TWO-STAGE MIXTURE LOGISTIC DISTRIBUTION

BY

IRWAN RIZADI

A Thesis Presented to the
DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

APPLIED STATISTICS

DECEMBER, 2016

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN - 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by **Irwan Rizadi** under the direction his thesis advisor and approved by his thesis committee, has been presented and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **Master of Science in Applied Statistics**.

Dr. Mohammad Hafidz Omar
(Advisor)

Dr. Hussain Salem Al-Attas
Department Chairman

Dr. Muhammad Riaz
(Member)

Dr. Salam A. Zummo
Dean of Graduate Studies

16/1/17

Date

Dr. Monjed H. Samuh
(Member)

Dedication to Allah SWT and Rasulullah SAW

My Mother, My Father, My Wife, and All My Family

Statistics World

ACKNOWLEDGEMENTS

With gratitude to Almighty Allah Who had made it possible for me to accomplish this study successfully. I would like to thank the King Fahd University of Petroleum and Minerals for the scholarship and opportunity to study Master of Science degree.

I wish to register my profound gratitude and unreserved appreciation to Dr. Mohammad Hafidz Omar who served as my thesis advisor. The success of this work is credited to his tireless support and priceless ideas. My special thanks also to my thesis committee members, Dr. Muhammad Riaz and Dr. Monjed H. Samuh for their constructive suggestion and contribution towards the success of this work.

I am also indebted to the chairman of Mathematics and Statistics Department Dr. Husain S. Al-Attas for his assistance and numerous help, especially in most difficult times. My appreciation also to all faculty members and staff in Department of Mathematics and Statistics for their direct and indirect support.

I would like to thank friends and colleagues, in particular, Md. Pear Hossain for the discussions, kindness and encouragements during the course of my study. Special thanks to the Indonesian student community for the kindness and for being my second family here.

Special thanks to my family most especially my mother, my father, my wife, my son, and my daughter, for their love and patience throughout the M.Sc. program. All of this remain nothing without full support from my family. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

1PL	: 1-Parameter Logistic
2PL	: 2-Parameter Logistic
3PL	: 3-Parameter Logistic
AIC	: Akaike Information Criterion
ARL	: Average Run Length
BIC	: Bayesian Information Criterion
CDF	: Cumulative Distribution Function
CF	: Characteristic Function
EM	: Expectation Maximization
IRT	: Item Response Theory
MGF	: Moment Generating Function
MLE	: Maximum Likelihood Estimation
MSE	: Mean of Squares Error
PDF	: Probability Density Function
SPC	: Statistical Process Control
SSE	: Sum of Squares Error

ABSTRACT

Full Name : Irwan Rizadi

Thesis Title : Two-stage Mixture Logistic Distribution

Major Field : Applied Statistics

Date of Degree : December 26, 2016

In this thesis, we developed a two-stage mixture logistic distribution and its properties as well as its application in social science. In the first stage, we mixed a point-degenerate distribution with the 2-parameter logistic distribution using a mixing proportion parameter λ and $(1-\lambda)$ respectively. Then, in the second stage, the resulting distribution has been mixed with the 1-parameter logistic distribution using the mixing proportion p and $(1-p)$. Some important properties of the two-stage mixture logistic distribution have been studied in this thesis. Some parameter estimations of the model also have been developed. The two-stage mixture logistic distribution found the application in the population growth model, item response theory, and statistical process control.

Keywords: Logistic function, mixture distribution, item response theory.

ملخص الرسالة

الاسم الكامل : إيروان ريزا دي

عنوان الرسالة : التوزيع اللوجستي المزدوج ذو المرحلتين

التخصص : الإحصاء التطبيقي

تاريخ الدرجة العلمية : ٢٦ ديسمبر ٢٠١٦

في هذه الأطروحة، قمنا بتطوير التوزيع اللوجستي المزدوج ذو المرحلتين وخصائصه وكذلك تطبيقاته في العلوم الاجتماعية. في المرحلة الأولى، قمنا بمزج توزيع النقطة الثابتة مع التوزيع اللوجيستي ذو المعلمين باستخدام معلمة المزج النسبي λ و $(1-\lambda)$ على التوالي. ثم في المرحلة الثانية، قمنا بمزج التوزيع الناتج من التوزيع اللوجستي ذو المعلمة الواحدة باستخدام معلمة المزج النسبي p و $(1-p)$. بعض الخصائص المهمة للتوزيع اللوجستي المزدوج ذو المرحلتين تم دراستها في هذه الاطروحة. كما تم تطوير بعض مقدرات المعالم للنموذج. التوزيع اللوجستي المزدوج ذو المرحلتين لها تطبيقات في نظرية استجابة البند، نموذج النمو السكاني، وضبط الجودة الاحصائية.

CHAPTER 1

INTRODUCTION

The logistic distribution is the family of the symmetrical distribution. It is popularly used in the population growth and Item Response Theory (IRT). We develop the two-stage mixture logistic distribution and provide some important properties of the mixture distribution. The research questions and objectives of the study are provided in this chapter.

1.1 Background of Study

The logistic growth function was first proposed as a tool for use in demographic studies in 1800's. Logistic distribution is the family of the symmetrical distribution. Figure 1.1 (a) shows that the logistic function has an "S" (sigmoid) curve in the Cumulative Distribution Function (CDF) and bell curve in the Probability Density Function (PDF) (see for example Cramer [1]). The "S" shaped curve and bell curve of logistic distribution with a different value of α and β was proposed by using the logistic function as follows.

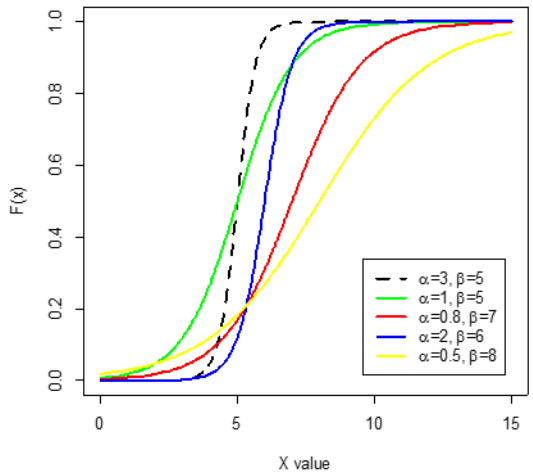
Definition 1.1 *The random variable X for logistic distribution with 2 parameters has the CDF*

$$F\left(x;\alpha,\beta\right)=\frac{1}{1+\exp\left[-\alpha\left(x-\beta\right)\right]}. \tag{1.1}$$

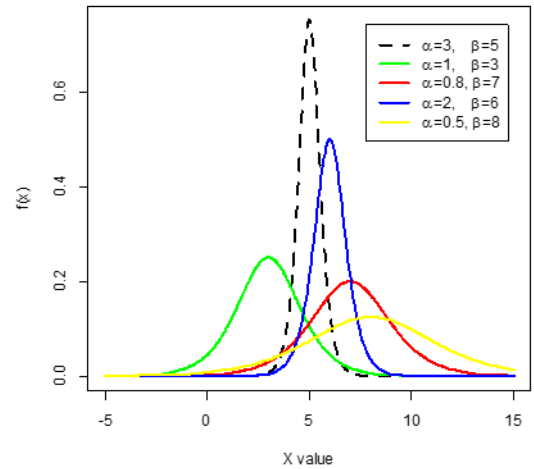
and PDF

$$f\left(x;\alpha,\beta\right)=\frac{\alpha\exp\left[-\alpha\left(x-\beta\right)\right]}{\left[1+\exp\left\{-\alpha\left(x-\beta\right)\right\}\right]^{2}}. \tag{1.2}$$

where β is the location parameter and α is the scale parameter.



(a). CDF for Logistic



(b). PDF for Logistic

Figure 1.1 Graph of Logistic Distribution

Figure 1.1 (b) provides the density function of logistic distribution with different values of α and β. The PDF of logistic distribution is a symmetrical distribution with midpoint zero. The $f(x)$ will peak at the midpoint when the random variable X moves through the real number. The logistic function is well applied in models for human population growth, as described by Oliver [2].

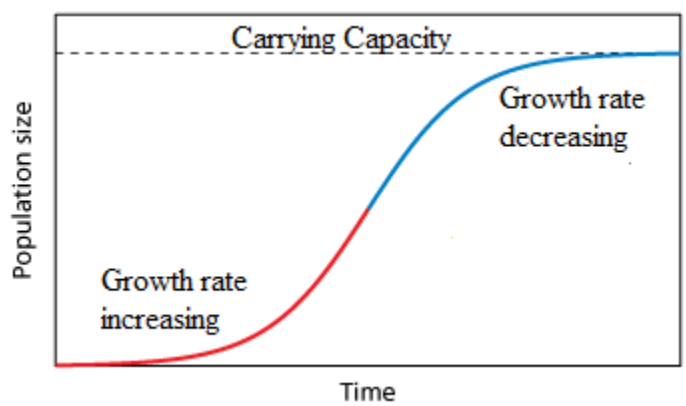


Figure 1.2 Logistic Population Growth Model

The population growth model has a characteristic "S" shaped curve. Refers to Figure 1.2, when there are a few individuals, the growth of population is slow. Otherwise, when there are more individuals, the growth of population is fast. Ultimately, the more

individuals there are in the environment, the slower the growth of the population due to the limitation of resources. In the logistic growth, a population continue to grow until it reaches the maximum number of individuals that an environment can support.

The logistic distribution is also popularly used in Item Response Theory (IRT). The 1-Parameter Logistic (1PL), 2-Parameter Logistic (2PL) and 3-Parameter Logistic (3PL) in IRT is built from the logistic distribution. Figure 1.3 shows the graph of 1PL.

Definition 1.2 *The random variable X for logistic distribution with 1 parameter has the CDF*

$$F(x;\beta) = \frac{1}{1+\exp\left[-(x-\beta)\right]} \tag{1.3}$$

and PDF

$$f(x;\beta) = \frac{\exp\left[-(x-\beta)\right]}{\left[1+\exp\left\{-(x-\beta)\right\}\right]^2}, \tag{1.4}$$

where β *is the location parameter.*

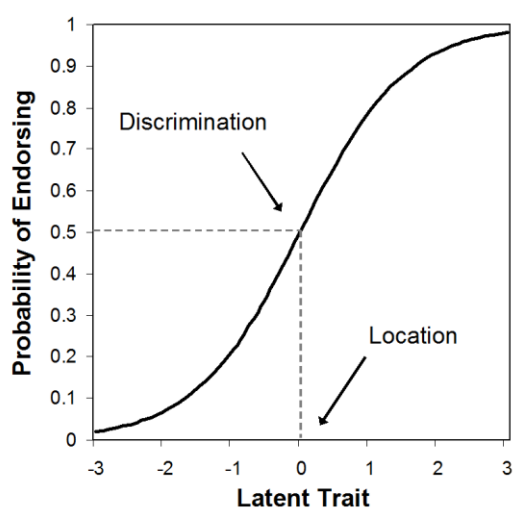
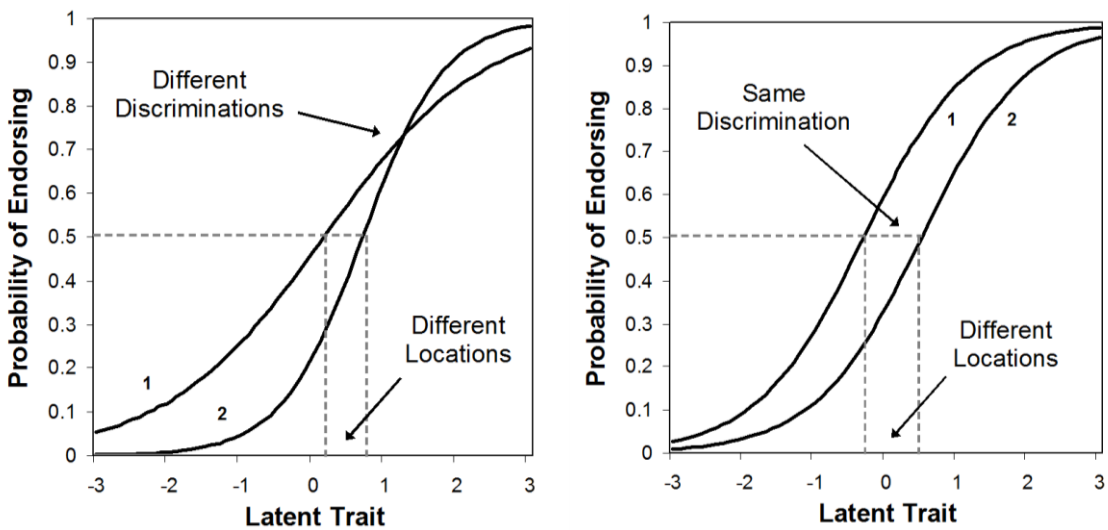


Figure 1.3 Graph of 1PL Model (Ainsworth [3])

Figure 1.3 shows the graph of 1PL model with a parameter β that refers to difficulty parameter, which is the point at which an examinee with latent trait or ability $x = \beta$ have the probability of correctly answering an item equal to 1/2. It is also the inflection point on the ability (x) metric. Questions with high values of β are difficult to test items, where examinees with low-ability have small probabilities of correctly answering the questions. On the other hand, items with small β values are easy, with most examinees have at least a moderate probability of correctly answering the question.

The 2PL model has 2 parameters α and β where α is a discrimination parameter and β is a difficulty parameter. The coefficient α refers to a discrimination parameter that permits items to differentially discriminate among examinees. Basically, α is defined as the slope of the item characteristic curve at the point of inflection β. The higher value of α, the more sharply the item discriminates among examinees.



(a). Graph of 2PL model with different discrimination and different locations

(b). Graph of 2PL model with same discrimination and different locations

Figure 1.4 Graph of 2PL Model in IRT (Ainsworth [3])

The 3PL model in IRT is also popularly used. It has the parameter λ as the third parameter where λ is a guessing item parameter. Parameter λ allows examinees, even ones with low ability, to have perhaps substantial probability of correctly answering even moderate or hard items. This is unlike the 2PL where guessing was not possible.

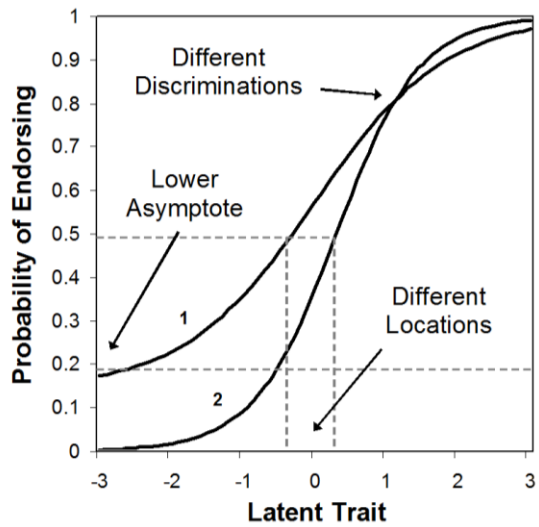


Figure 1.5 Graph of 3PL Model with Guessing Parameter (Ainsworth [3])

In this thesis, mixture distribution based on the logistic function is developed. Mixture distribution is defined as the probability distribution of a random variable that is derived from different subgroups. That is, first, select a subgroup randomly based on a given selection probabilities. Then, select a random variable from the chosen subgroup. Balakrishnan et al. [4] said that mixture distributions produces more flexible models with better probabilistic and statistical properties.

However, we investigate an extended mixture distribution of the logistic function, in particular, a two-stage mixture. In the first stage, we mix a point-degenerate distribution with the 2PL distribution using a mixing proportion parameters λ and $(1-\lambda)$ respectively. This gives us the 3PL distribution discussed earlier. Then, in the second stage,

the resulting distribution is mixed with the 1PL distribution respectively using the mixing proportion p and $(1-p)$. This defines our two-stage mixture logistic distribution.

The literature review for logistic and mixture distribution is discussed in Chapter 2. Some important properties of the two-stage mixture logistic distribution are studied in the Chapter 3. The properties include the Moment Generating Function (MGF), raw and centered moment, the characteristic function, hazard function and survival function. In Chapter 4, the parameter estimation of the model is developed in terms of method of moment estimation and Maximum Likelihood Estimation (MLE). In Chapter 5, we apply the two-stage mixture logistic distribution in human growth population, Item Response Theory (IRT), and Statistical Process Control (SPC).

1.2 Research Questions

The two-stage mixture logistic distribution is studied in this thesis. The research focus of this thesis is provided in the following three research questions:

Research Question 1 (Properties of distribution)

The properties of distribution are used to interpret statistical data more easily. What are some important properties of the two-stage mixture logistic distribution?

Research Question 2 (Parameter estimation of distribution)

An estimator gives a point estimate of the parameter(s) of interest. How we estimate the parameters of the two-stage mixture logistic distribution?

Research Question 3 (Application of distribution)

Logistic distribution was popularly used in social science. How do we apply the two-stage mixture logistic distribution in real life, especially in human population growth, IRT, and SPC?

1.3 Objectives of the Study

The main contribution of this research is to develop a two-stage mixture logistic distribution and its properties as well as its application in social science. The specific objectives as related to the previous research questions are as follows:

a) Research Question 1

1. Perform a comprehensive literature review on the mixture logistic distribution.
2. Develop a two-stage mixture logistic distribution.
3. Develop properties of the two-stage mixture logistic distribution.
   i. Moment generating function,
   ii. Moment of distribution,
   iii. Characteristic function,
   iv. Hazard and survival function.

b) Research Question 2

1. Develop parameter estimation for the two-stage mixture logistic distribution.
   i. Method of moment estimation,
   ii. Maximum likelihood estimation.

c) Research Question 3

1. Apply the two-stage mixture logistic distribution.

i. In human population growth,

ii. In item response theory,

iii. In statistical process control.

These objectives are the primary interest that motivated us to investigate more about logistic distribution, in particular, the two-stage mixture. We are interested in investigating some important properties of distribution as well as the parameter estimation. In Chapter 2, we discuss the literature review of logistic distribution in terms of a mixture distribution.

CHAPTER 2

LITERATURE REVIEW

In the recent years, most research publications in probability theory are talking about the mixture distribution, either the finite mixture or the infinite mixture distribution. Some statistical literature has received increasing attention for the density functions of the mixture distribution, partly because of interest in their mathematical properties and the considerable number of areas in which such density functions are encountered. In this chapter, comprehensive literature review of the previous works in the area of logistic distribution and mixture distribution is discussed.

2.1 Logistic Distribution

A logistic function was firstly proposed as a tool for use in demographic studies by Verhulst [5][6]. Verhulst assumed that the increase of the natural logarithm of a population size for a given geographic area as a function of the time is a constant minus a function which increases with the population. Verhulst derived his logistic function to describe the self-limiting growth of a biological population.

Other authors also applied the logistic function for estimating the growth of human population (see Pearl and Reed [7], Leach [8], and Oliver [2]). The initial stage of growth is approximately exponential. Then, as saturation begins, the growth slows and at maturity the growth stops. The application of the logistic curve as a model for population growth has been proposed in the 1920's by Pearl and Reed [7]. In other research, Leach [8] explained that the logistic function can give a more reliable forecast of total populations

and was supported by an examination of the growth population in Great Britain, USA, and Scotland. Three particular aspects of logistic function were considered by Oliver [2] to make it more useful in human population growth model which is the variance-covariance matrix for the estimates of his parameter, the logistic function is generalized and, both the function and its generalization are compared with an alternative parameterization.

The other interesting applications of the logistic model were applied in the area of public health by Dyke and Petterson [9] and Grizzle [10]. Grizzle [10] discussed analysis of public health data through the logistic model when the data can be arranged in a multi-way classification as a factorial arrangement. The data in Grizzle [10] consisted of school children with behavior problems who were analyzed based on whether or not their mothers had suffered previous infant losses.

Plackett [11] was first to use the logistic function in the analysis of survival data. He used ordered samples and expected values of order statistics. He discussed a method of estimating logistic parameters α and β from the k smallest observation in a sample of size n from the logistic distribution. He paid more attention to the efficiencies of the censored estimates relative to the whole samples.

Oliver [12] used the logistic function in the economic field to fit current consumption to current income. Fisk [13] used the logistic function in studying the income distributions. He said that the logistic distribution has certain useful characteristics such as simple Lorenz measures of inequality and a simple method of graphical analysis, which make it a useful tool in examining and comparing distributions of income. The differential function from which the logistic is derived can be varied to allow a wide range of different distributional forms to be fitted.

Another application of the logistic function is item response function which has been proposed by Lord, Novick, and Birnbaum [14]. Lord et al. used the logistic function for estimating and modelling mental test scores which is the foundation of IRT. Harris [15] considered the comparison of 1PL, 2PL, and 3PL model in the IRT. She illustrated the effect of changing parameter α, β, and λ in item characteristic curve.

Balakhrisnan [16] considered a general form of the logistic distribution and provided some basic characteristic and properties of the distribution. He used the CDF and PDF with parameters mean μ and variance σ^2. He also showed that the CDF of the standard normal variable is very close in shape to the CDF of the standard logistic variable.

Gupta [17] considered three methods of estimating the parameters (both location and scale) of the logistic distribution using sample quantiles and order statistics. Three type of the estimators are: (1) best linear unbiased estimators based on sample quantiles; (2) unbiased linear asymptotically best estimators of Blom [18]; and (3) asymptotically unbiased linear estimators of Jung [19]. The above three types of estimators are all asymptotically efficient. The quantile estimators are based on a fixed number of sample quantiles when the total sample size is very large.

Gupta et al. [20] elaborated more about best linear unbiased estimators of the parameters of the logistic distribution using order statistics. The paper deals with the problem of estimating the mean and the standard deviation of the logistic distribution by using ordered observations in small samples.

Stephens [21] discussed the goodness of fit tests for the logistic distribution, based on statistics calculated from the empirical distribution function. He discussed the test of fit of H_0 with possibly one or both parameters α and β unknown. He introduced three

statistics (W^2, U^2, and A^2) for which asymptotic percentage points are given. For each of the three statistics, one or both parameters of the distribution must be estimated from the data.

We completed the literature on logistic distributions, the properties, parameter estimation, and application of the logistic function. But because of our thesis is about mixture distribution, we also need to review what other researchers have studied in terms of the mixture distribution.

2.2 Mixture Distribution

Some examples of mixture discrete and mixture continuous distributions were explained clearly by Everitt and Hand [22]. Some important properties were also provided in case of a mixture of the normal distribution, exponential distribution, etc. Some parameter estimation methods were also described clearly. The most widely used of finite mixture distributions are those involving normal components, (see Medgyessy [23] and Gregor [24]). In Mixture of two normal distributions, Cohen [25] provided the parameter estimation, primarily in the method of moments.

Al-Hussaini and el-Hakim [26] presented the failure rate of the inverse Gaussian-Weibull mixture model. Their work was based on the different causes of failure of a given system could lead to different failure distributions. Sultan et al. [27] explained the properties and estimation of the mixture of two inverse Weibull distributions. A mathematical analysis for learning about the critical features of the hazard function was explained clearly by Chechile [28].

Taylor [29] proposed a semi-parametric generalization of the failure time mixture model. A mixture model analyzed failure time data in which there are thought to be two

groups of objects, those who could eventually develop the endpoint (or disease) and those who could not develop the endpoint. A logistic regression model is proposed for the incidence part of the model. The estimator arises naturally out of the EM algorithm approach for fitting failure time mixture models as described by Larson [30]. Farewell [31] used mixture models to analyze the survival data with long-term survivors.

Behboodian [32][33] considered a numerical method for computation of the Fisher information matrix of mixture distribution. It was shown by using a simple transformation which reduces the number of parameters and then the whole information matrix leads to the numerical evaluation of the integral.

Rost [34] firstly proposed mixture Rasch model or simply the 1PL model. Willse [35] provided the joint maximum likelihood estimation of mixture Rasch model. Hannah et al. [36] used Psychomix package to implement a more flexible Rasch model in the R software. Then, Leong, Mahdi, and Ling [37] provided some applications of mixture Rasch model in IRT. Most of the applications of mixture Rasch model in the IRT are studied by Reise and Revicki [38]. They developed fundamental concepts in mixture Rasch IRT and used these IRT models in applied problems.

2.2.1 Parameter Estimation

Since in most cases, parameter estimation may come across no analytical solution, Dempster [39] described clearly the Expectation-Maximization (EM) Algorithm for obtaining the maximum likelihood estimation of the mixture distribution. In another research, Leytham [40] considered the maximum likelihood estimates for the parameters of a mixture of two normal distributions in terms of an EM algorithm. Small sample properties of the parameter estimates were explored using Monte Carlo simulation. Agha

and Ibrahim [41] provide the algorithm to solve the MLE in case when the sample data are available only from the mixed distribution. Wengrzik [42] provided an easier way to understanding the EM Algorithm in terms of mixture distributions.

Alkasasbeh and Raqab [43] considered the maximum likelihood estimation of the different parameters of a generalized logistic distribution as well as other five estimations procedure. They compared the maximum likelihood estimators, the method of moment estimators, estimators based on percentiles, least squares estimators, weighted least squares estimators and the L-moments estimator by conducting extensive numerical simulations.

Now, we develop the mixture logistic distribution, in particular, the two-stage mixture and then provide the properties as well as the parameter estimation of the distribution.

CHAPTER 3

TWO-STAGE MIXTURE LOGISTIC DISTRIBUTION

In this chapter, we develop the two-stage mixture logistic distribution and provide some important properties. In the first stage mixture, we mix a point-degenerate distribution with a 2PL distribution as described in Chapter 1. Then, in the second stage mixture, the resulting distribution is mixed with 1PL distribution. The important properties of the two-stage mixture logistic distribution such as moment generating function, the moment, characteristic function, survival, and hazard function are discussed in this chapter.

3.1 Develop a Two-stage Mixture Logistic Distribution

Mixture distributions are increasingly used to model heterogeneous data in various important practical situations, where the data can be viewed as arising from two or more subpopulation components. Mixture distribution was divided into two part, finite mixture with k components and infinite mixture distribution. In this thesis, we develop a finite mixture of logistic distribution with components $k = 2$ in terms of two-stage mixture distributions.

3.1.1 First Stage Mixture Logistic Distribution

In the first stage, we mix a point-degenerate distribution with a 2PL distribution using a mixing proportion parameter λ and $(1-\lambda)$, respectively. The CDF and PDF of 2PL distribution were clearly defined in Definition 1.1. By using the CDF in Definition

1.1, we mix the 2PL distribution with a point-degenerate distribution using a mixing parameter λ and $(1-\lambda)$ respectively. Finally, we define the CDF and PDF of the first stage mixture distribution in Definition 3.1.

Definition 3.1 *The CDF of the first stage mixture logistic distribution is given by*

$$F(x;\alpha,\beta,\lambda)=\begin{cases}\lambda & \text{at } x=x^* \\ \lambda+\dfrac{(1-\lambda)}{1+\exp\left[-\alpha(x-\beta)\right]} & \text{if } x>x^*\end{cases}. \tag{3.1}$$

and the PDF is given by

$$f(x;\alpha,\beta,\lambda)=\begin{cases}\lambda & \text{at } x=x^* \\ \dfrac{(1-\lambda)\alpha\exp\left[-\alpha(x-\beta)\right]}{\left[1+\exp\left\{-\alpha(x-\beta)\right\}\right]^2} & \text{if } x>x^*.\end{cases} \tag{3.2}$$

The first stage mixture distribution is popularly known as the 3PL distribution, because the PDF of the distribution contains 3 parameters which are α, β, and λ, respectively representing the scale, location, and mixing parameters. We can also think about the 3PL as a modified 2PL distribution.

3.1.2 Second Stage Mixture Logistic Distribution

In the second stage, the resulting distribution of the first stage is mixed with the 1PL distribution using the mixing proportion p and $(1-p)$ respectively. The IPL distribution was clearly defined in Definition 1.2. We define the CDF and PDF of a second stage mixture logistic distribution in Definition 3.2.

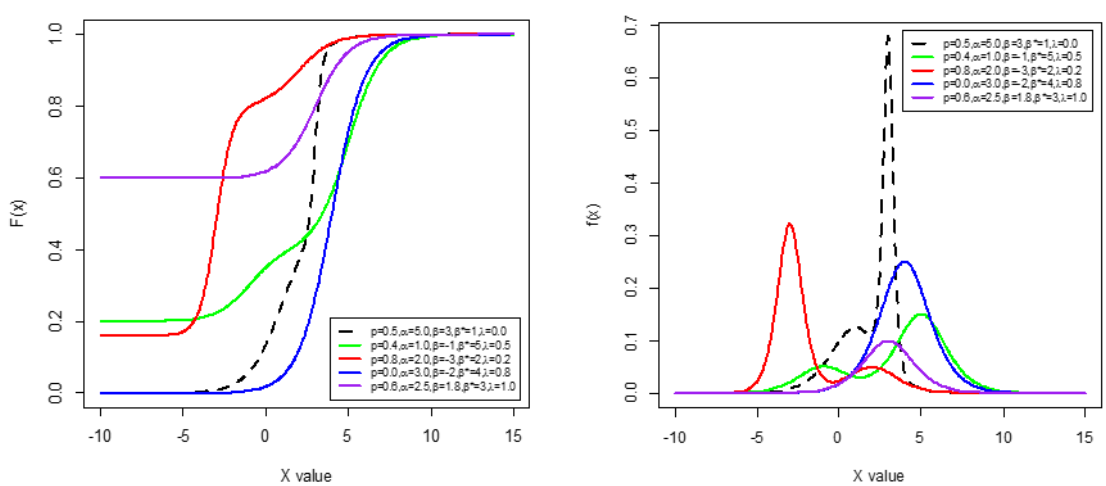
Definition 3.2 *The two-stage mixture logistic distribution has the CDF given by:*

$$F\left(x;\beta^*,\alpha,\beta,\lambda,p\right)=\begin{cases} p\lambda & \text{at } x=x^* \\ \dfrac{(1-p)}{1+\exp\left[-\left(x-\beta^*\right)\right]}+p\left(\lambda+\dfrac{(1-\lambda)}{1+\exp\left[-\alpha\left(x-\beta\right)\right]}\right) & \text{if } x>x^* \end{cases} \quad (3.3)$$

and the PDF is given by

$$f\left(x;\phi\right)=\begin{cases} p\cdot\lambda & \text{at } x=x^* \\ \dfrac{(1-p)\exp\left[-\left(x-\beta^*\right)\right]}{\left[1+\exp\left\{-\left(x-\beta^*\right)\right\}\right]^2}+\dfrac{p\alpha(1-\lambda)\exp\left[-\alpha\left(x-\beta\right)\right]}{\left[1+\exp\left\{-\alpha\left(x-\beta\right)\right\}\right]^2} & \text{at } x>x^*. \end{cases} \quad (3.4)$$

Figure 3.1 describes the CDF and PDF of the two-stage mixture distribution with different values of the parameters. The CDF has "*S*" curve form, but sometimes, forming some valley depends on the chosen parameter. The PDF shows that two-stage mixture logistic distribution is a symmetrical distribution, but possible to have 2 peaks (local maxima), depends on the chosen parameter.

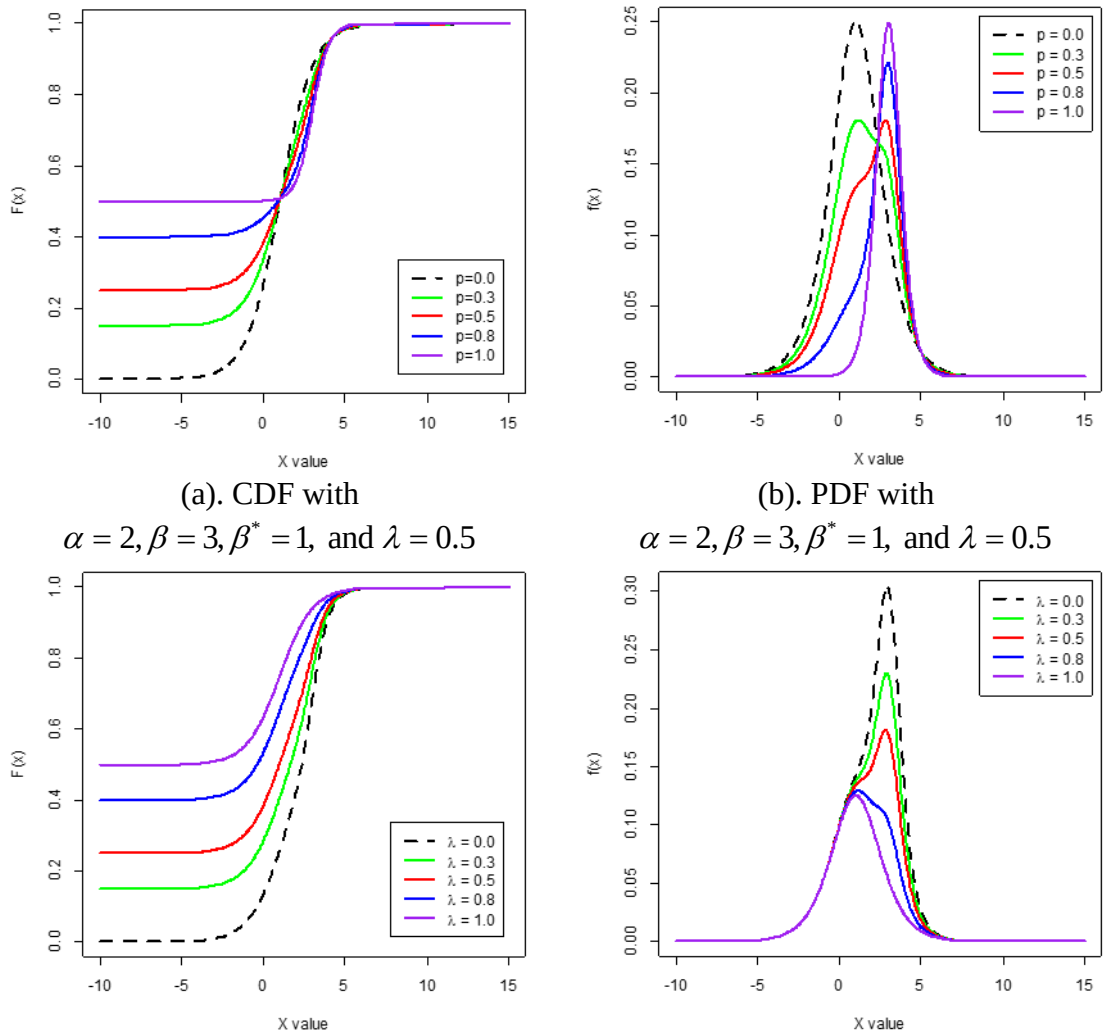


(a). CDF of Two-stage Mixture Logistic Distribution

(b). PDF of Two-stage Mixture Logistic Distribution

Figure 3.1 Graph of Two-stage Mixture Logistic Distribution

Figure 3.2 illustrates the CDF and PDF for different mixing proportions p and λ. The movement of p changes the dominant component of the CDF. In the PDF, when p is changed from 0 to 1, the peak of the graph move from left to the right side. From Figure 3.2 (d), we also see that the peak of the PDF will be lower when value of λ become higher.



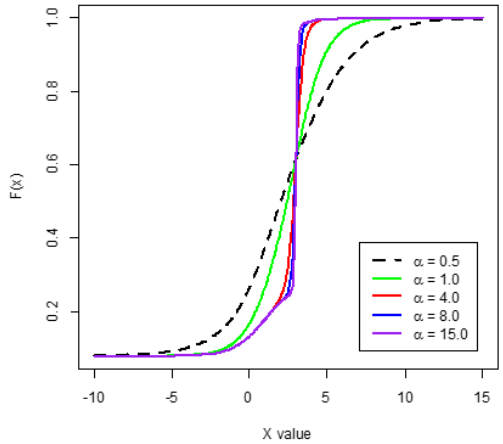
(a). CDF with $\alpha = 2, \beta = 3, \beta^* = 1,$ and $\lambda = 0.5$

(b). PDF with $\alpha = 2, \beta = 3, \beta^* = 1,$ and $\lambda = 0.5$

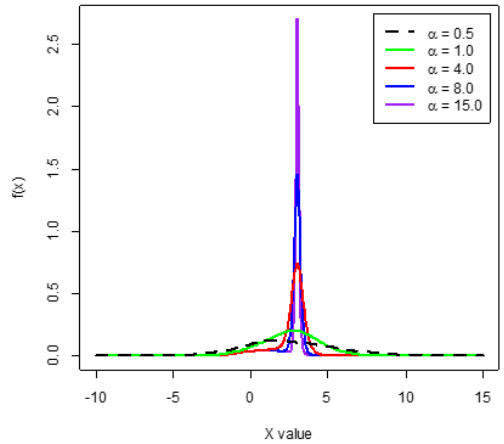
(c). CDF with $\alpha = 2, \beta = 3, \beta^* = 1,$ and $p = 0.5$

(d). PDF with $\alpha = 2, \beta = 3, \beta^* = 1,$ and $p = 0.5$

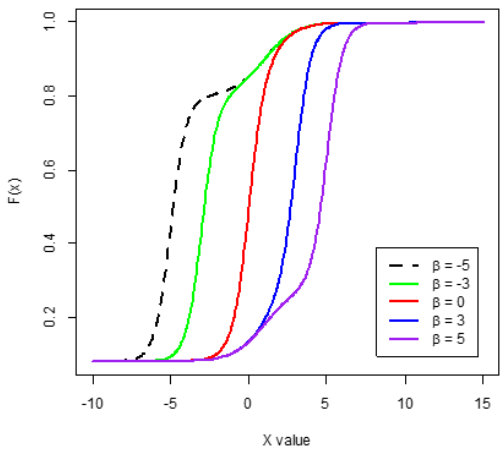
Figure 3.2 CDF and PDF with different mixing parameters of Two-stage Mixture Logistic Distribution



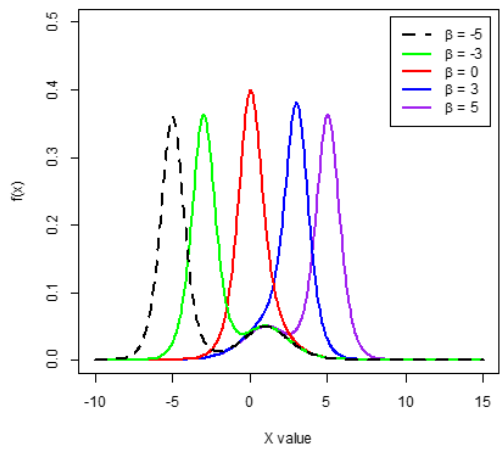
(a). CDF with
$p = 0.8, \beta = 3, \beta^* = 1,$ and $\lambda = 0.1$



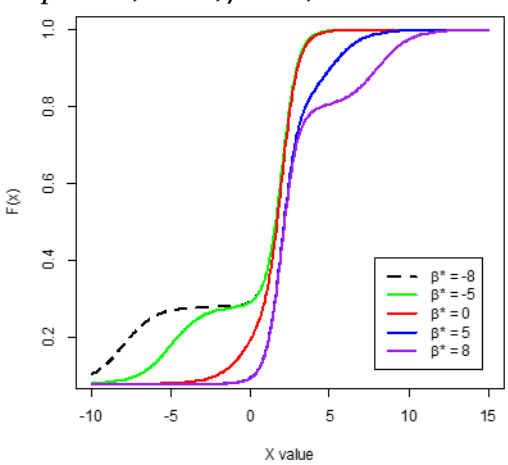
(b). PDF with
$p = 0.8, \beta = 3, \beta^* = 1,$ and $\lambda = 0.1$



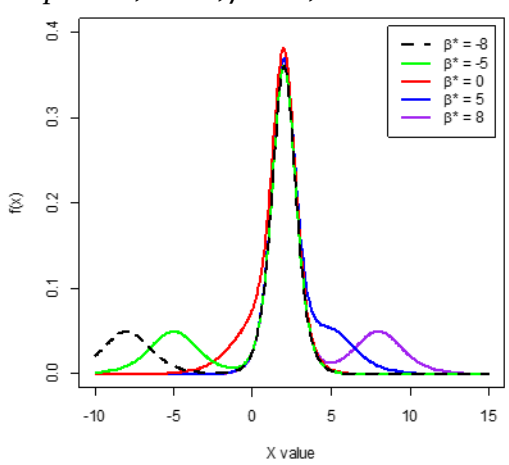
(c). CDF with
$p = 0.8, \alpha = 2, \beta^* = 1,$ and $\lambda = 0.1$



(d). PDF with
$p = 0.8, \alpha = 2, \beta^* = 1,$ and $\lambda = 0.1$



(e). CDF with
$p = 0.8, \alpha = 2, \beta = 2,$ and $\lambda = 0.1$



(f). PDF with
$p = 0.8, \alpha = 2, \beta = 2,$ and $\lambda = 0.1$

Figure 3.3 CDF and PDF with different component parameters of Two-stage Mixture Logistic Distribution

Figure 3.3 illustrates the CDF and PDF of the two-stage mixture model with different component parameters. Changing values of α to be larger shows steeper CDF in Figure 3.3 (a) and thus showing higher discrimination in Figure 3.3 (b). Changing values of β in Figure 3.3 (c) and (d) changes the location of the 3PL component of the model and changing values of β^* in Figure 3.3 (e) and (f) changes the location of 1PL component of the model.

Finally, we defined the second stage mixture logistic distribution with parameters $\phi = \left(\beta^*, \beta, \alpha, p, \lambda\right)$. This mixture distribution is basically mixing two logistic distributions. So, we can also call this distribution as finite mixture logistic distribution with $k = 2$. Some important properties of the two-stage mixture logistic distributions are provided in the next sections.

3.2 Moment Generating Function

The moment-generating function (MGF) of a random variable is an alternative specification of its probability distribution. It can be denoted by $M_X\left(t\right)$. We can obtain all the moments of X by differentiating $M_X\left(t\right)$ and then evaluating the result at $t = 0$. According to Ross [44], the MGF is defined as in Definition 3.3.

Definition 3.3 *The MGF of continuous random variable X for* $-h < t < h$ *where* $h \in \mathbb{R}$ *is defined by*

$$\begin{aligned} M_X\left(t\right) &= E\left[e^{tX}\right] \\ &= \int_{-\infty}^{\infty} e^{tX} f\left(x\right) dx. \end{aligned} \tag{3.5}$$

Definition 3.4 considered the moment generating function corresponding to the two components finite mixture model. Klugman et al. [45] and Ross [46] give the definition for the MGF for the k-components mixture distribution. In the following Definition 3.4, we used their definition when $k = 2$.

Definition 3.4 *The MGF corresponding to the two components finite mixture model is*

$$M(t) = pE\left[\exp(tX)\middle|J = 0\right] + qE\left[\exp(tX)\middle|J = 1\right] \tag{3.6}$$

Proof:

$$\begin{aligned} M(t) &= E\left[\exp(tX)\right] \\ &= E\left[\exp(tX)\middle|J = 0\right]P\left[J = 0\right] + E\left[\exp(tX)\middle|J = 1\right]P\left[J = 1\right] \\ &= pE\left[\exp(tX)\middle|J = 0\right] + qE\left[\exp(tX)\middle|J = 1\right]. \blacksquare \end{aligned}$$

Now, we provide the MGF of the two-stage mixture logistic distribution in Theorem 3.1.

Theorem 3.1 *For random variable X with density function* $f(x;\phi)$ *as defined in Equation* (3.4) *with parameter* $\phi = \left(\beta^*, \beta, \alpha, p, \lambda\right)$, *the MGF is given as*

$$\begin{aligned} M_X(t) = {} & \exp\left(tx^*\right)p\lambda + (1-p)\exp\left(t\beta^*\right)\Gamma(1-t)\Gamma(1+t) \\ & + p(1-\lambda)\exp(t\beta)\Gamma\left(1-\frac{t}{\alpha}\right)\Gamma\left(1+\frac{t}{\alpha}\right). \end{aligned} \tag{3.7}$$

Proof: First, we provide the MGF for the simple logistic distribution. We use this following result to get the result for two-stage mixture logistic distribution. For random variable X,

the PDF of the 2PL (see Johnson et al. [47]) as defined in Equation (1.2) with $\beta = 1/\alpha$ and $\alpha = \beta$. Then, we made the transformation $Y = \dfrac{X-\alpha}{\beta}$ and we obtained the PDF of Y is

$$f_Y(y) = \frac{\exp(-y)}{\left(1+\exp(-y)\right)^2} \tag{3.8}$$

and the CDF of Y is

$$F_Y(y) = \frac{1}{1+\exp(-y)}. \tag{3.9}$$

Now, we are interested in getting its moment generating function. Johnson et al. [47] provide the MGF of Y as:

$$M_Y(t) = B(1-t, 1+t), \tag{3.10}$$

where $B(.,.)$ is defined as a Beta function. Then, we transform the result with respect to X to get the MGF of X. Returning to the original form of the distribution and recalling that $X = \alpha + \beta Y$, we see that

$$\begin{aligned} M_X(t) &= E\left(\exp(tx)\right) = E\left(\exp\left(t(\alpha+\beta y)\right)\right) = E\left(\exp(t\alpha + t\beta y)\right) \\ &= E\left(\exp(t\alpha)\right)E\left(\exp(t\beta y)\right) \\ &= \exp(t\alpha)B(1-\beta t, 1+\beta t). \end{aligned} \tag{3.11}$$

Then, we derive the MGF of random variable from a two-stage mixture model defined in Equation (3.4).

$$M_X(t) = \int_{-\infty}^{\infty} \exp(tx) f(x)\,dx$$

$$= \sum_{x=x^*} \exp(tx) f(x) + \int_{x^*}^{\infty} \exp(tx) \frac{(1-p)\exp\left[-\left(x-\beta^*\right)\right]}{\left[1+\exp\left\{-\left(x-\beta^*\right)\right\}\right]^2}$$

$$+ \exp(tx) \frac{p\alpha(1-\lambda)\exp\left[-\alpha(x-\beta)\right]}{\left[1+\exp\left\{-\alpha(x-\beta)\right\}\right]^2} dx$$

$$= \exp\left(tx^*\right) P\left[X = x^*\right] + (1-p) \int_{-\infty}^{\infty} \exp(tx) \frac{\exp\left[-\left(x-\beta^*\right)\right]}{\left[1+\exp\left\{-\left(x-\beta^*\right)\right\}\right]^2} dx$$

$$+ p(1-\lambda) \int_{-\infty}^{\infty} \exp(tx) \frac{\alpha \exp\left[-\alpha(x-\beta)\right]}{\left[1+\exp\left\{-\alpha(x-\beta)\right\}\right]^2} dx$$

$$= \exp\left(tx^*\right) p\lambda + (1-p) M_1(t) + p(1-\lambda) M_2(t). \tag{3.12}$$

By using the result in Equation (3.12), then, we define $M_1(t)$ with refer to transformation $Y = \frac{X-\alpha}{\beta}$ equal to $Y = \left(X - \beta^*\right)$ in the first component and define $M_2(t)$ with referring $Y = \frac{X-\alpha}{\beta}$ equally to $Y = \alpha(X-\beta)$ in the second component.

Then, we obtain the result of $M_1(t)$ and $M_2(t)$:

$$M_1(t) = \exp\left(t\beta^*\right) B(1-t, 1+t) \text{ and}$$

$$M_2(t) = \exp(t\beta) B\left(1-\frac{t}{\alpha}, 1+\frac{t}{\alpha}\right). \tag{3.13}$$

Then, substituting Equation (3.13) into Equation (3.12), we finally obtained the result of the MGF of two-stage mixture distribution given by:

$$M_X(t) = \exp\left(tx^*\right) p\lambda + (1-p)\exp\left(t\beta^*\right) B(1-t, 1+t) + p(1-\lambda)\exp(t\beta) B\left(1-\frac{t}{\alpha}, 1+\frac{t}{\alpha}\right)$$

$$= \exp\left(tx^*\right) p\lambda + \left(1-p\right)\exp\left(t\beta^*\right) \frac{\Gamma\left(1-t\right)\Gamma\left(1+t\right)}{\Gamma\left(2\right)}$$

$$+ p\left(1-\lambda\right)\exp\left(t\beta\right)\frac{\Gamma\left(1-\frac{t}{\alpha}\right)\Gamma\left(1+\frac{t}{\alpha}\right)}{\Gamma\left(2\right)}$$

$$= \exp\left(tx^*\right) p\lambda + \left(1-p\right)\exp\left(t\beta^*\right) \Gamma\left(1-t\right)\Gamma\left(1+t\right)$$

$$+ p\left(1-\lambda\right)\exp\left(t\beta\right)\Gamma\left(1-\frac{t}{\alpha}\right)\Gamma\left(1+\frac{t}{\alpha}\right). \qquad \blacksquare$$

By differentiating $M_X\left(t\right)$ k times with respect to t and setting $t=0$, we can obtain the k-th moment about the origin. The moment of the two-stage mixture distribution is discussed in Section 3.3.

3.3 Moment of Distribution

Expected value of the random variable is one of the most important properties in probability theory. According to Ross [44], the expected value of random variable X is defined in Definition 3.5.

Definition 3.5 *The expected value of continuous random variable X is defined for all real values by*

$$E\left[X\right] = \int_{-\infty}^{\infty} xf\left(x\right)dx\,. \tag{3.14}$$

According to the expected value given in Equation (3.14), we considered the expected value related to the two components mixture model given by Klugman et al. [45] and Ross [46] in Definition 3.6.

Definition 3.6 *The expected value corresponding to the two components mixture model is given by*

$$E[X] = E[X|J=0]P[J=0] + E[X|J=1]P[J=1]$$

$$= pE[X|J=0] + qE[X|J=1]. \tag{3.15}$$

Definition 3.7 *The expected value for the k-moments corresponding to the two components mixture model is given by*

$$E(X^k) = E[X^k|J=0]P[J=0] + E[X^k|J=1]P[J=1]$$
$$= pE[X^k|J=0] + qE[X^k|J=1]. \tag{3.16}$$

The expected value of the two-stage mixture logistic distribution is defined in Theorem 3.2.

Theorem 3.2 *For random variable X with density function* $f(x;\phi)$ *as defined in Equation* (3.4) *with parameter* $\phi = (\beta^*, \beta, \alpha, p, \lambda)$, *the expected value is given as*

$$E[X] = x^* p\lambda + (1-p)\beta^* + p(1-\lambda)\beta. \tag{3.17}$$

Proof: We have two ways to give the proof of the expected value of *X*. First**,** we use the result of expected value for the 2PL model given by Johnson et al. [47] as:

$$E[X] = \beta. \tag{3.18}$$

Then, the expected value of X for the two-stage mixture logistic distribution is given by

$$
\begin{aligned}
E[X] &= \sum_{x=x^*} x f(x) + \int_{x^*}^{\infty} x f(x)\,dx \\
&= \sum_{x=x^*} x P\left[x = x^*\right] + \int_{x^*}^{\infty} x(1-p) f_1(x) + p(1-\lambda) f_2(x)\,dx \\
&= x^* p\lambda + (1-p)\int_{-\infty}^{\infty} x f_1(x)\,dx + p(1-\lambda)\int_{-\infty}^{\infty} f_2(x)\,dx \\
&= x^* p\lambda + (1-p) E_1(X) + p(1-\lambda) E_2(X). \qquad (3.19)
\end{aligned}
$$

This is the expected value of the two-stage mixture logistic distribution.

Then, by using Equation (3.18), we obtain the result of $E_1(X)$ and $E_2(X)$ as

$$
\begin{aligned}
E_1[X] &= \beta^* \text{ and} \\
E_2[X] &= \beta. \qquad (3.20)
\end{aligned}
$$

By substituting Equation (3.20) into Equation (3.19), we obtained the expected value of the two-stage mixture logistic distribution as given in Equation (3.17). ■

The second way to obtain the expected value of the distribution is by using the MGF result in Equation (3.7). We derive the first moment until the fourth moment of the distribution as well as other properties related to the moments.

To obtain the first moment or the expected value of the mixture distributions, we need to take the first derivative of the MGF and evaluate the result at $t = 0$.

$$
\begin{aligned}
M_X^{(1)}(t) &= \frac{\partial}{\partial t}\Big[\exp\left(tx^*\right) p\lambda + (1-p)\exp\left(t\beta^*\right)\Gamma(1-t)\Gamma(1+t) \\
&\quad + p(1-\lambda)\exp(t\beta)\Gamma\left(1-\frac{t}{\alpha}\right)\Gamma\left(1+\frac{t}{\alpha}\right)\Big] \\
&= p\lambda\frac{\partial}{\partial t}\exp\left(tx^*\right) + (1-p)\frac{\partial}{\partial t}\left[\exp\left(t\beta^*\right)\Gamma(1-t)\Gamma(1+t)\right] \\
&\quad + p(1-\lambda)\frac{\partial}{\partial t}\left[\exp(t\beta)\Gamma\left(1-\frac{t}{\alpha}\right)\Gamma\left(1+\frac{t}{\alpha}\right)\right]. \\
M_X^{(1)}(t)\big|_{t=0} &= x^* p\lambda + (1-p)\beta^* + p(1-\lambda)\beta.
\end{aligned}
$$

This is the first moment of the model which is the expected value given in (3.27). ■

After getting the expected value of the mixture model, we now define the variance of the distribution by using the second derivative of the MGF. The variance of two-stage mixture logistic distribution is defined in Theorem 3.3.

Theorem 3.3 *For random variable X with density function* $f\left(x;\phi\right)$ *as defined in Equation* (3.4) *with parameter* $\phi=\left(\beta^{*},\beta,\alpha,p,\lambda\right)$, *the variance is given by*

$$Var\left[X\right]=x^{*2}p\lambda+\left(1-p\right)\left[\frac{\pi^{2}}{3}+p\beta^{*}\left(\beta^{*}+2\beta\left(1-\lambda\right)\right)\right]$$
$$+\,p\left(1-\lambda\right)\left[\frac{\pi^{2}}{3\alpha^{2}}+\beta^{2}\left(1+p\left(1-\lambda\right)\right)\right].$$

Proof: By using the MGF result in Equation (3.7), we take the second derivative and setting the result at $t=0$ as follows.

$$M_{X}^{(2)}\left(t\right)=\frac{\partial^{2}}{\partial t^{2}}\left[\exp\left(tx^{*}\right)p\lambda+\left(1-p\right)\exp\left(t\beta^{*}\right)\Gamma\left(1-t\right)\Gamma\left(1+t\right)\right.$$
$$\left.+\,p\left(1-\lambda\right)\exp\left(t\beta\right)\Gamma\left(1-\frac{t}{\alpha}\right)\Gamma\left(1+\frac{t}{\alpha}\right)\right]$$

$$=p\lambda\frac{\partial^{2}}{\partial t^{2}}\exp\left(tx^{*}\right)+\left(1-p\right)\frac{\partial^{2}}{\partial t^{2}}\left[\exp\left(t\beta^{*}\right)\Gamma\left(1-t\right)\Gamma\left(1+t\right)\right]$$
$$+\,p\left(1-\lambda\right)\frac{\partial^{2}}{\partial t^{2}}\left[\exp\left(t\beta\right)\Gamma\left(1-\frac{t}{\alpha}\right)\Gamma\left(1+\frac{t}{\alpha}\right)\right]$$

$$M_{X}^{(2)}\left(t\right)\big|_{t=0}=x^{*2}p\lambda+\left(1-p\right)\left[\left(\beta^{*}\right)^{2}+\frac{\pi^{2}}{3}\right]+p\left(1-\lambda\right)\left[\beta^{2}+\frac{\pi^{2}}{3\alpha^{2}}\right]. \tag{3.21}$$

This is the second moment of the two-stage mixture logistic distribution.

By using the result of the second moment in Equation (3.21), we derive the variance of the two-stage mixture logistic distribution as follows.

$$Var(X) = E(X^2) - (E(X))^2$$

$$= x^{*2} p\lambda + (1-p)\left[\frac{\pi^2}{3} + (\beta^*)^2\right] + p(1-\lambda)\left[\frac{\pi^2}{3\alpha^2} + \beta^2\right] - \left[(1-p)\beta^* + p(1-\lambda)\beta\right]^2$$

$$= x^{*2} p\lambda + (1-p)\left[\frac{\pi^2}{3} + (\beta^*)^2\right] + p(1-\lambda)\left[\frac{\pi^2}{3\alpha^2} + \beta^2\right]$$
$$-\left[(1-p)^2(\beta^*)^2 + p^2(1-\lambda)^2\beta^2 + 2(1-p)\beta^* p(1-\lambda)\beta\right]$$

$$= x^{*2} p\lambda + (1-p)\left[\frac{\pi^2}{3} + (\beta^*)^2 + (1-p)(\beta^*)^2 + 2\beta^*\beta p(1-\lambda)\right]$$
$$+ p(1-\lambda)\left[\frac{\pi^2}{3\alpha^2} + \beta^2 + p(1-\lambda)\beta^2\right]$$

$$= x^{*2} p\lambda + (1-p)\left[\frac{\pi^2}{3} + p\beta^*\left(\beta^* + 2\beta(1-\lambda)\right)\right]$$
$$+ p(1-\lambda)\left[\frac{\pi^2}{3\alpha^2} + \beta^2\left(1 + p(1-\lambda)\right)\right].$$

Finally, this is the variance of the two-stage mixture logistic distributions. ■

We are also interested in getting the skewness of the two-stage mixture logistic distribution. To obtain the skewness of the distribution, we need to find the third moment of the distribution and then solve Equation (3.22).

$$Skew(X) = E\left[\left(\frac{X - E(X)}{\sigma}\right)^3\right] = \frac{E(X^3) - 3E(X)Var(X) - (E(X))^3}{(Var(X))^{\frac{3}{2}}} \tag{3.22}$$

The skewness of two-stage mixture logistic distribution is given in Theorem 3.4.

Theorem 3.4 *For random variable X with density function* $f\left(x;\phi\right)$ *as defined in Equation* (3.4) *with parameter* $\phi=\left(\beta^{*},\beta,\alpha,p,\lambda\right)$, *the skewness is given by*

$$Skew\left(X\right)=\frac{A-B+C}{\left(D\right)^{\frac{3}{2}}}, \tag{3.23}$$

where

$$A=p\lambda\left(x^{*}\right)^{3}+\left(1-p\right)\left(\pi^{2}\beta^{*}+\left(\beta^{*}\right)^{3}\right)+p\left(1-\lambda\right)\left(\frac{\pi^{2}\beta}{\alpha^{2}}+\beta^{3}\right),$$

$$B=3\left(p\lambda x^{*}+\left(1-p\right)\beta^{*}+p\left(1-\lambda\right)\beta\right)\left(p\lambda\left(x^{*}\right)^{2}+\left(1-p\right)\left(\frac{\pi^{2}}{3}+p\beta^{*}\left(\beta^{*}+2\beta\left(1-\lambda\right)\right)\right)\right),$$

$$C=p\left(1-\lambda\right)\left(\frac{\pi^{2}}{3\alpha^{2}}+\beta^{2}\left(1+p\left(1-\lambda\right)\right)\right)\Bigg)-\left(p\lambda x^{*}+\left(1-p\right)\beta^{*}+p\left(1-\lambda\right)\beta\right)^{3}, \text{ and}$$

$$D=p\lambda\left(x^{*}\right)^{2}+\left(1-p\right)\left(\frac{\pi^{2}}{3}+p\beta^{*}\left(\beta^{*}+2b\left(1-\lambda\right)\right)\right)+p\left(1-\lambda\right)\left(\frac{\pi^{2}}{3\alpha^{2}}+\beta^{2}\left(1+p\left(1-\lambda\right)\right)\right).$$

Proof: First, we derive the third moment of two-stage mixture logistic distribution by taking the third derivative of MGF in Equation (3.7) and evaluate the result at $t=0$.

$$\begin{aligned}M_{X}^{(3)}\left(t\right)=\frac{\partial^{3}}{\partial t^{3}}\Big[&\exp\left(tx^{*}\right)p\lambda+\left(1-p\right)\exp\left(t\beta^{*}\right)\Gamma\left(1-t\right)\Gamma\left(1+t\right)\\&+p\left(1-\lambda\right)\exp\left(t\beta\right)\Gamma\left(1-\frac{t}{\alpha}\right)\Gamma\left(1+\frac{t}{\alpha}\right)\Big]\end{aligned}$$

$$\begin{aligned}&=p\lambda\frac{\partial^{3}}{\partial t^{3}}\exp\left(tx^{*}\right)+\left(1-p\right)\frac{\partial^{3}}{\partial t^{3}}\left[\exp\left(t\beta^{*}\right)\Gamma\left(1-t\right)\Gamma\left(1+t\right)\right]\\&+\left(1-\lambda\right)\frac{\partial^{3}}{\partial t^{3}}\left[\exp\left(t\beta\right)\Gamma\left(1-\frac{t}{\alpha}\right)\Gamma\left(1+\frac{t}{\alpha}\right)\right].\end{aligned}$$

$$M_{X}^{(3)}\left(t\right)\big|_{t=0}=p\lambda\left(x^{*}\right)^{3}+\left(1-p\right)\left[\left(\beta^{*}\right)^{3}+\beta^{*}\pi^{2}\right]+p\left(1-\lambda\right)\left[\beta^{3}+\frac{\beta\pi^{2}}{\alpha^{2}}\right]. \tag{3.24}$$

This is the third moment of the two-stage mixture logistic distribution.

Then, substituting the expected value, the variance, and the third moment into the skewness expression in Equation (3.22), we get the skewness of two-stage mixture logistic distribution given in Equation (3.23). ■

As the family of symmetric distribution, logistic distribution has the value of skewness equal to 0. But data point may not be perfectly symmetric. So, an understanding of the skewness of the data set indicates whether deviations from the mean are going to be positive or negative. After getting the skewness of the distribution, we next define the fourth moment as well as the kurtosis of the distribution.

Kurtosis itself measures the "*tailedness*" of the probability distribution of a real value random variable. Kurtosis measures whether the data is peaked or flat relative to a normal distribution. That is, data set with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails. Data set with low kurtosis tend to have a flat top near the mean rather than a sharp peak. The kurtosis of a random variable is defined in Definition 3.8.

Definition 3.8 *The kurtosis of random variable X is given by*

$$Kurt(X) = E\left[\left(\frac{X - E(X)}{\sigma}\right)^4\right]$$
$$= \frac{E(X^4) - 4E(X^3)E(X) + 6E(X^2)(E(X))^2 - 3(E(X))^4}{(Var(X))^2}. \qquad (3.25)$$

To get the kurtosis of a two-stage mixture logistic distribution, we need to define the fourth moment of the distribution.

Theorem 3.5 *For random variable X with density function* $f\left(x;\phi\right)$ *as defined in Equation* (3.4) *with parameter* $\phi=\left(\beta^{*},\beta,\alpha,p,\lambda\right)$, *the kurtosis is given by*

$$Kurt\left(X\right)=\frac{1}{K}\left(L+M-N+O-P\right), \tag{3.26}$$

where

$$K=\left[p\lambda\left(x^{*}\right)^{2}+\left(1-p\right)\left(\frac{\pi^{2}}{3}+p\beta^{*}\left(\beta^{*}+2\beta\left(1-\lambda\right)\right)\right)+p\left(1-\lambda\right)\left(\frac{\pi^{2}}{3\alpha^{2}}+\beta^{2}\left(1+p\left(1-\lambda\right)\right)\right)\right]^{2},$$

$$L=p\lambda\left(x^{*}\right)^{4}+\left(1-p\right)\left[\frac{7\pi^{4}}{15}+2\pi^{2}\beta^{*2}+\left(\beta^{*}\right)^{4}\right],$$

$$M=p\left(1-\lambda\right)\left[\frac{7\pi^{4}}{15\alpha^{4}}+\frac{2\pi^{2}\beta^{2}}{\alpha^{2}}+\beta^{4}\right],$$

$$N=4\left(p\lambda x^{*}\left(1-p\right)\beta^{*}+p\left(1-\lambda\right)\beta\right)\left(p\lambda\left(x^{*}\right)^{3}\left(1-p\right)\left(\pi^{2}\beta^{*}+\left(\beta^{*}\right)^{3}\right)+p\left(1-\lambda\right)\left(\frac{\pi^{2}\beta}{\alpha^{2}}+\beta^{3}\right)\right),$$

$$O=6\left(p\lambda x^{*}+\left(1-p\right)\beta^{*}+p\left(1-\lambda\right)\beta\right)^{2}\left(p\lambda\left(x^{*}\right)^{2}+\left(1-p\right)\left(\frac{\pi^{2}}{3}+\left(\beta^{*}\right)^{2}\right)+p\left(1-\lambda\right)\left(\frac{\pi^{2}}{3\alpha^{2}}+\beta^{2}\right)\right),$$

$$P=3\left(p\lambda x^{*}+\left(1-p\right)\beta^{*}+p\left(1-\lambda\right)\beta\right)^{4}.$$

Proof: First, we derive the fourth moment of the mixture model by taking the fourth derivative of the MGF and evaluate the result at $t=0$.

$$\begin{aligned}M_{X}^{(4)}\left(t\right)=\frac{\partial^{4}}{\partial t^{4}}\Big[&\exp\left(tx^{*}\right)p\lambda+\left(1-p\right)\exp\left(t\beta^{*}\right)\Gamma\left(1-t\right)\Gamma\left(1+t\right)\\&+p\left(1-\lambda\right)\exp\left(t\beta\right)\Gamma\left(1-\frac{t}{\alpha}\right)\Gamma\left(1+\frac{t}{\alpha}\right)\Big]\end{aligned}$$

$$\begin{aligned}&=p\lambda\frac{\partial^{4}}{\partial t^{4}}\exp\left(tx^{*}\right)+\left(1-p\right)\frac{\partial^{4}}{\partial^{4}t^{4}}\left[\exp\left(t\beta^{*}\right)\Gamma\left(1-t\right)\Gamma\left(1+t\right)\right]\\&+p\left(1-\lambda\right)\frac{\partial^{4}}{\partial^{4}t^{4}}\left[\exp\left(t\beta\right)\Gamma\left(1-\frac{t}{\alpha}\right)\Gamma\left(1+\frac{t}{\alpha}\right)\right].\end{aligned}$$

$$M_X^{(4)}(t)\big|_{t=0} = p\lambda\left(x^*\right)^4 + (1-p)\left[\left(\beta^*\right)^4 + 2\pi^2\left(\beta^*\right)^2 + \frac{7\pi^4}{15}\right] + p(1-\lambda)\left[\beta^4 + \beta^2\frac{2\pi^2}{\alpha^2} + \frac{7\pi^4}{15\alpha^4}\right]. \tag{3.27}$$

This is the fourth moment of the two-stage mixture logistic distribution.

Then, by using the fourth moment of the mixture model, we derive the kurtosis of the model using Definition 3.8. By substituting the expected value, the variance, the second moment, the third moment, and the fourth moment into the kurtosis in Equation (3.25), the kurtosis of the model is given in Equation (3.26). ■

We have completed the moments of two-stage mixture logistic distribution as well as the related properties such as variance, skewness, and kurtosis. In Section 3.4, we discuss the characteristic function of the two-stage mixture logistic distribution.

3.4 Characteristic Function

Hazewinkel [48] explained that the characteristic function is the inverse of the Fourier transform of the probability density function. The characteristic function is the moment-generating function of i or the moment generating a function of X evaluated on the imaginary axis. Note however that the characteristic function of a distribution always exists, even when the probability density function or moment-generating function do not. According to Ross [44], a characteristic function can be defined as follows.

Definition 3.9 *The characteristic function of continuous random variable X is defined for all real values of t by*

$$\varphi(t) = E\left[\exp(itX)\right] = \int_{-\infty}^{\infty} \exp(itX)\, f(x)\, dx,\ t \in \mathbb{R},\ \text{and } i = \sqrt{-1}. \tag{3.28}$$

According to the characteristic function in Equation (3.28), we now consider the characteristic function related to the two components mixture model (see Klugman et al. [45] and Ross [46]) as follows.

Definition 3.10 *The characteristic function corresponding to the two components mixture model is given by*

$$\begin{aligned}\varphi(t) &= E\left[\exp(itX)\right] \\ &= E\left[\exp(itX)\middle|J=0\right]P\left[J=0\right]+E\left[\exp(itX)\middle|J=1\right]P\left[J=1\right] \\ &= pE\left[\exp(itX)\middle|J=0\right]+qE\left[\exp(itX)\middle|J=1\right].\end{aligned} \tag{3.29}$$

Now, we provide the characteristic function for the two-stage mixture logistic distribution as defined in Equation (3.4).

Theorem 3.6 *For random variable X with density function* $f(x;\phi)$ *as defined in Equation* (3.4) *with parameter* $\phi=\left(\beta^*,\beta,\alpha,p,\lambda\right)$*, the characteristic function is given by*

$$\begin{aligned}\varphi_X(t) &= \exp\left(itx^*\right)p\lambda+(1-p)\exp\left(it\beta^*\right)\frac{\pi t}{1.7\sinh(\pi t)} \\ &\quad + p(1-\lambda)\exp(it\beta)\frac{\pi t}{\alpha\sinh\left(\frac{\pi t}{\alpha}\right)}.\end{aligned} \tag{3.30}$$

Proof: First, we refer to the result of the characteristic function for the 2PL distribution given by Johnson et al.[47] as follows:

$$\varphi_X(t)=\exp(it\alpha)\frac{\pi\beta t}{\sinh(\pi\beta t)}. \tag{3.31}$$

Then, the characteristic function of the two-stage mixture logistic distribution is given by

$$\varphi_X(t) = \sum_{x=x^*} \exp(itx) f(x) + \int_{-\infty}^{\infty} \exp(itx) f(x) dx$$

$$= \sum_{x=x^*} \exp(itx) P\left[x = x^*\right] + \int_{-\infty}^{\infty} \exp(itx) \frac{(1-p)\exp\left[-\left(x-b^*\right)\right]}{\left[1+\exp\left\{-\left(x-b^*\right)\right\}\right]^2}$$

$$+ \exp(itx) \frac{pa(1-c)\exp\left[-a(x-b)\right]}{\left[1+\exp\left\{-a(x-b)\right\}\right]^2} dx$$

$$= \exp\left(itx^*\right) p\lambda + (1-p) \int_{-\infty}^{\infty} \exp(itx) \frac{\exp\left[-\left(x-b^*\right)\right]}{\left[1+\exp\left\{-\left(x-b^*\right)\right\}\right]^2} dx$$

$$+ p(1-c) \int_{-\infty}^{\infty} \exp(itx) \frac{a\exp\left[-a(x-b)\right]}{\left[1+\exp\left\{-a(x-b)\right\}\right]^2} dx$$

$$= \exp\left(itx^*\right) p\lambda + (1-p)\varphi_1(t) + p(1-c)\varphi_2(t). \tag{3.32}$$

By using the same procedure in getting the MGF of the mixture distribution, we use transformation procedure as defined in Section 3.2 for getting the characteristic function. Then, we obtain the result of $\varphi_1(t)$ and $\varphi_2(t)$ as follows.

$$\varphi_1(t) = \exp\left(it\beta^*\right) \frac{\pi t}{1.7\sinh(\pi t)} \text{ and}$$

$$\varphi_2(t) = p(1-\lambda)\exp(it\beta) \frac{\pi t}{\alpha \sinh\left(\frac{\pi t}{\alpha}\right)}. \tag{3.33}$$

By substituting Equation (3.33) into Equation (3.32), we obtain the characteristic function given in Equation (3.30). ■

Now, we have the characteristic function of two-stage mixture logistic distribution. Then, in Section 3.5, we discuss the survival function of two-stage mixture logistic distribution.

3.5 Survival Function

The survival function $S(t)$ gives the probability that a person survives longer than some specified time *t*, (see Kleinbaum & Klein [49]). Theoretically, as *t* ranges from 0 up to infinity, the survival function can be graphed as a smooth curved.

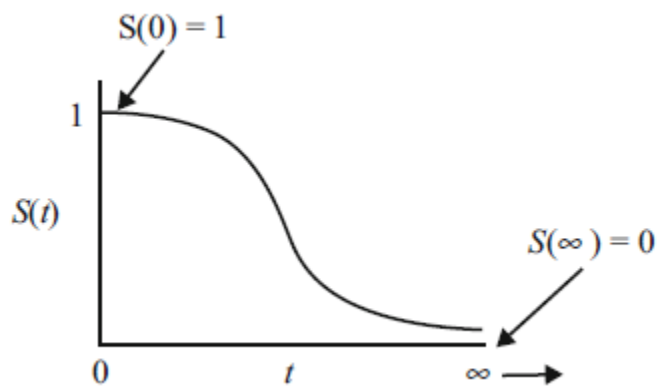


Figure 3.4 Graph of the Survival Function $S(t)$ (Kleinbaum and Klein [49])

Figure 3.4 illustrates that at the time $t=0$, $S(t)|_{t=0}=1$, that is, the probability of surviving past time 0 is one. Otherwise, at $t=\infty$, $S(t)|_{t=\infty}=0$, that is, if the study period increased without limit, eventually nobody would survive, so the survivor curve must eventually fall to zero. The survival function $S(t)$ corresponding to the mixture model is given by Al-Hussaini & Abd-El-Hakim [26] in Definition 3.11.

Definition 3.11 *The survival function of random variable X for two-component mixture distribution with CDF* $F_X(t)$ *and mixing proportion* p *is defined by*

$$S_X(t) = pS_1(t) + (1-p)S_2(t), \tag{3.34}$$

where $S_1(t) = 1 - F_1(t), S_2(t) = 1 - F_2(t)$, *but* $F_1(t)$ *and* $F_2(t)$ *are respectively the CDF of the first and the second component of the two-stage mixture logistic distribution.*

The survival function is a property of any random variable that explains a set of events, usually associated with mortality or failure of some system in the set of time. It describes the probability that the system will survive beyond a specified time. In terms of the two-stage mixture logistic distribution, we develop a survival function as follows.

Theorem 3.7 *For survival time t with distribution function* $F(t)$ *and parameter* $\phi = (\beta^*, \beta, \alpha, p, \lambda)$, *the survival function is given by*

$$S(t) = \frac{(1-p)\exp\left[-\left(t-\beta^*\right)\right]}{1+\exp\left[-\left(t-\beta^*\right)\right]} + \frac{p(1-\lambda)\exp\left[-\alpha(t-\beta)\right]}{1+\exp\left[-\alpha(t-\beta)\right]}. \tag{3.35}$$

Proof: Firstly, we derive $S_1(t)$ and $S_2(t)$ as follows:

$$\begin{aligned} S_1(t) = 1 - F_1(t) &= 1 - \frac{1}{1+\exp\left[-\left(t-\beta^*\right)\right]} \\ &= \frac{1+\exp\left[-\left(t-\beta^*\right)\right]-1}{1+\exp\left[-\left(t-\beta^*\right)\right]} \\ &= \frac{\exp\left[-\left(t-\beta^*\right)\right]}{1+\exp\left[-\left(t-\beta^*\right)\right]} \end{aligned} \tag{3.36}$$

and

$$\begin{aligned} S_2(t) = 1 - F_2(t) &= 1 - \left(\lambda + \frac{(1-\lambda)}{1+\exp\left[-\alpha(t-\beta)\right]}\right) \\ &= (1-\lambda) - \frac{(1-\lambda)}{1+\exp\left[-\alpha(t-\beta)\right]} \end{aligned}$$

$$= \frac{(1-\lambda)\left(1+\exp\left[-\alpha(t-\beta)\right]\right)-(1-\lambda)}{1+\exp\left[-\alpha(t-\beta)\right]}$$

$$= \frac{(1-\lambda)\left\{1+\exp\left[-\alpha(t-\beta)\right]-1\right\}}{1+\exp\left[-\alpha(t-\beta)\right]}$$

$$= \frac{(1-\lambda)\exp\left[-\alpha(t-\beta)\right]}{1+\exp\left[-\alpha(t-\beta)\right]}. \tag{3.37}$$

Then, the survival function is obtained by substituting $S_1(t)$ in Equation (3.36) and $S_2(t)$ in Equation (3.37) into the survival function for mixture distribution in Equation (3.34). Finally, we obtain the survival function of the two-stage mixture logistic distribution as described in Theorem 3.7. ■

3.5.1 Survival Odds

The survival odds is the odds of surviving beyond time *t*; i.e., $S(t)/(1-S(t))$ (see for example Kleinbaum & Klein [49]). This is the probability of not getting the event by time *t* divided by the probability of getting the event by time *t*. The failure odds is the reciprocal of the survival odds.

a. Survival Odds of Two-stage Mixture Logistic Model

The two-stage mixture logistic distribution with CDF defined in Equation (3.3) has survival odds as follows.

$$\frac{S(t)}{1-S(t)} = \frac{1-F(t)}{F(t)}$$

$$= \frac{\left[(1-p\lambda) - \frac{(1-p)}{\left(1+\exp\left[-\left(t-\beta^*\right)\right]\right)} - \frac{p(1-\lambda)}{\left(1+\exp\left[-\alpha(t-\beta)\right]\right)}\right]}{\left[p\lambda + \frac{(1-p)}{\left(1+\exp\left[-\left(t-\beta^*\right)\right]\right)} + \frac{p(1-\lambda)}{\left(1+\exp\left[-\alpha(t-\beta)\right]\right)}\right]}. \tag{3.38}$$

Then, for the log of survival odds of two-stage mixture logistic distribution, it has a complicated model. Hence, we are trying to define the special interest condition for the parameter.

b. Survival odds of 1PL model

For the condition $p = 0$, the two-stage mixture logistic distribution becomes the simple logistic distribution with 1 parameters β^*. The CDF of 1PL has defined in Equation (1.3) and the survival odds defined as follows.

Theorem 3.8 *The log of survival odds for the 1PL model can be given as* $\beta^* - t$ *where the slope is -1 and the intercept is* β^*.

Proof:

$$\frac{S(t)}{1-S(t)} = \frac{1-F(t)}{F(t)} = \frac{1-\frac{1}{1+\exp\left[-\left(t-\beta^*\right)\right]}}{\frac{1}{1+\exp\left[-\left(t-\beta^*\right)\right]}} = \exp\left[-\left(t-\beta^*\right)\right].$$

$$\text{Log (Survival odds)} = \ln\left(\exp\left[-\left(t-\beta^*\right)\right]\right) = \beta^* - t\,. \blacksquare \tag{3.39}$$

The log odds of survival of 1PL model is a linear function of the log of time with slope -1 and intercept β^*. This is a useful result enabling a graphical evaluation of the appropriateness of the 1PL. The other 1PL model for the parameter β can also be derived by defining a special condition for $p=1$, $\alpha=1$ and $\lambda=0$. Then, the log of survival odds is $\beta - t$.

c. Survival Odds of 2PL model

For the condition $p=1$ and $\lambda=0$, the two-stage mixture logistic distribution becomes the simple logistic distribution with 2 parameters. The CDF of 2PL has defined in (1.1) and the survival odds defined as follows.

Theorem 3.9 *The log of survival odds for 2PL model can be given as* $\alpha\beta - \alpha t$ *where the slope is* $-\alpha$ *and the intercept is* $\alpha\beta$.

Proof:

$$\frac{S(t)}{1-S(t)} = \frac{1-F(t)}{F(t)} = \frac{1-\dfrac{1}{1+\exp\left[-\alpha(t-\beta)\right]}}{\dfrac{1}{1+\exp\left[-\alpha(t-\beta)\right]}} = \exp\left[-\alpha(t-\beta)\right].$$

$$\text{Log (Survival odds)} = \ln\left(\exp\left[-\alpha(t-\beta)\right]\right) = \alpha\beta - \alpha t. \blacksquare \qquad (3.40)$$

The log odds of survival of 2PL model is a linear function of the log of time with slope $-\alpha$ and intercept $\alpha\beta$. This is a useful result enabling a graphical evaluation of the appropriateness of the 2PL.

d. Survival Odds of 3PL Model

For the condition $p=1$, the two-stage mixture logistic distribution becomes the simple logistic distribution with 3 parameters. The CDF of 3PL has defined in Equation (3.1) and the survival odds defined as follows.

Theorem 3.10 *The log of survival odds for 3PL model can be given by*

$$\ln\left(odds\right)=\ln\left(1-\lambda\right)+\alpha\beta-\alpha t-\ln\left(1+\lambda\exp\left[\alpha\left(\beta-t\right)\right]\right). \tag{3.41}$$

Proof: The survival odds of the 3PL model is given by

$$\frac{S\left(t\right)}{1-S\left(t\right)}=\frac{1-F\left(t\right)}{F\left(t\right)}=\frac{\dfrac{\left(1-\lambda\right)\exp\left[-\alpha\left(t-\beta\right)\right]}{1+\exp\left[-\alpha\left(t-\beta\right)\right]}}{\dfrac{1+\lambda\exp\left[-\alpha\left(t-\beta\right)\right]}{1+\exp\left[-\alpha\left(t-\beta\right)\right]}}=\frac{\left(1-\lambda\right)\exp\left[-\alpha\left(t-\beta\right)\right]}{1+\lambda\exp\left[-\alpha\left(t-\beta\right)\right]}. \tag{3.42}$$

By taking log of Equation (3.42), we define the log of survival odds as follows:

$$\begin{aligned}\text{Log (survival odds)} &= \ln\left(\frac{\left(1-\lambda\right)\exp\left[-\alpha\left(t-\beta\right)\right]}{1+\lambda\exp\left[-\alpha\left(t-\beta\right)\right]}\right)\\ &=\ln\left(1-\lambda\right)+\alpha\beta-\alpha t-\ln\left(1+\lambda\exp\left[\alpha\left(\beta-t\right)\right]\right)\\ &=\ln\left(1-\lambda\right)+\alpha\beta-\alpha t-\ln\left(1+\frac{\lambda\exp\left(\alpha\beta\right)}{\exp\left(\alpha t\right)}\right)\\ &=\ln\left(1-\lambda\right)+\alpha\beta-\ln\left(\exp\left(\alpha t\right)+\lambda\exp\left(\alpha\beta\right)\right).\end{aligned} \tag{3.43}$$

We also defined the modified survival odds of 3PL model which is the ratio of success to modified failure as follows.

$$\frac{S(t)}{(1-S(t))-\lambda}=\frac{1-F(t)}{F(t)-\lambda}=\frac{\dfrac{(1-\lambda)\exp\left[-\alpha(t-\beta)\right]}{1+\exp\left[-\alpha(t-\beta)\right]}}{\dfrac{1+\lambda\exp\left[-\alpha(t-\beta)\right]}{1+\exp\left[-\alpha(t-\beta)\right]}-\lambda}$$

$$=\frac{(1-\lambda)\exp\left[-\alpha(t-\beta)\right]}{1+\lambda\exp\left[-\alpha(t-\beta)\right]-\lambda\left(1+\exp\left[-\alpha(t-\beta)\right]\right)}$$

$$=\exp\left[-\alpha(t-\beta)\right]. \tag{3.44}$$

Now, the modified survival odds of 3PL model has the same result with the survival odds of 2PL model. We have completed the survival function and the odds of two-stage mixture logistic distribution. In Section 3.6, we discuss the Hazard function of two-stage mixture logistic distribution.

3.6 Hazard Function

After getting the survival function, now we are interested in finding the hazard function of the distribution. The hazard function, denoted by $h(t)$, gives the instantaneous potential per unit time for the event to occur, given that the individual has survived up to time *t*, (see Kleinbaum & Klein [49]). In contrast to the survival function, which focuses on not failing, the hazard function focuses on failing, that is on the occurring event.

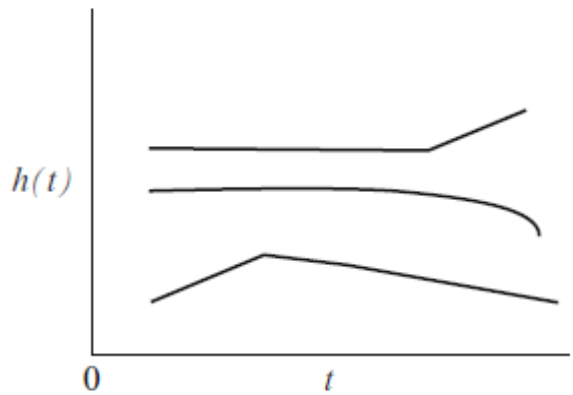


Figure 3.5 Graph of the Hazard Function $h(t)$ (Kleinbaum and Klein [49])

Figure 3.5 illustrates three different hazards. The graph of $h(t)$ does not have to start at 1 and go down to zero, but rather can start anywhere and go up and down in any direction over time. The hazard function always nonnegative and it has no upper bound. The hazard function of survival time *t* is given by Al-Hussaini & Abd-El-Hakim [26] in Definition 3.12.

Definition 3.12 *The hazard function of survival time t for two mixture distribution with PDF* $f(t)$ *is defined by*

$$h(t) = \frac{f(t)}{S(t)} = \frac{pf_1(t) + qf_2(t)}{pS_1(t) + qS_2(t)}. \tag{3.45}$$

The hazard function or usually we called failure rate is the frequency that expresses the component fails per unit of time *t*. In terms of two-stage mixture logistic distribution, we define the hazard function in Theorem 3.11.

Theorem 3.11 *For survival time t with density function* $f(t)$ *and parameter* $\phi = (\beta^*, \beta, \alpha, p, \lambda)$, *the hazard function is given by*

$$h(t) = \frac{\dfrac{(1-p)\exp\left[-\left(t-\beta^*\right)\right]}{\left[1+\exp\left\{-\left(t-\beta^*\right)\right\}\right]^2} + \dfrac{p(1-\lambda)a\exp\left[-\alpha(t-\beta)\right]}{\left[1+\exp\left\{-\alpha(t-\beta)\right\}\right]^2}}{\dfrac{(1-p)\exp\left[-\left(t-\beta^*\right)\right]}{1+\exp\left[-\left(t-\beta^*\right)\right]} + \dfrac{p(1-\lambda)\exp\left[-\alpha(t-\beta)\right]}{1+\exp\left[-\alpha(t-\beta)\right]}} \quad . \tag{3.46}$$

Proof: For getting the hazard function for mixture model, we apply the hazard function formula in Equation (3.45). Finally, we expressed the hazard function as described in Theorem 3.11. ■

CHAPTER 4

PARAMETER ESTIMATION OF TWO STAGE MIXTURE LOGISTIC DISTRIBUTION

In this chapter, we discuss research question 2 (parameter estimation of distribution). We derive the point estimators of the parameters of interest in the distribution. The parameters of the distribution are commonly called the characteristics of the population. Practically, the values of parameters are unknown, so, we need to estimate the population parameters for a random sample. We discuss two different methods of parameter estimation for two-stage mixture logistic distribution which are method of moment and maximum likelihood estimation.

4.1 Method of Moment Estimation

The method of moment estimation is corresponding to the law of large numbers. In the case of one parameter. Suppose $X_1, X_2, ..., X_n$ be random variables from the probability density $f_X\left(x|\phi\right)$ associated with an unknown parameter ϕ. If X_i are continuous random variables, then the law of large numbers is explained in Definition 4.1.

Definition 4.1 *The law of large number states that*

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \to \mu_X \text{ as } n \to \infty . \tag{4.1}$$

Thus, if the number of observations *n* is large, the distributional mean, μ_X should be well approximated by the sample mean. By setting $\bar{X} \approx \mu_X$ and solving for $\hat{\phi}$, the

method of moment estimator is obtained. For more than one parameter (for example p parameters), the method of moment estimator can be obtained by solving the Equation (4.2) equations simultaneously.

$$E\left[X^k\right] = m_k, \;\; k = 1, 2, \ldots p, \tag{4.2}$$

where $m_k = \frac{1}{n}\sum_{k=1}^{p} x^k$ is sample moment.

Since there are five unknown parameters $\phi = \left(\beta^*, \beta, \alpha, p, \lambda\right)$ of the model, we need to derive the first moment to the fifth moment of the two-stage mixture logistic distribution as defined in Equation (3.4). For the first to the fourth moment, we have already provided in Chapter 3. We derive the fifth moment of the mixture model by obtaining the fifth derivative of the MGF and then evaluate the result at $t = 0$.

Theorem 4.1 *For random variable X with density function* $f\left(x;\phi\right)$ *as defined in Equation* (3.4) *with parameter* $\phi = \left(\beta^*, \beta, \alpha, p, \lambda\right)$, *the fifth moment is given by*

$$\begin{aligned} E\left(X^5\right) = {} & p\lambda x^{*5} + \left(1-p\right)\left[\frac{7\beta^*\pi^4}{3} + \frac{10\beta^{*3}\pi^2}{3} + \left(\beta^*\right)^5\right] \\ & + p\left(1-\lambda\right)\left[\frac{7\beta\pi^4}{3\alpha^4} + \frac{10\beta^3\pi^2}{3\alpha^2} + \beta^5\right]. \end{aligned} \tag{4.3}$$

Proof: Take the fifth derivative of the MGF and then evaluate the result at $t = 0$

$$\begin{aligned} M_X^{(5)}\left(t\right) = \frac{\partial^5}{\partial t^5}\Big[& p\lambda \exp\left(tx^*\right) + \left(1-p\right)\exp\left(t\beta^*\right)\Gamma\left(1-t\right)\Gamma\left(1+t\right) \\ & + p\left(1-\lambda\right)\exp\left(t\beta\right)\Gamma\left(1-\frac{t}{\alpha}\right)\Gamma\left(1+\frac{t}{\alpha}\right)\Big] \end{aligned}$$

$$= p\lambda \frac{\partial^5}{\partial t^5} \exp\left(tx^*\right) + (1-p)\frac{\partial^5}{\partial t^5}\left[\exp\left(t\beta^*\right)\Gamma(1-t)\Gamma(1+t)\right]$$
$$+ p(1-\lambda)\frac{\partial^5}{\partial t^5}\left[\exp(t\beta)\Gamma\left(1-\frac{t}{\alpha}\right)\Gamma\left(1+\frac{t}{\alpha}\right)\right].$$

$$M_X^{(5)}(t)\big|_{t=0} = p\lambda x^{*5} + (1-p)\left[\left(\beta^*\right)^5 + \left(\beta^*\right)^3 \frac{10\pi^2}{3} + \left(\beta^*\right)\frac{7\pi^4}{3}\right]$$
$$+ p(1-\lambda)\left[\beta^5 + \beta^3 \frac{10\pi^2}{3\alpha^2} + \frac{7\beta\pi^4}{3\alpha^4}\right].$$

This is the fifth moment of the mixture model. ■

After obtaining all five moments of the mixture model, we now derive the method of moment estimators by equating the sample and population moments as given below:

$$E(X) = m_1:$$

$$p\lambda x^* + (1-p)\beta^* + p(1-\lambda)\beta = m_1,$$

$$E\left(X^2\right) = m_2:$$

$$p\lambda\left(x^*\right)^2 + (1-p)\left[\frac{\pi^2}{3} + \left(\beta^*\right)^2\right] + p(1-\lambda)\left[\frac{\pi^2}{3a^2} + \beta^2\right] = m_2,$$

$$E\left(X^3\right) = m_3: \tag{4.4}$$

$$p\lambda\left(x^*\right)^3 + (1-p)\left[\pi^2\beta^* + \left(\beta^*\right)^3\right] + p(1-\lambda)\left[\frac{\pi^2\beta}{\alpha^2} + \beta^3\right] = m_3,$$

$$E\left(X^4\right) = m_4:$$

$$p\lambda\left(x^*\right)^4 + (1-p)\left[\frac{7\pi^4}{15} + 2\pi^2\beta^{*2} + \left(\beta^*\right)^4\right] + p(1-\lambda)$$
$$\left[\frac{7\pi^4}{15\alpha^4} + \frac{2\pi^2\beta^2}{\alpha^2} + \beta^4\right] = m_4, \text{ and}$$

$$E\left(X^5\right) = m_5 :$$

$$p\lambda\left(x^*\right)^5 + (1-p)\left[\frac{7\beta^*\pi^4}{3} + \frac{10\beta^{*3}\pi^2}{3} + \left(\beta^*\right)^5\right] + p(1-\lambda)$$
$$\left[\frac{7\beta\pi^4}{3\alpha^4} + \frac{10\beta^3\pi^2}{3\alpha^2} + \beta^5\right] = m_5.$$

Then solve all the equalities in Equation (4.4) simultaneously for the five unknown parameters $\phi = \left(\beta^*, \beta, \alpha, p, \lambda\right)$. However, it is very difficult to have an analytical solution form for the two-stage mixture logistic distribution. This not surprising because Klugman et al [45] explained that there is no guarantee the moment equations will produce a solution or, if there is a solution, it will be unique. Therefore, there is no closed form solution for the method of moment estimators of the parameters of the two-stage mixture logistic distribution.

However, there are some solutions to estimate the method of moment for some interesting conditions. For the condition $p = 0$ where the mixture model become the 1PL distribution with parameter β^*.

Theorem 4.1 *Random variable X for two-stage mixture distribution with the special condition* $p = 0$ *has the method of moment estimator*

$$\hat{\beta}^* = m_1. \tag{4.5}$$

Proof: For the condition $p = 0$, the two-stage mixture logistic distribution becomes the simple logistic distribution with 1 parameter. The MGF of the logistic distribution with 1 parameter was described in Section 3.2 and the expected value discussed in Section 3.3.

Then, by solving $E(X)=m_1$ and put $E(X)=\beta^*$, we obtained the result of method of moment estimator. ■

For the condition $p=1$ where the two-stage mixture logistic model become the 3PL distribution with parameters β, λ, α. The estimator's solution for β, λ, and α are given in the Theorem 4.2.

Theorem 4.2 *Random variable X for two-stage mixture distribution with the condition* $p=1$ *has the method of moment estimators*

$$\hat{\beta}=\frac{3m_2\pm\sqrt{9m_2^2-8m_1m_3}}{4m_1},$$

$$\hat{\lambda}=\frac{-3m_1m_2+2m_3\pm m_1\sqrt{9m_2^2-8m_1m_3}}{2m_3}, \text{ where } 0<\lambda<1, \text{ and} \tag{4.6}$$

$$\hat{\alpha}=\frac{\sqrt{-3m_1m_2^2\pi^2+4m_1^2m_3\pi^2-m_1m_2\pi^2\sqrt{9m_2^2-8m_1m_3}}}{\sqrt{6\left(-m_2^2m_3+m_1m_3^2\right)}},\ \alpha>0.$$

Proof: For the condition $p=1$, the two-stage mixture logistic distribution becomes the simple logistic distribution with 3 parameters. The MGF of the logistic distribution with 3 parameters was described in Section 3.2. Then, we derived the first until the third moment of distribution due to there are 3 parameters. Finally, we got the result as follows.

$$E(X)=\hat{\beta}\left(1-\hat{\lambda}\right)=m_1,$$

$$E\left(X^2\right)=\hat{\beta}^2\left(1-\hat{\lambda}\right)+\frac{\left(1-\hat{\lambda}\right)\pi^2}{3\alpha^2}=m_2, \text{ and} \tag{4.7}$$

$$E\left(X^3\right)=\hat{\beta}^3\left(1-\hat{\lambda}\right)+\frac{\hat{\beta}\left(1-\hat{\lambda}\right)\pi^2}{3\hat{\alpha}^2}=m_3.$$

Then, solving these equations simultaneously gives the method of moment estimators as defined in Equation (4.6). ■

4.2 Maximum Likelihood Estimation

Keeping [50] discussed that method of maximum likelihood is a very useful method of parameter estimation. The general idea is to choose an estimator of the parameter ϕ that maximizes the likelihood function of the sample observations. In other words, for this value of ϕ the observed sample is also the most likely sample.

Definition 4.2 *If the continuous variable X has a probability density* $f\left(x|\phi,\phi_\alpha\right)$ *which depends on a parameter* ϕ *and possibly on other parameters represented jointly by* ϕ_α*, the likelihood of a set of sample values* $x_1, x_2, ..., x_N$ *is defined by*

$$L\left(x|\phi,\phi_\alpha\right) = f\left(x_1|\phi,\phi_\alpha\right)\cdot f\left(x_2|\phi,\phi_\alpha\right)\cdots f\left(x_N|\phi,\phi_\alpha\right)$$

$$= \prod_{i=1}^{n} f\left(x_i|\phi,\phi_\alpha\right). \tag{4.8}$$

Ultimately, the maximum likelihood estimator is given by solving Equation (4.9) with respect to the parameter ϕ:

$$\frac{\partial}{\partial\phi}\left(\log L\right) = 0\,. \tag{4.9}$$

Having $\frac{\partial^2}{\partial\phi^2}\left(\log L\right) < 0$ ensures that the estimator we obtained maximizes the likelihood function. Let $\phi = \left\{p, \beta, \beta^*, \lambda, \alpha\right\}$, then we apply Definition 4.2 for two-stage mixture logistic model and the likelihood function is defined as follows:

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

$$=\prod_{i=1}^{n}\left[\frac{(1-p)\exp\left[-\left(x_i-\beta^*\right)\right]}{\left[1+\exp\left\{-\left(x_i-\beta^*\right)\right\}\right]^2}+\frac{p(1-\lambda)\exp\left[-\alpha\left(x_i-\beta\right)\right]}{\left[1+\exp\left\{-\alpha\left(x_i-\beta\right)\right\}\right]^2}\right]. \tag{4.10}$$

Then, the log-likelihood function is given as follows:

$$\ell(x;\phi)=\ln\left(L(x;\phi)\right)=\ln\left(\prod_{i=1}^{n} f(x_i;\phi)\right)=\sum_{i=1}^{n}\ln\left(f(x_i;\phi)\right)$$

$$=\sum_{i=1}^{n}\ln\left(\frac{(1-p)\exp\left[-\left(x_i-\beta^*\right)\right]}{\left[1+\exp\left\{-\left(x_i-\beta^*\right)\right\}\right]^2}+\frac{p(1-\lambda)\exp\left[-\alpha\left(x_i-\beta\right)\right]}{\left[1+\exp\left\{-\alpha\left(x_i-\beta\right)\right\}\right]^2}\right). \tag{4.11}$$

Taking the derivative of the log-likelihood function (4.8) with respect to the parameter ϕ and set equating to zero.

$$\frac{\partial\ell(x;\phi)}{\partial\phi}=\frac{\partial\sum_{i=1}^{n}\ln\left(f(x_i;\phi)\right)}{\partial\phi}$$

$$=\sum_{i=1}^{n}\frac{\partial\ln\left(f(x_i;\phi)\right)}{\partial\phi}$$

$$=\sum_{i=1}^{n}\frac{1}{f(x_i;\phi)}\left[\frac{\partial\left(f(x_i;\phi)\right)}{\partial\phi}\right]=0. \tag{4.12}$$

Differentiate the log likelihood function in (4.12) for each parameter of two-stage mixture logistic distribution.

The derivative of log likelihood functions with respect to $\phi=\left\{p,\beta,\beta^*,\lambda,\alpha\right\}$ is given as follows:

$$\frac{\partial \ell(x;\phi)}{\partial \alpha} = \sum_{i=1}^{n} \frac{1}{f(x_i;\phi)} \left[\frac{p(\lambda-1)(\beta - x_i)\exp\left[\alpha(\beta + x_i)\right]\left[\exp(\alpha\beta) - \exp(\alpha x_i)\right]}{\left[\exp(\alpha\beta) + \exp(\alpha x_i)\right]^3} \right] \tag{4.13}$$

$$\frac{\partial \ell(x;\phi)}{\partial \lambda} = -\sum_{i=1}^{n} \frac{1}{f(x_i;\phi)} \left[\frac{p\exp\left[\alpha(\beta + x_i)\right]}{\left[\exp(\alpha x_i) + \exp(\alpha\beta)\right]^2} \right] \tag{4.14}$$

$$\frac{\partial \ell(x;\phi)}{\partial p} = -\sum_{i=1}^{n} \frac{1}{f(x_i;\phi)} \left[\frac{\exp\left(\beta^* + x_i\right)}{\left[\exp\left(\beta^*\right) + \exp(x_i)\right]^2} + \frac{(\lambda-1)\exp\left(\alpha(x_i + \beta)\right)}{\left[\exp(\alpha x_i) + \exp(\alpha\beta)\right]^2} \right] \tag{4.15}$$

$$\frac{\partial \ell(x;\phi)}{\partial \beta} = \sum_{i=1}^{n} \frac{1}{f(x_i;\phi)} \left[\frac{\alpha p(\lambda-1)\exp\left[\alpha(\beta + x_i)\right]\left[\exp(\alpha\beta) - \exp(\alpha x_i)\right]}{\left[\exp(\alpha\beta) + \exp(\alpha x_i)\right]^3} \right] \tag{4.16}$$

$$\frac{\partial \ell(x;\phi)}{\partial \beta^*} = -\sum_{i=1}^{n} \frac{1}{f(x_i;\phi)} \frac{(p-1)\exp\left(\beta^* + x_i\right)\left(\exp(x_i) - \exp\left(\beta^*\right)\right)}{\left(\exp(x_i) + \exp\left(\beta^*\right)\right)^3} \tag{4.17}$$

Now, we obtain random samples of size *n* from the two-stage mixture logistic population. Then, we use this random sample to calculate the MLE by solving the Equations (4.13) to (4.17) simultaneously.

For the two-stage mixture logistic model, we could not find an analytical solution for the MLE. Therefore, we require a numerical procedure to find the MLE. In this thesis, we choose the Expectation-Maximization (EM) algorithm as a numerical procedure to find the MLE solution. The advantage of the EM algorithm is that is easy to implement since it is not required to calculate the second derivatives. The estimators are also reliable up to the chosen convergence limit.

4.2.1 The EM Algorithm

The EM algorithm was described and introduced by Dempster et al. [39] as an algorithm for obtaining a maximum of the likelihood if an analytical calculation is not feasible. Dempster et al. [39] explained that in the EM algorithm, each iteration of the algorithm consists of an expectation step followed by a maximization step.

Each iteration of the EM algorithm involves two steps which we call the expectation step (E-step) and the maximization step (M-step). In general, the EM algorithm contains the following steps:

1. Defining a missing data and complete data.
2. Calculating the conditional expectation of the complete log-likelihood given the observed data using some initial estimation (E-step).
3. Maximizing the corresponding *Q*-functions to obtain a new estimate (M-step).
4. Iteratively replace the initial estimate in Step 2 with the new estimate in Step 3 and repeat Steps 2 and 3 until a stopping criterion is reached.

4.2.2 The EM Algorithm for Mixture distribution

As mentioned in Section 4.1, the maximum likelihood estimation of mixture distribution has no explicit solution. With an appropriate choice of the unobserved data, the EM algorithm can be used as described by McLachlan and Peel [51]. The EM algorithm comprises two steps, an E-step, and an M-step. The algorithm requires an initial choice of parameters, say $\phi^{(0)}$, and the MLE's are then determined by cycling between E and M steps until some convergence criteria are met.

Suppose that at iteration v we have a parameter set $\phi = \phi^{(v)}$. We have a sample of n observations $x = (x_1, x_2, \cdots, x_n)$ and each x_i is associated with one of two unobserved states due to the 2-stage mixture model. Define an unobserved vector $z = (z_1, z_2, \cdots, z_n)$, where z_i are an indicator vector with components zero and one, the component equal to one indicating the unobserved state associated with x_i. The complete data can thus be defined as $y = (x, z)$ and denoting $h\ x, z, \phi$ as joint density of the random variable *Y*.

Definition 4.3 *The finite mixture distribution of a random variable X with k components has the form*

$$g(x, \phi) = \sum_{j=1}^{k} p_j f(x, \theta_j), \tag{4.18}$$

where $\phi = (p_1, \ldots, p_k, \theta_1, \ldots, \theta_k)$.

The complete log-likelihood for the finite mixture distribution may be written by

$$\begin{aligned}
\log L_c(\phi, y) &= \log \prod_{i=1}^{n} h(y_i, \phi) = \sum_{i=1}^{n} \log h(x_i, z_i, \phi) \\
&= \sum_{i=1}^{n} \log\left(g(x_i, \phi \mid z_i) \cdot \tilde{g}(z_i, \phi)\right) \\
&= \sum_{i=1}^{n} \log\left(\prod_{j=1}^{k} f(x_i, \theta_j)^{z_{ij}} \prod_{j=1}^{k} p_j^{z_{ij}}\right) \\
&= \sum_{j=1}^{k} \sum_{i=1}^{n} z_{ij} \log\left(p_j f(x_i, \theta_j)\right).
\end{aligned} \tag{4.19}$$

Expectation-step

The E-step of the EM algorithm requires the calculation of the conditional expectation of the complete log-likelihood given the observed data *x* using ϕ^{v}, which is

$$
\begin{aligned}
Q\left(\phi,\phi^{(v)}\right) &= E_{\phi^{(v)}}\left[\log L_c\left(\phi,Y\right) \mid X = x\right] \\
&= E_{\phi^{(v)}}\left[\sum_{j=1}^{k}\sum_{i=1}^{n} Z_{ij}\log\left(p_j f\left(x_i,\theta_j\right)\right)\middle| X = x\right] \\
&= \sum_{j=1}^{k}\sum_{i=1}^{n}\left(\log\left(p_j f\left(x_i,\theta_j\right)\right)\right)\cdot E_{\phi^{(v)}}\left[Z_{ij}\,\middle|X_i = x_i\right]. \qquad (4.20)
\end{aligned}
$$

By applying the Bayes' theorem, the conditional expectation of Z_{ij} in Equation (4.20) at the v-th iteration is

$$
\begin{aligned}
E_{\phi^{(v)}}\left[Z_{ij}\,\middle|X_i = x_i\right] &= P\left(Z_{ij} = 1\middle|X_i = x_i\right) \\
&= \frac{P\left(X_i = x_i\middle|Z_{ij} = 1\right)P\left(Z_{ij} = 1\right)}{\sum_{r=1}^{k} P\left(X_i = x_i\middle|Z_{ri} = 1\right)P\left(Z_{ri} = 1\right)} \\
&= \frac{p_j^{(v)} f\left(x_i,\theta_j^{(v)}\right)}{g\left(x_i,\phi_j^{(v)}\right)} = w_{ij}^{(v)}, \qquad (4.21)
\end{aligned}
$$

for $i = 1,\ldots,n$ and $j = 1,\ldots,k$. Hence the E-step in Equation (4.20) can be written by

$$
\begin{aligned}
Q\left(\phi,\phi^{(v)}\right) &= \sum_{j=1}^{k}\sum_{i=1}^{n}\log\left(p_j f\left(x_i,\theta_j\right)\right)w_{ij}^{(v)} \\
&= \sum_{j=1}^{k}\sum_{i=1}^{n} w_{ij}^{(v)}\log p_j + \sum_{j=1}^{k}\sum_{i=1}^{n} w_{ij}^{(v)}\log f\left(x_i,\theta_j\right). \qquad (4.22)
\end{aligned}
$$

Maximization-step

The M-step now requires the maximization of the *Q*-function defined in Equation (4.22) with respect to ϕ. Since the p_j appears only in the first term and θ_j only in the second term, the maximization can be done separately by taking the derivative and evaluating the result equal to zero. Starting with the maximization of the first term of Equation (4.22), it is necessary to solve

$$\frac{\partial}{\partial p_j}\left(\sum_{j=1}^{k}\sum_{i=1}^{n} w_{ij}^{(v)} \log p_j + \tau\left[\sum_{j=1}^{k} p_j - 1\right]\right) = 0 \tag{4.23}$$

and

$$\frac{\partial}{\partial \lambda}\left(\sum_{j=1}^{k}\sum_{i=1}^{n} w_{ij}^{(v)} \log p_j + \tau\left[\sum_{j=1}^{k} p_j - 1\right]\right) = 0, \tag{4.24}$$

where τ denotes a Lagrange multiplier, since the constraint $\sum_{j=1}^{k} p_j = 1$ needs to hold. Equating Equation (4.24) to zero and solving yields

$$\tau = -n \tag{4.25}$$

and iterative estimator for p_j, which is

$$p_j^{(v+1)} = \frac{1}{n}\sum_{i=1}^{n} w_{ij}^{(v)}. \tag{4.26}$$

For the other parameters $\theta^{(v+1)}$, the maximization of Equation (4.22) with respect to θ_j can be defined by solving

$$\frac{\partial Q\left(\theta, \theta^{(v)}\right)}{\theta} = 0. \tag{4.27}$$

Then, replace the initial estimators with the new estimates found in Equation (4.22) and repeat until a stopping criterion is reached. For instance, this process delivers a sequence of values of the observed log-likelihood. To stop the iteration, the absolute difference $\left|\log L\left(\theta^{(v+1)}, x\right) - \log L\left(\theta^{(v)}, x\right)\right|$ is smaller than a small chosen value ε. If this occurs at the $(v+1)$-th iteration, the estimate of θ is $\hat{\theta} = \theta^{(v+1)}$.

4.2.3 The EM Algorithm for Two-stage Mixture Logistic Distribution

For the two-stage mixture logistic distribution, the x_i given z_i are conditionally independent with the conditional distributions

$$f_1\left(x_i \middle| \phi^{(v)}\right) = f\left[x \middle| z_i = (0,1), \phi\right] = \frac{\exp\left[-\left(x-\beta^*\right)\right]}{\left[1+\exp\left\{-\left(x-\beta^*\right)\right\}\right]^2} \tag{4.28}$$

and

$$f_2\left(x_i \middle| \phi^{(v)}\right) = f\left[x \middle| z_i = (1,0), \phi\right] = \frac{\alpha(1-\lambda)\exp\left[-\alpha(x-\beta)\right]}{\left[1+\exp\left\{-\alpha(x-\beta)\right\}\right]^2}. \tag{4.29}$$

We can think of the z_i as being based on the results of identical and independent Bernoulli trials, so that the z_i are drawn from the discrete distribution

$$\begin{aligned} P\left[z_i = (0,1)\right] &= 1-p = q \text{ and} \\ P\left[z_i' = (1,0)\right] &= p. \end{aligned} \tag{4.30}$$

E-step

The E-step of the EM algorithm requires the calculation of the conditional expectation of the complete log-likelihood given the observed data. For a mixture of two distributions, we need to use the Equation 4.31.

$$Q\left(\phi, \phi^{(v)}\right) = \log np\lambda + \sum_{j=1}^{2}\sum_{i=1}^{n} w_{ij}^{(v)} \log p_j + \sum_{j=1}^{2}\sum_{i=1}^{n} w_{ij}^{(v)} \log f\left(x_i, \theta_j\right). \tag{4.31}$$

M-step

The M-step now requires the maximization of the Q-function in Equation (4.31). We start with the maximization of the first term, which is p_j. Recalling from Equation (4.27) that $q = 1 - p$ and noting that $w_{i2} = 1 - w_{i1}$, the maximization step for p is

$$\begin{aligned}\frac{\partial Q}{\partial p} &= \frac{\partial}{\partial p}\log\left(np\lambda\sum_{i=1}^{n}\left[\frac{w_{i1}^{(v)}}{1-p} - \frac{1-w_{i1}^{(v)}}{p}\right]\right) \\ &= \frac{1}{p} + \sum_{i=1}^{n}\left[\frac{w_{i1}^{(v)}}{1-p} - \frac{1-w_{i1}^{(v)}}{p}\right],\end{aligned} \tag{4.32}$$

where

$$w_{i1}^{(v)} = \frac{\dfrac{\left(1-p^{(v)}\right)\exp\left[-\left(x-\beta^{*(v)}\right)\right]}{\left[1+\exp\left\{-\left(x-\beta^{*(v)}\right)\right\}\right]^2}}{\dfrac{\left(1-p^{(v)}\right)\exp\left[-\left(x-\beta^{*(v)}\right)\right]}{\left[1+\exp\left\{-\left(x-\beta^{*(v)}\right)\right\}\right]^2} + \dfrac{p^{(v)}\alpha^{(v)}\left(1-\lambda^{(v)}\right)\exp\left[-\alpha^{(v)}\left(x-\beta^{(v)}\right)\right]}{\left[1+\exp\left\{-\alpha^{(v)}\left(x-\beta^{(v)}\right)\right\}\right]^2}}. \tag{4.33}$$

Then for $\partial Q/\partial p = 0$,

$$\begin{aligned} p\sum_{i=1}^{n} w_{i1}^{(v)} + \frac{1}{p} &= \left(1-p\right)\sum_{i=1}^{n} 1 - w_{i1}^{(v)} \\ p^{(v+1)} &= \frac{1}{1+n}\sum_{i=1}^{n}\left(1-w_{i1}^{(v)}\right) \\ p^{(v+1)} &= \frac{1}{1+n}\sum_{i=1}^{n} w_{i2}^{(v)}. \end{aligned} \tag{4.34}$$

Now, we have the iterative estimator for p. The maximization step of the Q-function with respect to θ_j is solved by taking the derivative of the second term in Equation (4.22) with respect to θ_j where $\theta_j = \left(\beta, \beta^*, \alpha, \lambda\right)$ and setting the derivatives equal to zero.

Starting with a parameter β^*, the maximization includes calculating the derivative of Q with respect to β^* as follows.

$$\begin{aligned}
\frac{\partial Q}{\partial \beta^*} &= \frac{\partial}{\partial \beta^*} \sum_{i=1}^{n} w_{i1}^{(v)} \log f\left(x_i \middle| \beta^*\right) \\
&= \frac{\partial}{\partial \beta^*} \sum_{i=1}^{n} w_{i1}^{(v)} \log \left[\frac{\exp\left(x_i + \beta^*\right)}{\left[\exp\left(x_i\right) + \exp\left(\beta^*\right)\right]^2} \right] \\
&= \sum_{i=1}^{n} w_{i1}^{(v)} \frac{\partial}{\partial \beta^*} \left[\log\left[\exp\left(x_i + \beta^*\right)\right] - 2\log\left[\exp\left(x_i\right) + \exp\left(\beta^*\right)\right]\right] \\
&= \sum_{i=1}^{n} w_{i1}^{(v)} \left[1 - \frac{2\exp\left(\beta^*\right)}{\exp\left(x_i\right) + \exp\left(\beta^*\right)} \right] \\
&= \sum_{i=1}^{n} w_{i1}^{(v)} - 2\exp\left(\beta^*\right) \sum_{i=1}^{n} \frac{w_{i1}^{(v)}}{\exp\left(x_i\right) + \exp\left(\beta^*\right)} \qquad (4.35)
\end{aligned}$$

and setting the resulting derivative in Equation (4.35) equal to zero, we obtain

$$\sum_{i=1}^{n} w_{i1}^{(v)} - 2\exp\left(\beta^*\right) \sum_{i=1}^{n} \frac{w_{i1}^{(v)}}{\exp\left(x_i\right) + \exp\left(\beta^*\right)} = 0$$

$$\exp\left(\beta^*\right) = \frac{\sum_{i=1}^{n} w_{i1}^{(v)}}{2\sum_{i=1}^{n} \frac{w_{i1}^{(v)}}{\exp\left(x_i\right) + \exp\left(\beta^*\right)}}$$

$$\beta^{*(v+1)} = \log\left[\frac{\sum_{i=1}^{n} w_{i1}^{(v)}}{2\sum_{i=1}^{n} \frac{w_{i1}^{(v)}}{\exp\left(x_i\right) + \exp\left(\beta^{*(v)}\right)}} \right]. \qquad (4.36)$$

Now, we have the iterative estimator for the parameter β^* for the first component distribution. An iterative estimator for λ, β, and α in the second component can be obtained

by taking the derivative of the second term of Equation (4.22) with respect to λ, β, and α and setting the result to zero. For the parameter β, we need to solve the following formula

$$\begin{aligned}
\frac{\partial Q}{\partial \beta} &= \frac{\partial}{\partial \beta}\sum_{i=1}^{n} w_{i2}^{(v)} \log f\left(x_i \middle| \beta\right) \\
&= \frac{\partial}{\partial \beta}\sum_{i=1}^{n} w_{i2}^{(v)} \log\left[\frac{\alpha\left(1-\lambda\right)\exp\left(\alpha\left(x_i+\beta\right)\right)}{\left[\exp\left(\alpha x_i\right)+\exp\left(\alpha\beta\right)\right]^2}\right] \\
&= \sum_{i=1}^{n} w_{i2}^{(v)} \frac{\partial}{\partial \beta}\left[\log\left[\alpha\left(1-\lambda\right)\right]+\alpha\left(x_i+\beta\right)-2\log\left[\exp\left(\alpha x_i\right)+\exp\left(\alpha\beta\right)\right]\right] \\
&= \sum_{i=1}^{n} w_{i2}^{(v)}\left[0+\alpha-\frac{2\alpha\exp\left(\alpha\beta\right)}{\exp\left(\alpha x_i\right)+\exp\left(\alpha\beta\right)}\right] \\
&= \sum_{i=1}^{n} w_{i2}^{(v)} - 2\exp\left(\alpha\beta\right)\sum_{i=1}^{n}\frac{w_{i2}^{(v)}}{\exp\left(\alpha x_i\right)+\exp\left(\alpha\beta\right)}
\end{aligned} \tag{4.37}$$

and setting the resulting derivative in Equation (4.37) to zero, we obtain

$$\sum_{i=1}^{n} w_{i2}^{(v)} - 2\exp\left(\alpha\beta\right)\sum_{i=1}^{n}\frac{w_{i2}^{(v)}}{\exp\left(\alpha x_i\right)+\exp\left(\alpha\beta\right)} = 0$$

$$\exp\left(\alpha\beta\right) = \frac{\sum_{i=1}^{n} w_{i2}^{(v)}}{2\sum_{i=1}^{n}\frac{w_{i2}^{(v)}}{\exp\left(\alpha x_i\right)+\exp\left(\alpha\beta\right)}}$$

$$\beta^{(v+1)} = \frac{\log\left[\frac{\sum_{i=1}^{n} w_{i2}^{(v)}}{2\sum_{i=1}^{n}\frac{w_{i2}^{(v)}}{\exp\left(\alpha^{(v)}x_i\right)+\exp\left(\alpha^{(v)}\beta^{(v)}\right)}}\right]}{\alpha^{(v)}}. \tag{4.38}$$

Now, we have the iterative estimator for β given in Equation (4.38). For the parameter α, we need to solve the following formula

$$\frac{\partial Q}{\partial \alpha} = \frac{\partial}{\partial \alpha} \sum_{i=1}^{n} w_{i2}^{(v)} \log f\left(x_i \middle| \alpha\right)$$

$$= \frac{\partial}{\partial \alpha} \sum_{i=1}^{n} w_{i2}^{(v)} \log \left[\frac{\alpha(1-\lambda)\exp\left(\alpha\left(x_i+\beta\right)\right)}{\left[\exp\left(\alpha x_i\right)+\exp\left(\alpha\beta\right)\right]^2} \right]$$

$$= \sum_{i=1}^{n} w_{i2}^{(v)} \frac{\partial}{\partial \alpha} \left[\log\left[\alpha(1-\lambda)\right] + \alpha\left(x_i+\beta\right) - 2\log\left[\exp\left(\alpha x_i\right)+\exp\left(\alpha\beta\right)\right] \right]$$

$$= \sum_{i=1}^{n} w_{i2}^{(v)} \left[\frac{1}{\alpha} + x_i + \beta - \frac{2\left(\beta\exp\left(\alpha\beta\right)+x_i\exp\left(\alpha x_i\right)\right)}{\exp\left(\alpha x_i\right)+\exp\left(\alpha\beta\right)} \right]$$

$$= \left(\frac{1}{\alpha}+\beta\right)\sum_{i=1}^{n} w_{i2}^{(v)} + \sum_{i=1}^{n} w_{i2}^{(v)} \left[x_i - \frac{2\left(\beta\exp\left(\alpha\beta\right)+x_i\exp\left(\alpha x_i\right)\right)}{\exp\left(\alpha x_i\right)+\exp\left(\alpha\beta\right)} \right] \tag{4.39}$$

and setting the resulting derivative in Equation (4.39) equal to zero, which is

$$\left(\frac{1}{\alpha}+\beta\right)\sum_{i=1}^{n} w_{i2}^{(v)} + \sum_{i=1}^{n} w_{i2}^{(v)} \left[x_i - \frac{2\left(\beta\exp\left(\alpha\beta\right)+x_i\exp\left(\alpha x_i\right)\right)}{\exp\left(\alpha x_i\right)+\exp\left(\alpha\beta\right)} \right] = 0$$

$$\left(\frac{1}{\alpha}+\beta\right)\sum_{i=1}^{n} w_{i2}^{(v)} = -\sum_{i=1}^{n} w_{i2}^{(v)} \left[x_i - \frac{2\left(\beta\exp\left(\alpha\beta\right)+x_i\exp\left(\alpha x_i\right)\right)}{\exp\left(\alpha x_i\right)+\exp\left(\alpha\beta\right)} \right]$$

$$\frac{1}{\alpha} = -\frac{\sum_{i=1}^{n} w_{i2}^{(v)} \left[x_i - \frac{2\left(\beta\exp\left(\alpha\beta\right)+x_i\exp\left(\alpha x_i\right)\right)}{\exp\left(\alpha x_i\right)+\exp\left(\alpha\beta\right)} \right]}{\beta\sum_{i=1}^{n} w_{i2}^{(v)}}$$

$$\alpha^{(v+1)} = -\frac{\beta^{(v)}\sum_{i=1}^{n} w_{i2}^{(v)}}{\sum_{i=1}^{n} w_{i2}^{(v)} \left[x_i - \frac{2\left(\beta^{(v)}\exp\left(\alpha^{(v)}\beta^{(v)}\right)+x_i\exp\left(\alpha^{(v)}x_i\right)\right)}{\exp\left(\alpha^{(v)}x_i\right)+\exp\left(\alpha^{(v)}\beta^{(v)}\right)} \right]}. \tag{4.40}$$

Now, we have the iterative estimator for α given in Equation (4.40). Finally, we need to calculate the iterative estimator for the parameter λ by solving the following formula

$$\frac{\partial Q}{\partial \lambda} = \frac{\partial}{\partial \lambda}\left[\log np\lambda + \sum_{i=1}^{n} w_{i2}^{(v)} \log f\left(x_i \middle| \lambda\right)\right]$$

$$= \frac{\partial}{\partial \lambda}\log \lambda + \frac{\partial}{\partial \lambda}\sum_{i=1}^{n} w_{i2}^{(v)} \log\left[\frac{\alpha\left(1-\lambda\right)\exp\left(\alpha\left(x_i+\beta\right)\right)}{\left[\exp\left(\alpha x_i\right)+\exp\left(\alpha\beta\right)\right]^2}\right]$$

$$= \frac{1}{\lambda} + \sum_{i=1}^{n} w_{i2}^{(v)} \frac{\partial}{\partial \lambda}\left[\log\left[\alpha\left(1-\lambda\right)\right] + \alpha\left(x_i+\beta\right) - 2\log\left[\exp\left(\alpha x_i\right)+\exp\left(\alpha\beta\right)\right]\right]$$

$$= \frac{1}{\lambda} + \sum_{i=1}^{n} w_{i2}^{(v)}\left[\frac{-1}{\left(1-\lambda\right)}\right] \tag{4.41}$$

and setting the resulting derivative in Equation (4.41) equal to zero, which is

$$\frac{1}{\lambda} + \sum_{i=1}^{n} w_{i2}^{(v)}\left[\frac{-1}{\left(1-\lambda\right)}\right] = 0$$

$$\frac{1}{\left(1-\lambda\right)}\sum_{i=1}^{n} w_{i2}^{(v)} = \frac{1}{\lambda}$$

$$\lambda^{(v+1)} = \frac{1}{\sum_{i=1}^{n} w_{i2}^{(v)} + 1}. \tag{4.42}$$

Now, we have the iterative estimator for λ given in Equation (4.42). We have completed the EM Algorithm for the two-stage mixture logistic distribution. ■

4.3 Nonlinear Least Square Estimation

The nonlinear regression model is a generalization of the linear regression model in which the conditional mean of the response variable is not a linear function of the parameters. The logistic function for 1PL, 2PL, 3PL, and the mixture are the example of the nonlinear function. A nonlinear regression model has the form as follows:

$$Y_i = f\left(x_i, \theta\right) + \varepsilon_i, \quad i = 1, \ldots, n \tag{4.43}$$

where Y_i are response variables, f is a function of the covariate vector $x_i = \left(x_{i1},...,x_{ik}\right)^T$ and the parameter vector $\theta = \left(\theta_1,...,\theta_p\right)^T$, and ε_i are random errors. The ε_i are assumed to be uncorrelated with mean zero and constant variance.

The unknown parameter vector θ in the nonlinear regression model is estimated from the data by minimizing a suitable goodness of fit expression with respect to θ. The most popular criterion is the sum of squared residuals which is defined as follows.

$$\sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} \left[y_i - f\left(x_i, \theta\right) \right]^2 . \tag{4.44}$$

Estimation based on this criterion is known as nonlinear least squares. If the errors ε_i follow a normal distribution, then the least squares estimator for θ is also the maximum likelihood estimator.

Selecting an appropriate nonlinear regression model can be done by comparing the value of Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Mean Squared Error (MSE) of the model. The MSE is defined as the SSE in Equation (4.44) divided by the degree of freedom. Burnham and Anderson [52] was clearly explained about the understanding of AIC and BIC for selection model. The AIC and BIC are information based criteria that assess model fit based on the log likelihood information. The AIC and BIC are defined as follows:

$$AIC = \frac{-2LogLikelihood + 2k + 2k\left(k+1\right)}{n-k-1} \text{ and} \tag{4.45}$$

$$BIC = -2LogLikelihood + k\ln\left(n\right), \tag{4.46}$$

where k is the number of estimated parameters in the model and n is the number of observations in the dataset. The model having the smallest values of AIC and BIC is preferred as the best model.

CHAPTER 5

APPLICATION

The logistic function has found useful applications in a variety of fields. It was initially discovered and used as a model for the growth of human population by Verhulst [5]. Other authors also applied the logistic function for estimating the growth of human population (see Pearl and Reed [7], Leach [8], and Oliver [2]). Another application of the logistic function is Item Response Theory which has been pioneered by Lord, Novick, and Birnbaum [14] in 1968. Harris [15] considered the comparison of 1PL, 2PL, and 3PL model in the IRT. In the industry, the logistic function also can be used in Statistical Process Control, when the data are coming from logistic distribution.

In this chapter, we pay particular attention to three areas of application of the two-stage mixture logistic function:

1. Human population growth,
2. Item response theory, and
3. Statistical process control.

5.1 Human Population Growth

With few individuals, the growth of a population starts slow. When the number of individuals increases, the growth of the population is faster. However, the more individuals there are in the environment, the limited the resources. A population will continue to grow until it reaches the maximum number of individuals that environment can support.

Ultimately, the population saturates to a point where the growth becomes slower again. Thus, the population growth can usually be model by the logistic distribution.

An illustrative example of the usefulness of the logistic function to model growth is the fitting to Australia census population from 1788 to 2010. This data is available from Australia Bureau of Statistics [53]. Before 1788, the population of Australia was only from the Aborigines with estimates ranging from 300,000 to one million.

Historical population in Australia was divided into 3 terms, first was pre-1788, where the population of Australia was from the indigenous population, which is the Aborigines. Aboriginal Australians are legally defined as people who are members of the Aboriginal race of Australia (Indigenous to the Australian continent). Australian Aborigines migrated from somewhere in Asia at least 30,000 to 40,000 years ago. Recently, the Australian government counted approximately 606,164 aboriginal people or about 2.7% of its total population.

The second term was the settlement federation, starting from 1788 to 1900. The second term is the British colonization of Australia. It was started from the colonization from British Ships at Sydney in 1788.

The third term, from 1901 to the present is the post-Federation time. In the early 1900's, it was commonly believed that the indigenous population of Australia was going to become extinct. By 1900, the recorded Indigenous population of Australia had declined to approximately 93,000. The Indigenous population continued to decline to 1993, reaching a low of 74,000. By 1995, population numbers had climbed back to pre-colonization levels.

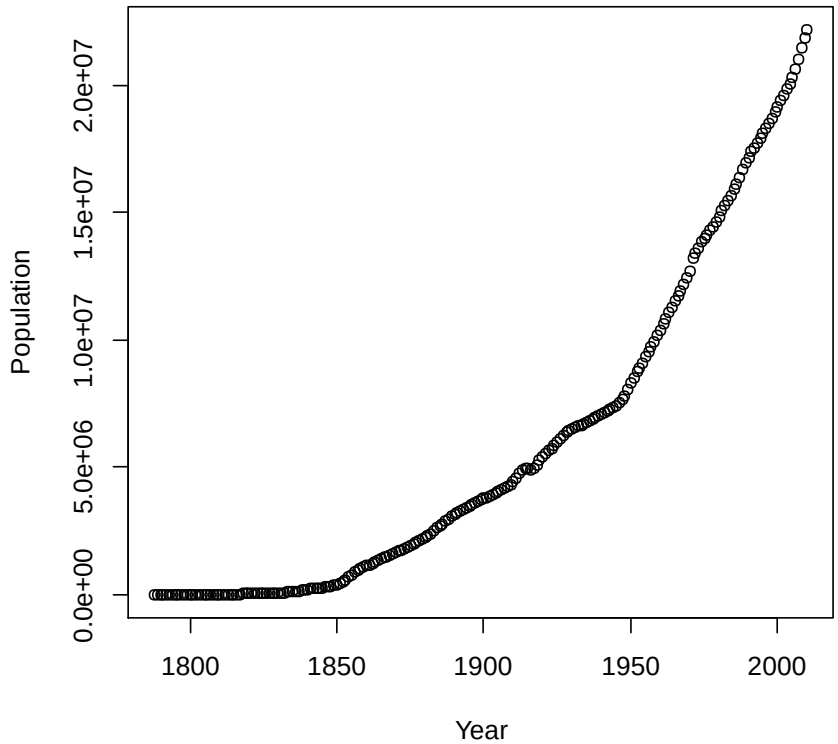


Figure 5.1 Graph of Australian Population Data 1788-2010

If we look at Figure 5.1, it seems like from 1788 to 1900 is the period of adaptation in the new land, where the population is growing. There were only a few of immigrants to Australia for the period 1788 to 1900. Therefore, the growth of population was slow in this period. In the period 1901 to 2010, there were more individuals immigrating. The federation of Australia has settled and the growth of population was fast.

Based on the Australian timeline history, we estimate the population growth of Australia using our logistic function. In particular, we estimate the parameters of the 1PL, 2PL, 3PL, and the mixture model by fitting these models to the Australian population.

5.1.1 The 1PL Nonlinear Regression

For the condition $p = 0$ in Equation (3.3), the two-stage mixture logistic distribution becomes the simple logistic distribution with 1 location parameter β^*. The 1PL

model has the CDF defined in Equation (1.3) and has log odds defined in Equation (3.44). By using nonlinear regression procedure, the nls subroutine of the stats package, in the statistical programming environment *R* version 3.2.5 [54], we have the output in Figure 5.2.

```
Formula: growth.new ~ 1/(1 + exp((b - years)))

Parameters:
  Estimate Std. Error t value Pr(>|t|)
b 1963.114      0.515    3812   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2102 on 222 degrees of freedom
```

Figure 5.2 Output of 1PL Nonlinear Regression

The column marked `Estimate` displays the least square estimates. Hence, we have the least square estimate of the parameter is $\hat{\beta}^* = 1963.114$. The *p*-values shown are based on asymptotic normality. The null hypothesis for testing the parameter of 1PL nonlinear regression is given by:

$$H_0: \beta^* = 0.$$

Since the *p*-value < $\alpha(0.05)$, we reject H_0 and conclude that the parameter β^* is very far from 0.

The column marked `Std. Error` displays the estimated standard error of these estimates. The small standard error for parameter β^* reflects the stability in the estimated β^* because the estimation error is small. The column `t value` shows the ratio of the parameter estimate to its standard error.

5.1.2 The 2PL Nonlinear Regression

For the condition $p = 1$ and $\lambda = 0$ in Equation (3.3), the two-stage mixture logistic distribution becomes the simple logistic distribution with 2 parameters (α and β). The 2PL model has the CDF defined in Equation (1.1) and has a log of odds defined in Equation (3.45). By using nonlinear regression in *R*, we have the output in Figure 5.3.

```
Formula: growth.new ~ 1/(1 + exp(a * (b - years)))

Parameters:
   Estimate Std. Error t value Pr(>|t|)
b 1.958e+03  5.120e-01 3823.38   <2e-16 ***
a 3.332e-02  5.651e-04   58.96   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.03664 on 221 degrees of freedom
```

Figure 5.3 Output of 2PL Nonlinear Regression

Hence, we have the least square estimates of the parameter are $\hat{\alpha} = 0.03332$ and $\hat{\beta} = 1958$. The null hypothesis for testing the parameters of 2PL nonlinear regression is given by:

$$H_0: \beta = \alpha = 0 .$$

Since all *p*-values < $\alpha(0.05)$, we reject H_0 and conclude that the parameters are very likely far from 0.

If the value of α is going to 1, the 2PL model become the 1PL. Since the value of $\hat{\alpha}$ is close to 0, the 2PL model is preferred over the 1PL model (see Figure 5.6).

5.1.3 The 3PL Nonlinear Regression

For the condition $p = 1$ in Equation (3.3), the two-stage mixture logistic distribution becomes the simple logistic distribution with 3 parameters $(\alpha, \beta,$ and $\lambda)$. The 3PL model has the CDF defined in Equation (3.1) and has a log of odds defined in Equation (3.46). By using nonlinear regression in *R*, we have the output in Figure 5.4.

```
Formula: growth.new ~ c + ((1 - c)/(1 + exp(a * (b - years))))

Parameters:
   Estimate Std. Error  t value Pr(>|t|)
c 1.369e-02  4.511e-03    3.034   0.0027 **
b 1.959e+03  6.132e-01 3194.328   <2e-16 ***
a 3.479e-02  7.442e-04   46.752   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0361 on 220 degrees of freedom
```

Figure 5.4 Output of 3PL Nonlinear Regression

Hence, we have the least square estimates of the parameter are $\hat{\alpha} = 0.03479$, $\hat{\beta} = 1959$ and $\hat{\lambda} = 0.01369$. The null hypothesis for testing the parameters of 3PL nonlinear regression is given by:

$$H_0: \beta = \alpha = \lambda = 0.$$

Since all *p*-values < $\alpha(0.05)$, we reject H_0 and conclude that the parameters are very likely different from 0. Although λ is significantly different from zero, the value of $\hat{\lambda}$ is still fairly close to 0. The closer the value of $\hat{\lambda}$ to zero, the more similar the 3PL to the 2PL model.

5.1.4 The Mixture Nonlinear Regression

The two-stage mixture logistic distribution with CDF defined in Equation (3.3) has 5 parameters (p , α , β^*, β, and λ) for us to estimate. By using nonlinear regression in *R*, we have the output in Figure 5.5.

```
Formula: growth.new ~ (1 - p)/(1 + exp(-(years - b1))) + (p * c) + (p *
    (1 - c)/(1 + exp(-a * (years - b2))))

Parameters:
    Estimate Std. Error  t value Pr(>|t|)
p  8.814e-01  5.885e-03  149.771  < 2e-16 ***
c  1.526e-02  3.057e-03    4.992 1.22e-06 ***
b1 1.874e+03  5.268e-01 3556.938  < 2e-16 ***
b2 1.968e+03  4.937e-01 3986.202  < 2e-16 ***
a  4.734e-02  1.011e-03   46.820  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02478 on 218 degrees of freedom
```

Figure 5.5 Output of Mixture Nonlinear Regression

Hence, we have the least square estimates of the parameters are $\hat{p} = 0.8814$, $\hat{\alpha} = 0.04734$, $\hat{\beta}^* = 1874$, $\hat{\beta} = 1968$, and $\hat{\lambda} = 0.01526$. The null hypothesis for testing the parameters of the mixture nonlinear regression is given by:

$$H_0\text{: } \beta = \beta^* = p = \alpha = \lambda = 0 .$$

Since all *p*-values < $\alpha(0.05)$, we reject H_0 and conclude that the parameters are very likely different from 0.

In this mixture model, it is possible to have two points of inflections. From Figure 5.5, the value of $\hat{p}$ is closer to 1, indicating the model has a dominant 3PL component compared to 1PL component. $\hat{\lambda} = 0.01526$ indicates that in this mixture model, the dominant 3PL component behaves fairly similarly to a 2PL component.

After getting all the parameter off the 1PL, 2PL, 3PL and the mixture model, then, we try to fit the model to data. Figure 5.6 describes the comparison of fitting the model to Australia population data from 1788 to 2010.

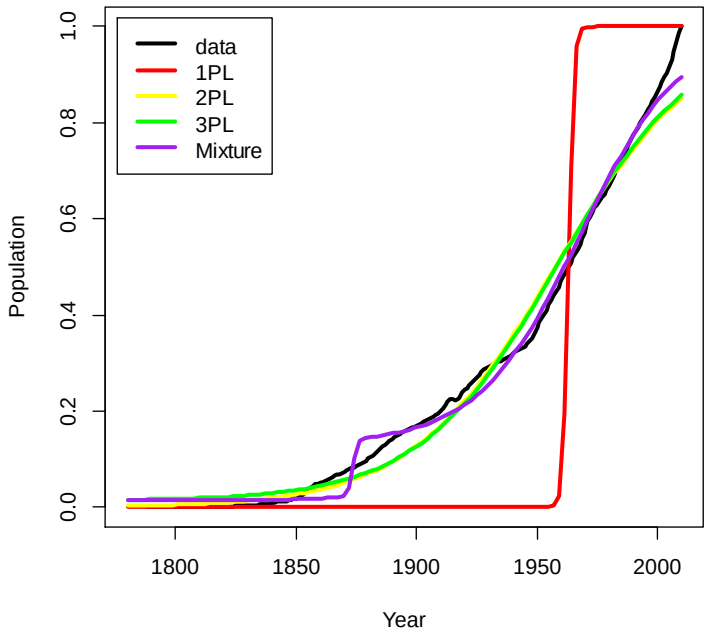


Figure 5.6 Logistic Growth Model of Australian Population

Figure 5.6 illustrates the comparison of the logistic model fit with the Australian population data. The black line shows the observed data of the Australia population in *Y* axis and year in the *X* axis. The 1PL model which is described in the red line looks very different with the data plot and others model. But, 1PL model is good to predict the current total of population. Looking at Figure 5.6 clearly, the 2PL (yellow line), the 3PL (green line), and the mixture model (purple line) are almost like each other. But, in the mixture model, the inflection point occurs in 2 points which are at 1874 and 1968.

We are interested to know what was happening in the mixture model at the inflection points as compared to the other models. For this, we can predict the population data around this inflection points as shown in Table 5.1.

Table 5.1 Comparison of Predicting Population of the Logistic Model

Year	Actual Data	Estimate				Error			
		1PL	2PL	3PL	Mixture	1PL	2PL	3PL	Mixture
1860	1,145,585.0	0.0	823,263.9	982,596.6	414,010.2	-1,145,585.0	-322,321.1	-162,988.4	-731,574.8
1870	1,647,756.0	0.0	1,132,216.3	1,252,928.3	548,681.1	-1,647,756.0	-515,539.7	-394,827.7	-1,099,074.9
1871	1,700,888.0	0.0	1,168,559.2	1,285,034.7	662,423.5	-1,700,888.0	-532,328.8	-415,853.3	-1,038,464.5
1872	1,742,847.0	0.0	1,206,001.8	1,318,174.2	917,754.4	-1,742,847.0	-536,845.2	-424,672.8	-825,092.6
1873	1,794,520.0	0.0	1,244,573.0	1,352,376.3	1,400,689.7	-1,794,520.0	-549,947.0	-442,143.7	-393,830.3
1874	1,849,392.0	0.0	1,284,302.2	1,387,671.4	2,050,479.1	-1,849,392.0	-565,089.8	-461,720.6	201,087.1
1875	1,898,223.0	0.0	1,325,219.4	1,424,090.1	2,611,202.3	-1,898,223.0	-573,003.6	-474,132.9	712,979.3
1876	1,958,679.0	0.0	1,367,354.8	1,461,664.0	2,939,512.0	-1,958,679.0	-591,324.2	-497,015.0	980,833.0
1877	2,031,130.0	0.0	1,410,739.3	1,500,425.0	3,093,689.2	-2,031,130.0	-620,390.7	-530,705.0	1,062,559.2
1880	2,231,531.0	0.0	1,548,701.0	1,624,159.5	3,218,602.6	-2,231,531.0	-682,830.0	-607,371.5	987,071.6
1890	3,151,355.0	0.0	2,103,075.8	2,127,879.5	3,397,922.9	-3,151,355.0	-1,048,279.2	-1,023,475.5	246,567.9
1900	3,765,339.0	0.0	2,828,677.2	2,800,375.1	3,671,671.0	-3,765,339.0	-936,661.8	-964,963.9	-93,668.0
1910	4,425,083.0	0.0	3,757,750.7	3,679,013.2	4,094,862.3	-4,425,083.0	-667,332.3	-746,069.8	-330,220.7
1920	5,411,297.0	0.0	4,914,450.1	4,795,146.5	4,735,242.7	-5,411,297.0	-496,846.9	-616,150.5	-676,054.3
1930	6,500,751.0	0.0	6,305,287.5	6,163,391.9	5,673,594.7	-6,500,751.0	-195,463.5	-337,359.1	-827,156.3
1940	7,077,586.0	0.0	7,909,349.1	7,769,360.6	6,985,698.5	-7,077,586.0	831,763.1	691,774.6	-91,887.5
1950	8,307,481.0	44.7	9,672,719.9	9,560,948.0	8,705,218.4	-8,307,436.3	1,365,238.9	1,253,467.0	397,737.4
1960	10,391,920.0	942,952.8	11,511,922.9	11,449,749.8	10,776,775.5	-9,448,967.2	1,120,002.9	1,057,829.8	384,855.5
1970	12,663,469.0	22,149,830.0	13,327,954.8	13,325,992.4	13,033,712.5	9,486,361.0	664,485.8	662,523.4	370,243.5
1980	14,807,370.0	22,172,470.0	15,026,655.3	15,082,753.1	15,238,531.6	7,365,100.0	219,285.3	275,383.1	431,161.6

The table shows the comparison of predicting the population data around this inflection points by using different logistic models. We also provide the actual data to see which model has the closest prediction to the actual data. From Table 5.1, the mixture model generally provides the smallest prediction error at the inflection points when compared to the other models.

Selecting an appropriate nonlinear regression model can be done by comparing the values of AIC, BIC, and MSE of the model. The definition of AIC, BIC and MSE are given in Section 4.3 earlier. Table 5.2 shows the model selection of nonlinear regression of Australia data.

Table 5.2 Model Selection of Nonlinear Regression of Australia Data

Model	# Parameter	Df	AIC	BIC	MSE
1PL	1	222	-59.710	-52.896	0.044198
2PL	2	221	-837.853	-827.631	0.001343
3PL	3	220	-843.580	-829.951	0.001303
Mixture	5	218	-1009.435	-988.992	0.000613

The smallest value of AIC, BIC, and MSE was the best model of nonlinear regression. Since the two-stage mixture logistic model has the smallest value of AIC, BIC, and MSE, hence, the two-stage mixture logistic model is very good to fit the Australian data instead of 1PL, 2PL, and 3PL.

5.2 Item Response Theory (IRT)

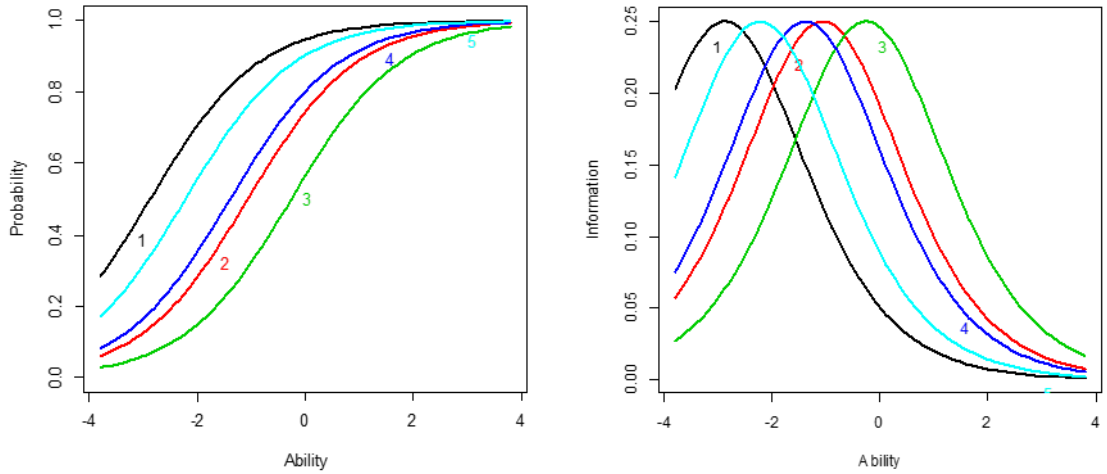
IRT is a psychometric theory with a family of associated mathematical models that relate the latent trait of interest to the conditional probability of responses to items on an assessment. The main purpose of IRT is to create a scale for the interpretation of assessments with useful properties. The logistic function is well used in the application of

IRT. In this section, we use the IRT model described in Equation (3.3) to fit the Law School Admission Test (LSAT) data.

LSAT data set is a classic example in educational testing for measuring ability traits. This test was designed to measure a single latent ability scale. This LSAT example is a part of data set given in Bock and Lieberman [55]. The LSAT data can be imported by using the *latent trait model* (ltm) package in *R* [56]. The LSAT data set represents responses of 1000 subjects from a large sample of students applying for admission to law schools at various university in the U.S. There are 5 questions available for the respondents. The "ltm" package is used to fit 1PL-IRT, 2PL-IRT, and 3PL-IRT model. For the mixture model, we use the "psychomix" package in *R* [57].

5.2.1 IRT in 1PL Model

We estimate the parameter of the 1PL model for the LSAT data by using the "ltm" package in *R*. In the 1PL model, only the difficulty parameter exists.



(a). Item Characteristic Curve (b). Item Information Curve

Figure 5.7 Curves of 1PL Model of LSAT Data

In Figure 5.7, there are 5 curves on the plot, one for each possible item on the test. The curves for items 1 to 5 are shown in different colors and were labeled by appropriate numbers. The person's ability is plotted along the horizontal axis and the vertical axis shows the probability of correct response for each item given the ability level.

Table 5.3 Summary of 1PL Model of LSAT Data

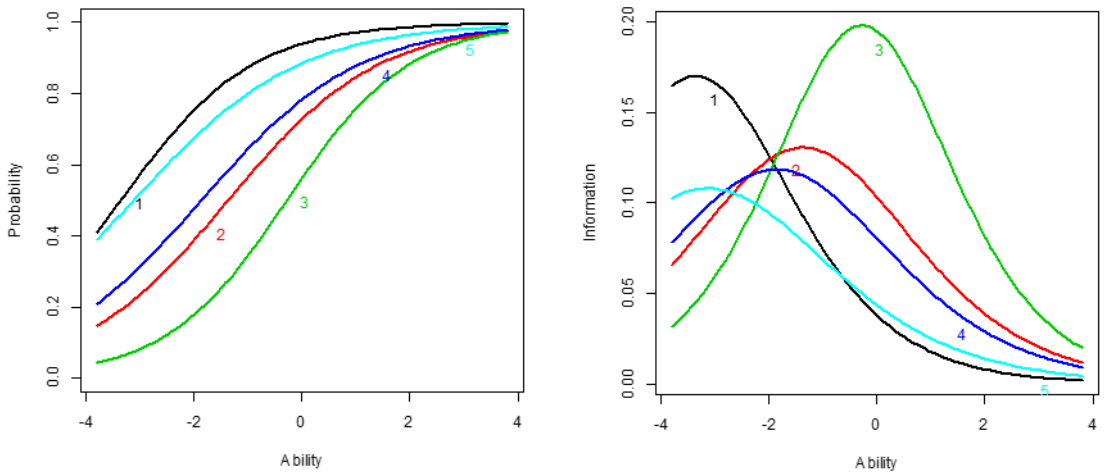
Item	**Difficulty Parameter**		
	Value	**Std. Error**	**Z. value**
1	-2.872	0.129	-22.307
2	-1.063	0.082	-12.946
3	-0.258	0.077	-3.364
4	-1.388	0.087	-16.048
5	-2.219	0.105	-21.166

All the items in the curve have high probability at the highest ability levels. In Table 5.3, item 1 has the highest probability of correct response at the lowest difficulty level of -2.872. We can also see the item difficulty in Figure 5.7 (a), where the highest curve is for item 1, signifying that it is the easiest item. The highest ability levels to answer an item correctly occurs in item 3 which has a difficulty level of -0.258. Item 3 is also described in Figure 5.7 (a) as green color as the lowest curve. In Figure 5.7 (b), the information for item 1 gradually declines after ability level of -2.872 as examinees become more knowledgeable. Generally, for all items, item information decreases after ability reaches the value of the item difficulty.

5.2.2 IRT in 2PL Model

We estimate the parameters of the 2PL model for the LSAT data by using the "ltm" package in *R*. In 2PL model, the discrimination parameter (α) is not fixed at 1 and thus

the model is a bit more flexible than the 1PL model. So, the correct response probability curves for the items are not necessarily parallel with one another.



(a). Item Characteristic Curve (b). Item Information Curve of 2PL

Figure 5.8 Curves of 2PL Model of LSAT Data

In Figure 5.8, there are 5 curves on the plot, one for each possible item on the test. The curves for items 1 to 5 are shown in different colors and were labeled by appropriate numbers. The person's ability is plotted along the horizontal axis and the vertical axis shows the probability of correct response for each item given the ability level.

Table 5.4 Summary of 2PL Model of LSAT Data

Item	**Difficulty Parameter**			**Discrimination Parameter**		
	Value	**Std. error**	**Z. value**	**Value**	**Std. error**	**Z. value**
1	-3.360	0.867	-3.875	0.825	0.258	3.198
2	-1.370	0.307	-4.457	0.723	0.187	3.872
3	-0.280	0.100	-2.808	0.891	0.233	3.828
4	-1.866	0.434	-4.298	0.689	0.185	3.719
5	-3.124	0.870	-3.590	0.658	0.210	3.131

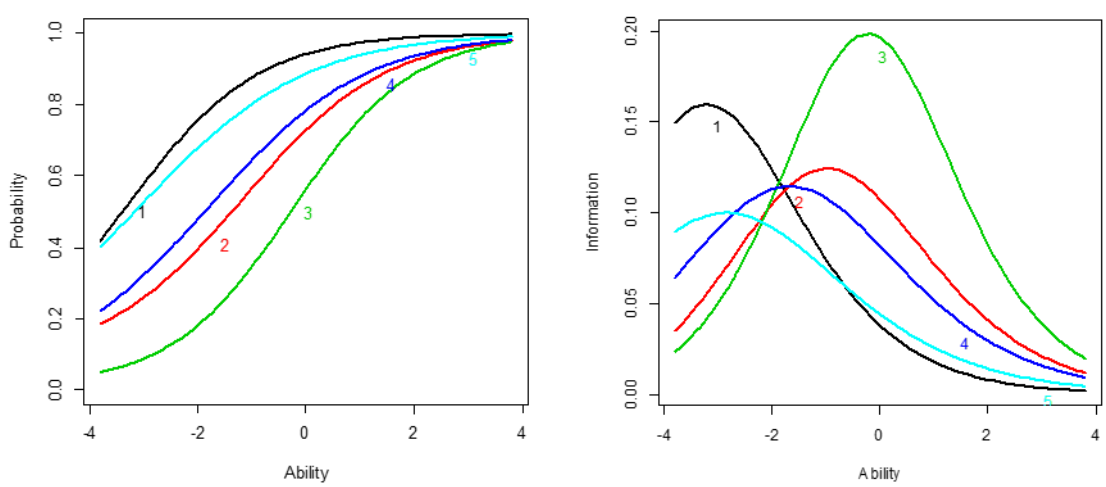
In Table 5.4, item 1 has the highest probability of correct response at the lowest difficulty level of -3.360. We can also see the item difficulty in Figure 5.8 (a), where the highest curve is for item 1, signifying that it is the easiest item. The highest ability levels

to answer an item correctly occurs in item 3 which has a difficulty level of -0.280. Item 3 is also described in Figure 5.8 (a) as green color as the lowest curve. In Figure 5.8 (b), the information for item 1 gradually declines after ability level of -3.360 as examinees become more knowledgeable. For all items in general, item information decreases after ability reaches the value of the item difficulty.

The item discriminations for the LSAT data are also provided in Table 5.4. Most of the items have discrimination (α) more than 0.5. The higher value of α, the more sharply the item discriminates among examinees. Item 3 has the highest discrimination value, which is around 0.891. The lowest discrimination value is in item 5, which is about 0.658.

5.2.3 IRT in 3PL Model

We estimate the parameters of the 3PL model for the LSAT data by using the "ltm" package in *R*. In 3PL model, a discrimination parameter is not fixed at 1 and the item guessing parameter λ as discussed in Section 1.1 exists.



(a). Item Characteristic Curve (b). Item Information Curve

Figure 5.9 Curves of 3PL Model of LSAT Data

In Figure 5.9, there are 5 curves on the plot, one for each possible item on the test. The curves for items 1 to 5 are shown in different colors and were labeled by appropriate numbers. The person's ability is plotted along the horizontal axis and the vertical axis shows the probability of correct response for each item given the ability level.

Table 5.5 Summary of 3PL Model of LSAT Data

Item	**Guessing Parameter**			**Difficulty Parameter**			**Discrimination Parameter**		
	value	**Std. error**	**Z. value**	**value**	**Std. error**	**Z. value**	**value**	**Std. error**	**Z. value**
1	0.037	0.865	0.043	-3.297	1.779	-1.853	0.829	0.288	2.880
2	0.078	2.528	0.031	-1.145	7.517	-0.152	0.760	1.377	0.552
3	0.012	0.282	0.042	-0.249	0.753	-0.331	0.902	0.419	2.152
4	0.035	0.577	0.061	-1.766	1.616	-1.093	0.701	0.257	2.722
5	0.053	1.560	0.034	-2.990	4.061	-0.736	0.666	0.328	2.028

In Table 5.5, item 1 has the highest probability of correct response at the lowest difficulty level of -3.297. We can also see the item difficulty in Figure 5.9 (a), where the highest curve is for item 1, signifying that it is the easiest item. The highest ability levels to answer an item correctly occurs in item 3 which has a difficulty level of -0.249. Item 3 is also described in Figure 5.9 (a) as green color as the lowest curve. In Figure 5.9 (b), the information for item 1 gradually declines after ability level of -3.297 as people become more knowledgeable. In general, information for all items decrease after ability reaches the value of the item difficulty.

The item discriminations for the LSAT data are also provided in Table 5.5. Most of the items have discrimination (α) more than 0.5. The higher value of α, the more sharply the item discriminates among examinees. Item 3 has the highest discrimination value, which is around 0.902. The lowest discrimination value is in item 5, which is about 0.666.

The guessing parameter λ allows examinees, even ones with low ability, to have perhaps substantial probability of correctly answering even moderate or hard items. In Table 5.5, most values of the guessing parameters are quite close to zero. This indicates that examines have low chance to correctly answer these LSAT items by guessing.

5.2.4 IRT in Two-stage Mixture Logistic Model

We estimate the parameters of the 2 components mixture logistic model for the LSAT data by using the "psychomix" package in *R* [57].

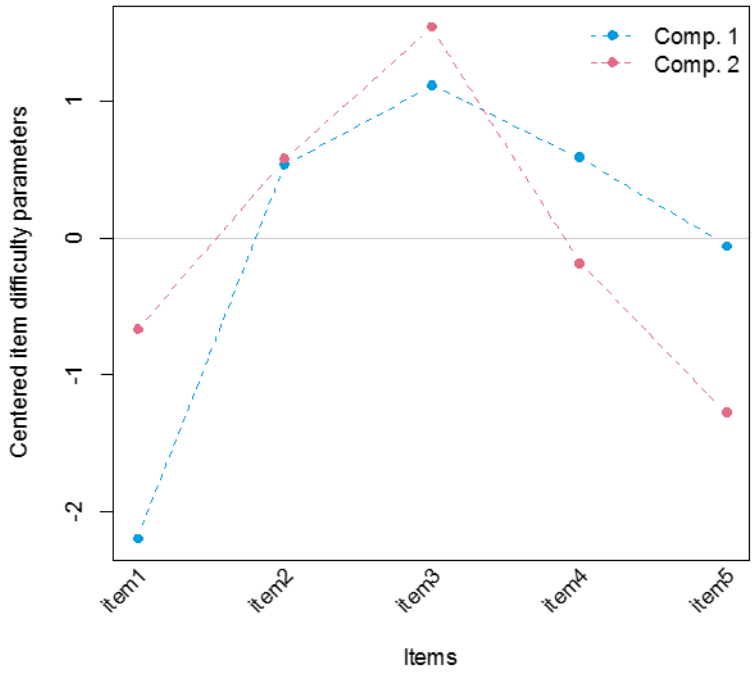


Figure 5.10 Graph of Estimated Item Parameters for Mixture Logistic Model

Figure 5.10 is the graph of the estimated item difficulty parameters given in Table 5.6. The estimated item difficulty parameters for component 1 are described by the blue dot and for component 2 the estimates are described by the red dot. The estimated difficulty parameter for item 1 in the first component is about -2.1997 and the second component estimate is about -0.6681. This means that for item 1, component 2 is more difficult than

component 1. This also applies to items 2 and item 3. But, for items 4 and 5 showed that component 1 is more difficult than component 2.

Table 5.6 Estimated Item Parameters for Two Stage Mixture Logistic distribution

Item	Component 1	Component 2
1	-2.1997	-0.6681
2	0.5387	0.5801
3	1.1171	1.5434
4	0.5971	-0.1773
5	-0.0531	-1.2780

Ultimately, selecting an appropriate IRT model that fits the data can be done based on the evaluation of model fit. Li et al. [58] compared between information criteria and Bayesian methods for mixture IRT model selection and found that BIC, selects the correct model under most simulated conditions. But, in this case, we are fitting the model using information of log likelihood, AIC, and BIC.

Table 5.7 Model Fitting for LSAT Data

Model	Log Likelihood	AIC	BIC
1 Parameter Logistic	-2473.054	4956.108	4980.646
2 Parameter Logistic	-2466.653	4953.307	5002.384
3 Parameter Logistic	-2466.660	4963.319	5036.935
Two-stage Mixture Logistic Model	-2464.505	4963.010	5040.354

According to the log likelihood results in Table 5.7, the mixture model is the best model. Since the mixture model has more parameters, it is conceivable that the log likelihood is highest for this model. But based on parsimony, AIC and BIC typically penalize for unnecessary parameters. Based on the BIC criteria, the 1PL model appears to be the best model. But, the 2PL model becomes the best model if we compare based on the value of AIC. Therefore, both the 1PL and 2PL model is the best and parsimonious model to fit the LSAT data.

5.3 Logistic EWMA Chart

The exponentially weighted moving average (EWMA) control chart is a good alternative to the Shewhart control chart when we are interested in detecting small shifts. The EWMA control chart was first introduced by Roberts [59]. The EWMA control chart is defined as:

$$z_i = \gamma x_i + (1-\gamma) z_{i-1} \tag{5.1}$$

where $0 < \gamma \leq 1$ is a constant and the starting value (required with the first sample at $i = 1$) is the process target, so that $z_0 = \mu_0$. Sometimes the average of preliminary data is used as the starting value of the EWMA, so that $z_0 = \bar{x}$.

For the normal distribution, the zero state (or initial state) control limits for the EWMA control chart are given as follows.

$$\text{UCL} = \mu_0 + L\sigma\sqrt{\frac{\gamma}{2-\gamma}\left[1-(1-\gamma)^{2i}\right]} \text{ and} \tag{5.2}$$

$$\text{LCL} = \mu_0 - L\sigma\sqrt{\frac{\gamma}{2-\gamma}\left[1-(1-\gamma)^{2i}\right]}, \tag{5.3}$$

where the factor L is the width of the control limits. L is usually set at 3 but can be adjusted to set the false alarm rate at some specified value.

According to Montgomery [60], the term $\left[1-(1-\gamma)^{2i}\right]$ in Equations (5.2) and (5.3) approaches unity as i becomes larger. This means that after EWMA control chart has been running for several time periods, the control limits will approach asymptotic steady state values given by

$$\text{UCL} = \mu_0 + L\sigma\sqrt{\frac{\gamma}{2-\gamma}} \text{ and} \tag{5.4}$$

$$\text{LCL} = \mu_0 - L\sigma\sqrt{\frac{\gamma}{2-\gamma}}\,. \tag{5.5}$$

If any point exceeds the control limits, the process is assumed to be out of control. The design parameters of the chart are the multiple of sigma used in the control limits (L) and the value of γ. As shown by Lucas and Saccucci [61], the zero state and steady state run lengths are very similar.

We could not use the standard EWMA to model the control limits for non-normal data. When the data is coming from logistic distribution, hence, we need to develop the EWMA control chart based on this distribution. By using the standard logistic distribution with mean 0 and variance 1, we develop the value of L at fixed $ARL_0 = 370$. The charting constant of Logistic EWMA (LEWMA) for different values of size n and γ are given in Table 5.8.

Table 5.8 Charting Constant of LEWMA at Fixed $ARL_0 = 370$

n	γ			
	0.1	**0.2**	**0.3**	**0.4**
1	2.770	3.020	3.180	3.305
5	2.730	2.900	2.985	3.040
10	2.720	2.880	2.958	3.000

We typically evaluate the decisions regarding sample size and sampling distribution through the Average Run Length (ARL) of the control chart. Essentially, the ARL is the average number of points that must be plotted before a point indicates an out of control condition. The value of ARL for the different shift in the population mean is given in Table 5.9.

Table 5.9 Out of Control ARL for Logistic EWMA

Shift (δ)	$n=1$				$n=5$				$n=10$			
	L=2.77	L= 3.02	L=3.18	L=3.305	L=2.73	L=2.9	L=2.985	L=3.04	L=2.72	L=2.88	L=2.958	L=3
	$\gamma=0.1$	$\gamma=0.2$	$\gamma=0.3$	$\gamma=0.4$	$\gamma=0.1$	$\gamma=0.2$	$\gamma=0.3$	$\gamma=0.4$	$\gamma=0.1$	$\gamma=0.2$	$\gamma=0.3$	$\gamma=0.4$
-1.00	8.01	10.27	14.02	20.47	2.16	2.33	2.47	2.59	1.37	1.44	1.49	1.52
-0.75	13.24	18.74	28.26	42.51	3.32	3.68	3.98	4.39	1.97	2.12	2.22	2.31
-0.50	27.65	44.88	69.76	100.24	6.42	7.38	8.67	10.56	3.61	3.98	4.34	4.76
-0.25	94.32	149.16	197.34	237.40	21.38	29.44	39.98	53.02	11.40	14.15	17.88	22.64
0.00	368.11	373.65	371.60	372.58	372.79	373.97	373.90	370.12	370.65	368.43	371.14	369.30
0.25	95.00	149.13	197.02	235.77	21.36	29.45	39.93	52.78	11.43	14.11	17.84	22.63
0.50	27.72	44.54	69.50	99.62	6.42	7.42	8.67	10.57	3.62	4.00	4.33	4.78
0.75	13.23	18.76	28.36	42.76	3.32	3.67	3.98	4.37	1.97	2.13	2.22	2.31
1.00	7.99	10.22	14.03	20.45	2.16	2.34	2.47	2.59	1.36	1.44	1.49	1.52
1.25	5.49	6.68	8.33	11.25	1.59	1.70	1.77	1.83	1.11	1.14	1.16	1.17
1.50	4.07	4.80	5.68	7.06	1.28	1.35	1.39	1.42	1.02	1.03	1.04	1.04
1.75	3.19	3.69	4.21	4.98	1.12	1.15	1.18	1.19	1.00	1.00	1.01	1.01
2.00	2.60	2.96	3.30	3.75	1.04	1.06	1.07	1.07	1.00	1.00	1.00	1.00
2.25	2.18	2.46	2.70	2.99	1.01	1.02	1.02	1.02	1.00	1.00	1.00	1.00
2.50	1.87	2.10	2.28	2.48	1.00	1.01	1.01	1.01	1.00	1.00	1.00	1.00
2.75	1.64	1.84	1.97	2.11	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
3.00	1.46	1.62	1.74	1.85	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 5.8 and Table 5.9 can be used to model the LEWMA control limits. We apply the LEWMA chart to the wind speed in Chicago, USA at 2009. This data was obtained from Mathematica version 10 [62]. The fitting of the logistic distribution to the histogram of wind speed data in Chicago is given in Figure 5.11.

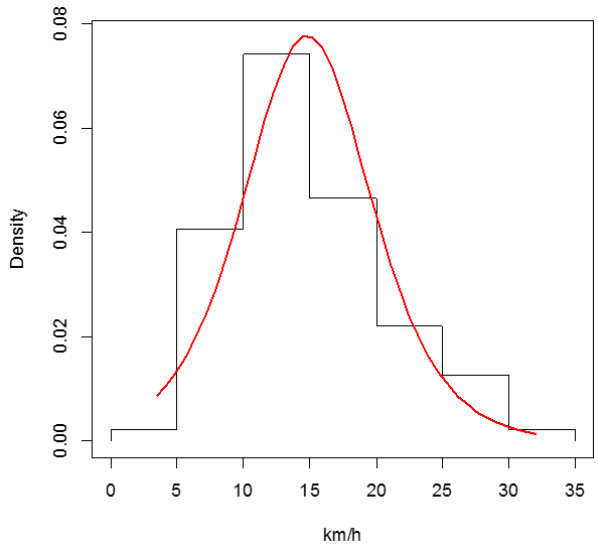
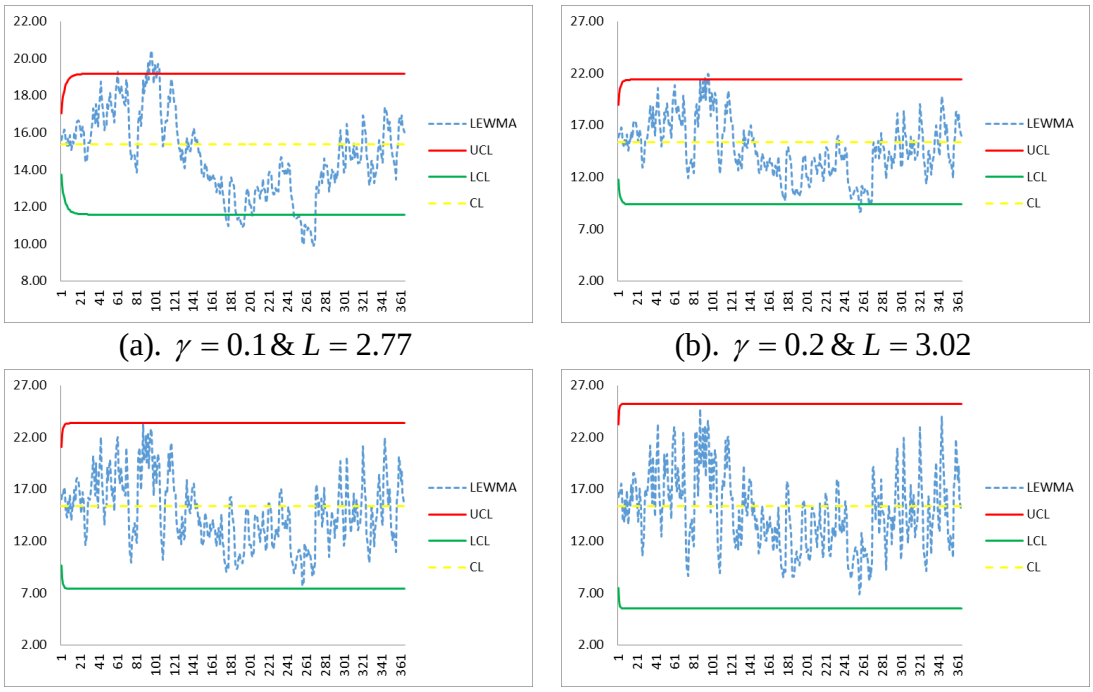


Figure 5.11 The Logistic PDF and Histogram of Wind Chicago

Figure 5.11 shows the logistic distribution can potentially be used to approximate the wind speed data in Chicago at 2009. The null hypothesis for testing the fit of the logistic distribution to the wind speed data in Chicago is given by:

H_0: Data are from logistic distribution.

To test this hypothesis, we use the Kolmogorov-Smirnov value of D for Chicago wind data which is 0.06635 against the critical value of is 0.071186 at $\alpha = 0.05$. Since value of $D <$ 0.071186, we fail to reject H_0 and conclude that Chicago wind data possibly come from the logistic distribution. Then, we develop the control limits of LEWMA for the wind speed of Chicago.



(a). $\gamma = 0.1$ & $L = 2.77$ (b). $\gamma = 0.2$ & $L = 3.02$

(c). $\gamma = 0.3$ & $L = 3.18$ (d). $\gamma = 0.4$ & $L = 3.305$

Figure 5.12 LEWMA Chart for Wind of Chicago 2009

Figure 5.12 illustrates the LEWMA chart for wind speed data in Chicago with different values of $\gamma = 0.1$, $\gamma = 0.2$, $\gamma = 0.3$, $\gamma = 0.4$, respectively. The number of out of control values in Figure 5.12 (a), (b), (c), (d) are 48, 7, 0, and 0, respectively. Figure 5.12 used value of L given in Table 5.8.

5.3.1 Mixture LEWMA Chart

The Mixture LEWMA chart (MLEWMA) is defined as:

$$z_i = \gamma y_i + (1-\gamma)\, z_{i-1}, \tag{5.6}$$

where $y_i = x_{1i}$ for $u_i < p$ and $y_i = x_{2i}$ for $u_i \geq p$, while u_i follows the Uniform distribution (0,1). The zero state (or initial state) control limits for the MLEWMA control chart are given by:

$$\text{UCL} = \mu_y + L\sigma_y\sqrt{\frac{\gamma}{2-\gamma}\left[1-(1-\gamma)^{2i}\right]} \text{ and} \tag{5.7}$$

$$\text{LCL} = \mu_y - L\sigma_y\sqrt{\frac{\gamma}{2-\gamma}\left[1-(1-\gamma)^{2i}\right]}, \tag{5.8}$$

where μ_y and σ_y are coming from mixture logistic distribution. Then, we develop tables of ARL for different value of mixing proportion *p*. Table 5.10, 5.11, 5.12, and 5.13 respectively provides values of ARL with γ values of 0.1, 0.2, 0.3, and 0.4, but for values of *L* of 2.77, 3.02, 3.18, and 3.305 from the simple logistic distribution. The tables also provide the ARL values for different mean shift ranging from -1 to 3.

Table 5.10 ARL for *L*=2.77 and γ = 0.1

Shift (δ)	**Simple Logistic**	**Mixture Logistic**						
		$p = 0.1$	$p = 0.2$	$p = 0.3$	$p = 0.4$	$p = 0.5$	$p = 0.7$	$p = 0.9$
-1.00	8.01	7.98	7.97	7.89	7.88	7.89	7.91	7.92
-0.75	13.24	13.20	13.13	13.16	13.15	13.16	13.14	13.19
-0.50	27.65	27.77	28.04	28.00	28.19	28.34	28.48	28.28
-0.25	94.32	96.93	98.96	102.05	103.20	105.42	107.52	102.46
0.00	368.11	364.36	361.29	353.69	347.31	336.18	316.61	313.49
0.25	95.00	92.74	90.49	89.08	87.71	86.18	85.54	89.19
0.50	27.72	27.66	27.55	27.65	27.57	27.72	28.15	28.78
0.75	13.23	13.27	13.29	13.38	13.39	13.49	13.73	13.70
1.00	7.99	8.05	8.06	8.13	8.18	8.20	8.28	8.30
1.25	5.49	5.52	5.54	5.57	5.60	5.64	5.70	5.66
1.50	4.07	4.09	4.10	4.14	4.16	4.17	4.19	4.17
1.75	3.19	3.21	3.21	3.23	3.25	3.26	3.27	3.25
2.00	2.60	2.61	2.61	2.63	2.64	2.65	2.66	2.63
2.25	2.18	2.19	2.20	2.20	2.21	2.22	2.23	2.21
2.50	1.87	1.88	1.89	1.90	1.91	1.91	1.91	1.89
2.75	1.64	1.65	1.66	1.67	1.68	1.68	1.67	1.65
3.00	1.46	1.47	1.48	1.49	1.49	1.49	1.47	1.46

Table 5.11 ARL for L=3.02 and γ = 0.2

Shift (δ)	Simple Logistic	Mixture Logistic						
		$p = 0.1$	$p = 0.2$	$p = 0.3$	$p = 0.4$	$p = 0.5$	$p = 0.7$	$p = 0.9$
-1.00	10.27	10.23	10.28	10.24	10.21	10.24	10.29	10.32
-0.75	18.74	18.90	19.08	19.22	19.36	19.56	19.82	19.46
-0.50	44.88	46.38	47.68	49.15	50.56	52.19	53.40	49.97
-0.25	149.16	157.70	167.30	175.00	181.80	187.60	186.00	166.04
0.00	373.65	366.90	354.50	338.60	319.70	296.20	262.72	257.28
0.25	149.13	140.00	132.60	124.90	118.20	112.70	105.01	113.53
0.50	44.54	43.68	42.77	41.870	41.24	40.72	40.50	43.37
0.75	18.76	18.65	18.60	18.54	18.54	18.63	19.15	19.68
1.00	10.22	10.28	10.38	10.39	10.47	10.59	10.82	10.77
1.25	6.68	6.69	6.73	6.76	6.85	6.90	7.00	6.93
1.50	4.80	4.82	4.85	4.87	4.92	4.95	5.00	4.94
1.75	3.69	3.70	3.73	3.75	3.76	3.79	3.81	3.77
2.00	2.96	2.98	2.99	3.01	3.02	3.03	3.05	3.02
2.25	2.46	2.48	2.47	2.49	2.51	2.52	2.53	2.50
2.50	2.10	2.11	2.12	2.13	2.14	2.14	2.15	2.13
2.75	1.84	1.84	1.85	1.86	1.86	1.87	1.87	1.84
3.00	1.62	1.63	1.64	1.65	1.65	1.65	1.65	1.63

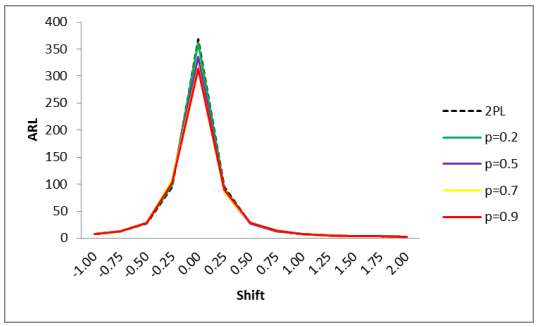
Table 5.12 ARL for L=3.18 and γ = 0.3

Shift (δ)	Simple Logistic	Mixture Logistic						
		$p = 0.1$	$p = 0.2$	$p = 0.3$	$p = 0.4$	$p = 0.5$	$p = 0.7$	$p = 0.9$
-1.00	14.02	14.17	14.36	14.52	14.71	14.85	15.07	14.68
-0.75	28.26	29.26	30.26	31.24	32.10	33.22	33.96	31.69
-0.50	69.76	73.84	78.93	83.38	87.68	91.17	94.06	82.87
-0.25	197.34	209.90	221.80	230.40	236.20	237.30	223.86	197.28
0.00	371.60	360.70	344.20	323.30	297.50	271.10	226.18	220.90
0.25	197.02	182.90	165.50	153.30	140.80	129.50	113.58	122.63
0.50	69.5	65.97	62.49	59.70	57.00	54.59	52.37	58.16
0.75	28.36	27.44	26.96	26.44	25.87	25.75	26.00	28.21
1.00	14.03	13.96	13.90	13.82	13.85	14.00	14.43	14.88
1.25	8.33	8.39	8.41	8.47	8.56	8.64	8.91	8.88
1.50	5.68	5.72	5.75	5.79	5.83	5.90	6.03	5.94
1.75	4.21	4.24	4.27	4.30	4.32	4.36	4.42	4.34
2.00	3.30	3.32	3.35	3.36	3.39	3.39	3.42	3.38
2.25	2.70	2.72	2.74	2.75	2.76	2.77	2.78	2.76
2.50	2.28	2.30	2.30	2.31	2.32	2.33	2.34	2.32
2.75	1.97	1.98	1.99	2.00	2.00	2.01	2.02	2.00
3.00	1.74	1.74	1.75	1.76	1.77	1.77	1.78	1.75

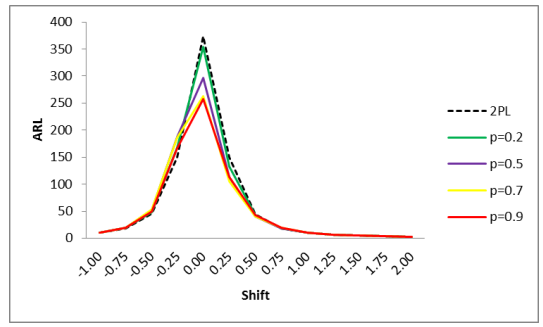
Table 5.13 ARL for L=3.305 and γ = 0.4

Shift (δ)	**Simple Logistic**	**Mixture Logistic**						
		$p = 0.1$	$p = 0.2$	$p = 0.3$	$p = 0.4$	$p = 0.5$	$p = 0.7$	$p = 0.9$
-1.00	20.47	21.28	22.07	22.82	23.61	24.33	24.97	23.14
-0.75	42.51	45.22	48.46	51.28	54.21	56.72	58.51	51.43
-0.50	100.24	107.58	115.12	123.42	129.29	134.85	135.28	117.28
-0.25	237.40	250.34	260.25	263.09	261.08	254.59	227.73	204.69
0.00	372.58	357.25	335.59	309.44	282.57	251.77	203.52	198.52
0.25	235.77	214.54	194.47	175.91	157.62	141.64	118.85	126.37
0.50	99.62	92.03	85.41	79.05	73.76	68.85	62.58	70.50
0.75	42.76	40.72	38.73	36.86	35.49	34.72	33.54	38.09
1.00	20.45	19.88	19.59	19.15	18.80	18.70	19.21	20.96
1.25	11.25	11.18	11.12	11.15	11.18	11.28	11.71	12.06
1.50	7.06	7.11	7.12	7.21	7.28	7.38	7.61	7.65
1.75	4.98	5.02	5.05	5.11	5.15	5.19	5.32	5.22
2.00	3.75	3.78	3.81	3.85	3.89	3.92	3.97	3.89
2.25	2.99	3.01	3.03	3.06	3.09	3.09	3.11	3.07
2.50	2.48	2.49	2.51	2.52	2.53	2.55	2.55	2.53
2.75	2.11	2.13	2.13	2.14	2.15	2.16	2.17	2.15
3.00	1.85	1.86	1.86	1.87	1.88	1.88	1.89	1.87

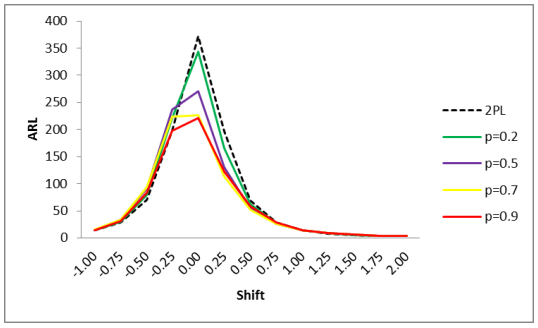
Looking at Tables 5.10, 5.11, 5.12, and 5.13, the value of ARL at $\delta = 0$ decreases when the mixing proportion p increases. This indicates the inappropriateness of using the L value from LEWMA instead of the MLEWMA values for mixture data situation. This can be seen clearly in Figure 5.13.



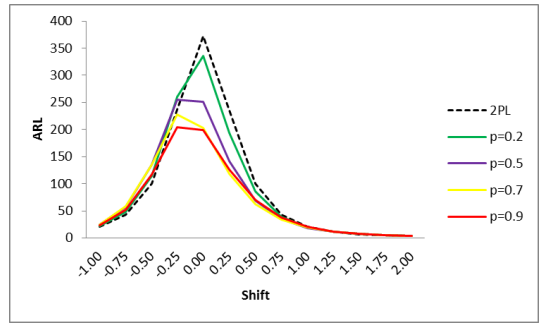
(a). $\gamma = 0.1$ and L=2.77



(b). $\gamma = 0.2$ and L=3.02



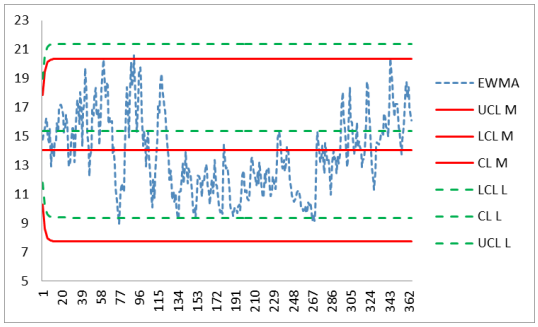
(c). $\gamma = 0.3$ and L=3.18



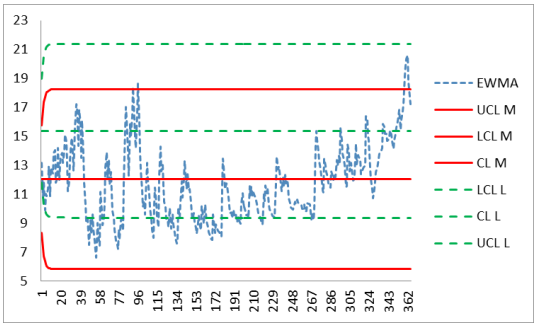
(d). $\gamma = 0.4$ and L=3.035

Figure 5.13 ARL Comparison at Fixed L from LEWMA

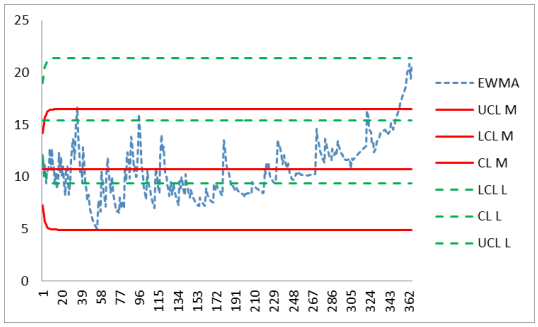
Then for comparison purposes, we create the MLEWMA against LEWMA control chart for $\gamma = 0.2$ and fix L at 3.02 as the value of L in simple logistic.



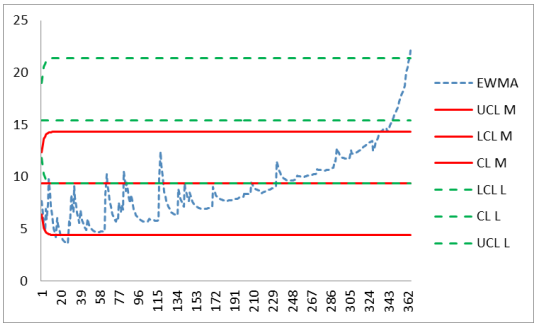
(a). $p = 0.2$



(b). $p = 0.5$



(c). $p = 0.7$



(d). $p = 0.9$

Figure 5.14 MLEWMA Chart with fix L=3.02 and $\gamma = 0.2$

Figure 5.14 illustrates the MLEWMA versus LEWMA chart for mixture wind speed data in Chicago and Florida with $\gamma = 0.2$ and fix L=3.02. The MLEWMA use the mixing proportion 0.2, 05, 07, and 0.9. The number of out of control values for mixture control chart in Figure 5.14 (a), (b), (c), and (d) are 2, 7, 14, and 40, respectively.

Due to misleading use of logistic L in charting the mixture logistic data, we now develop the exact value of L for MLEWMA at different mixing proportion p and smoothing constant γ as given in Table 5.14**.**

Table 5.14 Corresponding value of L for MLEWMA

γ	**Simple Logistic**	**Mixture Logistic**						
		$p = 0.1$	$p = 0.2$	$p = 0.3$	$p = 0.4$	$p = 0.5$	$p = 0.7$	$p = 0.9$
0.1	2.770	2.776	2.780	2.786	2.795	2.815	2.840	2.855
0.2	3.020	3.025	3.035	3.065	3.085	3.120	3.206	3.230
0.3	3.180	3.190	3.215	3.245	3.290	3.345	3.473	3.517
0.4	3.305	3.325	3.350	3.400	3.450	3.520	3.695	3.750

For corresponding value of L, we develop the table of ARL at smoothing constants γ = 0.1, γ = 0.2, γ = 0.3, and γ = 0.4 as given in Tables 5.15, 5.16, 5.17, and 5.18.

Table 5.15 ARL of MLEWMA for γ = 0.1

Shift (δ)	Simple Logistic	Mixture Logistic						
		p = 0.1	p = 0.2	p = 0.3	p = 0.4	p = 0.5	p = 0.7	p = 0.9
	L=2.77	**L=2.78**	**L=2.78**	**L=2.79**	**L=2.79**	**L=2.82**	**L=2.84**	**L=2.85**
-1.00	8.01	7.98	7.97	8.02	8.03	8.14	8.30	8.44
-0.75	13.24	13.22	13.28	13.29	13.40	13.64	13.92	14.24
-0.50	27.65	28.06	28.24	28.53	28.99	29.76	30.78	31.26
-0.25	94.32	97.41	100.46	103.84	108.60	114.80	122.21	119.90
0.00	368.11	370.89	368.20	369.77	367.60	372.47	367.43	372.41
0.25	95.00	93.49	91.91	90.57	91.04	91.35	93.89	102.05
0.50	27.72	27.88	27.86	27.98	28.40	29.03	30.42	31.46
0.75	13.23	13.32	13.41	13.51	13.67	14.00	14.52	14.72
1.00	7.99	8.06	8.15	8.20	8.33	8.50	8.71	8.80
1.25	5.49	5.55	5.59	5.63	5.70	5.80	5.92	5.97
1.50	4.07	4.11	4.14	4.17	4.22	4.28	4.37	4.40
1.75	3.19	3.21	3.22	3.27	3.29	3.33	3.40	3.41
2.00	2.60	2.61	2.62	2.65	2.68	2.71	2.76	2.76
2.25	2.18	2.19	2.21	2.22	2.24	2.27	2.30	2.30
2.50	1.87	1.88	1.90	1.92	1.93	1.95	1.98	1.97
2.75	1.64	1.66	1.67	1.68	1.70	1.71	1.73	1.72
3.00	1.46	1.48	1.49	1.50	1.50	1.52	1.52	1.52

Table 5.16 ARL of MLEWMA for γ = 0.2

Shift (δ)	Simple Logistic	Mixture Logistic						
		p = 0.1	p = 0.2	p = 0.3	p = 0.4	p = 0.5	p = 0.7	p = 0.9
	L=3.02	**L=3.02**	**L=3.03**	**L=3.06**	**L=3.08**	**L=3.12**	**L=3.21**	**L=3.23**
-1.00	10.27	10.25	10.37	10.65	10.81	11.23	12.23	12.43
-0.75	18.74	18.97	19.41	20.16	20.92	22.04	25.01	25.03
-0.50	44.88	46.66	48.77	52.88	56.77	61.72	73.91	71.20
-0.25	149.16	159.58	172.82	193.12	209.41	231.68	276.55	256.74
0.00	373.65	372.94	368.93	373.54	367.29	368.61	368.20	372.45
0.25	149.13	141.77	136.33	135.85	132.44	132.41	139.19	155.10
0.50	44.54	44.34	43.74	44.64	45.00	46.30	51.35	57.16
0.75	18.76	18.78	18.86	19.42	19.82	20.68	23.13	24.51
1.00	10.22	10.37	10.46	10.80	11.04	11.44	12.56	12.88
1.25	6.68	6.73	6.80	7.00	7.13	7.37	7.92	8.03
1.50	4.80	4.84	4.90	5.01	5.10	5.25	5.59	5.61
1.75	3.69	3.70	3.76	3.84	3.89	3.99	4.22	4.24
2.00	2.96	2.99	3.01	3.07	3.12	3.19	3.34	3.36
2.25	2.46	2.49	2.51	2.55	2.58	2.63	2.75	2.75
2.50	2.10	2.11	2.13	2.17	2.20	2.24	2.33	2.34
2.75	1.84	1.84	1.86	1.89	1.91	1.94	2.02	2.02
3.00	1.62	1.63	1.65	1.68	1.70	1.72	1.79	1.79

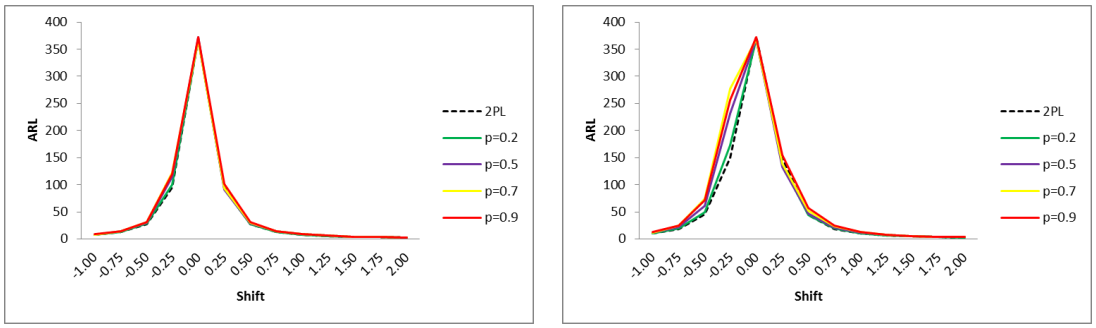
Table 5.17 ARL of MLEWMA for γ = 0.3

Shift (δ)	Simple Logistic	Mixture Logistic						
		$p = 0.1$	$p = 0.2$	$p = 0.3$	$p = 0.4$	$p = 0.5$	$p = 0.7$	$p = 0.9$
	L=3.18	**L=3.19**	**L=3.22**	**L=3.25**	**L=3.29**	**L=3.35**	**L=3.47**	**L=3.52**
-1.00	14.02	14.27	14.90	15.69	16.80	18.31	22.30	22.71
-0.75	28.26	29.51	31.81	34.75	38.83	43.60	56.85	55.61
-0.50	69.76	75.34	83.98	94.51	108.22	126.30	167.13	160.05
-0.25	197.34	214.62	240.05	264.41	296.79	331.37	393.38	376.16
0.00	371.60	371.66	371.57	369.60	367.34	372.59	368.93	372.52
0.25	197.02	186.44	178.67	172.92	170.12	171.18	176.93	197.37
0.50	69.50	67.10	66.26	66.18	67.35	69.90	77.74	91.45
0.75	28.36	27.73	28.20	28.71	29.85	31.75	36.99	42.81
1.00	14.03	14.05	14.37	14.89	15.65	16.54	19.61	21.63
1.25	8.33	8.42	8.67	8.94	9.36	9.95	11.44	12.08
1.50	5.68	5.74	5.89	6.06	6.33	6.65	7.47	7.66
1.75	4.21	4.25	4.35	4.46	4.62	4.82	5.31	5.39
2.00	3.30	3.33	3.41	3.49	3.59	3.73	4.02	4.09
2.25	2.70	2.73	2.78	2.84	2.92	3.01	3.22	3.27
2.50	2.28	2.31	2.34	2.39	2.44	2.52	2.68	2.71
2.75	1.97	1.99	2.02	2.05	2.10	2.16	2.28	2.32
3.00	1.74	1.75	1.78	1.81	1.85	1.90	2.00	2.02

Table 5.18 ARL of MLEWMA for γ = 0.4

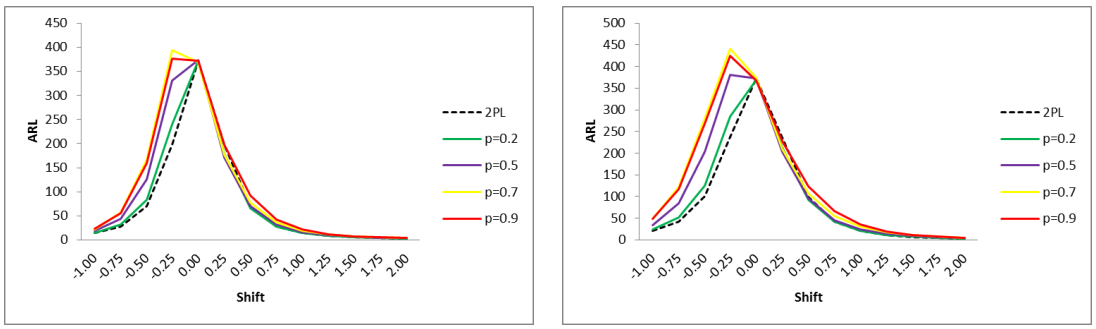
Shift (δ)	Simple Logistic	Mixture Logistic						
		$p = 0.1$	$p = 0.2$	$p = 0.3$	$p = 0.4$	$p = 0.5$	$p = 0.7$	$p = 0.9$
	L=3.305	**L=3.33**	**L=3.35**	**L=3.4**	**L=3.45**	**L=3.52**	**L=3.69**	**L=3.75**
-1.00	20.47	21.90	23.57	26.46	29.76	34.76	49.29	48.05
-0.75	42.51	47.18	52.15	60.61	70.55	84.40	122.22	117.55
-0.50	100.24	112.20	125.55	147.98	171.40	203.06	279.13	269.13
-0.25	237.40	260.26	284.60	316.93	346.85	381.09	441.41	425.27
0.00	372.58	372.90	369.82	373.41	370.49	371.89	374.50	367.19
0.25	235.77	224.86	212.70	209.07	202.56	203.10	210.38	226.98
0.50	99.62	95.89	92.84	92.96	93.04	95.69	107.02	124.20
0.75	42.76	41.88	41.68	42.67	43.83	46.20	55.18	67.11
1.00	20.45	20.53	20.69	21.63	22.67	24.39	30.10	36.24
1.25	11.25	11.47	11.73	12.34	12.99	13.98	17.49	19.92
1.50	7.06	7.25	7.44	7.84	8.24	8.94	10.88	11.76
1.75	4.98	5.08	5.20	5.47	5.73	6.13	7.27	7.53
2.00	3.75	3.82	3.92	4.08	4.26	4.51	5.15	5.28
2.25	2.99	3.04	3.11	3.23	3.35	3.51	3.92	3.97
2.50	2.48	2.53	2.57	2.65	2.74	2.84	3.13	3.19
2.75	2.11	2.15	2.18	2.24	2.30	2.39	2.61	2.65
3.00	1.85	1.87	1.90	1.95	2.00	2.07	2.23	2.27

Looking at Tables 5.15, 5.16, 5.17, and 5.18, the ARL at $\delta = 0$ are around 370 at corresponding value of γ and for different mixing proportion *p*. But, at negative shift, the ARL generally increases when the mixing proportion *p* increases. When the shift is positive, generally the ARL slightly increases as the mixing proportion *p* increases. This can be seen clearly in Figure 5.15.



(a). $\gamma = 0.1$

(b). $\gamma = 0.2$



(c). $\gamma = 0.3$

(d). $\gamma = 0.4$

Figure 5.15 ARL Comparison at Corresponding *L* from LEWMA and MLEWMA

Now, we proceed to develop the control chart for MLEWMA and compare with the LEWMA for real data set. We use the wind speed data in Chicago and Florida at 2009. Previously, we already provide the KS test for wind speed data in Chicago. The KS test for wind speed data in Florida is given by:

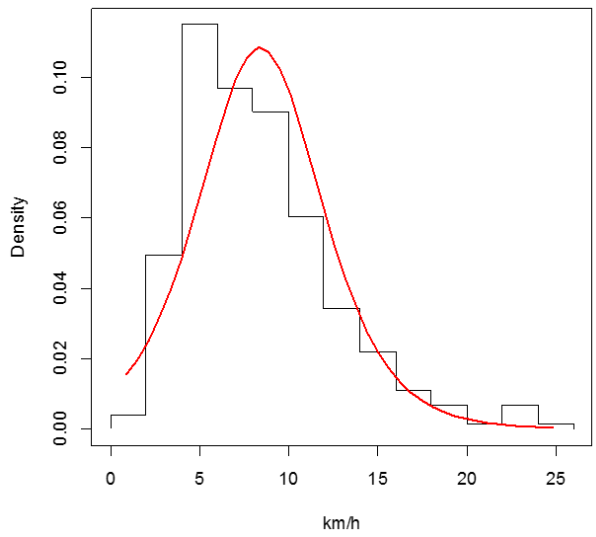
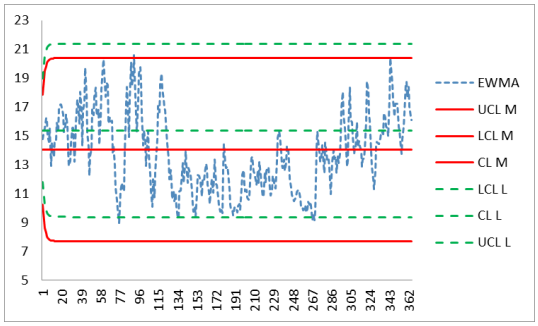


Figure 5.16 The Logistic PDF and Histogram of Wind Florida

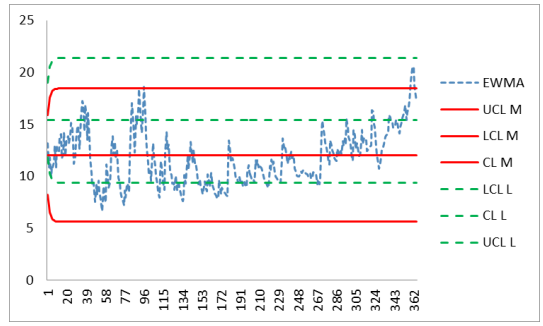
Figure 5.16 shows the logistic distribution can potentially be used to approximate the wind speed data in Florida at 2009. The null hypothesis for testing the fit of the logistic distribution to the wind speed data in Florida is given by:

H_0: Data are from logistic distribution.

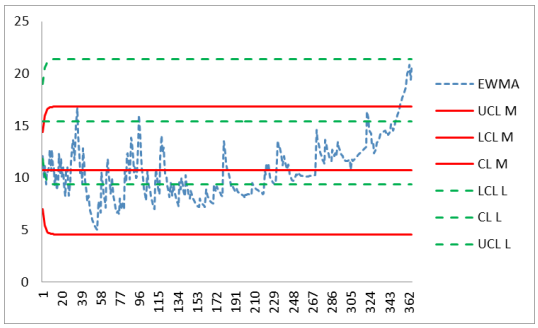
To test this hypothesis, we use the Kolmogorov-Smirnov value of D for Florida wind data which is 0.068101 against the critical value of is 0.071186 at $\alpha = 0.05$. Since value of $D < 0.071186$, we fail to reject H_0 and conclude that Florida wind data possibly come from the logistic distribution. Then, we develop the control limits of MLEWMA for the wind speed of Chicago and Florida.



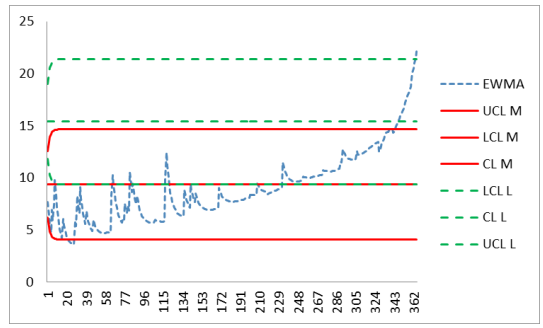
(a). $p = 0.2$ and L=3.035



(b). $p = 0.5$ and L=3.12



(c). $p = 0.7$ and L=3.206



(d). $p = 0.9$ and L=3.23

Figure 5.17 MLEWMA Chart for $\gamma = 0.2$ and Correct L for Mixture Data

Figure 5.17 illustrates the MLEWMA chart for mixture wind speed data in Chicago and Florida with $\gamma = 0.2$ and exact L value at specific mixing proportion. This MLEWMA use the mixing proportion 0.2, 05, 07, and 0.9. The number of out of control values for mixture control chart in Figure 5.17 (a), (b), (c), and (d) are 2, 6, 12, and 29, respectively. In contrast, the number of out of control values for simple logistic control chart in Figure 5.17 (a), (b), (c), and (d) are 5, 74, 132, and 224, respectively. The number of out of control values increases when the mixing proportion p increases. This means that the more the data comes from a mixture population, the more violation is detected by the inappropriate use of the simple logistic.

In comparison to Figure 5.14 for the simple logistic, the number of out of control values in Figure 5.17 decreases when the mixture data use the correct value of L from the

mixture logistic instead of the L from simple logistic distribution. This shows the inappropriateness of using the simple logistic L for mixture logistic data.

CHAPTER 6

CONCLUSION AND RECOMMENDATIONS

6.1 Summary and Conclusion

In this thesis, we have developed a two-stage mixture logistic distribution. Some important properties of the two-stage mixture logistic distribution have been studied in this thesis. Parameter estimation of the model also has been developed. The two-stage mixture logistic distribution found an application in the human population growth, IRT, and SPC.

In Chapter 2, we discussed the literature review of logistic distribution as well as the mixture distribution. We also discussed the previous works on the application of logistic distribution. We found that the two-stage mixture logistic distribution has not been studied yet in the literature.

In Chapter 3, we developed a two-stage mixture distribution. In the first stage, we mixed a point-degenerate distribution with the 2PL distribution using a mixing proportion parameter λ and $(1-\lambda)$ respectively. Then, in the second stage, the resulting distribution was mixed with the 1PL distribution using the mixing proportion p and $(1-p)$. We also discussed the moment generating function and characteristic of the mixture distribution. Then, by using the MGF, we developed the first moment until the fourth moment and the related properties of the moments such as expected value, variance, skewness and kurtosis of distribution. We also studied the survival function and hazard function and provide the survival odds for the model.

In Chapter 4, we developed the parameter estimation of the two-stage mixture logistic distribution. The method of moment estimation for the distribution parameters was studied in this chapter. We provided the solution for the special case of the distribution where $p = 0$ and $p = 1$ .We also studied the maximum likelihood estimation of the distribution and discussed its implementation using the EM algorithm.

In Chapter 5, we discussed some real-life applications of the distribution. For the first application, we used the model to estimate Australian population growth data. We derived the least square estimators of the model parameters by using nonlinear regression. Then we compared the results of the 1PL, 2PL, 3PL, and the mixture nonlinear regression models. In the second application, we applied the two-stage mixture logistic distribution in IRT to model the conditional probability of responses to items on LSAT data. In the final application, we also developed the Mixture Logistic EWMA charts, discussed the associated ARL properties, and compared the behavior of the MLEWMA charts to the LEWMA charts for mixture logistic wind speed data.

6.2 Recommendations

In this section, we recommend the use of the two-stage mixture logistic distribution for future research works as follows:

a. In this thesis, we mixed two logistic distributions with different parameters using the proper mixing proportion only. For the future, this distribution could also be mixed with other types of distributions where we could derive the properties as well as the parameter estimation.

b. We derived the important properties and parameter estimation of the two-stage logistic distribution in this thesis. We could also consider developing other parameter estimation methods such as Bayesian estimation of this distribution which has not yet been studied in this thesis.

c. We discussed in this thesis the application of the two-stage mixture distribution in the field of human growth population, item response theory, and statistical process control. We could also discuss other applications of two-stage mixture logistic distribution such as in survival analysis and in economic field.

REFERENCES

[1] J. S. Cramer, "The origins and development of the logit model," *Logit Model. from Econ. other fields*, no. August, pp. 149–158, 2003.

[2] F. R. Oliver, "Notes on the Logistic Curve for Human Populations," *J. R. Stat. Soc. Ser. A*, vol. 145, no. 3, pp. 359–363, 1982.

[3] A. Ainsworth, "Psychology 427: Intro to Psych Testing." [Online]. Available: http://www.csun.edu/~ata20315/psy427/.

[4] N. Balakrishnan, V. Leiva, A. Sanhueza, and E. Cabrera, "Mixture inverse Gaussian distributions and its transformations, moments and applications," *Statistics (Ber).*, vol. 43, no. 1, pp. 91–104, 2009.

[5] P. Ghent, X. Vol, D. M. Stasiw, and L. C. Ccrny, "P . F . Verhulst ' s ' Notice sur la loi que la populations suit dans son accroissement ' from Correspondence Mathematique et Phisique," vol. X, pp. 183–192, 1976.

[6] P. Verhulst, "Recherches mathématiques sur la loi d ' accroissement de la population," *Bruxelles*, vol. 18, pp. 1–38, 1845.

[7] R. Pearl and L. J. Reed, "On the Rate of Growth of the Population of the United States since 1790 and Its Mathematical Representation," *Proc. Natl. Acad. Sci.*, vol. 6, no. 6, pp. 275–288, 1920.

[8] D. Leach, "Re-Evaluation of the Logistic Curve for Human Populations," *J. R. Stat. Soc. Ser. A*, vol. 144, no. 1, pp. 94–103, 1981.

[9] D. G. V and P. H. D, "Analysis of Factorial Arrangements when the Data are Proportions," *Biometrics*, vol. 8, no. 1, pp. 1–12, 1952.

[10] J. E. Grizzle, "A New Method of Testing Hypotheses and Estimating Parameters for the Logistic Model," *Biometrics*, vol. 17, no. 3, pp. 372–385, 1961.

[11] R. L. Plackett, "The Analysis of Life Test Data," *Technometrics*, vol. 1, no. 1, pp. 9–19, 1959.

[12] F. R. Oliver, "Another Generalisation of the Logistic Growth Function," *Econometrica*, vol. 37, no. 1, pp. 144–147, 1969.

[13] P. R. Fisk, "The Graduation of Income Distributions," *Econometrica*, vol. 29, no. 2, pp. 171–185, 1961.

[14] F. M. Lord, M. R. Novick, and A. Birnbaum, *tatistical theories of Mental Test Scores*. MA: Addison-Wesley, 1968.

[15] D. Harris, "Comparison of 1-,2- and 3-parameter IRT Models," *Educ. Meas. Issues Pract.*, vol. 8, pp. 35–41, 1989.

[16] N. Balakrishnan, *Handbook of the Logistic Distribution*. New York: Marcel Dekker, Inc, 1991.

[17] S. S. Gupta, "Estimation of the parameters of the logistic distribution," *Biometrika*, vol. 53, no. 3 and 4, pp. 565–570, 1966.

[18] G. Blom, *Statistical Estimates and Transformed Beta Variables*. New York: John Wiley and Sons, 1958.

[19] J. Jung, "On linear estimates defined by a continuous weight function," *Ark. Mat.*, vol. 3, no. 1, pp. 199–209, 1955.

[20] S. Gupta, A. Qureishi, and B. Shah, "Best Linear Unbiased Estimators of the Parameters of the Logistic Distribution Using Order Statistics," *Technometrics*, vol. 9, no. 1, pp. 43–56, 1967.

[21] M. A. Stephens, "Tests of fit for the logistic distribution based on the empirical distribution function," *Biometrika*, vol. 66, no. 3, pp. 591–595, 1979.

[22] B. S. Everitt and D. J. Hand, *Finite Mixture Distributions*. London: Chapman and Hall, 1981.

[23] P. Medfyessy, *Decomposition of superpositions of distribution functions*. Budapest: [Wiley, International Biometric Society], 1961.

[24] G. J, "An algorithm for the Decomposition of a Distribution into Gaussian Components," *Biometrics*, vol. 25, no. 1, pp. 79–93, 1969.

[25] A. C. Cohen, "Estimation in Mixtures of Two Normal Distributions," *Technometrics*, vol. 9, no. 1, pp. 15–28, 1967.

[26] E. K. Al-Hussaini and N. S. Abd-El-Hakim, "Failure rate of the inverse Gaussian-Weibull mixture model," *Ann. Inst. Stat. Math.*, vol. 41, no. 3, pp. 617–622, 1989.

[27] K. S. Sultan, M. A. Ismail, and A. S. Al-Moisheer, "Mixture of Two Inverse Weibull distributions: Properties and estimation," *Comput. Stat. Data Anal.*, vol. 51, no. 11, pp. 5377–5387, 2007.

[28] R. A. Chechile, "Mathematical Tools for Hazard Function Analysis," *J. Math. Psychol.*, vol. 47, no. 5–6, pp. 478–494, 2003.

[29] J. M. G. Taylor, "Semi-Parametric Estimation in Failure Time Mixture Models," *Biometrics*, vol. 51, no. 3, pp. 899–907, 1995.

[30] M. G. Larson and G. E. Dinse, "A Mixture Model for the Regression Analysis of Competing Risks Data," *J. R. Stat. Soc. Ser. C (Applied Stat.*, vol. 34, no. 3, pp. 201–211, 1985.

[31] V. T. Farewell, "The Use of Mixhwe Models for the Analysis of Sureval Data with Long-Term Survivors," *Int. Biometric Soc.*, vol. 38, no. 4, pp. 1041–1046, 1982.

[32] J. Behboodian, “Information Matrix for a Mixture of Two Exponential Distribution,” *J. Stat. Comput. Simul.*, vol. 2, no. 1973, pp. 1–16, 1973.

[33] J. Behboodian, “Information Matrix for a Mixture of Two Normal Distributions,” *J. Stat. Comput. Simul.*, vol. 1, no. 1971, pp. 295–314, 1971.

[34] J. Rost, “Rasch Models in Latent Classes: An Integration of Two Approaches to Item Analysis,” *Appl. Psychol. Meas.*, vol. 14, no. 3, pp. 271–282, 1990.

[35] J. T. Willse, “Mixture Rasch Models With Joint Maximum Likelihood Estimation,” *Educ. Psychol. Meas.*, vol. 71, pp. 5–19, 2011.

[36] H. Frick, C. Strobl, F. Leisch, and A. Zeileis, “Flexible Rasch Mixture Models with Package psychomix,” *J. Stat. Softw.*, vol. 48, no. 1992, 2012.

[37] S. H. Leong, R. Mahdi, and S. E. Ling, “Mixture Rasch Model for Guessing Group Identification,” *AIP Conf. Proc.*, vol. 1522, pp. 1248–1253, 2013.

[38] S. P. Reise and D. A. Revicki, *Handbook of Item Response Theory Modeling: Applications to typical performance assessment.* New York: Routledge Taylor & Francis Group, 2015.

[39] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum Likelihood from Incomplete Data Via the EM Algorithm,” *J. R. Stat. Soc. Ser. B Methodol.*, vol. 39, no. 1, pp. 1–38, 1977.

[40] K. M. Leytham, “Maximum Likelihood Estimates for the Parameters of Mixture Distributions,” *Water Resour. Res.*, vol. 20, no. 7, pp. 896–902, 1984.

[41] M. Agha and M. T. Ibrahim, “Maximum Likelihood Estimation of Mixtures of Distributions,” *J. R. Stat. Soc. Ser. C (Applied Stat.*, vol. 33, no. 3, pp. 327–332, 1984.

[42] J. Wengrzik, “Parameter Estimation for Mixture Models Given Grouped Data,” Universitat Bremen, 2012.

[43] M. R. Alkasasbeh and M. Z. Raqab, “Estimation of the Generalized Logistic Distribution Parameters : Comparative study,” *Stat. Methodol.*, vol. 6, no. 3, pp. 262–279, 2009.

[44] S. Ross, *A First Course in Probability*, Eight Edit. New Jersey: Pearson Education, Inc, 2010.

[45] S. A. Klugman, H. H. Panjer, and G. E. Willmot, *Loss models: from data to decisions*. New Jersey: John Wiley & Sons, Inc, 2012.

[46] S. Ross, *Introduction to Probability Models*, 10th Ed. California: Academic Press, 2010.

[47] N. L. Johnson, S. Kotz, and N. Balakhrisnan, *Continuous Univariate Distributions*, vol. 2, no. 1. New York: John Wiley & Sons, Inc, 1995.

[48] M. Hazewinkel, *Encyclopaedia of Mathematics*, vol. III. Netherlands: Kluwer Academic Publisher, 2001.

[49] D. G. Kleinbaum and M. Klein, *Survival Analysis: A Self-Learning Text*, Third Edit. New York: Springer, 2012.

[50] E. . Keeping, *Introduction to Statistical Inference*. New York: Dover Publications, Inc, 2010.

[51] G. McLachlan and D. Peel, *Finite Mixture Models*. New York: Wiley Series in Probability and Statistics, 2000.

[52] K. P. Burnham and R. P. Anderson, "Multimodel Inference: Understanding AIC and BIC in Model Selection," *Sociol. Methods Res.*, vol. 33, no. 2, pp. 261–304, 2004.

[53] "Australian Bureau Statistics." [Online]. Available: http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3105.0.65.0012014?OpenDocument.

[54] R. C. Team, "R: A Language and Environment for Statistical Computing.," *R Found. Stat. Comput. Vienna, Austria.*, 2016.

[55] R. Darrell Bock and M. Lieberman, "Fitting a Response Model for n Dichotomously Scored Items," *Psychometrika*, vol. 35, no. 2, pp. 179–197, 1970.

[56] D. Rizopoulus, "ltm: An R package for Latent Variable Modelling and Item Response Theory Analyses," *J. Stat. Softw.*, vol. 17, no. 5, pp. 1–25, 2015.

[57] H. Frick, C. Strobl, F. Leisch, and A. Zeileis, "Psychometric Mixture Models," *J. Stat. Softw.*, 2015.

[58] F. Li, A. S. Cohen, S.-H. Kim, and S.-J. Cho, "Model Selection Methods for Mixture Dichotomous IRT Models," *Appl. Psychol. Meas.*, vol. 33, no. 5, pp. 353–373, 2009.

[59] S. W. Roberts, "Control Chart Tests Based on Geometric Moving Averages," *Technometrics*, vol. 42, no. 1, pp. 97–101, 1959.

[60] D. Montgomery, *Introduction to Statistical Quality Control*, 6th ed. New York: John Wiley & Sons, Inc, 2009.

[61] J. M. Lucas and M. S. Saccucci, "Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements," *Technometrics*, vol. 32, no. 1, pp. 1–12, Feb. 1990.

[62] "LogisticDistribution—Wolfram Language Documentation." [Online]. Available: https://reference.wolfram.com/language/ref/LogisticDistribution.html.

VITAE

PERSONAL DATA

Name	:	Irwan Rizadi
Date of Birth	:	May 24, 1989
Nationality	:	Indonesia
Address	:	KFUPM PO BOX 8122, Dhahran 31261, Saudi Arabia.
Phone Number	:	(966) 507485262
E-mail	:	irwanrizadi@gmail.com

EDUCATION

Master of Science (2016)

Applied Statistics, Department of Mathematics and Statistics, King Fahd University of Petroleum and Mineral, Dhahran, Saudi Arabia.
Advisor: Dr. Mohammad Hafidz Omar
Thesis: *Two-stage Mixture Logistic Distribution*

Bachelor of Science (2012)

Statistics, Department of Mathematics and Natural Science, Gadjah Mada University, Indonesia.
Advisor: Professor Sri Haryatmi Kartiko
Thesis: *Two-stage Cluster Sampling with Horvitz-Thompson Estimator Approach*

WORKING EXPERIENCES

1. Part time jobs at KFUPM, Saudi Arabia. **(2014 - present)**.
   - Writing a research paper under project (IN141013) with Dr. Khairul Saleh
     - Reporting research finding results
     - Statistical programming in MATLAB
   - Data analyst for ABET accreditation at Department of Architecture.
   - Grading statistics and actuarial student work (Course STAT 201, STAT 211, AS 381, and AS 482).
   - IT support in ITC Department.
2. Human Resource Development at PT. Trakindo Utama. **(2013-2014)**.
   - Giving training to senior high school students about technical skills.
   - Motivation training to students and teachers in senior high school.
   - Establishing communication and relationship with local government and schools.
3. Statistician at PT. Charoen Pokphand Indonesia, Tangerang. **(2012-2013)**.
   - Proficiency testing ISO 17525.
   - Laboratory data analysis.
   - Statistical programming with EXCEL macros.

- Reporting to General Manager.
- Supporting other departments in the company for any statistical analysis.

PUBLICATIONS

[1] Saleh, K. and **Rizadi, I**. (2017). *Exploited Stochastic Predator Prey System with Allee Effect and Constant Prey Refuge.* Under Preparation.

[2] Omar, M. H., **Rizadi, I**., and Riaz M. (2017). *Estimation of The Mean of a Mixture Logistic Distribution and Its Possible Application*. Under Preparation.

SEMINARS PRESENTED

1. (December, 2016). Thesis Defense at KFUPM. Two-stage Mixture Logistic Distribution.
2. (May, 2015). Math Seminar at KFUPM. Bivariate Poisson Gamma Distribution.
3. (June, 2012). Thesis Defense at UGM. Two-stage Cluster Sampling with Horvitz-Thompson Estimator Approach.

SEMINARS AND WORKSHOPS ATTENDED

1. (March, 2015). Introduction to MATLAB at KFUPM.
2. (December, 2014). Tips for Publishing Research Work in ISI Journal at KFUPM.
3. (August, 2013). Database Management using Statistical Analysis at Jakarta, Indonesia.
4. (May, 2009). Application of Statistics in Business and Industry at UGM, Indonesia.

SKILLS

1. Good leadership and communication skills.
2. Expertise in statistical data analysis using Eviews, SPSS, Minitab, R, Phyton, Mathematica, MATLAB, and many others.
3. Good knowledge in psychometric analysis using ltm and psychomix package in R.
4. Ability to automation EXCEL macros.
5. Advanced knowledge in Microsoft Office.
6. Ability to communicate fluently in English.