UNIVERSITYOFNAIROBI

NEGATIVE BINOMIAL-THREE PARAMETER LINDLEY DISTRIBUTION AND ITS PROPERTIES

BY

Wanjala Nanjala Anjela

156/38091/2020

A Thesis Submitted to the Department of Mathematics for Examination in Partial Fulfillment of the Requirements for the Award of Degree of Master of Science in Mathematical Statistics of the University of Nairobi

July 2023

Negative Binomial- Three Parameter Lindley Distribution and its properties.

Research Report in Mathematics, July 2023

Wanjala Nanjala Anjela

Department of Mathematics
Faculty of Science and Technology
Chiromo, off Riverside Drive
30197-00100 Nairobi, Kenya

Master Thesis

Submitted to the Department of Mathematics in partial fulfillment fora degree in Master of Science in Mathematical Statistics

Submitted to: The Department of Mathematics, University of Nairobi, Kenya

Abstract

In this study, we introduce a mixed distribution by mixing the distributions of negative binomial and three Parameter Lindley distribution This new mixed distribution has a thick tail and may be considered as an alternative for modelling count data with overdispersion. The properties and special cases of the compound distribution are studied. In addition, the parameter estimation for the compound distribution via the method of moments (MME) and the Maximum Likelihood Estimation (MLE) are provided. We present the performance of the Poisson, NB, L3,NBL, NBL2 and NBL3 distribution using real data in terms of log-likelihood, p-value, AIC (Akaike Information Criteria), BIC(Bayesian Information Criteria) and Kolmogorov-Smirnov statistic. It is shown that the negative binomial three parameter Lindley distribution provides a better fit compared to the other distributions under study for fitting over dispersed count data.

MasterThesis in Mathematics at the University of Nairobi, Kenya.
ISSN 2410-1397: **Research Report in Mathematics**

DISTRIBUTOR:Department of Mathematics, University of Nairobi, Kenya

Declaration and Approval

I the undersigned declare that this dissertation is my original work and to the best of my knowledge, it has not been submitted in support of an award of a degree in any other university or institution of learning.

	07-08-2023
Signature	Date

WANJALA NANJALA ANJELA
RegNo.156/38091/2020

In my capacity as a supervisor of the candidate's dissertation, I certify that this dissertation has my approval for submission.

	07-08-2023
Signature	Date

Dr George Muhua
Department of Mathematics,
University of Nairobi,
Box 30197, 00100 Nairobi, Kenya.
E-mail: muhuageorge@gmail.com

Dedication

I dedicate this project to Favour Bradley, Tasha Ashley and Talia Terryl.

Acknowledgements

All praise to Almighty God for granting me life and good health to further my studies. Special thanks to my Supervisor Dr. George Muhua for his continuous support, guidance, patience and prompt response throughout the process ensuring a successful project. My gratitude to the Department of Mathematics- University of Nairobi for giving me an opportunity to expand my knowledge and skills in the field of Statistics, as well as giving me a capable personnel to guide me through my academic life.
I wish to thank members of my group discussion for making studying more enjoyable.
Lastly, I thank my mother Eunice Nato for always encouraging me to carry on and My husband Chris Liambila for all the support you gave me throughout the entire period.

Wanjala Nanjala Anjela

July 2023.

Contents

Accronyms

pdf	Probability Density Function.
pmf	Probability Mass Function
cdf	Cumulative Distribution Function.
MLE	Maximum Likelihood Estimator
NB	Negative Binomial distribution
L3	Three Parameter Lindley distribution
BIC	Bayesian Information Criterion
AIC	Akaike Information Criterion
KS	Kolmogorov Smirnov Statistic

Chapter 1

GENERAL INTRODUCTION

1.1 Introduction

In this chapter, a background of studying mixture distributions is discussed. A brief description of the research problem is given, followed by the objectives of the study.

1.2 Background Information

Since 1894 the concept of mixture distribution was studied by a number of authors such as Blischke who defined mixture distribution as a weighted average of probability distribution with positive weights, which were probability distributions, called the mixing distributions that sum to one.

The Poisson distribution is typically used to fit count data when the number of events is randomly distributed over the time and/or space over which the event counts occur. In practice, however, observed count data often exhibit features such as overscattering and underscattering that commonly occur in applied data analysis(Rainer, 2000). Greenwood and Yule (1920) proposed a model with a Poisson mean gamma, or negative binomial (NB) distribution. The NB distribution is gaining popularity as a more flexible alternative to the Poisson distribution (Johnson et al.,)

In most cases, for mixed distributions, especially mixed Poisson binomial and mixed negative binomial, Increased customizability of count data compared to others handouts.

The overdispersion problem is usually solved by introducing mixed NBs. distribution. Several studies have shown that the mixed NB distribution fits the count data better. Compare with Poisson and NB distributions. These include the negative binomial beta exponential (Pudprommaratet al.2012) and the negative binomial Erlang (Kongrod et al. 2014). The Lindley distribution is It has been popularized by many researchers in recent years. Lindley distribution is a mixture of exponential distributions

The NB distribution is suitable for scattered count data that are not necessarily strong. A very heavy tail implies overdispersion, but the opposite is not true (Wang, 2011). Traditional statistical distributions and models, such as Poisson and NB distributions, cannot be used effectively for count data with strong tails. The Poissondistribution tends to underestimate the number of zeros given the mean of the data, whereas the NB distribution may overestimate the zeros and underestimate the observations to qualify as count data.

Many researchers have proposed mixed distributions. It is one of the most important methods for obtaining new probability distributions in applied probability and operations research (Gómez-Déniz et al., 2008). In this work, we consider mixed NB distributions as a more flexible alternative for analyzing count data, especially countdata with overdispersion. It is a combination of the NB distribution and the three-parameter Lindley distribution.

1.3 Research Problem

The traditional statistical distributions, such as the Poisson and Negative-Binomial distributions, cannot be used effectively for heavy-tail count data. we are therefore considering the mixture Negative Binomial-3 Parameter lindley distribution as a more flexible alternative to analyze overdispersed count data.

1.4 Ojectives

The main objective of this study is to propose a mixture of Negative Binomial and three parameter Lindley distributions that can be used to analyze count data with overdispersion.

1.4.1 Specific objectives

1. To study the properties of negative binomial three parameter lindley distribution and identify the special cases.

2. To estimate parameters of the distribution by using MLE and method of moments.

3. To compare efficiency of proposed distribution over Posisson, negative binomial, Three parameter Lindley, Negative Binomial Lindley and Negative Binomial weighted Lindley distributions.

Chapter 2

LITERATURE REVIEW

2.0.1 Introduction

In this chapter, we look at various work that has been done in the literature on mixtures of Negative Binomial and Lindley distributions.This chapter provides related literature useful in undertaking our research.It con- sists of researcher's critique and comparisons from other studies related to our area of interest.They include methodologies used,theoretical or conceptual framework,relationships and differences between studies.Critical aspects of reviewing the literature in clear and systematic way from existing studies has been done and missing gaps identified

2.0.2 Mixed Negative Binomial distributions

Lindley (1958) at first proposed a one-parameter distribution called the Lindley distribution which is a finite mixture of exponential (θ) and gamma ($2, \theta$) distributions. where θ represents the scaling parameter.

Bowman et al (1992) derived a large number of new Binomial mixtures distributions by assuming that the probability parameter p varied according to some laws, mostly derived from frullani integrals. They used the transformation $e =$

p^{-t} and considered various densities for the transformed variables. They also gave graphical representations for some of the more significant distributions.

Alanko and Duffy (1996) developed a class of Binomial mixtures arising from transformations of the Binomial parameter p as $1 - e^{-t}$ where λ was treated as a random variable. They showed that this formulation provided closed forms for the marginal probabilities in the compound distribution if the Laplace transform of the mixing distribution could be written in a closed form. They gave examples of the derived compound Binomial distributions; simple properties, and parameter estimates from moments and maximum likelihood estimation. They further illustrate the use of these models by examples from consumption processes.

G'omez (2006) proposed a new compound negative binomial distribution by mixing the p negative binomial parameter with inverse Gaussian distribution. Basic properties of the new distribution were given, three estimation of parameters method were given using method of moments, maximum likelihood method and zero proportion method. Finally, examples of application for both univariate and bivariate cases were given.

Zamani and Ismail (2010) came up with negative binomial – Lindley distribution which provides a better fit compared to the Poisson and the negative binomial for count data where the probability at zero has a large value. They gave simple properties, and parameter estimates from method of moments and maximum likelihood estimation. They also illustrated the use of model by examples from insurance count data.

Bodhisuwan and Zeephongsekul (2012) introduced a new distribution and a more flexible alternative to Poisson distribution when count data are overdispersed in the form of a Negative Binomial – Beta Exponential (NB – BE) distribution. They gave properties and parameters estimation using maximum likelihood method.

Shanker and Mishra (2013a, 2013b), Shanker and Amanuel (2013), and Shanker et al. (2013) Obtain various forms of the two-parameter Lindley distribution, they They discussed various properties such as skewness, kurtosis,hazard rate function, meanremaining lifetime function, probabilistic order, Mean Deviation, Stress Intensity Reliability.

Chapter 3

REVIEW OF POISSON,AND RELATED DISTRIBUTIONS

3.1 Introduction

In this chapter, a review of distributions used to fit count data is discussed. These distributions include Poisson, Negative Binomial distribution, One Parameter Lindley Distribution, Two Parameter Lindley Distribution, Three Parameter Lindley Distribution, Negative Binomial One Parameter Lindley Distribution, Negative Binomial two parameter Lindley Distribution.

3.2 Poisson Distribution

A random variable X is said to have a Poisson distribution with parameter μ if it takes integer values x = 0, 1, 2, . . . with probability

$$f(x;\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!};x = 0, 1, 2, ... \text{ and } \lambda > 0. \tag{3.2.1}$$

In actuality, however, observed count data frequently exhibit characteristics like over- or under-dispersion, which are typical in applied data analysis (Rainer, 2000).

Because the Poisson distribution has an equal mean and variance, any factor that influences one will also affect the other.

3.3 The Negative Binomial Distribution

As a more adaptable substitute for the Poisson distribution, the NB distribution has grown in popularity (Johnson et al., 2005).The Negative Binomial (NB) distribution is another distribution for count data. The NB distribution is often employed in case where a distribution is over-dispersed, i.e., its variance is greater than the mean which relaxes the equality of mean and variance property of the Poisson distribution. If X denotes a random variable distributed under a NB distribution with parameter r and p, then its probability mass function (pmf) is given by:

$$f(x;r,p) = \binom{x+r-1}{x} p^r(1-p)^x, x = 0, 1...., r > 0, 0 < p < 1 \quad (3.3.1)$$

Case 1 Let X= the total number of failures before the rth success. therefore x+r-1= the total number of trials before the rth success. Therefore,

$$P_k = Prob(X = k)$$

= [probability of having (r-1) successes out of (x+r-1) trials] x probability of achieving the rth success].

$$= [\begin{pmatrix} x+r-1 \\ r-1 \end{pmatrix} p^{r-1}q^{x}]p$$

$$= \begin{pmatrix} x+r-1 \\ r-1 \end{pmatrix} q^{x}p^{r}$$

$$= \begin{pmatrix} x+r-1 \\ k \end{pmatrix} q^{k}p^{r}, \text{k=0,1,2...}$$

Case 2 Let Y be the total number of trails required to achieve r successes. If Y=k then k-1= the number of trials required to obtain the first (r-1) successes

$$P_k = Prob(Y = k)$$

=[Probability of obtaining r-1 successes out of k-1 trials]x [probability of obtaining the rth success]

$$= [\begin{pmatrix} k-1 \\ r-1 \end{pmatrix} p^{r-1}q^{k-1}]p$$

$$= \begin{pmatrix} k-1 \\ r-1 \end{pmatrix} q^{k-r}p^{r}, \text{ k=r, r+1, r+2,....}$$

Its mean and variance are

$$E(X) = r(\frac{1-p}{p}) \tag{3.3. 2}$$

and

$$Var(X) = r(\frac{1-p}{p^2}) \tag{3.3.3}$$

respectively

For $X =\sim NB(r, p)$, the parameter p can be represented in terms of r, as $p =$

$\frac{r}{\mu+r}$ where μ is the mean response and r is the inverse of the dispersion parameter. Therefore the probability mass function of x can be written as follows:-

$$f(x;r,p) = \Gamma(r+x)(\frac{r}{\mu+r})^r[\frac{\mu}{\mu+r}]^x, x = 0, 1, 2, ...$$

where $\mu >= 0, r > 0\Gamma(.)$ is a complete gamma function. Then, its mean and variance can be defined as

$$E(x) = \mu$$

and

$$var(x) = \mu + (\frac{1}{r})\mu^2$$

We are going to apply the following identity to derive the moment generating function of the negative binomial distribution:-

$$\begin{pmatrix} -r \\ x \end{pmatrix} = (-1)^x \begin{pmatrix} x+r-1 \\ x \end{pmatrix}$$

we can prove that in this in the following way

$$\binom{-r}{x} = \frac{(-r)(-r-1).....(-r-x+1)}{x!}$$

$$= (-1)^x \frac{(r+x-1)....(r+1)r}{x!}$$

$$= (-1)^x \binom{x+r-1}{x}$$

Now,

$$M_{(t)} = \sum_{x=0}^{\infty} e^{tx} \binom{x+r-1}{x} (1-p)^x * p^r$$

Grouping terms and using the above identity we get,

$$M_{(t)} = p^r \sum_{x=0}^{\infty} \binom{x+r-1}{x} [e^t(1-p)]^x$$

$$= p^r \sum_{x=0}^{\infty} \binom{-r}{x} (-1)^x [e^t(1-p)]^x$$

$$= p^r \sum_{x=0}^{\infty} \binom{-r}{x} [-e^t(1-p)]^x$$

By using Newton's Binomial theorem,

$$(x+1)^r = \sum_{i=1}^{\infty} \binom{r}{i} x^i$$

provided that $| x |< 1$.

the last term becomes

$$M_{(t)} = \frac{p^r}{[1-(1-p)e^t]^r} \tag{3.1.4}$$

provided that $t < -log(1-p)$

3.4 One Parameter Lindley Distribution

This is a finite mixture of an exponential distribution and gamma distribution given as follows; Let,

$f_1(x) = \theta e^{-\theta x}\ x, \theta > 0$, which is exponential with parameter θ.

$f_2(x) = \theta^2 x$ which is gamma distribution with parameters 2 and θ

its pdf is given by

$$f(x;\theta) = \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x} \tag{3.4.1}$$

$x, \theta > 0$

It is clear that this distribution is a combination of the exponential distribution with parameter θ (or Gamma distribution(1, θ)) and Gamma distribution with parameters $(2, \theta)$ as follows:-

$$f(x;\theta) = \pi f1(x) + (1-\pi)f2(x)$$

where the mixing proportion π is equal to

$$\frac{\theta}{\theta+1}$$

3.5 The Negative Binomial Weighted Lindley distribution

Sunthree et al. (2018) created this brand-new mixed NB distribution, which is an NB-WL distribution created by combining the NB distribution with a WL distribution. This distribution consists three parameters, namely, r, θ and α.

TheoremLet $X \mid \lambda$ be a random variable following a NB distribution with parameters r and $p = \exp(-\lambda), X \mid \lambda \sim \text{NB}(r, p = \exp(-\lambda))$. If λ is distributed as the WL distribution with positive parameters θ and α, denoted by $\lambda \sim \text{WL}(\theta, \alpha)$), then X is called a NB-WL random variable.

Theorem 1. Let $X \sim \text{NB-WL}(r, \theta, \alpha)$. The pmf of X is given by

$$f(x; r, \theta, \alpha) = \binom{r+x-1}{x} \sum_{j=0}^{x} \binom{x}{j} (-1)^j \frac{(\theta-\alpha)(\theta-\alpha+r+j+1)}{(\theta-\alpha+1)(\theta-\alpha+r+j)^2}, \quad x = 0, 1, 2, \ldots \tag{3.4.1}$$

where $\theta > 0$ and $\theta > \alpha$

Proof. If $X \mid \lambda \sim \text{NB}(r, p = \exp(-\lambda))$ and $\lambda \sim \text{WL}(\theta, \alpha)$), then the pmf of X can be obtained by $f(x) = \int_0^\infty f_1(x \mid \lambda) g(\lambda; \theta, \alpha) d\lambda$, where $f_1(x \mid \lambda)$ is express as

$$f_1(x \mid \lambda) = \begin{pmatrix} r+x-1 \\ x \end{pmatrix} \exp(-\lambda r)(1-\exp(-\lambda))^x$$

$$= \begin{pmatrix} r+x-1 \\ x \end{pmatrix} \sum_{j=0}^{x} \begin{pmatrix} x \\ j \end{pmatrix} (-1)^j \exp(-\lambda(r+j))$$

By substituting $f_1(x \mid \lambda)$ into $f(x) = \int_0^\infty f_1(x \mid \lambda) g(\lambda;\theta,\alpha) d\lambda$, thus

$$f(x) = \begin{pmatrix} r+x-1 \\ x \end{pmatrix} \sum_{j=0}^{x} \begin{pmatrix} x \\ j \end{pmatrix} (-1)^j \left(\int_0^\infty \exp(-\lambda(r+j)) g(\lambda;\theta,\alpha) d\lambda \right)$$

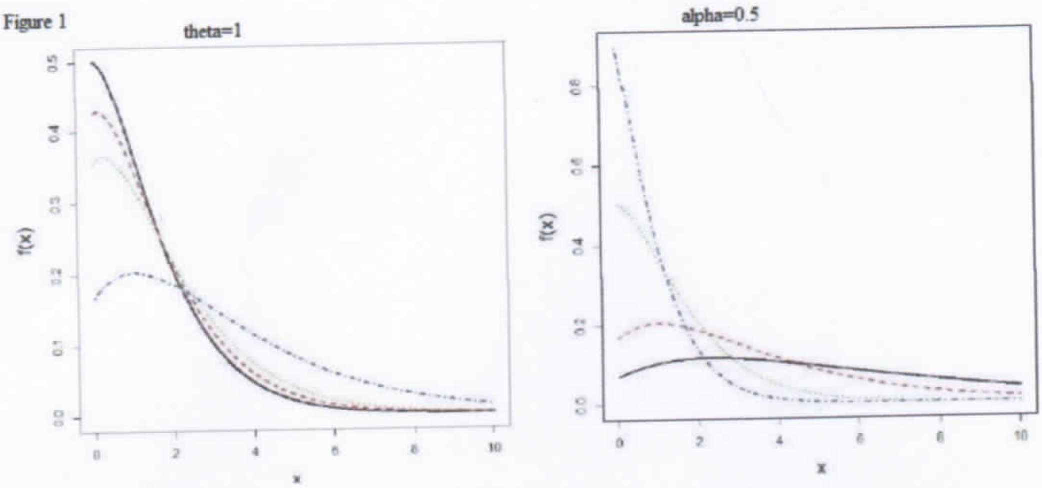
$$= \begin{pmatrix} r+x-1 \\ x \end{pmatrix} \sum_{j=0}^{x} \begin{pmatrix} x \\ j \end{pmatrix} (-1)^j M_\lambda(-(r+j))$$

Substituting $M_\lambda(-(r+j))$ the mgf of the WL distribution in the equation above, the pmf of the NB $-$WL(r,θ,α) is given as

$$f(x;r,\theta,\alpha) = \begin{pmatrix} r+x-1 \\ x \end{pmatrix} \sum_{j=0}^{x} \begin{pmatrix} x \\ j \end{pmatrix} (-1)^j \frac{(\theta-\alpha)(\theta-\alpha+r+j+1)}{(\theta-\alpha+1)(\theta-\alpha+r+j)^2},$$

The pmf of the NB-WL distribution of some specified values of r, θ and α are as shown

Figure 1



Theorem 2. If $X \sim \mathrm{NB-WL}(r, \theta, \alpha)$, the factorial moment of order a of X is

$$\mu'_a(X) = \frac{\Gamma(r+a)}{\Gamma(r)} \sum_{j=0}^{a} \binom{a}{j} (-1)^j \frac{(\theta-\alpha)(\theta-\alpha+r+j+1)}{(\theta-\alpha+1)(\theta-\alpha+r+j)^2}, \quad x = 0, 1, 2, \dots \tag{3.4.2}$$

for $\theta > 0$ and $\theta > \phi$.

Proof: According to Gómez-Déniz et al. (2008), the mixed NB distribution's factorial moment of order a can be written in terms of an elementary function by

$$\mu'_a(X) = E_\lambda\left(\frac{\Gamma(r+a)}{\Gamma(r)} \frac{(1-\exp(-\lambda))^a}{\exp(-\lambda a)}\right) = \frac{\Gamma(r+a)}{\Gamma(r)} E_\lambda(\exp(\lambda)-1)^a.$$

Using the binomial expansion of $(\exp(\lambda)-1)^a$, then $\mu'_a(X)$ can be written as

$$\mu'_a(X) = \frac{\Gamma(r+a)}{\Gamma(r)} \sum_{j=0}^{a} \begin{pmatrix} a \\ j \end{pmatrix} (-1)^j E_\lambda(\exp(\lambda(a-j))) = \frac{\Gamma(r+a)}{\Gamma(r)} \sum_{j=0}^{a} \begin{pmatrix} a \\ j \end{pmatrix} (-1)^j M_\lambda(a-j).$$

From the mgf of the NWL distribution with $t = a - j$, the $\mu'_a(X)$ is finally given as

$$\mu'_a(X) = \frac{\Gamma(r+a)}{\Gamma(r)} \sum_{j=0}^{a} \begin{pmatrix} a \\ j \end{pmatrix} (-1)^j \frac{(\theta-\alpha)(\theta-\phi+r+j+1)}{(\theta-\phi+1)(\theta-\phi+r+j)^2}.$$

Definition 2. Let $X \sim$ NB-WL(r, θ, α). some properties of X are as follows

1. The first two moments about zero of X are

$$E(X) = r(\varpi - 1),$$
$$E\left(X^2\right) = r(r+1)\varpi_2 - r(2r+1)\varpi_1 + r2,$$

2. The mean and variance of X respectively, are

$$\begin{aligned} E(X) &= r(\varpi - 1), \\ \mathrm{Var}(X) &= r(r+1)\varpi_2 - r(1 + r\varpi)\varpi_1. \end{aligned} \tag{3.4.3}$$

where $\varpi_k = \frac{(\theta-\alpha)(\theta-\alpha+k+1)}{(\theta-\alpha+1)(\theta-\alpha+k)^2}$.

3.6 Three Parameter Lindley Distribution

R. Shanker et al(2017) introduced a three parameter Lindley Distribution

$$f(x;\alpha,\beta,\theta) = \frac{\theta^2}{\theta\alpha+\beta}(\alpha+\beta\lambda)\mathrm{e}^{-\theta x} \tag{3.5.1}$$

$x, \beta, \alpha, \theta > 0, \alpha\theta + \beta > 0$

This distribution can easily be expressed as a mixture of exponential (θ) and gamma $(2, \theta)$ with mixing proportion

$$\frac{\alpha\theta}{\alpha\theta - \beta}$$

we have

$$f(x;\theta,\alpha,\beta) = pg_1(x) + (1-p)g_2(x)$$

$$p = \frac{\theta\alpha}{\theta\alpha+\beta}$$

$$g_1(x) = \theta e^{-\theta x}$$

$$g_2(x) = \theta^2 x e^{-\theta x}$$

Proof

$$f(x) = p_1 * \Gamma(1,\theta) + p_2 * \Gamma(2,\theta)$$

$$= \frac{\alpha\theta}{\alpha\theta+\beta}(\theta e^{-\theta x}) + \frac{\beta}{\alpha\theta+\beta}(\theta^{\beta} e^{-\theta x} x)$$

$$= \frac{\theta^2}{\alpha\theta+\beta}[\alpha + \frac{\beta\theta(\theta x)^{2-1}}{\theta^2}]e^{-\theta x}$$

$$= \frac{\theta^2}{\alpha\theta+\beta}[\alpha + \frac{\beta(\theta x)}{\theta}]e^{-\theta x}$$

$$= \frac{\theta^2}{\alpha\theta+\beta}[\alpha + \beta x]e^{-\theta x}$$

$$= \frac{\theta^2}{\alpha\theta+\beta}[\alpha + \beta x]e^{-\theta x} \quad (3.5.2)$$

The graph of the pdf of the Three parameter Lindley distribution for different values of θ, β, with a constant value of α are as shown below:-

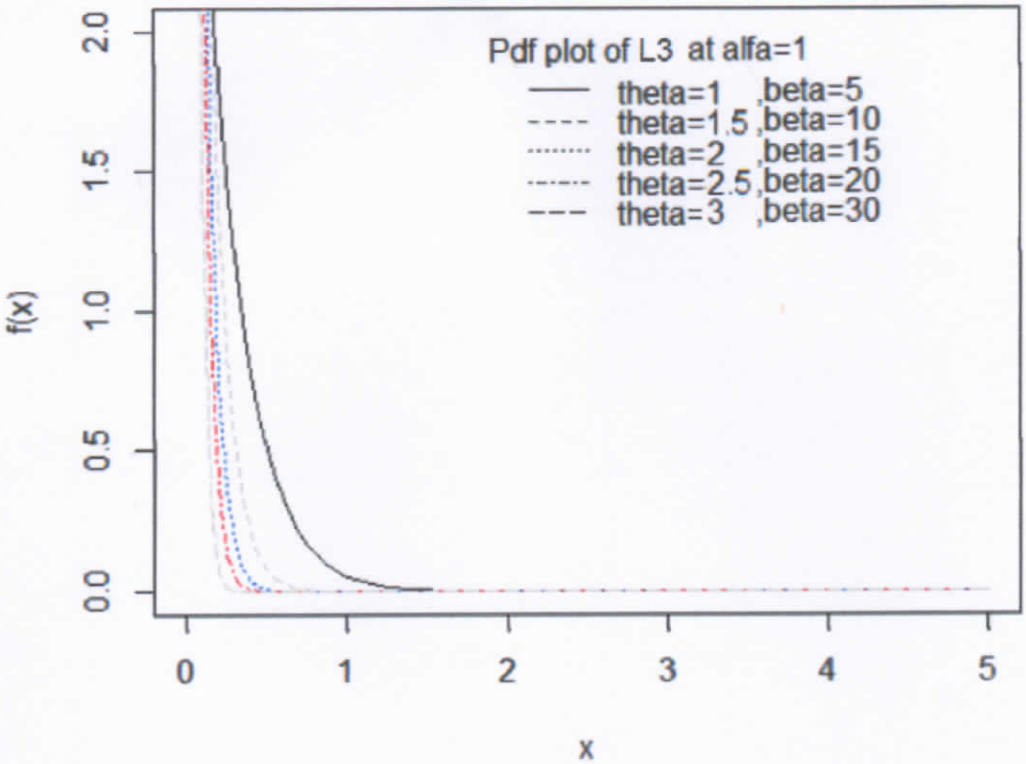


Figure 2.1 displays the probability density function's shapes. They claim that the parameter θ, whereas β is the weight parameter, is a shape parameter since various values of *theta* alter the overall shape.

The cumulative density function

$$F(x) = \int_0^\infty \frac{\theta^2}{\alpha\theta + \beta} f(x) dx$$

$$= \int_0^\infty \frac{\theta^2}{\alpha\theta + \beta} [\alpha + \beta x] \mathrm{e}^{-\theta x} dx$$

$$F(x) = \frac{\theta^2}{\alpha\theta + \beta} \int_0^\infty [\alpha + \beta x] \mathrm{e}^{-\theta x} dx$$

$$= \frac{\theta^2}{\alpha\theta + \beta} * I_l$$

$$I_l = \int_0^\infty (\beta x + \alpha) \mathrm{e}^{-\theta x} dx$$

$$= \beta x + \alpha = \frac{du}{dx} = \beta$$

$$dv = \mathrm{e}^{-\theta x} = v = \int_0^\infty \mathrm{e}^{-\theta x} dx = \frac{-\mathrm{e}^{-\theta x}}{\theta}$$

$$I_1 = [\frac{-(\beta x + \alpha)\mathrm{e}^{-\theta x}}{\theta}]_0^\infty - \int_0^\infty \frac{-\beta \mathrm{e}^{-\theta x}}{\theta} dx$$

$$I_2 = \int_0^\infty \frac{-\beta \mathrm{e}^{-\theta x}}{\theta} dx$$

$$u = -\theta x = \frac{du}{dx} = -\theta$$

$$= \frac{\alpha}{\theta^2} \int_0^\infty e^u du.$$

$$but \int_0^\infty e^u du = e^u$$

$$I_2 = \frac{\beta}{\theta^2} \int_0^\infty e^u du = \frac{\beta e^u}{\theta^2} = \frac{\beta \mathrm{e}^{-\theta x}}{\theta^2}$$

$$I_1 = \frac{-(\beta x + \alpha)\mathrm{e}^{-\theta x}}{\theta} - \frac{\beta \mathrm{e}^{-\theta x}}{\theta^2}$$

$$F(x) = \frac{\theta^2}{\theta\alpha + \beta} \int_0^\infty (\beta x + \alpha) \mathrm{e}^{-\theta x}$$

$$= \frac{-\theta(\beta x + \alpha)e^{-\theta x}}{\theta\alpha + \beta} - \frac{\beta e^{-\theta x}}{\theta\alpha + \beta}$$

$$= 1 - \frac{(\theta\beta x + \theta\alpha + \beta)e^{-\theta x}}{\theta\alpha + \beta}$$

$$F(x) = 1 - [1 + \frac{\theta\beta\alpha}{\theta\alpha + \beta}]e^{-\theta x}$$

$$F(x;\theta,\alpha,\beta) = 1 - [1 + \frac{\theta\beta x}{\theta\alpha + \beta}]e^{-\theta x}$$

$$x, \theta, \beta > 0, \theta\alpha + \beta > 0$$

$$F(x \mid \theta, \alpha, \beta) = 1 - \left(1 + \frac{\beta\theta x^{\alpha}}{\theta + \beta}\right) e^{-\theta x^{\alpha}}$$

Theorem Let λ be a random variable with the parameters α, β and θ in a three-parameter Lindley distribution. Then, λ's mgf is provided by

$$M_\lambda(t) = E(e^{t\lambda})$$

$= E(e^{t\lambda}) = \frac{\theta^2}{\theta\alpha+\beta}\left[\frac{\alpha(\theta-t)+\beta}{(\theta-t)^2}\right]$ where $t = -(r + j)$

$$M_\lambda(t) = E\left(e^{t\lambda}\right) = \int_0^\infty e^{t\lambda}\frac{\theta^2}{\theta\alpha + \beta}(\alpha + \beta\lambda)e^{-\theta\lambda}d\lambda$$

Proof:

$$= \frac{\theta^2}{\theta\alpha + \beta}\left[\alpha\int_0^\infty e^{-\lambda(\theta-t)}d\lambda + \beta\int_0^\infty \lambda e^{-\lambda(\theta-t)}d\lambda\right]$$

$$= \frac{\theta^2}{\theta\alpha + \beta}\left[\alpha\left(\frac{1}{\theta - t}\right) + \beta\left(\frac{1}{(\theta - t)^2}\right)\right] = \frac{\theta^2}{\theta\alpha + \beta}\left[\frac{\alpha(\theta - t) + \beta}{(\theta - t)^2}\right]$$

$$= \frac{w_1}{w_0} \tag{3.5.2}$$

where

$$w_2 = \frac{\alpha(\theta-2)+\beta}{(\theta-2)^2}$$

$$w_1 = \frac{\alpha(\theta-1)+\beta}{(\theta-1)^2}$$

$$w_0 = \frac{\alpha\theta+\beta}{\theta^2}, \theta \neq 0, 1, 2.$$

Chapter 4

THE NEGATIVE BINOMIAL-THREE PARAMETER LINDLEY DISTRIBUTION

4.1 Introduction

In this chapter, a new distribution which is a mixture of Negative Binomial and Three parameter Lindley distribution is Introduced. Its properties are studied, including its special cases. A simulation study is conducted to determine the consistency of the parameter estimates of the new distribution and finally data analysis is done to compare the performance of the new distribution over all other related distributions.

4.2 Definition

If the random variable X complies with the following stochastic formulation, it has a Negative Binomial-Three Parameter Lindley distribution.

$$X \mid r, \lambda \sim \mathrm{NB}\left(\mathrm{r}, \mathrm{p} = \mathrm{e}^{-\lambda}\right) \tag{4.2.1}$$

and

$$\lambda \sim L_3(\lambda \mid \alpha, \beta, \theta) \tag{4.2.2}$$

where,$x > 0$, $r > 0$ and $\theta, \alpha, \beta > 0$.

$$\lambda \sim f(\lambda \mid \alpha, \beta, \theta) = \frac{\theta^2}{\theta\alpha + \beta}(\alpha + \beta\lambda)\mathrm{e}^{-\theta x}$$

distribution refers to three Parameter Lindley distribution as proposed by Shanker et al (2017)

Theorem 4.2.1 Let X be a random variable which follows a Negative Binomial three parameter Lindley distribution with parameters r, α, β and θ. Then, the pmf of X is given by $\Pr(X = x \mid r, \alpha, \beta, \theta)$

$$= \binom{x+r-1}{x} \frac{\theta^2}{\alpha\theta + \beta} \sum_{j=0}^{x} \binom{x}{j} (-1)^j \left[\frac{\alpha(\theta + r + j) + \beta}{(\theta + r + j)^2}\right]; \quad x = 0, 1, \ldots \tag{4.2.3}$$

where $\alpha > 0, \beta > 0$ and $\theta > 0$.

Proof: If $X \mid r, \lambda \sim NB\left(r, p = e^{-\lambda}\right)$ and $\lambda \sim L_3(\alpha, \beta, \theta)$, then the marginal distribution for X can be obtained using

$\Pr(X = x \mid r, \alpha, \beta, \theta) = \int_0^\infty \Pr(X = x \mid r, \lambda) f(\lambda \mid \alpha, \beta, \theta) d\lambda$

Know that

$$\left(1 - e^{-\lambda}\right)^x = \sum_{j=0}^{x} \binom{x}{j} (-1)^j e^{-\lambda j} \tag{4.2.4}$$

Therefore,

$$\Pr(X = x \mid r, \lambda) = \binom{x+r-1}{x} \sum_{j=0}^{x} \binom{x}{j} (-1)^j e^{-\lambda(r+j)} \tag{4.2.5}$$

By using marginal distribution formula above, the pmf of the NBL_3 distribution can be obtained as

$$\Pr(X = x \mid r, \alpha, \beta, \theta)$$
$$= \binom{x+r-1}{x} \sum_{j=0}^{x} \binom{x}{j} (-1)^j \int_0^\infty e^{-\lambda(r+j)} f(\lambda \mid \alpha, \beta, \theta) d\lambda$$
$$= \binom{x+r-1}{x} \sum_{j=0}^{x} \binom{x}{j} (-1)^j M_\lambda[-(r+j)] \qquad 4.2.6$$

where $M_\lambda(t)$ is the moment generating function (mgf) of $L_3(\alpha, \beta, \theta)$.

Therefore the probability mass function of the three parameter Lindley Negative binomial distribution will be given by:-

$$\Pr(X = x \mid r, \alpha, \beta, \theta) = \binom{x+r-1}{x} \frac{\theta^2}{\alpha\theta+\beta} \sum_{j=0}^{x} \binom{x}{j} (-1)^j \left[\frac{\alpha(\theta+r+j)+\beta}{(\theta+r+j)^2}\right]; \qquad 4.2.7$$

$x = 0, 1, \ldots$ (2)

where $\alpha > 0, \beta > 0$ and $\theta > 0$.

and the corresponding cdf is

$$F_X(x) = \sum_{s \le x} f_x(x) = \sum_{s \le x} \binom{r+s-1}{s} \sum_{j=0}^{s} \binom{s}{j}$$
$$(-1)^j \frac{1}{\theta+1} \left(\frac{\theta^{\alpha+1}}{(\theta+r+j)^\alpha} + \frac{\theta^\beta}{(\theta+r+j)^\beta} \right) \qquad 4.2.8$$

4.3 Properties of Negative Binomial Three Parameter Lindley distribution

4.3.1 Shape of the Probability function

The graph of the pdf of the Negative Binomial three parameter Lindley distribution for different values of θ, β, α and r are as shown below:-

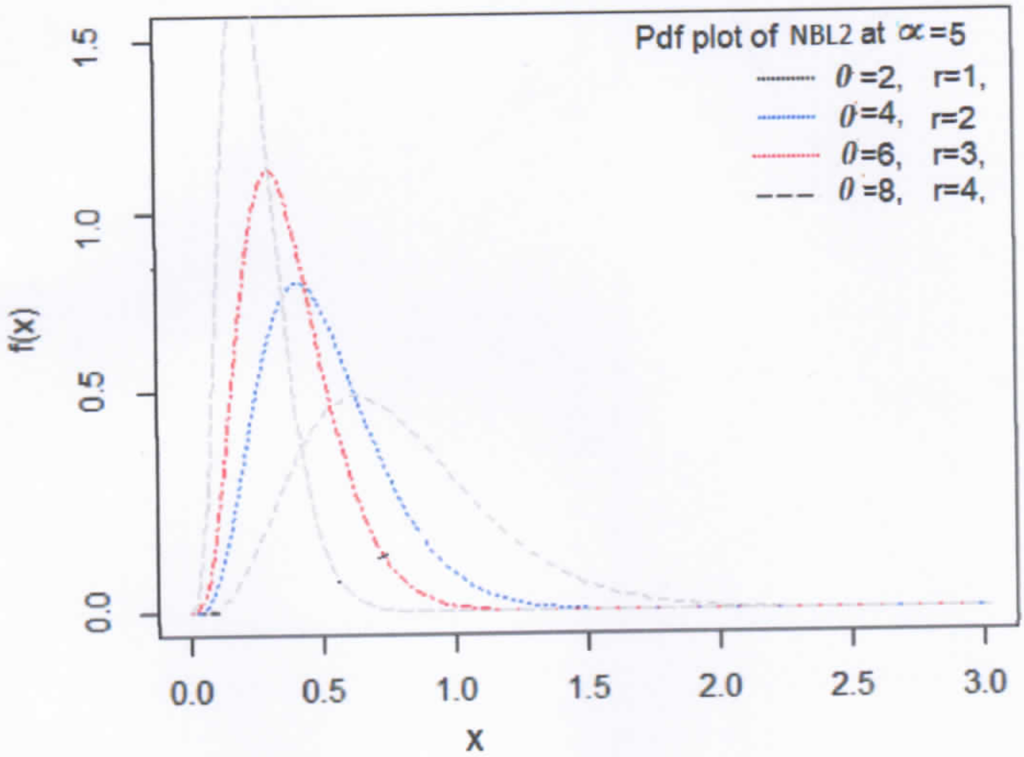
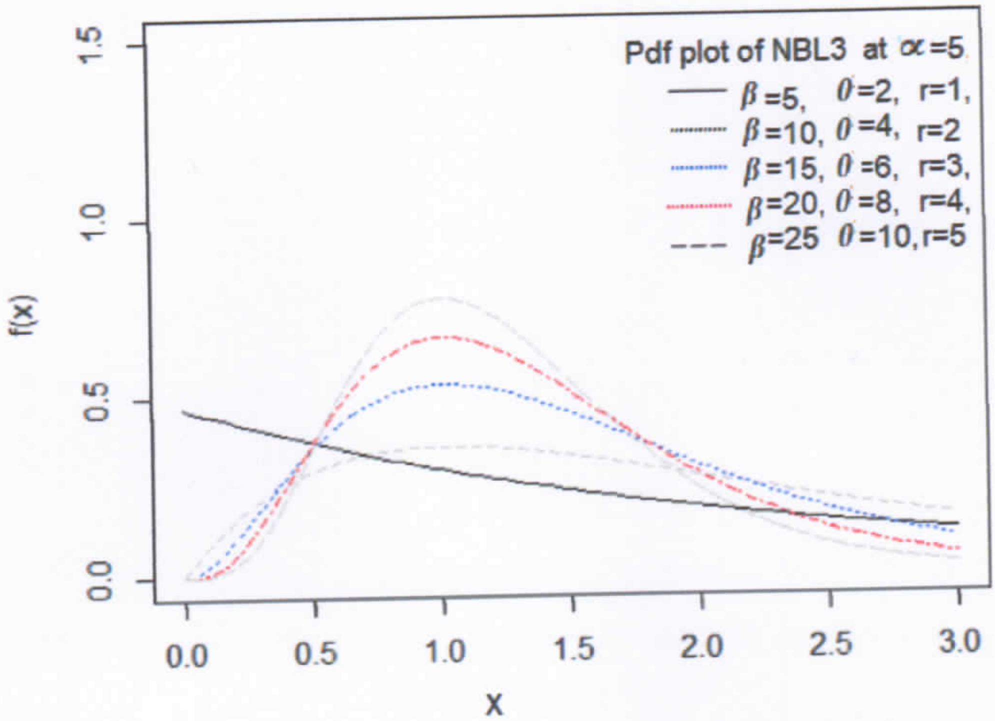

Pdf plot of NBL2 at α=5
θ=2, r=1,
θ=4, r=2
θ=6, r=3,
θ=8, r=4,
f(x)
0.0
0.5
1.0
1.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
x

Figure 4



Based on the figure, the pmf of NBL3 has the lowest mass at zero and the probability is significantly small at zero, when θ, β, r are large. However, when these parameters are small, the pmf of the NBL3 is right-tailed and has the highest mass at zero. Thus this proposed distribution is an alternative distribution to adequatelly fit the proportional data in the case where the probability at zero has either a small value or a large value. Also θ is a shape parameter since different values of θ change the overall shape of pmf hence determining the tail of the distribution. r is the dispersion parameter from Negative Binomial distribution, and it indicates whether the distribution is wide or narrow. From the pdf diagram, the large values of r indiate a narrow distribution whereas small values of r result to a wider disribution.

The Moment Generating Function

Given that $X \mid \lambda$ follows the NB distribution, $(r, p = e - \lambda)$, $X \mid \lambda$'s mgf will be

$$= p^r \sum_{x=0}^{\infty} \binom{-r}{x} [-e^t(1-p)]^x$$

$$= p^r \sum_{x=0}^{\infty} \binom{-r}{x} [-e^t + e^{t-\lambda}]^x$$

$$= p^r \sum_{x=0}^{\infty} \binom{-r}{x} (-e^t + e^(t - \lambda)x) \qquad 4.2.9$$

As a result, the law of total expectation may be used to obtain the moment generating function of the NBL3 distribution as shown below.

$$M_x(t) = E(e^{tx})$$

$$E_\lambda[e(e^{tx} \mid \lambda)]$$

$$= \int_0^\infty \sum_{x=0}^{\infty} \begin{pmatrix} -r \\ x \end{pmatrix} (e^t - 1)^x e^{-\lambda x} f_\lambda(\lambda) d\lambda$$

$$= \sum_{x=0}^{\infty} \begin{pmatrix} -r \\ x \end{pmatrix} (e^t - 1)^x \int_0^\infty e^{-\lambda x} f_\lambda(\lambda) d\lambda$$

$$= \sum_{x=0}^{\infty} \begin{pmatrix} -r \\ x \end{pmatrix} (e^t - 1)^x M_\lambda(-x)$$

$$= \sum_{x=0}^{\infty} \begin{pmatrix} -r \\ x \end{pmatrix} (e^t - 1)^x \frac{\theta^2}{\theta\alpha + \beta} \left[\frac{\alpha(\theta - t) + \beta}{(\theta - t)^2} \right]$$

$$= \sum_{x=0}^{\infty} \begin{pmatrix} -r \\ x \end{pmatrix} (e^t - 1)^x \frac{\theta^2}{\theta\alpha + \beta} \left[\frac{\alpha(\theta + r + j) + \beta}{(\theta +)^2} \right] \quad 4.3.1$$

4.3.2 Moments

If $X \sim NBL_3\ (r, \alpha, \beta, \theta)$, then the factorial moment of X at the k th degree becomes $\mu_{[\mathrm{k}]}(\mathrm{x}; \mathrm{r}, \alpha, \beta, \theta) = \frac{\Gamma(\mathrm{r}+\mathrm{k})}{(1+\theta)\Gamma(\mathrm{r})} \sum_{\mathrm{j}=0}^{\mathrm{k}} \begin{pmatrix} \mathrm{k} \\ \mathrm{j} \end{pmatrix} * (-1)^{\mathrm{j}} \left(\frac{\theta^{\alpha+1}}{(\theta-(\mathrm{k}-\mathrm{j}))^{\alpha}} + \frac{\theta^{\beta}}{(\theta-(\mathrm{k}-\mathrm{j}))^{\beta}} \right)$

$\mathrm{k} = 1, 2, 3, \ldots,$ and $\mathrm{r}, \alpha, \beta, \theta > 0$ 4.3.2

Proof. The k th factorial moment of X, i.e.,

$$\mu_{[k]}(\mathrm{x}; r, p) = E[X(X-1) \cdots (X-k+1)]$$

, is

$$\mu_{[k]}(x;r,p) = \frac{\Gamma(r+k)}{\Gamma(r)}\frac{(1-p)^k}{p^k}; k = 1,2,3,... \qquad 4.3.2$$

where $\Gamma(\cdot)$ is the complete gamma function, i.e.,

$$\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}dx, t > 0 \qquad 4.3.3$$

. where $p = e^{-\lambda}$ we can write it as follows:-

$$\mu_{[k]}\left(x;r,e^{-\lambda}\right) = E_\lambda\left[\frac{\Gamma(r+k)}{\Gamma(r)}\frac{\left(1-e^{-\lambda}\right)^k}{e^{-\lambda k}}\right] = \frac{\Gamma(r+k)}{\Gamma(r)}E_\lambda\left(e^\lambda - 1\right)^k \qquad 4.3.4$$

Using a binomial expansion in the term $\left(e^\lambda - 1\right)^k$ we can write as

$\mu_{[k]}\left(x;r,e^{-\lambda}\right) = \frac{\Gamma(r+k)}{\Gamma(r)}\sum_{j=0}^{k}(-1)^j E_\lambda\left(e^{\lambda(k-j)}\right)$

$= \frac{\Gamma(r+k)}{\Gamma(r)}\sum_{j=0}^{k}\binom{k}{j}(-1)^j M_\lambda(k-j).$

If

$$X \mid \lambda \sim NB\left(r, p = e^{-\lambda}\right)$$

and

$$\lambda \sim L_3(\alpha,\beta,\theta)$$

when substituting the mgf of λ as in (4) with $t = (k-j)$ into $\mu_{[k]}\left(x;r,e^{-\lambda}\right)$. The kth factorial moment of X is

$$\mu_{[k]}(x;r,\alpha,\beta,\theta) = \frac{\Gamma(r+k)}{(1+\theta)\Gamma(r)}\sum_{j=0}^{k}\binom{k}{j}(-1)^j\left(\frac{\theta^{\alpha+1}}{(\theta-(k-j))^\alpha} + \frac{\theta^\beta}{(\theta-(k-j))^\beta}\right). \qquad 4.3.5$$

We can determine the first four moments, or the variance and skewness, from the factorial moment of the $NB - L_3$ distribution. $E(X) = r(\pi_1 - 1)$

$E(X^2) = (r^2 + r)\pi_2 - (2r^2 + r)\pi_1 + r^2$

$E(X^3) = (r^3 + 3r^2 + 2r)\pi_3 - (3r^3 + 6r^2 + 3r)\pi_2 + (3r^3 + 3r^2 + r)\pi_1 - r^3$

$E(X^4) = (r^4 + 6r^3 + 11r^2 + 6r)\pi_4 - (4r^4 + 18r^3 + 26r^2 + 12r)\pi_3 r\pi_2 - (4r^4 + 6r^3 + 4r^2 + r)\pi_1 r^4,$

$V(X) = (r^2 + r)\pi_2 - r\pi_1(1 + r\pi_1)$

Skewness

$$= \{E(X^3) - 3E(X^2)E(X) + 2E(X^3)\}/\sigma_X^3$$

$$= \{(r^3 + 3r^2 + 2r)\pi_3 - (3r^3 + 6r^2 + 3r)\pi_2 + (3r^3 + 3r^2 + r)\pi_1 - r^3$$

$$-r(\pi_1 - 1)[3(r^2 + r)\pi_2 - (2r^2 + 3r + 2r^2\pi_1)\pi_1 + r^2]\}/\sigma_X^3$$

Kurtosis $= \{E(X^4) - 4E(X^3)(X^2)[E(X)]^2 - 3[E(X)]^4\}/\sigma_X^4 = \{(r^4 + 6r^3 + 11r^2 + 6r)\pi_4$

$$-(4r^4 + 18r^3 + 26r^2 + 12r)\pi_3 + (6r^4 + 18r^3 + 19r^2 + 7r)\pi_2$$

$$-(4r^4 + 6r^3 + 4r^2 + r)\pi_1 + r^4 - 4r(\pi_1 - 1)$$

$$[(r^3 + 3r^2 + 2r)\pi_3 - (3r^3 + 6r^2 + 3r)\pi_2 + (3r^3 + 3r^2 + r)\pi_1 - r^3]$$

$$+3r^2(\pi_1 - 1)^2[2(r^2 + r)\pi_2 - (2(r^2 + r) + r^2\pi_1)\pi_1 + r^2]\}/\sigma_x^4,$$

where $\pi_c = \frac{1}{1+\theta}\left(\frac{\theta^{\alpha+1}}{(\theta-c)^\alpha} + \frac{\theta\beta}{(\theta-c)^\beta}\right)$ and $\sigma_X = \sqrt{V(X)}$

Index of dispersion

The index of dispersion, which is also called variance-to-mean ratio, is a very useful tool to indicate a set of observations are clustered or dispersed compared to a standard statistical model. The index of dispersion of the LB3 distribution comes out to be

$$\begin{aligned}
\sigma &= \frac{\sigma^2}{\mu} \\
& \frac{E(X^2) - [E(X)]^2}{E(X)} \\
& \frac{E(X^2)}{E(X)} - E(X) \\
&= \frac{(r^2+r)\pi_2 - (2r^2+r)\pi_1 + r^2 - r(\pi_1 - 1)}{r(\pi_1 - 1)} \\
&= \frac{(r^2+r)\pi_2 - r\pi_1(1+r\pi_1)}{r(\pi_1 - 1)} \qquad 4.3.6
\end{aligned}$$

Given that the value or r is greater than one, the dispersion index will be greater than one,implying that the distribution is overdispersed

4.3.3 Order Statistic density function

Let $X_1, X_2, \dots, X_n$ be n independent and identically distributed (iid) random variables defined on Ω with the cumulative density function (cdf) $F_X(x)$ and the pmf $f_x(x)$. Let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ denote these random variables rearranged in non-descending order of magnitude. Thus, $X_{(k)}$ is the kth smallest number in the sample, $k = 1, 2, \dots, n$. Because order statistics are random variables, it is possible to compute probability values associated with values in their support. The kth order statistics density function of $X_{(k)}$ is (e.g., Casella and Berger, 2002) given by

$$f_{x_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} f_x(x)\,[F_x(x)]^{k-1}\,[1-F_x(x)]^{n-k}; x \in \Omega.$$

If $X_1, X_2, \ldots, X_n$ be n iid variables with the pmf $f_x(x)$ as in (5) and cdf $F_X(x)$ as follows

$$F_X(x) = \sum_{s \le x} f_x(x) = \sum_{s \le x} \binom{r+s-1}{s} \sum_{j=0}^{s} \binom{s}{j} (-1)^j \frac{1}{\theta+1} \left(\frac{\theta^{\alpha+1}}{(\theta+r+j)^{\alpha}} + \frac{\theta^{\beta}}{(\theta+r+j)^{\beta}} \right)$$

Definition Let X ~ NB-L3(r, α, β, θ). Then, the order statistic density function of X is

$$f_{x_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!(} \binom{r+x-1}{x}$$

$$\sum_{j=0}^{x} \binom{x}{j} (-1)^j \frac{1}{\theta+1} \left(\frac{\theta^{\alpha+1}}{(\theta+r+j)^{\alpha}} + \frac{\theta^{\beta}}{(\theta+r+j)^{\beta}} \right)$$

$$\times \left\{ \sum_{s \le x} \binom{r+s-1}{s} \sum_{j=0}^{s} \binom{s}{j} (-1)^j \frac{1}{\theta+1} \left(\frac{\theta^{\alpha+1}}{(\theta+r+j)^{\alpha}} + \frac{\theta^{\beta}}{(\theta+r+j)^{\beta}} \right) \right\}^{k-1}$$

$$\times \left\{ 1 - \sum_{s \le x} \binom{r+s-1}{s} \sum_{j=0}^{s} \binom{s}{j} (-1)^j \frac{1}{\theta+1} \left(\frac{\theta^{\alpha+1}}{(\theta+r+j)^{\alpha}} + \frac{\theta^{\beta}}{(\theta+r+j)^{\beta}} \right) \right\}^{n-k},$$

where $k = 1, 2, \ldots, n, s, x = 0, 1, 2, \ldots$ for $s \leq x$ and $r, \alpha, \beta, \theta > 0$.

4.3.4 Special Cases of Three Parameter Lindley Negative Binomial Distribution

Corollary 4.3.1 Let $X \sim NBL_3(r, \alpha, \beta, \theta)$. If $\alpha = 1$ and $\beta = 1$,The negative binomial-Lindley (NB-L) distribution is obtained using positive r and θ parameters. Its pmf is

$$f(x; r, \theta) = \frac{\theta^2}{\theta+1}\begin{pmatrix} r+x-1 \\ x \end{pmatrix}\sum_{j=0}^{x}\begin{pmatrix} x \\ j \end{pmatrix}(-1)^j\frac{\alpha(\theta+r+j)+1}{(\theta+r+j)^2}; x = 0, 1, 2, \ldots$$

where the NB-L distribution was proposed by Zamani and Ismail (2010).

Proof. If $X \sim NB - L_3(r, \alpha, \beta, \theta)$. and substituting $\alpha = 1$ and $\beta = 1$ in equation above, then pmf of X is given by

$$f(x; r, \theta) = \frac{\theta^2}{1+\theta}\begin{pmatrix} r+x-1 \\ x \end{pmatrix}\sum_{j=0}^{x}\begin{pmatrix} x \\ j \end{pmatrix}(-1)^j\frac{(\theta+r+j)+1}{(\theta+r+j)^2}; x = 0, 1, 2, \ldots$$

$$= \begin{pmatrix} r+x-1 \\ x \end{pmatrix}\sum_{j=0}^{x}\begin{pmatrix} x \\ j \end{pmatrix}(-1)^j\frac{(\theta+r+j)+1}{(\theta+r+j)^2}; x = 0, 1, 2, \ldots$$

which is the pmf of NB-L distribution.

Corollary 4.3.2. If $\alpha = 1$. we get

$$f(x; r, \theta, \beta) = \frac{\theta^2}{\theta+\beta}\begin{pmatrix} r+x-1 \\ x \end{pmatrix}\sum_{j=0}^{x}\begin{pmatrix} x \\ j \end{pmatrix}(-1)^j\frac{(\theta+r+j)+\beta}{(\theta+r+j)^2}; x = 0, 1, 2, \ldots$$

which Adil Rashid et al. (2020) call the Negative Binomial Two Parameter Lindley Distribution. **Corollary 4.3.3** If $\alpha = 1$ and $\beta = 0$. we get

$$f(x; r, \theta) = \theta\begin{pmatrix} r+x-1 \\ x \end{pmatrix}\sum_{j=0}^{x}\begin{pmatrix} x \\ j \end{pmatrix}(-1)^j\frac{(\theta+r+j)}{(\theta+r+j)^2}; x = 0, 1, 2, \ldots$$

$$f(x; r, \theta) = \theta\begin{pmatrix} r+x-1 \\ x \end{pmatrix}\sum_{j=0}^{x}\begin{pmatrix} x \\ j \end{pmatrix}(-1)^j\frac{1}{(\theta+r+j)}; x = 0, 1, 2, \ldots$$ This is

the exponential distribution.This probability distribution was obtained by Willmot et al (1981)

Corollary 4.3.4. If $\alpha = 1$ and $r = 1$. we get

$$f(x;\beta,\theta) = \frac{\theta^2}{\theta+\beta}\sum_{j=0}^{x}\binom{x}{j}(-1)^j\frac{(\theta+1+j+\beta)}{(\theta+1+j)^2}; x = 0,1,2,\dots,\theta,\beta > 0$$

This is the geometric distribution with 2 parameter Lindley Distribution.

Corollary 4.3.5. If $\alpha = 1, r = 1$ and $\beta = 1$. we get

$$f(x;\theta) = \frac{\theta^2}{\theta+1}\sum_{j=0}^{x}\binom{x}{j}(-1)^j\frac{(\theta+2+j)}{(\theta+1+j)^2}; x = 0,1,2,\dots,\theta,\beta > 0$$ which is a compound of geometric distribution with one Parameter Lindley Distribution.

Corollary 4.3.6 If $\alpha = 1, r = 1$ and $\beta = 0$. we get

$$f(x;\theta) = \theta\sum_{j=0}^{x}\binom{x}{j}(-1)^j\frac{(\theta+1+j)}{(\theta+1+j)^2}; x = 0,1,2,\dots,\theta,\beta > 0$$

$$f(x;\theta) = \theta\sum_{j=0}^{x}\binom{x}{j}(-1)^j\frac{1}{(\theta+1+j)}; x = 0,1,2,\dots,\theta,\beta > 0$$

This is a compound of geometric distribution with exponential distribution.

4.4 Estimation

The Maximum Likelihood Estimate

Here, we calculate the parameter point and interval estimates using the maximum likelihood procedure. In the NBL_3 distribution, the log-likelihood function, l, is given as

$$l = \ln L(r,\alpha,\beta,\theta) = \sum_{i=1}^{n} \ln\left[\Pr\left(X_i = x_i \mid r,\alpha,\beta,\theta\right)\right] = \sum_{x=0}^{\infty} n_x \ln[\Pr(X = x \mid r,\alpha,\beta,\theta)]$$

$$= \sum_{x=0}^{\infty} n_x \left\{ \ln\left(\begin{matrix} x+r-1 \\ x \end{matrix}\right) + 2\ln\theta - \ln(\alpha\theta+\beta) + \ln\left[\sum_{j=0}^{x} \left(\begin{matrix} x \\ j \end{matrix}\right) (-1)^j \frac{\alpha(\theta+r+j)+\beta}{(\theta+r+j)^2}\right]\right\} \quad 4.4.1$$

where the frequency for data with a x value is denoted by n_x. The first partial derivative is provided as the following with regard to all four parameters:

$$\frac{\partial l}{\partial \theta} = n\left(\frac{2}{\theta} - \frac{\alpha}{\alpha\theta+\beta}\right) + \sum_{x=0}^{\infty} n_x \left[\frac{\sum_{j=0}^{x} \left(\begin{matrix} x \\ j \end{matrix}\right) (-1)^{j+1} \frac{\alpha(\theta+r+j)+2\beta}{(\theta+r+j)^3}}{\sum_{j=0}^{x} \left(\begin{matrix} x \\ j \end{matrix}\right) (-1)^j \frac{\alpha(\theta+r+j)+\beta}{(\theta+r+j)^2}}\right] \quad 4.4.2$$

$$\frac{\partial l}{\partial \alpha} = -\frac{n\theta}{\alpha\theta+\beta} + \sum_{x=0}^{\infty} n_x \left[\frac{\sum_{j=0}^{x} \left(\begin{matrix} x \\ j \end{matrix}\right) (-1)^{j+1} \frac{1}{\theta+r+j}}{\sum_{j=0}^{x} \left(\begin{matrix} x \\ j \end{matrix}\right) (-1)^j \frac{\alpha(\theta+r+j)+\beta}{(\theta+r+j)^2}}\right] \quad 4,4,3$$

$$\frac{\partial l}{\partial \beta} = -\frac{n}{\alpha\theta+\beta} + \sum_{x=0}^{\infty} n_x \left[\frac{\sum_{j=0}^{x} \left(\begin{matrix} x \\ j \end{matrix}\right) (-1)^{j+1} \frac{1}{(\theta+r+j)^2}}{\sum_{j=0}^{x} \left(\begin{matrix} x \\ j \end{matrix}\right) (-1)^j \frac{\alpha(\theta+r+j)+\beta}{(\theta+r+j)^2}}\right] \quad 4.4.4$$

$$\frac{\partial l}{\partial r} = \frac{\partial}{\partial r}\left\{\sum_{x=0}^{\infty} n_x \left[\ln\left(\begin{matrix} x+r-1 \\ x \end{matrix}\right)\right]\right\} + \frac{\partial}{\partial r}\left\{\sum_{x=0}^{\infty} n_x \ln\left[\sum_{j=0}^{x} \left(\begin{matrix} x \\ j \end{matrix}\right) (-1)^j \frac{\alpha(\theta+r+j)+\beta}{(\theta+r+j)^2}\right]\right\} \quad 4.4.5$$

Let the representation of the expression in the first term of the partial derivative with regard to r, *partiall/* in (10) be

$$A(r) = \sum_{x=0}^{\infty} n_x \left[\ln \begin{pmatrix} x+r-1 \\ x \end{pmatrix} \right] \tag{4.4.6}$$

The phrase $A(r)$'s derivative can be expressed as (Klugman, Panjer, and Willmot 2008).

$$\frac{\partial}{\partial r}[A(r)] = \frac{\partial}{\partial r}\left\{ \sum_{x=0}^{\infty} n_x \left[\ln \begin{pmatrix} x+r-1 \\ x \end{pmatrix} \right] \right\} = \sum_{x=0}^{\infty} n_x \frac{\partial}{\partial r}\left[\ln \frac{(x+r-1)(x+r-2)\ldots r}{x!} \right] \tag{4.4.7}$$

A simplified version of the formula above yields

$$\frac{\partial}{\partial r}[A(r)] = \sum_{x=0}^{\infty} n_x \frac{\partial}{\partial r} \ln \prod_{m=0}^{x-1}(r+m) = \sum_{x=0}^{\infty} n_x \frac{\partial}{\partial r} \sum_{m=0}^{x-1} \ln(r+m) = \sum_{x=0}^{\infty} n_x \sum_{m=0}^{x-1} \frac{1}{r+m}$$

Maximum likelihood estimates are obtained by

$$\frac{\partial \ell}{\partial \mathrm{r}} = 0, \frac{\partial \ell}{\partial \alpha} = 0, \frac{\partial \ell}{\partial \beta} = 0, \frac{\partial \ell}{\partial \theta} = 0. \tag{4.4.8}$$

But solving these equations is complicated and difficult. So these equations are solved numerically using Newton Raphson method.

4.5 Simulation

This part simulates the generation of random variables from the NBL3 $(r, theta, alpha, beta)$. The created sample is then used to determine the MLE of the parameters. To determine the consistency of the estimates, the Bias and mean squared error (MSE) of the MLE of the parameters are determined. The algorithm for the requested simulation research is shown below.

1. Making a randon sample with NBL3 r, alpha, beta, and theta

step(i) The U(0,1) distribution is used to create a random variable.

step (ii) If the $x_i, i >= 0$, are arranged in the following order: $x_0 x_1 x_2..$ and if we lex

$$F_X(x) = \sum_{s \leq x} f_x(x) = \sum_{s \leq x} \binom{r+s-1}{s} \sum_{j=0}^{s} \binom{s}{j} (-1)^j \frac{1}{\theta+1} \left(\frac{\theta^{\alpha+1}}{(\theta+r+j)^{\alpha}} + \frac{\theta^{\beta}}{(\theta+r+j)^{\beta}} \right)$$

define the distribution function of X.

If $F(x_i - 1), = uF(x_i), i = 0, 1, 2, ...,$ then $X = x_l$.

As many times as the desired sample size is, steps (i) and (ii) are performed. The discrete inverse transform method for generating X is what the aforementioned technique is called.

2. Getting the parameters' MLE

For the produced sample collected in the preceding phase, the maximum likelihood equations (MLE) of $r, \theta, \alpha, and \beta$ are solved.

3.Bias and MSE of the MLEs are calculated. Assume that the MLE is $theta*$ and that the parameter $theta$'s true value is $theta_0$. When predicting $theta_0$, the bias of $theta*$ is thus given by

$$Bias(\theta^*) = E(\theta^* - \theta_0)$$

Regarding the NBL3 mass function, the expectation is $(r, \theta, \beta, \alpha)$. Similar to this,

$$MSE(\theta*) = E[(\theta * -\theta_0)2$$

is used to calculate the MSE of θ_0.

Average values of bias and variance for $r^*\theta^*, \alpha^*$ and β^* for NBL3

Sample Size	θ=3		$\alpha = 4$		$\beta = 2$		r=5	
	Bias (θ)	Var(θ)	Bias (α)	Var (α)	Bias (β)	Var (β)	Bias (r)	Var(r)
50	0.64619	0.51673	1.87777	3.52602	1.20045	1.46335	1.27721	1.56335
100	0.58956	0.37721	1.60015	2.65715	1.19996	1.44537	1.15996	1.44557
200	0.23850	0.06977	1.03593	1.23732	0.69749	0.50399	0.65745	0.50399
300	0.21259	0.05419	0.90122	0.82120	0.49964	0.25864	0.45964	0.25564

The Monte Carlo approximation method, using T=1000 repetitions, approximates the Bias and MSE of the MLE of θ. The Bias and MSE of the MLE of r, α, and β are determined in a similar way. If the Bias lowers (gets closer to zero) with an increase in sample size and the MSE also decreases, the MLE is considered to be consistent. The values of the Bias and MSE of the MLE of $r, \beta, \alpha, and \theta$ for the various sample sizes are displayed in Table 1. Table 1 shows that as the sample size is increased, the bias and MSE of r, beta, alpha, and theta decrease and eventually reach zero.

Average values for VAR and Bias for λ^* for Poisson

Sample Size	λ=3	
	Bias (λ)	Var (λ)
50	0.66619	0.71673
100	0.59779	0.39721
200	0.26850	0.09977
300	0.23732	0.05419

Average values of bias and Var for β^*,$\alpha^*$and r^* for NBL2

Sample Size	α=3		$\beta=4$		r= 2	
	Bias (α)	Var (α)	Bias (β)	Var (β)	Bias (r)	Var (r)
50	0.69693	0.61673	1.87777	3.52693	1.20045	1.66335
100	0.58956	0.37793	1.66015	2.65715	1.19996	1.44537
200	0.41850	0.41977	1.03593	1.41732	0.69749	0.41399
300	0.23732	0.05419	0.90122	0.92120	0.49964	0.23732

Although MSE and bias generated by poisson and negative binomial weighted lindley distributions decrease with increase sample size, their values are slightly higher than those frpm NBL3 .

The MLE estimations are therefore consistent and accurate in determining the true value of the parameters, it can be said. R software, version 3.4.4, is used to do calculations related to the study with the aid of self-programmed codes. The parameters from NBL3 ($r, \beta, \alpha, and\theta$) are estimated using the maximum likelihood method using the R software's maxLik package (Henningsen, Arne and Toomet, Ott, 2011).

4.6 Application to real data

This is to show the importance of NBL3 for count data analysis.We consider a real data set. Distributions such the Poisson, NB, L3, NBL, NB-NWL, and NBL3 distributions are utilized to fit this data set. To determine the significance of the K-S statistics, we utilized the ks.test function in the dgof package of R (Marsaglia et al., 2003; R Core Team, 2018). The lesser statistic result of the KS test corresponds to the best distribution.

The dataset consists of the total claims made under the third responsibility auto insurance policy taken into account in Wang's (2011) study.

We contrast the NBL3 distribution's fit with that of the following distributions:

The Negative Binomial

$$f(x;r,p) = \binom{x+r-1}{x} p^r(1-p)^x dx, x = 0, 1...., r > 0, 0 < p < 1$$

The poisson distribution

$$\mathrm{f}(\mathrm{x};\lambda) = \frac{\mathrm{e}^{-\lambda}\lambda^{\mathrm{x}}}{\mathrm{x}!}; \mathrm{x} = 0, 1, 2, ... \text{ and } \lambda > 0.$$

The Three parameter Lindley (L3) distribution

$$f(x;\alpha,\beta,\theta) = \frac{\theta^2}{\theta\alpha+\beta}(\alpha+\beta\lambda)\mathrm{e}^{-\theta x}$$

$x, \beta, \alpha, \theta > 0, \alpha\theta + \beta > 0$ as introduced by R. Shanker et al(2017)

The Negative Binomial - Two parameter Lindley distribution

$$f(x;r,\theta,\alpha)=\begin{pmatrix} r+x-1 \\ x \end{pmatrix}\sum_{j=0}^{x}\begin{pmatrix} x \\ j \end{pmatrix}(-1)^{j}\frac{(\theta-\alpha)(\theta-\alpha+r+j+1)}{(\theta-\alpha+1)(\theta-\alpha+r+j)^{2}},\quad x=0,1,2,...$$

where $\theta > 0$ and $\theta > \alpha$

The MLE and nlm functions in R are used to retrieve the estimated parameters for each distribution. We take into account the $\hat{\ell}$, which is the maximal value of the log-likelihood function under the studied distributions, to compare the parameter estimate of each distribution.

The dataset is overdispersed, as evidenced by the ratio of variance to mean, which indicates the dispersion of the data, which is 1.78.

Table 4: Observed values, expected values, and statistics of each distribution for Insurance claims data

Number of claims	Observed values	Observed values of fitting with distributions					
		Poisson	NB	NBL	NBL2	L3	NBL3
0	7,840	7,635.27	7,846.64	7,763.54	7,796.23	7,811.55	7,837.40
1	1,317	1,637.00	1,288.58	1,356.39	1,361.03	1,346.35	1,326.20
2	239	175.49	256.64	241.98	205.80	244.61	226.34
3	42	12.54	54.10	44.29	30.25	46.67	48.74
4	14	0.67	11.72	8.31	4.45	9.32	14.03
5	4	0.03	2.58	1.60	0.66	1.94	4.97
6	4	0.00	0.57	0.31	0.10	0.42	1.95
7	1	0.00	0.13	0.06	0.02	0.10	0.79
Estimated value of parameters		$\hat{\lambda}$=0.2144	$\hat{p}$=0.7659 $\hat{r}$=0.7015	$\hat{r}$=30.4775 $\hat{\theta}$=143.966	$\hat{r}$=21.7015 $\hat{\alpha}$=1.4737 $\hat{\theta}$=162.372	$\hat{\beta}$=21.7876 $\hat{\alpha}$=1.4737 $\hat{\theta}$=162.372	$\hat{r}$=4.8261 $\hat{\alpha}$=1.8355 $\hat{\beta}$=12.1718 $\hat{\theta}$=48.1251
$-\hat{\ell}$		5,590.78	5,396.31	5,396.29	5,431.77	5,348.97	5,341.93
KS test		0.0216	0.0022	0.0081	0.0031	0.0066	0.0007
(p-value)		(0.0003)	(0.9999)	(0.5669)	(0.8039)	(0.9999)	(0.9999)
AIC		857.49	857.41	857.32	857.31	857.01	855.06
BIC		865.78	865.71	864.67	864.15	864.14	859.20

Futhermore, to check the appropriate distribution,we use Akaike Information Criterion (AIC) and Baysian Information Criteria(BIC).

(i) AIC:The Akaike Information Criterion (AIC) is used to compare various semi-parametric and parametric models. Akaike makes the AIC suggestion (Akaike, 1974). It evaluates how well a predicted statistical distribution fits the data. This study's distribution's AIC is calculated using the following formula.

AIC= $2k - 2log(L)$

where L is the value of the likelihood function calculated using the parameter estimations and k is the number of estimated parameters.

(ii) BIC: The Baysian Information Criteria (BIC) is given by Schwarz (Schwarz, 1978). It is computed as follows

BIC=$klog(n) - 2log(L)$

where L is the value of the likelihood function calculated at the parameter estimates, k is the number of estimated parameters, and n is the number of observations.

We offer the values of the Kolmogorov-Smirnov (KS) statistic for further discussion.

The distribution with the lowest $-2log(L)$, AIC, BIC, and KS is regarded as the best model for fitting a particular data set when comparing lifespan distributions. The likelihood of choosing a distribution with a small number of parameters as the best model will, however, improve in terms of $-2log(L)$, AIC, and BIC. We use goodness of fit test statistic such as Komolgorov- Smirnov test statistics to validate the superiority of NBL3 distribution for the number of claims of the third liability vehicle insurance dataset. Consequently, Table 3 show that the NBL3 has the least value of $-2log(L)$, AIC, BIC, AD and KS , which indicates that the NBL3 demonstrates superiority over the poisson, NB,NBL, L3 and NBWL in modeling the the number of claims of the third liability vehicle insurance data.

Based on the results from the table, the NBL3 distribution is the best model in describing the number of claims of the third liability vehicle insurance as the model gives the smallest AIC, BIC and KSvalues, and thus is selected as the best model with fitted function given as

$$\Pr(X = x \mid r, \alpha, \beta, \theta) = \binom{x+r-1}{x} \frac{\theta^2}{\alpha\theta + \beta} \sum_{j=0}^{x} \binom{x}{j} (-1)^j \left[\frac{\alpha(\theta + r + j) + \beta}{(\theta + r + j)^2}\right]; \quad x = 0, 1, \ldots$$

where $\hat{r} = 4.8261, \hat{\beta} = 12.1718, \hat{\theta} = 48.1251$ and $\hat{\alpha} = 1.8355$

Chapter 5

Summary, Conclusion and Reccomendations

5.1 Introduction

in this Chapter, a summary of what was discussed is given. This will be followed by the conclusion on the findings and lastly reccomendation for future research work will be given.

5.2 Summary

In chapter one, a background of studying mixture distributions was discussed. A brief description of the research problem is given, followed by the objectives of the study.In the second chapter, we looked at various work that has been done in the literature on mixtures of Negative Binomial and Lindley distributions.This chapter provided related literature useful in undertaking our research. In chapter 3, a review of distributions used to fit count data was discussed and finally in the 4th chapter, the NBL3 distribution was constructed .Some of the properties of the new compound distribution were studies. A simulation study was conducted to determine the consistency of the parameter estimates of the new distribution and finally data analysis was done to compare the perfomance of the new distribution over all other related distributions.

5.3 Conclusion

The proposed distribution which is named NBL3 distribution, is a generalization for two and three-parameter negative binomial-Lindley distributions. The proposed distribution is quite general where its special cases resulted in several types of mixed negative binomial distributions such as negative binomial-generalized exponential distributions. Some statistical properties for the proposed distribution has been studied to understand the NBL3 distribution. The properties studied include the kth factorial moment and the dispersion index. Based on these statistical properties, one can easily obtain the measures of skewness, kurtosis and higher order moments for NBL3: From the dispersion index, it can be concluded that the data generated from the NBL3 distribution can have the properties of overdispersion. The derivation of maximum likelihood estimators of the parameters of NBL3 distribution is also presented. The adequacy of the model for NBL3 is significantly improved compared to those for Poisson , negative binomial,three parameter lindley distribution and Negative binomial new weighted lindley distribution suggesting that the NBL3 can be considered in fitting overdispersed count data.

5.4 Reccomendation

A higher parameter mixture of Negative Binomial and Lindley distribution can be explored.

References

1. Casella, G. & R. Berger. (2002). Statistical Inference, 2nd ed. Duxbury Press, Pacific Grove, CA.

2. Elbatal, I., Merovci, F. & Elgarhy, M. (2013). A new generalized Lindley distribution. Mathematical Theory and Modeling, 3(13), 30-47.

3. Flynn, M. & Francis, L. A. (2009). More flexible GLMs zero-inflated models and hybrid models. Casualty Actuarial Soc, 148-224.

4. Gençtürk, Y. & Yiğiter, A. (2016). Modelling claim number using a new mixture model: negative binomial gamma distribution. Journal of Statistical Computation and Simulation, 86(10), 1829-1839.

5. Ghitany, M. E., Atieh, B. & Nadarajah, S. (2008). Lindley distribution and its application. Mathematics and Computers in Simulation, 78(4), 493-506.

6. Gómez-Déniz, E., Sarabia, J. M. & Calderín-Ojeda, E. (2008). Univariate and multivariate versions of the negative binomial-inverse Gaussian distributions with applications. Insurance: Mathematics and Economics, 42(1), 39-49.

7. Greenwood, M. & Yule, G. U. (1920). An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. Journal of the Royal Statistical Society, 83(2), 255-279.

8. Gupta, R. D. & Kundu, D. (1999). Theory & methods: Generalized exponential distributions. Australian & New Zealand Journal of Statistics, 41(2),

173-188.

9. Hamid, H. (2014). Integrated Smoothed Location Model and Data Reduction Approaches for Multi Variables Classification (Unpublished doctoral dissertation). Universiti Utara Malaysia, Kedah, Malaysia.

10. Haight, F. (1967). Handbook of the Poisson distribution. John Wiley and Sons, New York.

11. Jambunathan M. B. (1954). Some properties of beta and gamma distributions. The annals of mathematical statistics. 25(2), 401-405.

12. Johnson, N. L., Kemp, A. W. & Kotz, S. (2005). Univariate discrete distributions, 444, John Wiley & Sons.

13. Klugman, S.A., Panjer H.H. & Willmot G.E. (2008). Loss Models: From Data to Decision. 3rd Edn., John Wiley and Sons, USA : 101-159.

14. Kongrod, S., Bodhisuwan, W., & Payakkapong, P. (2014). The negative binomial-Erlang distribution with applications. International Journal of Pure and Applied Mathematics, 92(3), 389-401.

15. Lord, D. & Geedipally, S. R. (2011). The negative binomial-Lindley distribution as a tool for analyzing crash data characterized by a large amount of zeros. Accident Analysis & Prevention, 43(5), 1738-1742.

16. Lindley, D. V. (1958). Fiducial distributions and Bayes' theorem. Journal of the Royal Statistical Society. Series B (Methodological), 20(1), 102-107.

17. Pudprommarat, C., Bodhisuwan, W., & Zeephongsekul, P. (2012). A new mixed negative binomial distribution. Journal of Applied Sciences(Faisalabad), 12(17), 1853-1858.

18. Marsaglia, G., Tsang, W. W. & Wang, J. (2003), Evaluating Kolmogorov's distribution. Journal of Statistical Software, 8(18), 1-4.

19. Panjer, H. H. & Willmot, G. E. (1981). Finite sum evaluation of the negative binomial-exponential model. ASTIN Bulletin: The Journal of the IAA, 12(2), 133-137.

20. R Core Team. (2018). R: A language and environment for statistical computing.

21. Rainer, W. (2000). Econometric Analysis of Count Data, 3rd edn, Springer Verlag, Berlin, Germany.

22. Sankaran, M. (1970). The discrete Poisson-Lindley distribution, Biometrics, 26(1), 145-149.

23. Wang, Z. (2011). One mixed negative binomial distribution with application. Journal of Statistical Planning and Inference, 141(3), 1153-1160.

24. Zamani, H. & Ismail, N. (2010). Negative binomial-Lindley distribution and its application. Journal of Mathematics and Statistics, 6(1), 4-9. Arnold, T. B., & Emerson, J. W. (2011). Nonparametric Goodness-of-Fit Tests for Discrete Null Distributions. *R* Journal, 3(2), 34-39.