5 Objectives	8
1.5.1 General Objective	8
1.5.2 Specific objectives	8
1.6 Significance of the Study	8
1.7 Outline of the project	9
2 LITERATURE REVIEW	10
2.1 Chapter Overview	10
2.2 Hybrid Censoring Scheme	10
2.3 Progressive Type-II Censoring Scheme	11
2.4 Progressive Type-II Hybrid Censoring Scheme	12
2.5 Lomax Distribution	13

3	LIKELIHOOD FUNCTIONS AND PROPOSED MLES VIA EM ALGO-	
	RITHM	15
3.1	Chapter overview	15
3.2	Progressive Type II Hybrid Censoring	15
3.3	Model Description	16
3.4	NR Algorithm	18
3.5	EM Algorithm	22
4	RESULTS AND DISCUSSIONS	28
4.1	Chapter Overview	28
4.2	Description of the Study	28
4.3	Numerical Outcomes and Analysis	31
4.4	Real Data Analysis	37
5	SUMMARY, CONCLUSIONS AND RECOMMENDATIONS	42
5.1	Introduction	42
5.2	Summary	42
5.3	Conclusions	43
5.4	Recommendations	44
	5.4.1 For use	44
	5.4.2 For Further study	44

List of Tables

4.1	<i>The MLEs via NR and EM algorithms , mean squared error and mean bias for the parameters of Lomax distribution under censoring scheme 1 when $\lambda = 1.5$ and $\theta = 0.8$</i>	31
4.2	<i>The MLEs via NR and EM algorithms , mean squared error and bias for the parameters of Lomax distribution under censoring scheme 2 when $\lambda = 1.5$ and $\theta = 0.8$</i>	33
4.3	<i>The MLEs via NR and EM algorithms, mean squared error and bias for the parameters of Lomax distribution under censoring scheme 3 when $\lambda = 1.5$ and $\theta = 0.8$</i>	35
4.4	The progressively type-II hybrid censored sample for the 128 bladder cancer patients when $m = 42, T = 4.50$, $R_1 = \dots = R_{41} = 2, R_{42} = 4$ is described below;	37
4.5	<i>Survival Estimates and their corresponding estimated standard errors</i>	38
4.6	The progressively type-II hybrid censored sample for the 128 bladder cancer patients when $m = 42, T = 3.67$, $R_1 = \dots = R_{40} = 2, R_{42} = 8$ is described below;	39
4.7	<i>Survival Estimates and their corresponding estimated standard errors</i>	40

List of Abbreviations and Acronyms

EM	Expectation-Maximization
NR	Newton Raphson
MLE	Maximum Likelihood Estimator
CDF	Cumulative Distribution Function
PDF	Probability Density Function
X_i	Lifetime of the i th item; $i = 1, \dots, n$.
$X_{i:m:n}$	the i^{th} observed failure time
R_i	the number of units removed at the time of i^{th} failure.
m	the number of failures observed before termination of experiment for Case I.
J	the number of failures observed before termination for Case II
D	the number of failures, i.e. $D = m$ for Case I and $D = J$ for Case II
R_D^*	the number of remaining units left at the time point T for Case I.
R_J^*	the number of remaining units left at the time point T for Case II.

Abstract

Lomax distribution is an important lifetime distribution. The process of obtaining estimates of parameters for different lifetime distributions under various schemes still remains an area of interest. In Lomax distribution, parameter estimates have been obtained using ordinary procedures like Newton Raphson when the test units follows a progressive Type-II hybrid censoring scheme whereby within most cases the obtained values do not converge easily to the true value. The MLEs are observed to be generally difficult to obtain in a closed format. As a result, we recommend to employ EM algorithm procedures in order to attain the MLEs of the parameters of Lomax distribution build onto progressive Type-II hybrid which amalgamates Type-II and hybrid censoring schemes. Performance of these obtained MLEs is compared with those obtained using NR methods and EM algorithm for different censoring schemes with regard to their mean bias and MSE at fixed parameter values of λ and θ . Simulation studies reveal that the MLEs via EM algorithm performs better than those obtained via NR method. The results of the obtained estimators are illustrated on real data.

CHAPTER ONE

1 INTRODUCTION

1.1 Background of the Study

Experiments based on various synopsis in life experiments and reliability data, have units going astray or being withdrawn prior to the failure occurring. The occurrence of the loss in the study may be unintentional or pre-designed. However, to maximize on the related cost and time linked to the testing, the withdrawal of units prior to the failure is planned beforehand. There are many different forms of schemes of censoring that have widely been discussed within the earlier studies. There are two most usual kinds of censoring schemes; Type-I censoring scheme, whereby an experiment goes on as far as time T specified in advance and Type-II censoring scheme whereby the test persists till a prevalence of a number of debacles specified in advance. The Type-I and Type-II censoring schemes mixed together makes hybrid censoring scheme.

A typical disadvantage with regular Type-I, Type-II and hybrid censoring schemes remains their inability to permit intermediate withdrawal of elements during the experiment except at the terminal point. However, progressive Type-II censoring scheme utilizes that benefit; (Balakrishnan et al., 2000, 2004 and 2007).

Nevertheless, progressive Type-II censoring scheme has a major problem in that an experiment can take a very long time. To remedy this predicament, Kundu et al.(2006) lodged a fresh censoring scheme termed Progressively Type-II Hybrid Censoring which guarantees that the pre-specified time point is not exceeded during the experiment. For more detailed description regarding benefits of progressively Type-II hybrid censoring, make reference to (Kundu et al., 2006) as well as (Childs et al., 2007). The authors in the Both publications have presumed the lifetimes possess exponential distributions.

Mokhtari et al. (2011) studied deductions for Weibull distribution under this censoring scheme.

Salem and Abo-Kasem (2011) considered approximation of parameters of Exponential-Weibull distribution procured from progressive type -II hybrid censored samples.

Clearly from these, we observe that very little research has been conducted grounded on progressive type-II hybrid censoring schemes considering that there exists quite a number of life time distributions. Therefore, in this project the focus is to estimate the MLEs of the parameters of Lomax distribution grounded onto progressive type-II hybrid censored samples. Maximum likelihood estimators (MLEs) are obtained via the EM algorithm and the outcomes compared with those obtained via Newton-Raphson method.

Lomax distribution emerged first as subsequent type of the Pareto distribution according to submission by Lomax (1954). Hassan and Al-Ghamdi (2009) applied Lomax distribution to life testing and reliability modeling. Harris (1968), Atkinson and Harrison (1978) used income and wealth distribution data to get estimates of Lomax distribution . (Corbelini et al., 2007) applied Lomax distribution to firm sizes.

More so, the distribution has also wide applications in the study of living things, modeling in computers among many others, see Holland et al. (2006). Bryson (1974), suggested utilization of Lomax distribution when the data is heavy-tailed as an alternate method to exponential distribution. The distribution also fosters as an utilitarian model in the research related to biological analysis and queuing theory. Details on the genesis, properties and other characteristics as well as significance of this distribution, refer Arnold (1983) and Johnson et al (1994).

1.2 Terminologies

Survival analysis - refers to the study of data that correlates with time from a well-designated time dawn until some well-defined event occurs or end point is attained. (Failure time)

Censoring - In survival analysis, this is a common feature. The existence time of a character is said to be censored when the endpoint of concern is not yet discerned for a character.

There exists three censoring classes;

Right censoring, Left censoring and Interval censoring.

Right censoring- It's one of the most frequently encountered type, at least in medical practices. Some participants exit an investigation earlier to the happening of an event, so we only know that their event time T_i lies in an interlude (t, ∞) . Several different study designs can lead to right censoring:

- (i) A fixed number of participants have a treatment administered concurrently, and the research is aborted at some time specified beforehand. Animal investigations are often concluded by eliminating any existing animals after a definite time. This is termed as **Type 1 censoring**.
- (ii) Subjects enter an investigation at interludes, and are all discerned to a fixed time. This is termed as **generalized Type 1 censoring**. An assessment of time from a distinct genesis for each subject is determined.
- (iii) Each individual may have a different fixed censoring time; it may be too costly to maintain all the participants in an animal research living for the entire time, so some may be sacrificed prematurely. This is termed **progressive type 1 censoring**.
- (iv) All subjects infiltrate the research concurrently and the investigation is concluded following a fixed number of events. Such investigations are frequently used when examining machine components for malfunctioning. This type is termed as **Type II censoring**.
- (v) **Progressive type II censoring** occurs during that time a research ceases in phases: following the first r_1 incident some research units are withdrawn, then following the next r_2 incidents further units are withdrawn and so on.
- (vi) Sometimes subjects are disconnected from an investigation due to outside incidents rather than at the climax of the research: patients in an investigation of medical analysis could perish in a traffic accident. Such censoring is called **random censoring**. It continually happens alongside type 1 censoring.

(vii) An amalgamation of Type-I and type-II births a **hybrid censoring scheme**.

(a) **Type-I hybrid**- The investigator fixes the termination time of the test $T = \min(X_{(R)}, \tau)$ and in this case far fewer than r failures may be observed upto the pre-fixed time.

(b) **Type-II hybrid**- Termination time $T = \max(X_{(R)}, \tau)$ is unknown to the investigator though this type guarantees more than r failures are observed.

Left censoring - This is slightly much more encountered: some participants have come across the incidents before comprehensive examination commences. Hence, the moment their event happens fall within the interlude $(0, T)$. It is important to note that the possibility of right and left censoring occurring in the same statistical set is very high.

Interval censoring - It materializes at that time when there lacks clear information on the occurrence of events and the only know-how is that the incident happened at a particular time interlude (T_1, T_2) . This form of censoring frequently emerges when observations are exceptional, and the incident has occurred amidst two noticeable moments.

The Expectation-Maximization (EM) algorithm; It's a repetitive procedure for calculating Maximum Likelihood Estimators. The procedure is applicable in situations where the observed data seemingly gets incomplete in some manner. Dempster et al (1977) introduced the procedure which is being widely applied in various lifetime distributions.

An EM algorithm constitutes of two paces:

The Expectation step (E-Step); over here, a prediction of the total figures log-likelihood contingent on the surveyed statistics is calculated.

The Maximization step (M-step); in this level the presupposition of the whole statistics likelihood attained in the E-step is maximized to find a new parameter estimate (improved estimate).

Newton-Raphson Algorithm The Newton-Raphson iteration is frequently used in statistics in maximum-likelihood estimation. The probability (or density) of the observed data which is viewed as a corollary of the variable is known as likelihood function.

Let θ be a scalar parameter, $l(\theta, y)$ the log-likelihood function. An assumption is made here that $l(\theta, y)$ is twice-differentiated with regard to θ .

Define the total function $U(\theta, y) = \frac{\delta l(\theta, y)}{\delta \theta}$

The observed information $I(\theta, y) = -\frac{\delta^2 l(\theta, y)}{\delta \theta^2}$.

To find the maximum likelihood estimate, the equation $U(\theta, y) = 0$ and the procedure below implemented.

Suppose that θ is 1-dimensional.

- (i) Select a primary estimate, $\theta^{(0)}$
- (ii) Determine $\theta^{(i)} = \theta^{(i-1)} + \frac{U(\theta^{(i-1)}; x)}{I(\theta^{(i-1)}; x)}$
- (iii) Repeat (ii) until convergence at some iteration k, i.e. where $U(\theta^{(k)}; y)$

1.3 Statement of the Problem

The field of survival analysis has greatly advanced over time with a number of statistical methods being employed to obtain the likelihood functions and compute the MLEs based on various censoring schemes for data drawn from different lifetime distributions. Newton Raphson and EM algorithm are some of the methods being widely used. In an experiment involving life testing as well as reliability studies, complete information may not be obtained always for all the units on failure times. Hence, data obtained from this type of experiments are known as censored data. In censoring, cost effectiveness and minimizing the total time for testing is very crucial.

The Type-I and Type-II besides a hybrid censoring scheme have an orthodox disadvantage. They do not permit the withdrawal of units occasionally except at the end of the investigation. Consequently, the three schemes do not help minimize the total time taken for testing and the related cost.

However, there exist many scenarios in life testing as well as in reliability experiments when

this allowance is desirable and inevitable particularly when the items under test are very difficult to obtain or very expensive. Because the three censoring schemes have this disadvantage of not allowing the units to be withdrawn at subsequent paces, progressive type- II right censoring scheme was floated which appears more general as well as flexible. See Balakrishnan (2007) and Balakrishnan et al. (2000) as for the comprehensive probe into the written work on progressive censoring.

As much as progressive censoring has this flexibility to remove the elements from the test at subsequent steps before termination of the test, it has turned out to be an inefficient scheme in some experiments which take much longer time before the required sample size is realized. Consequently, Kundu and Joarder (2006) suggested a fresh censoring scheme mentioned as progressive type-II hybrid censoring scheme. Experiment involving viability investigative units within this scheme is aborted at a time T specified beforehand.

Yongming Ma and Yimin Shi (2013) studied inferences in regard to Lomax distribution build onto type-II progressively hybrid censored statistics. However, they attained MLEs of the variables through normal iterative procedure which are relatively not robust against the initial values.

In this work, we shall obtain the MLEs of the parameters for data drawn from Lomax distribution build onto the progressive type-II hybrid censoring scheme via EM procedures. An EM algorithm is used because it has been observed to be a very effective procedure when dealing with data sets that are incomplete and have stray values, or dummies with distributions that are truncated, Dempster et al (1977). The MLEs obtained using EM algorithm are also relatively robust as they are not sensitive to their initial values like those obtained using normal iterative procedures, Ng et al (2002).

1.4 Justification of the Study

Censoring is a common feature of survival data. It has many types which can be used in analysis of different kinds of data representing various real life circumstances.

Experiments involving life testing as well as reliability studies are very demanding and the desire of the experimenter is to use a censoring scheme which is impartial to both the several elements used in the test and the total time spent for the same as well as the the effectiveness of the statistical deductions grounded on outcomes of the investigation.

In this project, use of progressive type- II hybrid censoring scheme over type-I, type -II, hybrid and progressive censoring schemes is being employed because it favors experiments that require units to be dismantled at various stages of failure before an appropriate intended sample size is realized and also has pre-specified time within which the experiment can occur.

The obtained MLEs via EM Algorithm for parameters of Lomax distribution are more robust against the initial values as they are not sensitive to their initial values like those obtained using normal iterative procedures.

1.5 Objectives

1.5.1 General Objective

The main objective of this research is to obtain MLEs of the parameters of Lomax distribution based onto a progressively Type-II hybrid censored data using an EM algorithm.

1.5.2 Specific objectives

- (i) To obtain MLE's of the parameters of the Lomax distribution under a progressive Type-II hybrid censoring scheme.
- (ii) To employ EM algorithm to compute the MLEs of the Lomax distribution under progressive type-II hybrid censoring scheme.
- (iii) To compare performance of the obtained MLEs by way of an EM Algorithm against those obtained via NR algorithm with regard to their mean bias and MSE for distinct number of censoring schemes at fixed parameter values of θ and λ .

1.6 Significance of the Study

In this work , the paramount concern is to compute the MLEs of Lomax distribution grounded on progressive Type -II hybrid scheme. The EM procedure through simulation study on various data sets will show how the estimate values obtained converges faster to the true values. By using an example drawn from a real data set, the procedure also reveals its significance in solving problems in fields of study like medicine among others. To this extent, the progressive type-II hybrid scheme and EM procedure contributes extensively to work related to survival analysis in regard to estimation of parameters of Lomax distribution which has not been studied under this censoring scheme using EM method before.

1.7 Outline of the project

In Chapter Two, a review of censoring schemes and work done on Lomax distribution is presented. In Chapter Three, the likelihood functions for the Lomax distribution are obtained.

The MLEs via EM and NR algorithms for Lomax distribution build onto progressive type-II hybrid censoring scheme is also demonstrated.

Chapter Four consists of results and discussion and Finally in Chapter Five, summary, conclusion, and recommendations are made.

CHAPTER TWO

2 LITERATURE REVIEW

2.1 Chapter Overview

Within this chapter, reconsideration on progressive censoring and progressive hybrid censoring schemes is presented. Lomax distribution and related work on the distribution is also presented.

2.2 Hybrid Censoring Scheme

Epstein (1954) started a hybrid censoring scheme, an amalgamation of type-I and type-II censoring schemes. These two types, I and II have been discussed extensively by various writers including Harter (1970), Mann et al. (1974), Cohen and Whitten (1988), Cohen (1963,1966,1991), Bain (1978) Bain and Engelhardt (1991). These authors have all attempted to consider lifetime studies in actuarial as well as industrial contexts, for parametric hand in hand with non-parametric cases.

Within Type I censoring, time T is pre-specified not subject to the failure times to the extend that no failures will be seen beyond this time, in other words at time T the experiment is terminated. See Cohen (1963 and 1966). In Type II censoring, several failures detected are definite say for example $m(m \leq n)$, such that at the juncture of the m^{th} abortion, experiment ceases with $n - m$ comparatively discerned abortion schedules being left. Within this case, the test time is unspecified meaning the distribution will simply be the distribution of the m^{th} order statistics coming from a prototype of size n drawn out of an assortment.

Hybrid censoring schemes further is designated as; Type-I hybrid and Type-II hybrid. A description is presented on the two types as follows:

Let us consider n items and indicate the prescribed failure times of the items as $X_{(1)} < X_{(2)} < \dots < X_{(n)}$. In type-I hybrid censoring, an investigation continues until a time τ specified earlier

is reached or a number $R < n$ specified beforehand fails. Thus a test ceases at an arbitrary time $T = \min(X_{(R)}, \tau)$. The termination time at this point is at most T . The work by Bartholomew (1963), Barlow et al (1968), and Chen et al (1988) really is a motivating factor since they considered type-I hybrid censored data drawn from an exponential distribution and provided an accurate less high confidence band for a parameter. Attributes on type-I hybrid censoring, check Kundu and Gupta (1988), Ebrahimi (1992) and Jeong et al. (1996) among others.

Similar to standard type I censoring, principal impediment due to type-I hybrid censoring remains that only few failures might be happening up to the pre-specified time. Within type II hybrid censoring, the point of time of closure of the test $T = \max(X_{(R)}, \tau)$ provides at least R failures being detected as far as the ending time. It is perceived that by taking $R = n$ and $\tau \rightarrow \infty$ the type I and type II forms a unique case of hybrid censoring scheme.

Hybrid censoring has become more accepted and many authors have discussed statistical inference for various distributions under hybrid censoring scheme in the reliability literature; for example, Draper and Guttman (1987), Childs et al. (2003), Kundu (2007), Banerjee et al. (2008), Pradhan et al (2009), Dube et al. (2010), Balakrishnan et al (2013) and Asgharzadeh et al (2013) among others.

2.3 Progressive Type-II Censoring Scheme

Type-1 and type-II are the ordinary censoring schemes. The Type-I makes provision for the experiments on life tests to terminate at a time T specified beforehand while Type-II censoring the experiments on life tests terminates immediately upon the $r - th$ abortion (r specified in advance). The type-I and type-II mixed together just as mentioned earlier results to a hybrid censoring scheme.

However, their flexibility to enable an element to be withdrawn occasionally except at the end of the test is not achievable. Hence, a censoring scheme widely termed Progressive Type II

right censoring was floated and so far quite extensive work on this has been undertaken on how to calculate the estimates of the Parameters for various lifetime distributions by several writers. see Asgharzadeh and Balakrishnan (2005), Gupta and Kundu (1996), Ng et al. (2002 and 2014), Childs et al. (2003), Soliman (2005), Wu (2003), Kus (2007), Childs and Balakrishnan (2000), Sarhan et al (2008) and Wu et al (2017) amongst other writers.

2.4 Progressive Type-II Hybrid Censoring Scheme

Even though supposition on progressive Type-II right censored trials has continued to be examined in the past researches quite extensively for a while, Type-II progressive censoring scheme owns one major problem; the time taken to perform the experiment can be very immense.

As a result of this, Kundu et al (2006) pioneered a fresh censoring scheme known as a Progressive type-II Hybrid Censoring, an amalgamation of type-II progressive and hybrid censoring schemes. The progressive Type-II hybrid censoring scheme allows for the life investigating tests to conclude at a time T specified earlier.

Extensive study in regard to progressive Type-II hybrid censoring and its significance, allude to Kundu et al (2006) and Childs et al (2007).

This progressively hybrid scheme in the last few years has also become more accepted in studying reliability and life-investigating tests. Mokhtari et al (2011) carried out a deduction on Weibull distribution beneath progressively Type-II hybrid censored data. Also Lin et al. (2011) estimated parameters of Generalized Rayleigh distribution in regard to progressive Type-II hybrid censored data. Salem and Abo-Kasem (2011) considered approximation of parameters of Exponential - Weibull distribution based on progressive type-II hybrid censored prototypes.

Recently, Ma Yongming and Shi Yimin (2013) studied the inference of Lomax distribution

grounded onto progressive type-II hybrid censoring scheme. They evaluated estimates due to the parameters employing the maximum likelihood technique and made a comparison with those obtained using the Bayesian approaches.

For more literature and outcome in regard to progressive type-II hybrid censoring schemes; refer to Hemmati and Khorram (2013), Bhattacharya et al (2014), Cho et al (2015), Chan et al (2015), Li and Huang (2011), Gorny and Cramer (2016) among others.

2.5 Lomax Distribution

Lomax (1954) introduced the distribution as a framework for failing data drawn from businesses. Since then, Lomax distribution continues to be widely utilized for life testing and modelling in reliability ; check Balkema et al. (1974) Raqab et al (2010). The distribution forms part of the diminishing non-fulfilling rate distributions, Chahkandi et al. (2009).

In addition, Lomax distribution is observed as a strong option for many other lifetime distributions for example exponential, Weibull, or gamma distributions especially if the considered population under test by the experimenter comes from a heavy-tailed distribution, Bryson (1974).

Giles et al., (2013) studied the bias of maximum likelihood estimator for the two parameters of Lomax distribution.

Ahsanulla (1991) analyzed the annals values of Lomax distribution.

Ahsanulla and Balakrishnan (1994) pioneered the repetitiveness relations across the moments of record values from Lomax distribution.

speculative affairs on Lomax distribution constructed from censored statistics are forwarded by Childs et al. (2001) whereby he examined the sequence facts and figures from non-homogeneous right-truncated Lomax distribution. Howlader et al., (2002) tabled estimates on the survival function of the Lomax distribution using Bayesian method.

Ghitany et al., (2007) reviewed Marshall-Olkin undertaking and stretched this to Lomax distribution as well as its relevance to censored data.

Asgharzadeh and Valiollahi (2011) obtained the Bayesian estimator of the scale parameter of the Lomax distribution with regard to progressive type-II censoring.

Raqab and Madi (2011) made Inference for the generalized Rayleigh distribution based on Progressively censored data. Cramer and Schmiedt (2011) examined the most appropriate censoring scheme for computing the parameters of the Lomax distribution by employing progressive type-II censoring.

Panahi and Asadi (2011), discussed deductions on stress-strength model for a Lomax distribution.

Helu et al (2013 and 2015) evaluated the parameters of Lomax distribution build on progressive censoring applying EM algorithm.

Bander and Marshall (2013) presented an approach of estimating the reliability function for Lomax distribution based on general progressive censoring scheme.

Recently, Yongming Ma and Yimin Shi (2013) also considered parameter estimation for Lomax distribution build onto type-II progressively hybrid censored data using ordinary MIEs. Therefore, as far as we know, no investigation has been carried out on the maximum likelihood estimation of parameters of Lomax distribution founded on progressive Type-II hybrid censoring scheme via EM Algorithm, hence our main objective in this study.

CHAPTER THREE

3 LIKELIHOOD FUNCTIONS AND PROPOSED MLEs VIA EM ALGORITHM

3.1 Chapter overview

In this chapter, an analysis of progressive type II hybrid censoring and derivation procedures for parameter estimator for Lomax distribution is carried out.

3.2 Progressive Type II Hybrid Censoring

Assume n independent identical components are set under an investigation . An integer $m < n$ is definite at the commencement of the test. Fixed also in advance is the time point T and m specified integers beforehand that is R_1, R_2, \dots, R_m satisfying $R_1 + R_2 + \dots + R_m + m = n$.

At the point of moment the initial failures, $X_{1:m:n}$ occurs, R_1 of the existing elements is haphazardly taken away.

Correspondingly, at the point of moment of the subsequent failures, the $X_{2:m:n}$, R_2 of the existing elements is taken away and so forth.

Whenever the m^{th} failure, $X_{m:m:n}$ happens earlier than the time phase T , an investigation terminates at this phase of time $X_{m:m;n}$. Alternatively, whenever the m^{th} failure does not happen earlier than this point of moment T and exclusively J failure occurs earlier than time point T , where $0 \leq J \leq T$, it follows that at this moment of time T the entire surviving $R_J^* = n - (R_1 + R_2 + \dots + R_J) - J$ elements are taken away with the investigation stopping at this point of moment T .

The two predicaments are identified as Case I and Case II . Thus , within an existence of progressively Type-II hybrid censoring schemes, two kinds of discernment are made as indicated below ;

CASE I: $X_{1:m;n}, \dots, X_{m:m;n}$ *if* $X_{m:m;n} < T$.

CASE II: $X_{1:m;n}, \dots, X_{J:m;n}$ *if* $X_{J:m;n} < T < X_{J+1:m;n}$.

See Kundu and Joarder (2006).

3.3 Model Description

Let D be the number of failures, i.e $D = m$, for Case I and $D = J$ for Case II. Let $X_{1;m;n}, X_{2;m;n}, \dots, X_{D;m;n}$ represent the progressive type -II hybrid right censored statistics from a population with probability density function (pdf), cumulative distribution function (cdf) and survival function be illustrated appropriately as shown:

$$f(x; \lambda, \theta) = \lambda\theta(1 + \theta x)^{-\lambda-1}, x > 0, \lambda, \theta > 0 \quad (3.1)$$

$$F(x; \lambda, \theta) = 1 - (1 + \theta x)^{-\lambda}, x > 0 \quad (3.2)$$

$$S(x) = (1 - \theta x)^{-\lambda}, x > 0, \lambda, \theta > 0 \quad (3.3)$$

The corresponding likelihood functions will be specified as:

$$L(x; \theta, \lambda,) = \prod_{i=1}^n f(x_i; \lambda, \theta)$$

$$L(x; \theta, \lambda,) = (\lambda\theta)^n \prod_{i=1}^n (1 + \theta x_i)^{-\lambda-1}$$

But based on this censoring scheme, the likelihood function takes this form:

$$L(x; \theta, \lambda) \propto \prod_{i=1}^n f(x_i; \lambda, \theta) [1 - F(x_i; \theta, \lambda)]^{R_i}$$

See Li and Huang (2011) as well as Raqab and Madi (2011).

CASE I, $D = m$

The likelihood function is designated as :

$$L(\theta, \lambda, x_{1;m:n} \dots x_{m;m:n}) \propto (\lambda\theta)^m \prod_{i=1}^m (1 + \theta x_{i;m:n})^{-\lambda-1} (1 + \theta x_{i;m:n})^{-\lambda R_i}$$

$$L(\theta, \lambda, x_{1;m:n} \dots x_{m;m:n}) \propto (\lambda\theta)^m \exp\left\{(-\lambda - 1) \sum_{i=1}^m \ln(1 + \theta x_{i;m:n}) + \lambda \sum_{i=1}^m -R_i \ln(1 + \theta x_{i;m:n})\right\} \quad (3.4)$$

we acquire the log-likelihood function as follows;

$$l(\theta, \lambda, x_{1;m:n} \dots x_{m;m:n}) \propto m \ln(\lambda\theta) - (\lambda + 1) \sum_{i=1}^m \ln(1 + \theta x_{i;m:n}) + \lambda \sum_{i=1}^m -R_i \ln(1 + \theta x_{i;m:n}) \quad (3.5)$$

CASE II, $D = J$

The likelihood function is stated as;

$$L(\theta, \lambda, x_{i;m:n} \dots x_{J;m:n}) \propto (\lambda\theta)^J \exp\left\{(-\lambda - 1) \sum_{i=1}^J \ln(1 + \theta x_{i;m:n}) + \lambda \sum_{i=1}^J -R_i \ln(1 + \theta x_{i;m:n}) - R_j^* \ln(1 + \theta T)\right\} \quad (3.6)$$

and its corresponding log-likelihood function designated as;

$$l(\theta, \lambda, x_{i;m:n} \dots x_{J;m:n}) \propto J \ln(\lambda\theta) - (\lambda+1) \sum_{i=1}^J \ln(1+\theta x_{i;m:n}) + \lambda \sum_{i=1}^J -R_i \ln(1+\theta x_{i;m:n}) - R_J^* \ln(1+\theta T) \quad (3.7)$$

The log-likelihood functions in equations (3.5) and (3.7) is amalgamated as shown below;

$$l(\theta, \lambda, x_i) \propto D \ln(\lambda\theta) - (\lambda + 1)D_\theta(y) + \lambda T_\theta(y) \quad (3.8)$$

where

$$D = m, \quad D_\theta(y) = \sum_{i=1}^m \ln(1 + \theta x_{i;m:n}), \quad T_\theta(y) = \sum_{i=1}^m -R_i \ln(1 + \theta x_{i;m:n})$$

for case I.

and

$$D = J, \quad D_\theta(y) = \sum_{i=1}^J \ln(1 + \theta x_{i;m:n}), \quad T_\theta(y) = \sum_{i=1}^J -R_i \ln(1 + \theta x_{i;m:n}) - R_J^* \ln(1 + \theta T)$$

for case II.

3.4 NR Algorithm

In line with this method, MLEs of the parameters θ and λ are obtained by differentiating equation (3.5) and (3.7) in regard to θ and λ then equating the normal equations to 0 as illustrated;

CASE I

$$\frac{\partial l}{\partial \lambda} = \frac{m}{\lambda} - \sum_{i=1}^m \ln(1 + \theta x_{i:m:n}) - \sum_{i=1}^m R_i \ln(1 + \theta x_{i:m:n}) = 0$$

This implies that

$$\hat{\lambda} = \lambda(\theta) = \frac{m}{\sum_{i=1}^m \ln(1 + \theta x_{i:m:n}) + \sum_{i=1}^m R_i \ln(1 + \theta x_{i:m:n})} \quad (3.9)$$

similarly:

$$\frac{\partial l}{\partial \theta} = \frac{m}{\theta} - (\lambda + 1) \sum_{i=1}^m \frac{x_{i:m:n}}{1 + \theta x_{i:m:n}} - \lambda \sum_{i=1}^m \frac{R_i x_{i:m:n}}{1 + \theta x_{i:m:n}} = 0 \quad (3.10)$$

Plugging in equation (3.9) into equation (3.10), the equation will be reduced to a one dimensional non-linear normal equation of θ . i.e

$$\frac{m}{\theta} - (\lambda(\theta) + 1) \sum_{i=1}^m \frac{x_{i:m:n}}{1 + \theta x_{i:m:n}} - \lambda(\theta) \sum_{i=1}^m \frac{R_i x_{i:m:n}}{1 + \theta x_{i:m:n}} = 0 \quad (3.11)$$

We employ NR algorithm to work out the equation (3.11) and obtain the MLE of θ . To achieve this, we implement the following procedure;

let

$$U(\theta, x) = \frac{m}{\theta} - (\lambda(\theta) + 1) \sum_{i=1}^m \frac{x_{i:m:n}}{1 + \theta x_{i:m:n}} - \lambda(\theta) \sum_{i=1}^m \frac{R_i x_{i:m:n}}{1 + \theta x_{i:m:n}}$$

$$U'(\theta, x) = \frac{-m}{\theta^2} - \lambda'(\theta) \sum_{i=1}^m \frac{x_{i:m:n}}{1 + \theta x_{i:m:n}} - (\lambda(\theta) + 1) \sum_{i=1}^m \frac{x_{i:m:n}^2}{(1 + \theta x_{i:m:n})^2} - \lambda'(\theta) \sum_{i=1}^m \frac{R_i x_{i:m:n}}{1 + \theta x_{i:m:n}}$$

$$-\lambda(\theta) \sum_{i=1}^m \frac{R_i x_{i:m:n}^2}{(1 + \theta x_{i:m:n})^2}$$

$$I(\theta, x) = \frac{m}{\theta^2} + \lambda'(\theta) \sum_{i=1}^m \frac{x_{i:m:n}}{1 + \theta x_{i:m:n}} + (\lambda(\theta) + 1) \sum_{i=1}^m \frac{x_{i:m:n}^2}{(1 + \theta x_{i:m:n})^2} + \lambda'(\theta) \sum_{i=1}^m \frac{R_i x_{i:m:n}}{1 + \theta x_{i:m:n}} \\ + \lambda(\theta) \sum_{i=1}^m \frac{R_i x_{i:m:n}^2}{(1 + \theta x_{i:m:n})^2} \quad (3.12)$$

where

$$\lambda'(\theta) = m \left\{ \sum_{i=1}^m \frac{x_{i:m:n}}{1 + \theta x_{i:m:n}} + \sum_{i=1}^m \frac{R_i x_{i:m:n}}{1 + \theta x_{i:m:n}} \right\}^{-2}$$

Next we obtain the MLE of θ by computing:

$$\theta^{(i)} = \theta^{(i-1)} + \frac{U(\theta^{(i-1)}; x)}{I(\theta^{(i-1)}; x)} \quad (3.13)$$

until there's no useful change.

The value obtained is then inserted in equation (3.9) to get the MLE of λ .

CASE II

Within this case, MLEs of the parameters θ and λ are attained by differentiating equation (3.7) in regard to θ and λ then equating the normal equations to 0 as shown;

$$\frac{\partial l}{\partial \lambda} = \frac{J}{\lambda} - \sum_{i=1}^J \ln(1 + \theta x_{i:m:n}) - \sum_{i=1}^J R_i \ln(1 + \theta x_{i:m:n}) = 0$$

This implies therefore

$$\widehat{\lambda} = \lambda(\theta) = \frac{J}{\sum_{i=1}^J \ln(1 + \theta x_{i:m:n}) + \sum_{i=1}^J R_i \ln(1 + \theta x_{i:m:n})} \quad (3.14)$$

similarly:

$$\frac{\partial l}{\partial \theta} = \frac{J}{\theta} - (\lambda + 1) \sum_{i=1}^J \frac{x_{i:m:n}}{1 + \theta x_{i:m:n}} - \lambda \sum_{i=1}^J \frac{R_i x_{i:m:n}}{1 + \theta x_{i:m:n}} - \frac{R_J^* T}{1 + \theta T} = 0 \quad (3.15)$$

Plugging in equation (3.14) into equation (3.15), the equation will be reduced to a one dimensional non-linear normal equation of θ . i.e

$$\frac{J}{\theta} - (\lambda(\theta) + 1) \sum_{i=1}^J \frac{x_{i:m:n}}{1 + \theta x_{i:m:n}} - \lambda(\theta) \sum_{i=1}^J \frac{R_i x_{i:m:n}}{1 + \theta x_{i:m:n}} - \frac{R_J^* T}{1 + \theta T} = 0 \quad (3.16)$$

Using NR procedure, we obtain;

$$U(\theta, x) = \frac{J}{\theta} - (\lambda(\theta) + 1) \sum_{i=1}^J \frac{x_{i:m:n}}{1 + \theta x_{i:m:n}} - \lambda(\theta) \sum_{i=1}^J \frac{R_i x_{i:m:n}}{1 + \theta x_{i:m:n}} - \frac{R_J^* T}{1 + \theta T}$$

$$U'(\theta, x) = \frac{-J}{\theta^2} - \lambda'(\theta) \sum_{i=1}^J \frac{x_{i:m:n}}{1 + \theta x_{i:m:n}} - (\lambda(\theta) + 1) \sum_{i=1}^J \frac{x_{i:m:n}^2}{(1 + \theta x_{i:m:n})^2} - \lambda'(\theta) \sum_{i=1}^J \frac{R_i x_{i:m:n}}{1 + \theta x_{i:m:n}}$$

$$\begin{aligned}
& -\lambda(\theta) \sum_{i=1}^J \frac{R_i x_{i:m:n}^2}{(1 + \theta x_{i:m:n})^2} + \frac{R_J T^2}{(1 + \theta T)^2} \\
I(\theta, x) = & \frac{J}{\theta^2} + \lambda'(\theta) \sum_{i=1}^J \frac{x_{i:m:n}}{1 + \theta x_{i:m:n}} + (\lambda(\theta) + 1) \sum_{i=1}^J \frac{x_{i:m:n}^2}{(1 + \theta x_{i:m:n})^2} + \lambda'(\theta) \sum_{i=1}^J \frac{R_i x_{i:m:n}}{1 + \theta x_{i:m:n}} \\
& + \lambda(\theta) \sum_{i=1}^J \frac{R_i x_{i:m:n}^2}{(1 + \theta x_{i:m:n})^2} - \frac{R_J T^2}{(1 + \theta T)^2} \tag{3.17}
\end{aligned}$$

where

$$\lambda'(\theta) = J \left\{ \sum_{i=1}^J \frac{x_{i:m:n}}{1 + \theta x_{i:m:n}} + \sum_{i=1}^J \frac{R_i x_{i:m:n}}{1 + \theta x_{i:m:n}} \right\}^{-2}$$

The MLE of θ is then obtained the same way as in equation (3.13).

3.5 EM Algorithm

In(1977), Dempster et al launched EM algorithm which was later used by Ng et al (2002). Yamaguchi and Watanabe (2004) also carried out an experiment to compute MLEs based on multidimensional statistics which had figures lacking employing NR and EM algorithm. Most recently, EM algorithm was used to obtain maximum likelihood estimators of the parameters of Generalized Rayleigh distribution based on progressive type-II hybrid censoring scheme. See Li and Lina (2015).

Within this work, we employ EM algorithm procedure to obtain MLEs of all undetermined parameters as described below.

Assume the life variable trails the Lomax distribution (3.1). A viable investigative experiment is performed hinged on progressive type-II hybrid censoring. Then the complete failure times

of D items $X_{1:m:n}, X_{2:m:n}, \dots, X_{D:m:n}$ are noted. Hence, this statistics that has been viewed can simply be designated as $X = (X_{1:m:n}, X_{2:m:n}, \dots, X_{D:m:n})$.

We denote $Z = \{Z_{ij}, j = 1, 2, \dots, R_D; i = 1, 2, \dots, D\} \cup \{Z_{Tj}, j = 1, 2, \dots, R_D^*\}$, where $\{Z_{ij}, j = 1, 2, \dots, R_D; i = 1, 2, \dots, D\}$ represents the j^{th} censored elements at the abortion time $X_{i:m:n}$, and $\{Z_{Tj}, j = 1, 2, \dots, R_D^*\}$ indicates the j^{th} censored elements at the cut time T . Subsequently all $Z = \{Z_{ij}, j = 1, 2, \dots, R_D; i = 1, 2, \dots, D\} \cup \{Z_{Tj}, j = 1, 2, \dots, R_D^*\}$ are unobserved. Consequently, we can denote exclusively the missing integers as Z and the entire statistics as $W = (X, Z)$.

Thus, the log-likelihood function for Case I and Case II constructed onto W is recorded as shown below,

$$H(w; \theta, \lambda) \propto n \ln \lambda + n \ln \theta - (\lambda + 1) \sum_j^m \ln(1 + \theta x_{j;m:n}) - (\lambda + 1) \sum_j^m \sum_k^{R_j} \ln(1 + \theta z_j) \quad (3.18)$$

and

$$H(w; \theta, \lambda) \propto n \ln \lambda + n \ln \theta - (\lambda + 1) \sum_j^J \ln(1 + \theta x_{j;m:n}) - (\lambda + 1) \sum_j^J \sum_k^{R_j} \ln(1 + \theta z_j) + \lambda \sum_j^{R_j^*} \ln(1 + \theta z_j) \quad (3.19)$$

Within an E-step, it necessitates that one determines the pseudo-likelihood function which is obtained from $H(w; \theta, \lambda)$ by substituting any function of Z_{jk} (say $g(z_{jk})$) with $E\{g(z_j/z_{jk} > x_{j;m:n})\}$, and $g(Z_{TK})$ by $E\{g(Z_{TK})/Z_{TK} > T\}$.

Therefore the pseudo-likelihood functions for the two cases are

$$H^*(w; \theta, \lambda) \propto n \ln \lambda + n \ln \theta - (\lambda + 1) \sum_{j=1}^m \ln(1 + \theta x_{j;m:n}) - (\lambda + 1) \sum_{j=1}^m \sum_{k=1}^{R_j} E\{\ln(1 + \theta z_j)/z_j > x_{j;m:n}\} \quad (3.20)$$

and

$$\begin{aligned}
H^*(w; \theta, \lambda) \propto n \ln \lambda + n \ln \theta - (\lambda + 1) \sum_j^J \ln(1 + \theta x_{j:m:n}) - (\lambda + 1) \sum_{j=1}^J \sum_{k=1}^{R_j} E\{\ln(1 + \theta z_j) / z_j > x_{j:m:n}\} \\
+ \lambda \sum_{j=1}^{R_j^*} E\{\ln(1 + \theta z_{T_k}) / z_{T_k} > T\}
\end{aligned} \tag{3.21}$$

respectively.

Considering $X_{j:m:n} = x_{j:m:n}$, the contingent distribution of Z_{jk} follows a truncated Lomax distribution accompanied by left truncation at, $x_{j:m:n}$.

Implying,

$$f_{z/y}(z_j/y) = f_w(z_j) / [1 - F_w(x_{j:m:n})], z_j > x_{j:m:n} \tag{3.22}$$

see Ng et al (2002).

Therefore, the last expressions within the pseudo-likelihood functions are estimated as shown:

From equations (3.1) and (3.2), we obtain:

$$f(z_j; \lambda, \theta) = \lambda \theta (1 + z_j)^{-\lambda-1}, 0 < z_j < 1, \lambda, \theta > 0$$

$$F(x_j; \lambda, \theta) = 1 - (1 + \theta x_j)^{-\lambda}, x_j > 0$$

$$1 - F(x_j; \lambda, \theta) = 1 - \{1 - (1 + \theta x_j)^{-\lambda}\} = (1 + \theta x_j)^{-\lambda}$$

so that

$$f(z_j/z_j = x_j; \theta, \lambda) = \frac{\lambda \theta (1 + z_j)^{-\lambda-1}}{(1 + \theta x_j)^{-\lambda}}, z_j > x_j \tag{3.23}$$

Then the expectation in equations (3.20) and (3.21) can be obtained as follows;

$$E[\ln(1 + \theta z_j) / z_j > x_{j:m:n}] = \int_{x_j}^1 \frac{\lambda \theta (1 + z_j)^{-\lambda-1}}{(1 + \theta x_j)^{-\lambda}} \ln(1 + \theta z_j) dz_j$$

$$E[\ln(1 + \theta z_j)/z_j > x_{j:m:n}] = \frac{\lambda\theta}{(1 + \theta x_j)^{-\lambda}} \int_{x_j}^1 (1 + z_j)^{-\lambda-1} \ln(1 + \theta z_j) dz_j \quad (3.24)$$

Next we integrate the integral in equation (3.24) by parts:

$$\text{let } u = \ln(1 + \theta z_j) \quad \text{and} \quad du = \frac{\theta}{1 + \theta z_j} dz_j$$

and

$$dv = (1 + z_j)^{-\lambda-1} dz_j \quad \text{and} \quad v = -\frac{1}{\lambda}(1 + z_j)^{-\lambda}$$

Using $\int u dv = uv - v \int du$ and substituting the limits, the integral in equation (3.24) is obtained as follows;

$$\int_{x_j}^1 (1 + z_j)^{-\lambda-1} \ln(1 + \theta z_j) dz_j = -\frac{1}{\lambda}(1 + x_j)^{-\lambda} \ln(1 + \theta x_j) + \frac{1}{\lambda}(1 + x_j)^{-\lambda} \int_{x_j}^1 \frac{\theta}{1 + \theta z_j} dz_j$$

$$\int_{x_j}^1 (1 + z_j)^{-\lambda-1} \ln(1 + \theta z_j) dz_j = -\frac{1}{\lambda}(1 + x_j)^{-\lambda} \ln(1 + \theta x_j) + \frac{1 - x_j}{\lambda}(1 + x_j)^{-\lambda} \ln(1 + \theta x_j)$$

which results to:

$$\int_{x_j}^1 (1 + z_j)^{-\lambda-1} \ln(1 + \theta z_j) dz_j = -\frac{x_j}{\lambda}(1 + x_j)^{-\lambda} \ln(1 + \theta x_j). \quad (3.25)$$

substituting equation (3.25) in (3.24) we obtain;

$$E[\ln(1 + \theta z_j)/z_j > x_{j:m:n}] = -\frac{\lambda \theta x_j}{(\lambda)(1 + \theta x_j)^{-\lambda}}(1 + x_j)^{-\lambda} \ln(1 + \theta x_j) \quad (3.26)$$

hence

$$E(v; \theta, \lambda) = E[\ln(1 + \theta z_j)/z_j > v] = -\frac{\lambda \theta v}{(\lambda)(1 + \theta v)^{-\lambda}}(1 + v)^{-\lambda} \ln(1 + \theta v) \quad (3.27)$$

The M-step necessitates an augmentation of the pseudo-likelihood function by substituting $E(V; \theta, \lambda)$ in equations (3.20) and (3.21). Assume that at the k^{th} stage, the estimates of (θ, λ) are $(\theta^{(k)}, \lambda^{(k)})$, then $(\theta^{(k+1)}, \lambda^{(k+1)})$ for the two cases can be obtained by maximizing equations (3.20) and (3.21) with regard to θ and λ appropriately as apprehended;

$$H^*(w; \theta, \lambda) \propto n \ln \lambda + n \ln \theta - (\lambda + 1) \sum_{i=1}^m \ln(1 + \theta x_{j:m:n}) - (\lambda + 1) \sum_{i=1}^m R_j E(x_{j:m:n}; \theta^{(k)}, \lambda^{(k)}) \quad (3.28)$$

and

$$\begin{aligned} H^*(w; \theta, \lambda) \propto n \ln \lambda + n \ln \theta - (\lambda + 1) \sum_{i=1}^J \ln(1 + \theta x_{j:m:n}) - (\lambda + 1) \sum_{i=1}^J R_j E(x_{j:m:n}; \theta^{(k)}, \lambda^{(k)}) \\ + \lambda R_j^* E(T; \theta^{(k)}, \lambda^{(k)}) \end{aligned} \quad (3.29)$$

Then the estimates of θ and λ can be obtained respectively as follows;

Case I:

$$\frac{\partial H^*(w; \theta, \lambda)}{\partial \theta} = \frac{n}{\theta} - (\lambda + 1) \sum_{i=1}^m \frac{x_{j:m:n}}{1 + \theta x_{j:m:n}} = 0$$

$$\hat{\theta}(\lambda) = \frac{n}{(\lambda + 1) \sum_{i=1}^m \frac{x_{j:m:n}}{1 + \theta x_{j:m:n}}} \quad (3.30)$$

$$\frac{\partial H^*(w; \hat{\theta}(\lambda), \lambda)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^m \ln(1 + \hat{\theta}(\lambda)x_{j:m:n}) - \sum_{j=1}^m R_j E(x_{j:m:n}, \theta^{(k)}, \lambda^{(k)}) = 0$$

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^m \ln(1 + \hat{\theta}(\lambda)x_{j:m:n}) - \sum_{j=1}^m R_j E(x_{j:m:n}, \theta^{(k)}, \lambda^{(k)})} \quad (3.31)$$

Case II:

$$\frac{\partial H^*(w; \theta, \lambda)}{\partial \theta} = \frac{n}{\theta} - (\lambda + 1) \sum_{i=1}^J \frac{x_{j:m:n}}{1 + \theta x_{j:m:n}} = 0$$

$$\hat{\theta}(\lambda) = \frac{n}{(\lambda + 1) \sum_{i=1}^J \frac{x_{j:m:n}}{1 + \theta x_{j:m:n}}} \quad (3.32)$$

and

$$\frac{\partial H^*(w; \hat{\theta}(\lambda), \lambda)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{j=1}^J \ln(1 + \hat{\theta}(\lambda)x_{j:m:n}) - \sum_{j=1}^J R_j E(x_{j:m:n}, \theta^{(k)}, \lambda^{(k)}) + R_j^* E(T, \theta^{(k)}, \lambda^{(k)}) = 0$$

$$\hat{\lambda} = \frac{n}{\sum_{j=1}^J \ln(1 + \hat{\theta}(\lambda)x_{j:m:n}) + \sum_{j=1}^J R_j E(x_{j:m:n}, \theta^{(k)}, \lambda^{(k)}) - R_j^* E(T, \theta^{(k)}, \lambda^{(k)})} \quad (3.33)$$

In an instant $\lambda^{(k)}$ is secured, $\theta^{(k+1)}$ is obtained as $\theta^{(k+1)} = \hat{\theta}(\lambda^k)$

CHAPTER FOUR

4 RESULTS AND DISCUSSIONS

4.1 Chapter Overview

In this chapter, the study investigates the behavior of proposed MLEs via EM algorithm on simulated and real data in a comparison to the MLEs obtained via NR algorithm.

4.2 Description of the Study

Assume $X_i, i = 1, 2, \dots, n$ are n identically independently distributed prototypes created out of Lomax model (3.1) and episodes beneath progressive type-II hybrid censored data are examined . $\lambda = 1.5$ and $\theta = 0.8$ are considered as the true values of the parameters of the progressive type-II hybrid censored samples for Lomax distribution. In this work, samples of sizes 40 and 80 are used under different censoring schemes described below. See Wu et al (2017).

$$\text{scheme 1 } R_1 = n - m, R_2 = \dots = R_m = 0$$

$$\text{scheme 2 } R_1 = 0, R_2 = n - m, R_3 = \dots = R_m = 0$$

$$\text{scheme 3 } R_1 = \dots = R_{m-1} = 0, R_m = n - m$$

To be able to contrast the outcomes for distinct censoring schemes, different T values: $T_1 = X_{[\frac{m}{2}]:m:n}$, $T_2 = X_{[\frac{4m}{5}]:m:n}$, $T_3 = X_{m:m:n} + 2$ respectively, are taken into account, where $[x]$ indicates the integral part of a decisive number x .

To create a type-II progressive hybrid censored illustration from Lomax model (3.1), we employ the procedures suggested in Balakrishnan and Aggarwala (2000) and also used in Kundu and Joarder (2006), which necessitates a number of stages :

(i) Generate m independent and identically distributed (iid) random numbers $U_1; U_2; \dots; U_m$ from the standard uniform distribution $U[0; 1]$.

(ii) Set

$$Z_i = \frac{\left(\frac{U_i}{\lambda\theta}\right)^{-\frac{1}{\lambda+1}} - 1}{\theta},$$

so that Z_i 's are iid standard Lomax distribution variate.

(iii) Given $n; m$ and the censoring scheme $R = (R_1; R_2; \dots; R_m)$, let $Y_1 = Z_1 = m$ and for $i = 1; \dots; m$

$$Y_i = Y_{i-1} + \frac{Z_i}{n - \sum_{j=1}^{i-1} R_j - i + 1}$$

Then, we can secure a progressive type-II censored sample $(Y_1; Y_2; \dots; Y_m)$ that comes from usual Lomax distribution with censoring scheme $R = (R_1; R_2; \dots; R_m)$.

(iv) Set

$$W_i = 1 - \lambda\theta(1 + \theta Y_i)^{-\lambda+1}$$

so that W_i 's form a progressive type-II censored sample from uniform distribution $U[0; 1]$.

(v) Set

$$X_{i:m:n} = F^{-1}(W_i) = \frac{(1 - W_i)^{-\frac{1}{\lambda}} - 1}{\theta}$$

so that X_i 's form a type-II progressive censored sample from Lomax model (3.1), where $F(x)$ is its cdf.

(vi) If $X_{m:m:n} \leq T$, it follows that the commensurate progressive type-II hybrid censored sample of Lomax is the progressive type-II censored sample $\{(X_{1:m:n}; R_1), \dots, (X_{m:m:n}; R_m)\}$ and $D = m, R_D^* = 0$ in this case. If $X_{m:m:n} > T$, we can find J such that $X_{J:m:n} < T < X_{J+1:m:n}$, then the corresponding progressive type-II hybrid censored sample is

$\{(X_{1:m:n}; R_1), \dots, (X_{J:m:n}; R_J)\}$ and $D = J$, $R_D^* = R_J^*$ in this case, where R_j^* is the same as defined earlier.

Once the samples have been generated, we obtain the EM estimates by repeating the process $s=500$ times for different sample sizes $n = 40$ and $n = 80$ respectively. In this work, in each of the 500 times, the estimates are obtained when the absolute difference between the log-likelihood functions is less than 0.00001.

Denoting the parameter estimates of the k^{th} experiment as $\widehat{\Theta}^{(k)} = (\widehat{\lambda}^k, \widehat{\theta}^k)$, ($k = 1, \dots, s$), the final mean squared errors and bias of the estimates are given respectively by:

$$MSE(\Theta_j) = \frac{1}{s} \sum_{k=1}^s (\widehat{\Theta}_j^k - \Theta_j)^2$$

and

$$MeanBias(\Theta_j) = \frac{1}{s} \sum_{k=1}^s (\widehat{\Theta}_j^k - \Theta_j)$$

All the computational outcomes as per distinct figures of n, m, T with various censoring schemes were computed using R statistical software.

4.3 Numerical Outcomes and Analysis

Table 4.1: *The MLEs via NR and EM algorithms , mean squared error and mean bias for the parameters of Lomax distribution under censoring scheme 1 when $\lambda = 1.5$ and $\theta = 0.8$*

T	n	m	NR						EM					
			Estimated value		MSE		Mean Bias		Estimated value		MSE		Mean Bias	
			$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$
T_1	40	30	2.1184	1.5166	0.38242	0.51352	0.6184	0.7166	1.9265	1.3155	0.18190	0.26574	0.4265	0.5155
		36	1.9161	1.4972	0.17314	0.48609	0.4161	0.6972	1.7242	1.2961	0.05027	0.24612	0.2242	0.4961
	80	50	1.9044	1.2890	0.16354	0.23912	0.4044	0.4890	1.7125	1.0879	0.04516	0.08289	0.2125	0.2879
		70	1.8822	1.1997	0.14612	0.15976	0.3822	0.3997	1.69034	0.9986	0.03623	0.03944	0.19034	0.1986
T_2	40	30	2.0257	1.5104	0.27636	0.50467	0.5257	0.7104	1.8338	1.3093	0.11142	0.25939	0.3338	0.5093
		36	1.8727	1.4507	1.3891	0.42341	0.3727	0.6507	1.6808	1.2496	0.03269	0.20214	0.1808	0.4496
	80	50	1.7192	1.2015	0.04805	0.16120	0.2192	0.4015	1.5273	1.0004	0.00075	0.04016	0.0273	0.2004
		70	1.6998	1.1047	0.03992	0.09284	0.1998	0.3047	1.5079	0.9036	0.00006	0.01073	0.0079	0.1036
T_3	40	30	1.9155	1.4888	0.17264	0.47444	0.4155	0.6888	1.7236	1.2877	0.04999	0.23785	0.2236	0.4877
		36	1.8528	1.2769	0.12447	0.22743	0.3528	0.4769	1.6609	1.0758	0.02589	0.07607	0.1609	0.2758
	80	50	1.7093	1.1854	0.04381	0.14853	0.2093	0.3854	1.5174	0.9843	0.00030	0.03397	0.0174	0.1843
		70	1.6864	1.0559	0.03475	0.06548	0.1864	0.2559	1.4945	0.8548	0.00003	0.00300	-0.0055	0.0548

From Table 4.1 , we observe that:

The estimates obtained using EM algorithm and their corresponding MSEs and mean bias are generally smaller compared to those obtained via NR algorithm. The results displayed indicates that the estimates of λ and θ obtained diminishes as the prototype size n decreases for definite time point T_1 , T_2 , and T_3 .

It is also observed that estimates obtained under fixed time point T_3 approaches true value faster the than those obtained under the fixed time point T_1 and fixed time point T_2 . For fixed time point T , an increase in the prototype size n generally gives estimates that performs better in terms of reduced MSEs and mean bias values.

For a fixed prototype size n , when the number of the discerned failures m increases, the results indicates that the estimates for the proposed EM methodology performs generally well with regard to their reduced values of MSEs and mean bias compared to the estimates for NR method under all the fixed time point T_1 , T_2 and T_3 .

When the figures of the discerned failures m is definite at different point of moment T an increase in the prototype size n leads to reduced estimates that are by comparison the same but with those for the EM methodology giving smaller values that appears to converge faster to the true value of λ and θ .

Table 4.2: The MLEs via NR and EM algorithms , mean squared error and bias for the parameters of Lomax distribution under censoring scheme 2 when $\lambda = 1.5$ and $\theta = 0.8$

T	n	m	NR						EM					
			Estimated value		MSE		Mean Bias		Estimated value		MSE		Mean Bias	
			$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$
T_1	40	30	2.0943	1.5233	0.35319	0.52316	0.5943	0.7233	1.9024	1.3222	0.16193	0.27269	0.4024	0.5222
		36	1.915	1.4554	0.17223	0.42955	0.415	0.6554	1.7231	1.2543	0.04977	0.20639	0.2231	0.4543
	80	50	1.8809	1.4145	0.14509	0.37761	0.3809	0.6145	1.6789	1.2134	0.03572	0.17089	0.1890	0.4134
		70	1.8296	1.3437	0.10864	0.29561	0.3296	0.5437	1.6377	1.1426	0.01896	0.11737	0.1377	0.3426
T_2	40	30	2.0227	1.4999	0.27322	0.48986	0.5227	0.6999	1.8308	1.2988	0.10943	0.2488	0.3308	0.4988
		36	1.8273	1.4032	0.10713	0.36385	0.3273	0.6032	1.6354	1.2021	0.01833	0.16168	0.1354	0.4021
	80	50	1.7382	1.2634	0.05674	0.21474	0.2382	0.4634	1.5463	1.0623	0.00214	0.0688	0.0463	0.2623
		70	1.7057	1.126	0.04231	0.10628	0.2057	0.326	1.5138	0.9249	0.00019	0.0156	0.0138	0.1249
T_3	40	30	1.9158	1.3814	0.17289	0.33803	0.4158	0.5814	1.7239	1.1803	0.05013	0.14463	0.2239	0.3803
		36	1.7703	1.329	0.07306	0.27984	0.2703	0.529	1.5784	1.1279	0.00615	0.10752	0.0784	0.3279
	80	50	1.726	1.2442	0.05108	0.19731	0.2260	0.4442	1.5341	1.0433	0.00116	0.05909	0.0341	0.2431
		70	1.6017	1.1112	0.01034	0.09685	0.1017	0.3112	1.4098	0.9101	0.00814	0.01212	-0.0902	0.1101

Table 4.2 displays the performance of the two method under scheme 2. The obtained estimated values of parameters under EM methodology and their corresponding MSEs and mean bias are generally smaller compared to those obtained via NR algorithm.

From the table 4.2 it can clearly be observed that just like in scheme 1, when the sample size n increases under fixed time point T , the estimates decrease consistently with those under the proposed EM method appearing to converge faster to the true value.

For different time point T , when the number of discerned failures m is definite, an escalation in the sample size n generally indicate that the estimates for the proposed EM method performs well in regard to their reduced values of MSEs and mean biases.

Table 4.3: *The MLEs via NR and EM algorithms, mean squared error and bias for the parameters of Lomax distribution under censoring scheme 3 when $\lambda = 1.5$ and $\theta = 0.8$*

T	n	m	NR						EM					
			Estimated value		MSE		Mean Bias		Estimated value		MSE		Mean Bias	
			$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\theta}$
T_1	40	30	2.1453	1.5109	0.41641	0.50538	0.6453	0.7109	1.9534	1.3098	0.20557	0.25989	0.4534	0.5098
		36	1.9235	1.4517	0.17935	0.42471	0.4235	0.6517	1.7316	1.2506	0.05364	0.20304	0.2316	0.4506
	80	50	1.7837	1.3974	0.08049	0.35689	0.2837	0.5974	1.5918	1.1963	0.00843	0.15707	0.0918	0.3963
		70	1.7321	1.2386	0.05387	0.19237	0.2321	0.4386	1.5402	1.0375	0.00162	0.05641	0.0402	0.2375
T_2	40	30	2.0703	1.3693	0.32524	0.3241	0.5703	0.5693	1.8784	1.1682	0.14319	0.13557	0.3784	0.3682
		36	1.9027	1.3143	0.16217	0.2645	0.4027	0.5143	1.7108	1.1132	0.04444	0.09809	0.2108	0.3132
	80	50	1.7278	1.1908	0.05189	0.15273	0.22780	0.3908	1.5359	0.9897	0.00129	0.03598	0.0359	0.1897
		70	1.6934	1.1011	0.03740	0.09060	0.1934	0.3010	1.5015	0.8999	0.00002	0.00998	0.0015	0.0999
T_3	40	30	2.0462	1.2048	0.2983	0.16386	0.5462	0.4048	1.8543	1.0037	0.12553	0.04149	0.3543	0.2037
		36	1.8686	1.1802	0.13587	0.14455	0.3686	0.3802	1.6767	0.9791	0.03122	0.03208	0.1767	0.1791
	80	50	1.7291	1.1344	0.05249	0.11182	0.2291	0.3344	1.5372	0.9333	0.00138	0.01777	0.0372	0.1333
		70	1.6632	1.0206	0.02663	0.04866	0.1632	0.2206	1.4713	0.8195	0.00082	0.00038	-0.0287	0.0195

From Table 4.3 , we can clearly see that those obtained estimated values of parameters under EM algorithm in terms of their corresponding MSEs and mean bias are generally smaller compared to those obtained via NR algorithm.

The estimates obtained under fixed time point T_3 , are closer to the true values than those obtained under fixed time point T_1 and fixed time point T_2 . For a definite point time T , an increase in the sample size n , results in generally better estimates with regard to reduced values of MSEs and mean biases.

The estimates for the parameters obtained using the proposed EM algorithm tend to converge towards the true values faster compared to those obtained using NR algorithm with regard to their corresponding mean squared errors and mean biases which are relatively smaller in all the three censoring schemes.

4.4 Real Data Analysis

Within this particular section, consideration on a simple application involving a set of data that has been utilized formerly with Lomax distribution is made. Facts and figures correlating to remission periods (in months) of a random prototype of 128 bladder cancer patients is considered as illustrated below; See Lee and Wang, (2003).

0.08, 0.20, 0.40, 0.50, 0.51, 0.81, 0.90, 1.05, 1.05, 1.19, 1.26, 1.35, 1.40, 1.46, 1.76, 2.02, 2.02, 2.07, 2.09, 2.23, 2.26, 2.46, 2.54, 2.62, 2.64, 2.69, 2.69, 2.69, 2.75, 2.83, 2.87, 3.02, 3.25, 3.31, 3.36, 3.36, 3.48, 3.52, 3.57, 3.64, 3.70, 3.82, 3.88, 4.18, 4.18, 4.23, 4.26, 4.33, 4.34, 4.40, 4.50, 4.51, 4.87, 4.98, 5.06, 5.09, 5.17, 5.32, 5.32, 5.34, 5.34, 5.41, 5.41, 5.49, 5.62, 5.71, 5.85, 6.25, 6.54, 6.76, 6.93, 6.94, 6.97, 7.09, 7.26, 7.28, 7.32, 7.39, 7.59, 7.59, 7.62, 7.63, 7.66, 7.87, 7.93, 8.26, 8.37, 8.53, 8.65, 8.66, 9.02, 9.22, 9.47, 9.74, 10.06, 10.34, 10.66, 10.66, 10.75, 11.25, 11.64, 11.76, 11.98, 12.02, 12.03, 12.07, 12.63, 13.11, 13.29, 13.80, 14.24, 14.76, 14.77, 14.83, 15.96, 16.62, 17.12, 17.14, 17.36, 18.10, 19.13, 20.28, 21.73, 22.69, 23.63, 25.74, 25.82, 26.31, 32.15, 34.26, 36.66, 36.66, 43.01, 46.12, 79.05

Table 4.4: The progressively type-II hybrid censored sample for the 128 bladder cancer patients when $m = 42, T = 4.50$, $R_1 = \dots = R_{41} = 2, R_{42} = 4$ is described below;

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
R_i	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
$X_{i:m:n} \setminus T$	0.08	0.20	0.40	0.50	0.51	0.81	0.90	1.05	1.05	1.19	1.26	1.35	1.40	1.46
	15	16	17	18	19	20	21	22	23	24	25	26	27	28
R_i	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
$X_{i:m:n} \setminus T$	1.76	2.02	2.02	2.07	2.09	2.23	2.26	2.46	2.54	2.62	2.64	2.69	2.69	2.69
	29	30	31	32	33	34	35	36	37	38	39	40	41	42
R_i	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	4.00
$X_{i:m:n} \setminus T$	2.75	2.83	2.87	3.02	3.25	3.31	3.36	3.36	3.48	3.52	3.57	3.64	3.70	3.82

The maximum likelihood estimates in regard to the above sample statistics were found to be $\hat{\lambda} = 0.942868$, $\hat{\theta} = 0.196449$ using an EM algorithm. The NR algorithm estimates turned out to be $\hat{\lambda} = 0.01233$, $\hat{\theta} = 1.4865$.

To construct the confidence interval for the survivor function, we make use of Greenwoods formula given below to estimate the standard error.

$$S.E(S_t) = S_t \sqrt{\sum \frac{D_t}{N_t(N_t - D_t)}} \quad (4.1)$$

where N_t is the number of patients at risk of dying at time t , D_t is the number of deaths at time t and S_t is the survival probability at time t .

Using a class size of 10, we put the data in groups form and obtain the survival estimates and their corresponding standard errors. The results are tabulated as demonstrated below;

Table 4.5: *Survival Estimates and their corresponding estimated standard errors*

Interval (months)	N_t	D_t	$\frac{D_t}{N_t(N_t - D_t)}$	$\sum \frac{D_t}{N_t(N_t - D_t)}$	EM		NR	
					S(x)	S.E	S(x)	S.E
0-10	128	92	0.019965	0.019965	0.28125	0.005615	0.32333	0.0064553
10-20	36	22	0.043651	0.063616	0.10938	0.0069583	0.15347	0.0097631
20-30	14	7	0.071429	0.135045	0.054688	0.0073853	0.093071	0.0125688
30-40	7	4	0.190476	0.325521	0.023437	0.0076292	0.057645	0.0187645
40-50	3	2	0.666667	0.992188	0.007812	0.007751	0.020026	0.0198696
50-60	1	0	0.000000	0.992188	0.007812	0.007751	0.020026	0.0198696
60-70	1	0	0.000000	0.992188	0.007812	0.007751	0.020026	0.0198696
70-80	1	1			0.00000	0.00000	0.00000	0.00000

Clearly, Table 4.5 shows that on average, the standard errors associated with the survival estimates obtained via EM estimated values are much smaller than those obtained via NR.

Table 4.6: The progressively type-II hybrid censored sample for the 128 bladder cancer patients when $m = 42, T = 3.67$, $R_1 = \dots = R_{40} = 2, R_{42} = 8$ is described below;

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
R_i	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
$X_{i:m:n}\backslash T$	0.08	0.20	0.40	0.50	0.51	0.81	0.90	1.05	1.05	1.19	1.26	1.35	1.40	1.46
	15	16	17	18	19	20	21	22	23	24	25	26	27	28
R_i	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
$X_{i:m:n}\backslash T$	1.76	2.02	2.02	2.07	2.09	2.23	2.26	2.46	2.54	2.62	2.64	2.69	2.69	2.69
	29	30	31	32	33	34	35	36	37	38	39	40	T	
R_i	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	8.00	
$X_{i:m:n}\backslash T$	2.75	2.83	2.87	3.02	3.25	3.31	3.36	3.36	3.48	3.52	3.57	3.64	3.70	

Based on this sample data on Table 4.6 above, maximum likelihood estimates were determined to be $\hat{\lambda} = 1.472485$, $\hat{\theta} = 0.1086397$ using the EM algorithm and $\hat{\lambda} = 4.31625$, $\hat{\theta} = 0.174976$ via NR algorithm.

Likewise as in Table 4.4, the standard errors associated with the survival estimates are computed as shown below in Table 4.7.

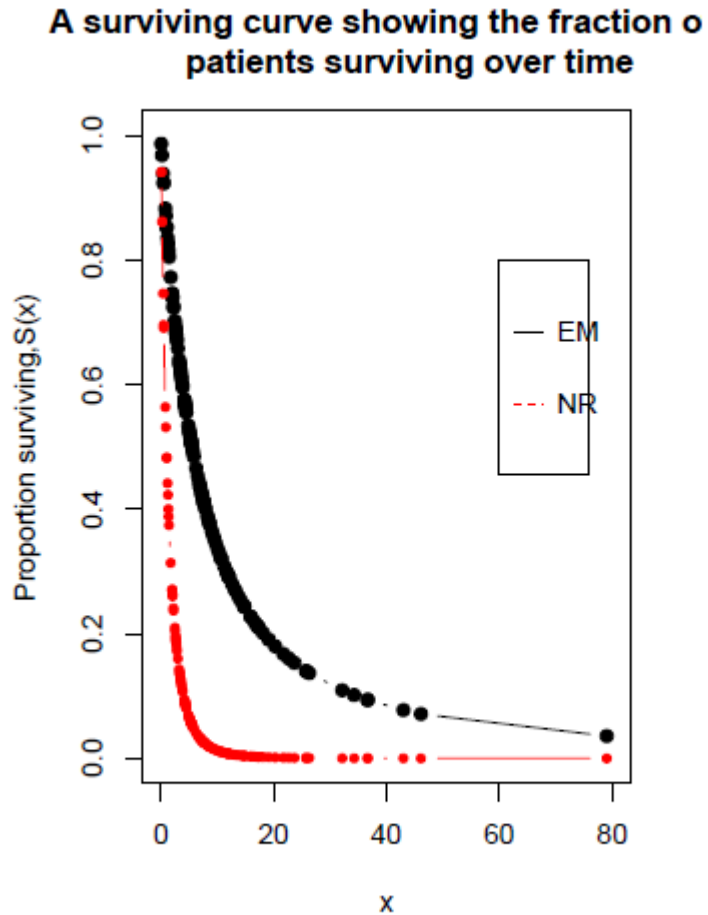
The outcomes in Table 4.7 also show that on average the standard errors associated with EM are much smaller than those associated with NR.

Table 4.7: *Survival Estimates and their corresponding estimated standard errors*

Interval (months)	N_t	D_t	$\frac{D_t}{N_t(N_t-D_t)}$	$\sum \frac{D_t}{N_t(N_t-D_t)}$	EM		NR	
					S(x)	S.E	S(x)	S.E
0-10	128	92	0.019965	0.019965	0.48115	0.009606	0.62035	0.0123853
10-20	36	22	0.043651	0.063616	0.30838	0.019618	0.37048	0.023569
20-30	14	7	0.071429	0.135045	0.165688	0.022375	0.194812	0.026308
30-40	7	4	0.190476	0.325521	0.094537	0.030774	0.110312	0.035909
40-50	3	2	0.666667	0.992188	0.01337	0.013266	0.047081	0.046714
50-60	1	0	0.000000	0.992188	0.01337	0.013266	0.047081	0.046714
60-70	1	0	0.000000	0.992188	0.01337	0.013266	0.047081	0.046714
70-80	1	1			0.00000	0.00000	0.000000	0.00000

Survival curves are plotted as shown below to show survival experience of the 128 bladder cancer patients over time:

Figure 1: A Surviving curve showing the fraction of patients surviving over time



graph.PNG

Both curves in Figure (1) above approaches zero. This implies that at most very few patients get cured of cancer. Between the period 10 – 80 months, the NR curve shows that the chances of survival of patients are relatively low (slightly above zero percent). However, the EM curve shows much higher chances of surviving within the same periods though decreasing with time.

CHAPTER FIVE

5 SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

In this chapter, summary , conclusion , recommendations and areas for further investigations are presented.

5.2 Summary

The main inspiration of this work was to determine the parameters of Lomax distribution based on progressive type-II hybrid censoring scheme. The maximum likelihood estimators for the parameters were attained using Expectation-Maximization algorithm along with Newton Raphson algorithm.

A simulation analysis was executed to collate the performance of the obtained MLEs using EM algorithm in estimating the parameters of Lomax distribution based on progressive type -II hybrid to those obtained via NR among three distinct censoring schemes.

For all the three given censoring schemes, the results have shown that :

- (i) For fixed prototype size n and fixed number of the discerned failures m , as fixed time point T increases, MSEs of estimated values of the parameters decrease.
- (ii) For fixed prototype size n and definite time point T , as the fixed number of the detected failures m increases, MSEs decrease.
- (iii) For fixed number of detected failures m and definite time point T , as prototype size n increases, MSEs decrease.

- (iv) Estimates obtained under fixed time point T_3 performs generally better compared to those obtained under fixed time point T_1 and fixed time point T_2 .

In addition, a factual data lay out has been considered with the study showing that the estimates obtained using the EM algorithm performs generally better as depicted by the small standard errors associated with survival estimates obtained using EM algorithm compared to those obtained by NR algorithm.

5.3 Conclusions

In this study, we have used EM and NR procedures to determine the MLE's of the parameters of Lomax distribution grounded onto progressive Type-II hybrid censored statistics. A comparison amid the estimates of the MLEs obtained using EM and NR in terms of bias and MSEs using a simulated data has been done. In addition, an investigation has also been done on real data by plotting the survival function curves using the parameter values obtained via EM and NR procedures. It is perceived that the estimates obtained under the submitted EM method generally performs well in the three schemes compared to those obtained under NR method in terms of yielding relatively low values of MSEs and mean Bias. Studies done in the recent past have also depicted similar results such as that done by Li and Lina (2015).

5.4 Recommendations

5.4.1 For use

In this study , we have obtained the MLEs of the parameters of Lomax distribution based onto progressive type-II hybrid censoring scheme. The MLEs attained using the EM algorithm has shown more robustness with regard to the obtained estimates converging to the true value faster. We therefore recommend the use of EM procedures based on Progressive type-II hybrid censoring on data related to life testing, reliability modeling as well as biological analysis.

5.4.2 For Further study

In this study we have presumed that the lifetime distributions are based on Lomax, further study could be extended to mixed models like Lomax-Poisson distribution, Lomax-logarithmic distribution among others. Real data set drawn from other areas like income distribution data can also be considered under Lomax distribution based on this censoring scheme and the performances of the MLE's compared with those obtained based on other different schemes.

References

- [1] Ahsanullah, M. (1991). Record values of Lomax distribution, *Statistica Neerlandica*, **41**, 21-29.
- [2] Ahsanulla, M. and Balakrishnan, N. (1994). Recurrence relations between the moments of record values from Lomax distribution. *The Indian Journal of Statistics, Series B (1960-2002)*. **56** (2), 140-146.
- [3] Arnold, B.C. (1983). *The Pareto distributions*. Fairland, MD: International Co-operative Publishing House.
- [4] Asgharzadeh, A. and Balakrishnan, N. (2005). Inference for the scaled half-logistic distribution based on progressively Type-II censored samples. *Communications in Statistics and Theory Methods*, **34**, 73-87.
- [5] Asgharzadeh, A. and Valiollahi, R. (2011). Estimation of the scale parameter of the Lomax distribution under progressive censoring. *International Journal of Statistics and Economics*, **6**, 37-48.
- [6] Asgharzadeh, A., Kazemi, M. and Kus, C. (2013). Analysis of the hybrid censored data from the logistic distribution. *Journal of Probability and Statistical Science*, **11**, 183-198.
- [7] Atkinson, A. B. and Harrison, A. J. (1978). *Distribution of Personal Wealth in Britain*. Cambridge University Press, Cambridge.
- [8] Bain, L.J. (1978). *Statistical Analysis of Reliability and Life-testing Models-Theory and Practice*, Marcel Dekker, New York.
- [9] Bain, L.J. and Engelhardt, M. (1991), *Statistical Analysis of Reliability and Life-Testing Models*, Second edition, Marcel Dekker, New York.
- [10] Balkema, A. and De Haan, L. (1974). Residual life time at great age. *Annals of Probability*, **2**, 792 - 804.

- [11] Balakrishnan, N. and Aggarwala, R. (2000). *Progressive censoring theory, methods and applications*. Boston, MA: Birkhuser.
- [12] Balakrishnan, N., Kannan, N., Lin, C.T. and Wu, S.J.S. (2004). Inference for the extreme value distribution under progressive type-II censoring, *Journal of Statistical Computation and Simulation*, **74**, 25-45.
- [13] Balakrishnan, N. (2007). Progressive Censoring methodology; *An appraisal (with discussion)*, TEST, **16**, 211-289.
- [14] Balakrishnan, N. and Kundu, D. (2013). Hybrid censoring: models, inferential results and applications (with discussion). *Computational Statistics and Data Analysis*, **59**, 166-209.
- [15] Bander, A. and Marshall, A. (2013). An approach of estimating the reliability function for the Lomax distribution based on general progressive censoring scheme. *Journal of quality and reliability engineering*.
- [16] Banerjee, A. and Kundu, D. (2008). Inference based on type-II hybrid censored data from a Weibull distribution. *IEEE Transactions on Reliability*, **57**, 369-378.
- [17] Barlow, R., Madansky, A., Proschan, F. and Scheuer, E. (1968). Statistical estimation procedures for the Burn-in process. *Technometrics*, **10** , 51-62.
- [18] Bartholomew, D. (1963). The sampling distribution of an estimate arising in life testing. *Technometrics*, **5**, 361-374.
- [19] Bhattacharya, S., Pradhan, B. and Kundu, D. (2014), *Analysis of hybrid censored competing risks data*. *Statistics*, **48** (5), 1138-1154.
- [20] Bryson, M.C. (1974). Heavy-tailed distributions; properties and tests. *Technometrics*, **16**, 61-68.
- [21] Chahkandi, M. and Ganjali, M. (2009). On some lifetime distributions with decreasing failure rate. *Computer Statistics Data Analysis*, **53**, 4433-4440.

- [22] Chan, P., Ng, H. and Su, F. (2015). *Exact likelihood inference for the two-parameter exponential distribution under Type-II progressively hybrid censoring*, *Metrika*, **78**, 747-770.
- [23] Chen, S. and Bhattacharya, G.K. (1988). Exact confidence bounds for an exponential parameter under hybrid censoring. *Communications in Statistics-Theory and Methods*, **17**, 1857-1870.
- [24] Childs, A., Balakrishnan, N. and Moshref, M. (2001). Order statistics from non-identical right truncated Lomax random variables with applications. *Statistical Papers*, **42**, 187-206.
- [25] Childs A., Chandrasekhar B., Balakrishnan N. and Kundu D. (2003). Exact likelihood inference based on type-I and type-II hybrid censored samples from the exponential distribution. *Annals of the Institute of Statistical Mathematics*, **55**, 319-330.
- [26] Childs, A., Chandrasekhar, B. and Balakrishnan, N. (2007). Exact likelihood inference for an exponential parameter under progressively hybrid schemes. *Statistical models and methods for biomedical and technical systems*. F.Vonta, M. Nikulin, N. Limnios, and C.Huber-Carod(Eds.) Boston, M.A: Birkhauser.
- [27] Cho, Y., Sun, H. and Lee, K. (2015), Exact likelihood inference for an exponential parameter under generalized progressive hybrid censoring scheme, *Statistical Methodology*, **23**, 18-34.
- [28] Cohen, A.C. (1963). Progressively censored samples in life testing, *Technometrics*, **5**, 327 - 329.
- [29] Cohen, A.C. (1966), Life-testing and early failure, *Technometrics*, **8**, 539-549.
- [30] Cohen , A.C. and Whitten, B.J. (1988). *Parameter Estimation in Reliability and Life Span Models*, Marcel Dekker, New York.

- [31] Cohen, A.C (1991). *Truncated and censored samples*, Marcel Dekker, New York.
- [32] Corbellini, A., Crosato, L., Ganugi, P. and Mazzoli, M. (2007). Fitting Pareto II distributions on firm size: *Statistical methodology and economic puzzles*. Paper presented at the International Conference on Applied Stochastic Models and data Analysis, Chania, Crete.
- [33] Cramer, E. and Schmiedt, A.B. (2011). Progressively type II censored competing risks data from Lomax distributions, *Comput Stat Data Anal*, **55**, 1285-1303.
- [34] Dempster, A.P., Laird, N.M. and Rubin, D.D. (1977). Maximum likelihood from incomplete data via EM algorithm. *Journal of the Royal Statistical Society, Series B*, **39**, 1-38.
- [35] Draper, N. and Guttman, I. (1987). Bayesian analysis of hybrid life tests with exponential failure times. *Annals of the Institute of Statistical Mathematics*, **39**, 219-225.
- [36] Dube, S., Pradhan, B. and Kundu, D. (2010). Parameter estimation of the hybrid censored log normal distribution. *Journal of Statistical Computation and Simulation*, **81**, 275-287.
- [37] Epstein, B. (1954). Truncated life tests in the exponential case. *Annals of Mathematical Statistics*, **25**, 555-564.
- [38] Ebrahimi, N. (1992). Prediction intervals for future failures in the exponential distribution under hybrid censoring. *IEEE Transactions on Reliability*, **41**, 127-132.
- [39] Ghitany, M.E., Al-Awadhi, F.A. and Alkhalfan, L.A. (2007). Marshal-Olkin extended Lomax distribution and its application to censored data. *Communications in Statistics Theory and Methods*, **36**, 1855-1866.
- [40] Giles, D., Feng, H. and Gadwin, R. (2013). On the bias of the maximum likelihood estimator for the two parameter Lomax distribution. *Computational Statistics Theory and Methods*, **42** (11): 1934-1950.

- [41] Gorny, J. and Cramer, E. (2016), Exact likelihood inference for exponential distribution under generalized progressive hybrid censoring schemes, *Statistical Methodology*, **29**, 70-94.
- [42] Gupta, R.D. and Kundu, D. (1996). Estimation of parameters from lifetime distribution based on progressive type II censored samples. *IEEE Transactions on Reliability*, **34**, 201-206.
- [43] Hassan, A. S. and Al-Ghamdi, A. S. (2009). Optimum step stress accelerated life testing for Lomax distribution. *Journal of Applied Sciences Research*, **5**, 2153-2164.
- [44] Harter, H. L. (1970). Order Statistics and their use in Testing and Estimation *U.S. Government Printing Office*, **2**. Washington, D.C.
- [45] Harris, C. M. (1968). The Pareto distribution as a queue service discipline. *Operations Research*, **16**, 307-313.
- [46] Helu, A. Samawi, H. and Raqab, M.Z. (2013). Estimation on Lomax progressive censoring using the EM algorithm. *Journal of Statistical Computation and Simulation*, 837-861.
- [47] Helu, A. Samawi, H. and Raqab, M.Z. (2015). Estimation on Lomax progressive censoring using the EM algorithm. *Journal of Statistical Computation and Simulation*, **85**, (5): 1035-1052.
- [48] Hemmati, F. and Khorram, E. (2013). Statistical analysis of log-normal distribution under type-II progressive hybrid censoring schemes, *Communications in Statistics-Simulation and Computation*, **42**, 52-75.
- [49] Holland, O., Golaup, A. and Aghvami, A.H. (2006). Traffic characteristics of aggregated module downloads for mobile terminal reconfiguration. *IEE Proceedings Communications*, **135**, 683-690

- [50] Howlader, H. and Hossain, A. (2002). Bayesian survival estimation of Pareto distribution of the second kind based on failure censored data. *Computational Statistics and Data Analysis*, **38**: 301-314.
- [51] Jeong, H.S., Park, J.I. and Yum, B.J. (1996). Development of (r, T) hybrid sampling plans for exponential lifetime distributions. *Journal of Applied Statistics*, **23**, 601-607.
- [52] Johnson, N. Kotz, S. and Balakrishnan, N. (1994). Continuous Univariate Distribution, *John Wiley and Sons*, **1**.
- [53] Li, J. and Lina, M. (2015). Inference for the Generalized Rayleigh Distribution Based on Progressively Type II hybrid Censored data. *Journal of Information and Computational Science*, 1101-1112.
- [54] Kundu, D. and Gupta, R.D. (1988). Hybrid censoring schemes with exponential failure distribution, *Communications in Statistics Theory and Methods*, **27**, 3065-3083.
- [55] Kundu, D. and Joarder, A. (2006). Analysis of Type II progressive hybrid censored data. *Computational Statistics and Data Analysis*, **50**, 2509-2528.
- [56] Kundu, D. (2007). On hybrid censored Weibull distribution. *Journal of Statistical Planning and Inference*, **137**, 2127-2142.
- [57] Kundu, D. and Pradhan, B. (2009). Estimating the parameters of the generalized exponential distribution in the presence of hybrid censoring *Communications in Statistics-Theory and Methods*, **38**, 2030-2041.
- [58] Kundu, D., Joarder, A. and Krishna, H. (2009). On Type-II Progressively Hybrid Censoring, *Journal of Modern Applied Statistical Methods*, **8**, Iss.2, Article **18**.
- [59] Kus, C. (2007). A new lifetime distribution, *Computational Statistics and Data Analysis*, **51**, 4497-4509.

- [60] Lee, E.T. and Wang, J.W (2003). *Statistical methods for survival data Analysis*, Wiley, New York. **3rd edition**.
- [61] Li, Z.H. and Huang, W.Y. (2011). Bayesian Estimation for Parameter of Lomax model under progressively Type II hybrid censored samples. *Proceedings of 2011 4th ICC SIT*.
- [62] Lin, C.T., Huang, Y.L. and Balakrishnan, N. (2011). Exact Bayesian variable sampling plans for the exponential distribution with progressive hybrid censoring, *Journal of Statistical Computation and Simulation*, **81**, 873-882.
- [63] Lomax, K. S. (1954). Business failures; another example of the analysis of failure data. *Journal of American Statistical Association*, **49**, 847-852.
- [64] Mann, N.R., Schafer, R.E. and Singpurwalla, N.D. (1974). *Methods for Statistical Analysis of Reliability and Life Data*. John Wiley and Sons, New York.
- [65] Mokhtari, E.B., Rad, A.H. and Yousefzadeh, F. (2011). Inference for Weibull distribution based on progressively type II hybrid censored data. *Journal of Statistical Planning and Inference*, **141** (8), 2824-2838.
- [66] Ng, H.K.T., Chan, C.S. and Balakrishnan, N. (2002). Estimation of Parameters from Progressively Censored Data Using EM algorithm. *Computational Statistics and Data Analysis*, **39**, 371-386.
- [67] Ng, H.K.T., Kinaci, I., Akdogan, Y. and Kus, C. (2014). Statistical Inference for Weibull distribution based on a modified progressive type II censoring scheme. *Sri Lankan Journal of Applied Statistics (special issue)*.
- [68] Panahi, H. and Asadi, S. (2011). Inference of stress-strength model for a Lomax distribution. *Proceedings of World Academy of Science, Engineering and Technology*, **7** (55), 275-278.

- [69] Raqab, M.Z., Asgharzadeh, A. and Valiollahi R. (2010). Prediction for Pareto distribution based on progressively Type-II censored samples. *Computational Statistics and Data Analysis*, **54**, 1732-1743.
- [70] Raqab, M. Z. and Madi, M. T. (2011). Inference for the generalized Rayleigh distribution based on Progressively censored data. *Journal of Statistical Planning and Inference*, **141** (10), 3313-3322.
- [71] Salem, A. M. and Abo-Kasem, O.E. (2011). Estimation for the Parameters of the exponentiated Weibull distribution based on progressive hybrid censored samples. *International Journal of Contemporary Mathematics and Sciences*, **6**, 1713-1724.
- [72] Sarhan, A. M. and Abuammoh, A. (2008). Statistical Inference Using Progressively Type-II Censored Data with Random Scheme, *International Mathematical Forum*, **35** (3), 1713-1725.
- [73] Soliman, A. A. (2005). Estimation of Parameters of life from progressively censored data using Burr-XII model. *IEEE Transactions on Reliability*, **54**, 34-42.
- [74] Wu, S. J. (2003). Estimation for the two -parameter Distribution under Progressive censoring with Uniform Removals, *Journal of statistical computation and Simulation*. **73** (2), 125-134.
- [75] Wu, M., Shi, Y. and Wang, Y. (2016). Statistical analysis for competing risks model from a weibull distribution under progressively hybrid censoring
- [76] Wu, M., Shi, Y. and Sun, Y. (2017). Statistical analysis for competing risks model from a weibull distribution under progressively hybrid censoring. *Communications in Statistics-Theory and Models*, **46** (2), 75-86
- [77] Yamaguchi, K. and Watanabe, M. (2004). *The EM algorithm and related statistical models*, New York, Basel. 21-36

- [78] Yongming, M. and Yimin, S. (2013). Inference for Lomax distribution based on type II progressively hybrid censored data. *Journal of Physical Sciences*, **17**, 33-41.