ESTIMATION OF PARAMETERS OF THE MCDONALD GENERALIZED BETA - BINOMIAL DISTRIBUTION USING ESTIMATING FUNCTIONS

NTHIWA JANIFFER MWENDE

A Thesis Submitted to the Graduate School in Partial Fulfillment for the Requirements of the award of Master of Science Degree in Statistics of Egerton University

EGERTON UNIVERSITY

APRIL, 2015

DECLARATION AND RECOMMENDATION

DECLARATION

This thesis is my original work and has not been submitted or presented in part or whole for examination in any institution.

Sign: ______________________________ Date: ______________________________

Nthiwa J. M.

SM12/3330/12

RECOMMENDATIONS

This research thesis has been submitted for examination with our approval as university supervisors.

Sign: ______________________________ Date: ______________________________

Prof. Ali Salim Islam

Mathematics department

Egerton University

Sign: ______________________________ Date: ______________________________

Dr. Orawo Luke

Mathematics department

Egerton University

COPY RIGHT

DEDICATION

To

My dad Timothy Nthiwa and mum Dorcus

ACKNOWLEDGEMENT

I would like to thank my God Almighty, for graciously letting me pursue my ambition when all was bleak and for granting me good health, sane mind, strength and sufficient grace to see me through this study. I offer deepest appreciation and love for a lifetime. Thank you and Glory be to You Lord forever and ever! Secondly, I would like to appreciate my supervisors. I would not have completed this thesis if Professor Ali Islam and Dr. Orawo Luke, my research supervisors, had not offered such infinite patience, critical commentaries and productive arguments. It was my good fortune to have benefited greatly from their excellent supervision. I wouldn't forget to thank the Chairman, Department of Mathematics, Dr. Gichuki for providing the necessary research materials I needed for this work. To my lecturers in the department, thank you all for your support and the humble and friendly environment you provided. To my colleagues; we've been through challenges that we've had to break up and make up. We've learnt a lot and maintained the friendship and have encouraged one another move on even when all seemed difficult to bear. Thank you for your support. I must thank all my wonderful friends in Egerton University Mathematics department for they were so generous with their times and recollections. To my beloved family; mum and dad, you've been my encouragement all the way. My brothers and sisters, Peninah, Muthoka, Kioko, Eunice, Alex, Anna and Rebecca, thank you for the financial and moral support you have accorded me. I am also greatly indebted to Mr. Stephen M. for his moral support when all seemed bleak. Thank you and God bless you. To my church members and friends; thanks for every contribution you made in one way or another.

Thank you and God bless you all!

ABSTRACT

Recently, there has been a considerable attention on modeling overdispersed binomial data that occur in toxicology, biology, clinical medicine, epidemiology and other related fields using a class of Binomial mixture distribution. Specifically, Beta-Binomial (BB) and Kumaraswamy-Binomial distribution (KB) in this class have been extensively used to model the overdispersed Binomial outcomes. A new three parameter binomial mixture distribution namely, McDonald Generalized Beta-Binomial (McGBB) distribution which is superior to KB and BB has been developed. The study on Point estimation for McGBB distribution model has been done using Maximum Likelihood estimates (MLEs) and shown to give better fit than the KB and BB distribution on both real life data set and on the extended simulation study in handling over-dispersed binomial data. However, MLEs are quite intensive in computation and not robust to variance misspecification. Estimating functions have for sometimes now been a key concept and subject of inquiry in research as a more general method of estimation which are robust to variance misspecification. This thesis considered estimation of parameters of the McGBB model using Estimating Functions based on Quasi-likelihood (QL) and Quadratic estimating equations (QEEs), which have not been developed and both of which are robust to variance structure misspecification. By varying the coefficients of the QEE's, four sets of estimating equations, denoted as GL, M1, M2 and M3, were obtained. This study then compared the small sample relative efficiency of the four sets of estimates obtained by the QEE's and the QL estimates with the MLEs based on a real life data sets arising from alcohol consumption practices and a simulated data. These comparisons show that estimates, using optimal QEEs and estimates of QL are highly efficient and are the best among all estimates investigated. Thus, this thesis has provided an estimation procedure based on the QL and QEEs for estimating the parameters of McGBB distribution which is superior to the Maximum likelihood method.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

BB	Beta-Binomial
EF	Estimating functions
EQL	Extended Quasi-likelihood
GL	Gaussian likelihood
KB	Kumaraswamy-Binomial
McGBB	McDonald Generalized Beta-Binomial
MLE	Maximum Likelihood Estimation
MM	Method of moments
QL	Quasi-likelihood
QEE	Quadratic Estimating Equations

CHAPTER ONE
INTRODUCTION

1.1 Background of the Study

Estimating functions have for sometimes been a key concept and subject of inquiry in research and it is known to provide a more flexible approach to estimation than ML method. The basis of this method is a set of simultaneous equations involving both the data and the unknown model parameters. Estimating functions provide more robust estimates, such as moment estimates, quasi-likelihood estimates (Breslow, 1990; Moore and Tsiatis, 1991), extended quasi-likelihood estimates (Nelder and Pregibon, 1987), the Gaussian likelihood estimates (Whittle, 1961; Crowder, 1985), estimates based on the pseudo-likelihood estimating equations (Davidian and Carrol, 1987) and estimates based on quadratic estimating functions (Crowder, 1987; Godambe and Thompson, 1989). The estimates were studied by Paul and Islam (1998) and the small and large sample efficiency and bias properties of these estimates were compared with the maximum likelihood estimate and showed that they performed better than the MLEs.

To obtain an estimator, the estimating function is equated to zero and then solve the resulting equation with respect to the parameter in order to obtain parameter estimate. Estimating equations are not quite intensive in computation and are robust to variance misspecification unlike MLEs. Moreover, the MLE estimators are based on the assumption that the distribution is known, however an estimating equation is free of such assumptions.

In real world applications, binomial outcome data are widely encountered like in toxicology, biology, clinical medicine, epidemiology and other similar fields. However, binomial distribution fails to model these binomial outcomes which results to over-dispersion because variance of the observed data is greater than the nominal variance. Use of continuous distribution defined on the standard unit interval is one way of handling over-dispersion to model the success probability p of the Binomial distribution. A class of Binomial mixture distribution which includes BB and KB has recently been used to model these binomial outcomes. A new three parameter binomial mixture distribution namely, McGBB distribution has been developed which is superior to KB and BB. It was developed by mixing the McDonald's generated beta distribution of the first kind with the success probability of binomial distribution. The parameters of McGBB Distribution were then estimated by maximum likelihood technique. The study has shown that additional parameter in McDonald

Generalized Beta-Binomial distribution allows accommodating wide range of shapes (Chandrabose et *al*. 2013).

However studies have shown that maximum likelihood estimates are quite intensive in computation and not robust to variance misspecification (Paul and Islam, 1994). In addition, study by Paul (2009) shows that MLE procedure may produce inefficient or biased estimates when the parametric model does not fit the data well hence MLE method is a biased estimate. A more serious objection to MLE approach was raised by Neyman and Scott (1948), who showed that when the number of nuisance parameters increases with the sample size, the MLE of a parameter of interest could be inefficient or even inconsistent. Thus this study derived the estimating equations for McGBB distribution based on QL, GL, M1, M2, and M3 which has not been done and are both robust to variance misspecification.

1.2 Statement of the Problem

Over-dispersion phenomenon occurs in binomial data when the variance of the observed binomial outcome exceeds variance of the nominal Binomial distribution. However, Binomial distribution often fails to model these binomial outcomes. A class of binomial mixture distribution which includes BB and KB has been used to model over-dispersed binomial outcomes but a new three parameter binomial mixture distribution known as McGBB distribution has been developed which has been found to be superior to KB and BB. The point estimation for parameters of McGBB distribution has been done using MLE and shown to give better fit than the KB and BB distribution on both real life data set and on the simulated data in handling over-dispersion in binomial outcome data. However, MLEs are quite intensive in computation and not robust to variance misspecification unlike Estimating functions. Thus, in this study the parameters of McGBB distribution were estimated by estimating functions based on Quasi-likelihood technique and Quadratic estimating equations which had not been done and are more efficient than the MLEs.

1.3 Objectives

1.3.1 General Objective

To estimate the parameters of McDonald Generalized Beta-Binomial (McGBB) distribution by using the estimating functions approach.

1.3.2 Specific Objectives

(i) To derive the Quasi-likelihood and Quadratic estimating equations for Mcdonald Generalized Beta Binomial Distribution.

(ii) To compare the performance in terms of efficiency of estimators based on the maximum likelihood method and estimating functions using simulation.

(iii) To compare the performance in terms of efficiency of estimators based on estimating functions and maximum likelihood method using real data.

1.4 Assumption

This study assumed that the data is binomial over-dispersed.

1.5 Justification

The proposed study drew its significance from the fact that it derives estimating equations based on QL and QEEs for McGBB distribution which has not been done. Studies have shown that estimating equations often give better estimates than MLEs by providing an estimator that is consistent and efficient, regardless of the true distribution. Also, it has been shown that when the number of nuisance parameters increases with the sample size, the MLE of a parameter of interest could be inefficient or even inconsistent. Thus, this study will provide the researchers with better estimates for the parameters of the McGBB distribution based on the QL, GL, M1, M2 and M3 than the MLEs.

This work can be applied in the agricultural set-up. For example, in Kenya, bee farming can be improved based on the knowledge from this distribution. One may access forage preferences in the different kinds of bees (bees that live in hives, ant-holes and tree barks in forests). Someone may be interested in investigating the behavior of bees among different colour of flowers and modeling the pattern of visitation as a random movement. This will be a test that will be used to advise farmers on the colour of flowers to plant depending on the kinds of bees reared in their farms.

In a medical setup, we may want to model the percentage of patients who have successfully undergone a particular medication procedure. One may want to assess whether the success probabilities are equal among a number of hospitals. Given the existence of some un-predetermined excess variation among the different hospitals, the information obtained would have a lot of effect on policy implications.

CHAPTER TWO
LITERATURE REVIEW

2.1 Introduction

McGBB Distribution is a new parametric model obtained by mixing McDonald Generalized Beta- Binomial Distribution of the first kind and binomial success probability p of binomial distribution. The study on Point estimation for McGBB distribution model has been done using MLEs and shown to give better fit than the KB and BB distribution on both real life data set and on the simulation study in handling over-dispersed binomial outcome data. However, MLEs are not robust to variance misspecification. Estimating functions have for sometimes been a key concept and subject of inquiry in research as a more general method of estimation which are robust to variance misspecification. Thus, this thesis provides the estimates of the parameters of McGBB distribution by estimating functions based on QL, GL, M1, M2 and M3. It also presents a comparison of these five methods with MLE using relative efficiencies which have not been done.

2.2 McDonald Generalized Beta-Binomial distribution

In this section McGBB distribution is defined with some properties of the same distribution given. McGBB Distribution is a binomial mixture distribution.

2.2.1 McDonald Generalized Beta-Binomial Distribution of the first kind

Let p be a random variable following McDonald's Generalized Beta-Binomial Distribution of the first kind (McDonald, 1984; McDonald and Xu, 1995) with three parameters, α, β and γ. The probability density function of p is then given by

$$f_{GB1}(p;\alpha,\beta,\gamma) = \frac{\gamma}{B(\alpha,\beta)} p^{\alpha\gamma-1}(1-p^{\gamma})^{\beta-1};\ 0 \le p \le 1 \text{ and } \alpha,\beta,\gamma > 0 \quad (1)$$

The r^{th} moment of the McDonald Generalized Beta-Binomial Distribution of the first kind is given by

$$E(P^r) = \frac{B(\alpha+\beta, \frac{r}{\gamma})}{B(\alpha, \frac{r}{\gamma})} \quad (2)$$

2.2.2 Definition of McGBB Distribution

In general, a Binomial mixture is obtained through an integration approach as follows. Suppose conditional on p, Y follows a binomial distribution given by Bin (n, p), which is denoted by Y/p ~ Bin (n, p), then unconditional Probability Mass Function of the Y can be obtained by evaluating the integral,

$$P_y\;y \;=\; P\;Y\;P\;f_p\;p\backslash\Theta\;dp. \tag{3}$$

where $y = 0, 1, \ldots, n$ and Θ is the parameter space of the mixing distribution.

A random variable Y is said to have McDonald Generalized Beta-Binomial (McGBB) Distribution with parameter n, α, β and γ if and only if it satisfies the following stochastic representation; Y/p ~ Bin (n, p) and p ~ GB1 (α, β, γ), where α, β and γ are positive real numbers. This distribution was denoted as, Y~ McGBB (n, α, β, γ).

2.2.3 McDonald Generalized Beta-Binomial) distribution Properties

Chandrabose *et al*. (2013) argues that by letting y be a discrete random variable that follows a McDonald Generalized Beta-Binomial (McGBB) distribution as defined above then the following results hold:

(i) The probability mass function of McGBB (α, β, γ) distribution is given by,

$$P_{McGBB}\;y;\alpha,\beta,\gamma \;=\; \binom{n}{y}\frac{\gamma}{B\;\alpha,\beta}\sum_{i=0}^{\infty} -1^{\;i}\binom{\beta-1}{i}\;B\;y+\alpha\gamma+\gamma i, n-y+1 \tag{4}$$

where, $y = 0, 1, \ldots, n$ and $\alpha, \beta, \gamma > 0$

(ii) A rearranged probability mass function of McGBB (n, α, β, γ) distribution is given by,

$$P_{McGBB} = \binom{n}{y}\frac{1}{B\;\alpha,\beta}\sum_{j=0}^{n-y} -1^{\;j}\binom{n-y}{j}\;B\;\frac{y}{\gamma}+\alpha+\frac{j}{\gamma},\beta\;, \tag{5}$$

where, $y = 0, 1, \ldots, n$ and $\alpha, \beta, \gamma > 0$

(iii) The r^{th} moment of McGBB (n, α, β, γ) distribution is given by,

$$E\;Y^r \;= n\,\frac{B(\alpha+\beta,{}^{r}\,\gamma)}{B(\alpha,{}^{r}\,\gamma)}. \tag{6}$$

Then the mean and variance of McGBB (n, α, β, γ) distribution are given respectively by,

$$E\;Y \;= n\pi \tag{7}$$

and

$$\operatorname{var}(Y) = n\pi(1-\pi)\left(1+(n-1)\rho\right), \tag{8}$$

where

$$\pi = \frac{B\left(\alpha+\beta, \frac{1}{\gamma}\right)}{B\left(\alpha, \frac{1}{\gamma}\right)} \tag{9}$$

and ρ is the over-dispersed parameter of McGBB distribution given as,

$$\rho = \frac{\frac{B\left(\alpha+\beta, \frac{2}{\gamma}\right)}{B\left(\alpha, \frac{2}{\gamma}\right)} - \left(\frac{B\left(\alpha+\beta, \frac{1}{\gamma}\right)}{B\left(\alpha, \frac{1}{\gamma}\right)}\right)^2}{\frac{B\left(\alpha+\beta, \frac{1}{\gamma}\right)}{B\left(\alpha, \frac{1}{\gamma}\right)} - \left(\frac{B\left(\alpha+\beta, \frac{1}{\gamma}\right)}{B\left(\alpha, \frac{1}{\gamma}\right)}\right)^2} \tag{10}$$

Chandrabose *et al*. (2013) showed that additional parameter in McDonald Beta-Binomial distribution allows it to accommodate a wide range of shapes. Maximum likelihood technique was used to estimate the three parameters and showed that the new model provides a better fit to BB and KB models (Chandrabose *et al*. 2013). In this study the main objective was to obtain the estimates of the parameters of McDonald Beta-Binomial distribution via estimating functions and compare them with those obtained via maximum likelihood technique.

2.2.4 Maximum likelihood Estimation of parameters of McGBB

The three unknown parameters of McGBB distribution have been estimated using the maximum likelihood estimation technique. Let $Y = (y_1, y_2, \ldots, y_N)$ be a random sample of size N from a McGBB distribution with unknown parameter vector $\Theta = (\alpha, \beta, \gamma)^T$, then the log-likelihood function for Θ can be defined as,

$$l(\Theta) = \sum_{k=0}^{N} log \binom{n}{y} \sum_{k=0}^{N} log \left(\frac{1}{B(\alpha,\beta)}\right) + \sum_{k=0}^{N} log \sum_{j=0}^{n-y_k} (-1)^j \binom{n-y_k}{j} B\left(\frac{y_k}{\gamma} + \alpha + \frac{j}{\gamma}, \beta\right) \tag{11}$$

The score function $U(\Theta)$ is defined as the gradient of $l(\Theta)$, derived by taking the partial derivatives of $l(\Theta)$ with respect to α, β and γ. The components of the score function: $U(\Theta) = U_\alpha(\Theta)$, $U_\beta(\Theta)$, and $U_\gamma(\Theta)$ which are given as follows;

$$U_\alpha(\Theta) = \frac{\partial l(\Theta)}{d\alpha} = N\left(\Psi(\alpha+\beta) - \Psi(\alpha)\right) + \sum_{k=0}^{N}\frac{1}{c_k}\sum_{k=0}^{n-y_k}(-1)^j\binom{n-y_k}{j}D_{\alpha k} \quad (12)$$

$$U_\beta(\Theta) = \frac{\partial l(\Theta)}{d\beta} = N\left(\Psi(\alpha+\beta) - \Psi(\beta)\right) + \sum_{k=0}^{N}\frac{1}{c_k}\sum_{k=0}^{n-y_k}(-1)^j\binom{n-y_k}{j}D_{\beta k}, \quad (13)$$

$$U_\gamma(\Theta) = \frac{\partial l(\Theta)}{d\gamma} = \sum_{k=0}^{N}\frac{1}{c_k}\sum_{k=0}^{n-y_k}(-1)^j\binom{n-y_k}{j}D_{\gamma k}, \quad (14)$$

$$c_k = \sum_{j=0}^{n-y_k}(-1)^j\binom{n-y_k}{j}B\left(\frac{y_k}{\gamma}+\alpha+\frac{j}{\gamma},\beta\right), \quad (15)$$

$$D_{\alpha k} = B\left(\frac{y_k}{\gamma}+\alpha+\frac{j}{\gamma},\beta\right)\left[\Psi\left(\frac{y_k}{\gamma}+\alpha+\frac{j}{\gamma}\right) - \Psi\left(\frac{y_k}{\gamma}+\alpha+\frac{j}{\gamma}+\beta\right)\right], \quad (16)$$

$$D_{\beta k} = \left(\frac{y_k}{\gamma}+\alpha+\frac{j}{\gamma},\beta\right)\left[\Psi(\beta) - \Psi\left(\frac{y_k}{\gamma}+\alpha+\frac{j}{\gamma}+\beta\right)\right], \quad (17)$$

$$D_{\gamma k} = \frac{y_k+j}{\gamma^2}B\left(\frac{y_k}{\gamma}+\alpha+\frac{j}{\gamma},\beta\right)\left[\Psi\left(\frac{y_k}{\gamma}+\alpha+\frac{j}{\gamma}+\beta\right) - \Psi\left(\frac{y_k}{\gamma}+\alpha+\frac{j}{\gamma}\right)\right], \quad (18)$$

and Ψ (.) is the digamma function.

The MLEs are obtained by solving the three simultaneous equations obtained by equating $U(\Theta) = 0$.

The alcohol data consumption was modeled by means of McGBB model and maximum likelihood method was used to obtain the estimates of parameters of McGBB distribution. The initial values for α and β were taken from BB model while for γ was taken as 1.

2.3 Estimating Functions

2.3.1 Introduction

The idea of Estimating functions was first introduced by Pearson (1894) and has become a useful tool for constructing estimators. The most important advantage of estimating functions method is that it does not require knowledge of the full model, but rather of some functions, such as moments. The basis behind estimating functions is to have, or to find, a set of simultaneous equations involving both the sample data and the unknown model parameters which are to be solved in order to find the estimates of the parameters.

The estimates obtained by Estimating functions, such as moment estimates, quasi-likelihood estimates (Breslow, 1990; Moore and Tsiatis, 1991), extended quasi-likelihood estimates (Nelder and Pregibon, 1987), the Gaussian likelihood estimates (Whittle, 1961; Crowder, 1985), estimates based on the pseudo-likelihood estimating equations (Davidian and Carrol, 1987) and estimates based on quadratic estimating functions (Crowder 1987; Godambe and Thompson 1989) provide a more robust estimates than MLEs when used in the estimation of parameters of distributions.

Two independent works of Durbin (1960) and Godambe (1960) showed that if we let $f(x;\ \theta, \phi)$ be the probability density function for a random variable x, indexed by two sets of parameters θ of p - dimensional and ϕ of q- dimensional then, an alternative approach to maximum likelihood method in estimation is to focus on functions $g(x, \theta)$ of the data x and the unknown set of parameters θ and to study the estimators as solutions of $g\ \ x;\ \theta, \phi\ \ = 0$.

2.3.2 Quasi-likelihood

Quasi-likelihood function was first introduced by Nelder and Wedderburn (1972). The basis of the method of estimating equations is a set of simultaneous equations involving both the data and the unknown model parameters. These equations are then solved in order to obtain estimates of the parameters. To define a quasi-likelihood function we need only to specify a relation between the mean and variance and the quasi-likelihood can then be used for estimation (Wedderburn, 1974). This means that the quasi-likelihood technique is used for estimating regression coefficients without fully specifying the distribution of the observed data, hence provides a more flexible approach to estimation than MLE estimation. QL is used to allow for over/under dispersion relative to the Poisson or binomial distribution and to estimate regression coefficients when it's not clear how to specify the distribution of the observed data.

Quasi-likelihood estimation is one way of allowing for over-dispersion, that is, greater variability in the data than would be expected from the statistical model used. It is most often used with models for count data or grouped binary data, i.e. data that can otherwise be modeled using the Poisson or binomial distribution.

Study by Wedderburn, (1974) presented quasi-likelihood to describe a function which has similar properties to the log-likelihood function, except that a quasi-likelihood function is not the log-likelihood corresponding to any actual probability distribution. Quasi-likelihood models can be fitted using a straight forward extension of the algorithms used to fit generalized linear models. Instead of specifying a probability distribution for the data, only a relationship between the mean and the variance is specified in the form of a variance function giving the variance as a function of the mean. Generally, this function is allowed to include a multiplicative factor known as the over-dispersion parameter or scale parameter that is estimated from the data. Most commonly, the variance function is of a form such that fixing the over-dispersion parameter at unity results in the variance-mean relationship of a probability distribution such as the binomial or Poisson.

The advantage of quasi-likelihood technique is that it doesn't need to fully specify a distribution for the response variable when one is uncertain about the random mechanism by which the data were generated. Paul and Islam (1994) showed that for the estimation of the mean (regression) parameters the Quasi-likelihood procedure performs the best among all the estimating functions investigated, hence it was considered in this study.

Suppose we have iid observations $y_1, y_2, \ldots, y_n$ with common mean vector u and variance $\sigma^2\,\mathrm{v}(u)$ and $U(u;Y)$ defined as;

$$U(u;Y) = \frac{Y_i - u}{\sigma^2\,\mathrm{v}(u)} \tag{19}$$

Then, the quasi-likelihood function is given by

$$Q(U,y) = \int_y^U \frac{y-t}{\sigma^2 v_t}\,dt, \text{ with}$$

$$E\left(U(u;Y)\right) = 0,\ var\left(U(u;Y)\right) = \frac{n}{\sigma^2\,\mathrm{v}(u)},\quad -E\left(\frac{\partial U}{\partial u}\right) = Var\left(U(u;Y)\right) \tag{20}$$

By following the above formulation, the QL estimating equation was obtained using the mean and variance in section 2.2.3 equation (7) and (8) respectively. The QL estimating equation; $U(u;Y) = 0$ was obtained which was then used to obtain QL estimates in chapter four.

2.3.3 The Extended Quasi-likelihood

Extended quasi-likelihood has been recommended for the estimation of dispersion parameters when there is no likelihood available. In the recent past there has been substantial interest in the joint fitting of mean and dispersion parameters. Wedderburn (1974) has argued that when an exact likelihood is not available, maximum likelihood (ML) estimation cannot be used. He showed that assuming the first two moments of responses only, estimating equations for the regression parameter can be given. For estimating dispersion parameters there are two general approaches. One can use either the extended quasi-likelihood (EQL) of Nelder and Pregibon (1987) or the pseudo-likelihood (PL), based on a normal likelihood. Davidian and Carroll (1987) adds that maximum pseudo-likelihood (MPL) are asymptotically consistent as the sample size increases while maximum extended quasi-likelihood (MEQL) estimators are consistent in the limit for a given sample size as parameter values tend to certain values.

In finite samples the mean square error (MSE) criterion seems to be the more relevant measure of the efficiency of estimators. Nelder and Lee (1992) showed that the sample size inconsistency of the MEQL estimator would often be offset by the small MSE in finite samples. Nelder and Lee (1992) showed that MEQL estimators work well for estimating dispersion parameters in models, such as negative binomial models. However, EQL for marginal models with over-dispersion may not be defined (McCullagh and Nelder, 1989).

An extension of the EQL to correlated errors expressed as the sum of independent error distributions can maintain high efficiency, providing a better approximation. This approach can also be used for a wider class of hierarchical generalized linear models (HGLMs) characterized solely by the first two moments.

2.3.4 Quadratic Estimating Equations

Godambe and Thompson (1989) and Crowder (1987) developed a general class of optimal Quadratic Estimating equations (QEEs) for both the mean and the dispersion parameters. The QEEs require the knowledge of the skewness and kurtosis of the population distribution, which in practical situations are unlikely to be known exactly, and estimation of these requires much more data than are usually available in practice. The study by Davidian and Carrol (l987) shows that when both skewness and kurtosis are equated to zero the QEEs forms Gaussian estimating equations.

Studies have found that Gaussian and quasi-likelihood estimates for count data are efficient compared to the maximum likelihood estimates (Dean and Lawless, 1989). They suggested that combining the quasi-likelihood estimating equations for the mean parameters and the optimal estimating equation of Crowder (1987) for dispersion parameter after setting the skewness and kurtosis to zero.

Estimation of the parameters by the quadratic estimating equations overcomes the failure of the maximum quasi-likelihood estimation to give reasonable results (Crowder, 1987). Study by Crowder (1987) showed that if we let Y_i to be the number of successes in n_i binomial trials and the random variable defined as $z_i = \frac{y_i}{n_i}$ to be binomial $(n_i;\ p)$, for i = 1, $\ldots, n$: with the binomial success probability p being distributed as a beta distribution with mean π_i and variance $\sigma_{i\lambda}^2$, then by considering estimating functions quadratic in z_i the QEEs has general form as $g_\lambda = \sum_i^n a_{i\lambda}(z_i - \pi_i) + b_{i\lambda}\{(z_i - \pi_i)^2 + \sigma_{i\lambda}^2\}$, where $a_{i\lambda}$ and $b_{i\lambda}$ are specified nonstochastic functions of λ.

When finding the QEEs based on McGBB distribution suggestions by Sudhir (2001) of varying the coefficients of QEEs were followed. By setting both skewness and kurtosis to zero and finding the higher moments for McGBB Distribution, four QEEs were obtain as shown in section (4.2.2) which were then used to find the estimates of parameters of McGBB.

2.3.5 Method of moments

The general idea behind the method of moments is to equate population moments to the corresponding sample moments. The moment estimates for the parameters α, β, γ and dispersion parameter, ρ of the McGBB distribution is obtained by equating sample moments to the corresponding population moments. Lee (2003) derived expressions for the mean and over-dispersion parameter estimates as follows:

$$\hat{\phi} = \frac{s^2 - \hat{\mu}}{\hat{\mu}^2} \quad (21)$$

where

$$S^2 = \sum_{j=1}^n (y_j - \overline{y})^2 / (n-1)$$

$$\hat{\mu} = \overline{y}$$

and $\overline{y}$ is the sample mean

2.3.6 The Gaussian Likelihood

Gaussian estimation procedure was introduced by Whittle (1961) which uses the normal log likelihood, without assuming that the data are normally distributed. Other studies have found that Gaussian and quasi-likelihood estimates for count data are more efficient than the maximum likelihood estimates (Dean and Lawless, 1989). On the other hand Crowder (1985) showed that Gaussian likelihood estimates are the estimates of choice if the interest is on estimation of the dispersion parameter or the joint estimation of the regression and the dispersion parameters.

CHAPTER THREE
METHODS

3.1 Introduction

This chapter describes a simulation study for assessing the performance of the estimates for the McGBB distribution parameters based on MLE, QL, GL, M1, M2 and M3. The derivation of Estimating Equations and the first derivatives of these estimating equations are shown in section 4.2. These equations are then integrated into the subroutines during simulation in order to obtain the estimates of parameters of McGBB distribution results shown in chapter four.

3.2 Simulation study

R language is truly one of the finest available software that allows for the statistical computation of files and graphics management. Thus, this study used this software to simulate data for the varying values of α, β and γ parameters.

McGBB distribution was used to generate over-dispersed data from McGBB distribution. Data simulated from the McGBB distribution was then used with the three estimation procedures to generate small sample results on parameters estimation based on the six estimation procedures under study. The relative efficiencies of the estimates α, β and γ obtained by the six estimation procedures using weekly (7 days) alcohol consumption survey data and simulated data was compared with MLEs. Along with estimates of the parameters, the estimated Relative efficiencies by all the methods including the maximum likelihood method was found. Estimated Relative efficiencies of α is $var\ \alpha_{ML}\quad var\ \alpha_t$ where t = QL, GL, M1, M2, M3. In the situation where relative efficiency is greater than one, then the procedure with its efficiency as the denominator is preferred than the "gold standard" ML. Taking $N = 399$, several sets of combination of α, β and γ parameters were chosen: $\gamma = 1,\ \beta = 0.5, \alpha = 0.0,\ \ldots, 1.0$ and $\gamma = 1, \alpha = 0.7$, $\beta = 0.0,\ \ldots, 1.0$. For each combination of α, β and γ parameters 5,000 samples were simulated from the McGBB distribution. During simulation, all the parameters α, β and γ were estimated for all the six procedures including maximum likelihood and their efficiencies and subsequently their relative efficiencies.

3.3 Maximum likelihood Estimation

McGBB distribution is a new distribution obtained by mixing McDonald beta-binomial distribution and success probability p of binomial distribution. If we let Y_i to be the number of successes in n_i binomial trials and the random variable defined as $z_i = y_i \; n_i$ to be binomial $n_i; \; p$, for i = 1, ...,n: with the binomial success probability p being distributed as a McDonald generalized beta distribution with mean π_i and variance $\pi \; 1-\pi \;\; 1+ \; n-1 \; \rho$, where ρ is the dispersion parameter, then the mean structure π_i is given by the logistic model

$$\pi_i \; X_i \,; \beta \; = \frac{e^{X_i\beta}}{1+e^{X_i\beta}} \tag{22}$$

where $X_i\beta = X_{i1}\beta_1 + \cdots + X_{ik}\beta_k$ and $X_1, \ldots, X_k$ are k explanatory variables, $\beta_1, \ldots, \beta_k$,are the k regression parameters.

The parameter ρ is overdispersion parameter. Maximum likelihood estimate (MLE) of $\beta_1, \ldots, \beta_k$ and ρ can be obtained by solving the maximum likelihood estimating equations. A probability mass function of McGBB (n, α, β, γ) distribution and the log-likelihood function for Θ are given by equations 5 and 9 respectively.

Let x_i be the number of successes observed in n_i clusters and p_i be the proportion of the successes, where $y_i = \frac{x_i}{n_i}$, then the conditional variation on observed y_i is given by:

$$Var \; y_i \; = Var \; E(y_i \;\; P_i) \; + E \; Var \; y_i \;\; P_i$$

$$Var \; y_i \; = Var \;\; E \; \frac{x_i}{n_i} \;\; + E \;\; Var \;\; \frac{x_i}{n_i}$$, where;

$$E \; \frac{x_i}{n_i} \; = n\pi_i \; 1-\pi_i \quad \text{and}$$

$$Var \; E \; \frac{x_i}{n_i} \; = n_i^2 Var \; \pi_i \; = n_i^2 \rho \; \pi_i \; 1-\pi_i$$

$$= n_i(E \; \pi_i \; - Var \; \pi_i \; - E \; \pi_i \; ^2 + n_i Var \; \pi_i \;)$$

$$= n_i \rho \pi_i \; 1-\pi_i \; + n_i \; n_i \pi_i \; - n_i(\rho \pi_i \; 1-\pi_i \; + n_i^2 \pi_i^2)$$

$$= n_i \pi_i \; 1-\pi_i \;\; 1+ \; n_i - 1 \; \rho \; , \tag{23}$$

where $\rho = \frac{\left(\frac{B(\alpha+\beta, \frac{2}{\gamma})}{B(\alpha, \frac{2}{\gamma})} - \frac{B(\alpha+\beta, \frac{1}{\gamma})}{B(\alpha, \frac{1}{\gamma})}\right)^2}{\left(\frac{B(\alpha+\beta, \frac{1}{\gamma})}{B(\alpha, \frac{1}{\gamma})} - \frac{B(\alpha+\beta, \frac{1}{\gamma})}{B(\alpha, \frac{1}{\gamma})}\right)^2}$, with $0 \le \pi \le 1$,

$$\pi_i = \frac{B\left(\alpha + \beta, \frac{1}{\gamma}\right)}{B\left(\alpha, \frac{1}{\gamma}\right)} \text{ and } E\left(\pi_i^2\right) = Var(\pi_i) + \left(E(\pi_i)\right)^2$$

The parameter π_i and ρ are not orthogonal except when $\rho = 0$. This is the extended McDonald generalized beta-binomial model which takes into consideration ρ as positive or negative based on whether the data is over-dispersed or under-dispersed. The MLEs are obtained by solving the three simultaneous equations obtained by equating $U(\Theta) = 0$ in section 2.2.4 for α, β and γ parameters. This equation 9 was solved by numerical methods to yield estimates of α, β and γ respectively. Taking second derivatives of this equation with respect to the parameters, the elements of the variance covariance matrix **I** was obtained. The efficiencies of MLE method was obtained using the asymptotic variance –covariance matrix of MLE's which was obtained by inverting the expected Fisher information matrix, where,

$$I = \begin{pmatrix} I_{BB} & I_{B\theta} \\ I_{\theta B} & I_{\theta\theta} \end{pmatrix}$$

$$I_{BB} = \left(-\frac{\partial^2 l}{\partial \beta_j \partial \beta_s}\right)_{kxk}$$

$$I_{B\theta} = I_{\theta B} = E\left(-\frac{\partial^2 l}{\partial \beta_j \partial \theta}\right)_{kxI}$$

$$I_{\theta\theta} = E\left(-\frac{\partial^2 l}{\partial^2 \theta}\right)$$

The maximum likelihood estimates are hereby denoted by λ_{ML}

3.4 The Inagaki results

By denoting the unbiased estimating equations obtained by the method of moments and other semi-parametric procedures by $u_1, u_2, ..., u_k$ and u_{k+1}, where u_j, $j = 1, 2, ..., k$

represent unbiased estimates for β_j and u_{k+1} represents the unbiased estimating equation for ρ Let $\hat{\lambda}$ be an estimate for $\lambda = \beta_1, \beta_2, \dots, \beta_k, \rho$. Using the method of moments or any semi-parametric procedure, the Inagaki (1973) result obtained under the usual regularity conditions, such as the finite dimensional parameter space, the expected values are continuously differentiable, while variance for Estimating Functions is given by Inagaki 1973 as;

$$Var(\hat{\lambda}) = \left\{A(\hat{\lambda})\right\}^{-1} B(\hat{\lambda}) \left[\left\{A(\hat{\lambda})\right\}^{-1}\right]^T \tag{24}$$

where A and B are square matrices of order $k+1$ with entries.

$$A_{k+1,j} = E\left(-\frac{\partial U_{k+1}}{\partial \beta_j}\right) \qquad A_{k+1,k+1} = E\left(-\frac{\partial U_{k+1}}{\partial \phi}\right)$$

$$A_{j,s} = E\left(-\frac{\partial U_j}{\partial \beta_s}\right) \qquad A_{j,k+1} = E\left(-\frac{\partial U_j}{\partial \phi}\right)$$

$$\beta_{j,s} = E\left(U_j U_s\right) \qquad \beta_{j,k+1} = \beta_{k+1,j} = E\left(U_j U_{k+1}\right)$$

$$\beta_{k+1,k+1} = E\left(U_{k+1}^2\right)$$

This is for all, $j, s = 1, \dots, k$

3.5 Comparison of the Estimation Methods

This was done through both real data and a simulation study. Several data sets were simulated from a parametric model; McGBB model for a specified parameter set and the performance of the ML with QEEs and QL methods was examined using the mean square error. In estimating the McGBB Parameters, data was simulated several times according to McGBB model to determine the efficiency of the true parameter values. A graphical comparison of the estimation methods was also done by plotting the graphs from a simulated data set for all the six procedures.

CHAPTER FOUR
RESULTS AND DISCUSSION

4.1 Introduction

This chapter displays the findings on McGBB distribution parameters estimates α, β and γ. Detailed discussion of the results is given based on observation on the tables displayed for MLE, QL, GL, M1, M2 and M3 estimates. Tables 2 and 3 display results for MLE, QL, GL, M1, M2 and M3 estimates based on real data set while tables 4 and 6 display results for MLE, QL, GL, M1, M2 and M3 estimates for varied values of α and β, respectively. Moreover, results of point parameter estimates of the McGBB using the Rodríguez-Avi *et al*. (2007) real data set and simulated data were obtained and are displayed in table 2 and table 5 respectively.

The estimates for the displayed results for various combinations of the distribution parameters can be replicated by running the subroutines displayed in the appendix A, B and C which was carried out in this study.

4.2 Parameter Estimation

4.2.1 The Quasi-Likelihood Estimation.

The quasi-likelihood (Wedderburn, 1974) is based on the knowledge of the form of first two moments of the random variable $z_i = Y_i / n_i$, defined in section 3.3.1 with,

$E(z_{i,}) = \pi_i$ and $var(z_i) = \pi(1-\pi)\{1+(n-1)\rho\}$ while $E(Y) = n\pi_i$ and $var(Y) = n\pi(1-\pi)\{1+(n-1)\rho\}$,

The quasi-likelihood with the above mean and variance is given by,

$$Q(z_i; \pi_i, \rho) = \int_{z_{i,}}^{\pi_i} \frac{(z_{i,}-\pi_i)n_i}{\pi_i(1-\pi_i)\{1+(n_i-1)\rho\}} d\pi_i \qquad (25)$$

This study also found out that from equation (25) two estimating equations arise; given any value of ρ the unbiased estimating equation for β_j and unbiased estimating equation that can be obtained by using the quasi likelihood when the $k\ \beta's$ parameters are estimated are given in equation 26 and 27 respectively.

$$U_j(\beta, \rho) = \frac{\partial Q}{\partial \beta_j} = \sum_{i=1}^{n} \frac{(z_{i,}-\pi_i)n_i\, d_{ij}(\beta)}{\pi_i(1-\pi_i)\{1+(n_i-1)\rho\}} = 0 \qquad (26)$$

where $i = 1, \dots, n$; $j = 1, \dots, k,$ and $0 \leq \pi \leq 1$

In this case $d_{ij}\beta$ is given as $d_{ij}\beta = \frac{\partial \pi_i}{\partial \beta_j} = \pi_i (1 - \pi_i) X_{ij}$.

$$U_{k+1}(\beta, \rho) = \frac{\partial Q}{\partial \beta_j} = \sum_{i=1}^{n} \frac{(z_{i,} - \pi_i)^2 n_i}{\pi_i (1 - \pi_i)(1 + (n_i - 1)\rho)} - m - k = 0 \qquad (27)$$

Given $Z = \frac{y}{n}$ through partial fraction and integration by parts equation 25 becomes

$$Q = \sum_{y=0}^{n} \frac{1}{1 + (n-1)\rho} \left[\log \left(\frac{\pi}{Z}\right) + (n - y) \log \left(\frac{1-\pi}{1-Z}\right) \right] \qquad (28)$$

where, $$\rho = \frac{\frac{B(\alpha+\beta, \frac{2}{\gamma})}{B(\alpha, \frac{2}{\gamma})} - \left(\frac{B(\alpha+\beta, \frac{1}{\gamma})}{B(\alpha, \frac{1}{\gamma})}\right)^2}{\frac{B(\alpha+\beta, \frac{1}{\gamma})}{B(\alpha, \frac{1}{\gamma})} - \left(\frac{B(\alpha+\beta, \frac{1}{\gamma})}{B(\alpha, \frac{1}{\gamma})}\right)^2} \quad \text{and} \quad \pi = \frac{B(\alpha+\beta, \frac{1}{\gamma})}{B(\alpha, \frac{1}{\gamma})}$$

Then the partial derivatives from equation (28) for the three parameters α, β, γ given ρ was also obtained as follows;

$$\frac{\partial Q}{\partial \alpha} = \sum_{y=0}^{n} \frac{1}{1 + (n-1)\rho} \left(y - \frac{\pi (n-y)}{1-\pi} \right) \left[\psi(\alpha + \beta) + \psi\left(\alpha + \frac{1}{\gamma}\right) - \psi(\alpha) + \psi\left(\alpha + \beta + \frac{1}{\gamma}\right) \right] \qquad (29)$$

$$\frac{\partial Q}{\partial \beta} = \sum_{y=0}^{n} \frac{1}{1 + (n-1)\rho} \left(y - \frac{\pi (n-y)}{1-\pi} \right) \left[\psi(\alpha + \beta) - \psi\left(\alpha + \beta + \frac{1}{\gamma}\right) \right] \qquad (30)$$

$$\frac{\partial Q}{\partial \gamma} = \sum_{y=0}^{n} \frac{1}{1 + (n-1)\rho} \left(\frac{y}{\gamma^2} - \frac{\pi (n-y)}{\gamma^2 (1-\pi)} \right) \left[\psi\left(\alpha + \beta + \frac{1}{\gamma}\right) - \psi\left(\alpha + \frac{1}{\gamma}\right) \right] \qquad (31)$$

$$\frac{\partial^2 Q}{\partial \alpha^2} = \sum_{y=0}^{n} \frac{n}{1 + (n-1)\rho} \left(\frac{\pi}{1-\pi} \right) \left[\psi(\alpha + \beta) + \psi\left(\alpha + \frac{1}{\gamma}\right) - \psi(\alpha) - \psi(\alpha + \beta) + \psi\left(\alpha + \frac{1}{\gamma}\right) - \psi(\alpha) - \psi\left(\alpha + \beta + \frac{1}{\gamma}\right) \right]^2 \qquad (32)$$

$$\frac{\partial^2 Q}{\partial\alpha\partial\gamma} = \sum_{y=0}^{n} \frac{n}{1+\ n-1\ \rho} \quad \frac{\pi}{\gamma^2\ 1-\pi} \quad \psi\ \ \alpha+\beta+\ \frac{1}{\gamma} \quad -\psi\ \ \alpha+\ \frac{1}{\gamma} \qquad \psi\ \alpha+\beta\ \ +$$

$$\psi\ \ \alpha+\ \frac{1}{\gamma} \quad -\psi\ \alpha\ \ -\psi\ \ \alpha+\beta+\ \frac{1}{\gamma} \tag{33}$$

$$\frac{\partial^2 Q}{\partial\alpha\partial\beta} = \sum_{y=0}^{n} \frac{n}{1+\ n-1\ \rho} \quad \frac{\pi}{1-\pi} \quad \psi\ \alpha+\beta\ \ +\psi\ \ \alpha+\ \frac{1}{\gamma} \quad -\psi\ \alpha\ \ -\psi\ \ \alpha+\beta+$$

$$\frac{1}{\gamma} \qquad \psi\ \alpha+\beta\ \ -\psi\ \ \alpha+\beta+\ \frac{1}{\gamma} \tag{34}$$

$$\frac{\partial^2 Q}{\partial\beta\partial\gamma} = \sum_{y=0}^{n} \frac{n}{1+\ n-1\ \rho} \quad \frac{\pi}{\gamma^2\ 1-\pi} \quad \psi\ \alpha+\beta\ \ -\psi\ \ \alpha+\beta+\ \frac{1}{\gamma} \qquad \psi\ \ \alpha+\beta+\frac{1}{\gamma}\ \ -$$

$$\psi\ \ \alpha+\ \frac{1}{\gamma} \tag{35}$$

$$\frac{\partial^2 Q}{\partial\beta^2} = \sum_{y=0}^{n} \frac{n}{1+\ n-1\ \rho} \quad \frac{\pi}{1-\pi} \quad \psi\ \alpha+\beta\ \ -\psi\ \ \alpha+\beta+\ \frac{1}{\gamma} \quad ^{2} \tag{36}$$

$$\frac{\partial^2 Q}{\partial\gamma^2} = \sum_{y=0}^{n} \frac{n}{1+\ n-1\ \rho} \quad \frac{\pi}{\gamma^4\ 1-\pi} \quad \psi\ \ \alpha+\beta+\ \frac{1}{\gamma} \quad -\psi\ \ \alpha+\ \frac{1}{\gamma} \quad ^{2} \tag{37}$$

4.2.2 The Quasi-Likelihood variance

Based on these estimating equations (24) and (25), the Inagaki (1973) gave the expressions for the asymptotic variances of the quasi-likelihood estimate whose elements, $\mathrm{A}\left(\hat{\lambda}\right)$ and $B\left(\hat{\lambda}\right)$ which are $k\times 1$ by $k\times 1$ matrix and are components of the Hessian matrix. The following expressions were obtained:

$$A_{js} = E\ \ -\frac{\partial U_j}{\partial\beta_s}\ \ =\ \ \sum_{i=1}^{n}\frac{n_i x_{ij}}{(1+\ n_i-1\ \rho)}\ \ ,\qquad j,s=1,\ldots,k$$

$$A_{j,k+1} = E\ \ -\frac{\partial U_j}{\partial\rho}\ \ =0\quad ,\qquad j=1,\ldots,k$$

$$a_j = A_{\ k+1,j} = E\ \ -\frac{\partial U_{k+1}}{\partial\beta_j}\ \ =\ \ \sum_{i=1}^{n}\ 1-2\pi_i\ \ X_{ij}\ ,\ j=1,\ldots,k$$

$$a_{k+1} = A_{k+1,k+1} = E\left(-\frac{\partial U_{k+1}}{\partial \rho}\right) = \sum_{i=1}^{n} \frac{n_i-1\ \rho)}{1+(n_i-1)\rho}\ ,\ j = 1, \dots, k$$

$$B_{j,s} = E\left(U_j U_s\right) = \sum_{i=1}^{n} \frac{n_i \pi_i\,(1-\pi_i)\, x_{ij} x_{is}}{1+(n_i-1)\rho} = A_{j,s}\ ,\ j = 1, \dots, k$$

$$b_j = B_{k+1,j} = B_{j,k+1} = E\left(U_{k+1} U_j\right) = \sum_{i=1}^{n} \frac{(1-2\pi_i)(1+\ 2n_i-1\ \rho X_{ij}.)}{(1+\rho)\ 1+(n_i-1)\rho}\ ,\ j = 1, \dots, k$$

$$B_{k+1,k+1} = E\left(U_j U_{k+1}\right) = \sum_{i=1}^{n} \frac{n_i^2 (n_i-1)^2) E(z_i-\pi_i)^4}{n_i \pi_i\,(1-\pi_i)^{\ 2}\ 1+(n_i-1)\rho^{\ 4}} - \frac{(n_i-1)^2}{1+(n_i-1)\rho^{\ 2}}$$

$$B_{k+1,k+1} = b_{k+1}$$

The entries of matrices A and B are;

$$\text{A}=\begin{pmatrix} A & \underline{0} \\ a' & a_{k+1} \end{pmatrix} \qquad \text{B}=\begin{pmatrix} A & b \\ b' & b_{k+1} \end{pmatrix}$$

Thus, substituting the above matrices into the asymptotic variance equation (23) yields the expression below which is used in finding relative efficiencies given in this chapter:

$$\text{var}(\hat{\lambda}_{Qm}) = \begin{pmatrix} A & \underline{0} \\ a' & a_{k+1} \end{pmatrix}^{-1} \begin{pmatrix} A & b \\ b' & b_{k+1} \end{pmatrix} \left(\begin{pmatrix} A & \underline{0} \\ a' & a_{k+1} \end{pmatrix}^{-1} \right)^{T} \quad (38)$$

Where A's and B's in this case are the first and second derivatives of the estimating equations respectively.

4.2.3 Quadratic Estimating Equations

In this section, suggestions by Sudhir (2001) were followed. By considering estimating functions quadratic in z_i the QEEs have a general form given as $g_{\lambda} = \sum_{i}^{m} a_{i\lambda}\ z_i - \pi_i\ + b_{i\lambda}\{\ z_i - \pi_i\ ^2 + \sigma_{i\lambda}^2\}$ Crowder (1987), where $a_{i\lambda}$ and $b_{i\lambda}$ are specified nonstochastic functions of λ. Thus, through derivation the unbiased quadratic estimating equations for parameters; α, β and γ of McGBB distribution was found as follows;

The unbiased quadratic estimating equations for α, β and γ and ρ have the form

$$U_j(\beta,\rho) = \sum_{i}^{n} \left[a_{i\beta_j}(z_i-\pi_i) + b_{i\beta_j}\{(z_i-\pi_i)^2 + \sigma^2{}_{i\lambda}\}\right] = 0, \quad (39)$$

By taking $a_{i\beta_j} = \frac{1}{\sigma^2{}_{i\lambda}} + \frac{(1-2\pi_i)^2 n_i}{2(1-\rho)\pi_i^2(1-\pi_i)^2} d_{ij}\beta$ and $b_{i\beta_j} = \frac{-n_i\rho(1-2\pi_i)d_{ij}(\beta)}{2(1-\rho)\pi_i(1-\pi_i)\sigma^2{}_{i\lambda}}$, then a Gaussian estimating equations denoted by $\boldsymbol{\lambda_{GL}}$, was obtained as,

$$U_j(\beta,\rho) = \sum_{i}^{n} \left[\left(\frac{1}{\sigma^2{}_{i\lambda}} + \frac{(1-2\pi_i)^2 n_i}{2(1-\rho)\pi_i^2(1-\pi_i)^2} d_{ij}\beta\right)(z_i-\pi_i) + \frac{-n_i\rho(1-2\pi_i)d_{ij}(\beta)}{2(1-\rho)\pi_i(1-\pi_i)\sigma^2{}_{i\lambda}}\{(z_i-\pi_i)^2 + \sigma_{i\lambda}\}\right]. \quad (40)$$

Secondly, by taking $a_{i\beta_j} = \frac{d_{ij}(\beta)}{\sigma^2{}_{i\lambda}}$, and $b_{i\beta_j} = 0$ unbiased estimating equations (QEE's) for McDonald Generalized Binomial Distribution was obtained. The equation below was denoted by $\boldsymbol{\lambda_{M1}}$

$$U_j(\beta,\rho) = \sum_{i}^{n} \left[\frac{d_{ij}\beta}{\sigma^2{}_{i\lambda}}(z_i-\pi_i) + b_{i\beta_j}\{(z_i-\pi_i)^2 + \sigma^2{}_{i\lambda}\}\right]$$

This simplifies to;

$$U_j(\beta,\rho) = \sum_{i}^{n} \frac{d_{ij}\beta}{\sigma^2{}_{i\lambda}}(z_i-\pi_i) \quad (41)$$

For $a_{i\beta_j} = \frac{-\Upsilon_{2i\lambda}+2+\Upsilon_{1i\lambda}(1-2\pi_i)\sigma_\lambda\,[\pi_i(1-\pi_i)]\,d_{ij}(\beta)}{\sigma^2{}_{i\lambda}\Upsilon_{i\lambda}}$, and $b_{i\beta_j} = \frac{[\Upsilon_{1i\lambda}-(1-2\pi_i)\sigma_\lambda(\pi_i(1-\pi_i))]\,d_{ij}(\beta)}{\sigma^3{}_{i\lambda}\Upsilon_{i\lambda}}$,

where $\Upsilon_{i\lambda} = \Upsilon_{2i\lambda} + 2 - \Upsilon^2{}_{1i\lambda}$ an optimal quadratic estimating equations was obtained. It was noted that the forms of the skewness $\Upsilon_{1\lambda}$ and the kurtosis $\Upsilon_{2\lambda}$ are not known. However, this was taken based on the second, third and fourth moments of the McDonald generalized beta-binomial distribution, which are:

$$\mu_{2i} = \pi_i(1-\pi_i)[1+(n_i-1)\rho]/n_i,$$

$$\mu_{3i} = \mu_{2i}(1-2\pi_i)[1+(2n_i-1)\rho]/n_i(1+\rho),$$

and, $\mu_{4i} = \mu_{2i}\left[\frac{(1+(2n_i-1)\rho)(1+(3n_i-1)\rho)(1-3\pi_i(1-\pi_i))}{1-\rho} + (n_i-1)\rho + 3n_i\mu_{2i}\right]\frac{1-\rho}{(1+\rho)(1+2\rho)n_i^2}$

The estimates obtained by solving these optimal quadratic estimating equations was denoted by $\boldsymbol{\lambda_{M2}}$. Further, the estimates obtained by solving the optimal quadratic estimating equations with $\Upsilon_{1i\lambda} = \Upsilon_{2i\lambda} = 0$ was denoted by $\boldsymbol{\lambda_{M3}}$. Note the estimate $\boldsymbol{\lambda_{M3}}$ is also obtained by using the pseudo-likelihood estimating equations of Davidian and Carrol (1987).

$$U_j(\beta,\rho) = \sum_i^n \frac{-\Upsilon_{2i\lambda}+2+\Upsilon_{1i\lambda}(1-2\pi_i)\sigma_\lambda \; \pi_i\,(1-\pi_i)\; d_{ij}(\beta)}{\sigma^2{}_{i\lambda}\Upsilon_{i\lambda}}(z_i-\pi_i) +$$

$$\frac{\Upsilon_{1i\lambda}-(1-2\pi_i)\sigma_\lambda \; \pi_i\,(1-\pi_i)\; d_{ij}(\beta)}{\sigma^3{}_{i\lambda}\Upsilon_{i\lambda}}\{(z_i-\pi_i)^2+\sigma^3{}_{i\lambda}\} = 0. \quad (42)$$

4.3 Small Sample Relative Efficiency

This study compares the small sample relative efficiency of the estimates $\boldsymbol{\lambda}$ obtained by the six estimation procedures; ML, *QL, GL, M1*, *M2* and *M3*. The estimating equations for all the five estimates given in this thesis (section 4.2.2) have the general form of equation (39) with specific expressions for $a_{i\beta_j}$ and $b_{i\beta_j}$ for each method.

A simulation study was conducted taking N reasonably small. For each sample McGBB parameters were estimated by the six procedures including the MLE procedure and small efficiency are calculated as follows.

Relative Efficiency $= var(\alpha_{ML}) \;\; var(\alpha_t)$ where t = QL, GL, M1, M2, M3

The relative efficiency results for McGBB parameters are given in tables 4 and 6 and plotted in figure 1, 2, 3 and 4 for simulated data.

4.4 Estimation Results

4.4.1 Real Data Results

Table 1: Alcohol consumption data

y	0	1	2	3	4	5	6	7
n	47	54	43	40	40	41	39	95

The data set above was used by Alanko and Lemmens (1996), Rodríguez-Avi *et al*. (2007), and Chandrabose *et al*. (2013) in the study of handling over-dispersion. It shows the number of days an individual consumes alcohol y, out of $n = 7$ days in N=399, where y= Number of days. For this data sets, the estimates for α, β and γ by different methods and estimated relative efficiencies are given in tables 2 and 3 respectively**.**

Table 2: The McGBB parameter estimates by MLE, QL, GL, M1, M2 and M3 for real data

Parameters		A	β	γ
	MLE	0.0333	0.1797	26.7312
	λ_{QL}	0.0281	0.1502	25.5541
	λ_{GL}	0.0301	0.1671	25.8127
	λ_{M1}	0.0287	0.1655	24.8421
	λ_{M2}	0.0312	0.1671	25.4523
	λ_{M3}	0.0282	0.1611	24.6746

Table 3: Relative efficiencies of McGBB parameters by MLE, QL, GL, M1, M2 and M3 for real data

Parameters		α	β	γ
	MLE	1.000	1.000	1.000
	λ_{QL}	1.0912	0.9085	0.9010
	λ_{GL}	1.0424	0.9831	0.8742
	λ_{M1}	0.6121	0.5192	0.5320
	λ_{M2}	0.9352	0.9811	0.9381
	λ_{M3}	0.3292	0.3615	0.3510

The table 3 presents the relative efficiencies of the real data given in table 1. It shows that $\boldsymbol{\lambda}_{GL}$ and $\boldsymbol{\lambda}_{QL}$ have higher efficiency in comparison with the other four methods for parameter α.

4.4.2 Simulation Results

Table 4: The estimated Relative efficiencies of McGBB parameters by MLE, QL, GL, M1, M2 and M3 methods for $\gamma = 1$, $\beta = 0.5$ and α varied

	Estimated Relative Efficiencies				
α varied	λ_{QL}	λ_{GL}	λ_{M1}	λ_{M2}	λ_{M3}
0.0	0.980	0.950	0.591	0.933	0.454
0.1	0.990	0.967	0.638	0.938	0.519
0.2	0.998	0.985	0.490	0.859	0.586
0.3	1.014	1.000	0.592	0.928	0.609
0.4	1.052	1.053	0.617	1.247	0.425
0.5	1.095	1.025	0.706	0.990	0.438
0.6	1.148	1.135	0.669	1.552	0.411
0.7	1.131	1.021	0.655	0.839	0.415
0.8	1.035	1.010	0.592	0.982	0.298
0.9	1.001	0.989	0.529	0.952	0.216
1.0	0.998	0.993	0.389	0.941	0.201

The table 6 presents the relative efficiencies of the simulated data with varied beta. It shows that λ_{GL} and λ_{QL} have higher efficiency in comparison with the other four methods followed by λ_{M2} λ_{M1} and least being λ_{M3}. λ_{GL} performs better between 0 to 0.9 showing higher relative efficiencies. These results are plotted in figure 1 and 3.

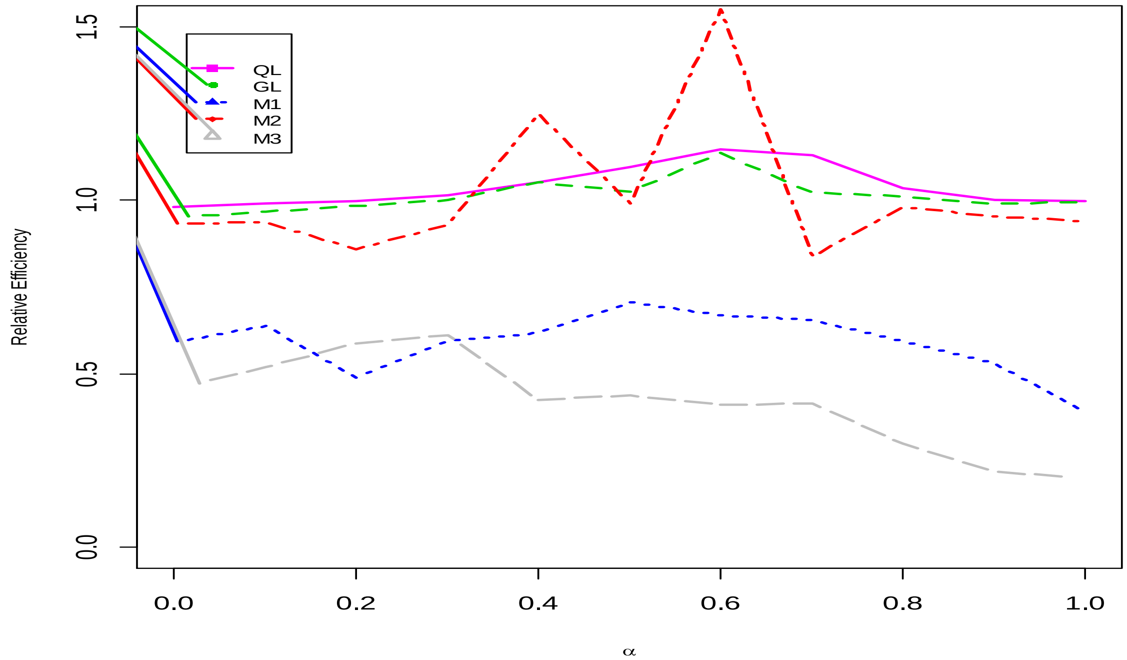

Figure 1: Plot of relative efficiencies comparison for various estimators relative to that of the MLE under McGBB model for α varied when $\beta = 0.5$ and $\gamma = 1$ for all procedures.

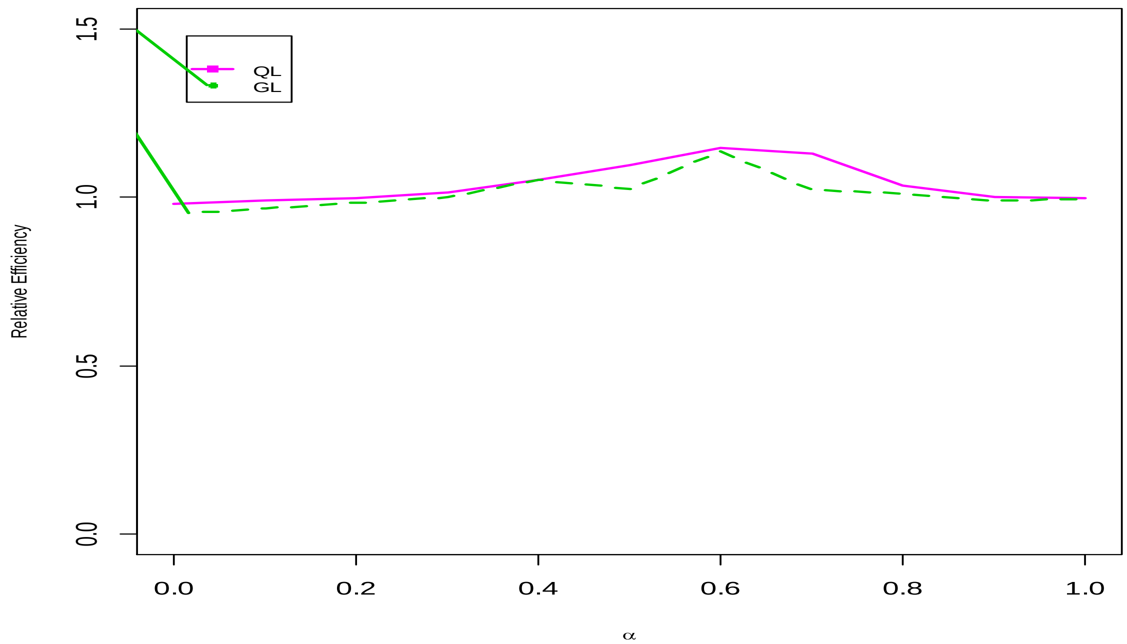

Figure 2: Plot of relative efficiencies comparison for estimators relative to that of the MLE under McGBB model for α varied when $\beta = 0.5$ for GL and QL procedures.

Table 5: The parameter estimates of McGBB by MLE, QL, GL, M1, M2 and M3 methods for $\alpha = 0.03,\ \beta = 0.15$ and $\gamma = 25$

Parameters		α	β	γ
	MLE	0.0399	0.1795	26.7311
	λ_{QL}	0.0331	0.1695	25.9599
	λ_{GL}	0.0310	0.1678	25.8135
	λ_{M1}	0.0291	0.1390	23.8428
	λ_{M2}	0.0302	0.1659	25.4519
	λ_{M3}	0.0254	0.1215	23.6751

The table 5 presents the parameter estimates for the simulated data using the real data initials estimates as the initial values in table 2. From this table it's clear that the simulated data estimates are near the real data estimate. This shows that the simulation predicts the real data well.

Table 6: The estimated Relative efficiencies of McGBB parameters by MLE, QL, GL, M1, M2 and M3 methods for $\gamma = 1$, $\alpha = 0.7$ and varied β

	Estimated Relative Efficiencies					
β varied	MLE	λ_{QL}	λ_{GL}	λ_{M1}	λ_{M2}	λ_{M3}
0.0	1.000	1.080	0.950	0.431	0.946	0.589
0.1	1.000	1.109	0.997	0.488	0.908	0.429
0.2	1.000	1.210	0.989	0.549	0.885	0.348
0.3	1.000	1.214	1.130	0.627	0.941	0.411
0.4	1.000	1.252	1.289	0.717	1.493	0.495
0.5	1.000	1.350	1.325	0.796	0.898	0.517
0.6	1.000	1.348	1.305	0.739	1.541	0.524
0.7	1.000	1.343	1.216	0.715	0.953	0.459
0.8	1.000	1.235	1.101	0.69	0.894	0.398
0.9	1.000	1.191	0.995	0.652	0.958	0.306
1.0	1.000	0.908	0.985	0.479	0.902	0.297

The table 6 presents the relative efficiencies of the simulated data with varied beta. It shows that λ_{GL} and λ_{QL} have higher efficiency in comparison with the other four methods followed by λ_{M2} and λ_{M1} and the least being λ_{M3}. Between the values 0 to 0.9 λ_{GL} performs better while λ_{QL} performs better between the values 0.3 to 0.8 since they have higher relative efficiencies in respect to other methods. These results are plotted in figure 1 and 3.

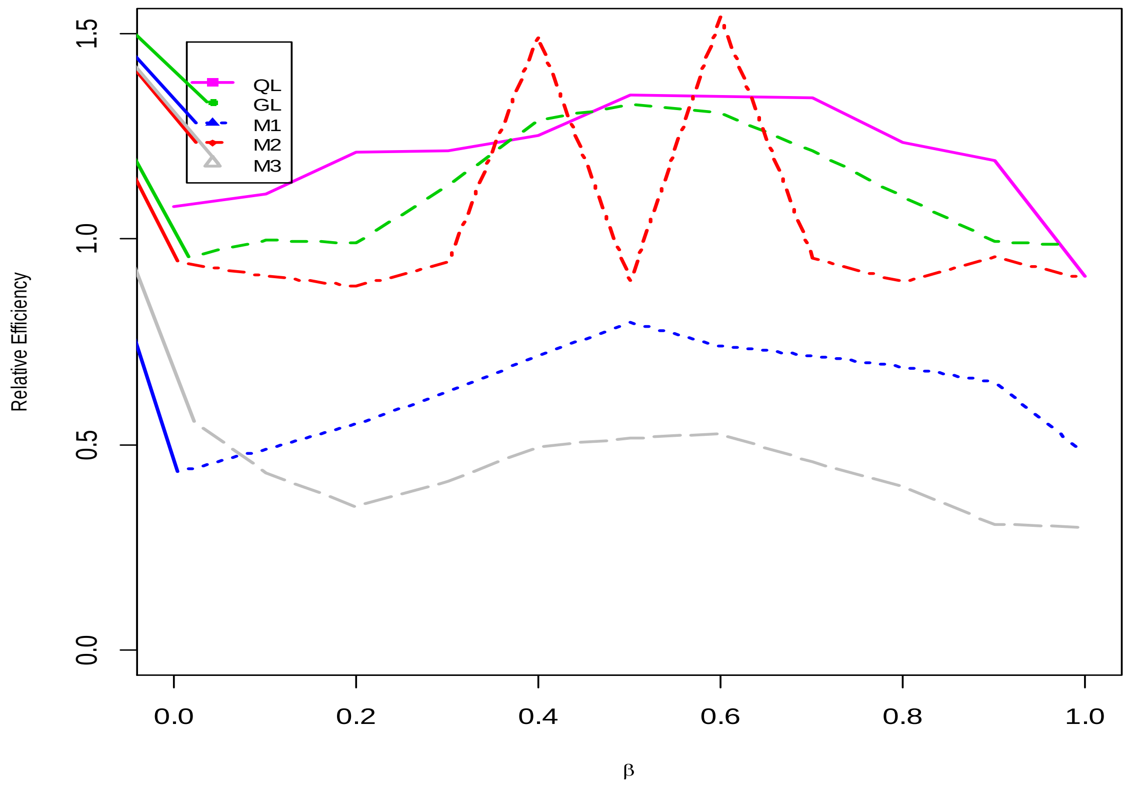

Figure 3: Plot of relative efficiencies comparison for various estimators relative to that of the MLE under McGBB model for β varied when $\gamma = 1$ and $\alpha = 0.7$ for all procedures.

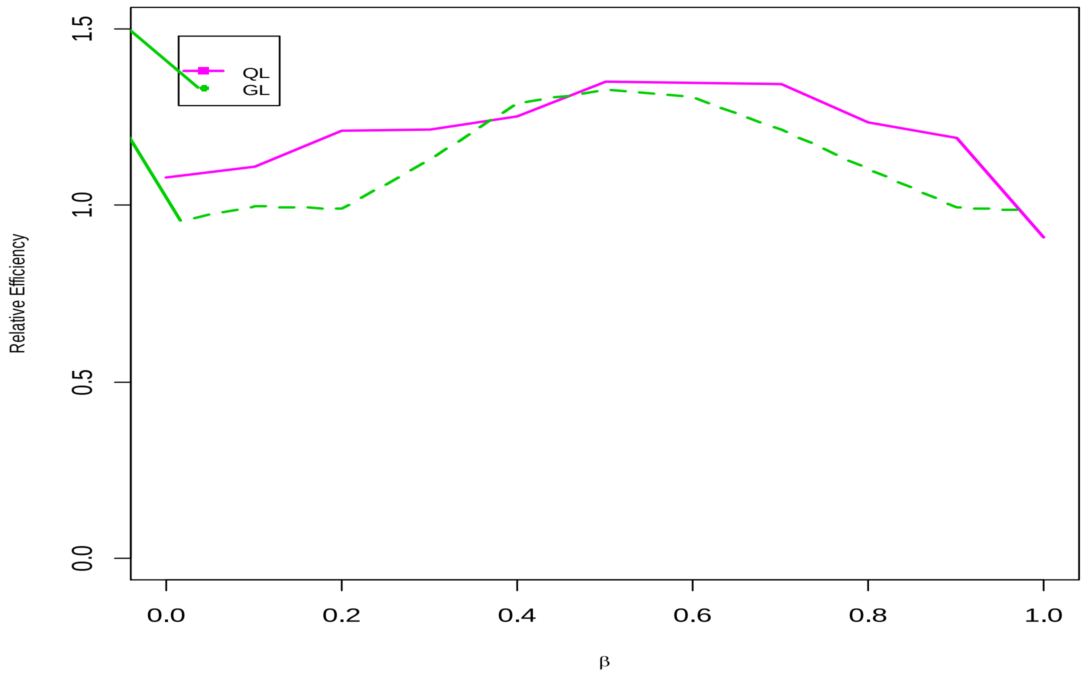

Figure 4: Plot of relative efficiencies comparison for estimators relative to that of the MLE under McGBB model for β varied when $\alpha = 0.7$ for GL and QL procedures.

4.5 Discussion

Maximum likelihood procedure relative efficiency comparison with the five procedures; QL, GL, M1, M2 and M3 was done as shown in respective tables and figures. Figure 1, shows relative efficiencies comparison when $\gamma = 1$ and $\beta = 0.5$ and α is varied for all procedures. These results present QL and GL as the best methods. While figure 3 shows β varied when $\alpha = 0.7$ and fix $\gamma = 1$ for all procedures and figure 4; β varied when $\alpha = 0.7$ and fix $\gamma = 1$ for GL and QL procedures under simulated data.

The relative efficiency results for the parameters are summarized in table 3 for the real data and those for simulated data are summarized in tables 4 and 6. Tables 3, 4 and 6 shows that the methods *QL; GL; M*2*;* all consistently provide high efficiency (never below 0.83) and method Q*L* is consistently the best. The next best appears to be the G*L* method followed by M2 and the least being M3. Estimates of parameters by all methods have high efficiencies. The good behaviour of the Gaussian likelihood estimator may be due to the fact that the Gaussian likelihood is a proper likelihood and the distribution of the data does not depend on a specific departure from the binomial distribution.

. In this thesis it is clear that the estimates based on the Gaussian likelihood estimates (the *GL* method); the estimates based on the combination of the quasi-likelihood estimates for the regression parameters and the optimal quadratic estimating equation for parameters after setting the skewness and kurtosis to zero (the *M*1 method) and the estimates based on the pseudo-likelihood estimating equations of Davidian and Carrol (1987) (the *M*3 method) are all special cases of the quadratic estimating equations while M2 is a QEE. It is noted that GL combine the good behaviour of Quasi- likelihood estimating Equation which has proved to provide consistent estimates in this thesis research. Many studies carried out on Estimating functions for example the study by Paul and Islam (1998) has proved superiority in parameter estimation as compared to MLE in Beta-Binomial Distribution. The present study echo this findings and shows that the estimates based on GL and QL provide best estimates through small sample Relative efficiency. Neither of these two methods requires the knowledge of third and the fourth moments of the McGBB distribution. The third best was optimal quadratic estimating equations using the third and the fourth moments of the McGBB distribution (the *M*2 method).The above three procedures have proved to be consistent. The least methods at the cost of some loss of efficiency are the *M1* and *M*3 which have proved to be inconsistent. Generally when data follow a McGBB distribution QL, GL and M2 methods are expected to have high efficiency.

CHAPTER FIVE
SUMMARY, CONCLUSION AND RECOMMENDATION

5.1 Introduction

This section gives the summary of the findings of research, conclusion of the study. The conclusion is given based on each specific objective given in this study. Recommendations for further study and areas of application of study based on the results are also given.

5.2 Summary and Conclusion

In this thesis the interest was in the analysis of the estimation of the McGBB distribution parameters. The dispersion parameter in this thesis plays the role of a nuisance parameter. However, in some instances like Toxicology and other similar fields, the dispersion parameter or the intraclass correlation parameter is of primary interest. The three objectives were achieved as follows;

- In this work, the estimate based on MLE and the estimating equations for QL, GL, M1, M2 and M3 procedures have been derived. The estimation was done using Quasi-likelihood and quadratic estimating equations (QEEs). By varying the coefficients of the QEEs four sets of estimating equations was obtained.
- A comparison of small sample relative efficiency of the five sets of estimates obtained by the QL, GL, M1, M2 and M3 estimates with the maximum likelihood estimates was done. Estimated small sample relative efficiencies of these estimates were also compared for a real life data sets arising from alcohol consumption practices.
- The performance of the estimates using R (2.13.0) software simulation technique was also examined. These comparisons results show that estimates, using optimal quadratic estimating equations are highly efficient and are the best among all estimates investigated.

5.3 Recommendation and Further Research

This study has investigated the Efficiency of the Maximum Likelihood estimators with QL, GL, M1, M2 and M3 through the construction of estimates for the parameters of McGBB Distribution. Future research may consider the construction of estimates on estimating functions based on method of moments, Gaussian likelihood for the McGBB

Distribution and a comparison be done with the MLEs. Secondly, further studies may use large sample in estimation of McGBB distribution parameters. A comparison of the large sample relative efficiencies based on estimating functions and Maximum Likelihood estimates can also be considered. The idea of Estimation equations can also be applied to other distributions apart from the McGBB distribution so as to construct better parameter estimates.

Thirdly, further study may consider interval estimation for McCBB dispersion parameter and researcher may also consider a robustness study for the efficiency property of these methods which is necessary when data come from other over/under-dispersed binomial distribution such as the probit normal binomial and the logit normal binomial distribution.

REFERENCES

Alanko, T., & Lemmens, P. H. (1996). Response effects in consumption surveys: an application of the betabinomial model to self-reported drinking frequencies. *Journal of Official Statistics*, **12:** 253 - 273.

Alexander, C., Cordeiro, G. M., Ortega, E. M., and Sarabia, J. M. (2012). Generalized beta-generated distributions. *Computational Statistics and Data Analysis,* **56**: 1880 - 1897.

Breslow, N.E. (1990). Tests of hypothesis in over-dispersed poisson regression and other quasi-likelihood models. *J. Am. Statist. Assoc.*, **85**: 565 - 571.

Chandrabose, M., Pushpa, W., and Roshan D. (2013). The McDonald Generalized Beta-Binomial distribution: A New Binomial Mixture Distribution and simulation based comparison with its nested distributions in handling over-dispersion. *International journal of statistics and probability*; **2**: 213 - 223.

Crowder, M. J. (1978). Beta-binomial Anova for proportions. *Journal of the Royal Statistical Society. Series C (Applied Statistics),* **27:** 34 - 37.

Crowder, M. J. (1985). Gaussian estimation for correlated binomial data, *Journal of the Royal Statistical Society*, Series B, **47**: 229 - 237.

Crowder, M .J. (1987). On linear quadratic estimating functions, *Biometrika,* **74**: 591 - 597

Davidian, M., and Carrol, R. J (1987). Variance function estimation*; Journal of the American statistical association.***82:** 1079 - 1091.

Dean, C**.,** and Lawless, J.F. (1989). Discussion on an extension of quasi-likelihood estimation (by V.P. Godambe and M.E. Thompson). J. *Statist. Plann. Inference***, 22**: 155 - 158.

Durbin, J. (1960). Estimation of parameters in time series regression models. *J. Roy. Stat. Soc*. Ser, B, **22**: 139 - 53.

Ennis, D. M., & Bi, J. (1998). The beta-binomial model: accounting for inter-trial variation in replicated difference and preference tests. *Journal of Sensory Studies,* **13:** 389 - 412.

Godambe, V.P. (1960). An optimum property of regular maximum likelihood estimation. *Ann. Math. Stat*., **31**: 1208 - 12.

Godambe, V. P., and Thompson, M. E. (1989). An extension of Quasi-likelihood estimation, *Journal of statistical planning and inference*, **22**: 137 - 152.

Inagaki, N. (1973). Asymptotic relations between the likelihood estimating functions and the maximum likelihood estimator, *Annals of the institute of statistical mathematics*, **25**: 1 - 26.

Islam, A. S. (1994). Parametric and semi-parametric models for the analysis of proportions in presence of over/under dispersion. PhD thesis, University of Windsor

Lee, Y. (2004). Estimation of parameters, *Biometrics*, **88**: 987 - 1006.

Lee, S. (2003). Analysis of Binary data in one way layout, *Biometrics*, **45**:195 - 206.

Li, X. H., Huang, Y. Y., and Zhao, X. Y. (2011). The Kumaraswamy Binomial Distribution. *Chinese Journal of Applied Probability and Statistics*, **27**: 511 – 521.

McCullagh, P., and Nelder, J. A. (1989). *Generalized Linear Models*, 2nd ed. Chapman & Hall, London.

McDonald, J. B. (1984). Some generalized functions for the size distribution of income. *Econometrica: Journal of the Econometric Society*, **52:** 647 - 663.

McDonald, J. B., and Xu, Y. J. (1995). A generalization of the beta distribution with applications *Journal of Econometrics*, **66**: 133 - 152

Moore, D. F. (1986). Asymptotic properties of moment estimators for Over-dispersed Counts and Proportions, *Biometrics*, **73**: 583 - 588.

Moore, D. F., and Tsiatis, M. (1991). Robust estimation of the standard error in moment methods for extra binomial/ extra Poisson variation, *Biometrika*, **47**: 383 - 401.

Nelder, J.A., and Lee, Y. (1992). Likelihood, quasi-likelihood and pseudo-likelihood: Some comparisons. J. *Roy***.** *Statist..* Soc**.** *Ser. B*, **54**: 273 - 284.

NeIder, J. A., and Pregibon, D. (1987). An extended quasi-likelihood function. *Biometrika*, **74:** 221 - 232.

Nelder J. A., and Wedderburn, R. W. (1972). Generalized linear models, *J. Roy statistical society*. *Ser.* A, **135**: 370 - 384.

Neyman, J., and Scott, E. L., (1948). Consistent estimates based on partially consistent observations, *Econometrica*, **16**: 1 - 32.

Paul, S. R. (2009). Estimation of regression and dispersion parameters in the analysis of proportions, *Canadian journal of statistics*, **25***:* 283 - 303.

Paul, S. R., and Islam, A. S. (1998). Joint estimation of the mean and dispersion parameters in the analysis of proportions: a comparison of efficiency and bias, *The Canadian*

Journal of Statistics, **26:** 83 – 94.

Pearson, K., (1894). Contribution to the mathematical theory of evolution, Philo-sophical Transactions of the Royal Society of London **185**: 71 - 110.

R Development Core Team. (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing (Version 2.10.0), Vienna, Austria.

Rodr´ıguez-Avi, J., Conde-S´anchez, A., S´aez-Castillo, A. J., & Olmo-Jim´enez, M. J. (2007). A generalization of the beta-binomial distribution. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **56:** 51 - 61.

RStudio. (2012). RStudio: Integrated development environment for R (Version 0.95.256), Boston, MA, USA.

Sudhir, R. P. (2001). Quadratic estimating equations for the estimation of regression and dispersion parameters in the analysis of proportions, *The Indian Journal of Statistics* **63:** 43–55.

Wedderburn, R. M. (1974). Quasi-likelihood functions, generalized linear models and the Gauss Newton method, *Biometrics*, **61:** 439 - 447.

Whittle, P. **(**1961**).** Gaussian estimation in stationary time series. *Bull. Internat. Stafist. Inst.*, **39**: 1 - 26.

APPENDIX

Appendix A: Simulation of the McGBB Distribution

```
dMcGBB<-function(x){
        n<-7
        a<-0.722
        b<-0.581
        c<-1
     j<-0:(n-x)
        term<-sum(((-1)^j)*(choose(n-x,j))*(beta(((x/c)+a+(j/c)),b)))
        return(choose(n,x)*(1/beta(a,b))*term)}
pj<-rep(0,8)
for (k in 1:8)  pj[k]<-dMcGBB(k-1)
qj<-c(1/8,1/8,1/8,1/8,1/8,1/8,1/8,1/8)
N<-400
Xobs<-rep(0,N)
for (i in 1:N){
    u2<-4; d<-2
    while(u2>d){        u1<-runif(1);
u2<-runif(1)

        y<-trunc(8*u1)+1            d<-pj[y]/0.22}

        Xobs[i]<-y-1            }
table(Xobs)
```

Appendix B: Maximum Likelihood Estimates of parameters of McGBB Distribution

library(bbmle)

```
###### MGeneralized Beta- binomial Negative Log Likelihood declaration
MGenBetaBinNLL<-function(x,a,b,c,fre,n){
  density<-c()
```

```
  for( i in 0:n){
    j <- 0:(n-i)
    term<-sum((((-1)**j)*(choose(n-i,j))*(beta(((i/c)+a+(j/c)),b)))
    vector.density<-choose(n,i)*(1/beta(a,b))*term
    density[i+1]<-vector.density  }
  MGBBLL<-sum(fre*log(density))
  return(-GBKBLL)}
```

Appendix C: Relative Efficiencies

```
##FInal Graph - Simulation
xlab.names<-expression(hat(beta))
main.names<-expression(paste("(c) Simulation: Estimated Relative Efficiencies of QL and GL
vs ", hat(beta) , ",  " , alpha,"=0.7", " using McGBB distribution"))
win.graph()
par(mfrow=c(1,2))
ylim1<-seq(0,1.5,0.1)
values<-c(0.0, 0.1,  0.2,  0.3,  0.4,      0.5,      0.6,      0.7,      0.8,      0.9,      1)
REFS1<-c(1.080, 1.109, 1.210,         1.214,  1.252,  1.350,  1.348,  1.343,  1.235,  1.191,  0.908)
REFS2<-c(0.95,  0.997, 0.989,  1.130,          1.289,  1.325,  1.305,  1.216,  1.101,  0.995,  0.985)
plot(values,REFS1,lwd=2,lty=1,type="l",pch=15,xlim=c(0,1),ylim=c(0,1.5),font.main=3,
cex.main=0.8,xlab=xlab.names,ylab="Relative Efficiency", main=main.names)
lines(values,REFS2,lwd=2,lty=2,type="l",pch=16,xlim=c(0,1),ylim=c(0,1.5),xlab=xlab.names,yl
ab="Relative Efficiency")legend("topleft",inset=.05,lwd=2,pch=c(15,16),lty=c(1,2),cex=0.8,
title="", c("QL","GL"),horiz=F)
REFS3<-c(0.431,  0.488,       0.549,  0.627,  0.717,  0.796,  0.739,  0.715,  0.690,  0.652,  0.479)
REFS4<-c(0.946,  0.908,       0.885,  0.942,  1.493,  0.898,  1.541,  0.953,  0.894,  0.958,  0.902)
REFS5<-c(0.589,  0.429,       0.348,  0.411,  0.495,  0.517,  0.524,  0.459,  0.398,  0.306,  0.297)
main.names1<-expression(paste("(d) Simulation: Estimated Relative Efficiencies of QL, GL,
M1, M2 and M3 vs ", hat(beta), ",  ", alpha,"=0.7", " using McGBB distribution"))
```

```
plot(values,REFS1,lwd=2,lty=1,type="l",pch=15,xlim=c(0,1),ylim=c(0,1.5),font.main=3,
cex.main=0.8,xlab=xlab.names,ylab="Relative Efficiency", main=main.names1)
lines(values,REFS2,lwd=2,lty=2,type="l",pch=16,xlim=c(0,1),ylim=c(0,1.5),xlab=xlab.names,yl
ab="Relative Efficiency")
lines(values,REFS3,lwd=2,lty=3,type="l",pch=17,xlim=c(0,1),ylim=c(0,1.5),xlab=xlab.names,yl
ab="Relative Efficiency")
lines(values,REFS4,lwd=2,lty=4,type="l",pch=18,xlim=c(0,1),ylim=c(0,1.5),xlab=xlab.names,yl
ab="Relative Efficiency")
lines(values,REFS5,lwd=2,lty=5,type="l",pch=24,xlim=c(0,1),ylim=c(0,1.5),xlab=xlab.names,yl
ab="Relative Efficiency")
legend("topleft",inset=.05,lwd=2,pch=c(15,16,17,18,24),lty=c(1,2,3,4,5), cex=0.8,title="",
c("QL","GL","M1","M2","M3"),horiz=F)
```