

Estimating Equations for Estimation of McDonald Generalized Beta—Binomial Parameters

Nthiwa M. Janiffer*, Ali Islam, Orawo Luke

Department of Mathematics, Egerton University, Njoro, Kenya

Email: *janienthiwa@yahoo.com, asislam54@yahoo.com, orawo2000@yahoo.com

Received 4 August 2014; revised 9 September 2014; accepted 25 September 2014

Copyright © 2014 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

There has been a considerable recent attention in modeling over dispersed binomial data occurring in toxicology, biology, clinical medicine, epidemiology and other similar fields using a class of Binomial mixture distribution such as Beta Binomial distribution (BB) and Kumaraswamy-Binomial distribution (KB). A new three-parameter binomial mixture distribution namely, McDonald Generalized Beta Binomial (McGGB) distribution has been developed which is superior to KB and BB since studies have shown that it gives a 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. The dispersion parameter will be treated as nuisance in the analysis of proportions since our interest is in the parameters of McGGB distribution. In this paper, we consider estimation of parameters of this McGGB model using Quasi-likelihood (QL) and Quadratic estimating functions (QEEs) with dispersion. By varying the coefficients of the QEE's we obtain four sets of estimating equations which in turn yield four sets of estimates. We compare small sample relative efficiencies of the estimates based on QEEs and quasi-likelihood with the maximum likelihood estimates. The comparison is performed using real life data sets arising from alcohol consumption practices and simulated data. These comparisons show that estimates based on optimal QEEs and QL are highly efficient and are the best among all estimates investigated.

Keywords

Maximum Likelihood, McDonald Generalized Beta Binomial, Simulation, Quadratic Estimating Equations, Quasi-Likelihood

*Corresponding author.

1. Introduction

Estimating functions have for sometimes been a key concept and subject of inquiry in research and it is known to be the most general method of estimation. The basis of this method is a set of simultaneous equations involving both the data and the unknown model parameters. 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 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. The usual procedure is to take a parametric model, such as, the McDonald Generalized beta-binomial model to allow over as well as under dispersion and obtain maximum likelihood estimates of the parameters McDonald Generalized Beta Binomial (McGGB) distribution is a three-parameter distribution which is superior to KB in handling over dispersed binomial data. This procedure may produce inefficient or biased estimates when the parametric model does not fit the data well. Alternatively, more robust estimates, such as moment estimates, quasi-likelihood estimates (Breslow, 1990 [1]; Moore and Tsiatis, 1991 [2]), extended quasi-likelihood estimates (Nelder and Pregibon, 1987 [3]), the Gaussian likelihood estimates (Whittle, 1961 [4]; Crowder, 1985 [5]), estimates based on the pseudo-likelihood estimating equations of Davidian and Carroll (1987) [6] and estimates based on quadratic estimating functions of Crowder (1987) [7] and Godambe and Thompson (1989) [8] can be considered. In this paper we consider estimating the parameters of McDonald Generalized Beta Binomial by the quadratic estimating equations (QEE's) of Crowder (1987) [7] and Godambe and Thompson (1989) [8] and compared the small sample efficiency and bias properties of these estimates with the maximum likelihood estimates. By varying the coefficients of the QEE's we obtain four sets of estimating equations. We compare the small sample efficiency of the five sets of estimates obtained by the QEE's and the quasi-likelihood estimates with the maximum likelihood estimates. We compare estimated relative efficiencies of the estimates for two sets of real life data arising from alcohol consumption practices and simulation study. Estimation of the parameters by the six methods is discussed in Section 3. In Section 4 we compare small sample relative efficiencies. This study shows that if interest is on the point parameters then the GL is the method of choice followed by QL.

2. 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 [9]; McDonald and Xu, 1995 [10]) with three parameters, α , β and γ . The probability density function of p is then given by

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

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

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

McDonald Generalized Beta Binomial Distribution

A random variable Y is said to have McDonald Generated Beta Binomial (McGGB) Distribution with parameter α , β and γ if and only if it satisfies the following stochastic representation. $Y/p \sim \text{Bin}(n, p)$ and $p \sim \text{GB1}(\alpha, \beta, \gamma)$, where α , β and γ are positive real numbers. This distribution was denoted as, $Y \sim \text{McGGB}(n, \alpha, \beta, \gamma)$.

In general, a Binomial mixture is obtained through an integration approach. Suppose Y follows a binomial distribution given by $\text{Bin}(n, p)$ and $Y/p \sim \text{Bin}(n, p)$. Unconditional PMF of the Y can be obtained by evaluating the integral

$$P_y(y) = \int P_{Y/p} f_p(p \setminus \Theta) dp \quad (3)$$

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

3. Estimation of Parameters of McDonald Generalized Beta-Binomial Distribution

3.1. Maximum Likelihood Method

The three unknown parameters of McGBB distribution have been estimated using the maximum likelihood estimation technique. Let $Y = (y_1, y_2, \dots, y_N)$ 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,

$$l(\Theta) = \sum_{k=0}^N \log \binom{n}{y} + \sum_{k=0}^N \log \left(\frac{1}{B(\alpha, \beta)} \right) + \sum_{k=0}^N \log \left[\sum_{j=0}^{n-y_k} (-1)^j \binom{n-y_k}{j} B \left(\frac{y_k}{\gamma} + \alpha + \frac{j}{\gamma}, \beta \right) \right]. \tag{4}$$

3.2. Quasi-Likelihood

The quasi-likelihood (Wedderburn, 1974 [11]) is based on the knowledge of the form of first two moments of the random variable $z_i = Y_i/n_i$. Where $E(z_i) = \pi_i$, and $\text{var}(z_i) = \pi(1-\pi)\{1+(n-1)\rho\}$. While with $Y_i = z_i n_i$ then, $E(Y) = n\pi_i$ and $\text{var}(Y) = n\pi(1-\pi)\{1+(n-1)\rho\}$.

The quasi-likelihood with the above mean and variance is given by $Q = \sum_{i=1}^n Q(z_i, \pi_i, \rho)$

where, $Q = (z_i, \pi_i, \rho) = \int_{z_i}^{\pi_i} \frac{(z_i - t)n_i}{t(1-t)\{1+(n_i-1)\rho\}} dt$.

By virtue of independence between samples, the quasi-likelihood with the above means and variance is given by:

$$Q = (z_i, \pi_i, \rho) = \int_{z_i}^{\pi_i} \frac{(z_i - \pi_i)n_i}{\pi_i(1-\pi_i)\{1+(n_i-1)\rho\}} d\pi_i \text{ and } U_j(\beta, \rho) = \frac{\partial Q}{\partial \beta_j} = \sum_{i=1}^n \frac{(z_i - \pi_i)n_i}{\pi_i(1-\pi_i)\{1+(n_i-1)\rho\}} = 0, \tag{5}$$

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

We denote Equation (5) by λ_{QL} . In this case $d_{ij}\beta$ is given as $d_{ij}\beta = \frac{\partial \pi_i}{\partial \beta_j} = \pi_i(1-\pi_i)X_{ij}$. Given $Z = \frac{y}{n}$ we have:

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

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

Then the partial derivatives for the three parameters α, β, γ given ρ as also obtained as follow:

$$\frac{\partial Q}{\partial \alpha} = \sum_{y=0}^n \left\{ \left(\frac{1}{(1+(n-1)\rho)} \right) \left[\left(y - \left(\frac{\pi(n-y)}{(1-\pi)} \right) \right) \left(\psi(\alpha + \beta) + \psi\left(\alpha + \left(\frac{1}{\gamma}\right)\right) - \psi(\alpha) + \psi\left(\alpha + \beta + \left(\frac{1}{\gamma}\right)\right) \right) \right] \right\} \tag{6}$$

$$\frac{\partial Q}{\partial \beta} = \sum_{y=0}^n \left\{ \left(\frac{1}{(1+(n-1)\rho)} \right) \left[\left(y - \left(\frac{\pi(n-y)}{(1-\pi)} \right) \right) \left(\psi(\alpha + \beta) - \psi\left(\alpha + \beta + \left(\frac{1}{\gamma}\right)\right) \right) \right] \right\} \tag{7}$$

$$\frac{\partial Q}{\partial \gamma} = \sum_{y=0}^n \left\{ \left(\frac{1}{(1+(n-1)\rho)} \right) \left[\left(\frac{y}{\gamma^2} - \left(\frac{\pi(n-y)}{\gamma^2(1-\pi)} \right) \right) \left(\psi\left(\alpha + \beta + \left(\frac{1}{\gamma}\right)\right) - \psi\left(\alpha + \left(\frac{1}{\gamma}\right)\right) \right) \right] \right\}. \tag{8}$$

3.3. Quadratic Estimating Equations

By considering estimating functions quadratic in z_i the QEEs has general form a

$g_\lambda = \sum_i^m \left[a_{i\lambda} (z_i - \pi_i) + b_{i\lambda} \left\{ (z_i - \pi_i)^2 + \sigma_{i\lambda}^2 \right\} \right]$, Crowder (1987) [7], 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 for McGGB distribution is found as follows.

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

$$U_j(\beta, \rho) = \sum_i^n \left[a_{i\beta_j} (z_i - \pi_i) + b_{i\beta_j} \left\{ (z_i - \pi_i)^2 + \sigma_{i\lambda}^2 \right\} \right] = 0. \tag{9}$$

If we take

$$a_{i\beta_j} = \left[\frac{1}{\sigma_{i\lambda}^2} + \frac{(1-2\pi_i)^2 n_i}{2(1-\rho)\pi_i^2(1-\pi_i)^2} d_{ij} \beta \right],$$

$$b_{i\beta_j} = \frac{-n_i \rho (1-2\pi_i) d_{ij} (\beta)}{2(1-\rho)\pi_i(1-\pi_i)\sigma_{i\lambda}^2}.$$

We obtain the Gaussian estimating equations. We denote this Equation (10) by λ_{GL}

$$U_j(\beta, \rho) = \sum_i^n \left[\left[\frac{1}{\sigma_{i\lambda}^2} + \frac{(1-2\pi_i)^2 n_i}{2(1-\rho)\pi_i^2(1-\pi_i)^2} d_{ij} \beta \right] (z_i - \pi_i) + \frac{-n_i \rho (1-2\pi_i) d_{ij} (\beta)}{2(1-\rho)\pi_i(1-\pi_i)\sigma_{i\lambda}^2} \left\{ (z_i - \pi_i)^2 + \sigma_{i\lambda}^2 \right\} \right] = 0 \tag{10}$$

If we take $a_{i\beta_j} = \frac{d_{ij} \beta}{\sigma_{i\lambda}^2}$, and $b_{i\beta_j} = 0$.

Then we obtain the unbiased estimating equations (QEE's) for McDonald Generalized Binomial Distribution. These equations were obtained by combining the quasi-likelihood estimating equations for the regression parameters and the optimal quadratic estimating equations of Crowder (1987) [7] for the dispersion parameter after setting Υ_1 and Υ_2 to zero.

We denote the estimates so obtained from Equation (11) by λ_{M1}

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

This simplifies to,

$$U_j(\beta, \rho) = \sum_i^n \left[\frac{d_{ij} \beta}{\sigma_{i\lambda}^2} (z_i - \pi_i) \right]. \tag{11}$$

For

$$a_{i\beta_j} = \frac{[-\Upsilon_{2i\lambda} + 2 + \Upsilon_{1i\lambda} (1-2\pi_i) \sigma_\lambda / \pi_i (1-\pi_i)] d_{ij} (\beta)}{\sigma_{i\lambda}^2 \Upsilon_{i\lambda}}$$

$$b_{i\beta_j} = \frac{[\Upsilon_{1i\lambda} - (1-2\pi_i) \sigma_\lambda / \pi_i (1-\pi_i)] d_{ij} (\beta)}{\sigma_{i\lambda}^3 \Upsilon_{i\lambda}}, \text{ where } \Upsilon_{i\lambda} = \Upsilon_{2i\lambda} + 2 - \Upsilon_{1i\lambda}^2.$$

We obtain the optimal quadratic estimating equations. We note that the forms of the skewness $\Upsilon_{1\lambda}$ and the kurtosis $\Upsilon_{2\lambda}$ are not known. We then take these 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 $\mu_{4i} = \mu_{2i} \left[\frac{\{1 + (2n_i - 1) \rho\} \{1 + (3n_i - 1) \rho\} \{1 - 3\pi_i (1 - \pi_i)\}}{1 - \rho} + (n_i - 1) \{\rho + 3n_i \mu_{2i}\} \right] \frac{(1 - \rho)}{(1 + \rho)(1 + 2\rho)n_i^2}.$

We denote the estimates obtained by solving these optimal quadratic estimating equations by λ_{M2} . Further we also denote the estimates obtained by solving the optimal quadratic estimating equations with $Y_{1i\lambda} = Y_{2i\lambda} = 0$ by λ_{M3} . Note the estimates λ_{M3} are also obtained by using the pseudo-likelihood estimating equations of Davidian and Carroll (1987) [7].

$$U_j(\beta, \rho) = \sum_i^n \left[\frac{\left[-Y_{2i\lambda} + 2 + \frac{Y_{1i\lambda}(1-2\pi_i)\sigma_\lambda}{\pi_i(1-\pi_i)} \right] d_{ij}(\beta)}{\sigma_{i\lambda}^2 Y_{i\lambda}} (z_i - \pi_i) + \frac{\left[Y_{1i\lambda} - \frac{(1-2\pi_i)\sigma_\lambda}{\pi_i(1-\pi_i)} \right] d_{ij}(\beta)}{\sigma_{i\lambda}^3 Y_{i\lambda}} \left\{ (z_i - \pi_i)^2 + \sigma_{i\lambda}^2 \right\} \right] = 0 \quad (12)$$

$$\sum_i^n \left[\frac{[-2] d_{ij}(\beta)}{\sigma_{i\lambda}^2 Y_{i\lambda}} (z_i - \pi_i) + \frac{[-(1-2\pi_i)\sigma_\lambda / \pi_i(1-\pi_i)] d_{ij}(\beta)}{\sigma_{i\lambda}^3 Y_{i\lambda}} \left\{ (z_i - \pi_i)^2 + \sigma_{i\lambda}^2 \right\} \right] = 0 \quad (13)$$

4. Small-Sample Relative Efficiency

The asymptotic relative efficiency may not be very useful when comparing different estimators in small samples. So we conducted a simulation study using relatively small n alongside the real data. We compare the small sample relative efficiency of the estimates obtained by the five estimation procedures: QL; GL; M1; M2; M3 with the MLE. The estimated Relative efficiency of α is $\text{var}(\alpha_{ML}) / \text{var}(\alpha_t)$ where $t = \text{QL}, \text{GL}, \text{M1}, \text{M2}, \text{M3}$. 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. The relative efficiency results for the McGGB parameters are summarized in Table 2 for the real data and those for simulated data are summarized in Table 3 and Table 4 and plotted in Figure 1 for simulated data.

5. Estimation

Table 1 shows the data set used by Alanko and Lemmens (1996) [12], Rodríguez-Avi *et al.* (2007) [13], and Chandrabose *et al.* (2013) [14] 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, $n =$ frequency of consumption. We used this data in Table 1 to obtain the estimates for α, β and γ and estimated relative efficiencies by the six different procedures as given in Table 2.

6. Simulation

We compare the relative efficiency of the estimates α, β and γ obtained by the six estimation procedures

Table 1. Number of alcohol consumption days and the frequency of consumption.

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

Table 2. The estimate α, β and γ and their estimated Relative efficiencies by MLE QL M1 M2 and M3 methods for the real data.

Parameters	Parameter estimates						Estimated relative efficiencies					
	MLE	QL	λ_{GL}	λ_{M1}	λ_{M2}	λ_{M3}	MLE	QL	λ_{GL}	λ_{M1}	λ_{M2}	λ_{M3}
α	0.0333	0.0281	0.0301	0.0287	0.0312	0.0282	1.000	1.0912	1.0424	0.6121	0.9352	0.3292
β	0.1797	0.1502	0.1671	0.1655	0.1671	0.1611	1.000	0.9085	0.9831	0.5192	0.9811	0.3615
γ	26.7312	25.5541	25.8127	24.8421	25.4523	24.6746	1.000	0.9010	0.8742	0.5320	0.9381	0.3510

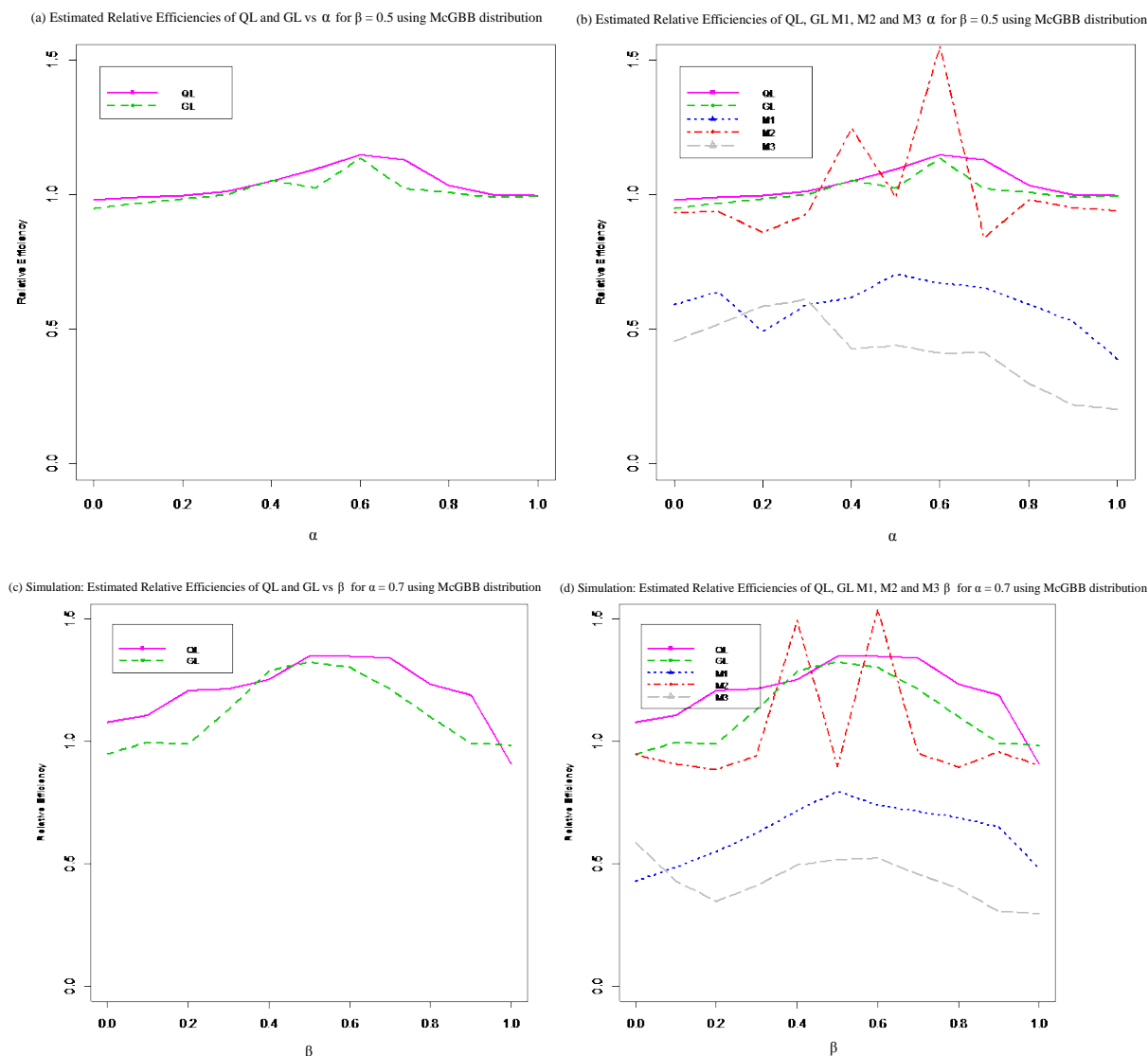


Figure 1. Plot of relative efficiencies for various estimators relative to that of the MLE under McDonald Generalized Beta-Binomial model: for (a) relative efficiency comparison for α varied when $\beta = 0.5$ for QL and GL procedures while (b) α varied when $\beta = 0.5$ for simulated data for all procedures; for (c) relative efficiency comparison for β varied when $\alpha = 0.7$ for QL and GL procedures while (d) β varied when $\alpha = 0.7$ for the simulated data for all procedures.

using weekly (7 days) alcohol consumption survey data and simulated data for the survey of weekly alcohol consumption for a small time frame ($n = 7$ days) along with estimates of the parameters of the maximum likelihood method. Estimated Relative efficiency of α is $\text{var}(\alpha_{ML})/\text{var}(\alpha_t)$ 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. Using the combination of α, β and γ parameters. We simulated 5000 samples from the MacDonal generalized Beta-Binomial distribution using the weekly alcohol consumption data. 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 for the six procedures. **Figure 1:** Maximum likelihood procedure relative efficiency comparison for (a) when we fix $\gamma = 1$ and $\beta = 0.5$ then α varied for GL and QL procedures, (b) α varied when $\beta = 0.5$ for all procedures. While (c) β varied when $\alpha = 0.7$ and fix $\gamma = 1$ for GL and QL procedures and (d) β varied when $\beta = 0.5$ for all procedures under simulated data.

7. Discussion

From **Table 2**, **Table 3** and **Table 4** we see that the methods QL, GL and $M2$ all consistently provide high efficiency (never below 0.83). Efficiency of parameters by the method GL is consistently the best. 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. Generally the estimates of parameters by all estimating functions methods have high efficiencies. In this paper we showed that the estimates obtained through small sample parameter estimates and efficiencies obtained during data analysis are the best for GL followed by QL and then $M2$ method (estimates based on the optimal quadratic estimating equations with the third and the fourth moments of the McGGB distribution) are consistent. The next best, at the cost of some loss of efficiency, are the $M1$ and then $M3$ seems to be the least method. Therefore, when data follow a McGGB distribution, these methods are expected to have high efficiency as compared to MLEs.

Table 3. Relative efficiencies for various estimators for α varied when $\beta = 0.5$ and $\gamma = 1$ for simulated data for all procedures.

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

Table 4. Relative efficiencies for various estimators for β varied when $\alpha = 0.7$ and $\gamma = 1$ for simulated data for all procedures.

β varied	Estimated relative efficiencies					
	MLE	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	1.35	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

8. Conclusion

The estimation functions are based on the knowledge of moments and one of the advantages of this approach is that it is robust to model misspecification. The comparison results in this paper indicate that the Estimating Equations are superior to MLE. The small relative efficiency for the estimates results also shows that estimates using optimal quadratic estimating functions of Crowder (1987) are highly efficient and are the best among all estimates investigated followed by Quasi-likelihood. Thus, we propose quadratic estimating function for estimation of point parameters of any model inclusive of McDonald Generalized Beta-Binomial instead of MLEs since they are consistent and robust to variance misspecification.

References

- [1] Breslow, N.E. (1990) Tests of Hypothesis in Over-Dispersed Poisson Regression and Other Quasi-Likelihood Models. *Journal of the American Statistical Association*, **85**, 565-571. <http://dx.doi.org/10.1080/01621459.1990.10476236>
- [2] Moore, D.F. and Tsatis, M. (1991) Robust Estimation of the Standard Error in Moment Methods for Extra Binomial/Extra Poisson Variation. *Biometrika*, **47**, 383-401. <http://dx.doi.org/10.2307/2532133>
- [3] NeIder, J.A. and Pregibon, D. (1987) An Extended Quasi-Likelihood Function. *Biometrika*, **74**, 221-232. <http://dx.doi.org/10.1093/biomet/74.2.221>
- [4] Whittle, P. (1961) Gaussian Estimation in Stationary Time Series. *Bulletin of the International Statistical Institute*, **39**, 1-26.
- [5] Crowder, M.J. (1985) Gaussian Estimation for Correlated Binomial Data. *Journal of the Royal Statistical Society, Series B*, **47**, 229-237.
- [6] Davidian, M. and Carrol, R.J. (1987) Variance Function Estimation. *Journal of the American Statistical Association*, **82**, 1079-1091. <http://dx.doi.org/10.1080/01621459.1987.10478543>
- [7] Crowder, M.J. (1987) On Linear Quadratic Estimating Functions. *Biometrika*, **74**, 591-597. <http://dx.doi.org/10.1093/biomet/74.3.591>
- [8] Godambe, V.P. and Thompson, M.E. (1989) An Extension of Quasi-Likelihood Estimation. *Journal of Statistical Planning and Inference*, **22**, 137-152. [http://dx.doi.org/10.1016/0378-3758\(89\)90106-7](http://dx.doi.org/10.1016/0378-3758(89)90106-7)
- [9] McDonald, J.B. (1984) Some Generalized Functions for the Size Distribution of Income. *Econometrica: Journal of the Econometric Society*, **52**, 647-663. <http://dx.doi.org/10.2307/1913469>
- [10] McDonald, J.B. and Xu, Y.J. (1995) A Generalization of the Beta Distribution with Applications. *Journal of Econometrics*, **66**, 133-152. [http://dx.doi.org/10.1016/0304-4076\(94\)01612-4](http://dx.doi.org/10.1016/0304-4076(94)01612-4)
- [11] Wedderburn, R.M. (1974) Quasi-Likelihood Functions, Generalized Linear Models and the Gauss Newton Method. *Biometrics*, **61**, 439-447.
- [12] Alanko, T. and 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.
- [13] Rodríguez-Avi, J., Conde-Sánchez, A., Sáez-Castillo, A.J. and Olmo-Jiménez, M.J. (2007) A Generalization of the Beta-Binomial Distribution. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, **56**, 51-61. <http://dx.doi.org/10.1111/j.1467-9876.2007.00564.x>
- [14] 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 Overdispersion. *International Journal of Statistics and Probability*, **2**, 213-223.